P. J. GIBLIN

Graphs, Surfaces and Homology

CHAPMAN AND HALL
SECOND EDITION

Graphs, Surfaces and Homology

An Introduction to Algebraic Topology

CHAPMAN AND HALL
MATHEMATICS SERIES

Edited by Professor R. Brown
Head of the Department of Pure Mathematics,
University College of North Wales, Bangor

Graphs, Surfaces and Homology

An Introduction to Algebraic Topology

P. J. GIBLIN

Department of Pure Mathematics
University of Liverpool

SECOND EDITION

Preface to the
second edition

I am glad to take the opportunity of a second edition to add a
number of references and change some discussions in the light
of work done or published since the first edition. The most
notable example is of course the four-colour conjecture, which
became a theorem (proved partly by massive computer calculations)
shortly after my book was published. I have also been told of
facts and articles of which I was ignorant, and these have been
included in the appropriate places. The new references have
all been included in the list on pages 309-12 but have not
generally been incorporated into the index. A completely wrong
proof and a number of smaller errors have been corrected. For
their helpfulness in pointing out some of these, or for
providing me with references, I am especially grateful to
Mr S. Wylie, Dr E.Kronheimer and Mr S.L. Wilson.

<div align="right">P.J.G. 18 June 1981</div>

Preface to the first edition

Topology is pre-eminently the branch of mathematics in which other mathematical disciplines find fruitful application. In this book the algebraic theory of abelian groups is applied to the geometrical and topological study of objects in euclidean space, by means of homology theory. Several books on algebraic topology contain alternative accounts of homology theory; mine differs from these in several respects.

Firstly the book is intended as an undergraduate text and the only mathematical knowledge which is explicitly assumed is elementary linear algebra. In particular I do not assume, in the main logical stream of the book, any knowledge of point-set (or "general") topology. (There are a number of tributaries, not part of the main stream but I hope no less logical, about which more in a moment.) A reader who is familiar with the concept of a continuous map will undoubtedly be in a better position to appreciate the significance of homology theory than one who is not but nevertheless the latter will not be at a disadvantage when it comes to understanding the proofs.

The avoidance of point-set topology naturally imposes certain limitations (in my view quite appropriate to a first course in homology theory) on the material which I can present. I cannot, for example, establish the topological invariance of

homology groups.   A weaker result, sufficient nevertheless for
our purposes, is proved in Chapter 5, where the reader will
also find some discussion of the need for a more powerful in-
variance theorem and a summary of the proof of such a theorem.

Secondly the emphasis in this book is on low-dimensional
examples - the graphs and surfaces of the title - since it
is there that geometrical intuition has its roots.   The goal
of the book is the investigation in Chapter 9 of the properties
of graphs in surfaces;   some of the problems studied there are
mentioned briefly in the Introduction, which contains an in-
formal survey of the material of the book.   Many of the results
of Chapter 9 do indeed generalize to higher dimensions (and the
general machinery of simplicial homology theory is available
from earlier chapters) but I have confined myself to one
example,   namely the theorem that non-orientable closed surfaces
do not embed in three-dimensional space.   One of the principal
results of Chapter 9, a version of Lefschetz duality, certainly
generalizes, but for an effective presentation such a general-
ization needs cohomology theory.   Apart from a brief mention
in connexion with Kirchhoff's laws for an electrical network I
do not use any cohomology here.

Thirdly there are a number of digressions, whose purpose
is rather to illuminate the central argument from a slight dis-
tance, than to contribute materially to its exposition.   (To
change the metaphor, these are the tributaries mentioned
earlier.   Some of them would turn into deep lakes if we were
to pursue them far.)   The longest digression concerns planar
graphs, and is to be found in Chapter 1, but there are others
on such topics as Kirchhoff's laws (also Chapter 1), minimal
triangulations, embedding and colouring problems (Chapter 2),
orientation of vector spaces (Chapter 3) and a game due to
J. H. Conway (Chapter 9).   In addition, there are a good many
examples which extend the material of the text in a more or
less informal way.

Material which is not directly relevant to the logical
development is indicated by an asterisk *, or occasionally by
some cautionary words at the outset.   Thus the digressions in
question can be regarded as optional reading, but even the
reader intent on getting to the results may find the more
relaxed atmosphere of some of the starred items congenial.
Occasional use is made there of concepts and theorems which
are related to the material of the text, but whose full expo-
sition would take us too far afield.   As an example, the proof
of Fáry's theorem on straightening planar graphs (Theorem 1.34)
assumes several plausible (but profound) results of point-set
topology.   The inclusion of this theorem makes possible several
useful discussions later in the book; I have been careful to
warn the reader that these discussions, resting as they do on a
result whose complete proof involves concepts not explained
here, must remain to some extent "informal".   Investigations
of this kind, and others in which an admission of informality
is made, are in fact reminiscent of the style of argument
usually adopted in more advanced mathematical work.   The tacit
assumption is present, that the author (and hopefully the
reader) could, if he were so minded, fill in all the details
and make the arguments formal, but that to do so in print would
be,   in this conservation-conscious age, a misuse of paper.

The fourth way in which my book differs from others of its
kind is more technical in nature.   I use collapsing as a
major tool, rather than the Mayer-Vietoris sequence, which
appears in a minor role in Chapter 7.   My reason for using
collapsing is that it leads straight to the calculation of the
homology groups of surfaces with only a minimal need for invar-
iance properties.

---

* thus asterisks do not direct the reader to footnotes, the
symbol for which is '.

The theory of abelian groups enters into homology theory
in two ways.   The very definition of homology groups is framed
in terms of quotient groups, and many theorems are expressed by
the exactness of certain sequences.   The abelian group theory
used in this book is gathered together in an Appendix, in which
proofs will be found of all the results used in the text.

About three-quarters of the material in this book was in-
cluded in a course of around forty lectures I gave to third
year undergraduates at Liverpool University.   Included in the
forty lectures were a few given by the students themselves.
Some of these were on topics which now appear as optional sec-
tions, while others took their subject matter from research
articles.   The material now included in the text and which I
omitted from the course comprises some of the work on planar
graphs, all the work on invariance, the homology sequence of a
triple, a few of the applications in Chapter 9, and several
small optional sections.   There are many other ways of cutting
down the material.   For example it is possible to restrict
attention entirely to 2-dimensional simplicial complexes: to
facilitate this suggestion I have tried to ensure that by judi-
cious skipping in Chapter 3 all mention of complexes of dimen-
sion higher than two can be suppressed.   Thereafter the results
can simply be restricted to this special case, and in Chapter 9
most of the work is 2-dimensional anyway.   Alternatively,
Chapters 5 and 7 can be omitted, together with optional material
according to the taste (or distaste) of the lecturer.   I think
it would be completely mistaken to give a course in which the
students sharpened their teeth on all the techniques, only to
leave without a sniff at the applications: better to omit
details from Chapters 1 - 8 in order to include at any rate a
selection of material from Chapter 9.

I have kept two principles before me, which I think should
apply to all university courses in mathematics, and especially
ones at a fairly high level.   Firstly the course should have

something to offer for those students who are going to take
the subject further - it should be a stepping-stone to "higher
things".   But secondly, and I think this is sometimes forgotten,
the course should exhibit a piece of mathematics which is
reasonably complete in itself, and prove results which are
interesting for their own sake.   I venture to hope that the
study of graphs in surfaces is interesting for its own sake,
and that such a study provides a good introduction to the power-
ful and versatile techniques of algebraic topology.   Bearing
in mind that a proportion of students will become school-
teachers, I have been careful to include a number of topics,
such as planar graphs in Chapter 1, colouring theorems in
Chapter 2 and "Brussels sprouts" in Chapter 9, which, suitably
diluted and imaginatively presented, could stimulate the inte-
rest of schoolchildren.

    I have found it impossible in practice to separate worked
examples, which often have small gaps in the working, from
exercises, which often have hints for solution.   All these are
gathered together under the heading of "Examples", but to help
the reader in selecting exercises I have marked by X those
examples where a substantial amount is left to the reader to
verify.   Two   X's indicate an example that I found difficult.
Some examples are starred: this indicates that their subject
matter is outside the main concern of the book.

    References to other books and to articles are given by the
name of the author followed by the date of publication in
brackets.   A list of references appears on pp. 309-12. Some
of the references within the text are to research articles and
more advanced textbooks where the reader can find full details
of subjects which cannot be explored here.   Inevitably these
references are often technical and not for the beginner unless
he is very determined.   I have also included among the refer-
ences several books which, though they are not specifically
referred to in the text, can be regarded as parallel reading.

There is a list of notation on pp. 313-7. My only con-
scious departure from generally received notation is the use,
in these days a little eccentric, of the declaration Q.E.D. to
announce the end of a proof.

P. J. Giblin

May 1976

ACKNOWLEDGEMENTS

To the students who attended my lectures,
     not only for attending,
          but also for improving many arguments;

In particular to Pierre Damphousse,
     whose well-made proof of Fáry's theorem is in
     Chapter 1;

To Dr C.G. Gibson and Dr H.R.Morton,
     who read the book in manuscript,
          and by their gentle but persuasive criticism
          improved it at many points;

To Mr J.E. Reeve,
     whose lecture notes developed into parts of Chapter 2;

To Mr J.W. Bruce,
     for his patient and invaluable help with proofreading;

To Miss Ann Garfield,
     for transforming so unprepossessing a manuscript into
     so agreeable a typescript;

To Professor R. Brown,
     both for the suggestion that I should write up my
     lecture notes for publication,
          and for his friendly encouragement and sound
          advice while the writing up was in progress;

        I offer my warmest thanks

        And I dedicate this book
             to my wife Rachel,
                  for everything else

Contents

Introduction

Algebraic methods have often been applied to the study of geometry. The algebraic reformulation of euclidean geometry as "coordinate geometry" by Descartes was peculiarly success-ful in that algebra and geometry exactly mirrored each other. The objects studied in this book - graphs, surfaces and higher dimensional "complexes" - are much more complicated than the lines and circles of euclidean geometry and it is not to be expected that any reasonable algebraic machinery can hope to describe all their features. We must expect many distinct geometrical situations to be described by the same piece of algebra. That is not necessarily a disadvantage for not all features are equally interesting and it can be useful to have a systematic way of ignoring some of the less interesting ones.

Throughout the book we shall be concerned with subsets X of a euclidean space (such as ordinary space $\mathbb{R}^3$) and shall attempt to clarify the structure of the sets X by looking at what are called <u>cycles</u> on them. For the present let us consider one-dimensional cycles, or 1-cycles, which have their origin in the idea of simple closed curve (that is, a closed curve without self-intersections). Thus consider the three graphs drawn on the next page. It is fairly evident that no simple closed curves at all can be drawn on the left-hand graph.

1

On the middle graph a simple closed curve can be drawn going
once round the triangle, while on the right-hand graph there
are three triangles, two right-angled and one equilateral,
which can be circumnavigated.   The first hint of algebra comes
when we try to combine two cycles (in this case two simple
closed curves) together.   We want to say something like

The way in which we make the vertical edge cancel out is by
giving orientations (directions) to the
edges of the triangles and taking all
cycles to be, say, clockwise.   Then we
have
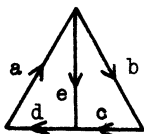
$$a + b + c + d = (a + e + d) + (b + c - e)$$

where the -e occurs because e is going the wrong way for the
right-hand cycle.   The sum now looks like a calculation in an
abelian group, and indeed in Chapter 1 we make the set of
1-cycles into an abelian group which measures the "number of
holes" in the graph, or the "number of independent loops" on
the graph.   This number is respectively 0, 1 and 2 for the
three graphs above, and the corresponding groups are 0, $\mathbb{Z}$,
$\mathbb{Z} \oplus \mathbb{Z}$.   The amount of group theory needed to read Chapter 1
is minimal, barely more than the definitions of abelian group
and homomorphism.

When we come to consider 1-cycles on higher-dimensional
objects the situation becomes more interesting.   The diagrams
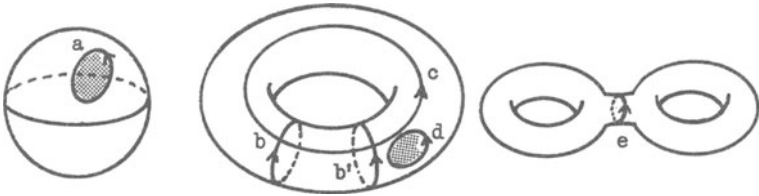
represent a spherical surface, a torus (or "inner tube") and
a double torus or pretzel, and a, b, b', c, d, e are 1-cycles.
Now a and d are each the boundary of a region - such a
region is shaded, but it would also be possible to take the
"complementary" region occupying the remainder of the surface.
On the other hand b, b' and c do not bound a region.   The
fact that the simple closed curve b fails to bound a region
reveals the presence of a hole in the torus (where the air
goes in an inner tube): "b goes round the hole".   Similarly
c reveals the presence of another hole (where the axle of the
wheel goes).   Since b and c are 1-dimensional we could
call the holes 1-dimensional too.

In order to use cycles for gathering geometric information
we shall declare those which, like a and d, bound a region,
to be "zero" - homologous to zero is the official phrase,
denoted ~ 0.   Now b and b' between them bound a region;
once the idea of orientation is made precise this implies that
$b - b' \sim 0$ (rather than $b + b' \sim 0$).   Rearranging, we have
$b \sim b'$, which gives formal expression to the fact that b
and b' go round the same hole in the torus.   There is a
technique available for making the bounding cycles zero: they
form a subgroup (the boundary subgroup) of the group of cycles
and what is required is the quotient group cycles/boundaries.
This quotient group, when the cycles are 1-cycles, is called
the first homology group.   The size of this group measures,
roughly speaking, the number of 1-dimensional holes the
object has.   The sphere has none and the torus has two; their
first homology groups are respectively 0 and $\mathbb{Z} \oplus \mathbb{Z}$.

Notice in passing that a and d can each be shrunk to
a point on its respective surface whereas e, on the double
torus, cannot.   Nevertheless e does bound a region (namely
one half of the surface) so e ~ 0.

The whole of, for example, the spherical surface can be
considered as a two-dimensional cycle, or 2-cycle, since it is
"closed"; certainly this cycle is not the boundary of anything
since there is nothing three-dimensional in the surface for it
to bound.   If on the other hand we replace the sphere by the
solid ball obtained by filling in the inside of the sphere,
then the 2-cycle given by the sphere becomes a boundary: in the
solid ball this 2-cycle is a boundary.   As before we can form
the homology group, this time the second homology group; for
the sphere this is $\mathbb{Z}$ and for the solid ball it is 0.

Notice that if we consider a solid torus (like a solid
tyre) then b and b' become homologous to zero, while c
does not.   In fact the first homology group changes from
$\mathbb{Z} \oplus \mathbb{Z}$ to $\mathbb{Z}$.

It is important to realize that the homology groups of an
object depend only on the object and not on the way in which
it is placed in space.   Thus if we cut the torus along b,
tie it in a knot and then re-stick along b, the resulting
"knotted torus" has the same homology groups as the original
one.

Stepping down one dimension we could consider the same
graph being placed in different ways in a surface, such as the
torus in the diagrams below:
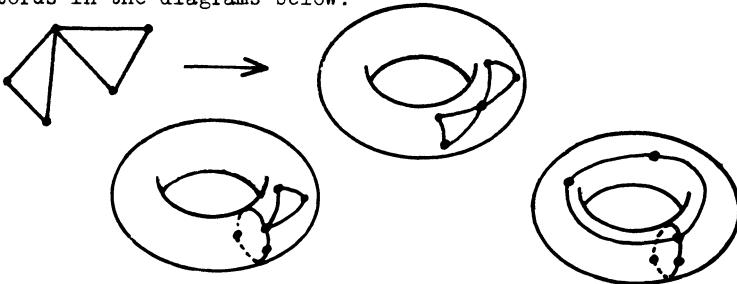
These three situations are quite different - for instance the
number of regions into which the torus is divided is different
in each case. We can study "embedding" or "placement"
problems by means of <u>relative homology groups</u>, which take into
account both graph and torus at the same time. The ordinary
groups and the relative groups are introduced in Chapter 4.

Here are some of the problems which are studied in Chap-
ter 9 by means of relative homology groups.

1. Given a graph in a surface, what can be said about
the number of regions into which the graph divides the surface
and about the nature of those regions? Much of Chapter 9 is
concerned with this problem.

2. What is the "largest" graph which can be embedded in
a surface without separating it into two regions? Of course
it is necessary to choose a good definition of largeness first;
here is an alternative formulation of the question. How many
simple closed cuts can be made on a surface before it falls
into two pieces? For example on a double torus the answer is
four - see the diagram on the left. For a "Klein bottle",
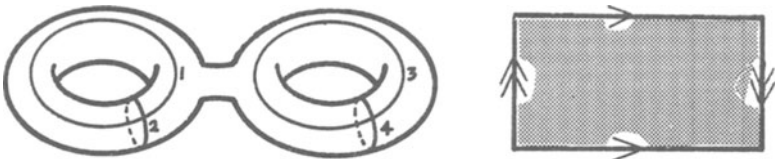
obtained from a plane rectangle by identifying the sides in
pairs as indicated on the diagram on the right, the answer is
two; can the reader find two suitable curves? The general
answer is in 9.20. The Klein bottle cannot be constructed in
three-dimensional space, and a proof that this is so is
sketched in 9.18. It can be constructed in $\mathbb{R}^4$ (p. 73).

3. What can be said about a cycle given by a simple
closed curve on, say, the torus? The simple closed curve b
in the diagram on page 3 winds round one way and the simple
closed curve c winds round the other way. Can a simple
closed curve wind round, say, twice the b-way and not go round

the c-way at all?   Certainly the most obvious candidates are
not simple closed curves:
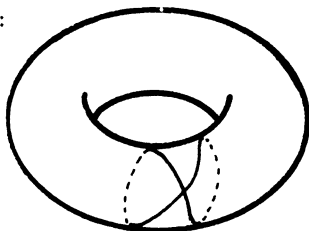
The general answer is in 9.22.

Curiously, although cycles were studied from the beginnings
of algebraic topology in the writings of H. Poincaré at the end
of the nineteenth century, no explicit mention was made of homo-
logy groups until much later, around the early 1930s.   What
were used instead were numerical invariants of the groups (the
so-called rank and torsion coefficients; see A.26), which
accounts for the lack of formal recognition given to many of
the basic results.   For example the exact homology sequence of
a pair (Chapter 6) did not appear explicitly until 1941.   The
reformulation of homology theory in terms of abelian groups not
only produced a substantial increase in clarity but made avail-
able powerful tools with which the scope of the theory could be
broadened.[†]

This is an example - one of many - where modern algebraic
formalism has come to the rescue of an attractive but elusive
geometrical idea.   Of course any rescue operation has its
dangers; in this case that formalism should cease to be sub-
servient to geometry and become an end in itself.   I have
tried very hard to avoid such a course and to keep constantly
in mind that the goal of the book is the geometrical applica-
tions of Chapter 9.

The techniques developed in this book do not enable us to
study "curved surfaces" such as the torus or sphere directly.

[†]   For more historical information see Lefschetz (1949), Pont
(1974), Lakatos (1977), Kline (1972), Eilenberg & Steenrod (1952),
Biggs, Lloyd and Wilson (1976).

(The techniques of differential topology are more suited to such
a direct study; for a beautiful introduction to that subject see
Milnor (1965).)   Instead we "triangulate" these surfaces, that
is, divide them up into (for the moment curved) triangular
regions in such a way that two regions which intersect do so
either at a single vertex or along a single edge.   Triangu-
lating a surface reduces the calculation of homology groups to
a finite procedure, which is explained in Chapter 4; it has the
additional advantage that it is possible to give a complete
prescription for constructing triangulations of surfaces.
This is given in Chapter 2; as an example consider the diagrams
below, showing a tetrahedron and an octahedron each inside a
sphere.   The hollow tetrahedron can be projected outwards
$(P \to P')$ on to the sphere which surrounds it, thus giving a
triangulation of the sphere.   Likewise the hollow octahedron
can be projected outwards $(P \to P')$ to give another triangu-

lation.   Indeed the hollow tetrahedron and octahedron can
themselves be regarded as triangulated surfaces in which it so
happens that the triangles are flat and the edges straight.
These rectilinear triangulations are referred to as simplicial
complexes and are studied in considerable generality in Chap-
ter 3.   Their importance lies in the fact that any triangu-
lated surface is homeomorphic to one in which the triangles
are flat and the edges straight.   (Two subsets X and Y of
euclidean space are called homeomorphic if there exists a con-
tinuous, bijective map f : X $\to$ Y with continuous inverse
f^{-1}.   Such a map f is called a homeomorphism.)   The maps
P $\to$ P' above are both homeomorphisms.   The kind of geometrical
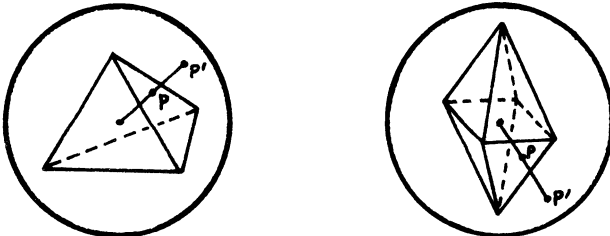
information which we are trying to capture - that is, infor-
mation of a topological nature - is so basic that it is un-
altered by homeomorphisms.   Thus it is no restriction to
concentrate entirely on rectilinear triangulations.

That is what we do in this book: all triangulations have
flat triangles with straight edges.   It should be noted,
however, that often when drawing surfaces we shall draw them
curved.   This is for two reasons: the pictures are more
recognizable and easier to draw; and they help to remind us
that the triangulations are a tool to help us gather infor-
mation, not an intrinsic part of the geometrical data.

The remarkable thing is that homology groups, in spite
of being defined via triangulations, really do measure some-
thing intrinsic and geometrical.   To be more precise: if X
and Y are homeomorphic, then the homology groups of X are
isomorphic to those of Y.   This is true not only if X and
Y are surfaces, but if they are any simplicial complexes
which may include, besides triangles, solid tetrahedra and
higher dimensional "simplexes".   This result is called the
topological invariance of homology groups, and it has the
following equivalent formulation: if, for some p, the pth
homology group of X is not isomorphic to the pth homology
group of Y, then X and Y are not homeomorphic.   Thus
homology groups can be used to distinguish objects from each
other.   It is not true in general that if the homology groups
of X and Y are isomorphic then X and Y are homeomor-
phic: many different objects can have the same groups.   This
brings us back to the point made at the beginning of this
Introduction, that many distinct geometrical situations are
described by the same piece of algebra, and shows that homo-
logy groups are not the ultimate tool in distinguishing
objects from one another.

It happens that the applications in this book do not
require the full invariance theorem quoted above, and

accordingly this theorem is not proved here.   Enough, in fact
more than enough, is proved in Chapter 5 for our purposes.
A sketch of the argument leading to the theorem of topological
invariance is also included, together with references for the
full proof.

The finite procedure for calculating homology groups can
in practice be very lengthy, and for this reason general
theorems are a help.   A number of such theorems are given
in Chapter 6: these enable the homology groups of all (closed)
surfaces to be calculated without difficulty.   Two more
theorems, slightly more technical in nature and not strictly
necessary for the applications, are given in Chapter 7.

Some of the work of Chapter 9 needs an "unoriented" hom-
ology theory which differs in detail from the oriented homology
theory studied in Chapters 1 - 7.   This is presented in
Chapter 8.

An appendix contains all the group theory needed to read
the book, presented in a fairly condensed form.   Virtually
no group theory is needed for Chapters 1 - 3.

Finally it should be emphasized that there are other
approaches to homology theory.   It is possible to define
homology groups for far more general objects than simplicial
complexes, in fact for arbitrary "topological spaces" which
may not be contained in a euclidean space at all.   There are
several methods for doing this, described as "singular" or
"Čech" or "Alexander-Spanier" homology.   They all have the
property of topological invariance, so that they can sometimes
be used to distinguish two spaces from one another - indeed
this is one of the principal applications of homology theory.
Again the method is not infallible: for example the three
objects pictured on the next page (cylinder, Möbius band,
circle) all have the same groups in any homology theory, but
no two are homeomorphic.   Accounts of other theories can be
found in, for example, Spanier (1966) and in Eilenberg and

Steenrod (1952).   The latter book also presents a most attrac-
tive account of homology theory from the axiomatic point of
view.
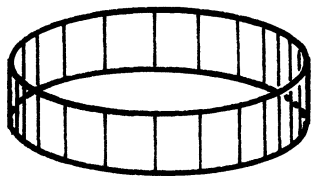
cylinder
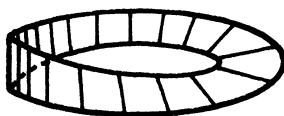
Möbius band
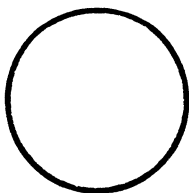
circle

Graphs

The theory of graphs, otherwise known as networks, is a branch
of mathematics which finds much application both within mathe-
matics and in science.   The reader is referred to books of
Harary (1967), Harris (1970) and Busacker and Saaty (1965)
for evidence to support this claim, since our motivation for
touching on the subject here is different.   Many of the ideas
which we shall encounter later can be met, in a diluted form,
in the simpler situation of graph theory, and that is the
reason why we study graphs first.   The problems studied by
graph theorists are usually specifically applicable to graphs,
and do not have sensible analogues in the more general theory
of "simplicial complexes" which we shall study.   It is hardly
surprising, therefore, that the theory we shall develop has
little to say of interest to graph theorists, and it is partly
for this reason that an occasional excursion is made, in this
chapter and elsewhere, into genuine graph-theoretic territory.
This may give the reader some idea of the difficult and inter-
esting problems which lie there, but which the main concern
of this book, namely homology theory, can scarcely touch.

ABSTRACT GRAPHS AND REALIZATIONS

A graph is intuitively a finite set of points in space, called
the vertices of the graph, some pairs of vertices being joined
by arcs, called the edges of the graph.   Two arcs are assumed
to meet, if at all, in a vertex, and it is also assumed that
no edge joins a vertex to itself and that two vertices are
never connected by more than one edge.   That is, we do not
allow

The positions of the vertices and the lengths of the
edges do not concern us; what is important is the number of
vertices and the pairs of vertices which are connected by an
edge.   Thus the only information we need about an edge is
the names of the vertices which it connects.   This much in-
formation can be summed up in the following "abstract" defi-
nition; we shall return to "reality" shortly.

1.1 Definition   An <u>abstract graph</u> is a pair (V, E) where
V is a finite set and E is a set of unordered pairs of
distinct elements of   V.   Thus an element of E is of the
form $\{v, w\}$ where v and w belong to V and $v \neq w$.
The elements of V are called <u>vertices</u> and the element $\{v, w\}$
of E is called the <u>edge</u> joining v and w   (or w and v).

Note that E may be empty: it is possible for no pair
of vertices to be joined by an edge.   Indeed V may be empty,
in which case E certainly is - this is the empty graph with
no vertices and no edges.   In any case if V has n elements
then E has at most $\binom{n}{2} = \frac{1}{2}n(n - 1)$ elements.   An abstract
graph with the maximum number of edges is called <u>complete</u>:
every pair of vertices is joined by an edge.

As an example of an abstract graph let $V = \{u, v, w\}$
and $E = \{\{u, v\}, \{u, w\}\}$. This
abstract graph (V, E) can be pictured,
or "realized" as we shall say in a
moment, by the diagram.
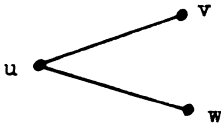
In many circumstances it is convenient to speak of graphs
in which each edge has a direction or orientation (intuitively
an arrow) attached to it. For example in an electrical
circuit carrying direct current each edge (wire) has a direc-
tion, namely the direction in which current flows.

1.2 Definition   An abstract oriented graph is a pair (V, E)
where V is a finite set and E is a set of ordered pairs of
distinct elements of E with the property that, if $(v, w) \in E$
then $(w, v) \notin E$. The elements of V are called vertices
and the element (v, w) of E is called the edge from v to **w.**

Associated with any abstract oriented graph there is an
abstract graph, obtained by changing ordered to unordered
pairs.   The two graphs have the same number of vertices, and
the same number of edges on accou... of the assumption that
(v, w) and (w, v) never both belong to E.

We now return to graphs "realized" in space.

1.3 Definition   Let (V, E) be an abstract graph. A real-
ization of (V, E) is a set of points in a real vector space
$\mathbb{R}^N$, one point for each vertex, together with straight segments
joining precisely those pairs of points which correspond to
edges.   The points are called vertices and the segments edges;
the realization is called a graph.   We require two "inter-
section conditions" to hold: (i) two edges meet, if at all,
in a common end-point, and (ii) no vertex lies on an edge
except at one of its end-points.   The edge joining v and w
will be denoted (vw), or (wv).

Note that the edges of a graph are straight; that this is
no restriction for graphs in $\mathbb{R}^3$ follows from Theorem 1.4

below.   We shall turn to the question of whether it is a res-
triction for graphs in $\mathbb{R}^2$ (planar graphs) in 1.34.

A realization of an abstract oriented graph, called an
oriented graph, is just a realization of the associated abs-
tract graph, together with the ordering on each edge of the
graph.   In practice this ordering is indicated by an arrow
from the first to the second vertex of the edge:

v •————————→————• w

is the edge from v to w,
which will be denoted (vw).
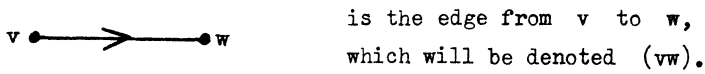
The complete abstract graphs with up to four vertices
have realizations in $\mathbb{R}^2$:
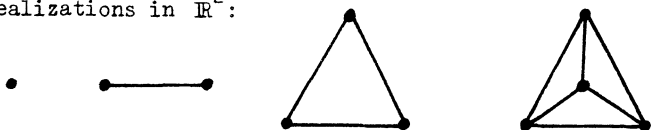
The reader can probably convince himself that the complete
abstract graph with 5 vertices cannot be realized in $\mathbb{R}^2$ (see
Proposition 1.32); however it can be realised in $\mathbb{R}^3$, as indeed
can every abstract graph.

1.4 Theorem   Every abstract graph can be realized in $\mathbb{R}^3$.

Proof   Let C be the subset of $\mathbb{R}^3$ defined by

$$C = \{(x, x^2, x^3) : x \in \mathbb{R}\}.$$

(C is called a "twisted cubic".)   I claim that no four
points of C are coplanar (it follows from this that no three
points of C are collinear).   To prove the claim take four
points P_1, P_2, P_3, P_4 of C, given respectively by distinct
real numbers x_1, x_2, x_3, x_4.   They are coplanar if and only if
the vectors $P_1 - P_2$, $P_1 - P_3$, $P_1 - P_4$ are linearly dependent,

i.e.
$$\begin{vmatrix} x_1 - x_2 & x_1^2 - x_2^2 & x_1^3 - x_2^3 \\ x_1 - x_3 & x_1^2 - x_3^2 & x_1^3 - x_3^3 \\ x_1 - x_4 & x_1^2 - x_4^2 & x_1^3 - x_4^3 \end{vmatrix} = 0.$$

However this determinant vanishes if and only if the x's are
not distinct, and this proves the claim.

To realize an abstract graph (V, E) in $\mathbb{R}^3$ take n
distinct points $P_1, \ldots, P_n$ of C, one for each vertex,
and join precisely those pairs of points which correspond to
elements of E by straight segments. It remains to verify
that the intersection conditions hold. Let $P_i P_j$ and $P_k P_\ell$
be two (distinct) segments; certainly $i \neq j$ and $k \neq \ell$ and
we can assume $j \neq \ell$. If $i = k$, so that the segments have
a common end-point, it is still true that P_i, P_j and P_ℓ
are distinct and so not collinear; hence the segments do not
have any more points in common. If on the other hand $i \neq k$
then all four P's are distinct and so not coplanar; hence
the segments do not meet at all. This shows that (i) of 1.3
holds; (ii) follows immediately from the fact that no three
points of C are collinear.   Q.E.D.

1.5 <u>Remarks</u> (1) When drawing graphs it is often convenient
to draw a figure in the plane with heavy dots for vertices and
straight lines (or even curved ones) for edges, possibly inter-
secting at points other than end-points. For example the
complete graph with 5 vertices might be drawn as in the dia-

gram on the left. This is to be
regarded as a recipe for constructing
an abstract graph which Theorem 1.4
assures us can be realized in $\mathbb{R}^3$.
When the edges are straight, as in
the diagram, it can also be regarded
as the projection into the plane of a realization in $\mathbb{R}^3$.
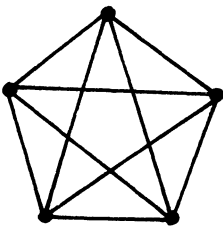
(2) The proof of 1.4 still works if the definition of
"abstract graph" is relaxed to the extent of allowing the set
V of vertices to be countable.

Of course the same abstract graph can be realized in
many ways. The graph in $\mathbb{R}^3$ formed by the edges and vertices

of a cube, and the planar graph (see diagram) are realizations
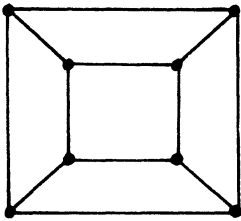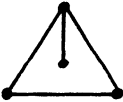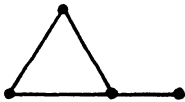
of the same abstract graph.
This, and the fact that the
precise names given to the ver-
tices of an abstract graph are
unimportant, suggests the
following definition.

1.6 Definition   Two abstract graphs (V, E) and (V', E')
are called <u>isomorphic</u> if there is a bijective map f : V → V'
such that

$$\{v, w\} \in E \quad \Longleftrightarrow \quad \{f(v), f(w)\} \in E' .$$

(Thus f is simply a relabelling of vertices which preserves
edges.)   Two graphs are called <u>isomorphic</u> if they are real-
izations of isomorphic abstract graphs.

The corresponding definitions for oriented graphs are
obtained by replacing unordered pairs { , } by ordered pairs
(,).

Thus and are isomorphic
graphs but

 and

are not isomorphic oriented graphs.   The Christmas tree and
Cossack dancer are isomorphic.
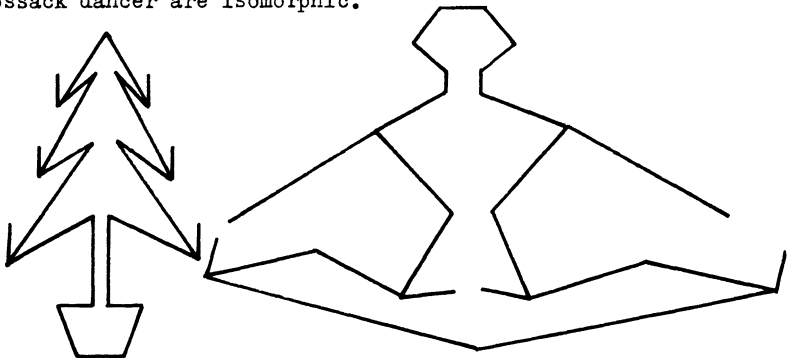
(For once, vertices are not emphasized with dots.)

To prove two graphs are isomorphic we have to construct
an explicit map from the vertex set of one to the vertex set
of the other which is bijective and preserves edges.   To
prove two graphs are not isomorphic we look at properties
of graphs which are invariant under isomorphisms and try to
find such a property possessed by one graph but not by the
other.   The simplest such property is the number of vertices
or edges; next simplest is the orders of the vertices.   The
order of a vertex is the number of edges with the vertex as
an end-point, and clearly two isomorphic graphs must have the
same number of vertices of any given order.   Slightly more
subtle methods are needed for the following two examples.

1.7 Examples   (1X) Exactly two of the following three
graphs are isomorphic.   Which two?   (Compare remark 1.5(1).)
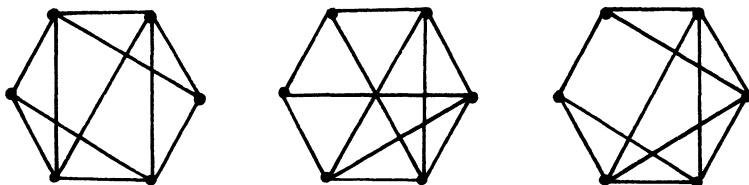
[Hint.   Look at the vertices of order 3.]

   (2X)   The first two graphs below are isomorphic, and the
vertices are labelled to show this.   Which other pairs are
isomorphic?   (Again see 1.5(1).)
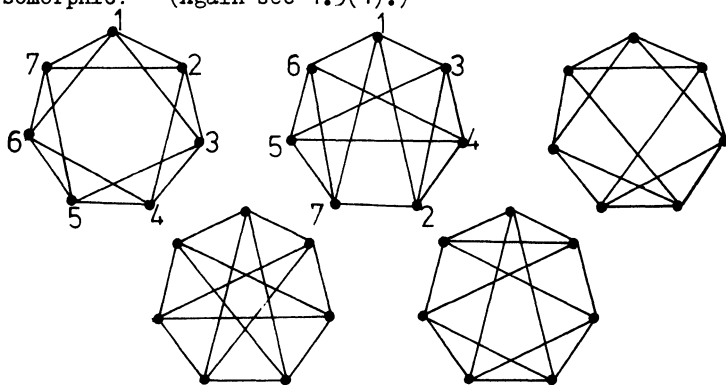
[Hint.   Count triangular paths.]

*KIRCHHOFF'S LAWS

The discussion which follows is intended to show how one
of the fundamental properties of a graph, its "cyclomatic num-
ber", arises naturally in the study of direct current (D.C.)
electrical circuits.   The development of graph theory proper
continues at 1.8.

   We can associate a graph with a D.C. circuit, replacing
terminals by vertices and wires by edges.   An edge joining
a vertex to itself or two or more edges connecting the same
pair of vertices can easily be eliminated by the addition of
extra vertices which do not essentially change the graph from
an electrical point of view.   We can also assume that the
graph is all in one piece, since separate D.C. circuits can
be treated separately.   Each edge has associated with it a
positive real number called the resistance, and some edges
will in addition contain sources of power.   The edges are
oriented (arbitrarily) and the object is to calculate the
current in each edge; a negative answer indicates that the
current is flowing against the chosen orientation.

   Kirchhoff's Laws provide linear equations for the currents.
These equations fall into two types:

   (K1)   There is one equation for each closed loop in the
graph.   A loop is first given a direction, one way round or
the other, and in the resulting equation the only currents
which appear are those along edges in the loop.   The coeffi-
cient of a current is the resistance of the edge if the orien-
tation of the edge agrees with that of the loop; otherwise it
is minus the resistance.   The "right-hand side" of the equa-
tion is the total electro-motive force (e.m.f.) in the loop,
with an appropriate convention on signs.

   (K2)   There is one equation for each vertex, expressing
the fact that the total current flowing into the vertex is
zero (a current flowing out being counted negatively).

Suppose that there are α_0 vertices and α_1 edges; then the α_0 equations (K2) are not independent since if we add them all up the sum vanishes identically.   However it is not hard to verify that any $\alpha_0 - 1$ of them are independent. Since there are α_1 unknowns it follows that in order to determine the currents we shall need at least $\alpha_1 - \alpha_0 + 1$ independent equations (K1) - possibly more, since there may be relations between these and the (K2) equations.   In fact, there are exactly $\alpha_1 - \alpha_0 + 1$ independent equations (K1) (that is, $\alpha_1 - \alpha_0 + 1$ "independent loops") and these are also independent of (K2) and so (K1) and (K2) between them uniquely determine the currents.   A proof of these facts is sketched in an appendix at the end of the chapter, where we have more notation at our disposal.   In the meantime here is a simple proof for the case of a graph with a single loop. For convenience assume the edges e^1, ..., e^n of the loop are all oriented in the same direction round the loop.   Let v^1, ..., v^n be the vertices round the loop; there may be
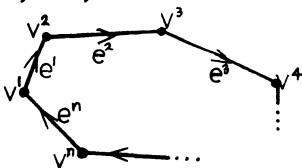
other vertices and edges not in the loop but these do not affect the argument.

Suppose that the (K1) equation of the loop is a linear combination of (K2) equations, the coefficient of the (K2) equation for v^i being λ_i, $i = 1$, ..., n.   The current along edge e^1 can occur only in the (K2) equations for v^1 and v^2, so that the coefficient of this current in the (K1) equation of the loop must be $\lambda_2 - \lambda_1$.   Since this is positive (resistances are positive and e^1 is oriented round the loop) we have $\lambda_2 > \lambda_1$.   In a similar way $\lambda_3 > \lambda_2$, $\lambda_4 > \lambda_3$, ..., $\lambda_n > \lambda_{n-1}$, $\lambda_1 > \lambda_n$.   Hence $\lambda_1 > \lambda_1$, a contradiction.

The "number of independent loops" $\alpha_1 - \alpha_0 + 1$ arising from Kirchhoff's analysis of electrical circuits was called the "cyclomatic number" of the circuit by James Clerk Maxwell.
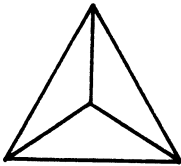
As a simple example, the graph on the
left has cyclomatic number 3 and it
is certainly very plausible that every
loop is a "combination" of the three
loops which bound the regions inside
the big triangle.

MAXIMAL   TREES   AND   THE   CYCLOMATIC   NUMBER

1.8 Definitions   A path on a graph G from v^1 to v^{n+1}
is a sequence of vertices and edges

$$v^1 e^1 v^2 e^2 \ldots v^n e^n v^{n+1}$$

where $e^1 = (v^1 v^2)$, $e^2 = (v^2 v^3)$, ..., $e^n = (v^n v^{n+1})$.
(See 1.3 for notation.   The path goes from v^1 to v^{n+1}
along the edges e^1, ..., e^n.)   Here $n \geqslant 0$: for $n = 0$ we
just have v^1 and no edges.   If G is oriented, we only
require that $e^i = (v^i v^{i+1})$ or $(v^{i+1} v^i)$ for $i = 1$, ..., n;
that is, the edges along the path do not have to point in the
"right" direction.   The path is called simple if the edges
e^1, ..., e^n are all distinct, and the vertices v^1, ..., v^{n+1}
are all distinct except that possibly $v^1 = v^{n+1}$.   If the
simple path has $v^1 = v^{n+1}$ and $n > 0$ it is called a loop;
thus in fact a loop always has $n \geqslant 3$.   We shall regard two
loops as equal when they consist of the same vertices and the
same edges.

A graph G is called connected if, given any two ver-
tices v and w of G there is a path on G from v to
w.   For any non-empty graph G, a component of G consists
of all the edges and vertices which occur in paths starting
at some particular vertex of G.   Thus a non-empty connected
graph has exactly one component.   A graph which is connected
and has no loops is called a tree.

1.9 Examples

(1)
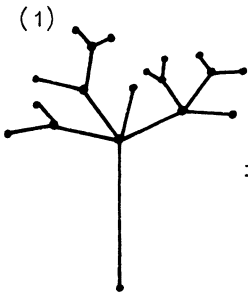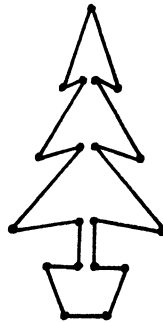
is a tree, but   is not.

(2) Given n points in space - say n points on the
twisted cubic of the proof of 1.4 - we can ask how many trees
there are with those points as vertices.   For instance when
n = 3, the answer is 3:
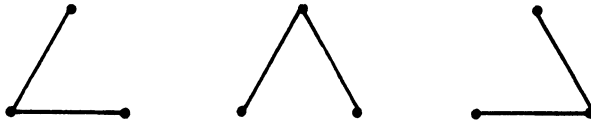
However these are all isomorphic, and a more interesting ques-
tion is: how many trees are there no two of which are iso-
morphic?   The former question has the answer n^{n-2} (n $\geqslant$ 1),
but the general answer to the latter question is unknown.
(See Ore (1962), p.59 for a proof of the first result, known
as Cayley's Theorem.)

(3) If there is a path on G from v to w then by
elimination of any repetitions it is easy to construct a
simple path on G from v to w.

(4X) An edge e of a graph is part of some loop on the
graph if and only if the removal of e does not increase
the number of components of the graph.

1.10 Definitions   Given a graph G, a graph H is called
a subgraph of G if the vertices of H are vertices of G
and the edges of H are edges of G.   H is called a proper
subgraph of G if in addition H $\neq$ G.   Any graph G will
have a subgraph which is a tree (e.g. the empty subgraph is
a tree!) so that the set $\mathcal{J}$ of subgraphs of G which are

trees will have maximal elements. That is, there will be at least one $T \in \mathcal{J}$ such that T is not a proper subgraph of any $T' \in \mathcal{J}$. Such a T is called a <u>maximal tree</u> for G. (Some authors say <u>spanning tree</u>.)

Maximal trees are mainly of interest for connected graphs, and we have the following result.

<u>1.11 Proposition</u>   Let G be a connected graph. A subgraph T of G is a maximal tree for G if and only if T is a tree and contains all the vertices of G.

<u>Proof</u>   We may suppose G is non-empty. Suppose that T is a maximal tree for G. Certainly T is then a non-empty tree; suppose that a vertex v of G fails to belong to T. Choose a vertex w of T and a path from v to w on G. Let v^{i+1} be the first vertex of this path which belongs to T; then the previous vertex v^i and the edge $e^i = (v^i v^{i+1})$ will not belong to T. Adding v^i and e^i to T we obtain say T', a larger subgraph of G than T. Furthermore, e^i cannot be part of any loop on T' since one of its end-points, v^i, does not belong to any edge of T' other than e^i. Thus T' is a tree and T is not maximal.

Suppose conversely that T is a subgraph of G which is a tree and contains all the vertices of G. Suppose that H is a subgraph of G with T a proper subgraph of H. Since T already has all the vertices of G there must be an edge, $e = (vw)$ say, in H but not in T. Since T is connected there is a simple path on T from v to w (compare 1.9 (3)); adding e to the end of this path gives a loop in H. Hence H is not a tree and T is maximal.
                                                                Q.E.D.

<u>1.12 Examples</u>   (1X) Given a connected graph G and a subgraph T which is a tree, T is a subgraph of a maximal tree for G. Indeed, given a subgraph H of G which

contains no loops (but is not necessarily connected), H is
a subgraph of a maximal tree for G. (Use the method of the
first paragraph of the proof of 1.11.)

1.12(2X) A vertex v of an edge (vw) in a graph G is
called free (or terminal) if v is not a vertex of any edge
of G other than e. An inner edge of a graph is an edge
neither of whose end-points is a free vertex. Thus in the
diagram e^1 and e^2 are inner
edges; e^3 is not. v is a
free vertex of e^3. A graph
with at least one edge, all of
whose edges are inner, is
bound to have a loop (since the number of edges is finite);
it follows that a tree with at least one edge always possesses
a non-inner edge. Removing a non-inner edge e of a graph,
together with a free vertex of e, is called an elementary
collapse, and a sequence of these is called a collapse. (We
shall meet this idea again in 3.30.) Since a collapse takes
a tree to a smaller tree it is easy to prove that any tree
collapses to a single vertex. Indeed, the converse is true:
if a graph collapses to a single vertex, then it must have
been a tree.
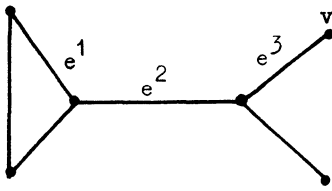
The following definition and theorem connect the idea of
maximal tree with the number $\alpha_1 - \alpha_0 + 1$ introduced during
the discussion of Kirchhoff's laws.

1.13 Definition   Let G be a connected graph with
$\alpha_0 = \alpha_0(G)$ vertices and $\alpha_1 = \alpha_1(G)$ edges. The cyclomatic
number of G is the integer

$$\mu = \mu(G) = \alpha_1 - \alpha_0 + 1 .$$

1.14 Theorem   Let G be a graph with cyclomatic number μ.
Then

(1) $\mu \geqslant 0$;

(2) $\mu = 0$ if and only if G is a tree;

(3) every maximal tree for G has $\alpha_1 - \mu$ edges. That is, every maximal tree for G can be obtained from G by removing μ suitable edges (and no vertices). These edges must be inner edges since T is connected;

(4) removing fewer than μ edges from G cannot produce a tree; removing more than μ edges always disconnects G.

Proof   The proof of (1) and (2) is by induction on α_1. Clearly (1) and (2) are true when $\alpha_1 = 0$; suppose they are true when $\alpha_1 < k$, where $k \geqslant 1$, and let G have k edges.

(i) Suppose G is a tree. Then G has a non-inner edge (compare 1.12 (2)), e say. Removing e and a free vertex of e from G leaves a tree G' with $k - 1$ edges. By the induction hypothesis, $\mu(G') = 0$, and by construction $\mu(G') = \mu(G)$. Hence $\mu(G) = 0$.

(ii) Suppose G is not a tree. Then G has an inner edge, namely any edge of a loop of G; let G' be obtained by removing this edge (and no vertices). Removing an edge from a loop cannot disconnect G, so that G' is connected and by the induction hypothesis $\mu(G') \geqslant 0$. By construction $\mu(G') = \mu(G) - 1$; hence $\mu(G) \geqslant 1$.

This completes the proof by induction of (1) and (2).

To prove (3), let T be a maximal tree for G, so that by 1.11 $\alpha_0(G) = \alpha_0(T)$. Using this and $\mu(T) = 0$ it follows easily that $\alpha_1(T) = \alpha_1(G) - \mu(G)$.

To prove (4), note that removing p edges (and no vertices) from G leaves either a disconnected graph or else a connected graph with cyclomatic number $\mu(G) - p$. From this and (1) and (2) we can deduce (4).   Q.E.D.

1.15 Remarks   (1) Removing μ edges at random may not leave a connected graph, hence certainly not a tree. However if it does leave a connected graph then this must be a maximal

tree: it will contain all the vertices and have cyclomatic
number $\mu - \mu = 0$.

1.15(2) Let G be a connected graph with $\mu(G) = 1$. Then
G has exactly one loop. (Proof: G has at least one loop
since $\mu(G) > 0$. If there are two loops let e be an edge
of one loop not in the other. Removing e from G leaves
G connected, and still containing a loop, but it also lowers
μ by one to zero. This is a contradiction.)

1.16 The basic loops   Let G be a connected graph with
cyclomatic number μ and let $e^1, \ldots, e^\mu$ be edges whose
removal from G leaves a maximal tree T (see 1.14 (3)).
We shall denote by $T + e^i$ the graph obtained from T by
replacing e^i. Then $\mu(T + e^i) = 1$, and by 1.15 (2) there
is a unique loop ℓ^i on $T + e^i$. We shall prove that in
some sense these loops "generate" all loops on G; one con-
sequence of this is that the Kirchhoff equations for
$\ell^1, \ldots, \ell^\mu$ are enough to determine all equations of type
(K1) (page 18). Before doing this we shall generalize the
idea of loop.

CHAINS AND CYCLES ON AN ORIENTED GRAPH

Let G be an oriented graph.   Given a path

$$v^1 e^1 v^2 e^2 \ldots v^n e^n v^{n+1}$$

on G we can associate with it a formal sum

$$\epsilon_1 e^1 + \epsilon_2 e^2 + \ldots + \epsilon_n e^n$$

where $\epsilon_i = +1$ if $e^i = (v^i v^{i+1})$ and $\epsilon_i = -1$ if
$e^i = (v^{i+1} v^i)$. Thus the formal sum is obtained by travelling
along the path from v^1 to v^{n+1} and writing down the edges
encountered, with a minus sign if they are oriented against
the direction of travel.   More generally, we have the following.

1.17 Definition   Let G be an oriented graph with edges $e^1, \ldots, e^r$. A 1-chain on G is a formal sum

$$\lambda_1 e^1 + \lambda_2 e^2 + \ldots + \lambda_r e^r$$

where each λ_i is an integer. When writing down 1-chains we omit any edges which have coefficient 0. The 1-chain in which all coefficients are zero is denoted 0. We define the sum of the 1-chains $\Sigma \lambda_i e^i$ and $\Sigma \lambda_i' e^i$ to be

$$\Sigma \lambda_i e^i + \Sigma \lambda_i' e^i = \Sigma(\lambda_i + \lambda_i')e^i$$

which is also a 1-chain. With this definition the set of 1-chains on G is an abelian group, denoted $C_1(G)$. (The verification that the 1-chains form an abelian group is left to the reader.)

In a similar way, a 0-chain on G is a formal sum

$$\lambda_1 v^1 + \ldots + \lambda_m v^m$$

where $v^1, \ldots, v^m$ are the vertices of G. 0-chains are added by adding coefficients, and the set of 0-chains is an abelian group, denoted $C_0(G)$.

1.18 Example   The 1-chain associated to a path is obtained is the manner described before 1.17. Note that the edges in a path need not be distinct (unless the path is simple), so
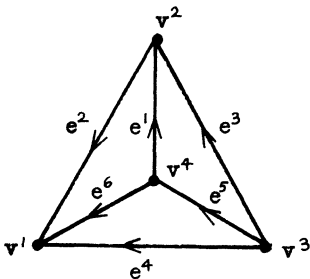
that the final coefficient of an edge may be a number other than 0, 1 or -1. For instance, the 1-chain associated with the path $v^4 e^1 v^2 e^2 v^1 e^6 v^4 e^1 v^2 e^3 v^3$ on the graph illustrated

is $e^1 + e^2 - e^6 + e^1 - e^3 = 2e^1 + e^2 - e^3 - e^6$.

<u>1.19</u> <u>Definition</u>   Let G be an oriented graph.   For any
edge e $=$ (vw) of G (regarded as a 1-chain), we define
the <u>boundary</u> of e to be ∂e $=$ $w - v$ (regarded as a 0-chain).
The <u>boundary homomorphism</u>

$$\partial \; : \; C_1(G) \; \to \; C_0(G)$$

is defined by $\partial(\Sigma \lambda_i e^i) = \Sigma \lambda_i \partial(e^i)$.   A <u>1-cycle</u> on G is
an element $c \in C_1(G)$ such that $\partial c = 0$.   The group of
1-cycles on G is denoted $Z_1(G)$.

As an example, the 1-chain associated to any loop on G
is always a 1-cycle; e.g. for the loop $v^1 e^2 v^2 e^1 v^4 e^5 v^3 e^4 v^1$
on the graph on the previous page, the associated 1-cycle is
$-e^2 - e^1 - e^5 + e^4$ and

$$\partial(-e^2 - e^1 - e^5 + e^4)$$
$$= \; -(v^1 - v^2) - (v^2 - v^4) - (v^4 - v^3) + (v^1 - v^3) \; = \; 0.$$

Thus cycles are generalizations of loops; the following theorem
describes cycles completely on a connected graph.

Now let G be a connected, oriented graph with cyclo-
matic number μ.   Let z be a 1-cycle on G, and let
$\ell^1, \ldots, \ell^\mu$ be the loops described in 1.16.   We denote the
1-cycle associated to ℓ^i by z^i, the direction of travel
round ℓ^i being chosen so that e^i occurs in z^i with
coefficient $+ 1$.

<u>1.20</u> <u>Theorem</u>   There are unique integers $\lambda_1, \ldots, \lambda_\mu$ such
that

$$z \; = \; \sum_{i=1}^{\mu} \lambda_i z^i$$

(if $\mu = 0$ this is to be read $z = 0$).   In fact, λ_i is the
coefficient of the edge e^i (see 1.16) in z.

<u>Proof</u>   First suppose that G is a tree, i.e. that

$\mu = 0$ and let z be a 1-cycle on G. We have to prove that $z = 0$, i.e. that there are no nontrivial 1-cycles on a tree. (This generalizes very minutely the fact that there are no loops on a tree.) The result is clear if G has no edges, so suppose it true for trees with $< k$ edges $(k \geqslant 1)$ and let G have k edges. Let e be a non-inner edge of G, and v a free vertex of e (see 1.12 (2)). The coefficient of e in z must be zero, since otherwise the coefficient of v in ∂z would not be zero. Hence z is a 1-cycle on the
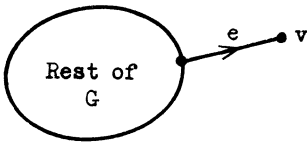
graph obtained by removing e and v from G. This graph is a tree with $k - 1$ edges; by induction $z = 0$.

Now suppose $\mu > 0$, and let z be a 1-cycle on G. Let λ_i be the coefficient of e^i in z, with e^i as in 1.16, and consider

$$z' = z - (\lambda_1 z^1 + \ldots + \lambda_\mu z^\mu).$$

Certainly z' is a 1-cycle, since z and the z^i are. Also the coefficient of each e^i in z' is 0, by choice of the λ^i and since the coefficient of e^i in z^i is $+1$ while e^i does not occur in z^j for $j \neq i$. Thus z' is actually a 1-cycle on the graph obtained by removing the edges $e^1, \ldots, e^\mu$ from G. This graph is a tree (maximal for G) and so by the first part of the proof $z' = 0$. This proves the required formula.

It remains to verify that the given expression for z is unique. To see this, suppose that

$$\sum_{i=1}^{\mu} \lambda_i z^i = \sum_{i=1}^{\mu} \lambda'_i z^i.$$

Comparing coefficients of e^i on the two sides gives $\lambda_i = \lambda'_i$.

Q.E.D.

1.21 Remarks and Examples   (1X) Theorem 1.20 will be of
considerable significance later when we come to calculate
homology groups.   For the present it shows that all 1-cycles,
and hence all loops, on a connected graph, are in a precise
sense combinations of the μ basic loops of 1.16.   It follows
that all the Kirchhoff equations for loops are consequences
of the equations for these basic loops.

(2X) A basis for the 1-cycles on a connected graph G
is a collection of 1-cycles of G in terms of which every
1-cycle on G can be expressed uniquely as a linear combina-
tion with integer coefficients.   Thus $z^1, \ldots, z^\mu$ in 1.20
is a basis.   But not every basis arises in this way from a
maximal tree.   For instance,
it is impossible to find
edges e^1, e^2, e^3, e^4 of
the graph shown, such that
the resulting cycles
z^1, z^2, z^3, z^4 are the basis
given by the boundaries of
the four small equilateral
triangles.
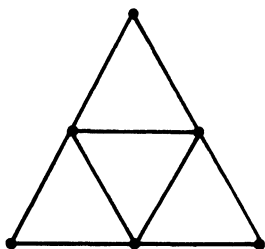
(3) 1-cycles are defined only on an oriented graph.
Suppose that G is an unoriented graph and G_1 and G_2 are
oriented graphs each obtained from G by orienting the edges
of G arbitrarily.   Then there is a natural isomorphism

$$C_1(G_1) \; \to \; C_1(G_2)$$

defined by $(vw) \to (vw)$ or (wv) according as the edge
(vw) of G receives the same or opposite orientations in
G_1 and G_2.   This clearly takes cycles to cycles, so that
$Z_1(G_1) \cong Z_1(G_2)$.   Orientation is a technical device intro-
duced to enable the boundary homomorphism (1.19) to be defined,
but the group of 1-cycles does not depend, up to isomorphism,

on the orientation.   In fact an equally satisfactory theory
can be developed for unoriented graphs using 1-chains and
0-chains with coefficients in the group $\mathbb{Z}_2$: this will be
dealt with in a more general context in Chapter 8.   The "$\mathbb{Z}_2$
theory" is, however, inadequate for some situations we shall
meet later, and that is why the other theory has been introduced
from the start.

1.21 (4X) Suppose that G is a connected graph with cyclo-
matic number μ.   Suppose that there exist edges $e^1, \ldots, e^\mu$
and loops $\ell^1, \ldots, \ell^\mu$ such that, for each i, e^i is an
edge of ℓ^i but not of ℓ^j for $j \neq i$.   Show that removing
$e^1, \ldots, e^\mu$ from G leaves a maximal tree.

(5X) A graph G is called __bipartite__ if its vertex set
can be partitioned into two disjoint subsets V_1, V_2 in such
a way that every edge of G joins a vertex in V_1 to a ver-
tex in V_2.   Show that G is bipartite if and only if every
loop in G has an even number of edges.   (Hence a tree is
certainly bipartite!) [Hint for 'if'.   Concentrate on one
component G_1 of G.   Define a homomorphism $\eta : C_1(G_1) \to \mathbb{Z}_2$
by $\eta c = 0$ or 1 according as the sum of coefficients in
c is even or odd.   Show (using 1.20) that $\eta z = 0$ for any
1-cycle z on G_1.   Now look at the 1-chains associated to
paths from a fixed vertex v of G_1 to an arbitrary vertex
w of G_1.   A direct proof can also be given by induction on
$\mu(G_1)$.]

(6XX) Let $\ell^1, \ldots, \ell^k$ be loops on an oriented connected
graph and let $z^1, \ldots, z^k$ be the 1-cycles given by these
loops with a choice of direction.   Let us call the loops
"independent" if, for all integers $\lambda_1, \ldots, \lambda_k$, we have

$$\lambda_1 z^1 + \ldots + \lambda_k z^k = 0 \implies \lambda_1 = \ldots = \lambda_k = 0 .$$

Find loops ℓ^1, ℓ^2, ℓ^3 on the complete graph with four
vertices which are independent, but such that z^1, z^2, z^3 do

not form a basis for the 1-cycles.

Show that two loops are independent if and only if they are distinct, and that three loops are independent if and only if none of the z^i is an integral linear combination of the other two.

1.21 (7X) Let G be an oriented graph with k vertices containing no oriented loop (an oriented loop is one in which all the edges point the same way round the loop). Show that it is possible to label the vertices of G by $v^1, \ldots, v^k$ in such a way that any oriented edge of G is of the form $(v^i v^j)$ with i > j. (Hint. Use induction on k. Note that, unless G has no edges, there must exist a vertex of G which is the beginning point of some edge but not the endpoint of any edge, for otherwise G would certainly contain an oriented loop.) The above result is used in Bernstein, Gel'fand and Ponomarev (1973) on p.24 - note that they use "cycle" in the sense of our word "loop". This paper contains deep connexions between graph theory and linear algebra.

Having investigated the kernel of the boundary homomorphism (1.19), namely the 1-cycles, we shall say something about its image.

1.22 Definition   Let G be an oriented graph.   The augmentation map is the homomorphism

$$\epsilon : C_0(G) \rightarrow \mathbb{Z}$$

given by

$$\epsilon(\Sigma \lambda_i v^i) = \Sigma \lambda_i .$$

1.23 Theorem   Let G be a connected oriented graph.   Then the sequence

$$C_1(G) \xrightarrow{\partial} C_0(G) \xrightarrow{\epsilon} \mathbb{Z}$$

is exact, that is to say   Image ∂ = Kernel ϵ .

Proof   To show that   Image $\partial \subset$ Kernel ϵ it is
enough to verify that, for any edge   e   of   G, $\epsilon \partial e = 0$.
This is immediate.   For the converse, any element of Kernel ϵ
can be written in the form

$$\sum_{i=1}^{m-1} \lambda_i (v^i - v^m)$$

(where $v^1, \ldots, v^m$ are the vertices of   G), each λ_i being
an integer.   Since   G   is connected there is a path from v^m
to v^i and the 1-chain c^i associated to this path has
$\partial c^i = v^i - v^m$.   Thus the above 0-chain equals

$$\partial (\Sigma \lambda_i c^i) \in \text{ Image } \partial \quad . \qquad \text{Q.E.D.}$$

*PLANAR   GRAPHS

A planar graph is a graph which is isomorphic to a graph in
the plane $\mathbb{R}^2$.   In this section we shall investigate some
properties of planar graphs, mostly connected with the idea
of cyclomatic number.   At the end of the section we shall
take up the question of whether our insistence that graphs
have straight edges places any essential restriction on the
class of planar graphs.

    Some proofs will be omitted from this section and others
will not be given in full detail; this is in order to avoid
devoting undue space to what is, in fact, a subject of peri-
pheral interest to the study of homology theory.   Besides,
we shall study graphs in more general "surfaces" in Chapter 9.
At the end of the section the reader will need to be familiar
with some point-set topology if he is to fill in all the
details in the proof of 1.34.   I hope, however, that the
general idea of the proof will be clear to readers without
such technical knowledge.

1.24 Definitions   A polygonal arc joining two points   a
and   b in $\mathbb{R}^N$ is a sequence of straight segments placed
end-to-end, starting at   a and finishing at   b.   A poly-
gonal arc is called simple if it has no points of self-inter-
section except that possibly   a = b;   thus a simple poly-
gonal arc with a $\neq$ b   can be thought of as a realization
of the abstract graph given, for some   n $\geqslant$ 2,   by

$$V = \{v^1, \ldots, v^n\} \quad \text{(n distinct vertices)}$$

$$E = \{\{v^1, v^2\}, \{v^2, v^3\}, \ldots, \{v^{n-1}, v^n\}\}.$$

A simple closed polygon is a simple polygonal arc with   a = b;
equivalently it is a realization of the abstract graph given,
for some   n $\geqslant$ 3,   by the same   V and E as above, except
that   E has an additional edge $\{v^n, v^1\}$.

polygonal arc   simple poly-   simple closed poly-
                       gonal arc   gon
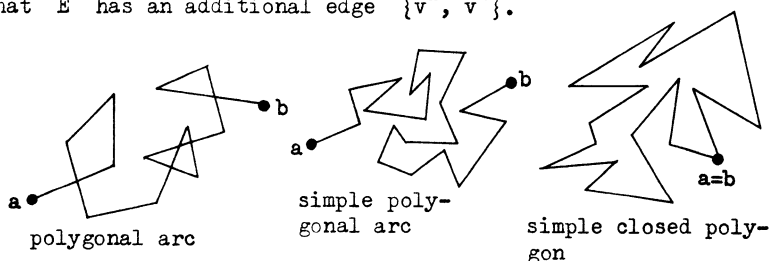
        It is very plausible that a tree   T   realized in $\mathbb{R}^2$
will not separate the plane - i.e.,   that any two points of
the plane not on   T   can be joined by a polygonal arc which
does not intersect   T.   Indeed this can be proved quite
easily by induction on the number of vertices of   T,   and
the same proof works when   T   is replaced by a disjoint union
of trees (a "forest").   It is also very plausible that a
simple closed polygon does separate the plane, and this is
the content of the next theorem.

        We call a subset of the plane bounded if it can be con-
tained inside a suitable circle.   Otherwise we call it un-
bounded.   We say that a subset   P of the plane separates

the plane into two disjoint subsets I and O if the
following hold.

(i) $I \cup O = \mathbb{R}^2 \setminus P$, $I \cap O = \emptyset$.

(ii)   Any two points of I (resp. O) can be joined by
a polygonal arc, hence by a simple polygonal arc, in I (resp. O).

(iii)   Any polygonal arc joining a point of I to a point
of O intersects P.

(These say that $\mathbb{R}^2 \setminus P$ has precisely two "path compo-
nents".)

<u>1.25 Jordan Curve Theorem for polygons</u>   Let P be a simple
closed polygon in the plane. Then P separates the plane
into two disjoint subsets I and O of which precisely one,
O say, is unbounded. (We call I the "inside" and O
the "outside" of P.)   Furthermore if e is any edge of P
and x and y are points of $\mathbb{R}^2 \setminus P$ sufficiently close to
e, on opposite sides of it and away from the end-points,
then precisely one of x, y belongs to I and (hence) the
other belongs to O.

A proof of 1.25 can be found in Courant and Robbins (1947),
pp. 267-9.   The last sentence of 1.25 follows from their
proof, or from the rest of the theorem together with the
remarks on trees above.

A graph G in $\mathbb{R}^2$ will divide the plane into "regions",
two points not on G being in the same region if and only if
they can be joined by a polygonal arc in $\mathbb{R}^2$ not intersecting
G.   Exactly one of these regions will be unbounded.   (The
graph G can be entirely contained inside a suitable circle,
and any unbounded region will contain points outside this
circle; on the other hand it is clear that any two points
outside the circle belong to the same region.)   Thus accord-
ing to 1.25 a simple closed polygon will divide the plane

into two regions.   Our first objective is to count the regions
for an arbitrary graph G.

Let e be an edge of a graph G in $\mathbb{R}^2$.   Points near
e and on one side of e will belong to one region, and e
is called <u>adjacent</u> to that region; thus e is adjacent either
to one region or to two different regions.   It is clear that
if e is part of a loop on G, then by 1.25 the regions on
the two sides of e will be distinct: e will be adjacent
to two regions.   The converse is also true: if e is not
part of any loop on G then e is adjacent to only one
region.   A proof is suggested in 1.30 (3) below.   Combining
this with the fact, already noted in 1.9 (4) for arbitrary
graphs, that an edge of a graph belongs to a loop if and only
if removing the edge does not increase the number of compon-
ents of the graph, we have the following result.

1.26 Proposition   Removing an edge e from a graph in
$\mathbb{R}^2$ leaves the number of components of the graph unchanged
if and only if it reduces the number of regions by one.

It is now an easy matter to count the regions.

1.27 Theorem   Let G be a graph in $\mathbb{R}^2$ with $\alpha_0(G)$ ver-
tices, $\alpha_1(G)$ edges and $c(G)$ components.   Then G divides
the plane into $r(G) = \alpha_1(G) - \alpha_0(G) + c(G) + 1$ regions.

Proof   Let us remove the edges from G one at a
time, leaving the vertices untouched, and note the effect on
the value of $\alpha_1 - \alpha_0 + c + 1 - r$.   In fact 1.26 says pre-
cisely that the value remains constant, and since when all
the edges have gone it is $0 - \alpha_0(G) + \alpha_0(G) + 1 - 1 = 0$
it must have been 0 to begin with.   This is the result.
$$\text{Q.E.D.}$$
The formula for $r(G)$ given in 1.27 is known, at least
for connected graphs $(c = 1)$, as <u>Euler's formula</u>.   For
$c = 1$ it reads

<u>1.28</u> $\alpha_0(G) - \alpha_1(G) + r(G) = 2$

or equivalently

<u>1.29</u> $r(G) = \mu(G) + 1 .$

<u>1.30 Remarks and Examples</u>   (1X) The most noticeable thing
about the formula in 1.27 is that it exists at all: the number
of regions depends only on the abstract structure of G and
not on the way in which G is placed in the plane.   It is
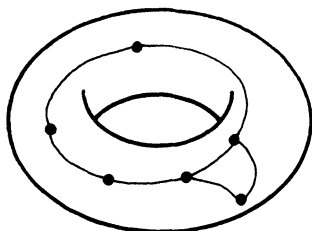
possible to draw "graphs" in surfaces
other than the plane, for example
in the torus surface.   The edges
will be curved lines on the surface,
and again we can count regions.
For the graph illustrated, $\alpha_0 = 6$,
$\alpha_1 = 7$,   c = 1 and   r = 2.   However it is possible to find
isomorphic graphs in the torus (these will have the same α_0,
α_1 and c), for which r = 1 or r = 3.   Thus the number
of regions does not depend only on the abstract structure of
the graph.   This phenomenon will be investigated in Chapter 9.
What goes wrong when the argument leading to 1.26 is applied
to graphs in the torus?

     (2X) On the circumference of a circle in the plane n
points are chosen and the chords joining every pair of points
drawn.   Assuming that no three of these chords are concurrent
(except at one of the n points) show that the interior of
the circle is divided into $1 + \binom{n}{2} + \binom{n}{4}$ regions.
[Hint.   Use Euler's formula.   The number of points of inter-
section of chords inside the circle is $\binom{n}{4}$ since any 4 of
the n points uniquely determine such a point of intersection.
It is amusing to note that for n = 1, 2, 3, 4, 5, 6 the
number of regions is 1, 2, 4, 8, 16, 31.]

     (3X) Here is a suggestion for proving that, if an edge
e of a graph G in $\mathbb{R}^2$ is not part of any loop on G,   then

the regions on the two sides of e are the same.   First
dispose of the case when G is a tree (e.g. using the
remarks preceding 1.25), and then proceed by induction on the
number of edges in G.   Remove an edge e' from a loop ℓ
on G, noting that e $\neq$ e' and e, apart possibly from its
end-points, is entirely inside or entirely outside ℓ.
Choose a polygonal arc from one side of e to the other,
using the induction hypothesis.   The only snag is that this
arc may cross e'.   It is not difficult to change the arc
so that it doesn't cross e'; a hint is given in the diagram.
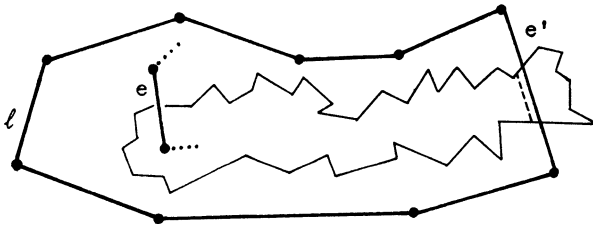
We shall look more closely at the formula 1.29 for con-
nected graphs in $\mathbb{R}^2$.   We already know, from 1.20, that there
is a basis for the 1-cycles
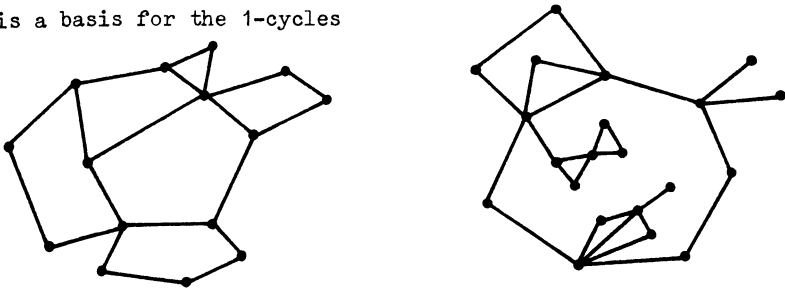
on G which consists of $\mu(G)$ elements.   It is a consequence
of a theorem of group theory (see A.30) that every basis has
this number of elements.   Now in the case of the left-hand
illustration above (where $\mu(G) = 6$) it is very plausible
that the boundaries of the six bounded regions give six cycles
forming a basis for the 1-cycles on G.   (It is at least
"clear" that every loop on G can be obtained by combining

these six loops suitably.)   Perhaps this is not so obvious
in the case of the right-hand illustration simply because it
is not so clear which are the loops to take.   If we can
associate one basic loop with each bounded region, then this
will confirm that $r(G)$ = $\mu(G)$ + 1,   the "1" on the right-
hand side coming from the unbounded region.   In fact such an
association is possible, as stated in the next proposition,
which will be given without proof.

Let   G   be a connected oriented graph in $\mathbb{R}^2$,   and let
R   be one of the regions into which   G   divides $\mathbb{R}^2$.   The
frontier $\mathcal{F}(R)$ of   R   consists of those edges of   G   which
are adjacent to   R,   and their end-points.   Thus $\mathcal{F}(R)$ is
a subgraph of   G.   Now suppose that   R   is a bounded region,
and consider just the graph $\mathcal{F}(R)$ in $\mathbb{R}^2$,   deleting the
remainder of   G.

1.31 Proposition   The edges of $\mathcal{F}(R)$ which are adjacent
to the unbounded region (when the remainder of   G   is deleted),
together with their end-points, form a simple closed polygon.
This gives us a loop on   G;   the loops associated in this way
to the   various bounded regions give a basis for the 1-cycles
on   G.

Note that the basis provided in this way may be different
from any given by a maximal tree; see 1.21 (2).   As an example

of 1.31, the frontier of one
of the   regions in the right-
hand illustration on the pre-
vious page is drawn on the
left, with the associated loop
drawn heavily.

Another noteworthy point about the basis described in
1.31 is that any edge of   G   belongs to at most two of the
basic loops.   A remarkable theorem of S. MacLane (1937)
asserts that the converse is true: if a connected graph   G

possesses loops a^1, ..., a^μ, whose corresponding 1-cycles
form a basis for the 1-cycles on G and which have the pro-
perty that each edge of G belongs to at most two of the
loops a^i, then G is planar. Two other criteria for the
planarity of a graph are known; one is due to Kuratowski
(see Berge (1962), p. 211), and the other to Whitney (1932).

    We can make a further deduction from 1.31. Suppose
that G (in $\mathbb{R}^2$) is connected, with $\mu(G) \geqslant 1$. Each
loop associated to a bounded region will have at least 3
edges, so that the sum of the numbers of edges in these loops
is at least 3μ. Also the frontier of the unbounded region
will contain at least 3 edges, and adding these in gives a
sum of at least $3\mu + 3$. This sum is $\leqslant 2\alpha_1(G)$ since an
edge of G belongs to at most two of the loops and to at
most one if the edge is adjacent to the unbounded region.
From this we deduce the following.

1.32 Proposition   If G is a connected graph in $\mathbb{R}^2$
satisfying $\mu(G) \geqslant 1$, then $3\alpha_0(G) - \alpha_1(G) \geqslant 6$.

    The case $\mu(G) = 0$ is left to the reader. As an
example of the proposition let G be the complete graph with
5 vertices. For this G, $\alpha_0 = 5$, $\alpha_1 = 10$. (Also
$\mu = 6 \geqslant 1$.) Thus $3\alpha_0 - \alpha_1 < 6$, and G is not planar.

1.33 Example   (X) Show that if G is a connected bipar-
tite graph (1.21 (5)) in $\mathbb{R}^2$ satisfying $\mu(G) \geqslant 1$, then
$2\alpha_0(G) - \alpha_1(G) \geqslant 4$. Find a bipartite graph which is not
planar.

    Given a graph G we can "subdivide" it by adding
extra vertices at interior points of edges; this automatically
increases the number of edges by the same amount, thus leav-
ing $\mu(G)$ unchanged. For example if we subdivide the com-
plete graph with 5 vertices by adding a single vertex on one
of the edges the new graph has $\alpha_0 = 6$, $\alpha_1 = 11$ and it is

no longer true that $3\alpha_0 - \alpha_1 < 6$.   Thus 1.32 does not show
that the subdivided graph is non-planar.   This indicates, of
course, how feeble the result 1.32 is; all the same it is
a priori conceivable that by adding enough extra vertices on
the edges of a non-planar graph we could make it planar -
this is rather like trying to draw it in the plane with curved
edges, since these could be approximated by broken lines.

A theorem of I. Fáry (1948) implies that this does not
happen: if some subdivision of a graph G is planar, then G
is planar.   It would take us too far out of our way to give
a full and rigorous proof of Fáry's theorem, since this would
involve the formal introduction of some notions of point-set
topology with which we are not otherwise concerned.   However
there follows a proof that is complete except for these
details.   It is a variant of Fáry's original proof, and was
found by P. Damphousse.   In fact Damphousse's argument proves
a stronger result than that proved by Fáry (this stronger
result is stated in Fáry's paper).

We shall consider "curved graphs" in the plane which
differ from the graphs considered so far only in that their
edges are allowed to be curved lines (precisely: homeomorphic
images of the closed interval $[0, 1]$) instead of straight
segments.   Thus a <u>curved graph</u> $\mathcal{G}$ consists of a finite set
of points (vertices) in the plane, some pairs of distinct
vertices being joined by curved lines (edges) which do not
intersect except possibly at end-points.   An edge does not
intersect itself either: it is a simple curve.   Furthermore
any particular pair of vertices is joined by at most one
edge.   It is possible to formulate and prove Euler's formula
1.27 for curved graphs: both formulation and proof require
the most general form of the Jordan Curve Theorem which states,
roughly speaking, that a simple closed curve always separates
the plane into two regions.   (See Newman (1951), p. 115 or
Lefschetz (1949), p. 61.   There are also several proofs in

Wall (1972).) In this sketch we shall assume the validity
of Euler's formula for curved graphs, and of some other
"plausible" results to be mentioned later.

The weak form of Fáry's theorem asserts that the abstract
graph underlying g can be realized in $\mathbb{R}^2$ (see 1.3), i.e.
that there is a (straight) graph G in $\mathbb{R}^2$ with the same
abstract structure as g. The strong form is as follows.

1.34 Fáry's Theorem   With the above notation, there is a
homeomorphism of the plane (that is, a bijective continuous
map $\mathbb{R}^2 \to \mathbb{R}^2$ with continuous inverse) mapping g onto a
(straight) graph G in such a way that vertices of g are
taken to vertices of G and edges to edges.

Proof   The proof is in several parts. First we
make changes to g to simplify its structure (I); (II)-(IV)
are consequences of these changes. (V) is the central argu-
ment (by induction on the number of vertices of g) and (VI)
is a small result needed in (V). We assume throughout that
g has at least four vertices, the other cases being easily
disposed of.

(I) We turn g into a <u>triangular</u> graph g^Δ, that is
a curved graph for which each one of the regions into which
g^Δ divides the plane has frontier (see before 1.31) a simple
closed curve consisting of precisely three edges of g^Δ.
This is achieved by a sequence of operations on g, as follows.
Select any two vertices of g which are on the frontier of
the same region and which are not already connected by any
edge of g; connect them by a simple curve within the region
to give a graph g'. Now select a pair of vertices of g'
which are on the frontier of the same region defined by g'
and are not connected by any edge of g' and connect them
within the region. This procedure is continued until it is
no longer possible to find a pair of vertices of the sort
described - this must eventually happen since no new vertices

are introduced and consequently the number of possible edges
is bounded.   An example is drawn below; the dotted edges are
the ones introduced.
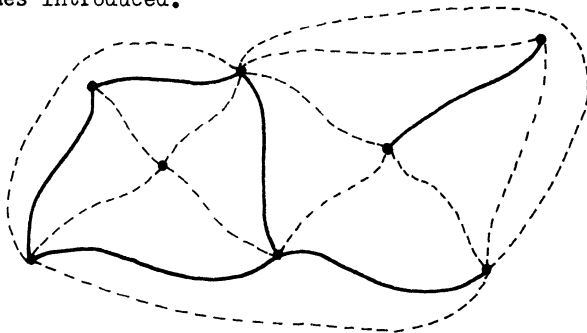
       To see that the result is triangular, suppose that a
region has frontier consisting of say four edges, as in the
diagram below.   By the construction A and C must be joined
by an edge outside the region and so must B and D.   But
this is impossible in $\mathbb{R}^2$.   (This is "obvious" but formally
requires the Jordan Curve Theorem.)

       (II) In any triangular graph with at least four vertices
there can be no vertex of order 0, 1 or 2.   To see, for
example, that there can be no vertex of order 2 suppose that
there is one - B say, in the diagram above.   As there is at
least one vertex other than A, B, C there must be such a
vertex on the frontier of one of the regions adjacent to the
arc ABC.   But this vertex is not joined to B, so (I) has
not been completed.

       (III) In any triangular graph with at least four vertices
there are at least four vertices of order $\leqslant 5$.   From the
triangularity it follows that g is connected and $3r = 2\alpha_1$

(r being the number of regions) and this together with Euler's
formula 1.28 shows $6\alpha_0 - 2\alpha_1 = 12$. Now writing N_i for
the number of vertices of order i (so that $N_0 = N_1 = N_2 = 0$
from (II)), we have $\alpha_0 = N_3 + N_4 + N_5 + \ldots$ and
$2\alpha_1 = 3N_3 + 4N_4 + 5N_5 + \ldots$ so that $12 = 6\alpha_0 - 2\alpha_1 =$
$3N_3 + 2N_4 + N_5 - N_7 - 2N_8 - \ldots \leq 3N_3 + 2N_4 + N_5$. Hence
$N_3 + N_4 + N_5 \geq \frac{1}{3}(3N_3 + 2N_4 + N_5) \geq 4$, as required.

(IV) There is at least one vertex v of g^Δ, of order
≤ 5 and not on the frontier of the unbounded region. This
is because there are precisely three vertices of g^Δ on the
frontier of the unbounded region, and by (III) at least four
vertices of g^Δ of order ≤ 5. Now consider the edges with
v as one end-point; they can be ordered, say anticlockwise,
round v. (This statement is not as innocent as it looks,
for the edges could for example spiral round v in a compli-
cated manner. What they cannot do is cross each other, and
that is what makes the result, which follows from the Jordan
Curve Theorem, true.) Consecutive pairs of these edges will
terminate in pairs of vertices, and these pairs will be joined
by edges of g^Δ which will form a simple closed curve $\mathcal{C}$
enclosing v.

not
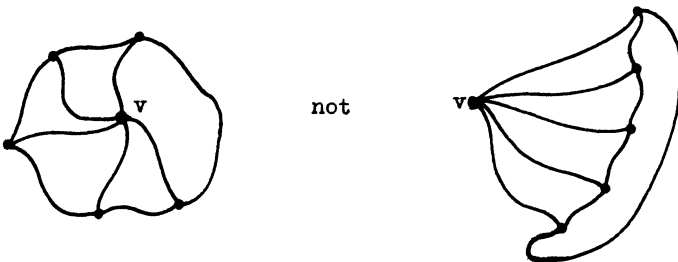
(It is not hard to see that the right-hand picture can only
arise when v is on the frontier of the unbounded region.)

(V) This is the induction argument by which the theorem
is proved. Suppose 1.34 is true whenever g has $< k$ ver-
tices ($k \geq 4$) and let g have k vertices. (I) - (IV)

apply to g, which turns into a triangular graph g^Δ; we then
remove the v of (IV) together with its attendant edges,
leaving a graph g_1, say. Now use the induction hypothesis
to choose a homeomorphism h of the plane to itself mapping
g_1 on to a straight graph G_1. This will take e (see (IV))
to a simple closed polygon C say, with 3, 4 or 5 sides and
with no vertex inside it since h will preserve the interior
of a simple closed curve. It remains to "fill in" the miss-
ing vertex and edges of g and change h inside e so that
it maps them to a point P inside C and straight segments
joining P to the vertices around C. Thus we need, in the
first place, the existence of a point P within C from
which all vertices of C are "visible". This is dealt with
in (VI).

Having found such a P it is possible to redefine h
inside e as above. (First we ensure that $h(v) = P$, then
that the curved edges through v are mapped to the straight
edges through P. It then remains to extend h to the int-
eriors of the "curved triangles" inside e. This is a stan-
dard procedure in point set topology and we shall not go into
the details.) Since g is a subgraph of g^Δ, this completes
the induction, for h will automatically take g on to a
straight graph.

(VI) Within any simple closed polygon C with 3, 4 or
5 sides there is always a point P from which all the ver-
tices of C are visible. The proof of this is left as an
exercise, with the following hint for a pentagon. Divide
the interior of C into three triangles by two straight seg-
ments connecting vertices of C and lying entirely inside C.
(It can be proved, in fact, that the interior of any simple
closed polygon with n sides can be divided into n - 2
triangles by drawing n - 3 such segments, intersecting only
at vertices of the polygon.) Since there are three triangles
one of them must be "central" in the sense that it is joined

to each of the other two along an edge.   In the diagrams the
central triangle is shaded.

It is not difficult to find points within the central triangle
close to the vertex A from which all vertices of the penta-
gon are visible.

Note that pentagons are a critical case: there exist hex-
agons within which there is no point from which all six ver-
tices are visible.

This completes the sketch of a proof of Fáry's Theorem.

*APPENDIX ON KIRCHHOFF'S EQUATIONS

In this appendix we shall present a proof of the fact that
Kirchhoff's equations (see p. 18) determine the currents
uniquely for any circuit.   Consider the basic loops $\ell^1, \ldots, \ell^\mu$
described in 1.16.   The (K1) equation of the loop ℓ^i will
be written $L_i = b_i$ where L_i is the part involving the
(unknown) currents and b_i is the total e.m.f. in the loop.
As has already been noted in 1.21 (1) the (K1) equation of
any loop is a consequence of these μ equations; it was also
noted during the earlier discussion that all but one of the
(K2) equations are independent.   Writing $V_i = 0$ for the
equation of the vertex v^i we have the following, which est-
ablishes the required result.

1.35 <u>Proposition</u>   The $\alpha = \alpha_1$ equations

$$L_1 = b_1, \; \ldots, \; L_\mu = b_\mu, \quad V_2 = 0, \; \ldots, \; V_n = 0 \quad (n = \alpha_0)$$

are independent.

<u>Proof</u>   Suppose that

$$\lambda_1 L_1 + \dots + \lambda_\mu L_\mu \equiv \nu_2 V_2 + \dots + \nu_n V_n. \qquad (*)$$

We prove that $\lambda_1 = \dots = \lambda_\mu = 0$ which, in view of the inde-
pendence of $V_2, \dots, V_n$, implies the result.

We introduce some new notation at this point. It is not
strictly necessary to the argument, but allows a more succinct
presentation. The ideas of 'cochain' and 'coboundary' are of
great importance in algebraic topology, but we shall not have
occasion to make use of them apart from in this one case.

In order to allow the resistances in the wires to be real
numbers and not just integers, it is convenient to work with
"real chains", that is linear combinations of vertices or edges
in which the coefficients are real numbers. In (1.17) the λ_i
become elements of $\mathbb{R}$, the real numbers, and $C_0(G)$ and $C_1(G)$
become real vector spaces. Likewise in (1.19) the boundary
homomorphism ∂ becomes a linear map. We shall keep to the old
notation, but, just for this proof, work over $\mathbb{R}$ instead of $\mathbb{Z}$.

A <u>0-cochain</u> is a linear map $\gamma^0 : C_0(G) \to \mathbb{R}$ and a <u>1-cochain</u>
is a linear map $\gamma^1 : C_1(G) \to \mathbb{R}$. The value of a cochain γ on
a chain c is denoted $\langle \gamma, c \rangle \in \mathbb{R}$, and the i-cochains form a
real vector space $C^i(G)$ $(i = 0, 1)$ with addition and scalar
multiplication defined by

$$\langle \gamma + \gamma', c \rangle = \langle \gamma, c \rangle + \langle \gamma', c \rangle; \qquad \langle \lambda\gamma, c \rangle = \lambda\langle \gamma, c \rangle.$$

There is a <u>coboundary homomorphism</u>

$$\delta : C^0(G) \to C^1(G)$$

defined by

$$\langle \delta\gamma^0, c_1 \rangle = \langle \gamma^0, \partial c_1 \rangle.$$

Now writing x_j for the (unknown) current in the oriented
edge e^j and r_j for the resistance of this (unoriented) edge,

L_i will have the form

$$L_i \equiv \sum_{j=1}^{\alpha} \epsilon_{ij} r_j x_j \quad (\epsilon_{ij} = 1, \ 0 \ \text{ or } \ 1).$$

Taking the direction round the loop ℓ^i to be given by the
orientation of the edge e^i, it follows that $\epsilon_{ij} = 0$ if
$i \neq j$ and $j \leqslant \mu$, while $\epsilon_{ii} = +1$. This is because e^j does
not occur in any of the loops ℓ^i for $i \neq j \leqslant \mu$. Hence the
$\mu \times \alpha$ matrix $A = (\epsilon_{ij})$ has the $\mu \times \mu$ identity matrix as its
first μ columns. This will be of significance later.

With L_i we associate the 1-cochain, also called L_i,
which takes the edge e^j to $\epsilon_{ij} r_j \in \mathbb{R}$. In the same way we
can associate a 1-cochain with V_i, and it is easy to check that
$V_i = \delta v^i$, where v^i is the 0-cochain taking the value 1 on
the vertex v^i and 0 on every other vertex. Thus the equation
(∗) above says precisely that $\lambda_1 L_1 + \ldots + \lambda_\mu L_\mu$ is in the image
of δ.

It follows from the formula for δ that if we evaluate
$\lambda_1 L_1 + \ldots + \lambda_\mu L_\mu$ on any cycle the result will be zero. Let
us evaluate it on the cycles $z^1, \ldots, z^\mu$ given by the basic
loops $\ell^1, \ldots, \ell^\mu$. The result is μ linear equations in
$\lambda_1, \ldots, \lambda_\mu$ and a little calculation reveals that the coeffi-
cient matrix is $(<L_i, z^j>)$. With A as above and
$R = \text{diag}(r_1, \ldots, r_\alpha)$ this coefficient matrix is precisely
ARA^T (T for transpose). We shall show that this matrix is
nonsingular.

Now each resistance is strictly positive, so we can define
a nonsingular matrix $S = \text{diag}(\sqrt{r_1}, \ldots, \sqrt{r_\alpha})$. Clearly
$ARA^T = (AS)(AS)^T$. However A has rank μ since it contains
the $\mu \times \mu$ identity matrix. Hence AS has rank μ. It is
now a standard fact of linear algebra that, AS being real,
the product with its transpose is nonsingular. (One way to
prove this is to note that $(AS)(AS)^T$, being symmetric, is
congruent to a diagonal matrix. Compare MacLane and Birkhoff

(1967), p. 385.)   Thus $\lambda_1 = \ldots = \lambda_\mu = 0$, as required.

<div align="right">Q.E.D.</div>

It is not difficult to work this argument backwards and show that, if ARA^T is singular, then the Kirchhoff equations are dependent or contradictory.   (The backward argument uses the fact that a cochain which vanishes on every cycle is necessarily a coboundary.   This holds over $\mathbb{R}$, but not over $\mathbb{Z}$.)   Thus the nonsingularity of ARA^T is necessary and sufficient for the Kirchhoff equations to determine the currents uniquely.   It is a curious fact that this condition is much weaker than the condition "$r_i > 0$ for all i" which was used at the end of the proof.   For example when $u = 1$ it says that the sum of the resistances in the unique loop should be nonzero.

In Kirchhoff's original 1847 paper, he shows how to determine the currents (in the physical situation $r_i > 0$) and also shows that all the $(K1)$ equations depend on the μ equations for the basic loops.   Parts of this paper are readily available in translation on pages 133-5 of Biggs, Lloyd and Wilson (1976). See also Lefschetz (1975), Chapter VI.

Closed surfaces

A surface is usually thought of as something smooth and
rounded, like the surface of a sphere or a torus.   In order
to study surfaces systematically we shall, as described in the
Introduction, assume that they are divided up into triangles.[†]
It is therefore necessary at the outset to pinpoint the deci-
sive property which an object must possess in order to be
called a surface, and to interpret this as a property of the
triangulation.

The decisive property is this.   Take a point P on the
surface; then the points of the surface close to P make up
a little patch on the surface, this patch being essentially a
two-dimensional disk.   The disk may be bent or stretched,
but must not be torn (that is, "essentially" means, in tech-
nical language, "homeomorphic to").   In terms of triangu-
lations this means that the edges opposite P in the triangles

---

[†]   This is no restriction for compact surfaces (e.g. surfaces
which are closed subsets of a bounded region of euclidean
space).   Indeed with the more general concept of infinite
triangulation a wider class of surfaces can be triangulated.
These matters are dealt with in detail in Chapter I of Ahlfors
and Sario (1960); the triangulation theorem is on p. 105.
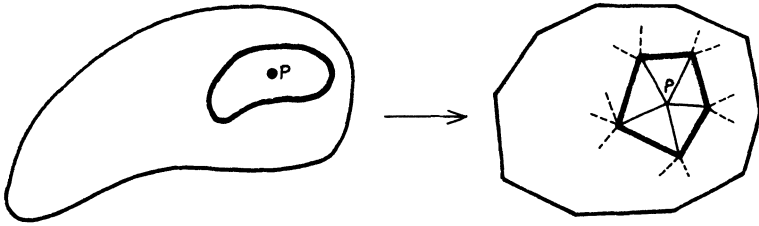
with P as a vertex form a simple closed polygon.   A defi-
nition along these lines is made precise in 2.1.   Note that
the definition excludes surfaces like the Möbius band and
cylinder which have boundary edges or rims; to emphasize that
our surfaces have no boundary we call them "closed surfaces".
The more general concept of "surface with boundary" is explained
and used in Chapter 9.

In the formal definition of closed surface (2.1) the
triangles will be ordinary flat triangles with straight edges.
This is simply because flat triangles are easier to define
and to work with than curvy ones.   The reader should not
feel uneasy because the smooth rounded surface has become a
somewhat angular and kinky object; he is quite at liberty to
continue drawing rounded surfaces - and indeed the author
will exercise that liberty too.   In Chapter 1 we saw that no
generality was lost by specifying that the edges of a graph
should be straight.   In a similar way every surface with
curved triangles can be realized (at any rate in $\mathbb{R}^5$) with
flat triangles, but the question of realization is not taken
up until 3.19.

The first objective of this chapter is to construct all
closed surfaces and to exhibit them in a convenient way by
means of a "polygonal representation".   From this represen-
tation it is possible, by a sequence of "scissors and glue"
operations to put every closed surface into exactly one of
a list of standard forms - this list appears in 2.8.   We can

then go on to derive a geometrical description of closed sur-
faces in terms of three basic ingredients - sphere, torus and
projective plane - and one mixing operation - connected sum.
In this way we can picture what all closed surfaces "look like".

Finally we show how to distinguish different closed sur-
faces from one another, and in particular how to prove that no
two closed surfaces in the list 2.8 are the same.

There is an intrinsic interest in trying to describe all
possible surfaces, but the longer term purpose so far as this
book is concerned is to provide plenty of good and geometrical
examples with which to illustrate later theorems.

CLOSED   SURFACES   AND   ORIENTABILITY

A graph in which there are no isolated vertices consists of a
collection of closed segments (edges) which satisfy an inter-
section condition, namely that two edges intersect, if at all,
in a common end-point (a vertex).   We now introduce a two-
dimensional building block and start to make two-dimensional
objects, again restricting the possibilities by imposing an
intersection condition.   The new building blocks are trian-
gles,which may be any shape or size.   Note that any triangle
comes equipped with three vertices and three straight edges.
The <u>intersection condition</u> we impose is this:

      Two triangles either

            (i)   are disjoint
or
            (ii) have one vertex in common
or
            (iii) have two vertices, and consequently the
entire edge joining them, in common.
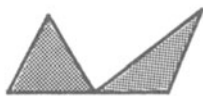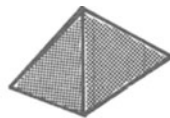
        (i)   (ii)   (iii)

Some configurations which are not allowed are pictured below:
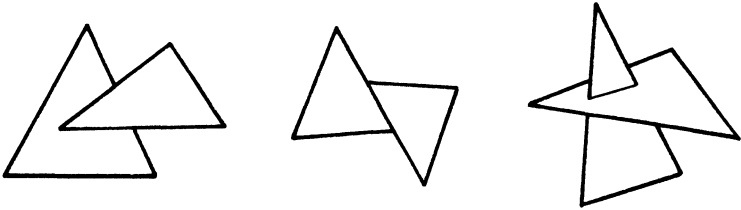
(In the terminology of Chapter 3, a collection of triangles
satisfying the intersection condition is said to form a pure
two-dimensional simplicial complex.)

As in the case of graphs, we have the concept of connect-
edness: a collection M of triangles satisfying the inter-
section condition is called <u>connected</u> if there is a path along
the edges of the triangles from any vertex to any other ver-
tex.   Indeed the set of edges and vertices of triangles in
M forms a graph M^1 and we have just defined "M is conn-
ected" to mean "M^1 is connected".

There is one other concept needed in order to define
closed surfaces.   Consider a vertex v of some triangle of
a collection satisfying the intersection condition.   The
edges opposite v in the triangles of M having v as a
vertex will form a graph, called the <u>link</u> of v.

2.1 <u>Definition</u>   A <u>closed surface</u> is a collection M of
triangles (in some euclidean space) such that
    (i) M satisfies the intersection condition (see above)
    (ii) M is connected
    (iii) for every vertex v of a triangle of M, the link
of v is a simple closed polygon.

2.2 <u>Remarks and Examples</u>
    (1) A hollow tetrahedron is a closed surface, the link
of any vertex being the three edges of the opposite face.
However two hollow tetrahedra with a common vertex do not
form a closed surface since the link of the common vertex is
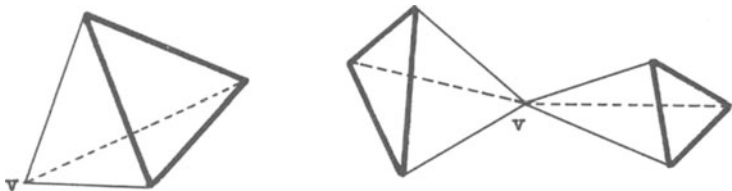
two simple closed polygons and not one.

2.2 (2) It is not difficult to show that any edge of a
closed surface M is on exactly two triangles of M. (Con-
sider the link of one of the end-points of the edge.)   Note
that two hollow tetrahedra with a vertex in common have this
latter property, but, as pointed out in (1), do not form a
closed surface.   Two hollow tetrahedra with an edge in common
do not satisfy even the weaker condition that every edge is
on exactly two triangles.

(3)   The diagram shows a schematic representation of a
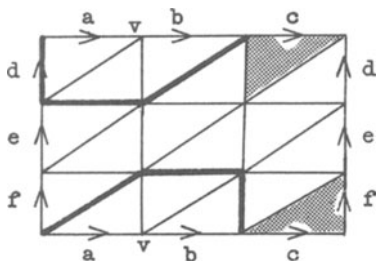
closed surface containing
18 triangles.   The pairs
of similarly marked edges
are to be "identified"
with arrows corresponding
so that, for example, the
two shaded triangles
actually have the edge c

in common.   The link of any vertex is then a simple closed
polygon - the link of v is drawn heavily in the diagram,
the two halves fitting together to form a simple closed hexa-
gon.   The closed surface so described is called a triangula-
tion of the <u>torus</u>(sometimes abbreviated to "torus") since,
informally, if we stick together the similarly marked edges
in $\mathbb{R}^3$, the result is a torus.

(These diagrams are not to scale!)
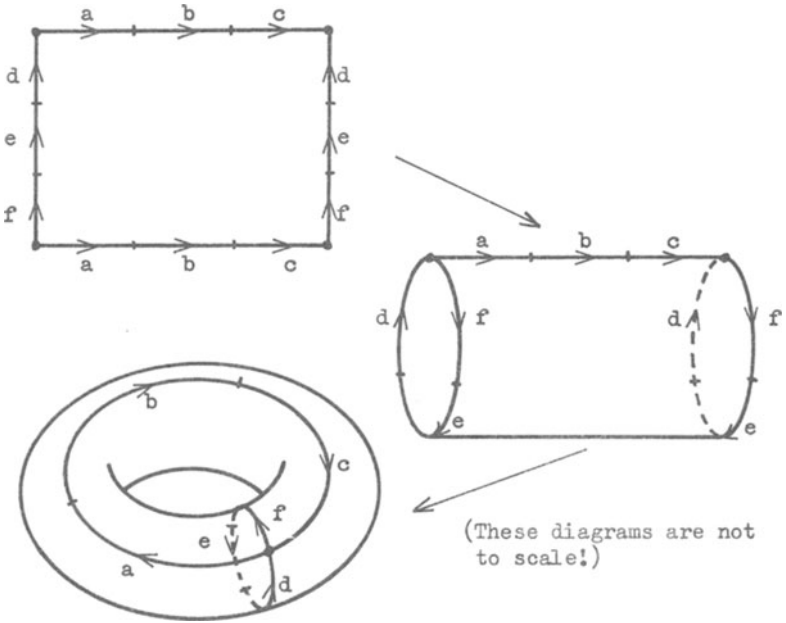
(4) The left-hand diagram below represents a closed sur-
face which is a triangulation of the "Klein bottle" and the
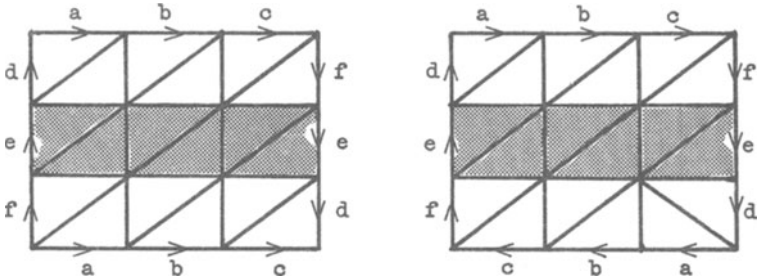
right-hand diagram represents a triangulation of the "projec-
tive plane".   (The shading in these diagrams is not relevant
until later.)   Neither of these surfaces can be constructed
in 3-dimensional space (see 9.18).   A picture of the Klein
bottle, in which there is a self-intersection along a circle,
appears on the next page.   For similar pictures of the pro-
jective plane, see p. 64 ; see also Hilbert and Cohn-Vossen

(1952), pp. 313-321.
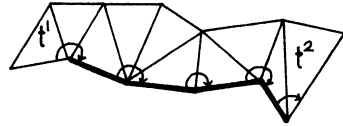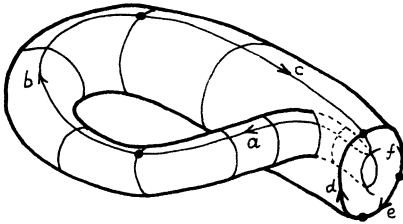
The arrows show the way in
which a sequence of tri-
angles is constructed (see
(5X)) using in turn the links
of vertices on the path.

2.2 (5X) Let t^1 and t^2 be two triangles of a closed sur-

face M.   Join some vertex of t^1 to some vertex of t^2 by

a simple path along the edges of M.   Using the fact that the

link of each vertex along this path is a simple closed poly-

gon it is possible to construct a sequence of triangles conn-

ecting t^1 and t^2, any two consecutive triangles in the

sequence having an edge in common.   Note that such a sequence

cannot always be constructed on the two tetrahedra with a

common vertex, but it can on the two tetrahedra with a common

edge.

(6) The definition of closed surface does not allow the
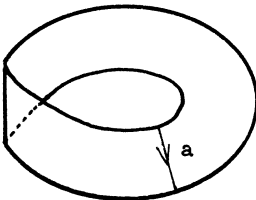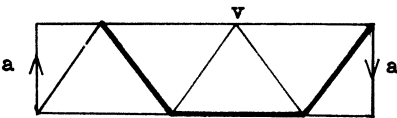
collection of triangles
on the left, since the
link of the vertex v
is not a simple closed
polygon (it is a poly-
gonal arc).   The dia-
gram shows a triangula-
tion of the **Möbius band**,
drawn underneath, which

is obtained by sticking together the ends of a rectangle after

a twist of 180°.   The Möbius band is a surface in a more

general sense which will be explained in 9.8.   Reversing one

of the arrows in the diagram we obtain a **cylinder**, in which

the ends of the rectangle are identified without a twist.

The major division amongst closed surfaces is into orient-
able and non-orientable closed surfaces.   Once again we shall
meet the concept of orientation more formally in the next
chapter; for the present we shall use the following informal
idea.   Suppose that on each triangle of a closed surface   M
we draw a circular arrow (called an orientation), as in the
diagrams.

In the left-hand diagram, the two triangles with a common
edge are said to be coherently oriented, while in the right-
hand diagram the two triangles are not coherently oriented.

2.3 Definition   A closed surface   M   is called (coherently)
orientable if all the triangles of   M   can be given an orient-
ation in such a way that two triangles with a common edge are
always oriented coherently.   Otherwise   M   is called non-
orientable.

2.4 Example (X) The torus in 2.2.(3) is orientable, but the
Klein bottle and projective plane in 2.2 (4) are non-orient-
able.   Note that the latter two closed surfaces both contain
Möbius bands (see 2.2 (6) and the shading in 2.2 (4)).   In
fact it will become clear that a closed surface is non-
orientable if and only if it contains a Möbius band.   In the
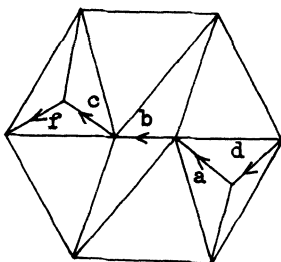
case of the Klein bottle,
removing the shaded
triangles in 2.2 (4)
leaves another Möbius band,
as can be seen by identi-
fying the two a's b's
and c's.   On the other
hand the diagram

illustrates removal of the shaded triangles from the projective
plane in 2.2 (4) (allowing a little distortion of the triangles).
This certainly does not contain any more Möbius bands.   Thus
the Klein bottle is "more non-orientable" than the projective
plane.   In fact it is just twice as non-orientable, as we
shall see following 2.9.

POLYGONAL REPRESENTATION   OF A CLOSED SURFACE

Let us start with a closed surface M, orient the edges
arbitrarily and then label the triangles t^1, ..., t^n and
their edges e^1, ..., e^m say.   (In fact $3n = 2m$; in part-
cular n is even.)   The information needed to construct M
comprises the n triangles and, for each triangle, the three
oriented edges which belong to it.   For example the tetra-
hedron on the left can be constructed from the four triangles
on the right.
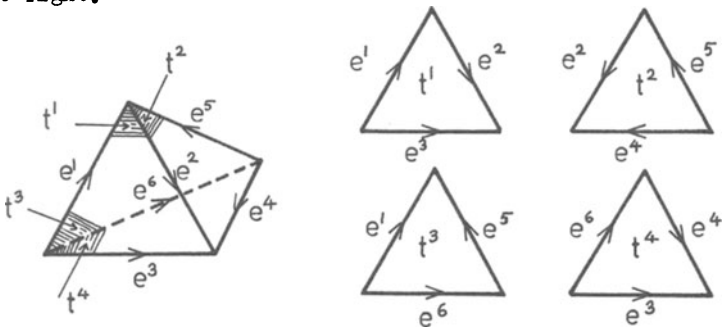
Our immediate object is to produce a more convenient plane
representation of M than the collection of n disjoint
triangles with labelled edges.   This is achieved, in the case
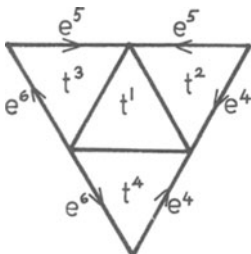
of the tetrahedron above,
by partially assembling
M in the plane: we could
stick t^2 to t^1 along
e^2, then t^3 to t^1 along
e^1, then t^4 to t^1

along e^3. The result is a diagram which completely describes
the way in which the triangles of M intersect one another.
Other examples of such a plane representation are in 2.2 (3),
(4) and (6) above.

The same procedure can be followed for any surface M.
Again we start with n disjoint triangles with labelled edges
and assemble them one at a time. Triangles which have been
assembled at a given stage will be called "used" and the re-
mainder "unused". It is essential to identify at each stage
an edge of an unused triangle to an edge of a used one; the
boundary of the plane region covered by the used triangles
will then always be a simple closed polygon. There is one
slight snag: we may find that the unused triangle which is
added at some stage overlaps the used triangles, as in the
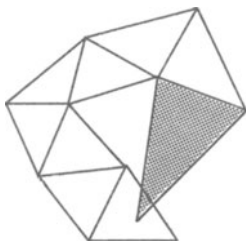
diagram, where the shaded
triangle was the last to be
added. To overcome this
we must abandon hope of
having the triangles in the
plane actually congruent to
the triangles of M.

Instead, as each triangle is used, we distort it (keeping the
edges straight) so that the triangles in the plane do not over-
lap - for example by making the boundary polygon given by the
used triangles always convex. The final representation of
M will still faithfully record all the intersections of
triangles of M (that is, it determines M "up to isomorphism"
in the language of Chapter 3). It remains to prove that,
until all triangles have been used, there is always an unused
triangle with an edge in common with the boundary polygon
given by the used triangles. There is certainly an unused
triangle with a vertex v in common with some used triangle,
and the construction ensures that all vertices are on the
boundary polygon. It is now easy to verify, using the fact

that the link of v in M is a simple closed polygon, that
there is an unused triangle with v and another vertex (and
hence an edge) in common with a used triangle. (Note that
the construction would break down at some stage if we started
with two tetrahedra with a common vertex.)

Continuing until all the triangles of M are used, we
obtain a <u>polygonal representation</u> of M as the region bounded
by a simple closed polygon. This region is broken up into n
triangles which correspond to the triangles of M and the
boundary polygon has n + 2 edges which occur in equally
labelled pairs. The identifications necessary to recover M
can be described by a <u>symbol</u>: start at any vertex and read
round the boundary polygon, writing an edge or its 'inverse'
according as its orientation goes with or against the direc-
tion of travel. A symbol for the tetrahedron above is

$$(e^5)^{-1}e^4(e^4)^{-1}(e^6)^{-1}e^6e^5 .$$

Similarly the torus in 2.2 (3), which could not arise directly
from this construction since it has vertices not on the boun-
dary polygon, has a symbol $abcd^{-1}e^{-1}f^{-1}c^{-1}b^{-1}a^{-1}fed$, while
the Klein bottle in 2.2 (4) has a symbol $abcfedc^{-1}b^{-1}a^{-1}fed$.
Note that a symbol is essentially cyclic: we could start read-
ing it off at a different place; also read round the other way.

The triangles of a polygonal representation of a closed
surface M can be oriented coherently (all clockwise, say,
but the sceptical reader can prove it by induction on the
number of triangles); whether this gives a coherent orientation
of M itself depends on the disposition of arrows round the
boundary polygon. If the symbol contains ... a ... a^{-1} ...
then the orientation will be coherent across the edges making
up a; if it contains ... b ... b ... then the orientation
will fail to be coherent across the edges making up b. In
fact, we have the following result.

<u>2.5</u> <u>Proposition</u>   Suppose that a closed surface M is rep-
resented by a symbol, as above.   Then M is orientable if
and only if, for every letter occurring in the symbol, its
inverse also occurs.

*Suppose we are given a simple closed polygon with an even
number of sides, the plane region bounded by the polygon being
divided into triangles in such a way that the sides of the
polygon are edges of triangles and two triangles intersect,
if at all, in a vertex or an edge.   Now mark the sides of the
polygon with letters and arrows, each letter being used exactly
twice as in the examples of 2.2 (3), (4).   The diagram can be
interpreted as a 'schema' for sticking triangles together, the
similarly marked edges being identified with arrows correspon-
ding.   When is it possible to realize this schema by a closed
surface?   The problem of realizing 'schemas' of this sort
will be treated in more detail in Chapter 3 (see in particular
3.20); for the present we shall be content with stating the
following result.   The identification of edges entails iden-
tification of certain vertices of the boundary polygon, and
with this in mind we shall label the vertices of the diagram,
those to be identified being given the same label, as in the
diagrams below.
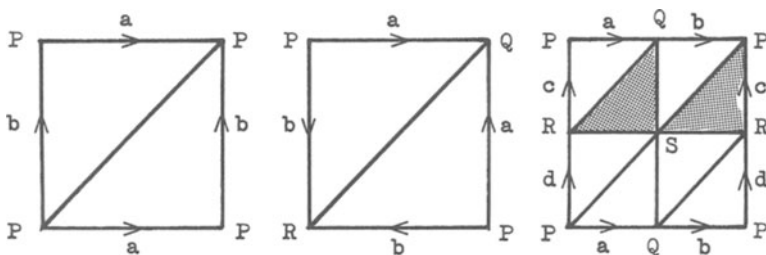
Note that none of these three determines a closed surface.
For the left-hand and centre diagrams the two triangles would
have all three edges in common, while for the right-hand dia-
gram the two shaded triangles would have two vertices but no

edge in common (see the intersection condition preceding 2.1).

2.6 Proposition   A schema of the kind described above is
realizable as a closed surface if and only if

> (i)   the end-points of any edge are labelled
>   differently,

and   (ii)   no two edges (unless they are to be identified)
               have end-points labelled the same,

and   (iii) no two triangles have vertices labelled the
               same.

The diagram is then called a polygonal representation of the
closed surface.

   The polygonal representations described before are a
special case of these; the identifications necessary to re-
cover M can again be described by a symbol and the orient-
ability criterion 2.5 still holds.

   Note that (i) - (iii) imply that the link of any vertex
will be a simple closed polygon.   The reader may care to
verify this, and also that the only case in which (i) and
(ii) hold but (iii) does not is the centre diagram on the
previous page.

   It is worth noting that any diagram of the kind described
before 2.6 can be subdivided into one which represents a closed
surface: subdivide every triangle into six by drawing the
medians, and then subdivide every small triangle into six
again by drawing the medians.   One such subdivision suffices
for the centre and right-hand diagrams, but both are needed
for the left-hand diagram.   The subdivision obtained by
drawing the medians of all triangles is called barycentric
subdivision and will be encountered again in Chapter 5.
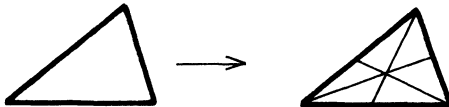
TRANSFORMATION OF CLOSED SURFACES TO STANDARD FORM

We shall describe a method, traditionally known as "scissors
and glue", of reducing any polygonal representation of
a surface to precisely one of a list of standard forms.
During the course of this it may be necessary - or at any
rate convenient - to subdivide M barycentrically a number
of times; this is in order to ensure that our scissors always
cut along edges of triangles of M. Thus the final symbol
will represent either M or some repeated barycentric sub-
division of M. Note that if M is a closed surface and
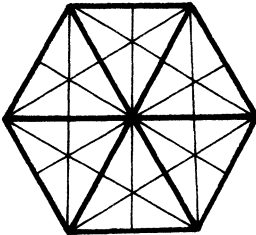
M' is obtained from M by
subdividing barycentrically
then M' is still a closed
surface. In the diagram, the
link of the central vertex
contains six edges of M and
twelve edges of M'.

    The reduction to standard form is in several steps,
numbered as the parts of 2.7. The reader who turns straight
to the answer (2.8) will not find himself much penalized in
the sequel, since we shall not often make use of the precise
nature of the steps used in the reduction to standard form;
the main exception is the proof of 2.14.

2.7 (1) First, a piece of notation. If a succession of
oriented edges occurs twice on the boundary polygon, either
in the same or opposite directions, then we shall always
denote the succession of edges by a single letter. For
example the diagram of 2.2 (3) has symbol $xyx^{-1}y^{-1}$ where
$x = abc$, $y = d^{-1}e^{-1}f^{-1}$. In what follows both x and x^{-1}
(for example) are called <u>letters</u>. Note that, as in group
theory, $(abc)^{-1} = c^{-1}b^{-1}a^{-1}$.

    (2) There is a further analogy with group theory in
that aa^{-1} can be "cancelled" unless this would result in

the symbol's disappearing entirely. For the part aa^{-1} of
a symbol containing at least four letters can be eliminated by
"closing up":
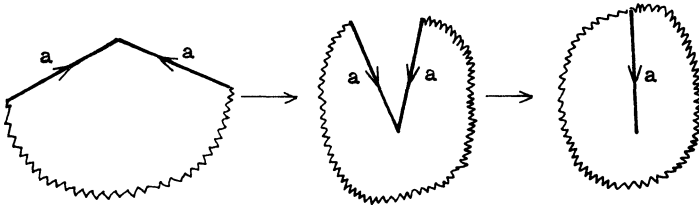
   It may look as though this procedure will inevitably
introduce curved edges into the diagram. In fact the result-
ing diagram will still be essentially a graph in the plane,
and by Fary's theorem (1.34) we can redraw it with straight
edges and with the same triads of edges spanning triangles
as before. In any case the diagram should really be treated
as a schema for showing intersections of triangles, and it
does not matter whether the edges are straight or curved.
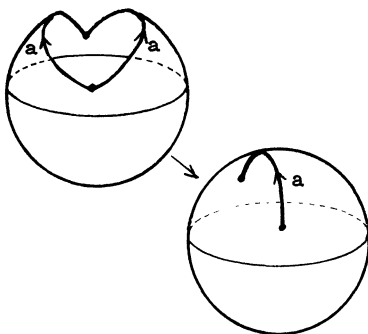
The closing
up operation is performed
whenever possible during
the succeeding stages.
As for the symbol aa^{-1}
itself, this is one of the
final list of standard forms,
and is to be regarded as a
sphere (see diagrams).

2.7 (3) A pair .....a.....a..... is called a pair of
the first kind and a pair .....a.....a^{-1}..... is called
a pair of the second kind. A pair of the first kind can be
brought together by the following rearrangement:
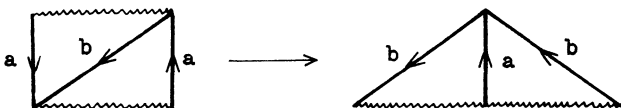
which replaces ..... a ..... a ..... by ..... bb ..... .
This process repeated produces

$$a_1 a_1 a_2 a_2 \cdots a_p a_p X$$

where X consists entirely of pairs of the second kind, or
is empty, and we say that p = 0 if there are no pairs of the
first kind at all.   The details of the repetition will be
left to the reader, guided by the following hints:

aab ... b ... can be turned into aacc ... ;

aab ... b^{-1} ... can be turned into aab ... b ... provided
there is a pair of the first kind besides aa.

The symbol aa represents the closed surface obtained
from a circular disk by identifying together each pair of
diametrically opposite points on the boundary circle.   This
cannot be done in $\mathbb{R}^3$ (compare 2.16 (5) and 9.18).   A sur-
face given by aa is known as a projective plane or sphere
with a cross-cap.   Any attempt to draw the surface in $\mathbb{R}^3$
will result in self-intersections, and two such attempts are
pictured below.   In each case the self-intersection is along
the segment XY.
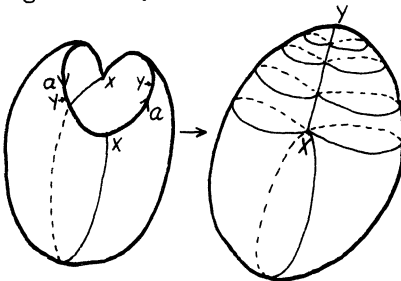
In both models, a is bent
double, to lie along XYX.
But the two halves, XY and
YX, are on different sheets
of the surface which inter-
sect accidentally.
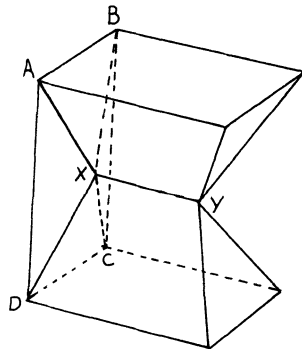
This surface consists of
three rectangles (top, bot-
tom and one end), four trap-
-ezia and four triangles (ABX and CDX are missing).

2.7 (4) If the symbol X in (3) above contains one pair
of the second kind then either the whole symbol is aa^{-1} or,
in view of (2), there is another pair of the second kind.
Choosing $\ldots a \ldots a^{-1} \ldots$ to be the pair which is closest
together there must then be two "interlocking" pairs
$\ldots a \ldots b \ldots a^{-1} \ldots b^{-1} \ldots$ (the b^{-1} could also come
before a) and this can be turned into $\ldots cdc^{-1}d^{-1} \ldots$ by
the following:
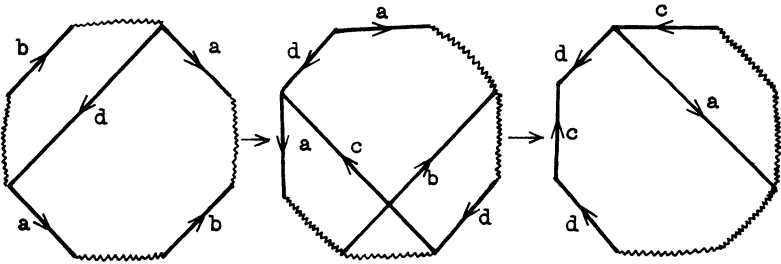
Note that no new pairs of the first kind can appear during
this transformation, since no section of the figure, after
cutting, is turned over before glueing back.   In this way
the symbol is turned into[†]

$$a_1 a_1 \ldots a_p a_p c_1 d_1 c_1^{-1} d_1^{-1} \ldots c_q d_q c_q^{-1} d_q^{-1} .$$

The symbol $cdc^{-1}d^{-1}$ represents a <u>torus</u> or <u>sphere with
handle</u> (compare 2.2 (3), where c appears as abc and d
as fed,   reading anticlockwise from bottom left).

(5) Finally, if there is at least one pair of the first
kind then every $cdc^{-1}d^{-1}$ (handle) can be turned into two
pairs of the first kind (cross-caps):

[†] The symbol given here has the property that the vertices
which occur at the ends of the letters in the symbol all
represent the same vertex of the closed surface.   This comes
about because at various points of the reduction pairs aa^{-1}
are eliminated (compare (2) above), and a vertex which was on
the boundary polygon becomes interior.
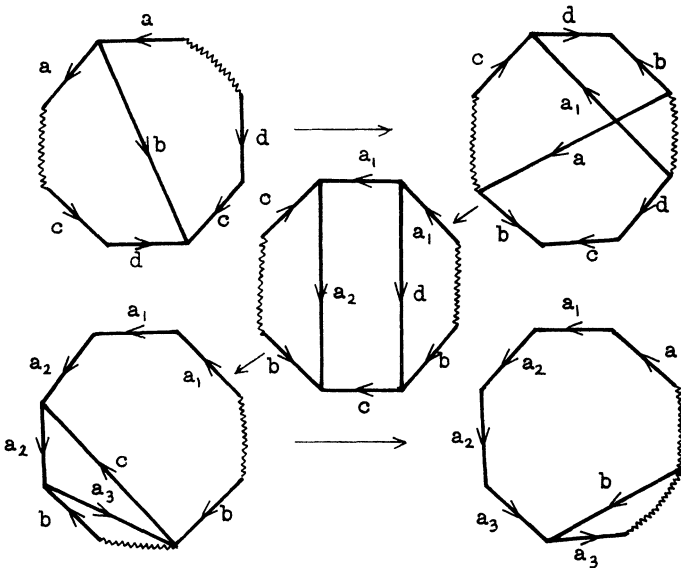
2.8 Theorem   Every closed surface can (possibly after bary-centric subdivisions) be represented by one of the following symbols:

    (1) aa^{-1} (sphere; orientable)

    (2) $a_1a_1a_2a_2 \cdots a_ka_k$, $k \geqslant 1$ (sphere with k cross-caps; non-orientable)

    (3) $a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1} \cdots a_hb_ha_h^{-1}b_h^{-1}$, $h \geqslant 1$
                             (sphere with h handles; orientable)

The number k or h is called the genus of the closed sur-face; the genus of a sphere is 0.

It is proved in 2.14 that any given closed surface can be reduced to only one of the forms given in 2.8.

When a closed surface is represented by one of the above symbols we say that it is in standard form. Two closed sur-faces M and N which reduce to the same standard form are called equivalent, and we write M ≈ N. Of course two equi-valent closed surfaces may be broken up into triangles

in quite different ways.   What M ≈ N says is that, forgett-
ing about the ways in which M and N are broken into tri-
angles, we can obtain each of them from a plane polygonal
region with the same number of sides by making the same iden-
tifications, given by the symbols.   Thus M and N "look
the same" - are in fact homeomorphic (see Chapter 5).

Any closed surface with standard form aa^{-1} is called
a (triangulation of the) <u>sphere</u>; likewise closed surfaces
with standard forms aa and $aba^{-1}b^{-1}$ are called (triangu-
lations of) <u>projective plane</u> and <u>torus</u> respectively.

There is a simple geometrical interpretation for the
juxtaposition of symbols, such as the succession of torus
or 'handle' symbols $aba^{-1}b^{-1}$ which occurs in 2.8 (3).   Let
M and N be closed surfaces given by symbols X and Y
respectively (not necessarily in standard form).   We shall
construct a new surface with symbol XY obtained by justa-
posing X and Y.   The new surface is far from being uniquely
determined by M and N, but nevertheless its standard form
is uniquely determined by those of M and N.   This follows
from 2.17 below.

<u>2.9 Construction</u>   Choose a triangle of M meeting the
boundary polygon of the plane representation with symbol X
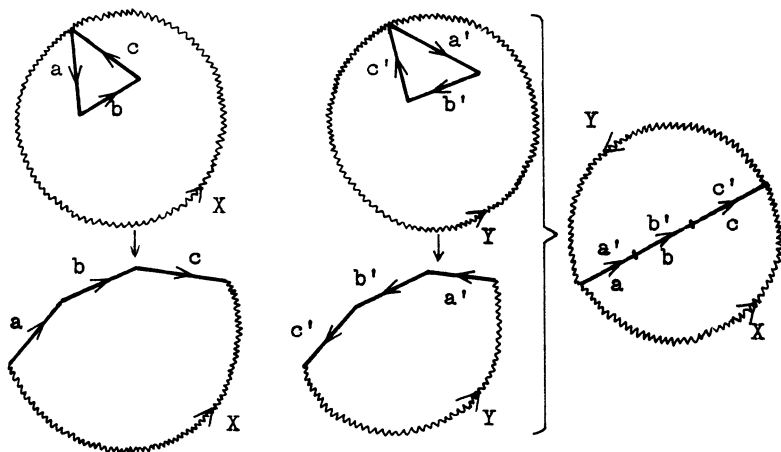
precisely in a vertex of M at the start of X (if necessary
barycentrically subdividing M), and similarly with N.
Then remove these two triangles and glue together as in the
diagram.

    It is easy to check that the result is a closed surface,
orientable if and only if M and N are both orientable.
It is called a connected sum of M and N.   Informally we
cut a hole in each surface and glue them together along the
boundary rims.   Thus for a connected sum of two tori there
appear to be two choices:
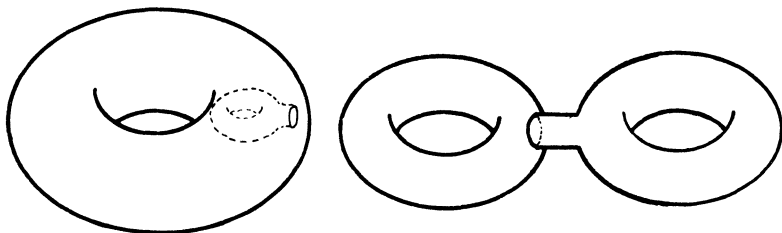
In fact, as already claimed, these have the same standard
form (indeed the little torus inside can be pulled through
the hole to turn the left-hand picture into the right-hand
one!).   The surface is called a "double torus", or "pretzel".
Thus the surface of 2.8 (3) is an "n-fold torus" which looks
like n tori joined by n - 1 tubes, and 2.8 (2) is an
"n-fold projective plane".   In particular the Klein bottle,
which has the usual symbol (compare 2.2 (4)) $xyx^{-1}y$, has
standard form $a_1a_1a_2a_2$, a "double projective plane".

    Let M and N be closed surfaces.   The uniqueness
property mentioned before 2.9 (and proved in 2.17) amounts
to saying that any two connected sums of M and N are
equivalent: we can speak up to equivalence of the connected
sum of M and N.

2.10 Notation   We write M # N   for any connected sum
M and N.

Thus any two choices of M # N (for given M and N)
are equivalent. Indeed, more is true: if M and N are
each allowed to vary within an equivalence class then all the
possible connected sums are still equivalent. This too is
proved in 2.17.

The notation can be extended to several "summands";
again M_1 # ... # M_n is well-defined up to equivalence.
We write nM for M # ... # M with n summands. Writing
T for any closed surface with standard form $aba^{-1}b^{-1}$, i.e.
for any torus, P for any projective plane and S for any
sphere, the statement of 2.8 can be rephrased as follows:
every closed surface is equivalent to exactly one of S, kP
or hT.

It can be shown that, for any closed surface M, we have
M # S $\approx$ M $\approx$ S # M. This is most easily done using Euler
characteristics (described below), but the reader can pro-
bably convince himself by drawing pictures. The same goes
for the statement P # 2P $\approx$ P # T (this is essentially shown
in 2.7 (5)); note that 2P $\not\approx$ T so that "cancellation" is
not possible. Since 2P can be replaced by T provided
this leaves at least one P, we can also rephrase 2.8 by
saying that every closed surface is equivalent to exactly
one of

$$S, \quad hT, \quad P \text{ \# } \frac{k-1}{2}T \text{ (k odd)}, \quad K \text{ \# } \frac{k-2}{2}T \text{ (k even , } \geqslant 2)$$

where K = 2P is a Klein bottle.

EULER CHARACTERISTICS

We shall prove that a given closed surface M cannot be
reduced to two different standard forms 2.8. In addition to
this we shall present a method by which the standard form of
any surface given by a polygonal representation (compare 2.6)
can be found very quickly and without actually performing

the scissors and glue operations of 2.7.

2.11 Definition   Let K be a collection of triangles
satisfying the intersection condition given before 2.1.
Writing $\alpha_0(K)$, $\alpha_1(K)$ and $\alpha_2(K)$ for, respectively, the
number of vertices, edges and triangles of K, the Euler
characteristic of K is defined to be

$$\chi(K) \;=\; \alpha_0(K) - \alpha_1(K) + \alpha_2(K).$$

2.12 Proposition   The Euler characteristic of K is un-
affected by barycentric subdivision.

Proof   Write the α s after one subdivision as α';
then

$$\alpha_0' = \alpha_0 + \alpha_1 + \alpha_2, \qquad \alpha_1' = 2\alpha_1 + 6\alpha_2, \qquad \alpha_2' = 6\alpha_2 .$$

Hence

$$\alpha_0' - \alpha_1' + \alpha_2' \;=\; \alpha_0 - \alpha_1 + \alpha_2 . \qquad\qquad \text{Q.E.D.}$$

Our immediate aim is to calculate the Euler characteristic
of the closed surface corresponding to any given symbol.

2.13 Theorem   Suppose that a closed surface M has a symbol
containing n letters (n is necessarily even), and is rep-
resented by a plane polygonal region bounded by a simple
closed polygon with n + r sides (each side of the polygon
corresponding to an edge of M - recall that a letter in a
symbol can stand for a succession of edges). The n vertices
of the polygon at beginnings or ends of letters will represent
vertices of M; let the number of distinct vertices of M
amongst these be m .
    Then
$$\chi(M) \;=\; m - \tfrac{1}{2}n + 1 .$$

Proof   The set of triangles assembled in the plane
to give the plane representation of M is itself a simplicial
complex of dimension two, which we shall call D (for disk).

The Euler characteristic of D is easily evaluated by Euler's formula (1.28): the $\alpha_0(D)$ vertices and $\alpha_1(D)$ edges of triangles making up D form a graph which divides the plane into $\alpha_2(D) + 1$ regions (one region being unbounded). Hence by Euler's formula $\chi(D) = 1$. (An alternative proof of this result can be given by the method of "collapsing"; see 3.33.)

The number of vertices of D on the boundary polygon is $n + r$ while the number of vertices of M on the boundary polygon is $m + \frac{1}{2}r$. Since vertices inside the polygon count equally in D and in M, it follows that $\alpha_0(M) - \alpha_0(D) = m - n - \frac{1}{2}r$. Similarly $\alpha_1(M) - \alpha_1(D) = -\frac{1}{2}n - \frac{1}{2}r$ and of course $\alpha_2(M) - \alpha_2(D) = 0$. Thus $\chi(M) - \chi(D) = m - \frac{1}{2}n$ and this, together with $\chi(D) = 1$, gives the result.   Q.E.D.

2.14 Corollary   The Euler characteristic of any closed surface with standard form

$$2.8 \ (1) \quad \text{is} \quad 2 \quad ;$$
$$2.8 \ (2) \quad \text{is} \quad 2 - k \ ;$$
$$2.8 \ (3) \quad \text{is} \quad 2 - 2h \ .$$

The standard form of any closed surface M is unique, and is determined by $\chi(M)$ and the orientability or otherwise of M.

Proof   For the first statement, it is enough to calculate the Euler characteristics of the closed surfaces given by 2.8, since the scissors and glue operations of 2.7 do not affect the Euler characteristic. It is easy to check that in case (1), $m = 2$, while in cases (2) and (3), $m = 1$. The result now follows from the theorem.

For the second statement, note that the scissors and glue of 2.7 does not affect orientability either. Since no two standard forms agree in point of orientability and Euler characteristic, the result follows.   Q.E.D.

2.15 Remark   When calculating the number m of 2.13 the
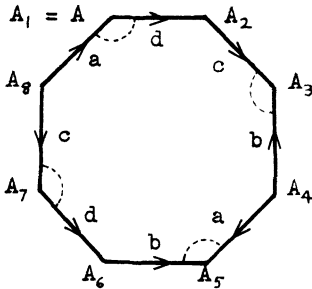
procedure of the following example is used. Choose any vertex (at beginning or end of a letter) and call it A, say $A_1 = A$.

A_1 is at the end of a; so is A_5. Hence in M, $A_5 = A$.

A_5 is at the end of b; so is A_3. Hence in M, $A_3 = A$.

A_3 is at the end of c; so is A_7. Hence in M, $A_7 = A$.

A_7 is at the beginning of d; so is A_1: we are back where we started.

The important thing is that no other vertex can be A, since in M the link of A is one simple closed polygon, which is indicated by the dotted lines in the diagram.   How start at another vertex, say $A_2 = B$ and deduce $A_4 = A_6 = A_8 = B$. Hence m = 2, so that $\chi(M) = -1$ and since M is non-orientable (it contains ...b...b...), M has standard form 3P, i.e. $a_1a_1a_2a_2a_3a_3$. Thus M is a sphere with three cross-caps.

2.16 Examples   Check the following, (1) - (4).

(1X) abcbca is a Klein bottle, i.e. has standard form 2P.

(2X) $abca^{-1}cb^{-1}$ has standard form 3P.

(3X) $abcdefe^{-1}db^{-1}afc$ has standard form 6P.

(4X) $ae^{-1}a^{-1}bdb^{-1}ced^{-1}c^{-1}$ has standard form 2T.

*(5X) Since any orientable closed surface is equivalent to a connected sum of tori (see p. 69) it is clear that a representative of each equivalence class can be realized in $\mathbb{R}^3$. In Chapter 9 (see 9.18) we shall sketch a proof of the theorem

that no nonorientable closed surface can be realized in $\mathbb{R}^3$.
However they can all be realized in $\mathbb{R}^4$, as follows. Since
every non-orientable closed surface is equivalent to a conn-
ected sum of tori and either a projective plane P or a
Klein bottle K (see p. 69) it is enough to check that P
and K can be realized in $\mathbb{R}^4$.

For P, start with a Möbius band, triangulated as in
the diagram. It is
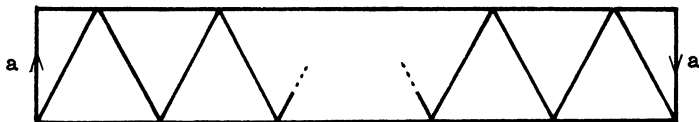
clear that with enough triangles this can be realized in $\mathbb{R}^3$
(can the reader discover the minimum number needed?). Now
take coordinates (x, y, z, t) in $\mathbb{R}^4$ and let the Möbius
band be realized in the space t = 0. We add triangles as
follows: for each pair v, w of vertices consecutive along
the rim of the band, add the triangle with vertices v, w
and (0, 0, 0, 1). The result is a closed surface (for
example the link of (0, 0, 0, 1) is the rim of the band,
which is a simple closed polygon) and the Euler characteristic
is easily verified to be 1. Hence it is a projective plane.
We have "added a cone" to the boundary rim of the band.
Cones will be used on several occasions later in the book;
see the index for references.

The Klein bottle K can be constructed in $\mathbb{R}^4$ in a
similar way, starting with a triangulation of the "double
Möbius band" - two bands with parts of their rims identified.
The boundary rim of the double band is still a single closed
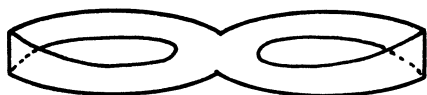
curve, and adding a cone
gives a non-orientable
closed surface of Euler
characteristic 0, hence
a Klein bottle. An

alternative approach is to start with the picture of a Klein
bottle on p. 55, where there is self-intersection along a
circle.   Using the fourth spatial dimension one of the inter-
secting tubes can be "moved aside" to miss the other.   (This
is analogous to taking two intersecting lines in a horizontal
plane and moving one of them slightly upwards to miss the other.)

<u>2.16</u> *(6X) Let M be a closed surface in $\mathbb{R}^3$, not passing
through the origin 0, and with the property that every half-
line through 0 meets M precisely once.   Then M is a triangu-
lation of the sphere, i.e. has standard form aa^{-1}.   One way to
see this is as follows.   For each vertex v of M let v_1 be
the (unique) intersection of the half-line 0v with a fixed
sphere centred at 0.   Thus $v \to v_1$ projects the vertices of M
on to the sphere.   Similarly the edges of M project to arcs of
great circles on the sphere and triangles to spherical triangles.
Since half-lines meet M exactly once we obtain a collection of
(spherical) triangles covering the sphere, with intersections
exactly as in M.   To recover the standard form aa^{-1} we can re-
move one spherical triangle t and flatten separately t, and the
rest of the sphere, on to a plane (e.g. by stereographic projec-
tion).   This gives a plane representation of M by t and a tri-
angular region which is itself broken up into triangles.   Re-

identifying one edge of t
with one boundary edge of the
triangular region gives the
standard form aa^{-1}.   It is
possible to avoid curved tri-
angles in this argument,
though of course in the plane
they can in any case be
straightened by Fáry's theorem
(1.34).
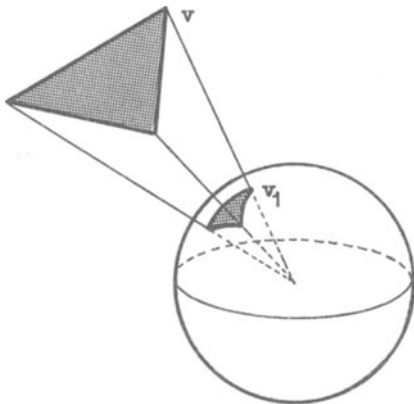
*(7X) Suppose that the closed surface M_1 in (6) has the
property that it is "symmetric about 0", i.e. that it

is left unchanged by reflexion through the point O.   Then,
at any rate provided M_1 has enough triangles, we can form
a closed surface M_2 by identifying opposite triangles of
M_1.   Since $\chi(M_1) = 2$ it is clear that $\chi(M_2) = 1$, so
that M_2 must be a triangulation of the projective plane.
To see this informally, think of identifying diametrically
opposite points all over the sphere.   We can throw away the
southern hemisphere provided we leave the equator and iden-
tify opposite points of it.   Flattening out the hemisphere
we get a disk, or a square, with opposite points of its
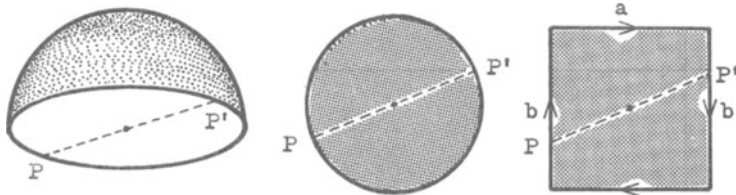
Each diametrically opposite pair P,P'
is identified to a point

boundary identified.   The latter is one of the usual pictures
of a projective plane (2.2 (4)).

<u>2.16</u> *(8XX) Let M be a closed surface.   We form a <u>sub-</u>
<u>division</u> M_1 of M, by dividing up every triangle of M
into smaller triangles.   We require the subdivisions of any
two adjacent triangles to be consistent on their common edge
as in the diagram.   A special subdivision is the barycentric

subdivision used in
this chapter. (See **also**
Chapter 5.)   Let
$\overline{\alpha}_0(M_1)$, $\overline{\alpha}_1(M_1)$ denote
the numbers of vertices
and edges of M_1 lying
in edges of M.   Show that $\overline{\alpha}_0(M_1) - \overline{\alpha}_1(M_1) = \alpha_0(M) - \alpha_1(M)$.
Note that, if D denotes the part of M_1 lying within a
single triangle of M, then $\chi(D) = 1$;   compare the proof

of 2.13.   Deduce that $X(M) = X(M_1)$: this says that the
Euler characteristic is invariant under subdivision.

Now consider the effect of making a connected sum (2.9)
on Euler characteristics.   Since barycentric subdivision does
not affect Euler characteristics (2.12) it is immediate from
the definition that $X(M_1 \# M_2) = X(M_1) + X(M_2) - 2$ for
any connected sum $M_1 \# M_2$ of two closed surfaces M_1 and
M_2.   It is also true that $M_1 \# M_2$ is orientable if and
only if M_1 and M_2 are both orientable: thus orientability
and Euler characteristic for $M_1 \# M_2$ are determined by
M_1 and M_2.   By 2.14 we have:

<u>2.17</u> <u>Proposition</u> (1)   Any two connected sums of M_1 and
M_2 are equivalent;

(2)   If $M_1 \approx N_1$ and $M_2 \approx N_2$ then

$$M_1 \# M_2 \approx N_1 \# N_2 .$$

*MINIMAL   TRIANGULATIONS

Among all triangulations of, say, a torus - that is, among all
closed surfaces M with standard form $aba^{-1}b^{-1}$ - there will
be some which have the fewest possible triangles.   In this
section we investigate such "minimal triangulations" briefly
and relate them to a famous problem of graph theory.

Let us fix attention on a particular standard form, which
has Euler characteristic X given by 2.14.   Any closed sur-
face M with this standard form has the following properties:

<u>2.18</u> $\begin{cases} \alpha_1(M) \leqslant \frac{1}{2}\alpha_0(M)(\alpha_0(M) - 1) & \text{(see the note following 1.1)} \\ 3\alpha_2(M) = 2\alpha_1(M) & \text{(2.2 (2))} \\ \alpha_0(M) - \alpha_1(M) + \alpha_2(M) = X & \text{(2.11)} \end{cases}$

These give $\alpha_0^2 - 7\alpha_0 + 6X \geqslant 0$.   With the exceptions
noted below, this is equivalent to

2.19 $\alpha_0 \geqslant \frac{1}{2}(7 + \sqrt{(49 - 24\chi)})$.

The exceptions are as follows:

> If $\chi = 1$ there is a solution $(\alpha_0, \alpha_1, \alpha_2) = (1, 0, 0)$
>
> If $\chi = 2$ there are solutions $(\alpha_0, \alpha_1, \alpha_2) = (2, 0, 0)$
>
> and $(3, 3, 2)$.

However none of these exceptions corresponds to a genuine surface. The only one that makes sense at all is $(3, 3, 2)$ for the sphere, and this is two distinct triangles with the same three vertices and the same three edges, a configuration which cannot be realized with flat triangles and which violates the intersection condition given before 2.1.

Note that $\alpha_0(M)$ is minimal if and only if $\alpha_2(M)$ is minimal, and also that equality occurs in 2.19 if and only if it occurs in the first relation of 2.18.

2.20 <u>Examples</u> (1) Consider the standard form aa^{-1}: the sphere. Then $\chi = 2$ and from 2.19, $\alpha_0(M) \geqslant 4$. In fact there is a triangulation, namely the hollow tetrahedron, with exactly four vertices.

(2) Consider the standard form aa: the projective plane. Then $\chi = 1$ and $\alpha_0(M) \geqslant 6$. A triangulation with six vertices is drawn in 4.17 (5).

(3) Consider the standard form $aba^{-1}b^{-1}$: the torus. Then $\chi = 0$ and $\alpha_0(M) \geqslant 7$. A triangulation with seven vertices is drawn in 3.22 (1). There is another standard form with $\chi = 0$, namely $a_1 a_1 a_2 a_2$: the Klein bottle. Is there a triangulation with seven vertices? In fact not: the smallest possible number is eight, and it is not hard to find triangulations with eight vertices. This is a curiously exceptional example, for in every other case where the right hand side of 2.19 is an integer there does exist a triangulation with exactly that number of vertices. Not surprisingly, this is far from easy to prove; we shall shortly sketch a

proof which depends on a very difficult theorem in graph
theory, and also on a result to be proved in Chapter 9.

2.20 (4X) Show from 2.19 that

$$\alpha_0(M) \;\geqslant\; [\tfrac{1}{2}(9 + \sqrt{(48 - 24X)})]$$

where $[x]$ means the greatest integer $\leqslant x$.

(5) In the non-orientable case, Ringel (1955) has deter-
mined precisely the number of vertices in a minimal triangu-
lation. It is $[\sqrt{(12 - 6X)} + \tfrac{5}{2}] + 2$ for $X \neq 0, -1$; and
8 for $X = 0$ (Klein Bottle); 9 for $X = -1$.

Given a graph G we can enquire, for example, "Is it
possible to embed G in the torus?" This means "Is there
some triangulation M of a torus with the property that the
graph consisting of the set of vertices and edges of M con-
tains a subgraph isomorphic to G?" Thus for example let
G_n denote the complete graph with n vertices (every pair
of vertices being joined by an edge); then for $n \leqslant 7$, G_n
can be embedded in a torus, for the triangulation drawn in
3.22 (1) has the property that its set of vertices and edges
is a G_7. Indeed, whenever we have a triangulation for which
equality holds in 2.19 it will satisfy $\alpha_1 = \tfrac{1}{2}\alpha_0(\alpha_0 - 1)$,
and therefore the set of vertices and edges of the triangu-
lation will be a complete graph G_{α_0}.

There is an alternative, less formal, way in which we
can think of embedding a graph in a closed surface. We can
think of the surface in its usual rounded shape and the graph
drawn with curved edges in the surface. For example the
left-hand diagram on the following page shows G_5 in the
torus. In fact, the two ideas are equivalent: a graph can
be embedded in this way in a curved surface if and only if
it can be embedded in the way previously defined as a straight
graph contained in some triangulation. Roughly speaking,
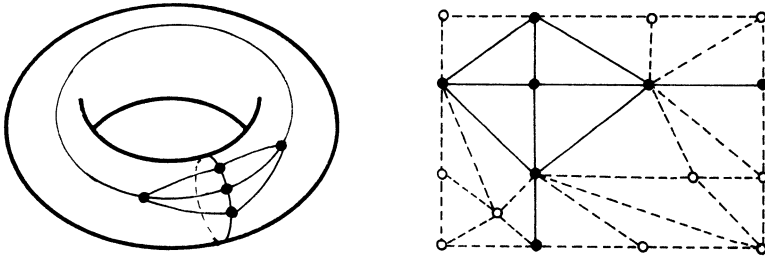
given a graph in a curved surface, we add extra vertices and
edges away from the graph until the surface is divided up
into curved triangles, and then regard this as a schema for
the construction of a genuine (straight) triangulation (see
2.6).   This is analogous to the construction of a triangular
graph in the plane (1.34 (I)) except for the introduction of
extra vertices.   One way of carrying this out for G_5 in
the torus is illustrated above.

In the spirit of the preceding paragraph, let us see how
to embed an arbitrary graph in a closed surface of sufficiently
high genus (2.8).   Given a graph G in $\mathbb{R}^3$ (see 1.4),
project G from some point in $\mathbb{R}^3$ on to a plane, choosing
the point of projection O in such a way that the ray through
O and a vertex of G does not meet G again, while the
ray through O and any other point of G meets G again at
most once.   For example, starting with G_5 we might obtain
the diagram in 1.5 (1) in the plane.   At each crossing point
(not the projection of a vertex) we stick a small handle on
the plane:
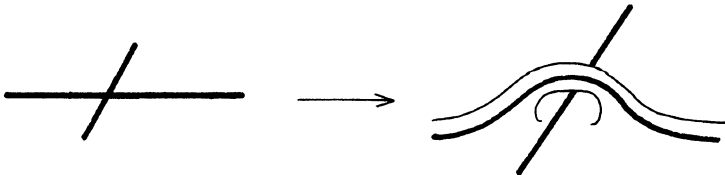
In this way all the crossings are eliminated, and enclosing
the projected graph in a large circle which is then capped
by a hemisphere we obtain a graph isomorphic to G realized

in a sphere with handles (or with crosscaps, if a crosscap is
added somewhere).   Applying this procedure to 1.5 (1) we can
embed G_5 in an orientable surface of genus 5 - far from
the best result, for it also embeds in the torus, which has
genus 1.

We are interested, in what follows, not in whether a
graph embeds in a particular surface, but in whether it embeds
in some surface in a given equivalence class - the notion of
equivalence being that given after 2.8.   Thus unless other-
wise stated we regard all surfaces as variable within a def-
inite equivalence class.   This most accurately reflects the
informal idea of drawing the graph on a curved surface.

It is not an easy matter to determine the surface of
smallest genus in which a given graph can be embedded.   A
remarkable theorem, due essentially to G. Ringel and J. W. T.
Youngs, but in which several other people had a hand, deter-
mines just this for the complete graph G_n.   (See Ringel &
Youngs (1968), Youngs (1967), Youngs (1970) and, especially,
Ringel (1974).)   The following statement of the theorem is
directed at the opposite problem: given a surface (variable
of course within an equivalence class) what is the largest
complete graph which can be embedded in it?

<u>2.21</u> <u>Theorem</u> (Ringel-Youngs)   The largest complete graph
which will embed in a closed surface, other than the Klein
bottle, of Euler characteristic χ, is the one with

$$H(\chi) = [\tfrac{1}{2}(7 + \sqrt{(49 - 24\chi)})]$$

vertices.   In the case of the Klein bottle, the largest com-
plete graph is G_6, not G_7.

Note the anomalous position of the Klein bottle, and also
the similarity between 2.21 and 2.19.   I claim, in fact, the
following.

2.22 __Theorem__[†]   Suppose that $\frac{1}{2}(7 + \sqrt{(49 - 24\chi)})$ is an
integer, λ say, where χ is an integer $\leqslant 2$.   Then there
is a closed surface with λ vertices and Euler characteristic
χ.   Indeed if $\chi < 0$ is even then there are two such, one
orientable and one non-orientable.   If $\chi = 0$ then there is
only one, which is a torus.   The set of vertices and edges
of the surface is a complete graph G_λ.

Note that here the claim is that G_λ embeds in a parti-
cular closed surface, not merely one of a given equivalence
class.   Note also that the triangulation is minimal by 2.19.
In fact the theorem above asserts that the lower bound of
2.19 is attained provided $\frac{1}{2}(7 + \sqrt{(49 - 24\chi)})$ is an integer.

Here is a sketch of a proof of 2.22.   It follows from
the Ringel-Youngs theorem that, with $\chi \neq 0$, G_λ embeds in
any closed surface with characteristic χ but not in any with
characteristic $> \chi$.   The case $\chi = 0$ requires special
attention.   From this it can be shown that whenever G_λ is
embedded in a closed surface of characteristic χ, the graph
will divide it up into triangles; that is, G_λ forms the
vertices and edges of triangulation in which the triangles
are the triangular regions into which G_λ divides the surface.
A proof of this is sketched in the last paragraph but one of
9.28 (7), which depends on several results of Chapter 9.
Using the fact that G_λ is a graph there is no difficulty in
checking that the triangles satisfy the intersection condition
given before 2.1.   (The only possible exception is G_3 embed-
ded in the sphere to give two (curved) triangles meeting in all
three edges, but fortunately $\lambda = 3$ is not a possible value.

[†] My aim in this brief exposition is to display some connex-
ions between the Ringel-Youngs theorem, minimal triangulations
and (shortly) colouring problems.   I hasten to point out that
existing proofs of the Ringel-Youngs theorem proceed on a case
by case basis, and that one case is close in content to the
present theorem.

This anomalous case turned up in the discussion of 2.19 as well.)

2.23 Examples

(1) Take $\chi = -10$ in 2.22. Then $\lambda = 12$; hence there are surfaces equivalent to 6T and to 12P (both of which have characteristic -10) having 12 vertices. In each case the set of vertices and edges is G_{12}: this result says that starting with G_{12} we can build up both an orientable and a non-orientable closed surface by filling in suitable triangles. (The G_{12} had better be in a high dimensional space to avoid accidental intersections of triangles. In fact a theorem in Chapter 3 (3.19) says that five dimensions will do.)

(2X) For which values of χ is $\frac{1}{2}(7 + \sqrt{(49 - 24\chi)})$ an integer? (Answer: $\chi = (7n - n^2)/6$ where $n \equiv 0, 1, 3$ or 4 (mod 6).)

(3X) Use 2.21 to find explicitly the Euler characteristics of the orientable and non-orientable surfaces of lowest genus into which it is possible to embed G_n.

(4XX) It is a curious fact that any two minimal triangulations of the sphere are isomorphic, and the same goes for the projective plane and the torus. Here two closed surfaces in the sense of 2.1 are called isomorphic if there is a bijection f from the set of vertices of one to the set of vertices of the other which has the property that both f and f^{-1} carry triads of vertices spanning triangles into triads of vertices spanning triangles.

To see this in the case of the projective plane we try to construct a closed surface with 6 vertices, every pair of vertices being joined by an edge. This implies that the link of each vertex contains all the others. Starting with an arbitrary vertex, label it
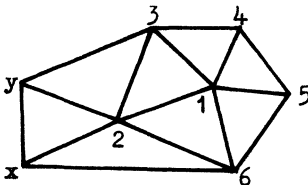
1 and the vertices in its link 2, ..., 6.   Two vertices x
and y in the link of 2 are unnamed, but x cannot be 5,
since 5 can only occur once in the link of 6;   hence
x = 4, y = 5.   In this way the closed surface can be deter-
mined uniquely.

The result for the torus amounts to this: the vertices
of the tesselation of the plane by equilateral triangles can
be labelled with 1, ..., 7 in such a way that every number
is always surrounded by the remaining six, and this labelling
can be done in an essentially unique way.

The isomorphism property between minimal triangulations
does not appear to extend further than this.   G. Edmunds
wrote a computer programme to search for minimal triangula-
tions of the orientable surface of genus 6 (this has $X = -10$,
which gives $\lambda = 12$ in 2.22), and he found at least fifty
non-isomorphic ones!

It is worth pointing out the connexion between 2.21 and
the _colouring problem_ for graphs in a closed surface.   A
graph is n-_colourable_ if each vertex van be assigned one of
n colours in such a way that two vertices at the ends of an
edge are always coloured differently.   Thus for example a
graph with n vertices is certainly n-colourable, and a com-
plete graph G_n is not m-colourable for any m < n.   Now
consider all graphs which can be embedded in a given closed
surface (or, as usual, in any equivalent one).   What is the
smallest n for which all these graphs are n-colourable?
This number is called the _chromatic number_ of the closed
surface.   Clearly, from 2.21 the chromatic number of any
closed surface of Euler characteristic X is at least $H(X)$
(except for the Klein bottle, when it is at least 6.)   In
fact Heawood (1890) proved that the chromatic number is
$\leqslant H(X)$, for any closed surface except the sphere.   (See
9.28 (8).)   Heawood actually believed, incorrectly, that he

had proved the chromatic number equal to $H(X)$, and the
equality became known as <u>Heawood's conjecture</u>.   Franklin (1934)
found a counterexample, when he showed that the chromatic num-
ber of the Klein bottle is 6, not $H(0) = 7.$ The Ringel-Youngs
theorem showed, astonishingly, that Franklin's counterexample
was the only one.   Thus the chromatic number is

> $H(X)$ for any closed surface except Klein bottle or sphere

>   6   for the Klein bottle

> $\geqslant 4$ for the sphere.

The sphere is not covered by Heawood's theorem: in fact in
his 1890 paper Heawood showed only that the chromatic number of
a sphere is $\leqslant 5.$ It had been conjectured around 1852 that the
chromatic number was 4 and indeed another function of Heawood's
paper was to expose a fallacious proof of this conjecture.   For
over 100 years the notorious "four colour conjecture" remained
unsolved.   A proof was found in June 1976 and published the
following year, but even then the proof used high speed computers
and it is a matter for speculation whether a purely conceptual
proof that can be checked by a human being will ever be found.
At any rate the <u>four colour theorem</u> asserts that the chromatic
number of the sphere is 4.

For an account of the proof by the original authors see
Appel and Haken (1977); see also Saaty and Kaimer (1978).
For the five colour theorem see e.g. Berge (1962), p. 213
or Wilson (1972), p. 85.   For the history of the four colour
conjecture see Biggs, Lloyd and Wilson (1976).   For the connex-
ion between colouring graphs and colouring geographical maps see
the latter book or Wilson (1972), §19.

Simplicial complexes

The reader may have noticed that although surfaces are con-
structed out of triangles I did not define the word "triangle"
in Chapter 2.   Doubtless the reader has come across triangles
before.   Nevertheless a precise definition in terms suited to
our purpose is given below, where triangles appear under the
alias of "2-simplexes".   The precise definition makes it
clear that "triangle" is a good way to continue the sequence
"point, segment, ..." (which becomes "0-simplex, 1-simplex,
.."") and suggests that the fourth term should be "solid
tetrahedron": this is our 3-dimensional building block.
We shall need building blocks in higher dimensions too, but
most examples in the text will be, as hitherto, 2-dimensional.
Algebraic concepts closely analogous to the cycles of Chapter 1
will be introduced in Chapter 4.

It is possible, by judicious skipping in this chapter,
to avoid contact with simplexes of dimension higher than two.
All initial definitions are stated for the case of two dimen-
sions as well as in general; from 3.9 just assume $n \leqslant 2$.

Everything in the chapter takes place in a real vector
space $\mathbb{R}^N$.   Elements of $\mathbb{R}^N$ are called <u>points</u> or <u>vectors.</u>
Unless otherwise stated the only restriction on N is that
it should be large enough for the discussion to make sense;

thus if we speak of four non-coplanar points then obviously N must be at least 3.

SIMPLEXES

It is clear that two points v^0 and v^1 are the end-points of a segment if and only if they are distinct. The segment then consists of all points $(1 - \lambda)v^0 + \lambda v^1$ where $0 \leqslant \lambda \leqslant 1$, i.e. it is the set of points

$$\{\lambda_0 v^0 + \lambda_1 v^1 \;:\; \lambda_0 + \lambda_1 = 1, \quad \lambda_0 \geqslant 0, \quad \lambda_1 \geqslant 0\} \;.$$

Similarly three points v^0, v^1, v^2 are the vertices of a triangle if and only if they are not collinear, i.e. if and only if $v^1 - v^0$ and $v^2 - v^0$ are linearly independent. In the left-hand diagram $P = \lambda v^0 + (1 - \lambda)v^1$ $(0 \leqslant \lambda \leqslant 1)$
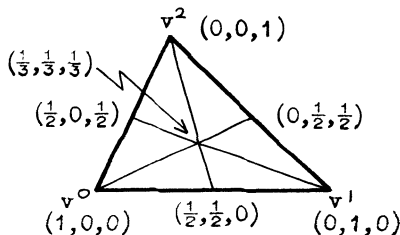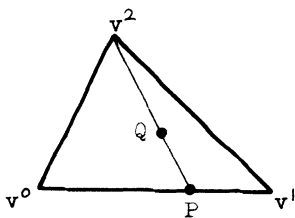
and $Q = \mu(\lambda v^0 + (1 - \lambda)v^1) + (1 - \mu)v^2$ $(0 \leqslant \mu \leqslant 1)$. Hence

$$Q = \lambda_0 v^0 + \lambda_1 v^1 + \lambda_2 v^2$$

where $\lambda_0 = \lambda\mu$, $\lambda_1 = (1 - \lambda)\mu$ and $\lambda_2 = 1 - \mu$. Notice that the λ_i are $\geqslant 0$ and have sum 1. Conversely given non-negative λ_0, λ_1, λ_2 with sum 1 we can solve for λ and μ; hence the triangle is the set of points

$$\{\lambda_0 v^0 + \lambda_1 v^1 + \lambda_2 v^2 \;:\; \lambda_0 + \lambda_1 + \lambda_2 = 1,$$

$$\lambda_i \geqslant 0 \text{ for } i = 1, 2, 3\}.$$

The advantage of the λ_i over λ and μ is that any given point Q has uniquely determined λ_i (compare 3.5) whereas v^2 has $\mu = 0$ but arbitrary λ. The numbers λ_0, λ_1, λ_2 are called the <u>barycentric coordinates</u> of Q. The barycentric coordinates of some points in the triangle are given in the right-hand diagram on the previous page.

It can be shown that if we drop the condition of non-negativeness on the λ_i the above set of points is enlarged to include all those in the plane of the triangle. Points with all $\lambda_i > 0$ are inside the triangle, points with two $\lambda_i > 0$ and one zero are on an edge, and points with two λ_i equal to zero are at a vertex.

<u>3.1 Definition</u> Let v^0, ..., v^n be $n + 1$ vectors in $\mathbb{R}^N$, $n \geqslant 1$. They are called <u>affinely independent</u> (<u>a-independent</u>) if $v^1 - v^0$, ..., $v^n - v^0$ are linearly independent. By convention if $n = 0$ then the vector v^0 is always a-independent (even if it is zero).

<u>3.2 Examples</u> (1) v^0 and v^1 are a-independent if and only if $v^0 \neq v^1$; v^0, v^1, v^2 are a-independent if and only if they are not collinear.

(2X) Let v^0, ..., v^n be in $\mathbb{R}^N$. They are a-independent if and only if, for all real numbers λ_0, ..., λ_n, we have

$$\left(\begin{array}{l} \lambda_0 v^0 + \ldots + \lambda_n v^n = 0 \\ \lambda_0 + \ldots + \lambda_n = 0 \end{array} \right) \Rightarrow \lambda_0 = \ldots = \lambda_n = 0 .$$

(Hint: $\lambda_0 v^0 + \ldots + \lambda_n v^n = \lambda_1(v^1 - v^0) + \ldots + \lambda_n(v^n - v^0)$ when $\Sigma\lambda_i = 0$.) This makes it clear that the definition of a-independence does not in reality depend on the ordering of the v^i: we could equally well define it by the linear

independence of $v^0 - v^1$, $v^2 - v^1$, $v^3 - v^1$, ..., $v^n - v^1$.

Also, adding a constant to all the v^i does not affect their a-independence or otherwise. (In contrast, even if v and w are linearly independent, $v - v$ and $w - v$ are not.) Thus a-independence does not depend on the choice of origin.

3.2 *(3) Let U be a subspace of $\mathbb{R}^N$ and let $b \in \mathbb{R}^N$. the set of vectors

$$U + b = \{u + b : u \in U\}$$

is called an <u>affine subspace</u> of $\mathbb{R}^N$ of <u>dimension</u> equal to that of U. We say that $U + b$ is <u>parallel</u> to U (it certainly has no points in common with U provided $b \notin U$, and if $b \in U$ then $U + b = U$).
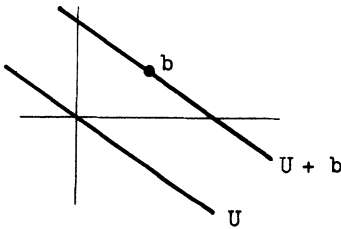
For example, let $N = 2$ and let U be a line through the origin. Then $U + b$ is the line parallel to U, through b.

Notice that, given $U + b$, we can recover U:

$$U = \{v - w : v, w \in U + b\}.$$

On the other hand b is not uniquely determined by $U + b$ since in fact

$$U + b = U + b' \iff b - b' \in U.$$

(For the example above, these say that we can recover U by moving $U + b$ parallel to itself until it passes through the origin; and that replacing b by any $b' \in U + b$ gives the same line parallel to U.)

If v^0, ..., v^n are a-independent then $v^1 - v^0$, ..., $v^n - v^0$, being linearly independent, are a basis for a sub-

space U of $\mathbb{R}^N$ of dimension n. The affine subspace
$U + v^0$ then has dimension n and contains all of $v^0, \ldots, v^n$.
Notice that any affine subspace $U' + b$ containing
$v^0, \ldots, v^n$ must have dimension $\geqslant n$, for U' will contain
the n linearly independent vectors $v^1 - v^0, \ldots, v^n - v^0$.
Hence the smallest affine subspace containing $v^0, \ldots, v^n$
has dimension n.

Definition 3.1 can be reformulated: $v^0, \ldots, v^n$ are
a-independent if and only if the smallest affine subspace of
$\mathbb{R}^N$ containing them has dimension n.

<u>3.3 Definition</u>   Let $v^0, \ldots, v^n$ be vectors in $\mathbb{R}^N$.   A
vector v is said to be <u>affinely dependent</u> (<u>a-dependent</u>)
on them if there exist real numbers $\lambda_0, \ldots, \lambda_n$ such that

$$\lambda_0 + \ldots + \lambda_n = 1 \quad \text{and} \quad v = \lambda_0 v^0 + \ldots + \lambda_n v^n.$$

<u>3.4 Examples</u>   (1X) If $v^0, \ldots, v^n$ are a-independent
then none of them is a-dependent on the rest. If v is
a-dependent on $v^0, \ldots, v^n$ then $v^0, \ldots, v^n, v$ are not
a-independent.

*(2) The set of vectors a-dependent on $v^0, \ldots, v^n$ is
the smallest affine subspace of $\mathbb{R}^N$ containing $v^0, \ldots, v^n$
and is called the <u>affine span</u> of $v^0, \ldots, v^n$.

<u>3.5 Proposition</u>   Let $v^0, \ldots, v^n$ be a-independent and
let v be a-dependent on them. Then there exist unique
real numbers $\lambda_0, \ldots, \lambda_n$ such that $\Sigma \lambda_i = 1$ and $v = \Sigma \lambda_i v^i$.
The λ's are called the <u>barycentric coordinates</u> of v with
respect to $v^0, \ldots, v^n$.

<u>Proof</u>   The λ's exist by 3.3.   To prove uniqueness
suppose that $\Sigma \lambda_i v^i = \Sigma \mu_i v^i$ where $\Sigma \lambda_i = \Sigma \mu_i = 1$ (all
sums run from 0 to n).   Then $\Sigma \nu_i v^i = 0$, where
$\nu_i = \lambda_i - \mu_i$ so that $\Sigma \nu_i = 0$.   Hence by 3.2(2) we have
$\nu^0 = \ldots = \nu^n = 0$, i.e. $\lambda_i = \mu_i$ for $i = 0, \ldots, n$.   Q.E.D.

By analogy with the discussion preceding 3.1 we can now define the general concept of simplex.

3.6 Definition   Let $v^0, \ldots, v^n$ be a-independent. The (closed) <u>simplex</u> with vertices $v^0, \ldots, v^n$ is the set of points a-dependent on $v^0, \ldots, v^n$ and with every barycentric coordinate $\geqslant 0$. The simplex is also said to be <u>spanned</u> by $v^0, \ldots, v^n$. The points with all barycentric coordinates > 0 are said to be <u>interior</u> to the simplex, and the set of interior points is sometimes called the <u>open simplex</u> with vertices $v^0, \ldots, v^n$. The <u>boundary</u> of the simplex consists of those points which are not interior, i.e. which have at least one barycentric coordinate $= 0$.

3.7 Notation   In this book, the unqualified word "simplex" will mean "closed simplex". The simplex with a-independent vertices $v^0, \ldots, v^n$ will be denoted by $(v^0 \ldots v^n)$ and often represented by a symbol like s_n or just s. The integer n is called the <u>dimension</u> of $(v^0 \ldots v^n)$ (compare 3.2 (3)). A simplex of dimension n is called an <u>n-simplex.</u> The boundary of s_n will be denoted by $\dot{s}_n$, and the interior by $\overset{o}{s}_n$.

3.8 Examples   (1) For any vector v^0 the 0-simplex (v^0) is just the set $\{v^0\}$. We often denote (v^0) by merely v^0. The open 0-simplex with vertex v^0 is also $\{v^0\}$. The boundary of (v^0) is empty.

For any two distinct vectors v^0, v^1 the 1-simplex $(v^0 v^1)$ is the closed <u>segment</u> with end-points v^0, v^1. The boundary of $(v^0 v^1)$ is $\{v^0, v^1\}$, and the interior is the open segment with end-points v^0, v^1.

For any three non-collinear vectors v^0, v^1, v^2 the 2-simplex $(v^0 v^1 v^2)$ is the <u>triangle</u> with vertices v^0, v^1, v^2. The boundary of $(v^0 v^1 v^2)$ is the set of points on the edges of the triangle, i.e. $(v^0 v^1) \cup (v^1 v^2) \cup (v^2 v^0)$, and the interior is the rest of $(v^0 v^1 v^2)$.

0-simplex   1-simplex   2-simplex
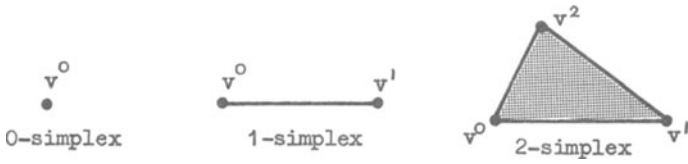
For any four non-coplanar vectors v^0, v^1, v^2, v^3 the 3-simplex $(v^0 v^1 v^2 v^3)$ is the <u>solid tetrahedron</u> with vertices v^0, v^1, v^2, v^3. Its boundary consists of four triangles.

3.8 *(2) A closed simplex is a closed subset of $\mathbb{R}^N$ in the usual sense that it contains all its limit points. This can be seen by means of barycentric coordinates: the ith barycentric coordinate of each point of a sequence of points in the simplex will be a real number $\geqslant 0$, so that the ith barycentric coordinate of the limit point of a convergent sequence will also be $\geqslant 0$. The interior of a closed simplex is the corresponding open simplex, which is not an open subset of $\mathbb{R}^N$ unless $n = N$.

(3X) A simplex determines its vertices - that is, if $(v^0 \ldots v^n) = (w^0 \ldots w^m)$ then $m = n$ and the v's are merely a permutation of the w's. To prove this we must find a characterization of the vertices of a simplex purely in terms of the set of points in the simplex. (This is perhaps a slightly unfamiliar situation - in Euclidean geometry a triangle is defined in terms of its sides, and the sides intersect in pairs at the vertices, which are therefore determined. Here, on the other hand, a triangle is defined by the set of points inside it and on its boundary.) Such a characterization, which is easy to establish using barycentric coordinates, is this: a point $v \in (v^0 \ldots v^n)$ is a vertex (i.e. is one of the points $v^0, \ldots, v^n$) if and only if it is not the mid-point of any segment joining two distinct points of $(v^0 \ldots v^n)$.

3.8 (4X) A simplex is <u>convex,</u> that is if v and w are
points of the simplex then every point of the (straight) seg-
ment joining v and w belongs to the simplex.   Indeed the
simplex is the smallest convex set containing its vertices,
which is expressed by saying that it is the <u>convex hull</u> of its
vertices.

(5X) Let $w \in (v^0 \ldots v^n)$ and let $v \neq w$ be an inter-
ior point of $(v^0 \ldots v^n)$.   Show that the segment from v to
w , produced beyond w if necessary, meets the boundary of
the simplex in precisely one point.   (This amounts to showing
that there is precisely one point of the simplex of the form
$(1 - \alpha)v + \alpha w$ with $\alpha \geqslant 1$ and having at least one barycentric
coordinate zero.)

(6X) Let v be an interior point of the simplex
$(v^0 \ldots v^n)$.   Show that $v^0, \ldots, v^{n-1}$, v are a-independent.

3.9 <u>Definition</u>   Let $s_n = (v^0 \ldots v^n)$ be a simplex.
A <u>face</u> of s_n is a simplex whose vertices form a (nonempty)
subset of $\{v^0, \ldots, v^n\}$.   If the subset is proper (i.e. not
the whole of $\{v^0, \ldots, v^n\}$) we say that the face is a <u>proper</u>
<u>face</u>.   If s_p is a face of s_n we write $s_p < s_n$ or
$s_n > s_p$; note that $s_n < s_n$ for any simplex s_n.

3.10 <u>Remarks and Examples</u>   (1) The boundary of s_n (3.6)
is the union of the proper faces of s_n.

(2) A face of s_n consists of those points of s_n for
which the barycentric coordinates attached to vertices not in
the face are zero.   It follows at once that two faces of s_n
either are disjoint or else intersect in a face of s_n.

(3) Strictly speaking a vertex v^i of s_n is not a
face of s_n.   What is a face is the 0-simplex $(v^i) = \{v^i\}$.
It does little harm, however, to fail to distinguish between
v^i and (v^i).

3.10 (4X) An n-simplex has $\binom{n+1}{p+1}$ faces of dimension p
and $2^{n+1} - 1$ faces altogether.

(5) The 0- and 1-dimensional faces of an n-simplex form
a complete graph on n + 1 vertices.

ORDERED SIMPLEXES AND ORIENTED SIMPLEXES

A fairly detailed treatment of orientation is given in an
appendix to this chapter, which is intended to motivate
Definition 3.11 below and relate it to the concept of orien-
tation in vector spaces.   It is not necessary to read the
appendix in order to follow this section.   Orientations of
2-simplexes have already been introduced informally in con-
nexion with closed surfaces (see 2.3).   The algebraic theory
developed from Chapter 4 onwards depends strongly on the con-
cept of orientation, which is used to define a "boundary"
homomorphism analogous to that introduced in 1.19 for oriented
graphs.   (Compare 1.21 (3))

According to definition 3.6, permutation of the vertices
of a simplex leaves the simplex unchanged.   If on the other
hand we insist on a definite ordering of the vertices the
result is called an <u>ordered simplex</u>.   Thus associated with
any n-simplex there are (n + 1)! distinct ordered simplexes.
In practice we do not wish to distinguish between all of these.

Consider for example a 2-simplex $(v^0v^1v^2)$.   The six
ordered simplexes associated with it are those in which the
ordering is as follows:

$$(v^0v^1v^2) \qquad (v^0v^2v^1)$$
$$(v^1v^2v^0) \qquad (v^2v^1v^0)$$
$$(v^2v^0v^1) \qquad (v^1v^0v^2)$$
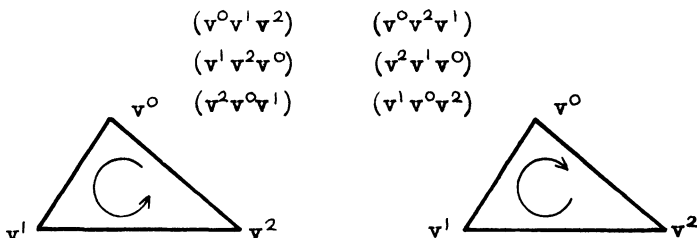
The three even permutations of v^0, v^1, v^2 all order the vertices "anticlockwise" (taking this as the direction for v^0, v^1, v^2 itself) while the three odd permutations order them "clockwise". This suggests the following general definition.

3.11 <u>Definition</u>   Let $s_n = (v^0 \ldots v^n)$ be an n-simplex. An <u>orientation</u> for s_n is a collection of orderings for the vertices consisting of a particular ordering and all even permutations of it. An <u>oriented n-simplex</u> σ_n is an n-simplex s_n together with an orientation for s_n.

3.12 <u>Notation</u>   We write $\sigma_n = (v^0 \ldots v^n)$ to mean that σ_n is the oriented n-simplex with vertices $v^0, \ldots, v^n$ and orientation given by the ordering displayed and all even permutations of it. It will always be clear from the context whether $(v^0 \ldots v^n)$ means a simplex or an oriented simplex. When $n > 0$ we write $-\sigma_n$ for the oriented simplex consisting of the same n-simplex and the collection of other possible orderings of the vertices. Thus by definition $-(-\sigma_n) = \sigma_n$.

3.13 <u>Remarks</u>   (1) A 0-simplex has a unique orientation; an n-simplex for $n \geqslant 1$ has two orientations.

(2) A 1-simplex has two orderings and two orientations: ordering and orientation are the same in this case. They are represented diagramatically by an arrow from the first to the second vertex (compare p.14). An orientation of a triangle is indicated by a circular arrow as in the above diagrams.

(3) For any permutation π of $\{0, \ldots, n\}$ the oriented simplexes $\sigma_n = (v^0 \ldots v^n)$ and $\tau_n = (v^{\pi(0)} \ldots v^{\pi(n)})$ satisfy $\sigma_n = \tau_n$ if π is even, and $\sigma_n = -\tau_n$ if π is odd.

SIMPLICIAL   COMPLEXES

Simplicial complexes are the most general geometrical objects
we study in this book.   They include as special cases the
graphs of Chapter 1 and the closed surfaces of Chapter 2, and
they provide a good framework for developing an algebraic
theory analogous to the cycles of Chapter 1.   As before we
impose "intersection conditions" (compare 1.3, 2.1);   now that
we have the ideas of simplex and face at our disposal it is
possible to state these very succinctly.

3.14 Definitions   A simplicial complex is a finite set K
of simplexes in $\mathbb{R}^N$ with the following two properties:

    (1) if $s \in K$ and $t < s$ then $t \in K$

    (2) intersection condition: if $s \in K$ and $t \in K$ then
$s \cap t$ is either empty or else a face both of s and of t.

The dimension of K is the largest dimension of any simplex
in K.   The underlying space of K, denoted $|K|$, is the
set of points in $\mathbb{R}^N$ which belong to at least one simplex
of K.   That is, $|K|$ is the union of the simplexes of K.
Occasionally, if Σ is merely a set of simplexes we use $|\Sigma|$
to denote the union of the simplexes in Σ.

    Notice that if $s \in K$ then all the faces of s are also
elements of K.   It may seem odd that we should explicitly
include these faces in K since they are, as subsets of $\mathbb{R}^N$,
contained in s: the faces of s do not contribute anything
extra to the underlying space.   The reason is to be found in
the next chapter, where algebraic apparatus is introduced
which requires the presence of all these faces.   It is worth
emphasizing at this point that a simplicial complex is a
finite set of simplexes in $\mathbb{R}^N$ while the portion of $\mathbb{R}^N$
which these simplexes together occupy (i.e. the underlying
space) is generally an infinite set of points in $\mathbb{R}^N$.   (The
underlying space $|K|$ is finite if and only if K has

dimension 0.)

When drawing simplicial complexes it is sometimes help-
ful to indicate vertices by a heavy dot and to shade in 2-
simplexes which are part of the complex. We shall occasion-
ally make use of this in what follows. But for the majority
of the diagrams we just use the convention that straight seg-
ments are 1-simplexes, intersections of segments (and no
other points) are 0-simplexes, and all triangular regions in
the diagram are in fact spanned by 2-simplexes. From now on
the reader should assume this latter convention unless there
is indication (usually by selective shading) to the contrary.

3.15 Examples   (1) A graph is a simplicial complex of
dimension 1. This agrees with 1.3, although admittedly no
explicit distinction is made there between the finite set of
vertices and edges (the simplicial complex) and the, usually
infinite, underlying space.

(2) The empty collection of simplexes is a simplicial
complex whose underlying space is empty. This simplicial
complex does not have a dimension, unless we make a special
convention to cover it. A popular choice is -1.

(3) A closed surface M as defined in 2.1 consists of
triangles (i.e. 2-simplexes) satisfying three conditions one
of which is an intersection condition. If we let M^+ be the
set of triangles, edges and vertices in M then M^+ is a
simplicial complex of dimension 2: thus a closed surface is
a simplicial complex provided we explicitly mention the ver-
tices and edges as well as the triangles. In the future we
shall assume this has been done. A closed surface is then a
special kind of simplicial complex called "pure": a simplicial
complex K of dimension n is called pure if every simplex
of K is a face of some n-simplex of K. An example of a
non-pure simplicial complex is $K = \{(v^0v^1), (v^0), (v^1), (v^2)\}$

where v^0, v^1, v^2 are distinct.   Here dim $K = 1$ but (v^2)
is not a face of any 1-simplex of K.

<u>3.15</u> (4X) The three diagrams below do not represent simplicial
complexes (Why?), but below them are diagrams for genuine sim-
plicial complexes with, in each case, the same underlying
space.   The triples in brackets are the numbers of 0, 1 and
2-simplexes respectively.

(5,4,0)   (6,9,4)   (7,9,3)
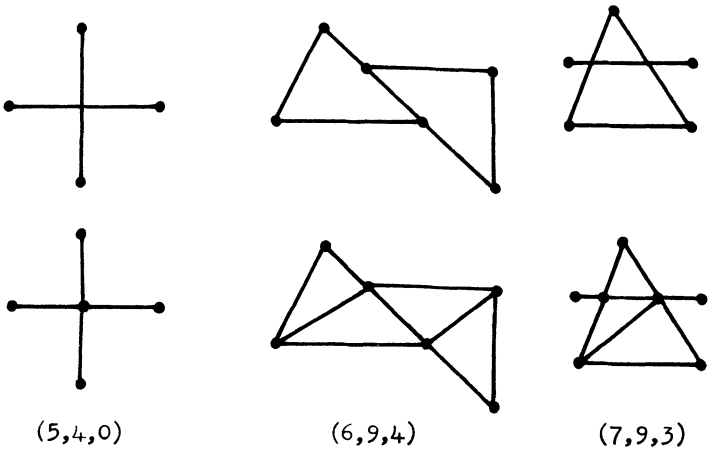
(5X) The set of all faces of an n-simplex s_n is a sim-
plicial complex of dimension n (see 3.10 (2)), which is
denoted by $\bar{s}_n$.   The set of all proper faces of s_n is a
simplicial complex of dimension n - 1 (even for n = 0 if
we adopt the convention of (2) above).   It is denoted $\mathring{s}_n$
and is called the <u>boundary</u> of s_n.   Thus (compare 3.10 (1))
the "boundary of s_n" refers either to the set of proper
faces of s_n or to the union of the proper faces of s_n.
This will not cause difficulty.

(6XX) The conditions (1) and (2) of 3.14 are equivalent
to (1) and the following condition (2)'.

(2)' If s and $t \in K$, where $s \neq t$, then **the**

open simplexes $\overset{\circ}{s}$ and $\overset{\circ}{t}$ with the same vertices as s and
t are disjoint.

It is not too hard to prove (1) & (2) => (1) & (2)'.
The proof of the converse implication is more difficult. One
method is to use induction, assuming the result true whenever
K contains only simplexes of dimension < n, and proving it
when K contains only simplexes of dimension ≤ n. The in-
duction hypothesis then applies to the subset of K consisting
of simplexes of dimension < n and since open simplexes are
disjoint we have s ∩ t = $\overset{\circ}{s}$ ∩ $\overset{\circ}{t}$ for any s, t ∈ K. Hence
by the induction hypothesis s ∩ t is a union of common faces
of s and t. However it is also convex (see 3.8 (4)) and
it is not hard to show that a convex union of (closed) faces of
a simplex is itself a face.

3.16 Definitions   (1) A subcomplex of a simplicial complex
K is a subset L of K which is itself a simplicial complex.
(This is so if and only if L has the property that

$$s \in L \ \& \ t < s \ \Rightarrow \ t \in L,$$

since 3.14 (2) is automatically satisfied.) The subcomplex
L is called proper if L ≠ K.

(2) An oriented simplicial complex is a simplicial com-
plex in which every simplex is provided with an orientation.

Thus every simplicial complex can certainly be oriented,
usually in many different ways. This is not to be confused
with the coherent orientation of a closed surface, as in 2.3,
which is not always possible.

By orienting a simplicial complex each subcomplex becomes
oriented in a natural way. Orientations of 1-simplexes and
2-simplexes will be indicated by arrows and circular arrows
respectively.

<u>3.17</u> <u>Examples</u>   (1X) If K_1 and K_2 are subcomplexes of
K then so are $K_1 \cup K_2$ and $K_1 \cap K_2$. How about $K_1 \setminus K_2$?

(2XX) Let K_1 and K_2 be simplicial complexes in $\mathbb{R}^N$
and suppose that $\left| K_1 \right| \cap \left| K_2 \right| = \left| L \right|$, where L is a sub-
complex both of K_1 and of K_2. Show that $K_1 \cup K_2$ is a
simplicial complex. (There is no difficulty showing that
(1) of 3.14 holds, and the only tricky case of (2) is when
$s \in K_1$ and $t \in K_2$. It is probably easiest to prove that,
supposing $s \neq t$, the open simplexes determined by s and t
are disjoint, and then to use 3.15 (6).)

(3) Let r be an integer $\geqslant 0$. The <u>r-skeleton</u> of a
simplicial complex K is the set of simplexes of K with
dimension $\leqslant r$; it is denoted K^r. Clearly K^r is a sub-
complex of K. The 1-skeleton K^1 of any simplicial complex
K is a graph.

(4) A simplicial complex K is called <u>connected</u> if K^1
is a connected graph, i.e. if there is a path along the edges
of K^1 connecting any two given vertices. For any K, a
<u>component</u> of K is a maximal connected subcomplex; thus K
is the (disjoint) union of its components, and K is connected
if and only if it has exactly one component.

ABSTRACT   SIMPLICIAL   COMPLEXES   AND   REALIZATIONS

A simplicial complex K is determined by (i) its set V of
vertices and (ii) the subsets of V which span simplexes of
K. This information can be presented abstractly, and the
result is an "abstract simplicial complex" which can be
"realized" to give K. The definitions are parallel to those
given for graphs in Chapter 1, and again there is a theorem
that realization is always possible. We shall make use of
these ideas in the next section, in an attempt to clarify the
meaning of certain diagrams representing simplicial complexes.

3.18 Definition   An abstract simplicial complex is a pair
$X = (V, S)$ where V is a finite set whose elements are
called the vertices of X and S is a set of non-empty sub-
sets of V.   Each element $s \in S$ is called a simplex of X,
and if $s \in S$ has precisely $n + 1$ elements $(n \geqslant 0)$, s is
called an n-simplex.   (Thus an "abstract" simplex is merely
the set of its vertices.)   S is required to satisfy the
following two conditions:

 (1) $v \in V$ => $\{v\} \in S$
 (2) $s \in S$, $t \subset s$, $t \neq \emptyset$ => $t \in S$.

The dimension of X is, so long as $V \neq \emptyset$, the largest n
for which S contains an n-simplex.   (The dimension of
$(\emptyset, \emptyset)$ is defined to be -1.)
    Two abstract simplicial complexes $X = (V, S)$ and
$X' = (V', S')$ are said to be isomorphic if there is a
bijection $f : V \to V'$ with the property that

$$\{v^0, \ldots, v^n\} \in S \iff \{f(v^0), \ldots, f(v^n)\} \in S'.$$

It is clear that any simplicial complex K determines an
abstract simplicial complex $X = (V, S)$: take V to be the
set of vertices of K and let $\{v^0, \ldots, v^n\} \in S$ if and only
if $(v^0 \ldots v^n) \in K$.   Two simplicial complexes are called
isomorphic if the abstract simplicial complexes they determine
are isomorphic.   A realization of an abstract simplicial
complex X is a simplicial complex K whose corresponding
abstract simplicial complex is isomorphic to X.   Thus the
vertices of K can be so labelled that a set of vertices
of K spans an n-simplex of K if and only if the corres-
ponding set of vertices of X is an n-simplex of X.

    This idea generalizes 1.3 and the following theorem gen-
eralizes 1.4.   Readers interested only in 2-dimensional sim-
plicial complexes should substitute $n = 2$ throughout the

theorem and proof. The proof could in any case safely be
omitted on a first reading.

3.19 Theorem   Every abstract simplicial complex of dimen-
sion n has a realization in $\mathbb{R}^{2n+1}$. (Note that any two
realizations of the same abstract simplicial complex are by
definition isomorphic.)

Sketch of proof   Consider the "rational normal curve"

$$C \;=\; \{(x,\, x^2,\, \ldots,\, x^{2n+1}) \;\in\; \mathbb{R}^{2n+1} \;:\; x \in \mathbb{R}\}.$$

I claim that any 2n + 2 distinct points of C are a-inde-
pendent. The proof is similar to that in 1.4 of the case
n = 1: it amounts to checking that a certain determinant (a
Vandermonde determinant) does not vanish.

Now let $X = (V,\, S)$ be an abstract simplicial complex
and take points $P_1, \ldots, P_k$ on C, one for each element
of V $(k = |V|)$. If $s \in S$ then $s \subset V$ and
$|s| \leqslant n + 1 < 2n + 2$ so the P's corresponding to ele-
ments of s are a-independent and span a $(|s| - 1)$-simplex.
Let K be the set of such simplexes for all $s \in S$. It
remains to prove that K is a simplicial complex for, if it
is, then the underlying abstract simplicial complex is cer-
tainly X again.

Take two simplexes in K and suppose that their sets of
vertices overlap. Then we can take them to be

$$(P_1 \ldots P_\alpha \ldots P_\beta) \quad \text{and} \quad (P_\alpha \ldots P_\beta \ldots P_\gamma)$$

say, where $\gamma < 2n + 2$ since dim X = n, so that $P_1, \ldots, P_\gamma$
are a-independent. Using the definition of a-independence
and 3.2 (2) it is not hard to show that these simplexes inter-
sect precisely in $(P_\alpha \ldots P_\beta)$ which is a common face. If
the simplexes have disjoint vertex sets it is equally easy to
show that they are disjoint.   Q.E.D.

Since orientation is defined entirely in terms of vertices, there is no difficulty in transferring the concept of orientation from simplicial complexes to abstract simplicial complexes.

TRIANGULATIONS AND DIAGRAMS OF SIMPLICIAL COMPLEXES

A subset U of $\mathbb{R}^N$ which is the underlying space (3.14) of some simplicial complex is called a <u>polyhedron</u>. Of course there will almost certainly be many simplicial complexes whose underlying space is a given polyhedron U (the only exceptions are when U is finite); any of these is called a <u>triangulation</u> of U.

Something more will be said about triangulations in Chapter 5. For the present we are concerned with diagrammatic representation, which we shall discuss by means of an example which sufficiently illustrates what is involved, and in particular clarifies the meaning of diagrams such as those in 2.2 (3), (4) and preceding 2.6.
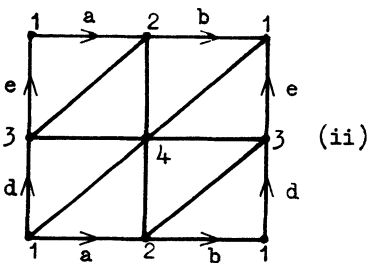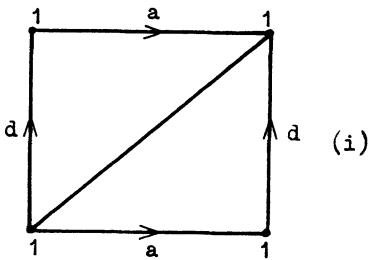
If we cut a torus along two circles and open it out as in 2.2 (3) the result is a plane rectangle with opposite sides similarly marked in order to remind us that they came from the same place on the torus. Consider the adjacent three diagrams (which occurred in Chapter 2). We want, if possible, to regard these as recipes or
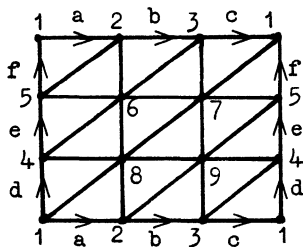
"schemas" for the con-
struction of simplicial
complexes in which dots,
edges and triangles rep-
resent 0-, 1- and 2-
simplexes respectively.
To do this, we first
regard each of them as a schema for constructing an abstract
simplicial complex X = (V, S).

First label the vertices (dots) consistently, so that for
example in any one diagram the beginning of a has the same
label wherever it occurs.   For notational convenience they
have been labelled with numerals.   Then take V to be the
set of vertices, so that V = {1} in (i), V = {1, 2, 3, 4}
in (ii) and V = {1, ..., 9} in (iii).

Next define S as follows: a subset s ⊂ V belongs to
S if and only if

$$s = \{v\} \quad \text{for some} \quad v \in V$$

or   s is the set of vertices of an edge in the diagram

or   s is the set of vertices of a triangle in the diagram.

It is easy to see that this defines an abstract simplicial
complex X = (V, S), which by 3.19 has a realization, but
there are two questions to ask.

Qu. 1.   Do edges and triangles in the diagram represent
1- and 2-simplexes of X?

Qu. 2.   What connexion, if any, is there between a
realization of X and the torus from which we started?

3.20   The answer to Qu. 1 is "yes" if and only if any given
edge in the diagram has differently labelled end-points (which
implies that any given triangle has three differently labelled
vertices).

This is clear from the definition of an n-simplex of X
as a set of vertices with precisely n + 1 elements.   Thus
(i) does not define an abstract simplicial complex of the kind
required.   Each edge and each triangle of (i) gives only a
0-simplex of X, and a realization of X is just one point.
On the other hand for (ii) and (iii) the answer to Qu. 1 is
"yes".   (In (ii), it does not matter that a and b have
the same set of end-points;   what matters is that the end-
points of a are different and the end-points of b are
different.)

As for Qu. 2, this demands a less formal answer.   When
(ii) or (iii) is realized two edges with the same vertices will
automatically be realized in the same place, i.e. will give
rise to the same 1-simplex.   The question is: does the real-
ization look anything like the original torus?

3.21   Suppose that two edges (or two triangles) of the dia-
gram have the same vertices if and only if they came from the
same position on the torus before it was cut open.   Then we
shall regard a realization of X as a triangulation of the
torus.

This means that when X is realized only the necessary
identifications take place, and no extra ones.   (It follows
that the underlying space of the realization is homeomorphic
to the original torus.)   When the X coming from (ii) is
realized a great many extra identifications take place (for
example a and b^{-1} are identified), and the reader may like
to try drawing the result.   (Hint: the bottom half of the
diagram is a repetition of the top half, and can be discarded.)

3.22 Examples   (1X) Theorem 3.19 assures us that the X
coming from (iii) can be realized in $\mathbb{R}^5$ (see also 2.16 (5)),
whereas the original torus was in $\mathbb{R}^3$.   Can this X in fact
be realized in $\mathbb{R}^3$?   A harder question (for the answer see

Gardner (1975, 1978)) is whether the diagram below, which also
gives a torus, can be realized in $\mathbb{R}^3$.
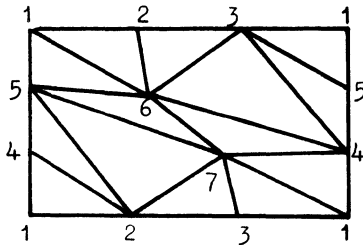
3.22 (2) In 2.6, (i) ensures that the answer to Qu. 1 is
"yes" while (ii) and (iii) relate to Qu. 2.

(3X)
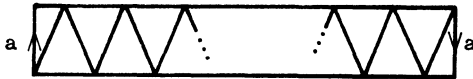
This diagram represents a Möbius band (compare 2.2 (6)) if
and only if there are at least five triangles. Reversing
one of the arrows it represents a cylinder (obtained by iden-
tifying the ends of a rectangle without a twist) if and only
if there are at least six triangles.

   (4X) Find a triangulation of the object obtained by
identifying all three sides of a triangle, as in the diagram
(this is sometimes referred to as a dunce hat).
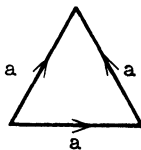
   (5X) Cut a double torus (see p.68) along suitable curves
until it can be opened out flat, and hence find a triangulation.
(You could use the method of 2.13 to check that your opened
out diagram represents the double torus, which is orientable
and has Euler characteristic -2.)

<u>3.22</u>(6)   Any diagram of the type exemplified by (i)-(iii) above can be subdivided into one in which the answer to Qu. 1 (p. 103) is "Yes" and in which the answer to Qu. 2 is that the realization is homeomorphic to the torus (or whatever it was we started from).   Two barycentric subdivisions are enough for (i), and one is enough for (ii).   (Compare p. 61.)
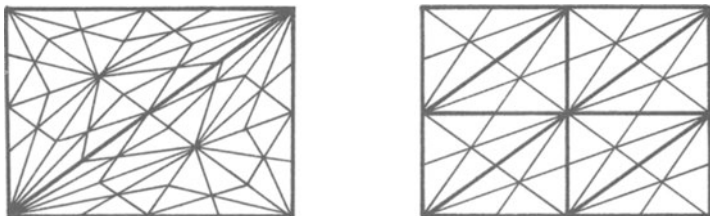
*(7)   Diagrams such as (i)-(iii) above can be interpreted other than as recipes for the construction of abstract simplicial complexes.   For example (i) can be interpreted as a covering of the torus by two triangles which are wrapped around the torus in such a way that each edge becomes a circle and all the
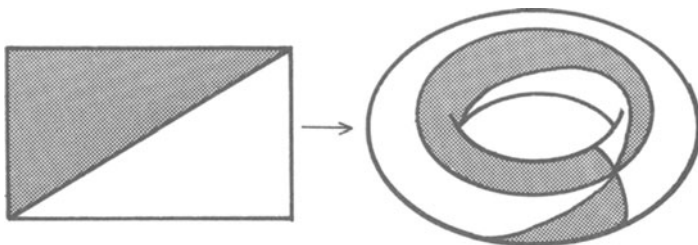
vertices land on the same point.   Thus identifications take place on the boundaries of the triangles, but notice that there are no identifications of interior points.   This kind of construction cannot be made with flat triangles, and the intersection condition 3.14(2) is hopelessly violated.

Likewise (ii) can be interpreted as eight triangles on the torus, and again the intersection condition is violated, though in this case each individual triangle lands on the torus without identifications.   Of course with (iii) the intersection con-

dition is satisfied.

These coverings by curved triangles which may have iden-
tifications on their boundaries and may violate the intersection
condition are called <u>cell complexes</u>.   It is possible to use
much more general building blocks than triangles: for example
the diagram makes the torus into a cell complex with one 2-cell

(the rectangle), two
1-cells (a and b) and
one 0-cell (v).
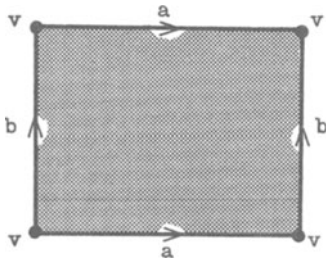
This is certainly a
desirable situation:
calculations with cells
are easier than with
simplexes because there
are fewer of them.

But you never get something for nothing: the definition of
cell complex is technically much more complicated than that of
simplicial complex, and requires considerable preparation of a
"general topology" kind.   One reason for this is that flat
triangles can be defined simply by a little linear algebra,
whereas for curvy triangles one cannot get away from the con-
cept of continuous mappings between subsets of, at least,
euclidean spaces.   For more information on cell complexes,
see for example Brown (1968), p. 121.   An account of the homo-
logy theory of quite a general class of cell complexes can be
found in Cooke and Finney (1967).

STARS JOINS AND LINKS

In this section we shall introduce some standard terms used in connexion with simplicial complexes, and state a few results for the reader to verify. The section is intended primarily as a source of examples from which the reader may develop a greater facility in working with and thinking about simplicial complexes. Only 3.23, 3.29 (2) and 3.29 (5) are of significance later. The first two are straightforward while the third is a fairly hard exercise giving the construction of "stellar subdivisions" which are used in Chapter 5, itself mostly "optional reading". Enough is said in Chapter 5 to enable the reader who has not worked through 3.29 (5) to follow the arguments, at any rate for the case of 2-dimensional simplicial complexes.

Throughout, K is a simplicial complex.

3.23 Closure Let Σ be a set of simplexes of K. The closure of Σ (in K) is the smallest subcomplex of K which contains Σ, i.e. it is defined by

$$\overline{\Sigma} = \{s \in K : s < t \text{ for some } t \in \Sigma\} \ .$$

Other notations are $C\ell(\Sigma)$ and $C\ell(\Sigma, K)$. If $s \in K$ and $\Sigma = \{s\}$, then we shall write $\bar{s}$ rather than $\overline{\{s\}}$ for the closure of Σ. (Compare 3.15 (5)) Note that $|\bar{s}| = s$ (3.14); indeed $|\Sigma| = |\overline{\Sigma}|$ for any Σ.

3.24 Star Let Σ be a set of simplexes of K. The star of Σ (in K) is defined by

$$St(\Sigma) = St(\Sigma, K) = \{s \in K : s > t \text{ for } \\ \text{for some } t \in \Sigma\} \ .$$

The **closed star** of Σ (in K) is the closure of the star of Σ, and is denoted $\overline{St}(\Sigma)$ or $\overline{St}(\Sigma, K)$. Thus

$$\overline{St}(\Sigma) = \{u \in K : \exists\, t \in \Sigma \text{ and } s \in K$$

$$\text{with } s > u \text{ and } s > t\}$$

and $\overline{St}(\Sigma)$ is a subcomplex of K. The most important example
is $\Sigma = \{s\}$: we write $St(s)$ and $\overline{St}(s)$ rather than $St(\{s\})$
and $\overline{St}(\{s\})$.

3.25 Examples   (1X) Let s_0 be a 0-simplex of K. Then
$s_0 \in St(\Sigma)$ <=> $s_0 \in \Sigma$. It follows from this that $St(\Sigma)$
is a subcomplex of K if and only if, for each component K_1
of K, Σ contains either all or none of the 0-simplexes of K_1.
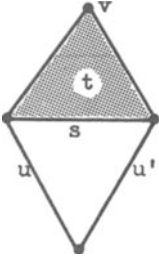
(2) Let K be the simpli-
cial complex in the diagram,
and let $\Sigma = \{s\}$. Then
$\overline{St}(s) = \overline{t}$ where t is
the 2-simplex. On the
other hand
$St(\overline{s}) = (\overline{t} \setminus \{v\}) \cup \{u, u'\}$.
Thus $\overline{St}(\Sigma)$ and $St(\overline{\Sigma})$ may be sets neither of which contains
the other.

(3X) For $s \in K$,

$t \in \overline{St}(s)$ <=> s and t are both faces of
the same simplex of K.

Thus the relation between s and t defined by $t \in \overline{St}(s)$ is
symmetric (and reflexive). Is it transitive?

(4X) If $\Sigma_1 \subset \Sigma_2$ then $St(\Sigma_1) \subset St(\Sigma_2)$. Is the converse
true?

3.26 Join   Let $s_p = (v^0 \ldots v^p)$ and $s_q = (w^0 \ldots w^q)$
be simplexes in $\mathbb{R}^N$. They are called joinable if the
$p + q + 2$ vectors $v^0, \ldots, v^p, w^0, \ldots, w^q$ are affinely
independent, and in that case their join $s_p s_q$ is the
$p + q + 1 -$ simplex $(v^0 \ldots v^p w^0 \ldots w^q)$. (If they are not

joinable then their join is undefined.)   Thus $s_p s_q = s_q s_p$ when either side is defined.

Two simplicial complexes K and L in $\mathbb{R}^N$ are called joinable if, for each $s \in K$ and $t \in L$, s and t are joinable, and if any two such joins, st and s't', meet, if at all, in a common face of st and s't'. The join KL of K and L is then the simplicial complex

$$KL = K \cup L \cup \{st : s \in K \text{ and } t \in L\}.$$

3.27 Examples   (1XX) Let $a \in s_p$ and $b \in s_q$; then the segment joining a and b consists of points $\lambda a + (1 - \lambda)b$ for $0 \leqslant \lambda \leqslant 1$. We have: s_p and s_q are joinable if and only if any two such segments intersect at most in an end-point, and in that case

$$s_p s_q = \{\lambda a + (1 - \lambda)b : a \in s_p, \ b \in s_q, \ 0 \leqslant \lambda \leqslant 1\} .$$

Likewise two simplicial complexes K and L are joinable if and only if the segments $\{\lambda a + (1 - \lambda)b : 0 \leqslant \lambda \leqslant 1\}$ are disjoint except for end-points for all $a \in |K|$ and $b \in |L|$, and $|KL|$ is then the union of all such segments. It follows that if $|K| = |K_1|$ and $|L| = |L_1|$ and K and L are joinable, then so are K_1 and L_1 and $|KL| = |K_1 L_1|$.

(2) The simplicial complexes (in $\mathbb{R}^2$) on the left are not joinable, whereas those in the centre are, and their join is drawn on the right.
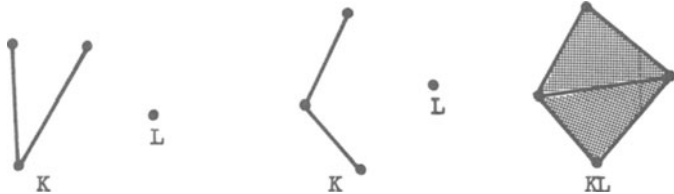
(3) Any K and the empty simplicial complex $\emptyset$ are joinable, and $K\emptyset = K$.

3.27 (4) If L consists of a single 0-simplex (v), and K
and L are joinable, then KL is also called the cone on K
with vertex v.   We shall meet this idea again in 6.8.

(5X) Writing α_p for the number of p-simplexes of a
simplicial complex, and supposing KL is defined,

$$\alpha_p(KL) = \alpha_p(K) + \left(\sum_{q=0}^{p-1} \alpha_{p-q-1}(K)\alpha_q(L) \right) + \alpha_p(L)$$

(the middle sum being absent when p = 0).   In particular let
K and L be 0-dimensional, consisting of m and n 0-sim-
plexes respectively.   Then KL is a graph with cyclomatic
number $(m - 1)(n - 1)$.   (Clearly KL is bipartite (1.21 (5))).

(6X) If the simplexes in 3.26 are regarded as oriented,
then the orientation of $\sigma_p \sigma_q$ can be defined by the ordering
$(v^0 \dots v^p w^0 \dots w^q)$.   In that case,

$$\sigma_p \sigma_q = (-1)^{(p+1)(q+1)} \sigma_q \sigma_p .$$

3.28 Link   Let s ∈ K.   The link of s in K is defined
by

$$Lk(s) = Lk(s, K) = \{t \in K : st \in K\}.$$

Recall that st ∈ K means (i) s and t are joinable, and
(ii) the join st belongs to K.   Lk(s) is a subcomplex of
K.

3.29 Examples   (1) In the case of the simplicial complex
of 3.25 (2), Lk(t) = $\emptyset$, Lk(s) = $\{v\}$, Lk(v) = $\bar{s}$.

(2X) Let K be a simplicial complex of dimension 2.
Then the link of a 0-simplex v^0 consists of all the 0-sim-
plexes v^1 for which $(v^0 v^1) \in K$ and all the 1-simplexes
$(v^1 v^2)$ for which $(v^0 v^1 v^2) \in K$.   The link of a 1-simplex
$(v^0 v^1)$ consists of all 0-simplexes v^2 for which $(v^0 v^1 v^2) \in K$.
If we suppose that the link of every 0-simplex is a simple
closed polygon, then it follows that the link of every 1-simplex

is two 0-simplexes and that K is pure (every simplex is a
face of a 2-simplex).   Thus K consists precisely of the
triangles of K together with their edges and vertices.
Hence the definition of a closed surface given in 2.1 can be
rewritten:

A <u>closed surface</u> is a connected 2-dimensional simplicial
complex in which the link of every 0-simplex is a simple
closed polygon.

<u>3.29</u>(3X) Let s_0 be a 0-simplex of K.   Then

$$Lk(s_0) = \overline{St}(s_0) \setminus St(s_0).$$

Is this ever true of an n-simplex for n > 0?

(4X) Let s ∈ K.   Then $\bar{s}$ and Lk(s) are joinable,
and $\bar{s}Lk(s)$ = $\overline{St}(s)$.   In particular if s is a 0-simplex
s_0, then $\overline{St}(s_0)$ is the cone on $Lk(s_0)$ with vertex s_0.

(5XX) Let s ∈ K and let v be an interior point of s,
i.e. a point of s not on a proper face of s.   Let v also
denote the simplicial complex whose only simplex is (v).
Then v and $\overset{\bullet}{s}$ (the set of proper faces of s, 3.15 (5))
are joinable, and their join $v\overset{\bullet}{s}$ is joinable to Lk(s, K).
Further, $\left|v\overset{\bullet}{s}\right|$ = s, and $\left|(v\overset{\bullet}{s})Lk(s, K)\right|$ = $\left|\overline{St}(s, K)\right|$.
(3.8 (5) is relevant to the proof of $\left|v\overset{\bullet}{s}\right|$ = s; see also
3.27 (1).)   Finally, $(K \setminus \overline{St}(s, K)) \cup (v\overset{\bullet}{s})Lk(s, K)$ is a
simplicial complex with the same underlying space as K.   It
is said to be obtained from K by <u>starring</u> s <u>at</u> v.
Repeated application of this procedure yields a <u>stellar sub-</u>
<u>division</u> of K.   Some examples of starring are drawn below.
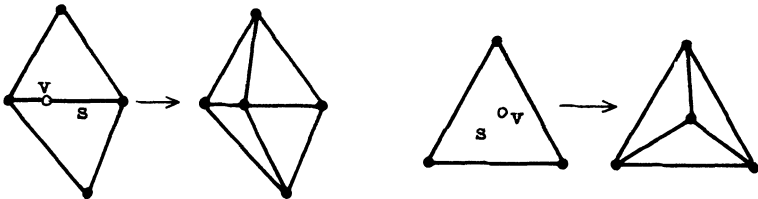
COLLAPSING

The technique of collapsing is a simple geometrical idea which
will be exploited in Chapter 6.   Since, however, it does not
assume any of the theory of homology groups, we shall intro-
duce the idea here and give one or two mild applications.

<u>3.30 Definitions</u>   Let K be a simplicial complex.   A
<u>principal simplex</u> of K is a simplex of K which is not
a proper face of any simplex of K.   For any $s \in K$, a
<u>free face</u> of s is a proper face t of s such that t is
not a proper face of any simplex of K besides s.   (Thus
a free face of s has dimension dim s - 1 and if s has
a free face then s must be principal.)
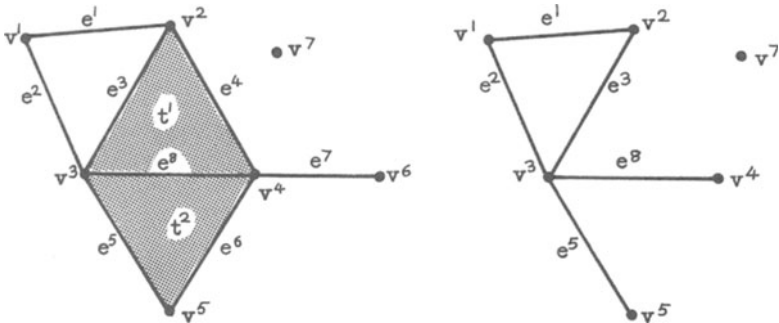
In the left-hand diagram, the principal simplexes are:
$$v^7;\ e^1,\ e^2,\ e^7;\ t^1,\ t^2.$$
Of these, only e^7, t^1, t^2 have free faces, namely v^6 for
e^7, e^3 and e^4 for t^1, and e^5 and e^6 for t^2.

   Suppose that K has a principal simplex s which has a
free face t.   Then the subset K_1 of K obtained by re-
moving s and t is a subcomplex of K, and is said to be
obtained from K by an <u>elementary collapse</u> (which we describe
as being a <u>collapse of</u> s <u>across</u> t).   The notation is

$K \searrow^e K_1$. The right-hand diagram above shows the result of three elementary collapses (t^1 across e^4, t^2 across e^6 and e^7 across v^6); two further elementary collapses, of e^5 across v^5 and e^8 across v^4, now become possible.

A sequence $K \searrow^e K_1 \searrow^e K_2 \searrow^e \ldots \searrow^e K_m$ of elementary collapses is called a <u>collapse</u> and is denoted by $K \searrow K_m$. (A special case is the empty sequence by which $K \searrow K$!) If K_m consists of a single vertex we say that K <u>collapses to a point</u>.

This concept has already been encountered for graphs in 1.12 (2): a graph collapses to a point if and only if it is a tree.

Let $\mathcal{D}$ be the plane region (a <u>disk</u>) enclosed by a simple closed polygon (see 1.25), and let D be a triangulation of $\mathcal{D}$, i.e. a simplicial complex (of dimension two) such that $|D| = \mathcal{D}$. Our object is to prove that D collapses to a point. A naïve argument would run: "Collapse one triangle of D across a free face on the boundary polygon. What is left is a triangulation D' of a region $\mathcal{D}'$ enclosed by a simple closed polygon, and by induction D' $\searrow$ point, Q.E.D."
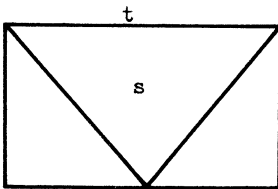
Unfortunately, this will not do, as can be seen by collapsing s across t in the diagram: what is left is not enclosed by a simple closed polygon. The trouble is that s meets the boundary polygon in an edge and an isolated vertex.

<u>3.31</u> <u>Lemma</u>   Let us denote the simple closed polygon enclosing $\mathcal{D}$ by ∂D. Suppose that D contains at least two triangles; then D contains at least two triangles meeting ∂D in an arc, i.e. in an edge or two edges. (see 3.22 (3) for a picture which, on removing the arrows, gives an example where

D contains exactly two such triangles.)

<u>Proof</u>   The proof is by induction on $\alpha_2(D)$, the num-
ber of triangles of D: clearly it is true when $\alpha_2 = 2$ so
suppose it holds when $\alpha_2 < n$ $(n \geqslant 3)$ and let $\alpha_2(D) = n$.
Let s be a triangle of D meeting ∂D in at least an edge -
clearly D must contain at least two such triangles.   Then
s $\cap$ ∂D is

(1)   (2)   (3)
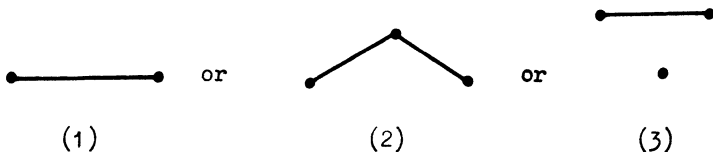
If there are no triangles of type (3) in D then all choices of
s meet ∂D in an arc, so the result is proved.   If there is
a triangle s of type (3), proceed as follows.   Label the

vertices of ∂D by
$P_1, P_2, P_3, ..., P_m$ where
s has vertices $P_1 P_2 P_i$
for some i, $4 \leqslant i \leqslant m - 1$.
Then $P_2 P_3 \cdot ..P_i P_2$ and
$P_1 P_i P_{i+1}...P_m P_1$ are
simple closed polygons enclosing plane regions $\mathcal{D}'$, $\mathcal{D}''$ tri-
angulated by D', D'', say. If $\alpha_2(D') = \alpha_2(D'') = 1$, then
the unique triangles in D', D'' meet ∂D in an arc. If
$\alpha_2(D') > 1$, $\alpha_2(D'') = 1$, then by the induction hypothesis
choose two triangles of D' meeting $\partial D'$ in an arc. At
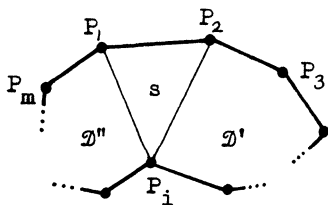most one of these can contain the edge $P_2 P_i$; the other[†],
together with the unique triangle in D'', both meet ∂D in
an arc.   The other cases are treated similarly; in every
case D contains at least two triangles meeting ∂D in an
arc.   Q.E.D.

[†] Notice that the proof would break down at this point if we
were trying to prove the weaker assertion that D always con-
tains at least one triangle meeting ∂D in an arc. This is
a not uncommon feature of proofs by induction.

<u>3.32</u>   <u>Theorem</u>   With the above notation, any triangulation
D of $\mathcal{D}$ collapses to a point.

<u>Proof</u>   Clearly the theorem holds when $\alpha_2(D) = 1$, the
collapse then being

Suppose the theorem is true when $\alpha_2 < n$ $(n \geqslant 2)$ and let
$\alpha_2(D) = n$. By the lemma choose a triangle s of D meeting
∂D in an arc, and collapse as follows:

or

In either case $D \searrow D_1$ where D_1 is a triangulation of the
plane region bounded by a simple closed polygon, and
$\alpha_2(D_1) = n - 1$.   Thus $D \searrow D_1 \searrow$ point, by the induction
hypothesis.   Q.E.D.

       The <u>Euler characteristic</u> of a simplicial complex K of
dimension n is $\chi(K) = \alpha_0(K) - \alpha_1(K) + \alpha_2(K) - \ldots +$
$(-1)^n \alpha_n(K)$, α_i being the number of simplexes of dimension i
in K, i = 0, ..., n.   (Compare 2.11)   Clearly if $K \searrow L$
then $\chi(K) = \chi(L)$ (consider the effect of an elementary

collapse).   Hence:

<u>3.33</u> <u>Corollary</u>   If D is a triangulation of the plane
region $\mathcal{D}$ enclosed by a simple closed polygon, then $\chi(D) = 1$.
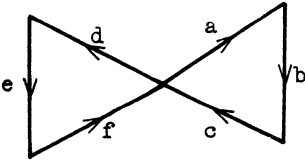(Compare the proof of 2.13.)

There is another consequence of 3.31 which will be use-
ful later.

<u>3.34</u> <u>Theorem</u>   Let D be a triangulation of the plane region
$\mathcal{D}$ enclosed by a simple closed polygon ∂D, and let the sim-
plicial complex K be obtained from D by deleting one
2-simplex.   Then $K \searrow \partial D$.

<u>Proof</u>   There is nothing to prove if $\alpha_2(D) = 1$, for
then $K = \partial D$ and the empty sequence of elementary collapses
will work !   Suppose the result is true when $\alpha_2 < n$ $(n \geqslant 2)$
and that $\alpha_2(D) = n$.   By the lemma (3.31) K has at least
one triangle s meeting ∂D in an arc.   Collapse as in the
proof of 3.32; let this take D to D' and K to K'.
Then D' is a triangulation of a disk containing n - 1
triangles, and K' is obtained from D' by removing a single
triangle s; hence by the induction hypothesis $K' \searrow \partial D'$.
Now none of the elementary collapses comprising this collapse
can be across a face of s, since the face or faces of s
in K' are still present at the end, in $\partial D'$.   Hence
$K = K' \cup \bar{s} \searrow \partial D' \cup \bar{s}$ where $\bar{s}$ denotes, as usual, the set
of all faces of s including s itself.   Since $\partial D' \cup \bar{s} \searrow \partial D$
this completes the induction.   Q.E.D.

<u>3.35</u> <u>Examples</u> (1) This result can be used to given an alt-
ernative proof of 2.13 without using $\chi(D) = 1$.   For removing
a single triangle from a closed surface gives a simplicial
complex which collapses to the 1-dimensional subcomplex which
appears, in a polygonal representation of the surface in the
plane, as the boundary of the polygon.   The Euler character-
istic of this graph is easy to calculate, and the removal of a

triangle at the beginning just lowers the Euler characteristic
by 1.   For example the triangulation of a torus in 2.2 (3)
collapses, after removal of a triangle, to the **graph** on the
left which has Euler character-
istic -1.   Thus the original
simplicial complex has Euler
characteristic 0.

3.35 (2) The triangulations of a cylinder and a Möbius band
in 3.22 (3) collapse to simple closed polygons (sometimes
called spines of the original simplicial complexes).

    $(3X)$ Let M be a closed surface and let K be a sub-
complex of M which does not contain all the triangles of
M.   Supposing that K contains at least one triangle, then
it contains at least one triangle with a free face, and there-
fore K collapses to a subcomplex with fewer triangles and so
by induction to a graph.   (To see that K must have a
triangle with a free face, use 2.2 (5) to show that otherwise
K would have to contain all the triangles of M.)

    *(4) It is perhaps worth noting that 3.32 does not gener-
alize to higher dimensions.   A simplicial complex in $\mathbb{R}^3$
whose underlying space is homeomorphic to the unit sphere
$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ in $\mathbb{R}^3$ will enclose a
region known as a 3-ball.   It is possible to find such a sim-
plicial complex and a triangulation of the resulting 3-ball
which does not collapse to a point.   See Bing (1964).

*APPENDIX ON ORIENTATION

The passage from simplexes to oriented simplexes is an essen-
tial one for setting up the homology theory described in this
book.   The concept of orientation arises also in the study of
real vector spaces, and in this appendix we shall relate these
two ideas as well as proving the basic property of orienta-
tions - that there are always exactly two of them.

Let V be a real vector space of dimension $n \geqslant 1$ provided with a definite basis, called the <u>preferred basis</u>, and suppose that $e^1, \ldots, e^n$ and $f^1, \ldots, f^n$ are <u>ordered bases</u> of V, that is they are bases in which the vectors come in a prescribed order. (Such a basis is sometimes called an <u>n-frame</u>.) We want to compare the "orientations" of these two bases.

Take for example $V = \mathbb{R}^2$. Then geometrically e^1, e^2 and f^1, f^2 have the same orientation if and only if they are both "clockwise" or both "anticlockwise" - more precisely this means that the rotations from e^1 to e^2 and from f^1 to f^2, in each case through an angle $< 180°$, are both in the same direction. Thus in the diagram the first and second bases have the same orientation, not shared by the third basis. When $V = \mathbb{R}^3$ then geometrically we require that the two bases should both be "right-handed" or both "left-handed". Thus for example e^1, e^2, e^3 and e^2, e^1, e^3 have opposite orientation while e^1, e^2, e^3 and $-e^2, e^1, e^3$ have the same orientation.
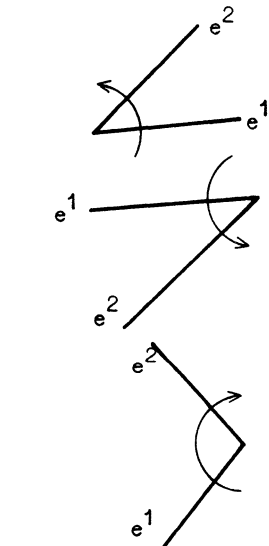
Now we ask the question: can we find a continuous family[†] of ordered bases $u^1(t), \ldots, u^n(t)$, depending on a parameter

---

[†] Write $u^j(t) = (a_{1j}(t), \ldots, a_{nj}(t))$, where these are the coordinates of $u^j(t)$ relative to the preferred basis. Then we require that the n^2 functions $a_{ij}(t)$ should be continuous. (This definition is in fact independent of the choice of preferred basis.)

$t \in [0, 1]$, such that for $t = 0$ this basis is $e^1, \ldots, e^n$
and for $t = 1$ it is $f^1, \ldots, f^n$? (The important thing is
that for every $t \in [0, 1]$ we must get a basis; we can turn
e^1, e^2 (an ordered basis of $\mathbb{R}^2$) into e^2, e^1 by the family
$(1 - t)e^1 + te^2$, $te^1 + (1 - t)e^2$, but for $t = \frac{1}{2}$ these vec-
tors are not linearly independent.) The reader can probably
convince himself that for $V = \mathbb{R}^2$ and $V = \mathbb{R}^3$ such a family
exists if and only if the two ordered bases have the same
orientation in the intuitive sense described above.

Any linear isomorphism $V \to V$ determines, relative to
the preferred basis in each copy of V, a nonsingular real
$n \times n$ matrix. The isomorphism is called <u>orientation
preserving</u> if the determinant of this matrix is positive, and
<u>orientation reversing</u> if it is negative. Note that this
definition does not in reality depend on the choice of pre-
ferred basis, for if A is the matrix relative to the preferred
basis then the matrix relative to any other basis is of the
form $P^{-1}AP$ for some nonsingular P, and this has the same
determinant as A. I now claim the following.

<u>3.36 Theorem</u>   Let $e^1, \ldots, e^n$ and $f^1, \ldots, f^n$ be ordered
bases of V. The following statements are equivalent:

(i) There exists a continuous family of ordered bases
taking $e^1, \ldots, e^n$ to $f^1, \ldots, f^n$;

(ii) The unique isomorphism $V \to V$ taking e^i to f^i
for $i = 1, \ldots, n$ is orientation preserving.

When these hold we say that the two bases have the <u>same orien-
tation</u>, otherwise that they have <u>opposite orientation</u>.

The theorem will be proved at the end of this appendix.
It says that ordered bases can be divided into two orientation
classes in such a way that two bases in the same class can be
connected by a continuous family and two bases in different
classes cannot. The classes can be distinguished by means

of a determinant.

<u>3.37</u> <u>Examples</u>   (1X) $e^1, \ldots, e^n$ and $e^{\pi(1)}, \ldots, e^{\pi(n)}$,
where π is a permutation of the set $\{1, \ldots, n\}$ have the
same orientation if and only if π is an even permutation.
It is enough to check that if π is a transposition then the
two bases have opposite orientation, and this is most easily
proved by taking (as we may) $e^1, \ldots, e^n$ itself to be
the preferred basis.

(2X) $e^1, \ldots, e^n$ and $-e^1, \ldots, -e^n$ have the same
orientation if and only if n is even.

Now let us turn to <u>orientation of simplexes</u>. Let
$(v^0 \ldots v^n)$ and $(v^{\pi(0)} \ldots v^{\pi(n)})$ be two ordered simplexes,
where π is a permutation of $\{0, \ldots, n\}$. The vectors
$v^0, \ldots, v^n$ lie in an affine subspace U^A of dimension n
in $\mathbb{R}^N$, while the vectors $v^i - v^j$ lie in a linear space U
of dimension n parallel to U^A. Consider the following two
bases of U.
$$v^1 - v^0, \ldots, v^n - v^0$$
and $$v^{\pi(1)} - v^{\pi(0)}, \ldots, v^{\pi(n)} - v^{\pi(0)}.$$

<u>3.38</u> <u>Definition</u> $(v^0 \ldots v^n)$ and $(v^{\pi(0)} \ldots v^{\pi(n)})$ have
the <u>same</u> or <u>opposite orientation</u> according as these two bases
of U have the same or opposite orientation.

<u>3.39</u> <u>Examples</u>
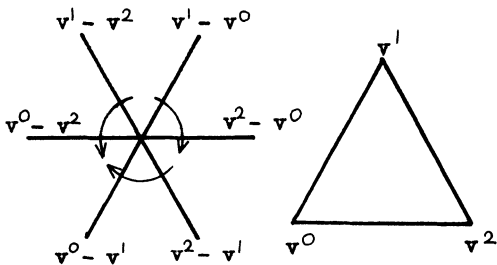
$(v^0 v^1 v^2) = (v^1 v^2 v^0)$
(same orientation)

$(v^0 v^1 v^2) = -(v^2 v^1 v^0)$
(opposite orientation)

$$v^1 - v^0, \quad v^2 - v^0, \quad v^3 - v^0$$

is "right-handed";

$$v^3 - v^2, \quad v^1 - v^2, \quad v^0 - v^2$$

is "left-handed".

Hence

$$(v^0 v^1 v^2 v^3) = -(v^2 v^3 v^1 v^0).$$

<u>3.40 Proposition</u> $(v^0 \dots v^n)$ and $(v^{\pi(0)} \dots v^{\pi(n)})$ have
the same orientation if and only if π is an even permutation.

<u>Proof</u>   It is enough to check that if π is a trans-
position then the two simplexes have opposite orientation.
Assuming that π is a transposition leaving 0 fixed the
result follows from 3.37 (1).   Assuming on the other hand
that π transposes 0 and i for some $1 \leqslant i \leqslant n$, the
matrix, referred to the first basis above, of the isomorphism
taking the first basis to the second, has the following form.

$$i \cdots \begin{pmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ -1 & -1 & \cdots & -1 & -1 & -1 & \cdots & -1 \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & & 1 \end{pmatrix}$$

Here the blank entries are all zero.   This clearly has
determinant -1.   Q.E.D.

    To conclude this appendix, here is a proof of Theorem
3.36.

<u>Proof of 3.36</u>   Any basis of V determines a (real) nonsing-
ular n x n matrix whose j$^{\text{th}}$ column consists of the coordinates
of the j$^{\text{th}}$ basis vector relative to the preferred basis.   If

E and F are the matrices determined in this way by the
given bases, then the matrix, relative to the preferred basis,
of the unique isomorphism taking e^i to f^i for
i = 1, ..., n is FE^{-1}. Thus statements (i) and (ii) of
the theorem are respectively equivalent to

(i)' There is a continuous family of nonsingular n × n
matrices taking E to F;

and (ii)' $|E|$ and $|F|$ have the same sign.

The implication (i)' => (ii)' is a straightforward con-
sequence of the fact that the determinant of a matrix, being a
polynomial in the entries of the matrix, is a continuous func-
tion of the entries. Thus if the general matrix of the
family is A_t (so that $A_0 = E$ and $A_1 = F$) then $|A_t|$,
being continuous and nonzero for $0 \leqslant t \leqslant 1$, must have the
same sign for t = 0 and t = 1.

To prove (ii)' => (i)' we must exhibit a continuous
family taking E to F. It is clearly enough to construct
a continuous family taking E to one of the two diagonal
matrices diag (1, 1, ..., 1, ± 1). The family we shall con-
struct will be made up of a sequence of families, corresponding
to a sequence of operations which reduce E to one of these
two matrices. Let A_0 be any nonsingular n × n matrix;
consider the following two operations.

I.   Multiply all the elements of one row (or column)
      of A_0 by $\lambda > 0$.

II.   Add to one row (or column) of A_0 any multiple μ
      of a different row (or, respectively, column).

Note that neither of these affects the sign of $|A_0|$. Suppose
A_1 is obtained from A_0 by an operation I. Then the general
member of a continuous family taking A_0 to A_1 is A_t,
where A_t is obtained from A_0 by an operation I, replacing

λ by $1 - t + \lambda t$. Similarly an operation II can be achieved via a continuous family.

Now using operations I and II any nonsingular $n \times n$ matrix $E = (a_{ij})$ can be reduced to a diagonal matrix with 1's and -1's on the diagonal. We show how to turn E into a matrix of the form

$$\begin{pmatrix} \pm 1 & 0 & . & . & . & . & 0 \\ 0 & & & & & & \\ . & & & & & & \\ . & & & & & & \\ 0 & & & & & & \end{pmatrix}$$

where the box contains an $(n - 1) \times (n - 1)$ nonsingular matrix; the proof then proceeds by induction. In fact, if $a_{11} = 0$ then choose a nonzero entry in the first row or column (they can't all be zero); say this is the entry a_{i1} in the ith row. Then by an operation II, with $\mu = 1/a_{i1}$, we obtain a matrix with top left-hand entry 1. If $a_{11} \neq 0$ multiply the first row by $1/|a_{11}|$ (an operation I), obtaining a matrix with top left-hand entry $\pm$ 1. Now the other entries in the first row and then the first column can easily be made zero by operations II.

Next, two -1's on the diagonal can be turned into $+1$'s:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Each operation is of type II. Similarly a 1 and a -1 can be interchanged by a sequence of operations of type II. This completes the proof that every E can be reduced to one of diag $(1, 1, \ldots, 1, \pm 1)$.   Q.E.D.

An alternative statement of the theorem is that the general linear group $GL(n, \mathbb{R})$ contains precisely two path components for any $n \geqslant 1$. The proof above constructs a polygonal path joining any two matrices in this group with determinants of

equal sign, regarding $GL(n, \ \mathbb{R})$ as being contained in $\mathbb{R}^{n^2}$ in the natural way.

Homology groups

In this chapter we shall formalize the ideas sketched in the
Introduction and define homology groups for an arbitrary
oriented simplicial complex K. There is one group, $H_p(K)$,
in each dimension p with $0 \leqslant p \leqslant \dim K$; the group $H_p(K)$
measures, roughly speaking, the number of "independent
p-dimensional holes" in K. If K is an oriented graph then
$H_1(K)$ is isomorphic to the group $Z_1(K)$ of 1-cycles on K
(see 1.19).

In order to read this chapter it is necessary to know about
quotient groups and exact sequences, and to know a little about
presentations. All these are explained in the appendix to the
book.

CHAIN GROUPS AND BOUNDARY HOMOMORPHISMS

The chain groups associated with an oriented simplicial complex
are exactly analogous to those associated with an oriented
graph (1.17), and are defined as follows.

4.1 Definition Let K be an oriented simplicial complex of
dimension n, and let α_p be the number of p-simplexes of K.
For $0 \leqslant p \leqslant n$, let $\sigma_p^1, \ldots, \sigma_p^{\alpha_p}$ be the oriented p-simplexes
of K; for such p the pth <u>chain group</u> of K, denoted $C_p(K)$,

is the free abelian group on the set $\{\sigma_p^1, \ldots, \sigma_p^{\alpha_p}\}$ (A.11).
Thus an element of $C_p(K)$ is a linear combination

$$\lambda_1 \sigma_p^1 + \ldots + \lambda_{\alpha_p} \sigma_p^{\alpha_p}$$

with the λ's integers. This is called a p-_chain_ on K and
two p-chains are added by adding corresponding coefficients.
When writing down p-chains it is customary to omit p-simplexes
whose coefficient is zero - unless they are all zero, when we
just write 0. We also write σ for 1σ . For p > n or
p < 0 the group $C_p(K)$ is defined to be 0. (Notice that
$C_p(\emptyset) = 0$ for all p, since $\emptyset$ has no simplexes at all.)

We shall shortly introduce a "boundary homomorphism" which
generalizes that for graphs (1.19). Recall that the boundary
$\partial(v^0 v^1)$ of an oriented 1-simplex $(v^0 v^1)$ is the 0-chain
$v^1 - v^0$. Now consider an oriented 2-simplex $(v^0 v^1 v^2)$. It
is certainly desirable that the boundary should be the 1-chain

$$\partial(v^0 v^1 v^2) \;=\; (v^1 v^2) + (v^2 v^0) + (v^0 v^1).$$

However, if the 2-simplex belongs to an oriented simplicial com-
plex K then the three edges of the triangle will also have
specific orientations as elements of K, and it might be that,
for example, the bottom edge has orientation $(v^2 v^1)$. In that
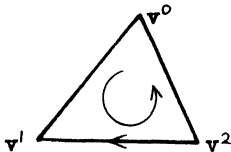
case $(v^1 v^2) \notin K$ so that the
above formula does not in fact
give a 1-chain: we should
really write

$$\partial(v^0 v^1 v^2) \;=\; -(v^2 v^1) + \ldots \;.$$

Alternatively - and this is the course we adopt - we can
make the convention that

$(v^1 v^2)$ means the 1-chain $(v^1 v^2)$ if $(v^1 v^2) \in K$

and (v^1v^2) means the 1-chain $-(v^2v^1)$ if $(v^2v^1) \in K$.

Then the above formula for $\partial(v^0v^1v^2)$ always gives a 1-chain.
In addition, this convention enables us to preserve the natural
order of indices and write

$$\partial(v^0v^1v^2) = (v^1v^2) - (v^0v^2) + (v^0v^1) .$$

More generally, the convention is as follows.

4.2 Convention   Let $(v^0 \dots v^p)$ be an oriented p-simplex
$(p \geqslant 1)$. Let K be an oriented simplicial complex and suppose
that the oriented simplex $\sigma = -(v^0 \dots v^p)$, i.e.
$\sigma = (v^{\pi(0)} \dots v^{\pi(p)})$ for an odd permutation π, belongs to K.
Then we make the convention that $(v^0 \dots v^p)$ means the p-chain
$-\sigma$, i.e. the p-chain on K in which the coefficient of σ
is -1 and every other coefficient is 0. Thus for an odd
permutation π

$$(v^0 \dots v^p) = -(v^{\pi(0)} \dots v^{\pi(p)})$$

both as oriented simplexes (see 3.12) and as p-chains on K.

4.3 Definition   Let $\sigma = (v^0 \dots v^p)$ be an oriented p-simplex
of K (K as in 4.1) for some $p > 0$. The **boundary** of σ
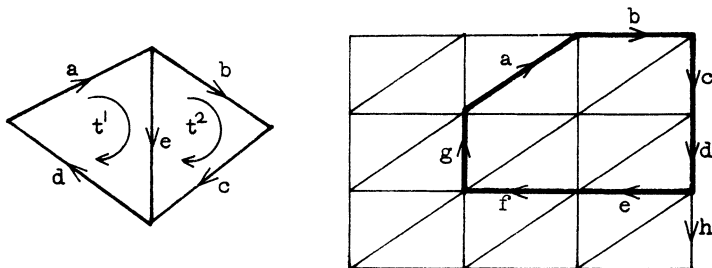is the $(p-1)$-chain

$$\partial\sigma = \partial_p\sigma = \sum_{i=0}^{p} (-1)^i (v^0 \dots \hat{v^i} \dots v^p)$$

where the hat over v^i indicates that it is omitted. Note
that we are using the convention 4.2 here, for the face of σ
with v^i omitted may or may not have the right orientation to
be a $(p-1)$-simplex of K. For $p = 0$, $\partial_p\sigma$ is defined to be
zero. The **boundary homomorphism** $\partial = \partial_p : C_p(K) \to C_{p-1}(K)$
is defined by

$$\partial_p(\Sigma\lambda_i\sigma_p^i) = \Sigma\lambda_i\partial_p(\sigma_p^i) \qquad \text{(compare A.12)}$$

for $0 \leqslant p \leqslant n$, and defined to be the trivial homomorphism
otherwise.

As an example, consider the two coherently oriented tri-

angles in the diagram on the left.   We have
$$\partial(t^1 + t^2) = (a + e + d) + (b + c - e) = a + b + c + d,$$
corresponding to the geometrical idea that the boundary of the
whole area is the polygon $a + b + c + d$.   Notice that
$\partial(a + b + c + d) = 0$.   Similarly, orienting all the triangles
in the right-hand diagram clockwise, the boundary of the sum of
the seven triangles enclosed by the heavy line is
$a + b + c + d + e + f + g$, which is a 1-chain with boundary 0.
If we make the usual identifications to turn the simplicial
complex on the right into a torus (see 2.2. (3)) then
$\partial(c + d + h) = 0$, but it can be shown quite easily that
there is no 2-chain with boundary $c + d + h$: the simple closed
polygon $c + d + h$ goes round a hole in the torus and does not
bound any region.   Thus, on the torus, $a + \ldots + g$ bounds a
region (is "homologous to zero" as we shall say shortly) where-
as $c + d + h$ does not.   The homology groups we study pick
out the simple closed polygons - more generally the 1-cycles -
which do not bound any 2-chain by taking the group of 1-cycles
and factoring out by the subgroup consisting of those which do
bound 2-chains.

<u>4.4</u> <u>Remark</u>   It is far from obvious that the above definition

of $\partial\sigma$ makes sense - i.e. that it depends only on the
oriented simplex σ and not on the particular ordering of
vertices given. For instance, $(v^0 v^1 v^2) = (v^1 v^2 v^0)$ (as
oriented simplexes), and

$$\partial(v^0 v^1 v^2) = (v^1 v^2) - (v^0 v^2) + (v^0 v^1);$$
$$\partial(v^1 v^2 v^0) = (v^2 v^0) - (v^1 v^0) + (v^1 v^2).$$

These expressions are indeed equal (bearing 4.2 in mind), but
perhaps some general reassurance is necessary. (The reassur-
ance which follows could be omitted on a first reading.) We
shall also check that the formula is compatible with the nota-
tion (3.12) $-\sigma$ for the same simplex as σ but with opposite
orientation. Both these things are contained in the statement

$$\partial\left(v^{\pi(0)} \dots v^{\pi(n)}\right) = \pm \, \partial(v^0 \dots v^n)$$

where the sign is + or - according as π is even or odd,
and it is enough to verify this when π is a transposition
(which will be odd).

Suppose, then, that $n \geqslant 1$ and π transposes i and j
where $i < j$. Consider the face of σ not containing the
vertex v^j. This occurs in the formula for $\partial(v^0 \dots v^n)$ as

$$(-1)^j (v^0 \dots v^i \dots \hat{v^j} \dots v^n) = (-1)^j \tau$$

say and in the formula for $\partial(v^{\pi(0)} \dots v^{\pi(n)})$ as

$$(-1)^i (v^0 \dots \hat{v^j} \dots v^i \dots v^n) = (-1)^i \tau' \quad \text{say.}$$

Now moving v^i to the left $j - i - 1$ places turns τ' into
τ. This amounts to effecting $j - i - 1$ transpositions, so

$$(-1)^i \tau' = (-1)^i (-1)^{j-i-1} \tau = -(-1)^j \tau \, .$$

Similarly the face of σ not containing v^i occurs in the two
formulae with opposite sign, while a face of σ containing

both v^i and v^j occurs with the same power of -1 but with v^i and v^j interchanged - hence altogether with opposite sign. This completes the verification.

4.5 Definition   The <u>augmentation</u>

$$\epsilon : C_0(K) \rightarrow \mathbb{Z}$$

is the homomorphism defined by

$$\epsilon(\Sigma\lambda_i \sigma_0^i) = \Sigma\lambda_i$$

(compare 1.22).

The sequences (not in general exact)

$$\ldots 0 \rightarrow 0 \rightarrow C_n(K) \overset{\partial}{\rightarrow} C_{n-1}(K) \rightarrow \ldots \rightarrow C_1(K) \overset{\partial}{\rightarrow} C_0(K) \rightarrow 0 \rightarrow 0 \ldots$$

and

$$\ldots 0 \rightarrow 0 \rightarrow C_n(K) \overset{\partial}{\rightarrow} C_{n-1}(K) \rightarrow \ldots \rightarrow C_1(K) \overset{\partial}{\rightarrow} C_0(K) \overset{\epsilon}{\rightarrow} \mathbb{Z} \rightarrow 0 \ldots$$

are called the <u>chain complex</u> of K and the <u>augmented chain complex</u> of K, respectively.

The main property of the boundary homomorphisms is the following.

4.6 <u>Proposition</u>   For any p, the homomorphism

$$\partial \circ \partial = \partial_{p-1} \circ \partial_p : C_p(K) \rightarrow C_{p-2}(K)$$

is trivial. Also $\epsilon \circ \partial_1 : C_1(K) \rightarrow \mathbb{Z}$ is trivial.

<u>Proof</u>   The last statement is easy to check; for the first one it is enough to prove that, for any p-simplex $\sigma = (v^0 \ldots v^p)$ of K, with $p \geqslant 2$,

$$\partial_{p-1}(\partial_p \sigma) = 0.$$

Now

$$\partial_{p-1}(\partial_p \sigma) = \sum_{i=0}^{p} (-1)^i \partial_{p-1}(v^0 \ldots \hat{v}^i \ldots v^p).$$

This is a sum of $p(p + 1)$ terms in which each $(p - 2)$-dimensional face of σ occurs exactly twice. In fact, consider the $(p - 2)$-dimensional face $(v^0 \ldots \hat{v^i} \ldots \hat{v^j} \ldots v^p)$, where $i < j$. This occurs once in

$$(-1)^i \partial_{p-1}(v^0 \ldots \hat{v^i} \ldots v^p), \text{ with sign } (-1)^i(-1)^{j-1}$$

and once in

$$(-1)^j \partial_{p-1}(v^0 \ldots \hat{v^j} \ldots v^p), \text{ with sign } (-1)^j(-1)^i.$$

The total coefficient of this $(p - 2)$-simplex in $\partial_{p-1}(\partial_p \sigma)$ is therefore $(-1)^i(-1)^{j-1} + (-1)^j(-1)^i = 0$.   Q.E.D.

4.7 Corollary   For any p,

$$\text{Im } \partial_{p+1} \subset \text{ker } \partial_p .$$

Also $\text{Im } \partial_1 \subset \text{ker } \epsilon .$

HOMOLOGY GROUPS

4.8 Definitions   Consider the sequence

$$C_{p+1}(K) \xrightarrow{\partial_{p+1}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K).$$

Ker ∂_p is denoted by Z_p, or $Z_p(K)$, and elements of Z_p are called p-cycles (compare 1.19). Z_p is a free f.g. abelian group.   (See A.33)

Im ∂_{p+1} is denoted by B_p, or $B_p(K)$, and elements of B_p are called p-boundaries (we also say that these elements bound). B_p is a free f.g. abelian group.

Thus the first statement of 4.7 can be summed up by the motto: "every boundary is a cycle". The converse, "every cycle bounds" is false in general - see the discussion preceding 4.4.

Since $B_p \subset Z_p$ there is a quotient group

$H_p(K) = Z_p(K)/B_p(K)$, called the pth <u>homology group</u> of K.
It measures the extent to which K has non-bounding p-cycles
i.e. the extent to which K has "p-dimensional holes" - com-
pare the Introduction, where some examples with p = 1 were
given. Note that $H_p(K)$ is a finitely generated abelian
group.

The quotient group Ker $\epsilon/B_0(K)$ is denoted by $\tilde{H}_0(K)$ and
is called the <u>reduced</u> 0th <u>homology group</u> of K.

<u>4.9</u> <u>Further terminology and remarks</u> (1) Let K^r denote the
r-skeleton of K - that is the subcomplex of K consisting of
all simplexes of dimension $\leqslant r$. Then clearly $C_p(K) = C_p(K^r)$
and $Z_p(K) = Z_p(K^r)$ provided $r \geqslant p$, and $H_p(K) = H_p(K^r)$
provided $r \geqslant p + 1$.

(2) Taking r = 1 in (1) we have $Z_1(K) = Z_1(K^1)$.
Since K^1 is an oriented graph we can use the method of Chap-
ter 1 to write down a basis for $Z_1(K^1)$. Explicitly, when K
is connected, a basis for $Z_1(K) = Z_1(K^1)$ is obtained as
follows (compare 1.20): choose a maximal tree T in K^1 and
let $e^1, \ldots, e^\mu$ be the edges of K not in T. Let
z^i (i = 1, $\ldots$, μ) be the 1-cycle associated to the unique
loop on T + e^i, the coefficient of e^i in z^i being +1.
Then $z^1, \ldots, z^\mu$ is the required basis. If K is not con-
nected, then a basis is given by finding a basis for the cycles
in each component.

Notice that, according to 1.23, $\tilde{H}_0(K) = 0$ when K is
connected.

(3) If K is n-dimensional, then $C_{n+1}(K) = 0$ so that
$B_n(K) = 0$ and $H_n(K) = Z_n(K)/0 \cong Z_n(K)$. Perhaps it is
being too pedantic to distinguish between Z_n and $Z_n/0$, so
let us write $H_n(K) = Z_n(K)$. Note that Z_n is free abelian
since it is a subgroup of C_n (see A.33): thus <u>top dimensional</u>
<u>homology groups are free abelian.</u>

In the case n = 1, i.e. K is an oriented graph, we have

$H_1(K) = Z_1(K)$ so that, if K is connected, $H_1(K) \cong \mathbb{Z}^{\mu(K)}$
and free generators of $H_1(K)$ are given by the basic cycles
described in (2) above. These are the "independent loops" or
"independent holes" in the graph.

For any K of dimension n, we have $H_p(K) = 0$ for
$p < 0$ and for $p > n$.

4.9 (4) An element of $H_p(K)$ is a coset $\bar{z} = z + B_p(K)$
where $z \in Z_p(K)$. We shall use the notation $\{z\}$ for this
coset; it is called the __homology class__ of the cycle z. (If
several simplicial complexes are under discussion at the same
time, the notation $\{z\}_K$ will be used.) Any cycle z' in
$\{z\}$ - that is, any cycle $z' = z + b$ for $b \in B_p(K)$ - is
called a __representative cycle__ for $\{z\}$. We say that z and
z' are __homologous__ if $z' - z \in B_p(K)$, and write $z \sim z'$;
in particular if $z \in B_p(K)$ (i.e. if z bounds) we write
$z \sim 0$: z is homologous to zero. Thus $z \sim z' \iff \{z\} = \{z'\}$.
An element of $C_p(K)$ which is not a cycle does not represent
any homology class. Nevertheless we could (but won't) con-
sider two chains to be homologous if their difference bounds.
Note that a chain which is homologous to zero is necessarily a
cycle, by 4.7.

(5) As in the case of graphs (see 1.21 (3)) the particular
orientation given to a simplicial complex does not affect the
homology groups up to isomorphism. The orientation is intro-
duced in order to define the boundary homomorphism; neverthe-
less a theory does exist for unoriented simplicial complexes
and this theory is explained in Chapter 8. For simplicial
complexes of dimension > 1 the unoriented theory is a weaker
tool than the oriented theory.

Zero[th] homology groups   Let K be an oriented simplicial
complex; the object is to calculate $H_0(K)$. A vertex v of
K can also be regarded as a 0-cycle (with the coefficient of
v equal to 1 and other coefficients zero), and there is a

simple geometrical interpretation of homology between vertices:

<u>4.10 Proposition</u> $v \sim v'$ $\iff$ v and v' lie in the same
component of K. (See 3.17 (4))

<u>Proof</u>   If v and v' lie in the same component of K,
then a path from v' to v will have an associated 1-chain
(1.18) c with the property $\partial c = v - v'$; consequently
$v \sim v'$. For the converse, we have to use $v \sim v'$ to construct
a path from v' to v. The reader can probably convince him-
self that this is possible by drawing a few pictures: alter-
natively here is a sketch of a proof by induction.   Suppose
that $v - v' = \partial c$, and let us write, temporarily, $|c|$ for
the sum of the moduli of the coefficients in c; the induction
is on $|c|$ and the case $|c| = 1$ is left to the reader.
Suppose, to simplify notation, that all 1-simplexes of K with
v as an end-point are oriented towards v. Then c must
contain at least one such 1-simplex, e say, with positive
coefficient. Writing $e = (v''v)$, we have $|c - e| < |c|$
and $\partial(c - e) = v'' - v'$ so that by induction there is a path
from v' to v'', and hence via e to v.   Q.E.D.

Now let K be a non-empty oriented simplicial complex with
say k components $K_1, \ldots, K_k$, and let v^i be a vertex in
K_i for $i = 1, \ldots, k.$

<u>4.11 Proposition</u> $H_0(K)$ is freely generated by the homology
classes $\{v^i\}$ for $i = 1, \ldots, k.$ Thus $H_0(K) \cong \mathbb{Z}^k$ (also
$H_0(\emptyset) = 0$ since $C_p(\emptyset) = 0$ for all p).

<u>Proof</u>   The implication "$\Leftarrow$" of 4.10 shows that
$\{v^1\}, \ldots, \{v^k\}$ generate $H_0(K) = C_0(K)/B_0(K)$ (note that
$Z_0(K) = C_0(K)$). To show that they are free generators sup-
pose $\lambda_1 v^1 + \ldots + \lambda_k v^k = \partial c$ for some $c \in C_1(K)$.   Write
$c = c^1 + \ldots + c^k$ where c^i contains only 1-simplexes in K_i
$(i = 1, \ldots, k)$. Then ∂c^i contains only vertices in K_i,
so that we deduce $\lambda_i v^i = \partial c^i$ for $i = 1, \ldots, k.$ Taking

the augmentation (4.5) of both sides, and using (4.6) $\epsilon \circ \partial = 0$
we find $\lambda_i = 0$ for $i = 1, \ldots, k$.   Q.E.D.

This shows that $H_0(K)$ measures a very fundamental geo-
metrical property of K, namely the number of separate pieces
(components) it has.

4.12 Remark   In fact it is always true that, in the above
notation, $H_p(K) \cong H_p(K_1) \oplus \ldots \oplus H_p(K_k)$ for all p. This
is a special case of a theorem (7.2) to be proved later on homo-
logy groups of unions of not necessarily disjoint simplicial
complexes. It is also not hard to give a direct proof (start
with $k = 2$).

Now consider the sequence

$$C_1(K) \xrightarrow{\partial} C_0(K) \xrightarrow{\epsilon} \mathbb{Z} \ .$$

We know that $\operatorname{Im} \partial \subset \operatorname{Ker} \epsilon \subset C_0(K)$ by 4.6, and applying
A.16, Ex(7), this gives a short exact sequence

$$0 \to \frac{\operatorname{Ker} \epsilon}{\operatorname{Im} \partial} \to \frac{C_0}{\operatorname{Im} \partial} \to \frac{C_0}{\operatorname{Ker} \epsilon} \to 0$$

i.e.

$$0 \to \widetilde{H}_0(K) \to H_0(K) \to \frac{C_0}{\operatorname{Ker} \epsilon} \to 0 \ .$$

Now suppose that K is non-empty; then ϵ is epic and
$C_0/\operatorname{Ker} \epsilon \cong \operatorname{Im} \epsilon = \mathbb{Z}$. Hence the sequence is split (A.19),
and the following result holds.

4.13 Proposition   For a non-empty K,

$$H_0(K) \cong \widetilde{H}_0(K) \oplus \mathbb{Z} \ .$$

Also $H_0(\emptyset) = \widetilde{H}_0(\emptyset) = 0 \ .$

Note that this confirms that, for a connected K,
$\widetilde{H}_0(K) = 0$ (4.9 (2)).   Indeed the main reason for introducing

$\tilde{H}_0(K)$ is that it is often 0 and this makes calculations with exact sequences more straightforward. In the notation of 4.11, the classes represented by $v^i - v^1$ for $i = 2, \ldots, k$ freely generate $\tilde{H}_0(K)$, when $k \geqslant 2$.

First homology groups   Given an oriented connected simplicial complex K we now seek to calculate $H_1(K)$. When K is a graph (dim $K = 1$) there is no problem: we just calculate $Z_1(K)$ as in Chapter 1 (see 4.9 (2) and (3)). Thus in this case $H_1(K) \cong \mathbb{Z}^{\mu(K)}$ with generators $\{z^1\}, \ldots, \{z^\mu\}$.

In general there is a nontrivial subgroup $B_1(K)$ of $Z_1(K)$ to factor out:

$$C_2(K) \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K)$$

$$H_1(K) = \text{Ker } \partial_1/\text{Im } \partial_2 = Z_1(K)/B_1(K).$$

To calculate this quotient in a particular instance we use the method of presentations; see the Appendix but note that the examples we actually meet are very simple and require only a minimum of technique. The example in A.25 is a typical calculation. First we find a basis for $Z_1(K) = Z_1(K^1)$ by the method of 4.9(2), then we find generators for $B_1(K)$. The generators we use are $\partial_2 t$ where t runs over the oriented 2-simplexes of K. A presentation of $H_1(K)$ is, when K is connected,

4.14
$$\left(\begin{array}{l} \text{cycles } z^1, \ldots, z^\mu, \;\; \mu = \mu(K^1), \text{ associated to} \\ \qquad \text{basic loops on } K^1 \text{ (see 4.9 (2))} \\ \\ \text{a relation } \partial_2 t = 0 \text{ for each oriented} \\ \qquad \text{2-simplex of } K. \end{array} \right.$$

Direct calculations using 4.14 are rather tedious, and we shall not need to make many of them. Here is an example.

4.15 Möbius Band   The diagram shows a triangulation K of the Möbius band, and the heavily drawn edges are ones whose removal leaves a maximal tree in K^1. It is a good idea to
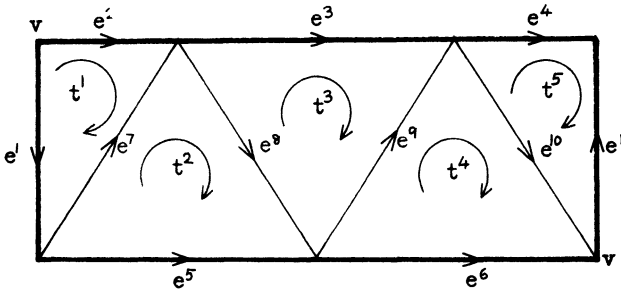
label these edges e^1, ..., e^μ (in this example $\mu = 6$) and
the remaining edges $e^{\mu+1}$, ..., e^α $(\alpha = \alpha_1(K))$. The reason
for this is as follows. Suppose that z is any 1-cycle on K;
then by 1.20

$$\underline{4.16} \quad \left(\begin{array}{l} z = \lambda_1 e^1 + \ldots + \lambda_\mu e^\mu + \lambda_{\mu+1} e^{\mu+1} + \ldots + \lambda_\alpha e^\alpha \\[2mm] \Rightarrow \quad z = \lambda_1 z^1 + \ldots + \lambda_\mu z^\mu \end{array} \right.$$

where z^i $(i = 1, \ldots, \mu)$ is the 1-cycle associated with the
unique loop on (maximal tree + e^i), the 1-cycle being oriented
so that it contains e^i with coefficient $+ 1$. This is of
great assistance when expressing $\partial_2 t$ as a linear combination
of z^1, ..., z^μ.

   Let z^1, ..., z^6 be as above. Thus
$z^1 = e^1 + e^7 + e^8 + e^9 + e^{10}$, $z^2 = e^2 + e^8 + e^9 + e^{10}$ etc.,
but there is no need to write them all out. $B_1(K)$ is gen-
erated by the following elements of $Z_1(K)$

$$\partial_2 t^1 = e^2 - e^7 - e^1 = z^2 - z^1 \qquad \text{(see 4.16 above)}$$

$$\partial_2 t^2 = e^7 + e^8 - e^5 = -z^5$$

$$\partial_2 t^3 = e^3 - e^9 - e^8 = z^3$$

$$\partial_2 t^4 = e^9 + e^{10} - e^6 = -z^6$$

$$\partial_2 t^5 = e^4 - e^1 - e^{10} = z^4 - z^1 .$$

Thus $H_1(K)$ has presentation

$$\left(\begin{array}{l} z^1,\; z^2,\; z^3,\; z^4,\; z^5,\; z^6 \\ z^2 - z^1 \;=\; -z^5 \;=\; z^3 \;=\; -z^6 \;=\; z^4 - z^1 \;=\; 0 \;. \end{array}\right.$$

(Remember that the coset symbol $\{\ \}$ is conventionally omitted when writing down presentations; see A.21.)   This shows that all the generators can be expressed in terms of z^1, and (compare A.24), $H_1(K)$ is isomorphic to the abelian group with presentation

$$\left(\begin{array}{l} z^1 \\ \\ \\ \end{array}\right. \qquad\qquad \cong\ \mathbb{Z}$$

Furthermore it shows that $\{z^1\}_K$ generates $H_1(K)$.   Using the relations alternative generators are $\{z^2\}$, or $\{z^4\}$, or $\{z^2 + z^3\}$ ... .

Notice that the cycle z^1 goes "once round" the Möbius band.   Thus the result says that a 1-cycle going "once round" does not bound anything, and that any 1-cycle is, up to homology, a multiple of this one.   For example $e^2 + e^3 + e^4 + e^5 + e^6$ is a 1-cycle going round the rim of the band, and

$$\begin{aligned} \{e^2 + e^3 + e^4 + e^5 + e^6\} &= \{z^2 + z^3 + z^4 + z^5 + z^6\} \qquad \text{by 4.16} \\ &= \{z^2\} + \{z^3\} + \{z^4\} + \{z^5\} + \{z^6\} \\ &= \{z^1\} + 0 + \{z^1\} + 0 + 0 \\ &\qquad\qquad\qquad \text{from the presentation of } H_1(K) \\ &= 2\{z^1\}. \end{aligned}$$

Indeed it is geometrically clear that the rim is a circle going twice round before joining up.   The following picture shows more informally that the rim a is homologous to twice a circle b going round the middle of the band.
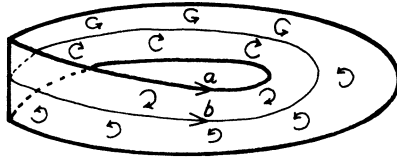
The orientations shown are coherent (2.3) except across the
centre line b.   Thus   a - 2b bounds the whole oriented area
(i.e. if the area were broken up into triangles, oriented as
shown, the sum of their boundaries would be   a - 2b), so that
a $\sim$ 2b.

We known from 4.11 that $H_0(K) \cong \mathbb{Z}$, with generator the
homology class of any vertex.   To find $H_2(K)$ we need
Ker $(\partial_2 : C_2(K) \to C_1(K))$, since $C_3(K) = 0$.   Now a general
element of $C_2(K)$ is $c = \sum_{i=1}^{5} \lambda_i t^i$; suppose that $\partial_2 c = 0$.
The edge e^2 occurs only in $\partial_2 t^1$; hence t^1 must have
coefficient zero in c.   Similar arguments show that t^2, t^3,
t^4, t^5 all have coefficient zero in c, so that
$c = 0 : H_2(K) = 0$.   An alternative way of doing this is to
write down ∂c explicitly, using $\partial t^1 = e^2 - e^7 - e^1$ etc.,
and to deduce from $\partial c = 0$ that $\lambda_1 = \ldots = \lambda_5 = 0$.

Thus $H_0(K) \cong \mathbb{Z}$ generator $\{v\}$

$H_1(K) \cong \mathbb{Z}$ generator $\{e^1 + e^7 + e^8 + e^9 + e^{10}\}$

$H_2(K) = 0$.

4.17 Examples   (1X) Triangulate the cylinder (3.22(3)) and
show that the homology groups are

$H_0 \cong \mathbb{Z}$; $H_1 \cong \mathbb{Z}$ generator a cycle going round
one end of the cylinder

$$H_2 = 0.$$

Thus up to isomorphism the groups are the same as for the Möbius band. This fact becomes obvious by the method of collapsing (3.30) applied to homology groups (6.6). See 6.7(1).

4.17(2X) Show that, for any p,

$$C_p(K) \cong Z_p(K) \oplus B_{p-1}(K).$$

(Hint. Form a short exact sequence from $\partial_p : C_p(K) \rightarrow C_{p-1}(K).$)

(3X) Using $B_p(K) \subset Z_p(K) \subset C_p(K)$ and A.16, Ex. (7) show that

$$C_p(K)/B_p(K) \cong H_p(K) \oplus B_{p-1}(K) .$$

This shows that the torsion subgroup of $H_p(K)$ (the subgroup of $H_p(K)$ consisting of those elements which are of finite order) is isomorphic to the torsion subgroup of C_p/B_p (since B_{p-1} is free abelian). In particular the torsion subgroup of $H_p(K)$ is determined by the homomorphism $\partial_{p+1} : C_{p+1}(K) \rightarrow C_p(K).$
Calculate the group $C_1(K)/B_1(K)$ when K is an n-sided polygon. Note that the result depends on n, whereas $H_1(K) \cong \mathbb{Z}$ for any n since $\mu(K) = 1$ (see 4.9 (3)).

(4X) Calculate the homology groups of the torus given in 2.2(3), or in 3.22(1). (Of course you must first orient the 1- and 2-simplexes, but it doesn't matter how - see 4.9(5).) This homology calculation is about as complicated as one normally wants to do "from first principles". Verify that $H_0 \cong \mathbb{Z}$; $H_1 \cong \mathbb{Z} \oplus \mathbb{Z}$ with generators $\{a + b + c\}$ and $\{d + e + f\}$ in 2.2(3); $H_2 \cong \mathbb{Z}$ with generator the sum of all triangles of the torus, coherently oriented.

4.17(5X) Orient the simplicial complex shown, which is a pro-

jective plane (2.7(3)), and verify that the homology groups are $H_0 \cong \mathbb{Z}$; $H_1 \cong \mathbb{Z}_2$ with generator $\{a + b + c\}$; $H_2 = 0$.
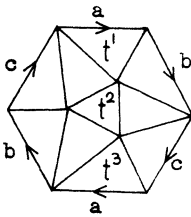
Note that H_1 is a finite group here; in fact $2\{a + b + c\} = 0$, i.e. $2(a + b + c) \sim 0$. This can be seen directly by adding up the boundaries of the ten triangles in the picture, all oriented clockwise. (The three labelled tri-angles are for later reference.)

We shall not have to calculate, directly from the defini-tion, many homology groups $H_p(K)$ for $p > 1$. There is one case of interest, however, in connexion with the orientability of closed surfaces.

4.18 Proposition   Let K be a closed surface (2.1 or 3.29 (2)).   Then

$$H_2(K) \cong \begin{cases} 0 & \text{if } K \text{ is non-orientable} \quad (2.3) \\ \mathbb{Z} & \text{if } K \text{ is orientable.} \end{cases}$$

A generator for $H_2(K)$, when K is orientable, is given by the homology class of the sum of all 2-simplexes of K, co-herently oriented.

To be precise: K is already oriented, and, apart from this, the 2-simplexes can be coherently oriented. The gener-ator of $H_2(K)$ is $\Sigma(\pm t^i)$, the sum being over all oriented 2-simplexes t^i of K, with sign + or − according as the orientation of t^i as a simplex of K does or does not agree with that in the chosen coherent orientation. Note that $H_2(K) = Z_2(K)$ since K is 2-dimensional (see 4.9(3)). A generator of $H_2(K)$, when K is an orientable closed surface, is called a _fundamental cycle_, or _fundamental class_, for K.

Up to sign it is determined by K.

Proof   Suppose first that K is orientable. For sim-
plicity let us suppose that the orientation of K is such that
the 2-simplexes receive orientations which are already coherent.
We seek $Z_2(K)$ (compare 4.9(3)), so let $z \in Z_2(K)$, and suppose
$z \neq 0$ so that some oriented 2-simplex t occurs with coeffi-
cient $\lambda \neq 0$ in z. Let t' be any other oriented 2-simplex.
Then (see 2.2(5)) there is a sequence of triangles of K
connecting t and t', two consecutive members of the sequence
always having an edge in common. From this and $\partial z = 0$ it
follows that every triangle in the sequence, and in particular
t', also has coefficient λ in z. Hence z is a multiple
of the sum of all triangles, coherently oriented, and $Z_2(K)$
is infinite cyclic with this sum as generator.

Now suppose that $H_2(K) \neq 0$ and let z be a non-zero
element of $Z_2(K)$. Exactly as above we can deduce that the
coefficient of each oriented 2-simplex in z must be, up to
sign, the same non-zero integer (only up to sign because the
triangles of K are no longer assumed coherently oriented).
Now define a new orientation of the triangles of K by changing
the orientation of those with negative coefficient in z. It
is easy to see (using $\partial z = 0$) that this coherently orients the
triangles of K : K is orientable.   Q.E.D.

Thus H_2 distinguishes algebraically between orientable
and non-orientable surfaces. Fundamental cycles for the tori
drawn in 2.2(3) and 3.22(1) are obtained by orienting all the
triangles in the pictures clockwise, or alternatively all anti-
clockwise. Of course this doesn't work for the Klein bottle
or projective plane of 2.2(4), for such an orientation will not
be coherent across, for example, the edge d.

RELATIVE HOMOLOGY GROUPS

In Chapter 9 we shall investigate some of the properties of a
graph which is realized as a subcomplex of a surface. This is
a situation involving a pair of simplicial complexes - or, as we
shall say, a relative situation - and calls for algebraic tools
which involve both complexes at once. In particular consider
the graph in a torus drawn below. There is a region bounded by
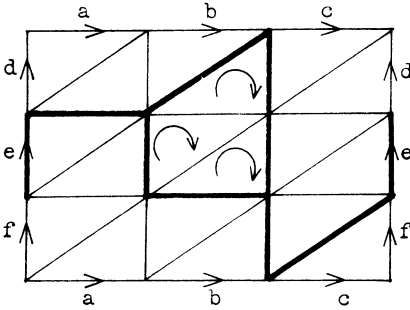
five edges and contain-
ing the three oriented
triangles. The sum of
the boundaries of these
three triangles is cer-
tainly not zero, but it
does contain only edges
in the graph, and the
sum of the three trian-

gles is called a relative cycle. Thus relative cycles have a
connexion with regions, and this is exploited in Chapter 9.

4.19 Definition   Let K be an oriented simplicial complex of
dimension n and L be a subcomplex of K (L is therefore orien-
ted in a natural way by K). For any p, $0 \leqslant p \leqslant n$, the pth
chain group of K modulo L, or the pth relative chain group (of
the pair (K, L)) is the subgroup of $C_p(K)$ consisting of those
p-chains on K in which the coefficient of every simplex of L is
zero. It is denoted $C_p(K, L)$, and an element of $C_p(K, L)$ is
called a p-chain of K modulo L, or a relative p-chain. (There
is always a choice of terminology between "relative" and "modulo".
In the definitions below we shall usually give the former and
leave the reader to give expression to the latter. Also "modulo",
when it occurs, will always be abbreviated to "mod".) For p > n
or p < 0, $C_p(K, L)$ is defined to be zero.
        Note that $C_p(K, \emptyset) = C_p(K)$. The following definition
of "relative boundary" also reduces to the old definition of

boundary when $L = \emptyset$.

4.20 <u>Definition</u>   With the notation of 4.19, we define

$$j = j_q : C_q(K) \rightarrow C_q(K, L)$$

to be the homomorphism, which changes to zero (or leaves at zero) the coefficient of every simplex in L (for $q < 0$ or $q > n$, $j_q = 0$).   Now define

$$\partial = \partial_p : C_p(K, L) \rightarrow C_{p-1}(K, L)$$

by $\partial c = j_{p-1}(\partial_p c)$ for all $c \in C_p(K, L)$.

∂ is called the <u>relative boundary homomorphism</u>.   Thus if i_p denotes the inclusion of $C_p(K, L)$ in $C_p(K)$, then $\partial_p = j_{p-1} \circ \partial_p \circ i_p$.   We may say that j "forgets L" and then the relative boundary of a chain is the ordinary boundary followed by forgetting L.

4.21 <u>Proposition</u>   For any p, the homomorphism

$$\partial \circ \partial = \partial_{p-1} \circ \partial_p : C_p(K, L) \rightarrow C_{p-2}(K, L)$$

is trivial.

The proof is similar to that of 4.6 and will be omitted.

The remaining definitions are precisely analogous to those for the "absolute" case $L = \emptyset$.   They are summarized below.

4.22 <u>Definitions</u>   <u>Chain complex</u> of (K, L):

$$\ldots \; 0 \rightarrow C_n(K, L) \overset{\partial}{\rightarrow} C_{n-1}(K, L) \overset{\partial}{\rightarrow} \ldots$$

$$\overset{\partial}{\rightarrow} C_1(K, L) \overset{\partial}{\rightarrow} C_0(K, L) \rightarrow 0.$$

(There is no augmented relative chain complex.)

Group of <u>relative p-cycles</u> $= Z_p(K, L) = \text{Ker } \partial_p$.

Group of <u>relative p-boundaries</u> $= B_p(K, L) = \text{Im } \partial_{p+1}$.

By 4.21, $B_p(K, L) \subset Z_p(K, L)$;   the quotient group

$$H_p(K, L) = Z_p(K, L)/B_p(K, L)$$

is the pth relative homology group of the pair (K, L).   All
these reduce to the previous definitions when $L = \emptyset$.

If $z \in Z_p(K, L)$ then $z + B_p(K, L) \in H_p(K, L)$ is denoted
by $\{z\}$ or $\{z\}_{K,L}$ and is called the relative homology class
of z.   Any element of $\{z\}$, i.e. any $z + b$ for
$b \in B_p(K, L)$, is called a representative relative cycle for
$\{z\}$.   Two relative cycles z and z' are said to be homo-
logous mod L (written $z \overset{L}{\sim} z'$) if $\{z\} = \{z'\}$.

4.23 Remarks and Examples   (1) If $K = L$ then all the
groups and homomorphisms are trivial.   If dim $K = n$, then
$H_p(K, L) = 0$ for $p < 0$ or $p > n$, and $H_n(K, L) = Z_n(K, L)$
is free abelian.

(2) $H_p(K, L)$ is not in general a subgroup of $H_p(K)$.
For example let K be a 1-simplex $e = (vw)$ together with
its end-points and let L be the end-points.   Then $H_1(K) = 0$
since K is a tree (see
the preamble to 4.14), but
$H_1(K, L) \cong \mathbb{Z}$, with gen-
erator $\{e\}$.   For
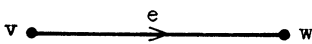
$C_0(K, L) = 0$, so $\partial_1 = 0$ and $Z_1(K, L) = C_1(K, L)$ is
infinite cyclic with generator e.

(3) More generally than (2), let K be any oriented
graph and $L = K^0$, the set of vertices of K.   Then
$C_0(K, L) = 0$, $\partial_1 = 0$ and $Z_1(K, L) = C_1(K, L)$ is a free
abelian group with one generator for each edge of K.

If K is any oriented connected graph and L is a maximal
tree (1.10) for K, then L contains all the vertices of K
and so $C_0(K, L) = 0$, $\partial_1 = 0$ and $Z_1(K, L) = C_1(K, L)$ is a
free abelian group with one generator for each edge of K not
in L - hence of rank $\mu(K)$.

<u>4.23</u> (4) Let K^r denote the r-skeleton of K.   Then the
chain complex of (K^r, K^{r-1}) is

$$\dots \to 0 \to C_r(K^r, K^{r-1}) \to 0 \to \dots$$

with only one non-zero group.   Thus $H_p(K^r, K^{r-1})$ is zero
when $p \neq r$, and is free abelian of rank $\alpha_r(K)$ (the number
of r-simplexes of K) when $p = r$.   As an example, let K
be a 2-simplex and all its faces, and $r = 2$, so that $K^r = K$,
K^{r-1} = the set of edges and vertices of K.   Then $H_2(K, K^1)$
is free abelian of rank 1, generated by the homology class of
the 2-simplex.   In this case $H_1(K, K^1) = 0 = H_0(K, K^1)$.

The next proposition enables relative cycles and boundaries
to be recognized as such without undue difficulty.   It is con-
venient, now and later, to make the following convention.
When L is a subcomplex of K, we regard $C_p(L)$ as a sub-
group of $C_p(K)$, namely the subgroup consisting of those
p-chains on K in which the coefficient of every p-simplex
not in L is zero.   (Thus, with this convention, $C_p(K)$ is
the internal direct sum of $C_p(L)$ and $C_p(K, L)$.)   In order
that the convention shall not lead us into error, it is essen-
tial that the boundary of a p-chain on L should be the same
whether we regard it in its own right or according to the con-
vention as a special kind of p-chain on K.   That this is in-
deed the case follows from the fact that L is a subcomplex
of K and that therefore if $\sigma \in L$ then all the faces of σ
also belong to L.

<u>4.24</u> <u>Proposition</u>   (1) Let, in the usual notation,
$c \in C_p(K, L)$.   Then c is a relative cycle, i.e. belongs to
$Z_p(K, L)$, if and only if $\partial c \in C_{p-1}(L)$.

(2) Let $c \in C_p(K, L)$.   Then c is a relative boundary,
i.e. $c \in B_p(K, L)$, if and only if there is an (absolute)
p-chain $c' \in C_{p+1}(K)$ such that $\partial c' - c \in C_p(L)$.

Proof   (1) is virtually immediate from the definitions
(and the above convention), so the proof will be omitted.
(2) can be a very confusing statement to verify, so here is a
full proof.

Suppose that $c \in B_p(K, L)$, i.e. $c = \partial c'$ for some
$c' \in C_{p+1}(K, L)$. We show that this c' (which does of course
belong to $C_{p+1}(K)$) satisfies $\partial c' - c \in C_p(L)$. In fact

$$jc = c = \partial c' = j\partial c', \quad \text{so that} \quad j(c - \partial c') = 0 .$$

From the definition of j, this implies $\partial c' - c \in C_p(L)$.

Suppose conversely that c', as given in (2), exists.
Then let $c'' = jc' \in C_{p+1}(K, L)$. We show $\partial c'' = c$, which
proves that $c \in B_p(K, L)$.

We shall use the following simple but useful trick. Let
$\rho : C_{p+1}(K) \to C_{p+1}(L)$ denote the "restriction" homomorphism
which puts at zero the coefficient of any simplex not in L.
Then clearly, for any $c' \in C_{p+1}(K)$, $c' = jc' + \rho c'$.

We are given $j(\partial c' - c) = 0$; also $jc = c$ since
$c \in C_p(K, L)$. Hence

$$c = jc = j\partial c' = j\partial jc' + j\partial \rho c' = j\partial jc' \quad \text{since} \quad \partial \rho c' \in C_p(L)$$
$$= \partial c'' \text{ by definition. Q.E.D.}$$

Zeroth Relative homology groups   As an example of 4.24, let
K be connected and L non-empty. We show $H_0(K, L) = 0$,
i.e. $B_0(K, L) = C_0(K, L)$. Let v be a vertex of K not
in L (if no such vertex exists then the result is true since
$C_0(K, L) = 0$), and w a vertex of L. Join w to v by a
path in K; the associated 1-chain (1.18) c' has
$\partial c' = v - w$, so that $\partial c' - v \in C_0(L)$, i.e. (by 4.24)
$v \in B_0(K, L)$. Hence every vertex of K not in L is a rela-
tive boundary, which implies the result.

The general result on $H_0(K, L)$ is as follows. The
details of proof are left to the reader.

4.25 Proposition   Let K be an oriented simplicial complex
and L a subcomplex of K.   Then $H_0(K, L) = 0$ provided L
has at least one vertex in every component of K, but $H_0(K, L)$
acquires an infinite cyclic summand for every component of K
not containing any vertex of L.   Free generators for $H_0(K, L)$
are the relative homology classes of one vertex from each of
these latter components.

4.26 Remark   (compare 4.12)   Let K have components
$K_1, \ldots, K_k$ and write $L_1 = L \cap K_1, \ldots, L_k = L \cap K_k$.   Then
it can be shown that, for all p,

$$H_p(K, L) \cong H_p(K_1, L_1) \oplus \ldots \oplus H_p(K_k, L_k).$$

First relative homology groups   To calculate $H_1(K, L)$ we
need a basis for $Z_1(K, L)$ and generators for $B_1(K, L)$.   The
method of presentations will then give the result as in the
absolute case.   The easiest case is when all the vertices of
K belong to L, for then $C_0(K, L) = 0$ so that
$Z_1(K, L) = C_1(K, L)$ and a basis is given by the edges of K
not in L.   Generators for $B_1(K, L)$ are simply $\partial_2 t$ where
t runs through the oriented triangles of K not in L.

For completeness, here is a method for writing down a basis
of $Z_1(K, L)$ in general, followed by a very brief sketch of
the proof.   We assume K is connected (compare 4.26) to
simplify the statement.

Let K be a connected oriented simplicial complex and L
a subcomplex of K, with components $L_1, \ldots L_\ell$ say.   Let
T_α be a maximal tree in the 1-skeleton of L_α $(\alpha = 1, \ldots, \ell)$
and let T be a maximal tree in the 1-skeleton of K which
contains $T_1, \ldots, T_\ell$ (compare 1.12(1)).   For each edge e^i
of K which is in neither T nor L let z^i be the 1-cycle
on K given by the unique loop on $T + e^i$ such that the
coefficient of e^i in z^i is +1.   Thus the number of z^i is

$$m = \mu(K) - \mu(L_1) - \ldots - \mu(L_\ell).$$

Next let v^1, ..., v^ℓ be vertices of K in, respectively, the components L_1, ..., L_ℓ of L, and let c^α ($\alpha = 2$, ..., ℓ) be the 1-chain associated to some path in K from v^1 to v^α. (Thus if $\ell = 0$ or $\ell = 1$ there is no c^α to consider.)

*4.27 Proposition   The following elements of $Z_1(K, L)$ are free generators:

$$jz^1, \ldots, jz^m, \; jc^2, \ldots, jc^\ell .$$

<u>Sketch of proof</u>   The proof depends on the following sequence.

$$0 \to Z_1(L) \xrightarrow{i} Z_1(K) \xrightarrow{j} Z_1(K, L) \xrightarrow{\partial_*} H_0(L) \xrightarrow{i_*} H_0(K) \to 0$$

where i = inclusion, j is the j of 4.20 restricted to $Z_1(K)$, $\partial_* z = \{\partial z\}_L$, $i_*\{w\}_L = \{w\}_K$. The sequence is a special case of the "homology sequence of a pair" (see 6.3(6)) and is exact, provided L is non-empty. The case $L = \emptyset$ ($\ell = 0$) is covered by 1.20. The proof of exactness will be left to the reader - or he can wait for the general theorem. From the exact sequence we deduce two short exact sequences:

$$0 \to \text{Im } j \xrightarrow{i} Z_1(K, L) \xrightarrow{\partial_*} \text{Im } \partial_* \to 0 \qquad (*)$$

and $\qquad 0 \to Z_1(L) \xrightarrow{i} Z_1(K) \xrightarrow{j} \text{Im } j \to 0 \qquad (**)$

(each i being an inclusion).

Free generators for $Z_1(L)$ are obtained by putting back, one by one, the edges of L_1 not in T_1, then the edges of L_2 not in T_2, etc. Free generators for $Z_1(K)$ are all of these, plus z^1, ..., z^m. It follows from $(**)$ that free generators for $\text{Im } j$ are jz^1, ..., jz^m.

Free generators for $\text{Im } \partial_* = \text{Ker } i_*$ are $\{v^2 - v^1\}$, $\{v^3 - v^1\}$, ..., $\{v^\ell - v^1\}$, since $i_*(\lambda_1\{v^1\}_L + \ldots + \lambda_\ell\{v^\ell\}_L) = (\lambda_1 + \ldots + \lambda_\ell)\{v^1\}_K$. Finally, $\partial_* jc^\alpha = \partial_*(c^\alpha - \rho c^\alpha)$ (where ρ is the "restriction" which puts at zero the coefficients of simplexes not in L) =

$\{\partial c^{\alpha}\}_L - \{\partial \rho c^{\alpha}\}_L = \{v^{\alpha} - v^1\}_L - 0$, since $\rho c^{\alpha} \in C_1(L)$ so that $\{\partial \rho c^{\alpha}\}_L = 0$. Thus $jc^2, \ldots, jc^{\ell}$ are elements of $Z_1(K, L)$ which go, under ∂_*, to the free generators of Im ∂_*. The result follows from (*) and the last sentence of A.20(3).   Q.E.D.

*4.28 Remark   The following method for expressing a given $z \in Z_1(K, L)$ as an integer linear combination of the free generators listed in 4.27 is a straightforward consequence of the above proof.

Coefficient of jz^i = coefficient of e^i in z

$\qquad\qquad\qquad\qquad\qquad\qquad (i = 1, \ldots, m)$

Coefficient of jc^{α} = coefficient of $\{v^{\alpha}\}$ in $\partial_* z$

$\qquad\qquad\qquad\qquad\qquad\qquad (\alpha = 2, \ldots, \ell)$

$\qquad\qquad\quad$ = the sum of the coefficients of those

$\qquad\qquad\qquad\quad$ vertices of L^{α} in $\partial z \in C_0(L)$.

4.29 Examples   (1) Let K be the cylinder drawn below and L the (heavily drawn) subcomplex consisting of the two ends.
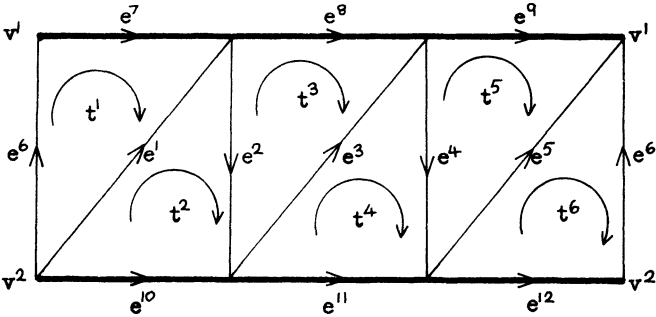
In this case L contains all the vertices of K, so that $C_0(K, L) = 0$, and $Z_1(K, L) = C_1(K, L)$ with basis $e^1, \ldots, e^6$. Generators for $B_1(K, L)$ are $\partial t^1, \ldots, \partial t^6$ and a presentation for $H_1(K, L)$ is

$$\begin{pmatrix} e^1, \ldots, e^6 \\ -e^1 + e^6 = e^1 + e^2 = -e^2 - e^3 = e^3 + e^4 = -e^4 - e^5 = e^5 - e^6 = 0 \end{pmatrix}$$

Thus $H_1(K, L)$ is infinite cyclic, with generator $\{e^6\}$ for example.

*If we use 4.27 to obtain a basis for $Z_1(K, L)$, we have $\ell = 2$ and taking L_1 to be the top of the cylinder and L_2 the bottom, we can take $T_1 = \{e^7, e^8, v^1, v^3, v^4\}$, $T_2 = \{e^{10}, e^{11}, v^2, v^5, v^6\}$ and $T = T_1 \cup T_2 \cup \{e^6\}$. Then $m = 5$ and the cycles $z^1, \ldots, z^5$ are obtained from T by replacing , respectively say, the edges $e^1, \ldots, e^5$. The first five generators of $Z_1(K, L)$ are then $e^1 - e^6$, $e^2 + e^6$, $e^3 - e^6$, $e^4 + e^6$, $e^5 - e^6$. The remaining generator is, say, e^2, which is jc^2 where $c^2 = e^7 + e^2 - e^{10}$ is associated to a path from $v^1 \in L_1$ to $v^2 \in L_2$. To express say $\partial t^4 = e^3 + e^4$ in terms of these we use 4.28: the coefficients of jz^3 and jz^4 are both 1, and the coefficient of jc^2 is 0 since $\partial(e^3 + e^4) = v^6 - v^5$, with sum of coefficients 0.

I leave it to the reader to verify that $H_2(K, L) \cong \mathbb{Z}$ with generator $\{t^1 + \ldots + t^6\}$. Of course $H_0(K, L) = 0$ (4.25).

<u>4.29</u> *(2X) Take K as in (1), but L this time to be one end, say the top, of the cylinder. Here L does not contain all the vertices of K, so some technique, e.g. that of 4.27, is really needed to write down generators for $Z_1(K, L)$. Verify that $H_1(K, L)$ and $H_2(K, L)$ are both trivial.

(3X) Take K to be the Möbius band given in 4.15 and L to be the rim of the band consisting of the edges $e^2, \ldots, e^6$ and their endpoints. Verify that $H_1(K, L) \cong \mathbb{Z}_2$, with generator $\{e^1\}$. Notice that $e^1 + e^1 \overset{L}{\sim} 0$; to see this
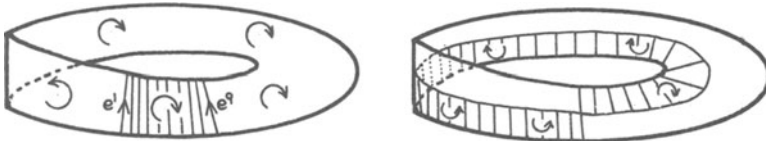
geometrically refer to the left-hand diagram on the previous page. Now $e^1 \overset{L}{\sim} e^9$ since the boundary of the shaded area (oriented as shown) is, apart from pieces of L, $e^1 - e^9$. But also $e^1 \overset{L}{\sim} -e^9$ since the boundary of the rest of the band, oriented as shown, is $e^1 + e^9$ apart from pieces of L. Hence $e^1 \overset{L}{\sim} -e^1$.

Verify that $e^1 + e^7 + e^8 + e^9 + e^{10}$, which goes "once round" the band is homologous mod L to e^1. This can be seen geometrically by considering the boundary of the shaded region in the right-hand diagram.

4.29 (4X)   Consider the homomorphism $j : C_1(K) \to C_1(K, L)$. If $z \in Z_1(K)$ then $jz \in Z_1(K, L)$, for

$$\hat{\partial} jz = j\partial jz = j\partial(z - \rho z) \quad (\rho : C_1(K) \to C_1(L) \text{ puts}$$
$$\text{coefficients of simplexes not}$$
$$\text{in L equal to 0)}$$
$$= 0 - j\partial\rho z \text{ since } \partial z = 0$$
$$= 0 \text{ since } \partial\rho z \in C_0(L).$$

Thus there is an induced homomorphism, which we still call   j,

$$j : Z_1(K) \to Z_1(K, L)$$

(this appeared in the proof of 4.27).   Suppose that L is connected (it does not matter whether K is connected).   Show that this j is epic, i.e. that every relative 1-cycle can be "extended", by adding edges of L, to an absolute 1-cycle. This is plausible from the picture, where the relative cycle **w** can be extended to an absolute cycle by adding the part of L between the endpoints **v** and **v'** of **w**. (Hint for the formal proof: let $\epsilon_K : C_0(K) \to \mathbb{Z}$ and
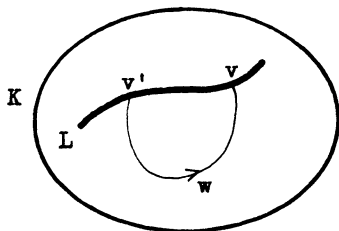
$\epsilon_L : C_0(L) \to \mathbb{Z}$ be the augmentations (4.5). Show that
$\epsilon_L \partial w = \epsilon_K \partial w = 0$ (see 4.6) and then use 1.23, which says that
$C_1(L) \overset{\partial}{\to} C_0(L) \overset{\epsilon_L}{\longrightarrow} \mathbb{Z}$ is exact.)

The result is false when L is not connected. For
example the relative cycle e^1 on the cylinder in (1) above
cannot be extended to an absolute cycle by adding edges in L.

<u>4.29</u> (5X) Take K to be the torus given in 2.2(3), and L to
be some particular subcomplex of K.   Verify that $H_2(K, L)$ is
free of rank equal to the number of regions into which L div-
ides K (in the obvious intuitive sense of "region").   Note
that when L is the 1-skeleton of K this is a special case
of 4.23(4), for the regions are then the 2-simplexes.

This will not work in general for a non-orientable surface
such as the Klein bottle. For example take $L = \emptyset$; then there
is one region but $H_2(K, L) = H_2(K) = 0$ by 4.18.   More of
this in Chapter 9.

THREE HOMOMORPHISMS

We shall define three homomorphisms

$$i_* : H_p(L) \to H_p(K)$$
$$j_* : H_p(K) \to H_p(K, L)$$
$$\partial_* : H_p(K, L) \to H_{p-1}(L)$$

which, in the next chapter, will be fitted together to form an
exact sequence.   An alternative approach is sketched in an
Appendix to this chapter. As usual K will be an oriented
simplicial complex and L a subcomplex of K.

<u>4.30</u> <u>The homomorphism</u> i_* Define $i_* : H_p(L) \to H_p(K)$

$$\text{by } i_*\{z\}_L = \{z\}_K \quad (z \in Z_p(L))$$

$$\text{i.e. } i_*\left(z + B_p(L)\right) = z + B_p(K) .$$

Thus we simply regard $z \in Z_p(L)$ as a cycle on K and take its homology class there. Of course $\{z\}_K$ can be zero when $\{z\}_L$ is non-zero. For example let K be a 2-simplex and all its faces, and L the boundary of the 2-simplex (the set of proper faces).
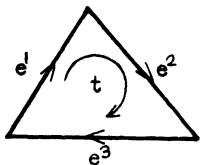
Then $\{e_1 + e_2 + e_3\}_L$ generates $H_1(K) \cong \mathbb{Z}$, but $\{e^1 + e^2 + e^3\}_K = 0$ since $\partial t = e^1 + e^2 + e^3$.

Several things must be checked before the definition of i_* is acceptable. These are

(1) If $z \in Z_p(L)$ then $z \in Z_p(K)$ (recall the preamble to 4.24)

(2) If $z \in B_p(L)$ then $z \in B_p(K)$

(3) $\{z + z'\}_K = \{z\}_K + \{z'\}_K$ where $z, z' \in Z_p(L)$.

(1) and (2) say that i_* is well-defined, and (3) say that i_* is a homomorphism.

All these assertions are easy to check.

4.31 Examples   (1) Let K be the Möbius band given in 4.15 and let L be the rim of the band. Then $H_1(L) \cong \mathbb{Z}$, with generator $\{e^2 + e^3 + e^4 + e^5 + e^6\}_L$, and $H_1(K) \cong \mathbb{Z}$, with generator $\{z^1\}_K = \{e^1 + e^7 + e^8 + e^9 + e^{10}\}_K$ (see 4.15). The homomorphism i_* is determined by its effect on the generator of $H_1(L)$, and according to the calculation given in 4.15,

$$i_*\{e^2 + e^3 + e^4 + e^5 + e^6\}_L = \{e^2 + e^3 + e^4 + e^5 + e^6\}_K$$

$$= 2\{z^1\}_K .$$

Thus i_* takes a generator to twice a generator: it is "multiplication by two".

(2X) Let K be the cylinder given in 4.29 and let L be

(i) both ends; (ii) one end of the cylinder. Show that
i_* takes (i) each of the generators $\{e^7 + e^8 + e^9\}_L$ and
$\{e^{10} + e^{11} + e^{12}\}_L$ of $H_1(L) \cong \mathbb{Z} \oplus \mathbb{Z}$ to a generator of
$H_1(K)$; (ii) the generator $\{e^7 + e^8 + e^9\}_L$ of $H_1(L) \cong \mathbb{Z}$
to a generator of $H_1(K)$. What is Ker i_* in case (i) ?

<u>4.31</u> (3X) Suppose that K has components $K_1, \ldots, K_k$ and
that $L \cap K_1, \ldots, L \cap K_k$ are connected (possibly empty).
Show that $i_* : H_0(L) \rightarrow H_0(K)$ is a monomorphism. When
is it an isomorphism? (In fact it is an isomorphism precisely
when $H_0(K, L) = 0$; in particular when K and L are both
connected and non-empty.)

(4X) Let K be the oriented simplicial complex obtained
from the projective plane drawn in 4.17(5) by removing the
three 2-simplexes t^1, t^2, t^3 (and then orienting the remaining
simplexes). Let L be the subcomplex of K consisting of
the eight 1-simplexes and five 0-simplexes on the boundaries
of these three 2-simplexes. Find four 1-cycles z^1, z^2, z^3,
z^4 which freely generate $Z_1(L)$. Calculate $H_1(K)$, including
generators, and express in terms of them the four homology
classes $\{z^\alpha\}_K$, $\alpha = 1, 2, 3, 4$. Hence find $i_*: H_1(L) \rightarrow H_1(K)$;
in particular show that it is epic.

<u>4.32</u> <u>The homomorphism</u> j_* Define $j_* : H_p(K) \rightarrow H_p(K, L)$

by $j_*\{z\}_K = \{jz\}_{K,L}$ for $z \in Z_p(K)$.

This time the following have to be checked:

(1) If $z \in Z_p(K)$ then $jz \in Z_p(K, L)$. (see 4.29(4))

(2) If $z \in B_p(K)$ then $jz \in B_p(K, L)$

(3) $\{j(z + z')\}_{K,L} = \{jz\}_{K,L} + \{jz'\}_{K,L}$

(The proof of (2) is very similar to that of 4.24(2).)

<u>4.33</u> <u>Examples</u>   (1X) Take K and L as in 4.31(1). Then
$H_1(K, L) \cong \mathbb{Z}_2$ with generator $\{e^1\}$ (see 4.29(3)). Thus

$j_*\{e^1 + e^7 + e^8 + e^9 + e^{10}\}_K = \{e^1\} + \{e^7\} + \{e^8\} + \{e^9\} + \{e^{10}\}$

(all on (K, L)) $= 5\{e^1\} = \{e^1\}$. Thus j_* takes generator

to generator. What is $j_*\{e^6 + e^2 + e^3 + e^4 + e^5\}_K$?

<u>4.33</u> (2X) Take K and L to be as in (i) of 4.31(2). Show
that $j_* : H_1(K) \to H_1(K, L)$ is trivial.

<u>4.34</u> <u>The homomorphism</u> ∂_* Define $\partial_* : H_p(K, L) \to H_{p-1}(L)$

$$\text{by} \qquad \partial_*\{z\}_{K,L} = \{\partial z\}_L \qquad z \in Z_p(K, L).$$

This time the following have to be checked:

(1) If $z \in Z_p(K, L)$ then $\partial z \in Z_{p-1}(L)$

(2) If $z \in B_p(K, L)$ then $\partial z \in B_{p-1}(L)$

(3) $\{\partial z + \partial z'\}_L = \{\partial z\}_L + \{\partial z'\}_L$.

From 4.24(1), $\partial z \in C_{p-1}(L)$; however $\partial\partial z = 0$ so $\partial z \in Z_{p-1}(L)$.
Notice, incidentally, that although ∂z is a boundary, $\{\partial z\}_L$
is not automatically zero. This is because ∂z is not
usually the boundary of something in L — though of course it
might be, as in the picture of 4.29(4), where z appears as
w. For (2), suppose $z = \hat\partial w$ ($w \in C_{p+1}(K, L)$); show that
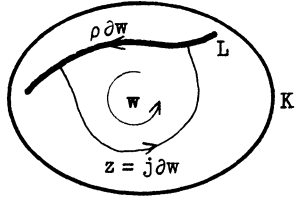
$\partial z = \partial(-\rho\partial w)$, where
$\rho : C_p(K) \to C_p(L)$ puts coeffi-
cients of simplexes not in L
equal to 0. (3) is immediate.

<u>4.35</u> <u>Examples</u>   (1) Take K, L as in 4.31(1). Then
$\partial_* : H_1(K, L) \to H_0(L)$ is trivial — hardly surprising, for
any homomorphism $\mathbb{Z}_2 \to \mathbb{Z}$ is trivial.

    (2X) Take K, L as in (i) of 4.31(2). Then
$\partial_* : H_2(K, L) \to H_1(L)$ takes the generator $\{t^1 + \ldots + t^6\}_{K,L}$
to $\{e^7 + e^8 + e^9\}_L - \{e^{10} + e^{11} + e^{12}\}_L$, while
$\partial_* : H_1(K, L) \to H_0(L)$ takes $\{e^6\}_{K,L}$ to $\{v^1\}_L - \{v^2\}_L$.

<u>4.35</u>(3X) The reader may care to verify that the sequence

$$H_2(L) \xrightarrow{i_*} H_2(K) \xrightarrow{j_*} H_2(K, L) \xrightarrow{\partial_*} H_1(L) \xrightarrow{i_*} H_1(K)$$

$$\xrightarrow{j_*} H_1(K, L) \xrightarrow{\partial_*} H_0(L) \xrightarrow{i_*} H_0(K) \xrightarrow{j_*} H_0(K, L) \to 0$$

is exact in each of the cases (1), (2) above. (Compare 6.1)

(4X) If K and L are both connected, then

∂_* : $H_1(K, L) \to H_0(L)$ is always trivial (compare (1) above).

[*]APPENDIX : CHAIN COMPLEXES

There is a systematic approach to the construction of homomor-
phisms like the three, i_*, j_* and ∂_*, constructed above,
and this approach will be sketched here. An elegant generali-
zation of this approach has been given by D. Puppe (1962).

<u>4.36</u> <u>Definitions</u>   A <u>chain complex</u> (C_p, d_p) is a collection
(C_p) of abelian groups, one for each $p \in \mathbb{Z}$, together with a
homomorphism $d_p : C_p \to C_{p-1}$ for each p,   such that, for all
p, $d_{p-1} \circ d_p = 0$. The examples to keep in mind are
$C_p = C_p(K)$, $d_p = \partial_p$ where K is an oriented simplicial
complex, and $C_p = C_p(K, L)$, $d_p = \hat{\partial}_p$ where L is a sub-
complex of the oriented simplicial complex K.   A <u>chain map</u>
(f_p) from a chain complex (C_p, d_p) to another, (C'_p, d'_p)
is a collection of homomorphisms $f_p : C_p \to C'_p$, such that
$d'_p \circ f_p = f_{p-1} \circ d_p$ for all p. This means that the dia-
gram

$$\cdots \to C_{p+1} \xrightarrow{d_{p+1}} C_p \xrightarrow{d_p} C_{p-1} \to \cdots$$

$$\downarrow f_{p+1} \qquad \downarrow f_p \qquad \downarrow f_{p-1}$$

$$\cdots \to C'_{p+1} \xrightarrow{d'_{p+1}} C'_p \xrightarrow{d'_p} C'_{p-1} \to \cdots$$

"commutes", i.e. for each square the two routes between the top left and bottom right are in fact the same homomorphism.

We can carry over the terminology of simplicial complexes to the general situation and speak of p-<u>cycles</u> (elements of $Z_p = \text{Ker } d_p$), p-<u>boundaries</u> (elements of $B_p = \text{Im } d_{p+1}$) and <u>homology groups</u> $H_p = Z_p/B_p$, which exist because $d_{p-1} \circ d_p = 0$.

Thus a chain map takes cycles to cycles and boundaries to boundaries. It follows easily that there is an induced homomorphism

<u>4.37</u>
$$\left(\begin{array}{c} f_* : H_p \to H'_p \\[2mm] \text{defined by } f_*(z + B_p) = f(z) + B'_p \quad (z \in Z_p) \end{array} \right.$$

or, as we may write, using $\{ \; \}$ ambiguously,

$$f_*\{z\} = \{f(z)\}.$$

<u>4.38</u> <u>Examples</u>   (1X) Taking, in the usual notation

$$C_p = C_p(L), \quad d_p = \partial_p : C_p(L) \to C_{p-1}(L)$$

$$C'_p = C_p(K), \quad d_p = \partial_p : C_p(K) \to C_{p-1}(K)$$

and $i_p : C_p(L) \to C_p(K)$ to be the inclusion, (i_p) is in fact a chain map, giving an induced homomorphism

$$i_* : H_p(L) \to H_p(K)$$

as in 4.30.

(2X) Taking

$$C_p = C_p(K), \quad d_p = \partial_p : C_p(K) \to C_{p-1}(K)$$

$$C'_p = C_p(K, L), \quad d_p = \hat{\partial}_p : C_p(K, L) \to C_{p-1}(K, L)$$

and $j_p : C_p(K) \to C_p(K, L)$ to be the usual j_p (4.20),

(j_p) is in fact a chain map giving an induced homomorphism

$$j_* \; : \; H_p(K) \; \to \; H_p(K, L)$$

as in 4.32.

4.38 (3X) Show that the "restriction" ρ of the proof of 4.24(2) does not, except in special cases like $L = K$ or $L = \emptyset$, give a chain map between the appropriate chain complexes.

(4XX) Suppose that each f_p in 4.36 is an epimorphism. Show that

$$z' + f_p Z_p \; \to \; \{z'\} + f_* H_p \qquad (z' \in Z_p')$$

defines an isomorphism

$$Z_p' / f_p Z_p \; \to \; H_p' / f_* H_p \; .$$

(In particular this is true when $f_p = j_p$.)

(5X) Let $f_p : C_p(K, L) \to C_{p-1}(L)$ be the composite

$$C_p(K, L) \xrightarrow{\text{inclusion}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K) \xleftarrow{} C_{p-1}(L).$$

Taking $C_p = C_p(K, L)$, $d_p = \hat{\partial}_p$; $C_p' = C_{p-1}(L)$, $d_p' = \partial_{p-1}$ it is not quite true that (f_p) is a chain map from (C_p, d_p) to (C_p', d_p'), for in fact $d_p' \circ f_p = -f_{p-1} \circ d_p$. Nevertheless f_p does take cycles to cycles and boundaries to boundaries, and there is an induced homomorphism $H_p(K, L) \to H_{p-1}(L)$, which is in fact the ∂_* of 4.34.

The notation for chain complexes and chain maps is easier to handle when the chain groups C_p are collected together into one big group C, as follows.

4.39 <u>Definitions</u> Let (C_p, d_p) be a chain complex. The <u>direct sum</u> $C = \oplus C_p$ of the groups C_p is the abelian group whose elements are families (c_p), $c_p \in C_p$, $p \in \mathbb{Z}$, with

all but a finite number of c_p equal to 0, and the operation $(c_p^1) + (c_p^2) = (c_p^1 + c_p^2)$. Thus if $C_p = 0$ when $p < 0$ or $p > n$, as is the case in the geometrical examples, then $C \cong C_0 \oplus C_1 \oplus \ldots \oplus C_n$ and an element of C can be regarded as an n-tuple $(c_0, c_1, \ldots, c_n)$.

The collection of homomorphisms (d_p) can be brought together into a single homomorphism $d : C \to C$, defined by: the element in the pth place of the family $d(c_p)$ is $d_{p+1} c_{p+1} \in C_p$.

$$
\begin{array}{ccccccccc}
C & \cdots & \oplus & C_{p+1} & \oplus & C_p & \oplus & C_{p-1} & \oplus & \cdots \\
\downarrow d & & & & d_{p+1} & & d_p & & d_{p-1} & \\
C & \cdots & \oplus & C_{p+1} & \oplus & C_p & \oplus & C_{p-1} & & \cdots
\end{array}
$$
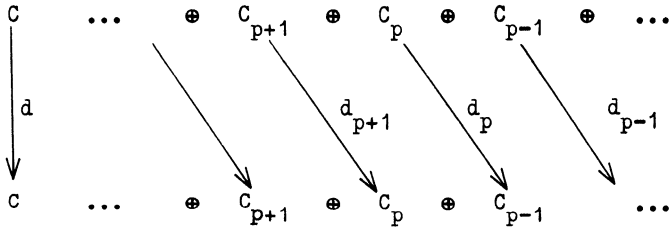
Then the condition $d_{p-1} \circ d_p = 0$ for all p is equivalent to $d \circ d = 0$, i.e. Im $d \subset$ Ker d. Also a chain map (f_p) can be thought of as a single homomorphism $f : C \to C'$ such that $d' \circ f = f \circ d$.

C, together with $d : C \to C$ is sometimes called a **differential graded group**, d being the differential and the decomposition of C as a direct sum being the grading. For further information see Hilton and Wylie (1960).

The question
of invariance

This chapter has a carefully chosen title: we are concerned
here primarily with the question, and not with the answer.
The question is this: suppose that K and K_1 are simplicial
complexes triangulating the same object, i.e. with $|K| = |K_1|$.
Is it true that $H_p(K) \cong H_p(K_1)$ for all p ?   The answer
is in fact "yes" — indeed a stronger assertion (5.13) is true,
namely that the isomorphisms hold if $|K|$ is merely supposed
homeomorphic to $|K_1|$, i.e. if there exists a continuous and
bijective map f : $|K| \rightarrow |K_1|$.   (The inverse map f^{-1} will
automatically be continuous by a theorem of general topology,
using the "compactness" of $|K|$.   See Kelley (1955), p.141.)
The stronger assertion is referred to as the topological invar-
iance of homology groups: it says that homology groups do not
depend on particular triangulations but only on the underlying
topological structure of $|K|$.

The topological invariance result can be restated in a
negative form: if there exists an integer p such that $H_p(K)$
and $H_p(K_1)$ are not isomorphic, then $|K|$ and $|X_1|$ are not
homeomorphic.   Using results of Chapter 6 this in turn implies
that no two inequivalent closed surfaces (2.8) have homeomorphic
underlying spaces.

But we shall not prove the topological invariance theorem

5.13, since it would divert attention from the developments of
succeeding chapters.   The only invariance result actually
needed is a much weaker one, and that one is proved in full;
see 5.7.

Though this chapter up to 5.9 forms an integral part of
the logical structure of the book, it could be omitted on a
first reading.   It might be a good idea to read the first
part of the section "Triangulations, simplicial approximation
and topological invariance", say up to 5.10, which expands on
the general statements above.

INVARIANCE   UNDER   STELLAR   SUBDIVISION

Let K be a simplicial complex.   A subdivision of K
is a simplicial complex K_1 with the property that $|K_1| = |K|$
and given $s_1 \in K_1$ there exists $s \in K$ such that $s_1 \subset s$.
Thus the simplexes of K_1 are contained in simplexes of K
but K_1 and K triangulate the same subset of $\mathbb{R}^N$.   An
example is provided by the barycentric subdivision of a sur-
face described on p.61, where each triangle is divided into
six by drawing the medians, and each edge is thereby halved;
more of this anon.   Another example is the stellar subdivision
described in 3.29(5): a reader who does not want to work
through that example should read what follows by specializing
always to the case where K has dimension 2.   That will suf-
fice for the applications in this book and will give a per-
fectly adequate idea of the techniques involved.   A stellar
subdivision is the result of a sequence of subdivisions, each
of these being starring of a simplex s at an interior point
v.   When K has dimension 2, the following pictures show the
result of starring s at an interior point v.   When s is
a vertex no change takes place.   In general, we remove
$\overline{St}(s, K)$ and replace it by the join $v \dot{s} Lk(s, K)$; compare
3.24, 3.26.

$s = (v^0 v^1 v^2)$

$s = (v^0 v^1)$

Three triangles with a common face, s
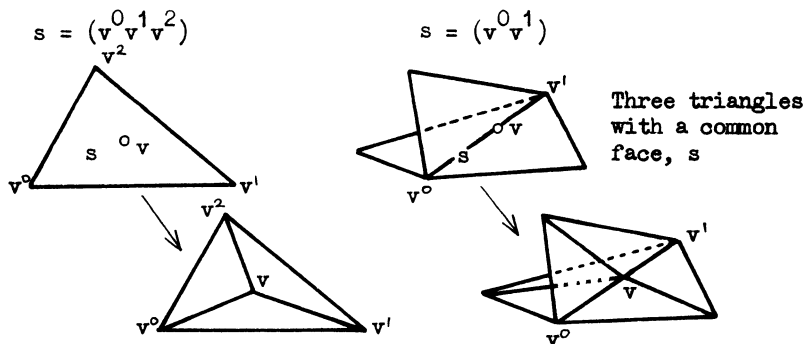
When K has dimension 2 we can produce the first bary-
centric subdivision of K by starring each 2-simplex at its
centroid and then starring each 1-simplex of the original sim-
plicial complex at its mid-point.
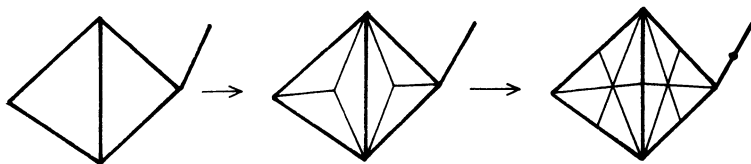
In fact the procedure of starring by descending dimension always
produces the first barycentric subdivision; a general defini-
tion of barycentric subdivision is given in an appendix to this
chapter for those who wish to verify this fact.

Let K_1 be obtained from K by starring $s = (v^0 \ldots v^n)$
at v.   Then all simplexes of K with s as a face are
deleted, and certain simplexes, listed below, are introduced.
If K is oriented, then K_1 can be oriented in a natural way,
and the orientations are indicated in the list, where it is
assumed that $(v^0 \ldots v^n)$ gives the correct orientation for s.

5.1
$$\begin{cases} (v), \text{ often written just } v \\ (v^0 \ldots v \ldots v^n) \text{ where } v \text{ replaces } v^i \text{ in } s \ (i = 0, \ldots, n) \\ \pm (v w^0 \ldots w^m) \\ \pm (v^0 \ldots v \ldots v^n w^0 \ldots w^m) \\ v \text{ replacing } v^i \text{ in } s \ (i = 0, \ldots, n) \end{cases}$$

where $\pm(v^0 \ldots v^n w^0 \ldots w^m)$
(same sign as on the
left) belongs to K

The reader should check in some simple cases that these give the "right" orientations for simplexes of K_1. Note that

$$(v^0 \ldots v \ldots v^n) = (-1)^i (vv^0 \ldots \hat{v}^i \ldots v^n)$$

where v replaces v^i in s on the left of the equation.

Some notation will be useful in what follows. Write σ for the oriented simplex $(v^0 \ldots v^n)$, σ^i for $(v^0 \ldots \hat{v}^i \ldots v^n)$, τ for $(w^0 \ldots w^m)$ and τ^j for $(w^0 \ldots \hat{w}^j \ldots w^m)$. Joins are denoted as usual by concatenation of symbols ($\sigma\tau$ etc.) and $v\sigma$, for example, denotes $(v)\sigma = (vv^0 \ldots v^n)$. The notation can be extended to chains: for example $(\partial\sigma)\tau$ means $\Sigma (-1)^i (\sigma^i \tau) \in C_{m+n}(K)$.

Note that, in this notation, we have the following formula for the boundary of a join:

5.2 $\partial(\sigma\tau) = (\partial\sigma)\tau - (-1)^n \sigma(\partial\tau), \quad n = \dim \sigma.$

(If $\dim \sigma = 0$ then $(\partial\sigma)\tau$ is to be read as τ, and similarly if $\dim \tau = 0$ then $\sigma(\partial\tau) = \sigma$.)

Now we can define a homomorphism

$$\alpha_p : C_p(K) \to C_p(K_1)$$

as follows. Let $\sigma\tau = (v^0 \ldots v^n w^0 \ldots w^m)$ be a simplex in $St(s, K)$; then

$$\alpha_p(\sigma\tau) = v(\partial\sigma)\tau = \sum (v^0 \ldots v \ldots v^n w^0 \ldots w^m)$$

where v replaces v^i and the sum is from $i = 0$ to $i = n$. This replaces $\sigma\tau$ by the sum of simplexes into which it is broken, with appropriate orientations. Also define $\alpha_p\sigma = v(\partial\sigma) = \sum (v^0 \ldots v \ldots v^n)$ (see diagram).
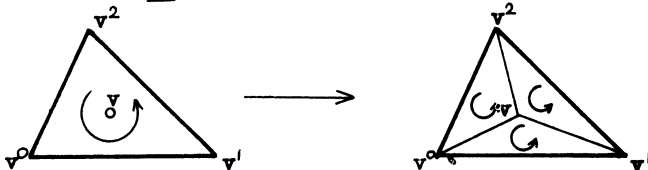
Other simplexes, not in $St(s, K)$, α_p leaves unchanged.

<u>5.3</u> <u>Proposition</u> α_p takes cycles to cycles and boundaries
to boundaries and so induces a homomorphism

$$\alpha_* : H_p(K) \to H_p(K_1)$$

defined by $\alpha_*\{z\} = \{\alpha_p z\}$.

<u>Proof</u>   In fact the collection of homomorphisms α_p is
a chain map (4.36), i.e. writing ∂ for the boundary homo-
morphism in either K or K_1, $\partial \circ \alpha_p = \alpha_{p-1} \circ \partial$ for all p.
The result follows easily from this (compare 4.37). To prove
the assertion about (α_p) it is enough to check that
$\partial \circ \alpha_p(\sigma\tau) = \alpha_{p-1} \circ \partial(\sigma\tau)$ whenever $\sigma\tau \in K$, and also
$\partial \circ \alpha_p \sigma = \alpha_{p-1} \circ \partial \sigma$. Using the formula for α_p above,
5.2 and the fact that $\partial \circ \partial = 0$ it follows without much
difficulty that

$$\partial \circ \alpha_p(\sigma\tau) = \alpha_{p-1} \circ \partial(\sigma\tau) = (\partial\sigma)\tau - (-1)^n v(\partial\sigma)(\partial\tau)$$

where $n = \dim \sigma$. The other assertion follows similarly.

<div align="right">Q.E.D.</div>

Next we shall define a homomorphism

$$\beta_p : C_p(K_1) \to C_p(K).$$

To do this we consider a more general situation.

<u>5.4</u> <u>Definition</u>   Let K_1 and K be any two simplicial com-
plexes. A <u>simplicial map</u>

$$\beta : K_1 \to K$$

is a map from the set of vertices of K_1 to the set of ver-
tices of K with the property that

$$(w^0 \ldots w^p) \in K_1 \implies \{\beta w^0, \ldots, \beta w^p\} \text{ is the set of ver-}$$
$$\text{tices of a simplex of } K.$$

Note that βw^0, ..., βw^p need not be distinct: for example β could map all vertices of K_1 to the same vertex of K.

Given a simplicial map $\beta : K_1 \to K$ as above we can define a homomorphism

$$\beta_p \; : \; C_p(K_1) \; \to \; C_p(K)$$

by

$$\beta_p(w^0 \dots w^p) \; = \; \begin{cases} (\beta w^0 \dots \beta w^p) & \text{if these } p+1 \text{ vertices} \\ & \text{of } K \text{ are distinct} \\ 0 & \text{otherwise} \end{cases}$$

<u>5.5</u> <u>Proposition</u>   With the above notation, β_p induces a homomorphism

$$\beta_* \; : \; H_p(K_1) \; \to \; H_p(K)$$

defined by $\beta_*\{z\} \; = \; \{\beta_p z\}$.

<u>Proof</u>   Again (β_p) is a chain map. The proof will be left to the reader, with the following hint: consider in turn the cases where βw^0, ..., βw^p contains $p+1$ distinct vertices of K, has one repetition, and has more than one repetition. In the first case $\beta_{p-1} \circ \partial(w^0 \dots w^p) = \partial \circ \beta_p(w^0 \dots w^p)$ since the formulae for the two sides are identical. In the other cases both sides are zero.   Q.E.D.

Returning to the case where K_1 is obtained from K by starring $s = (v^0 \dots v^n)$ at v, there is a simplicial map $\beta : K_1 \to K$ defined by $\beta v = v^0$ and $\beta w = w$ for $w \neq v$. By 5.5 this gives a homomorphism $\beta_* : H_p(K_1) \to H_p(K)$.

<u>5.6</u> <u>Theorem</u>   With the notation just given, and that of 5.3, α_* and β_* are inverse isomorphisms, so that

$$\alpha_* \; : \; H_p(K) \; \cong \; H_p(K_1).$$

<u>Proof</u>   It is immediate from the definitions that $\beta_p \circ \alpha_p = $ identity; hence

$\beta_* \circ \alpha_* \{z\} = \beta_* \{\alpha_p z\} = \{\beta_p \circ \alpha_p z\} = \{z\}$ for any $\{z\} \in H_p(K)$, i.e. $\beta_* \circ \alpha_* =$ identity.

Next note that, for any $c \in C_p(K_1)$, $\alpha_p \circ \beta_p c - c$ contains only simplexes in $\overline{St}(v, K_1)$, since β_p does not change anything outside $\overline{St}(v, K_1)$, and α_p does not change anything outside $\overline{St}(s, K)$. If $z \in Z_p(K_1)$, it follows that $\alpha_p \circ \beta_p z - z$ is a cycle on $\overline{St}(v, K_1)$, since α_p and β_p take cycles to cycles. Now $\overline{St}(v, K_1)$ is the join of $\{v\}$ to the link $Lk(v, K_1)$ (see 3.29(4)): i.e. the closed star is a cone on the link. Now any p-cycle $(p > 0)$ on a cone is homologous to zero. This follows at once from 6.8, or alternatively from the following direct argument.

Let $z' \in Z_p(\overline{St}(v, K_1))$ and write L for $Lk(v, K_1)$. Then there are chains $c' \in C_{p-1}(L)$ and $c'' \in C_p(L)$ such that $z' = vc' + c''$ (thus c' contains the simplexes in z' which have v as a vertex). Hence $0 = \partial z' = c' - v(\partial c') + \partial c''$ from 5.2, and this implies that $c' + \partial c'' = 0$ and $\partial c' = 0$. Hence

$$\partial(vc'') = c'' - v(\partial c'') = c'' + vc' = z'$$

so that $z' \sim 0$.

This shows that $\alpha_p \circ \beta_p z \sim z$, and so $\alpha_* \circ \beta_* =$ identity for $p > 0$. For $p = 0$, we have $\alpha_0 \circ \beta_0 v - v = v^0 - v \sim 0$ (since $(v^0 v) \in K_1$) and $\alpha_0 \circ \beta_0 w = w$ for $w \neq v$; hence the result, namely $\alpha_* \circ \beta_* =$ identity for all p.   Q.E.D.

5.7 Corollary (Invariance under stellar subdivision)
Suppose that K_1 is a stellar subdivision of K, i.e. is obtained from K by a sequence of starrings. Then $H_p(K_1) \cong H_p(K)$. In particular the isomorphism holds if K_1 is obtained from K by one or more barycentric subdivisions.

Proof   This follows immediately by repeated application of the theorem.   Q.E.D.

5.8 The relative case   In order to avoid overburdening the above proofs with notation, the absolute case only was

considered. Nevertheless the same arguments can be used for
a pair (K, L) consisting of a simplicial complex K and a
subcomplex L. Thus (K_1, L_1) is a subdivision of (K, L)
if K_1 subdivides K and L_1 subdivides L. A stellar
subdivision K_1 of K will automatically give a stellar sub-
division L_1 of L, and in the case of a single starring the
homomorphisms α_* and β_* can be defined as before, but this
time between $H_p(K, L)$ and $H_p(K_1, L_1)$. (Note that a sim-
plicial map $\beta : (K_1, L_1) \to (K, L)$ is a simplicial map
$K_1 \to K$ such that $(w^0 \dots w^p) \in L_1 \implies \{\beta w^0, \dots, \beta w^p\}$ are
the vertices of a simplex of L.) As before, α_* and β_*
are inverse isomorphisms, and it follows that
$H_p(K_1, L_1) \cong H_p(K, L)$ whenever (K_1, L_1) is obtained from
(K, L) by successive starrings.

5.9 Remark   Two simplicial complexes K and L are called
piecewise linearly (p.l.) equivalent if there are subdivisions
K_1 of K and L_1 of L such that K_1 and L_1 are isomor-
phic. For example, any two closed surfaces with the same
standard form are p.l. equivalent (as well as being equivalent
in the sense of 2.8). To see this informally superimpose
their standard polygons (assumed regular) and subdivide to make
a common subdivision of each; the diagram indicates roughly how
this is done for two overlapping triangles. The converse is
also true: if two surfaces are p.l. equivalent then they have
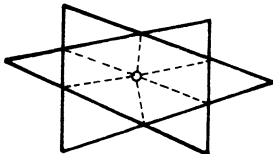
the same standard form.
Compare 2.16(8), which says
that Euler characteristic is
unaltered by subdivision.

    The same sort of subdivision construction as that illus-
trated above can be used to show that any two triangulations
of the same polyhedron, i.e. any two simplicial complexes in
$\mathbb{R}^N$ with the same underlying space, have a common subdivision,
and hence are p.l. equivalent.

TRIANGULATIONS SIMPLICIAL APPROXIMATION

AND TOPOLOGICAL INVARIANCE

In this section we shall sketch very briefly the route to a
more general invariance theorem than 5.7.   Full details appear
in many textbooks of algebraic topology, e.g. Maunder (1970),
Hocking and Young (1961) or Hilton and Wylie (1960).   There
is also a pleasant account of the 2-dimensional case in Ahlfors
and Sario (1960).

Let X be a subset of a euclidean space $\mathbb{R}^N$.   Then cer-
tainly X may fail to be the underlying space of any simplicial
complex - for example if X is a circle - but the next best
thing to hope for is that there is some simplicial complex K
such that $|K|$ is homeomorphic to X.   Thus taking X to be
the unit circle in $\mathbb{R}^2$
and K to be the set
of edges and vertices
of the square in $\mathbb{R}^2$
with vertices $\left(\pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2}\right)$
a homeomorphism
$f : |K| \to X$ can be
defined by radial pro-
jection from the origin.
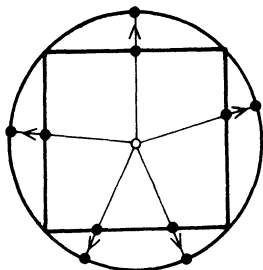
Let $X \subset \mathbb{R}^N$.   Given a simplicial complex K and a homeo-
morphism $h : |K| \to X$ we call K (or, more precisely, h) a
triangulation of X.
In this sense the closed
surfaces of Chapter 2
are triangulations of
appropriate "curved
surfaces" in euclidean
space.   For example the
hollow tetrahedron is a
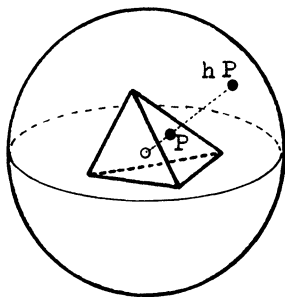
triangulation of a 2-sphere, h being given say by radial
projection.

As another example, the unit disk D^2 in $\mathbb{R}^2$ is the set
of points (x, y) with $x^2 + y^2 \leqslant 1$; thus a triangulation of
D^2 is a homeomorphism $h : |K| \to D^2$ where K is a simplicial
complex. Now K must be 2-dimensional (since h is a homeo-
morphism); the image of the 1-skeleton of K in D^2 will be
a curved graph in the sense of the discussion leading up to
Fáry's theorem 1.34. By that theorem there is a homeomorphism
$f : \mathbb{R}^2 \to \mathbb{R}^2$ taking this curved graph to a straight graph,
so that $f \circ h$ defines an isomorphism between K and a sim-
plicial complex in $\mathbb{R}^2$. Thus every triangulation of the disk
D^2 can be realized in $\mathbb{R}^2$. With slightly more trouble one
can prove, also using Fáry's theorem, that every triangulation
of the 2-sphere $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ can be
realized in $\mathbb{R}^3$.

When K is a triangulation of X we cannot immediately
define the homology groups of X to be those of K, since
there could well be another simplicial complex K_1 and a
homeomorphism $h_1 : |K_1| \to X$; we do not know that
$H_p(K_1) \cong H_p(K)$ for all p. As an example, K_1 could be
any subdivision of K, in which case h_1 could equal h, and
then we know the existence of an isomorphism only for stellar
subdivisions.

What is needed is a result which asserts the following:
Suppose that K_1 and K are oriented simplicial complexes
and that $f : |K_1| \to |K|$ is a homeomorphism. Then
$H_p(K_1) \cong H_p(K)$ for all p. Granted this and given h_1 and
h as above, we simply take $f = h^{-1} \circ h_1$ to show that the
two triangulations of X have isomorphic homology groups.
There follows a brief summary of the steps needed to obtain
the above result.

It is worth noting that it is not necessarily true that
K_1 and K as above are p.l. equivalent (5.9). At one time

it was conjectured that any two triangulations of homeomorphic
polyhedra were p.l. equivalent; this was known as the Haupt-
vermutung (main conjecture).   The first counterexample
(6-dimensional polyhedra) was found by Milnor (1961), and more
recently counterexamples have been constructed for more res-
tricted conjectures in which special properties of the poly-
hedra are assumed.

Let K_1 and K be any two simplicial complexes and let
$f : |K_1| \to |K|$ be a continuous map. We "approximate" f by
a simplicial map $\beta : K_1 \to K$.

<u>5.10</u> <u>Definition</u>   A simplicial map $\beta : K_1 \to K$ is called a
<u>simplicial approximation</u> to $f : |K_1| \to |K|$ provided that,
for each vertex v of K_1,

$$f\left(\operatorname{St}(v, K_1)\right) \subset \operatorname{St}(\beta v, K).$$

As an example, the diagram shows a hexagon K_1 (vertices
1, ..., 6) and a triangle K (vertices 1,3,5) inscribed in a
circle.   A continuous map $|K_1| \to |K|$ is defined by pro-
jection towards the centre of the circle, and a simplicial
approximation to this is $\beta : K_1 \to K$ where β takes
1,2,3,4,5,6 to, respectively, 1,3,3,5,5,1.   On the other
hand consider the inverse map $g : |K| \to |K_1|$ given by outward
projection.   There is no simplicial map $K \to K_1$ which is a
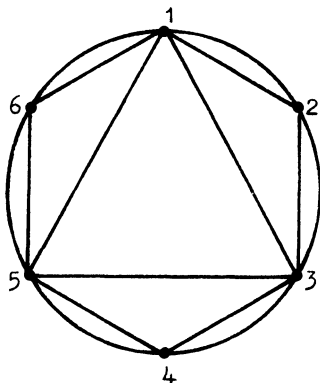
simplicial approximation
to g, since $g(\operatorname{St}(1, K))$
is not contained in the
star of any vertex of K_1.
However K can be bary-
centrically subdivided -
once is enough, giving
K' - so that the map
$g : |K'| = |K| \to |K_1|$
does have a simplicial

approximation. In general, the following result holds.

5.11 Simplicial Approximation Theorem   Given simplicial

complexes K_1 and K and a continuous map $f : |K_1| \to |K|$
there exists an integer r and a simplicial map $\beta : K_1^{(r)} \to K$
(the superscript denoting r^{th} barycentric subdivision) which
is a simplicial approximation to f.

Suppose, then, that K_1, K and f are as in 5.11; by
5.7 there is an isomorphism $\gamma_* : H_p(K_1) \to H_p(K_1^{(r)})$ and
so a homomorphism

$$f_* : H_p(K_1) \xrightarrow{\gamma_*} H_p(K_1^{(r)}) \xrightarrow{\beta_*} H_p(K)$$

for each p. One vital, and far from obvious, fact about f_*
is that it does not depend on r or β, so long as β is a
simplicial approximation to f. Thus it really depends only
on f, and the notation f_* is justified. Another important
fact is this:

5.12 Lemma   Let $g : |K_2| \to |K_1|$ and $f : |K_1| \to |K|$ be

continuous maps, where K_2, K_1 and K are (oriented) simpli-
cial complexes. Then $(f \circ g)_* = f_* \circ g_* : H_p(K_2) \to H_p(K)$
for each p.

The result required now follows at once by taking $K_2 = K$
and f a homeomorphism with inverse g. Hence:

5.13 Topological invariance of homology groups   Let K_1 and

K be (oriented) simplicial complexes and $f : |K_1| \to |K|$ be
a homeomorphism. Then $f_* : H_p(K_1) \to H_p(K)$ is an isomorphism
for each p.

*APPENDIX ON BARYCENTRIC SUBDIVISION

Let s be an n-simplex, $s = (v^0 \ldots v^n)$. The _barycentre_ of
s is the point $\hat{s} = \frac{1}{n+1}(v^0 + v^1 + \ldots + v^n)$. Thus the
barycentre of a 0-simplex (v) is v, the barycentre of a

1-simplex $(v^0 v^1)$ is its mid-point, and the barycentre of a
2-simplex $(v^0 v^1 v^2)$ is its centroid.

The first barycentric subdivision K' of a simplicial
complex K is the simplicial complex with vertex set
$\{\hat{s} : s \in K\}$, and with simplexes $t = (\hat{s}^0 \hat{s}^1 \dots \hat{s}^q)$ where
$s^0 < s^1 < \dots < s^q$ and $s^0, \dots, s^q$ are simplexes of K.
(The superscript is just for indexing, and does not necessarily
give the dimension of s^i.) Notice that $t \subset s^q$; to see that
K' is a subdivision of K we must therefore show that
$|K| \subset |K'|$. If $x \in |K|$ then $x \in s$ for some $s \in K$ - we
can take the s of lowest dimension for which this is true,
so that if $s = (v^0 \dots v^p)$ then $x = \lambda_0 v^0 + \dots + \lambda_p v^p$
where the barycentric coordinates are all > 0. We can also
assume $\lambda_0 \geqslant \lambda_1 \geqslant \dots \geqslant \lambda_p$. Let $s^i = (v^0 \dots v^i)$
$(i = 0, \dots, p)$; then it is easy to check that $x \in (\hat{s}^0 \hat{s}^1 \dots \hat{s}^p)$:
in fact

$$x = (\lambda_0 - \lambda_1)\hat{s}^0 + 2(\lambda_1 - \lambda_2)\hat{s}^1 + \dots +$$
$$p(\lambda_{p-1} - \lambda_p)\hat{s}_{p-1} + (p + 1)\lambda_p \hat{s}_p .$$

If L is a subcomplex of K then L is automatically
subdivided into L' when K is subdivided into K'. The
reader may like to check the following assertions.

(1) $\hat{s}$ is a vertex of L' if and only if $s \in L$.

(2) If t is a simplex of K' such that all vertices of
t belong to L', then $t \in L'$.

(3) Suppose that $(vw) \in K'$, where v is a vertex of L'
and w is not a vertex of L'; then w is not a vertex of K
(hence is the barycentre of a simplex of K of dimension > 0).

(4) Suppose that $s \in K$ and that no vertex of s belongs
to L. Then there is no edge $(vw) \in K'$ with $v \in L'$ and
w the barycentre of a face of s.

Some general theorems

In this chapter and the next we shall prove some general results which will find application in Chapter 9. For the present the motivation lies in the evident difficulty of calculating homology groups directly from the definition, and the consequent need for some help from general theorems. Enough is proved here to enable us to calculate the homology groups of all the closed surfaces described in Chapter 2, without difficulty. We shall also calculate the homology groups of some other standard simplicial complexes such as cones (6.8) and spheres (6.11(3)), and show (6.10) that relative homology groups can be defined in terms of ordinary homology groups.

The method presented here for the calculation of homology groups of closed surfaces is based on the idea of collapsing which was introduced in 3.30. This is not the only available method. From the point of view of this book, the advantage of using it is that, having established the invariance of homology groups under barycentric subdivision (5.7), we are in possession of a rapid and rigorous approach to the results. Another method of wide application comes from the Mayer-Vietoris sequence, presented in Chapter 7, where indications are given as to how the homology groups of closed surfaces can be re-calculated by use of the sequence. For effective and

unfettered use of the Mayer-Vietoris sequence one really needs
to use something like the topological invariance theorem stated
in 5.13 - that is, it is better to forget about particular
triangulations altogether.   Some uninhibited calculations
using the Mayer-Vietoris sequence appear in the next chapter,
starting at 7.5.

THE   HOMOLOGY   SEQUENCE   OF   A   PAIR

Let K be an oriented simplicial complex of dimension n and
L a subcomplex of K.   Recall from 4.30, 4.32 and 4.34 the
homomorphisms i_*, j_* and ∂_* ;   the following sequence is
called the homology sequence of the pair (K, L).

$$0 \to H_n(L) \xrightarrow{i_*} H_n(K) \xrightarrow{j_*} H_n(K, L) \xrightarrow{\partial_*} H_{n-1}(L) \xrightarrow{i_*} H_{n-1}(K) \xrightarrow{j_*} \ldots$$

$$\ldots \xrightarrow{j_*} H_1(K, L) \xrightarrow{\partial_*} H_0(L) \xrightarrow{i_*} H_0(K) \xrightarrow{j_*} H_0(K, L) \to 0$$

6.1 Theorem   The homology sequence of the pair (K, L) is
exact.

Proof   As usual, we regard $C_p(L)$ as a subgroup of
$C_p(K)$ (see the note preceding 4.24), and denote by
$\rho : C_p(K) \to C_p(L)$ the "restriction" which puts coefficients
of simplexes not in L equal to 0.   We have to prove exact-
ness at three places - $H_p(L)$, $H_p(K)$, $H_p(K, L)$ - and each of
these involves two proofs of inclusion.

        (1) Exactness at $H_p(L)$.   Let $z \in Z_{p+1}(K, L)$;   then
$i_* \partial_* \{z\}_{K,L} = \{\partial z\}_K$.   However $z \in C_{p+1}(K)$,   so $\partial z \in B_p(K)$,
i.e. $\{\partial z\}_K = 0$.   Hence $\mathrm{Im}\,\partial_* \subset \mathrm{Ker}\,i_*$.
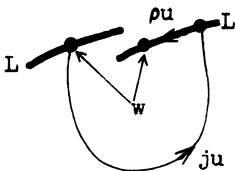
Suppose, conversely,
that $w \in Z_p(L)$ and
$i_*\{w\}_L = \{w\}_K = 0$.   Then
$w = \partial u$ for some

$u \in C_{p+1}(K)$.   Writing $u = ju + \rho u$ we have
$w = \partial u = \partial ju + \partial \rho u$.   Hence $\partial ju \in C_p(L)$, i.e. $ju \in Z_{p+1}(K, L)$.
Since $\partial_*\{ju\}_{K,L} = \{\partial ju\}_L = \{w\}_L - \{\partial \rho u\}_L = \{w\}_L$ $(\partial \rho u \in B_p(L))$
we have $\{w\}_L \in \text{Im } \partial_*$,   so Ker $i_* \subset \text{Im } \partial_*$.   Notice that
although ∂ju may not equal w (see figure), it is homologous
to w, since $\partial \rho u \sim 0$ in L.

(2) <u>Exactness at</u> $J_p(K)$.   Let $z \in Z_p(L)$;   then
$j_* i_* \{z\}_L = \{jz\}_{K,L} = \{0\} = 0$.   Hence Im $i_* \subset$ Ker j_* .
(This is the only one of the three cases where the composite
homomorphism is trivial even at the chain level.)

Conversely let $w \in Z_p(K)$ be such that
$j_*\{w\}_K = \{jw\}_{K,L} = 0$.   Then $jw = \hat{\partial}u = j\partial u$ for some
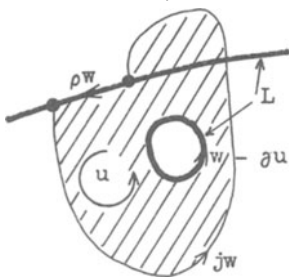
$u \in C_{p+1}(K, L)$.   Hence
$w - \partial u \in C_p(L)$;   indeed
$\partial(w - \partial u) = \partial w = 0$
so $w - \partial u \in Z_p(L)$.
Finally
$i_*\{w - \partial u\}_L = \{w - \partial u\}_K$
$= \{w\}_K$ (since
$\partial u \in B_p(K))$, so that
$\{w\}_K \in \text{Im } i_*$.   This

shows Ker $j_* \subset \text{Im } i_*$.

(3) <u>Exactness at</u> $H_p(K, L)$.   Let $z \in Z_p(K)$;   then
$\partial_* j_*\{z\}_K = \{\partial jz\}_L = \{\partial(z - \rho z)\}_L = \{-\partial \rho z\}_L$ $(\partial z = 0) = 0$
since $\rho z \in C_p(L)$.   (Same diagram as in (2), with z replacing
w.)   Hence Im $j_* \subset$ Ker ∂_* .

Conversely, let $w \in Z_p(K, L)$ be such that
$\partial_*\{w\}_{K,L} = \{\partial w\}_L = 0$.   Thus $\partial w = \partial u$, for some $u \in C_p(L)$,
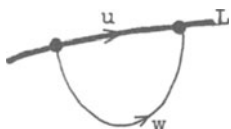
so that $\partial(w - u) = 0 : w - u \in Z_p(K)$.
Further $j_*\{w - u\}_K = \{jw - ju\}_{K,L}$
$= \{w\}_{K,L}$.   Thus $\{w\}_{K,L} \in \text{Im } j_*$;
hence Ker $\partial_* \subset \text{Im } j_*$.   Q.E.D.

6.2 Remark   There is a slightly different form of the sequence in 6.1 which is sometimes more convenient on account of the fact that it contains reduced zero-dimensional homology groups, and these are zero for connected simplicial complexes $(4.9(2))$. Recall (4.8), $\tilde{H}_0(K) = \text{Ker } \epsilon_K / \text{Im } \partial_1$, where $\epsilon_K : C_0(K) \to \mathbb{Z}$ is the augmentation. For $z \in \text{Ker } \epsilon_K$ write
$$[z] = [z]_K = z + \text{Im } \partial_1 \in \tilde{H}_0(K).$$
    Define $\tilde{\partial}_* : H_1(K, L) \to \tilde{H}_0(L)$

by $\{z\}_{K,L} \to [\partial z]_L$

        $\tilde{i}_* : \tilde{H}_0(L) \to \tilde{H}_0(K)$

by $[z]_L \to [z]_K$

and $\tilde{j}_* : \tilde{H}_0(K) \to H_0(K, L)$

by $[z]_K \to \{jz\}_{K,L}$.

It is left to the reader to check that these really do define homomorphisms, and that the reduced sequence of the pair (K, L), identical with the above sequence but ending
$$\ldots \to H_1(K, L) \xrightarrow{\tilde{\partial}_*} \tilde{H}_0(L) \xrightarrow{\tilde{i}_*} \tilde{H}_0(K) \xrightarrow{\tilde{j}_*} H_0(K, L) \to 0$$

is exact provided L is non-empty.

6.3 Examples   (1) The exact sequence of (K, L) can sometimes be used to derive the relative homology groups of (K, L) from knowledge of the homology groups of L, those of K, and the homomorphisms i_*. The technique is to get $H_p(K, L)$ in the middle of a short exact sequence:

$$0 \to \text{Ker } \partial_* \to H_p(K, L) \xrightarrow{\partial_*} \text{Im } \partial_* \to 0$$

$$\begin{array}{cc} \| & \| \\ \text{Im } j_* & \text{Ker}\left(i_* : H_{p-1}(L) \to H_{p-1}(K)\right) \\ \wr\| & \\ H_p(K)/\text{Im}\left(i_* : H_p(L) \to H_p(K)\right) & \end{array}$$

If this sequences splits (A.19) then $H_p(K, L)$, including generators, can be calculated - see the examples below.

<u>6.3</u> (2) Let K and L be as in 4.29(3) (Möbius band and its rim). Using the calculations of 4.15, the groups and homomorphisms written in under the following reduced homology sequence are known:

$$H_2(L) \to H_2(K) \to H_2(K, L) \to H_1(L) \overset{i_*}{\to} H_1(K) \overset{j_*}{\to} H_1(K, L)$$

$$\qquad 0 \qquad\quad 0 \qquad\qquad\qquad\qquad \mathbb{Z} \overset{\times 2}{\to} \mathbb{Z}$$

$$\qquad \to \tilde{H}_0(L) \;\to\; \tilde{H}_0(K) \;\to\; H_0(K, L) \;\to\; 0$$

$$\qquad\qquad 0 \qquad\qquad 0$$

Hence $H_2(K, L) \cong \operatorname{Ker} i_* = 0$

$\qquad H_1(K, L) \cong H_1(K)/\operatorname{Im} i_* \cong \mathbb{Z}_2$, with generator the image under j_* of a generator of $H_1(K)$

$\qquad\quad H_0(K, L) = 0.$

(If we use the unreduced sequence it is only necessary to see that $i_* : H_0(L) \to H_0(K)$ is an isomorphism (compare 4.31(3)), which implies that the homomorphisms before and after this i_* are trivial.)

(3X) Use the (unreduced) homology sequence to calculate the relative homology groups of (cylinder, both ends) - see 4.29(1) for diagram. In this case $i_* : H_0(L) \to H_0(K)$ has infinite cyclic kernel, generated by $\{v^1 - v^2\}_L$. Also $i_* : H_1(L) \to H_1(K)$ takes each generator $\{e^7 + e^8 + e^9\}_L$, $\{e^{10} + e^{11} + e^{12}\}_L$ to a generator of $H_1(K)$; in particular this i_* is epic. Use the method of (1) above.

(4X) Take K = the cylinder of 4.29(1) and L = one end. This time the reduced homology sequence of (K, L) (as in (2)

above) and the fact that $i_* : H_1(L) \to H_1(K)$ is an isomorphism show $H_2(K, L)$ and $H_1(K, L)$ are both trivial.

<u>6.3</u> (5X) Show that, for any (K, L),

$$H_p(K, L) = 0 \text{ for all } p \iff i_* : H_p(L) \to H_p(K) \text{ is an}$$
$$\text{isomorphism for all } p.$$

(6) Let K^1 be the 1-skeleton of K; then $H_1(K^1) = Z_1(K^1) = Z_1(K)$ by 4.9(1) and (3). Similar remarks apply to L and (K, L); applying 6.1 to the pair (K^1, L^1) we obtain an exact sequence

$$0 \to Z_1(L) \to Z_1(K) \to Z_1(K, L) \to H_0(L) \to H_0(K)$$
$$\to H_0(K, L) \to 0.$$

This was used in the proof of 4.27.

(7X) Let K be a connected oriented graph and L a sub-complex of K with components $L_1, \ldots, L_\ell$. Write $\mu(L_\alpha) = \mu_\alpha$ $(\alpha = 1, \ldots, \ell)$, thus $\mu_\alpha = \text{rank } Z_1(L_\alpha) = \text{rank } H_1(L_\alpha)$; and $\mu(K) = \mu$. Noting that $H_1(K, L)$ is free abelian (4.23(1)), show that $H_1(K, L)$ has rank $\geq \ell - 1$, and that, if the rank is $\ell - 1$ then $\mu = \mu_1 + \ldots + \mu_\ell$. (Compare 4.27.) Also, if $\mu = \mu_1 + \ldots + \mu_n$, then $i_* : H_1(L) \to H_1(K)$ is an isomorphism.

(8X) In the example of 4.31(4), calculate $H_1(K, L)$ and $H_2(K, L)$ from the reduced homology sequence of (K, L). (First verify that $H_2(K) = 0$.)

(9X) Let T be a maximal tree in the 1-skeleton of a connected simplicial complex K. Use the reduced homology sequence of (K, T) to prove that

$$j_* : H_1(K) \to H_1(K, T)$$

is an isomorphism. Show that the edges of K not in T form a basis for $Z_1(K, T)$. Thus if $e^1, \ldots, e^\mu$ are these edges,

a presentation of $H_1(K, T)$ is

$$\begin{pmatrix} e^1, \ \ldots, \ e^\mu \\ \hat\partial t \ = \ 0 \quad \text{for each 2-simplex } t \text{ of } K. \end{pmatrix}$$

Given generators for $H_1(K, T)$ how does one obtain generators for $H_1(K)$? What is the connexion between this method and that of 4.14 ?

<u>6.3</u> (10X) Use 4.38(4) to show that, in the usual notation,

$$j\Big(Z_p(K)\Big) \ = \ Z_p(K, L) \quad <=> \quad i_* \ : \ H_{p-1}(L) \ \to \ H_{p-1}(K)$$

$$\text{is a monomorphism.}$$

Taking $p = 1$, compare with 4.29(4) and 4.31(3).

THE EXCISION THEOREM

The next theorem will be used mainly in Chapter 9, though a few applications will be given here. It derives its name from the fact that it compares the relative homology groups of a pair $(L_1 \cup L_2, L_1)$ with those of a pair $(L_2, L_1 \cap L_2)$ obtained by "excising" $L_1 \setminus L_2$.
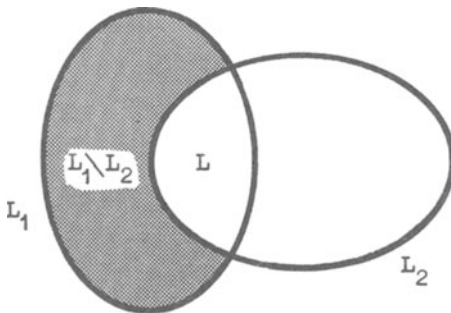
Let K be an oriented simplicial complex and L_1, L_2 subcomplexes of K with $K = L_1 \cup L_2$. Write $L = L_1 \cap L_2$. As usual we regard all chain groups as subgroups of those of K.

<u>6.4</u> <u>Theorem</u> $H_p(K, L_1) \ = \ H_p(L_2, L)$ for all p.

<u>Proof</u>   The groups $C_p(K, L_1)$ and $C_p(L_2, L)$ are clearly the same subgroup of $C_p(K)$. That is not quite the end of the

proof, because we must check that the relative boundary homo-
morphisms

$$C_p(K, L_1) \rightarrow C_{p-1}(K, L_1) \text{ and } C_p(L_2, L) \rightarrow C_{p-1}(L_2, L)$$

coincide. The result will then follow from the definition of
relative homology groups. Let

$$j : C_{p-1}(K) \rightarrow C_{p-1}(K, L_1) \text{ and } j' : C_{p-1}(L_2) \rightarrow C_{p-1}(L_2, L)$$

be the usual homomorphisms: in each case the homomorphism puts
at zero the coefficient of any simplex not in L_1, so that if
$c \in C_{p-1}(L_2)$ then $jc = j'c \in C_{p-1}(K, L_1) = C_{p-1}(L_2, L)$.
Hence, for any $c' \in C_p(K, L_1) = C_p(L_2, L)$, we have
$j\partial c' = j'\partial c'$. This $\underline{is}$ the end of the proof.   Q.E.D.

<u>6.5</u> <u>Example</u>   Let M be a 2-dimensional simplicial complex
and s a 2-simplex of M. Then removing s from M leaves
a subcomplex K say. We ask the following question: when
does $H_1(K)$ depend, up to isomorphism, only on M and not on
the particular 2-simplex s which was removed?   Now $M = K \cup \bar{s}$,
$\dot{s} = K \cap \bar{s}$, where, as usual, $\bar{s}$ and $\dot{s}$ are the set of faces
and proper faces of s, respectively. By excision
$H_p(M, K) = H_p(\bar{s}, \dot{s})$, which is isomorphic to $\mathbb{Z}$ if p = 2
and to 0 if p = 1, by an easy calculation (or see 4.23(4)).
    Consider the exact sequence (6.1)

$$H_2(K) \rightarrow H_2(M) \rightarrow H_2(M, K) \rightarrow H_1(K) \rightarrow H_1(M) \rightarrow H_1(M, K)$$
$$\mathbb{Z} \qquad\qquad\qquad\qquad\qquad 0$$

If $H_1(M)$ is free abelian and $H_2(M) = 0$ then it follows from
the exact sequence that $H_1(K) \cong H_1(M) \oplus \mathbb{Z}$ and therefore that
$H_1(K)$ does not depend on the choice of s. These conditions
are not necessary, however, for if M is a closed surface the
same result holds (this is proved in Chapter 6 in the course of
proving 6.13). (XX) The reader is invited to find an M for

which $H_1(K)$ does depend on s and to explain this dependence
from the exact sequence.

COLLAPSING REVISITED

The idea of collapsing was introduced in 3.30, but the reason
for introducing it is the following result.

6.6 Theorem   Suppose that an oriented simplicial complex K
collapses to a subcomplex L.   Then the homomorphism

$$i_* \; : \; H_p(L) \; \rightarrow \; H_p(K)$$

is an isomorphism for all p.

Proof   Suppose first that K collapses to L by an
elementary collapse, so that $K = L \cup \{s, t\}$ say where s   is
a principal q-simplex of K and t is a free face of s.
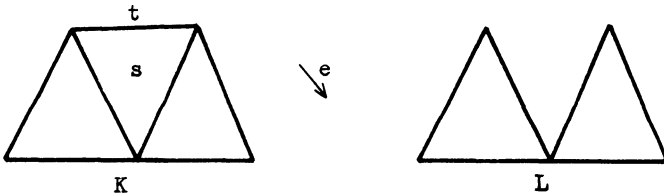
We shall prove that $H_p(K, L) = 0$ for all p;   the result then
follows immediately from the homology sequence of (K, L).

The chain complex of (K, L) contains only two nonzero
relative chain groups, namely those in dimensions q and q - 1,
and therefore takes the form

$$\ldots \; \rightarrow \; 0 \; \rightarrow \; C_q(K, L) \; \overset{\hat{\partial}}{\rightarrow} \; C_{q-1}(K, L) \; \rightarrow \; 0 \; \rightarrow \; \ldots$$

Now $C_q(K, L)$ is infinite cyclic, generated by the oriented
simplex σ corresponding to s, and $C_{q-1}(K, L)$ is also in-
finite cyclic, generated by the oriented simplex τ correspon-
ding to t.   Furthermore $\hat{\partial}\sigma = \pm \tau$ so that $\hat{\partial}$ is an iso-
morphism.   It follows from the definition of relative homology

groups that $H_p(K, L) = 0$ for all p.

For the general case, consider a sequence of elementary collapses $K \searrow^e K_1 \searrow^e K_2 \searrow^e \ldots \searrow^e K_m = L$. It is easy to verify that the homomorphism

$$i_* : H_p(L) \to H_p(K)$$

is the composite

$$H_p(K_m) \to H_p(K_{m-1}) \to \ldots \to H_p(K_1) \to H_p(K)$$

where each homomorphism is given by the inclusion of a subcomplex. Consequently i_* is an isomorphism by the first part.

<div align="right">Q.E.D.</div>

Note that this theorem does not just tell us $H_p(K)$ up to isomorphism: it says that p-cycles whose homology classes on L generate $H_p(L)$ have homology classes on K which generate $H_p(K)$.

6.7 Examples   (1) The results in 4.15 and 4.17(1) (the homology groups of Möbius band and cylinder) become obvious, using 6.6, since each simplicial complex collapses to a simple closed polygon.

(2) As a special case of 6.6, if K collapses to a point, then $H_p(K) = 0$ (p > 0); $H_0(K) \cong \mathbb{Z}$. If K is a graph, then the converse is true - see 1.12(2). However there exist 2-dimensional simplicial complexes which satisfy the homological condition just given but which cannot be collapsed at all since there is no principal simplex with a free face. An example is given by any triangulation of the "dunce hat"; see 3.22(4).

Before going on to calculate the homology groups of closed surfaces by collapsing, here is a simple application to cones.

Let K be a simplicial complex in $\mathbb{R}^N$ and let $v \in \mathbb{R}^N$ be a vector not affinely dependent on the vertices of K (if necessary add one to N in order to find such a v). The cone on K with vertex v, denoted CK, is the simplicial complex

whose simplexes are of the form

$$(v) \quad \text{or} \quad (vv^0 \ldots v^p) \quad \text{or} \quad (v^0 \ldots v^p)$$

where $(v^0 \ldots v^p) \in K$. (K is called the base of the cone.)
Thus CK is the join of K to the simplicial complex $\{v\}$
(see 3.26), and, up to isomorphism, is independent of the
choice of v. If K is oriented then CK can be oriented by
regarding the above elements of CK as oriented simplexes.

6.8 Theorem   (Homology groups of a cone)   With the above
notation,

$$H_p(CK) = 0 \quad \text{for} \quad p > 0$$

while $$H_0(CK) \cong \mathbb{Z}.$$

Proof   We prove CK collapses to a point - in fact
to v - and the result then follows from 6.6. If K is
empty then $CK = \{v\}$, so there is nothing to collapse (or to
prove). Otherwise we proceed by induction on the dimension
of K. If dim K = n, then any (n + 1)-simplex $(vv^0 \ldots v^n)$
of CK is principal, and $(v^0 \ldots v^n)$ is a free face of it.
Hence all the (n + 1)-simplexes of CK can be collapsed,
removing the n-simplexes of K and therefore leaving a cone on
a simplicial complex of dimension n - 1. If dim K = 0 the
same procedure collapses CK directly to the vertex of the
cone.   Q.E.D.

6.9 Corollaries   (1) Suppose that L is a subcomplex of
K and that L is a cone. Then

$$j_* : H_p(K) \to H_p(K, L) \quad p > 0$$

and $$\tilde{j}_* : \tilde{H}_0(K) \to H_0(K, L) \quad \text{(see 6.2)}$$

are isomorphisms.

    (2) For any K,

$$H_p(CK, K) \cong H_{p-1}(K) \quad p > 1$$

$$H_1(CK, K) \cong \tilde{H}_0(K)$$

6.9 (3) (<u>Homology groups of a simplex</u>)   Let s be an n-simplex $(n \geqslant 0)$ and $K = \bar{s}$ (oriented arbitrarily).   Then $H_p(K) = 0$ for $p > 0$ and $H_0(K) \cong \mathbb{Z}$.

<u>Proofs</u>   (1) The homology sequence of the pair (K, L) contains the section

$$H_p(L) \rightarrow H_p(K) \overset{j_*}{\rightarrow} H_p(K, L) \rightarrow H_{p-1}(L) \; .$$

For $p > 1$, both end-groups are zero, by 6.8, and the result follows from exactness.   For $p = 0$, $p = 1$ the result follows from a similar application of the reduced homology sequence.

(2) The proof is similar to that of (1).

(3) K is in fact a cone, with empty base if $n = 0$ and with base the closure of an $(n - 1)$-dimensional face of s if $n > 0$.   Thus the result follows from 6.8.   Q.E.D.

Using the excision theorem and 6.8 we can establish the following interpretation of relative homology groups as the absolute homology groups of an associated simplicial complex. The idea is that the relative homology groups of the pair (K, L) in some sense disregard L, and homologically this can be done by adding a cone on L.   Thus let $K \cup CL$ be obtained by
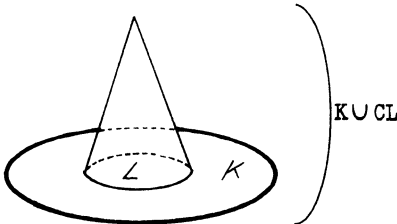
taking the union of K and a cone CL on L, the cone meeting K precisely on L as in the sketch.

6.10 <u>Theorem</u>   With the above notation

$$H_p(K, L) \cong H_p(K \cup CL) \quad p > 0$$

and $$H_0(K, L) \cong \tilde{H}_0(K \cup CL)$$

(N.B. If L is empty, then CL is a single point not in $|K|$.)

<u>Proof</u>   By the excision theorem 6.4,

$$H_p(K \cup CL, CL) \; = \; H_p(K, K \cap CL) \; = \; H_p(K, L)$$

for all p. The result follows from 6.9(1).   Q.E.D.

<u>6.11 Examples</u>   (1) Let K be the Möbius band given in
4.15 and let L be the rim of the band. Adding a cone on L
gives precisely the projective plane in 4.17(5). The relative
homology groups of (K, L) are given in 4.29(3). (It is
difficult to visualize K ∪ CL in this case because a Möbius
band cannot be realized in $\mathbb{R}^3$ in such a way that the circular
rim is a plane polygon with the band all on one side of it.
With a possibly larger number of triangles in the Möbius band
we can do the construction in $\mathbb{R}^4$; compare 2.16(5).)

    (2X) Let K be connected, and let the subcomplex L have
components $L_1, \ldots, L_\ell$. Then we can form a simplicial com-
plex M by adding disjoint cones on the subcomplexes
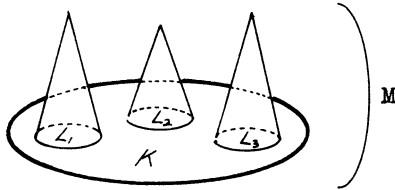
$L_1, \ldots, L_\ell$ –
this differs from
K ∪ CL in that
the vertices of
the cones are all
distinct.

Show that $H_p(K, L) \; \cong \; H_p(M)$ p > 1

and $H_1(K, L) \; \cong \; H_1(M) \; \oplus \; \mathbb{Z}^{\ell-1}$.

    As an example of this let K be the cylinder given in
4.29(1) and L_1 and L_2 the two ends. Then M is a closed
surface which is in fact a sphere, i.e. has standard form
aa^{-1} (2.8). The relative homology groups of $(K, L_1 \cup L_2)$
are in 4.29(1) and the homology groups of any closed surface
are given in 4.18 and 6.13 below.

6.11 (3X) Suppose that K, L and M are simplicial complexes
of dimension 0, each consisting of $n \geqslant 1$ vertices. Suppose
that K and L are joinable, and also that KL and M are
joinable (3.26); denote the join of KL and M by KLM.
(Thus when $n = 2$ the join KLM is an octahedron.) Show
that $H_1(KLM) = 0.$ (Hint. Take a maximal tree in the
graph KL and extend it to a maximal tree in the 1-skeleton of
KLM. Show, using 6.8, that the basic 1-cycles (4.9(2)) are
homologous to zero.)

(4) Let S^n denote the boundary (i.e. the set of proper
faces) of an $(n + 1)$-simplex $\sigma_{n+1}.$ Thus S^0 is two points,
S^1 is a "hollow" triangle, S^2 is a hollow tetrahedron.
(In general, S^n is homeomorphic to
$\{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + \ldots + x_{n+1}^2 = 1\}$ and can be
called an n-_sphere_.) The homology groups of S^0 present no
trouble, so suppose $n > 0$ and consider the following two sub-
complexes of $S^n.$ Let v be a vertex of σ_{n+1} and let σ_n
be the n-simplex whose vertices are the remaining $n + 1$ ver-
tices of $\sigma_{n+1}.$

Define L_1 = the set of all faces of $\sigma_n.$

L_2 = the cone with vertex v on the set of all
proper faces of $\sigma_n.$ (The diagram shows
$n = 2.$)
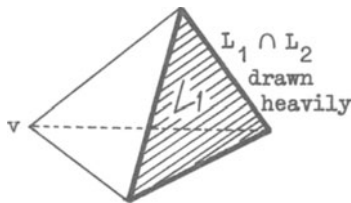
Then $S^n = L_1 \cup L_2,$
$S^{n-1} = L_1 \cap L_2.$
By the excision theorem,

$$H_p(S^n, L_1) \cong H_p(L_2, S^{n-1})$$

$$\cong \begin{cases} H_{p-1}(S^{n-1}) & p > 1 \\ \tilde{H}_0(S^{n-1}) & p = 1 \end{cases}$$

by 6.9(2). Also $\tilde{H}_0(S^{n-1}) = 0$ $(n > 1)$; $\cong \mathbb{Z}$ $(n = 1)$.
Using 6.9(1) and the fact that L_1 is a cone, it follows from
the above results that

$$H_p(S^n) \cong H_{p-1}(S^{n-1}) \qquad p > 1$$

$$H_1(S^n) \cong \begin{cases} 0 & n > 1 \\ \mathbb{Z} & n = 1 \end{cases}$$

The second of these results, together with $H_p(S^1) = 0$ $(p > 1)$,
fills in the two lines in the table; the remaining entries follow from the first result. We have: for $n \geqslant 1$,
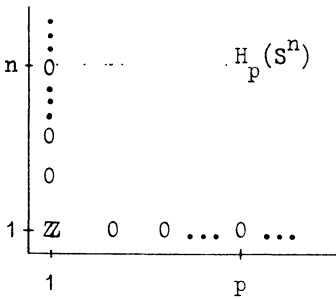
$$n \begin{cases} \vdots \\ 0 \quad \cdots \qquad\qquad H_p(S^n) \\ \vdots \\ 0 \\ 0 \end{cases}$$
$$1 \mid \mathbb{Z} \quad 0 \quad 0 \ \ldots \ 0 \ \ldots$$
$$\qquad 1 \qquad\qquad p$$

$$H_p(S^n) = \begin{cases} 0 & 0 < p < n \\ \mathbb{Z} & p = 0 \text{ or } n \ . \end{cases}$$

An n-cycle whose homology class generates $H_n(S^n)$ contains all
the n-simplexes of S^n, oriented "coherently", i.e. in such
a way that their boundaries add up to zero. For example, S^2
is a closed surface, and such an n-cycle is a fundamental cycle
for the surface (4.18).

<div align="center">HOMOLOGY   GROUPS   OF   CLOSED   SURFACES</div>

We shall now use the method of collapsing to calculate the first
homology groups of all the closed surfaces found in Chapter 2.
This cannot be done directly, since a closed surface has no
principal simplex with a free face.   However we can remove
a 2-simplex and, using 3.34, collapse what is left to a graph,
whose first homology group is easy to calculate.   We need to
know the effect of removing a 2-simplex, and that is the content of the next Lemma.

    Let M be a closed surface, and let K be obtained from
M by removing a single 2-simplex s; thus K is a subcomplex

of M. Let z be the 1-cycle on K going once round the
boundary of s and let A be the subgroup of $H_1(K)$ generated
by $\{z\}_K$. (Of course $\{z\}_M = 0$ since z bounds s.)

6.12 <u>Lemma</u> (1) Suppose that M is orientable. Then

$$i_* : H_1(K) \to H_1(M)$$

is an isomorphism.

(2) Suppose that M is non-orientable. Then there is a
short exact sequence

$$0 \to A \xrightarrow{\text{inclusion}} H_1(K) \xrightarrow{i_*} H_1(M) \to 0 .$$

It follows that $H_1(M) \cong H_1(K)/A$ and generators for
$H_1(M)$ can be obtained by taking the images under i_* of gen-
erators for $H_1(K)$.

<u>Proof</u>   Consider the homology sequence of the pair
(M, K):

$$H_2(M) \xrightarrow{j_*} H_2(M, K) \xrightarrow{\partial_*} H_1(K) \xrightarrow{i_*} H_1(M) \to H_1(M, K)$$

$$\| \qquad\qquad\qquad\qquad\qquad\qquad\qquad \|$$

$$H_2(\bar{s}, \dot{s}) \cong \mathbb{Z} \qquad\qquad\qquad\qquad H_1(\bar{s}, \dot{s}) = 0$$

Here $\bar{s}$ and $\dot{s}$ are respectively the set of faces of s and
the set of proper faces - the boundary - of s. The equal-
ities follow by excision, and by 4.23(4), which gives the homo-
logy groups of $(\bar{s}, \dot{s})$.

(1) Suppose that M is orientable. Let σ be the
oriented 2-simplex of M corresponding to s. Now M has a
coherent orientation in which s receives the orientation of
σ; also $H_2(M) \cong \mathbb{Z}$, with generator the sum of all coherently
oriented 2-simplexes of M (4.18). Since j_* suppresses all
2-simplexes of M but σ, it follows that j_* will take the
generator of $H_2(M)$ to $\{\sigma\}$, the homology class of σ in

$H_2(M, K) = H_2(\bar{s}, \mathcal{S})$. However $\{\sigma\}$ generates $H_2(\bar{s}, \mathcal{S})$ so j_* is an isomorphism. Hence $\partial_* = 0$ and i_* is an isomorphism by exactness.

(2) Suppose that M is non-orientable. Then $H_2(M) = 0$ (4.18).   As in (1), $H_2(M, K) \cong \mathbb{Z}$, with generator $\{\sigma\}$. Now $\partial_*\{\sigma\} = \{\partial\sigma\}_K = \{\pm z\}$, i.e. Im $\partial_* = A$. This proves the result, since Im $\partial_* = $ Ker i_*.   Q.E.D.

Now consider a closed surface M represented in the standard form of 2.8. According to 3.34, removing a 2-simplex produces a simplicial complex K which collapses to the graph G in M given by the boundary of the polygonal region. Note that this collapse really does occur in K since the only identifications to be made in passing from the polygonal representation to the closed surface are identifications round the boundary polygon, and this is not touched during the collapse. It would not be valid, for instance, to invoke 3.32 to prove that M collapses to a point, since the collapsing of a disk starts from the boundary polygon, and the edges of M around the polygon are not in reality free.

We must now identify the graph G. Let us consider the case 2.8(2) for definiteness. Then the symbol is $a_1 a_1 \ldots a_k a_k$. Each a_i will contain several vertices and edges of M, but the 2k vertices of the polygon at the endpoints of the a_i are all the same vertex v of M. (For $v_1 = v$ because they are both at the beginning of a_1, $v_2 = v_1$ because they are both at the end of a_1, and so on.) Thus each a_i gives a loop on M and all these loops have the vertex v of M in common. The case 2.8(3) is similar, and 2.8(1), the sphere aa^{-1},
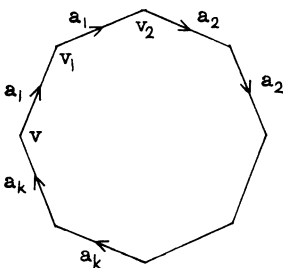
differs only in that the end-points of a are distinct ver-
tices of M.   Thus the graph G to which K collapses is as
follows:

(1) a simple polygonal arc in the case of the sphere
aa^{-1};

(2) k loops, given by a_1, ..., a_k, with a single ver-
tex in common, in the case of the non-orientable closed surface
$a_1 a_1 \ldots a_k a_k$;

(3) 2h loops, given by a_1, b_1, ..., a_h, b_h, with a
single vertex in common, in the case of the orientable closed
surface $a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_h b_h a_h^{-1} b_h^{-1}$.   (See 3.35(1) for the
case h = 1.)

Since K collapses to G, the map $i_* : H_1(G) \to H_1(K)$
is an isomorphism by 6.12.   The group $H_1(G)$ is isomorphic to
$\mathbb{Z}^{\mu(G)}$ and has generators given by the basic loops on G (see
4.9(3)).   Thus

$$H_1(K) \cong \begin{cases} 0 & \text{in Case 1} \\ \mathbb{Z}^k, \text{ and is generated by } \{a_1\}, \ldots, \{a_k\} & \text{in Case 2} \\ \mathbb{Z}^{2h}, \text{ and is generated by } \{a_1\}, \{b_1\}, \ldots, \{a_h\}, \{b_h\} & \\ & \text{in Case 3} \end{cases}$$

When M is non-orientable (Case 2), we need to know the
homology class on K represented by the 1-cycle going round the
boundary of a triangle s in M.   Then we can apply 6.12(2).
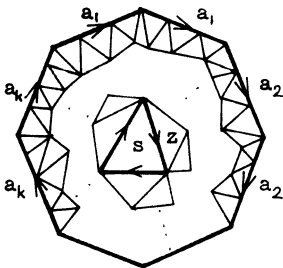
The inside of the polygon
is broken up into tri-
angles, and by adding up
the boundaries of all the
triangles except s,
oriented clockwise in the
picture, we have

$$a_1 + a_1 + a_2 + a_2 + \ldots + a_k + a_k \sim z \text{ in K.}$$

Thus the group A of 6.12(2) is generated by
$2\{a_1\} + \ldots + 2\{a_k\} \in H_1(K)$. (There is   little difficulty
proving that the triangles of a triangulation of a disk can be
oriented coherently in this way: an induction similar to that
in 3.32 will work. Of course this does not coherently orient
M — an impossibility — since the orientations are not co-
herent across the edges of M making up any of the a_i.)

6.13 Theorem   Let M be a closed surface. Then $H_1(M)$ is
as follows.

    (1) $H_1(M) = 0$ when M has standard form aa^{-1}.

    (2) $H_1(M) \cong \mathbb{Z}^{k-1} \oplus \mathbb{Z}_2$ for the non-orientable closed
surface $a_1 a_1 \ldots a_k a_k$. The classes $\{a_1\}, \ldots, \{a_k\}$ on M
generate $H_1(M)$ and the unique element of order two is
$\{a_1\} + \ldots + \{a_k\}$.

    (3) $H_1(M) \cong \mathbb{Z}^{2h}$ for the orientable closed surface
$a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_h b_h a_h^{-1} b_h^{-1}$. The classes $\{a_1\}, \{b_1\}, \ldots, \{a_h\}, \{b_h\}$
freely generate $H_1(M)$.

Proof   The cases (1) and (3) follow at once from 6.12(1)
and the group $H_1(K)$, given above. Now for the case (2).
The exact sequence of 6.12(2) shows that $H_1(M)$ is generated
by the images under i_* of the generators of $H_1(K)$, i.e.
by $i_*\{a_j\}_K = \{a_j\}_M$ for $j = 1, \ldots, k$. Write $\alpha_j = \{a_j\}_K$.
Now $H_1(M) \cong H_1(K)/A$ and this is the quotient group of the
free abelian group generated by $\alpha_1, \ldots, \alpha_k$, by the subgroup
generated by 2α, where $\alpha = \alpha_1 + \ldots + \alpha_k$. Changing gener-
ators for $H_1(K)$ to $\alpha_1, \ldots, \alpha_{k-1}, \alpha$, the quotient has
presentation
$$\left(\begin{matrix} \alpha_1, \ldots, \alpha_{k-1}, \alpha \\ 2\alpha = 0 \end{matrix} \right.$$

and is therefore isomorphic to $\mathbb{Z}^{k-1} \oplus \mathbb{Z}_2$. The unique element
of order two is the coset $\alpha + A$ corresponding to α, so the
unique element of order two in $H_1(M)$ is

$$i_* \alpha = \{a_1 + \ldots + a_k\}_M = \{a_1\}_M + \ldots + \{a_k\}_M . \qquad \text{Q.E.D.}$$

Thus two closed surfaces M_1 and M_2 have the same standard form (i.e. are equivalent) if and only if $H_1(M_1) \cong H_1(M_2)$: the first homology group is a sufficiently delicate invariant to distinguish the closed surfaces from one another. (In contrast, it does not distinguish a cylinder from a Möbius band, nor either of these from a simple closed polygon!)

6.14 Examples   (1X) If M is a sphere with one crosscap (projective plane) then $H_1(M) \cong \mathbb{Z}_2$. This agrees with 4.17(5). For M a sphere with two crosscaps (Klein bottle), $H_1(M) \cong \mathbb{Z} \oplus \mathbb{Z}_2$. The usual representation of a Klein bottle is shown on the left (compare 2.2(4)) and the standard one on the right. Cutting along the diagonal b in the right-hand diagram and sticking together the two a_1's produces the left-hand picture with $c = a_2$ and $d = b$. Also $b \sim a_1 + a_2$ (add up the boundaries of triangles in the "top half" of the right-hand picture), so by 6.13(2) $H_1(M)$ is generated by $\{c\}$ and $\{d\}$, and $\{d\}$ is the unique element of order two. To see that $d + d \sim 0$, note that $d \sim d'$ from the left half of the left-hand picture and $d' \sim -d$ from the right half. (Compare 4.29(3).)
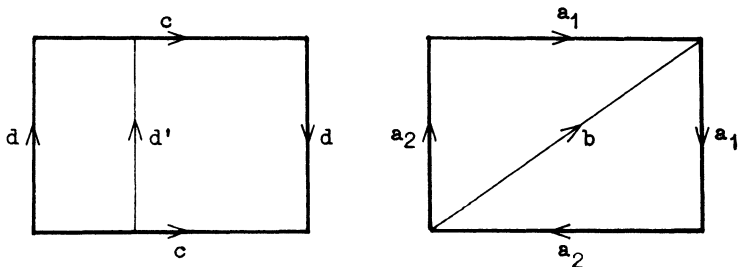
Let z be the 1-cycle going across the left-hand diagram between the mid-points of d, parallel to c. What is $\{z\}$ in terms of $\{c\}$ and $\{d\}$? What is $2\{z\}$? (In particular $2\{z\} \neq 0$.)

<u>6.14</u> (2) Taking h = 2 in 6.13(3), possible positions for
a_1, b_1, a_2, b_2 on a double torus are shown in the diagram on
the left.   These 1-cycles are homologous to the four shown in
the diagram on the right. These four accordingly represent free
generators for $H_1(M)$.
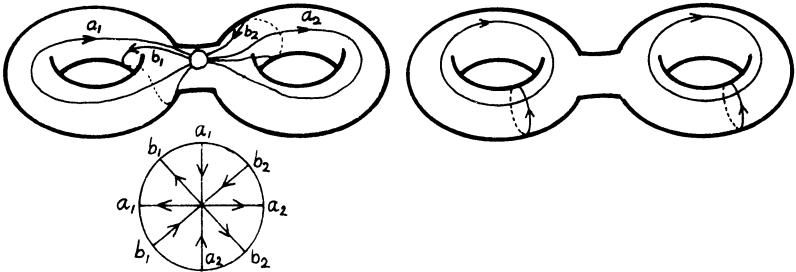
Similarly the 1-cycles drawn below, for arbitrary h, are homo-
logous to $a_1, b_1, \ldots, a_h, b_h$ and so represent free generators
for $H_1(M)$.
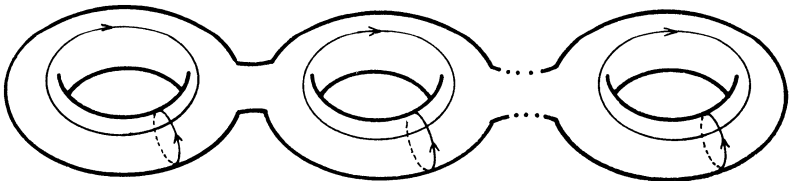
(3X) The method used to prove 6.13 can of course be used
to calculate the first homology groups of a closed surface
given in any polygonal representation, not just the standard
one.   The reader may care to try the examples of 2.16(1)-(4)
directly from the representations given there.

THE EULER CHARACTERISTIC

In this section we shall establish a connexion between the
Euler characteristic of a simplicial complex (i.e. the alter-
nating sum of the numbers of p-simplexes) and the homology
groups of the simplicial complex. In particular this implies
using the invariance theorem 5.13, that the Euler character-
istic is a topological invariant of a simplicial complex: it
is unaffected by homeomorphisms. It is interesting to note
that though the Euler characteristic can be defined without
mentioning homology groups, the only known way of proving its
topological invariance is by the use of homology theory. The
invariance theorem says that two simplicial complexes with
different Euler characteristics cannot have homeomorphic under-
lying spaces. The converse it not true, even for closed sur-
faces, for hT and 2hP have the same characteristic (namely
2 - 2h) for any h > 0.

6.15 Definition   Let K be a simplicial complex of dimension
n and L a subcomplex of K. Let $\alpha_p(K)$ and $\alpha_p(K, L)$
denote the numbers of p-simplexes in K and in $K \setminus L$ respec-
tively. Then

$$\chi(K) = \sum_{p=0}^{n} (-1)^p \alpha_p(K) \quad \text{and} \quad \chi(K, L) = \sum_{p=0}^{n} (-1)^p \alpha_p(K, L)$$

are called the Euler (or Euler-Poincaré) characteristics of K
and of K mod L respectively. (Compare 2.11)

The connexion between these characteristics and the homol-
ogy groups of K and of (K, L) is given in the next theorem,
in which the K of 6.15 is assumed oriented. Let

$$\beta_p(K) = \text{rank } H_p(K) \quad \text{and} \quad \beta_p(K, L) = \text{rank } H_p(K, L) .$$

(See A.29 for the definition of rank of a finitely generated
abelian group.) These numbers are called the pth Betti numbers

of K and of K mod L respectively.

6.16 Theorem

$$\chi(K) = \sum_{p=0}^{n} (-1)^p \beta_p(K) \quad ; \quad \chi(K, L) = \sum (-1)^p \beta_p(K, L).$$

Proof   The result is a purely algebraic one.   Given
any sequence of f.g. abelian groups

$$0 \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} \dots \to C_p \xrightarrow{d_p} \dots$$
$$\to C_1 \xrightarrow{d_1} C_0 \xrightarrow{d_0} 0$$

where $d_{p-1} \circ d_p = 0$ for $p = 1, \dots, n$, we can define
$H_p = \operatorname{Ker} d_p / \operatorname{Im} d_{p+1}$ $(p = 0, \dots, n)$.   (Compare 4.36)
Writing $\alpha_p = \operatorname{rank} C_p$, $\beta_p = \operatorname{rank} H_p$, we shall prove
$\sum_{p=0}^{n} (-1)^p \alpha_p = \sum_{p=0}^{n} (-1)^p \beta_p$.   The results required are then
immediate.

    We have short exact sequences

$$0 \to \operatorname{Ker} d_p \to C_p \xrightarrow{d_p} \operatorname{Im} d_p \to 0$$

and $0 \to \operatorname{Im} d_{p+1} \to \operatorname{Ker} d_p \to H_p \to 0$.

Thus by A.31, $\alpha_p = \operatorname{rank} \operatorname{Ker} d_p + \operatorname{rank} \operatorname{Im} d_p$,

and $\beta_p = \operatorname{rank} \operatorname{Ker} d_p - \operatorname{rank} \operatorname{Im} d_{p+1}$

so that $\alpha_p - \beta_p = \operatorname{rank} \operatorname{Im} d_p + \operatorname{rank} \operatorname{Im} d_{p+1}$ and
$\sum_{p=0}^{n} (-1)^p (\alpha_p - \beta_p) = 0$ since d_0 and d_{n+1} are both trivial.
                                                          Q.E.D.

    Notice that A.31(3) is the special case of this result
which arises when the long sequence is exact, in which case
$H_p = 0$ for all p.

<u>6.17</u> <u>Examples</u>   (1) From the definitions of chain groups it
follows that, in the notation of 6.15, $\chi(K) = \chi(L) + \chi(K, L)$.
This says something not quite so obvious about homology groups,
in view of 6.16.

Similarly if L_1 and L_2 are subcomplexes of a simplicial
complex K then clearly

$$\chi(L_1 \cup L_2) = \chi(L_1) + \chi(L_2) - \chi(L_1 \cap L_2).$$

(2X) Using 6.16 and the calculated homology groups of
closed surfaces in 4.18 and 6.13 we can check the values of
their Euler characteristics given in 2.14.

(3X) In the notation of 6.11(3), we can calculate the
Euler characteristic of KLM directly.   In fact (compare
3.27(5))

$$\alpha_0(KL) = 2n \; ; \; \alpha_1(KL) = n^2$$

$$\alpha_0(KLM) = 3n \; ; \; \alpha_1(KLM) = n^2 + 2n^2 = 3n^2 \; ; \; \alpha_2(KLM) = n^3 \, .$$

Hence $\chi(KLM) = 3n - 3n^2 + n^3$. Since $H_0(KLM) \cong \mathbb{Z}$ and
$H_1(KLM) = 0$, it follows from 6.16 that rank $H_2(KLM) = (n-1)^3$.
However $H_2(KLM)$ is free abelian since KLM is 2-dimensional
(compare 4.9(3)), so that $H_2(KLM) \cong \mathbb{Z}^{(n-1)^3}$.

(4X) Let s be an n-simplex and let $K = \bar{s}$.   Denote as
usual by K^q the q-skeleton of K.   Now from 6.9(3),
$H_p(K) = 0$ if $0 < p < n$; consequently (4.9(1)) $H_p(K^q) = 0$
if $0 < p < q \leqslant n$.   Also $\alpha_p(K) = \binom{n+1}{p+1}$ since every choice
of p + 1 vertices from among the vertices of s spans a
p-simplex of K.   In an analogous way to (3) above, show that
$H_q(K^q)$ is a free abelian group of rank $\binom{n}{q+1}$ if $0 < q < n$.
In particular putting q = n - 1 we get an alternative proof
of 6.11(4): $H_{n-1}(S^{n-1}) \cong \mathbb{Z}$ for all n > 1.

Two more general theorems

The theorems proved in this chapter are separated from those of Chapter 6 on account of their slightly more technical nature and the fact that we do not use them in an essential way in what follows. Both theorems assert the exactness of certain sequences. The first attempts to answer the question: given the homology groups of two subcomplexes L_1, L_2 of a simplicial complex K (not necessarily disjoint), what are the homology groups of $L_1 \cup L_2$? They obviously depend on the homology groups of $L_1 \cap L_2$ as well, and moreover on the way in which L_1 and L_2 are stuck together - for example two cylinders can be stuck together to give either a torus or a Klein bottle (see 7.5). The result is not an explicit formula for $H_p(L_1 \cup L_2)$ but an exact sequence which, with luck, will give a good deal of information. An example where it does not give quite enough information to determine a homology group of $L_1 \cup L_2$ is mentioned in 7.6(3).

The other theorem proved in this chapter, the exactness of the "homology sequence of a triple", is a generalization of the homology sequence of pair (see 6.1).

THE MAYER-VIETORIS SEQUENCE

Let K be an oriented simplicial complex of dimension n, and let L_1, L_2 be subcomplexes of K with $L_1 \cup L_2 = K$. Write $L_1 \cap L_2 = L$ and, as usual, regard $C_p(L)$, $C_p(L_1)$ and $C_p(L_2)$ as subgroups of $C_p(K)$, for each p.

7.1 Definitions

(1) Define a homomorphism

$$\phi : H_p(L) \to H_p(L_1) \oplus H_p(L_2)$$

by
$$\{z\}_L \to (\{z\}_{L_1}, -\{z\}_{L_2})$$

(2) Define a homomorphism

$$\psi : H_p(L_1) \oplus H_p(L_2) \to H_p(K)$$

by
$$(\{z^1\}_{L_1}, \{z^2\}_{L_2}) \to \{z^1 + z^2\}_K .$$

(3) Suppose that $z \in Z_p(K)$, and that $z = c^1 + c^2$ where $c^1 \in C_p(L_1)$ and $c^2 \in C_p(L_2)$. Then $0 = \partial z = \partial c^1 + \partial c^2$, so that $\partial c^1 = -\partial c^2$ and both ∂c^1 and ∂c^2 must belong to $C_{p-1}(L)$ — indeed to $Z_{p-1}(L)$ since $\partial \circ \partial = 0$.
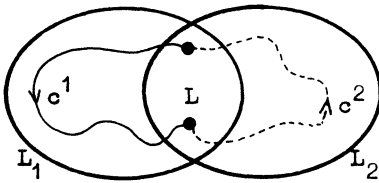
Define a homomorphism

$$\Delta : H_p(K) \to H_{p-1}(L)$$

by
$$\{z\}_K \to \{\partial c^1\}_L .$$

To prove this makes sense, suppose that $\{w\}_K = \{z\}_K$, and that $w = c^3 + c^4$ where $c^3 \in C_p(L_1)$ and $c^4 \in C_p(L_2)$. We have to show $\{\partial c^1\}_L = \{\partial c^3\}_L$. Now $z - w = \partial c$ for some $c \in C_{p+1}(K)$, so that writing $c = c^5 + c^6$ ($c^5 \in C_{p+1}(L_1)$,

$c^6 \in C_{p+1}(L_2))$, we have $c^1 - c^3 - \partial c^5 = -(c^2 - c^4 - \partial c^6)$.
Accordingly both sides of this equation belong to $C_p(L)$,
whence

$$\partial(c^1 - c^3 - \partial c^5) = \partial c^1 - \partial c^3 \in B_{p-1}(L),$$

i.e. $\{\partial c^1\}_L = \{\partial c^3\}_L$ as required.

The verifications that (1) and (2) also make sense are
left to the reader. We now fit the three homomorphisms
ϕ, ψ, Δ into an exact sequence, called the **Mayer-Vietoris**
Sequence of the triad $(K; L_1, L_2)$.

7.2 Theorem   The following sequence is exact.

$$0 \to H_n(L) \overset{\phi}{\to} H_n(L_1) \oplus H_n(L_2) \overset{\psi}{\to} H_n(K) \overset{\Delta}{\to} H_{n-1}(L)$$

$$\overset{\phi}{\to} H_{n-1}(L_1) \oplus H_{n-1}(L_2) \overset{\psi}{\to} \ldots \overset{\Delta}{\to} H_0(L) \overset{\phi}{\to} H_0(L_1) \oplus H_0(L_2)$$

$$\overset{\psi}{\to} H_0(K) \to 0.$$

Proof   (1) Exactness at $H_p(L)$. Let $z \in Z_{p+1}(K)$
and $z = c^1 + c^2$ where $c^1 \in C_{p+1}(L_1)$, $c^2 \in C_{p+1}(L_2)$. Then

$$\phi\Delta\{z\} = (\{\partial c^1\}_{L_1}, -\{\partial c^1\}_{L_2}) = (\{\partial c^1\}_{L_1}, \{\partial c^2\}_{L_2}) = (0, 0).$$

Hence $\operatorname{Im} \Delta \subset \operatorname{Ker} \phi$.
    Conversely, suppose that $u \in Z_p(L)$ satisfies
$\phi\{u\} = (0, 0)$, i.e. $\{u\}_{L_1} = 0$ and $\{u\}_{L_2} = 0$. Thus
$u = \partial x^1 = \partial x^2$, so $\partial(x^1 - x^2) = 0$, i.e. $x^1 - x^2 \in Z_p(K)$.
Furthermore $\Delta\{x^1 - x^2\}_K = \{\partial x^1\}_L = \{u\}_L$. Hence
$\operatorname{Ker} \phi \subset \operatorname{Im} \Delta$.

    (2) Exactness at $H_p(L_1) \oplus H_p(L_2)$. Let $z \in Z_p(L)$;
then $\psi\phi\{z\} = \{z - z\}_L = 0$. Hence $\operatorname{Im} \phi \subset \operatorname{Ker} \psi$.
    Conversely, suppose that $(\{u^1\}, \{u^2\}) \in H_p(L_1) \oplus H_p(L_2)$
lies in $\operatorname{Ker} \psi$, i.e. that $\{u^1 + u^2\}_K = 0$. Thus
$u^1 + u^2 = \partial x$ for some $x \in C_{p+1}(K)$. Let

$\rho_\alpha : C_{p+1}(K) \to C_{p+1}(L_\alpha)$, $\alpha = 1, 2$ be the restriction homomorphisms which change to zero the coefficient of any $(p + 1)$-simplex not in L_α ; I claim that $u^1 - \partial\rho_1 x \in Z_p(L)$ and that $\phi\{u^1 - \partial\rho_1 x\} = (\{u^1\}, \{u^2\})$. To see this, note that

$$x - \rho_1 x - \rho_2 x \in C_{p+1}(L)$$

so that
$$\partial x - \partial\rho_1 x - \partial\rho_2 x \in B_p(L)$$

i.e.
$$(u^1 - \partial\rho_1 x) + (u^2 - \partial\rho_2 x) \in B_p(L) .$$

But the first bracket belongs to $C_p(L_1)$ and the second to $C_p(L_2)$; hence both belong to $C_p(L)$ and indeed since u^1 and u^2 are cycles both brackets belong to $Z_p(L)$. Finally $u^\alpha - \partial\rho_\alpha x \sim u^\alpha$ on L_α $(\alpha = 1, 2)$ and $-(u^1 - \partial\rho_1 x) \sim u^2 - \partial\rho_2 x$ on L, and hence on L_2. Applying the definition of ϕ the required result follows. Thus Ker $\psi \subset$ Im ϕ . (The reader may like to draw a diagram to illustrate this inclusion, say in the case $p = 1$.)

(3) Exactness at $H_p(K)$. Let $(\{z^1\}, \{z^2\}) \in H_p(L_1) \oplus H_p(L_2)$. Then $\Delta\psi(\{z^1\}, \{z^2\})$ $= \{\partial z^1\}_L = 0$ since z^1 is a cycle. Hence Im $\psi \subset$ Ker Δ.

Conversely, suppose that $u \in Z_p(K)$ satisfies $\Delta\{u\} = 0$, i.e. writing $u = c^1 + c^2$ in the usual way, that $\{\partial c^1\}_L = 0$. This means that $\partial c^1 = \partial x$ for some $x \in C_p(L)$. Then $c^1 - x \in Z_p(L_1)$ and $c^2 + x \in Z_p(L_2)$ since u is a cycle. Since

$$\psi(\{c^1 - x\}_{L_1}, \{c^2 + x\}_{L_2}) = \{c^1 + c^2\}_K = \{u\}_K$$

this shows that Ker $\Delta \subset$ Im ψ .   Q.E.D.

7.3 Remark   There is a reduced form of the Mayer-Vietoris sequence in which 0-dimensional homology groups are replaced by reduced groups (4.8) and Δ, ϕ and ψ are suitably altered at the lower end of the sequence. The resulting sequence is still exact.

<u>7.4</u> <u>Examples</u>   (1) If L is empty, then K is the union
of disjoint subcomplexes L_1, L_2 and from the Mayer-Vietoris
sequence

$$H_p(K) \cong H_p(L_1) \oplus H_p(L_2)$$

for all p.   This result extends in an obvious way to the
union of any number of disjoint subcomplexes, and it is not
hard to give a direct proof.

(2X) Suppose that L is connected.   Then

$$\phi : H_0(L) \rightarrow H_0(L_1) \oplus H_0(L_2)$$

is monic.   (If L is empty there is nothing to prove; other-
wise let v be a vertex of L so that $\{v\}$ generates the
infinite cyclic group $H_0(L)$.   Now

$$\phi\{v\} = (\{v\}_{L_1}, -\{v\}_{L_2}) \neq (0, 0)$$

and since $H_0(L_1) \oplus H_0(L_2)$ has no non-zero element of finite
order this is the result.)

(3X) Suppose that L consists of a single point.   Then
$H_p(L) = 0$ for $p > 0$ so that from the Mayer-Vietoris sequence
$H_p(K) \cong H_p(L_1) \oplus H_p(L_2)$ for $p > 1$.   From (2) it follows
that this holds for $p = 1$ as well.   What about $p = 0$?

(4X) Suppose that L consists of two points, v and w
say.   Show that

$$\phi : H_0(L) \rightarrow H_0(L_1) \oplus H_0(L_2)$$

has infinite cyclic kernel generated by $\{v - w\}$ if v and
w lie in the same component of L_1 and in the same component
of L_2, and is monic otherwise.   (Note that if v and w
lie in different components of L_1 then $\lambda v + \mu w \sim 0$ on
L_1 implies $\lambda = \mu = 0$.   Compare the proof of 4.11.)   Hence

show that

$$H_p(K) \cong H_p(L_1) \oplus H_p(L_2), \qquad p > 1$$

$$H_1(K) \cong H_1(L_1) \oplus H_1(L_2) \oplus A$$

where $A = 0$ or $A \cong \mathbb{Z}$. In particular $A \cong \mathbb{Z}$ when K is connected, and the diagram shows where the extra 1-cycle comes from. Find examples to illustrate the case $A = 0$.
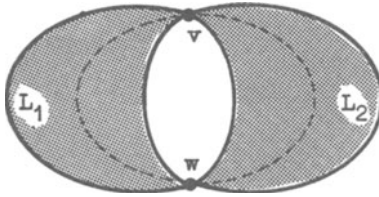
$\underline{7.4}$ (5X) Some earlier results can be reproved using the Mayer-Vietoris sequence, for instance 6.11(4) (homology groups of S^n), 6.6 (collapsing) and 6.12 (effect of removing a 2-simplex from a closed surface). The reader is invited to try these examples, and to look for others in Chapter 6 amenable to this alternative treatment.

(6) Let K be an oriented simplicial complex in $\mathbb{R}$, entirely contained in the subspace $x_1 = 0$ (where $(x_1, \ldots, x_N)$ are coordinates in $\mathbb{R}^N$). Let C_1 be the cone on K with vertex $(1, 0, \ldots, 0)$ and C_2 the cone on K with vertex $(-1, 0, \ldots, 0)$, so that $C_1 \cap C_2 = K$. We call $C_1 \cup C_2$ a $\underline{\text{suspension}}$ of K and denote it by SK. Show that

$$H_p(SK) \cong H_{p-1}(K) \quad (p > 1)$$

$$H_1(SK) \cong \tilde{H}_0(K).$$
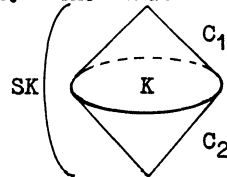
Which closed surfaces, if any, are suspensions of graphs?

The remaining examples will be of a more informal nature than those given above, in that we shall not worry very much about the triangulations of the objects in question but rather witness the Mayer-Vietoris sequence in action by sticking together simple objects to make more complicated ones. In this

way I hope that the geometrical content of the sequence will become clearer.

<u>7.5 Example</u>   If we identify the ends of two cylinders, the result can be either a torus or a Klein bottle.
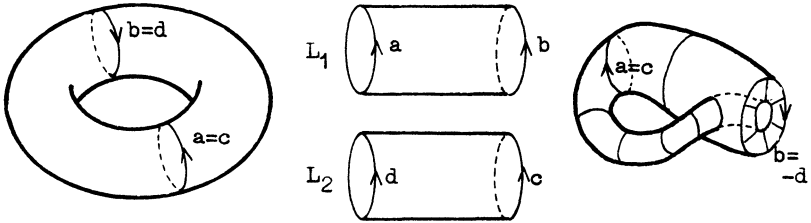
We shall assume the following: any triangulation L_1 of a cylinder has $H_1(L_1) \cong \mathbb{Z}$, with generator a cycle going round one end; $H_2(L_1) = 0$. (Compare 4.17; the result could indeed be proved by a collapsing argument.) Now let us see whether the Mayer-Vietoris sequence enables us to calculate the homology groups of torus and Klein bottle, which are already known from 6.13. Note that in either case $L_1 \cap L_2 = L$ consists of two disjoint circles.

$$H_2(L_1) \overset{}{\oplus} H_2(L_2) \overset{\psi_2}{\to} H_2(K) \overset{\Delta_2}{\to} H_1(L) \overset{\phi_1}{\to} H_1(L_1) \oplus H_1(L_2) \overset{\psi_1}{\to}$$
$$\quad 0 \qquad\qquad 0 \qquad\qquad\qquad\qquad \mathbb{Z} \oplus \mathbb{Z} \qquad \mathbb{Z} \ \oplus \ \mathbb{Z}$$

$$\overset{\psi_1}{\to} H_1(K) \overset{\Delta_1}{\to} H_0(L) \overset{\phi_0}{\to} H_0(L_1) \oplus H_0(L_2) \overset{\psi_0}{\to} H_0(K) \to 0.$$
$$\qquad\qquad\quad \mathbb{Z} \oplus \mathbb{Z} \qquad\quad \mathbb{Z} \ \oplus \ \mathbb{Z}$$

Consider ϕ_1 : $H_1(L) \to H_1(L_1) \ \oplus \ H_1(L_2)$.

The groups are $\mathbb{Z} \oplus \mathbb{Z}$ $\mathbb{Z} \ \oplus \ \mathbb{Z}$

with generators $\{a\}_L,\ \{b\}_L$ $(\{a\}_{L_1},\ 0),\ (0,\ \{c\}_{L_2})$

$$= \quad \alpha_1 \quad , \qquad \alpha_2 \quad \text{say.}$$

Now $\phi_1\{a\}_L = (\{a\}_{L_1}, -\{a\}_{L_2}) = (\{a\}_{L_1}, -\{c\}_{L_2}) = \alpha_1 - \alpha_2$

$\phi_1\{b\}_L = (\{b\}_{L_1}, -\{b\}_{L_2}) = (\{a\}_{L_1}, \pm\{c\}_{L_2}) = \alpha_1 \pm \alpha_2$

where the sign is + or - for the Klein bottle or torus respectively.

From this it follows that

$$H_2(K) \cong \text{Ker } \phi_1 \cong \begin{cases} \mathbb{Z} \text{ generator } \{a - b\}_L \text{ for} \\ \qquad\qquad\qquad\qquad\qquad \text{torus} \\ 0 \text{ for Klein bottle.} \end{cases}$$

A generator for $H_2(K)$ is any element going to $\{a - b\}_L$ under Δ_2. Such an element is the homology class of the sum of all 2-simplexes of K, coherently oriented.

A similar examination of ϕ_0 reveals that $\text{Im } \Delta_{1_2} = \text{Ker } \phi_0 \cong \mathbb{Z}$, with generator $\{v^1 - v^2\}_L$ where v^1, v^2 are vertices in the two components of L.

Now there is a short exact sequence

$$0 \to \text{Im } \psi_1 \xrightarrow{\text{inclusion}} H_1(K) \xrightarrow{\Delta_1} \text{Im } \Delta_1 \to 0 \text{ .}$$

Since $\text{Im } \Delta_1 \cong \mathbb{Z}$, this sequence splits (A.19). Also $\text{Im } \psi_1 \cong (H_1(L_1) \oplus H_1(L_2))/\text{Im } \phi_1$ by the homomorphism theorem (A.15). This latter group has presentation

$$\begin{pmatrix} \alpha_1, \ \alpha_2 \\ \alpha_1 - \alpha_2 \\ \alpha_1 \pm \alpha_2 \end{pmatrix} \quad (\text{+ for Klein bottle, - for torus})$$

and so is isomorphic to $\mathbb{Z}_2$ (generator $\alpha_1 + \text{Im } \phi_1$) or $\mathbb{Z}$ (generator $\alpha_1 + \text{Im } \phi_1$) according as the sign is + or -. A generator for $\text{Im } \psi_1$ is therefore $\psi_1(\alpha_1)$ in either case, i.e. $\{a\}_K$. An element of $H_1(K)$ going under Δ_1 to $\{v^1 - v^2\}_L$ is $\{z\}$, where z is indicated heavily on the diagrams below.
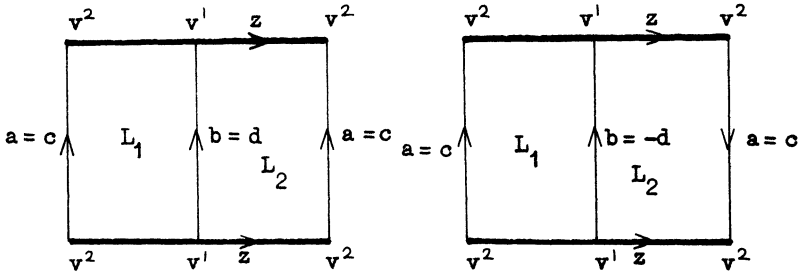
Thus $H_1(\text{torus}) \cong \mathbb{Z} \oplus \mathbb{Z}$; $H_1(\text{Klein bottle}) \cong \mathbb{Z} \oplus \mathbb{Z}_2$ with generators $\{z\}$ and $\{a\}$ in each case.

<u>7.6</u> <u>Examples</u>   (1X) In the spirit of 7.5, split a Klein bottle into two Möbius bands and calculate the homology groups of the Klein bottle from those of the Möbius bands.
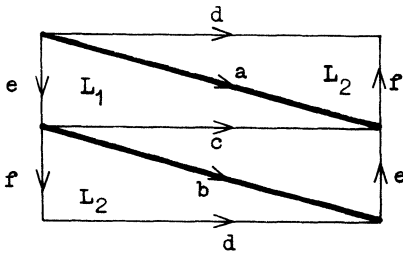
L is a single circle, drawn heavily.

Assume that $H_1(L_1) \cong \mathbb{Z}$ (generator $\{c\}$); $H_1(L_2) \cong \mathbb{Z}$ (generator $\{d\}$); $H_1(L) \cong \mathbb{Z}$ (generator $\{a + b\}$); $a + b \sim 2c$ on L_1; $a + b \sim 2d$ on L_2.

(2X) Let M_1 and M_2 be two orientable closed surfaces and let L_1, L_2 be obtained by removing a 2-simplex from, respectively, M_1 and M_2. Then $M = L_1 \cup L_2$ is a connected sum $M_1 \# M_2$ (2.9). Use the Mayer-Vietoris sequence to show that

$$H_1(M) \cong H_1(M_1) \oplus H_1(M_2).$$

(Hints. Show that $H_1(L_1) \cong H_1(M_1)$, either by quoting 6.12 or by using the Mayer-Vietoris sequence. Then use the sequence again to show $H_1(M) \cong H_1(L_1) \oplus H_1(L_2)$.)   Thus once the

first homology group of any torus is known this reproduces
6.13(2), at least up to isomorphism.

7.6 (3X) The reason why (2) works is that the boundary of
the "missing" 2-simplex is homologous to zero on L_1 and on
L_2. If we stick together a torus and a projective plane, say,
this is no longer the case.   Make the following assumption:
L_2 = triangulation of a projective plane, minus a 2-simplex,
has $H_1(L_2) \cong \mathbb{Z}$ and the boundary of the "missing" 2-simplex
represents twice a generator.   Now use the Mayer-Vietoris
sequence to calculate the first homology group of the connected
sum in question.   The above assumption follows, of course,
from the collapsing arguments of Chapter 6 - but then so does
the final result.   Removing a 2-simplex from a triangulation
of a projective plane does in fact give a triangulation of a
Möbius band, so the assumption is related to a calculation in
4.15.

It is worth noting that the Mayer-Vietoris sequence cannot
by itself be used to prove the assumption, nor even that
$H_1(L_2) \cong \mathbb{Z}$ given $H_1(P) \cong \mathbb{Z}_2$ and $H_2(P) = 0$.   For the
sequence only provides us with a short exact sequence

$$0 \to \mathbb{Z} \to H_1(L_2) \to \mathbb{Z}_2 \to 0$$

from which $H_1(L_2)$ could be isomorphic either to $\mathbb{Z}$ or to
$\mathbb{Z} + \mathbb{Z}_2$.   Compare A.20(2).

By writing a general non-orientable closed surface kP as

$$P \# \tfrac{1}{2}(k - 1)T \quad \text{k odd}$$
$$\text{or} \quad 2P \# \tfrac{1}{2}(k - 2)T \quad \text{k even}$$

(compare p. 69) the above result can be used to calculate
$H_1(kP)$, making use of the above assumption and an analogous
one for 2P = Klein bottle.   Can the Mayer-Vietoris sequence
be used to calculate $H_1(2P)$ using just the above assumption
about P ?

<u>7.6</u> (4X) Again in the spirit of 7.5, calculate the homology groups of the following objects.
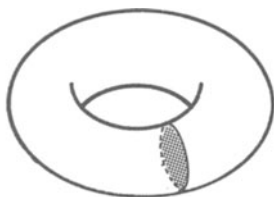
Union of a torus and a disk spanning a meridian circle. (The disk has the homology groups of a point, since (3.32) it collapses to a point.)
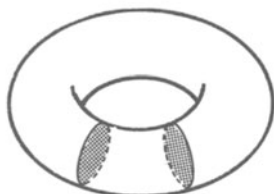
Union of a torus and two disks spanning parallel meridian circles.
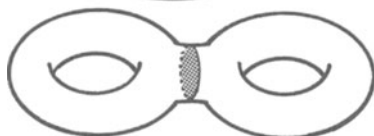
Union of a double torus and a disk spanning the "waist".

*HOMOLOGY SEQUENCE OF A TRIPLE

The theorem about to be proved generalizes 6.1, the exactness of the homology sequence of a pair. It will be used in Chapter 9, but not in an essential way, and is included here for completeness. There are several ways of proving the theorem: undoubtedly the "right" way is to use some purely algebraic machinery such as that in 4.36 et seq to deduce the result from an exact sequence of chain complexes; or an ad hoc proof analogous to that of 6.1 can be constructed. We shall adopt a third alternative, and deduce it from 6.1 itself, together with one of the more picturesque results in the theory of exact sequences, known as the Braid Theorem (see 7.9).

The notation here is chosen to be close to that in Chapter 9, where the theorem will be used. Let K and N be subcomplexes of a simplicial complex M, and suppose that $K \subset N$.
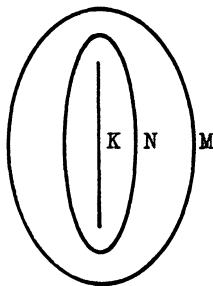
Thus $K \subset N \subset M$. There are three pairs here: (N, K), (M, K) and (M, N). The object is to fit their relative homology groups into an exact sequence. As usual, all chain groups are regarded as sub-groups of the chain groups of M.

7.7 Definitions   Define $\alpha = \alpha_p : H_p(N, K) \to H_p(M, K)$

by $\alpha\{z\}_{N,K} = \{z\}_{M,K}$ for $z \in Z_p(N, K)$.

Define

$$\beta = \beta_p : H_p(M, K) \to H_p(M, N)$$

by $\beta\{z\}_{M,K} = \{jz\}_{M,N}$ $z \in Z_p(M, K)$

where $j : C_p(M) \to C_p(M, N)$

is the usual homomorphism.

Define

$$\gamma = \gamma_p : H_p(M, N) \to H_{p-1}(N, K)$$

by $\gamma\{z\}_{M,N} = \{j'\partial z\}_{N,K}$ $z \in Z_p(M, N)$

where $j' : C_{p-1}(N) \to C_{p-1}(N, K)$

is the usual homomorphism.

The verifications that these are well-defined are left to the reader.

7.8 Theorem   If M has dimension n, then the sequence

$$0 \to H_n(N, K) \overset{\alpha}{\to} H_n(M, K) \overset{\beta}{\to} H_n(M, N) \overset{\gamma}{\to} H_{n-1}(N, K) \overset{\alpha}{\to} \cdots$$

$$\cdots \to H_1(M, N) \overset{\gamma}{\to} H_0(N, K) \overset{\alpha}{\to} H_0(M, K) \overset{\beta}{\to} H_0(M, N) \to 0$$

is exact. It is called the homology sequence of the triple

(M, N, K).

Note that if $K = \emptyset$ this reduces to the homology sequence of the pair (M, N). As mentioned above the proof will depend on 6.1 and a purely algebraic result on exact sequences, which is as follows.

Consider the diagram of abelian groups and homomorphisms below.

$$
\begin{array}{c}
\nu_{p+1} \qquad \xi_p \qquad \phi_p \qquad \gamma_p \\
F_{p+1} \quad A_p \quad B_p \quad E_p \quad F_{p-1} \\
\alpha_{p+1} \; \zeta_{p+1} \quad \lambda_p \; \theta_p \quad \eta_p \; \beta_p \quad \psi_p \; \mu_{p-1} \\
D_{p+1} \quad C_p \quad D_p \quad C_{p-1} \\
\eta_{p+1} \; \beta_{p+1} \; \psi_{p+1} \; \mu_p \; \alpha_p \; \zeta_p \; \lambda_{p-1} \; \theta_{p-1} \\
B_{p+1} \quad E_{p+1} \quad F_p \quad A_{p-1} \quad B_{p-1} \\
\phi_{p+1} \qquad \gamma_{p+1} \qquad \nu_p \qquad \xi_{p-1}
\end{array}
$$
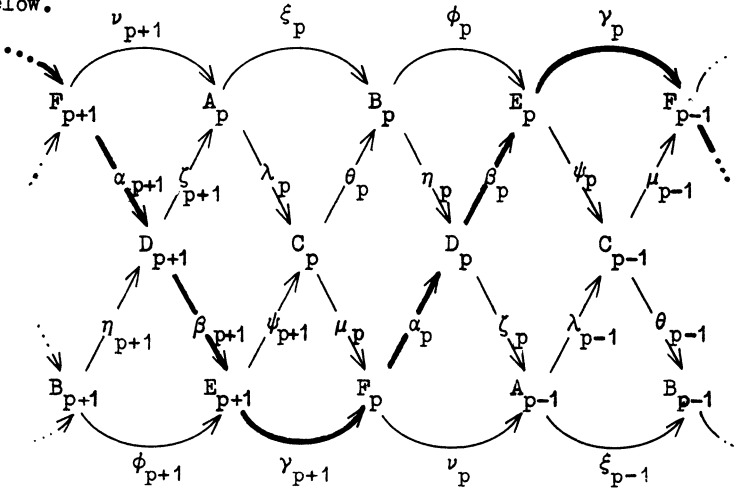
There are four "sine-wave" sequences in the diagram, in which the homomorphisms are, respectively, labelled λ, μ, ν ; ξ, η, ζ; θ, ϕ, ψ; and α, β, γ. Apart from $\alpha, \beta,$ and γ, the notation in this theorem has no connexion with previous usage - sheer pressure of notation has necessitated the re-use of several symbols. The "regions" of the diagram are three- or four-sided, and to say that the diagram "commutes" is to say that two alternative routes from one corner group of a region to another are always the same homomorphism - thus $\eta_p \circ \theta_p = \alpha_p \circ \mu_p$ for all p, etc.

<u>7.9 Braid Theorem</u>   Suppose that the above diagram commutes and that the first three sequences mentioned above are exact. Suppose also that $\beta_p \circ \alpha_p = 0$ for all p. Then the sequence of α's, β's and γ's is exact.

Proof   The proof is of a kind called __diagram-chasing,__
and we shall only give about half of it, leaving the reader to
pursue the rest. It is enough to prove exactness at F_p, D_p
and E_p.

   (1) Exactness at F_p

$$\alpha_p \circ \gamma_{p+1} = \alpha_p \circ \mu_p \circ \psi_{p+1} = \eta_p \circ \theta_p \circ \psi_{p+1} = 0$$

                              by exactness of θ, ϕ, ψ sequence

Hence Im $\gamma_{p+1} \subset$ Ker α_p.
   Conversely suppose $\alpha_p f = 0$, $f \in F_p$. Then
$0 = \zeta_p \alpha_p f = \nu_p f$; hence $f = \mu_p c$ for some $c \in C_p$.
Thus $0 = \alpha_p \mu_p c = \eta_p \theta_p c$, so $\theta_p c \in$ Ker η_p = Im ξ_p.
Hence $\theta_p c = \xi_p a$ for some $a \in A_p$. We now have
$\theta_p(\lambda_p a) = \xi_p a = \theta_p c$; hence $c - \lambda_p a \in$ Ker θ_p = Im ψ_{p+1}
This gives $c - \lambda_p a = \psi_{p+1} e$ for some $e \in E_{p+1}$, and

$$\gamma_{p+1} e = \mu_p \psi_{p+1} e = \mu_p(c - \lambda_p a) = \mu_p c = f .$$

Hence Ker $\alpha_p \subset$ Im γ_{p+1}.

   (2) Exactness at D_p.   One half of this is given; for
the other half suppose $\beta_p d = 0$ for some $d \in D_p$. Then
$0 = \psi_p \beta_p d = \lambda_{p-1} \zeta_p d$. Hence $\zeta_p d \in$ Ker λ_{p-1} = Im ν_p,
so $\zeta_p d = \nu_p f$ for some $f \in F_p$. Thus $\zeta_p \alpha_p f = \nu_p f = \zeta_p d$,
so $d - \alpha_p f \in$ Ker ζ_p = Im η_p. This gives $d - \alpha_p f = \eta_p b$
for some $b \in B_p$, and since $\phi_p b = \beta_p \eta_p b = 0$ (since
$\beta_p d = 0$), $b \in$ Ker ϕ_p = Im θ_p, and $b = \theta_p c$ for some
$c \in C_p$. Thus

$$\alpha_p(\mu_p c + f) = \eta_p \theta_p c + \alpha_p f = \eta_p b + d - \eta_p b = d.$$

Hence Ker $\beta_p \subset$ Im α_p.
The proof of exactness at E_p is left to the reader.   Q.E.D.

   To deduce that the homology sequence of a triple is exact

we make the following substitutions:

$$A_p = H_p(K), \quad B_p = H_p(M), \quad C_p = H_p(N), \quad D_p = H_p(M, K),$$

$$E_p = H_p(M, N), \quad F_p = H_p(N, K).$$

The homomorphisms being the usual ones, and α_p, β_p, γ_p being as in 7.7, the three sequences named first in the preamble to 7.9 become the homology sequences of the pairs (N, K), (M, K) and (M, N) respectively, while the fourth sequence becomes the homology sequence of the triple (M, N, K). The details of checking that the hypotheses of the Braid Theorem are satis-field are left to the reader.

7.10 <u>Example</u>   (X) In the notation of 7.8, suppose that $N \searrow K$. Using $H_p(N, K) = 0$ (see the proof of 6.6) and 7.8 deduce that $H_p(M, K) \cong H_p(M, N)$ for all p. What are the corresponding results (i) if $M \searrow N$; (ii) if $M \searrow K$?

Homology modulo 2

The homology theory developed so far applies only to oriented simplicial complexes. In this chapter we shall show how to make adjustments in the theory so that it applies to unoriented simplicial complexes. Roughly speaking if an oriented simplex σ and the same simplex with opposite orientation, $-\sigma$, are to be regarded as giving the same chain, then $-1 = +1$ and so the coefficients must be regarded as lying in the group (or field) $\mathbb{Z}_2$ instead of in the group $\mathbb{Z}$. A satisfactory theory can be developed on these lines, though it is weaker than the theory with integral coefficients: for example the closed surfaces nT and $2nP$ have the same "modulo 2" homology groups for any $n \geqslant 1$, and so cannot be distinguished by means of these groups.

The point of including this theory here is that with its aid certain results in Chapter 9 on the embedding of graphs in surfaces can be proved for both orientable and non-orientable closed surfaces.

<u>8.1</u> <u>Definitions</u>   Let K be a simplicial complex of dimension n, and let $s_p^1, \ldots, s_p^{\alpha_p}$ be the p-simplexes of K for $p = 0, \ldots, n$. A p-<u>chain</u> <u>mod 2</u> on K is a formal sum

$$\lambda_1 s_p^1 + \ldots + \lambda_{\alpha_p} s_p^{\alpha_p}$$

where the coefficients λ_i are elements of the field $\mathbb{Z}_2$ so that each coefficient is 0 or 1. The set of p-chains mod 2 is denoted by $C_p(K ; 2)$ (or $C_p(K ; \mathbb{Z}_2)$) and it is a vector space over the field $\mathbb{Z}_2$ with the definitions of <u>addition</u> and <u>scalar multiplication</u> as follows:

$$\sum \lambda_i s_p^i + \sum \lambda_i' s_p^i = \sum (\lambda_i + \lambda_i') s_p^i$$

$$\lambda \left(\sum \lambda_i s_p^i \right) = \sum \lambda \lambda_i s_p^i , \qquad \lambda \in \mathbb{Z}_2 .$$

For values of p other than 0, ..., n we define $C_p(K ; 2) = 0$.

The <u>boundary homomorphism mod 2</u>

$$\partial = \partial_p : C_p(K ; 2) \rightarrow C_{p-1}(K ; 2)$$

is defined for $0 < p \leqslant n$ by

$$\partial \left(\sum \lambda_i s_p^i \right) = \sum \lambda_i \partial s_p^i ,$$

where $\qquad \partial(v^0 \ldots v^p) = \sum_{i=0}^{p} (v^0 \ldots \hat{v}^i \ldots v^p) ,$

and is defined to be trivial otherwise. In fact ∂ is a linear map.

The <u>augmentation mod 2</u> is the linear map

$$\epsilon : C_0(K ; 2) \rightarrow \mathbb{Z}_2$$

defined by $\qquad \epsilon \left(\sum \lambda_i s_0^i \right) = \sum \lambda_i .$

Thus, in the diagram, we have (all mod 2)

$$\partial t^1 = a + e + d$$
$$\partial t^2 = b + c + e$$
$$\partial(t^1 + t^2) = a + b + c + d + 2e$$
$$= a + b + c + d \quad \text{since} \quad 2e = 0$$

$$\partial a = u + v, \text{ etc.}$$

$$\partial(a + b + c + d) = u + v + v + w + w + x + x + u$$

$$= 0.$$
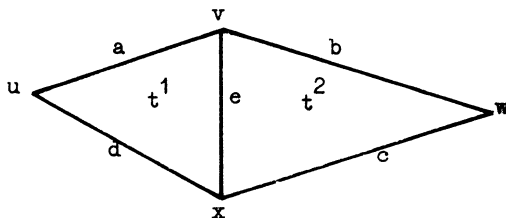
As another example, consider the diagram of a Möbius band in 4.15. Removing all arrows which indicate orientations we have

$$\partial(t^1 + t^2 + t^3 + t^4 + t^5) = e^2 + e^3 + e^4 + e^5 + e^6 \quad (\text{mod } 2);$$

that is "the mod 2 boundary of the whole band is its rim". Putting it the other way round, the rim is, mod 2, a boundary. This is false in ordinary homology: there is no 2-chain on the oriented K whose boundary is $e^2 + \ldots + e^6 \in C_1(K)$. This suggests that, for simple closed polygons, the property of enclosing a region is closely connected with being a boundary mod 2. (This will be made precise and proved in 9.19.)

An example where the Möbius band occurs as a region in something bigger is given by the Klein bottle of 2.2(4), where the shaded triangles form a Möbius band. In this case the rim of the band represents, in ordinary homology, $\pm 2\{a + b + c\}$, which is not zero (compare the last sentence of 6.14(1)). When we pass to homology mod 2 the corresponding class is zero, since the rim is, mod 2, a boundary.

8.2 <u>Remarks</u>   *(1) We can form the group of p-chains with coefficients in any abelian group A in an analogous way to that given above for $A = \mathbb{Z}_2$, and this group is denoted by $C_p(K ; A)$. The case $A = \mathbb{Z}$ is the one used hitherto. In the formula for ∂, a sign $\pm$ must be introduced, + if i is

even and - if i is odd, in order that $\partial \circ \partial$ should be
zero. If A is a field then $C_p(K ; A)$ is a vector space
over A, using the same definitions of scalar multiplication
and addition as for $A = \mathbb{Z}_2$; furthermore in that case ∂
is a linear map.

8.2 (2) Suppose that $f : A \rightarrow B$ is a map between vector
spaces over $\mathbb{Z}_2$, and that $f(a_1 + a_2) = f(a_1) + f(a_2)$ for
all a_1, a_2 in A. Then f is automatically a linear map,
i.e. f also satisfies $f(\lambda a) = \lambda f(a)$ for all $\lambda \in \mathbb{Z}_2$ and
$a \in A$. (There are only two cases, $\lambda = 0$ and $\lambda = 1$, to
check!)

Chain complexes mod 2, augmented chain complexes mod 2,
boundaries and cycles are all defined as in the case of $\mathbb{Z}$
coefficients (4.5 and 4.8). We also have the following result.

8.3 <u>Proposition</u>   For any p,

$$\partial_{p-1} \circ \partial_p \; : \; C_p(K ; 2) \; \rightarrow \; C_{p-2}(K ; 2)$$

is the trivial homomorphism. Furthermore

$$\epsilon \circ \partial_1 \; : \; C_1(K ; 2) \; \rightarrow \; \mathbb{Z}_2$$

is trivial.

Thus $\text{Im } \partial_{p+1} \subset \text{Ker } \partial_p$ for all p, and we can at any
rate form a quotient group

$$H_p(K ; 2) \; = \; \text{Ker } \partial_p / \text{Im } \partial_{p+1} \; = \; Z_p(K ; 2)/B_p(K ; 2)$$

where $Z_p(K ; 2)$ and $B_p(K ; 2)$ are regarded as just abelian
groups. $H_p(K ; 2)$ is called the pth <u>homology group of</u> K
mod 2 (or <u>with</u> $\mathbb{Z}_2$ <u>coefficients</u>). In fact $H_p(K ; 2)$ is a
vector space over $\mathbb{Z}_2$ with scalar multiplication defined by

$$\lambda\Big(z + B_p(K ; 2)\Big) \; = \; \lambda z + B_p(K ; 2)$$

where $z \in Z_p(K ; 2)$ and $\lambda \in \mathbb{Z}_2$. Thus as a vector space, and hence as an abelian group also,

$$H_p(K ; 2) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \ldots \oplus \mathbb{Z}_2 .$$

The number of summands is the dimension of the vector space $H_p(K ; 2)$, and is called the pth <u>connectivity number</u> of K. It is denoted $\hat{\beta}_p(K)$.

There is no analogue of "torsion coefficients", for every vector space is free in the sense that

$$\lambda x = 0 \quad (\lambda \in \text{field}, \ x \in \text{vector space}) \Rightarrow \lambda = 0$$
$$\text{or } x = 0 .$$

Of course considered just as an abelian group, $\mathbb{Z}_2 \oplus \ldots \oplus \mathbb{Z}_2$ is far from free, since $\lambda x = 0$ for any $x \in \mathbb{Z}_2 \oplus \ldots \oplus \mathbb{Z}_2$ when $\lambda = 2 \in \mathbb{Z}$.

Two elements z, z' in $Z_p(K ; 2)$ are called <u>homologous mod 2</u> if $z - z' \in B_p(K ; 2)$. We then write $z \sim z'$ (mod 2). The homology class mod 2 of $z \in Z_p(K ; 2)$, namely $z + B_p(K ; 2)$, is denoted $\{z\}$ or $\{z\}_K$.

We also define the <u>reduced group</u> (again a vector space)

$$\tilde{H}_0(K ; 2) = \text{Ker } \epsilon / \text{Im } \partial_1 .$$

<u>8.4</u> <u>Examples</u>   (1) Let M be a closed surface and let $c \in C_2(M ; 2)$ be the sum of all the triangles of M. Then c is a cycle, because every edge of M is on exactly two triangles of M. Notice that this does not assume M to be orientable. Using a similar argument to that in 4.18 it follows that $H_2(M ; 2) \cong \mathbb{Z}_2$, with c (strictly $\{c\}$) as a generator. We call c the <u>fundamental cycle mod 2</u> on M. Thus $H_2(M ; 2)$ does not distinguish between orientable and non-orientable closed surfaces. (But see 8.9(3).)

8.4 (2X) Suppose that K is connected. A basis for $Z_1(K \; ; \; 2)$ is obtained in the following way, exactly analogous to that for $Z_1(K)$ in 4.9(2). Let T be a maximal tree in the 1-skeleton K^1 of K, and let $e^1, \ldots, e^\mu$ be the edges of K^1 not in T. Then for each $i = 1, \ldots, \mu$, $T + e^i$ contains a unique loop and the sum of the edges in this loop is a 1-cycle z^i mod 2. We have: $z^1, \ldots, z^\mu$ is a basis for $Z_1(K \; ; \; 2)$. If K is not connected put together bases for the 1-cycles in each component of K.

In particular if K is a connected graph $(K^1 = K)$ then $H_1(K \; ; \; 2) \cong Z_2 \oplus \ldots \oplus Z_2$ (μ summands) with basis $z^1, \ldots, z^\mu$. If K has several components then $H_1(K \; ; \; 2)$ is the direct sum of the $H_1(K_i \; ; \; 2)$ where K_i runs through the components of K.

(3X) Let K be obtained from a closed surface M by removing a single 2-simplex s, and let z be the 1-cycle mod 2 on K given by the sum of the edges of s. Then $z \sim 0$ (mod 2) on K whether or not M is orientable - compare 7.6(3).

(4) $H_0(K \; ; \; 2)$ has dimension equal to the number of components of K, with basis consisting of the homology classes mod 2 of one vertex from each component. The proof given in 4.11 of the corresponding statement about $H_0(K)$ also proves the present statement, subject only to the obvious changes from Z to Z_2 in various places. Similarly, if K is non-empty,

$$H_0(K \; ; \; 2) \cong \tilde{H}_0(K \; ; \; 2) \oplus Z_2.$$

Again the proof given before (4.13) is valid, remembering that short exact sequences of vector spaces and linear maps always split (see after A.19).

(5X) Let K be the triangulation of a Möbius band given in 4.15. The calculation of $H_1(K \; ; \; 2)$ proceeds in exactly the same way as that for $H_1(K)$, except that all arrows can be

removed and minus signs changed to plus signs. We can either
regard the presentations as defining abelian groups in which,
additionally to the relations given, every element has order
two, or better, as defining vector spaces over $\mathbb{Z}_2$ as quotients
of vector spaces by subspaces. In the latter case the theory
is exactly parallel to that expounded in the Appendix and we
can manipulate the presentations until a basis for the quotient
vector space emerges. Of course these manipulations are
enormously simplified by the fact that $1 + 1 = 0$, and at the
end of the calculation there will be generators but no relations
in the presentation.

The result for the Möbius band is

$$H_1(K ; 2) \cong \mathbb{Z}_2 \text{ with basis } \{z^1\} .$$

Note that $\{e^2 + e^3 + e^4 + e^5 + e^6\} = 2\{z^1\} = 0 \pmod 2$
(see the discussion before 8.2).

The definitions of <u>relative homology groups mod</u> 2 and
<u>relative boundary homomorphisms mod</u> 2 are exactly parallel to
those in Chapter 4, integer coefficients being replaced by
coefficients in $\mathbb{Z}_2$. These relative groups are also vector
spaces over $\mathbb{Z}_2$ and the homomorphisms are linear maps (see
8.2(2)). The statement and proof of 4.24 are still valid :
given $c \in C_p(K, L ; 2)$ then $c \in Z_p(K, L ; 2)$ if and only
if $\partial c \in C_{p-1}(L ; 2)$; and $c \in B_p(K, L ; 2)$ if and only if
$\partial c' - c \in C_p(L ; 2)$ for some $c' \in C_{p+1}(K ; 2)$.

The homomorphisms i_*, j_* and ∂_* defined in 4.30, 4.32
and 4.34 carry over to the mod 2 situation and the <u>mod 2 homo-
logy sequence of a pair</u> (K, L) is exact. The statement and
proof in 6.1 apply to the present case merely by changing inte-
gers to integers mod 2 throughout. The <u>reduced</u> sequence is
also exact provided $L \neq \emptyset$.

The <u>excision theorem</u> is also true in mod 2 homology and
the proof given in 6.4 carries over. The same is true of

<u>collapsing</u> (6.6): if $K \searrow L$ then $i_* : H_p(L ; 2) \to H_p(K ; 2)$
is an isomorphism for all p.   It follows from the proof of
6.8 that the mod 2 homology groups of any cone CK are
$H_p(CK ; 2) = 0$, $p > 0$; $H_0(CK ; 2) \cong \mathbb{Z}_2$.

All the invariance properties of Chapter 5 carry over to
the mod 2 case.

The calculation of $H_1(M ; 2)$ for a closed surface M is
simpler than that of $H_1(M)$ since the analogue of 6.12 is:

Let K be obtained from M by removing a single
2-simplex.   Then $i_* : H_1(K ; 2) \to H_1(M ; 2)$ is an isomor-
phism.

The proof is provided by the proof of 6.12(1), which
carries over whether M is orientable or not.   Hence we have
the following result (compare 6.13).

<u>8.5</u>   <u>Theorem</u>   Let M be a closed surface.   Then $H_1(M ; 2)$
is as follows:

(1) $H_1(M ; 2) = 0$ when M has standard form aa^{-1}

(2) $H_1(M ; 2) \cong (\mathbb{Z}_2)^k$ when M has standard form
$a_1 a_1 \ldots a_k a_k$.   A basis is $\{a_1\}$, ..., $\{a_k\}$.

(3) $H_1(M ; 2) \cong (\mathbb{Z}_2)^{2h}$ when M has standard form
$a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_h b_h a_h^{-1} b_h^{-1}$.   A basis is $\{a_1\}$, $\{b_1\}$, ...,
$\{a_h\}$, $\{b_h\}$.   (Cycles representing this basis are obtained by
removing the arrows from those drawn in 6.14(2).)

Notice that $H_1(2nP ; 2) \cong H_1(nT ; 2)$ for any integer
$n \geqslant 1$.   Thus 2nP and nT have the same homology groups,
mod 2, for any $n \geqslant 1$ (see 8.4(1)).   Not surprisingly, the
mod 2 groups are a weaker tool than the groups with integer
coefficients.

The theorem corresponding to the Euler characteristic for-
mula of 6.16 is the following.   Let K be a simplicial complex
of dimension n.   Recall that $\hat{\beta}_p = \dim H_p(K ; 2)$ and that

α_p = number of p-simplexes of K = dim $C_p(K ; 2)$.

8.6 Theorem

$$\chi(K) \;=\; \sum_{p=0}^{n} (-1)^p \alpha_p \;=\; \sum_{p=0}^{n} (-1)^p \hat{\beta}_p \;.$$

The proof is simpler than that of 6.16 since it depends only
on the "rank plus nullity" theorem of linear algebra rather
than on the more difficult theorem A.31. (The rank plus
nullity theorem states that if f : V → W is a linear map
between vector spaces, and V is finite dimensional, then
dim(Im f) + dim(Ker f) = dim V.)

8.7 Alternating sum of dimensions theorem. Suppose that

$$0 \;\rightarrow\; A_n \;\rightarrow\; A_{n-1} \;\rightarrow\; \ldots \;\rightarrow\; A_1 \;\rightarrow\; A_0 \;\rightarrow\; 0$$

is an exact sequence of vector spaces and linear maps. Then

$$\sum_{i=0}^{n} (-1)^i \dim A_i \;=\; 0. \qquad (\text{Compare A.32(3)})$$

8.8 Examples   (1X) Let K and L be as in 4.29(3) (Möbius
band and its rim).   Then $H_1(K, L ; 2) \cong \mathbb{Z}_2$ with generator
$\{e^1\}$ and $H_2(K, L ; 2) \cong \mathbb{Z}_2$ with generator
$\{t^1 + t^2 + t^3 + t^4 + t^5\}$.   These follow either from a direct
calculation as in 4.29(3) or from the mod 2 homology sequence
of (K, L) and 8.4(5) above (the mod 2 homology groups of K).
Compare 6.3(2), which does the calculations with ordinary
homology.

   (2X) Suppose that L is a connected subcomplex of K.
Show that $j(Z_1(K ; 2)) = Z_1(K, L ; 2)$. Is this ever true
if L is not connected?   (Compare 4.29(4).)

   (3X) Repeat example 6.5 (removal of a 2-simplex from a
2-dimensional simplicial complex) with homology mod 2.

<u>8.8</u> (4X) Let S^n denote the set of proper faces of an $(n+1)$-simplex $(n \geqslant 1)$. Then $H_p(S^n ; 2) = 0$ $(p \neq 0, n)$; $H_0(S^n ; 2) \cong H_n(S^n ; 2) \cong \mathbb{Z}_2$. (Compare 6.11(4).)

The statement and proof (7.1, 7.2) of the exactness of the <u>Mayer-Vietoris sequence</u> of a triad $(K ; L_1, L_2)$ go over readily to the mod 2 case. Of course there is no longer any need for the minus sign in the definition of ϕ. All the examples of 7.4 work equally well with mod 2 homology groups (of course $\mathbb{Z}$ must be changed to $\mathbb{Z}_2$ in (4)), and the reader is invited to try them. Likewise the <u>homology sequence of a triple</u> (M, N, K) is exact in mod 2 theory. (See 7.7, 7.8)

<u>8.9</u> <u>Examples</u> (1X) Let M_1 and M_2 be closed surfaces and let K_1 and K_2 be obtained from them by removing a 2-simplex from each. Use the Mayer-Vietoris sequence (mod 2) to show that

$$i_* : H_1(K_\alpha ; 2) \rightarrow H_1(M_\alpha ; 2)$$

is an isomorphism for $\alpha = 1, 2$. Hence use the sequence (mod 2) to show that

$$H_1(M_1 \# M_2 ; 2) \cong H_1(M_1 ; 2) \oplus H_1(M_2 ; 2) .$$

Compare 7.6(2)-(3).

(2) In the calculation, given in 7.5, of the first homology groups of triangulations of torus and Klein bottle by sticking together two cylinders, the only difference between the two cases is a difference of sign, and this disappears when we pass to mod 2 homology groups. Thus, as noted in 8.5 above, these two closed surfaces have isomorphic first homology groups.

(3X) It might be supposed that there was no way of distinguishing between an orientable and a non-orientable closed surface using exclusively mod 2 homology. This is not so. Let M be a closed surface and let M' be the first barycentric

subdivision of M.   Clearly M' is also a closed surface
equivalent to M and M' is orientable if and only if M is
orientable.   Let w_p (p = 0, 1, 2) be the element of
$C_p(M' ; 2)$ given by the sum of all the p-simplexes of M'.
It is easy to verify that w_1 is a 1-cycle mod 2 - this just
uses the fact that every vertex of M' is an end-point of an
even number of edges.   In fact

$$\{w_1\} = 0 \quad <=> \quad M' \quad \text{(and hence } M) \text{ is orientable.}$$
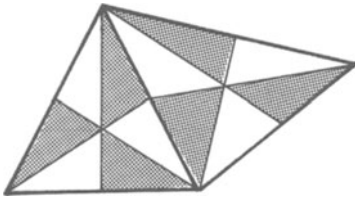
The reader may like
to find an informal
proof of this.   The
diagram is intended
to be a hint.

Notice that w_0 and w_2 are also cycles.   The elements
$\{w_p\} \in H_p(M' ; 2)$,   p = 0, 1, 2 are called the Whitney (or
Stiefel-Whitney) homology classes of M.   They are an example
of "characteristic classes" which are of great importance in
algebraic topology.   (See for example Spanier (1966), and
Milnor and Stasheff (1974).   A modern reference for the homo-
logy classes mentioned above is Halperin and Toledo (1972).)

8.9*(4)   If the ordinary homology groups $H_p(K)$ of a simpli-
cial complex K are known, then it is actually possible to
derive the groups $H_p(K ; 2)$.   The formula is

$$H_p(K ; 2) \cong \left(H_p(K) \otimes \mathbb{Z}_2 \right) \quad \oplus \quad \left(\text{Tor}(H_{p-1}(K), \mathbb{Z}_2) \right).$$

We shall not go into the details of the tensor product $\otimes$ and
the torsion product Tor here, but the answer can be worked
out for any given K using the following rules.   (See for
example Hilton and Wylie (1960), Ch. 5, especially p.189;   note
that they use * for Tor.   The formula is known as the
Universal Coefficients Theorem for homology.)

   (i)   Direct sums can be "multiplied out", e.g.

$$(A_1 \oplus A_2) \otimes (B_1 \oplus B_2)$$

$$\cong (A_1 \otimes B_1) \oplus (A_1 \otimes B_2) \oplus (A_2 \otimes B_1) \oplus (A_2 \otimes B_2) ,$$

$$\mathrm{Tor}(A_1 \oplus A_2,\ B_1 \oplus B_2)$$

$$\cong \mathrm{Tor}(A_1,\ B_1) \oplus \mathrm{Tor}(A_1,\ B_2) \oplus \mathrm{Tor}(A_2,\ B_1) \oplus \mathrm{Tor}(A_2,\ B_2).$$

(ii) $A \otimes B \cong B \otimes A$; $\mathrm{Tor}(A, B) \cong \mathrm{Tor}(B, A)$

(iii) $\mathbb{Z} \otimes \mathbb{Z}_k \cong \mathbb{Z}_k$; $\mathbb{Z} \otimes \mathbb{Z} \cong \mathbb{Z}$; $A \otimes 0 = 0$;

$\mathbb{Z}_m \otimes \mathbb{Z}_n \cong \mathbb{Z}_{(m,n)}$ where (m, n) = h.c.f.   of   m
and   n.

(iv) $\mathrm{Tor}(\mathbb{Z}, \mathbb{Z}_k) = 0$; $\mathrm{Tor}(\mathbb{Z}, \mathbb{Z}) = 0$; $\mathrm{Tor}(A, 0) = 0$;

$\mathrm{Tor}(\mathbb{Z}_m, \mathbb{Z}_n) \cong \mathbb{Z}_{(m,n)}$.

Thus for example take $K = P$, the projective plane, so that
$H_0(P) \cong \mathbb{Z}$, $H_1(P) \cong \mathbb{Z}_2$, $H_2(P) = 0$.   Then,

$$H_0(P\ ;\ 2) \cong (\mathbb{Z} \otimes \mathbb{Z}_2) \oplus \mathrm{Tor}(0,\ \mathbb{Z}_2) \cong \mathbb{Z}_2$$

$$H_1(P\ ;\ 2) \cong (\mathbb{Z}_2 \otimes \mathbb{Z}_2) \oplus \mathrm{Tor}(\mathbb{Z},\ \mathbb{Z}_2) \cong \mathbb{Z}_2$$

$$H_2(P\ ;\ 2) \cong (0 \otimes \mathbb{Z}_2) \oplus \mathrm{Tor}(\mathbb{Z}_2,\ \mathbb{Z}_2) \cong \mathbb{Z}_2 .$$

Graphs in surfaces

It has already been suggested, in the preamble to 4.19, that relative cycles have some connexion with the regions into which a graph divides a surface.   To be a little more precise, let M be a closed surface and let K be a graph which is a sub-complex of M.   If K is removed from M, what is left is a number of disjoint subsets of M which we refer to as the "regions" into which K divides M.   (Alternatively we may picture M as being cut along K: it falls into separate pieces, one for each region.)   Consider one of these regions, and take out all the 2-simplexes from it - in the diagram one
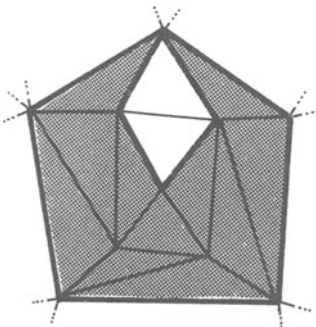
region is shaded and K drawn heavily.   Now put back the 2-simplexes one by one.   It is very plaus-ible that the sum of the 2-simplexes replaced will have mod 2 boundary lying in K, i.e. will be an element of the relative cycle group $Z_2(M, K ; 2)$, when all the 2-simplexes have been replaced and not before.   Thus there may be some connexion between the number of regions and the dimension of

229

$Z_2(M, K ; 2) = H_2(M, K ; 2)$.

In order to begin establishing such a connexion we must have some interpretation of the number of regions in terms of homology theory. It is no good saying that since the regions are the "components" of $M \setminus K$ the number of regions is the rank of "$H_0(M \setminus K)$" for $M \setminus K$ is not a simplicial complex unless $K = \emptyset$. We shall adopt a roundabout approach which replaces the regions by subcomplexes of the second barycentric subdivision of M (see p. 61).

Here is a rapid application of one of the results we are aiming for, namely that the number of regions is the dimension of $H_2(M, K ; 2)$. Let M be a triangulation of the 2-sphere so that M has standard form aa^{-1} and $H_1(M ; 2) = 0$, $H_2(M, 2) \cong \mathbb{Z}_2$ (see 8.4(1) and 8.5). Let K be a graph which is a subcomplex of M. Consider the homology sequence mod 2 of the pair (M, K):

$$H_2(K ; 2) \rightarrow H_2(M ; 2) \rightarrow H_2(M, K ; 2) \rightarrow H_1(K ; 2) \rightarrow H_1(M ; 2)$$
$$\underset{0}{\|} \qquad \underset{\mathbb{Z}_2}{\|} \qquad \qquad \qquad \qquad \underset{0}{\|}$$

Now $H_1(K ; 2)$ is isomorphic to the direct sum of the first homology groups of the components of K. If K_1 is such a component, the dimension of $H_1(K_1 ; 2)$ is $\mu(K_1) = \alpha_1(K_1) - \alpha_0(K_1) + 1$ (compare 8.4(2)). Hence the dimension of $H_1(K ; 2)$ is $\alpha_1(K) - \alpha_0(K) + k$ where k is the number of components of K. Thus by the alternating sum of dimensions theorem

$$r = \text{number of regions} = \alpha_1(K) - \alpha_0(K) + k + 1 .$$

Thus

9.1 $$\alpha_0(K) - \alpha_1(K) + r = k + 1 .$$

(Compare Euler's formula 1.27. Note that one of the ingredients of 1.27 was the Jordan Curve Theorem for polygons, so our

result includes that theorem.)

If M were some closed surface other than the 2-sphere, $H_1(M ; 2)$ would not be 0, and instead of 9.1 we should get

9.2 $\alpha_0(K) - \alpha_1(K) + r = k + 1 - d$

where d is the dimension of the image of
$i_* : H_1(K ; 2) \to H_1(M ; 2)$.   The reader should check this informally for the three graphs in the torus drawn below.
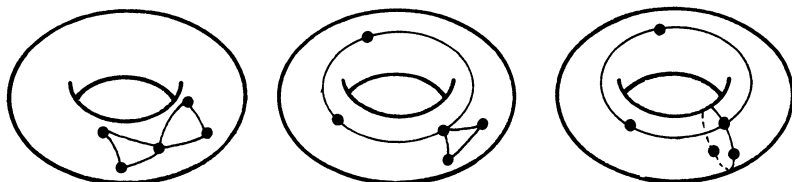
(In the right-hand diagram, for example, the two 1-cycles given by the loops on K are a basis for $H_1(M ; 2) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$, so that d = 2 and 9.2 reads 5 - 6 + 1 = 1 + 1 - 2.)

It is worth pointing out at this stage the connexion between "graphs in the sphere" and "graphs in the plane". This must be to some extent an informal discussion since it uses Fáry's theorem (1.34) which was not proved in full detail. Let M be a sphere, that is a closed surface with standard form aa^{-1}, and let G be a graph which is a subcomplex of M. Possibly some barycentric subdivisions of M are necessary in passing to the standard form; remove one triangle from the subdivided M and proceed as in the diagrams.   The result will be that some subdivision of G is realized as a curved graph in $\mathbb{R}^2$.   Now use Fáry's theorem to realize the same subdivision of G as a straight graph in $\mathbb{R}^2$.   Finally another application of the same theorem realizes G itself as a graph with straight edges in $\mathbb{R}^2$.   Thus every graph in the sphere
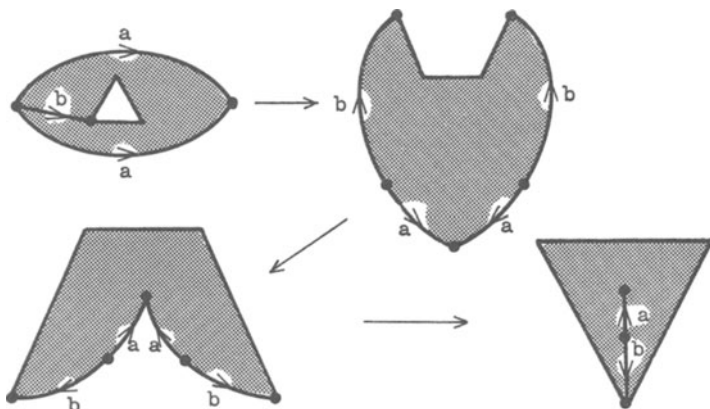
can be realized in the plane.   A converse statement is also
true - compare the diagrams in the discussion preceding 9.18.

As was pointed out above, $M \setminus K$ is not usually a sim-
plicial complex.   We get around this difficulty by "thickening"
K a very little and then taking away not just K but the in-
side of the thickening as well.

Everything that follows can be generalized to the study of
simplicial complexes K in "combinatorial manifolds" M, which
are higher dimensional analogues of surfaces, though the results
are not usually as complete in higher dimensions as they are
here.   One such generalization is given in 9.18 where we out-
line a proof that non-orientable closed surfaces cannot be real-
ized in $\mathbb{R}^3$.   Information on higher dimensions can be found
in, for example, Glaser (1970).

REGULAR   NEIGHBOURHOODS

Let M be a closed surface and K a subcomplex of M; for
the moment K need not be a graph.   Let M" and K" be the
second barycentric subdivisions of M and K respectively,
so that K" is a subcomplex of M".   Consider the following
subcomplexes of M":

9.3 $\left(\begin{array}{l} N = \text{Closure } \{s \in M'' : s \text{ has at least one vertex in } K''\} \\ V = \{s \in M'' : s \text{ has no vertex in } K''\}. \end{array}\right.$

(Thus N is the closed star of K" in M" - see 3.24.)
N is called a <u>regular neighbourhood</u> of K in M, and it
surrounds tne various components of K in such a way that the
pieces of N surrounding different components do not inter-
sect.   (See the diagram on p. 234).

     This non-intersection property would not be guaranteed by
taking only first barycentric subdivisions, as the following
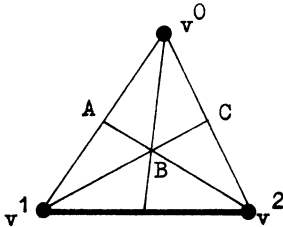
example shows.   In the
diagram, $(v^0 v^1 v^2)$ is a
triangle of M; suppose
that K consists of
(v^0), (v^1), (v^2) and
$(v^1 v^2)$.   Then replacing M"
and K" by M' and K' in
the definition of N above,
the pieces of N surrounding the two components of K inter-
sect along ABC.

     Now return to the genuine N defined by 9.3.   Since N
keeps as close as possible to K - no vertex of N is more
than "one step away" from K" along the edges of M" - the
following are plausible:

     (i)   N and K have the same number of components (this
follows from 9.5 below);

     (ii)   each region contains precisely one component of V
(this component is a slightly shrunk down version of the region).

     (Indeed taking a component V_i of V we can consider the
subset of M" defined by

$$V_i^+ = V_i \cup \{s \in M'' : s \text{ has a vertex in } K'' \text{ and a vertex in } V_i\},$$
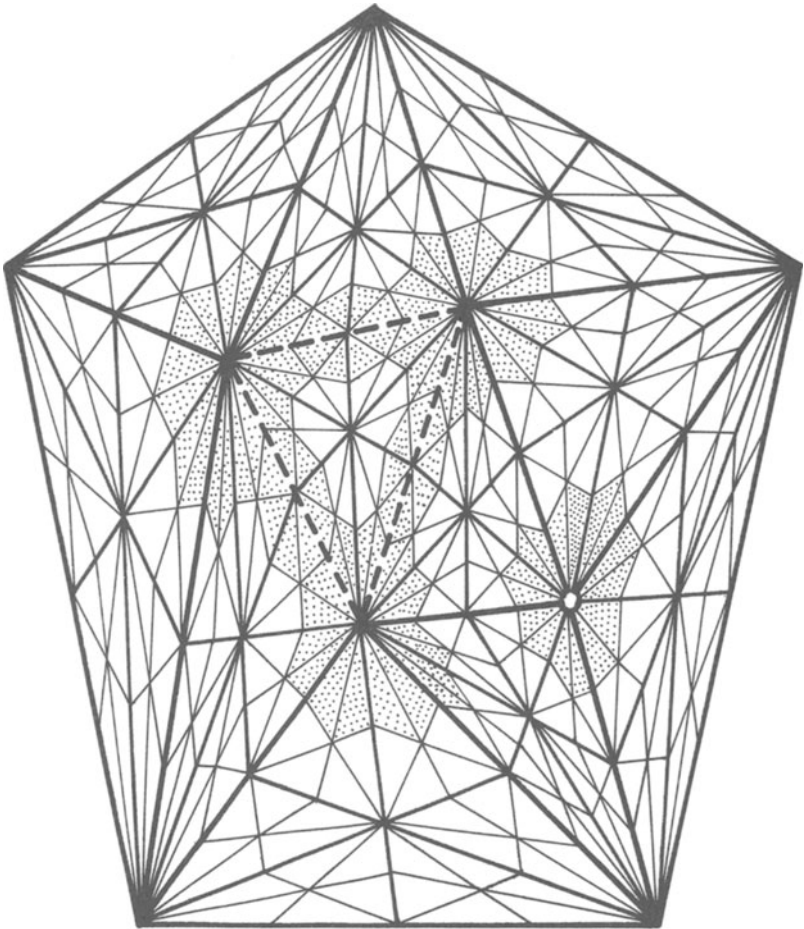
The diagram shows 11 triangles of a closed surface M, each sub-
divided into 6 triangles of M' and 36 triangles of M". K is
the subcomplex of M consisting of the 3 broken edges and their
end-points, together with the isolated white vertex of M. Thus
K is automatically subdivided into K' and K". The shaded tri-
angles and their faces make up the <u>regular neighbourhood</u> N of K
in M: each shaded triangle has at least a vertex in common with
K". Note that N has two components: each surrounds a compon-
ent of K. The white triangles and their faces make up the "com-
plementary" complex V: thus N∩V is 3 simple closed polygons.

and call this, or possibly $\left|V_i^+\right|$, the <u>region determined by</u> V_i.
It is not hard to prove that the V_i^+ are disjoint and that
their union is $M \setminus K$. But they are not subcomplexes of M''
(unless $K = \emptyset$ or $K = M$) and that is the reason for using
the V_i instead. As we shall see, the V_i have pleasant
properties - they are "surfaces" - which makes them compara-
tively easy to study.)

In view of (ii) and the fact that V is a simplicial com-
plex we now have a measure of the number of regions, and we
frame this as a definition.

<u>9.4</u> <u>Definition</u>   Let V be as in 9.3.   The <u>number of regions</u>
into which K divides M is the number of components of V.

The object, then, is to calculate $H_0(V \; ; 2)$ (or $H_0(V)$);
for good measure we shall also calculate $H_1(V \; ; 2)$, and,
when M is orientable, $H_1(V)$. This gives additional infor-
mation about the regions, as we shall see. The first step is
the following result.

<u>9.5</u> <u>Lemma</u> N collapses to K''.   Hence $i_* : H_p(K'') \to H_p(N)$
is an isomorphism for each p, and $H_p(N, K'') = 0$ for each p.
The corresponding results with mod 2 homology groups also hold.

<u>Proof</u>   Consider a 2-simplex $(v^0 v^1 v^2)$ of the first
barycentric subdivision M' of M. The number of vertices
of this 2-simplex which belong to K' is either 0, 1, 2 or 3,
and the corresponding intersections of N with $(v^0 v^1 v^2)$ are
shaded in the top line of diagrams below. (Compare the dia-
gram opposite.) The lower diagrams show, in the second and
third cases, the result of collapsing the 2-simplexes and a
1-simplex of $N \setminus K''$ which lie inside $(v^0 v^1 v^2)$. In this
way N collapses to K'' together with some 1-simplexes of N,
and a similar consideration of the 1-simplexes of M' shows
that these 1-simplexes of $N \setminus K''$ can also be collapsed away.
Hence $N \searrow K''$.
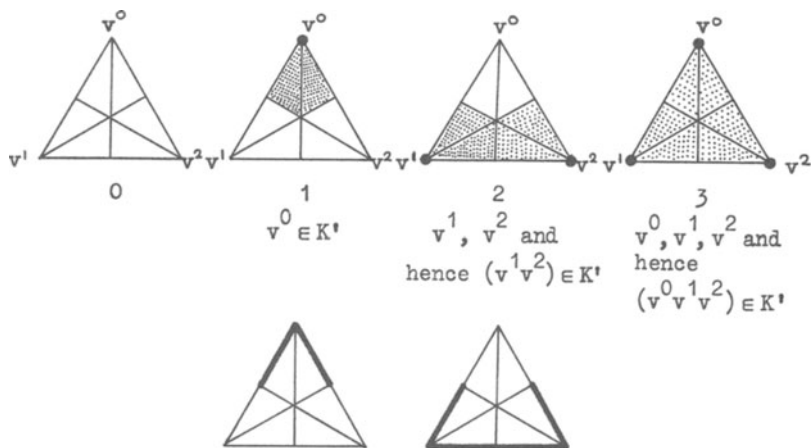
The remaining statements of the lemma follow from the invariance of homology groups under collapsing (6.6), the homology sequence of (N, K'') (compare 6.3(5)) and the corresponding results for homology groups mod 2.   Q.E.D.

9.6 Corollary $H_p(M'', K'') \cong H_p(M'', N)$ for all p.   The same result holds using homology mod 2.

Proof   This is an immediate consequence of the exactness of the homology sequence of the triple (M'', N, K'') and 9.5. (See 7.8, and replace M by M'', K by K''.)   Alternatively a direct argument can be given, and we sketch one here.   As usual, all p^{th} chain groups are assumed to be subgroups of $C_p(M'')$.

Let $\beta : H_p(M'', K'') \to H_p(M'', N)$ be defined by $\beta\{z\} = \{jz\}$, where $z \in Z_p(M'', K'')$ and $j : C_p(M'') \to C_p(M'', N)$ is the usual homomorphism which "forgets N".   It is easy to show that β is well-defined and a homomorphism;   we need to show that it is an isomorphism.   This will use

$$H_{p-1}(N, K'') = 0 = H_p(N, K'').$$
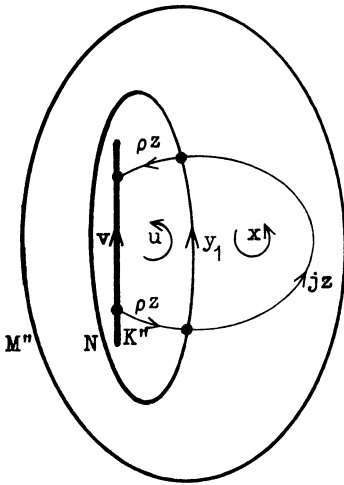
To show β is monic,
suppose $\{jz\} = 0$, i.e.
$jz = \partial x + y$ where
$x \in C_{p+1}(M'')$ and $y \in C_p(N)$.
(Compare 4.24(2).) We can
clearly assume $x \in C_{p+1}(M'', N)$
by adjusting y; let us also
write $y = y_1 + y_2$ where
$y_1 \in C_p(N, K'')$ and
$y_2 \in C_p(K'')$. In the diagram,
$y_2 = 0$. Further write
$\rho z \in C_p(N, K'')$ for the part
of z lying in N; thus
$z = jz + \rho z$.

Now $\rho z + y_1 = z - \partial x - y_2 \in C_p(N, K'')$
and $\partial(\rho z + y_1) = \partial z - 0 - \partial y_2 \in C_{p-1}(K'')$
so that $\rho z + y_1 \in Z_p(N, K'') = B_p(N, K'')$ since $H_p(N, K'') = 0$.
Thus $\rho z + y_1 = \partial u + v$ say where $u \in C_{p+1}(N, K'')$,
$v \in C_p(K'')$. As the diagram suggests, to prove $\{z\} = 0$ we
use $u + x$: certainly $u + x \in C_{p+1}(M'', K'')$, and

$$\partial(u + x) = \rho z + y_1 - v + jz - y_1 - y_2 = z - (v + y_2).$$

Since $v + y_2 \in C_p(K'')$ this shows that $z \in B_p(M'', K'')$, as
required.

The proof that β is epic is easier, and depends on taking
$z \in Z_p(M'', N)$ and writing $\partial z = x + y$ where $x \in C_{p-1}(N, K'')$
and $y \in C_{p-1}(K'')$. Then $x \in Z_{p-1}(N, K'') = B_{p-1}(N, K'')$ and
writing $x = \partial u + v$ in an analogous way to that above, it
turns out that $\beta\{z - u\} = \{z\}$. The reader should try draw-
ing a diagram of this in the case $p = 1$.   Q.E.D.

Now $M'' = N \cup V$, so that by excision (6.4)
$H_p(M'', N) \cong H_p(V, N \cap V)$ for all p. Also

$H_p(M'', K'') \cong H_p(M, K)$ by the invariance of homology groups under barycentric subdivision (5.7, 5.8), so that we can combine the foregoing results into the following proposition.

9.7 Proposition $H_p(M, K) \cong H_p(V, N \cap V)$ for all p;

$$H_p(M, K ; 2) \cong H_p(V, N \cap V ; 2)   \text{for all} p.$$

In order to identify the groups $H_p(V, N \cap V)$ we must digress a little on the subject of surfaces.

SURFACES

The class of surfaces includes that of closed surfaces (chapter 2) and also includes such objects as triangulations of a Möbius band which fail to be closed surfaces because there exist vertices (on the rim) whose link is not a simple closed polygon but a simple polygonal arc.   (See 2.2(6).)

9.8 Definition   A surface is a connected simplicial complex of dimension 2 in which the link of each vertex is either a simple closed polygon or a simple polygonal arc.

It is not difficult to see from this that every edge of a surface M is a face of either one or two triangles of M (compare 2.2(2)).   For example if an edge were a face of three triangles then the link of either end-point of the edge would contain which cannot be part of any simple closed polygon or simple polygonal arc.

9.9 Definition   The boundary ∂M of a surface M is the closure of the set of edges of M which are a face of exactly one triangle of M. (Thus ∂M consists of these edges of M and their end-points, and is a subcomplex of M.)   Note that a vertex v of M belongs to ∂M if and only if the link of v in M is a simple polygonal arc.

A surface is called orientable if its triangles can be oriented in such a way that two triangles with a common edge

are always oriented coherently.   (Compare 2.3.)   Otherwise
it is <u>non-orientable</u>.

<u>9.10</u> <u>Proposition</u>   (1) For any surface M, the boundary ∂M
consists of a collection (possibly empty) of simple closed poly-
gons.

   (2) Given any two triangles t^1, t^2 of M there is a
chain of triangles connecting them, any two consecutive members
of the chain having an edge in common.

<u>Proof</u>   (1) We need only verify that any vertex v of
∂M is an end-point of precisely two edges of ∂M.   The link
of v in M is a simple
polygonal arc, with end-
vertices u and u'
say, so that (vu) and
(vu') both belong to
∂M.   If (vw) is a
further edge of M with
end-point v then by
definition $w \in Lk(v, M)$; this implies that (vw) $\notin$ ∂M since
(vw) will be a face of two triangles of M.
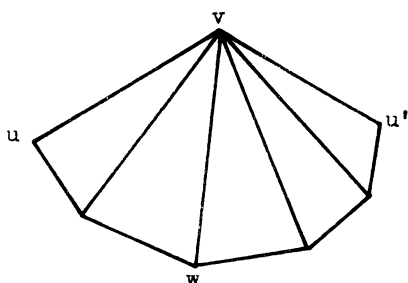
   (2) The argument suggested in 2.2(5) for the case of
closed surfaces still works, since it only uses the fact that
the link of each vertex is connected.   Q.E.D.

<u>9.11</u> <u>Examples</u>   (1) Any closed surface is a surface with
empty boundary.

   (2) The triangulations of Möbius band and cylinder in
3.22(3) are surfaces.   The boundaries consist of one and two
simple closed polygons respectively.

   (3X) Let M be obtained from a closed surface by removing
a single triangle.   Then M is a surface.   What happens if
two triangles are removed?

   (4X) Let M be a surface with boundary ∂M consisting

of n ⩾ 1 simple closed polygons. By adding a cone on each
boundary component, the cones having different vertices and
intersecting M precisely along the components of ∂M, we
obtain a closed surface M̄. The link of a vertex of ∂M is
"completed" to a simple closed polygon by the addition of two
edges and a vertex of a cone, while the link of the vertex of
one of the cones is a component of ∂M, which is a simple
closed polygon. (Compare 6.11(1), (2).) Thus a surface can
be thought of as a "closed surface with disks cut out".

    The genus of M is defined to be the genus (number of
handles or crosscaps, see 2.8) of M̄. Thus the genus of a
cylinder is 0 and the genus of a Möbius band is 1. Purists
may like to use this construction to define cylinders and Möbius
bands: a surface M is a (triangulation of a) cylinder if it
has two boundary components and M̄ has standard form aa^{-1},
etc.

    Show that $\chi(\bar{M})$ = $\chi(M)$ + n (see the second formula
(Euler characteristic of a union) of 6.17(1) and recall that a
cone collapses to a point (6.8) and therefore has Euler charac-
teristic 1). Show that M is orientable if and only if M̄
is orientable.

 9.11 (5) According to 3.35(3), a subcomplex of a closed surface
which does not contain all the triangles of the closed surface
collapses to a subcomplex of the 1-skeleton. Applying this to
the subcomplex M of the closed surface M̄ constructed in (4)
this shows that M collapses to a graph. Hence: for any
surface M with non-empty boundary, $H_2(M)$ = 0 and $H_1(M)$ is
free abelian. A graph to which M collapses is sometimes
called a spine of M.

    (6X) Let M be a surface as in (4) and let M_1 and M_2
be surfaces each isomorphic (as simplicial complexes) to M.
Form the simplicial complex DM (a "double" of M) from the
disjoint union $M_1 \cup M_2$ by identifying each oriented simplex

of ∂M_1 with the corresponding oriented simplex of ∂M_2.
Thus starting with a cylinder the picture is as follows:
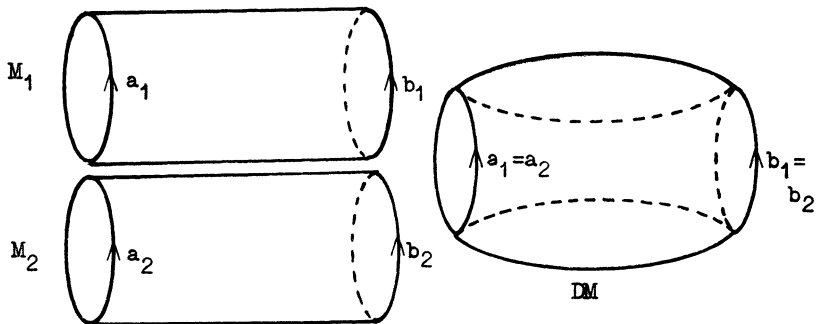
In this case DM is a torus.   Show that DM is always a
closed surface, orientable if and only if M is orientable.

Show that $\chi(DM) = 2\chi(M)$.   As in (4), we take the genus
of M to be that of $\bar{M}$; find the genus of DM in terms of
that of M.   (Consider separately the two cases M orientable
and M non-orientable.)

It is also possible to formulate a definition of DM in
which oriented boundary components are identified with opposite
orientations (in fact this is the more usual definition).   This
requires a more complicated definition, but granted that it can
be made does the change affect the genus of DM?

It was proved in (5) above that $H_2(M) = 0$ for any surface
with non-empty boundary.   It follows similarly that
$H_2(M ; 2) = 0$ in the same circumstances.   A more interesting
group to consider is $H_2(M, \partial M)$, for when M is orientable
the sum of all 2-simplexes, coherently oriented, is a nonzero
element of this group.   In fact the proof (4.18) in the case
of a closed surface can be imitated here to show $H_2(M, \partial M) \cong \mathbb{Z}$
with the element already mentioned as generator.   (A more pre-
cise definition of this generator is given by the formula in
4.18.)   In the non-orientable case $H_2(M, \partial M) = 0$ (as for
closed surfaces);   on the other hand with homology mod 2 the
sum of all triangles always gene⸱ᵤtes $H_2(M, \partial M ; 2) \cong \mathbb{Z}_2$.

These results are collected in the following proposition.

9.12 Proposition   Let M be a surface with non-empty boundary ∂M.

	$H_2(M)$	$H_2(M ; 2)$	$H_2(M, \partial M)$	$H_2(M, \partial M ; 2)$
M orientable	0	0	$\mathbb{Z}$	$\mathbb{Z}_2$
M non-orientable	0	0	0	$\mathbb{Z}_2$

(See 4.18 for the case $\partial M = \emptyset$.)

9.13 Proposition   (1) Let M be a surface.   Then

$$H_1(M ; 2) \cong H_1(M, \partial M ; 2).$$

(2) Let M be an orientable surface.   Then

$$H_1(M) \cong H_1(M, \partial M).$$

Proof   (1) Consider the homology sequence, mod 2, of (M, ∂M), and assume $\partial M \neq \emptyset$ since otherwise there is nothing to prove. (The notation " ; 2" is omitted here to save space.)

$$H_2(M) \rightarrow H_2(M, \partial M) \rightarrow H_1(\partial M) \rightarrow H_1(M) \rightarrow$$

$$\quad 0 \qquad\qquad 1 \qquad\qquad n$$

$$H_1(M, \partial M) \rightarrow H_0(\partial M) \rightarrow H_0(M) \rightarrow H_0(M, \partial M)$$

$$\qquad\qquad\qquad n \qquad\qquad 1 \qquad\qquad 0$$

The numbers are the dimensions of the vector spaces in the sequence, n being the number of components of ∂M (n $\geq$ 1). By the alternating sum of dimensions theorem (8.7), $H_1(M ; 2)$ and $H_1(M, \partial M ; 2)$ have the same dimension, and so are isomorphic.

(2) A similar argument using the homology sequence with $\mathbb{Z}$ coefficients and the alternating sum of ranks theorem (A.32(3)) shows that $H_1(M)$ and $H_1(M, \partial M)$ have equal rank. Now $H_1(M)$ is free abelian, by 9.11(5), so the result will follow once it is known that $H_1(M, \partial M)$ is free abelian. Let $\bar{M}$ be constructed as in 9.11(4) and let C be the union of the n cones on boundary components of M. By excision, $H_1(\bar{M}, C) \cong H_1(M, \partial M)$. Now from the reduced homology sequence of the pair $(\bar{M}, C)$ and the fact that $H_1(C) = 0$ it follows that $H_1(\bar{M}, C) \cong H_1(\bar{M}) \oplus \mathbb{Z}^{n-1}$. The result now follows by using 6.13, which says that $H_1(\bar{M})$ is free abelian.   Q.E.D.

9.14 Remarks   (1) The proof of 9.13 is, to say the least, unnatural. A more illuminating version of 9.13 can be given using "cohomology" groups, which we do not use in this book — though "cochains" are mentioned in the appendix to Chapter 1. The version given here will suffice for our purposes.

(2X) The homomorphism j_* between the first and second groups in each part of 9.13 is not in general an isomorphism. In fact using information already given, and the homology sequence of $(M, \partial M)$, the reader should be able to prove the following. Assume $\partial M \neq \emptyset$, then:

j_* : $H_1(M ; 2) \rightarrow H_1(M, \partial M ; 2)$ is an isomorphism if and
only if ∂M is connected
$(n = 1)$.

j_* : $H_1(M) \rightarrow H_1(M, \partial M)$ is an isomorphism if and only if
M is orientable and ∂M is
connected.

LEFSCHETZ DUALITY

Now we return to the relative homology groups of $(V, N \cap V)$
which appear in 9.7, and justify the digression on surfaces.

9.15 Proposition   In the notation of 9.3, each component of
V is a surface whose boundary is the intersection of the com-
ponent with N.   Provided $K \neq \emptyset$ and $K \neq M$ this boundary
is never empty.   The same result holds with V and N inter-
changed.

Proof   We shall prove the first two sentences and leave
the third as an (easy) exercise.   (The reader may find the
diagram on p.234 helpful in following the proof.)   For the
first sentence we need to show the following: if v is a
vertex of V then

(i) $v \notin N$ => $Lk(v, V)$ is a simple closed polygon

(ii) $v \in N$ => $Lk(v, V)$ is a simple polygonal arc.

(i) Suppose $v \notin N$.   Certainly $Lk(v, M'')$ is a simple
closed polygon, for M'' is a closed surface.   Furthermore
$Lk(v, V) \subset Lk(v, M'')$ so we just need to prove the converse
inclusion.   For this suppose that $s \in Lk(v, M'')$, so that
$sv \in M''$.   If a vertex of s failed to belong to V then it
would belong to K'' and so sv and hence v would belong to
N, contrary to assumption.   Thus s, and so sv, belong to
V, and this shows $s \in Lk(v, V)$.

(ii) Suppose $v \in V \cap N$.   Then $v \notin K''$ but there is a
1-simplex $(vw) \in M''$ with $w \in K''$.
    Suppose v is a vertex of M'.   Then w, being one step
away from v along an edge of M'', is the barycentre of a
simplex $s \in M'$ of dimension 1 or 2.   Also $s \in K'$ since
$w \in K''$ (recall w is a vertex of K'' <=> w is the bary-
centre of a simplex of K').   Finally v is a vertex of s
so $v \in K'$, a contradiction.

Hence v must be the barycentre of a 1- or 2-simplex of M' and we take these cases in turn.

<u>Case 1</u> $v = \hat{s}$ (barycentre of s) where s is a 1-simplex of M'. From the diagram, $Lk(v, M'')$ has four vertices $\hat{a}, \hat{b}, \hat{c}, \hat{d}$.
Now if $a \in K'$ then $s \in K'$ so $v \in K''$, a contradiction.
Hence $a \notin K'$, i.e. $\hat{a} \notin K''$, i.e. $\hat{a} \in V$.
Similarly $\hat{c} \in V$.
Thus the w mentioned above must be either
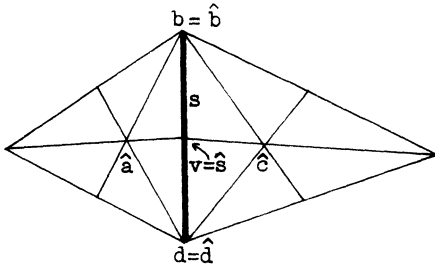
b = $\hat{b}$

s

$\hat{a}$ v=$\hat{s}$ $\hat{c}$

d=$\hat{d}$

a, b, c, d are simplexes of M'

$\hat{b}$ or $\hat{d}$, say $\hat{b}$ which then belongs to K'. If $\hat{d}$ also belonged to K' then $s = (\hat{b}\hat{d})$ would belong to K' (the simplexes of K' are precisely those of M' all of whose vertices belong to K') and so $v \in K''$, a falsehood. The conclusion is that $Lk(v, V)$ is

$\hat{a}$ •   • $\hat{c}$

• $\hat{d}$
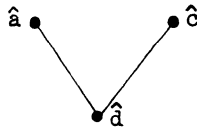
<u>Case 2</u> $v = \hat{s}$ where s is a 2-simplex of M'. Let $s = (v^1 v^2 v^3)$, as in the diagram. Now at most two of v^1, v^2, v^3 can belong to K' since if all three did then s would belong to K' (same reason as in Case 1), and so v would belong to K''. If, say, v^1 and v^2 belong to K', then just v^3, v^4, v^5 belong to V and
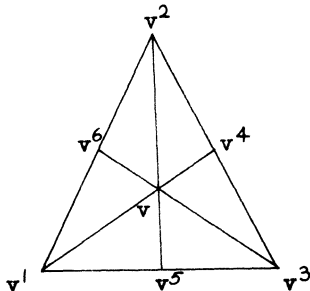
v^2

v^6 v^4

v

v^1 v^5 v^3

$Lk(v, V)$ is as in the left-hand diagram; if on the other hand
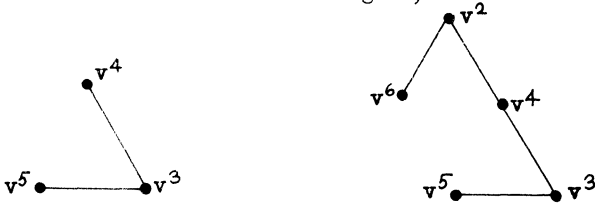
just $v^1 \in K'$ then $Lk(v, V)$ is as in the right-hand diagram.

Finally we must check that, for any component V_1 of V, $V_1 \cap N$ is non-empty, assuming $K \neq \emptyset$ and $K \neq M$. Now these two restrictions on K imply, respectively, that $N \neq \emptyset$ and $V \neq \emptyset$, so choose a vertex $v \in V_1$ and a vertex $w \in N$ and join them by a path $v \ldots w$ in M''. Let w' be the first vertex in this path which belongs to N; then all previous vertices, and hence edges, in the path belong to V, and there-fore to V_1. Furthermore $w' \in V$ since otherwise $w' \in K''$ and this implies that the vertex preceding w' in the path belongs to N. Hence $V_1 \cap N \neq \emptyset$.   Q.E.D.

9.16 <u>Remarks</u>   (1) The above analysis applies to a subcomplex K of a closed surface M. It is not difficult to generalize it to the case where M has non-empty boundary. The defini-tions of V and N remain the same, and 9.15 remains true except that the boundary of a component of V is the inter-

section of the component with $(N \cup \partial M)$. The situation is indicated in the diagram. Like-wise each component of N is a surface with boundary the intersec-tion of the component with $(V \cup \partial M)$.
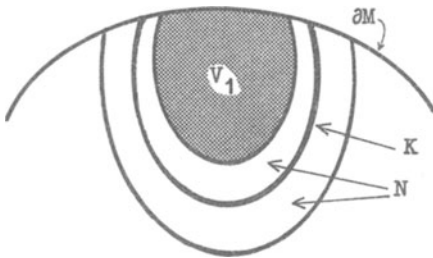
(2) In the notation introduced on p. 233 , let V_1^+ be the region determined by the component V_1 of V. Then in fact

$|V_1^+|$ is homeomorphic to $|V_1 - \partial V_1|$: the region can be
continuously shrunk down to the interior of the surface V_1.
In particular if V_1 is a triangulation of a disk (i.e. of
the plane region bounded by a simple closed polygon) then $|V_1^+|$
is homeomorphic to the open disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$.
We say then that "the region is an open disk". In fact this
happens if and only if $H_1(V_1 ; 2) = 0$ (see 9.21). Compare
the discussion of triangulating a disk in Chapter 5, p. 173.

Of course the closure $\mathrm{Cl}(V_1^+)$ is a subcomplex of M''.
Note that it is not necessarily a surface and $|\mathrm{Cl}(V_1^+)|$ is not
necessarily homeomorphic to $|V_1|$. An example appears in the
first diagram of this chapter.

<u>9.17 Theorem</u>   (A version of "Lefschetz duality".)

(1) Let M be a closed surface and K a subcomplex of
M, and let V be given by 9.3.   Then

$$H_0(V ; 2) \cong H_2(M, K ; 2) \quad \text{and} \quad H_1(V ; 2) \cong H_1(M, K ; 2) .$$

In particular the number of regions into which K divides M
is $\dim(H_2(M, K ; 2))$.

(2) Let M be an orientable closed surface and K a
subcomplex of M.   Then

$$H_0(V) \cong H_2(M, K) \quad \text{and} \quad H_1(V) \cong H_1(M, K).$$

In particular the number of regions into which K divides M
is rank $(H_2(M, K))$.

<u>Proof</u>   (1) By 9.7, $H_2(M, K ; 2) \cong H_2(V, N \cap V ; 2)$.
But V is a disjoint union of surfaces and $N \cap V$ is the
union of their boundaries, by 9.15. Thus by 9.12 (and 7.4(1)
which gives the homology of a disjoint union), $H_2(V, N \cap V ; 2)$
has dimension equal to the number of components of V, and
hence is isomorphic to $H_0(V ; 2)$. Likewise,
$H_1(M, K ; 2) \cong H_1(V, N \cap V ; 2)$ by 9.7 and this is isomor-

phic to $H_1(V ; 2)$ by 9.13(1) (and 7.4(1)) since V is a
union of surfaces as above.

(2) The proof is parallel to that of (1), using the fact
that since M is orientable so is M" and all components of
V (and N).   Q.E.D.

Using "cohomology" the Lefschetz duality theorem, like
9.13, can be given a more natural look.   In particular it is
possible to describe explicit isomorphisms between the groups
involved.   Several books give the details (which require much
more machinery than has been set up here).   See for example
Cairns (1961), Ch. 7, or Spanier (1966), Ch. 6.

*A 3-DIMENSIONAL SITUATION

Before going on to applications of the Lefschetz duality theorem
proved above here is an informal application of the corresponding
theorem in the next dimension up.   We shall not prove this
theorem - though a proof closely analogous to that for 9.17
can be constructed.

The 3-sphere is the set of points in $\mathbb{R}^4$ with coordinates
(x_1, x_2, x_3, x_4) satisfying $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$.   A
particular triangulation of the 3-sphere is the simplicial com-
plex S^3 considered in 6.11(4): the set of proper faces of a
4-simplex.   Let us assume that any triangulation M of the
3-sphere will have homology groups isomorphic to those of S^3
(this is a consequence of the invariance theorem 5.13), i.e.
$H_0(M) \cong \mathbb{Z}$, $H_1(M) = 0 = H_2(M)$, $H_3(M) \cong \mathbb{Z}$ and the same with
$\mathbb{Z}_2$ coefficients except that the $\mathbb{Z}$'s are replaced by $\mathbb{Z}_2$'s.

Now let K be a subcomplex of M.   Then it can be shown
that the number of (3-dimensional) regions into which K divides
M is the dimension of $H_3(M, K ; 2)$.   It is also equal to the
rank of $H_3(M, K)$, since M is orientable in a sense analogous
to that for surfaces.

I claim the following: if a simplicial complex K can be realized in $\mathbb{R}^3$ then it is (isomorphic to) a proper subcomplex of some triangulation M of the 3-sphere. Here is an informal proof that this is so. Suppose K is realized in $\mathbb{R}^3$; then enclose it inside a big tetrahedron and triangulate the (solid) tetrahedron with K as a subcomplex (this can always be done). Now regard the solid tetrahedron as one 3-dimensional face of a 4-simplex and triangulate the 3-skeleton S^3 of the 4-simplex so that the triangulated solid tetrahedron is a subcomplex. This makes K a subcomplex of a subdivision of S^3. The sequence of steps one dimension down is illustrated below.

| K in $\mathbb{R}^2$ | K inside triangle | triangle triangulated | Hollow tetrahedron triangulated. |
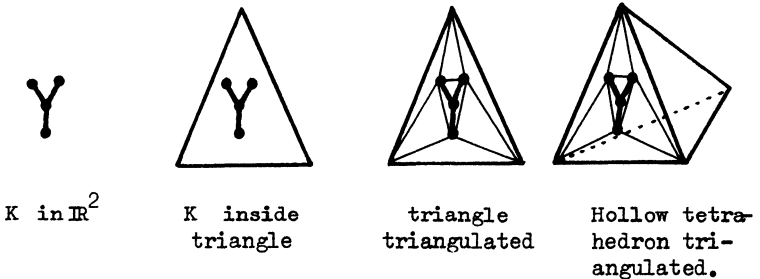
(In this example there is no need to divide up the other three faces of the hollow tetrahedron.)

Suppose now that K is a non-orientable closed surface and is a subcomplex of M. Consider the exact sequence

$$H_3(K) \rightarrow H_3(M) \rightarrow H_3(M, K) \rightarrow H_2(K) \rightarrow H_2(M) .$$

Since $H_2(K) = 0$ and $H_3(K) = 0$ this gives the rank of $H_3(M, K)$ as one: there is one region, i.e. K does not "separate" M into two or more regions. The same sequence with $\mathbb{Z}_2$ coefficients, using $H_2(K ; 2) \cong \mathbb{Z}_2$, $H_3(K ; 2) = 0$

gives the dimension of $H_3(M, K ; 2)$ as two: there are two regions. This contradiction shows that K cannot be a subcomplex of M. Using the result above, this gives:

9.18   Theorem   Let K be a non-orientable closed surface. Then K cannot be realized in $\mathbb{R}^3$.

We have seen, in 2.17(5), that for every orientable closed surface there is an equivalent closed surface realizable in $\mathbb{R}^3$. It follows from the exact sequence used above that, when an orientable closed surface is realized in $\mathbb{R}^3$, it separates $\mathbb{R}^3$ into two regions (the regions "inside" and "outside" the surface).

An orientable closed surface K in $\mathbb{R}^3$ is also said to be two-sided. To explain the meaning of this term take a point x, interior to one of the triangles of K, and construct a very short segment $x'x''$ normal (i.e. perpendicular) to the triangle and with x as its mid-point. Thus x' and x'' are "on opposite sides" of K. We call K two-sided in $\mathbb{R}^3$ if it is impossible to go from x' to x'' along a path which always keeps very close to K and never crosses K. Clearly if such a path existed then K could not separate $\mathbb{R}^3$ into two regions, so indeed an orientable closed surface in $\mathbb{R}^3$ is two-sided. (The definition can be made precise in terms of regular neighbourhoods; we really want K as a subcomplex of a triangulation of the 3-sphere, and then require that the boundary of a regular neighbourhood is not connected.)

Some people say that a Klein bottle in $\mathbb{R}^3$ is one-sided (i.e. not two-sided), but as a Klein bottle cannot be realized in $\mathbb{R}^3$ this statement must be treated with caution. Perhaps the reader can decide for himself whether he wants to call the "surface" with self-intersection drawn on p. 55 one-sided or not. One-sidedness of a surface in $\mathbb{R}^4$ is not possible, and a surface cannot separate $\mathbb{R}^4$.

It does make good sense to say that a Möbius band in $\mathbb{R}^3$

is one-sided and a cylinder in $\mathbb{R}^3$ is two-sided, but the inter-
pretation of these statements needs a little care on account of
the boundary rims.   The path from x' to x" has to keep
away from the rims.

It is possible to realize closed surfaces inside other
spaces besides euclidean spaces.   Consider first a solid ball
(a 2-sphere plus all the points inside it in $\mathbb{R}^3$) and form a

new object by identifying
together each pair b, b'
of diametrically opposite
points on the boundary
sphere.   The result is
called "3-dimensional
projective space" by
analogy with the pro-
jective plane (see
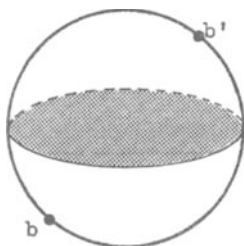
2.7(3)).   Indeed the equatorial disk (shaded in the diagram)
turns into a projective plane since opposite points on its
boundary circle are identified.   Thus the projective plane is
realized inside 3-dimensional projective space, and it is easy
to see that it is one-sided (and hence cannot separate the
space).

The 3-dimensional projective space is an example of a 3-
dimensional manifold in which the points close to any given
point p form a "neighbourhood" of p which is essentially a
solid ball - compare the corresponding idea for closed sur-
faces, where "solid ball" is replaced by "disk", on p. 49.
For points b = b' as above the neighbourhood of b = b' is
in two parts before identifications take place, but these fit
together to form a solid ball when diametrically opposite points
are all identified in pairs.

Another example of a 3-dimensional manifold is obtained by
starting with a solid torus (a torus T plus all the points
inside it in $\mathbb{R}^3$) and identifying together each pair c, c' of

points which are diametrically opposite on any of the "vertical"
circles lying on T.
Call the result M.
Thus each vertical
disk, such as the one
shaded, turns into a
projective plane P
in M.   It is not
difficult to see that
P (which is non-orientable) is two-sided in M; however P
does not separate M.   The torus T also turns into a closed
surface T' in M.   If we slice T into two equal pieces by
a horizontal plane then T' is obtained from one of these
pieces by identifying together the two end-points on each ver-
tical semicircle.   Thus T' is another torus, and it is not
hard to see that the orientable surface T' is one-sided in M
(and hence cannot separate M into two regions).
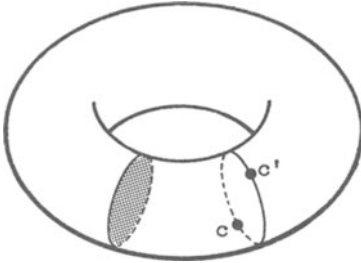
In fact, by taking suitable 3-dimensional manifolds, any
combination of orientable/non-orientable, two-sided/one-sided,
separates/does not separate the manifold, is possible, except
that, as pointed out above, a surface which separates is always
two-sided.   (That leaves six combinations; can the reader find
a non-orientable surface which does separate, and so is two-
sided?)

SEPARATING SURFACES BY GRAPHS

Let K be a graph which is a subcomplex of a closed surface M.
We can ask the question: when will K separate M, i.e. when
will it divide M into r regions where r > 1?   This can be
thought of more picturesquely as follows: Suppose we cut M
along the edges of K; when will M fall apart?   Imagining
that our scissors are rather blunt we can even picture them
removing a regular neighbourhood of K so that M drops into
a collection of surfaces $V_1, \ldots, V_r$, the components of V.

First let   K   be a simple closed polygon, so that
$H_1(K ; 2) \cong \mathbb{Z}_2$.   Consider the exact sequence

$$H_2(K; 2) \rightarrow H_2(M; 2) \rightarrow H_2(M, K; 2) \rightarrow H_1(K; 2) \overset{i_*}{\rightarrow} H_1(M; 2)$$

$$\quad 0 \qquad\qquad 1 \qquad\qquad r \qquad\qquad 1$$

The numbers are dimensions of homology groups, using 9.17.   If
$i_* = 0$ then exactness implies that $r = 2$;   if $i_* \neq 0$,   so
that the 1-cycle mod 2 on   M   given by the simple closed poly-
gon   K   is not homologous to zero, then exactness implies that
$r = 1$.   Hence:

9.19 Proposition   A simple closed polygon separates a closed
surface if and only if the 1-cycle mod 2 given by the polygon is
homologous to zero on the closed surface.   (Compare 8.4(3).)

A general graph   K   will separate   M   if and only if   K
contains a simple closed polygon of the kind described in 9.19.

Now let us ask what is the largest graph   K   which can be
embedded in a closed surface   M   without separating it.   Here
we allow   M   to be replaced by any equivalent closed surface
(one with the same standard form), so that we are asking ques-
tions like: Of all the graphs which are subcomplexes of some
triangulation of a torus and which fail to separate, which are
the "largest"?   (Compare p. 78.)   As a measure of largeness
we take $\dim H_1(K ; 2)$,   which is the sum of the cyclomatic
numbers of the components of   K.   Consider the exact sequence,
where $\mathbb{Z}_2$ coefficients are assumed:

$$H_2(K) \rightarrow H_2(M) \overset{j_*}{\rightarrow} H_2(M, K) \overset{\partial_*}{\rightarrow} H_1(K)$$

$$\overset{i_*}{\rightarrow} H_1(M) \rightarrow H_1(M, K) \rightarrow \tilde{H}_0(K) \,.$$

Now $H_2(K ; 2) = 0$, $H_2(M ; 2) \cong \mathbb{Z}_2$.

Suppose that K does not separate M. By Lefschetz duality (9.17) $H_2(M, K ; 2) \cong \mathbb{Z}_2$ and j_* , being a monomorphism, must be an isomorphism, so that $\partial_* = 0$ and i_* is monic. Hence

$$\dim H_1(K ; 2) \leqslant \dim H_1(M ; 2), \text{ with equality if and only if } i_* \text{ is an isomorphism.}$$

Conversely if i_* is an isomorphism then $\partial_* = 0$, so that j_* is an isomorphism and $H_2(M, K ; 2) \cong \mathbb{Z}_2$. Hence:

9.20 Proposition   K is a maximal non-separating graph for M (maximal in the sense of having the largest possible $\dim H_1(K ; 2)$) if and only if $i_* : H_1(K ; 2) \rightarrow H_1(M ; 2)$ is an isomorphism.

Examples of maximal non-separating graphs on torus and double-torus are indicated below. (Cycles mod 2 representing a basis for $H_1(M ; 2)$, when M is orientable, are obtained
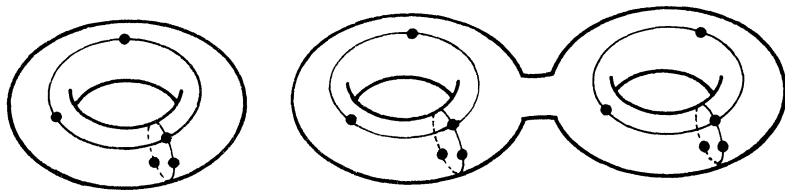
by removing the arrows from the diagram in 6.14(2).)

If we suppose that K is connected, then $\tilde{H}_0(K ; 2) = 0$ and i_* is an isomorphism implies that $H_1(M, K ; 2) = 0$. Since in this case V is connected, Lefschetz duality implies that $H_1(V ; 2) = 0$. What does this tell us about V? It says, in the first place, that ∂V is connected, for using the mod 2 homology sequence of $(V, \partial V)$ we find $H_1(\partial V ; 2) \cong \mathbb{Z}_2$. The corresponding closed surface $\bar{V}$ (9.11(4)) is therefore the

union of V with a single cone, so that $\chi(\bar{V}) = \chi(V) + 1 = 2$
since $\chi(V)$ is the alternating sum of the dimensions of the
mod 2 homology groups of V (8.6) and this is 1. Now the only
closed surfaces with $\chi = 2$ are triangulations of a 2-sphere
(standard form aa^{-1}), so $\bar{V}$ is one of these. Returning to
V amounts to removing the star of one vertex from a closed
surface with standard form aa^{-1}, and this gives a triangu-
lation of a disk (the plane region bounded by a simple closed
polygon). Any doubts that this can be done while keeping the
edges straight are dispelled by Fáry's Theorem (1.34); if the
reader does not want to accept what was not fully proved in
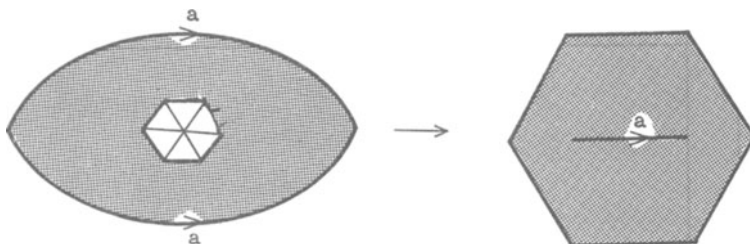
Chapter 1 then he should be content with knowing that V is
obtained from a triangulation of a 2-sphere by removing the
star of a vertex (by "punching a hole in the sphere"). Hence:

9.21 Proposition   Let U be a surface with non-empty boun-
dary satisfying $H_1(U ; 2) = 0$. Then U is isomorphic to a
triangulation of the plane region bounded by a simple closed
polygon - a disk.

Returning to the previous discussion, this shows that a
connected maximal non-separating graph K in a closed surface
M has the property that the surface V (defined by 9.3) is a
triangulation of a disk: cutting along K the closed surface
M turns into a disk. If M is represented by a polygonal
region in the plane (see the polygonal representations of
Chapter 2) such a graph K is given by the boundary polygon.

REPRESENTATION OF HOMOLOGY ELEMENTS BY

SIMPLE CLOSED POLYGONS

Given a closed surface M, which elements of $H_1(M)$ and
$H_1(M ; 2)$ are the homology classes of 1-cycles given by simple
closed polygons which are subcomplexes of M ?

   Take M to be a torus, so that $H_1(M) \cong \mathbb{Z} \oplus \mathbb{Z}$.   Con-
sider the exact sequence (K being a simple closed polygon con-
tained in M)

$$H_1(K) \overset{i_*}{\to} H_1(M) \overset{j_*}{\to} H_1(M, K) \to \tilde{H}_0(K) = 0.$$

Now $H_1(M, K) \cong H_1(V)$ by Lefschetz duality (9.17) and, V
being a disjoint union of surfaces with non-empty boundary,
$H_1(V)$ is free abelian (see 9.11(5)).   Hence by exactness,
$H_1(M)/\text{Im } i_*$ is free abelian.   Suppose that the simple closed
polygon K represents (when oriented) $p\alpha + q\beta$ where p and
q are integers and α and β are free generators for $H_1(M)$.
Thus $\text{Im } i_*$ is infinite cyclic with generator $p\alpha + q\beta$, unless
p = q = 0 in which case $\text{Im } i_* = 0$.   In any case $H_1(M)/\text{Im } i_*$
has presentation

$$\left(\begin{matrix} \alpha, \ \beta \\ p\alpha + q\beta \ = \ 0 \end{matrix} \right.$$

so that (compare A.22), $H_1(M)/\text{Im } i_* \cong \mathbb{Z} \oplus \mathbb{Z}_h$ where h is the
highest common factor of p and q, or $\mathbb{Z} \oplus \mathbb{Z}$ if p = q = 0.
Since this is free abelian, either p = q = 0 or h = 1.   Hence:

9.22 Proposition   Let M be a torus and let α, β be free
generators for $H_1(M)$.   Then an element of $H_1(M)$ represent-
able by a simple closed polygon must be either 0 or $p\alpha + q\beta$
where p and q are coprime.   (If p = 0 this says q = 0, 1
or -1.)

   Taking α and β to be the standard generators the reader
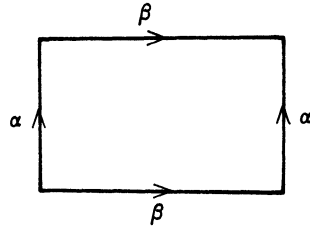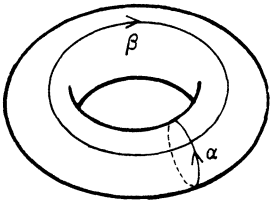
can probably convince himself informally that, say, 2α cannot
be represented by a simple closed polygon: "it is impossible
to draw a simple curve on the surface which goes round twice
one way but not at all the other way before joining up". And
given a triangulation there may be elements which are of the
form $p\alpha + q\beta$ with p and q coprime but which cannot be
represented by simple closed polygons: $n\alpha + \beta$ where n is
the number of edges in the triangulation is a pretty safe bet.
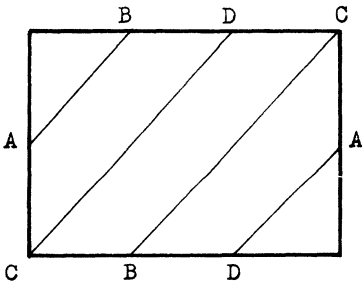
Nevertheless it is not
difficult to believe
(and it is true) that
any element of the given
form can be represented
on M or on some bary-
centric subdivision of
M. Thus to this extent
the condition of 9.22
is both necessary and
sufficient. For example $3\alpha + 2\beta$ can be indicated informally
as in the diagram. The curve ABCDA can be approximated by
a graph on a suitably fine barycentric subdivision of any given
triangulation.

The construction for general coprime p and q can be
seen informally as follows. Draw the grid of lines in the
plane given by $x = \lambda$ and $y = \mu$ for all integers λ and μ.
A torus surface is obtained from the plane by identifying
(x, y) with (x', y') whenever $x - x'$ and $y - y'$ are both
integers, for the identification means that each (x, y) in the

plane is identified with a point in the grid square with cor-
ners $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, and the edges of
this grid square are then identified in the usual way to yield
a torus.   (The identification can be pictured as rolling up
the plane to give an infinite cylinder, and then rolling at
right-angles to give a torus.)   Now draw the straight line
from $(0, 0)$ to (p, q).   It is a pleasant exercise to check
that no two points on this line, except the two end-points, are
identified in passing to the torus; hence on the surface the
line becomes a simple closed curve which clearly winds p times
round one way and q times the other.
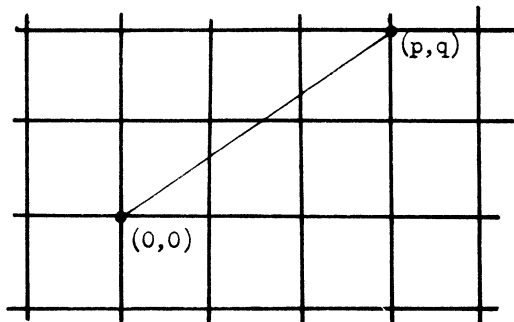
The situation on an arbitrary orientable closed surface is
similar.   The argument leading to 9.22 yields the following
result.

9.23 Proposition   Let M be an orientable closed surface.
Suppose that an element ξ of $H_1(M)$ is the homology class of
a cycle given by a simple closed polygon.   Then ξ is indivi-
sible, i.e. not a multiple of any element of $H_1(M)$ besides
$\pm \xi$, or else $\xi = 0$.

In fact the exact sequence used to prove 9.22 shows that
$H_1(M)/\langle\xi\rangle$ is free abelian, $\langle\xi\rangle$ being the subgroup of $H_1(M)$
generated by ξ, and the result follows easily from this.   It
is also true (see A.34(4)) that if ξ is indivisible or zero
then $H_1(M)/\langle\xi\rangle$ is free abelian, and indeed it can be shown
that such ξ are always representable on M or some bary-

centric subdivision of M by simple closed polygons.[†] Granted
this, it follows that every element of $H_1(M \; ; \; 2)$ is represent-
able on M or some barycentric subdivision of M by a simple
closed polygon.

Note that Lefschetz duality gives no information when M
is non-orientable. The reader may like to consider informally
such questions as: Can every element of $H_1(M \; ; \; 2)$, where M
is non-orientable, be represented on M or some barycentric
subdivision of M by a simple closed polygon? (The answer
is in fact "yes", and the reader is left with no hints beyond
the diagram in 9.25(2), where the arrows should be removed from
the dotted lines. (Next, try $a_1 + a_2$!))

ORIENTATION PRESERVING AND REVERSING LOOPS

Let K be a simple closed polygon which is a subcomplex of a
closed surface M, and consider the regular neighbourhood N
of K in M (9.3). Certainly N is connected, since $N \searrow K''$
by 9.5; hence N is a surface (9.15) with non-empty boundary
$N \cap V$. Consider the corresponding closed surface $\bar{N}$ (9.11(4))
obtained by adding a cone on each boundary component of N.
Thus if N has n boundary components, $\chi(\bar{N}) = \chi(N) + n = n$
since $\chi(N) = \chi(K'') = 0$, and since $\chi(\bar{N}) \leqslant 2$ (2.14) this
implies n = 1 or 2. Correspondingly $\chi(\bar{N}) = 1$ or 2 so
$\bar{N}$ is a projective plane or sphere. The N we started with
is therefore a Möbius band or cylinder (compare 9.11(4)).

9.24 Definition   The simple closed polygon K in the above
discussion - or the loop on M which it determines - is
called orientation preserving or reversing according as the
regular neighbourhood is a cylinder or Möbius band.

The geometrical meaning of this definition is that a small

[†] This result was pointed out to me by H. R. Morton; see
Poincaré (1904) p. 70, or Schafer (1976).

oriented circle lying in the surface, whose centre traverses
the loop once, returns with its orientation preserved or re-
versed according as N is a cylinder or Möbius band.

9.25 **Examples**   (1) If M is orientable, so are N and $\bar{\bar{N}}$;
hence every loop on an orientable surface is orientation pre-
serving.

   (2) Let M be non-orientable.   Then $H_1(M) \cong \mathbb{Z}^{k-1} \oplus \mathbb{Z}_2$
where k is the genus of M (6.13).   The diagram shows (in
the case k = 3) how to represent the unique element of order
2 in $H_1(M)$ by a simple closed polygon - it may be necessary
to subdivide M in
order to get a gen-
uine polygon which
is a subcomplex of
M.   (To see why it
represents
$\{a_1 + a_2 + a_3\}$, note
that $PQ + QR + RP$
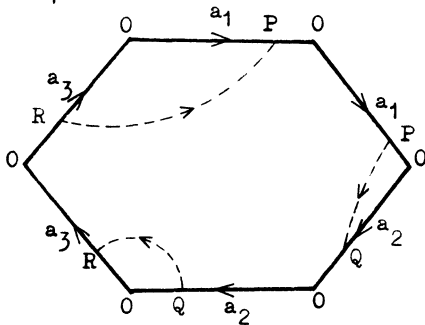
$\sim PO + OQ + QO + OR + RO + OP = OP + PO + OQ + QO + OR + RO$
$= a_1 + a_2 + a_3$.)   Suppose, then, that K is any simple
closed polygon in M which gives a 1-cycle representing the
unique element of order 2 in $H_1(M)$.   Is K orientation pre-
serving or reversing?   Now reducing mod 2 it follows that K
gives a 1-cycle mod 2 homologous (mod 2) to $a_1 + \ldots + a_k$
and therefore not representing $0 \in H_1(M ; 2)$ (see 8.5).
Hence K does not separate M (9.19): in the usual notation
V is connected.   Also from the homology sequence of (M, K)
it follows that $H_2(M, K) \cong \mathbb{Z}$ (note the integer coefficients
here), so that $\mathbb{Z} \cong H_2(M'', K'') \cong H_2(M'', N) \cong H_2(V, \partial V)$ (see
9.6, 9.7 which do not depend on M being orientable).   Since
V is a single surface this shows that it is orientable (see
9.12): to put it more picturesquely, <u>cutting</u> M <u>along</u> K <u>it</u>
<u>turns into an orientable surface.</u>

Since M" $= N \cup V$, and $N \cap V$ consists of simple
closed polygons, which have Euler characteristic 0, we have
$2 - k = X(M) = X(M") = X(N) + X(V) = X(V)$ since
$X(N) = X(K) = 0$.   Further $X(\overline{V}) = X(V) + n$ (n = number of
boundary components of N, or equally well of V) and $X(\overline{V})$
is even since $\overline{V}$ is orientable.   Hence $2 - k + n =$ even
number, and therefore n is even if and only if k is even:
K is orientation preserving if and only if k is even.

In particular for a Klein bottle k = 2, and the loop is
orientation preserving in this case.   The surface V is a
cylinder - the Klein bottle becomes a cylinder when cut along K.

9.25 (3X) In the notation of (2), show that if K gives a
cycle representing any nonzero element of $H_1(M)$ besides the
one of order 2 then V is a non-orientable surface.

(4X) Show that if K gives a cycle which is homologous
to zero (mod 2) on M, then K is always orientation preser-
ving.   (Recall that such a K separates M, so that V has
two components.)

(5X) Suppose that K_1 and K_2 are disjoint simple closed
polygons which are subcomplexes of some triangulation M of a
torus.   It is very plausible that they cannot give cycles which
freely generate $H_1(M)$: such cycles "ought to intersect".   To
prove this, note first that by (1), the regular neighbourhood
of $K = K_1 \cup K_2$ in M has 4 boundary components.   Supposing
that K_1 and K_2 do give cycles representing free generators
for $H_1(M)$ it follows from the homology sequence of (M, K)
and Lefschetz duality that, in the usual notation, V is conn-
ected.   Since $X(V) = X(M) = 0$ (compare the last paragraph
but one of (2) above), we have $X(\overline{V}) = X(V) + 4 = 4$, which
is impossible for a closed surface $\overline{V}$.

(6X) Suppose that K is a simple closed polygon (sub-
complex of a closed surface M) and that K is a boundary com-
ponent of some surface S contained in M.   Then K is

orientation preserving. (For, writing N for a regular neigh-
bourhood of K in M, ∂N will have one component in S" and
one outside S".) Thus, for example, the simple closed poly-
gons bounding the shaded regions in the Klein bottle and pro-
jective plane of 2.2(4) are orientation preserving.

Now suppose that K and L are disjoint simple closed
polygons in M. Let a and b be the corresponding 1-cycles
mod 2 and suppose that $a \sim b$ (mod 2). Then $a + b = \partial c$
for some $c \in C_2(M ; 2)$ and it is not hard to check that the
2-simplexes of M occurring in c with coefficient 1, to-
gether with their edges and vertices, form a surface with
boundary $K \cup L$. Hence, by the above result, K and L are
both orientation preserving.

One consequence of this is that if two simple closed poly-
gons are orientation reversing and represent the same element
of $H_1(M ; 2)$ then they must intersect. The reader should
examine some simple closed curves on a Klein bottle to convince
himself that this result is plausible.

9.25*(7XX) The concept of orientation preservation or reversal
depends only on the homology class mod 2 of the polygon. More
precisely, let K and L be simple closed polygons, subcom-
plexes of a closed surface M, and let the corresponding 1-
cycles mod 2 be a and b. If $a \sim b$ (mod 2) then K and
L are both orientation preserving or both orientation reversing.
Here is a very brief and informal outline of a possible argument
for this. All homology is mod 2 in this example.

Suppose that, as above, $a \sim b$. First we want to change
K and L so that they meet only in points at which they cross.
This is called moving K and L to "general position". By
subdividing M barycentrically enough times we can move K and
L to general position by arbitrarily small moves, affecting
neither the homology classes they represent nor their properties
or orientation preservation or reversal. (It would not be
difficulty to give an elementary proof of this; the proof would

however be full of rather intricate detail, and we shall assume
the result here.)   It is therefore enough to prove the above
underlined result when K and L are in general position.

   Now a ~ b so a + b = ∂c for some c ∈ C_2(M ; 2).
Consider the part of M, U say, covered by the 2-simplexes of
M with coefficient 1 in c.   At a crossing of K and L
vertically opposite segments belong to U, as in the left-hand
diagram, where U is shaded.   For each crossing we perform
the operation shown (perhaps subdividing M first) turning U
into U' and K ∪ L into B say.   Then U' is a collection
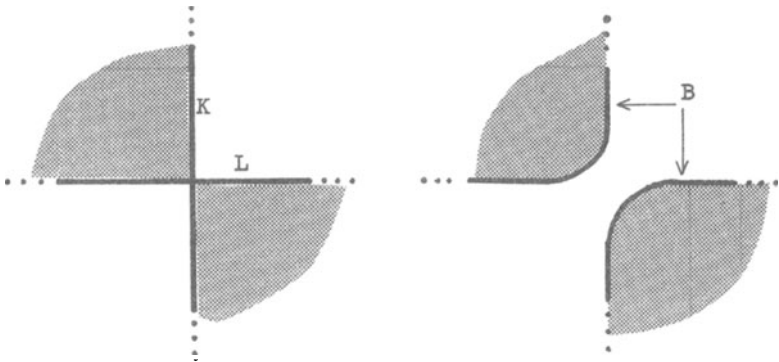
of disjoint surfaces with total boundary B.   Each component
of B will be orientation preserving, by the first part of (6)
above.   Let N be a regular neighbourhood of B; we are in-
terested in the number of boundary components of N.   By the
property of B just mentioned this number will certainly be
even.

   Now suppose that K is orientation preserving and L is
orientation reversing.   Then a regular neighbourhood of K has
boundary with two components and a regular neighbourhood of L
has connected boundary.   Of course these neighbourhoods will
intersect near the points of K ∩ L but let us imagine instead
that one passes "under" the other.   Then we have two surfaces
(the now disjoint neighbourhoods of K and L) with 1 + 2 = 3

boundary components and a sequence of operations corresponding
to that pictured above turns the surface into N.   It is not
hard to verify that each operation changes the number of boun-
dary components by an even number, possibly 0.   Hence N has
3 + even number of boundary components, which will be odd.
This contradiction proves the result.

Here is one consequence.   On a closed surface kP take
the basis $\{a_1\}$, ..., $\{a_k\}$ for $H_1(kP ; 2)$ (compare 8.5(2),
2.8).   Note that a_1, ..., a_k are all orientation reversing;
in fact we have the following.   A simple closed polygon in kP
is orientation reversing if and only if it represents
$\lambda_1\{a_1\} + ... + \lambda_k\{a_k\} \in H_1(kP ; 2)$ with $\lambda_1 + ... + \lambda_k = 1$
(mod 2).   Compare (2) above, and also p. 259.

9.25 (8XX) Show that, when K is a graph in a closed surface
M, the number of orientable components of V is the rank of
$H_2(M , K)$.   (Compare part of (2) above.)   Now let M be a
projective plane and denote by n the rank of $H_1(K)$ (or
equally well the dimension of $H_1(K ; 2)$).   Show that the num-
ber of orientable components of V is always n and that there
is never more than one non-orientable component of V.

More generally, show that if M = kP then there are never
more than k non-orientable components of V.   (Hint.   First
show that the number of orientable components is $\geqslant n + 1 - k$
and then that the total number of components is $\leqslant n + 1$.)

A GENERALIZATION OF EULER'S FORMULA

9.26 Theorem   Let K be a non-empty graph contained in a
closed surface M, and let V_1, ..., V_r be the components of
V (in the usual notation).
Then

$$\alpha_0(K) - \alpha_1(K) + r = \chi(M) + \sum_{i=1}^{r} \hat{\beta}_1(V_i)$$

where, as usual, $\hat{\beta}_1$ denotes the dimension of a first homology

group with $\mathbb{Z}_2$ coefficients.

<u>Proof</u>   Now M" = V $\cup$ N and V $\cap$ N is a union of
simple closed polygons, so $\chi(M) = \chi(M") = \chi(V) + \chi(N)$
$= \chi(V) + \chi(K)$ since N $\searrow$ K". The result now follows because
$\chi(V_i) = 1 - \hat{\beta}_1(V_i)$, V_i being a surface with non-empty boun-
dary (since K $\neq \emptyset$) for i = 1, ..., r.   Q.E.D.

<u>9.27 Corollary</u>   ("Euler's inequality")   In the notation of
9.26,

$$\alpha_0(K) - \alpha_1(K) + r \geqslant \chi(M)$$

with equality if and only if each V_i is a triangulation of a
disk.

<u>Proof</u>   This follows from 9.26 and 9.21.   Q.E.D.

        This result generalizes Euler's formula (1.27), which
refers to the case when M is a sphere, or equally well to the
case of a planar graph (see the discussion on p.231).   When
K is connected and M is a sphere it follows easily from the
homology sequence (mod 2) of (M, K) and Lefschetz duality
that $H_1(V ; 2) = 0$.

<u>9.28 Examples</u>   (1) Let K be the 1-skeleton of M.   Then
equality holds in 9.27, which is just the definition of $\chi(M)$.

        (2X) Suppose that there is some way of realizing a graph
K as a subcomplex of a closed surface so that all the compon-
ents of V are triangulations of disks (i.e. "all regions are
open disks" - 9.16(2)).   Show from 9.27 that the number of
regions is minimal, i.e. $\leqslant$ the number of regions in any other
way of realizing K as a subcomplex of M.   (Compare the dia-
grams after 9.2.)

        *(3X) Suppose that K is a connected non-empty graph con-
tained in a closed surface M and that some component U of
V (usual meaning for V) is not a triangulation of a disk.

Let ∂U be the union of $p > 0$ simple closed polygons
$S_1, \ldots, S_p$.   Remove U from M" and replace it by disjoint
cones on $S_1, \ldots, S_p$.   The resulting simplicial complex M_1
is connected (since K is connected) and is in fact a closed
surface with K" as a subcomplex.   Also $\chi(M_1) = \chi(M) + p - \chi(U)$.
Now $\chi(U) \leqslant 1$ (since $\partial U \neq \emptyset$) and indeed $\chi(U) < 1$ since
U is not a triangulation of a disk (compare 9.21).   Hence
$\chi(M_1) > \chi(M)$.   This shows that K" is realized in a closed
surface of higher Euler characteristic than M.   According to
a general theorem on graphs in surfaces this implies that K
can also be realized in a closed surface of Euler characteristic
$> \chi(M)$ (compare (4) below).   Hence we have the following
result.

> Suppose that the connected graph K can be realized
> as a subcomplex of a closed surface M but not as a sub-
> complex of any closed surface N with $\chi(N) > \chi(M)$.
> Then, however K is realized in M, all regions will be
> open disks (i.e. all components of V will be triangu-
> lated disks; compare 9.16(2).)

Since M orientable implies M_1 orientable we also have:

> Suppose that the connected graph K can be realized
> in an orientable closed surface M but not in any orien-
> table closed surface N of lower genus than M.   Then
> the same conclusion holds.   (Is this true with "non-
> orientable" replacing "orientable"?)

As an example of this result, the complete graphs G_5, G_6,
G_7 can be realized in a torus, but not in a sphere (compare
1.32).   Hence no matter how they are realized in a torus, the
regions will be open disks.

9.28 [*](4X)Here is a sketch of a proof of the following result.

> If some subdivision K_1 of a graph K is a sub-

complex of a closed surface M_1 then K is a sub-
complex of an equivalent closed surface.

This resembles a version of Fáry's theorem (1.34), but is much
easier to prove.   (Compare pp. 40 and 78.)   It is enough to
consider the case when the subdivision K_1 is obtained by
adding one vertex v on an edge (uw) of K.   We consider the
link of v in M_1, drawn in the left-hand picture.
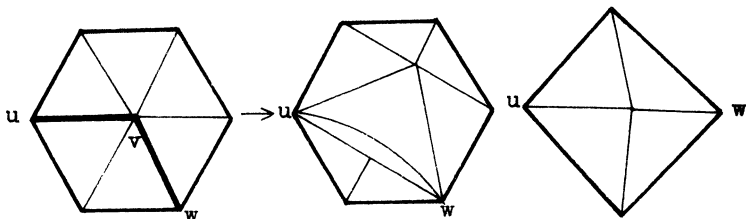
Suppose for the moment that u and w are not adjacent in the
link.   Now (uv) and (vw) belong to K_1 but the other
edges radiating from v cannot belong to K_1 since $v \notin K$.
Transform as shown in the middle picture, removing v and intro-
introducing two new vertices.   It is possible that the trans-
formed M_1 now contains two 1-simplexes (uw) - the one drawn
curved, and, possibly, one already present in M_1.   In that
case transform the one already present (which cannot belong to
K_1) as shown in the right-hand diagram.   This approach, fol-
lowed by the same transformation, also deals with the case when
u and w are adjacent on the link of v.   Finally straighten
everything by appeal to the general realization theorem of 3.19.
The result is K as a subcomplex of a closed surface which is
clearly equivalent to M_1.   (It is even more clearly p.l.
equivalent to M_1 - see 5.9.)

Why does this approach not prove the corresponding result
for planar graphs?

9.28(5X) The converse result to that in (3) - "all regions
are open disks implies that the graph cannot be realized in a

closed surface of higher Euler characteristic" - is false.
Find a counterexample.

9.28 (6X) It has already been noted that for planar graphs
the removal of an edge decreases the number of regions by one
if and only if the edge belongs to a loop (see 1.26). Show
that for a graph in an arbitrary closed surface, removing an
edge from the graph decreases the number of regions (by one)
if and only if the edge belongs to a loop which is homologous
to zero (mod 2).

*(7X) One feature of the inequality 9.27 is that both
sides are invariant under subdivision of M.   (Compare 2.16(8).)
We might say that 9.27 expresses something intrinsic about the
situation, not dependent on the particular triangulation.
An inequality of a different stamp is, in the same notation,
$3r \leqslant 2\alpha_1(K)$.   Here, the right-hand side increases with sub-
division but the left-hand side remains fixed.   Other pecu-
liarities of this inequality are its extreme weakness
$(2\alpha_1(K) - 3r$ tends to be not merely positive but very large),
which makes it difficult to prove by induction, and the anomaly
that it is false when $\alpha_1 = 0$ or 1.

It is very plausible that each region has at least three
edges of K enclosing it (at any rate when K has at least
three edges), and certainly each edge of K is adjacent to at
most two regions.   So the sum of the numbers of edges of K
adjacent to the various regions should be $\geqslant 3r$ and $\leqslant 2\alpha_1$,
whence $3r \leqslant 2\alpha_1$.   The argument below makes this more precise.

Consider a 1-simplex e of K.   Running along the two
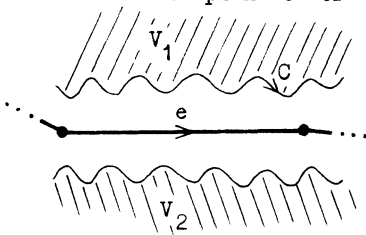
sides are boundary
components of surfaces
V_1 and V_2 which are
components of V
(possibly $V_1 = V_2$).
Call e adjacent to

these boundary components and also to V_1 and V_2.

Suppose $V_1 \neq V_2$. Traverse the boundary component C of V_1 adjacent to e in a particular direction, noting the succession of edges of K which are adjacent and giving them as they occur the direction of C. (These edges are the ones immediately, say, "on the right" of C; note that by 9.25(6) C is orientation preserving!) Since $V_1 \neq V_2$ we eventually return to e and assign it the same direction as before. Hence the edges of K adjacent to C contain a loop. (If $V_1 = V_2$ this result may or may not be true; the reader should look for examples of both events.)

Use the result just given and induction on $\alpha_1(K)$ to prove that, provided K is not a disjoint union of trees, every component of V has a loop of edges adjacent to it. In particular, provided K is not a disjoint union of trees, every component of V has at least three edges of K adjacent to it. What if K is a disjoint union of trees? (Note that in that case there is only one region.) Now use the fact that each edge of K is adjacent to at most two components of V to prove that $3r \leqslant 2\alpha_1(K)$ provided $\alpha_1(K) \geqslant 2$.

If K satisfies $3r = 2\alpha_1(K)$ then it follows that every component of V has precisely three edges of K adjacent to it, forming a loop. Hence, provided K has no isolated vertices, every region is bounded by a triangle (see p. 235). This proves something stated on p. 81: in the notation used there, every region into which G_λ divides the surface is a disk, by (3) above; hence by 9.27 we have the equality $\alpha_0(G_\lambda) - \alpha_1(G_\lambda) + r = \chi$. Using this, and $\alpha_0 = \frac{1}{2}\left(7 + \sqrt{(49 - 24\chi)}\right)$, $\alpha_1 = \frac{1}{2}\alpha_0(\alpha_0 - 1)$ it follows that $3r = 2\alpha_1$, so that G divides the surface into triangles.

The above analysis shows that a graph G in a surface S divides the surface into triangles if and only if, in the usual notation, $\alpha_0 - \alpha_1 + r = \chi$ and $3r = 2\alpha_1$. It is worth noting that these two conditions can be combined into a single

condition, namely $3\alpha_0 - \alpha_1 = 3\chi$. For we always have
$\alpha_0 - \alpha_1 + r \geqslant \chi$ and $2\alpha_1 \geqslant 3r$, so that
$3\alpha_0 - \alpha_1 \geqslant 3\alpha_0 - 3\alpha_1 + 3r \geqslant 3\chi$. Hence $3\alpha_0 - \alpha_1 = 3\chi$
if and only if both the preceding inequalities are equalities.
Compare Ringel (1974), p. 57.

9.28[*](8X) The result $3r \leqslant 2\alpha_1$ of (7), together with "Euler's
inequality", can be used to give a proof, essentially Heawood's,
of Heawood's Theorem, quoted at the end of Chapter 2: Every
graph in a closed surface, other than the sphere, of Euler
characteristic χ can be coloured with $H(\chi)$ colours, where

$$H(\chi) = [\tfrac{1}{2}(7 + \sqrt{(49 - 24\chi)})] .$$

Here are some hints. Prove that the average of the orders of
the vertices of K is $2\alpha_1(K)/\alpha_0(K)$ and that this is
$\leqslant 6 - (6\chi(M)/\alpha_0(K))$. From this it follows that there is at
least one vertex where $\leqslant 6 - (6\chi(M)/\alpha_0(K))$ edges meet.

Suppose now that $\chi(M) \leqslant 0$, and prove the following claim
by induction on $\alpha_0(K)$: If c is a positive integer satisfying
the inequality

$$6 - (6\chi(M)/c) \leqslant c - 1$$

then any graph in M can be coloured with c colours. (Note
that the conclusion is obvious if $\alpha_0(K) \leqslant c$, so start the
induction at c.) Now prove that the smallest c satisfying
the hypothesis of the claim is $H(\chi)$.

Finally consider $\chi(M) = 1$ (projective plane). When
$\chi(M) = 2$ the method only gives the uninteresting result that
every graph in the sphere can be coloured with 6 colours. The
stronger result with 5 in place of 6 is known to be true –
this is the Five Colour Theorem; see Berge (1962), p. 213 or
Wilson (1972), p. 85.

*BRUSSELS SPROUTS

As a final example of graphs in surfaces we shall consider an
amusing game due to J . H. Conway, and called (by him) <u>Brussels</u>
<u>Sprouts</u>.   (One case of this game, together with another more
serious game called Sprouts, are described in Gardner (1967).)
This is a pencil and paper, or pencil and surface, game which
involves drawing "graphs" with, probably, curved edges in a
curved surface or in some plane representation of one.   Thus
we shall not deal with explicit triangulations, bu⁺ hope instead
that the reader is prepared to believe what was claimed in
Chapter 2, namely that a "curved graph" can be drawn in a
"curved surface" if and only if the corresponding straight
graph can be realized as a subcomplex of a (triangulated) closed
surface in the appropriate equivalence class.   (The sceptical
reader can, of course, avoid this by always playing the game
on a (preferably very fine) triangulation!)

    The game starts with $n \geqslant 1$ crosses in a closed surface,
and there are two players who move alternately.   Each cross
provides 4 arms or "sprouts" and a move is to join two sprouts
with a simple arc (one that does not intersect itself) and draw
a small crossing line somewhere on the arc.   Two sample first
moves with $n = 2$ are drawn below.
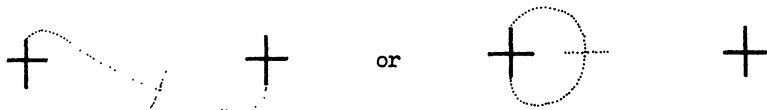
The two sprouts at the ends of the arc are not available for
future moves, but two new ones have appeared, at the ends of the
crossing line.   No arc is allowed to intersect any of the rest
of the figure except at the beginning and end of the arc, which
must be the end-points of previously unused sprouts.   Eventually,

as we shall see, no more moves are possible; the last person
to move is the winner.

At every stage of the game we have a curved graph in a
closed surface M and after p moves $(p \geqslant 0)$ the graph, K
say, has $\alpha_0(K) = 5n + 3p$
and $\alpha_1(K) = 4n + 4p$.
Furthermore since a
sprout is added on each
side of every arc drawn
during the game it is clear that at least one sprout must point
into every region into which K divides M.   Since the number
of sprouts is constant at 4n we have that r, the number of
regions, is $\leqslant 4n$.   Applying Euler's inequality, this gives
$p \leqslant 5n - \chi(M)$.   Hence:
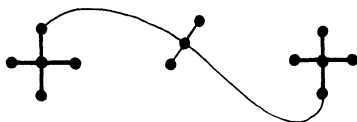
9.29 Proposition   The game of Brussels sprouts on a closed
surface of Euler characteristic χ must end after at most
$5n - \chi(M)$ moves.

Presumably the game can always be played so that it ends
after exactly $5n - \chi(M)$ moves.   Note that at the end of such
a "maximal" game all the regions into which K divides M
would be disks, by 9.27.

Now let us look more closely at the position when the game
has finished, which happens, say, after m moves.   At the end
there will be exactly 4n regions, with a sprout pointing into
each one (otherwise a further move is possible).   Taking a
regular neighbourhood N of K and complementary union of sur-
faces V as usual, it follows that the boundary of each compon-
ent of V is connected.   (For if a component V_1 of V had
disconnected boundary, the region corresponding to V_1 would
have a sprout pointing into it near each component of ∂V_1, and
a further move would be possible.)

Suppose now that M is orientable.   Then each component
of V is an orientable surface with connected boundary.   By

adding a cone on this boundary to obtain an orientable closed surface, which will have even Euler characteristic, it follows that $\chi(V_i)$ is odd for each component V_i of V; consequently $\hat{\beta}_1(V_i) = \dim H_1(V_i ; 2)$ is even. Hence from 9.26,

$$\alpha_0(K) - \alpha_1(K) + r - \chi(M) \quad \text{is even}$$

which gives, $\chi(M)$ being even,

$$n - m \quad \text{is even.}$$

This proves the following result.

9.30 Theorem   For the game of Brussels sprouts on an orientable closed surface, the first player wins (i.e. m is odd) if and only if the number of crosses to begin with is odd.

It is not difficult to prove that

$$5n - 2 \leqslant m \leqslant 5n - \chi(M)$$

whether M is orientable or not. Presumably when M is orientable m can take any value congruent to n (mod 2) between these bounds. If M is a sphere, $\chi(M) = 2$ so the actual number of moves is determined by n: this is the game played in the plane.

The case when M is non-orientable is more tricky, and only a few remarks on the subject will be made here.

9.31 Proposition   On a non-orientable surface M with standard from $a_1a_1 \ldots a_ka_k$ (sphere with k crosscaps), starting with n crosses, there is a strategy for the first player to win unless n and k are both even.

Hints for proof   Letting m be the number of moves in a game, we have, from 9.26

$$5n - m = 2 - k + \sum \hat{\beta}_1(V_i)$$

(compare the argument leading to 9.30).

<u>Case I</u>   k even, n odd.   Then the first player wins
if and only if $\sum \hat{\beta}_1(V_i)$ is even.   He can ensure this by
drawing a loop on M which represents the unique element of
order 2 in $H_1(M ; 2)$; the game is then essentially on an
orientable surface (compare 9.25(2)), and this implies that
each $\hat{\beta}_1(V_i)$ is even.

<u>Case II</u>   k odd, n odd.   The first player requires
$\sum \hat{\beta}_1(V_i)$ to be odd, and he can ensure this by enclosing all
the crosses in a disk, leaving one sprout pointing out of it:

$(n = 5)$
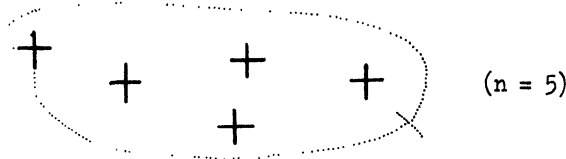
(Then all V_i but one are orientable.)

<u>Case III</u>   k odd, n even.   As in Case I.   Q.E.D.

    The remaining case, k and n both even, is more diffi-
cult and I claim that there is a strategy for the second player
to win.   Take k = 2 (Klein bottle); then in the notation of
9.2 (and letting m be the total number of moves),

$$m = 5n + d - 2.$$

Note that $0 \leqslant d \leqslant 2$ and that the first player wins (n being
even) if and only if d = 1.   Thus I claim that by playing
suitably the second player can prevent the game from ending
with d = 1.

    For general n and k (both even), the reader may like to
prove the following.   The first player's first move must be a
<u>closed</u> curve (i.e. both ends must be at sprouts on the same
cross) which <u>separates</u> the closed surface into two <u>non-orientable</u>
surfaces in <u>both</u> of which further play in possible.   (There are
four assertions here, given by the underlined words; the word

"must" means that if the first player fails to do any of these
four things then the second player can ensure a win by his
first move.   If the first player does do these and continues
to play rationally then the second player must keep on his toes
at every subsequent move in order to win.)

Appendix:
abelian groups

This appendix presents an account of the elementary theory of abelian groups which is heavily biased towards quotient groups, exact sequences and presentations, these being the main ideas from the theory which are used in the body of the book.   In fact scarcely anything is said which is not strictly relevant to the material of the book, the only exceptions being a few partially worked examples.   The reader who wants an unbiased account of abelian group theory should consult Fuchs (1970, 1973) or Kaplansky (1954).   Also full details of the classification theorem for finitely generated abelian groups, stated but not proved in A.27, appear in Hartley and Hawkes (1970). Every result from abelian group theory which is actually used in the book is proved in this appendix.

One piece of notation: here and in the text the value of a map f at an element x in its domain is denoted by fx rather than $f(x)$, unless there is a special need for bracketing as in $f(x+y)$.

BASIC DEFINITIONS

A.1   An <u>abelian group</u> is a set A, together with a binary operation $+$ defined on A and satisfying the following four axioms:

(1) For all a, b, c in A, $a + (b + c) = (a + b) + c$.

(2) There exists $0 \in A$ such that, for all $a \in A$,

$$a + 0 = 0 + a = a.$$

(3) For each $a \in A$ there exists $b \in A$ such that

$$a + b = b + a = 0.$$

(4) For all a, b in A, $a + b = b + a$.

It follows from these axioms that the 0 of (2) is unique and
that, in (3), b is uniquely determined by a. We write $-a$
for this unique b and abbreviate $c + (-a)$ to $c - a$ for any
$c \in A$.

Throughout this appendix the letters A, B, C will, unless
otherwise stated, stand for arbitrary abelian groups. As is
customary, we refer to a group using just the name of its under-
lying set.

A.2   A <u>subgroup</u> of A is a subset of A which is also an
abelian group under the same binary operation. Thus $B \subset A$ is
a subgroup of A if and only if B is nonempty and, for all
$a, b \in B$, we have $a - b \in B$.

A.3   Any group with precisely one element is denoted 0
and called a <u>trivial</u> or <u>zero group</u>.

A.4   Let $a \in A$ and $n \in \mathbb{Z}$. By na we mean 0 if
$n = 0$; $a + a + \ldots + a$ (n summands) if $n > 0$; and
$(-a) + \ldots + (-a)$ $(-n$ summands) if $n < 0$. Thus
$(-n)a = -(na)$ for any n.

A.5   Let $a \in A$. If, for all $n \in \mathbb{Z}$, $n \neq 0$, we have
$na \neq 0$ then a is said to have <u>infinite order</u>. Otherwise,
the <u>order</u> of a is the smallest integer $n > 0$ such that
$na = 0$. Thus a has order 1 if and only if $a = 0$ and,
for any a, a and $-a$ have the same order, or both have
infinite order. The <u>order</u> $|A|$ <u>of a group</u> A is simply the
number of elements in A.

A.6   A is called <u>cyclic</u> if there exists $a \in A$ suchthat any
$b \in A$ is of the form na for some $n \in \mathbb{Z}$. Such an a is
called a <u>generator</u> of A.   Note that -a is also a generator.

Examples: (1) The integers $\mathbb{Z}$ under addition is a cyclic
group with generator 1 (or, alternatively, -1), which has
infinite order.   (2) The set $\{0, 1, \ldots, k - 1\}$ $(k \geqslant 1)$
under addition modulo k is a cyclic group $\mathbb{Z}_k$ with generator
1 (or, in fact, any element coprime to k), which has order k.
Note that $\mathbb{Z}_1 = 0$.   (3) For any A and $a \in A$ the sub-
set $\{na : n \in \mathbb{Z}\}$ is a subgroup of A which is cyclic, a
being a generator.   It is called the subgroup <u>generated by</u> a
and has order equal to the order of a (infinite order if a
has infinite order).

A.7   A <u>homomorphism</u> f from A to B is a map
$f : A \to B$ such that, for all a_1, $a_2 \in A$, $f(a_1 + a_2) = fa_1 + fa_2$.
(Here the binary operations in both groups are denoted + .)
This implies $f0 = 0$ and $f(-a) = -(fa)$ for any $a \in A$.

Examples: (1) The <u>trivial</u> homomorphism $f : A \to B$ is
defined by $fa = 0$ for all $a \in A$; we write $f = 0$.   (Thus
"0" now has three meanings besides its customary one of "the
integer zero".)   Any homomorphism $0 \to A$ or $A \to 0$ is trivial.
(2) The identity map $1_A : A \to A$ is a homomorphism.   If A
is a subgroup of B then the inclusion $i : A \to B$ defined by
$ia = a$ for all $a \in A$ is a homomorphism.   (3) If
$f : A \to B$ and $g : B \to C$ are homomorphisms then so is the
composite $g \circ f : A \to C$, defined by $(g \circ f)a = g(fa)$ for
all $a \in A$.

A.8   Let $f : A \to B$ be a homomorphism.   We define

    <u>Kernel</u> of f = Ker f = $\{a \in A : fa = 0\}$;
    <u>Image</u> of f = Im f = $\{b \in B : b = fa$ for some $a \in A\}$.

These are subgroups of A and of B respectively.

We call f <u>monic</u> or a <u>monomorphism</u> if Ker f = 0.

We call f <u>epic</u>   or an <u>epimorphism</u> if Im f = B.

We call f an <u>isomorphism</u> if both Ker f = 0 and Im f = B.

An isomorphism is bijective and conversely any bijective homo-
morphism is an isomorphism. If f is an isomorphism then so
is its inverse f^{-1}.   A and B are called <u>isomorphic,</u> denoted
A $\cong$ B, if there exists an isomorphism f : A → B.   The rela-
tion $\cong$ is an equivalence relation.   Note that, in accordance
with A.3, we write A = 0   rather than A $\cong$ 0.

Example: Any cyclic group is isomorphic either to $\mathbb{Z}_k$
for some k $\geqslant$ 1 or to $\mathbb{Z}$ (see A.6).

FINITELY GENERATED (f.g.) AND FREE ABELIAN GROUPS

<u>A.9</u>   Suppose that a_1, ..., a_n are elements of A with
the property that every a ∈ A can be expressed in the form

$$a = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n$$

with λ_1, ..., λ_n integers (compare A.4).   Then A is called
<u>finitely generated (f.g.)</u> with <u>generators</u> a_1, ..., a_n.   The
expression for a is called an <u>integral linear combination</u> of
a_1, ..., a_n.

<u>A.10</u>   Suppose that A is f.g. with generators a_1, ..., a_n.
Suppose in addition that, for all integers λ_1, ..., λ_n ,

$$\lambda_1 a_1 + \dots + \lambda_n a_n = 0 \implies \lambda_1 = \dots = \lambda_n = 0 .$$

Then A is called a <u>free f.g. abelian group</u> and a_1, ..., a_n
are called <u>free generators</u> or a <u>basis</u> for A. (Since all groups
considered in the text are f.g. we usually refer there to <u>free</u>
<u>abelian groups</u>, omitting "f.g." from the description.)   By
convention 0 is free with empty basis.   It is proved in A.30
that any two bases of the same free f.g. abelian group contain

the same number of elements.

Thus A is free and f.g. with basis $a_1, \ldots, a_n$ if and
only if every $a \in A$ can be expressed uniquely as an integral
linear combination of $a_1, \ldots, a_n$. The free f.g abelian
groups are analogous to finite dimensional vector spaces, with
the field over which the vector spaces are defined being re-
placed by the group $\mathbb{Z}$ of integers under addition. Some
results about free f.g. abelian groups are in fact derived from
the corresponding results about vector spaces at the end of
this appendix.

Examples: (1) $\mathbb{Z}$ is f.g.; for generators we can take,
for example, 2, 3, or alternatively just 1, or alternatively -1.
The generator 1 is a basis; so is -1; but 2, 3 are not.
(2) $\mathbb{Z}_k$ is f.g. but not free, since for any $a \in \mathbb{Z}_k$ we have
$ka = 0$. (3) Let $\mathbb{Z}^n$ be the groups of n-tuples of integers
$(\lambda_1, \ldots, \lambda_n)$ under the operation

$$(\lambda_1, \ldots, \lambda_n) + (\mu_1, \ldots, \mu_n) = (\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n).$$

Then $\mathbb{Z}^n$ is a free f.g. abelian group with basis

$$(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1).$$

This is often called the <u>standard basis</u> for $\mathbb{Z}^n$.   *(4) The
group $\mathbb{Q}$ of rational numbers under addition is not f.g. For
let $a_1 = p_1/q_1, \ldots, a_n = p_n/q_n$ be rational numbers, with
the p_i and q_i integers. Then any integral linear combina-
tion of $a_1, \ldots, a_n$, when reduced to its lowest terms, must
have denominator a factor of $q_1 \ldots q_n$. Since there exist
rational numbers without this property, e.g. $1/2q_1 \ldots q_n$,
the given rational numbers do not generate $\mathbb{Q}$.

<u>A.11 Free abelian group on a set</u>   The most important example
for our purposes, arising in the text when defining chain groups,
is as follows. Let $S = \{x_1, \ldots, x_n\}$ where the x's are

distinct. Let FS be the set of expressions

$$\sum \lambda_i x_i \;=\; \lambda_1 x_1 + \ldots + \lambda_n x_n$$

where $\lambda_1, \ldots, \lambda_n$ are integers. Two elements of FS,
$\sum \lambda_i x_i$ and $\sum \mu_i x_i$ say, are defined to be equal if and only
if $\lambda_i = \mu_i$ for $i = 1, \ldots, n$. Addition is defined by

$$\sum \lambda_i x_i + \sum \mu_i x_i \;=\; \sum (\lambda_i + \mu_i) x_i \;.$$

Then FS is an abelian group; in fact it is free and f.g. with
basis $1x_1 + 0x_2 + \ldots + 0x_n$, $0x_1 + 1x_2 + 0x_3 + \ldots + 0x_n$,
$0x_1 + \ldots + 0x_{n-1} + 1x_n$. If we agree to omit terms with coef-
ficient zero and to write $1x_i$ as x_i then this basis becomes
just $x_1, \ldots, x_n$. We call FS the _free abelian group on_ S.
(Note that FS depends, strictly speaking, on the ordering of
S. This dependence can be eliminated by disregarding the order
in which the terms of $\sum \lambda_i x_i$ are written.) Conventionally
$F\emptyset = 0$. Notice that $FS \cong \mathbb{Z}^n$, an isomorphism being given
by $\lambda_1 x_1 + \ldots + \lambda_n x_n \rightarrow (\lambda_1, \ldots, \lambda_n)$.

A.12 Proposition   Let A be free and f.g. with basis
$a_1, \ldots, a_n$ and let B be an abelian group. Given elements
$b_1, \ldots, b_n$ in B there is a unique homomorphism $f : A \rightarrow B$
with $fa_i = b_i$ for $i = 1, \ldots, n$.

Proof   The conditions on f imply that
$f(\lambda_1 a_1 + \ldots + \lambda_n a_n) = \lambda_1 b_1 + \ldots + \lambda_n b_n$ for any integers
$\lambda_1, \ldots, \lambda_n$. It is easy to check that, A being free with
basis $a_1, \ldots, a_n$, this does define a homomorphism.   Q.E.D.

    For example, the boundary homomorphism $\partial_p : C_p(K) \rightarrow C_{p-1}(K)$
(see 1.19, 4.3) is completely specified by its values on the p-
simplexes of K. In this example both groups are free and f.g.

QUOTIENT GROUPS

The definition of homology groups (4.8) depends on the idea of
quotient group. Given a subgroup A of an abelian group B
the quotient group B/A is a device for "ignoring" or "making
zero" the elements of A. In the text A is the boundary

subgroup of a cycle group B. Both of these are free abelian,
but the following definition does not assume this.

A.13   Let A be a subgroup of B. For each $b \in B$ the
subset

$$\bar{b} = b + A = \{b + a : a \in A\}$$

is called a <u>coset of</u> A <u>in</u> B. (From Chapter 4 onwards we
use the standard, but rather unfortunate, notation {b} for
b + A. It would be too confusing to use it here.) The set
of cosets of A in B is denoted B/A (pronounced "B over
A") and it is an abelian group under the operation

$$\bar{b}_1 + \bar{b}_2 = \overline{b_1 + b_2} \; .$$

This group is called the <u>quotient group</u> of B by A.

Apart from checking the group axioms it is necessary to
check that the binary operation is well-defined, i.e. that if
$\bar{b}_1 = \bar{b}_3$ and $\bar{b}_2 = \bar{b}_4$ then $\overline{b_1 + b_3} = \overline{b_2 + b_4}$. This is
easily done using the following fact: $\bar{b} = \bar{b}'$ if and only if
$b - b' \in A$.

Examples: (1) Let $B = \mathbb{Z}$ and $A = \{nk : n \in \mathbb{Z}\}$ where
k is a fixed integer $\geqslant 1$. Thus A is the set of multiples
of k. Then $f : \mathbb{Z}_k \to B/A$ defined by $fn = \bar{n}$ for
$n = 0, 1, \ldots, k - 1$, is an isomorphism. (2) B/A = 0 if
and only if B = A. (In particular 0/0 = 0.) (3) If B is
finite then $|B/A| = |B|/|A|$ since every coset has $|A|$ ele-
ments, the cosets are disjoint and there are $|B/A|$ of them.
In particular $|A|$ divides $|B|$: this is Lagrange's theorem
in the abelian group case.

A.14   The <u>natural projection</u> $\pi : B \to B/A$ (where A is a
subgroup of B) is defined by $\pi b = \bar{b}$ for all $b \in B$.
Clearly π is an epimorphism with kernel A.

A.15   <u>The homomorphism theorem</u>   Let $f : G \to H$ be a homo-

morphism of abelian groups. Then $G/\text{Ker } f \cong \text{Im } f$, an iso-

morphism h being defined by $\bar{x} \to fx$ for all $x \in G$.

<u>Proof</u>   h is well defined since

$$\bar{x} = \bar{y} \Rightarrow x - y \in \text{Ker } f \Rightarrow f(x - y) = 0 \Rightarrow fx = fy$$

and a homomorphism since

$$h(\bar{x} + \bar{y}) = h(\overline{x + y}) = f(x + y) = fx + fy = h\bar{x} + h\bar{y}.$$

Finally h is monic since $fx = 0 \Rightarrow x \in \text{Ker } f \Rightarrow \bar{x} = 0$ and

epic since $\text{Im } f = \{fx : x \in G\}$.   Q.E.D.

Examples: (1) Take $f : \mathbb{Z} \to \mathbb{Z}_k$ to be defined by

fn = residue of n mod k.   Then f is a homomorphism with

kernel the multiples of k.   Compare example (1) in A.13.

*(2) Let G be the group of sequences $(\lambda_1, \lambda_2, \lambda_3, \ldots)$

where the λ_i are integers, under the operation

$(\lambda_1, \lambda_2, \ldots) + (\mu_1, \mu_2, \ldots) = (\lambda_1 + \mu_1, \lambda_2 + \mu_2, \ldots)$.

(Compare $\mathbb{Z}^n$ in example (3) of A.10.)   Then G is an abelian

group, not f.g. (in fact uncountable).   Define $f : G \to G$ by

$f(\lambda_1, \lambda_2, \lambda_3, \ldots) = (\lambda_2, \lambda_3, \ldots)$.   Then f is epic, with

kernel K isomorphic to $\mathbb{Z}$.   Hence by the homomorphism theorem

$G/K \cong G$: a nontrivial subgroup can be factored out without

altering G up to isomorphism.   It can be shown using A.27 that

that if A is f.g. and $A/B \cong A$ then B = 0 (see A.34(5)).

EXACT   SEQUENCES

<u>A.16</u>   The "sequence"

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

of abelian groups and homomorphisms is called <u>exact at</u> B if

$\text{Im } f = \text{Ker } g$. (Note that this implies $g \circ f = 0$.)   A

sequence is called <u>exact</u> if it is exact at each group which is

not at either end of the sequence.   Exactness at an end group

has no meaning.   A <u>short exact sequence</u> is an exact sequence
of the form

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

where, of course, the first and last homomorphisms are trivial.
(The word "exact" was first used in this sense in Eilenberg and
Steenrod (1952).   There is some connexion with the concept of
exact differential.)   Examples:

(1) $0 \to A \to 0$ is exact (i.e.   exact at A) if and only
if $A = 0$.

(2) $0 \to A \overset{f}{\to} B$ is exact if and only if f is monic.

(3) $A \overset{f}{\to} B \to 0$ is exact if and only if f is epic.

(4) $0 \to A \overset{f}{\to} B \to 0$ is exact (at A and at B) if and
only if f is an isomorphism.

(5)   For any homomorphism f : A → B,   there is a short
exact sequence

$$0 \to \operatorname{Ker} f \overset{i}{\to} A \overset{f'}{\to} \operatorname{Im} f \to 0.$$

Here i is the inclusion (ix = x for all   x ∈ Ker f), and
f'a = fa for all a ∈ A.

(6)   If A is a subgroup of B then

$$0 \to A \overset{i}{\to} B \overset{\pi}{\to} B/A \to 0$$

is exact, where i = inclusion and π = natural projection
(A.14).   Conversely if

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

is exact, then $C \cong B/\operatorname{Im} f$ by the homomorphism theorem A.15.

(7)   If $A \subset B$ and B is a subgroup of C, then A is
a subgroup of C and B/A, C/A, C/B are defined.   There is
a short exact sequence

$$0 \to B/A \overset{i}{\to} C/A \overset{p}{\to} C/B \to 0$$

where $i(b + A) = b + A$ and $p(c + A) = c + B$. By the homomorphism theorem A.15, $C/B \cong (C/A)/(B/A)$: the cancellation rule for subgroups.

A.16 (8) If

$$0 \to A \to B \to C \to 0$$

is exact and B is finite, then A and C are also finite, and $|B| = |A| \cdot |C|$.   (Compare A.13 Ex (3) and use $C \cong B/A$.)

DIRECT SUMS AND SPLITTING

A.17   Let $A_1, \ldots, A_n$ be abelian groups. The direct sum $A_1 \oplus \ldots \oplus A_n$ of the groups is the set of n-tuples $(a_1, \ldots, a_n)$ where $a_1 \in A_1, \ldots, a_n \in A_n$, under the binary operation

$$(a_1, \ldots, a_n) + (a_1', \ldots, a_n') = (a_1 + a_1', \ldots, a_n + a_n').$$

Then $A_1 \oplus \ldots \oplus A_n$ is an abelian group.

The most important case is $n = 2$, and there are "associative" properties such as

$$A \oplus B \oplus C \cong (A \oplus B) \oplus C \cong A \oplus (B \oplus C)$$

which are very easy to check.

Examples: (1) The $\mathbb{Z}^n$ of A.10, Ex (3) is equal to $\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}$ (n summands). (2) If A and B are f.g. with generators $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$ respectively, then $A \oplus B$ is f.g. with generators $(a_1, 0), \ldots, (a_n, 0)$, $(0, b_1), \ldots, (0, b_m)$. If A and B are free and f.g. with the given generators as bases then $A \oplus B$ is free and f.g. with the given generators as basis. Similarly with more than two summands. *(3) Suppose that p and q are coprime integers $\geqslant 1$; consider $(1, 1) \in \mathbb{Z}_p \oplus \mathbb{Z}_q$. Now $n(1, 1) = (n, n)$ ($n \in \mathbb{Z}$) and $(n, n) = (0, 0) \Longleftrightarrow p|n$ and $q|n \Longleftrightarrow pq|n$ since p and q are coprime. Therefore $(1, 1)$ has order pq and $\mathbb{Z}_p \oplus \mathbb{Z}_q$, having order pq, is

cyclic with $(1, 1)$ as generator. (Conversely if p and q have a common factor > 1, then $\mathbb{Z}_p \oplus \mathbb{Z}_q$ is not cyclic.)

<u>A.18</u>   Let A and B be subgroups of C such that $A \cap B = \{0\}$. The <u>internal direct sum</u> $A \oplus B$ of A and B is the subgroup of C given by

$$A \oplus B = \{a + b : a \in A \text{ and } b \in B\} .$$

The clash of notation with A.17 is not serious since $(a, b) \to a + b$ defines an isomorphism between the direct sum and the internal direct sum of A and B, assuming of course $A \cap B = \{0\}$. Note that given c in the internal direct sum there exist unique $a \in A$ and $b \in B$ such that $c = a + b$.

Several theorems in the text are expressed by the exactness of certain sequences of homology groups. When applying these theorems, especially in Chapters 6, 7 and 9, the following result is often used.

<u>A.19</u>   <u>Splitting Lemma</u>   Suppose that

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

is exact.

(1) If there exists a homomorphism $h : C \to B$ such that $g \circ h$ is the identity on C, then B is the internal direct sum (A.18) of $\text{Im } f$ and $\text{Im } h$, and furthermore $B \cong A \oplus C$. We call h a <u>splitting homomorphism</u> and say that the short exact sequence is <u>split</u>.

(2) If C is free and f.g., a splitting homomorphism always exists: the sequence is always split.

<u>Proof</u>   (1) Suppose that such an h exists. Then, for any $b \in B$, $(g \circ h \circ g)b = gb$, so that $b - (h \circ g)b \in \text{Ker } g = \text{Im } f$. Hence b can be written as the sum of elements from $\text{Im } f$ and $\text{Im } h$. To see that

$\text{Im } f \cap \text{Im } h = \{0\}$ suppose that $fa = b = hc$ ($a \in A$, $b \in B$, $c \in C$).   Then

$$0 = (g \circ f)a = gb = (g \circ h)c = c$$

by exactness and the given property of h.   This gives the result, $b = 0$.   Since f is monic, $\text{Im } f \cong A$ (for example using the homomorphism theorem A.15); likewise h is monic since $g \circ h = $ identity, so that $\text{Im } h \cong C$.   Hence $B = \text{Im } f \oplus \text{Im } h \cong A \oplus C$.

(2)   Let $c_1, \ldots, c_n$ be a basis for C and let $b_i \in B$ be chosen such that $gb_i = c_i$, $i = 1, \ldots, n$ (g is epic). Let $c \in C$.   Then there exist unique integers $\lambda_1, \ldots, \lambda_n$ such that $c = \lambda_1 c_1 + \ldots + \lambda_n c_n$;   define $hc = \lambda_1 b_1 + \ldots + \lambda_n b_n \in B$.   Then h is a homomorphism and $g \circ h = $ identity.   Q.E.D.

We shall make use at the end of this appendix of the following variant of the Splitting Lemma.   Suppose for a moment that A, B, C are vector spaces over the same field and that f, g are linear maps.   It still makes sense to say that

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

is exact: it means that f is injective, $\text{Im } f = \text{Ker } g$ (both are subspaces of B) and g is surjective.   Suppose that C is finite-dimensional and let $c_1, \ldots, c_n$ be a basis for C. As in (2) above we can construct a linear map $h : C \to B$ such that $g \circ h = $ identity, and as in (1) we can show that this implies $B = \text{Im } f \oplus \text{Im } h \cong A \oplus C$, where this is the usual direct sum of vector spaces.   (In fact the same holds even if C is infinite-dimensional but the finite-dimensional case will suffice for the applications here.)

A.20   Examples   (1)   The sequence

$$0 \to A \overset{f}{\to} A \oplus C \overset{g}{\to} C \to 0$$

where $fa = (a, 0)$ and $g(a, c) = c$, is split by
$h : C \to A \oplus C$ defined by $hc = (0, c)$.   This applies whether
C is free or not.

$\underline{A.20}$ (2) Suppose that

$$0 \to \mathbb{Z} \xrightarrow{f} B \xrightarrow{g} \mathbb{Z}_2 \to 0$$

is exact. Let $a = f1$ and let $b \in B$ satisfy $gb = 1 \in \mathbb{Z}_2$.
Then $g(2b) = 1 + 1 = 0$, so that $2b \in \text{Ker } g = \text{Im } f$ and
$2b = \lambda a$ for some integer λ. $\underline{\text{Case I}}$: λ is even.   Say
$\lambda = 2n$; then $h : \mathbb{Z}_2 \to B$ given by $h0 = 0$, $h1 = b - na$
is a splitting homomorphism; hence $B \cong \mathbb{Z} \oplus \mathbb{Z}_2$.
$\underline{\text{Case II}}$: λ is odd.   Say $\lambda = 2n + 1$; then $a = 2(b - na)$
and $b = (2n + 1)(b - na)$.   Hence B is cyclic, generated by
$b - na$ (consider separately elements of B which go to 0 and
1 under g).   But a has infinite order, so B is infinite:
$B \cong \mathbb{Z}$ and f takes 1 to twice a generator of B.   Note
that any attempt to split the sequence will fail, for a split-
ting homomorphism h would have to satisfy
$h1 + h1 = h(1 + 1) = 0$, and there is no element in $\mathbb{Z}$,
besides 0, of finite order.

   (3) Suppose that

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is exact and that A and C are f.g. with generators
$a_1, \ldots, a_n$ and $c_1, \ldots, c_m$.   Let $b_i \in B$ satisfy $gb_i = c_i$,
$i = 1, \ldots, m$.   Then $fa_1, \ldots, fa_n$, $b_1, \ldots, b_m$ generate
B, which is accordingly f.g.

(Proof: let $b \in B$; then $gb = \lambda_1 c_1 + \ldots + \lambda_m c_m$ for suit-
able integers $\lambda_1, \ldots, \lambda_m$.   Hence $b - \lambda_1 b_1 - \ldots - \lambda_m b_m \in$
$\text{Ker } g = \text{Im } f$.   Since $\text{Im } f$ is generated by $fa_1, \ldots, fa_n$
this completes the proof.)   Thus if the sequence is split, and
A and C are known, then B is determined by the splitting
lemma and generators for B are given by the above result.

A more precise formulation when C is free is given in (4)
below.

If A and C are both free and f.g. and the above gen-
erators are bases then B is free with basis fa_1, ..., fa_n,
b_1, ..., b_m.

A.20 (4) In the notation of the first part of (3) suppose
that C is free, so that the sequences splits, and that
c_1, ..., c_m is a basis for C.   Then there is an isomorphism
$\theta : B \rightarrow A \oplus C$ for which $\theta(fa_i) = (a_i, 0)$, $i = 1, ..., n$;
$\theta b_i = (0, c_i)$, $i = 1, ..., m$.   (Proof below.)   Thus
fa_1, ..., fa_n generate the subgroup of B corresponding under
θ to $A \oplus 0$ while b_1, ..., b_m generate the subgroup of B
corresponding under θ to $0 \oplus C$.   We often sum this up by :
$B \cong A \oplus C$ with generators fa_1, ..., fa_n and b_1, ..., b_m.
Other examples of this usage occur in A.22 and A.25.   (Proof
of result: Let $b \in B$.   Then there exist integers λ_i, μ_j
such that $b = \sum \lambda_i fa_i + \sum \mu_j b_j$ by (3).   Write
$\theta b = \left(\sum \lambda_i a_i, \sum \mu_j c_j \right)$. It has to be checked that this is
well-defined, and is an isomorphism.   The first amounts to
checking that if $b = 0$ then $\theta b = (0, 0)$.   Now if $b = 0$
then $gb = 0$; since $g \circ f = 0$ this implies $\sum \mu_j c_j = 0$
and so $\mu_1 = ... = \mu_m = 0$ by the basis property.   Since f
is monic this gives $\sum \lambda_i a_i = 0$.   Hence θ is well-defined.
The rest of the proof is left to the reader.)   The result may
be false if C is not free, even if the sequence splits.
E.g. in (1) above take $A = \mathbb{Z}$, $C = \mathbb{Z}_2$, and put $a_1 = 1$,
$c_1 = 1$, $b_1 = (1, 1)$.

PRESENTATIONS

Let A be a f.g. abelian group with generators a_1, ..., a_n.
Let F be a free abelian group of rank n with basis
x_1, ..., x_n and let $t : F \rightarrow A$ be the epimorphism defined
by $t(\lambda_1 x_1 + ... + \lambda_n x_n) = \lambda_1 a_1 + ... + \lambda_n a_n$ for all integers

$\lambda_1, \ldots, \lambda_n$. (Note that this t is well-defined, since any element of F is uniquely expressible as an integer linear combination of $x_1, \ldots, x_n$.) The short exact sequence

$$0 \to \mathrm{Ker}\, t \xrightarrow{i} F \xrightarrow{t} A \to 0$$

(i being the inclusion) is sometimes called a "presentation" of A; however we shall use this word in a different (but closely related) sense below. Note that by the homomorphism theorem we have $A \cong F/\mathrm{Ker}\, t$. Note also that there is usually a wide choice available for $a_1, \ldots, a_n$ and in particular for n.

The homology groups studied in the text do in fact arise in precisely this way, as the quotient of a free abelian group by a subgroup. It is worth looking more closely at what this means. Each element of the free abelian group F with basis $x_1, \ldots, x_n$ can be expressed uniquely as an integer linear combination of $x_1, \ldots, x_n$. Given a subgroup K of F, it follows from a later result (A.33) that K is free and f.g.; however in practice we are just given generators for K, so let $y_1, \ldots, y_m$ be any generators, not necessarily free, for K. Any element of F/K can be expressed as $\lambda_1 \bar{x}_1 + \ldots + \lambda_n \bar{x}_n$ for integers $\lambda_1, \ldots, \lambda_n$, but the expression is not unique (unless $K = 0$): in fact

$$\sum \lambda_i \bar{x}_i = \sum \mu_i \bar{x}_i \iff \sum (\lambda_i - \mu_i)\bar{x}_i = 0 \iff \sum (\lambda_i - \mu_i)x_i \in K.$$

Now $x_1, \ldots, x_n$ is a basis for F so there are uniquely determined integers α_{ij} ($i = 1, \ldots, m$; $j = 1, \ldots, n$) such that

$$y_i = \alpha_{i1}x_1 + \ldots + \alpha_{in}x_n \qquad (i = 1, \ldots, m)$$

and an element of F belongs to K if and only if it is an integer linear combination of the y's. The group F/K is then completely specified by the information that

$$\bar{x}_1, \ldots, \bar{x}_n \quad \text{generate it}$$

and that $\quad \alpha_{i1}\bar{x}_1 + \ldots + \alpha_{in}\bar{x}_n = 0 \quad$ for $i = 1, \ldots, m;$

for the equations tell us precisely when two integer linear combinations of the $\bar{x}$'s give the same element of F/K.

The equations are called <u>relations</u> and the above description of F/K is called a <u>presentation</u> (by generators and relations). In practice we drop the bars, which are a notational inconvenience, and call

<u>A.21</u>
$$\begin{pmatrix} x_1, \ldots, x_n \\ \alpha_{i1}x_1 + \ldots + \alpha_{in}x_n = 0 \quad (i = 1, \ldots, m) \end{pmatrix}$$

a <u>presentation</u> of F/K.

The formulation just given is designed for situations where, as with the definition of homology groups, we are given a free abelian group F and a subgroup K and wish to identify F/K. It is also possible to start with a f.g. abelian group A, choose F and t as at the beginning of the section, and let $K = \text{Ker } t$. Then replacing each x_i by $tx_i = a_i \in A$ in A.21 we obtain what can be called a presentation of A. For example let $A = \mathbb{Z} \oplus \mathbb{Z}_2$, $a_1 = (1, 0)$, $a_2 = (0, 1)$. Then

$$\begin{pmatrix} a_1, \quad a_2 \\ \quad 2a_2 = 0 \end{pmatrix}$$

is in this sense a presentation of $\mathbb{Z} \oplus \mathbb{Z}_2$. (Compare A.28(1).) We shall stick to the former interpretation in what follows.

<u>A.22</u> <u>Example</u>   Consider the abelian group F/K with presentation

$$\begin{pmatrix} x, \quad y \\ \alpha x + \beta y = 0 \end{pmatrix}$$

where α and β are integers, not both zero. Let h be the

highest common factor of α and β (the highest common factor of 0 and an integer k is, naturally, defined to be $|k|$), and let $\alpha = \alpha_1 h$, $\beta = \beta_1 h$ so that α_1 and β_1 are coprime. Hence there exist integers γ, δ such that $\alpha_1 \gamma + \beta_1 \delta = 1$; write

$$\alpha_1 x + \beta_1 y = u \quad \text{and} \quad -\delta x + \gamma y = v.$$

Then $x = \gamma u - \beta_1 v$; $y = \delta u + \alpha_1 v$ so that u and v generate the free abelian group F with basis x, y and so form a basis for it. (This can be checked directly; the general case follows from A.30 and A.34(2).) In terms of the basis u, v the relation becomes $hu = 0$, so that

$$\begin{pmatrix} u, & v \\ hu & = 0 \end{pmatrix}$$

is a presentation of the same abelian group F/K. Now this presentation is easy to interpret, for the epimorphism

$$F \to \mathbb{Z}_h \oplus \mathbb{Z}$$

given by $\lambda u + \mu v \to (\lambda, \mu)$ $(\lambda, \mu \in \mathbb{Z})$

has kernel the subgroup K generated by hu. Hence

$$F/K \cong \mathbb{Z}_h \oplus \mathbb{Z}.$$

Furthermore the elements of F/K which correspond to generators (1, 0) and (0, 1) of $\mathbb{Z}_h \oplus \mathbb{Z}$ are $\bar{u}$ and $\bar{v}$. We express this by saying $F/K \cong \mathbb{Z}_h \oplus \mathbb{Z}$ with generators $\bar{u}$ and $\bar{v}$. (Compare A.20(4).)

When we are given a presentation of an abelian group as in the above example, the problem is to write the group is some recognizable form. This involves changing the generators and relations, without changing the group which they present, until they become simple enough to spot the group up to isomorphism. Here are three ways in which generators and relations presenting

a group F/K can be changed.

A.23 (1) We can take a new basis for F and re-express the
relations in terms of the new basis. (If F has a basis con-
taining n elements, then any n elements of F which gener-
ate F automatically form a basis - see A.30, A.34(2).)
This was done in A.22, where x and y became u and v.
Neither F nor K is changed, so F/K certainly remains the
same.

(2) We can change the generators for K. This amounts
to replacing the relations by an equivalent set - that is, one
which implies and is implied by the old set. For instance we
could make a succession of changes of the following kind.

$$i^{th}\ \text{relation}\ \rightarrow\ \pm\,i^{th}\ \text{relation}\ +\ \lambda(j^{th}\ \text{relation})$$

where λ is an integer and $i \neq j$. Certainly the new rela-
tions will be equivalent to the old ones and the "left-hand
sides" will define the same K. Note that the relation $0 = 0$
can be removed if ever it occurs.

(3) Consider the presentations

$$\begin{pmatrix} x_1,\ \ldots,\ x_k \\ \\ r_1 = 0,\ \ldots,\ r_m = 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_1,\ \ldots,\ x_k,\ \ x_{k+1},\ \ldots,\ x_n \\ \\ r_1 = 0,\ \ldots,\ r_m = 0 \\ x_{k+1} = 0,\ \ldots,\ x_n = 0 \end{pmatrix}$$

(Here the r's are of course integer linear combinations of
$x_1,\ \ldots,\ x_k$.) All that has happened here is that some extra
generators have been added, and put equal to zero - given with
one hand and taken away with the other. Writing the groups
presented as F_1/K_1 and F/K respectively, we can regard F_1
as a subgroup of F; then K_1 is a subgroup of K and the
inclusion $F_1 \subset F$ gives an isomorphism $F_1/K_1 \cong F/K$.

A.24   The next change of presentation involves elements from

all the above. Suppose we have a presentation of F/K with generators

$$x_1, \ldots, x_k, \ x_{k+1}, \ldots, x_n$$

and relations

(R_1)
$$\begin{aligned} x_{k+1} &- (\text{integer linear combination of } x_1, \ldots, x_k) = 0 \\ &\vdots \\ x_n &- (\text{integer linear combination of } x_1, \ldots, x_k) = 0 \end{aligned}$$

(R_2) Other relations, possibly involving all of $x_1, \ldots, x_n$.

Then (R_1) tells us that $x_{k+1}, \ldots, x_n$ are redundant generators; we proceed to eliminate them. The relations (R_2) can be expressed entirely in terms of $x_1, \ldots, x_k$ by substituting from (R_1) for $x_{k+1}, \ldots, x_n$. Let the result be (R_2'). By A.23(2) the presentation with generators $x_1, \ldots, x_n$ and relations (R_1) and (R_2') defines the same F/K. Now let $x_{k+1}', \ldots, x_n'$ denote the left-hand sides of the relations (R_1). Then $x_1, \ldots, x_k, \ x_{k+1}', \ldots, x_n'$ are also a basis for F so that using (1) the following presentation gives the same F/K.

$$\left(\begin{aligned} &x_1, \ldots, x_k, \ x_{k+1}', \ldots, x_n' \\[4pt] &(R_1') \quad x_{k+1}' = 0, \ldots, x_n' = 0 \\[4pt] &(R_2') \quad \text{As } (R_2), \text{ but with } x_{k+1}, \ldots, x_n \text{ eliminated by} \\ &\hspace{10.5cm} \text{using } (R_1). \end{aligned} \right.$$

Finally we invoke (3) to say that F_1/K_1 defined by

$$\left(\begin{aligned} &x_1, \ldots, x_k \\[4pt] &(R_2') \end{aligned} \right.$$

is isomorphic to F/K, an isomorphism being induced by the
inclusion $F_1 \subset F$.

Here is an example which is typical of the calculations in
the text.

A.25 Example   Consider F/K defined by twelve generators
$x_1, \ldots, x_{12}$ and eleven relations as follows:

$$x_1 - x_8 = 0 \qquad\qquad x_4 + x_5 = 0$$

$$x_2 - x_8 + x_9 = 0 \qquad\qquad x_5 + x_6 = 0$$

$$x_9 - x_{10} - x_{11} = 0 \qquad\qquad x_6 + x_7 - x_{12} = 0$$

$$x_{11} - x_{12} = 0 \qquad\qquad x_1 - x_7 = 0$$

$$x_2 + x_3 = 0 \qquad\qquad x_{10} = 0$$

$$x_3 + x_4 = 0$$

First write the relations in the following equivalent form.

$$x_8 = x_1 \qquad\qquad x_6 = x_2$$

$$x_7 = x_1 \qquad\qquad x_{10} = 0$$

$$x_3 = -x_2 \qquad\qquad x_{11} = x_9$$

$$(\text{from} \quad x_{11} = x_9 - x_{10} = x_9)$$

$$x_4 = x_2$$

$$(\text{from} \quad x_4 = -x_3 = x_2) \qquad\qquad x_{12} = x_9$$

$$(\text{from} \quad x_{12} = x_{11} = x_9)$$

$$x_5 = -x_2$$

$$x_6 + x_7 - x_{12} = 0$$

$$x_2 - x_8 + x_9 = 0$$

This makes it clear that all generators except x_1, x_2 and x_9
are redundant, and invoking A.24 (the first nine relations are
(R_1) and the last two are (R_2)) we are reduced to

$$\left(\begin{array}{l} x_1, \ x_2, \ x_9 \\[4pt] x_2 + x_1 - x_9 \ = \ 0 \\[4pt] x_2 - x_1 + x_9 \ = \ 0 \end{array}\right.$$

Eliminating x_9 in the same way we obtain

$$\left(\begin{array}{l} x_1, \ \ x_2 \\[4pt] x_2 - x_1 + (x_2 + x_1) \ = \ 0, \end{array}\right. \qquad \text{i.e.} \qquad \left(\begin{array}{l} x_1, \ \ x_2 \\[4pt] 2x_2 \ = \ 0 \end{array}\right.$$

The group F_1/K_1 now presented is clearly $\mathbb{Z} \oplus \mathbb{Z}_2$; there is an exact sequence

$$0 \ \to \ K_1 \ \overset{i}{\to} \ F_1 \ \overset{p}{\to} \ \mathbb{Z} \oplus \mathbb{Z}_2 \ \to \ 0$$

where i is inclusion and $p(x_1) = (1, 0)$, $p(x_2) = (0, 1)$. Since $F_1 \subset F$ induces $F_1/K_1 \cong F/K$ we deduce that $F/K \cong \mathbb{Z} \oplus \mathbb{Z}_2$ and that an isomorphism $F/K \to \mathbb{Z} \oplus \mathbb{Z}_2$ is given by $\bar{x}_1 \to (1, 0)$, $\bar{x}_2 \to (0, 1)$, bars denoting cosets with respect to K. We express this by saying "$F/K \cong \mathbb{Z} \oplus \mathbb{Z}_2$ with generators $\bar{x}_1$ and $\bar{x}_2$". (Compare A.20(4).)

What has been said about presentations so far should enable the reader to follow the comparatively few direct calculations of homology groups from presentations in the text, e.g. 4.15, 4.29. It is possible to give a procedure for reducing to a "standard form" any given presentation; this procedure is based on A.22 - 24 and is usually described in terms of row and column operations on the "relation matrix" (α_{ij}) of A.21. For the kind of calculations required in this book the general method is unnecessarily cumbersome, and it is easier to operate directly on the relations as in A.25. However the general method enables one to prove general theorems. The following

theorem, which is given here without proof, receives full dis-
cussion in, for example, Hartley and Hawkes (1970), Chs. 7 and
10, Cairns (1961), pp. 221-30 and MacLane and Birkhoff (1967),
pp. 361-4.

A.26   Theorem   Let F be a free f.g. abelian group and K
a subgroup of F.   Then F/K has a presentation of the form

$$
\left(
\begin{array}{l}
x_1, \ \ldots, \ x_n \\
\delta_1 x_1 = 0, \ \ldots, \ \delta_k x_k = 0
\end{array}
\right.
$$

where $\delta_1, \ \ldots, \ \delta_k$ are integers > 1 and $\delta_1 | \delta_2, \ \delta_2 | \delta_3, \ \ldots,$
$\delta_{k-1} | \delta_k$.   (Of course $k \leqslant n$.   If there are no relations we
say $k = 0$.)   The integer $n - k$, and the δ_i, are determined
by the isomorphism class of F/K.

We call $n - k$ the <u>rank</u> of F/K and the δ_i are called
the <u>invariant factors</u> or <u>torsion coefficients</u> of F/K.   An
alternative approach to rank is described in the next section.

A.27   Corollary   (<u>Classification theorem for f.g. abelian</u>
<u>groups</u>.)

$$
F/K \ \cong \ \mathbb{Z}_{\delta_1} \ \oplus \ \ldots \ \oplus \ \mathbb{Z}_{\delta_k} \ \oplus \ \mathbb{Z} \ \oplus \ \ldots \ \oplus \ \mathbb{Z}
$$

where there are $n - k$ $\mathbb{Z}$'s, i.e. rank $F/K = n - k$.

Note that from the discussion leading to A.21 every f.g.
abelian group is isomorphic to such a quotient F/K.   The
corollary follows from the theorem by examining the short exact
sequence

$$
0 \ \rightarrow \ K \ \rightarrow \ F \ \overset{t}{\rightarrow} \ \mathbb{Z}_{\delta_1} \ \oplus \ \ldots \ \oplus \ \mathbb{Z}_{\delta_k} \ \oplus \ \mathbb{Z} \ \oplus \ \ldots \ \oplus \ \mathbb{Z} \ \rightarrow \ 0
$$

where $tx_i = (0, \ \ldots, \ 0, \ 1, \ 0, \ \ldots, \ 0)$ with the 1 in the
i^{th} position, for $i = 1, \ \ldots, \ n$.

A.28 Remarks and Examples   (1)   It is also possible to use
the method of presentations to calculate quotients B/A where

B is not free.   In fact if B is merely f.g. then, roughly
speaking, we take a presentation for B and add extra relations
to express the fact that generators for A are zero.   For
example take B = $\mathbb{Z} \oplus \mathbb{Z}_2$ and A = subgroup generated by
(3, 1).   Then writing a = (1, 0), b = (0, 1) we have the
presentation

$$\left(\begin{array}{c} a, \ b \\ 2b = 0 \\ 3a + \ b = 0 \end{array} \right.$$

for B/A.   This gives B/A $\cong \mathbb{Z}_6$ with generator a + A.
(Compare the remark following A.21.)

Calculations of the kind just given can be an aid when
using the homomorphism theorem A.15, for if g : B → C is a
homomorphism then Im g $\cong$ B/Ker g.   An example occurs towards
the end of 7.5, though it so happens that there B is free.

[*](2) The classification theorem A.27 implies at once that
a f.g. abelian group containing no element of finite order is
free.   This is not easy to prove directly - indeed it is a
good part of the content of A.27.   For given any abelian group
A the subset T of elements of finite order forms a subgroup,
called the torsion subgroup, of A.   There is a short exact
sequence

$$0 \ \to \ T \ \to \ A \ \to \ A/T \ \to \ 0 \ .$$

Certainly A/T has no element of finite order so, assuming that
A is f.g., the above result implies that A/T is free and the
sequence splits: A $\cong$ T $\oplus$ A/T.   This is the decomposition of
A.27 without the detailed structure of T.

RANK OF A f.g. ABELIAN GROUP

Some of the results below can be obtained from the classifica-
tion theorem A.27, but since we do not prove that theorem a

self-contained discussion is given here. Except for starred
examples, what follows assumes only A.1 - A.20.

Let A be (as usual) an abelian group. We shall con-
struct a vector space $\tilde{A}$ over the rational numbers Q which,
roughly speaking, suppresses the elements of finite order in A.
Some results needed in the text do not involve the elements of
finite order, and these results can be proved readily by this
means.   The construction is a special case of the "tensor
product" construction - see for example MacLane and Birkhoff
(1967), pp. 320-5 - but in the present case the details are
much more straightforward than in the general case, and in
fact follow precisely the standard construction of the rational
numbers from the integers.   In tensor notation, we are con-
structing $A \otimes Q$.

The elements of $\tilde{A}$ are symbols ("fractions") a/n, pro-
nounced "a over n", where $a \in A$ and n is a nonzero
integer. (The tensor notation is $a \otimes \frac{1}{n}$. Beware that in a
general tensor product $A \otimes B$ not every element can be written
in the form $a \otimes b$ with $a \in A$ and $b \in B$.) We identify
a/n with b/m if and only if, for some nonzero integers k, ℓ,
we have

$$ka = \ell b, \qquad kn = \ell m.$$

(This can be expressed alternatively by saying that we impose
the corresponding equivalence relation on pairs $(a, \frac{1}{n})$ and
then work with the equivalence classes.)   In particular
a/n = ka/kn for any nonzero integer k.

The rules for addition and scalar multiplication in $\tilde{A}$
are:

$$a_1/n_1 + a_2/n_2 \quad = \quad (n_2 a_1 + n_1 a_2)/n_1 n_2$$

$$r(a/n) \quad = \quad pa/qn \text{ where } r = p/q \text{ is a rational}$$
$$\text{number and } p, q \text{ are integers.}$$

These rules respect the identifications, and $\tilde{A}$ is a rational

vector space. The zero vector is $0/1$. Note the following
important fact: $a/n = 0/1$ if and only if a has finite
order. (Proof: if a has finite order $k \geqslant 1$, so that
$ka = 0$, then $a/n = ka/kn = 0/kn = 0/1$. If $a/n = 0/1$
then, for some nonzero integers k, l we have $ka = 0$, $kn = l$.
Hence a has finite order.) Thus passing from A to $\tilde{A}$
suppresses the elements of finite order.

A.29   Let A be f.g. with generators $a_1, \ldots, a_n$. Then
$a_1/1, \ldots, a_n/1$ generate $\tilde{A}$ as a rational vector space, and
accordingly $\tilde{A}$ is finite dimensional. The <u>rank</u> of A is
defined to be the dimension of $\tilde{A}$. (This is reconciled with
A.26 and A.27 in A.32(2) below.)

A.30   <u>Theorem</u>   Let A be free and f.g. with basis $a_1, \ldots, a_n$.
Then rank $A = n$ and every basis for A has n elements.

<u>Proof</u>   We show that $a_1/1, \ldots, a_n/1$ is a basis for
$\tilde{A}$, which accordingly has dimension n so that rank $A = n$.
The second result follows because every basis for $\tilde{A}$ contains
n elements.

   Certainly, as in A.29, the stated elements generate $\tilde{A}$.
For $a \in A$ has the form $\lambda_1 a_1 + \ldots + \lambda_n a_n$ for integers λ_i,
and we have $a/m = \lambda_1/m \cdot a_1/1 + \ldots + \lambda_n/m \cdot a_n/1$ so that
every element of $\tilde{A}$ is a rational linear combination of
$a_1/1, \ldots, a_n/1$. To see that these vectors are linearly inde-
pendent, suppose that, for some rational numbers $r_1, \ldots, r_n$,
we have $r_1(a_1/1) + \ldots + r_n(a_n/1) = 0/1$. Multiplying
through by the product, k say, of the denominators of the r_i
we obtain $kr_1(a_1/1) + \ldots + kr_n(a_n/1) = 0/1$, where the
coefficients are now integers. Hence $kr_1 a_1 + \ldots + kr_n a_n$
has finite order and, A being free, must in fact be 0.
Using $k \neq 0$ and the basis property we have
$r_1 = \ldots = r_n = 0$.   Q.E.D.

   Examples: $\mathbb{Z}^n$ has rank n; the trivial group has rank
0; $\mathbb{Z} \oplus \mathbb{Z}_2$ has rank 1.

Now let $f : A \to B$ be a homomorphism. Define
$\tilde{f} : \tilde{A} \to \tilde{B}$ by $\tilde{f}(a/n) = fa/n$. This is well-defined (i.e.
if $a/n = b/m$ then $fa/n = fb/m$) and is a linear map from
$\tilde{A}$ to $\tilde{B}$. (What we have now defined is a "functor" from the
"category" of abelian groups and homomorphisms to the "category"
of rational vector spaces and linear maps. Compare MacLane
and Birkhoff (1967), p.24. The next remark says that the
functor is "exact".)

Suppose that

$$A \overset{f}{\to} B \overset{g}{\to} C$$

is exact (at B). Then

$$\tilde{A} \overset{\tilde{f}}{\to} \tilde{B} \overset{\tilde{g}}{\to} \tilde{C}$$

is exact (at $\tilde{B}$), i.e. the subspaces $\text{Im } \tilde{f}$ and $\text{Ker } \tilde{g}$ of $\tilde{B}$
are equal.

(Proof: $\widetilde{gf}(a/n) = \tilde{g}(fa/n) = gfa/n = 0/n = 0/1$ so
$\text{Im } \tilde{f} \subset \text{Ker } \tilde{g}$. Suppose $\tilde{g}(b/n) = 0/1$. Then gb has
finite order k say, and $g(kb) = k(gb) = 0$. Hence
$kb \in \text{Ker } g = \text{Im } f$, i.e. $kb = fa$ for some $a \in A$. Then
$b/n = kb/kn = fa/kn = \tilde{f}(a/kn)$. Hence $\text{Ker } \tilde{g} \subset \text{Im } \tilde{f}$.)
Thus the operation $\sim$ takes exact sequences to exact sequences.
Recall from p. 288 that every short exact sequence of vector
spaces and linear maps is split.

A.31 Theorem   Suppose that

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

is an exact sequence of f.g. abelian groups. Then

$$\text{rank } B = \text{rank } A + \text{rank } C.$$

Proof   The corresponding short exact sequence of vector
spaces splits, so $\tilde{B} \cong \tilde{A} \oplus \tilde{C}$; hence

$$\text{rank } B = \dim \tilde{B} = \dim (\tilde{A} \oplus \tilde{C}) = \dim \tilde{A} + \dim \tilde{C}$$
$$= \text{rank } A + \text{rank } C. \qquad \text{Q.E.D.}$$

A.32 **Examples** (1) If A and C are f.g. then rank $(A \oplus C)$ = rank A + rank C. (Apply A.31 to the short exact sequence of A.20(1).)

(2) We are now in a position to see that, defining rank $A = \dim \tilde{A}$ as in A.29, the rank of the group

$$A = \mathbb{Z}_{\delta_1} \oplus \ldots \oplus \mathbb{Z}_{\delta_k} \oplus \mathbb{Z}^{n-k}$$

of A.27 really is $n - k$. For $A = T \oplus \mathbb{Z}^{n-k}$ where T is finite (i.e. $\tilde{T} = 0$) and we have

$$\text{rank } A = \text{rank } T + \text{rank } \mathbb{Z}^{n-k} \quad \text{by (1) above}$$

$$= 0 + (n - k) \quad \text{using } \tilde{T} = 0 \text{ and A.30.}$$

(3) Let

$$0 \rightarrow A_1 \overset{f_1}{\rightarrow} A_2 \overset{f_2}{\rightarrow} \ldots \rightarrow A_{n-1} \overset{f_{n-1}}{\longrightarrow} A_n \overset{f_n}{\rightarrow} 0$$

be an exact sequence of f.g. abelian groups. We can define $n - 1$ short exact sequences

$$0 \rightarrow \text{Ker } f_i \rightarrow A_i \rightarrow \text{Im } f_i \rightarrow 0 \quad (i = 1, \ldots, n-1)$$
$$\| $$
$$\text{Ker } f_{i+1}$$

(Compare A.16(5).) Applying A.31 to each of these yields the alternating sum of ranks theorem: $\sum_{i=1}^{n} (-1)^i \text{ rank } A_i = 0.$

A.33 **Theorem** Suppose that B is free and f.g. of rank m, and that A is a subgroup of B. Then A is free and f.g. of rank $\leq m$.

Proof We proceed by induction on m. The result for $m = 1$ follows from the fact that any nontrivial subgroup of $\mathbb{Z}$ consists of the multiples of a fixed nonzero integer and is

therefore itself infinite cyclic.   Now assume the result is
proved when B has rank < m (m ≥ 2) and choose a basis
b_1, ..., b_m for the free group B of rank m.   Define
t : B → $\mathbb{Z}$ by $t(\lambda_1 b_1 + ... + \lambda_m b_m) = \lambda_m$; then Ker t is
free of rank m - 1, with basis b_1, ..., b_{m-1}.   We shall
apply the induction hypothesis to Ker t by letting C = tA
and t' : A → C be the epimorphism given by t'a = ta for
all a ∈ A.   Then Ker t' is a subgroup of Ker t and is
therefore free of rank ≤ m - 1 by the induction hypothesis.
The short exact sequence

$$0 \;\to\; \text{Ker } t' \;\overset{i}{\to}\; A \;\overset{t'}{\to}\; C \;\to\; 0$$

is split (A.19) since C is a subgroup of $\mathbb{Z}$ and therefore
free of rank ≤ 1.   Thus A ≅ Ker t' ⊕ C is free of rank ≤ m.
(Once A is known to be f.g. the inequality of ranks follows
also from the fact that $\tilde{A}$ is a subspace of $\tilde{B}$.)   Q.E.D.

A.34 <u>Examples</u>   (1) Suppose that A and B are free and
f.g. of the same rank and that f : A → B is an epimorphism.
Then f is in fact an isomorphism.   (Consider the short exact
sequence 0 → Ker f $\overset{i}{\to}$ A $\overset{f}{\to}$ B → 0.   This splits by A.19
and therefore A ≅ Ker f ⊕ B.   Hence rank (Ker f) = 0 and
since Ker f is free by A.33 it must be zero, i.e. f is
monic.)   In contrast f : $\mathbb{Z}$ → $\mathbb{Z}$ defined by fn = 2n is
monic without being epic.

        (2) Suppose that A if free and f.g. of rank n and that
a_1, ..., a_n generate A.   Then in fact they are a basis for A.
(This is just (1) applied to the epimorphism $\mathbb{Z}^n$ → A given
by $(\lambda_1, ..., \lambda_n) \to \lambda_1 a_1 + ... + \lambda_n a_n$.)

        (3) Let A be free and f.g. and suppose that a and b
are elements of A with the property that αa = βb for some
integers α, β not both zero.   Then a and b have a "common
factor", i.e. there is an element c ∈ A of which a and b
are both multiples.   (When A = $\mathbb{Z}$ the hypothesis   αa = βb is

redundant since we can take β = a, α = b, unless a = b = 0
when we can take α = β = 1.) To see this suppose a $\neq$ 0, so
that $\beta \neq 0$, and let C be the subgroup of A generated by
a and b. Denoting by $\langle a \rangle$ the infinite cyclic subgroup of
C generated by a, it follows that $C/\langle a \rangle$ is finite, since
a nonzero multiple βb of b belongs to $\langle a \rangle$. Hence
rank C = 1 by A.31, i.e. C is infinite cyclic and a suit-
able c is a generator for C.

A.34 (4) A nonzero element a of an abelian group A is
called <u>indivisible</u> if it is not a multiple of any element of A
besides $\pm$ a (i.e. if a = λb, $\lambda \in \mathbb{Z}$, b $\in$ A => λ = $\pm$1).
Thus for example Q contains no indivisible element, and the
only indivisible elements of $\mathbb{Z}$ are $\pm$1. Let A be a free
f.g. abelian group, and let a $\in$ A be nonzero. Then the
following four statements are equivalent:

(i)   a is indivisible.

(ii) $A/\langle a \rangle$ is free abelian, $\langle a \rangle$ being the (infinite
cyclic) subgroup of A generated by a.

(iii)   a is an element of some basis for A.

(iv)   Given any basis x_1, ..., x_n for A the unique
expression a = $\lambda_1 x_1$ + ... + $\lambda_n x_n$ (λ_1, ..., $\lambda_n \in \mathbb{Z}$) has
the highest common factor of λ_1, ..., λ_n equal to 1.

(The proofs (i) => (ii) => (iii) => (iv) => (i) are fairly
straightforward. Here are some hints. (i) => (ii): Show
$A/\langle a \rangle$ has no nonzero element of finite order, perhaps using
(3) above. (ii) => (iii): Use the short exact sequence

0 $\to$ $\langle a \rangle$ $\to$ A $\to$ $A/\langle a \rangle$ $\to$ 0. (iii) => (iv): Let
a_1 = a, a_2, ..., a_n be a basis containing a. Let Λ be the
n x n integral matrix expressing the a's in terms of the
x's; thus the first row of Λ is λ_1, ..., λ_n. Since the
a's are a basis we can consider the matrix Λ^{-1} expressing

the x's in terms of the a's, and we have $\Lambda\Lambda^{-1} = I_n$, the
n x n identity matrix. Hence $\det \Lambda \cdot \det \Lambda^{-1} = 1$, so
$\det \Lambda = \pm 1$ (i.e. Λ is unimodular). Expanding $\det \Lambda$ by
the first row shows that the highest common factor of
$\lambda_1, \ldots, \lambda_n$ is 1. (iv) => (i): follows from the definition
of basis.)

<u>A.34</u>[*](5) Let $f : A_1 \to A_2$ be a homomorphism of abelian
groups and let T_i be the torsion subgroup of A_i (i = 1, 2),
that is the subgroup of elements of finite order. Then we
have $f(T_1) \subset T_2$ so f restricts to a homomorphism
$f' : T_1 \to T_2$ defined by $f't = ft$ for all $t \in T_1$. (This is
another example of a functor (compare p. 302), this time from
the category of abelian groups and homomorphisms to itself.
This functor is not exact, as can be verified by considering
the exact sequence $\mathbb{Z} \to \mathbb{Z} \oplus \mathbb{Z}_2 \to \mathbb{Z}$ where the homomorphisms
are $m \to (0, m \bmod 2)$ and $(n, p) \to n$. But see the next
remark.)

Suppose that A_1 is f.g. and Ker f is finite. Then
the exactness of

$$A_1 \xrightarrow{f} A_2 \xrightarrow{g} A_3$$

implies the exactness of

$$T_1 \xrightarrow{f'} T_2 \xrightarrow{g'} T_3 .$$

(Hints for proof. Clearly $g' \circ f' = 0$. Identify T_1
with $0 \oplus T_1$ and use A.27 to write $A_1 \cong F_1 \oplus T_1$ with F_1
free and f.g. For simplicity write $A_1 = F_1 \oplus T_1$. Now
$g't_2 = 0 \Rightarrow t_2 = f(x_1, t_1)$ for some $(x_1, t_1) \in A_1$.
Hence $t_2 = f(x_1, 0) + f(0, t_1)$. Deduce $f(x_1, 0)$ has
finite order, k say. Hence $(kx_1, 0) \in$ Ker f but has in-
finite order unless $x_1 = 0$.)

As an example suppose B is a subgroup of the f.g. abelian
group A and suppose $A/B \cong A$. We show B = 0. Using the
exact sequence

$$0 \to B \overset{i}{\to} A \overset{\pi}{\to} A/B \to 0$$

(compare A.16 Ex (6)), and A.31, it follows that rank B = 0, i.e. B is finite. The result above then shows that

$$0 \to B \to T \to T' \to 0$$

is exact where T, T' are the torsion subgroups of A and A/B respectively. But T and T' have the same number of elements, so B = 0.

A.34*(6) Consider an exact sequence

$$A \overset{f}{\to} A \oplus B \overset{g}{\to} B$$

(i) If A and B are finite, then f is monic and g is epic. (Construct two short exact sequences from f and g as in A.16 Ex(5) and use A.16 Ex(8).)

(ii) If A and B are f.g. then Ker f is finite. (Show this and B/Im g is finite by applying the alternating sum of ranks theorem (A.32(3)) to the exact sequence $0 \to$ Ker f $\to A \to A \oplus B \to B \to$ B/Im g $\to 0$.) Hence if A, B are f.g. then with f' as in (5) above we have Ker f = Ker f' = 0 by (5) and (i). Hence: if A and B are f.g. then f is monic. It is not true in general that g is epic.

(iii) Take B = 0 in (ii). We have: if A is f.g. then any epimorphism $A \to A$ is an isomorphism. (Compare A.34(1). It is easy to construct counterexamples when A is not f.g.)

(iv) If A is f.g. and B is finite then f is monic and g is epic.

References

Apart from references specifically cited in the text, the following list contains a few books which can be regarded as parallel, or further, reading (possibly both). In the latter respect the list is by no means exhaustive. Books which could be regarded as parallel reading are prefixed by †. More advanced books are prefixed by ††.

†Agoston, M. Algebraic Topology, Marcel Decker AG (1976)

††Ahlfors, L. V. & L. Sario, Riemann Surfaces, Princeton University Press (1960).

Appel, K. A. & W. H. Haken, "The solution of the four-color-map problem", Scientific American, Oct. 1977, p. 108.

†Arnold, B. H. Intuitive Concepts in Elementary Topology, Prentice-Hall (1962).

†Berge, C. (Trans. A. Doig) The Theory of Graphs and its Applications Methuen/Wiley (1962).

Bernstein, I. N., I. M. Gelfand & V. A. Ponomarev, "Coxeter Functors and Gabriel's Theorem", Russian Mathematical Surveys Vol. 28 (1973), pp. 17-32.

Biggs, N. L., E. K. Lloyd & R. J. Wilson, Graph Theory 1736-1936, Oxford University Press (1976).

Bing, R. H. "Some aspects of the topology of 3-manifolds" In Lectures on Modern Mathematics Vol. II, Wiley (1964)

†Brown, R. Elements of Modern Topology, McGraw-Hill (1968)

†Busacker, R. G. & T. L. Saaty, Finite Graphs and Networks: an introduction with applications. McGraw-Hill (1965)

†Cairns, S. S.   Introductory Topology,   Ronald Press Co.,
     New York (1961)

††Cooke, G. E. & R. L. Finney,   Homology of Cell Complexes
     Princeton University Press (1967)

†Courant, R. & H. E. Robbins,   What is Mathematics?   Oxford
     University Press (1947)

††Eilenberg, S. & N. E. Steenrod,   Foundations of Algebraic
     Topology,   Princeton University Press (1952)

Fáry, I.   "On straight line representations of planar graphs",
     Acta Sci. Math. Szeged   Vol. 11 (1948), pp. 229-233

†Flegg, H. G.   From Geometry to Topology,   English Universities
     Press (1974)

Franklin, P.   "A six colour problem",   Journal of Mathematics
     and Physics   Vol. 13 (1934), pp.363-9.

†Fréchet, M. & Ky Fan,   Initiation to Combinatorial Topology,
     Prindle, Weber and Schmidt (1967)

††Fuchs, L.   Infinite abelian groups,   Academic Press Vol. I
     (1970), Vol. II (1973)

Gardner, M.   "Of sprouts and Brussels sprouts, games with a
     topological flavor", Scientific American,   July 1967,
     p. 114.

Gardner, M.   "On the remarkable Császár polyhedron and its
     applications in problem solving",   ibid., May 1975,
     p. 102.

Gardner, M.   "In which a mathematical aesthetic is applied to
     modern minimal art", ibid., Nov. 1978, p. 20.

††Glaser, L. C.   Geometrical Combinatorial Topology, Vol. I,
     Van Nostrand Reinhold (1970)

††Gramain, A.   Topologie des Surfaces,   Presses Universitaires
     de France (1971)

†Griffiths, H. B.   Surfaces,   Cambridge University Press (1976)

Halperin, S. & D. Toledo,   "Stiefel-Whitney homology classes",
     Annals of Mathematics, Vol. 96 (1972), pp. 511-525

††Harary, F. (ed.)   Graph Theory and Theoretical Physics,
     Academic Press (1967)

††Harris, B. (ed.)   Graph Theory and its Applications,
     Academic Press (1970)

†Hartley, B. & T. O. Hawkes,   Rings, Modules and Linear Algebra
     Chapman & Hall (1970)

Heawood, P. J.   "Map colour theorem",   Quarterly Journal of
     Mathematics,   Vol. 24 (1890), pp. 332-338

†Hilbert, D.& S. Cohn-Vossen,   Geometry and the Imagination
        Chelsea Publishing Company, New York (1952)

††Hilton, P. J. & S. Wylie,   Homology Theory,   Cambridge
        University Press (1960)

††Hocking, J. G. & G. S. Young,   Topology,   Addison-Wesley
        (1961)

††Kaplansky, I.   Infinite abelian groups,   University of
        Michigan Press (1954)

††Kelley, J. L.   General Topology,   D. Van Nostrand (1955).

Kline, M.   Mathematical Thought from Ancient to Modern Times,
        Oxford University Press (1972).

Lakatos, I.   Proofs and Refutations,   Cambridge University
        Press (1977).

†Lefschetz, S.   Introduction to Topology,   Princeton University
        Press (1949).

†Lefschetz, S.   Applications of Algebraic Topology,   Springer-
        Verlag (1975).

MacLane, S.   "A combinatorial condition for planar graphs",
        Fundamenta Mathematicae   Vol. 28 (1937), pp. 22-32.

†MacLane, S. & G. D. Birkhoff,   Algebra,   Macmillan (1967).

†Massey, W. S.   Algebraic Topology: An Introduction,
        Harcourt Brace & World (1967).

††Maunder, C. R. F.   Algebraic Topology,   Van-Nostrand Reinhold
        (1970).

Milnor, J. W.   "Two complexes which are homeomorphic but combin-
        atorially distinct",   Annals of Mathematics   Vol. 74
        (1961), pp. 575-590.

†Milnor, J. W.   Topology from the Differentiable Viewpoint,
        University of Virginia Press (1965).

††Milnor, J. W. & J. D. Stasheff,   Characteristic Classes,
        Annals of Mathematics Studies No. 76,   Princeton
        University Press (1974).

†Newman, M. H. A.   Elements of the Topology of Plane Sets of
        Points,   Cambridge University Press, Second Edition
        (1951).

††Ore, O.   Theory of Graphs,   American Mathematical Society
        Colloquium Publications,   Vol. 38 (1962).

Poincaré, H.   "Cinquième complément à l'analysis situs",
        Rend. del Circolo Math. di Palermo   Vol. 18 (1904),
        pp. 45-110.

Pont, J.-C.   La Topologie Algébrique des Origines a Poincaré,
        Presses Universitaires de France (1974).

Puppe, D.   "Korrespondenzen in abelschen Kategorien",   Mathematische Annalen,   Vol. 148 (1962), pp. 1-30.

Ringel, G.   "Wie man die geschlossenen nichtorientierbaren Flächen in möglichst wenig Dreiecke zerlegen kann", Mathematische Annalen Vol. 130 (1955), pp. 317-326.

††Ringel, G.   Map Color Theorem,   Springer-Verlag (1974).

Ringel, G. & J. W. T. Youngs,   "Solution of the Heawood Conjecture", Proceedings of the National Academy of Sciences U.S.A.,   Vol. 60 (1968), pp. 438-445.

†Rourke, C. P. & B. J. Sanderson,   Introduction to Piecewise-Linear Topology,   Springer-Verlag (1972).

Saaty, T. L. & P. C. Kaimer,   The Four Color Problem: Assaults and Conquest,   McGraw-Hill (1978).

Schafer, J. A.   "Representing homology classes on surfaces", Canadian Mathematical Bulletin Vol. 19 (1976), pp. 373-4.

†Seifert, H. and W. Threlfall,   Lehrbuch der Topologie, Chelsea Publishing Company, New York (Reprint of the 1934 edition).

†Singer, I. M. and J. A. Thorpe,   Lecture Notes on Elementary Topology and Geometry,   Scott Foresman (1967).

††Spanier, E. H.   Algebraic Topology,   McGraw-Hill (1966).

†Wall, C. T. C.   A Geometric Introduction to Topology, Addison-Wesley (1972).

Whitney, H.   "Non-separable and planar graphs",   Transactions of the American Mathematical Society   Vol. 34 (1932), pp. 339-362.

†Wilson, R. J.   Introduction to Graph Theory,   Oliver and Boyd (1972).

Youngs, J. W. T.   "The Heawood map-colouring conjecture", pp. 313-354 in Harary (1967).

Youngs, J. W. T.   "The mystery of the Heawood conjecture", pp. 17-50 in Harris (1970).

List of
notation

313

III. <u>Greek characters</u>

Index

Not all phrases containing more than one keyword are indexed under each keyword. Thus "abstract graph" appears only under "graph", "homology group" only under "homology", etc. An underlined reference is to a definition or, failing that, a drawing.

vertex ctd.
  of graph <u>13</u>
  Kirchhoff equation for 18, 45
  order of <u>17</u>, 42
  of simplex 92
  terminal 23

Wall, C. T. C. 41
Whitney, H. 39
Whitney homology classes 227
Wilson, R. J. 84, 270
Wylie, S. 162, 171, 227

Young, G. S. 171
Youngs, J. W. T. 80