Mathematics 205B

Topology — II

Course Notes

Revised, Winter 2012

Department of Mathematics

University of California, Riverside

Table of Contents

Preface

This is the second of three entry level graduate courses in topology and geometry. The basic texts for this course are the following:

[**M**] **J. R. Munkres.** *Topology* (Second Edition), *Prentice-Hall, Saddle River NJ*, 2000. ISBN: 0–13–181629–2. [This is the text for the previous course in the sequence.]

[**H**] **A. Hatcher.** *Algebraic Topology* (Third Paperback Printing), *Cambridge University Press, New York NY*, 2002. ISBN: 0–521–79540–0.

This book can be legally downloaded from the Internet at no cost for personal use, and here is the link to the online version:

`www.math.cornell.edu/∼hatcher/AT/ATpage.html`

To a great extent these books correspond to the two main parts of the course:

1. A continuation of the material on fundamental groups and covering spaces, picking things up where the previous course ended.

2. A brief introduction to homology theory and its topological applications.

The material in the first part of the course has become fairly standard, but the material in the second part is much less so, particularly at the introductory level. Both parts of the course have a common theme: The creation of algebraic "pictures" of a topological space; these are obtained by studying algebraic constructions involving certain types of topological configurations in the space. Fundamental groups study the 1-dimensional configurations given by closed curves, and the basic idea of homology theory is to study some analogous configurations in higher dimensions. One way to compare these subjects is to describe the conclusions which follow from the respective methods:

(a) Using point set topology and fundamental groups, one can show that $\mathbb{R}$ and $\mathbb{R}^n$ are not homeomorphic if $n \geq 2$, and $\mathbb{R}^2$ and $\mathbb{R}^m$ are not homeomorphic if $m \neq 1$.

(b) Using homology theory, one can show that $\mathbb{R}^n$ and $\mathbb{R}^m$ are not homeomorphic if $m \neq n$.

Clearly the second statement is much stronger than the first, and this reflects the powerful impact that homological methods had on 20^{th} century mathematics; their influence has reached well beyond topology and geometry into many areas of algebra and even some areas of analysis. Pages 4 and 5 in the course directory file

`morgan-lamberson.pdf`

gives more specific information and also mentions the downside: Homology theory requires an extremely substantial amount of work to produce a mathematically complete treatment, and much of the time the motivation for the constructions is not at all apparent to someone who is just beginning to study the subject.

To a great extent this reflects a basic mathematical problem; namely, to create a rigorous logical setting for translating geometrical intuition into algebraic formalism which can be manipulated effectively and applied successfully to analyze geometrical questions. Without a logically rigorous framework the reliability of geometrical results cannot be guaranteed, but some intuition is often indispensable for making any sense out of the formal constructions.

Our approach to this dilemma will be to stress the motivation and to develop the subject so that the geometric applications have high priority; one can expand this into a logically complete, but reasonably accessible, treatment by following various links to other files in the course directory. As suggested above, one reason for choosing this approach is a view the effort needed to go through large amounts of formal technical machinery might be more bearable if one knows that it has substantial applications. We have already mentioned one such application, and needless to say there will be others.

Comments on the texts

Taking all things into account, the first part of Munkres (on point set topology) is one of the very best accounts of the subject, with an excellent balance of clear exposition, logical completeness and drawings to motivate the underlying geometrical content of the subject (there are some peculiar choices of terms and symbolism, and in a number of instances more motivation would help, but the perfect text is an ideal which is rarely if ever realized). The second part of Munkres comes close to meeting this standard, but there are numerous cases where more motivational comments and drawings are probably necessary, and sometimes the logical thoroughness of the exposition interferes with the clarity of the exposition. To its credit, the second part gives logically complete accounts of several basic applications of topology to basic geometrical results like the Fundamental Theorem of Algebra and the Jordan Curve Theorem (a simple closed curve in the plane separates it into two connected pieces), but the proofs really push the theory in the book to its limits, and consequently the reasoning is often very delicate and difficult to follow. We shall see that homology theory often yields much simpler and more conceptual proofs.

Hatcher's book begins by covering the same topics which appear in the second half of Munkres, and it proceeds to go much further in the subject. The challenges faced in covering the further material are much greater than the corresponding challenges in Munkres. In particular, the gap between abstract formalism and geometrical intuition is much greater, and it is not clear how well any single book can reconcile these complementary factors. More often than not, algebraic topology books stress the former at the expense of the latter, and one important strength of Hatcher's book is that its emphasis tilts very much in the opposite direction. The book makes a sustained effort to include examples that will provide insight and motivation, using pictures as well as words, and it also attempts to explain how working mathematicians view the subject. Because of these objectives, the exposition in Hatcher is significantly more casual than in most if not all other books on the subject. Unfortunately, the book's informality is arguably taken too far in numerous places, leading to significant problems in several directions; these include assumptions about prerequisites, clarity, wordiness, thoroughness and some sketchy motivations that are difficult for many readers to grasp. One goal of these notes is to address some issues along these lines.

Some additional references

Here are four other references; they are generally at a higher level than the present course, but they should be within the reach of students in this course. The first is a book that has been used as a text in the past, the second is a fairly detailed history of the subject during its formative years, and the last two are classic (but not outdated) books; the first and third also have detailed historical notes.

J. W. Vick. *Homology Theory*. (Second Edition). *Springer–Verlag, New York etc.*, 1994. ISBN: 3–540–94126–6.

J. Dieudonné. *A History of Algebraic and Differential Topology* (1900 – 1960). *Birkhäuser Verlag, Zurich etc.*, 1989. ISBN: 0–817–63388–X.

S. Eilenberg and N. Steenrod. *Foundations of Algebraic Topology*. (Second Edition). *Princeton University Press, Princeton NJ*, 1952. ISBN: 0–691–07965–X.

E. H. Spanier. *Algebraic Topology*, *Springer–Verlag, New York etc.*, 1994.

The `amazon.com` sites for Hatcher's and Spanier's books also give numerous other texts in algebraic topology that may be useful. Finally, there are two other books by Munkres that we may quote throughout these notes. The first will be denoted by [MunkresEDT] and the second by [MunkresAT]; if we simply refer to "Munkres," it will be understood that we mean the previously cited book, *Topology* (Second Edition).

J. R. Munkres. *Elementary differential topology*. (Lectures given at Massachusetts Institute of Technology, Fall, 1961. Revised edition. Annals of Mathematics Studies, No. 54.) *Princeton University Press, Princeton, NJ*, 1966. ISBN: 0–691–09093–9.

J. R. Munkres. *Elements of Algebraic Topology*. *Addison-Wesley, Reading, MA*, 1984. (Reprinted by Westview Press, Boulder, CO, 1993.) ISBN: 0–201–62728–0.

Prerequisites

The name "algebraic topology" suggests that the subject uses input from both algebra and topology, and this is in fact the case; since topology began as a branch of geometry, it is also reasonable to expect that some geometric input is also required. Our purpose here is to summarize the main points from prerequisite courses that will be needed. Additional background material which is usually not covered explicitly in the prerequisites will be described in the first unit of these notes.

Set theory

Everything we shall need from set theory is contained in the following online directory:

`http://math.ucr.edu/∼res/math144`

In particular, a fairly complete treatment is contained in the documents `setsnotes`n`.pdf`, where $1 \leq n \leq 8$.

There are two features of the preceding that are somewhat nonstandard. The first is the definition of a function from a set A to another set B. Generally this is given formally by the graph, which is a subset $G \subset A \times B$ such that for each $a \in A$ there is a unique $b \in B$ such that $(a, b) \in G$. Our definition of function will be a **triple** $f = (A, G, B)$, where $G \subset A \times B$ satisfies the condition in the preceding sentence. The reason for this is that we must specify the target set or **codomain** of the function explicitly; in fact, the need to specify the codomain has already arisen at least implicitly in prerequisite graduate topology courses, specifically in the definition of the fundamental group. A second nonstandard feature is the concept of **disjoint union** or **sum** of an indexed family $\{ X_\alpha \}$ of sets. The important features of the disjoint sum, which is written $\amalg_\alpha X_\alpha$, are that it is a union of subsets Y_α which are canonically in 1–1 correspondence with the sets X_α and that $Y_\alpha \cap Y_\beta = \emptyset$ if $\alpha \neq \beta$. Another source of information on such objects is Unit V of the online notes for Mathematics 205A which are cited below.

Topology

This course assumes familiarity with the basic material in the first graduate level topology course: This corresponds to material in the following sections of Munkres:

12–14, 17, 20, 22–26, 29–31, 51–54, 58–59, 67–71

(and also the Supplementary Exercises for Chapter 2)

Supplementary material for Sections 12–31 can be found in the following online directory:

`http://math.ucr.edu/∼res/math205A`

In particular, the file `gentopnotes2008.pdf` contains a fairly complete set of lecture notes for the portions of the course that do not involve fundamental groups and covering spaces. This material is based upon the textbook by Munkres cited in the Preface. Two major differences between the notes and Munkres appear in Unit V. The discussion of quotient topologies is somewhat different from that of Munkres, and in analogy with the previously mentioned discussion of set-theoretic disjoint sums there is a corresponding construction of disjoint sum for an indexed family of topological spaces. Some supplementary material for Sections 51-82 can be found in the following online directory for an older version of 205B.

`http://math.ucr.edu/~res/math205B`

In addition to the sections listed above, it might be worthwhile to look also at the supplementary exercises for Chapter 13.

Algebra

Some concepts in group theory are needed; most are at the undergraduate level. Several other concepts from group theory are presented in Munkres and will be covered in the course.

We shall also need the following basic result, which is proved in graduate level algebra courses:

STRUCTURE THEOREM FOR FINITELY GENERATED ABELIAN GROUPS. *Let G be a finitely generated abelian group (so every element can be written as a monomial in integral powers of some finite subset $S \subset G$). Then G is isomorphic to a direct sum*

$$(H_1 \oplus \cdots \oplus H_b) \oplus (K_1 \oplus \cdots \oplus K_s)$$

where each H_i is infinite cyclic and each K_j is finite of order t_j such that t_{j+1} divides t_j for all j. — For the sake of uniformity set $t_j = 1$ if $j > s$. Then two direct sums as above which are given by $(b; t_1, \cdots)$ and $(b'; t'_1, \cdots)$ are isomorphic if and only if $b = b'$ and $t_j = t'_j$ for all j.

A proof of this fundamental algebraic result may be found in Sections II.1 and II.2 of the following standard graduate algebra textbook:

> **T. Hungerford.** *Algebra.* (Reprint of the 1974 original edition, Graduate Texts in Mathematics, No. 73.) *Springer-Verlag, New York–Berlin–etc.*, 1980. ISBN: 0–387–90518–9.

Material from standard undergraduate linear algebra courses will also be used as needed.

Analysis

We shall assume the basic material from an upper division undergraduate course in real variables as well as material from a lower division undergraduate course in multivariable calculus through the theorems of Green and Stokes as well as the 3-dimensional Divergence Theorem. The classic text by W. Rudin (*Principles of Mathematical Analysis*, Third Edition) is an excellent reference for real variables, and the following multivariable calculus text contains more information on the that subject than one can usually find in the usual 1500 page calculus texts (the book is far from perfect, but especially at the graduate level it is useful as a background reference).

> **J. E. Marsden and A. J. Tromba.** *Vector Calculus* (Fifth Edition), *W. H. Freeman & Co., New York NY*, 2003. ISBN: 0–7147–4992–0.

The concept of a category does not appear explicitly in Munkres, but it is implicit in many places. Although we shall not need a formal treatment of category theory at the beginning of the course, eventually some of the basic ideas will be indispensable, so we shall describe the concepts that arise very early in the course. The file

`categories2012.pdf`

gives a more organized treatment of the topics treated here, and **reading the document is an implicit assignment for this course.**

A **category** is an abstract mathematical system which reflects some very basic features of many classes of mathematical objects and the well-behaved morphisms relating them. In set theory the objects and morphisms are sets and functions of sets, and in topology the most basic examples involve topological spaces and continuous mappings. There are many algebraic examples, including groups and morphisms, vector spaces over a fixed field $\mathbb{F}$ and $\mathbb{F}$-linear transformations, or partially ordered sets and monotonically increasing functions. Many other examples appear in the file cited above.

In all cases one has objects and morphisms from one object to another with specified domains (sources) and codomains (targets) which are written with notation like $f : A \to B$. There are also algebraic operations which behave like composition of functions in the following respects:

(i) The composition $g \circ f$ of g and f is defined if and only if the target of f is the source of g.

(ii) For each object X there is an "identity morphism" $1_X : X \to X$, and for each morphism $f : A \to B$ we have $1_B \circ f = f = f \circ 1_A$.

(iii) There is an associative law $h \circ (g \circ f) = (h \circ g) \circ f$ for threefold compositions.

The most important additional concept is that of an isomorphism between two objects A and B. This involves a pair of morphisms $f : A \to B$ and $g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. As elsewhere in mathematics, if one has such a pair of morphisms we say that f and g are inverse to each other (or inverses of each other).

The ubiquity of categories reflects a basic fact: If a class of mathematical objects is defined, it is usually possible to define a good concept of mappings or morphisms from one object to another without too much trouble. Thus it is natural to speculate about an appropriate notion of morphism relating one category to another. It turns out that there are **two** such notions called **contravariant functors** and **covariant functors**. A *covariant functor* is a system of transformations such that

(a) for each object X in the source category there is an associated object $T(X)$ in the target category,

(b) for each morphism $f : X \to Y$ in the source category there is an associated morphism $T(f) : T(X) \to T(Y)$ in the target category,

(c) the construction on morphisms preserves identity morphisms and compositions; the latter means that $T(g \circ f) = T(g) \circ T(f)$.

Here is an example involving topological spaces: If X is a topological space, let $T(X)$ be the set of continuous curves $\gamma : [0, 1] \to X$, and if $f : X \to Y$ is continuous define $T(f)\gamma = f \circ \gamma$. The fundamental group of a pointed space is a more sophisticated example of this sort going from pointed topological spaces to groups.

As noted above, there is also a dual concept of *contravariant functor* from one category to another; the main differences with covariant functors are that a morphism $f : A \to B$ is sent to $T(f) : T(B) \to T(A)$ (*i.e.*, the domain and codomain are switched) and the composition identity is $T(g \circ f) = T(f) \circ T(g)$ (*i.e.*, the order of composition is reversed).

One basic example of a contravariant functor is the dual space construction on a category of vector spaces over some field $\mathbb{F}$. Specifically, a vector space V is sent to the space V^* of $\mathbb{F}$-linear functionals $V \to \mathbb{F}$, and if $T : V \to W$ is a linear transformation then $T^* : W^* \to V^*$ sends a linear functional $f : V \to \mathbb{F}$ to the composite $T \circ f$.

Functors have a simple but far-reaching property which is fairly easy to prove: *If the morphisms f and g are inverse to each other and T is a functor (covariant or contravariant), then $F(f)$ and $F(g)$ are also inverse to each other.*

Since functors are mathematical objects, one can speculate even further about morphisms relating functors and whether such a notion is more than a formal curiosity. It turns out that there is a useful notion called a **natural transformation** of functors (where both the source and target have the same variance). Since we shall not need this concept until later in the course, we shall pass on discussing it here.

I. Further Properties of Covering Spaces

This course begins with three units which conclude the treatment of fundamental groups and covering spaces. The main objective of the first unit is to set up an important 1–1 correspondence between the isomorphism classes of connected covering spaces over a given "nice" space X (Hausdorff, connected, locally arcwise connected, with some sort of local simple connectivity condition) and subgroups of $\pi_1(X)$. The crucial step in formulating the correspondence is given by establishing a link between the fundamental group of X and suitably defined groups of automorphisms for connected covering spaces over X.

In several respects this correspondence is similar to the 1–1 correspondence in Galois Theory; namely, if $E \supset F$ is a Galois extension field and G is the group of all autormorphisms of E which are the identity on F, then there is a 1–1 order-reversing correspondence between the subgroups of G and the subfields L such that $E \supset L \supset F$, and a subgroup H is normal in G if and only if the corresponding subfield L_H is a Galois extension of F; in fact, it turns out that normal subgroups of $\pi_1(X)$ correspond to a special family of covering spaces called **regular coverings**. In the theory of Riemann surfaces, and more generally in algebraic geometry, the analogies between covering spaces and Galois Theory are more than formal coincidences.

IMPORTANT. Even though the word "covering" appears in the phrases "open covering" and "covering spaces," there is no direct connection between the two usages; however, in practice this ambiguity usually does not cause any difficulties.

Notational convention. In the textbook references for the various subsections, "**H**" indicates a reference from Hatcher and "**M**" indicates a reference from Munkres.

I.0 : Introduction

(**M**, §§51–55, 56; **H**, Ch. 0, Ch. 1 Introduction, §1.1)

We shall first give a very brief review of topics on fundamental groups and covering spaces from the previous course. In order to keep things from getting too routine, we shall add some comments that might help motivate or clarify the main ideas.

Topic 1. *Homotopy of mappings and homotopy classes.* To save time and space, we shall only mention many of the basic definitions and consequences; the former are easy to locate using the indices in Hatcher and Munkres, and the latter can be found in the same general parts of these books.

The single most important concept in this course and the last part of the preceding course is the notion of a homotopy relating two continuous mappings with the same domain X and codomain Y, and the most basic formal property of this concept is that it defines an equivalence relation on the continuous mappings from X to Y; it is customary to denote the associated set of equivalence classes (or *homotopy classes*) by $[X, Y]$. If two continuous mappings X and Y are homotopic then we write $f \simeq g$.

WARNING. If $f : X \to Y$ is a function (not necessarily continuous) and $Y \subset Y'$ (with the subspace topology if we are talking about topological spaces), then in elementary mathematics one often identifies f with the composite $i \circ f$ where i denotes the inclusion of Y in Y'; in other words, the target or codomain of the function is not viewed as a uniquely determined object. In this course it will be extremely important to view the codomain as uniquely determined for many reasons (see the discussion following Proposition 4 below).

One important property of the homotopy relation is that it is compatible with composition (if f and g are homotopic to f' and g', then $f \circ g$ and $f' \circ g'$ are also homotopic). This implies that the homotopy class construction is **functorial** in the sense that if $h : X' \to X$ and $k : Y \to Y'$ are continuous mappings then one has associated mappings of homotopy classes

$$h^* : [X, Y] \longrightarrow [X', Y] , \qquad k_* : [X, Y] \longrightarrow [X, Y']$$

sending a homotopy class $[f]$ to the classes $[f \circ h]$ and $[k \circ f]$ respectively; functoriality means that if h or k is the identity then so is h^* or k_*, and compositions are preserved in the senses that $(h_1 \circ h_2)^* = h_2^* \circ h_1^*$ (so the order of composition is *reversed*) and $(k_1 \circ k_2)_* = k_{1*} \circ k_{2*}$ (so the order of composition is *preserved*). This construction also has the following *homotopy invariance property*: *If $h_0 \simeq h_1$ or $k_0 \simeq k_1$, then $h_0^* = h_1^*$ or $k_0^* = k_1*$.*

In the same notation, if h or k is a homeomorphism then composition with h or k induces an isomorphism of continuous function sets $\mathcal{C}(X, Y) \cong \mathcal{C}(X', Y)$ or $\mathcal{C}(X, Y')$, and it follows similarly that one has associated isomorphisms of homotopy classes. In fact, h^* or k_* induces and isomorphism of homotopy classes under the weaker notion of homotopy equivalence; specifically, a continuous map $\varphi : A \to B$ is a homotopy equivalence if and only if there is a reverse map $\psi : B \to A$ (a homotopy inverse) such that $\psi \circ \phi$ and $\phi \circ \psi$ are homotopic to the identity maps of A and B respectively. If there is a homotopy equivalence from one space to another, we say that these spaces are *homotopy equivalent*; it is elementary to check that this defines an equivalence relation on topological spaces.

A space is said to be *contractible* if it has the homotopy type of a point; convex subsets of normed vector spaces are important examples of contractible spaces. If X is contractible and $h : X \to P$ is a homotopy equivalence where P consists of a single point, then it follows that for all spaces W we have isomorphisms

$$h_* : [W, X] \longrightarrow [W, P] \;\cong\; P$$

because there is a unique continuous map from W to P (the constant map sending everything to the unique point in P). To conclude this discussion, we note that the concept of homotopy equivalence would not be particularly useful if all spaces were contractible; particular results from the previous course imply that the standard unit circle $S^1 \subset \mathbb{R}^2$ is not (see the second consequence of Proposition 4 below).

Additional comments on homotopy of mappings and homotopy classes. Since the concept of homotopy is often introduced with a minimum of motivation, we shall give an important class of examples.

PROPOSITION 1. *Let X be a compact metric space, let U be an open subset of $\mathbb{R}^n$ for some n, and let $f : X \to U$ be continuous. Then there is some $\delta > 0$ such that if $g : X \to U$ is another continuous function satisfying $\mathbf{d}(\, f(\mathbf{x}),\, g(\mathbf{x})\,) < \delta$ for all $\mathbf{x}$, then for all such $\mathbf{x}$ the closed line segment joining $f(\mathbf{x})$ to $g(\mathbf{x})$, given by*

$$\{\ \mathbf{y} \mid \mathbf{y} = \ t\, f(\mathbf{x}) + (1-t) g(\mathbf{x}),\ \text{some } t \in [0,1]\ \}$$

lies entirely inside U. Consequently, the mapping $H : X \times [0,1] \to U$ defined by

$$H(x,t) \quad = \quad t\, f(\mathbf{x}) + (1-t) g(\mathbf{x})$$

is a homotopy from f to g.

In other words, for domains and codomains given as in the proposition, a mapping g is homotopic to f if g is sufficiently close to f with respet to the uniform convergence metric.

Sketch of proof of Proposition 1. We can define a continuous function $h : X \to \mathbb{R}$ by $h(\mathbf{x}) = \mathbf{d}(\, f(\mathbf{x}), \mathbb{R}^m - U\,)$. In fact, this function is positive valued because f maps X into U, and by the compactness of X it takes a minimum value δ. Therefore, if $\mathbf{x}$ is an arbitrary point in X and $\mathbf{d}(\, f(\mathbf{x}),\, \mathbf{v}\,) < \delta$, then the closed line segment joining $f(\mathbf{x})$ to $\mathbf{v}$ lies entirely in U. Consequently, if g satisfies the condition in the lemma for this choice of δ, the closed line segment joining $f(\mathbf{x})$ to $g(\mathbf{x})$ lies entirely in U for all $\mathbf{x} \in X$. If we define H as in the statement of the proposition, then the the image of H lies entirely in U, while the continuity of H and the identities $H(\mathbf{x}, 0) = f(\mathbf{x})$, $H(\mathbf{x}, 1) = g(\mathbf{x})$ follow immediately from the definitions.∎

A relatively small amount of additional work shows that if K is a compact subset of $\mathbb{R}^m$ for some m, then the set $[K, U]$ of homotopy classes of continuous maps from K to U is countable:

THEOREM 2. *If K is a compact subset of $\mathbb{R}^m$ for some m and U is an open subset of $\mathbb{R}^n$ for some n, then $[K, U]$ is countable.*

This result shows that the concept of homotopy shrinks the uncountably infinite set of continuous mappings from K to U into something countable that might be easier to analyze.

Sketch of proof of Theorem 2. Suppose that $f : K \to U$ as above is continuous, and let $\delta > 0$ be given as in Proposition 1. Denote the coordinate projections of f by f_i, where $1 \leq i \leq n$.

By the Stone-Weierstrass Approximation Theorem, there are polynomial functions p_i on $K \subset \mathbb{R}^m$ such that

$$|\,(p_i|K) - f_i\,| \quad < \quad \frac{\delta}{2\sqrt{n}}$$

for each i, and in fact we can also find polynomials g_i with rational coefficients such that

$$|\,(p_i|K) - (g_i|K)\,| \quad < \quad \frac{\delta}{2\sqrt{n}}\ .$$

If we let $g : \mathbb{R}^m \to \mathbb{R}^n$ be the function whose coordinates are given by the polynomials g_i, it follows that $|\, f - (g|K)\,| < \delta$.

Standard set-theoretic computations show that there are only countably many polynomials in m variables with rational coefficients, and thus there are only countably many choices for g.

Combining the preceding two paragraphs with Proposition 1, we conclude that f is homotopic to one of the countable family of continuous functions whose coordinates are given by polynomials in m variables with rational coefficients, and therefore the set $[K, U]$ is countable.∎

EXAMPLES. The simplest way to see that the set $[K, U]$ can be infinite for some choices of K and U is to let K consist of a single point and take U to be an infinite union of pairwise disjoint open disks in $\mathbb{R}^n$. In this case the homotopy classes are in 1–1 correspondence with the arc components of U, and by construction there are infinitely many of the latter. In fact, one has examples where K and U are both arcwise connected (see the first consequence of Proposition 4 below).

Topic 2. *Topological constructions for spaces with basepoints.* In many situations it is necessary or helpful to work with **pointed sets or spaces** of the form (X, x_0) where $x_0 \in X$; the point x_0 is called the **basepoint** of the (pointed) object. A (basepoint preserving) mapping of pointed objects $f : (X, x_0) \to (Y, y_0)$ is defined to be a morphism $f : X \to Y$ such that $f(x_0) = y_0$, and a basepoint preserving homotopy of basepoint preserving continuous maps is a homotopy $H : X \times [0,1] \to Y$ such that $H(x_0, t) = y_0$ for all $t \in [0,1]$. One can then discuss basepoint preserving homotopy classes of basepoint preserving continuous mappings as in Topic 1. Given two pointed spaces (X, x_0) and (Y, y_0), there is a natural forgetful map

$$[(X, x_0), (Y, y_0)] \longrightarrow [X, Y]$$

obtained by forgetting basepoint considerations; if X and Y are reasonably well-behaved arcwise connected spaces, then this forgetful map is onto, but it is not necessarily 1–1 (see the discussion at the end of this section).

Topic 3. *The fundamental group of a pointed space.* The basic undelying idea is that one can string together (or *concatenate*) two curves if the ending point of the first is the starting point of the second; to simplify the discussion we shall assume that all curves are parametrized by the closed interval $[0,1]$.. If α and β satisfy this condition, we shall use the notation $\alpha + \beta$ to be the curve whose first part is α and whose second part is β. The explicit parametrization for $\alpha + \beta$ is given in Munkres.

Strictly speaking, concatenation is not associative, but it is associative up to an endpoint preserving homotopy (*i.e.*, the homotopy is constant at both end points). Similarly, if C_0 and C_1 are the constant curves whose values are the initial and final values of α, then the curves α, $C_0 + \alpha$ and $\alpha + C_1$ are all endpoint preservingly homotopic to each other. Finally, if $-\alpha$ denotes the curve with the opposite parametrization

$$-\alpha(t) \quad = \quad \alpha(1-t)$$

then $\alpha + (-\alpha)$ is endpoint preservingly homotopic to C_0 while $(-\alpha) + \alpha$ is endpoint preservingly homotopic to C_1.

One justification for the "$\pm$" notation is that if α and β are piecewise smooth curves in an open subset of $\mathbb{R}^n$ then we have the following line integral identities:

$$\int_{\alpha+\beta} \rho(s)\, ds \quad = \quad \int_\alpha \rho(s)\, ds \; + \; \int_\beta \rho(s)\, ds \quad , \qquad \int_{-\alpha} \rho(s)\, ds \quad = \quad -\int_\alpha \rho(s)\, ds$$

WARNING. Although contatenation is associative up to endpoint preserving homotopy, one rarely has the strict associativity condition $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for curves. Furthermore, in general there is not even a homotopy commutativity property for the concatenation construction; in particular, if one of $\alpha + \beta, \beta + \alpha$ is definable then both composites are defined if and only if $\alpha(1) = \beta(0)$ and $\beta(1) = \alpha(0)$, and even if all four of these points are equal the curves need not be endpoint preservingly homotopic to each other; specific examples will be given in the next unit.

SECOND WARNING. Some books and papers define $\alpha+\beta$ so that the first part of the curve is β and the last part is α. Each convention has advantages and disadvantages, but in any case it is good to recognize which convention is used in a particular reference in order to avoid misinterpreting some statements.

If we are dealing with closed curves with $\alpha(0) = \alpha_1$ and fix the starting and ending point as a specific element $x_0 \in X$, then the preceding discussion implies that the set of basepoint preserving homotopy classes of such closed curves is a group which is called the fundamental group (or in some older writings the Poincaré group) of (X, x_0) and denoted by $\pi_1(X, x_0)$.

One can think of the fundamental group as a low resolution algebraic picture of a (pointed) space. This constrution extends to basepoint preserving continuous maps: If $f : (X, x_0) \to (Y, y_0)$ is such a mapping then there is a group homomorphism

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

such that $(g \circ f)_* = g_* \circ f_*$ and the identity map of point spaces induces the identity on fundamental groups, furthermore, if $f, g : (X, x_0) \to (Y, y_0)$ are basepoint preservingly homotopic mappings, then $f_* = g_*$. It follows that if f is a basepoint preserving homotopy equivalence then f_* is an isomorphism. In particular, if $K \subset \mathbb{R}^n$ is a convex set, then $\pi_1(K, k_0)$ is isomorphic to $\pi_1(\{k_0\}, k_0) \cong \{1\}$ for an arbitrary $k_0 \in K$ because the inclusion $\{k_0\} \subset K$ is a basepoint preserving homotopy equivalence (verify this). Another elementary property of the fundamental group is that it behaves well with respect to Cartesian products:

$$\pi_1\big(\, X \times Y, (x_0, y_0)\big) \quad \cong \quad \pi_1(X, x_0) \times \pi_1(Y, y_0)$$

Finally, we note that if (X, x_0) is a pointed space and $A \subset X$ is the arc component of x_0, then the inclusion mapping induces an isomorphism $\pi_1(A, x_0) \to \pi_1(X, x_0)$ because the image of a basepoint preserving closed curve or basepoint preserving homotopy is arcwise connected and hence must be contained in A. In particular, the fundamental group only yields information about the arc component of the basepoint, and therefore the fundamental group cannot be used by itself to obtain information about any other arc component of the space.

Topic 4. *The fundamental group of a circle and some far-reaching generalizations.* If the fundamental groups of every pointed space were trivial, then the concept would not be of much use, so one of the first things to check is that this group is nontrivial for an important example of a topological space, and if $S^1 \subset \mathbb{R}^2 \cong \mathbb{C}$ is the unit circle (centered at the origin) then the first result in this direction is that $\pi_1(S^1, 1) \cong \mathbb{Z}$. The proof of this result uses two fundamental properties of the exponential winding map

$$p : \mathbb{R} \longrightarrow S^1 \,, \qquad p(t) \;=\; \exp(2\pi\, it)$$

called the *Path Lifting Property* and the *Covering Homotopy Property*. It is useful and important to formulate abstract versions of these proofs, and this is done using the concept of a *covering space projection* $p : (E, e_0) \to (B, b_0)$; in practice one often suppresses the mapping p and the basepoints, saying that *E is a covering space of B*.

IMPORTANT DEFAULT HYPOTHESIS. The theory of covering spaces only works well for spaces that are Hausdorff and locally arcwise connected, so *from now on, unless stated otherwise we shall assume that all spaces under consideration have these properties.*

The proof that $\pi_1(S^1, 1) \cong \mathbb{Z}$ extends to covering space projections as follows: If (E, e_0) and (B, b_0) are connected spaces satisfying the Default Hypothesis and $p : (E, e_0) \to (B, b_0)$ is a covering space projection, then $p_* : \pi_1(E, e_0) \to \pi_1(B, b_0)$ is a 1–1 homomorphism.

IMPORTANT REMARK. The preceding discussion shows that there is generally no relation between the injectivity or surjectivity of a continuous map f and the analogous properties for the associated homomorphism f_* of fundamental groups. In particular, the covering space example shows that f_* can be 1–1 but not onto when f is onto but not 1–1, and the example $(S^1, 1) \subset (\mathbb{C}, 1)$ shows that f_* can be onto but not 1–1 when f is 1–1 but not onto.

It turns out that the cosets of $p_* [\pi_1(E, e_0)]$ in $\pi_1(B, b_0)$ also have a topological interpretation which is analogous to the proof that the fundamental group of the circle is $\mathbb{Z}$.

PROPOSITION 3. *Let $p : (E, e_0) \to (B, b_0)$ be a covering space projection, where E and B are connected spaces satisfying the Default Hypothesis, and let $F \subset E$ denote the inverse image $p^{-1}[\{b_0\}]$ (the* fiber *of the basepoint; note that $e_0 \in F$). Then there is a* right action *of the group $\pi_1(B, b_0)$ on F; in other words a set-theoretic mapping $\Phi : F \times \pi_1(B, b_0) \to F$ such that*

$$\Phi(x, 1) \quad = \quad x \ , \quad \Phi(x, gh) \quad = \quad \Phi\left(\Phi(x, g), h\right)$$

for all $x \in F$ and $g, h \in \pi_1(B, b_0)$. Every point in F has the form $\Phi(e_0, g)$ for some g, and $\Phi(e_0, h) = e_0$ if and only if h lies in the image of $\pi_1(E, e_0)$.

Usually we simplify notation and denote $\Phi(x, g)$ by $x \cdot g$ or simply xg. One advantage of this convention is that we can rewrite the defining identities as $x \cdot 1 = x$ and $(xg)h = x(gh)$.

Sketch of proof of Proposition 3. We construct Φ using the Path Lifting and Covering Homotopy Properties. If $x \in F$ and $g \in \pi_1(B, b_0)$, choose a basepoint preserving closed curve $\gamma : S^1 \to B$ representing g, let $\widetilde{\gamma}$ be the unique lifting to a curve $[0, 1] \to E$ such that $\widetilde{\gamma}(0) = x$ given by the Path Lifting Property, and provisionally take xg to be $\widetilde{\gamma}(1)$. In order for this provisional construction to be well-defined, we need to check that we get the same end point for all choices of γ, but this follows immediately by applying the Covering Homotopy Property to the basepoint preserving homotopy relating γ to the other representative. The displayed identities follow immediately (for $x\,1 = x$ the lifting of the constant curve at b_0 is the constant curve at x, while for $(xg)h = x(gh)$ one chooses representatives α and β for g and h, so that $\alpha + \beta$ is a representative for gh and the appropriate lifting of the latter curve ends at $(xg) \cdot h$, which verifies the property in question). Given x in F, there is a curve θ from e_0 to x in E, and if g represents the class of $p \circ \theta$ in the fundamental group, then by construction we have $x = e_0 g$. Finally, if $e_0 g = e_0$ then in the previous notation we have $\widetilde{\gamma}(1) = e_0$ so that $\widetilde{\gamma}$ is a closed curve representing some $g' \in \pi_1(E, e_0)$ such that $p_*(g') = g$; conversely, if g satisfies this condition and we take β to be a closed curve in E_0 representing g', then β is the associated lifting of $p \circ \gamma$, and this implies that $e_0 \cdot p_*(g') = e_0$ as required.■

Free homotopy classes and the role of basepoints. Two natural questions are why we take the trouble to introduce pointed spaces and whether the previously described forgetful map

$$\mathcal{F} : [(X, x_0), (Y, y_0)] \longrightarrow [X, Y]$$

is bijective. Our first result shows that the mapping is an isomorphism in at least some important special cases:

PROPOSITION 4. *For every pointed space (X, x_0) the forgetful map*

$$\mathcal{F} : [(X, x_0), (S^1, 1)] \longrightarrow [X, S^1]$$

is bijective.

Before proving this we shall mention two implications for questions that arose earlier.

First consequence. If we take $(X, x_0) = (S^1, 1)$, then it follows that $[S^1, S^1]$ is infinite and hence S^1 is not contractible.

Second consequence. Since $\mathbb{C} = \mathbb{R}^2$ is contractible it follows that $[S^1, \mathbb{R}^2]$ is a point and the map $[S^1 \S^1] \rightarrow [S^1, \mathbb{R}^2]$ induced by inclusion is not injective. In particular, the homotopy class construction does not send injective maps of spaces to injective maps of homotopy classes.

Sketch of proof of Proposition 4. We first show that every continuous map $X \rightarrow S^1$ is homotopic to a basepoint preserving map, and we do this with the group operation on S^1 given by multiplication of complex numbers. Given $f : X \rightarrow S^1$, we know that there is a continuous curve β in S^1 joining $f(x_0)$ to 1; if we define a homotopy by $H(x,t) = f(x) \cdot \gamma(1-t)^{-1}$, then H is a homotopy from f to a basepoint preserving map. Similarly, suppose that the basepoint preserving maps f and g are homotopic but the homotopy is not necessarily basepoint preserving, let $\alpha(t) = H(x_0, t)$ so that $\alpha(0) - \alpha(1) = 1$, and define a new homotopy by $H'(x,t) = H(x,t) \cdot \alpha(t)^{-1}$; it follows from the construction that H' is a basepoint preserving homotopy from f to g.■

Another issue involving basepoints is the extent to which $\pi_1(X, x_0)$ depends upon the choice of x_0. Clearly if we take a disconnected space such as the disjoint union of $\mathbb{R}$ and S^1, then the fundamental group will be trivial if the basepoint lies in $\mathbb{R}$ and will be nontrivial if the basepoint lies in S^1. Thus the question reduces to arcwise connected spaces or basepoints which lie in the same arc component of a space. In fact, if x_1 and x_0 lie in the same arc component and γ is a curve joining x_0 to x_1, then an isomorphism $\pi_1(X, x_0) \rightarrow \pi_1(X, x_1)$ can be defined so that it takes the class of a closed curve α in the first group to the class of the closed curve $(-\gamma) + (\alpha + \gamma)$ in the second. However, different choices of γ often yield different isomorphisms of these fundamental groups; in particular, if $x_0 = x_1$, then by construction the resulting automorphism is conjugation by the class of γ in $\pi_1(X, x_0)$ — in other words, the map sending u to $g^{-1}ug$, where $g = [\gamma]$ — and this isomorphism is always the identity if and only if the fundamental group is abelian (as shown in Munkres, Section 71, there are well behaved compact subsets of $\mathbb{R}^2$ for which the fundamental group is not abelian, so there are examples where the isomorphism depends upon the choice of path).

Further information on such issues appears on pages 332–333 of Munkres and Exercise 3 on page 335.

More generally, if X is arcwise connected one can ask about the relationship between the group $\pi_1(X, x_0)$ and the set of free homotopy classes $[S^1, X]$; we shall not need this in the course, but it seems worthwhile to describe the relationship as a motivation for introducing basepoints. A discussion of the forgetful map

$$\pi_1(X, x_0) \longleftarrow [S^1, X]$$

and more general questions of the same type appears in Section 4.A of Hatcher (in particular, see page 422, and see pages 341–342 for definitions of some of the concepts discussed there). It turns out that *the forgetful map $\mathcal{F}$ is always onto, and two elements in the fundamental group go to the same class under $\mathcal{F}$ if and only if they lie in the same conjugacy class in the fundamental group.*

It follows that *the role of the basepoint is to create a setting where one can directly apply the methods and results of group theory*, and the discussion would be more awkward (at best) if we tried to do everything without basepoints.

I.1 : Lifting criterion

(**M**, §79; **H**, §§1.1, 1.3)

Reminder. For the rest of this unit and the next, we assume all spaces under discussion are Hausdorff and locally arcwise connected (hence connectedness is equivalent to arcwise connectedness).

If we are given a covering space projection $p : (E, e_0) \to (B, b_0)$ where E and B are (arcwise) connected, and an arbitrary continuous mapping $f : (X, e_0) \to (B, b_0)$, one natural question is whether such a mapping can be lifted to a continuous mapping $F : (X, e_0) \to (B, b_0)$ and whether such a lifting is unique if there is at least one lifting. The existence of such liftings if $X = [0, 1]$ played a crucial role in the material discussed in Section I.0, and one might hope that information on the more general question might also have far-reaching consequences. Simple examples show that one cannot expect to have such liftings all the time; for example if we take the usual covering space projection from $\mathbb{R}$ to S^1, then no such lifting exists. More generally, if F exists then on the fundamental group level we have

$$f_* \;=\; (p \circ F)_* \;=\; p_* \circ F_*$$

and this implies that the image of f_* lies in the image of p_* if a lifting exists; this condition clearly fails for $\mathbb{R} \to S^1$. On the other hand, the main result about liftings is that they exist if the displayed condition holds:

THEOREM 1. *Suppose that we are given a covering space projection $p : (E, e_0) \to (B, b_0)$, where E and B are (arcwise) connected, and let $f : (X, e_0) \to (B, b_0)$ be a continuous mapping. Then there is a continuous lifting $F : (X, e_0) \to (B, b_0)$ such that $f = p \circ F$ if and only if the image of $f_* : \pi_1(X, e_0) \to \pi_1(B, b_0)$ is contained in the image of $p_* : \pi_1(E, e_0) \to \pi_1(B, b_0)$, and if such a lifting exists then it is unique.*

Proof. The idea of the construction is simple: Given a point $x \in X$, we take a curve γ in X joining x_0 to x, we lift $f \circ \gamma$ uniquely to some curve β in E starting at e_0, and we try to define $F(x)$ as $\beta(1)$. There are several things that need to be checked in order to conclude that this actually works. First, we must verify that the construction does not depend upon the choice of γ; we shall need the assumption on maps of fundamental groups in order to complete this step. Next, we need to verify that this construction is continuous. If we can complete these steps, then the construction yields the lifting identity $p \circ F$ immediately, and uniqueness follows by the same considerations proving uniqueness for the Path Lifting and Covering Homotopy Properties.

Suppose that α and β are continuous curves in X which join x_0 to x_1, and let γ be the closed curve $\beta + (-\alpha)$. Then $\gamma + \alpha$ is basepoint preservingly homotopic to β, and as in the proof of the Covering Homotopy Property it follows that the final points of the liftings of

$$f \circ (\gamma + \alpha) \;=\; f \circ \gamma + f \circ \alpha$$

and $f \circ \beta$ which start at e_0 must also have the same endpoints. Thus the mapping will be well-defined if we can show that the corresponding liftings of $f \circ \alpha$ and $f \circ \gamma + f \circ \alpha$ also have the same endpoints.

This is the point at which we use the hypotheses on images of fundamental groups, which implies that $f\circ\gamma$ is basepoint preservingly homotopic to $p\circ\theta$ for some closed curve θ in E; by uniqueness of path liftings we know that θ is the unique lifting of $p\circ\theta$ starting at e_0. Therefore it follows that the liftings of $p\circ\theta + f\circ\alpha$ and $f\circ\gamma + f\circ\alpha$ have the same endpoints. If φ denotes the unique lifting of $f\circ\alpha$ with starting point e_0, then the lifting of the first curve $p\circ\theta + f\circ\alpha$ is equal to $\theta + \varphi$ because θ is a closed curve; this implies that the endpoint of this lifting is the endpoint of φ. But the latter is just the lifting of $f\circ\alpha$, and if we combine this with the previous observations we see that the corresponding liftings of $f\circ\alpha$ and $f\circ\gamma + f\circ\alpha$ do have the same endpoints, which is what we needed to prove in order to show that F is well defined.

To prove continuity, let $x \in X$, let V be an evenly covered open neighborhood of $f(x)$ in B (*i.e.*, the inverse image of V is a disjoint union of homeomorphic copies of V), and let U be an arcwise connected open neighborhood of x such that $f(x) \in V$. Choose the sheet $V_0 \subset E$ mapping homeomorphically to V such that $F(x) \in V_0$, let γ be a curve joining x_0 to x, and let φ be the unique lifting of $f\circ\gamma$ to E starting at e_0; by construction the endpoint of φ is equal to $F(x)$. — Now let $y \in U$; the continuity of F will follow if we can prove that $F(y) \in V_0$. Let δ be a short curve in U starting at x and ending at y, and let ψ be the unique lifting of $f\circ\psi$ starting at $F(x)$. By connectedness the image of $f\circ\psi$ is contained in V_0; by construction $F(y)$ is the endpoint of the unique lifting of

$$f\circ(\gamma + \psi) \quad = \quad (f\circ\gamma) + (f\circ\psi)$$

which is the endpoint of ψ. But we have seen that this point lies in V_0, and therefore F is continuous. The uniqueness of F now follows from the same sort of argument employed for the Path Lifting Property.■

This result has numerous consequences, but for the time being we shall only mention one of them. Given two connected covering space projections $p : X \to W$ and $q : Y \to W$, we shall say that they are *equivalent* if there is a homeomorphism $h : X \to Y$ such that $q\circ h = p$ (hence also $p\circ h^{-1} = q$). It is elementary to check that equivalence of covering space projections determines an equivalence relation on coverings.

COROLLARY 2. *If $p : (X, x_0) \to (W, w_0)$ and $q : (Y, y_0) \to (W, w_0)$ are covering space projections with all spaces connected, then they are equivalent if and only if the images of $\pi_1(X, x_0)$ and $\pi_1(Y, y_0)$ in $\pi_1(W, w_0)$ are equal.*

Example. If W is not locally arcwise connected, then the uniqueness conclusion in Corollary 2 does not necessarily hold. Explicit examples are described in the document `polishcircle.pdf` (in the course directory).

Proof of Corollary 2. The lifting criterion implies that there are continuous basepoint preserving mappings h and k such that $q\circ h = p$ and $p\circ k = q$. These equations imply the identities $q\circ(h\circ k) = q$ and $p\circ(k\circ h) = p$. By uniqueness of liftings it follows that $h\circ k$ is the $k\circ h$ is the identity on Y, so that h and k are homeomorphisms which satisfy the conditions for equivalence of coverings.■

The preceding result is a major step in the classification of connected covering spaces, but it is incomplete in two respects. First, we would like to know if every subgroup of $\pi_1(X, x_0)$ can be realized as the fundamental group of some connected covering space. Second, we would like to know if there is something special about covering spaces whose fundamental groups map bijectively to normal subgroups of $\pi_1(X, x_0)$. We shall answer these questions in the next two sections.

Recognizing simply connected coverings

The preceding discussions also yield a criterion for recognizing when a connected covering space is simply connected; this will be useful in Section I.3.

PROPOSITION 3. *Let $p : X \to Y$ be a covering space projection where X and Y are connected, and assume that for each x in the fiber $p^{-1}[\{x_0\}]$ there is a unique $g \in \pi_1(X, x_0)$ such that $x = x_0 \cdot g$. Then X is simply connected.*

Proof. If h is in the image of $p_*; \pi_1(X, x_0) \to \pi_1(Y, y_0)$ then $x_0 \cdot h = x_0$, so the only way one can have uniquess is if this image is trivial. Since p_* is injective, this implies that the fundamental group of X must be trivial.■

Change of basepoints in covering spaces

If $p : (X, x_0) \to (Y, y_0)$ is a based covering space projection where X and Y are (arcwise) connected, then for each $g \in \pi_1(X, x_0)$ the map $p^g : (X, x_0 \cdot g) \to (Y, y_0)$, defined by the formula $p^g(x) = p(x)$, is also a based covering space projection. By the results of Section 0 we know that there is an isomorphism $\Phi(\gamma) : \pi_1(X, x_0) \to \pi_1(X, x_0 \cdot g)$ given by choosing a curve γ joining x_0 to x_1 and defining $\Phi(\gamma)$ as the map sending the class of a closed curve α in the first group to the class of the closed curve $(-\gamma) + (\alpha + \gamma)$ in the second. Therefore the images of p_* and p_*^g are isomorphic subgroups of $\pi_1(X, x_0)$. In fact, we can say considerably more;

THEOREM 4. *Suppose we are in the setting described above, and suppose that the image of p_* is the subgroup $H \subset \pi_1(X, x_0)$. Then the image of p_*^g is the subgroup $g^{-1} H \, g$.*

Proof. We need to compare the images of p_* and $\Phi(\gamma) \circ p_*^g$ in $\pi_1(Y, y_0)$. By the discussion preceding the statement of the theorem, the latter composite sends the class of the curve α to the class of the composite

$$p \circ \big((-\gamma) + (\alpha + \gamma 0 \big) \quad = \quad -(p \circ \gamma) \; + \; (p \circ \alpha + p \circ \gamma) \; .$$

Now the right hand side represents the class

$$[-(p \circ \gamma)] \cdot p_*([\alpha]) \cdot [p \circ \gamma] \quad = \quad g^{-1} \, p_*([\alpha]) \, g$$

in $\pi_1(Y, y_0)$, and since p_* corresponds to the inclusion $H \subset \pi_1(Y, y_0)$, it follows that the image of p_*^g is equal to $g^{-1} \, H \, g$.■

I.2 : Covering space transformations

(**M**, §§79, 81; **H**, §1.3)

One elementary property of the covering space projection

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

is that it is highly symmetric. If n is an arbitrary integer, then the translation map $\sigma_n(t) = t + n$ is a self-equivalence of the covering space which satisfies the condition $p \circ \sigma_n = p$. Since $\sigma_n \circ \sigma_m = \sigma_{m+n}$, these self-equivalences generate an infinite cyclic group. These examples can be generalized as follows:

Definition. Let $p : X \to Y$ be a covering space projection where X and Y are connected. A *covering space transformation* or *deck transformation* of $p : X \to Y$ is a homeomorphism $\sigma : X \to X$ such that $p \circ \sigma = p$. Observe that there is no condition regarding basepoints; the reason for this is implicit in the first paragraph of the proof of Theorem 1 below.

It follows immediately that the set of all covering space transformations is a group with respect to composition, and the structure of this group is given as follows:

THEOREM 1. *Let $p : X \to Y$ be a covering space projection where X and Y are connected, and let $H \subset G$ be the subgroup inclusion corresponding to the map of fundamental groups $p_*; \pi_1(x, x_0) \to \pi_1(Y, y_0)$. Then the group of covering transformations of $p : X \to Y$ is anti-isomorphic to $N(H)/H$, where $N(H)$ denotes the normalizer of H in G.*

A 1–1 correspondence of groups $h : \Gamma \to \Gamma'$ is said to be an **anti-isomorphism** if $h(g_1g_2) = h(g_2)h(g_1)$ for all $g_1, g_2 \in \Gamma$. Observe that h is an anti-isomorphism then so is h^{-1}, and the inverse map from any group to itself defines an anti-isomorphism. If $G^{\mathbf{OP}}$ is G with the order of multiplication reversed, then it is an elementary exercises to verify that $G^{\mathbf{OP}}$ is a group which is anti-isomorphic to G, and consequently an anti-isomorphism from Γ to G is equivalent to an isomorphism from Γ to $G^{\mathbf{OP}}$.

Example. If G is the group of invertible $n \times n$ matrices over some field (where $n > 1$ is an integer), then the transposition map defines an anti-isomorphism from G to itself.

Proof of Theorem 1. The first thing to note is that a covering space transformation σ is uniquely determined by its value at the basepoint x_0 because σ is a lifting of p. By results from Section I.0 we know that $\sigma(x_0) = x_0 \cdot g$ for some $g \in G$; in general g is not unique, for if $h \in H$ then $x_0 \cdot h = x_0$ and hence $x_1 = x_0 \cdot (hg)$.

The second step is to show that there is a covering space transformation σ such that $\sigma(x_0) = x_0 \cdot g$ if and only if g lies in the normalizer of H in G. By definition, the existence of such a mapping is equivalent to the existence of a lifting of $p : (X, x_0) \to (Y, y_0)$ to $p^g : (X, x_0 \cdot g) \to (Y, y_0)$, and by the criterion in the preceding section such a lifting exists if and only if the image of the induced fundamental group homomorphism p_* is contained in the image of the corresponding homeomorphism p^g. If we denote the image of p_* by H as in Theorem 1.4, then that result implies that the image of p^g_* is equal to $g^{-1}\, H\, g$, so the desired mapping σ exists if and only if $g^{-1}\, H\, g \subset H$. Since this holds if and only if g lies in the normalizer of H, this completes the proof of the second step.

Finally, we need to prove that the map σ as constructed above is a homeomorphism and that the construction defines a group which is anti-isomorphic to $N(H)/H$. The first step in doing so is to prove the following relationship:

$$\sigma(x_0) = x_0 g_1 \textit{ and } \tau(x_0) = x_0 g_2 \textit{ imply } \tau \circ \sigma(x_0) = x_0 g_1 g_2$$

In the notation of this display, we need to prove that $\tau(x_0 g_1) = x_0 g_1 g_2$. — Let α and β be curves joining x_0 to $x_0 g_1$ and $x_0 g_2$ respectively. Then $\alpha + (\sigma \circ \beta)$ is a lifting of $(p \circ \alpha) + (p \circ \beta)$ starting at x_0, and hence its endpoint is $x_0 g_1 g_2$. On the other hand, $\tau(x_0 g_1)$ is defined by taking the endpoint for unique lifting of $p \circ \beta$ with starting point $x_0 g_1$. Since this unique lifting is $\sigma \circ \beta$ it follows that the $\tau(x_0 g_1)$ must be equal to $x_0 g_1 g_2$, and this proves the identity we wanted.

Consider the map e which sends a covering transformation to the unique element $[g]$ of $N(H)/H$ such that

$$\sigma(x_0) \quad = \quad x_0 \cdot e(\sigma) \ .$$

We have shown that the map e from covering transformations to $N(H)/H$ is 1–1 and onto, and the results of the previous paragraph imply that $e(\tau \circ \sigma) = e(\sigma) \cdot e(\tau)$. This immediately implies that e is an anti-isomorphism, and it also shows that if $e(\sigma) = g$ and $e(\tau) = g^{-1}$ then

$$e(\sigma) \circ e(\tau) \quad = \quad \text{Identity} \quad = \quad e(\tau) \circ e(\sigma)$$

so that $e(\tau)$ is a continuous inverse to $e(\sigma)$.■

The following consequence of Theorem 1 is particularly noteworthy.

COROLLARY 2. *Let $p : X \to Y$ be a covering space projection where X and Y are connected, and let $H \subset G$ be the subgroup inclusion corresponding to the map of fundamental groups $p_*; \pi_1(X, x_0) \to \pi_1(Y, y_0)$. Then the following are equivalent:*

(i) H is a normal subgroup of G.

(ii) For each x in the fiber $p^{-1}[\{x_0\}]$ there is some (unique) covering transformation σ such that $x = \sigma(x_0)$.

We shall say that the covering is a **regular** covering if either (hence both) of these conditions holds.

Proof. By the theorem we know that x is $\sigma(x_0)$ for some x_0 if and only if $x = x_0 g$ for some g in the normalizer of H.

Suppose that (ii) holds. Then for each x in the fiber $p^{-1}[\{x_0\}]$ there is some $g \in G$ such that $x_0 g = x$ and a covering transformation σ such that $\sigma(x_0) = x_0 g$, so that $e(\sigma) = [g]$ in the terminology developed in the proof of Theorem 1. Since we know that the latter implies that g lies in the normalizer of H, it follows that G must be the normalizer of H, so that H is normal in G.

Conversely, if (i) holds and we take x as in (ii) so that $x = x_0 g$, then by Theorem 1 we know there is a covering transformation σ such that $e(\sigma) = [g]$, and by the definition of the latter we must have $\sigma(x_0) = x$.■

I.3 : Universal coverings and applications

(**M**, §§80, 82; **H**, §1.3)

In Section 1 we showed that, up to equivalence, there is at most one based covering space projection $p : (X, x_0) \to (Y, y_0)$ such that X is connected and the image of p_* is a given subgroup $H \subset \pi_1(Y, y_0)$. The purpose of this section is to show that, under a reasonable condition on Y, every subgroup $H \subset \pi_1(Y, y_0)$ can be realized in this manner. In principle, everything reduces to constructing a covering when H is the trivial group, so we begin by showing that the latter can always be realized.

Simply connected covering spaces

A connected, Hausdorff, locally arcwise connected topological space does not necessarily have a simply connected covering space, and in fact there are examples which are compact subsets of $\mathbb{R}^2$; details appear on pages 486–487 of Munkres. However, the following result shows that simply connected covering spaces exist for a reasonably broad family of spaces (which contains most objects of primary interest in algebraic and geometric topology):

THEOREM 1. *Let Y be a (Hausdorff, locally arcwise connected and) connected space, and assume that Y is* locally simply connected *in the sense that every point of Y has a neighborhood base of simply connected open neighborhoods. Then there is a simply connected covering space projection $p : X \to Y$.*

By the Lifting Criterion of Section 1, if $p : X \to Y$ is a covering space projection where X is simply connected, then for an arbitrary connected covering space projection $q : W \to Y$ there is a mapping $q' : X \to W$ such that $q \circ q' = p$, and hence we can think of X as sitting over every other covering space W. This is one reason why $p : X \to Y$ is often called a *universal covering space* if X is simply connected.

Sketch of proof. The proof is long and somewhat messy in places, and as with the real number system the basic formal properties of universal covering spaces turn out to be more important in everyday mathematical work than the explicit method of construction. Therefore we shall simply comment upon a few features of the proof and give a reference to Munkres, pp. 495–498, for details.

The first point we would like to make is that the formal description of points in the universal covering space is more or less forced upon us. If there is a simply connected covering X of the connected space Y, then each point $x \in X$ can be joined to the basepoint $x_0 \in X$ by a curve γ; furthermore, if $p \circ \gamma$ denotes the projection of γ onto Y, then the point x only depends upon the endpoint preserving homotopy class of $p \circ \gamma$. If U is a simply connected evenly covered neighborhood of $p(x)$ and V is a neighborhood of x which projects homeomorphically to U, then the points of V are endpoints of curves having the form $\gamma + \beta'$, where $\beta = p \circ \beta'$ is a curve in U starting at $p(x)$ and β' is a lifting to a curve in V starting at x. These considerations also show that the topology on X is also more or less forced upon us.

One important feature of the universal covering space construction is that the space X comes equipped with an action of $\pi_1(Y, y_0)$ by covering space transformations. The simple connectivity of X follows from this fact and Proposition 1.3.■

Recognizing Hausdorff quotients

Although quotient topologies play an important role in algebraic and geometric topology, there are numerous instances in which one must be careful about jumping to conclusions when using them. In particular, the quotient of a Hausdorff space need not be Hausdorff (for a simple example, take the unit interval with the eqivalence relation whose equivalence classes are $\{0\}$ and $(0, 1]$). Therefore we need some criteria for showing that certain quotient spaces we shall construct from Hausdorff will indeed be Hausdorff, and accordingly we shall derive some simple but useful criteria before proceeding with our discussion of covering spaces.

The quotient spaces of interest to us all have the property that the quotient map $X \to X/\mathcal{R}$ is open; basic results on quotient topologies show that a continuous open surjection $X \to Y$ is always a quotient map (see Munkres, Theorem 22.1, p.140).

PROPOSITION 2. *Suppose that $f : X \to Y$ is continuous, open and surjective, and assume further that the set*

$$\{(x_1, x_2 \in X \times X \mid f(x_1) \;=\; f(x_2)\}$$

is closed in $X \times X$. Then Y is Hausdorff.

Proof. The definition of the product topology implies that $f \times f$ is an open mapping. By assumption the set $W \subset X \times X$ of all (x_1, x_2) such that $f(x_1) \neq f(x_2)$ is open, and hence its image is open in $Y \times Y$. But this image is the complement of the diagonal $\Delta_Y \subset Y \times Y$ (first and second coordinates equal), and hence Δ_Y is closed. The latter condition is equivalent to the Hausdorff property for a topological space, and therefore Y must be Hausdorff.■

This is also a good point to include another fact that we shall need:

PROPOSITION 3. *Suppose that $f : X \to Y$ is continuous, open and surjective. If X is locally arcwise connected, then so is Y.*

The proof of this is straightforward and left to the reader as an exercise.■

Finally, we shall give an important way of constructing continuous open surjections using group actions and their orbit spaces (see Munkres, Exercise 8, p. 199 for definitions and a few basic results):

PROPOSITION 4. *Let X be a topological space, let G be a topological group with the discrete topology, and let $\alpha : G \times X \to X$ be a continuous action of G on X. Then the orbit space quotient projection $p : X \to X/G$ is open.*

Note that the orbit space projection is continuous and onto by construction.

Proof. Let $V \subset X$ be open, and let $V' \subset X/G$ be its image under p. By definition of the quotient topology, V' is open in Y if and only if its inverse image is open in X. But the inverse image of V' is equal to

$$\bigcup_{g \in G} g \cdot V$$

and the latter is open because each subset $g \cdot V$ is open, so that $V' = p[V]$ is in fact an open subset of Y.■

Covering spaces realizing subgroups of π_1

We can now use universal coverings to construct covering spaces associated to arbitrary fundamental groups.

THEOREM 4. *Let Y be a (Hausdorff, locally arcwise connected and) connected space, assume that Y is locally simply connected, and let H be a subgroup of $\pi_1(Y, y_0)$. Then there is a connected covering space projection $p : (X, x_0) \to (Y, y_0)$ such that the image of p_* is equal to H.*

Sketch of proof. The hypotheses imply the existence of a simply connected covering space projection $q : (W, w_0) \to (X, x_0)$. By the results of Section 2 the group of covering space transformations for q is anti-isomorphic to $\pi_1(Y, y_0)$. Let X be the quotient of W by the equivalence relation $w \sim w'$ if and only if there is some covering space transformation T such that $T(w_0) = w_0 \cdot h$ for some $h \in H$ and $T(w) = w'$. It is then a straightforward exercise to verify that the space X is a connected covering space of Y and that X is Hausdorff and locally arcwise connected (proofs of the latter use Propositions 2–4). Furthermore, it also follows that $w_0 \cdot g = w_0$ if and only if $g \in H$. Since the subgroup of all g with this property is $\pi_1(x, x_0)$, the results of Section 0 imply that H is isomorphic to $\pi_1(X, x_0)$.■

In most books the preceding result is stated and proved with the slightly weaker hypothesis that Y is *semilocally simply connected*, which means that every point has at least one simply connected open neighborhood. We have stated the result with the stronger condition because most of the spaces encountered in algebraic and geometric topology are locally simply connected; in fact, most have neighborhood bases of contractible open subsets.

We can summarize the results on classifying covering spaces as follows:

THEOREM 5. *Let Y be a locally simply connected space which is connected, and suppose that $y_0 \in Y$. Then there is a $1-1$ correspondence between equivalence classes of based connected covering space projections $p : (X, x_0) \to (Y, y_0)$ and subgroups of $\pi_1(Y, y_0)$. Under this correspondence a covering space is regular if and only if the associated subgroup is normal in $\pi_1(Y, y_0)$.*■

FINAL REMARKS. These will not be used elsewhere in the course, but they shed some light on the material presented thus far.

Example. If U is a connected open subset of $\mathbb{R}^2$ then U is locally simply connected, so that U has a simply connected universal covering space. Fundamental results in complex variable theory imply that this universal covering is homeomorphic to $\mathbb{R}^2$. Analogous results are false in higher dimensions; in particular, if $n \geq 3$ then $\mathbb{R}^n - \{\mathbf{0}\}$ is simply connected but we shall see that it is not homeomorphic to $\mathbb{R}^n$.

Problem. If U is connected and open in $\mathbb{R}^n$ where $n \geq 3$, is its universal covering homeomorphic to an open subset of $\mathbb{R}^n$? — I suspect the answer to this question is usually no, but there do not seem to be any simple examples.

II. Computing Fundamental Groups

In order to use fundamental groups effectively for studying geometrical questions, it is necessary to have some methods for describing them concretely, preferably in a group-theoretic manner. The purpose of this unit is to develop some basic techniques for doing so. One way of analyzing fundamental groups involves finding regular covering spaces whose groups of covering space transformations are anti-isomorphic to some fixed group (see Section 1), and another way is given by decomposing a space into smaller pieces with known fundamental groups and developing group-theoretic methods for recovering the fundamental group of the whole space (see Section 3); the latter requires algebraic concepts that usually do not appear in entry level graduate courses on group theory, and this material corresponds to Section 2.

II.1 : Orbit spaces

(**H**, §§1.3, 1.B)

The methods of this section will realize all finite cyclic groups as fundamental groups of reasonable spaces. By the structure theorem for finitely generated abelian groups, it will follow that every finitely generated abelian group is the fundamental group of some reasonable space. We shall also prove that at least some nonabelian finite groups can arise as fundamental groups. In fact, **every** group can be realized as the fundamental group of some space, but proving this requires more machinery that we shall develop in this course; one method of constructing examples appears in Section 1.B of Hatcher.

Generalities

We shall begin with an abstract formulation of some general properties of covering space transformations on a regular covering space:

Definition. We shall say that an action of a topological group G on a space X is a **free action** (or G acts freely) if for every $x \in X$ the only solution to the equation $g \cdot x = x$ is the trivial solutions for which $g = 1$.

Our first result essentially implies that free actions are the same as covering space transformations if G is finite (we shall not try to make this more precise because it will not be needed subsequently).

THEOREM 1. *Let G be a finite group which acts freely on the Hausdorff topological space X, and let $\pi : X \to X/G$ denote the orbit space projection. Then π is a covering space projection.*

Proof. Let $x \in X$ be arbitrary, and let $g \neq 1$ in G. Then there are open neighborhoods $U_0(g)$ of x and $V_0(g)$ of $g\,x$ that are disjoint. If we let $W(g) = U(g) \cap g^{-1}\,V(g)$ is another open set containing x, while $g\,W(g)$ is an open set containing $g\,x$, and we have $W(g) \cap g\,W(g) = \emptyset$. Let

$$W \quad = \quad \bigcap_{h \neq 1} \; W(h)$$

so that W is an open set containing x.

We claim that if $g_1 \neq g_2$, then $(g_1 W) \cap (g_2 W) = \emptyset$. If we know this, then it will follow immediately that $\pi[W]$ is an open set in X/G whose inverse image is the open subset of X given by $\cup_g \, g\,W$. This and the definition of the quotient topology imply that $\pi[W]$ is an evenly covered open neighborhood of x, and therefore it will follow that π is a covering space projection.

Thus it remains to prove the statement in the first sentence of the preceding paragraph. Note first that it will suffice to prove this in the special case where $g_1 = 1$; assuming we know this, in the general case we then have

$$g_1 W \cap g_2 W \quad = \quad g_1 \left(W \cap (g_1^{-1} g_2)\, W \right)$$

and the coefficient of g_1 on the right hand side is empty by the special case when $g_1 = 1$ and the fact that $g_1 \neq g_2$ implies $1 \neq g_1^{-1} g_2$. — But if $g \neq 1$ then we have $W \cap g\,W \subset W(g) \cap g\,W(g)$, and we know that the latter is empty by construction. Therefore $W \cap g\,W = \emptyset$, and as noted before this completes the proof.∎

COROLLARY 2. *In the setting of the theorem, if X is simply connected then the fundamental group of X/G is anti-isomorphic to G.*

Sketch of proof. It is straightforward to verify that the group of covering transformations of $X \to X/G$ is given by G.∎

Some important examples

The results of the previous course show that S^n is simply connected if $n \geq 2$, and if we combine this with Theorem 1 and Corollary 2 we see that *if G is a finite group which acts freely on S^n, then the quotient space S^n/G has a fundamental group which is anti-isomorphic to G. In particular, if G is abelian then the fundamental group is isomorphic to G.*

The last statement is true because the anti-isomorphisms and isomorphisms are the same for abelian groups.

Example 1. The *real projective plane.* — This space is denoted by $\mathbb{RP}^2$, and two equivalent constructions of it as a quotient space are described in Unit V and the accompanying exercises for the online 205A notes cited earlier. For our purposes here, it is convenient to think of $\mathbb{RP}^2$ as the quotient of S^2 by the equivalence relation which identifies $\mathbf{x}$ and $\mathbf{y}$ if and only if one of these unit vectors is ± 1 times the other. This corresponds to viewing $\mathbb{RP}^2$ as the quotient space of S^2 by the action of $\mathbb{Z}_2$ on S^2 which sends (n, x) to $(-1)^n x$, and therefore Theorem 1 implies that the orbit space projection $S^2 \to \mathbb{RP}^2$ is a 2-sheeted covering space. Since S^2 is simply connected, it follows that the fundamental group of $\mathbb{RP}^2$ must be isomorphic to $\mathbb{Z}_2$.

Of course, there are also similar examples for which S^2 is replaced by S^n for an arbitrary integer $n \geq 2$, and in this case the quotient space $S^n/\{\pm 1\}$ is called *real projective n-space* and denoted by $\mathbb{RP}^n$.

Example 2. *Simple lens spaces.* — A variation of the preceding construction yields spaces whose fundamental groups are cyclic of arbitrary order. Let $\mathbb{D}$ denote either the complex numbers or the quaternions, let d be the dimension of $\mathbb{D}$ as a real vector space, and let G be a finite subgroup of the group S^{dm-1} of elements of $\mathbb{D}$ with unit length. For example, if $\mathbb{D} = \mathbb{C}$ (the complex numbers), then G can be a cyclic group of arbitrary order, while if $\mathbb{D}$ is the quaternions then one also has some nonabelian possibilities, most notably the quaternion group of order 8 whose elements are given by ± 1, $\pm \mathbf{i}$, $\pm \mathbf{j}$, and $\pm \mathbf{k}$. If $\mathbb{D} = \mathbb{C}$ and $m > 1$, then the quotient spaces $S^{2m-1}/\mathbb{Z}_q$ (for $q > 1$) are the objects known as (simple) **lens spaces** (sometimes the case $q = 2$ is excluded because that quotient is the previously described real projective space); the reason for assuming $m > 1$ is that the corresponding quotient space for S^1 is homeomorphic to S^1. By Theorem 1 we know that the fundamental group of a simple lens space $S^{2m-1}/\mathbb{Z}_q$ is isomorphic to q; since $q > 1$ is arbitrary, it follows that every cyclic group of finite order is isomorphic to the fundamental group of some lens space.

If $\mathbb{D}$ is the quaternions, G is the nonabelian quaternion group of order 8 described above and $m = 1$, then the space S^3/G is called the 3-*dimensional quaternionic space form* associated to the group G. The fundamental group of this quotient is then anti-isomorphis to the (nonabelian) quaternion group.

Example 3. The *Klein bottle.* — Define an action of the finite group $\mathbf{Z}_2$ on the torus $T^2 = S^1 \times S^1$ so that the nontrivial element $T \in \mathbf{Z}_2$ satisfies $T \cdot (z, w) = (-z, \overline{w})$ where S^1 is viewed as the set of unit complex numbers and the bar denotes conjugation. This is a free action because $T(z, w) = (z, w)$ would imply $z = -z$, and we know this is impossible over the complex numbers. In this case the quotient space is the **Klein bottle**.

One of the homework exercises outlines a proof that the fundamental group of the Klein bottle is an infinite nonabelian group which contains an index 2 (normal) subgroup isomorphic to $\mathbb{Z} \times \mathbb{Z}$ but has no elements of finite order except the identity.

II.2 : Amalgamation constructions for groups

(**M**, §§67–69; **H**, §1.2)

This material was covered in the previous course during the Fall 2011 Quarter.

II.3 : The Seifert-van Kampen Theorem

(**M**, §70; **H**, §1.2)

This material was covered in the previous course during the Fall 2011 Quarter.

The course directory file `svkproof.pdf` contains an alternate proof of the hard part of the Seifert-van Kampen Theorem for spaces that are locally simply connected. In contrast to the proof in Munkres, the alternate proof provides a semi-explicit construction of the universal covering for a space $X = U \cup V$ (where U, V and $U \cap V$ are all arcwise connected, locally simply connected, and open subsets of X).

II.4 : Examples

(**M**, §§59, 71, 72; **H**, §1.2)

This material was covered in the previous course during the Fall 2011 Quarter.

III. Graph complexes

In this unit we shall study the fundamental groups for a special class of spaces which are built out of very simple pieces but turn out to be important in many branches of mathematics, and in some sense are "toy models" for the sorts of objects usually studied in algebraic and geometric topology. More precisely, these spaces (called finite graph complexes, edge-path graphs or more simply just *graphs*) are excellent test cases for applying the methods and results of this course.

Informally, a graph can be constructed by taking a finite collection of closed intervals and identifying their endpoints in a suitable fashion; following geometric intuition, the images of the intervals are called *edges* and the images of their endpoints are called *vertices*. Note that these are NOT graphs as defined and studied in coordinate geometry and calculus, but the name has stuck and become standard usage, both in mathematics and in its applications to numerous other subjects such as computer science, physics, chemistry, industrial engineering, the biological sciences and even to other areas of knowledge where it is useful to look at chains of relationships or passage from one state of a system to another. A fairly simple application of graph theory to a problem about relationships is given in the following online video:

`http://www.youtube.com/watch?v=b3lbjoiEAyA`

One of the main results in this unit is a complete description of the fundamental group of a connected finite graph using the numbers of edges and vertices. This result in turn leads to an algebraic theorem about free groups that is somewhat nonintuitive: If F is a free group on a finite number $n > 1$ of generators, then for each $m > n$ the group F contains a subgroup of finite index with m generators; in contrast, if A is a free abelian group on n generators, then every subgroup of A has at most n generators.

III.1 : Basic definitions

(**M**, §83; **H**, §1.A, Ch. 2 Introduction)

Since we have already described finite graphs intuitively, we shall proceed to the formal description.

Definition. A **finite edge-path graph complex** (more simply a *finite graph*) is a pair $(X, \mathcal{E})$ consisting of a compact Hausdorff space X and a finite family $\mathcal{E}$ of closed subsets with the following properties:

(1) Each subset $E \in \mathcal{E}$ is homeomorphic to the closed interval $[0, 1]$.

(2) The space X is the union of all the subspaces E in the family $\mathcal{E}$.

(3) If E_1 and E_2 are distinct subsets of $\mathcal{E}$, then either $E_1 \cap E_2$ is empty or else it is a single point corresponding to a vertex of each interval E_i.

COMMENTS ON THE DEFINITION. The endpoints of a set homeomorphic to $[0, 1]$ are topologically characterized by the fact that their complements are connected; for all other points, the complement has two components. As above, we shall say that a subset of $\mathcal{E}$ is an edge and an endpoint of an *edge* will be called a *vertex*.

The setting in Chapter 14 of Munkres is more general and includes examples where the set of edges is infinite but each vertex lies on only finitely many edges. We are restricting attention to examples with finitely many edges in order to simplify the discussion.

Examples. It is easy to draw many examples of graphs, and such drawings are extremely useful for understanding this concept. The file `graphpix1.pdf` contains a few examples, including some that will appear later in this course.

An alternate definition

Our definition of a graph assumes that two edges meet in just one endpoint, but in some situations it is convenient to consider examples for which the intersection of two edges is also allowed to be both vertices of the two edges as in the following illustration:

D

(Two vertices at the corners, two edges have these endpoints.)

We shall prove that every object of this more general type can be expressed as a graph in the sense of our definition.

LEMMA 1. *Let Γ be a system satisfying the conditions for an finite edge-vertex graph except that two edges may have both of their vertices, and let $\mathcal{E}$ be the collection of edges for this system. Then there is another family of closed subsets $\mathcal{E}'$ such that the following hold:*

(i) *The family $\mathcal{E}'$ is a collection of edges for a graph structure on Γ.*

(ii) *Each element of $\mathcal{E}'$ is contained in a unique element of $\mathcal{E}$ such that one endpoint of $\mathcal{E}'$ is also an endpoint for $\mathcal{E}$ but another is not, and each edge in $\mathcal{E}$ is a union of two edges in $\mathcal{E}'$.*

(iii) *The intersection of two distinct edges in $\mathcal{E}'$ is **at most** a common vertex.*

Proof. For each edge $E \in \mathcal{E}$, pick a point $b_E \in E$ that is not an endpoint. It follows that $E - \{b_E\}$ has two connected components, each of which contains exactly one endpoint of E. If x is an endpoint of E define the set $[x, E]$ to be the closure of the component of $E - \{b_E\}$ which contains x. If $\mathcal{E}'$ denotes the set of all such subsets $[x, E]$, then it follows immediately that $\mathcal{E}'$ has the properties stated in the lemma. Note that by construction the endpoints of a given edge $[x, E]$ are x and b_E.■

The family $\mathcal{E}'$ is frequently called the *derived* graph structure associated to $\mathcal{E}$.

As noted in one of the exercises, many examples of edge-vertex graphs are suggested by ordinary letters and numerals.

Subgraphs

Definition. Let $(X, \mathcal{E})$ be a finite edge-path graph. A *subgraph* $(X_0, \mathcal{E}_0)$ is given by a subfamily $\mathcal{E}_0 \subset \mathcal{E}$ such that X_0 is the union of all the edges in $\mathcal{E}_0$. It is said to be a *full subgraph* if two vertices $\mathbf{v}$ and $\mathbf{w}$ lie in X_0 and there is an edge $E \in \mathcal{E}$ joining them, then $E \in \mathcal{E}_0$.

PROPOSITION 2. *Let $(X, \mathcal{E})$ be a finite edge-path graph, and let $(X_0, \mathcal{E}_0)$ be a subgraph. Then the derived graph $(X_0, \mathcal{E}_0')$ is a full subgraph of $(X, \mathcal{E}')$.*

Proof. Suppose we are given an edge K in $(X, \mathcal{E}')$, so that its vertices must have the form $\mathbf{y}$, $\mathbf{m}_L$ where L is an edge in $(X, \mathcal{E})$ that has $\mathbf{y}$ as one of its endpoints and $\mathbf{m}_L$ is a non-vertex point in L. If both of these vertices belong to X_0, then the latter contains a point of L which is not a vertex, and since $(X_0, \mathcal{E}_0)$ is a subgraph it follows that L must be entirely contained in X_0. But this automatically implies that the edge in the derived complex with endpoints $\mathbf{y}$ and $\mathbf{m}_L$ must be contained in X_0.■

Connectedness

One immediate consequence of the definitions is that every point of a graph lies in the arc component of some vertex; specifically, if x lies on the edge E and the vertices of the latter are a and b, then x lies in the same arc component as both a and b. In fact, one can prove much stronger conclusions:

PROPOSITION 3. *If $(X, \mathcal{E})$ is a finite edge-path graph, then X is connected if and only if for each pair of distinct vertices $\mathbf{v}$ and $\mathbf{w}$ there is an* edge-path sequence *$E_1, \cdots, E_n$ such that $\mathbf{v}$ is one vertex of E_1, $\mathbf{w}$ is one vertex of E_n, for each k satisfying $1 < k \leq n$ the edges E_k and E_{k-1} have one vertex in common, and $\mathbf{v}$ and $\mathbf{w}$ are the "other" vertices of E_1 and E_n. Furthermore, X is a union of finitely many components, each of which is a full subgraph.*

IMPORTANT: In a general edge-path sequence defined as in the statement of the proposition, we do **NOT** make any assumptions about whether or not these two vertices are equal. If they are, then we shall say that the edge-path sequence is *closed* or that it is a *circuit* or **cycle**.

Proof. First of all, since every point lies on an edge, it follows that every point lie in the connected component of some vertex. In particular, there are only finitely many connected components. Define a binary relation on the set of vertices such that $\mathbf{v} \sim \mathbf{w}$ if and only if the two vertices are equal or there is an edge-path sequence as in the statement of the proposition. It is elementary to check that this is an equivalence relation, and that vertices in the same equivalence class determine the same connected component in X.

Given a vertex $\mathbf{v}$, let $Y_{\mathbf{v}}$ denote the union of all edges containing vertices which are equivalent to $\mathbf{v}$ in the sense of the preceding paragraph. If we choose one vertex $\mathbf{v}$ from each equivalence class, then we obtain a finite, pairwise disjoint family of closed connected subsets whose union is X, and it follows that these sets are must be the connected components of X. In fact, by construction each of these connected component is a full subgraph of $(X, \mathcal{E})$.■

Frequently it is convenient to look at edge-path sequences that are *minimal* or *simple* in the sense that one cannot easily extract shorter edge-path sequences from them. Here is a more precise formulation:

Definition. Let $E_1, \cdots, E_n$ be an edge-path sequence such that the vertices of E_i are $\mathbf{v}_{i-1}$ and $\mathbf{v}_i$. This sequence is said to be *reduced* if $\mathbf{v}_1, \cdots, \mathbf{v}_n$ are distinct and either $n \neq 2$ or else $\mathbf{v}_0 \neq \mathbf{v}_2$ (if $n = 2$ and $\mathbf{v}_0 = \mathbf{v}_2$, then the edge-path is just a sequence with $E_2 = E_1$, physically corresponding to going first along E_1 in one direction and then back in the opposite direction).

We then have the following result:

PROPOSITION 4. *If two distinct vertices $\mathbf{x}$ and $\mathbf{y}$ can be connected by an edge-path sequence, then they can be connected by a reduced sequence.*

Proof. Take a sequence with a minimum number of edges. We claim it is automatically reduced. If not, then there is a first vertex which is repeated, and a first time at which it is repeated. In other words, there is a minimal pair (i, j) such that $i < j$ and $\mathbf{v}_i = \mathbf{v}_j$, which means that if (p, q) is any other pair with this property we have $p \geq i$ and $q > j$. If we remove E_{i+1} through E_j from the edge-path sequence, we obtain a shorter sequence which joins the given two vertices.■

There may be several different reduced sequences joining a given pair of vertices. For example, take X to be a triangle graph in the plane whose vertices are the three noncollinear points **a**, **b** and **c**, and whose edges are the three line segments joining these pairs of points. Then **ab**, **bc** and **ac** are two reduced edge-path sequences joining **a** to **c**.

Definition. A circuit (or cycle) $E_1 \ , \cdots \ , E_n$ is called a a *simple circuit* or *simple cycle* if it is reduced.

COROLLARY 5. *Every simple circuit in a graph contains at least three edges.*■

Further topological properties of graphs

By definition and construction, a finite edge-path graph is compact Hausdorff, and in fact one can say considerably more:

PROPOSITION 6. *A finite edge-path graph is homeomorphic to a subset of* $\mathbb{R}^n$ *for some* n.

At the end of this section we shall prove that a graph is always homeomorphic to a subset of $\mathbb{R}^3$.

Proof. Suppose that the vertices are $\mathbf{v}_1, \ \cdots \ , \mathbf{v}_n$. Consider the graph in $\mathbb{R}^n$ whose vertices are the standard unit vectors $\mathbf{e}_i$ and whose edges are the closed line segments $A_{i,j}$ joining these vertices; the resulting compact subspace of $\mathbb{R}^n$ is a graph because two of these segments intersect in at most a common endpoint (use linear independence of the unit vectors to prove this). Define a continuous map f from original the graph X to the new graph Y such that if E is an edge with vertices $\mathbf{v}_i$ and $\mathbf{j}$ for $i < j$ and E is a homeomorphism from $[0, 1]$ to E such that $\mathbf{v}_i$ corresponds to 0 and $\mathbf{v}_j$ corresponds to 1, then $t \in [0, 1]$ is sent to

$$t\,\mathbf{e}_j \ + \ (1 - t)\,\mathbf{e}_i$$

(since E is homeomorphic to $[0, 1]$ and endpoints are topologically characterized by the property that their complements are connected, it follows that either $\mathbf{v}_i$ corresponds to 0 and $\mathbf{v}_j$ corresponds to 1 or vice versa; in the second case, if we compose the original homeomorphism with the reflection on $[0, 1]$ sending s to $1 - s$, then we obtain a homeomorphism for which the first alternative holds).

It is a routine exercise to verify that f is continuous and 1–1, and therefore it maps X homeomorphically onto its image.■

The next result implies that a graph has a simply connected (universal) covering space.

PROPOSITION 7. *If* $(X, \mathcal{E})$ *is a finite edge-path graph and* $x \in X$*, then* x *has a (countable) neighborhood base of contractible open subsets.*

Proof. Suppose first that x is a vertex of X, and view X as a subset of $\mathbb{R}^n$ using the previous result. Define $\mathbf{OpenStar}\,(x, \mathcal{E})$ to be the complement of the set of all points on edges E' which do not have x as a vertex. The described set is the union of all E' which do not have x as a vertex and hence is closed, so its complement **OpenStar** $(x, \mathcal{E})$ must be open. For every ε such that $0 < \varepsilon < \sqrt{2}$ let

$$\mathbf{OpenStar}\,(\varepsilon; x, \mathcal{E})$$

denote the points in **OpenStar** $(x, \mathcal{E})$ whose distance from x is less than ε. Then it follows that there is a straight line homotopy from the identity on **OpenStar** $(\varepsilon; x, \mathcal{E})$ to the constant map with value x, and therefore every such neighborhood is contractible. Since X is presented as a metric space, it follows that a suitably chosen countable of this neighborhood family will be the desired countable neighborhood base of x.

Now suppose that x is not a vertex of X, so that there is a unique edge E containing x; by assumption x lies in the complement of the end points in E, and the corresponding subset of E is homeomorphic to the open interval $(0, 1)$. Since every point in $(0, 1)$ has a neighborhood base of contractible open subsets, the conclusion to the proposition will follow if we know that E − endpoints is open in X. The complement to this set is the set of all points that are either vertices of E or else lie on some edge other than E. This is a finite union of closed sets and hence closed, and therefore the set $E - \{\text{endpoints}\}$ must be open in X as desired.■

Addendum: Embedding graphs in $\mathbb{R}^3$

For many purposes it is enough to know that a connected graph is always topologically embeddable in some $\mathbb{R}^n$, but in some contexts it is useful to know the smallest n for which this is possible. The methods of point set topology show that a connected compact subset of $\mathbb{R}$ is an interval, and the final result of this section shows that for all other graphs the minimum value of n is 2 or 3.

THEOREM 8. *Let $(X, \mathcal{E})$ be a connected graph. Then there is a $1 - 1$ continuous mapping $\varphi : X \to \mathbb{R}^3$ such that every edge in $\mathcal{E}$ is mapped linearly to the closed segment joining the images of the vertices (in other words, the embedding is* rectilinear*).*

There are obvious examples for which we can take $n = 2$ (in particular, the complete graphs on three or four vertices), and later in this course we shall prove that $n = 3$ for two specific graphs; in other words, these graphs are not homeomorphic to subsets of $\mathbb{R}^2$ and are said to be *nonplanar*. A celebrated theorem due to K. Kuratowski characterizes all nonplanar connected graphs in terms of the two specific nonplanar examples that we shall analyze in Section VII.4.

Proof of Theorem 8. The proof is built around the following two key observations:

(A) *There is an infinite set of isolated points in $\mathbb{R}^3$ such that no four lie in a single plane.*

(B) *For the sequence of points in the preceding observation, the intersections of two line segments joining two pairs of vertices is at most a common vertex.*

Before proving these, we shall indicate how they yield the theorem. Given the connected graph $(X, \mathcal{E})$, define φ on vertices so that each vertex goes to a point in the sequence of (A) and the mapping is 1–1. Next, if K is an edge of $\mathcal{E}$ with vertices a and b, extend φ to K by sending the latter homeomorphically to the closed line segment joining $\varphi(a)$ to $\varphi(b)$. Construct such an extension for each edge in the graph. Then observation (B) implies that if $K \neq K'$ are distinct edges of $\mathcal{E}$, it follows that $\varphi[K] \cap \varphi[K']$ is at most a common vertex of the two line segments. Since the map on vertices is 1–1, it follows that this common vertex is a common vertex of K and K', and therefore the mapping φ will be continuous and 1–1, so that ϕ maps X homeomorphically to its image because X is compact.

Verification of observation (A). — Obviously $\mathbb{R}^3$ contains a set of 4 such points; for example, take $\mathbf{0}$ and the three standard unit vectors. Suppose we know that there is a set of n points, say

Y, satisfying the property in (A); by induction, it will suffice to find a similar subset containing Y with one additional point. Let

$$P_i \ , \qquad 1 \ \leq i \ \leq \ \binom{n}{3}$$

be the planes determined by triples of points in Y. Then $\mathbb{R}^3 - \cup_i P_i$ is an open dense subset. Form Y' by adding a point z in the complement of $\cup_i P_i$ such that $|z| \geq n+1$. If we continue to define points recursively in this manner, we obtain the sort of subset described in (A), and in fact is it a closed subset of $\mathbb{R}^3$.

Verification of observation (B). — There are two cases to consider, depending upon whether or not the two edges have a common vertex. Both rely upon the following elementary fact from linear algebra: *If* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *and* $\mathbf{d}$ *are nonplanar points in* $\mathbb{R}^3$, *then* $\mathbf{b}-\mathbf{a}$, $\mathbf{c}-\mathbf{a}$ *and* $\mathbf{d}-\mathbf{a}$ *are linearly independent.* (If the points were coplanar, then there would be a 2-dimensional vector subspace containing the three difference vectors.)

CASE 1. Suppose that the segments have one endpoint in common, so that the endpoints of one segment are $\mathbf{x}$ and $\mathbf{y}$ while the endpoints of the other are $\mathbf{y}$ and $\mathbf{z}$. We claim that they cannot have any other points in common. If there were such a point then there would be $s, t \in [0,1]$ such that

$$t\mathbf{x} \ + \ (1-t)\mathbf{y} \ = \ s\mathbf{z} \ + \ (1-s)\mathbf{y}$$

so that $t(\mathbf{x}-\mathbf{y}) = s(\mathbf{z}-\mathbf{y})$. Since there is a fourth point $\mathbf{w}$ such that $\mathbf{y}$, $\mathbf{x}$, $\mathbf{z}$ and $\mathbf{w}$ are nonplanar, it follows that $\mathbf{x}-\mathbf{y}$ and $\mathbf{z}-\mathbf{y}$ are linearly independent, and consequently the only way the displayed equation can be valid is if $s = t = 0$, in which case we have

$$t\mathbf{x} + (1-t)\mathbf{y} \ = \ \mathbf{y} \ = \ s\mathbf{z} + (1-s)\mathbf{y}$$

so that the common point must be the common endpoint of the two segments.

CASE 2. Suppose that the segments have no endpoints in common and that they join $\mathbf{y}$ to $\mathbf{x}$ and $\mathbf{z}$ to $\mathbf{w}$ respectively. As in the preceding case, if the segments did have some point in common, then there would be $s, t \in [0,1]$ such that

$$t\mathbf{x} \ + \ (1-t)\mathbf{y} \ = \ s\mathbf{w} \ + \ (1-s)\mathbf{z}$$

and if we subtract one side of this equation from the other we obtain the equation

$$t\mathbf{x} \ + \ (1-t)\mathbf{y} \ - \ s\mathbf{w} \ + \ (s-1)\mathbf{z} \ = \ \mathbf{0} \ .$$

By linear independence this equation is satisfied if and only if each of $t, (1-t), -s, (s-1)$ is zero. However, this is impossible because it implies both $t=0$ and $t=1$. Therefore the two segments cannot have any points in common. — As indicated previously, this completes the proof of Theorem 8.■

III.2 : Maximal trees

(**M**, §84; **H**, §1.A))

The graph $(X, \mathcal{E})$ is said to be a *tree* if for distinct vertices u and v in X there is a UNIQUE reduced edge path sequence

$$E_1, \cdots, E_r$$

such that u is the "initial" endpoint of E_1 and v is the "final" endpoint of E_r.

A reduced edge path sequence is sometimes called a **simple chain** in the graph.

One can visualize many examples of trees by looking at letters of the alphabet; examples include the letters

E, F, H, I, K, M, N, V, W, X, Y, Z.

The numerals 4 and 5 as depicted on a standard calculator display (with an open top on the 4) also correspond to examples of trees. On the other hand, the linear graphs corresponding to triangles, rectangles, pentagons, etc. are not trees. Other nonexamples include the letter A, the numeral 4 as depicted in print with a closed top, and the numerals 6 and 8 as depicted on a standard calculator display.

PROPOSITION 1. *Every tree has a vertex that lies on only one edge.*

Proof. If the tree has only one edge, then the result is trivial. Assume now that the tree $(X, \mathcal{E})$ has $m \geq 2$ edges. We shall assume further that every vertex of $(X, \mathcal{E})$ lies on at least two distinct edges and derive a contradiction.

Let A_1 be an edge of $(X, \mathcal{E})$, and let v_0 and v_1 be its edges. Let A_2 be a second edge which has v_1 as a vertex, and let v_2 be its other vertex. Continuing in this manner, we obtain an infinite sequence of edges A_n with vertices v_{n-1} and v_n such that $A_n \neq A_{n-1}$. Since there are only finitely many edges and vertices, there must be some first value of k such that $v_k = v_{k+j}$ for some $j > 0$ (in other words, there is a first repeated value in the sequence). By construction, we must have $j \geq 3$. By construction, we know that A_{k+j} defines a simple chain joining $v_k = v_{j+k}$ to v_{j+k-1} and similarly the sequence $A_{k+1}, \cdots A_{k+j-1}$ defines a simple chain joining these two vertices. Now the first simple chain consists of one edge, while the second consists of at least two because $k + j - 1 \geq k + 2 > k + 1$, and thus we have constructed two simple chains joining these vertices. Since we are assuming the graph is a tree, this is impossible, and therefore it follows that there must be some vertex which lies on only one edge.∎

We shall also need the following companion result:

PROPOSITION 2. *Suppose $(X, \mathcal{E})$ is a tree and v_0 is a vertex which lies on only one edge, say E_0. Let $(X_0, \mathcal{E}_0)$ be the subgraph given by the union of all edges except E_0 (hence its vertices are all the vertices of the original graph except v_0). Then $(X_0, \mathcal{E}_0)$ is also a tree.*

Proof. Suppose that u and w are vertices of the subgraph and $A_1, \cdots, A_r$ is a simple chain connecting them. We claim that none of the edges A_i can be equal to E_0; if this is true then it will follow that the subgraph will be a tree (see the final step of the argument).

As usual, label the vertices of the edges A_i such that $u = a_0$, $w = a_r$, and the vertices of A_i are a_i and a_{i-1}. By hypothesis, $A_i \neq A_{i\pm1}$ for all i. Suppose that we have $A_j = E_0$ for some j; then either $v_0 = a_{i-1}$ or else $v_0 = a_i$. Let v_1 be the other vertex of E_0.

CASE 1: Suppose that $v_0 = a_{i-1}$. Since $a_0 = u \neq v_0$, it follows that $i > 0$. Since E_0 is the only edge containing the vertex v_0, it follows that $A_{i-1} = A_i$, with $v_1 = a_{i-2} = a_i$. This contradicts the definition of a simple chain, and hence we can conclude that $v_0 \neq a_{i-1}$. CASE 2. Suppose that $v_0 = a_i$. Since $a_r = w \neq v_0$, it follows that $i < r$. Since E_0 is the only edge containing the vertex v_0, it follows that $A_{i+1} = A_i$, with $v_1 = a_{i-1} = a_{i+1}$. This contradicts the definition of a simple chain, and hence we can conclude that $v_0 \neq a_i$. Combining these results, we can conclude that E_0 does not appear in the simple chain sequence $A_1, \cdots, A_r$, so that the latter is a simple chain in $(X_0.\mathcal{E}_0$. This simple chain is unique in the smaller complex by the uniqueness condition on the larger complex, and therefore the smaller complex must also be a tree.■

We are now ready to state one of the most important properties of trees:

THEOREM 3. *If $(T, \mathcal{E})$ is a tree and $\mathbf{v}$ is a vertex of this graph, then $\{\mathbf{v}\}$ is a strong deformation retract of X.*

Proof. This is trivial for graphs with one edge because they are homeomorphic to the unit interval. Suppose now that we know the result for trees with at most n edges, and suppose that $(T, \mathcal{E})$ has $n + 1$ edges.

By Lemma 84.2 we may write $T = T_0 \cup A$ where A is an edge and T_0 is a tree with n edges such that $A \cap T_0$ is a single vertex $\mathbf{w}$. Let $\mathbf{y}$ be the other vertex of A. The proof splits into cases depending upon whether or not the vertex $\mathbf{v}$ of T is equal to $\mathbf{y}$, $\mathbf{w}$ or some other vertex in T_0.

We shall need the following two results on strong deformation retracts; in both cases the proofs are elementary:

(1) *Suppose X is a union of two closed subspaces $A \cup B$, and let $A \cap B = C$. If C is a strong deformation retract of both A and B, then C is a strong deformation retract of X.*

(2) *Suppose X is a union of two closed subspaces $A \cup B$, and let $A \cap B = C$. If C is a strong deformation retract of B, then A is a strong deformation retract of X.*

Suppose first that the vertex is $\mathbf{w}$. Then $\{\mathbf{w}\}$ is a strong deformation retract of both A and T_0, so by the first statement above it is a strong deformation retract of their union, which is T.

Now suppose that the vertex is $\mathbf{y}$. Then the second statement above implies that A is a strong deformation retract of T. Since $\{\mathbf{y}\}$ is a strong deformation retract of A, it follows that $\{\mathbf{y}\}$ is also a strong deformation retract of T.

Finally, suppose that the vertex $\mathbf{v}$ lies in T_0 but is not $\mathbf{w}$. Another application of the second statement implies that T_0 is a strong deformation retract of T, and since $\{\mathbf{v}\}$ is a strong deformation retract of T_0, it follows that $\{\mathbf{v}\}$ is also a strong deformation retract of T.■

COROLLARY 4. *The fundamental group of a tree is trivial.*■

Definition. Let $(X, \mathcal{E})$ be a graph. A subgraph $M \subset X$ is a *maximal tree* in X if M is a tree and there is no tree M' in X which properly contains M.

It is fairly straightforward to show that maximal trees exist. First of all, X must contain subgraphs that are trees, for any subgraph consisting of a single edge is a tree. Because of this, it follows that there must be some tree in X with a maximal number of edges, and this will be a maximal tree.■

For the sake of completeness, we state the following elementary result:

LEMMA 5. *If $(X, \mathcal{E})$ is a graph with a maximal tree M and Y is a subgraph of X containing M, then M is a maximal tree in Y.*■

Finally, we shall need the following important property of maximal trees:

PROPOSITION 6. *If $(X, \mathcal{E})$ is a connected graph and $T \subset X$ is a maximal tree, then all the vertices of $(X, \mathcal{E})$ belong to T.*

Proof. As usual, assume the conclusion is false; then there is some vertex $v \notin T$. By connectedness there is an edge-path sequence joining v to some point in T, and among these sequences there is one $E_1, \cdots, E_n$ of minimum length. Since we have an edge-path sequence we can denote the vertices on the edges by $v_0, \cdots, v_n$ such that $v = v_0$, $v_n \in T$, and the edges of E_i are v_i and v_{i-1}. By the minimality of this sequence we know that $v_i \in T$ if and only if $i = n$.

Let $T_1 = T \cup E_n$. We claim that T_1 is also a tree. The key point in verifying this will be the following observation:

> If an edge-path sequence in T_1 contains E_n, then E_n is either the first edge or the last edge, and v_{n-1} is either the initial vertex or the final vertex.

This is true because the vertex v_{n-1} lies on E_n but not on any edges in T (if it did, then $v_{n-1} \in T$ and by our assumptions this is not the case). If E_n appeared in the middle of the sequence, the one of the two edges containing v_{n-1} would have to lie in T, and this would imply $v_{n-1} \in T$.

To prove that T_1 is a tree, consider an arbitrary pair of vertices w and w'. If they both lie in T, then there is a unique reduced edge-path in T joining them, and we claim that there is no other reduced edge-path in T_1 which joins them. Any such path would have to contain the edge E_n (the only edge not in T). Since a reduced edge-path containing E_n must start or end with E_n, such an edge-path cannot join two points in T. — Now consider reduced edge-path sequences joining v_{n-1} to some vertex w in T. Since $v_n \in T$, there is a unique edge-path sequence $K_1, \cdots, K_m$ joining v_n to w. If we insert E_n at the beginning of this sequence, we obtain a reduced edge-path sequence joining v_{n-1} to w in T_1. To see that this sequence is unique, note that every edge-path sequence joining v_{n-1} to w must start with E_n because no other edge in T_1 contains v_{n-1}. If we remove E_n from the sequence, we obtain a reduced edge-path sequence joining v_n to w, and since E_n does not appear in this sequence it must be an edge-path sequence in T. Therefore the sequence joining v_n to w must be the previously described edge-path sequence $K_1, \cdots, K_m$, and it follows that there is only one edge-path sequence in T_1 joining v_{n-1} to w.

The preceding shows that T_1 is a tree in X which properly contains T. Since T was assumed to be a maximal tree, this yields a contradiction, so our hypothesis about a vertex not in T must be false and hence T must contain all the vertices.■

III.3 : Fundamental groups of graphs

(**M**, §84; **H**, §1.A))

In this section we shall show that the fundamental group of a connected graph $(X, \mathcal{E})$ has a very simple description depending only upon the numbers of vertices and edges in $\mathcal{E}$.

We already know that the fundamental group of a tree is trivial, and the crucial step in proving the main result is to describe the fundamental group of a connected graph which is the union of a maximal tree and a single edge. A complete graph on three vertices has this property, and its fundamental group is infinite cyclic because such a graph is homeomorphic to S^1 (verify this), and the first result is a generalization of this fact to all graphs which are unions of a tree and a single edge.

PROPOSITION 1. *Suppose that the connected graph $(X, \mathcal{E})$ contains a maximal tree T such that X is the union of T with a single edge E^*. Then X is homotopy equivalent to S^1.*

The file `graphpix2.pdf` in the course directory contains a drawing and a simplified discussion of the main ideas in the proof.

Proof. Since T is a maximal tree, the vertices of E^* lie in T. If $\mathbf{a}$ and $\mathbf{b}$ are these vertices, then there is a reduced edge-path sequence $E_1, \cdots, E_n$ joining $\mathbf{a}$ to $\mathbf{b}$, and if we let Γ be the union of the $E - i$'s and E^*, it follows that Γ must be homeomorphic to S^1. By construction Γ determines a subgraph of X. For the sake of uniformity, set $\mathbf{v}_0 = \mathbf{a}$ and $\mathbf{v}_n = \mathbf{b}$.

We claim that Γ is a strong deformation retract of X. Let Y be the subgraph obtained by removing the edges E^* and E_i from $\mathcal{E}$, and for each i let Y_i be the component of $\mathbf{v}_i$. By our assumptions it follows that Y and the subgraphs Y_i are trees. It will suffice to prove that if $i \neq j$ then $\mathbf{v}_j \notin Y_i$, for then we have $Y_i \cap \Gamma = \{\mathbf{v}_i\}$ and we can repeatedly apply the criteria in the previous argument to show that Γ is a strong deformation retract of X.

Suppose now that $\mathbf{v}_j \notin Y_i$ for some $j \neq i$. Then there is some reduced edge-path sequence $F_1, \cdots, F_m$ joining $\mathbf{v}_i$ to $\mathbf{v}_j$ in Y_i. Since the vertices of the edges F_r contain at least one $\mathbf{v}_j$ other than $\mathbf{v}_i$, there is a first edge in the sequence F_s which contains such an edge, say $\mathbf{v}_k$. Of course, none of the edges F_r lies in Γ. However, we also know that there is a reduced edge path sequence in $\Gamma \cap T$ which joins $\mathbf{v}_j$ to $\mathbf{v}_k$, and we can merge this with the edge-path sequence $F_1, \cdots, F_s$ (whose edges lie in T but not Γ) to obtain a reduced cycle in T. Since T is a tree, this is a contradiction, and therefore we must have $Y_i \cap \Gamma = \{\mathbf{v}_i\}$. As noted before, this suffices to complete the proof.■

The preceding special case is a key step in proving the following general result:

THEOREM 2. *Let $(X, \mathcal{E})$ be a connected graph, let T be a maximal tree in X, and let p be a vertex of T. Then $\pi_1(X, x)$ is a free group on k generators, where k is the number of edges that are in X but not in T.*

Let T be a maximal tree in the connected graph X, and let $F_1, \cdots, F_b$ denote the edges of X which do not lie in T. Let $W \subset X$ be the open set obtained by deleting exactly one non-vertex point from each of the edges F_i, and let $U_j = W \cup F_j$. It then follows that each U_j is an open subset of X and if $i \neq j$ then $U_i \cap U_j = W$. Furthermore T is a strong deformation retract of W and for each subset of indices $i_1, \cdots, i_k$ the set $F_{i_1} \ \cup \ F_{i_k}$ is a strong deformation retract of $U_{i_1} \ \cup \ U_{i_k}$. In particular, we know that the sets W and U_i are all arcwise connected. By the preceding result we know that F_1 and U_1 are homotopy equivalent to S^1, and we claim by induction that the fundamental groups of $F_1 \cup \ \cdots \ \cup F_t$ and $U_1 \cup \ \cdots \ \cup U_t$ are free on t generators. For if the result is true for $t \geq 1$, then we have

$$\bigcup_{i \leq t+1} U_i \ = \ \left(\bigcup_{i \leq t} U_i \right) \ \cup \ U_{t+1} \ , \qquad W \ = \ \left(\bigcup_{i \leq t} U_i \right) \ \cap \ U_{t+1}$$

so that the Seifert-van Kampen Theorem implies that the fundamental group of $U_1 \cup \cdots \cup U_{t+1}$ is the free product of the fundamental groups of $U_1 \cup \cdots \cup U_t$ and U_{t+1}. By induction the group for the first space is free on t generators while the group for the second is infinite cyclic, and this completes the proof of the inductive step.■

The preceding results yield a few simple criteria for recognizing when a connected graph is a tree.

THEOREM 3. *If X is a connected graph, then the following are equivalent:*

(*i*) *X is a tree.*

(*ii*) *X is contractible.*

(*iii*) *X is simply connected.*

Proof. We already know that the first condition implies the second and the second implies the third, so it is only necessary to prove that (*iii*) implies (*i*). However, if T is a maximal tree in X and $T \neq X$, then we know that the fundamental group of X is a free group on k generators, where $k > 0$ is the number of edges which are in X but not in T. Therefore if X is simply connected we must have $T = X$.■

The Euler characteristic of a graph

If $(X, \mathcal{E})$ is a connected graph, then the preceding discussion shows that the fundamental group of X is a free group on a finite set of free generators. We would like to have a formula for the number of generators which can be read off immediately from the graph structure and does not require us to find an explicit maximal tree inside the graph.

Definition. The *Euler characteristic* of $(X, \mathcal{E})$ is the integer $\chi(X, \mathcal{E}) = v - e$, where e is the number of edges in the graph and v is the number of vertices.

If there is exactly one edge, then clearly $v = 2$, $e = 1$, and the Euler characteristic is equal to $1 = 2 - 1$. The first indication of the Euler characteristic's potential usefulness is an extension of this to arbitrary trees.

PROPOSITION 4. *If $(T, \mathcal{E})$ is a tree, then $\chi(T, \mathcal{E}) = 1$.*

Proof. Not surprisingly, this goes by induction on the number of edges. We already know the formula if there is one edge. As before, if we know the result for trees with n edges and T has $n+1$ edges we may write $T = T_0 \cup A$, where T_0 is a tree, A is a vertex, and their intersection is a single point. For each subgraph Y let $e(Y)$ and $v(Y)$ denote the numbers of edges and vertices in Y. Then we have $e(T) = e(T_0) + 1$, $v(T) = v(T_0) + 1$, and hence we also have

$$\chi(T) \quad = \quad v(T) - e(T) \quad = \quad [v(T_0) + 1] - [e(T_0) + 1] \quad = \quad v(T_0) - e(T_0) \quad = \quad 1$$

which is the formula we wanted to verify.■

THEOREM 5. *If $(X, \mathcal{E})$ is a connected graph, then the fundamental group of X is a free group on $1 - \chi(X, \mathcal{E})$ generators.*

Note that if G is a finite group on n generators, then there are exactly 2^n homomorphisms from G to $\mathbb{Z}_2$ (there are that many ways to define a function from the set of generators to $\mathbb{Z}_2$, and each such function extends uniquely to a group homomorphism). Therefore the number of free generators does not depend upon the choice of a generating set; more precisely, if m and n are positive integers such that G is free on sets of m and n generators, then $m = n$. — Similarly, a group G cannot be simultaneously free on both a finite and an infinite set of generators, for the number of homomorphisms into $\mathbb{Z}_2$ is finite if and only if the generating set is finite; finally, if α is a transfinite cardinal number, then a free group on a set of generators with cardinality α also has cardinality α (verify this), and therefore in all cases the cardinality of a set of free generators for a free group is independent of the choice of generators.

Proof of Theorem 5. We adopt the notational conventions in the preceding argument. Let T be a maximal tree in X, and suppose that there are k edges in X which are not in T, so that the fundamental group is free on k generators. By construction we know that $v(T) = v(X)$ and $e(X) = e(T) + k$, and by the preceding result we know that the Euler characteristic of T is 1. Therefore we have

$$\chi(X, \mathcal{E}) \quad = \quad v(X) - e(X) \quad = \quad v(T) - e(T) - k \quad = \quad 1 - k$$

so that $k = 1 - \chi(X, \mathcal{E})$ as required.■

In the exercises we note that the theorem is also valid for the edge-path graphs as defined in the files for this course.

COROLLARY 6. *If two connected graphs X and X' are base point preservingly homotopy equivalent as topological spaces, then they have the same Euler characteristics.*

In particular, the corollary applies if X and X' are homeomorphic. For this reason we often suppress the edge decomposition and simply use $\chi(X)$ when writing the Euler characteristic.

Proof. If X and X' are homotopy equivalent, then their fundamental groups are isomorphic, and hence they are both free groups with the same numbers of generators. Since the Euler characteristics can be expressed as functions of these numbers of generators, it follows that the Euler characteristics of X and X' must be equal.■

COROLLARY 7. *A connected graph X is a tree if and only if $\chi(X) = 1$.*

Proof. We know that $\chi(X) = 1$ if and only if X is simply connected.■

REMARK. More generally, one has the following criteria for recognizing whether two connected graphs X and Y are homotopy equivalent:

(1) *The connected graphs X and Y are homotopy equivalent if and only if their fundamental groups are isomorphic.*■

(2) *The connected graphs X and Y are homotopy equivalent if and only if their Euler characteristics are equal.*■

The results of this course show that the fundamental groups are isomorphic if and only if the Euler characteristics are equal, so (2) will follow from (1). Proving the latter is not all that difficult, but we shall not give the details here.

III.4 : Finite coverings of graphs

($\mathbf{M}$, §§83, 85; $\mathbf{H}$, §1.A)

As indicated at the beginning of this unit, we shall conclude our discussion of graphs with an application to some mildly counter-intuitive results on finite index subgroups of finitely generated free groups. Since each such group is the fundamental group of some graph X and finite index subgroups of $\pi_1(X)$ should correspond to finite coverings of X, we begin with an observation about finite covering spaces of graphs.

PROPOSITION 1. *Let $(X, \mathcal{E})$ be a connected finite graph, and suppose that $p : W \to X$ be a connected finite covering. Then there is a finite graph complex structure $\mathcal{E}'$ on W such that p maps each edge of $\mathcal{E}'$ homeomorphically to an edge of $\mathcal{E}$.*

Proof. Let k be the number of sheets in the covering space projection, so that each point in X has exactly k preimages in W. If E_α is an edge in $\mathcal{E}$ and $F_\alpha = p^{-1}[E_\alpha]$, then the restriction p_α of p to F_α defines an k-sheeted covering space projection over E_α; the space F_α is not necessarily connected, and in fact F_α is a finite disjoint union of connected covering spaces over X, with each such space corresponding to a component of F_α. Since E_α is simply connected, p_α must be a homeomorphism on each component of F_α, and it follows that F_α must be a union of pairwise disjoint compact subspaces $F_{\alpha,j}$ where $1 \leq j \leq k$ such that p_α maps each subspace homeomorphically onto E_α.

To show that the subsets $F_{\alpha,j}$ determine a graph structure on W, we need to look at the intersections

$$F_{\alpha,j} \ \cap \ F_{\beta,i}$$

where $(\alpha, j) \neq (\beta, i)$. If $\alpha = \beta$ then these intersections are empty by the reasoning in the preceding paragraph. If $\alpha \neq \beta$, then the relations

$$p\,[F_{\alpha,j} \ \cap \ F_{\beta,i}] \quad \subset \quad p\,[F_{\alpha,j}] \ \cap \ p\,[F_{\beta,i}] \quad = \quad E_\alpha \ \cap \ E_\beta$$

imply that $F_{\alpha,j} \ \cap \ F_{\beta,i}$ is empty if $E_\alpha \cap E_\beta$ is empty, while if the latter is not empty then the intersection is contained in the inverse image of the vertex common to $E_\alpha \cap E_\beta$. Since the restriction of p to each component of F_α is 1–1, it follows that $F_{\alpha,j} \cap F_{\beta,i}$ contains at most one point if $\alpha = \beta$, and if this happens then this point is a vertex of both $F_{\alpha,j}$ and $F_{\beta,i}$, proving that we have a graph structure on X.■.

We can use this result to prove the following purely algebraic result on subgroups of finite index:

PROPOSITION 2. *Let F be a free group on k generators, for some positive integer k, and let H be a subgroup of index n. Then H is a free group on $nk - n + 1$ generators.*

A standard result in algebra states that if M is a finitely generated free module on m generators over a principal ideal domain $\mathbb{D}$ and $N \subset M$ is a $\mathbb{D}$-submodule, then N is free on n generators for some $n \leq m$. In contrast, the result above says that a free subgroup of a free group may have more generators than the group containing it. After proving this result, we shall also describe an example to show that a finitely generated free group also contains a non-finitely generated free subgroup (which is not of finite index).

Proof. Let $(X, \mathcal{E})$ be a connected graph whose fundamental group is free on k generators; one method of constructing such a graph is to take edges A_i, B_i and C_i for $1 \leq i \leq k$, where the edges of A_i are x, p_i, and q_i, the edges of B_i are x, r_i, and s_i, and the edges of B_i are x, u_i, and v_i (topologically, X is a union of k circles such that each pair intersect at x and nowhere else). By the formula relating the number of generators for F and the Euler characteristic, we know that $k = 1 - \chi(X)$, or equivalently $\chi(X) = 1 - k$. Let Y be the connected covering space of X corresponding to the subgroup H. Then Y is a graph, and the fundamental group of Y is H, so that H is a free group.

We know that the number of free generators for H is given by $1 - \chi(Y)$, so it is only necessary to compute this Euler characteristic. Let e and v be the number of edges and vertices for $(X, \mathcal{E})$, so that $n = 1 - \chi(X)$, where $\chi(X) = v - e$. Since Y is an n-sheeted covering of X, if we take the associated edge decomposition of Y (such that each edge of Y maps homeomorphically to an edge of X) we see that the numbers of vertices and edges for Y are nv and ne respectively, so that

$$\chi(Y) \quad = \quad n \cdot \chi(X) \ .$$

Therefore the number of generators for the fundamental group of Y is given by

$$1 - \chi(Y) \quad = \quad 1 - n \cdot \chi(X) \quad = \quad 1 - n(1-k) \quad = \quad nk - n + 1$$

which is what we wanted to prove.■

COROLLARY 3. *If F is a free group on k generators for some $k \geq 2$, then F contains free subgroups on m generators for all $m \geq k$.*

QUESTION. Does this result extend to the case $k = 1$? Prove this or explain why it cannot be true.

Proof. It suffices to prove this result when $k = 2$ since F automatically contains a free subgroup with 2 generators (take a subset of some generating set for F).

Let X be a graph whose fundamental group is free on 2 generators u and v, and let Y_n be the n-sheeted covering space whose fundamental group is the (free) subgroup generated by by u and v^n for some $n \geq 2$. Then the fundamental group of Y_n is isomorphic to a free group on $n+1$ generators. It follows that for every positive integer $m \geq 3$ there is some n such that $\pi_1(Y_n)$ contains a free group on m generators.■

Example. The free group on two generators also contains a free subgroup with a countably infinite set of generators (hence the same is also true for every free group on more than two generators). Here is a sketch of the argument. Filling in the details is left to the reader as an exercise:

Let $X = S^1 \vee S^1$ with base point given by the common point of the two circles, and let u and v be free generators of $\pi_1(X)$ which are represented by the two circles. Let K denote the kernel of the homomorphism from $\pi_1(X)$ to $\mathbb{Z}$ which sends u to zero and v to a generator.

Let Y be the covering space of X whose fundamental group is isomorphic to K. It follows that Y is homeomorphic to a copy of the real line with a circle attached at each point $2q\pi$ where q runs through all integers (verify this!). An explicit model for Y is the set of all points $(x, y, z) \in \mathbb{R}^3$ such that either $(x, y) = (1, 0)$ (in other words, the line with parametric equations $(1, 0, t)$ for $t \in \mathbb{R}$) or $x^2 + y^2 = 1$ and $z = 2q\pi$ for some integer q. If we view X as the subset of $\mathbb{R}^2$ given by

$$\{ \ x^2 + y^2 \quad = \quad 1\} \ \cup \ \{ \ (x-2)^2 + y^2 = 1 \ \}$$

then the covering space projection corresponds to the map sending (x, y, z) to (x, y) on the first piece and sending $(1, 0, t)$ to

$$\left(2 - \cos t, -\sin t \right)$$

on the second.

Let $A_m \subset Y$ be the set of all points such that $|z| \leq m$. Then A_m consists of a closed line segment with $2m + 1$ circles attached symmetrically with respect to the end points. It follows that $\pi_1(A_m)$ is a free group on $2m + 1$ generators, and the inclusion of $\pi_1(A_m)$ in $\pi_1(A_{m+1})$ is a 1–1 map sending the free generators of the first group to a subset of a set of free generators for the second.

Since every compact subset of Y is contained in some A_m, it follows that the fundamental group of Y is an increasing limit of the fundamental groups of the subspaces A_m. Since this limit is a free group on a countably infinite set of generators, it follows that $\pi_1(Y)$ must have the same property.■

III.∞ : Infinite graphs

(**M**, §§83–85; **H**, §1.B)

If we are given a connected graph $(X, \mathcal{E})$ with a nontrivial (hence infinite) fundamental group, then by Proposition 1.7 and Theorem I.3.1 we know that X has a simply connected covering space, and the methods of Corollary I.1.2 show that the inverse image of an edge is a disjoint union of edges and hence the universal simply connected covering space $\widetilde{X}$ has a decomposition into subsets homeomorphic to intervals. It is natural to think of this as an infinite graph complex, and results from Munkres and Hatcher provide a mathematically precise setting for doing so. One particularly significant application of infinite graphs is the following result (which was originally proved by algebraic methods):

Munkres, Theorem 85.1, p. 514. *If G is a free group and $H \subset G$ is a subgroup, then H is also free.*

As noted at the end of the preceding section, if G has a finite set of free generators and is not cyclic, then the number of generators for H can be any number between 1 and $\aleph_0$ (the cardinal number of the integers).

Other uses of infinite graphs are discussed in Section 1.B of Hatcher, particularly in the section on graphs of groups. Such constructions are often extremely valuable sources of geometric insights into a group's algebraic structure. The Wikipedia articles

`http://en.wikipedia.org/wiki/Geometric_group_theory`

`http://en.wikipedia.org/wiki/Group_cohomology`

contain more detailed additional information (and further links) concerning the ways in which topology and group theory — particularly infinite group theory — have interacted with each other in mathematical research during the past century.

IV. Prelude to homology theory

This unit is intended as a bridge between the material on fundamental groups and covering spaces in the first three units and the material on homology theory in the remaining units of the course. The main link is an algebraic construction on the sets of vertices and edges of a connected graph which is called the (algebraic) **chain complex** of the graph. We shall illustrate how these chain groups relate to the **Königsberg bridge problem**, which is often regarded as the question which led to the development of algebraic topology. The next step is extending the notion of algebraic chains to higher dimensional objects which resemble graphs in the sense that they are constructed from standard building blocks, which turn out to be homeomorphic to n-fold products of the form $[0,1]^n$ for some positive integer n, in a fairly systematic manner. Finally, we shall consider 2-dimensional chains briefly and place the most basic algebraic properties of chain complexes into an abstract setting; in fact, this abstract study evolved into a major topic in algebra which goes by the name **homological algebra**, and there are a few comments about this in Section IV.∞.

There is a well-written motivational discussion for the basic concepts of homology theory in Chapter VI of the following book:

> **W. S. Massey.** *A Basic Course in Algebraic Topology*, Graduate Texts in Mathematics Vol. 127, *Springer-Verlag, New York NY*, 1991. ISBN: 0–387–97430–X.

This chapter is posted in the course directory as `massey-chapter6.pdf`; the information in this chapter might make the reasons behind the formal constructions in the next few units easier to understand.

IV.1 : Algebraic chains for graphs

(**H**, § 1.A)

Let $(X, \mathcal{E})$ be a finite graph complex, and let ω denote a linear ordering of the vertices (since the set of vertices is finite, the existence of such orderings is immediate); we shall often denote the combined graph structure and vertex ordering by symbolism such as $\mathcal{E}^\omega$. The **simplicial chain groups** $C_q(X, \mathcal{E}^\omega)$ are defined such that

> $C_1(X, \mathcal{E}^\omega)$ is free abelian on the edges in $\mathcal{E}$,
>
> $C_0(X, \mathcal{E}^\omega)$ is free abelian on the vertices in $\mathcal{E}$,
>
> $C_q(X, \mathcal{E}^\omega)$ is trivial if $q \neq 0, 1.$ in $\mathcal{E}$,

The associated **graph chain complex** $C_*(X, \mathcal{E}^\omega; d)$ consists of these simplicial chain groups together with **boundary homomorphisms**

$$d_q : C_q(X, \mathcal{E}^\omega) \longrightarrow C_{q-1}(X, \mathcal{E}^\omega)$$

such that $d_q = 0$ unless $q = 1$ and d_1 on a free generator corresponding to an edge E is given as follows: One of the two vertices of E precedes the other, and if $\partial_- E$ precedes $\partial_+ E$ let

$$d_1(E) \;=\; \partial_+ E \,-\, \partial_- E\ .$$

Since the chain groups are free abelian, this map of free generators extends to a homomorphism of chain groups.

Chains are often visualized as unions of simple edge paths in the graph, with $\pm E$ corresponding to an oriented path starting at $\partial_- E$ and ending at $\partial_+ E$.

The use of the word "chain" can be motivated geometrically as follows: Let $E_1, \cdots, E_n$ be an edge-path sequence such that the vertices of E_i are $\mathbf{v}_{i-1}$ and $\mathbf{v}_i$. Then we can associate a 1-chain to this path by the formula

$$c(E) \quad = \quad \sum_i \varepsilon_i \, E_i$$

where $\varepsilon_i \in \{\pm 1\}$ is defined to be $+1$ if $\mathbf{v}_{i-1}$ precedes $\mathbf{v}_i$ and -1 if $\mathbf{v}_i$ precedes $\mathbf{v}_{i-1}$. The signs have been chosen so that $d_1(c(E)) = \mathbf{v}_n - \mathbf{v}_0$. If the edge path is a **cycle** in the sense that the initial and final points agree, it follows that d_1 sends that chain to 0, and conversely if d_1 sends the chain for an edge path to zero then the edge path is a cycle. More generally, we define the **1-cycles** to be those 1-dimensional chains (or 1-chains) c such that $d_1(c) = 0$. Similarly, we say that a 0-dimensional chain b is a **boundary** if $b = d_1(c)$ for some 1-chain c. The **homology groups** $H_q(X, \mathcal{E}^\omega)$ are then given as follows:

$H_1(X, \mathcal{E}^\omega)$ is isomorphic to the kernel of d_1.

$H_0(X, \mathcal{E}^\omega)$ is isomorphic to the quotient of $C_0(X, \mathcal{E}^\omega)$ by the image of d_1.

We then have the following results relating the chain groups to the topological structure of the graph:

THEOREM 1. *Let $(X, \mathcal{E}^\omega; d)$ be a graph with a linear order of its vertices.*

(i) If the connected components of X are given by $(X_i, \mathcal{E}_i)$, then the chain groups $C_*(X, \mathcal{E}^\omega)$ and homology groups $H_*(X, \mathcal{E}^\omega)$ are isomorphic to direct sums of the corresponding groups for $(X_i, \mathcal{E}_i^\omega)$.

(ii) If X is connected then $H_0(X, \mathcal{E}^\omega) \cong \mathbb{Z}$ and $H_1(X, \mathcal{E}^\omega)$ is free abelian on $1 - \chi(X, \mathcal{E})$ generators, where $\chi(X, \mathcal{E})$ is the Euler characteristic given by number of vertices minus the number of edges.

This result implies that the structure of the homology groups is completely determined by the homotopy type of the underlying space X.

Proof. We begin with the first part. Since a graph is locally arcwise connected and compact, we know that it has finitely many components, and they are the same as the arc components. Therefore every edge of the graph is contained in a unique component, and the two vertices of the graph is also contained in that component. This means that the boundary map d_1 for the graph can be decomposed as a direct sum of boundary maps for the individual components

$$(d_1)_i : C_q(X_i, \mathcal{E}_i^\omega) \longrightarrow C_0(X_i, \mathcal{E}_i^\omega)$$

and it follows that one has a similar decomposition for the homology groups of $(X, \mathcal{E}^\omega)$.

Assume now that X is connected. The first step in proving (ii) is to prove the assertion about H_0. Define a map

$$\varepsilon : C_0(X, \mathcal{E}_i^\omega) \longrightarrow \mathbb{Z}$$

(called an *augmentation homomorphism*) such that its value at each vertex gnerator is equal to $+1$.

CLAIM 1. *The kernel of ε is equal to the image of d_1.*

To see that the kernel of ε contains the image of d_1 it suffices to show that $\varepsilon \circ d_1 = 0$, and since chain groups are free abelian it suffices to check this for a free generator corresponding to an arbitrary edge E. Since

$$\varepsilon \circ d_1(E) \quad = \quad \varepsilon\partial_+(E) \; - \; \varepsilon\partial_-(E) \quad = \quad 1 - 1 \quad = \quad 0 \; .$$

Suppose now that the chain $a = \Sigma_i \; n_i \, \mathbf{v}_i$ lies in the kernel of d_i, so that $\Sigma_i \; n_i = 0$ and

$$n_1 \quad = \quad - \sum_{i>1} n_i \qquad .$$

Since X is connected, for each $i > 1$ there is an edge path starting at $\mathbf{v}_1$ and ending at $\mathbf{v}_i$, and by the discussion above we have 1-chains c_i such that $d_1(c_i) = \mathbf{v}_i - \mathbf{v}_0$. Therefore we have

$$\sum_{i>1} n_i \, c_i \quad = \quad \sum_{i>1} n_i \; (\mathbf{v}_i - \mathbf{v}_0) \quad =$$

$$\sum_{i>1} n_i \, \mathbf{v}_i \; - \; \left(\sum_{i>1} n_i \right) \mathbf{v}_0$$

and since $\Sigma_{i \geq 1} \; n_i = 0$ the last expression is equal to a, so that a lies in the image of d_1.

The preceding discussion implies that H_0 is infinite cyclic and generated by the class of $\mathbf{v}$, where the latter is an arbitrary vertex of the graph, and the image of d_1 is the subgroup generated by $\mathbf{w} - \mathbf{v}^*$, where $\mathbf{v}^*$ is a fixed vertex and $\mathbf{w}$ runs through the remaining vertices. This means that the image of d_1 is a free abelian subgroup on $V - 1$ generators, where V denotes the set of vertices in the graph. Therefore d_1 defines a surjective homomorphism from C_1, which is free on E generators, to a free abelian group on $V - 1$ generators. The conclusion about the structure of H_1 will then be a consequence of the following algebraic result:

CLAIM 2. *If A and B are free abelian groups on α and β generators respectively and $\varphi : A \to B$ is surjective, then the kernel of φ is free abelian on $\beta - \alpha$ generators.*

This is true because A is isomorphic to the direct sum of B and the kernel of φ.

As noted above, this completes the proof of Theorem 1.■

Other coefficients. If $\mathbb{D}$ is an arbitrary commutative ring with unit, then one can define chains modules $C_*(X, \mathcal{E}^\omega; \mathbb{D})$ with coefficients in $\mathbb{D}$ if we replace the free abelian groups with the corresponding free modules over $\mathbb{D}$; the boundary homomorphism can be defined exactly as before, and it will be a $\mathbb{D}$-module homomorphism (hence the homology groups will also be $\mathbb{D}$-modules).

The Königsberg Bridge Problem

One particular graph that is historically noteworthy is the Königsberg Bridge Graph, in which the vertices correspond to four land masses in the city of Königsberg (now Kaliningrad, Russia) and the 1-cells (or *edges*) correspond to the seven bridges which joined pairs of land masses in the 18^{th} century (see `koenigsberg.pdf` for drawings). This configuration can be modeled by a graph with vertices $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ representing the land masses and edges representing one bridge each from $\mathbf{w}$ to $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ along with two bridges joining $\mathbf{y}$ to each of $\mathbf{x}$ and $\mathbf{z}$. This configuation is homotopic to a simplicial comples if we add extra vertices $\mathbf{u}_1$ and $\mathbf{u}_2$ on each of the bridges joining

$\mathbf{y}$ to $\mathbf{x}$ and $\mathbf{v}_1$ and $\mathbf{v}_2$ on each of the bridges joining $\mathbf{y}$ to $\mathbf{z}$. This will be our graph $(P, \mathbf{K})$, and we shall let C_* denote the ordered chain complex with $\mathbb{Z}_2$ coefficients which associated to some ordering of the vertices; since $1 = -1$ in $\mathbb{Z}_2$, one obtains the same boundary map for every ordering of the vertices.

The problem is to determine whether there is a path on this complex in which each bridge is crossed exactly once, and the first step is to formulate this in terms of the chain complex C_*. What we want is a 1-chain $\sum_{\mathbf{E}} \theta(\mathbf{E})\,\mathbf{E}$, where the sum runs over all free generators (or basis elements) of C_1 and $\theta_{\mathbf{E}}$ is nonzero for all $\mathbf{E}$, such that the boundary of this 1-chain has the form $\mathbf{p} + \mathbf{q}$ for two vertices in C_0 (the case $\mathbf{p} = \mathbf{q}$ is allowed). The problem is then to determine if such a 1-chain exists.

Euler's crucial insight into the problem can be stated algebraically as follows:

PROPOSITION 2. *Let $(X, \mathcal{E})$ be a connected graph, let $\gamma \in C_1(X, \mathcal{E}; \mathbb{Z}_2)$ be the 1-chain $\sum_{\mathbf{E}} \mathbf{E}$, where the sum runs over all free generators (or basis elements) of C_1, and write $d(\gamma) = \sum_{\mathbf{v}} n(\mathbf{v})\,\mathbf{v}$ for suitable mod 2 integers $n(\mathbf{v})$, where the sum runs over all vertices of $(P, \mathbf{K})$. Then $n(\mathbf{v})$ is the mod 2 reduction the number $m(\mathbf{v})$ of 1-simplices $\mathbf{E}$ that have $\mathbf{v}$ as one of their endpoints.*

Proof. An integer representing $n_{\mathbf{v}}$ is equal to the number (mod 2) of edges containing $\mathbf{v}$ as a vertex.∎

COROLLARY 3. *In the preceding setting, if there is a 1-chain γ such that $d(\gamma) = \mathbf{p} - \mathbf{q}$ where $\mathbf{p} = \mathbf{q}$ is possible, then $m_{\mathbf{v}}$ must be even if $\mathbf{v} \neq \mathbf{p}, \mathbf{q}$.*∎

Application to the Königsberg Bridge Problem. The impossibility of finding a suitable 1-chain for our Königsberg bridge network now follows by observing that $m = 3$ for $\mathbf{w}$, $\mathbf{x}$ and $\mathbf{z}$, while $m = 5$ for $\mathbf{y}$. In particular, if γ is a chain as in the statement of the theorem, then in $d(\gamma)$ the coefficients of all four of these vertices must be nonzero.∎

NOTE. We are not necessarily claiming that one needs to introduce chain groups in order to solve this problem (in fact, when Euler first solved the problem he did not view it as question in mathematics although he later modified his opinion). The purpose here is to illustrate how the use of chain complexes can provide a framework for finding solutions to this and similar questions.

It is left as an exercise for the reader to show that the homology groups for the Königsberg bridge graph are given by $H_1 \cong \mathbb{Z}^4$ and $H_0 \cong \mathbb{Z}$.

IV.2 : Triangulations and simplicial complexes

(**H**, §2.1)

We have seen that algebraic techniques work are fairly effective for studying graphs. These are 1-dimensional objects which are built from simple, well understood pieces which are either points or closed intervals. One natural question is whether we can develop similar techniques for studying higher dimensional spaces which are built from relatively simple pieces like points and finite cartesian products of a closed interval with itself. Since the n-dimensional sphere S^n is simply connected if $n \geq 2$, it is clear that we need something more than just the fundamental group to analyze questions in higher dimensions. Algebraic constructions known as **homology theories** turn out to be extremely useful generalizations of the fundamental groups of spaces and the previously defined homology groups of graphs, particularly for spaces which have decompositions into nice pieces. The goal of this section is to describe the sorts of decompositions that are needed and to discuss examples of spaces which have such decompositions.

One starting point is to consider subsets of the plane that are nicely behaved and are geometrically 2-dimensional. The simplest examples are closed polygonal regions which are bounded by a simple closed curve which is a broken line. Such regions include solid triangular regions, solid rectangular regions, and many other shapes that can be fairly irregular (see the first drawing in `triangulations1.pdf`). Suppose that we are given a closed polygonal region and we want to compute its area. If the region has an irregular shape, it is unlikely that we can use a closed formula to compute the area directly, but usually we can compute the area very efficiently by decomposing such a region into a finite union of nonoverlapping triangular regions for which there are good area formulas and then adding the areas of the triangular regions to obtain the area of the original region (see `triangulations1.pdf` again for examples of triangulating irregular regions). This suggests that our standard building blocks for 2-dimensional objects should be solid triangular regions in the plane, and they should fit together such that each pair meets in either a common edge or a common vertex. Further study suggests that every surface is a union of subspaces which are homeomorphic to triangles such that the intersection of any two corresponds to a common vertex or a common edge (see `triangulations2.pdf`).

Some additional thought is needed to decide on good examples of 3-dimensional building blocks. We would like to be able to use such blocks to build familiar sorts of figures like pyramids, prisms and rectangular solids (note that the latter are special cases of prisms); pictures of a few typical examples are on pages 4 and 5 of `triangulations1.pdf`. Since the base of a pyramid is a closed polygonal region and such regions can be decomposed into triangles, it follows that every pyramid has a nice decomposition into **tetrahedra**, which are pyramids with triangular bases (see the last two pages of `triangulations1.pdf`. Similarly, if we have a prism with a polygonal base, the decomposability of the latter into triangular regions means we only need to have decent building blocks for prisms with triangular bases. Fortunately, it turns out that a triangular prism can be decomposed into a union of triangular pyramids such that each pair intersect in a common face, edge or vertex (or maybe the empty set). The last page of `triangulations3.pdf` and the file `prism-dissection.pdf` indicate how this is done, and in the next section we shall describe this decomposition explicitly by means of linear equations and inequalities. The conclusion is that we can view tetrahedra as the fundamental building blocks for a large class 3-dimensional objects which arise in elementary(?) solid geometry.

In higher dimensions the appropriate building blocks will be analogs of triangular and tetrahedral regions. The n-dimensional analog is called an n-**simplex**; for $n = 1, 2, 3$ an n-simplex is simply a closed line segment, triangular region, or tetrahedral region respectively. Before proceeding further, we need to define the notion of simplex precisely, and this requires some background material on linear algebra and geometry.

Barycentric coordinates

To expedite the discussion, we begin with some online references for the notion of barycentric coordinates;

`en.wikipedia.org/wiki/Barycentric_coordinate_system(mathematics`

`mathworld.wolfram.com/BarycentricCoordinates.html`

A more leisurely and detailed discussion of barycentric coordinates, and more generally the use of linear algebra to study geometric problems, is contained in Section I.4 of the following online document:

`http://math.ucr.edu/~res/math133/geometrynotes1.pdf`

The file `math133exercises1.pdf` in the same directory has further material on these topics, and pages 13–30 of

`http://math.ucr.edu/~res/progeom/pgnotes02.pdf`

go further into the geometric uses of barycentric coordinates. Another standard reference is Chapter I of the following book:

> **J. F. P. Hudson.** *Piecewise Linear Topology. W. A. Benjamin, New York*, 1969. (Online: `http://www.maths.ed.ac.uk/~aar/surgery/hudson.pdf`)

An extremely detailed study of the topics in this section appears in the following online book:

`http://www.cis.penn.edu/~jean/gbooks/convexpoly.html`

Finally, Eilenberg and Steenrod also covers the portions this material needed for algebraic and geometric topology in greater detail.

Affine independence and barycentric coordinates. The crucial algebraic information is contained in the following result.

PROPOSITION 1. *Suppose that the ordered set of vectors* $\mathbf{v}_0, \cdots, \mathbf{v}_n$ *lie in some vector space* V. *Then the vectors* $\mathbf{v}_1 - \mathbf{v_0}, \cdots, \mathbf{v}_n - \mathbf{v}_0$ *are linearly independent if and only if every vector* $\mathbf{x} \in V$ *has at most one expansion of the form* $t_0\mathbf{v}_0 + \cdots + t_n\mathbf{v}_n$ *such that* $t_0 + \cdots + t_n = 1$.

A finite ordered set of vectors satisfying either (hence both) conditions is said to be *affinely independent*. Note that since the second condition does not depend upon the choice of ordering, a set of vectors is affinely independent if and only if for some arbitrary j the vectors $\mathbf{v}_i - \mathbf{v}_j$ (where $i \neq j$) is linearly independent. A linear combination in which the coefficients add up to 1 is called an *affine combination*.

Sketch of proof. To show the first statement implies the second, use the fact that $\mathbf{x} - \mathbf{v}_0$ has at most one expansion as a linear combination of $\mathbf{v}_1 - \mathbf{v_0}, \cdots, \mathbf{v}_n - \mathbf{v}_n$. To prove the reverse implication, show that if $\mathbf{x} - \mathbf{v}_0$ has more than one expansion as a linear combination of $\mathbf{v}_1 - \mathbf{v_0}, \cdots, \mathbf{v}_n - \mathbf{v}_n$, then $\mathbf{x}$ has more than one expansion as an affine combination of $\mathbf{v}_0, \cdots, \mathbf{v}_n$.■

COROLLARY 2. *If* $S = \{\mathbf{v}_0, \cdots, \mathbf{v}_n\}$ *is affinely independent, then every nonempty subset of* S *is affinely independent.*

This follows immediately from the uniqueness of expansions of vectors as affine combinations of vectors in S.■

The coefficients t_i are called **barycentric coordinates**. If we put physical weights of t_i units at the respective vertices $\mathbf{v}_i$, then the center of gravity for the system will be at the point $t_0\mathbf{v}_0 + \cdots + t_n\mathbf{v}_n$. If, say, $n = 2$, then this center of gravity will be inside the triangle with the given three vertices if and only if each t_i is positive, and it will be on the triangle defined by these vertices if and only if each t_i is nonnegative and at least one is equal to zero. A discussion of this physical interpretation in the 2-dimensional case appears in the file `centroids.pdf`. We should note that the discussion in this online reference can be extended to arbitrary (finite) dimensions.

More generally, if $\mathbf{v}_0, \cdots, \mathbf{v}_n$ are affinely independent then the **n-simplex** with vertices $\mathbf{v}_0, \cdots, \mathbf{v}_n$ is the set of all points expressible as affine combinations such that each coefficient is nonnegative (*i.e.*, **convex combinations**). For the record, we note that the standard plural form of simplex is simplices.

Frequently the n-simplex described above will be denoted by $\mathbf{v}_0 \cdots \mathbf{v}_n$. Note that if $n = 0$, then a 0-simplex consists of a single point, while a 1-simplex is a closed line segment, a 2-simplex is given by a triangle and the points that lie "inside" the triangle (also called a *solid triangular region*), and a 3-simplex is given by a pyramid with a triangular base (*i.e.*, a *tetrahedron*) together with the points inside this pyramid (also called a *solid tetrahedral region*). Illustrations of 3-simplices are given in the file `three-simplex.pdf`.

The following definition will also play an important role in our discussions.

Definition. If $\mathbf{v}_0, \cdots, \mathbf{v}_n$ form the vertices of a simplex $\mathbf{v}_0 \cdots \mathbf{v}_n$, then the **faces** of this simples are the simplices whose vertices are given by proper subsets of $\{\mathbf{v}_0, \cdots, \mathbf{v}_n\}$; note that such proper subsets are affinely independent by Corollary 2. If a proper subset $T \subset S$ has $k+1$ elements, then we shall say that the simplex $\Delta(T)$ whose vertices are given by T is a **k-face** of the original n-simplex, which in this notation is equal to $\Delta(S)$.

Definition. The *standard n-simplex* Δ_n is the set of all points $(t_0, \cdots, t_n) \in \mathbb{R}^{n+1}$ such that $t_j \geq 0$ for all j and $\sum_j t_j = 1$. Note that the set of unit vectors $\{\mathbf{e}_0, \cdots, \mathbf{e}_n\}$ is affinely independent because the set $\{\mathbf{e}_1 - \mathbf{e}_0, \cdots, \mathbf{e}_{n+1} - \mathbf{e}_0\}$ is linearly independent.

With the concept of an n-simplex at our disposal, we can define a suitable notion of polyhedron with arbitrary dimension.

Polyhedra and simplicial complexes

A subset $P \subset \mathbb{R}^m$ is a *polyhedron* if

(i) P is a finite union of simplices $A_1, \cdots, A_q$,

(ii) For each pair of indices $i \neq j$, the intersection $A_i \cap A_j$ is a common face.

The simplices $A_1, \cdots, A_q$ are said to form a *simplicial decomposition* of P, and if $\mathbf{K}$ is the collection of simplices given by the A_j and all their faces, then the ordered pair $(P, \mathbf{K})$ is called a (finite) *simplicial complex*.

If X is an arbitrary topological space, then a (finite) *triangulation* of X consists of a simplicial complex $(P, \mathbf{K})$ and a homeomorphism $t : P \to X$.

Numerous drawings of 2-dimensional examples appear in the file `triangulations3.pdf`. In these notes we shall concentrate on algebraic descriptions of basic examples.

ONE OF THE SIMPLEST EXAMPLES. Consider the solid rectangle in the plane given by $[a, b] \times [c, d]$, where $a < b$ and $c < d$. Everyday geometrical experience shows this can be split into two 2-simplices along a diagonal, and in fact it is the union of two 2-simplices, one with vertices (a, c), (a, d) and (b, d), and the other with vertices (a, c), (b, c) and (b, d). A point (x, y) which lies in the solid rectangle will be in the first simplex if and only if

$$(y - c)(b - a) \quad \leq \quad (d - c)(x - a)$$

and this point will be in the second simplex if and only if

$$(y - c)(b - a) \quad \geq \quad (d - c)(x - a)$$

Generalizations of this example will play an important role in the standard approach to algebraic topology.

If $(P, \mathbf{K})$ is a simplicial complex, then a subset $\mathbf{L} \subset \mathbf{K}$ is said to be a *subcomplex* if $\sigma \in \mathbf{L}$ implies that every face of σ also lies in $\mathbf{L}$. The union of the simplices in $\mathbf{L}$ is a closed subspace of P which is denoted by $|\mathbf{L}|$. With this notation we have $P = |\mathbf{K}|$.

LINEAR GRAPHS. The graphs of the previous unit are simple but important examples of 1-dimensional simplicial complexes; the only difference between the two concepts is that the latter may include some connected components which consist only of isolated vertices. One way of viewing this section and the next is to think of them as laying the foundations for effective study of similar objects in higher dimensions. As noted in Unit III, the study of such connected 1-dimensional complexes is the objective of *graph theory*, and the latter is significant for both its theory and applications; further study of this is well beyond the scope of the present course, but here are some written and electronic references:

J. A. Bondy and U. S. R. Murty. *Graph Theory: An Advanced Course*. Springer-Verlag, New York-*etc.*, 2008. ISBN: 1-846-28969-6.

G. Chartrand. *Introductory Graph Theory* [UNABRIDGED]. Dover Publications, New York, 1984. ISBN: 0-486-24775-9.

`http://en.wikipedia.org/wiki/Graph_theory`

`http://www.utm.edu/departments/math/graph/`

`http://www.math.fau.edu/locke/GRAPHTHE.HTM`

`http://www.math.uni-hamburg.de/home/diestel/books/`[*continue*]

`graph.theory/GraphTheoryIII.counted.pdf`

SIMPLICIAL COMPLEXES AND Δ-COMPLEXES. Our definition of simplicial complex is more restrictive than Hatcher's definition; this is explained on page 107 of Hatcher (see the third paragraph following Example 2.5). Each concept has its advantages and disadvantages. However, terms like Δ-complex or Δ-set are often also used for other mathematical constructions, and one should not assume that the meanings in other publications are "obviously" equivalent to the meaning in Hatcher.

Decompositions of prisms

The rectangular example has the following important generalization:

PROPOSITION 3. *Suppose that $A \subset \mathbb{R}^m$ is a simplex with vertices $\mathbf{v}_0, \cdots, \mathbf{v}_n$. Then $A \times [0,1] \subset \mathbb{R}^{m+1}$ has a simplicial decomposition with exactly $n+1$ simplices of dimension $n+1$.*

Proof. For each i between 0 and n let $\mathbf{x}_i = (\mathbf{v}_i, 0)$ and $\mathbf{y}_i = (\mathbf{v}_i, 1)$. We claim that the vectors

$$\mathbf{x}_0, \cdots, \mathbf{x}_i, \mathbf{y}_i \cdots, \mathbf{y}_n$$

are affinely independent and the corresponding simplices

$$\mathbf{x}_0 \cdots \mathbf{x}_i \mathbf{y}_i \cdots \mathbf{y}_n$$

(where $0 \leq i \leq n$) form a simplicial decomposition of $A \times [0,1]$.

An illustration for the case $n = 2$ is given in Figure 12 of `triangulations3.pdf`.

To prove affine independence, take a fixed value of i and suppose we have

$$\sum_{j<i} t_j\,\mathbf{x}_j \ \ + \ \ a\,\mathbf{x}_i \ + \ b\,\mathbf{y}_i \ \ + \ \ \sum_{j>i} t_j\,\mathbf{y}_j \ \ =$$

$$\sum_{j<i} t'_j\,\mathbf{x}_j \ \ + \ \ a'\,\mathbf{x}_i \ + \ b'\,\mathbf{y}_i \ \ + \ \ \sum_{j>i} t'_j\,\mathbf{y}_j$$

where the coefficients in each expression add up to 1; the summation will be taken to be zero if the limits reduce to $j < 0$ or $j > n$. If we view $\mathbb{R}^{m+1}$ as $\mathbb{R}^m \times \mathbb{R}$ and project down to $\mathbb{R}^m$ we obtain the equation

$$\sum_{j<i} t_j\,\mathbf{v}_j \ \ + \ \ (a+b)\,\mathbf{x}_i \ + \sum_{j>i} t_j\,\mathbf{v}_j \ \ = \ \ \sum_{j<i} t'_j\,\mathbf{v}_j \ \ + \ \ (a'+b')\,\mathbf{v}_i \ + \sum_{j>i} t'_j\,\mathbf{v}_j$$

and by the affine independence of the vectors $\mathbf{v}_k$ it follows that $t_j = t'_j$ if $j \neq i$ and also that $a + b = a' + b'$. On the other hand, if we project down to the second coordinate (the copy of $\mathbb{R}$), then we obtain

$$b \ + \sum_{j>i} t_j \ \ = \ \ b' \ + \sum_{j>i} t'_j$$

and since $t_j = t'_j$ for all j it follows that $b = b'$. Finally, since the sum of all the coefficients is equal to 1, the preceding observations imply that $1 - a = 1 - a'$, and therefore we also have $a = a'$. Therefore the vectors

$$\mathbf{x}_0, \ \cdots \ , \mathbf{x}_i, \mathbf{y}_i \ \cdots \ , \mathbf{y}_n$$

are affinely independent.

We shall next check that every point $(\mathbf{z}, u) \in A \times [0,1]$ lies in one of the simplices

$$\mathbf{x}_0 \ \cdots \ \mathbf{x}_i\mathbf{y}_i \ \cdots \ \mathbf{y}_n$$

listed above. Write $\mathbf{z} = \sum_j \ t_j\,\mathbf{v}_j$ where $t_j \geq 0$ for all j and $\sum \ t_j = 1$. It follows that $u \leq 1 = \sum_{j\geq 0} \ t_j$; let $i \leq n$ be the largest nonnegative integer such that $u \leq \sum_{j\geq i} \ t_j$. We claim that $(\mathbf{z}, u)$ lies in the simplex $\mathbf{x}_0 \ \cdots \ \mathbf{x}_i\mathbf{y}_i \ \cdots \ \mathbf{y}_n$. Let $b = \sum_{j\geq i} (t_j \ - \ u)$, and let $a = u - \left(\sum_{j>i} \ t_j\right) \ = \ t_i - b$. Then we have $a, \ b \geq 0$, and

$$(\mathbf{z}, u) \ \ = \ \ \sum_{j<i} t_j\,\mathbf{x}_j \ \ + \ \ a\,\mathbf{x}_i \ + \ b\,\mathbf{y}_i \ \ + \ \ \sum_{j>i} t_j\,\mathbf{y}_j$$

where all the coefficients are nonnegative and add up to 1.

To conclude the proof, we need to show that the intersection of two simplices as above is a common face. Suppose that $k < i$ and

$$(\mathbf{z}, u) \ \ \in \ \ \left(\mathbf{x}_0 \ \cdots \ \mathbf{x}_i\mathbf{y}_i \ \cdots \ \mathbf{y}_n\right) \ \cap \ \left(\mathbf{x}_0 \ \cdots \ \mathbf{x}_k\mathbf{y}_k \ \cdots \ \mathbf{y}_n\right) \ .$$

Then we must have

$$\sum_{j\leq i} p_j\,\mathbf{x}_j \ \ + \ \sum_{j\geq i} q_j\,\mathbf{y}_j \ \ = \ \sum_{j\leq k} p'_j\,\mathbf{x}_j \ \ + \ \sum_{j\geq k} q'_j\,\mathbf{y}_j$$

where all the coefficients are nonnegative and the coefficients on each side of the equation add up to 1. If we project down to $\mathbb{R}^m$ we obtain $p_j + q_j = p'_j + q'_j$ for all j (by convention, we take a coefficient to be zero if it does not lie in the corresponding summation as above). It follows immediately that $p_j = p'_j$ if $j < k$, while $p_j = q'_j$ if $k < j < i$ and $q_j = q'_j$ if $j > i$. Furthermore, if we project down to the last coordinate we see that

$$u \;=\; \sum_{j \geq i} q_j \;=\; \sum_{j \geq k} q'_k \;.$$

Since $q_j = q'_j$ if $j > i$, it follows that

$$q_i \;=\; \sum_{k \leq j \leq i} q'_j$$

and since all the coefficients are nonnegative, it follows that $q_i \geq q'_i$. On the other hand, we also have $q'_i = p'_i + q'_i = p_i + q_i$, and hence we conclude that $q_i = q'_i$ and $p_i = 0$. Applying the first of these, we see that

$$0 \;=\; \sum_{k \leq j < i} q'_j$$

and hence the nonnegativity of the coefficients implies that $q'_j = 0$ for all j such that $k \leq j < i$. We also know that $p'_j = 0$ for $j > k$, and therefore it follows that $p'_j + q'_j = 0$ when $k < j < i$ The equations $p_j + q_j = p'_j + q'_j$ and the nonnegativity of all terms now imply that $p_j = q_j = 0$ when $k < j < i$.

The conclusions of the preceding paragraph imply that the point $(\mathbf{z}, u)$ actually lies on the simplex

$$\mathbf{x}_0 \;\cdots\; \mathbf{x}_k \mathbf{y}_i \;\cdots\; \mathbf{y}_n$$

and since the latter is a common face of $\mathbf{x}_0 \;\cdots\; \mathbf{x}_i \mathbf{y}_i \;\cdots\; \mathbf{y}_n$ and $\mathbf{x}_0 \;\cdots\; \mathbf{x}_k \mathbf{y}_k \;\cdots\; \mathbf{y}_n$ it follows that the $(n+1)$-simplices

$$\mathbf{x}_0 \;\cdots\; \mathbf{x}_i \mathbf{y}_i \;\cdots\; \mathbf{y}_n$$

(where $0 \leq i \leq n$) form a simplicial decomposition of $A \times [0,1]$.■

COROLLARY 4. *If $P \subset \mathbb{R}^m$ is a polyhedron, then $P \times [0,1] \subset \mathbb{R}^{m+1}$ is also a polyhedron.*

Before discussing the proof of this we note one important special case.

COROLLARY 5. *For each positive integer m, the hypercube $[0,1]^m \subset \mathbb{R}^m$ is a polyhedron.*

Proof of Corollary 5 from Corollary 4. If $m = 1$ this follows because the unit interval is a 1-simplex; by Corollary 4, if the result is true for $m = k$ then it is also true for $m = k+1$. Therefore the result is true for all m by induction.■

Proof of Corollary 4. Let $\mathbf{K}$ be a simplicial decomposition for P, and let $\mathbf{K}^*$ be obtained from $\mathbf{K}$ by including all the faces of simplices in $\mathbf{K}$. Choose a linear ordering of the vertices in $\mathbf{K}^*$ (note that there are only finitely many). For each vertex $\mathbf{v}$ of $\mathbf{K}^*$, as before let $\mathbf{x} = (\mathbf{v}, 0)$ and $\mathbf{y} = (\mathbf{v}, 1)$. Then $P \times [0,1]$ is the union of all simplices of the form

$$\mathbf{x}_0 \;\cdots\; \mathbf{x}_i \mathbf{y}_i \;\cdots\; \mathbf{y}_n$$

where $\mathbf{v}_i < \mathbf{v}_{i+1}$ with respect to the given linear ordering of the vertices in $\mathbf{K}^*$, **and furthermore** the vertices $\mathbf{v}_i$ are the vertices of a simplex in $\mathbf{K}^*$. The set $P \times [0,1]$ is the union of these simplices by Proposition 3 and the fact that P is the union of the simplices $\mathbf{v}_0 \cdots \mathbf{v}_n$. The fact that these simplices form a simplicial decomposition will follow from the construction and the next result.

LEMMA 6. *Suppose that we have two polyhedra P_1 and P_2 in $\mathbb{R}^q$, and there exist simplicial decompositions $\mathbf{K}_1$ and $\mathbf{K}_2$ such that the following hold:*

(i) Both $\mathbf{K}_1$ and $\mathbf{K}_2$ are closed under taking faces of simplices.

(ii) The set $\mathbf{L}_1$ of all simplices in $\mathbf{K}_1$ contained in $P_1 \cap P_2$ equals the set $\mathbf{L}_2$ of all simplices in $\mathbf{K}_2$, and this collection determines a simplicial decomposition of $P_1 \cap P_2$.

Then $\mathbf{K}_1 \cup \mathbf{K}_2$ determines a simplicial decomposition of $P_1 \cup P_2$.

The hypothesis clearly applies to the construction in Proposition 3, so Corollary 4 indeed follows once we prove Lemma 6.■

Proof of Lemma 6. It follows immediately that $P_1 \cup P_2$ is the union of the points of the simplices in $\mathbf{K}_1 \cup \mathbf{K}_2$. Suppose now that we are given an intersection of two simplices in the latter. This intersection will be a common face if both simplices lie in either $\mathbf{K}_1$ or $\mathbf{K}_2$, so the only remaining cases are those where one simplex α lies in $\mathbf{K}_1$ and the other simplex β lies in $\mathbf{K}_2$.

We know that $\alpha \cap \beta$ is convex. Furthermore, by the hypotheses we know that $\alpha \cap \beta$ must be a union of simplices that are faces of both α and β. Therefore it follows that every point in $\alpha \cap \beta$ is a convex combination of the vertices which lie in $\alpha \cap \beta$, and consequently $\alpha \cap \beta$ is the common face determined by all vertices which lie in $\alpha \cap \beta$.■

GENERALIZATIONS — CONVEX LINEAR CELLS. [Also known as CONVEX POLYTOPES] These are closed bounded subsets of some $\mathbb{R}^n$ defined by a finite number of linear equations or inequalities. Note that sets defined by finite systems of this type are automatically convex. Prisms, simplices and cubes are obvious examples, but of course there are also many others, and there are numerous illustrations of more complicated examples in the file `convex-polytopes.pdf`.

For every such object, there is a finite set E of *extreme points* such that the cell is the set of all convex combinations of the extreme points; in other words, for each $\mathbf{x}$ in the cell and each extreme point $\mathbf{e}$ there are scalars $t_{\mathbf{e}}$ such that $t_{\mathbf{e}} \geq 0$, $\sum_{\mathbf{e}} t_{\mathbf{e}} = 1$, and $x = \sum_{\mathbf{e}} t_{\mathbf{e}} \mathbf{e}$. Visually these extreme points correspond to an informal notion of vertices for the examples of `convex-polytopes.pdf`, and in fact a basic theorem about convex linear cells states that *every example has a simplicial decomposition for which E is the set of vertices.* Proofs of this statement appear in [MunkresEDT] and the book by Hudson; we shall discuss some additional facts about such objects later in these notes.

The file `prismatoids.pdf` illustrates how simplicial decompositions of convex linear cells can be used to study a somewhat nontrivial problem in classical solid geometry which appeared in standard textbooks on that subject.

A few other easily stated but challenging problems on convex polytopes in $\mathbb{R}^3$ are contained in the files `wswGeometrytest*.pdf` where $* = 1, 2$, or 3 (the last two contain files with links in the first one), and solutions to these exercises using vector geometry are given in the companion file `wswvectorproofs.pdf`.

Finally, we note the following important fact:

Homeomorphism property of convex bodies. *Let P be a convex linear cell in $\mathbb{R}^n$ whose interior is nonempty. Then P is homeomorphic to the disk D^n such that the (point-set) boundary of P corresponds to S^{n-1}.*

In particular, it follows that all such subspaces are homeomorphic to the hypercube $[0,1]^n$ and to the standard simplex in $\mathbb{R}^n$ whose vertices are the zero vector and the n standard unit vectors.

A convex linear cell in $\mathbb{R}^n$ is said to be a **convex body** if it has a nonempty interior. The basic idea behind the proof of the homeomorphism property for convex bodies (which is describable as *radial projection from an interior point*) is fairly easy to grasp, but writing out all the details is a bit messy. Further information is given in the files `convexbodies.pdf` and `convexbodies2.pdf`.

DEFAULT HYPOTHESIS FOR SIMPLICIAL COMPLEXES. Unless specifically indicated otherwise, we shall assume that the set of simplices in a simplicial decomposition $\mathbf{K}$ is closed under taking faces. In order to justify this, we need to know that if $\mathbf{K}^*$ is obtained from $\mathbf{K}$ by adding all the faces of simplices in the latter, then the intersection of two simplices in $\mathbf{K}^*$ is a (possibly empty) common face. — To see this, suppose that α and β are simplices in $\mathbf{K}^*$, where α and β are faces of the simplices α' and β' in $\mathbf{K}$. If $\mathbf{x} \in \alpha \cap \beta$, then $\mathbf{x}$ is a convex combination of vertices in $\alpha' \cap \beta'$, and in fact these vertices must lie in both α and β. Since $\alpha \cap \beta$ is convex, it follows that $\alpha \cap \beta$ must be the simplex whose vertices lie in α and in β.■

IV.3 : Chain complexes and exact sequences

(**H**, §§2.1–2.2)

This section has two parts, the first of which is mainly geometric and the second of which is entirely algebraic. The starting point to find a generalization of the chain complex for a graph which works for 2-dimensional simplicial complexes; there should be groups of k-dimensional chains for $k = 0, 1, 2$ with boundary maps from k-dimensional chains to $(k-1)$-dimensional chains. If $k = 0$ or 1 then we can define everything exactly as in the case of graphs, and 2-dimensional chains will be integral linear combinations of the 2-simplices in the simplicial decomposition associated to the simplicial complex. Defining the boundary may be slightly less obvious, but it has a straightforward geometrical motivation; potential mathematical uses of 2-dimensional chains may be less obvious, so we shall indicate how they can be applied to proving versions of Green's Theorem which apply to very general sorts of closed regions in the plane (as opposed to the sorts of regions for which proofs are given in most of the usual textbooks for multivariable calculus). — In the second part of this section, we shall generalize some important properties of chain complexes for 1- and 2-dimensioanl simplicial complexes in a purely algebraic fashion and derive a few elementary consequences of the abstract definitions.

The chain complex of a 2-dimensional simplicial complex

Let $(P, \mathbf{K})$ be a dimensional simplicial complex of dimension ≤ 2. As in the case of graphs, the first step is to choose a partial ordering ω of the vertices. We shall use the usual notation $\mathbf{v} < \mathbf{w}$ to indicate that one vertex precedes another. For each integer $k = 0, 1, 2$ the *k-dimensional ordered simplicial chain group* of $(P, \mathbf{K})$, written $C_k(P, \mathbf{K}^\omega)$ is a free abelian group on all objects $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$, where $\mathbf{v}_0 < \ \cdots \ < \mathbf{v}_k$ and $\mathbf{v}_0, \ \cdots \ , \mathbf{v}_k$ are the vertices of a k-simplex in $\mathbf{K}$. By construction, it follows that $C_k(P, \mathbf{K}^\omega) = 0$ if $k < 0$ or $k > \dim \mathbf{K}$. The boundary homomorphism

$$d_k : C_k(P, \mathbf{K}^\omega) \ \longrightarrow \ C_{k-1}(P, \mathbf{K}^\omega)$$

is defined on free generators by the following formulas:

$$d_1(\mathbf{v}_0\mathbf{v}_1) \quad = \quad \mathbf{v}_1 - \mathbf{v}_0 \ , \qquad d_2(\mathbf{v}_0\mathbf{v}_1\mathbf{v}_2) \quad = \quad \mathbf{v}_0\mathbf{v}_1 + \mathbf{v}_1\mathbf{v}_2 - \mathbf{v}_0\mathbf{v}_2$$

The first boundary formula is generalizes the previous one for graphs, while the second is the algebraic chain for the edge path in the faces of the simplex $\mathbf{v}_0\mathbf{v}_1\mathbf{v}_2$ going first from $\mathbf{v}_0$ to $\mathbf{v}_1$ then from $\mathbf{v}_1$ to $\mathbf{v}_2$, and finally from $\mathbf{v}_2$ to $\mathbf{v}_0$.

One immediate consequence of the definitions is that $d_1 \circ d_0$ is the zero homomorphism, so that the kernel of d_1 contains the image of d_2; if we define $d_k = 0$ for $k \neq 0, 1$ (which is the only possible choice since $C_k = 0$ for $k \neq 0, 1$), then we obtain an identity of the form $d_k \circ d_{k+1}$ for all integers k and a similar sequence of inclusions

$$\text{Image}(d_{k+1}) \quad \subset \text{Kernel}(d_k) \ .$$

We then define the **k-dimensional homology group**

$$H_k(P, \mathbf{K}^\omega)$$

to be the chain group subquotients

$$\text{Kernel}(d_k)/\text{Image}(d_{k+1}) \ .$$

In analogy with simplicial chains on graphs, elements of the kernel of d_k are called **cycles** and elements of the image of d_{k+1} are called **boundaries**. Note that all 0-dimensional chains are cycles, and the only boundary chain in the top dimension is the zero chain, so that the H_0 is just the quotient $C_0/\text{Image}(d_1)$ and the top dimensional homology is just the subgroup of cycles.

We know that the union of the edges and vertices in a simplicial complex is a graph, and the simplicial complex is connected if and only if the this subcomplex is connected (note that every 2-simplex lies in the same component as each of its edges). By construction, the chain the groups of the simplicial complex and the edge-vertex subcomplex are the same in dimensions 0 and 1, so it follows immediately that *if the underlying space P is connected then $H_0(P, \mathbf{K}^\omega)$ is infinite cyclic and the class of an arbitrary vertex is a generator.* — A generalized version of these observations is one of the problems in `exercises03-2012.pdf`.

Another immediate consequence of the definitions is that if a simplicial complex $(P, \mathbf{K})$ of dimension ≤ 2 is written as a disjoint union of its components $(P_j, \mathbf{K}_j)$, then one has corresponding splittings

$$C_k(P, \mathbf{K}^\omega) \cong \oplus_j C_k(P, \mathbf{K}_j^\omega) \ , \qquad H_k(P, \mathbf{K}^\omega) \cong \oplus_j H_k(P, \mathbf{K}_j^\omega) \ .$$

Once again, there is a generalized version of this in `exercises03-2012.pdf`. If we combine this algebraic splitting theorem with the preceding paragraph, we conclude that *the homology group $H_0(P, \mathbf{K}^\omega)$ is a free abelian group on the set of components (equivalently, arc components) of P.*■

Examples. 1. Graph theory gives us many examples of 1-dimensional simplicial complexes with nontrivial 1-dimensional homology, and similar 2-dimensional examples can be constructed by a simple trick; namely, for each edge E in a graph attach a 2-simplex F_E which meets E in a single edge and does not meet any other points of the graph (see the drawing in `expansion.pdf` for a special case).

2. The boundary of a 3-simplex (in classical language, a tetrahedron) is the simplest example of a complex for which $H_2 \neq 0$. Specifically, if the vertices are given by $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ with the numerical ordering, then it is an elementary but slightly messy exercise to check that the chain

$$\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3 \; - \; \mathbf{v}_0\mathbf{v}_2\mathbf{v}_3 \; + \; \mathbf{v}_0\mathbf{v}_1\mathbf{v}_3 \; - \; \mathbf{v}_0\mathbf{v}_1\mathbf{v}_2$$

is a cycle (hence nonzero in homology) and every cycle is an integral multiple of this chain. In the next unit we shall justify these statements from a more systematic viewpoint.

3. For some 2-dimensional complexes the group H_1 contains nontrivial elements of finite order. The simplest example is the real projective plane, and an explanation of this is given in the file `rp2triangulation.pdf`. It is also possible to give examples for which H_1 is cyclic of order q for all integers $q \geq 3$, and we shall describe a few at a later point in the course.

Green's Theorem and 2-dimensional chains

Many books on multivariable calculus note that Green's Theorem can be generalized from the simple sorts of regions treated in all texts to a fairly general class of plane regions bounded by a finite number of reasonable closed curves. A typical example of such a discussion (from Marsden and Tromba) is posted in the file `greens-thm.pdf`. The file `green-chains.pdf` contains an illustrated and more detailed discussion which involves the concepts of 2-chains and their boundaries. In particular, if we are given a closed bounded **polygonal region** A in the plane which has a simplicial decomposition whose boundary is a union of finitely many closed curves, then the generalization of Green's Theorem to A depends upon the following result:

Existence of triangulating chains. *Suppose that we are given A as above, let $\mathbf{K}$ be a simplicial decomposition of A as union of 2-simplices, let ω be a linear ordering of the vertices, and denote the closed boundary curves of A by Γ_i where i runs from 1 to the number N of boundary components. Then there is a 2-chain Φ_A in $C_2(A, \mathbf{K}^\omega)$ such that the following hold:*

(i) The chain Φ_A is equal to $\sum \varepsilon_\sigma \sigma$, where the sum runs over the free generators σ corresponding to the 2-simplices of $\mathbf{K}$ and each coefficient ε_σ is ± 1.

(ii) The boundary chain $d_2(\Phi_A)$ is a sum of chains z_i coming from the subgroups $C_1(\Gamma_i) \subset C_1(A)$ such that z_i is a closed reduced edge path in Γ_i (so that z_i is a cycle expressible as a sum $\sum \varepsilon_\kappa \kappa$ where κ runs through all the edges in Γ_i and $\varepsilon_\kappa = \pm 1$).

This relates to the standard derivation of Green's Theorem in following way. The double integral over A is the sum of the double integrals over the 2-simplices into which A is decomposed, and the chain Φ_A is an algebraic model for this sum decomposition of the double integral. There is a partial explanation of the reasons for the $\pm$ signs in `green-chains.pdf`, and the need for signs traces back to the choice of ordering for the vertices.

The version of Green's Theorem in multivariable calculus texts implies that the result applies to each 2-simplex, showing that the double integrals over these pieces are equal to line integrals over the boundary curves and hence to a sum S of line integrals over various edges in $\mathbf{K}$ with suitable orientations. Some of these terms involve edges which are parts of the boundary curves Γ_i, and their contributions to S is a sum of $\pm$ the appropriate line integrals over the curves Γ_i. The key point in proving Green's Theorem for A and its boundary curves is that the line integrals over the remaining edges turn out to cancel in pairs, and the statement about $d_2(\Phi_A)$ is an abstract algebraic way of expressing this cancellation property. Intuitively speaking, the signs for the line integrals over the Γ_i's are chosen so that a unique outermost curve is taken in the counterclockwise sense and the remaining curve is taken in the clockwise sense, but we shall not attempt to make this precise.

As noted in `green-chains.pdf`, proving the existence of the chain Φ_A for an arbitrary polygonal region requires more machinery than we can develop in this course. If time permits, we shall post a file `green-chains2.pdf` which explains how one can prove the existence of Φ_A using material from Section 3.3 of Hatcher.

The algebraic framework for homology

This half of the section is basically algebraic, and at first the need for formally introducing some of the concepts may be unclear. However, the notions described here arise repeatedly in algebraic topology and other subjects.

Definition. Suppose we are given a diagram of the form

$$A \xrightarrow{f} B \xrightarrow{g} C$$

in which the objects are abelian groups (possibly with some additional structure) and the morphisms are abelian group homomorphisms (possibly preserving the extra structure). We shall say that the diagram is *exact* at B if the kernel of g is equal to the image of f.

More generally, if we are given a linear diagram such as

$$\cdots \longrightarrow Z \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow D \longrightarrow \cdots$$

we shall say that it is an *exact sequence* if it is exact at every object which is the domain of one morphism and the codomain of another.

Examples of exact sequences

There are many standard exact sequences in elementary algebra.

1. A *short exact sequence* is one having the form $0 \to A \to B \to C \to 0$. Exactness at A means that the kernel of $A \to B$ is the image of $0 \to A$, which is equivalent to saying that the map is injective. Similarly, exactness at C means that the kernel of $C \to 0$ is the image of $B \to C$, which is equivalent to saying that the map is surjective. The short exact sequence property is then equivalent to saying that $A \to B$ is injective, and C is isomorphic to the quotient of B by the image of A.

2. The *cokernel* of a homomorphism $f : A \to B$ is defined to be the quotient group $B/f[A]$. Given an arbitrary homomorphism $f : A \to B$, one then has the following *kernel – cokernel exact sequence*:

$$0 \longrightarrow \mathrm{Ker}(f) \longrightarrow A \longrightarrow B \longrightarrow \mathrm{Coker}(f) \longrightarrow 0$$

3. The following are elementary but extremely useful observations involving a module homomorphism $f : A \to B$:

3(a). f is 1–1 if and only if $0 \to A \to B$ is exact.

3(b). f is onto if and only if $A \to B \to 0$ is exact.

3(c). f is an isomorphism if and only if $0 \to A \to B \to 0$ is exact.

4. Let U be a connected open subset of $\mathbb{R}^2$, let $\mathbf{C}^\infty(U)$ denote the infinitely differentiable real valued functions on U, and let let $\mathbf{VF}(U)$ denote the infinitely differentiable (2-dimensional) vector fields on U in the sense of vector analysis. If we let $\mathbb{R} \to \mathbf{C}^\infty(U)$ denote the inclusion of the constant functions and take the gradient map from $\mathbf{C}^\infty(U)$ to $\mathbf{VF}(U)$, then it follows that the sequence $\mathbb{R} \to \mathbf{C}^\infty(U) \to \mathbf{VF}(U)$ is exact. Furthermore, if we take the map $\mathbf{VF}(U) \to \mathbf{C}^\infty(U)$ which sends a vector field $\mathbf{F} = (P, Q)$ to its "scalar curl" $Q_1 - P_2$, then the sequence $\mathbf{C}^\infty(U) \to \mathbf{VF}(U) \to \mathbf{C}^\infty(U)$ will be exact **provided** U is convex (or more generally star-shaped). — On the other hand, the second sequence is not exact if $U = \mathbb{R}^2 - \{\mathbf{0}\}$, for the previously described vector field on U with coordinate functions v/r and $-u/r$ has zero scalar curl but is not the gradient of any smooth function on U; this follows from Green's Theorem and the previous line integral calculation.

We can extend the preceding if U is a connected open set in $\mathbb{R}^3$ by considering the following sequence:

$$\mathbb{R} \quad \xrightarrow{\text{constants}} \quad \mathbf{C}^\infty(U) \quad \xrightarrow{\text{grad}} \quad \mathbf{VF}(U) \quad \xrightarrow{\text{curl}} \quad \mathbf{VF}(U) \quad \xrightarrow{\text{div}} \quad \mathbf{C}^\infty(U)$$

This is again exact at the left hand object $\mathbf{C}^\infty(U)$, and standard results in vector analysis imply that the kernel of the curl is contained in the image of the gradient, while the kernel of the divergence is contained in the image of the curl. If U is convex, then one can also show that the sequence is exact, but in general this is not true.

Examples. In particular, the vector field $\mathbf{F}$ on $\mathbb{R}^2 - \{\mathbf{0}\}) \times \mathbb{R}$ defined by

$$\mathbf{F}(u, v) \quad = \quad \left(\frac{v}{u^2 + v^2}, \frac{-u}{u^2 + v^2}, 0 \right)$$

satisfies $\nabla \times \mathbf{F} = \mathbf{0}$, but $\mathbf{F}$ cannot be expressed as a gradient over U. To see this, observe that the line integral of $\mathbf{F}$ over the counterclockwise unit circle in the xy-plane is equal to 2π, but if we could write $\mathbf{F} = \nabla g$ over U then the line integral over every closed curve in U would be zero. Similarly, the vector field $\mathbf{F}$ on $\mathbb{R}^3 - \{\mathbf{0}\}$ defined by $\mathbf{F}(\mathbf{x}) = |\mathbf{x}|^{-1}\mathbf{x}$ satisfies $\nabla \cdot \mathbf{F} = 0$, but $\mathbf{F}$ cannot be expressed as the curl of another vector field defined over all of U. One way of seeing this is to let Σ be the unit sphere in $\mathbb{R}^3$ and verify the surface integral computation

$$\iint_\Sigma \mathbf{F} \cdot d\mathbf{\Sigma} \quad = \quad 4\pi \ .$$

If $\mathbf{F}$ could be written as $\nabla \times \mathbf{G}$ over U then by Stokes' Theorem the surface integral would be zero (this is true because the sphere has no boundary curves).

Graded objects

The next concept is simple but indispensable.

Definition. Let A be a set, and let $\mathbf{C}$ be a category. A *graded object* over $\mathbf{C}$ with grading set A is a function X from A to the objects of $\mathbf{C}$. The object corresponding to a is generally denoted by X_a.

For example, one can construct a graded vector space over the reals with grading set the integers $\mathbb{Z}$ by taking $V_n = \mathbb{R}^n$ for $n \geq 0$ and setting V_n equal to the zero space if $n < 0$.

Another example is obtainable from an algebra of polynomials $\mathbb{R}[x_1, \cdots, x_n]$ in finitely many indeterminates. Here we can take V_n to be the set of all homogeneous polynomials of degree n together with the zero polynomial.

In this course we shall mainly be interested in *nonnegatively graded objects*, where the indexing set is $\mathbb{Z}$ and the object X_n is a suitable zero object if $n < 0$. For the categories of abelian groups or modules over some associative ring with unit, the meaning of zero object is obvious, and these categories are the only ones to be considered here.

Definition. If X and Y are nonnegatively graded objects over a category $\mathbf{C}$, then a *graded morphism of degree zero* or *grade preserving morphism* is a function f which assigns to each $n \in \mathbb{Z}$ a morphism $f_n : X_n \to Y_n$ in the category $\mathbf{C}$.

In the polynomial example, one can define a grade preserving homomorphism by sending the homogeneous polynomial $p(x_1, x_2, \cdots x_n)$ to the homogeneous polynomial $q(x_1, x_2, \cdots x_n) = p(x_1, x_1 + x_2, \cdots x_n)$. Obviously there are many other maps of this type.

The following observation is immediate:

PROPOSITION 1. *Given a category* $\mathbf{C}$*, the* $\mathbb{Z}$*-graded objects over* $\mathbf{C}$ *and graded morphisms of degree zero form a category.*■

In fact, this category has many structural properties that are direct analogs of properties that hold for $\mathbf{C}$ (for example, subobjects, quotient objects, direct products, and so on).

Abstrct chain complexes

We are now ready to formulate a purely algebraic version of the groups $C_*(P, \mathbf{K}^\omega)$ and their boundary homomorphisms.

Definition. Let $\mathbf{C}$ be the category of abelian groups and homomorphisms or a category of unital modules over an associative ring with unit R. A **chain complex** over $\mathbf{C}$ is a pair (A, d) consisting of a graded object A over $\mathbf{C}$ indexed by the integers together with morphisms $d_j : A_j \to A_{j-1}$ such that $d_{j-1} \circ d_j = 0$ for all j.

Here are a few simple examples.

1. Given an arbitrary graded module A, one can make it into a chain complex by taking $d_j = 0$ for all j. More generally, given a sequence of homomorphisms $f_{2j} : A_{2j} \to A_{2j-1}$, one can define a chain complex whose graded module is A with $d_{2j} = f_{2j}$ and $d_{2j-1} = 0$.
2. Suppose we are given three modules B, H, and B'. The we can define a chain complex with $A_2 = B$, $A_1 = B \oplus H \oplus B'$, and $A_0 = B'$ and $A_j = 0$ otherwise such that d_2 is injection into the first summand, d_1 is projection onto the third summand, and all other maps d_j must be zero (since either their domain or codomain is zero).
3. If U is open in $\mathbb{R}^2$, then one can obtain a chain complex from the previous sequence involving $\mathbf{C}^\infty(U)$ and $\mathbf{VF}(U)$, if one takes A_3 to be the reals, A_2 and A_0 to be the smooth functions, A_0 to be the vector fields, with morphisms given by inclusion of constants from A_3 to A_2, gradient from A_2 to A_1, scalar curl from A_1 to A_0, and with all other real vector spaces and morphisms equal to zero. Similarly, if U is open in $\mathbb{R}^3$ one has a system with A_4 equal to the reals, A_3 and A_0 equal to the smooth functions, A_2 and A_1 equal to the vector fields, with morphisms given by inclusion of constants from A_4 to A_3, gradient from A_3 to A_2, curl from A_2 to A_1, divergence from A_1 to A_0, and with all other real vector spaces and morphisms equal to zero.

The mapping d is often called a *differential*; the motivation is related to the preceding examples where the maps are given by some form of differentiation.

Definition. Given two chain complexes (A, d) and (B, e) a **chain map** $f : A \to B$ is a graded morphism such that for all integers j we have $e_j \circ f_j = f_{j-1} \circ d_j$. In other words, the following diagram is commutative:

$$\begin{array}{ccc} A_j & \xrightarrow{f_j} & B_j \\ \downarrow d_j & & \downarrow e_j \\ A_{j-1} & \xrightarrow{f_{j-1}} & B_{j-1} \end{array}$$

If the differential in a chain complex (A, d) is unambiguous from the context we shall frequently write A instead of (A, d).

The following consequences of the definitions are elementary but important.

PROPOSITION 2. *Given a category* $\mathbf{C}$*, the chain complexes over over* $\mathbf{C}$ *and chain complex morphisms form a category.*

PROPOSITION 3. *If* (A, d) *and* (B, e) *are chain complexes over* $\mathbf{C}$ *and* $f : (A, d) \to (B, e)$ *is a morphism of chain complex such that the mappings* f_j *are all isomorphisms, then the map* f^{-1} *of graded modules defined by* $(f^{-1})_j = f_j^{-1}$ *is also a chain map.*

Proof. To simplify the formulas let $g_j = f_j^{-1}$. The conclusion of the proposition is equivalent to the identities $d_j \circ g_j = g_{j-1} \circ e_j$ as maps from B_j to A_{j-1}.

Let $b \in B_j$ be arbitrary. Since f_{j-1} is injective, it follows that $d_j \circ g_j(b) = g_{j-1} \circ e_j(b)$ if and only if $f_{j-1} \circ d_j \circ g_j(b) = f_{j-1} \circ g_{j-1} \circ e_j(b)$. The left hand side is equal to

$$f_{j-1} \circ d_j \circ g_j(b) \quad = \quad e_j \circ f_j \circ g_j(b) \quad = \quad e_j(b)$$

by the defining identity for chain maps and the fact that g is inverse to f, and the latter fact also implies that the right hand side is equal to $e_j(b)$. Therefore it follows that the maps g_j satisfies the defining conditions for a chain map.■

As before, the category of chain complexes over $\mathbf{C}$ has many structural properties that are direct analogs of properties that hold for $\mathbf{C}$ and the category of graded objects over $\mathbf{C}$ (such as subobjects, quotient objects, direct products).

A few additional remarks about **subcomplexes** and **quotient complexes** of a chain complex seem worthwhile. If (A, d') is a chain subcomplex of (B, d), then it follows that $A_j \subset B_j$ for all j and that d_j maps A_j to A_{j-1} via d'_j. The quotient complex has chain groups which are quotients B_j/A_j and a differential d'' such that $d''_j[x] = [d_j x]$, where "$[\cdots]$" denotes the equivalence class in the appropriate quotient module. There is a well-defined map of this sort because d_j maps A_j into A_{j-1}.

HOMOLOGY OF CHAIN COMPLEXES. The natural sequel to the preceding discussion is to define the homology of an arbitrary chain complex (C_*, d_*) by the same type of formula used for the chain complex of a 2-dimensional simplicial complex:

$$H_k(C_*, d_*) \quad = \quad \text{Kernel}(d_k)/\text{Image}(d_{k+1})$$

Since this section has already considered a very wide range of concepts, we shall postpone further discsussion to the next unit.

V. Simplicial chain complexes and homology

The goal of this unit is to define a chain complex $C_*(P, \mathbf{K}^\omega)$ of abelian groups associated to a simplicial complex and a linear ordering ω of its vertices. This definition will extend the previous ones for complexes of dimension ≤ 2, and the homology groups of the chain complex $C_*(P, \mathbf{K}^\omega)$ will be called the **simplicial homology groups** of the complex (with respect to the given linear ordering of the vertices), and they will be denoted by $H_n(P, \mathbf{K}^\omega)$, where n runs through all the integers. As in the low-dimensional cases, both the chain and homology groups will be zero if n is negative.

Homology groups may be interpreted as furnishing an "algebraic picture" of the underlying topological space P. In order to develop the important properties of these groups it will be necessary to introduce some basic concepts and results from a subject called *homological algebra*, but efforts will be made to keep this to a minimum.

We have stated that the groups provide information about the underlying space P rather than the simplicial complex $(P, \mathbf{K})$ and vertex ordering ω because *these groups turn out to depend only upon P itself*. In most textbooks this is done by defining objects called **singular chain complexes** for which the independence of $\mathbf{K}$ and ω are true by simple formal properties of the chain complexes and the fact that they are defined for arbitrary topological spaces. We shall not follow this standard approach for two reasons:

1. Verifying the fundamental properties of singular chain complexes and singular homology groups requires more time than is left in the course, and it seems more important to illustrate how these homology groups are used to study some geometric and topological questions of independent interest.

2. The motivations for several of the constructions and proofs require a more extensive treatment of simplicial complexes than we can give within the time constraints of this course.

Our alternative approach in the next unit will be to give a set of axioms for a singular homology theory; subsequent parts of the course will depend formally on this axiomatic description, but the basic methods and proofs will be the same as if we had taken the more standard approach. These issues will be discussed further at the beginning of the next unit.

V.1 : Simplicial chains and homology

(**H**, §§1.A, 2.1)

The first part of this section continues the discussion of abstract chain complexes from the previous unit, and in the second part we shall use some of these general considerations when we define and study chain complexes associated to simplicial complex.

If (A, d) is a chain complex, then the condition $d_j \circ d_{j+1}$ implies that the kernel of d_j (the submodule of cycles) contains the image of d_{j+1} (the submodule of boundaries). The sequence determined by the chain complex is exact at A_j if and only if these two submodules are equal. One can view homology groups as measuring the extent to which a chain complex is not an exact sequence.

Formal Definition. Let (A, d) be a chain complex. The j^{th} (or j-dimensional) *homology group* $H_j(A) = H_j(A, d)$ is equal to the quotient module

$$(\text{Kernel } d_j)/(\text{Image } d_{j+1}) \ .$$

By the definitions, the sequence of morphisms determined by a chain complex (A, d) is exact at A_j if and only if $H_j(A) = 0$.

Computation of the homology groups for most examples in the previous unit is fairly straightforward.

1. If we take an arbitrary graded module A and make it into a chain complex by taking $d_j = 0$ for all j, then $H_j(A, 0) = A_j$. If we are given a sequence of homomorphisms $f_{2j} : A_{2j} \to A_{2j-1}$ and define a chain complex whose graded module is A with $d_{2j} = f_{2j}$ and $d_{2j-1} = 0$, then $H_{2j}(A) = \text{Kernel } d_{2j}$ and $H_{2j-1}(A) = A_{2j-1}/\text{Image } d_{2j}$.

2. In Example 2 from the previous section, the homology is zero if U is a convex open subset of $\mathbb{R}^2$ or $\mathbb{R}^3$.

The next result is fairly simple to prove but absolutely fundamental.

THEOREM 1. *If $f : (A, d^A) \to (B, d^B)$ is a map of chain complexes, then there are unique homomorphisms $f_* : H_k(A) \to H_k(B)$ such that if $u \in H_k(A)$ is represented by $z \in A_q$, then $f_*(u)$ is represented by $f_q(z)$. Furthermore, if f is an identity chain map then f_* is also the identity, and if $g : (B, d^B) \to (C, d^C)$ is another chain map, then $(g \circ f)_* = g_* \circ f_*$.*

The second sentence of the theorem implies that the construction sending f to f_* defines a covariant functor from chain complexes to graded modules. Thus the following is immediate.

COROLLARY 2. *In the setting above, if f is an isomorphism then so is f_*.*■

Proof of Theorem 1. The condition in the first sentence of the theorem implies uniqueness, and the formula for f_* immediately yields the functoriality properties in the second sentence. Thus everything reduces to showing that there is indeed a homomorphism f_* with the asserted property.

First of all, we must check that $f_q(z)$ is a cycle if z is a cycle. To see this note that

$$d^B_q \circ f_q(z) \quad = \quad f_{q-1} \circ d^A_q(z) \quad = \quad f_{q-1}(0) \quad = \quad 0$$

so there is no problem here. Next, we need to check that if z and w represent the same class in A_q, then $f_q(z)$ and $f_q(w)$ represent the same class in B_q. However, it z and w represent the same class, then $z - w = d_{q+1}(y)$, and hence we have

$$f_q(z) - f_q(w) \quad = \quad f_q(z - w) \quad = \quad f_q \circ d^A_{q+1}(y) \quad = \quad d^B_q \circ f_{q+1}(y)$$

so that the images of z and w represent the same class in $H_q(B)$. The identities $f_*(u_1 + u_2) = f_*(u_1) + f_*(u_2)$ and $f_*(r \cdot u) = r \cdot f_*(u)$ now follow immediately from the definition of f_* and the standard choices of representatives for $u_1 + u_2$ and $r \cdot u$.■

Chain complexes associated to simplicial complexes

Our next objective is to define chain complexes and homology groups for simplicial complexes of arbitrary dimension.

One central feature of algebraic topology is that there are usually several different chain complexes which yield the same homology groups, each of which has its own advantages and disadvantages. Our choice involves relatively small chain complexes.

Definition. Suppose that $(P, \mathbf{K})$ is a simplicial complex, and choose a linear ordering ω for the vertices of $\mathbf{K}$; we shall use the usual notation $\mathbf{v} < \mathbf{w}$ to indicate that one vertex precedes another. For each integer k, the *k-dimensional ordered simplicial chain group* of $(P, \mathbf{K}, \omega)$, written $C_k(P, \mathbf{K}^\omega)$, is a free abelian group on all formal symbols of the form $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$, where $\mathbf{v}_0 < \ \cdots \ < \mathbf{v}_k$. By construction, it follows that $C_k(P, \mathbf{K}^\omega) = 0$ if $k < 0$ or $k > \dim \mathbf{K}$. The boundary homomorphism

$$d_k : C_k(P, \mathbf{K}^\omega) \ \longrightarrow \ C_{k-1}(P, \mathbf{K}^\omega)$$

is defined on free generators by the formula

$$d_k(\, \mathbf{v}_0 \ \cdots \ \mathbf{v}_k \,) \quad = \quad \sum_{j=0}^{n} (-1)^j \, \mathbf{v}_0 \ \cdots \ \widehat{\mathbf{v}}_i \ \cdots \ \mathbf{v}_k$$

where $\widehat{\mathbf{v}}_i$ means that $\mathbf{v}_i$ is omitted; by the definition of free generators, it follows that there is a unique extension to the group $C_k(P, \mathbf{K}^\omega)$.

Since our purpose is to define homology groups, presumably we want to verify that the preceding data define a chain complex. For this purpose it will be helpful to introduce some additional definitions.

If $k > 0$ and $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ is as above, then the $i^{\rm th}$ **face operator** $\partial_i^{[k]}(\, \mathbf{v}_0 \ \cdots \ \mathbf{v}_k \,)$ is given by

$$\mathbf{v}_0 \ \cdots \ \widehat{\mathbf{v}}_i \ \cdots \ \mathbf{v}_k \ .$$

Frequently we shall suppress the superscript $[k]$ to simplify notation. The following identity for iterated faces is elementary but fundamentally important:

LEMMA 1. *If $k - 1 \geq j \geq i$, then $\partial_j^{[k-1]} \circ \partial_i^{[k]} = \partial_i^{[k-1]} \circ \partial_{j+1}^{[k]}$.*

The identity is true because the result of applying both composites to $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ is given by deleting $\mathbf{v}_i$ and $\mathbf{v}_{j+1}$.■

With Lemma 1, it is fairly easy to prove that the boundary maps d_k define a chain complex.

THEOREM 2. *In the setting above we have $d_{k-1} \circ d_k = 0$.*

The proof of this result is given in Lemma 2.1 on pages 105–106 of Hatcher and also in the course directory file `chainboundary.pdf`.■

We now define the k-dimensional simplicial homology group of $(P, \mathbf{K})$ for ordered simplicial chains, also called the k-dimensional ordered simplicial homology group and denoted by

$$H_k(P, \mathbf{K}^\omega)$$

to be the k-dimensional homology of the chain complex $C_*(P, \mathbf{K}^\omega)$, where the differential or boundary is given as above.

The preceding definition depends not only upon the choice of a simplicial decomposition but also upon choosing a linear ordering of the vertices. It turns out that the homology groups only depend upon the underlying topological space P, and we shall address this issue in the next unit. For the time being, we shall give a reference for a proof that different vertex orderings determine isomorphic homology groups. This can be found in pages 51–60 of the notes

`http://math.ucr.edu/~res/math246A/algtopnotes2010.pdf`

starting with the paragraph titled "Second definition" and continuing through the end of the proof of Theorem 6. However, we shall not need the independence result in this unit.

The following is an immediate consequence of the constructions in Section IV.3:

THEOREM 3. *Let $(P, \mathbf{K})$ be a simplicial complex, let $(Q, \mathbf{L})$ be a subcomplex, let $i : (Q, \mathbf{L}) \to (P, \mathbf{K})$ be the inclusion mapping, and let ω be a linear ordering for the vertices of $\mathbf{K}$ (and hence also for the vertices of $\mathbf{L}$). Then the map of graded objects*

$$i_\# : C_*(Q, \mathbf{L}^\omega) \longrightarrow C_*(P, \mathbf{K}^\omega)$$

sending a free generator $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ in $C_k(Q, \mathbf{L}^\omega)$ to its counterpart in $C_k(P, \mathbf{K}^\omega)$ is a chain complex inclusion and hence induces homomorphisms i_* of homology groups from $H_*(Q, \mathbf{L}^\omega)$ to $H_*(P, \mathbf{K}^\omega)$. Both the chain inclusions and homology maps define covariant functors on the category of simplicial complexes and subcomplex inclusions.

This is true because the boundary chain associated to a simplex $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ in $\mathbf{L}$ is an integral linear combination of simplices in $\mathbf{L}$.■

Examples. It is important to recognize that *the mapping in homology i_* is usually* not *injective.* — We shall see many examples of this, but for the time being it is enough to let $P = \Delta_2$ and let Q be the subcomplex of all edges. In this case $H_1(Q, \mathbf{L}^\omega)$ is infinite cyclic but $H_1(P, \mathbf{K}^\omega)$ is trivial; in fact, the boundary of the generator $\mathbf{e}_0\mathbf{e}_1\mathbf{e}_2$ is a 1-chain in $C_1(Q, \mathbf{L}^\omega)$ which generates $H_1(Q, \mathbf{L}^\omega)$, but the corresponding homology class in $H_1(P, \mathbf{K}^\omega)$ must be trivial. Informally speaking, the 1-chain is a cycle and a boundary in P, and hence it must also be a cycle in Q, but it is not a boundary in Q.

V.2 : Examples and special cases

(**H**, §§2.1, 2.2)

The first observation is an immediate consequence of the definitions.

PROPOSITION 0. *If P consists of a single point and $\mathbf{K}$ is the associated (trivial) simplicial decomposition of P as a single vertex, then $H_k(P, \mathbf{K}^\omega) \cong \mathbb{Z}$ if $k = 0$ and $H_k(P, \mathbf{K}) \cong 0$ if $k \neq 0$, where ω is the only possible vertex ordering.*

This is true because $C_k(P, \mathbf{K}^\omega) \cong \mathbb{Z}$ if $k = 0$ and $C_k(P, \mathbf{K}) \cong 0$ if $k \neq 0$ by construction, so that the boundary mappings must be zero, and if $(A_*, 0)$ is a chain complex with zero boundary mappings then $H_k(A) \cong A_k$ for all k (why?).■

Since a one point space is a 0-dimensional simplex, the proposition describes the homology groups of such a simplex. The next step is to prove that the simplicial homology groups of an arbitrary simplex are all isomorphic to the homology groups of a point. This will require an algebraic digression.

Acyclic complexes

Definition. An *augmented chain complex* over a ring R consists of a chain complex (C_*, d) and a homomorphism $\varepsilon : C_0 \to R$ (the augmentation map) such that ε is onto and $\varepsilon \circ d_1 = 0$.

All of the simplicial chain complexes defined above have canonical augmentations given by sending 0-chains of the form $\sum n_{\mathbf{v}} \, \mathbf{v}$ to the corresponding integers $\sum n_{\mathbf{v}}$.

Definition. A simplicial complex is said to be *acyclic* ("has no nontrivial cycles") if $H_j(P, \mathbf{K}) = 0$ for $j \neq 0$ and $H_0 \cong \mathbb{Z}$, with the generator in homology represented by an arbitrary vertex generator of $C_0(P, \mathbf{K})$.

There is a simple geometric criterion for a simplicial chain complex to be acyclic.

Definition. A simplicial complex $(P, \mathbf{K})$ is said to be *star shaped* with respect to some vertex $\mathbf{v}$ in $\mathbf{K}$ if for each simplex A in $\mathbf{K}$ either $\mathbf{v}$ is a vertex of A or else there is a simplex B in $\mathbf{K}$ such that A is a face of B and $\mathbf{v}$ is a vertex of B. If $(P, \mathbf{K})$ is star shaped with respect to $\mathbf{v}$, then a linear vertex ordering ω is said to be star shaped with respect to $\mathbf{v}$ if $\mathbf{v}$ is minimal in the ordering ω. We shall also say that $\mathbf{v}$ is the star vertex of $\mathbf{K}$ with the ordering ω.

Some examples are depicted in `starshaped.pdf`. One particularly important example for the time being is the standard simplex Δ_n with its standard decomposition; in this case the complex is star shaped with respect to very vertex.

Note that if P is star shaped with respect to some vertex $\mathbf{v}$, then $\{\mathbf{v}\}$ is a strong deformation retract of P and hence P is contractible. An explicit homotopy from the identity on P to the constant map is given by the straight line homotopy $H(x, t) = (1 - t)x + t\,\mathbf{v}$, whose image always lies in P.

THEOREM 1. *If the simplicial complex $(P, \mathbf{K})$ is star shaped with respect to some vertex $\mathbf{v}$ and the linear vertex ordering ω is star shaped with respect to this vertex, then the augmented chain complex $C_*(P, \mathbf{K}^\omega)$ is acyclic, and the map $i_* : H_*(\{\mathbf{v}\}, \mathbf{L}) \to H_*(P, \mathbf{K})$ induced by $\{\mathbf{v}\} \subset P$ is an isomorphism.*

Before proving this, we shall use it to extend the example at the end of the preceding section. Given a simplicial complex $(P, \mathbf{K})$, we have seen that it is isomorphic to a subcomplex of a simplex A whose vertices are the same as those of $\mathbf{K}$, and of course we can take this ordering for the vertices of A. If $\mathcal{F}_A$ is the simplicial decomposition of A given by its faces and P is connected, then the proposition implies that the homology maps $H_*(P, \mathbf{K}^\omega) \to H_*(A, \mathcal{F}_A^\omega)$ are isomorphisms in dimension zero and are zero homomorphisms in all other dimensions. In particular, if $H_k(P, \mathbf{K}^\omega) \neq 0$ for some $k > 0$, then this group is not a subgroup of $H_k(A, \mathcal{F}_A^\omega) = 0$.

Proof. We need to define a modified version of the simplicial chain complex $C_*(P, \mathbf{K}^\omega)$. As suggested by the following passage from `morgan-lamberson.pdf`, this sort of thing happens frequently in homology theory and sometimes makes the subject seem like a real-life version of the film *Groundhog Day* (see `http://www.imdb.com/title/t0107048`).

> The main trouble with algebraic topology is that there are many different approaches to defining the basic ... homology ... groups. Each approach brings with it a fair amount of required technical baggage ... one must pay a fairly high price ... as one slogs through the basic constructions and proves the basic results. Furthermore, possibly the most striking feature of the subject, the interrelatedness (and often equality) of the theories ... requires even more machinery.

The new chain complex $C_*^+(P, \mathbf{K}^\omega)$ will contain $C_*(P, \mathbf{K}^\omega)$ along with additional free generators $\mathbf{v}\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ in dimension $k+1$ for each simplex generator $\mathbf{v}_0 \ \cdots \ \mathbf{v}_k$ in dimension k such that $\mathbf{v}_0$ is the minimal star vertex $\mathbf{v}$. The boundary mapping on the extra generators is defined exactly as before in terms of formal faces for these generators, and this yields the desired chain complex. The 0-dimensional chain groups of both complexes are isomorphic, and one can check that the augmentation for $C_*(P, \mathbf{K}^\omega)$ is also an augmentation for $C_*^+(P, \mathbf{K}^\omega)$ (the only extra free generators in dimension 1 have the form $\mathbf{v}\mathbf{v}$ and these are all cycles, so we still have $\varepsilon \circ d_1 = 0$ in the larger chain complex).

The chain complex constructed in the previous paragraph has the following important property:

LEMMA 2. *There is a chain map*

$$\rho : C_*^+(P, \mathbf{K}^\omega) \to C_*(P, \mathbf{K}^\omega)$$

whose restriction to $C_*(P, \mathbf{K}^\omega)$ is the identity.

Proof of Lemma 2. Defining ρ on chains is easy. Send each free generator in the subcomplex to itself and send the extra free generators to zero. If j is the subcomplex inclusion, then it follows immediately that $\rho \circ j$ is the identity, and to complete the proof we only need to verify that ρ is a chain map; this amounts to showing that $\rho \circ d_{k+1}(\mathbf{v}\mathbf{v_0} \ \cdots \ \mathbf{v}_k = 0$ if $\mathbf{v} = \mathbf{v}_0$. By construction we know that $d_{k+1}(\,\mathbf{v}\mathbf{v}_0 \ \cdots \ \mathbf{v}_k\,)$ is equal to

$$\mathbf{v}\mathbf{v}_1 \ \cdots \ \mathbf{v}_k \ - \ \mathbf{v}_0 \ \cdots \ \mathbf{v}_k \ + \ \sum_{i=1}^{k} (-1)^{i+1}\, \mathbf{v}\mathbf{v}_0 \ \cdots \ \widehat{\mathbf{v}}_i \ \cdots \ \mathbf{v}_k \ .$$

The first two terms cancel because $\mathbf{v} = \mathbf{v}_1$ and the remaining summation is an integral linear combination of the extra generators, which are mapped to zero under ρ, and thus we have $\rho \circ (\mathbf{v}\mathbf{v_0} \ \cdots \ \mathbf{v}_k = 0$, which is what we needed to prove the lemma.■

Proof of Theorem 1 resumed. The underlying motivation is that the space P is contractible, and we want to define an algebraic analog of the contracting homotopy described before the statement of the theorem.

Formally, we first define a map of graded abelian groups $\eta : C_*(P, \mathbf{K}) \to C_*(P, \mathbf{K})$ such that η_q in dimension q is zero if $q \neq 0$ and η_0 sends a chain y to $\varepsilon(y)\,\mathbf{v}$. Then η is a chain map because $\varepsilon \circ d_1 = 0$.

We next define the *contracting chain homotopy homomorphisms* $D_q : C_q(P, \mathbf{K}) \to C_{q+1}(P, \mathbf{K})$ such that

$$d_{q+1} \circ D_q \quad = \quad \text{identity} \ - \ d_q \circ D_{q-1}$$

if q is positive and

$$d_1 \circ D_0 \quad = \quad \text{identity} \ - \ \eta_0$$

on C_0. We do this by setting $D_q(\mathbf{x}_0 \ \cdots \ \mathbf{x}_q) = \rho(\mathbf{v}\mathbf{x}_0 \ \cdots \ \mathbf{x}_q)$ and taking the unique extension which exists since the classes $\mathbf{x}_0 \ \cdots \ \mathbf{x}_q$ are free generators for C_q. Elementary calculations show that the mappings D_q satisfy the conditions given above (see the file `chaincontraction.pdf` for the details).

To see that $H_q(P, \mathbf{K}) = 0$ if $q > 0$, suppose that $d_q(z) = 0$. Then the first formula implies that $z = d_{q+1} \circ D_q(z)$. Therefore $H_q = 0$ if $q > 0$. On the other hand, if $z \in C_0$, then the second formula implies that $d_1 \circ D_0(z) = z - \varepsilon(z)\,\mathbf{v}$. Furthermore, since $\varepsilon \circ d_1 = 0$ and $d_0 = 0$, it follows that

(i) the map ε passes to a homomorphism from H_0 to $\mathbb{Z}$,

(ii) since $\varepsilon(\mathbf{v}) = 1$ this homomorphism is onto,

(iii) the multiples of the class $[\mathbf{v}]$ give all the classes in H_0.

Taken together, these imply that $H_0(P, \mathbf{K}^\omega) \cong \mathbb{Z}$, and it is generated by $[\mathbf{v}]$. This completes the computation of $H_*(P, \mathbf{K}^\omega)$.■

The next step is to find examples with nonzero homology groups in arbitrary positive dimensions. This is relatively easy now.

THEOREM 3. *If $n \geq 1$ and $\partial\Delta_{n+1}$ is the subcomplex given by the union of the simplices in the set $\mathcal{F}$ of proper faces (geometrically, the boundary of the simplex) and ω denotes the standard ordering of the unit vectors in $\mathbb{R}^{n+2}$, then $H_q(\partial\Delta_{n+1}, \mathcal{F}^\omega)$ is isomorphic to $\mathbb{Z}$ if $q = 0$ or n, and $H_q(\partial\Delta_{n+1}, \mathcal{F}^\omega)$ is zero otherwise.*

Proof. Consider the chain complex inclusion

$$C_*(\partial\Delta_{n+1}, \mathcal{F}^\omega) \quad \subset \quad C_*(\Delta_{n+1}, \mathcal{F}^\omega_+)$$

where $\mathcal{F}_+$ consists of all the faces of the simplex including itself (the only face not in $\mathcal{F}$). This subcomplex inclusion map is bijective except in dimension $n+1$, and this implies that both chain complexes have the same cycles and boundaries in all dimensions except $n+1$ and n. In these dimensions we have the following exceptional behavior:

(i) The cycles in dimension n are the same for both complexes.

(ii) There are no nonzero cycles or boundaries for either complex in dimension $n+1$ because d_{n+1} is injective on the larger complex and there are no nonzero $(n+1)$-chains in the smaller complex.

(iii) The cycles in dimension n form an infinite cyclic group generated by the class of

$$d(\mathbf{e}_0 \ \cdots \ \mathbf{e}_{n+1})$$

in the chain complex for Δ_{n+1} (because the latter is acyclic). Although this class is an n-chain for the subcomplex $\partial\Delta_{n+1}$, it cannot be a boundary in the chain complex for the subcomplex because the latter has no nonzero$(n+1)$-chains.

The equality of the chain groups in almost all dimensions implies that the q-dimensional homology groups of both chain complexes are the same provided $q \neq n, n+1$. Points $(i) - (iii)$ combine to show that $H_{n+1}(\partial\Delta_{n+1}, \mathcal{F}^\omega) = 0$ and $H_n(\partial\Delta_{n+1}, \mathcal{F}^\omega)$ must be isomorphic to $\mathbb{Z}$ such that the chain

$$\sum_{j=0}^{n} (-1)^j \, \mathbf{v}_0 \ \cdots \ \widehat{\mathbf{v}}_i \ \cdots \ \mathbf{v}_n$$

represents a generator.■

In particular, it follows that for each positive integer n there is a simplicial complex $(P, \mathbf{K})$ such that $H_n(P, \mathbf{K}^\omega)$ is nonzero.

The computation for the homology of $\partial\Delta_{n+1}$ indicates that sometimes one can compute the homology groups of a complex if one knows something about the homology groups of some subcomplex and vice versa. In the next two section we shall present systematic methods for doing similar computations in more general situations.

V.3 : Relative groups and exactness properties

(**H**, §2.1)

The final result of the previous section leads naturally to the following general question:

> *If $i : A \to B$ defines an inclusion of chain complexes, how can we analyze the kernel and cokernel of the homology maps $i_* : H_*(A) \to H_*(B)$ in a reasonably effective manner?*

As in many other instances, the answer to this question involves some additional constructions. Let $A \subset B$ be a chain complex inclusion, and consider the quotient complex B/A; let $i : A \to B$ denote the inclusion map, and let $j : B \to A/B$ denote the projection. We then have the following result:

PROPOSITION 1. *Let $i : A \to B$ and $j : A \to A/B$ be injection and projection maps of chain complexes as above. Then for each k there is a homomorphism $\partial : H_k(B/A) \to H_{k-1}(A)$ defined as follows: If $u \in H_k(B/A)$ and $x \in B_k$ is such that $j(x)$ represents u, then $\partial(u)$ is represented by $y \in A_{k-1}$ such that $i(y) = d(x)$. Furthermore, if we are given a second pair $i' : A' \to B'$ and $j' : B' \to B'/A'$ as above and a chain map $f : B \to B'$ such that f maps A to A' by a chain map g and $h : B/A \to B'/A'$ is the map given by passage to quotients, then the corresponding homomorphisms ∂ and ∂' satisfy $g_* \circ \partial = \partial' \circ h_*$.*

Proof. First of all, we should check that the definition makes sense. The first step in doing so is to verify that if we are given x there is always a suitable choice of y. In general the class x need not be a cycle, but we know that $j(x)$ is a cycle representing u, and therefore $0 = d \circ j(x) = j \circ d(x)$, which means that $d(x) = i(a)$ for some a. This element is a cycle; we know that $d(a) = 0$ if and only if $i \circ d(a) = 0$, and since $i \circ d(a) = d \circ i(a) = d \circ d(x) = 0$, it does follow that $d(a) = 0$ as required.

Next, we need to check that the construction is well defined when one passes to homology. Suppose that $j(x)$ and $j(x')$ represent the same class in $H_k(B/A)$. It then follows that $j(x - x')$ is a boundary, which means there is some $w \in B_{k+1}$ such that $d(w) - (x - x')$ lies in A, which is the image of i. Express the difference element as $i(z)$; then we have

$$i(dz) \quad = \quad d(iz) \quad = \quad d\big(\, d(w) - (x - x')\,\big) \quad = \quad d(x') - d(x)$$

so that $d(x) = i(a)$ and $d(x') = i(a')$ imply that $a' - a = d(z)$.

Next, we need to check that ∂ is a module homomorphism. Given classes u and u' represented by x and x', it follows that $x + x'$ represents $u + u'$, while $d(x) = i(a)$ and $d(x') = i(a')$ imply $d(x + x') = i(a + a')$. Thus $a + a'$ represents $u + u'$, showing that ∂ is additive. If $r \in R$, then similar considerations show that $\partial(r \cdot u)$ is represented by $r \cdot a$, and therefore ∂ is compatible with scalar multiplication.

Finally, suppose we have chain maps as described in the proposition, let $u \in H_k(B/A)$, and let $x \in B_k$ be such that $j(x)$ represents u. Then a representative for $g_*\partial(u)$ is given by $g(a)$, where $ia = dx$, while a representative for $\partial' h_*(u)$ is given by z such that $i'(z) = d'f(x)$. The right hand side equals $f \circ d(x) = f \circ i(a) = i' \circ g(a)$, and therefore we see that $z = g(a)$, which means that $g_*\partial(u) = \partial' h_*(u)$ as desired.■

We may now state and prove the following basic result:

THEOREM 2. (Long Exact Homology Sequence Theorem — Algebraic Version). *Let $i : A \to B$ and $j : A \to A/B$ be injection and projection maps of chain complexes as above. Then there is a long exact sequence of homology groups as follows:*

$$\cdots \quad H_{k+1}(B/A) \xrightarrow{\partial} H_k(A) \xrightarrow{i_*} H_k(B) \xrightarrow{j_*} H_k(B/A) \xrightarrow{\partial} H_{k-1}(A) \quad \cdots$$

This sequence extends indefinitely to the left and right. Furthermore, if we are given chain maps f, g and h as in Proposition 2, then we have the following commutative diagram in which the two rows are exact:

$$\begin{array}{ccccccccccc}
\cdots & H_{k+1}(B/A) & \xrightarrow{\partial} & H_k(A) & \xrightarrow{i_*} & H_k(B) & \xrightarrow{j_*} & H_k(B/A) & \xrightarrow{\partial} & H_{k-1}(A) & \cdots \\
\cdots & \downarrow h_* & & \downarrow g_* & & \downarrow f_* & & \downarrow h_* & & \downarrow g_* & \cdots \\
\cdots & H_{k+1}(B'/A') & \xrightarrow{\partial'} & H_k(A') & \xrightarrow{i'_*} & H_k(B') & \xrightarrow{j'_*} & H_k(B'/A') & \xrightarrow{\partial'} & H_{k-1}(A') & \cdots
\end{array}$$

A proof of this theorem appears on page 117 of Hatcher and in the course directory file `longexact.pdf`.■

Application to simplicial complexes

In order to apply the preceding algebraic results, we need to define *relative homology groups* associated to a *simplicial complex pair*

$$\big(\ (P, \mathbf{K}),\ (Q, \mathbf{L})\ \big)$$

consisting of a simplicial complex $(P, \mathbf{K})$ and a subcomplex $(Q, \mathbf{L})$. To simplify notation, we shall usually denote such a pair by $(\mathbf{K}, \mathbf{L})$. Furthermore, in order to simplify notation we shall frequently suppress the superscript associated to a linear vertex ordering ω, tacitly assuming that there is some default choice unless indicated otherwise.

Definition. In the setting above the *relative simplicial chain groups*, denoted by $C_*(\mathbf{K}, \mathbf{L})$, are given by the corresponding quotient complex

$$C_*(\mathbf{K})/C_*(\mathbf{L})\ .$$

If we are given a (commutative) diagram of simplicial complex inclusions

$$\begin{array}{ccc}
\mathbf{L} & \longrightarrow & \mathbf{K} \\
\downarrow & & \downarrow \\
\mathbf{L}' & \longrightarrow & \mathbf{K}'
\end{array}$$

then there are canonical (and functorial) chain maps

$$\varphi : C_*(\mathbf{K})/C_*(\mathbf{L}) \longrightarrow C_*(\mathbf{K}')/C_*(\mathbf{L}')$$

defined by passage to quotients. The *relative simplicial homology groups*, denoted by $H_*^{\text{ordered}}(\mathbf{K},\mathbf{L})$ and $H_*(\mathbf{K},\mathbf{L})$ respectively, are then defined to be the homology groups of the corresponding quotient complexes. We should note that the previously defined (absolute) chain groups may be viewed as special cases of this definition for which $\mathbf{L} = \emptyset$.

By Theorem 2 above, we have the following result:

THEOREM 3. (Long Exact Homology Sequence Theorem — Simplicial Version). *Let $i : \mathbf{L} \to \mathbf{K}$ denote a simplicial subcomplex inclusion, and suppose we also have $i' : \mathbf{L}' \to \mathbf{K}'$ such that everything fits into a commutative diagram as above. Then there are long exact sequences of homology groups, and they fit into the following commutative diagram, in which the rows are exact and the vertical arrows represent the canonical maps induced by inclusions of subcomplex pairs:*

$$\begin{array}{ccccccccccc}
\cdots & H_{k+1}(\mathbf{K},\mathbf{L}) & \xrightarrow{\partial} & H_k(\mathbf{L}) & \xrightarrow{i_*} & H_k(\mathbf{K}) & \xrightarrow{j_*} & H_k(\mathbf{K},\mathbf{L}) & \xrightarrow{\partial} & H_{k-1}(\mathbf{L}) & \cdots \\
\cdots & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
\cdots & H_{k+1}(\mathbf{K}',\mathbf{L}') & \xrightarrow{\partial} & H_k(\mathbf{L}') & \xrightarrow{i'_*} & H_k(\mathbf{K}') & \xrightarrow{j'_*} & H_k(\mathbf{K}',\mathbf{L}') & \xrightarrow{\partial'} & H_{k-1}(\mathbf{L}') & \cdots
\end{array}$$

This follows immediately from the definitions and Theorem 2.■

Another proof of Theorem V.2.3. The preceding result yields a much quicker proof of Theorem V.2.3, which computes the homology of $\partial\Delta_{n+1}$; in effect, the theorem is a machine for outsourcing the detailed examination of the chain groups and their interrelations.

First of all, since $n \geq 2$ we know that $\partial\Delta_{n+1}$ is connected and therefore the map

$$H_0(\partial\Delta_n,\ ...) \rightarrow H_0(\Delta_n,\ ...)$$

is an isomorphism (both groups are infinite cyclic and a generator is represented by the class of a vertex). Next, the relative homology groups

$$H_*(\Delta_{n+1}, \partial\Delta_{n+1},\ ...)$$

are trivial to compute because the relative chain complex

$$C_*(\Delta_{n+1}, \partial\Delta_{n+1},\ ...)$$

is $\mathbb{Z}$ in dimension $n+1$ and 0 otherwise; since there is only one nonzero group in this chain complex, all the boundary maps are zero and hence the homology and chain groups are the same. Therefore, if $q \neq n, n+1$ we have the following exact sequences:

$$0 = H_{q+1}(\Delta_{n+1}, \partial\Delta_{n+1}) \rightarrow H_q(\partial\Delta_{n+1}) \rightarrow H_q(\Delta_{n+1}) \rightarrow H_q(\Delta_{n+1}, \partial\Delta_{n+1}) = 0$$

Exactness of the first two morphisms imply that the map from $H_q(\partial\Delta_{n+1})$ to $H_q(\Delta_{n+1})$ is injective, and exactness of the last two morphisms shows that this map is also surjective, so that the homology groups of $\partial\Delta_n$ and Δ_{n+1} are isomorphic if $q \neq n, n+1$. In these cases the homology groups of Δ_n are zero unless $q = 0$ and $\mathbb{Z}$ in the latter case, so the same is true for $\partial\Delta_{n+1}$. We also know that

$H_{n+1}(\partial\Delta_{n+1}) = 0$ because the corresponding chain group is zero, so we are left with the following exact sequence:

$$0 = H_{n+1}(\Delta_{n+1}) \;\to\; H_{n+1}(\Delta_{n+1}, \partial\Delta_{n+1}) \;\to\; H_n(\partial\Delta_{n+1}) \;\to\; H_n(\Delta_{n+1}) = 0$$

As in the preceding discussion, this exact sequence implies that $H_{n+1}(\Delta_{n+1}, \partial\Delta_{n+1}) \cong \mathbb{Z}$ and $H_n(\partial\Delta_{n+1})$ are isomorphic.■

The Five Lemma

Theorem 3 provides one fundamental piece of algebraic input for relating the homology of a simplicial complex and the homology of a subcomplex. Another important piece of input for manipulating long exact sequence is given by the following result:

PROPOSITION 4. *Suppose we are given a commutative diagram of modules as below in which the rows are exact and the horizontal maps a, b, d and e are isomorphisms. Then the mapping c is also an isomorphism:*

$$\begin{array}{ccccccccccc} \cdots & A & \xrightarrow{f_*} & B & \xrightarrow{g_*} & C & \xrightarrow{j_*} & D & \xrightarrow{k_*} & E \\ & \downarrow a & & \downarrow b & & \downarrow c & & \downarrow d & & \downarrow e \\ & A' & \xrightarrow{f'_*} & B' & \xrightarrow{g'_*} & C' & \xrightarrow{h'_*} & D' & \xrightarrow{k'_*} & E' \end{array}$$

A proof of this theorem appears on page 129 of Hatcher.■

One important consequence of the Five Lemma is the following:

COROLLARY 5. *Let $i : A \to B$ and $j : A \to A/B$ be injection and projection maps of chain complexes as in Theorem 2 and likewise for $i' : A \to B$ and $j' : A \to A/B$, and let $f : B \to B'$ be a chain complex morphism which maps A to A'. Denote the induced maps of subcomplexes and quotient complexes by $g : A \to A'$ and $h : A/B \to A'/B'$. Then if any two of the graded maps of homology groups*

$$f_*\ , \quad g_*, \quad h_*$$

is an isomorphism (in all dimensions), then so is the third.■

V.4 : Computational techniques

(**H**, §2.2)

The long exact sequence of the previous section is often a useful tool for comparing the homology groups of a simplicial complex and a subcomplex, but in many cases we need something stronger. In this section we shall derive two important principles which are extremely useful for making quantitative comparisons.

NOTATIONAL CONVENTIONS. The general setting involves a simplicial complex $(P, \mathbf{K})$ which is a union of two subcomplexes $(P_1, \mathbf{K}_1)$ and $(P_2, \mathbf{K}_2)$, and it will also involve the intersection subcomplex

$$(P_1 \cap P_2, \mathbf{K}_1 \cap \mathbf{K}_2)\ .$$

Throughout this section we shall use such notation without further comment. We shall also assume that we are given some linear ordering of the vertices, but frequently we shall suppress the associated superscript in order to simplify the notation. Also, in some cases we shall simply use $\mathbf{L}$ to denote the simplicial complex $(Q, \mathbf{L})$.

Excision for simplicial homology

In the general setting described above, one approach to analyzing the homology of $(P, \mathbf{K})$ is to start with the homology groups of the subcomplex $(P_1, \mathbf{K}_1)$, the relative homology groups of the pair $(\mathbf{K}, \mathbf{K}_1)$ and the long exact homology sequence of the pair. The Excision Property for simplicial homology implies that the relative homology groups only depend upon data involving the somewhat complementary subcomplex $\mathbf{K}_2$.

THEOREM 1. (Simplicial Excision Property) *Suppose that the simplicial complex* $\mathbf{K}$ *is the union* $\mathbf{K}_1$ *and* $\mathbf{K}_2$*. If*

$$(\mathbf{K}_2, \mathbf{K}_1 \cap \mathbf{K}_2) \longrightarrow (\mathbf{K} = \mathbf{K}_1 \cup \mathbf{K}_2, \mathbf{K}_1)$$

is the inclusion map of pairs, then the induced homology maps e_* *are isomorphisms in all dimensions.*

Proof. In fact, we claim that the chain complex map is an isomorphism (hence the of homology groups is also an isomorphism). Recall that the relative chain groups $C_q(\mathbf{L}, \mathbf{M})$ of a subcomplex pair $(\mathbf{L}, \mathbf{M})$ are free abelian on the q-simplices of $\mathbf{L}$ that do not belong to $\mathbf{M}$.

By construction, in each dimension q the chain map

$$e_\# : C_*(\mathbf{K}_2, \mathbf{K}_1 \cap \mathbf{K}_2) \longrightarrow C_*(\mathbf{K}, \mathbf{K}_1)$$

takes each free generator representing a q-simplex in $\mathbf{K}_2$, but not in $\mathbf{K}_1 \cap \mathbf{K}_2$, to its counterpart representing a q-simplex in $\mathbf{K}$ but not in $\mathbf{K}_1$. Since the inclusion of $\mathbf{K}_2 - (\mathbf{K}_1 \cap \mathbf{K}_2)$ in $\mathbf{K} - \mathbf{K}_1$ is a 1–1 onto mapping under which the q-simplices correspond, it follows that $e_\#$ maps the free generators of one chain group to free generators of the other and hence is an isomorphism.■

Example. If $(\mathbf{K}, \mathbf{L})$ is a simplicial complex pair, then $\mathbf{K}$ is said to be an *elementary expansion* of $\mathbf{L}$ if $\mathbf{K}$ is obtained by adjoining a single simplex A such that exactly one maximal lower-dimensional face of A lies in $\mathbf{L}$. In the file `expansion.pdf`, the triangle together with any one of the 2-simplices is an elementary expansion of the triangle, the the entire complex depicted in that file is obtained from the triangle by a sequence of three elementary expansions (one along each edge).

CLAIM. *If* $\mathbf{K}$ *is an elementary expansion of* $\mathbf{L}$*, then the inclusion from* $\mathbf{L}$ *to* $\mathbf{K}$ *induces isomorphisms in homology.* — By the exact homology sequence, the conclusion will hold if $H_*(\mathbf{K}, \mathbf{L}) = 0$ in all dimensions. Let $\mathbf{A}$ denote the usual simplicial decomposition of A, and let $\mathbf{A}'$ be the corresponding decomposition $\mathbf{L} \cap \mathbf{A}$ of the face which is the intersection of the two subcomplexes. By excision the inclusion

$$(\mathbf{A}, \mathbf{A}') \longrightarrow (\mathbf{K}, \mathbf{L})$$

induces isomorphisms in homology, so everything reduces to proving that $H_*(\mathbf{A}, \mathbf{A}') = 0$. This will follow from the long exact homology sequence of the pair because $H_q(\mathbf{A}) = H_q(\mathbf{A}') = 0$ if $q \neq 0$ and the vertex inclusions represented by vertical arrows in the diagram below are isomorphisms; since the top map is the identity, it follows that the bottom map must also be an isomorphism.

$$\begin{array}{ccc}
\mathbb{Z} \cong H_0(\{\mathbf{v}\}) & \stackrel{=}{\longrightarrow} & H_0(\{\mathbf{v}\}) \cong \mathbb{Z} \\
\downarrow \cong & & \downarrow \cong \\
\mathbb{Z} \cong H_0(\mathbf{A}') & \stackrel{i_*}{\longrightarrow} & H_0(\mathbf{A}) \cong \mathbb{Z}
\end{array}$$

If we apply this result three times to the complex in `expansion.pdf`, we find that the inclusion of the triangle in the larger complex (with four 2-simplices) induces isomorphisms in homology.

Simplicial Mayer-Vietoris sequences

One variant of the long exact homology sequence, known as a **Mayer-Vietoris**, is a particularly effective tool for relating the homologies of $\mathbf{K} = \mathbf{K}_1 \cup \mathbf{K}_2$, $\mathbf{K}_1$, $\mathbf{K}_2$, and $\mathbf{K}_1 \cap \mathbf{K}_2$). It can be viewed as an analog of an elementary formula for counting elements of finite sets

$$\#(A \cup B) \quad = \quad \#(A) \;+\; \#(B) \;-\; \#(A \cap B)$$

and of the Seifert-van Kampen Theorem, which relates the fundamental group of an arcwise connected topological space X to the fundamental groups of arcwise connected open subsets U and V such that $U \cup V = X$ and $U \cap V$ is arcwise connected.

THEOREM 2. (Simplicial Mayer-Vietoris Theorem) *Let $\mathbf{K}$ be a topological space, and let $\mathbf{K}_1$ and $\mathbf{K}_2$ be open subsets such that $\mathbf{K} = \mathbf{K}_1 \cup \mathbf{K}_2$. Denote the inclusions of $\mathbf{K}$ and $\mathbf{K}$ in X by i_1 and i_2 respectively, and denote the inclusions of $\mathbf{K}_1 \cap \mathbf{K}_2$ in $\mathbf{K}_1$ and $\mathbf{K}_2$ by g_1 and g_2 respectively. Then there is a long exact sequence*

$$\cdots \to H_{q+1}(\mathbf{K}) \to H_q(\mathbf{K}_1 \cap \mathbf{K}_2) \to H_q(\mathbf{K}_1) \oplus H_q(\mathbf{K}_2) \to H_q(\mathbf{K}) \to \cdots$$

in which the map from $H_*(\mathbf{K}_1) \oplus H_*(\mathbf{K}_2)$ to $H_*(\mathbf{K})$ is given on the summands by $(i_1)_*$ and $(i_2)_*$ respectively, and the map from $H_*(\mathbf{K}_1 \cap \mathbf{K}_2)$ to $H_*(\mathbf{K}_1) \oplus H_*(\mathbf{K}_2)$ is given on the factors by $-(g_1)_*$ and $(g_2)_*$ respectively (note the signs!!).

There is a purely formal study of the relationship between long exact homology sequences and Mayer-Vietoris sequences in Chapter I.15 of Eilenberg and Steenrod. However, in our context it will be easier to prove the result using the definitions of the chain complexes.

Proof. The underlying chain complex inclusion mappings $(i_1)_\#$ and $(i_2)_\#$ yield a chain map

$$\nabla_\# : C_*(\mathbf{K}_1) \oplus C_*(\mathbf{K}_2) \longrightarrow C_*(\mathbf{K})$$

and this map is onto because every simplex of $\mathbf{K}$ is either in $\mathbf{K}_1$ or $\mathbf{K}_2$. Direct calculation shows that the kernel of $\nabla_\#$ is the image of the chain map

$$C_*(\mathbf{K}_1 \cap \mathbf{K}_2) \;\longrightarrow\; C_*(\mathbf{K}_1) \oplus C_*(\mathbf{K}_2)$$

which is given on the factors by $-(g_1)_\#$ and $(g_2)_\#$ respectively (equivalently, one could take the negatives of both maps and still have a correct conclusion, but it is necessary for the signs to be opposite). These maps fit together to form a short exact sequence of chain complexes, and the theorem follows by taking the long exact homology sequence associated to the given short exact sequence of chain complexes.■

A somewhat nontrivial but still fairly simple homology computation using Mayer-Vietoris sequences appears in the file `mv-example.pdf`.

V.∞ : Is homology topologically invariant?

(**H**, §2.1)

Although the results of Section IV.1 show that homeomorphic graphs have isomorphic homology groups. However, the proof of this fact is somewhat convoluted; one proves that H_1 is free abelian on the same number of generators that freely generate the nonabelian group π_1. It is not difficult to find plenty of examples suggesting that a similar result holds for arbitrary simplicial complexes, and at the beginning of the $20^{\rm th}$ century the issue of topological invariance was an urgent open question which had major implications for the development of geometric topology. Ultimately the invariance question was solved by proving something considerably stronger: *Homotopy equivalent simplicial complexes have isomorphic homology groups.* The first proofs used properties of simplicial chain complexes and homology that are beyond the scope of this course; the standard modern approach is based upon the **singular homology theory** developed by Eilenberg and Steenrod in the 1940s. We do not have enough time in this course to develop this theory, which is absolutely central to modern algebraic topology and is based upon ideas from simplicial homology that we do not have enough time to develop. However, following Eilenberg and Steenrod we shall give a list of axioms for singular homology (significantly longer than theirs!) which suffice to yield many basic applications of singular homology.

VI. Axiomatic singular homology

We shall begin by continuing the discussion in the final paragraph of Unit V.

The basic idea behind the construction of singular homology is simple; given a topological space X, one constructs a chain complex which is associated to X itself (without any decomposition or ordering data) which is covariantly functorial with respect to continuous functions and resembles the simplicial chain complex associated to a simplicial complex with a linear vertex ordering. A detailed construction of singular homology appears in Chapter 2 of Hatcher and Unit IV of the following notes:

`http://math.ucr.edu/~res/math246A/algtopnotes2010.pdf`

Of course, there are also accounts of singular homology in virtually every textbook on algebraic topology written during the past 60 years. As noted before, the construction of singular homology requires a substantial amount of information about simplicial complexes and chain complexes, including many topics omitted from the present course. One reference for this necessary background material is Sections III.3 – III.5 of the previously cited online notes.

We have already noted that there is not enough time in the course to present the construction of singular homology together with a reasonable amount of supporting material that will (a) explain the geometrical motivation for several key constructions, (b) apply the theory to questions of independent interest such as results in Sections 55-56 and 61-65 of Munkres' Book and their generalizations to higher dimensions. We shall deal with this problem by giving an axiomatic description of singular homology in this unit and discussing a few standard applications in the next unit. The possibility of taking such an approach is mentioned in the second full paragraph of page 98 in Hatcher, and in these notes our approach is near the far end of his contrast between emphasizing the explicit construction of singular theory and prioritizing its basic formal properties. However, we should also state very clearly that **the construction of singular homology also has a considerable amount of additional structure that is indispensable to any further study of algebraic topology**; in particular, Chapter 3 of Hatcher studies a major piece additional structure, and Section 4.L of Hatcher provides a glimpse into yet another key structural aspect of singular homology (relatively recent results due to M. Mandell show that elaborations of the structure in Section 4.L completely determine the homotopy type of "reasonably well-behaved" spaces — here is the reference: Institut des Hautes Études Scientifiques Publications Mathématiques **103** (2006), pp. 213–246).

Finally, we should note that our list of axioms is fairly long to say the least, and it turns out that there is a great deal of redundancy in our axioms. Of course, one would ultimately like to have a set of axioms that have little or no logical interdependence, but for our purposes it seems best to start with a set of axioms which can be used fairly simply and quickly to prove nontrivial results. At this point, proving that some of these axioms imply the others would cut seriously into the time available for discussing the applications we want to cover, but in any further study of the subject it would be necessary to deal with such issues in one way or another.

VI.1 : Primitive data and basic axioms

(**H**, Ch. 2 Introduction, §§2.1, 2.3)

In the previous unit it was useful to define relative homology groups for pairs $(\mathbf{K}, \mathbf{L})$ consisting of a simplicial complex $\mathbf{K}$ and a subcomplex $\mathbf{L}$. We need to begin by generalizing this concept of pairs to topological spaces.

Definition. A *pair of topological spaces* is an ordered pair (X, A) where A is a subspace of X. A (continuous) map of pairs $f : (X, A) \to (Y, B)$ is given by a continuous mapping $f : X \to Y$ such that $f[A] \subset B$. — It follows immediately that pairs of spaces and maps of pairs form a category.

We can embed the category of topological spaces and continuous mappings into the category of pairs by sending X to $(X, \emptyset)$, and often we shall use X to denote this pair. Given two pairs (X, A) and (Y, B), their product in the category of pairs is given by $(X \times Y,\ X \times B \cup A \times Y)$. With this definition, the cartesian product of two maps of pairs becomes a map of pairs (verify this!). In particular, if $B = \emptyset$ then we can write $(X, A) \times Y = (X \times Y, A \times Y)$.

There is also a natural embedding of the category of spaces with base points into the category of pairs of spaces. Specifically, if (X, x_0) is a space with base point we take the corresponding pair $(X, \{x_0\})$, and similarly base point preserving maps define maps of the associated pairs of spaces.

Another important construction in the category of pairs is the natural map $j_{(X,A)} : (X, \emptyset) \to (X, A)$ which is the identity on X (and the empty function on $\emptyset$). We can take this further by defining two **forgetful functors** on pairs of spaces by the rules

$$F_1\big((X, A)\big) \ = \ X \ , \qquad F_2\big((X, A)\big) \ = \ A$$

and extending them to morphisms in the obvious fashion. If we do this, then j becomes a natural transformation from F_1 to the identity functor, and similarly one can view the inclusion $i : A \subset X$ defines a natural transformation from F_2 to F_1.

The data for an abstract singular homology theory

Obviously we want to have homology groups for pairs of spaces, and we would also like to have an analog of the long exact homology sequence for a simplicial complex pair. The first of these explains the need to include homology groups into the axioms, and the second indicates that we need analogs of the connecting maps for long exact homology sequences.

1. For each pair of spaces (X, A) and each integer q we are given an abelian group $H_q(X, A)$ and a homomorphism $\partial_q : H_q(X, A) \to H_{q-1}(A)$.

We also want some analogs of Mayer-Vietoris exact sequences, and for these we need slightly different types of connecting homomorphisms for such sequences. One fundamental difference between singular and simplical homology is that the latter has Mayer-Vietoris sequences for complexes presented as unions of subcomplexes — which are closed subsets — but in general for singular homology it is necessary to limit ourselves to Mayer-Vietoris sequences for spaces presented as unions of **open** subsets. Since we have already noted that Mayer-Vietoris sequences are somehow analogous to the Seifert-van Kampen Theorem and the latter involves spaces presented as unions of open subsets, there is some precedent for restricting attention to such decompositions.

2. For each decomposition of a space $X = U \cup V$ such that Interior$(U) \cup$ Interior$(V) = X$ we are given a homomorphism $\Delta : H_q(X) \to H_{q-1}(U \cap V)$.

Note that the condition on U and V is automatically satisfied if U and V are both open in X, and in fact this is probably the most important special case for our purposes.

In our discussion of simplicial homology we noted the existence of homology homomorphisms $H_*(\mathbf{L}^\omega) \to H_*(\mathbf{K}^\omega)$ associated to a subcomplex inclusion $\mathbf{L} \subset \mathbf{K}$. In singular homology we want to have homology homomorphisms associated to arbitrary continuous mappings of pairs.

3. For each map of pairs $f : (X, A) \to (Y, B)$ and each integer a we are given a homomorphism $f_* : H_q(X, A) \to H_q(Y, B)$.

Sometimes such data are called a ∂-functor (pronounced "dell functor") from the category of pairs of topological spaces and maps of pairs to the category of abelian groups. The maps ∂ are frequently called *connecting homomorphisms* or *switchback homomorphisms*.

It is also useful (but actually redundant) to assume a relationship between the fundamental group and the first homology group which is a formal version of the relationship between the fundamental group and first homology group of a connected graph, and for this we need the following:

4. There is a family of group homomorphisms $h(X, x) : \pi_1(X, x) \to H_1(X, \{x\})$.

Finally, we want some sort of relationship between our axiomatized singular homology theory and the simplicial homology groups we have studied.

5. If $(P, \mathbf{K})$ is a simplicial complex (strictly speaking, with an ordering of the vertices) and $(Q, \mathbf{L})$ is a subcomplex of $(P, \mathbf{K})$, there is a sequence of homomorphisms $\theta_{(\mathbf{K}^\omega, \mathbf{L}^\omega)}$: $H_q(\mathbf{K}^\theta, \mathbf{L}^\theta) \to H_q(P, Q)$.

For the sake of notational simplicity, for the rest of this section we implicitly assume that we are given compatible linear orderings of simplicial complexes (in other words, subcomplexes inherit the induced ordering) and simplicial homology groups will be written without including the orderings explicitly.

Functoriality and naturality

One reason the fundamental group is so useful is that it is functorial with respect to continuous maps; if f and g are two composable maps of pointed spaces and f_* and g_* then $(g \circ f)_* = g_* \circ f_*$ and if f is an identity map then so is f_*. We want a similar sort of condition for homology groups.

(A.1) *If $f : (X, A) \to (Y, B)$ and $g : (Y, B) \to (Z, C)$ are maps of pairs, then $(g \circ f)_* = g_* \circ f_*$. If f is the identity map on (X, A), then f_* is the identity on $H_q(X, A)$ for all integers q.*

One important consequence of this is that if f is a homeomorphism of pairs, then the homology maps f_* are isomorphisms (if $g = f^{-1}$, then we have $g_* = (f_*)^{-1}$ by the same sort of argument which proves an analogous result for fundamental groups).

We have only defined maps in simplicial homology groups from $H_*(\mathbf{K}, \mathbf{L})$ to $H_*(\mathbf{K}', \mathbf{L}')$ when $\mathbf{K}$ and $\mathbf{L}$ are subcomplexes of $\mathbf{K}'$ and $\mathbf{L}'$ respectively, but we would like θ to be natural with respect to such maps of pairs.

(A.2) *If we are given inclusions as in the preceding paragraph and the underlying spaces are given by* (P, Q) *and* (P', Q') *respectively, then for each integer* q *the diagram*

$$\begin{array}{ccc} H_q(\mathbf{K}, \mathbf{L}) & \longrightarrow & H_q(\mathbf{K}', \mathbf{L}') \\ \downarrow \theta & & \downarrow \theta \\ H_q(P, Q) & \xrightarrow{\partial} & H_q(P', Q') \end{array}$$

is commutative, where the horizontal arrows come from subcomplex or subspace inclusions of pairs.

We also want the homology homomorphisms induced by a map f to be compatible with the maps ∂ in the following sense:

(A.3) *If* $f : (X, A) \to (Y, B)$ *is a map of pairs, then for each integer* q *the diagram*

$$\begin{array}{ccc} H_q(X, A) & \xrightarrow{\partial} & H_{q-1}(A) \\ \downarrow f_* & & \downarrow f_* \\ H_q(Y, B) & \xrightarrow{\partial} & H_{q-1}(B) \end{array}$$

is commutative.

We also want a similar property for the map $\Delta : H_{q+1}(U \cup V) \to H_q(U \cap V)$ given in the data. It turns out that this is a consequence of other axioms, but we shall not try to prove this.

(A.4) *If we are given spaces* $X_i = U_i \cup V_i$ *for* $i = 1, 2$*, where* $\mathrm{Interior}(U_i) \cup \mathrm{Interior}(V_i) = X_i$*, and* $f : X_1 \to X_2$ *is a continuous map which maps* U_1 *and* V_1 *into* U_2 *and* V_2 *respectively, then for all integers* q *the diagram*

$$\begin{array}{ccc} H_q(X_1) & \xrightarrow{\Delta} & H_{q-1}(U_1 \cap V_1) \\ \downarrow f_* & & \downarrow f_* \\ H_q(X_2) & \xrightarrow{\Delta} & H_{q-1}(U_2 \cap V_2) \end{array}$$

is commutative.

We also want the connecting homomorphisms from singular and simplicial homology to be compatible:

(A.5) *If* $(Q, \mathbf{L})$ *is a subcomplex of the simplicial complex* $(P, \mathbf{K})$*, the for all integers* q *the diagram*

$$\begin{array}{ccc} H_q(\mathbf{K}, \mathbf{L}) & \xrightarrow{\partial} & H_{q-1}(\mathbf{L}) \\ \downarrow \theta & & \downarrow \theta \\ H_q(P, Q) & \xrightarrow{\partial} & H_{q-1}(Q) \end{array}$$

is commutative.

Finally, we want a naturality property of the map h from fundamental groups to 1-dimensional homology.

(A.6) *If $f : (X,x) \to (Y,y)$ is a continuous base point preserving map of arcwise connected spaces, then the diagram*

$$\begin{array}{ccc} \pi_1(X,x) & \xrightarrow{f_*} & \pi_1(Y,y) \\ \downarrow h & & \downarrow h \\ H_1(X,\{x\}) & \xrightarrow{f_*} & H_1(Y,\{y\}) \end{array}$$

is commutative.

Usually a mapping like h is called a Hurewicz (hoo-RAY-vich) homomorphism.

VI.2 : Exactness, homotopy invariance and support properties

(**H**, §2.3)

We want singular homology groups to have strong versions of the properties that hold for simplicial complexes, and we would like to have some sort of relation between the singular homology of an arbitrary topological space and the the homology of simplicial complexes. Exactness and homotopy invariance are strengthened versions of the long exact simplicial homology sequence and the theorem stating that the simplicial homology of a star shaped complex is isomorphic to the singular homology of a point. The support property states that the singular homology of a space X is determined by the singular homology of its compact subspaces, and in fact it is determined by continuous maps from polyhedra into X.

Exactness

In simplicial homology we have a long exact sequence associated to a pair $(Q, \mathbf{L}) \subset (P, \mathbf{K})$, and we want a similar exact sequence in singular homology for the pair (P, Q). In fact, we want such a sequence for an arbitrary pair, and we want it to have good compatibility properties. We shall start with existence:

(B.1) *If (X,A) is a pair of topological spaces then there is a long exact sequence*

$$\cdots \quad H_{q+1}(X,A) \xrightarrow{\partial} H_q(A) \xrightarrow{i_*} H_q(X) \xrightarrow{j_*} H_q(X,A) \xrightarrow{\partial} H_{q-1}(A) \quad \cdots$$

which extends indefinitely to the left and to the right for all integers q. In this sequence i_* is induced by the inclusion map $A \to X$, and j_* is induced by the inclusion of pairs from X to (X,A).

We actually need two types of compatibility; namely, compatibility with respect to continuous maps of pairs and compatibility with the maps θ passing from simplicial to singular homology. These will be stated individually, and they all follow from (B.1), the first group of axioms, and the known properties of simplicial homology groups (hence they are redundant).

(B.2) *If we are given a continuous map of pairs $f : (X,A) \to (Y,B)$, then we have the following commutative ladder diagram in which the two rows are exact:*

$$\begin{array}{ccccccccccccc} \cdots & H_{q+1}(X,A) & \xrightarrow{\partial} & H_q(A) & \xrightarrow{i_*} & H_q(X) & \xrightarrow{j_*} & H_q(X,A) & \xrightarrow{\partial} & H_{q-1}(A) & \cdots \\ \cdots & \downarrow f_* & & \downarrow f_* & & \downarrow f_* & & \downarrow f_* & & \downarrow f_* & \\ \cdots & H_{q+1}(Y,B) & \xrightarrow{\partial} & H_q(B) & \xrightarrow{i_*} & H_q(Y) & \xrightarrow{j_*} & H_q(Y,B) & \xrightarrow{\partial} & H_{q-1}(B) & \cdots \end{array}$$

This statement turns out to be a fairly straightforward consequence of (A.1) and (B.1).

(B.3) *Let $(X, \mathbf{K})$ be a simplicial complex, and let $(A, \mathbf{L})$ be a subcomplex. Then there is a commutative ladder as below in which the horizontal lines represent the long exact homology sequences of pairs and the vertical maps are the natural transformations from simplicial to singular homology.*

$$\begin{array}{ccccccccccccc} \cdots & H_{q+1}(\mathbf{K},\mathbf{L}) & \xrightarrow{\partial} & H_q(\mathbf{L}) & \xrightarrow{i_*} & H_q(\mathbf{K}) & \xrightarrow{j_*} & H_q(\mathbf{K},\mathbf{L}) & \xrightarrow{\partial} & H_{q-1}(\mathbf{L}) & \cdots \\ \cdots & \downarrow\theta & & \downarrow\theta & & \downarrow\theta & & \downarrow\theta & & \downarrow\theta & \\ \cdots & H_{q+1}(X,A) & \xrightarrow{\partial} & H_q(A) & \xrightarrow{i_*} & H_q(X) & \xrightarrow{j_*} & H_q(X,A) & \xrightarrow{\partial} & H_{q-1}(A) & \cdots \end{array}$$

This statement turns out to be a fairly straightforward consequence of (A.2), (B.1) and the long exact simplicial homology sequence for the pair $(\mathbf{K}, \mathbf{L})$.

This is a good point at which to derive a few simple consequences of the axioms thus far.

PROPOSITION 1. *(i) For every space X the groups $H_q(X, X)$ are trivial.*

(ii) If X is the empty set then $H_q(X) = 0$ for all integers q.

Proof. The second statement follows from the first because $H_q(X) = H_q(X, \emptyset)$.

To prove the first statement, consider the long exact homology sequence for the pair (X, X), and especially the following piece:

$$H_q(X) \xrightarrow{i_*} H_q(X) \xrightarrow{j_*} H_q(X,X) \xrightarrow{\partial} H_{q-1}(X) \xrightarrow{i_*} H_{q-1}(X) \xrightarrow{j_*} H_{q-1}(X,X)$$

In this sequence i denotes the identity map on X so that i_* is the identity on $H_*(X)$. By exactness this implies that j_* and ∂ are both zero mappings (the first because i_* is onto, the second because i_* is 1–1). Now $j_* = 0$ implies that the map ∂ is 1–1 by exactness, and its image is isomorphic to $H_q(X, X)$. On the other hand, we also know that $\partial = 0$, and this implies that $H_q(X, X) = 0$.■

There will eventually be one more axiom concerning long exact sequences (the Mayer-Vietoris Sequence Axiom), but we we postpone its statement because if fits more naturally into another group of assumptions.

Homotopy invariance

We have seen that the simplicial homology of an n-simplex is isomorphic to the simplicial homology of a one point complex, with $H_0(P, \mathbf{K}) \cong \mathbb{Z}$ and $H_q(P, \mathbf{K}) = 0$ for $q \neq 0$, and our proof used an algebraic analog of the standard topological homotopy contracting the simplex into a vertex. Also, for spaces with base points we know that base point preservingly homotopic maps induce the same homomorphisms of fundamental groups. The Homotopy Invariance Axiom is a very strong analog of these phenomena. For the sake of completeness, we note that a two maps of pairs $f_0, f_1 : (X, A) \to (Y, B)$ are homotopic as maps of pairs if there is a continuous map

$$H : (X, A) \times [0, 1] \;=\; \big(X \times [0, 1],\, A \times [0, 1]\big) \;\longrightarrow\; (Y, B)$$

such that the restriction of H to $X \times \{i\}$ is f_i for $i = 0, 1$.

(C.1) *If $f_0, f_1 : (X,A) \to (Y,B)$ are homotopic as maps of pairs, then their induced homology homomorphisms $f_{0\,*}$ and $f_{1\,*}$ are equal.*

One important consequence of this is that if f is a homotopy equivalence of pairs in the appropriate sense, then the homology maps f_* are isomorphisms. This is analogous to a basic result for fundamental groups, and the proof is also analogous: Let g be a homotopy inverse so that $g \circ f$ and $f \circ g$ are homotopic to the respective identity maps. Then homotopy invariance and functoriality imply that $g_* \circ f_*$ and $f_* \circ g_*$ are identity maps in homology, and hence f_* is an isomorphism with inverse g_*.■

Compact supports and polyhedral generation

Two important properties of the fundamental group are that

(1) each class in $\pi_1(X,x)$ comes from $\pi_1(C,x)$ for some compact subset C containing x,

(2) if C is a compact subset of C containing x, then a class $\alpha \in \pi_1(C,x)$ maps to the trivial element of $\pi_1(X,x)$ (under the inclusion induced map) if and only if there is some compact subset D such that $C \subset D \subset X$ and α maps to the trivial element of $\pi_1(D,x)$ (under the inclusion induced map).

These results hold because the image of a closed curve is always compact subset, and similarly for a nullhomotopy of a closed curve. We need a similar property for singular homology which is sometimes called a *compact supports property*:

(C.2) *If $u \in H_q(X,A)$ then there is a pair of compact subsets (C,C') such that u is in the image of the map from $H_q(C,C')$ to $H_q(X,A)$ induced by inclusion. Furthermore, if (C,C') is a pair of compact subsets and $v \in H_q(C,C')$ maps trivially to $H_q(X,A)$ under the map induced by inclusion, then there is some compact pair (D,D') such that $C \subset D$, $C' \subset D'$, and v maps trivially to $H_q(D,D')$ under the map induced by inclusion.*

Although this does not correspond to any of the original axioms due to Eilenberg and Steenrod, it plays a very important role in the applications of singular homology theory and the proof of its uniqueness.

The compact supports property is adequate for most basic applications of singular homology theory, but in order to prove uniqueness we need a slightly stronger *polyhedral generation property*, which essentially says that the singular homology of a space is somehow given by the singular homology of simplicial complexes.

(C.3) *If $u \in H_q(X,A)$ then there is a simplicial complex pair $(P',\mathbf{K}') \subset (P,\mathbf{K})$ $(\mathbf{K},\mathbf{K}')$ and a continuous map of pairs*

$$f : (P,P') \longrightarrow (X,A)$$

such that u is in the image of the map f_* from $H_q(P,P')$ to $H_q(X,A)$. Furthermore, if $(\mathbf{K},\mathbf{K}')$ is a simplicial complex pair with underlying space pair (P,P') and $v \in H_q(P,P')$ maps trivially to $H_q(X,A)$ under the map f_*, then there is another simplicial complex pair $(Q',\mathbf{L}') \subset (Q,\mathbf{L})$ and a continuous function $g : (Q,Q') \to (X,A)$ such that the following hold:

(i) *The pair (P,P') is contained in (Q,Q') such that P' and P correspond to subcomplexes of $\mathbf{L}'$ and $\mathbf{L}$ respectively.*

(ii) *The restriction of g to (P,P') is equal to f.*

(*iii*) *The image of v in $H_q(Q, Q')$, under the map induced by inclusion, is zero.*

Note. Nothing is stated or assumed about the relationship between **K** and **L**; in particular, **K** is not necessarily a subcomplex of **L**.

It is at least somewhat known that the polyhedral exhaustion property holds for singular homology (in particular, this is an immediate consequence of results on geometric realizations of semisimplicial sets; one reference is J. P. May, *Simplicial Objects in Algebraic Topology*, University of Chicago Press, Chicago IL, 1982). Although one can give a fairly direct proof of (C.3) for the standard construction of singular homology using machinery from algebraic topology appearing in most textbooks, a direct reference for such a proof seems hard to find, so we shall give a proof of this sort in Section VI.∞.

The implication (C.3) $\Longrightarrow$ (C.2) is very short and simple. The continuous image of a polyhedron is a compact subset of the codomain, and hence in (C.2) we can take the compact sets to be suitable continuous images of polyhedra.∎

VI.3 : Normalization properties

(**H**, §§2.3, 2.A)

These properties are analogous to specifying the initial value for a solution to a differential equation; in a variety of special cases, singular homology groups should be given by objects and principles we already understand. First of all, we want singular homology to be the same as simplicial homology for polyhedral pairs, and the natural transformation from simplicial to singular homology is an obvious candidate for an isomorphism between the two types of homology groups. Furthermore, we also want homology groups to vanish in negative dimensions, we want the homology groups of a space to split into a directs sum of the homology groups of its arc components, and — as in the case of connected graphs — we want the first homology group of an arcwise connnected space to be the abelianization of the fundamental group (the last of these is the reason for introducing the natural homomorphism from π_1 to H_1 as part of the structure). We shall state the redundancies as we introduce the axioms, but we do not have the time to give any proofs.

The first axiom involves the equivalence of singular and simplicial homology.

(D.1) *If $(Q, \mathbf{L})$ is a subcomplex of $(P, \mathbf{K})$, is a linear ordering of the vertices in* **K**, *then for all integers q the map $\theta : H_q(\mathbf{K}, \mathbf{L}) \to H_q(P, Q)$ is an isomorphism.*

We should note that **the topological invariance of simplicial homology groups and their independence of the vertex ordering are immediate consequences of this axiom**, and if we combine these with homotopy invariance we conclude that simplicial homology groups depend only on the homotopy type of the underlying topological space.∎

The next axiom is a splitting principle for the homology of a space into a sum of the homology groups of the arc components. If $(P, \mathbf{K})$ is a polyhedron then we have seen that $H_0(\mathbf{K})$ is free abelian on the set of components (equivalently, arc components), and $H_*(\mathbf{K})$ splits into a direct sum of the homology groups for the various components, and we want a similar principle for singular homology:

(D.2) *If X is written as a union of its (pairwise disjoint) arc components X_α, then the inclusion maps $i_\alpha : X_\alpha \to X$ define an isomorphism from the (weak) direct sum $\oplus_\alpha\, H_*(X_\alpha)$ to $H_*(X)$.*

The weak direct sum of the abelian groups G_α consists of all elements of $\prod_\alpha G_\alpha$ such that only finitely many coordinates are nonzero (with addition defined coordinatewise). This coincides with the direct product if and only if the family $\{G_\alpha\}$ is finite.

The next assumption is that the homology of an arcwise connected space is always given by the same rule which applies to simplicial complexes.

(D.3) *If X is arcwise connected, then $H_0(X) \cong \mathbb{Z}$.*

The preceding two axioms imply that if X is arbitrary, then $H_0(X)$ is isomorphic to a free abelian group on the arc components of X.■

Here are a few specific consequences of (D.2) and (D.3) that are very useful in many situations.

PROPOSITION 1. *(i) If X is an arcwise connected space and $p \in X$, let $P = \{p\}$. Then the inclusion of P in X induces an isomorphism from $H_0(P)$ to $H_0(X)$.*

(ii) If X is a space whose arc components are the subspaces X_α, let $e_\alpha \in H_0(X)$ denote the image of the standard generator of $H_0(\{1\}) \cong \mathbb{Z}$ under the homomorphism induced by the composite $\{1\} \cong \{p_\alpha\} \subset X_\alpha$ where $p_\alpha \in X_\alpha$. Let $f : X \to Y$ be continuous, and define Y_β, e_β, p_β to be the analogs of X_α, e_α, p_α for Y. If f maps the arc component X_v of X into the arc component Y_w of Y, then f_* sends e_v to e_w.

Proof. Let $c : X \to P$ be the constant map. Then the composite $P \subset X \to P$ is the identity, so that the composite map in homology

$$\mathbb{Z} \;\cong\; H_0(P) \;\to\; H_0(X) \;\to\; H_0(P) \;\cong\; \mathbb{Z}$$

is also the identity. Since $H_0(X) \cong \mathbb{Z}$, it follows that the maps $H_0(P) \to H_0(X)$ and $H_0(X) \to H_0(P)$ must be isomorphisms.

The second part involves the following commutative diagram, in which we abuse notation and let f refer generically to various mappings defined using f:

$$\begin{array}{ccccc}
\{p_v\} & \stackrel{\subset}{\longrightarrow} & X_v & \stackrel{\subset}{\longrightarrow} & X \\
\downarrow f & & \downarrow f & & \downarrow f \\
\{f(p_v)\} & \stackrel{\subset}{\longrightarrow} & Y_v & \stackrel{\subset}{\longrightarrow} & Y
\end{array}$$

If we take the associated commutative diagram of homology groups, the conclusion of the second part drops out immediately.■

The next axiom is actually redundant, but we are assuming it for the sake of convenience:

(D.4) *If $q < 0$ then for all pairs (X, A) we have $H_q(X, A) = 0$.*

The final axiom in this group is also redundant, but it provides an important relationship between 1-dimensional homology and the fundamental group. An early version of this result was due to H. Poincaré.

(D.5) *If X is an arcwise connected space and $x \in X$, then the homomorphism $h(X, x) : \pi_1(X, x) \to H_1(X, \{x\})$ is onto and its kernel is the commutator subgroup.*

VI.4 : Excision and Mayer-Vietoris sequences

(**H**, §2.3)

These axioms are variants of the simplicial homology isomorphism

$$H_*(\mathbf{K}_1, \mathbf{K}_1 \cap \mathbf{K}_2) \longrightarrow H_*(\mathbf{K}_1 \cup \mathbf{K}_2, \mathbf{K}_1)$$

and the long exact Mayer-Vietoris that were derived in Unit V. One can prove each of the three statements below is basically equivalent to the others.

The first statement is analogous to excision in simplicial homology. However, one major difference is the restriction on pairs; in the simplicial excision result there is no hypothesis that the subcomplex have a nonempty interior, but in the singular excision axiom the subspace must have a nonempty interior.

(E.1) *Suppose that the space X can be written as a union of subsets $A \cup B$ such that the interiors of A and B form an open covering of X. Then the inclusion of pairs from $(B, A \cap B)$ to $(X = A \cup B, A)$ induces isomorphisms in homology.*

In particular, this axiom applies if A and B are open subsets of X.

The second excision axiom is an alternative form of the first one; it goes back to the work of Eilenberg and Steenrod, and it is the version most often found in textbooks.

(E.2) *Suppose that (X, A) is a pair of spaces and U is a subset of A such that the closure $\overline{U}$ is contained in the interior of A. Then the inclusion of pairs from $(X - U, A - U)$ to (X, A) induces isomorphisms in homology.*

One can derive (E.2) as a consequence of (E.1) by taking $B = X - U$ (note that the open set $X - \overline{U}$ is contained in $X - U$).

AS noted in the first chapter of Eilenberg and Steenrod, the preceding axioms imply the existence of corresponding Mayer-Vietoris sequences. As noted in Unit V, Mayer-Vietoris sequences may be viewed as analogs of the Seifert-van Kampen Theorem, which describes the fundamental group of a space X in terms of the fundamental groups of two open subspaces U and V such that $X = U \cup V$ and all the spaces X, U, V, $U \cup V$, $X = U \cap V$ are nonempty and arcwise connected. If X is the union of two open subsets U and V (with not restrictions involving arcwise connectedness), these Mayer-Vietoris sequences exhibit a corresponding relationship involving the homology groups of U, V, $X = U \cup V$ and $U \cap V$. Once again, to avoid lengthy digressions we shall assume the existence of such sequences.

(E.3) *Let X be a topological space with $X = U \cup V$ such that* $\mathrm{Interior}(U) \cup \mathrm{Interior}(V) = X$. *Denote the inclusions of U and V in X by i_U and i_V respectively, and denote the inclusions of $U \cap V$ in U and V by g_U and g_V respectively. Then there is a long exact sequence*

$$\cdots \to H_{q+1}(X) \to H_q(U \cap V) \to H_q(U) \oplus H_q(V) \to H_q(X) \to \cdots$$

in which the map from $H_*(U) \oplus H_*(V)$ to $H_*(X)$ is given on the summands by $(j_U)_*$ and $(j_V)_*$ respectively, the map from $H_{q+1}(X)$ to $H_q(U \cap V)$ is the map Δ in the axiomatic data, and the map from $H_*(U \cap V)$ to $H_*(U) \oplus H_*(V)$ is given in coordinates by $(i_U)_*$ and $-(i_V)_*$ respectively (note the signs!!).

As before, this axiom applies if U and V are open subsets of X.

We shall also need a naturality property for Mayer-Vietoris sequences with respect to suitably defined mappings of triads $(X;U,V)$. Specifically, let $X_i = U_i \cup V_i$, where $i = 1,2$ and U_i, V_i satisfy the condition in (E.3). If $f : X_1 \to X_2$ is a continuous mapping such that $f[U_1] \subset U_2$ and $f[V_1] \subset V_2$ (this is what we shall call map of triads), then we want this sequence to have the following naturality properties with respect to f:

(E.4) *In the setting of the preceding paragraph, assume we are given a map of triads f from $(X_1;U_1,V_1)$ to $(X_2;U_2,v_2)$. Then there is a commutative ladder as below in which the horizontal lines represent the long exact Mayer-Vietoris sequences of* (E.3) *and the vertical maps are all induced by f:*

$$\begin{array}{ccccccccccc} \cdots & \to & H_{q+1}(X_1) & \to & H_q(U_1 \cap V_1) & \to & H_q(U_1) \oplus H_q(V_1) & \to & H_q(X_1) & \to & \cdots \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ \cdots & \to & H_{q+1}(X_2) & \to & H_q(U_2 \cap V_2) & \to & H_q(U_2) \oplus H_q(V_2) & \to & H_q(X_2) & \to & \cdots \end{array}$$

In analogy with the naturality properties of (B.2) and (B.3), axiom (E.4) turns out to be a fairly straightforward consequence of (E.3) and (A.4).

VI.∞ : Existence and uniqueness theorems

(**H**, §2.3, 3.F)

TO BE COMPLETED

VII. Some elementary applications

The motivation for developing delicate and abstract topological machinery like singular homology is that such constructions are useful for answering mathematical questions that were interesting but difficult to handle with previously existing tools. One of the most obvious examples is the Jordan Curve Theorem, which states that the complement of a simple closed curve in the plane has two connected components, and the curve is the boundary of each component. Experience strongly suggests that such a result is true, but even in simple cases like regular smooth curves the proof is challenging (for example, see the proof in M. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, 1976). There is a proof of this result in Munkres which does not use homology theory, but it is long and delicate. We shall use homology theory to give a fairly short proof of the Jordan Curve Theorem and its higher dimensional generalizations; one needs the full force of homology theory for the latter, for they cannot be proved using the concepts in Munkres' book.

Likewise, homology theory provides a very simple proof that the coordinate spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic if $m \neq n$. Once again, the material in Munkres yields these results if $n = 1$ or 2 but cannot be used to draw any conclusions if $m, n \geq 3$. We shall also use homology theory to give alternate proofs for two results of Munkres about graphs which are not topologically embeddable in $\mathbb{R}^2$ (although we know that all graphs are nicely embeddable in $\mathbb{R}^3$). Finally, if time permits we shall use homology theory to derive a classical formula of R. Descartes and L. Euler relating the numbers of edges, vertices and faces in a polyhedron which bounds a convex linear cell in $\mathbb{R}^3$:

$$E + 2 = V + F$$

Many additional uses of homology theory are mentioned very briefly in `morgan-lamberson.pdf`.

VII.1 : Consequences of the axioms

(**H**, §§2.1–2.3, 2.B)

Our first objective is to show that the coordinate spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ are not homeomorphic if $m \neq n$. For the sake of clarity and convenience we begin by showing that certain convex sets in $\mathbb{R}^n$ are homeomorphic. We have already stated more general results and referred to files for their proofs, but it seems worthwhile to give direct, simple proofs for the specific examples of interest to us here.

Semi-explicit homeomorphisms of various convex sets

The sets of interest to us are the n-simplex E_n in $\mathbb{R}^n$ given by the inequalities

$$x_i \geq 0 \ , \qquad \sum_i x_i \geq 1$$

the hypercubes $[a, b]^n$ which are homeomorphic to each other because all closed intervals in $\mathbb{R}$ are homeomorphic, and the usual unit disk D^n.

THEOREM 1. *All of the sets listed above are homeomorphic such that interior points of one correspond to interior points of the other and boundary points of one correspond to boundary points of the other.*

Proof. We begin with the easiest pair; namely, the disk and the hypercube $[-1,1]^n$. Given a vector $x \in \mathbb{R}^n$, let $|x|_2$ denote its length with respect to the usual inner product and let $|x|_\infty$ be the maximum of the absolute values of the coordinates ($= \max_i |x_i|$). Both of these define norms on $\mathbb{R}^n$, and the unit disks with respect to these norms are D^n and $[-1,1]^n$ respectively. If one defines a map f of $\mathbb{R}^n$ to itself by $f(\mathbf{0}) = \mathbf{0}$ and by

$$f(x) \quad = \quad \frac{|x|_\infty}{|x|_2} \cdot x$$

if $x \neq \mathbf{0}$, then it follows that f is 1–1 onto and a homeomorphism except possibly at $\mathbf{0}$, and that for each $r > 0$ the map f sends points satisfying $|x|_2 = r$ to points satisfying $|x|_\infty = r$; one can check continuity of f and its inverse at $\mathbf{0}$ using the elementary inequalities

$$|x|_\infty \quad \leq \quad |x|_2 \quad \leq \quad n \cdot |x|_\infty \ .$$

It follows that f defines a homeomorphism from D^n to $[-1,1]^n$.

Since all n-dimensional hypercubes are homeomorphic, it will suffice to show that E_n is homeomorphic to the hypercube $[0,1]^n$ such that their boundaries correspond. For this we need the "taxicab norm" $|x|_1 = \sum_i |x_i|$. Let F_n be the unit disk with respect to this norm. Then E_n and $[0,1]^n$ are the intersections of the unit disks F_n and $[-1,1]^n$ with the *closed first orthant* in $\mathbb{R}^n$ defined by the inequalities $x_n \geq 0$. In analogy with the previous paragraph define a mapping g by $g(\mathbf{0}) = \mathbf{0}$ and

$$g(x) \quad = \quad \frac{|x|_\infty}{|x|_1} \cdot x$$

if $x \neq \mathbf{0}$. Then g has similar properties to f, with continuity at $\mathbf{0}$ is true because of the inequalities

$$|x|_\infty \quad \leq \quad |x|_1 \quad \leq \quad n \cdot |x|_\infty \ .$$

By construction, both f and g map the first orthant into itself such that the boundary points (those for which some coordinate $x_i = 0$) are sent to themselves. The boundaries of E_n and $[0,1]^n$ are given by their intersections with the orthants and their intersections with the sets $|x|_p = 1$ where $p = 1$ and ∞ respectively, and therefore it follows that g defines the desired homeomorphism from E_n to $[0,1]^n$.∎

Some nonhomeomorphic spaces

THEOREM 2. *If m and n are distinct positive integers, then S^m and S^n are not homeomorphic, and similarly $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic.*

Proof. By Theorem 1 we know that S_k is homeomorphic to the boundary of the simplex E_{k+1}, and hence $H_q(S^k) = \mathbb{Z}$ if $q = 0, k$ and zero otherwise. In particular, this means that the homology groups of S^m and S^n are not isomorphic if $m \neq n$, so the spaces cannot be homeomorphic.

If $\mathbb{R}^m$ and $\mathbb{R}^n$ were homeomorphic, then it follows that their one point compactifications would also be homeomorphic (verify this as a general statement about locally compact Hausdorff spaces!). Since these one point compactificatons are homeomorphic to S^m and S^n respectively, it follows that $\mathbb{R}^m$ and $\mathbb{R}^n$ cannot be homeomorphic if $m \neq n$.∎

Local homology at a point

Intuitively it is clear that a closed interval is not homeomorphic to a Y shaped graph because the latter has a vertex which lies on exactly three edges. Similarly, in Munkres it is noted that a figure eight space (8) is not homeomorphic to a figure theta (θ) space even though they are homotopy equivalent, and one expects this because the first space has a graph decomposition for which there is a vertex lying on four edges and the analogous statement for the second space appears to be false. These statements appear to reflect something about the topological nature of neighborhoods of points in a space. Local homology provides an efficient means for handling such problems.

Definition. Let X be a Hausdorff topological space, and let $x \in X$. The *local homology groups of x in X* are given by $H_*(X, X - \{x\})$.

These groups have the following important properties:

PROPOSITION 3. (Localization property) *Let $x \in X$ where X is Hausdorff, and let U be an open neighborhood of x. Then the inclusion map of pairs induces isomorphisms from $H_q(U, U - \{x\})$ to $H_q(X, X - \{x\})$ for all integers q.*

Proof. This is an immediate consequence of the excision axiom (E.1) to U and $V = X - \{x\}$, for then $X = U \cup V$ and $U \cap V = U - \{x\}$.∎

PROPOSITION 4. (Topological invariance) *If X and Y are Hausdorff spaces with $x \in X$, and if $f : X \to Y$ is a homeomorphism, then there is an isomorphism of local homology groups from $H_*(X, X - \{x\})$ to $H_*(Y, Y - \{f(x)\})$.*

Proof. The homeomorphism f induces a homeomorphism of pairs $(X, x - \{x\}) \cong (Y, Y - \{y\})$, so the associated homology groups must be isomorphic.∎

For computational purposes the following result is very helpful when working with local homology groups:

LEMMA 5. *Suppose that $B \subset A \subset X$ and B is a deformation retract of A. Then the inclusion map of pairs induces isomorphisms from $H_*(X, B)$ to $H_*(X, A)$.*

Proof. By the exactness axiom (B.2) we have the following commutative diagram in which the rows are long exact homology sequences:

$$\begin{array}{ccccccccccccc}
\cdots & H_{q+1}(X,B) & \xrightarrow{\partial} & H_q(B) & \xrightarrow{i_*} & H_q(X) & \xrightarrow{j_*} & H_q(X,B) & \xrightarrow{\partial} & H_{q-1}(B) & \cdots \\
\cdots & \downarrow g_* & & \downarrow f_* & & \downarrow = & & \downarrow g_* & & \downarrow f_* & \\
\cdots & H_{q+1}(X,A) & \xrightarrow{\partial} & H_q(A) & \xrightarrow{i_*} & H_q(X) & \xrightarrow{j_*} & H_q(X,A) & \xrightarrow{\partial} & H_{q-1}(A) & \cdots
\end{array}$$

The mappings f and g are the associated inclusions of spaces or pairs. Since B is a deformation retract of A the maps f_* are isomorphisms, and of course the identity maps on $H_*(X)$ are also isomorphisms. Therefore the Five Lemma implies that the mappings g_* are also isomorphisms.∎

Example. The hypotheses of the proposition do not imply that the inclusion $(X, B) \subset (X, A)$ is a homotopy equivalence of pairs; as noted in Hatcher, the inclusion $(D^n, S^{n-1}) \subset (D^n, D^n - \{0\})$ satisfies the hypothesis but this map is not a homotopy equivalence of pairs.

Application to graphs

If $(X, \mathcal{E})$ is a graph then it is easy to compute the local homology of X at all points.

THEOREM 6. *Let $(X, \mathcal{E})$ be a connected graph, and let $x \in X$. Then the local homology group $H_1(X, X - \{x\})$ is given as follows:*

(i) *If x is not a vertex for $\mathcal{E}$ then $H_1(X, X - \{x\}) \cong \mathbb{Z}$.*

(ii) *If x is a vertex which lies on exactly n edges, then $H_1(X, X - \{x\}) \cong \mathbb{Z}^{n-1}$.*

In particular, if $n_k(X, \mathcal{E})$ is the number of vertices which lie on k vertices and $k \neq 2$, then Theorem 6 implies that $n_k(X, \mathcal{E})$ depends only upon the topological space X because it is the number of points in X for which the 1-dimensional local homology is isomorphic to $\mathbb{Z}^{k-1}$. Stated differently, if the underlying spaces of the connected graphs $(X, \mathcal{E})$ and $(X', \mathcal{E}')$ are homeomorphic, then $n_k(X, \mathcal{E})$ and $n_k(X', \mathcal{E}')$. One can apply this very easily to determine whether graphs corresponding to various letters of the alphabet are homeomorphic to each other, and there is a problem of this type in the exercises.

Proof of Theorem 6. By the Localization Property it suffices to compute the relative groups $H_1(U, U - \{x\})$ where U is some open neighborhood of x.

(i) Suppose that x lies in the edge E but is not an endpoint, and let F be the union of all the edges except E together with the vertices of E, and let $U = X - F$; then U is open and contains x, and the pair $(U, U - \{x\})$ is homeomorphic to $(V, V - \{t\})$ where V is the open unit interval $(0, 1)$ and $t \in V$. The local homology of the latter pair can be studied using the tail end of the long exact homology sequence:

$$\cdots \to 0 = H_1(V) \to H_1(V, V - \{t\}) \to H_0(V - \{t\}) \to H_0(V) = \mathbb{Z}$$

The homology groups of V are given as in this sequence because V is convex and hence contractible. Since the space $V - \{t\}$ has two components, axiom (D.2) and Proposition VI.3.1 implies that $H_0(V - \{t\}) \cong \mathbb{Z}^2$ and each free generator of the latter maps onto a free generator of $H_0(V)$. It follows immediately that the local homology group must be isomorphic to $\mathbb{Z}$.

(ii) Let V be the open star on the vertex x as defined in Unit III. Then $\{x\}$ is a deformation retract of V by Proposition III.1.7 and $V - \{x\}$ is homeomorphic to a union of pairwise disjoint subsets $V_j = E_j - [\textit{endpoints}]$, where E_j runs through the n edges which have x as one of their endpoints. Since V is contractible one has an exact sequence for computing $H_1(V, V - \{t\})$ just like the one in the preceding paragraph. However, in this case we know that that $H_0(V - \{t\}) \cong \mathbb{Z}^n$ and each free generator of the latter maps onto a free generator of $H_0(V)$. It follows immediately that the local homology group must be isomorphic to $\mathbb{Z}^{n-1}$.■

Local homology also yields the following strengthening of Theorem 2.

THEOREM 7. (Invariance of dimension, L. E. J. Brouwer) *Let U and V be nonempty open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. If U is homeomorphic to V, then $m = n$.*

Proof. It suffices to prove that if W is a nonempty open subset of $\mathbb{R}^k$ then $H_k(W, W - \{p\}) \cong \mathbb{Z}$ and $H_j(W, W - \{p\}) = 0$ if $j \neq k$. Thus the local homology groups at points $x \in U$ and $y \in V$ are not isomorphic if $m \neq n$, and accordingly U and V cannot be homeomorphic in that case.

By the localization property it suffices to prove that the local homology groups $H_*(\mathbb{R}^k, \mathbb{R}^k - \{p\})$ are given as in the preceding paragraph. Furthermore, it suffices to consider the case where $p = \mathbf{0}$, for the translation map $T(x) = x + p$ is a homeomorphism which induces isomorphisms $H_*(\mathbb{R}^k, \mathbb{R}^k - \{\mathbf{0}\}) \cong H_*(\mathbb{R}^k, \mathbb{R}^k - \{p\})$.

It will be convenient to treat the case $k = 1$ separately. We can use the argument in the first part of Theorem 6 to prove that $H_1(\mathbb{R}, \mathbb{R} - \{0\}) \cong \mathbb{Z}$. Since $\mathbb{R} - \{0\}$ has the homotopy type of S^0

it follows that its homology vanishes in all dimensions except zero, so for each $q \geq 2$ we have the following exact sequence in which all terms except the relative group are known to be zero:

$$0 = H_q(\mathbb{R}) \to H_q(\mathbb{R}, \mathbb{R} - \{0\}) \to H_{q-1}(\mathbb{R} - \{0\}) \to H_{q-1}(\mathbb{R}) = 0$$

It follows that $H_q(\mathbb{R}, \mathbb{R} - \{0\})$ must also be zero if $q \geq 2$. Finally the 0-dimensional relative homology is given by the following piece of the long exact homology sequence

$$H_0(\mathbb{R} - \{0\}) \to H_0(\mathbb{R}) \to H_0(\mathbb{R}, \mathbb{R} - \{0\}) \to 0$$

since homology groups vanish in negative dimensions. We already know that the map at the left of this exact sequence is onto, and by exactness it follows that the second map is zero and the third is 1–1. These combine to imply that $H_0(\mathbb{R}, \mathbb{R} - \{0\}) = 0$.

Since S^{k-1} is a deformation retract of $\mathbb{R}^k - \{\mathbf{0}\}$ it follows that the homology groups of the latter are $\mathbb{Z}$ in dimensions $0, k-1$ and zero otherwise. If $q > 0$, then we have the exact homology sequence

$$0 = H_q(\mathbb{R}^k) \to H_q(\mathbb{R}^k, \mathbb{R}^k - \{\mathbf{0}\}) \to H_{q-1}(\mathbb{R}^k - \{\mathbf{0}\}) \to H_{q-1}(\mathbb{R}^k) \ .$$

If $q \geq 2$ then the groups at the end of this sequence are both zero, and therefore we have

$$H_q(\mathbb{R}^k, \mathbb{R}^k - \{\mathbf{0}\}) \quad \cong \quad H_{q-1}(\mathbb{R}^k - \{\mathbf{0}\})$$

for $q \geq 2$. This yields the conclusion of the theorem except in the cases $q = 0, 1$. The 0-dimensional case can be established by the same argument employed in the previous paragraph, so we are left with the 1-dimensional case, for which we have the following exact sequence:

$$0 = H_1(\mathbb{R}^k) \to H_1(\mathbb{R}^k, \mathbb{R}^k - \{\mathbf{0}\}) \to H_0(\mathbb{R}^k - \{\mathbf{0}\}) \to H_0(\mathbb{R}^k) = \mathbb{Z}$$

Since $k \geq 2$ the space $\mathbb{R}^k - \{\mathbf{0}\}$ is connected and hence the map at the right is an isomorphism. This implies that the map in the middle is zero and hence the map on the left is onto. Since the domain of the latter map is zero, it follows that $H_1(\mathbb{R}^k, \mathbb{R}^k - \{\mathbf{0}\}) = 0$.∎

Further nonhomeomorphism theorems

We know that two spheres and two Euclidean spaces of different dimensions cannot be homeomorphic, and it is natural to ask similar questions about other familiar pairs like D^m and D^n or $\mathbb{R}^m_+$ and $\mathbb{R}^n_+$ where $\mathbb{R}^k_+$ denotes the points in $\mathbb{R}^k$ whose first coordinate is nonnegative. Invariance of domain provides effective criteria for dealing with such questions, and the following result will show that the paired spaces cannot be diffeomorphic if their dimensions are unequal:

THEOREM 8. *Suppose that X_m and X_n are subspaces of some $\mathbb{R}^p$ and for $k = m, n$ the set X_k has an open dense subset which is homeomorphic to an open subset of $\mathbb{R}^k$. If X_m and X_n are homeomorphic, then $m = n$.*

Proof. Let $U_k \subset X_k$ be the open dense subset, and let $h : X_m \to X_n$ be the homeomorphism. Then $h^{-1}[U_n]$ is dense in X_m because h is a homeomorphism, and it is open by the continuity of h, and $h[U_m]$ has analogous properties. It follows that the intersections $U_m \cap h^{-1}[U_n]$ and $h[U_m] \cap U_n$ are open subsets of U_m and U_n respectively and hence are homeomorphic to open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. These sets are homeomorphic because h maps the first to the second, and therefore Invariance of Dimension implies that $m = n$.∎

VII.2 : Nonretraction and fixed point theorems

(**H**, §2.B; **M**, §55)

Recall that a continuous mapping $f : X \to Y$ is a **retract** if there is a continuous mapping $g : Y \to X$ such that $g \circ f$ is the identity on X, and a continuous mapping $p : X \to Y$ is a **retraction** if there is a continuous mapping $q : Y \to X$ such that $p \circ q$ is the identity on Y. It follows that the mappings q and f are 1–1 and the mappings p and g are onto; also, q is a retract and g is a retraction. Many but not all subspace inclusion mappings are retracts, and the following result shows the special natures of retracts.

PROPOSITION 1. *Suppose that the mapping $f : X \to Y$ is a retract. Then the induced maps in homology f_* are injections onto direct summands.*

Proof. The composite $g_* \circ f_* = (g \circ f)_*$ is the identity on $H_*(X)$. Therefore the identity $g_* f_*(a) = a$ implies that the kernel of f_* is zero so that f_* is 1–1. One can then check directly that each group $H_q(Y)$ is isomorphic to the direct sum of the image of f_* and the kernel of g_*.■

COROLLARY 2. *For every $k \geq 2$, the sphere S^{k-1} is not a retract of the disk D^k.*

This is true because the homology groups of D^k vanish in all positive dimensions but the group $H_{k-1}(S^{k-1})$ is nonzero.■

The Brouwer Fixed Point Theorem

At this point it is almost traditional to state and prove the Brouwer Fixed Point Theorem.

THEOREM 3. (Brouwer Fixed Point Theorem) *For all $n \geq 0$ every continuous map $f : D^n \to D^n$ has a* fixed point*; in other words, there is a point $\mathbf{x}$ in D^n such that $f(\mathbf{x}) = \mathbf{x}$.*

Proof. If $n = 1$ this is a fairly simple exercise in point set topology. Suppose that $i : S^0 \to D^1$ is the inclusion mapping and $r : D^1 \to S^0$ is such that $r \circ i$ is the identity on S^0. As noted before, it follows that r is onto; since D^1 is connected but S^0 is not, this is impossible and consequently there cannot be a continuous mapping r such that $r \circ i$ is the identity.

In the remaining cases, the standard proof is analogous to the argument on page 32 of Hatcher for the case $n = 2$. Assume that there is a continuous mapping $f : D^n \to D^n$ with no fixed point, so that $f(x) \neq x$ for all x. For each point $x \in D^n$, let $r(x) \in S^{n-1}$ be the unique point where the ray from $f(x)$ through x meets the boundary sphere. By construction $r(x) = x$ if $x \in S^{n-1}$, and if r is continuous then it follows that S^{n-1} is a retract of D^n because $r \circ i$ is the identity.

The argument in Hatcher states that the "continuity of r is clear"; although this seems reasonable, it is still necessary to check the continuity of the geometrically described mapping r retraction explicitly. For the sake of completeness we have written up the details in `brouwer.pdf`.

An application to matrices with nonnegative entries

The Brouwer Fixed Point Theorem and its generalizations play important roles in many branches of mathematics and their applications to other subjects. Here is one online reference for further information:

`http://en.wikipedia.org/wiki/Brouwer_fixed-point_theorem`

In these notes we shall only use the theorem to prove a result on eigenvalues and eigenvectors of matrices. The first step is the following elementary fact:

LEMMA 4. *Let X be a topological space which is homeomorphic to D^n for some $n \geq 0$. Then every continuous map $f : X \to X$ has a fixed point.*

Proof. Let $f : X \to X$ be continuous, and let $h : X \to D^n$ be a homeomorphism. Then $h \circ f \circ h^{-1}$ is a continuous map from D^n to itself and thus has a fixed point $\mathbf{p}$ by Brouwer's Theorem. In other words we have $h \circ f \circ h^{-1}(\mathbf{p}) = \mathbf{p}$. If we take $\mathbf{q} = h(\mathbf{p})$, straightforward computation shows that $f(\mathbf{q}) = \mathbf{q}$.■

THEOREM 5. (Perron – Frobenius) *Let $n > 1$, and let A be an $n \times n$ matrix which is invertible and has nonnegative entries. Then A has a positive eigenvalue λ such that λ has a nonzero eigenvector with nonnegative entries.*

Proof. Recall that the 1-norm on $\mathbf{R}^n$ is defined by $|\mathbf{x}|_1 = \sum_j |x_j|$, where the coordinates of $\mathbf{x}$ are given by $x_1, \cdots, x_n$. For each $\mathbf{x} \in \Delta_n$ (the standard simplex whose vertices are the unit vectors), define

$$f(\mathbf{x}) \quad = \quad (\,|A\mathbf{x}|_1)^{-1} \cdot A\mathbf{x} \; .$$

Observe that the coordinates of $A\mathbf{x}$ are all nonnegative because the entries of A and the coordinates of $\mathbf{x}$ are nonnegative, this vector is nonzero because A is invertible, and if $\mathbf{y}$ is a nonzero vector with nonnegative entries then $|\mathbf{y}|_1^{-1}\mathbf{y}$ must lie in Δ_n. Therefore we indeed have a continuous map f from the simplex to itself.

By the lemma, we know that f has a fixed point; in other words, there is some $\mathbf{v} \in \Delta_n$ such that

$$\mathbf{v} \quad = \quad (\,|A\mathbf{v}|_1)^{-1} \cdot A\mathbf{v}$$

and since the latter is equivalent to saying that $A\mathbf{v}$ is a positive multiple of $\mathbf{v}$, this completes the proof.■

COROLLARY 6. *In the setting of the theorem above, if all the entries of the matrix A are positive, then the eigenvector has positive entries.*

Proof. Let $\mathbf{y}$ be the eigenvector obtained in the theorem. Since $A\mathbf{y}$ is a positive scalar multiple of $\mathbf{y}$, it will suffice to prove that the entries of $A\mathbf{y}$ are all positive. But these entries are given by expressions of the form

$$z_i \quad = \quad \sum_j \; a_{i,j} y_j$$

and if we choose k such that $y_k \neq 0$ then it follows that $z_i \geq a_{i,k} y_k$; the right hand side is a product of

VII.3 : Separation and invariance theorems

(**H**, §2.B; **M**, §63)

Most of this has become standard in algebraic topology texts, and we shall quote Hatcher as appropriate. The following result corresponds to the first half of Proposition 2B.1 on page 169 of that reference.

PROPOSITION 1. *If $A \subset S^n$ is homeomorphic to D^k for some $k < n$, then the $H_i(S^n - A)$ is infinite cyclic if $i = 0$ and trivial otherwise.*

Note. The hypotheses imply that A must be a proper subset of $\mathbb{R}^n$ because the homology groups of A and S^n are not isomorphic.

Proof. The proof proceeds by induction on k. If $k = 0$ then $S^n - A$ is homeomorphic to $\mathbb{R}^n$ and the conclusion in this case follows immediately. Assume now that the result is known whenever a subset A is homeomorphic to D^{k-1} for some k satisfying $1 \leq k \leq n$, and assume now that $A \subset S^n$ is homeomorphic to D^k.

The homeomorphisms of Section VII.1 imply that D^n is homeomorphic to $D^{n-1} \times [0,1]$. If $t \in [0,1]$ let $A_t \subset A$ correspond to $D^{n-1} \times \{t\}$ under some fixed homeomorphism $A \cong D^{n-1} \times [0,1]$, and if $a < b$ let $A[a,b] \subset A$ correspond to $D^{n-1} \times [a+, b-]$, where $a+$ is the larger of a and 0, and $b-$ is the smaller of b and 1.

Suppose now that $u \in H_q(S^n - A)$ lies in the kernel of the homomorphism $c_* : H_q(S^n - A) \to H_q(P)$, where P is a one point space and $c : S^n - A \to P$ is the constant map. We want to show that $u = 0$; the compact supports property implies the existence of a compact subset $L \subset S^n - A$ such that u lies in the image of the map $H_q(L) \to H_q(A)$ induced by inclusion; let u' be a class which maps to u in this fashion. If

$$j_t : S^n - A \longrightarrow S^n - A_t$$

is the inclusion mapping, then the inductive hypothesis implies that $j_{t*}(u) = 0$ for each $t \in [0,1]$.

By the compact supports property, for each t there is some compact subset K_t such that $L \subset K_t \subset S^n - A_t$ such that u maps to zero under the homology map associated to the inclusion $H_q(L) \to H_q(K_t)$. Since A_t and K_t are disjoint compact subsets of S^n, there is some $\varepsilon(t) > 0$ such that K_t and $A_{[t-\varepsilon(t), t+\varepsilon(t)]}$ are disjoint. It follows that the image of u in the homology of $H_q(A[t - \varepsilon(t), t + \varepsilon(t)])$ is zero.

The open or half open intervals $(t - \varepsilon(t), t_\varepsilon(t)) \cap [0,1]$ form an open covering of $[0,1]$, so by the Lebesgue Covering Lemma there is some $M > 0$ such that every closed interval of length $\leq 1/M$ lies in some subset of this open covering. It follows that u maps to zero in each of the sets $H_q(S^n - A[j - 1/M, j/M])$ where $j = 1, ..., M$. The next objective is to show by induction on j that u maps to zero in $H_q(S^n - A[0, j/M])$; the case $j = 1$ is known by the preceding discussion, and when $j = M$ the set $A[0, j/M]$ is all of A.

Assume now that u maps to zero in $H_q(S^n - A[0, j/M])$ where $1 \leq j \leq M - 1$. Then the identities

$$A[0,(j+1)/M] \;=\; A[0,j/M] \cup A[j/M,(j+1)/M]\,, \qquad A_{j/M} \;=\; A[0,j/M] \cap A[j/M,(j+1)/M]$$

and their complementary analogs

$$S^n - A[0,(j+1)/M] \;=\; (S^n - A[0,j/M]) \cap (S^n - A[j/M,(j+1)/M])$$

$$S^n - A_{j/M} \;=\; (S^n - A[0,j/M]) \cup (S^n - A[j/M,(j+1)/M])$$

and the latter determine a long exact Mayer-Vietoris sequence. Consider the following piece of that sequence:

$$H_{q+1}(S^n - A_{j/M}) \to H_q(S^n - A[0,(j+1)/M]) \to H_q(S^n - A[0,j/M]) \oplus H_q(S^n - A[j/M,(j+1)/M])$$

If $u' \in H_q(S^n - A[0,(j+1)/M])$ is the image of u under the map induced by the inclusion $S^n - A \subset S^n - A[0,(j+1)/M]$, then the inductive hypotheses and the previous arguments show that u' maps to zero in each of the groups $H_q(S^n - A[0,j/M])$ and $H_q(S^n - A[j/M,(j+1)/M])$. Therefore by exactness u' lies in the image of $H_{q+1}(S^n - A_{j/M})$. Since the latter group vanishes, it follows that u' must be zero, completing the inductive argument with respect to j. As noted in the preceding paragraph, this implies that $u = 0$ and completes the inductive argument with respect to the dimension k such that $A \cong D^k$.■

The Jordan-Brouwer Separation Theorem

If we remove a circle from the plane, we obtain two connected regions — an interior and an exterior region — and mathematically these regions are defined by the inequalities $|x - a| < r$ and $|x - a| > r$, where a is the center of the circle and r is its radius. Similarly, if we are given a relatively simple example of a simple closed curve in the plane, it is generally easy to see that the complement is a union of two disjoint connected components, and it is natural to conjecture that the same is true for an arbitrary simple closed curve in the plane. However, as curves become more complicated it becomes increasingly difficult to verify this explicitly for examples like the curve in the file `fishmaze.pdf` (note that the bounded component of the curve's complement is the pink shaded region in `fishmaze2.pdf`). The online article

`http://en.wikipedia.org/wiki/Jordan_curve_theorem`

discusses the history of this result fairly extensively (and corrects some widely circulated misinformation), and the article is definitely worth reading. Everyday experience with geometric objects in 3-space strongly suggests that there are analogous results for suitably defined closed surfaces in $\mathbb{R}^3$ which include subsets homeomorphic to S^2, and the Jordan-Brouwer Separation Theorem generalizes the Jordan Curve Theorem to subsets of $\mathbb{R}^n$ which are homeomorphic to S^{n-1} for all values of n.

Our statement the Jordan-Brouwer Theorem contains a somewhat stronger conclusion than the version in Hatcher; the case $n = 2$ is the classical Jordan Curve Theorem.

THEOREM 2. (Jordan-Brouwer Separation Theorem.) *Let $n \geq 2$, and suppose that $A \subset S^n$ is homeomorphic to S^{n-1}. Then $S^n - A$ contains two components, and A is the frontier of each component.*

Note. In the discussion preceding the statement of this theorem we have considered compact subsets $A \subset \mathbb{R}^n$ which are homeomorphic to S^{n-1}, and in fact the analogous conclusion for subsets of $\mathbb{R}^n$ follows from the theorem by passing to one point compactifications. In fact, one can say slightly more; namely, exactly one of the components of $\mathbb{R}^n - A$ must be a bounded open subset (specifically, the component not containing the point at infinity in the one point compactification of $\mathbb{R}^n$ under the identification of the latter with S^n).

It is natural to ask if one also has similar results if A is homeomorphic to some other closed surface such as the torus T^{n-1}, and the answer is that similar conclusions hold more generally. In particular this follows from results on the homology of compact manifolds and the Alexander Duality Theorem in Section 3.3 of Hatcher.

The standard textbook proof of the Jordan-Brouwer Separation Theorem involves proving the following complementary result on the homology of subsets of S^n which are homeomorphic to spheres of dimension $\leq n - 2$.

PROPOSITION 3. *Let $A \subset S^n$ be homeomorphic to S^k where $0 \leq k \leq n - 2$. Then the homology groups of $S^n - A$ are homeomorphic to the homology groups of S^{n-k-1}.*

It is important to note that the complement does not necessarily have the homotopy type of S^{n-k-1}. In particular, there are simple closed (knotted) curves K in $\mathbb{R}^3$ and S^3 for which the fundamental group is nonabelian and hence not isomorphic to $\pi_1(S^1)$; there is an extensive theory of knotted curves in 3-space which goes far beyond the scope of this course, and currently there is a high level of activity aimed at answering many open questions about such curves.

On the other hand, for the standard linear embedding of S^k in S^n corresponding to

$$S^k \subset \mathbb{R}^{k+1} = \mathbb{R}^{k+1} \times \{\mathbf{0}\} \subset \mathbb{R}^{k+1} \times \mathbb{R}^{n-k} \cong \mathbb{R}^{n+1}$$

(note that the image is contained in S^n), the complement is homeomorphic to $S^{n-k-1} \times \mathbb{R}^{k+1}$ such that $S_{n-k-1} \times \{\mathbf{0}\}$ corresponds to the unit sphere in $\{\mathbf{0}\} \times \mathbb{R}^{n-k}$ (see the file `sphere-complement.pdf`).

Proof of Proposition 3. The proof proceeds by induction on k, so we need to start by verifying the result in that case, in which A consists of two points. Since S^n is highly symmetric we can assume that one of the points is the unit vector $\mathbf{e}_{n+1} \in \mathbb{R}^{n+1}$, which implies that if one point of A is removed the complement is homeomorphic to $\mathbb{R}^n$. If we now remove the second point we are left with a subset homeomorphic $\mathbb{R}^n - \{p\}$, and since the latter is homeomorphic to $S^{n-1} \times \mathbb{R}$ the conclusion about homology groups in this case follows immediately.

Suppose now that the result is known for subsets homeomorphic to S^{k-1}, where $1 \leq k \leq n-2$. Let A be a subset which is homeomorphic to S^n, and let $A_\pm \subset A$ be the subspace corresponding to the hemisphere $D^n_\pm \subset S^n$ defined by the coordinate inequalities $x_{n+1} \geq 0$ (for D^n_+) and $x_{n+1} \leq 0$ (for D^n_-). Let $A_0 = A_+ \cap A_-$, so that A_0 is homeomorphic to S^{k-1}. Consider now the Mayer-Vietoris sequence for the decomposition

$$S^n - A_0 \quad = \quad (S^n - A_+) \cup (S^n - A_-) \;, \qquad \text{where} \quad S^n - A \quad = \quad (S^n - A_+) \cap (S^n - A_-) \;.$$

We are particularly interested in the following pieces of this exact sequence:

$$H_{q+1}(S^n - A_+) \oplus H_{q+1}(S^n - A_-) \to H_{q+1}(S^n - A_0) \to H_q(S^n - A) \to H_q(S^n - A_+) \oplus H_q(S^n - A_-)$$

If $q > 0$ then the first and last terms of this exact sequence are zero by Proposition 1, and hence the map from the second term to the third is an isomorphism. By induction we know that $H_{q+1}(S^n - A_0)$ is trivial for all $q \geq 1$ except $q+1 = n-(k-1)-1$ and is infinite cyclic in the latter case. Therefore we have $H_q(S^n - A) = 0$ if $q > 0$ and $q \neq n-k-1$, and in the latter case we have $H_q(S^n - a) \cong \mathbb{Z}$.

Suppose now that $q = 0$ in the displayed exact sequence. Then $H_1(S^n - A_0) = 0$ and $H_0(S^n - A_0) = \mathbb{Z}$ because $k - 2 \leq n$, and therefore the extended Mayer-Vietoris sequence reduces to

$$0 = H_1(S^n - A_0) \to H_0(S^n - A) \to H_0(S^n - A_+) \oplus H_0(S^n - A_-) \cong \mathbb{Z} \oplus \mathbb{Z} \to H_0(S^n - A_0) = Z$$

where the map at the right is surjective, so that its kernel is isomorphic to $\mathbb{Z}$. By exactness this kernel is the image of $H_0(S^n - A)$, and the mapping from the latter onto the kernel is 1–1, so that we have $H_0(S^n - A) \cong \mathbb{Z}$. This completes the proof of the inductive step.∎

Proof of the Jordan-Brouwer Separation Theorem. The first step is to prove that the complement has exactly two components. Let $A_\pm$ and A_0 be defined as in the preceding proposition and consider the corresponding Mayer-Vietoris sequence; in particular, we are interested in the following piece:

$$0 = H_1(S^n - A_+) \oplus H_1(S^n - A_-) \to H_1(S^n - A_0) \to H_0(S^n - A) \to \;\; (\textit{next line})$$

$$H_0(S^n - A_+) \oplus H_0(S^n - A_-) \cong \mathbb{Z} \oplus \mathbb{Z} \to H_0(S^n - A_0) = Z$$

In this case we know that $H_1(S^n - A_0) \cong \mathbb{Z}$ and hence the latter maps injectively into $H_0(S^n - A)$. Furthermore, we can use the same argument as in Proposition 3 to conclude that the image of $H_0(S^n - A)$ in the direct sum is also isomorphic to $\mathbb{Z}$, and therefore by exactness we must have $H_0(S^n - A) \cong \mathbb{Z} \oplus \mathbb{Z}$, so that $S^n - A$ has exactly two components.

It remains to prove that points of A are limit points of each component. Suppose that $S^n - A$ is the union of the two open, connected, disjoint subsets U and V.

Assume that not every point of A is a limit point of both U and V. Without loss of generality, it is enough to consider the case where $x \in A$ is not a limit point of V. Since $x \notin V$, it follows that there is some open set W_0 in S^n such that $x \in W_0$ and $W_0 \cap V = \emptyset$.

Consider the open set $W_0 \cap A$ in A; since the latter is homeomorphic to S^{n-1}, it follows that there is a subneighborhood of the form $A - E$, where $E \subset A$ is homeomorphic to a closed $(n-1)$-disk and $A - E$ is homeomorphic to an open $(n-1)$-disk centered at x. If $W = W_0 \cap S^n - E$, then W is still open in S^n and we still have $x \in W$ and $W \cap V = \emptyset$.

By construction we have $S^n - E = U \cup A - E \cup V$ where the pieces are pairwise disjoint. Furthermore, we have $A - E \subset W$ and hence $U \cup W$ is an open set of $S^n - E$ which is disjoint from V and contains U and $A - E$. Therefore it follows that $S^n - E$ is a union of the nonempty disjoint open sets $U \cup W$ and V and hence is disconnected. On the other hand, since E is homeomorphic to a closed disk we know that $S^n - E$ is connected, so we have a contradiction. The source of this contradiction was our assumption that x was not a limit point of V, and hence this must be false. Therefore x must be a limit point of V, and as noted above it follows that every point of A is a limit point of both U and V.■

The Mayer-Vietoris sequence in Theorem 2 also has the following implication; details of the proof are left to the reader (remember that the homology groups of $S^n - A_\pm$ and $S^n - A_0$ in positive dimensions are known to vanish except for $H_1(S^n - A_0)$):

COROLLARY 4. *In the setting of Theorem* 2 *the homology groups of each component of* $S^n - A$ *are zero in every positive dimension.*■

If $n = 2$ a remarkable theorem of A. Schönflies yields a much stronger conclusion: If U is a component of $S^2 - A$ then its closure $\overline{U}$ is homeomorphic to D^2 such that A corresponds to S^1 (it is also possible to use results from complex variable theory to the prove weaker result that the open set U is simply connected — see the file `ahlfors.pdf`). On the other hand, if $n \geq 3$ then a component U of $S^n - A$ need not even be simply connected. The standard example when $n = 3$ is the Alexander Horned Sphere discussed in Example 2.B.2 on pages 170–172 of Hatcher. The following online site has an interesting video showing the recursive construction of the Alexander sphere:

`http://www.youtube.com/watch?v=Pe2mnrLUYFU`

With the preceding results at our disposal, we can prove the following basic result exactly as in Hatcher:

THEOREM 5. (Invariance of Domain, Brouwer) *Let* U *be an open subset of* $\mathbb{R}^n$ *for some* $n \geq 2$, *and let* $h : U \to \mathbb{R}^n$ *be continuous and* $1-1$. *Then* h *is an open mapping, the image* $h[U]$ *is an open subset of* $\mathbb{R}^n$, *and* h *maps* U *homeomorphically onto* $h[U]$.

The name of the result refers to the fact that if V is homeomorphic to an open subset of $\mathbb{R}^n$, then V must also be an open subset of $\mathbb{R}^n$.

Proof. It will suffice to prove that h is an open mapping, and to prove the latter it will suffice to show that if $D \subset U$ is an ordinary closed disk of some radius about a point of U and ∂D is the boundary sphere of D, then $f[D - \partial D]$ is an open subset of $\mathbb{R}^n$ (since every open subset of U is a union of open disks that are interiors of closed disks). Since f is 1–1 it follows that f maps D and ∂D homeomorphically onto their images.

As usual, view $\mathbb{R}^n$ as $S^n - \{p\}$ via one point compactification. Then the preceding results imply that $S^n - f[D]$ is a connected open subset and $S^n - f[\partial D]$ is an open subset with two components, say W_1 and W_2; label these so that $S^n - f[D] \subset W_1$. Now we also have

$$S^n - f[\partial D] \;=\; (S^n - f[D]) \,\cup\, (S^n - f[D - \partial D])$$

and the subsets on the right hand side are disjoint; since $f[D - \partial D]$ is connected it is contained in one of the components W_1, W_2. If $f[D - \partial D]$ were contained in W_1 then we would have $S^n - f[\partial D] \subset W_1 \subset S^n - f[\partial D]$ so that the two sets would be equal, contradicting the fact that $S^n - f[\partial D]$ is disconnected. Therefore $f[D - \partial D]$ must be contained in W_2. This gives us the chain of inclusions

$$(S^n - f[D]) \,\cup\, (S^n - f[D - \partial D]) \;\subset\; W_1 \cup W_2 \;\subset\; S^n - f[\partial D]\,, \quad \text{where}$$

$$(S^n - f[D]) \,\cap\, (S^n - f[D - \partial D]) \;=\; \emptyset \;=\; W_1 \cap W_2\,.$$

Since $S^n - f[D] \subset W_1$ and $S^n - f[D - \partial D] \subset W_2$, the set-theoretic relations combine to imply that $S^n - f[D] = W_1$ and $S^n - f[D - \partial D] = W_2$. This proves that $f[D - \partial D]$ is an open subset of $\mathbb{R}^n$ (hence also of U), by the statement at the beginning of the proof this also completes the proof of the theorem.∎

We shall limit ourselves to one simple consequence.

COROLLARY 6. *If $\mathbb{R}^n_+$ is defined to be the set of all points whose last coordinate is nonnegative, then $\mathbb{R}^n_+$ is not homeomorphic to $\mathbb{R}^m$ for any positive integer m.*

Proof. We first consider the cases where $n \le m$. In these cases the sets cannot be homeomorphic by Invariance of Domain because $\mathbb{R}^n_+$ is not an open subset of $\mathbb{R}^m$ (as usual, we identify $\mathbb{R}^n$ with the set of all points in $\mathbb{R}^m$ whose last $m - n$ coordinates are all zero).

Suppose now there is a homomorphism f from $\mathbb{R}^m$ to $\mathbb{R}^n_+$ where $m < n$. If H is the hyperplane in $\mathbb{R}^n$ of all points whose last coordinate is zero and $W = \mathbb{R}^m - f^{-1}[H]$, then f defines a homeomorphism from W to $\mathbb{R}^n_+ - H \cong \mathbb{R}^n$. This is impossible by invariance of dimension, and therefore $\mathbb{R}^n_+$ cannot be homeomorphic to $\mathbb{R}^m$ if $m < n$.∎

VII.4 : Nonplanar graphs

(**M**, §64)

We have already seen that every graph has a nice rectilinear embedding in $\mathbb{R}^3$. In this section we shall use homology theory to prove that some graphs do not admit any topological embeddings into $\mathbb{R}^2$. We shall treat two examples, and at the end of this section we shall explain why they are particularly important. The approach in this section is close to that in Munkres, the main difference being that we use homology theory to give simpler proofs of some key steps in the arguments.

This is a fairly well-known example with three vertices a, b, c representing houses and another three vertices g, w, e representing gas, water and electricity utilities. There are nine edges which they join the individual houses to each of the three utilities, and the question is whether this can be done on a flat surface with none of the lines crossing over or under each other. This example is depicted in Figure 1 of `graphpix4.pdf`, and in the literature of graph theory it is often called $K_{3,3}$. In mathematical terms, here is what we what to prove:

THEOREM 1. *The utilities network $K_{3,3}$ is not homeomorphic to a subset of S^2.*

In fact, one has the same conclusion if S^2 is replaced by $\mathbb{R}^2$ because $K_{3,3}$ and S^2 are not homeomorphic — the quickest way to see this is to note that $H_2(K_{3,3}) = 0$ for dimensional reasons but $H_2(S^2) \cong \mathbb{Z}$.

As suggested by Figure 1 in `graphpix4.pdf`, it is fairly easy to embed the subgraph of $K_{3,3}$ by removing one edge; the point of the proof is that there cannot be some clever way of inserting the remaining edge.

The proof of Theorem 1 involves separation theorems that are similar to the Jordan Curve Theorem but are somewhat more complicated to state and prove. The first of these involves *theta spaces* which can be expressed as unions of three subsets E_1, E_2, E_3 which are all homeomorphic to $[0, 1]$ and whose intersections are given by their endpoints. Figure 2 in `graphpix4.pdf` is a simple but typical example. We want to prove that every theta space in S^2 has the separation properties which are apparent in the figure. This can be stated formally as follows:

PROPOSITION 2. *If $X \subset S^2$ is a theta space with edges E_1, E_2, E_3 meeting at the common endpoints $\{A, B\}$, then $S^2 - X$ has three connected components U, V, W such that*

the boundary of U is $E_1 \cup E_2$,

the boundary of V is $E_2 \cup E_3$,

the boundary of W is $E_1 \cup E_3$.

Note that we can make X into a graph by taking the derived decomposition that we defined in Unit III.

Proof. There are three main steps. First, we prove that $S^2 - X$ has exactly three components. Next, we prove that $E_1 \cup E_3$ is the boundary of one of these components. Finally, we use the same sort of argument to obtain similar conclusions for $E_2 \cup E_3$ and $E_1 \cup E_2$. Since the simple closed curves given by $E_1 \cup E_2$, $E_2 \cup E_3$, are distinct, it follows that they bound distinct components of $S^2 - X$, and since there are exactly three components in the latter, it follows that each is the boundary of one of the given simple closed curves.

By the preceding discussion, we need only show the assertions that $S^2 - X$ has three components and one component of $S^2 - X$ has $E_1 \cup E_3$ as its boundary. We shall begin by proving the first statement, and it will be convenient to introduce some notation for certain open subsets of S^2. For $i = 1, 2, 3$ let $U_i = S^2 - E_i$ and if $i \neq j$ let $U_{i,j}$ be

$$S^2 \ - \ (E_i \cup E_j) \ \ = \ \ U_i \cap U_j \ .$$

Finally, let $U_{1,2,3}$ be $S^2 - X$ and note that the latter is equal to $U_1 \cap U_2 \cap U_3$. Consider the Mayer-Vietoris exact sequence associated to the decomposition $U_3 = U_{1,3} \cup U_{2,3}$, noting that $U_{1,2,3} =$

$U_{1,3} \cap U_{2,3}$. Since U_3 has the homology of a point by Proposition 1 of the preceding section, the final nontrival terms in the Mayer-Vietoris sequence are given as follows:

$$0 = H_1(U_3) \to H_0(U_{1,2,3}) \to H_0(U_{1,2}) \oplus H_0(U_{2,3}) \to H_0(U_3) \cong \mathbb{Z}$$

By the Jordan Curve Theorem the direct sum isomorphic to $\mathbb{Z}^4$, and the axiom regarding 0-dimensional homology implies that the standard free generators for this direct sum all map to the standard free generator of $H_0(U_3) \cong \mathbb{Z}$. Therefore the kernel of the map from the direct sum into $H_0(U_3)$ is isomorphic to a free abelian group on three generators, and by exactness this group is isomorphic to the image of the map φ from $H_0(U_{1,2,3})$ to $H_0(U_{1,2}) \oplus H_0(U_{2,3})$. Since $H_1(U_3) = 0$ it also follows that φ is 1–1, and therefore we have shown that $H_0(U_{1,2,3} = S^2 - X)$ is a free abelian group on three generators. This computation implies that the open set $U_{1,2,3}$ has exactly three components, so we have completed the first step of the proof.

The only point remaining is to prove that $E_1 \cup E_3$ is the boundary of one component in $S^2 - X$. By the Jordan Curve Theorem, the set $U_{1,3} = S^2 - (E_1 \cup E_3)$ has two components and $E_1 \cup E_2$ is the boundary of each one. Denote these components by V and W, and notice that one of them must contain the connected set $E_2 - \{A, B\}$. Without loss of generality, we may assume that this component is W (if not, reverse the roles of V and W in the discussion which follows). We then have

$$U_{1,2,3} \quad = \quad V \ \cup (W - (E_2 - \{A, B\})\,) \ .$$

Each of the summands on the right is an open and closed subset of $U_{1,2,3}$, and therefore each component of $U_{1,2,3}$ is contained in V or W. Now we know that $V \subset U_{1,2,3}$, and V must be a component of $U_{1,2,3}$ because V is a maximal connected subset of $U_{1,3}$, which contains $U_{1,2,3}$, and hence V is also a maximal connected subset of $U_{1,2,3}$. By construction the boundary of V is $E_1 \cup E_3$, and thus we have shown that the latter bounds one component of $U_{1,2,3} = S^2 - X$. Since we have already noted that this assertion (plus the one about three components) imply the conclusion of the proposition, this completes the proof.■

At a later point we shall also need information about the higher homology groups of the (components of the) space $S^2 - X$ when X is a theta space. The result is analogous to Corollary VII.3.4.

COROLLARY 3. *If $X \subset S^2$ is a theta space and U is a component of $S^2 - X$, then $H_i(U) = 0$ for all $i > 0$, and likewise for $H_i(S^2 - X)$.*

Proof. The proof of the proposition shows that the boundary of each component is a simple closed curve, and thus we can apply Corollary VII.3.4 directly to find the higher dimensional homology of U. The statement about $S^2 - X$ follows because this space is locally arcwise connected and hence its homology is the direct sum of the homology of its components.■

We are now ready to prove that the graph $K_{3,3}$ is not topologically embeddable in $\mathbb{R}^2$.

Proof of Theorem 1. We shall assume that there is a topological embedding of the graph in S^2 and derive a contradiction. It may be worthwhile to look at Figure 3 in `graphpix4.pdf` in order to visualize the steps in the argument.

Let X be a graph, and let $X_0 \subset X$ be the subgraph consisting of all edges that do not have e as a vertex. If p and q are vertices which are endpoints of some edge, denote that edge by E_{pq}. Then X_0 is a theta space with edges

$$L_1 \quad = \quad E_{ag} \cup E_{aw} \ , \qquad L_2 \quad = \quad E_{bg} \cup E_{bw} \ , \qquad L_3 \quad = \quad E_{cg} \cup E_{cw} \ .$$

Then Proposition 2 implies that $S^n - X_0$ has three components, and the remaining vertex $e \in X$ must lie in one of them, say U. It follows that each of the half-open intervals

$$E_{ae} - \{a\} \ , \quad E_{be} - \{b\} \ , \quad E_{ce} - \{c\}$$

must be contained in the component U because each is connected and contains e. Therefore each of a, b, c must lie in the closure $\overline{U}$ of U.

Trial and error suggests that the conclusion of the preceding sentence is impossible, and we shall now give mathematical reasons for this. The endpoints of L_1, L_2 and L_3 are g and w, and we also know that $a \in L_1$, $b \in L_2$ and $c \in L_3$ but none of these points can be endpoints of an edge L_i. Proposition 2 implies that the boundary of U is the union of exactly two of these edges, so only two of the points in $\{a, b, c\}$ can lie in $\overline{U}$, and thus we have derived a contradiction. The source of this contradiction was the assumption that X could be topologically embedded in S^2, and therefore we know this assumption is false. As noted earlier, this suffices to complete the proof of the theorem.■

The complete graph on 5 *vertices*

We now proceed to the next example. Recall that the complete graph on n vertices is a graph with n vertices such that for each pair of vertices $\{p, q\}$ there is an edge whose endpoints are p and q.

THEOREM 3. *The complete graph on* 5 *vertices is not homeomorphic to a subset of* S^2.

We have already noted that the complete graph on 4 vertices can be embedded in S^2, and the standard embedding is given in Figure 4 of `graphpix4.pdf`. The vertices of this graph are denoted by A, B, C, D, and the 6 edges will be labeled lexicographically (alphabetical order) as follows:

$$E_1 = AB \ , \quad E_2 = AC \ , \quad E_3 = AD \ , \quad E_4 = BC \ , \quad E_5 = BD \ , \quad E_6 = CD$$

The first step in the proof of Theorem 3 is to prove that an arbitrary topological embedding of the complete graph on 4 vertices into S^2 has the same separation properties that evidently hold in Figure 4:

THEOREM 4. *Let* $X \subset \S^2$ *be homeomorphic to the graph described above, and label its edges and vertices as in the preceding discussion. Then* $S^2 - X$ *has four components* $U, V, W, \mathcal{O}$ *such that*

the boundary of U *is* E_1, E_3 *and* E_5,

the boundary of V *is* E_2, E_3 *and* E_6,

the boundary of W *is* E_4, E_5 *and* E_6,

the boundary of $\mathcal{O}$ *is* E_1, E_2 *and* E_4.

Proof of Theorem 4. The strategy is similar to the method for proving Proposition 2:

(1) Prove that $S^2 - X$ has four components.

(2) Prove that $\Gamma = E_1 \cup E_2 \cup E_4$ is the boundary of one component.

(3) Use similar arguments to show that the other three triangular graphs in the theorem statement bound components of $S^2 - X$. As before, each of the four triangular graphs bounds a component, and since there are exactly four components it follows that each component is the boundary of one such graph,

To prove the first step, let X_0 be obtained from X by deleting the interior points of the edge E_4 (see Figure 5 in `graphpix4.pdf`), and let X_1 be the triangle graph whose edges are E_4, E_5 and E_6. Then $X_0 \cup X_1 = X$ and $X_0 \cap X_1 = E_5 \cup E_6$; note that the latter is homeomorphic to a closed interval. Consider the Mayer-Vietoris exact sequence for the decomposition

$$S^2 - (E_5 \cup E_6) \quad = \quad (S^2 - X_0) \ \cup \ (S^2 - X_1) \ , \qquad S^2 - X \quad = \quad (S^2 - X_0) \ \cap \ (S^2 - X_1) \ .$$

By construction X_0 is a theta space and X_1 is a simple closed curve, so the homology groups of $S^2 - X_0$ and $S^2 - X_1$ are known (and likewise for the homology groups of $S^2 - (E_5 \cup E_6)$ by Proposition VII.3.2). If we feed this into the long exact Mayer-Vietoris sequence, we find that the final nontrivial terms of the Mayer-Vietoris sequence are given as follows:

$$0 = H_1(S^2 - (E_5 \cup E_6)) \to H_0(S^2 - X) \to H_0(S^2 - X_0) \oplus H_0(S^2 - X_1) \to H_0(S^2 - (E_5 \cup E_6)) = \mathbb{Z}$$

The results of this section and the preceding ones imply that the direct sum is isomorphic to $\mathbb{Z}^3 \oplus \mathbb{Z}^2 \cong \mathbb{Z}^5$; furthermore, the standard free generators of this group map to the standard free generator of $H_0(S^2 - (E_5 \cup E_6)) = \mathbb{Z}$. As before, it follows that $H_0(S^2 - X)$ is a free abelian group on 4 generators, which means that $S^2 - X$ has four components, completing the proof of the first step.

In the second step we are interested in the complement of the triangular subgraph $\Gamma = E_1 \cup E_2 \cup E_4$. By the Jordan Curve Theorem $S^2 - \Gamma$ has two components, say G and H. One of these components contains the remaining vertex D; as before, without loss of generality we might as well assume that $D \in H$.

The half open intervals $E_3 - \{A\}, E_5 - \{B\}, E_6 - \{C\}$ are all connected, disjoint from Γ, and contain D, so they are all contained in H. Then we have

$$S^2 \ - \ X \quad = \quad G \ \cup \ (\, H - (E_1 \cup E_2 \cup E_3)\,)$$

where G and $(\, H - (E_1 \cup E_2 \cup E_3)\,)$ are nonempty, open and disjoint (hence both are also closed in $S^2 - X$).

Let Q_1, Q_2, Q_3, Q_4 be the components of $S^2 - X$, and number them such that $G \subset Q_1$. Since G is open and closed in $S^2 - X$, every connected subset of the latter is either contained in G or disjoint from it. In particular, since G is contained in Q_1 we know that these two sets are not disjoint and hence the connected component Q_1 must be contained in G, so that the two sets are equal. Since the boundary of G is $E_1 \cup E_2 \cup E_4$, this proves the statement needed to complete the second step of the argument. As noted at the beginning of this proof, the third step follows once we have completed the first two, and therefore we have completed the proof of the theorem.■

Proof of Theorem 3. Assume that the complete graph on 5 vertices is homeomorphic to some subset $Y \subset S^2$, and let a, b, c, d, e be its vertices. Let $X \subset Y$ be the subgraph of all edges which do not have e as an endpoint, so that X is homeomorphic to a complete graph on 4 vertices. We shall now use Theorem 4 to analyze $S^2 - X$.

Let E_{uv} be the edge joining the vertices u and v in Y. Without loss of generality, we can assume that e lies in the component of $S^2 - X$ whose boundary is $E_{ab} \cup E_{bc} \cup E_{ac}$. In any case, the vertex e lies in one component of $S^2 - X$, and we can treat the other cases by permuting the roles of a, b, c, d. Note that d does not lie in the closure $\overline{U}$ of U by the proof of Theorem 4.

Now each of the sets $(E_{xe} - \{x\})$ — where $x = a, b, c, d$ — is connected and contains e, so each of these connected sets must be contained in U. This implies that each boundary endpoint x

of $E_{xe} - \{x\}$ must be contained in $\overline{U}$. However, we have already observed that d does not lie in this subset, and therefore we have a contradiction. The problem arises from our assumption that $Y \subset S^2$ is homeomorphic to a complete graph on 5 vertices, and consequently no such subset can exist.■

Kuratowski's Theorem

The results of this section lead to the more general question of determining which connected graphs are not topologically embeddable in $\mathbb{R}^2$. Clearly a graph which contains a subgraph isomorphic to the utilities network or the complete graph on 5 vertices cannot be homeomorphic to a subset of $\mathbb{R}^2$. The end of Section 64 in Munkres mentions a remarkable converse to this result attributed to C. Kuratowski (1896–1980): *Every graph which is not homeomorphic to subset of* $\mathbb{R}^2$ *must contain a subgraph homeomorphic to* either *the utilities network* or *the complete graph on five vertices.* Here is an online reference for the proof:

`http://cs.princeton.edu/~ymakaryc/papers/kuratowski.pdf`

The file `kuratowski.pdf` contains clickable links to other proofs and further information, including independent discoveries of this result by others.

VII.5 : Rationalizations of abelian groups

(**H**, §2.2)

Frequently it is useful to begin the analysis of fundamental groups by considering their abelianizations, which are the corresponding 1-dimensional homology groups; one reason for this is that the structure theory of finitely generated abelian groups is completely understood while the theory of finitely presented — and not necessarily abelian — groups is not (in fact, their are theorems stating that certain basic problems about groups cannot be solved by systemaic recursive processes). Similarly, since the structure and morphism theory of finite-dimensional vector spaces over a field is much simpler than the structure and morphism theory of finitely generated abelian groups, and there are many situations in which it is useful to work with versions of homology theory that take values in some category of vector spaces over some field **k** and **k**-linear transformations. The purpose of this section is to prove the existence of an axiomatic homology theory valued in the category of rational vector spaces. It turns such a theory can be constructed out of a theory valued in the category of abelian groups by purely algeraic means. Accordingly, we being with a method for converting abelian groups into rational vector spaces; the construction is a straightforward generalization of the standard way to construct the rationals out of the integers using formal fractions.

Modules of quotients

Despite the similarity of names, modules of quotients are quite different from quotient modules. In a very precise sense, modules of quotients resemble the field of quotients associated to an integral domain, while quotient modules correspond to quotient rings associated to an integral domain.

The constructions described in these notes can be carried out in far greater generality than the situations we consider, but we specialize here in order to simplify the discussion.

Definition. Let G be an abelian group. The *rationalization* or G, or the *localization of* G *over the rationals* is formed by a construction very similar to the construction of the rationals from the integers. One starts with ordered pairs (g,r) where $g \in G$ and r is a nonzero integer, and one identifies (g,r) with (h,s) if there is a nonzero integer t such that $t(sg-rh) = 0$ (this is slightly stronger than the condition in the construction of $\mathbb{Q}$ from $\mathbb{Z}$, in which t is always 1). This condition defines an equivalence relation on the set of all ordered pairs, and we let $G_{(0)}$ denote the set of equivalence classes. Formally, the class of (g,r) is supposed to represent an object of the form $r^{-1} \cdot g$, and motivated by this we define addition and multiplication by a rational number as follows:

$$[g,r] \ + \ [h,s] \quad = \quad [sg+rh, rs] \ , \qquad pq^{-1}[g,r] \quad = \quad [pg, qr]$$

At this point it is necessary to verify that our definitions of sums and scalar products do not depend upon the choices of representatives for equivalence classes; this is elementary and entirely similar to the corresponding proof for the formal definition of rational numbers in terms of integers. The following result is also elementary:

THEOREM 1. *The object $G_{(0)}$ constructed above is a rational vector space, and the construction also has the following properties:*

(i) If $g_1, \cdots, g_m$ generate G, then their images under j_G span the rational vector space $G_{(0)}$.

(ii) For each abelian group G there is a group homomorphism $j_G : G \to G_{(0)}$ sending $g \in G$ to the equivalence class $[g,1]$. This map is an isomorphism if G is a rational vector space.

(iii) If $f : G \to H$ is a homomorphism then there is an associated linear transformation of rational vector spaces $f_{(0)} : G_{(0)} \to H_{(0)}$ such that the constructions sending an object or morphism Γ to $\Gamma_{(0)}$ define an ADDITIVE *covariant functor (call it $\mathcal{L}$ for the sake of definiteness) from the category of abelian groups and homomorphisms to the category of rational vector spaces and linear transformations. Furthermore, the maps j_G define a natural transformation from the identity to this functor $\mathcal{L}$.*

(iv) The construction sends the infinite cyclic group $\mathbb{Z}$ to $\mathbb{Q}$ and it sends every finite cyclic group to **0**. *Furthermore, for all abelian groups G and H we have $[G \oplus H]_{(0)} \cong G_{(0)} \oplus H_{(0)}$, and likewise for (weak)* infinite *direct sums.*

In particular, if G is a finitely generated abelian group which is the direct sum of β infinite cyclic groups and several finite cyclic groups, then $G_{(0)}$ is a rational vector space whose dimension is equal to β.

Comments on the proof. Most of the verifications are extremely straightforward and left to the reader, so we shall simply note a few key features. First of all, scalar multiplication by a rational number n/m (where $m \neq 0$) is given by

$$(n/m) \cdot [g,r] \quad = \quad [ng, mr]$$

and similarly the mapping $g_{(0)}$ is defined by the formula

$$f_{(0)}[g,r] \quad = \quad [\, f(g), \, r \,] \ .$$

We shall need the second formula for our next result.∎

The following property of the rationalization construction is somewhat less trivial, and it has far-reaching consequences.

THEOREM 2. *The functor $\Gamma \to \Gamma_{(0)}$ sends exact sequences to exact sequences.*

Proof. Every exact sequence is essentially built from short exact sequences; for example, if $A \to B \to C$ is an exact sequence involving $f : A \to B$ and $g : B \to C$, then the sequence is given by fitting together the following sequences:

$$0 \to \mathrm{Ker}(f) \to A \to \mathrm{Image}(f) = \mathrm{Kernel}(g) \to 0$$

$$0 \to \mathrm{Image}(f) = \mathrm{Kernel}(g) \to B \to \mathrm{Image}(g) \to 0$$

$$0 \to \mathrm{Image}(g) \to C \to \mathrm{Cokernel}(g) \to 0$$

Therefore it will be enough to prove the result for short exact sequences. In other words, if $0 \to A \to B \to C \to 0$ is exact, we need to prove the same holds for $0 \to A_{(0)} \to B_{(0)} \to C_{(0)} \to 0$.

We shall only prove that the sequence is exact at the middle object; the proofs at the other two objects are similar and left to the reader. Suppose that $f : A \to B$ is 1–1 and $g : B \to C$ is onto such that the image of f is the kernel of g. Then $g \circ f = 0$ and additivity imply that $g_{(0)} \circ n_{(0)} = 0$, and therefore it follows immediately that the image of $n_{(0)}$ is contained in the kernel of $g_{(0)}$. Suppose now that $[b, t]$ lies in the kernel of $g_{(0)}$. By definitions it follows that there is a nonzero integer s such that $s \cdot g(b) = 0$. By exactness of the original sequence, there is some $a \in A$ such that $f(a) = sb$, and we claim that $n_{(0)}$ maps $[a, st]$ to $[b, t]$. To see this, note that $n_{(0)}[a, st] = [sb, st]$ and the right hand side is equal to to $[b, t]$ because $stb - tsb = 0$.■

The preceding results have the following implication for chain complexes.

COROLLARY 3. *Let (C, d) be a chain complex of abelian groups. Then rationalization defines a chain complex $(\,C_{(0)}, d_{(0)}\,)$ of rational vector spaces, and the homology of this chain complex is isomorphic to the rationalized homology groups $H_*(C)_{(0)}$.*■

Application to homology theories

Since the functor $\Gamma \to \Gamma_{(0)}$ sends exact sequences to exact sequences and sends $\mathbb{Z}$ to $\mathbb{Q}$, we have the following:

Construction of rational simplicial homology. *The rationalized simplicial chain groups $C_*(\mathbf{K}, \mathbf{L})_{(0)}$ form chain complexes whose homology groups are the rationalized simplicial homology groups $H_*(\mathbf{K}, \mathbf{L})_{(0)}$. These rationalized groups have the same exactness and excision properties as ordinary homology groups, and they have the analogous properties except for the results on H_0 of an arcwise connected space except that if $\mathbf{K}$ is starshaped with respect to a vertex then $H_0(\mathbf{K})_{(0)} \cong \mathbb{Q}$.*■

Construction of rational singular homology. *Suppose that we are given the data for an axiomatic singular homology as in Unit* VI. *If we apply the functor $\Gamma \to \Gamma_{(0)}$ to these data, the result is data which satisfy all the axioms of Unit* VI *with the following modifications:*

(i) *In axiom* (D.3), *the statement should be changed as follows:* If X is arcwise connected then $H_0(X)_{(0)} \cong \mathbb{Q}$.

(ii) *Axiom* (D.4) *is excluded.*■

Frequently it is much simpler to work with the rationalized theory because (i) the rationalized homology groups are rational vector spaces, and the isomorphism type of a vector space is given by its dimension, (ii) linear transformations of rational vector spaces are completely determined by their ranks. In particular, this eliminates issues involving nonzero elements of finite order in abelian groups and proper subgroups of finite index. We shall study one application for which this simplification is particularly useful. It turns out that there are many other places in algebraic topology where it is much easier to "work over the rational numbers," and there is an extensive body of work on this topic. Here are two general references:

`http://en.wikipedia.org/wiki/Rational_homotopy_theory`

P. J. Hilton, Serre's contribution to the development of algebraic topology. Expositiones Mathematicæ **22** (2004), 375–383.

VII.6 : Cell decompositions and Euler's Formula

(**H**, Ch. 0, §2.2, Appendix)

In this final section we shall use ideas from homology theory to derive a well-known formula relating the number of vertices, edges and faces of a convex 2-dimensional polyhedron in the sense of elementary solid geometry; in our terminology, this is the boundary of a convex linear cell with a nonempty interior. By the results of `convexbodies.pdf` and `convexbodies2.pdf`, each of these objects is homeomorphic to S^2.

EULER'S POLYHEDRON FORMULA. *Suppose that $P \subset \mathbb{R}^3$ is the boundary of a convex linear cell with a nonempty interior, so that P has a decomposition into closed regions congruent to convex polygonal regions (faces), every pair of which meets in either one common edge or one common vertex. Then the numbers V, E and F of vertices, edges and faces satisfy the equation*

$$V - E + F = 2 .$$

The online site

`http://www.ics.uci.edu/~eppstein/junkyard/euler/`

contains further information about this result and its history.

Regular cell complexes

We shall need a generalization of the notion of simplicial complex called a *regular cell complex*. Such objects can be described recursively as follows:

Definition. Let k be a nonnegative integer. If (X, A) is a pair of spaces, we shall say that X *is obtained from A by regularly attaching a k-dimensional cell* if there is a 1–1 continuous mapping $f : (D^k, S^{k-1}) \to (X, A)$ such that $X = f[D^k] \cup A$ and $f[S^{k-1}] = f[D^k] \cap A$; this differs from the usual definition of cell attachment because it assumes that the restriction of f to S^{k-1} is 1–1 (see Hatcher for further information and comments).

A *regular cell decomposition* of a compact (Hausdorff) space X is a finite family $\mathcal{E} = \{\mathcal{E}_\alpha$ of closed subsets E_α (called cells) such that

(*i*) each E_α is homeomorphic to a closed disk $D^{k(\alpha)}$ for some nonnegative integer $k(\alpha)$ called the *dimension* of E_α.

(*ii*) for each cell E_α, the boundary set $\partial E_\alpha \cong S^{k(\alpha)-1}$ is a union of cells whose dimensions are less than $k(\alpha)$ called faces of E_α,

(*iii*) for each pair of distinct cells E_α and E_β, the intersection $E_\alpha \cap E_\beta$ is a common face.

The simplicial decomposition of a simplicial complex is a special type of regular cell complex, but there are many other examples, the most obvious of which are convex polygonal regions in the plane and 3-dimensional objects like a solid cube or pyramid, for which there is one 3-dimensional cell and whose boundary cells are the usual concept of face (for example, in the cube these are the six squares on the boundary). More generally, results in [MunkresEDT] and the book by Hudson yield a similar result for convex linear cells.

STANDARD DECOMPOSITION OF CONVEX LINEAR CELLS. *Let E be a convex linear cell in $\mathbb{R}^n$ with nonempty interior. Then E has a regular cell decomposition with exactly one n-dimensional cell such that each cell in the boundary is also a convex linear cell.*

This provides the geometric input that we need; the next step will involve constructions from algebraic topology.

Homology and cell attachment

The next step is to examine the significance of cell attachment in algebraic topology. Here is the main result:

THEOREM 1. *Suppose that the pair (X, A) is obtained by regularly attaching a k-cell to A, and let $D \subset X$ denote the image $f[D^k]$, and let $S \subset X$ denote the image $f[S^{k-1}]$. Then the inclusion of (D, S) in (X, A) induces isomorphisms of homology groups from $H_*(D, S)$ to $H_*(X, A)$.*

Proof. As suggested in the statement of the theorem, let $f : (D^k, S^{k-1}) \to (D, S)$ be the homeomorphism describing the cell attachment. Define subsets F_0 and G_0 of D^k by the inequalities $|v| > \frac{3}{4}$ and $|v| \geq \frac{1}{2}$ respectively (see the drawing in `cell-add.pdf`), let $F = f[F_0]$ and $G = f[G_0]$, and define new subsets $B = A \cup F$, $C = A \cup G$. Observe that C is a closed subset of X and B is an open subset (its complement is the image of the closed disk of radius $\frac{1}{2}$). Furthermore, the closure $\overline{B}$ is contained in the interior of C.

Consider the following commutative diagram:

$$\begin{array}{ccc} H_*(D-F, G-F) & \xrightarrow{p_*} & H_*(D, G) \\ \Big\downarrow g_* & & \Big\downarrow f_* \\ H_*(X-B, C-B) & \xrightarrow{q_*} & H_*(X, C) \end{array}$$

The mappings p_* and q_* are induced by inclusions, and g is a homeomorphism of pairs given by f. It follows that g_* is an isomorphism, and the excision axiom implies that p_* and q_* are also isomorphisms. Therefore the commutativity of the diagram implies that the map $H_*(X-B, C-B) \to H_*(X, C)$ is also an isomorphism.

Now S is a strong deformation retract of G, and this implies that A is a strong deformation retract of C. We CLAIM that the homology mappings

$$H_*(D, S) \to H_*(D, G) \ , \qquad H_*(X, A) \to H_*(X, C)$$

are also isomorphisms. prove.

Consider the commutative diagrams of long exact homology sequences given by axiom (B.2) and the pair inclusions $(D,S) \to (D,G)$, $(X,A) \to (X,C)$. In the first diagram the inclusion mappings induce homology isomorphisms $H_*(X) \to H_*(X)$ (which are identity maps) and $H_*(S) \to H_*(G)$ (which are isomorphisms since S is a strong deformation retract of G). Since all these maps are isomorphisms, the Five Lemma (Proposition V.3.4) implies the maps $H_*(D,S) \to H_*(D,G)$ are also isomorphisms. Similar considerations imply that the maps $H_*(X,A) \to H_*(X,C)$ are also isomorphisms.

To conclude the proof, consider the following commutative diagram:

$$\begin{array}{ccc} H_*(D,S) & \longrightarrow & H_*(D,G) \\ \downarrow & & \downarrow \\ H_*(X,A) & \longrightarrow & H_*(X,C) \end{array}$$

The immediately preceding discussion implies that the horizontal arrows are isomorphisms, and we had previously shown that the left hand vertical arrows are isomorphisms; combining these with a diagram chase, we see that the right hand vertical arrows are also isomorphisms.∎

This result has a simple but important application to (finite) regular cell complexes.

PROPOSITION 2. *If X has a finite regular cell complex structure with cells of dimension $\leq n$, then $H_q(X)$ is finitely generated for all q and $H_q(X) = 0$ for $q > n$.*

Proof. The definitions imply that X has an increasing finite sequence of subspaces

$$\emptyset \;=\; X_{-1} \;\subset\; \cdots \; X_N \;=\; X$$

such that each is obtained from the preceding one by attaching a cell, and if $p < q$ all the p-cells are attached before any q-cells are attached. The conclusion of the proposition is trivial for X_{-1}; assume by induction that the conclusion is true for X_{k-1}. We want to prove that the result is also true for X_k; this will suffice to prove the result for $X = X_N$.

The key step is the following observation: *If $A \to B \to C$ is exact and both A and C are finitely generated abelian groups, then B is also a finitely generated abelian group.* — To see this, note that B lies inside a short exact sequence

$$0 \to A' \to B \to C' \to 0$$

where A' is a quotient of A and C' is a subgroup of C. Since subgroups and quotient groups of finitely generated abelian groups are also finitely generated, it follows that A' and C' are finitely generated, and this implies that B must also be finitely generated.

We now apply this to the exact sequences $H_q(X_{k-1}) \to H_q(X_k) \to H_q(X_k, X_{k-1})$. The first group is finitely generated by the induction hypothesis, and the second is finitely generated by Theorem 1, so the preceding paragraph implies that $H_q(X_k)$ is also finitely generated. Furthermore, if $q > n$ then $H_q(X_{k-1}) = 0$ by induction and $H_q(X_k, X_{k-1}) = 0$ by Theorem 1, and these plus exactness imply that $H_q(X_k) = 0$.∎

Euler characteristics

At this point we generalize the Euler characteristic of a graph to finite regular cell complexes in two ways; these invariants are integer valued, and we shall show that these are equal.

Definition. Let X be a topological space such that $H_q(X)$ is a finitely generated abelian group for all $q > 0$. The q^{th} *Betti number* $\beta_q(X)$ is equal to the rank of $H_q(X)$; by definition the rank of a finitely generated free abelian group A is equal to the number of infinite cyclic generators in a standard decomposition (this number depends only upon the group)and is equal to the dimension of the rational vector space $A_{(0)}$.

Definition. Suppose that X has a finite regular cell complex structure, so that Proposition 2 applies. The *homological Euler characteristic* $\chi^H(X)$ is defined to be the alternating sum

$$\sum_q (-1)^q \, \beta_q(X) \ .$$

This sum is actually finite because $H_q(X)$ is trivial for all sufficiently large values of q (and is zero if $q < 0$).

Definition. Suppose that X has a finite regular cell complex structure $\mathcal{E}$, and for each nonnegative integer q let $c_q(X, \mathcal{E})$ denote the number of q-cells in $\mathcal{E}$. The *cellular Euler characteristic* $\chi^C(X, \mathcal{E})$ is defined to be the alternating sum

$$\sum_q (-1)^q \, c_q(X, \mathcal{E}) \ .$$

Euler's Polyhedron Formula — and its analogs in higher dimensions — will follow quickly from the next result.

THEOREM 3. *Suppose that X has a finite regular cell complex structure $\mathcal{E}$. Then the two Euler characteristics $\chi^H(X)$ and $\chi^C(X, \mathcal{E})$ are equal.*

Proof. We shall use the same increasing chain of subspaces

$$\emptyset \ = \ X_{-1} \ \subset \ \cdots \ X_N \ = \ X$$

described in the proof of Proposition 2.

Once again, the result is is trivial for X_{-1}, once again we assume by induction that the conclusion is true for X_{k-1}, and once again it will suffice to prove the conclusion of the theorem for X_k.

The crucial steps are to study what happens to both Euler characteristics when one adds a single cell of dimension r to form X_k from X_{k-1}. It is easy to see what happens to the cellular Euler characteristic; since we are adding a single cell in dimension r we have

$$\chi^C(X_{k-1}, \mathcal{E}_{k-1}) \ + \ (-1)^r \ \ = \ \ \chi^C(X_k, \mathcal{E}_k) \ + \ (-1)^r$$

where $\mathcal{E}_j$ denotes the induced cell structure on X_j.

The analysis of the homological Euler characteristic requires a closer examination of the exact homology sequence of the pair (X_k, X_{k-1}) and its rationalization. Since there is only one nonzero homology group of the latter pair and it is in dimension r, it follows that the inclusion map $H_q(X_{k-1}) \to H_q(X_k)$ is an isomorphism if $q \neq r, r-1$, so that $\beta_q(X_{k-1}) = \beta_q(X_k)$ for $q \neq r, r-1$. To compare the remaining two Betti numbers, we need to look at the nontrivial part of the rationalized exact homology sequence:

$$0 \to H_r(X_{k-1})_{(0)} \to H_r(X_k)_{(0)} \to H_r(X_k, X_{k-1})_{(0)} \cong \mathbb{Q} \to H_{r-1}(X_{k-1})_{(0)} \to H_{r-1}(X_k)_{(0)} \to 0$$

There are two cases depending upon whether or not the map ∂ from $\mathbb{Q} = H_r(X_k, X_{k-1})_{(0)}$ to $H_{r-1}(X_{k-1})_{(0)}$ is trivial or nontrivial. If ∂ is trivial then by exactness we have

$$\beta_{r-1}(X_k) \quad = \quad \beta_{r-1}(X_{k-1}) \ , \qquad \beta_r(X_k) \quad = \quad \beta_r(X_{k-1}) \ + \ 1$$

and thus we also have

$$\chi^H(X_k) \ - \ \chi^H(X_{k-1}) \quad = \quad (-1)^r \quad = \quad \chi^C(X_k, \mathcal{E}_k) \ - \ \chi^C(X_{k-1}, \mathcal{E}_{k-1})$$

which combines with $\chi^H(X_{k-1}) = \chi^C(X_{k-1}, \mathcal{E}_{k-1})$ to imply that $\chi^H(X_k) = \chi^C(X_k, \mathcal{E}_k)$. On the other hand, if ∂ is nontrivial then by exactness we have

$$\beta_{r-1}(X_k) \ + \ 1 \quad = \quad \beta_{r-1}(X_{k-1}) \ , \qquad \beta_r(X_k) \quad = \quad \beta_r(X_{k-1})$$

and thus we also have

$$\chi^H(X_{k-1}) \ - \ \chi^H(X_k) \quad = \quad (-1)^{r-1} \quad = \quad \chi^C(X_{k-1}, \mathcal{E}_{k-1}) \ - \ \chi^C(X_k, \mathcal{E}_k)$$

which combines with $\chi_H(X_{k-1}) = \chi^C(X_{k-1}, \mathcal{E}_{k-1})$ to imply that $\chi^H(X_k) = \chi^C(X_k, \mathcal{E}_k)$.∎

Euler's Polyhedral Formula now follows as the specialization of the final result to $n = 2$.

COROLLARY 4. *Let P be the boundary of a convex linear cell in $\mathbb{R}^{n+1}$ with nonempty interior, and for each integer k between 0 and n let V_k denote the number of k-dimensional faces in the standard regular cell decomposition. Then the alternating sum*

$$\sum_k \ (-1)^k \, V_k$$

is 0 if k is odd and 2 if k is even.

Proof. The homological Euler characteristic of P is $1 + (-1)^n$ which is 0 if k is odd and 2 if k is even. Therefore the result follows from Theorem 3.∎