The submitted manuscript has been authored by a contractor of the U.S. Government under contract No. W31 109-ENG 38. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

ANL-HEP-PR-78-23
May, 1978

"INTRODUCTORY LECTURES ON FIBRE BUNDLES AND TOPOLOGY FOR PHYSICISTS"*

Gerald H. Thomas
High Energy Physics Division
Argonne National Laboratory
Argonne, Illinois 60439

ABSTRACT

These lectures may provide useful background material for understanding gauge theories, particularly the non-perturbative effects such as instantons, and monopoles. These notes are designed to introduce high energy physicists to the mathematical language of topology and fibre bundles.

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

*Work performed under the auspices of the United States Department of Energy

TABLE OF CONTENTS

	Page
I. INTRODUCTION	6
A. Apologies	6
B. Path Integrals and Homotopy Classes	7
C. Homotopy Classes and Fibre Bundles	17
1. Inequivalent vacua	17
2. Magnetic monopoles	21
II. LANGUAGE OF TOPOLOGY	30
A. Topological Spaces	30
B. Continuous Maps	32
C. Product Spaces, Quotient Spaces, and Subspaces	36
D. Connectivity	39
E. Pasting	41
F. Compactness	43
G. Global v.s. Local	45
III. FIBRE BUNDLES	48
A. Informal Discussion	48
B. Precise Definition of Fibre Bundle	50
C. Bundle Maps	57
D. Bundle Transformation Theorem	60
E. Bundle Construction Theorem	65
IV. BUNDLES FROM TOPOLOGICAL GROUPS	71
A. Topological Groups	71
B. Quotient Space G/H	72

	Page
C. Transformation Group on X	73
D. $X \sim G/H$	75
E. Bundle Structure Theorem	77
1. Local cross section assumption	77
2. Statement of theorem	78
3. p is a projection	79
4. Coordinate neighborhoods	80
5. Pasting	81
6. ϕ_g are coordinate functions	84
7. H/K_0 is group of bundle	85
8. Structure is independent of local cross section	86
9. Bundle maps by left translations	87
F. Perspective	88
V. DIFFERENTIABLE MANIFOLDS AND TENSOR BUNDLES	90
A. Differentiable Manifolds	90
B. Coordinate Transformations	93
C. Tensor Bundles	95
VI. COVERING HOMOTOPY THEOREMS	102
A. Bundle Homotopy	102
B. Covering Homotopy Theorem for Product Bundle	104
C. Covering Homotopy Theorem for X Compact	105
1. Choosing the right cover	105
2. Digression on separation properties for compact X	107
a. Compact is regular	108
b. Compact is normal	108
c. Urysohn's theorem	109
3. Completing the choice of cover for 2-fold induction step	109
4. Induction step	113

	Page
D. Covering Homotopy Theorem for Non-Compact X	115
1. Choosing the right cover	116
2. Completing the choice of cover for induction step	117
3. Induction step	118
4. Loose ends	124
E. Appendix on Digression on Separation Properties	127
1. Hausdorff	127
2. Regular	127
3. Normal	128
4. Proof that compact is Hausdorff, regular and normal	128
5. Proof that metric spaces are normal	129
6. Proof of Urysohn's theorem	131
VII. THE HOMOTOPY GROUPS	137
A. Homotopies	137
B. Group Structure of Homotopy Classes	141
C. Homomorphisms Between Homotopy Groups	146
D. Basic Properties	148
1. Identity	148
2. Products	148
3. Commutativity of ∂	149
4. Exactness	149
5. Homotopy	155
6. Trivial space	156
7. Fundamental theorem	156
E. Exact Sequence for Bundles	161
VIII. THE HOMOLOGY GROUPS	164
A. Simplexes and Mod 2 Homology Groups	165
1. Simplexes	165
2. Complexes	166

	Page
3. Chain groups	167
4. Boundary homomorphism ∂	168
5. p -boundary and p -cycle groups	169
6. p -homology groups	170
7. Examples	171
B. Oriented Simplexes and Homology Groups with Coefficient G	177
1. Oriented simplex and complex	177
2. Boundary operator ∂	180
3. Chain group	181
4. p -boundary, p -cycle and p -homology groups	183
5. Example	183
6. Euler and Betti numbers	186
7. Another example; torsion numbers	187
C. Relative Homology Theory	194
1. Simplicial maps	194
2. Commutativity with ∂	195
3. ϕ_p induces homomorphisms between homology	195
4. Pairs of spaces: chain, boundary, cycle and homology groups	196
5. Exact sequences	197
D. Singular Homology Theory	198
1. Axioms	199
2. Singular p -simplexes and chain groups	200
3. Boundary, cycle and homology groups	201
4. Relative chain, boundary, cycle and homology groups	202
5. Relation to the axioms	202
E. Connection Between Homology and Homotopy	203

ACKNOWLEDGMENTS

These lectures would not have been written without the interest of my colleagues in learning the mysteries of topology. I am especially grateful to Stanley Fenster, Evelyn Monsay, Murray Peshkin and Dennis Sivers for patiently listening to my attempts at explaining a difficult subject, for what they taught me about the subject, and for their editorial suggestions on these notes. I also thank Louis Michel for getting me interested in the subject of topology and for teaching me it has important applications in physics.

I. INTRODUCTION

A. Apologies

There have been a number of papers in the last few years on topological aspects of gauge field theories. Unfortunately topology is not a standard part of a high energy physicists' training. Moreover, learning topology from mathematics textbooks can be a formidable task. There is some evidence, though, that the work necessary to learn topology is justified by the increased understanding it makes possible of the gauge theories for strong interactions.

These lecture notes were written to help make the mathematical texts more accessible to those people willing to attempt the study of topology and its applications to physics. As such, they do not replace the many excellent textbooks written on the subject. Indeed the material for these lectures has been drawn from these sources. I have written out the proofs in great detail, assuming the reader has no particular background in mathematics. This makes the exposition considerably longer than what exists in the original sources. I hope the longer exposition will help the reader to learn the mathematical language, which, once learned, allows an elegant summary of the ideas of topology. I hope also that these notes will suggest to the reader that the subject is in fact not overly difficult, and the strangeness of the arguments are due only to lack of familiarity. The reader should look up the original sources after mastering some of the language, since they contain an order of magnitude more results than are

included in these notes, and their discussions are decidedly more elegant.

Although these lectures are primarily on the mathematics of topology, it is perhaps a useful introduction to sketch a few examples of the topological reasoning one is starting to find in the physics literature. For more details, one can consult the excellent set of notes by [S. Coleman] and by [V. N. Popov].

B. Path Integrals and Homotopy Classes

Let us calculate the elementary Green's function for a free particle by Feynman's path integral. [R. P. Feynman and A. R. Hibbs]. To start, consider a particle of mass m moving freely in 1 dimension. The classical action for such a system is

$$S[x] = \int_{t_1}^{t_2} dt \frac{1}{2} m \dot{x}^2$$

where we assume the particle is at $x(t_1) = x_1$ at time $t = t_1$, $i = a, b$. The quantum mechanical amplitude for the particle to propagate from x_a to x_b in time $t_b - t_a$ is obtained by summing $e^{\frac{i}{\hbar} S[x]}$ over all possible paths $x(t)$:

$$K(x_b, t_b; x_a, t_a) = \sum_{\text{paths}} e^{\frac{i}{\hbar} S[x]}$$

According to Feynman, this sum can be performed by considering the limit

$$\lim_{N \rightarrow \infty} \frac{1}{A} \int \frac{dx_1}{A} \dots \int \frac{dx_N}{A} \exp \frac{im}{2\hbar} \Delta t \left\{ \left(\frac{x_1 - x_a}{t_1 - t_a} \right)^2 + \left(\frac{x_2 - x_1}{t_2 - t_1} \right)^2 + \dots + \left(\frac{x_N - x_{N-1}}{t_N - t_{N-1}} \right)^2 + \left(\frac{x_b - x_N}{t_b - t_N} \right)^2 \right\}$$

where all time intervals are equal to $\Delta t = \frac{t_b - t_a}{N}$, and the normalization Feynman finds to be

$$A = \left(\frac{2\pi i \hbar \Delta t}{m} \right)^{\frac{1}{2}}$$

The integrals over x_k are Gaussians and give

$$\begin{aligned} K(x_b t_b; x_a t_a) &= \lim_{N \rightarrow \infty} \left(\frac{2\pi i \hbar (t_b - t_a)}{m} \right)^{-\frac{1}{2}} \exp \left\{ \frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)} \right\} \\ &= \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-\frac{1}{2}} \exp \left(\frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)} \right) \end{aligned}$$

[This is for $t_b > t_a$; by assumption $K=0$ for $t_b < t_a$.] You may be more familiar with this in the time-momentum representation:

$$\begin{aligned} \tilde{K}(p_b, t_b; p_a, t_a) &= \int_{-\infty}^{\infty} dx_b e^{\frac{-ip_b x_b}{\hbar}} \int_{-\infty}^{\infty} dx_a e^{\frac{ip_a x_a}{\hbar}} K(x_b t_b; x_a t_a) \\ &= (2\pi \hbar) \delta(p_b - p_a) e^{\frac{-i}{\hbar} \left(\frac{p_b^2}{2m} \right) (t_b - t_a)} \end{aligned}$$

In the energy-momentum representation we have

$$\begin{aligned}
 k(p_b, E_b; p_a, E_a) &= \int_{-\infty}^{\infty} dt_b \int dt_a \tilde{K}(p_b t_b; p_a t_a) \frac{1}{\hbar} (E_b t_b - E_a t_a) \\
 &= (2\pi\hbar)^2 \delta(p_b - p_a) \delta(E_b - E_a) \frac{1/\hbar}{E_b - p_b^2/2m + i\epsilon}
 \end{aligned}$$

which is perhaps the most familiar form.

Now we can solve a different 1-dimensional problem by supposing the particle is constrained to travel on a closed ring (a one dimensional problem with periodic boundary conditions). Let s be a measure of the arc-length so that the action is

$$S[s] = \int_{t_a}^{t_b} dt \frac{1}{2} m \dot{s}(t)^2$$

where now $s(t_b) - s(t_a) = \bar{s} + nL$, L is the total arc length of the ring and $0 \leq \bar{s} < L$ is the distinguishable distance the particle can travel. We assume that no observable depends upon the winding number n ; we therefore must sum over all winding numbers before computing the propagator. We don't sum over the equivalent initial points $s(t_a)$. The free propagator is

$$\begin{aligned}
 K_L(\bar{s}, t_b - t_a) &= \sum_{n=-\infty}^{\infty} \lim_{N \rightarrow \infty} \frac{1}{A} \int_{-\infty}^{\infty} \frac{ds_1}{A} \dots \int \frac{ds_N}{A} \\
 &\exp \frac{im}{2\hbar} \Delta t \left\{ \left(\frac{s_1 - 0}{t_1 - t_a} \right)^2 + \dots + \left(\frac{\bar{s} + nL - s_N}{t_b - t_N} \right) \right\}
 \end{aligned}$$

$$= \sum_{n=-\infty}^{\infty} \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-\frac{1}{2}} \exp \left(\frac{i m (\bar{s} + nL)^2}{2\hbar (t_b - t_a)} \right),$$

which is the sum over just the free particle propagators which are assumed to be indistinguishable.

We can go now to momentum space, where p is the variable canonical to $\bar{s}$:

$$\begin{aligned} \tilde{K}_L(p_b, t_b - t_a) &= (2\pi\hbar) \delta(p_b - p_a) \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-\frac{1}{2}} \\ &\sum_{n=-\infty}^{\infty} \int_0^L d\bar{s} \exp \left(\frac{i m (\bar{s} + nL)^2}{2\hbar (t_b - t_a)} \right) \exp \left(\frac{-i p_b \bar{s}}{\hbar} \right). \end{aligned}$$

Changing variables to $s = \bar{s} + nL$, the propagator is

$$\begin{aligned} \tilde{K}_L(p_b, t_b - t_a) &= 2\pi\hbar \delta(p_b - p_a) \left[\frac{2\pi i \hbar (t_b - t_a)}{m} \right]^{-\frac{1}{2}} \\ &\sum_{n=-\infty}^{\infty} e^{\frac{i p_b nL}{\hbar}} \int_{nL}^{(n+1)L} ds e^{\frac{-i p_b s}{\hbar}} \exp \frac{i m s^2}{2\hbar (t_b - t_a)} \end{aligned}$$

It is interesting to note that when $p_b L = N\hbar$, $N = 0, \pm 1, \pm 2, \dots$ then since $e^{\frac{-i p_b s}{\hbar}} = 1$,

$$\tilde{K}_L\left(\frac{N\hbar}{L}, t_b - t_a\right) = 2\pi\hbar \delta(p_b - p_a) \exp \left(\frac{-1}{\hbar} \frac{(N\hbar/L)^2}{2m} (t_b - t_a) \right)$$

is the free particle propagator in 1-dimension with energy $E = \frac{(N\hbar/L)^2}{2m}$.

Proof of Quantization

Now let us get the full propagator in the energy-momentum basis:

$$k(p_b E_b; p_a E_a) = (2\pi\hbar)^2 \delta(p_b - p_a) \delta(E_b - E_a) F(E_b, p_b)$$

where

$$\begin{aligned} F(E, p) &= \int_0^L ds e^{-ips/\hbar} \int_0^\infty dt \left(\frac{2\pi i \hbar t}{m} \right)^{-\frac{1}{2}} e^{\frac{iEt}{\hbar}} \\ &\times \sum_{n=-\infty}^{\infty} e^{\frac{im(s+nL)^2}{2\hbar t}} \\ &= \int_0^L ds e^{-ips/\hbar} \sum_{n=-\infty}^{\infty} \frac{m}{k+i\epsilon} \left[\theta(s+nL) e^{\frac{(s+nL)^2}{\hbar}(k+i\epsilon)} \right. \\ &\quad \left. + \theta(-s-nL) e^{\frac{-1}{\hbar}(s+nL)(k+i\epsilon)} \right] \end{aligned}$$

where $k = \sqrt{2mE}$. Since $s+nL > 0 \Rightarrow n \geq 0$ and $s+nL < 0 \Rightarrow n < 0$,

F breaks up into two terms $F_+ + F_-$ where

$$\begin{aligned} F_+(E, p) &= \int_0^L ds e^{-ips/\hbar} \left(\frac{m}{k+i\epsilon} \right) \sum_{n=0}^{\infty} e^{\frac{1}{\hbar}(s+nL)(k+i\epsilon)} \\ F_-(E, p) &= \int_0^L ds e^{-ips/\hbar} \frac{m}{k+i\epsilon} \sum_{n=1}^{\infty} e^{\frac{-1}{\hbar}(s-nL)(k+i\epsilon)} \end{aligned}$$

Evaluating F_+ first we have

$$F_+(E, p) = \frac{m}{k+i\epsilon} \int_0^L ds e^{\frac{-ips}{\hbar}} e^{\frac{1}{\hbar}s(k+i\epsilon)} \sum_{n=0}^{\infty} e^{\frac{inL(k+i\epsilon)}{\hbar}}$$

$$\begin{aligned}
&= \frac{m}{k+i\epsilon} \frac{\hbar}{1} \frac{1}{k-p+i\epsilon} \left[e^{\frac{iL}{\hbar}(k-p+i\epsilon)} - 1 \right] \frac{1}{1 - e^{\frac{iL}{\hbar}(k+i\epsilon)}} \\
&= \frac{m}{k+i\epsilon} \frac{i\hbar}{k-p+i\epsilon} e^{\frac{-iLp}{2\hbar}} \frac{\sin \frac{L}{2\hbar}(k-p)}{\sin \frac{L}{2\hbar}(k+i\epsilon)}
\end{aligned}$$

Next, evaluate $F_-(E,p)$:

$$\begin{aligned}
F_-(E,p) &= \frac{m}{k+i\epsilon} \int_0^L ds e^{\frac{-ips}{\hbar}} e^{\frac{-is}{\hbar}(k+i\epsilon)} \sum_{n=1}^{\infty} e^{\frac{1}{\hbar} nL(k+i\epsilon)} \\
&= \frac{m}{k+i\epsilon} \frac{\hbar}{-i} \frac{e^{\frac{iL}{\hbar}(k+i\epsilon)}}{k+p+i\epsilon} \frac{\left(e^{\frac{-iL}{\hbar}(k+p+i\epsilon)} - 1 \right)}{1 - e^{\frac{iL}{\hbar}(k+i\epsilon)}} \\
&= \frac{m}{k+i\epsilon} \frac{i\hbar}{k+p+i\epsilon} e^{\frac{-iLp}{2\hbar}} \frac{\sin \frac{L}{2\hbar}(k+p)}{\sin \frac{L}{2\hbar}k}
\end{aligned}$$

Thus, collecting terms,

$$F(E,p) = \frac{m}{k+i\epsilon} \frac{ine^{\frac{-iLp}{2\hbar}}}{\sin \frac{L}{2\hbar}(k+i\epsilon)} \left\{ \frac{\sin \frac{L}{2\hbar}(k-p+i\epsilon)}{k-p+i\epsilon} + \frac{\sin \frac{L}{2\hbar}(k+p+i\epsilon)}{k+p+i\epsilon} \right\}$$

The function $F(E,p)$ has singularities whenever

$$\frac{L(k+i\epsilon)}{2\hbar} = n\pi \quad n = 0, \pm 1, \pm 2, \dots$$

This means that states of definite momentum (in the momentum-time basis) propagate with a single energy only if $p_n = \frac{2\pi n\hbar}{L}$,

for here the propagator

$$F(E, p_n) = \frac{i\hbar}{E - \frac{p_n^2}{2m} + i\epsilon},$$

Otherwise a definite momentum state propagates as a superposition of many energy states. The argument is therefore that states having both a definite energy and momentum related by $E = p^2/2m$ must have a quantized value of $\frac{pL}{2\pi\hbar} = n$.

Following [V. N. Popov], we can arrive at the same result in a shorter way. We evaluate K_L as follows, where $s = \bar{s}$ and $t = t_b - t_a$:

$$\begin{aligned} K_L(s, t) &= \sum_{n=-\infty}^{\infty} \left(\frac{2\pi i \hbar t}{m} \right)^{-\frac{1}{2}} e^{\frac{im}{2\hbar t} (s+nL)^2} \\ &= \int \frac{dp}{2\pi\hbar} \sum_{n=-\infty}^{\infty} e^{\frac{ip(s+nL)}{\hbar} - \frac{it}{\hbar} \frac{p^2}{2m}} \\ &= \int \frac{dp}{2\pi\hbar} \left(\sum_{n=-\infty}^{\infty} e^{\frac{ipnL}{\hbar}} \right) e^{\frac{ips}{\hbar} - \frac{itp^2}{2m\hbar}}. \end{aligned}$$

Use the identity

$$\begin{aligned} \sum_{n=-\infty}^{\infty} e^{\frac{ipnL}{\hbar}} &= \sum_{n=-\infty}^{\infty} e^{i \left(\frac{pL}{2\pi\hbar} \right) 2\pi n} \\ &= \sum_{m=-\infty}^{\infty} \delta \left(\frac{pL}{2\pi\hbar} - m \right) \\ &= \frac{2\pi\hbar}{L} \sum_{m=-\infty}^{\infty} \delta \left(p - \frac{2\pi m \hbar}{L} \right) \end{aligned}$$

to obtain the result

$$K_L(s, t) = \frac{1}{L} \sum_{m=-\infty}^{\infty} e^{\frac{is}{\hbar} \frac{2\pi m \hbar}{L}} e^{\frac{-it}{\hbar 2m} \left(\frac{2\pi m \hbar}{L} \right)^2}.$$

This shows the only allowed states have quantized values of $\frac{pL}{2\pi\hbar}$. If the closed curve is a circle then $\frac{pL}{2\pi\hbar}$ is the angular momentum, and we have thus shown that the angular momentum is quantized.

Now we check that this method gives the same answer in the energy-momentum plane as the previous method:

$$\begin{aligned} F(E, p) &= \int_0^L ds e^{\frac{-isp}{\hbar}} \int_0^\infty dt e^{\frac{iEt}{\hbar}} K_L(s, t) \\ &= \frac{1}{L} \sum_{n=-\infty}^{\infty} \int_0^L ds e^{\frac{-is}{\hbar} \left(p - \frac{2\pi n \hbar}{L} \right)} \int_0^\infty dt e^{\frac{1}{\hbar} t \left(E - \frac{(2\pi n \hbar)^2}{2m} \right)} \\ &= \sum_{n=-\infty}^{\infty} \frac{\frac{i\hbar}{\left(\frac{(2\pi n \hbar)^2}{L} \right)^2}}{E - \frac{(2\pi n \hbar)^2}{2m} + i\epsilon} \frac{1}{\frac{Lp}{\hbar} - 2\pi n} \left(e^{\frac{-iLp}{\hbar}} - 1 \right). \end{aligned}$$

The series can be evaluated using the psi functions

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \frac{1}{\hat{p} - n} \frac{1}{\hat{k} - n} \frac{1}{\hat{k} + n} &= \frac{1}{\hat{p}\hat{k}^2} \\ &+ \frac{1}{2\hat{k}} \frac{1}{\hat{p} - \hat{k}} \left\{ \psi(1-\hat{k}) - \psi(1+\hat{k}) - \psi(1-\hat{p}) + \psi(1+\hat{p}) \right\} \\ &+ \frac{1}{2\hat{k}} \frac{1}{\hat{p} + \hat{k}} \left\{ \psi(1-\hat{k}) - \psi(1+\hat{k}) + \psi(1-\hat{p}) - \psi(1+\hat{p}) \right\} \end{aligned}$$

Then using the identity $\Psi(1-z) - \Psi(1+z) = -\frac{1}{z} + \pi \cot \pi z$ one derives

$$\sum_{n=-\infty}^{\infty} \frac{1}{\hat{p}-n} \frac{1}{\hat{k}-n} \frac{1}{\hat{k}+n} = \frac{\pi}{2\hat{k} \sin \pi \hat{k} \sin \pi \hat{p}} \left\{ \frac{\sin \pi(\hat{k}-\hat{p})}{\hat{k}-\hat{p}} + \frac{\sin \pi(\hat{k}+\hat{p})}{\hat{k}+\hat{p}} \right\}.$$

Using this result, one can then substitute in and obtain

$$F(E,p) = \frac{m}{k+i\epsilon} \frac{i\hbar e^{\frac{-iLp}{2\hbar}}}{\sin \frac{L(k+i\epsilon)}{2\hbar}} \left\{ \frac{\sin \frac{L}{2\hbar}(k-p+i\epsilon)}{k-p+i\epsilon} + \frac{\sin \frac{L}{2\hbar}(k+p+i\epsilon)}{k+p+i\epsilon} \right\}$$

where $k = \sqrt{2mE}$. This is identical to the propagator previously obtained.

Conclusion

The point of this exercise is to show in a very simple problem how topological properties manifest themselves. For a particle which goes around a closed loop of length L , the winding number n of the number of times the particle does a complete loop is a topological invariant of the space the particle travels in. It should be no surprise that since one has periodic boundary conditions, the momenta and energy are quantized.

We see, at least non-rigorously, that the underlying cause of this quantization is the topological structure of the space. At $t=0$, the wave function with a definite momentum is single valued as s goes around the loop. This means it is a continuous map from the loop into $U(1)$, the set of possible phases an amplitude can have. Topologically a loop is a circle S^1 , and the set

of phases $U(1)$ is also S^1 . The possible maps of S^1 into S^1 can be broken up into homotopy classes, i.e. classified by their winding number. Each momentum wave function $e^{\frac{ips}{\hbar}}$ must belong to one of these classes in order to be single valued, hence the condition $\frac{pL}{2\pi\hbar} = n$ is obtained.

Note that the topological considerations would be used also if the particle moving on a loop does so through a potential field.

There exists other examples of this same phenomena in field theory problems where the relevant topological spaces are the internal symmetry group and the space time manifold with certain boundary conditions imposed. One is invited to study the lectures of Coleman and of Popov for details.

C. Homotopy Classes and Fibre Bundles

1. Inequivalent vacua

There are various problems in physics which lead to fibre bundle structures. As our first example, suppose we have a charged particle constrained to move on a circular ring X .

The motion of the particle will depend upon the kind of external field $A_\mu(x)$ which is present. For our example we assume the field is static, and only A_i , $i=1,2,3$ are non-zero. In other words, only a magnetic field is possible.

The first possibility is the magnetic field is 0. One says that A_i is a "pure gauge"

$$A_i(x) = -i\lambda(x)^{-1}\partial_i\lambda(x)$$

where $\lambda(x)$ is a phase, an element of the internal symmetry group $U(1)$. The phase $\lambda(x)$ is a continuous, single valued function on the ring X , since we assume in this example that no magnetic fields are present. It is necessarily true that $A_i(x)$ is also a continuous (vector valued) function on X .

This simple example illustrates a topological property. We can ask how many types of functions λ map X into $U(1)$ continuously. We call two maps $\lambda_0(x)$ and $\lambda_1(x)$ homotopic (write $\lambda_1 \sim \lambda_0$) if we can find a continuous family of maps $\lambda_t(x) = h(t, x)$ such that $h: I \times X \rightarrow U(1)$ is continuous for $I = [0, 1]$, $h(0, x) = \lambda_0(x)$ and $h(1, x) = \lambda_1(x)$. The map h is called the homotopy, and we say $\lambda_t(x)$ is a continuous deformation of $\lambda_0(x)$ into $\lambda_1(x)$. This subject is treated in great detail in chapter VI.

We can ask to what extent do maps $\lambda(x)$ which are homotopic describe the same physical situation? The homotopy relation $\lambda_0 \sim \lambda_1$ is an equivalence relation, so that the maps $X \rightarrow U(1)$ are broken up into equivalence classes. When X is a circle, the set of equivalence classes is denoted $\pi_1(U(1))$. We now ask how many elements are there in $\pi_1(U(1))$.

We answer this by noting the topology of $U(1)$ is that of a circle, so $\pi_1(U(1))$ is the set of homotopically inequivalent maps of a circle onto itself. Maps with the same winding number are homotopic, so the elements of $\pi_1(U(1))$ can be thought of as the set of integers Z which specify the possible winding numbers.

We can understand the significance of the winding number by computing the flux, $\oint \underline{A} \cdot d\mathbf{l}$. Let $\lambda(x) = \exp i\phi(x)$, then $\underline{A} = \nabla\phi$ and $\oint \underline{A} \cdot d\mathbf{l} = \phi(x'_0) - \phi(x_0)$. Here x'_0 differs from x_0 by one complete rotation. In order for $\lambda(x) = \exp(i\phi(x))$ to be single valued, $\phi(x'_0) - \phi(x_0) = 2\pi n$ for some integer n : hence $\frac{1}{2\pi} \oint \underline{A} \cdot d\mathbf{l} = n$. Different values of n correspond to different amounts of magnetic flux trapped inside the ring. Maps $\lambda(x)$ which are homotopic represent the same flux inside the ring. The electron's wave

function depends only on the flux and hence only on the winding number n of $\lambda(x)$.

In fancy language, the electron on the ring sees topologically different vacua depending on the topological charge n . For all of these cases, we have united the internal space $U(1)$ with X as a product space. Call $B = X \times U(1)$ the bundle space. The continuous function $\lambda: X \rightarrow U(1)$ defines a cross section $f: X \rightarrow B$ as $f(x) = (x, \lambda(x)^{-1})$. The cross section f is continuous, and satisfies the property $pf(x) = x$, where p is the natural projection $p: B \rightarrow X$ defined by $p(x, y) = x$. The number of homotopy classes of cross sections is again $\pi_1(U(1))$. [We shall define all these concepts properly in subsequent lectures. The rest of the discussion on inequivalent vacua can be skipped until one has read chapters II-III].

We can do the same problem, but assume X is covered by two patches V_1 and V_2 . The two patches intersect in two disjoint open sets W_1 and W_2 . Again suppose there is a pure gauge in each patch:

$$iA_k^{(1)}(x) = \lambda_1(x)^{-1} \partial_k \lambda_1(x)$$

$$iA_k^{(2)}(x) = \lambda_2(x)^{-1} \partial_k \lambda_2(x) .$$

The wave function for the electron is

$$\psi^{(1)}(x) = \lambda_1(x)^{-1} \psi_0 \quad x \in V_1$$

$$\psi^{(2)}(x) = \lambda_2(x)^{-1} \psi_0 \quad x \in V_2 ,$$

in the two patches. The relationship between the two descriptions is a gauge transformation

$$g_{12}(x) = \lambda_1(x)^{-1} \lambda_2(x)$$

where

$$\psi^{(1)}(x) = g_{12}(x) \psi^{(2)}(x)$$

This construction (as we shall see) is that of a coordinate bundle with base X , fibre $U(1)$, group $U(1)$, coordinate neighborhoods V_1 and V_2 and coordinate transformation functions $g_{12}(x)$ and $g_{21}(x) = g_{12}(x)^{-1}$. The bundle construction theorem will assure us that there exists a bundle space B , projection p , and coordinate functions ϕ_j with the required properties to make a coordinate bundle. The coordinate transformation theorem will tell us that this coordinate bundle is equivalent to a product bundle.

We also note that we can define a cross section $f: X \rightarrow B$ such that $pf(x) = x$. Define

$$f(x) = \phi_1(x, \lambda_1(x)^{-1}) \quad x \in V_1$$

$$= \phi_2(x, \lambda_2(x)^{-1}) \quad x \in V_2$$

We can check that when $x \in V_1 \cap V_2$, the two definitions agree:

$$\begin{aligned} \phi_1(x, \lambda_1(x)^{-1}) &= \phi_2(x, g_{21}(x) \lambda_1(x)^{-1}) \\ &= \phi_2(x, \lambda_2(x)^{-1} \lambda_1(x) \lambda_1(x)^{-1}) \\ &= \phi_2(x, \lambda_2(x)^{-1}) \end{aligned}$$

There is no new physics added in this example by considering more patches. We were able to transform the representation using several patches into the representation having only one patch.

The one patch example above gives a bundle space which is a torus. The cross sections are functions which rotate n times around the fibre as one makes one revolution around the base. The two patch example gives a bundle space which is a torus twisted a multiple of 2π times before being pasted together. The cross sections are again functions which rotate n times around the fibre for one revolution around the base. The equivalence between the two examples is that we can undo the twist of the torus by a suitable redefinition of the cross section (cross section will change homotopy classes).

2. Magnetic monopoles

More interesting phenomena can occur if we go to a more complicated base space, and consider other than pure gauges. An example is [Wu and Yang's] treatment of magnetic monopoles. Let us recall that a monopole solution is described by a $\underline{B}$ field

$$\underline{B} = \frac{g}{r^2} \hat{e}_r$$

defined over all of space excluding the origin. It is topologically equivalent to a space X considered as the surface of a sphere of radius r . If the physics depends not on $\underline{B}$ but on the vector

potential, we need an $\underline{A}$ such that $\underline{\nabla} \times \underline{A} = \underline{B}$. In spherical coordinates, the curl of $\underline{A}$ is

$$(\underline{\nabla} \times \underline{A})_r = \frac{1}{r \sin \theta} \left(\frac{\partial \sin \theta A_\phi}{\partial \theta} - \frac{\partial A_\theta}{\partial \phi} \right)$$

$$(\underline{\nabla} \times \underline{A})_\theta = \frac{1}{r \sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{1}{r} \frac{\partial (r A_\phi)}{\partial r}$$

$$(\underline{\nabla} \times \underline{A})_\phi = \frac{1}{r} \left(\frac{\partial (r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right)$$

One choice for $\underline{A}$ is

$$A_\phi^I = g \frac{1 - \cos \theta}{r \sin \theta}, \quad A_\theta = A_r = 0.$$

Another equally good choice is

$$A_\phi^{II} = -g \frac{1 + \cos \theta}{r \sin \theta}, \quad A_\theta = A_r = 0.$$

We note that A_ϕ^I is singular only at $\theta = \pi$ and A_ϕ^{II} is singular only at $\theta = 0$. This suggests we describe the physics of the monopole in two overlapping patches V_I and V_{II} where V_I excludes $\theta = \pi$, V_{II} excludes $\theta = 0$, and $V_I \cap V_{II}$ is open and includes the region $\theta = \pi/2$.

This description is forced on us, once we go to the vector potentials, since we require the flux $\oint \underline{dr} \cdot \underline{A} = \int_S \underline{ds} \cdot \underline{B}$ to be correctly given by A . For A_ϕ^I , the flux is

$$\oint d\phi r \sin \theta A_\phi^I = 2\pi g(1 - \cos \theta)$$

which is the flux through the upper cap of the sphere. As θ approaches π , the total flux through the sphere approaches $4\pi g$. This is the expected value of

$$\oint \underline{B} \cdot \underline{ds} = 4\pi g$$

using Gauss' law. However, if we use A_ϕ^I to get the flux through the lower cap, we get

$$- \oint d\phi r \sin \theta A_\phi^I = -2\pi g(1 - \cos \theta)$$

which should go to zero as θ goes to π . Evidently it does not. The correct calculation is to use A_ϕ^{II} for the lower cap, in which case the flux is

$$- \oint d\phi r \sin \theta A_\phi^{II} = 2\pi g(1 + \cos \theta).$$

For any angle θ , the flux through upper and lower caps is $4\pi g$, the correct answer.

As discussed by [Wu and Yang], the description of a charge particle on the sphere is given by the path dependent phase factors

$$\phi_{PQ} = \exp i e \int_P^Q \underline{A} \cdot \underline{dr}.$$

Since we have two choices for $\underline{A}$, we get two descriptions when $\underline{P}$ and $\underline{Q}$ lie in $V_I \cap V_{II}$:

$$\phi_{PQ}^I = \exp(i e \int_P^Q \underline{dr} \cdot \underline{A}^I)$$

$$\phi_{PQ}^{II} = \exp(i e \int_P^Q \underline{dr} \cdot \underline{A}^{II}) .$$

Of course these two descriptions should describe the same physics when PQ describes a closed path in $V_I \cap V_{II}$, since the phase reflects the amount of flux flowing through the surface enclosed by the path. As an example, let PQ describe a path with constant θ from $\phi = 0$ to ϕ . The two phases are

$$\phi_{PQ}^I = \exp(i e g \phi (1 - \cos \theta))$$

$$\phi_{PQ}^{II} = \exp(-i e g \phi (1 + \cos \theta)) .$$

These two phases are related by a gauge transformation $g_{I,II}(\phi)$ defined by

$$\phi^I = g_{I,II}(\phi) \phi^{II}$$

$$g_{I,II}(\phi) = \exp(i 2 e g \phi) .$$

For the closed path $\phi = 0$ to $\phi = 2\pi$, the phases ϕ^I and ϕ^{II} should be equal, which constrains $g_{I,II}(\phi)$ to be single valued. This implies the quantization condition on g :

$$4 \pi e g = 2 \pi n$$

$$g = n/2e$$

More generally, we see $\bar{x}^I = g_{I,II} \bar{x}^{II}$ and $g_{I,II}(x)$ should be a continuous single valued function of $x \in V_I \cap V_{II}$.

The rest of this discussion can be skipped until one is familiar with the results of chapters II-III.

We shall learn that the above construction gives a non-trivial coordinate bundle with base space X (the sphere), fibre $U(1)$, group $U(1)$, coordinate neighborhoods V_I and V_{II} , and coordinate transformations $g_{I,II}(x)$. The bundle construction theorem will define up to equivalence a bundle space B , projection p and coordinate functions ϕ_j . As shown in [Steenrod], chapter 18, the number of inequivalent coordinate bundles is classified by the number of homotopy classes of maps $g: S^1 \rightarrow U(1)$ (recall S^1 is the circle and $U(1)$ is topologically a circle) which can be the restriction of a single valued coordinate transformation from $V_I \cap V_{II}$ to the circumference of the sphere X . That is, the set $\pi_1(U(1))$ tells us how many inequivalent coordinate bundles can be formed on X . This set is isomorphic to the integers Z , and we see that the quantization number n of the monopole labels the inequivalent bundle structures.

We note that for a pure gauge and one coordinate neighborhood covering X , the set of homotopy classes of cross sections $X \rightarrow X \times U(1)$ is $\pi_1(S^2) = 0$. (S^2 is the surface of a sphere). There is essentially only one vacuum gauge. If one uses two coordinate

neighborhoods V_I and V_{II} to X , the requirement that

$$\phi^I = \lambda_I(x)^{-1}$$

$$\phi^{II} = \lambda_{II}(x)^{-1}$$

be pure gauges means $\lambda_i(x)$ are continuous over V_i , for $i = I$ and II . The coordinate transformation

$$g_{I,II}(x) = \lambda_I(x)^{-1} \lambda_{II}(x)$$

has the form of a coordinate bundle which is equivalent to a product bundle. We can thus define a cross section

$$\begin{aligned} f(x) &= \phi_I(x, \lambda_I(x)^{-1}) & x \in V_I \\ &= \phi_{II}(x, \lambda_{II}(x)^{-1}) & x \in V_{II} \end{aligned}$$

and show as before that the two definitions agree whenever $x \in V_I \cap V_{II}$. Since the coordinate bundle B one obtains is equivalent to a product bundle B_0 , there exists a bundle map $h: B_0 \rightarrow B$ which induces the identity map on the base space X . For each cross section $f(x)$ on B there is an induced cross section on B_0 ,

$$f_0(x) = h_x^{-1} f(x).$$

It is not difficult to show that f_0 is unique, continuous, and a cross section on B_0 . Conversely, given any cross section on B_0

there is a unique cross section on B . Thus it is quite plausible that the cross sections one obtains for B with two patches are broken up into the same equivalence classes as those one obtains on B_0 with one patch (the product bundle). So it seems again true that equivalent coordinate bundles allow equivalent physical descriptions.

A second note on gauge theories on a sphere X concerns what happens if the internal symmetry group of the gauge transformations is not $U(1)$ but some other group G . Our arguments suggest that $\pi_1(G)$ classifies the "monopole" solutions. For $G = SU_2$, the group space is simply connected and $\pi_1(SU_2) \approx 0$ has one homotopy class. All maps are deformable into each other, and no monopole solutions exist. For $G = SO(3)$, $\pi_1(SO(3)) \approx \mathbb{Z}_2$ the integers modulo 2. There are two distinct homotopy classes, one of which is a monopole solution. For further reading consult [Coleman's] lectures, or various discussions in the literature (for example, [Tze and Ezawa]).

Bibliography

- S. Coleman, "Classical Lumps and their Quantum Descendants," Lectures at 1975 "Ettore Majorana" summer school, Harvard preprint 12/75.
- S. Coleman "Uses of Instantons", Harvard Preprint, HUTP-78/A004.
- W. Drechsler and M. E. Mayer, Fibre Bundle Techniques in Gauge Theories (Springer Verlag, Berlin, 1977).
- R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, 1965).
- G. 't Hooft, Nucl. Physics B79, 276 (1974).
- E. Lubkin, Annals of Physics 23, 233 (1963).
- A. M. Polyakov, JETP Lett. 20, 194 (1974).
- V. N. Popov "Functional Integrals in Quantum Field Theory" CERN Th 2424
- N. Steenrod, The Topology of Fibre Bundles (Princeton University Press, Princeton, N.J., 1951).
- A. Trautman, Rep. on Math. Physics 1, 29 (1970).
- H. C. Tze and Z. F. Ezawa, Phys. Rev. D14, 1006 (1976).
- T. T. Wu and C. N. Yang, Phys. Rev. D12, 3845 (1975).

Some Recent Instanton Applications

G. 't Hooft, Phys. Rev. 14, 3432 (1976).

P. Carruthers, S. S. Pinsky, and F. Zachariasen, CALTECH preprint CALT-68-598.

J. Hietarinta, W. F. Palmer and S. S. Pinsky, Ohio State Preprint C00-1545-227.

L. S. Brown, R. D. Carlitz, D. B. Creamer and C. Lee, Phys. Rev. 17, 1583 (1978).

C. G. Callan, R. Dashen, D. J. Gross, F. Wilczek and A. Zee, Princeton Preprint C00-2220-132.

L. Baulieu, J. Ellis, M. K. Gaillard, and W. J. Zakrzewski, CERN preprint Th.2482.

R. D. Carlitz and D. B. Creamer, Univ. of Pittsburg, Preprint PITT-199.

Bibliography

S. Coleman, "Classical Lumps and their Quantum Descendants," Lectures at 1975 "Ettore Majorana" summer school, Harvard preprint 12/75.

S. Coleman "Uses of Instantons", Harvard Preprint, HUTP-78/A004.

W. Drechsler and M. E. Mayer, Fibre Bundle Techniques in Gauge Theories (Springer Verlag, Berlin, 1977).

R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, 1965).

G. 't Hooft, Nucl. Physics B79, 276 (1974).

E. Lubkin, Annals of Physics 23, 233 (1963).

A. M. Polyakov, JETP Lett. 20, 194 (1974).

V. N. Popov "Functional Integrals in Quantum Field Theory" CERN Th 2424

N. Steenrod, The Topology of Fibre Bundles (Princeton University Press, Princeton, N.J., 1951).

A. Trautman, Rep. on Math. Physics 1, 29 (1970).

H. C. Tze and Z. F. Ezawa, Phys. Rev: D14, 1006 (1976).

T. T. Wu and C. N. Yang, Phys. Rev. D12, 3845 (1975).

Some Recent Instanton Applications

G. 't Hooft, Phys. Rev. 14, 3432 (1976).

P. Carruthers, S. S. Pinsky, and F. Zachariasen, CALTECH preprint CALT-68-598.

J. Hietarinta, W. F. Palmer and S. S. Pinsky, Ohio State Preprint C00-1545-227.

L. S. Brown, R. D. Carlitz, D. B. Creamer and C. Lee, Phys. Rev. 17, 1583 (1978).

C. G. Callan, R. Dashen, D. J. Gross, F. Wilczek and A. Zee, Princeton Preprint C00-2220-132.

L. Baulieu, J. Ellis, M. K. Gaillard, and W. J. Zakrzewski, CERN preprint Th.2482.

R. D. Carlitz and D. B. Creamer, Univ. of Pittsburg, Preprint PITT-199.

II. LANGUAGE OF TOPOLOGY

A. Topological Spaces

According to the Encyclopedia Britannica, "topology is a branch of mathematics that deals with selected properties of collections of related physical or abstract elements. It had its beginnings in the study of those properties that endure when the collection undergoes distortion but remains intact." It is the task of this introductory lecture to 1) indicate how the qualitative nature of topology rests on a precise mathematical framework, and 2) suggest that such reasoning can have profound implications.

Mathematicians deal with sets of objects whose exact nature need not be specified. What is of interest is the possible structures such sets of objects might have. We shall initially be interested in topological structures, though for later applications we shall be interested in sets having both topological and algebraic structures. The definition of a topological structure is deceptively simple, so after giving it, we shall spend some time seeing to what extent it coincides with any intuitive preconceptions we might have.

The following definition can be found in standard references, e.g. [Bourbaki]:

"A topological structure (or more briefly, a topology) on a set X is a structure given by a set $\mathcal{D}$ of subsets of X having the following properties (called axioms of topological structures):

- I. Every union of sets of $\mathcal{D}$ is a set of $\mathcal{D}$.
- II. Every finite intersection of sets of $\mathcal{D}$ is a set of $\mathcal{D}$.

The sets of $\mathcal{D}$ are called open sets of the topological structure defined by $\mathcal{D}$ on X A topological space is a set endowed with a topological structure." We write the union of A and B as $A \cup B$; intersection of A and B as $A \cap B$; union of a family of sets $\{A_\alpha\}_{\alpha \in F}$ as $\bigcup_{\alpha \in F} A_\alpha$ and intersection of that family as $\bigcap_{\alpha \in F} A_\alpha$.

It helps to visualize X as a subset of the two dimensional plane. Let $\mathcal{D}$ be the set of all subsets of the plane formed by the intersection of X with any circle in the plane. Obviously some circles do not meet (intersect) the subset X at all, so $\emptyset$, the empty set is in $\mathcal{D}$. Also there exist some circles which completely contain X , so X is also in $\mathcal{D}$. This will be generally true since the sets X and $\emptyset$ are meant to be included in $\mathcal{D}$ by axioms I and II. "Every" union is defined to include the union of the empty subset of $\mathcal{D}$, i.e. $\emptyset$; The intersection of the empty subset of $\mathcal{D}$ is defined to be the set X . These peculiar definitions are supposed to make the axioms easier to remember. If they don't help you, you can remember separately that $\emptyset$ and X must be open sets.

Fig. 1

Let us suppose the set X is that given by Fig.1 and A is a subset of X as drawn. If X were drawn on a rubber sheet and stretched and pulled, X and consequently A could be terribly distorted. Intuitively, the only distinctive property about A which remains is that A has an interior, a boundary and an exterior. Each of these concepts can be formulated in terms of the open sets. Note that circles lying entirely within A are necessarily in $\mathring{A}$. The interior $\mathring{A}$ of A is defined as those points x of A which lie in some open set V entirely contained in A . So any point x in $\mathring{A}$ has friends very close to it which are also in A , namely those in some (possibly tiny) circle about x .

Now those points which are exterior to A are just those points in the interior of the complementary set $X \setminus A$ (points of X not in A). Both $\mathring{A}$ and $\mathring{X \setminus A}$ are unions of open sets, and hence are both open, and their union is open. The boundary of A consists of those points of X which are neither in the interior nor the exterior of A ; the boundary is thus the complement of an open set.

Complements of open sets arise frequently and are called closed sets. One way of generating a closed set $\bar{A}$ from any given set A is to consider the complement of the exterior of A : $\bar{A} = X \setminus \mathring{X \setminus A}$. $\bar{A}$ is the closure of A . $\bar{A}$ is the smallest closed set containing A and $\mathring{A}$ is the largest open set contained in A .

B. Continuous Maps

Now what do we have so far. Well, if someone distorts our drawing in Fig.1, any subset of points interior to A stays interior; any points exterior to A stay exterior and any points on the

boundary stay on the boundary. If we want to be precise, we should replace the idea of distortion above by homeomorphism, a 1-1 mapping of X onto some other topological space Y such that open sets of X correspond to open sets of Y and vice versa. We say that two homeomorphic spaces have equivalent topological structures.

If we wanted only to make precise the intuitive notion of what property of a collection remains invariant under homeomorphism, we would be finished. Any collection of objects can be made into many possible topological spaces, depending on how the open sets are chosen. Homeomorphism divides the possible spaces into disjoint equivalence classes; elements in the same class then have the same topological attribute. It has also been found practical to consider transformations other than homeomorphisms.

Specifically, we want to express in topological language the idea of a continuous transformation. Intuitively we mean a function $f:X \rightarrow Y$ (read this as " f maps X into Y ") with the property that for each $x_0 \in X$ (read this as " x_0 is an element of X "), $f(x)$ is "as close as we wish" to $f(x_0)$ for x "sufficiently close" to x_0 . The mathematicians have made this notion precise by the use of the concept of a neighborhood. A neighborhood of x is a set V containing an open set which contains x and lies entirely in V . A property which holds for x "sufficiently close" to x_0 is defined as one which holds for every x in some neighborhood of x_0 . A property is valid for any y "as close as we wish" to y_0 if it is true for any y contained in every neighborhood of y_0 .

The definition of a continuous function $f:X \rightarrow Y$ at x_0 becomes: for every neighborhood V of $f(x_0)$ there exists a neighborhood U of

x_0 such that $x \in U$ implies $f(x) \in V$. Thus V can be a neighborhood of points "as near as we please" to $f(x_0)$, and U is a neighborhood of points "sufficiently close" to x_0 . Points $x \in U$ will then have their image $f(x)$ in V . The set of images $f(x)$ for $x \in U$ is written $f(U)$.

We can restate the definition slightly. Since $x \in U$ implies $f(x) \in V$, this is defined to be the statement $f(U) \subset V$ (read this as "every point of $f(U)$ is in V "), which again is defined as $U \subset f^{-1}(V)$ ($f^{-1}(V)$ is the set of points x such that $f(x)$ is in V). Since U is a neighborhood of x_0 , it contains an open set W containing x_0 . The open set W is necessarily in $f^{-1}(V)$, so we can also say $f^{-1}(V)$ is a neighborhood of x_0 . Combining these verbal manipulations, we get that f is continuous at x_0 if the inverse image of every neighborhood of $f(x_0)$ is a neighborhood of x_0 .

Open sets containing x_0 are necessarily neighborhoods of x_0 , and it turns out to be a useful question to ask what the inverse images of open sets are for continuous functions. Now an open set V is obviously a neighborhood of each and every one of its points. The inverse image $f^{-1}(V)$ is likewise a neighborhood of every one of its points. For if x is any element of $f^{-1}(V)$, then V is a neighborhood of $f(x)$. Since f is continuous we deduce $f^{-1}(V)$ is a neighborhood of x .

Let us ask what consequence follows from $f^{-1}(V)$ being a neighborhood of each of its points. Let U be the union of every open set in $f^{-1}(V)$, so by the axioms U is also an open set. Of course, by construction $U \subset f^{-1}(V)$. We want to show $f^{-1}(V) \subset U$, i.e. $U = f^{-1}(V)$. To do this, consider any point $x \in f^{-1}(V)$; $f^{-1}(V)$ is a neighborhood of x by assumption, so there exists an open set

containing x which lies in $f^{-1}(V)$. By definition of U , this open set lies also in U . Thus $x \in U$ and we obtain the important result that $f^{-1}(V)$ is open whenever V is open, assuming f is continuous. [N.B. if V is not in $f(X)$, then $f^{-1}(V)$ is empty and hence open].

The impressive fact is that the converse also holds: if $f^{-1}(V)$ is open whenever V is open, then f is continuous. We understand how this comes about as follows. Suppose $f: X \rightarrow Y$, and if W is any open set of Y , then $f^{-1}(W)$ is open. Consider any point $f(x_0) \in W$, and a neighborhood V [of points "as close as we wish"] to $f(x_0)$. There exists an open set W containing $f(x_0)$ with $W \subset V$. By assumption $U = f^{-1}(W)$ is open, and is thus a neighborhood [of points x "sufficiently close"] to x_0 such that $f(x) \in V$ whenever $x \in U$.

At a deep level then, we take a continuous function to be one such that inverse images of open sets are open. This shows in what sense a continuous function preserves the topology of X . For if $f: X \rightarrow Y$, X and Y topological spaces with open sets $\{U_\alpha\}$ and $\{V_\beta\}$ respectively, then f defines a new topology on X . This topology is $\{f^{-1}(V_\beta)\}$, and we can verify for any f (continuous or not) that unions and finite intersections of $\{f^{-1}(V_\beta)\}$ are again members of this topology. Continuous functions are special in that this new topology is contained in the given topology of X since each $f^{-1}(V_\beta)$ is open.

The final comment on this subject is that since homeomorphic maps $f: X \rightarrow Y$ have an inverse, and relate open sets of X to open sets of Y uniquely, both f and f^{-1} are continuous. The converse is also true.

C. Product Spaces, Quotient Spaces, and Subspaces

The most familiar examples of topological spaces are those from analysis and from geometry. We shall discuss some of these now, to give the formal concepts more substance.

The set of real numbers is denoted by R . The usual topological structure assigned to R is formed by taking open sets to be all the bounded open intervals. Functions which map R into itself are then continuous in the topological sense if and only if they are continuous in the usual sense of analysis.

Recall that the plane R^2 is formed by taking the direct product of the line with itself. The process of constructing a new space as the product of two given spaces occurs frequently. The product topology of the product $X \times Y$ of two topological spaces is built from the base of all open sets $U \times V$, where U is open in X and V is open in Y . The open sets of $X \times Y$ consist of all possible unions of sets in this base. For the plane with the product topology, the open sets are thus unions of open rectangles. We note in passing that the plane could be assigned the same topology if the open sets were unions of open circles [points x, y : $(x-x_0)^2 + (y-y_0)^2 < r_0^2$ for each fixed x_0, y_0, r_0]. There is, in general, nothing unique about the choice of the base of a topology except convenience.

We can think of each point of R^n as a function f from the integers $I_n = \{1, \dots, n\}$ to the real numbers R . The values $(f(1), \dots, f(n))$ are the usual coordinates of R^n . The base of the (usual) topology of R^n consists of sets $U_1 \times \dots \times U_n$ where U_k are open in R for $k = 1, \dots, n$. The idea of a product space can be

extended to an infinite index set I by letting R^I be the set of all functions $f: I \rightarrow R$. The product topology of R^I is formed from the base of sets of infinite products $\prod_{t \in I} U_t$ such that U_t is open in R for each $t \in I$ and $U_t = R$ for all but a finite number of indices $t \in I$. An element in the base consists of all functions $f: I \rightarrow R$ subject to the restriction that for a finite number of points $t \in I$, U_t is not R and $f(t) \in U_t$.

The consideration of product sets allows one to rapidly build quite sophisticated topological spaces. One can also build new spaces from an essentially inverse operation, that of taking quotients. Roughly speaking, one constructs a new space by identifying points of the old space as equivalent. Sometimes such spaces are called identification spaces. Recall that the quotient of two numbers x and y is really a reduction of pairs of numbers (u, v) into equivalence classes:

all values of (u, v) such that $xv = uy$ are identified with the quotient class x/y . More generally, the quotient set X/R of a set X is constructed relative to an equivalence relation R which identifies points of $X: x \equiv y \pmod{R}$ is the way we will write that x is identified with y modulo the equivalence R . This way of writing emphasizes that equivalence relations are reflexive $[x \equiv x \pmod{R}]$, symmetric $[x \equiv y \pmod{R} \Rightarrow y \equiv x \pmod{R}]$ and transitive $[x \equiv y \pmod{R} \text{ and } y \equiv z \pmod{R} \Rightarrow x \equiv z \pmod{R}]$. Because of these

properties, R breaks X up into disjoint (equivalence) classes. These disjoint classes form the points of the quotient set X/R .

If X is a topological space, there is a natural topology to assign to X/R , so that points near identified points of X will be nearby in the quotient space. Let $p: X \rightarrow X/R$ be the natural map, defined by letting $p(x)$ be the class of points identified with x module R . The open sets U of X/R are those such that $p^{-1}(U)$ is open in X . With this definition the natural map is a continuous map.

As a simple example, consider the points of the real line which are identified by translation by an integer: $x \equiv y \pmod{R}$ if and only if (iff) $x-y$ is an integer. The quotient space so constructed is homeomorphic to a circle. The torus can be similarly constructed from the relation $(x,y) \equiv (u,v) \pmod{R}$ iff $x-u$ any $y-v$ are integers. The torus is homeomorphic to R^2/R . [q.v. Fig.2].

Fig.2 A square with edges identified is a torus.

A more sophisticated example is the sphere S^4 in 4-dimensions with antipodal points identified. The sphere S^4 is the topology of the transformation group SU_2 , and the quotient space formed by identifying opposite points is homeomorphic to $O(3)$, the group of orthogonal transformations in 3-dimensions.

Another general construction of new topological spaces from old is that of a subspace. Given a topological space X , and a subset $A \subset X$, the natural topology one assigns to A consists of taking the open sets to be all subsets of A formed from the intersection of each open set U of X with A , $U \cap A$. With this topology, the inclusion map $i: A \rightarrow X$ defined as $i(x) = x$ for each $x \in A$, is continuous.

D. Connectivity

Lots of examples come to mind if one considers subspaces of R^2 . One can remove open or closed circles at different points, getting spaces looking like swiss cheese. One can instead consider the plugs one has removed, obtaining disconnected pieces. The kinds of connectivity of these spaces are topological invariants, i.e. they remain under homeomorphic transformations. As such, they are important concepts to keep in mind. The examples

Fig. 5 Subspace of R^2 formed by removing two discs.

Fig. 4 The subspace of two discs.

illustrated above display two different kinds of connectivity. The subspace X of two discs can be written as the disjoint union of two sets open in X . In general, a space X is called connected as long as one cannot write X as the union of disjoint open sets.

The subspace Y formed by removing two discs is obviously connected. However, one cannot remove the holes by any homeomorphic mapping. One way to find such holes is by considering closed curves in Y . A closed curve is a continuous map $g: I = [0, 1] \rightarrow Y$ such that $g(0) = g(1) = y_0 \in Y$. The interval I with ends identified

is a circle S^1 . As seen in the figure, there are essentially three kinds of closed curves: those which circle neither hole, those which circle one, and those which circle the other. To make precise the equivalence of each of these kinds of curves, one uses the notion of homotopy (discussed at great length in chapter VII). A closed curve $g_0:S^1 \rightarrow Y$ is homotopic to the closed curve $g_1:S^1 \rightarrow Y$ if there is a continuous map $h:S^1 \times I \rightarrow Y$ such that $h(x,0) = g_0(x)$ and $h(x,1) = g_1(x)$. Intuitively, we see that h deforms the curve g_0 continuously into g_1 . The relation " g_0 is homotopic to g_1 " is an equivalence relation on the set of continuous maps of $S^1 \rightarrow Y$. The set of equivalence classes $\pi_1(Y)$ is a topological invariant of the space Y . For our example, $\pi_1(Y)$ has three elements, two classes other than the identity class (maps homotopic to the constant map $i:S^1 \rightarrow y_0$). Spaces in which $\pi_1(Y)$ has only the identity class are called simply connected. The Lie group SU_2 considered as a topological space is connected and simply connected; the Lie group $O(3)$ (considered as a topological space) is connected, but not simply connected.

E. Pasting

There is one more standard way of constructing new spaces from old, what [Bourbaki] calls "pasting together of topological spaces." This construction is used to construct gauge theories, fibre bundles and differentiable manifolds. As the name implies, what is involved is like pasting several scraps of paper together.

Imagine that we have been given a family $\{X_i\}_{i \in I}$ of topological spaces, or patches. To paste these patches together, we need a set of coordinate transformations g_{ij} which identify points of X_i with points of X_j along the region of pasting. Let A_{ij} be the set of points of X_i which are to be pasted with points of X_j , for each pair of indices $i, j \in I$. We will always paste X_i with itself by the identity map, $g_{ii}(x) = x$ for $x \in X_i$. The other assumption one makes is that when considering three patches with $i, j, k \in I$ pasting X_j with X_k , followed by pasting X_i is the same as pasting X_i with X_k directly. In terms of the coordinate transformations, when $x \in A_{ij} \cap A_{ik}$, then $g_{ik}(x) = g_{ij}(g_{jk}(x))$.

It is now an exercise in topological reasoning to show how to construct a topological space from the pasted patches. The first step is to consider the sum of spaces

Fig.5. Patches i and j

$\{X_i\}$ thought of as disjoint patches. The technical device is to construct the product space $X_i \times \{i\}$ for each $i \in I$. Note $\{i\}$ is the

topological space with one element, and hence the open sets are $\emptyset$ and $\{i\}$. The union of these sets $X = \bigcup_{i \in I} X_i \times \{i\}$ is a disjoint union of sets $X_i \times \{i\}$. The topology of X is defined by taking as open sets, those whose intersection with each $X_i \times \{i\}$ is open in $X_i \times \{i\}$. In particular, since $X_i \times \{i\}$ is itself open, X is the disjoint union of open sets.

Now we use the coordinate transformations g_{ij} to identify points of X . The resultant quotient space is the space formed by pasting together the patches. Define $(x, i) \equiv (y, j) \pmod{R}$ iff $x \in A_{ij}$ and $y = g_{ij}(x)$. The requirements on g_{ij} insure R is an equivalence relation: Reflexive since $x = g_{ii}(x)$, $x \in X_i \Rightarrow (x, i) \equiv (x, i)$; Symmetric because of the identity $g_{ji} = g_{ij}^{-1}$, transitive because of the transitivity of the g_{ij} . The quotient set X/R which results from pasting together all the patches can be assigned the quotient topology.

Let ϕ be the canonical (natural) map of $X \rightarrow X/R$; that is, $\phi(x)$ is the equivalence class of points of X which are identified with x . If ϕ is restricted to $X_i \times \{i\}$, then ϕ is a 1-1 map of $X_i \times \{i\} \rightarrow X/R$. If we assume that each A_{ij} is open in X_i , and that the coordinate transformations are homeomorphisms, then ϕ maps $X_i \times \{i\}$ homeomorphically onto $\phi(X_i \times \{i\})$. Thus each patch has a homeomorphic image in X/R .

F. Compactness

An important attribute of a topological space concerns the value taken by an infinite intersection of closed sets. Suppose $\{F_\alpha\}_{\alpha \in F}$ is a family of closed subsets of X . It has the finite

intersection property if the intersection of every finite subfamily of F is non-empty. The space X is called compact if the intersection $\bigcap_{\alpha \in F} F_\alpha$ is non-empty for any family F of closed subsets with the above finite intersection property. For example consider the closed subsets of the plane $\mathbb{R}^2$ formed as the complement of open circles about the origin. This family has the finite intersection property, but the infinite intersection is empty since the plane does not have a point at infinity. The plane is not compact. The complex plane with ∞ added is compact, and hence so is the sphere since the stereographic projection maps the complex plane homeomorphically onto the sphere.

The above statement of compactness can be converted to a statement about open sets using the duality between open and closed sets. Rephrasing the above statement, every family F with the finite intersection property has non-zero intersection. Equivalently, if F has zero intersection, then there exists some finite subfamily with zero intersection. A family $\{F_\alpha\}$ with zero intersection implies the set $\{U_\alpha\}_{\alpha \in F}$ formed by the complements $U_\alpha = X \setminus F_\alpha$, $\alpha \in F$ is an open covering of X : $\bigcup_{\alpha \in F} U_\alpha = X$ and $\{U_\alpha\}$ are open sets. Thus compactness means every open covering of X has a finite subcovering.

Some authors use compactness only if the space X satisfies the additional property that any two distinct points of X have disjoint neighborhoods. A space with this separation property is called Hausdorff, and is an attribute of most spaces one deals with. In these lectures we will always make the assumption that compact spaces are Hausdorff.

Compactness and connectedness are properties transmitted by continuous functions: the continuous image of a compact (connected) space is compact (connected). Products of compact (connected) spaces are again compact (connected). Quotients of connected spaces are again connected. However, quotients of Hausdorff spaces are not necessarily Hausdorff, nor are quotients of compact spaces necessarily compact. Other more sophisticated properties and examples can be found in the textbooks.

G. Global v.s. Local

In addition to the global concepts of compactness and connectedness, there are local concepts which are topological properties. A space X is locally connected if for each $x \in X$, every neighborhood of x contains a connected neighborhood. The space X is locally compact if X is Hausdorff and if every $x \in X$ has a compact neighborhood. It is a theorem that every neighborhood of x will then have a compact neighborhood. The real line R is connected, locally connected, locally compact but not compact.

A locally finite family of sets $\{A_i\}_{i \in I}$ is intuitively one which, for each $x \in X$, has only a finite number of sets A_i which intersect points sufficiently close to x . This can be rephrased as the requirement for every $x \in X$, points sufficiently close to x fail to intersect A_i for all but a finite number of indices $i \in I$. In terms of neighborhoods, the definition is that $\forall x \in X$, ("for every x in X ") $\exists$ ("there exists") a neighborhood V of x so that $V \cap A_i = \emptyset$ for all but a finite number of indices $i \in I$. The last

formulation is the translation into topological language of the intuitive concept of locally finite families.

The concept of locally finite families is used in the definition of paracompact spaces: X is paracompact if it is Hausdorff and satisfies the axiom: every open covering of X has a refinement (collection of open sets formed from finite intersections of the given open covering) which is locally finite. In general relativity, the manifolds one uses are assumed to be paracompact. More is said on this subject in modern treatments of differential geometry and general relativity, such as [S. W. Hawking and G. F. R. Ellis].

Bibliography

N. Bourbaki, Elements of Mathematics, General Topology (Addison-Wesley Publishing Company, Inc. Reading Mass. 1966), vol.I.

S. W. Hawking and G. F. R. Ellis, The Large Scale Structure of Space Time, (Cambridge University Press, Cambridge 1973).

J. G. Hocking and G. S. Young, Topology, (Addison-Wesley Publishing Company, Inc., Reading Mass., 1961).

J. L. Kelley, General Topology, (D. Van Nostrand Company Inc. New York, 1955).

L. S. Pontryagin, Topological Groups second edition (Gordon and Breach, Science Publishers, Inc. New York 1966).

III. FIBRE BUNDLES

For various reasons, we have decided to learn about fibre bundles. Perhaps the strongest motivation is that a fibre bundle is the mathematical structure we call a gauge theory, and such theories are current candidates for a unified view of all interactions in physics. Such theories are necessarily complex, and we hope to benefit from the knowledge accumulated by the mathematicians.

In chapter II, I gave the elementary definitions and properties of topological spaces. In this chapter, we study a richer structure, whose elements are topological spaces, continuous maps and topological groups. This richness of structure makes it possible to connect fibre bundles to gauge theories.

A. Informal Discussion

Before diving into the mathematics, I would first like to give an intuitive idea of fibre bundles, due to [A. Trautman]. Let T denote time and X denote ordinary space. We can think of mechanics as a theory which tries to unify these two spaces. One way is the theory of statics. The position of a particle in space is independent of time. Time and space form a trivial 4-dimensional space $T \times X$, the direct product space. For Galilean theories, the major change is that we allow motion, a continuous map x of time T into space X . The path in $T \times X$ is the map $t \rightarrow (t, x(t))$.

When we go to special relativity, space and time are united

in a more complicated fashion. The motion of a particle, then, is also a more complicated expression. It is enough to note at this time that gauge theories unite space-time with internal symmetry spaces. It is plausible that the study of methods which unite two spaces will be useful. Let us give somewhat more elementary examples of uniting two spaces. We will always call one of the spaces X the base, and the other space Y the fibre. We picture the unified space as one with Cartesian coordinates X and Y (Figure 1).

Fig.1.

The unified space we call the bundle space B . The sense in which X is a coordinate axis of B is that we require a projection $p:B \rightarrow X$ which is continuous. Clearly all of these concepts are true for product spaces $[B = X \times Y]$.

Let us consider a cylinder. This is a direct product space of a circle X with a line segment Y . The circle can be thought of as the unit interval I with the ends identified. The fibre can be taken

Fig.2.

as the unit interval. The cylinder can be obtained by making first the square, then identifying the left and right edges of the figure.

We obtain a different space by doing the same construction for the cylinder, except for the last step. Suppose we identify edges by introducing a single twist: " 1 " is identified with " 0 " and " 0 " with " 1 ". This bundle space is not a cylinder, but a Möbius strip. The Möbius strip is not a product space, yet the base space and fibre space are the same as those for the cylinder which is a product space.

B. Precise Definition of Fibre Bundle

Now that we know a couple of examples, we take up the study of fibre bundles in the sense of [Steenrod]. He gives a very good treatment, and his book is highly recommended for study, even though it is somewhat difficult. I hope these lectures will make his book more accessible.

To avoid ambiguity, one introduces several terms to distinguish similar, but not identical structures. Steenrod uses coordinate bundle B to mean the following structure of spaces (all spaces will be topological spaces) and maps (all maps will be continuous functions):

Coordinate bundle

- (1) A space B called the bundle space.
- (2) A space X called the base space.
- (3) A continuous map $p: B \rightarrow X$ called the projection.
- (4) A space Y called the fibre.
- (5) A space G , whose elements form a group with group product $\alpha: G \times G \rightarrow G$ ($\alpha(g_1, g_2) \in G$) such that α is continuous and $g \mapsto g^{-1}$ defines a continuous map. A space G with this group structure is a topological group. The product $\alpha(g_1, g_2)$ is written $g_1 g_2$. In addition G is required to act continuously on the fibre Y : there must be a continuous function $\eta: G \times Y \rightarrow Y$ such that $\eta(e, y) = y$ (e is group identity) for every $y \in Y$, and $\eta(g_1 g_2, y) = \eta(g_1, \eta(g_2, y))$ for every $g_1, g_2 \in G$ and $y \in Y$. When this is true, G is also said to be a topological transformation group of Y relative to the map η . When there is no confusion, $\eta(g, y)$ is written as $g \cdot y$. This product satisfies $e \cdot y = y$ and $(g_1 g_2) \cdot y = g_1 \cdot (g_2 \cdot y)$. Finally, G is required to act effectively on Y : if $g \cdot y = y$ for every y then $g = e$. That is, the only

element in G which leaves every element of Y invariant is the identity. With certain technical complications, this can always be made to be true.

- (6) An open covering $\{V_j\}_{j \in J}$ of X , the elements V_j of the covering are called coordinate neighborhoods and J is the index set. The open sets $\{V_j\}$ cover X in the sense that $\bigcup_{j \in J} V_j = X$.
- (7) $\forall j \in J$, coordinate functions $\phi_j: V_j \times Y \rightarrow p^{-1}(V_j)$ such that ϕ_j are homeomorphisms. These show that locally, B is homeomorphic to a product space. [Recall that the notation $p^{-1}(V_j)$ represents the set of all points in B whose images under p lie in V_j .]
- (8) The coordinate function must satisfy $p\phi_j(x,y) = x$, $x \in V_j$.
- (9) The coordinate function defines $\phi_{j,x}: Y \rightarrow p^{-1}(x)$, $x \in V_j$ by $\phi_{j,x}(y) = \phi_j(x,y)$. It is required that $\phi_{j,x}^{-1}\phi_{i,x}: Y \rightarrow Y$ for $x \in V_i \cap V_j$ act on Y like an element of G .
- (10) Defining $g_{ji}(x) = \phi_{j,x}^{-1}\phi_{i,x}$, $g_{ji}: V_i \cap V_j \rightarrow G$ is required to be continuous.

Exercises in the use of the definitions

These definitions provide a unification of two spaces X and Y where the fibres over x , $p^{-1}(x)$ are related to each other in a continuous way by means of an effective, continuous group G which acts on Y . To learn more about this structure, let us investigate further the picture we have of B , as a space with coordinate axes X and Y .

Fig. 3.

Even though B is not a product space of X and Y , locally it is a product $V_j \times Y$. The point (x, y) of our figure corresponds to $b = \phi_j(x, y)$. Conversely, if we are given any $b \in B$, then $x = p(b)$ must be in at least one neighborhood V_k since they cover X . To complete the picture, we must find the $y \in Y$ such that $\phi_k(x, y) = b$. Since ϕ_k is a homeomorphism, such a y exists:

$$\phi_{k,x}(y) = \phi_k(x, y) = b$$

$$\Rightarrow y = \phi_{k,x}^{-1}(b) .$$

It is helpful to give a name p_k to $\phi_{k,x}^{-1}$ since it projects out the y component of b ; the subscript reminds us this projection is only good locally ($x \in V_k$).

A key role in the theory of these coordinate bundles is played by the coordinate transformations g_{ij} [recall $g_{ij}(x) = \phi_{i,x}^{-1} \phi_{j,x}$, $x \in V_i \cap V_j$]. In a gauge theory, these are the gauge transformation. (See [Wu and Yang's] treatment; they give a

dictionary between gauge theories and fibre bundles, and give also some simple examples.) The coordinate transformations satisfy a number of easily derived properties. Later we will see that when we are given functions with some of these properties, we can show they serve as a way to define coordinate bundles. The basic property is that for $x \in V_i \cap V_j \cap V_k$,

$$g_{ij}(x)g_{jk}(x) = g_{ik}(x) .$$

This is evident from the definition of g_{ij} ,

$$\phi_{i,x}^{-1} \phi_{j,x} \phi_{j,x}^{-1} \phi_{k,x} = \phi_{i,x}^{-1} \phi_{k,x} .$$

Because of this property $g_{ii}(x)g_{ii}(x) = g_{ii}(x)$, which implies from group theory that $g_{ii}(x) = e, \forall x \in V_i$. Because of $g_{ij}(x)g_{ji}(x) = g_{ii}(x) = e$, then $g_{ij}(x)^{-1} = g_{ji}(x), \forall x \in V_i \cap V_j$.

The coordinate transformations, the projection p and the y -projector p_k are intimately related. We already noted that p gives the x -coordinate and p_k gives the y coordinate:

$$p\phi_j(x,y) = x$$

$$p_j\phi_j(x,y) = y .$$

Another way to say this is for any $b \in B$, $p(b) = x$ and $p_k(b) = y$ for $x \in V_k$, so that

$$b = \phi_j(p(b), p_k(b)) .$$

The coordinate transformation comes in when $p(b)$ is in two neighborhoods V_j and V_k . There are now two y coordinates, $p_j(b)$ and $p_k(b)$. For our picture to be consistent, these two y coordinates must be related. This is easy to show:

$$\begin{aligned} p_j(b) &= \phi_{j,x}^{-1}(b) = \phi_{j,x}^{-1} \phi_{k,x} \phi_{k,x}^{-1}(b) \\ &= g_{jk}(x) \cdot p_k(b) \end{aligned}$$

where $x = p(b) \in V_j \cap V_k$. Thus two points (x, y) and (x', y') correspond to the same point b in the bundle space iff

$$x' = x$$

$$y' = g_{jk}(x) \cdot y$$

where $x \in V_k$, $x' \in V_j$ ($x = x' \Rightarrow x \in V_j \cap V_k$).

Equivalences and Fibre Bundles

Suppose we have defined one coordinate bundle B with elements $\{B, X, p, Y, G, \phi_j, V_j, J\}$. This notation of a coordinate bundle is somewhat cumbersome, but it serves to remind us of all the elements which make up the bundle structure. Namely, there is a bundle space B , base space X , fibre space Y , projection $p: B \rightarrow X$ and group G which acts effectively on Y . In addition, there are the coordinate neighborhoods V_j , and coordinate functions ϕ_j indexed over the set J which serve to make precise the idea that the bundle space is locally a product space. What we will now consider is two different local structures; we would like to know when they are equivalent. So we consider another coordinate

bundle $\{B, X, p, Y, G, \phi_j', V_j', J'\}$ which differs from the first in the local structure. Steenrod gives a natural requirement that the two coordinate bundles are strictly equivalent iff the union of the two sets of coordinate functions is a coordinate function of a bundle. That is, form the bundle from the elements $\{B, X, p, Y, G\}$ and use $\{V_j\}_{j \in J} \cup \{V_{j'}'\}_{j' \in J'}$ as the open covering of X , and the ϕ_j, ϕ_j' , defined on $V_j \times Y$ and $V_{j'}' \times Y$ respectively. For these coordinate functions to be elements of a coordinate bundle, they must satisfy

$$g_{jk}(x) = \phi_{j,x}^{-1} \phi_{k,x} \in G \quad x \in V_j \cap V_k$$

$$g_{j',k'}'(x) = \phi_{j',x}'^{-1} \phi_{k',x}' \in G \quad x \in V_{j'}' \cap V_{k'}'$$

$$\bar{g}_{j',k}(x) = \phi_{j',x}'^{-1} \phi_{k,x} \in G \quad x \in V_{j'}' \cap V_k$$

Only the last requirement is new, and thus provides the main restriction. Also we must have

$$\bar{g}_{j',k}: V_{j'}' \cap V_k \rightarrow G$$

a continuous function; the continuity of g_{ij} and $g_{i',j}'$, we have by assumption. To sum up, the two coordinate bundles are strictly equivalent if $\bar{g}_{j',k}(x) = \phi_{j',x}'^{-1} \phi_{k,x}$ defines a continuous map of $V_{j'}' \cap V_k$ into an element coinciding with an operation of G .

The notion of strict equivalence allows one to introduce the notion of a fibre bundle. If we are given a coordinate bundle B , the set of all coordinate bundles strictly equivalent to B

is a fibre bundle. B is called a fibre bundle with coordinates; the fibre bundle is independent of coordinates. We almost never use fibre bundles, but instead their representative coordinate bundles.

C. Bundle Maps

The understanding of fibre bundles (i.e. their representative coordinate bundles) is greatly enhanced if one can compare two such structures via a map which preserves some of the bundle structure. To this end, one introduces the notion of a bundle map. For coordinate bundles with the same fibre and group, Steenrod defines a bundle map $h: B \rightarrow B'$ as follows:

- (1) $h: B \rightarrow B'$ is a continuous map of the bundle spaces.
- (2) h is fibre preserving.

By fibre preserving it is meant that a fibre $Y_x = p^{-1}(x)$ over x is mapped homeomorphically onto some fibre $Y_{x'} = p^{-1}(x')$ over x' . [This evidently induces a function $\bar{h}: X \rightarrow X'$ since for each $x \in X$, Y_x is mapped onto a unique fibre so a unique x' is defined. By definition, $h(b) = b'$, $\bar{h}p(b) = p'h(b)$ and so the above figure is commutative:

$$\bar{h}p = p'h.$$

It is also easy to show $\bar{h}$ is continuous. According to Chapter II we need to show inverse images of open sets are open. Suppose

U' is any open set in X' . Since both p' and h are continuous, $h^{-1}p'^{-1}(U')$ is open. By the commutativity of the above figure $p^{-1}\bar{h}^{-1}(U') = h^{-1}p'^{-1}(U')$ so

$p^{-1}\bar{h}^{-1}(U')$ is also open (and a subset of B). The finite intersection of open sets is open so $p^{-1}\bar{h}^{-1}(U') \cap p^{-1}(V_j)$ is open for each $j \in J \Rightarrow p^{-1}(V_j \cap \bar{h}^{-1}(U'))$ is open for each $j \in J$. Using the homeomorphism between $p^{-1}(V_j)$ and $V_j \times Y$ we conclude $V_j \cap \bar{h}^{-1}(U')$ is open for each $j \in J$. The union of all such open sets $\bar{h}^{-1}(U')$ is open by the topology axiom and so $\bar{h}$ is indeed continuous.]

- (3) the bundle structure $\bar{h}$ induces on B, p, X, Y, G is a coordinate bundle strictly equivalent to B . [We note that $\bar{h}$ being continuous defines an open covering $\{\bar{h}^{-1}(V'_j)\}_{j \in J'}$ using the coordinate neighborhoods V'_j of B' . Denoting h_x as the map h restricted to Y_x , h_x is a homeomorphism. We can define a coordinate function on $\bar{h}^{-1}(V'_j) \times Y$ as

$$\psi_j(x, y) = h_x^{-1} \phi'_j(x', y)$$

$$x' = \bar{h}(x) \in V'_j.$$

We verify that $\psi_j(x, y)$ is a coordinate function by noting

$$a) \quad p\Psi_j(x,y) = ph_x^{-1}\phi'_j(x',y) = x.$$

This result follows since $\phi'_j(x',y)$ is an element in b' whose inverse image corresponds uniquely to some point $b \in p^{-1}(x)$, so $p(b) = x$.

$$b) \quad \text{Using } \Psi_{j,x} = h_x^{-1}\phi'_{j,x}, \text{ for } x \in \bar{h}^{-1}(V'_j) \cap \bar{h}^{-1}(V'_k) \\ \text{then } \Psi_{j,x}^{-1}\Psi_{k,x} = \phi'_{j,x}{}^{-1}h_x h_x^{-1}\phi'_{k,x} = g'_{jk}(x') = g'(\bar{h}(x)) \\ \text{coincides with an element of } G \text{ (since } g'_{jk}(x') \\ \text{does by assumption) and}$$

$$c) \quad x \rightarrow \Psi_{j,x}^{-1}\Psi_{k,x} \text{ is continuous since it is the composition of } \bar{h} \text{ and } g'_{jk}, \text{ two continuous functions.}$$

$$d) \quad \Psi_j \text{ defines a homeomorphism. First, } \Psi_j \text{ is continuous in } x,y \text{ (since } \phi'_j \text{ is), and } h_x \text{ is a homeomorphism; because } p\Psi_j(x,y) = x, \Psi_j: \bar{h}^{-1}(V'_j) \times Y \rightarrow p^{-1}(\bar{h}^{-1}(V'_j)). \text{ To show a continuous map from } p^{-1}(\bar{h}^{-1}(V'_j)) \text{ into } \bar{h}^{-1}(V'_j) \times Y, \text{ define}$$

$$\eta_j = \Psi_{j,x}^{-1} \\ = p'_j h_x.$$

$$\text{This is continuous and } \eta_j \Psi_j(x,y) = p'_j \phi'_j(x',y) = y.$$

$$\text{Thus } b \mapsto (p(b), \eta_j(b)) \text{ is the continuous map}$$

$$\text{which is inverse to } \Psi_j. \text{ The elements } \{B, X, p, Y, G, \Psi_j, \bar{h}^{-1}(V'_j)\} \text{ form a coordinate bundle}$$

$$\text{which we have called here the } \underline{\text{induced bundle.}}$$

To verify the induced bundle is strictly equivalent to

$\mathcal{B}$ we need to check that

$$\begin{aligned}\bar{g}_{kj}(x) &= \psi_{k,x}^{-1} \phi_{j,x} \\ &= \phi_{k,x}'^{-1} h_x \phi_{j,x}\end{aligned}$$

coincides with an element of G , and

$$\bar{g}_{kj} : \bar{h}^{-1}(V'_k) \cap V_j \rightarrow G$$

is continuous.

The bundle map allows one to compare two coordinate bundle structures. One can use such maps to introduce a second notion of equivalence. Steenrod considers two coordinate bundles B and B' with the same fibre, group and base. They are called equivalent if there exists a bundle map $h: B \rightarrow B'$ such that the induced map $\bar{h}: X \rightarrow X$ is the identity. Fibre bundles are equivalent if representative coordinate bundles of each are equivalent.

2. Bundle Transformation Theorem

The other important role bundle mappings play is to switch the emphasis of coordinate bundles from the consideration of bundle spaces and fibres to the coordinate transformations (gauge transformations). The main result we want to obtain (bundle transformation theorem) is that two coordinate bundles are equivalent iff there exist continuous functions $\lambda_j: V_j \rightarrow G$ defined for each $j \in J$ such that

$$\begin{aligned}g'_{jk}(x) &= \lambda_j(x)^{-1} g_{jk}(x) \lambda_k(x), \\ x &\in V_j \cap V_k.\end{aligned}$$

Here we assume, of course, that B and B' have the same base space, fibre space and group. If they have different coordinate neighborhoods $\{V'_j\}$ and $\{V''_k\}$ for B and B' respectively, consider a new open covering $\{V'_j \cap V''_k = V_\ell\}$. The coordinate bundle B is strictly equivalent to a coordinate bundle with coordinate neighborhoods V_ℓ and ϕ_ℓ defined as ϕ_j restricted to $V'_j \cap V''_k$. Similarly, B' is strictly equivalent to a coordinate bundle with the same neighborhoods by restricting ϕ'_k to $V'_j \cap V''_k = V_\ell$. Thus, there is no loss in the generality of the theorem to use the same covering of X for the coordinate neighborhoods of B and B' .

We shall suppose B and B' are equivalent. Then, there is a bundle map $h: B \rightarrow B'$ such that the induced map $\bar{h}: X \rightarrow X$ is the identity. There is also a set of continuous maps $\bar{g}_{kj}$ into G defined by

$$\bar{g}_{kj}(x) = \phi'_{k,x}{}^{-1} h_x \phi_{j,x} \quad x \in \bar{h}^{-1}(V_k) \cap V_j.$$

Since $\bar{h}$ is the identity,

$$\bar{g}_{kj}(x) = \phi'_{k,x}{}^{-1} h_x \phi_{j,x} \quad x \in V_k \cap V_j,$$

and the $\bar{g}_{kj}(x)$ satisfy

$$\bar{g}_{ki}(x) = \bar{g}_{kj}(x) g_{ji}(x) \quad x \in V_i \cap V_j \cap V_k$$

$$\bar{g}_{lj}(x) = g'_{lk}(x) \bar{g}_{kj}(x) \quad x \in V_\ell \cap V_j \cap V_k,$$

which follow using the definition of $g_{ij}(x)$, $g'_{ij}(x)$ in terms of $\phi_{i,x}$ and $\phi'_{i,x}$. But these relations define a $\lambda_j(x) = \bar{g}_{jj}(x)^{-1}$

which satisfies

$$\lambda_k(x)^{-1} = \bar{g}_{kj}(x)g_{jk}(x)$$

$$\lambda_j(x)^{-1} = g'_{jk}(x)\bar{g}_{kj}(x)$$

$$= g'_{jk}(x)\lambda_k(x)^{-1}g_{jk}(x)^{-1},$$

so we have

$$g_{jk}(x) = \lambda_j(x)g'_{jk}(x)\lambda_k(x)^{-1}$$

or

$$g'_{jk}(x) = \lambda_j(x)^{-1}g_{jk}(x)\lambda_k(x).$$

This demonstrates half of the theorem.

To prove the second half of the theorem we assume $g'_{kj}(x)$ and $g_{kj}(x)$ are related by

$$g'_{kj}(x) = \lambda_k(x)^{-1}g_{kj}(x)\lambda_j(x), \quad x \in V_k \cap V_j$$

where $\lambda_k(x)$ are continuous functions from V_k to G . We want to show the corresponding B and B' are equivalent. Then let us define a map

$$\bar{g}_{kj}(x) = \lambda_k(x)^{-1}g_{kj}(x).$$

It follows that

$$\begin{aligned}
\bar{g}_{kj}(x)^{-1} \bar{g}_{ki}(x) &= g_{kj}(x)^{-1} \lambda_k(x) \lambda_k(x)^{-1} g_{ki}(x) \\
&= g_{jk}(x) g_{ki}(x) \\
&= g_{ji}(x)
\end{aligned}$$

so that

$$\bar{g}_{ki}(x) = \bar{g}_{kj}(x) g_{ji}(x).$$

Similarly,

$$\begin{aligned}
\bar{g}_{lj}(x) \bar{g}_{kj}(x)^{-1} &= \lambda_l(x)^{-1} g_{lj}(x) g_{jk}(x) \lambda_k(x) \\
&= \lambda_l(x)^{-1} g_{lk}(x) \lambda_k(x) \\
&= g'_{lk}(x)
\end{aligned}$$

so that

$$\bar{g}_{lj}(x) = g'_{lk}(x) \bar{g}_{kj}(x).$$

The above properties we shall show allow us to demonstrate that $\mathcal{B}$ and $\mathcal{B}'$ are equivalent.

We must show the existence of a bundle map of $\mathcal{B}$ to $\mathcal{B}'$ which induces the identity on X . Let us define as a candidate map,

$$h_{kj}(b) = \phi'_k(x, \bar{g}_{kj}(x) \cdot p_j(b))$$

for $x = p(b)$, $x \in V_k \cap V_j$. If we show that the value of $h_{kj}(b)$ does not depend upon the local coordinate neighborhoods V_j and V_k , we will have defined a single valued function. Suppose,

therefore, that $x \in V_j \cap V_k \cap V_l \cap V_m$:

$$\begin{aligned}
 h_{kj}(b) &= \phi'_k(x, \bar{g}_{kj}(x) \cdot p_j(b)) \\
 &= \phi'_k(x, \bar{g}_{kj}(x) g_{jl}(x) \cdot p_l(b)) \\
 &= \phi'_k(x, \bar{g}_{kl}(x) \cdot p_l(b)) \\
 &= h_{kl}(b) .
 \end{aligned}$$

We also have

$$\begin{aligned}
 h_{kl}(b) &= \phi'_m(x, g'_{mk}(x) \bar{g}_{kl}(x) \cdot p_l(b)) \\
 &= \phi'_m(x, \bar{g}_{ml}(x) \cdot p_l(b)) \\
 &= h_{ml}(b) ,
 \end{aligned}$$

so that $h_{kj}(b) = h_{ml}(b)$. The functions $h_{kj}(b)$ are defined and continuous on the open domains $p^{-1}(V_k \cap V_j)$, and agree when the domains overlap. This guarantees that the $h_{kj}(b)$ define a single valued continuous function h on B . By definition $p^*h(b) = x = \bar{h}p(b)$ where $\bar{h}$ is the identity map, and is evidently induced by h . To check that h is a bundle map, we note that

$$\begin{aligned}
 \phi'_{k,x}{}^{-1} h_x \phi_{j,x} \cdot y &= \phi'_{k,x}{}^{-1} \phi'_k(x, \bar{g}_{kj}(x) \cdot y) \\
 &= \phi'_{k,x}{}^{-1} \phi'_{k,x} \bar{g}_{kj}(x) \cdot y \\
 &= \bar{g}_{kj}(x) \cdot y
 \end{aligned}$$

so that $\phi_{k,x}^{-1} h_x \phi_{j,x}$ defines a continuous map from $V_j \cap V_k$ to G , by the construction of $\bar{g}_{kj}(x)$. Our arguments show that there is a bundle map h from B to B' which induces the identity map on X , starting from the assumption of continuous maps $\lambda_j: V_j \rightarrow G$ such that $g'_{jk}(x) = \lambda_j(x)^{-1} g_{jk}(x) \lambda_k(x)$. Therefore, we have proved B is equivalent to B' iff this relation holds.

Before going on, we should notice that whenever two coordinate bundles B and B' are equivalent, the bundle spaces B and B' are homeomorphic. This follows from arguments similar to the one above. The converse is not true; B can be homeomorphic to B' but the coordinate bundle B is not equivalent to B' . An example is the torus (B) and the twisted torus (paste two ends of a cylinder after rotating one end by 180°) where the group G has two elements, the identity and the rotation by 180° . The bundle spaces are homeomorphic, but the coordinate bundle structures are not equivalent in the sense of Steenrod.

E. Bundle Construction Theorem

From the bundle transformation theorem we have just proved, we see that any two coordinate bundles with the same coordinate functions $g_{jk}(x)$ are equivalent, assuming they have the same base space, fibre and group. The next result is used quite often, and shows that if we are given a topological transformation group G on a space Y , a space X , an open covering $\{V_j\}_{j \in J}$ of X and functions g_{jk} satisfying

$$(1) \quad g_{jk}: V_j \cap V_k \rightarrow G \text{ is continuous}$$

$$(2) \quad g_{jk}(x) g_{kl}(x) = g_{jl}(x) \text{ for } x \in V_j \cap V_k \cap V_l$$

then there is a unique (up to bundle equivalence) coordinate bundle with base space X , group G , fibre Y , coordinate neighbor-

hoods $\{V_j\}_{j \in J}$ and coordinate transformations g_{jk} . We will refer later to this as the bundle construction theorem.

The essential part of the proof of this theorem is to form the bundle space by pasting together (q.v. Chapter II) the various local product spaces $V_j \times Y$.

To paste together such spaces we form the union of the product spaces $V_j \times Y \times j$, which are disjoint open sets of $X \times Y \times J$. (For the purist, J is given the discrete topology that every subset of J is open.) The union of these disjoint open sets Steenrod calls T :

$$T = \bigcup_{j \in J} V_j \times Y \times j$$

The pasting together of these spaces is accomplished by identifying points:

$$(x, y, j) \equiv (x', y', k) \pmod{R} \text{ iff}$$

$$x = x' \text{ and } y' = g_{kj}(x) \cdot y.$$

The relation R allows us to define the quotient space T/R , whose points are the equivalence classes defined by R . We call the

space T/R which we have pasted together B , and will show it is the bundle space. The natural map of $T \rightarrow T/R$ which assigns to each element (x,y,j) the equivalence class $[x,y,j]$ it belongs to we call q :

$$q(x,y,j) = [x,y,j] .$$

One defines the quotient topology such that q is continuous.

The projection $p:B \rightarrow X$ we define in the natural way

$$p[x,y,j] = x,$$

which is unique since all elements of $[x,y,j]$ have the same x . To show p is continuous, we have to demonstrate that inverse images of open sets are open. Following Steenrod let W be any open set of X . Now the function $p \circ q$ takes T into X . The inverse image of W by this function, $q^{-1}p^{-1}(W)$ is first of all in $W \times Y \times J$ and second of all in T , both of which are open sets, so their intersection is open. Every element of the intersection is necessarily in $q^{-1}p^{-1}(W)$, so $q^{-1}p^{-1}(W)$ is open. The topology on $T/R = B$ is defined (Chapter II) by demanding that V is open in B whenever $q^{-1}(V)$ is open in T . This is the quotient topology. Therefore, $q^{-1}(p^{-1}(W))$ open implies $p^{-1}(W)$ is open, so p is continuous.

Now it remains for us to construct coordinate functions which generate the given coordinate transformations $g_{kj}(x)$. These coordinate functions are supposed to be homeomorphisms from $V_j \times Y$ to $p^{-1}(V_j)$ for each $j \in J$. The obvious candidate is

$$\phi_j(x, y) = q(x, y, j)$$

which for each $x \in V_j$, $y \in Y$ gives a point in B , is jointly continuous in x and y , and satisfies $p\phi_j(x, y) = x$. Of course one might worry that there are points in $p^{-1}(V_j)$ which are not images of $V_j \times Y$. To take care of this worry let $b = [x, y, k]$ be any point in $p^{-1}(V_j)$ so $x \in V_j \cap V_k$. The point (x, y, k) is equivalent to $(x, g_{jk}(x) \cdot y, j)$, which has the image in B of

$$b = q(x, g_{jk}(x) \cdot y, j) = \phi_j(x, g_{jk}(x) \cdot y)$$

showing that b is the image of a point in $V_j \times Y$.

Another hurdle a homeomorphism must jump is that it must be a 1-1 map. That is, if $\phi_j(x, y) = \phi_j(x', y')$ then $x = x'$ and $y = y'$. But such an equality says $(x, y, j) \equiv (x', y', j) \pmod{R}$, so that $x = x'$ and $y' = g_{jj}(x) \cdot y = y$ since $g_{jj}(x) = e$. We have demonstrated that ϕ_j is a 1-1 map of $V_j \times Y$ onto $p^{-1}(V_j)$. If ϕ_j^{-1} is continuous, then ϕ_j is a homeomorphism.

Because ϕ_j is 1-1 and onto, ϕ_j^{-1} is a uniquely defined function from $p^{-1}(V_j)$ onto $V_j \times Y$. We have to show that for every open set W in $V_j \times Y$, the inverse image $(\phi_j^{-1})^{-1}(W) = \phi_j(W)$ is open in B . Now recalling the quotient topology, open sets of B are those whose inverse projection in T are open, so we want to show that $q^{-1}\phi_j(W)$ is open. The standard trick is to consider not $q^{-1}\phi_j(W)$, but the intersection of this with each open set $V_k \times Y \times k$, since these sets cover T . If each intersection is open then the union will be open and we will have the required result. The intersection $q^{-1}\phi_j(W) \cap (V_k \times Y \times k)$ is certainly in $(V_j \cap V_k) \times Y \times k$,

itself an open set in T . We need to show $q^{-1}\phi_j(W) \cap (V_k \times Y \times k)$ is open in $(V_j \cap V_k) \times Y \times k$. Let us define $\bar{q}$ as the function q restricted to act on $(V_j \cap V_k) \times Y \times k$. On this restricted set, $\bar{q}$ can be built up in two stages: first the map

$$r: (V_j \cap V_k) \times Y \times k \rightarrow V_j \times Y$$

defined by $r(x, y, k) = (x, g_{jk}(x) \cdot y)$ so r is continuous. (note $g_{jk}(x) \cdot y$ is continuous in x and y , and going to the product space $(x, g_{jk}(x) \cdot y)$ is also continuous); second, the continuous map

$$\phi_j: V_j \times Y \rightarrow B,$$

since $\phi_j(x, g_{jk}(x) \cdot y) = [x, g_{jk}(x) \cdot y, j] = [x, y, k] = \bar{q}(x, y, k)$. Now we saw r is continuous; also $r^{-1} = \bar{q}^{-1}\phi_j$ so that $r^{-1}(W) = \bar{q}^{-1}(\phi_j(W))$ is open whenever W is open. Since showing $\bar{q}^{-1}(\phi_j(W)) = q^{-1}(\phi_j(W)) \cap (V_k \times Y \times k)$ open was our task, we thus demonstrate that ϕ_j^{-1} is continuous and hence ϕ_j is a homeomorphism.

Now that we know ϕ_j is a homeomorphism, we can define $\phi_{j,x}(y) = \phi_j(x, y)$ as before and ask whether $\phi_{j,x}^{-1}\phi_{k,x}: Y \rightarrow Y$ coincides with an element of G , for $x \in V_j \cap V_k$. Let y' be related to y by this map

$$y' = \phi_{j,x}^{-1}\phi_{k,x}(y).$$

This says $\phi_{j,x}(y') = \phi_{k,x}(y)$, or equivalently $\phi_j(x, y') = \phi_k(x, y)$. By definition this means $(x, y', j) \equiv (x, y, k) \pmod{R}$ iff $x = x$ and $y' = g_{jk}(x) \cdot y$. Thus $\phi_{j,x}^{-1}\phi_{k,x}$ has the same action as $g_{jk}(x)$, an element of G . Furthermore, by assumption $g_{jk}: V_j \cap V_k \rightarrow G$ is con-

tinuous.

We have shown in detail that the $g_{jk}(x)$ define a coordinate bundle. We also know (bundle transformation theorem) any two coordinate bundles with the same coordinate transformations are equivalent. This finishes Steenrod's proof of the bundle construction theorem.

Bibliography

- Norman Steenrod, The Topology of Fibre Bundles, (Princeton University Press, Princeton, New Jersey, 1951).
- A. Trautman, "Fibre Bundles associated with Space-time," Reports on Mathematical Physics 1, 29 (1970).
- T. T. Wu and C. N. Yang, Phys. Rev. D12, 3845 (1975).

IV. BUNDLES FROM TOPOLOGICAL GROUPS

Gauge theories provide only one of many examples in physics of unifying structures. In the next two lectures two general classes of examples will be given and shown to be fibre bundle structures. This lecture will be devoted to topological groups; the next lecture to tensors over differentiable manifolds.

A. Topological Groups

We have come upon topological group structures earlier. Recall that a topological group G is a set of elements with a topology defined, where the elements form a group such that the group operations are continuous: 1) $(g_1, g_2) \rightarrow g_1 g_2$ is continuous for all $g_1, g_2 \in G$ and ii) $g \rightarrow g^{-1}$ is continuous for all $g \in G$. Under appropriate circumstances, the topological group may have more attributes. As an example, consider the group T^2 of translations in the plane. This is a Lie group whose generators are the momenta p_x and p_y . As a topological space, T^2 is homeomorphic to the plane R^2 . Now T^2 has several subgroups, and we will consider for our example the discrete subgroup Z^2 generated by unit translations along the x or the y axis.

This subgroup is that which leaves invariant the lattice with

unit square. Each unit square is identical, so the inequivalent points of the lattice are those of a square ABCD whose side AD is identified with CB

(this forms a cylinder), and whose side DC is identified with AB. The topological space formed by this identification is a torus, X .

Thus, starting with a topological group T^2 , we have identified two separate spaces Z^2 (the topology of Z^2 is such that all subsets of Z^2 are open) and the torus X . Moreover, the torus is formed as the quotient space of T^2 and the equivalence relation defined by Z^2 ; it is common to denote X by T^2/Z^2 . This notation suggests taking T^2 to be the product of X and Z^2 . Thus the group T^2 unifies a (base) space X with a (fibre) space $Y = Z^2$. The statement correctly suggests that a coordinate bundle structure is formed by the space T^2 as bundle space, X as base space, $p: T^2 \rightarrow X$ the natural map as projection, and Z^2 as fibre and as group. We will work out the coordinate neighborhoods and coordinate functions when we treat the general case below.

B. Quotient Space G/H

We now consider the general case of a topological group G and a subgroup H , which is a closed subspace. We can use H as a way of identifying points for an equivalence relation: $g \equiv g' \pmod{H}$ iff $g^{-1}g' \in H$. The quotient set formed by this identification is denoted G/H , and the natural map $p: G \rightarrow G/H$ assigns to

$g \in G$ the equivalence class $[g]$ of which g is a representative member, $p(g) = [g]$. As usual with quotient sets, we assign G/H the topology formed by taking $U \subset G/H$ to be open whenever $p^{-1}(U)$ is open. This makes G/H a topological space, called the quotient space.

The elements of G/H are subsets of G . For example, the equivalence class of the identity element e of G is $[e] = H$, the elements of the subgroup. The equivalence class of $g \in G$ is the set of all g' such that $g^{-1}g' \in H$. This means for each $h \in H$, there is a $g' = gh$ in the class $[g]$. The set of all such g' is also called a left coset of H , and denoted gH . Thus the class $[g]$ is the left coset gH . This leads to the alternative name for G/H of coset space.

C. Transformation Group on X

Our interest in a topological group G is that it can occur as a transformation group on a space X . That is there is a product " $\cdot$ " such that $(g, x) \rightarrow g \cdot x$ defines a continuous map from $G \times X$ to X and $\forall g, g' \in G, \forall x \in X, g \cdot (g' \cdot x) = (gg') \cdot x$ and $e \cdot x = x$ for the identity e of G . For example G might be the Galilean group of transformations and X the physical space of 3 dimensions. Such a group has an additional property that any point $x_0 \in X$ can be transformed to any other point $x \in X$ by some operation $g \in G$: $x = g \cdot x_0$. When G has this property, one says G is transitive.

For transitive groups G , there arises a natural subgroup. Choose any point $x_0 \in X$, and call it the base point. Now consider all the elements $h \in G$ which leave this point fixed, $h \cdot x_0 = x_0$. The set of such elements H form a subgroup of G , since 1) whenever

$h_1 \cdot x_0 = x_0$ and $h_2 \cdot x_0 = x_0$, then $(h_1 h_2) \cdot x_0 = h_1 \cdot (h_2 \cdot x_0) = h_1 \cdot x_0 = x_0$ and
 2) whenever $h \cdot x_0 = x_0$, then $h^{-1} \cdot x_0 = h^{-1} \cdot (h \cdot x_0) = (h^{-1} h) \cdot x_0 = e \cdot x_0 = x_0$.

To discuss the topological properties of H , we note that relative to our base point x_0 , we can define a function $p': G \rightarrow X$ by $p'(g) = g \cdot x_0$, which is continuous since the product "." was defined to be continuous. Now the set H is the set of all $g \in G$ such that $p'(g) = x_0$. If X has the Hausdorff separation property (distinct points of X have disjoint neighborhoods), then we can show that H is a closed subset of G .

Define a function $i: G \rightarrow X$ which is $i(g) = x_0$ for all $g \in G$. Since for each open set $U \subset X$, $i^{-1}(U)$ is either empty or X depending on whether $x_0 \notin U$ or $x_0 \in U$, then $i^{-1}(U)$ is open. Thus i defines a continuous map. By the above arguments, H is the set of points such that $i(g) = p'(g)$. Consider the mapping $r(g) = (i(g), p'(g))$ of G into $X \times X$, which is evidently continuous. Let $D = \{(x, x) : x \in X\}$ be the diagonal subset of $X \times X$. Then H is just the inverse image $r^{-1}(D)$ of this diagonal. Since r is continuous, if D is closed, so is $r^{-1}(D) = H$. The demonstration rests on showing that when X is Hausdorff, D is closed.

Let (x, x') be any point not in D ; necessarily $x \neq x'$ and so there exist disjoint open sets U and V such that $x \in U$ and $x' \in V$. This means $U \times V$, which is an open set of $X \times X$, is not in D . By forming the union of all such sets $U \times V$ over all points (x, x') not in D , we obtain an open covering of $X \times X \setminus D$. Thus the set of points of $X \times X$ not in D form an open set, so D is a closed set. Note the converse is also true. If D is a closed set, then for any two points x, x' such that $x \neq x'$, $(x, x') \notin D$ so there exists

neighborhoods U and V of x and x' respectively such that $(U \times V) \cap D = \emptyset$. This implies $U \cap V = \emptyset$.

D. $X \sim G/H$

The interest in the quotient space G/H is enhanced because under certain circumstances we can identify it topologically with the space X on which G acts. Now we will assume X is Hausdorff. This we saw implied H is a closed subgroup. We now show that if H is closed, G/H is Hausdorff. This is a necessary condition to identify G/H with X .

We first note the natural projection $p: G \rightarrow G/H$ takes open sets into open sets. Let U be any open set of G ; $p(U)$ is the corresponding set of equivalence classes, and $p^{-1}(p(U))$ consists of all points of G whose images lie in $p(U)$. Let $g \in U$, then $p(g) = gH$ and $p^{-1}(p(g)) = gH$, the set of all points in G of the form gh for every $h \in H$. Thus $p^{-1}(p(U)) = \{gh: g \in U \text{ and } h \in H\}$, which is the union over points of H of sets Uh . If the sets Uh are open, then this union is open and $p^{-1}(p(U))$ is open. By definition of the quotient topology this implies $p(U)$ is open. The set Uh is the set of products gh for $g \in U$. Define a map $r(g) = gh$ for fixed h , which maps $G \rightarrow G$. Since the group operation is continuous, r is continuous. Moreover, there is a continuous inverse for r , $r^{-1}(g) = gh^{-1}$, since $r^{-1}(r(g)) = r^{-1}(gh) = (gh)h^{-1} = g(hh^{-1}) = g$. This shows r is a homeomorphic map, and therefore takes open sets into open sets. In particular, $r(U) = Uh$ is open. This completes the demonstration that p takes open sets into open sets. Mappings with this property are called open maps or interior maps.

Now we show G/H is Hausdorff if H is closed. Take two distinct points x, x' in G/H , and choose a representative point g in the class x , and g' in the class x' . Since g and g' belong to

different equivalence classes $g^{-1}g' \notin H$. Because H is closed, its complement $G \setminus H$ is open, and $g^{-1}g'$ is in $G \setminus H$. Let W be an open set in $G \setminus H$ containing $g^{-1}g'$; by construction W and H are disjoint. Now $g^{-1}g'$ is the continuous image of $(g^{-1}, g') \in G \times G$ so for the neighborhood W of $g^{-1}g'$, there must exist a neighborhood $U^{-1} \times V$ of (g^{-1}, g') whose image is in W . In particular, U is an open neighborhood of g and V is an open neighborhood of g' and $U^{-1}V \subset W$. If we show $p(U)$ and $p(V)$ have no point in common, they provide disjoint open neighborhoods of x and x' respectively, proving G/H is Hausdorff. Suppose, on the contrary, $x'' \in p(U) \cap p(V)$, and g'' is a representative element of G in the equivalence class x'' . There is some element h of H such that $g''h \in U$ (since $p(g'') \in p(U)$) and an h' of H such that $g''h' \in V$. Since $U^{-1}V \subset W$, $(g''h)^{-1}g''h'$ must be in W ; that is, we have a point $h^{-1}h'$ of H in W . The set W was a neighborhood disjoint from H , so we have a contradiction. Thus $p(U)$ and $p(V)$ have no point in common, and so G/H is Hausdorff.

We now raise the question of when will the quotient space G/H be equivalent to X . Now the map $p': G \rightarrow X$ ($p'(g) = g \cdot x_0$) can be used to define a map from G/H to X . For any $x \in X$, $p'^{-1}(x)$ is the set of all group elements g such that $g \cdot x_0 = x$. Since G is transitive, there exists a g_0 such that $x = g_0 \cdot x_0$. Thus the set of elements leaving x fixed satisfies $g \cdot x_0 = g_0 \cdot x_0$, or equivalently $(g^{-1}g_0) \cdot x_0 = x_0$. Thus these are all the points equivalent to g_0 , which is a point in the quotient space G/H . To each x , there is a unique class of elements $[g_0] \in G/H$. Conversely, to each class of elements $[g] \in G/H$, there is a unique element $x \in G/H$

$$q[g] = g \cdot x_0$$

and this defines a unique 1-1 function from G/H to X . To show q is continuous, let U be an open set of X ; since p' is continuous, $p'^{-1}(U)$ is open in G . The inequivalent classes of $p'^{-1}(U)$ are just the elements of $q^{-1}(U)$. Moreover, these classes are the natural projection of $p'^{-1}(U)$ by p : $p(p'^{-1}(U)) = q^{-1}(U)$. Since p is an open map (we showed p takes open sets into open sets), we find that $q^{-1}(U)$ is open whenever U is open. Therefore, q is continuous.

This establishes a 1-1 continuous map q of G/H onto X . If q^{-1} were also continuous, G/H would be homeomorphic to X . This need not be true. For the result to be true, it is sufficient either for G to be compact or for $p': G \rightarrow X$ to map open sets into open sets. (For the proofs see [Steenrod] and [Bourbaki].)

E. Bundle Structure Theorem

Thus under reasonable conditions, the study of certain topological spaces reduces to the study of quotient spaces of a group with respect to closed subgroups. We now want to show how a topological group G , closed subspace H and projection p of G onto G/H serve to define a bundle structure with base space G/H , fibre H , group G , projection p , and bundle space G .

1. Local cross section assumption

To show the bundle structure, one needs some guarantee that the bundle space G is locally a product space between the base G/H and fibre H . In other words, about each point $x \in G/H$, there should be a neighborhood V of x such that $p^{-1}(V) \approx V \times H$. If this were true, there would evidently exist local cross sections

$f: V \rightarrow G$ such that f is continuous and $pf(x) = x$ for every $x \in V$.

One could take any continuous function $r: V \rightarrow H$ and define $f(x) = \phi(x, r(x))$ where ϕ is the homeomorphism between $V \times H$ and $p^{-1}(V)$ (i.e. the coordinate function for neighborhood V) satisfying $p\phi(x, y) = x$.

We quote without proof the theorem that for any Lie group G , closed subgroup H , there exists a local cross section of H in G . In other words, denoting the coset H by x_0 , there exists an open neighborhood V of x_0 and a continuous function $f: V \rightarrow G$ such that $pf(x) = x$ for every $x \in V$. The function f is called the local cross section. Most of our interest is in Lie groups, so there is no important restriction in what follows to consider only groups G and closed sets H in which there is a local cross section of H in G . See [Steenrod] for further discussion, and [Chevalley] for the proof.

2. Statement of theorem

With this restriction on G , we now prove the Bundle Structure Theorem of Steenrod. If the closed subgroup H of G admits a local cross section f , and $p: G \rightarrow G/H$ is the natural map, then there is a bundle structure with G as bundle space, G/H as base space, p as projection, H as fibre and G as group. The theorem can be generalized

slightly by considering a closed subgroup K of H . Now let $p: G/K \rightarrow G/H$ be the map $p(gK) = gH$ which takes the coset gK of G/K into gH for each $g \in G$. Then there is a bundle structure with bundle space G/K , base space G/H , projection p , fibre H/K , and group G/K_0 . The subgroup K_0 is the largest subgroup of K invariant in G (we will define this below). Furthermore, for any two local cross sections, the resultant coordinate bundles are equivalent. Finally, left translations of G/K by G define bundle mappings of the bundle onto itself. The above figure is a way of picturing the bundle space, base space and fibre. The third axis, as we shall see, denotes the set used to index the coordinate neighborhoods.

1. p is a projection

We now go over Steenrod's proof. First, we introduce the natural maps p_1 and p_2 of G onto G/K and G/H respectively. The mapping diagram illustrates the relationship between p , p_1 and p_2 .

By this definition, for $g \in G$, $p_1(g) = gK$, $p_2(g) = gH$ and $p(gK) = gH$; $pp_1 = p_2$. By the general arguments p_1 and p_2 are continuous. To show p is continuous, one takes any open set $U \subset G/H$ so $p_2^{-1}(U)$ is open since p_2 is continuous. Using $p_2^{-1} = p_1^{-1}p^{-1}$ we have that $p_1^{-1}(p^{-1}(U)) = p_2^{-1}(U)$ is open. By definition of the quotient topology on G/K , whenever $p_1^{-1}(p^{-1}(U))$ is open, $p^{-1}(U)$ is open. Thus, p is a continuous function. [Of course, when $K = \{e\}$, there would

be nothing to prove here since $p = p_2$.]

4. Coordinate neighborhoods

The next part of the proof involves defining what the coordinate neighborhoods are, and constructing coordinate functions from the local cross section f that we are given. Let G be the indexing set for the bundle. Since there exists a local cross section of H in G , there is an open neighborhood V of x_0 (x_0 is the coset H in G/H) on which a continuous function $f:V \rightarrow G$ is defined satisfying $p_2 f(x) = x$, $\forall x \in V$. Now define for each $g \in G$, other open sets

$$V_g = g \cdot V,$$

by which we mean the homeomorphic transformation of V into another open set of G/H defined by each element $g \in G$. Specifically, define a map r which takes each class $[g']$ into $[gg']$:

$$r[g'] = g \cdot [g']$$

$$= [gg']$$

Since we have seen that $g' \rightarrow gg'$ defines a homeomorphism of G onto itself, the natural projection onto G/H induces a homeomorphism r of G/H onto G/H . In particular r maps V onto V_g showing V_g is an open set.

Now each point $x \in G/H$ has a representative element g where $x = [g]$. Moreover, for this g , $x = g \cdot x_0$ so $x \in V_g$. Thus the $\{V_g\}_{g \in G}$ is an open covering of G/H (which we are thinking of as a space X).

Next, we can "translate" the local cross section f to each of the open sets V_g by defining

$$f_g(x) = g f(g^{-1} \cdot x)$$

$$\forall x \in V_g.$$

Thus every $x \in V_g = g \cdot V$ satisfies $g^{-1} \cdot x \in V$; $f(g^{-1} \cdot x)$ is an element of G , then so is $g f(g^{-1} \cdot x)$. Each of these operations is continuous, so $f_g: V_g \rightarrow G$ is continuous. Now if we project $f_g(x)$ onto G/H , we obtain the coset $f_g(x)H = (g f(g^{-1} \cdot x))H = g(f(g^{-1} \cdot x)H) = g(p_2 f(g^{-1} \cdot x)) = g \cdot (g^{-1} \cdot x) = (g g^{-1}) \cdot x = e \cdot x = x$ using the definitions. Thus we have $p_2 f_g(x) = x$, and f_g is a local cross section in the neighborhood V_g .

5. Pasting

Given the local cross sections, and coordinate neighborhoods for each $g \in G$, the idea of the proof is to take the $V_g \times H/K$ and paste them together to form the bundle space G/K . Now we recall (chapter II) that to paste patches together we first form

$$T = \bigcup_{g \in G} (V_g \times H/K \times \{g\})$$

the disjoint union of open patches. We then identify points in T by defining a continuous map

$$\phi : T \rightarrow G/K$$

so that the equivalence class defined by $\phi(x, y, g) = \phi(x', y', g')$ is identified with a point of G/K . To carry out this program, let us define for each $g \in G$

$$\phi_g(x, y) = f_g(x) \cdot y$$

for each $x \in V_g$ and $y \in H/K$. The function ϕ above will then be

defined as

$$\phi(x, y, g) = \phi_g(x, y) \quad .$$

The function $\phi_g(x, y)$ is the continuous operation of $f_g(x)$, an element of G , on the coset y of H/K . So if hK is the coset y , the multiplication by an element of G will in general give a coset of G/K . Therefore, ϕ_g is a continuous map of $V_g \times H/K$ into G/K .

Now we ask over what points of G/H are the images $\phi_g(x, y)$. For this we compute $p\phi_g(x, y) = p(f_g(x) \cdot y) = p(f_g(x) \cdot hK) = p((f_g(x)h) \cdot K) = (f_g(x)h) \cdot H = f_g(x) \cdot (hH) = f_g(x) \cdot H = p_2 f_g(x) = x$, where we have written y as the coset hK . The result that $p\phi_g(x, y) = x$ shows $\phi_g: V_g \times H/K \rightarrow p^{-1}(V_g)$.

Next, we construct a function which will take each point of $p^{-1}(V_g)$ into the "y-axis", H/K . Define for each $z \in p^{-1}(V_g)$,

$$p_g(z) = f_g(p(z))^{-1} \cdot z \quad .$$

The assumption $z \in p^{-1}(V_g)$ implies $p(z) \in g \cdot V$ or $g^{-1} \cdot p(z) \in V$. This shows $p(z)$ is in the domain of f_g , that p_g is well defined, and $p_g(z)$ is continuous. Each z has the form $g'K$ since z is a coset of G/K . Thus

$$\begin{aligned} p_g(z) &= f_g(p(g'K))^{-1} \cdot g'K \\ &= f_g(g'H)^{-1} \cdot g'K \\ &= (f((g^{-1}g') \cdot H)^{-1} g^{-1}) \cdot g'K \\ &= [f((g'^{-1}g)^{-1} \cdot H)^{-1} (g^{-1}g')] \cdot K \end{aligned}$$

$$= f_{g, -1_g(H)}^{-1} \cdot K$$

Writing H as x_0 , $f_{g, -1_g}(x_0)$ satisfies $p_2 f_{g, -1_g}(x_0) = x_0$, so that $f_{g, -1_g}(x_0) = h$, some element of H . From above this implies $p_g(z) = hK$, a coset of H/K . Thus p_g is a continuous map of G/K into H/K .

Moreover, we have that for $x \in V_g$, $y \in H/K$,

$$p_g \phi_g(x, y) = f_g(p \phi_g(x, y))^{-1} \cdot \phi_g(x, y)$$

$$= f_g(x)^{-1} \cdot \phi_g(x, y)$$

$$= f_g(x)^{-1} \cdot (f_g(x) \cdot y)$$

$$= (f_g(x)^{-1} f_g(x)) \cdot y$$

$$= e \cdot y$$

$$= y$$

Thus we can define a continuous function $\phi_g^{-1}(z) = (p(z), p_g(z))$ of G/K into $V_g \times H/K$ which is inverse to $\phi_g(x, y)$:

$$\phi_g \phi_g^{-1}(z) = \phi_g(p(z), p_g(z))$$

$$= f_g(p(z)) \cdot p_g(z)$$

$$= f_g(p(z)) \cdot (f_g(p(z))^{-1}) \cdot z$$

$$= (f_g(p(z)) f_g(p(z))^{-1}) \cdot z$$

$$= e \cdot z$$

$$= z$$

Also, $\phi_g^{-1} \phi_g(x, y) = (p\phi_g(x, y), p_g\phi_g(x, y)) = (x, y)$. Therefore, ϕ_g is a homeomorphism of $V_g \times H/K$ onto $p^{-1}(V_g)$. Thus one can use ϕ_g to define a continuous function $\phi: T \rightarrow G/K$, which will paste together the patches $V_g \times H/K \times \{g\}$ to form G/K . We needn't do this explicitly, however, if we demonstrate that the ϕ_g satisfy the proper conditions to be coordinate functions.

6. ϕ_g are coordinate functions

To verify $\phi_g(x, y)$ are the coordinate functions, we must check that

$$g_{g_1 g_2}(x) = \phi_{g_1, x}^{-1} \phi_{g_2, x}$$

coincides with an operation of the group of the fibre. We must identify the group which operates effectively on the fibre, and show $g_{g_1 g_2}$ maps $V_{g_1} \cap V_{g_2}$ continuously into this group. Consider the action of $g_{g_1 g_2}(x)$ on y :

$$\begin{aligned} g_{g_1 g_2}(x) \cdot y &= (\phi_{g_1, x}^{-1} \phi_{g_2, x}) \cdot y \\ &= \phi_{g_1, x}^{-1} (\phi_{g_2}(x, y)) \\ &= p_{g_1}(\phi_{g_2}(x, y)) \\ &= (f_{g_1}(x)^{-1} f_{g_2}(x)) \cdot y. \end{aligned}$$

The action of $g_{g_1 g_2}(x)$ on y coincides with an operation of $f_{g_1}(x)^{-1} f_{g_2}(x)$. Since $x = p_2 f_{g_1}(x) = p_2 f_{g_2}(x)$, then $f_{g_1}(x)H = f_{g_2}(x)H$, or $f_{g_1}(x)^{-1} f_{g_2}(x) \in H$. The function $g_{g_1 g_2}$ is a continuous map of

$V_{g_1} \cap V_{g_2}$ into H , since both $f_{g_1}(x)$ and $f_{g_2}(x)$ are continuous functions of x .

7. H/K_0 is group of bundle

Now we would expect that H is the group of the bundle (which it is when $K = \{e\}$). The only thing that could go wrong is that H does not act effectively on H/K . Recall that a group action is effective if whenever $g \cdot y = y \ \forall y \in H/K$ then $g = e$. So let us suppose for $g \in H$,

$$g \cdot y = y \quad \forall y \in H/K.$$

Stating this in terms of the coset notation, we have

$$g \cdot (hK) = hK \quad \forall h \in H.$$

In other words, $ghK = hK$, or $gh \in hK$, or $g \in hKh^{-1}$ for every $h \in H$. So we must determine the number of elements in the set K_0 defined as the intersection of sets hKh^{-1} over all elements h of H . The set hKh^{-1} consists of all points of H of the form hkh^{-1} for any $k \in K$. By its definition, $\forall h \in H$, $hK_0h^{-1} = K_0$, and so K_0 is called an invariant subgroup of H . It is the largest invariant subgroup of H inside K . We have proved that any g which leaves every element $y \in H/K$ fixed, is an element of K_0 .

Conversely, suppose that g is any element of K_0 . Then for every $h \in H$, $g \in hKh^{-1}$ by definition of K_0 . Thus $gh \in hK$, or $ghK = hK$, and so g leaves fixed each element $hK = y \in H/K$.

This leads us to consider the group H/K_0 . To show this quotient space is also a group, we define multiplication by $hK_0 * h'K_0 = hh'K_0$, $\forall h, h' \in H$. This is equivalent to $(hK_0)(h'K_0) = (hh'K_0h'^{-1})h'K_0 = hh'K_0(h'^{-1}h')K_0 = hh'K_0K_0 = hh'K_0$. One can

verify this defines a topological group structure on H/K_0 . To each element h of H there corresponds a natural element $\pi(h)$ of H/K_0 (the coset hK_0). Define the action of $\pi(h)$ on H/K by

$$\pi(h) \cdot y = h \cdot y$$

$$\forall y \in H/K.$$

In this way we see that H/K_0 is an effective action on H/K .

Thus to each element $g_{g_1 g_2}(x)$ corresponds continuously an element $\pi(g_{g_1 g_2}(x)) \in H/K_0$, and $g_{g_1 g_2}(x)$ coincides with the action of $\pi(g_{g_1 g_2}(x))$:

$$g_{g_1 g_2}(x) \cdot y = \pi(g_{g_1 g_2}(x)) \cdot y.$$

From this we conclude that H/K_0 is the group of the bundle, and ϕ_g are the coordinate functions.

8. Structure is independent of local cross sections

Now if two different local cross sections f and f' were given, defined on open neighborhoods V and V' of x_0 , then we could repeat the above demonstration, putting primes on f , ϕ_g and p_g . Both coordinate bundles have the same base space, fibre, bundle space and group. We need to check that

$$\bar{g}_{g_1 g_2}(x) \cdot y = \phi_{g_1, x}^{-1} \phi_{g_2, x}(y)$$

coincides with an operation of H/K_0 , and $\bar{g}_{g_1 g_2}$ maps $V'_1 \cap V_{g_2}$ continuously into H/K_0 . We compute that

$$\begin{aligned} \bar{g}_{g_1 g_2}(x) \cdot y &= p'_1 \phi_{g_2}(x, y) \\ &= (f'_{g_1}(x)^{-1} f_{g_2}(x)) \cdot y. \end{aligned}$$

As before we note $x = p_2 f'_{g_1}(x) = p_2 f'_{g_2}(x)$ implies that $f'_{g_1}(x)^{-1} f'_{g_2}(x)$ is in H , and maps $V'_{g_1} \cap V_{g_2}$ continuously into H . Identifying $\bar{g}_{g_1 g_2}(x)$ with $\pi(f'_{g_1}(x)^{-1} f'_{g_2}(x))$ we prove the result that the bundle structures are strictly equivalent.

4. Bundle maps by left translations

The last statement of the proof involves defining for each $g \in G$ the left translations of G/K :

$$g \cdot (g'K) = (gg')K, \quad \forall g' \in G.$$

This defines continuously for each $z \in G/K$, the point $g \cdot z \in G/K$. Now $g \cdot z$ is in the fibre of $p(g \cdot z)$. Writing z as $g'K$, we see that $p(g \cdot z) = p(gg'K) = gg'H$ and $g \cdot p(z) = g \cdot (g'H) = gg'H$. Thus we have $p(g \cdot z) = g \cdot p(z)$. This shows the induced map on the base G/H is also defined by g , and that left translation by g preserves fibres. To check it gives a bundle map, we consider

$$\begin{aligned} \bar{g}_{g_1 g_2}(x) \cdot y &= \phi_{g_1, x}^{-1}(g \cdot \phi_{g_2, x}(y)) \\ &= (f'_{g_1}(x')^{-1} g f'_{g_2}(x)) \cdot y \end{aligned}$$

$$x' = g \cdot x \in V_{g_1}$$

$$x \in V_{g_2}$$

This shows $\bar{g}_{g_1 g_2}(x)$ coincides with the operation of $\pi(f'_{g_1}(x')^{-1} g f'_{g_2}(x))$ and is a continuous map of $(g^{-1} \cdot V_{g_1}) \cap V_{g_2}$ into H/K_0 .

This proves that left translations by g maps the bundle onto itself.

P. Perspective

With some effort, we have verified that any Lie group G , and closed subgroups H of G and K of H define a coordinate bundle structure. Let us take $K = \{e\}$, then the bundle space is G , base space G/H , fibre H and group H . The coordinate functions are calculable up to bundle equivalence from any local cross section of H in G . As an example consider G to be the Galilean group of transformations on 3-space X and H to be the subgroup leaving an arbitrary point $x_0 \in X$ fixed. Then we identify X with G/H ; H is the group of rotations.

As another example let G be the Galilean group, but H a crystallographic group: elements of H leave invariant a given crystal. Again G , G/H , H form elements of a coordinate bundle. A nice summary is given by [Michel].

One can clearly provide whole classes of interesting examples. So if we are able to show there are calculational techniques of such things as homotopy groups of base spaces, bundle spaces, and fibre spaces for coordinate bundles, then we can obtain useful results for our favorite physics problem.

Another less apparent benefit is that the bundle properties of topological groups can be used to simplify more general bundle structures by, in some cases, finding equivalent bundles with smaller groups. Steenrod discusses this in great detail.

Bibliography

N. Bourbaki, op. cit.

C. Chevalley, Theory of Lie Groups (Princeton University Press, 1946, Princeton).

L. Michel, "Topological Classification of Symmetry Defects in ordered media", Talk given at the VIth International Colloquium on Group Theoretical Methods in Physics Tübingen, July 18-22, 1977.

N. Steenrod, op. cit.

V. DIFFERENTIABLE MANIFOLDS AND TENSOR BUNDLES

Of all the examples of fibre bundle structures, the most familiar to physicists is the tensor bundle. Of course one needn't recognize it as such in order to use it. In this lecture we will show how tensor fields become identified with bundle structures.

A. Differentiable Manifolds

We know that tensors are constructed on spaces with certain differentiable properties, so it is natural to begin the discussion with the topological properties of these spaces. An n -dimensional manifold X is a topological space which is locally homeomorphic to an open set of Euclidean n -space R^n . The only manifolds to be discussed are those which as topological spaces are Hausdorff (distinct points have disjoint open neighborhoods), connected (the space can not be written as the disjoint union of two open sets), metric (there is a distance function defined for each pair of points of X , and this distance function is used to define open sets), and separable (X has a countable base Σ , which is a countable collection of open sets such that every open set of X is the union of open sets of Σ).

An example of a 2-manifold is the sphere which is locally homeomorphic to an open set of the plane. If one thinks about navigating

Fig. Each open patch on the sphere is homeomorphic to an (open) square of the 2-dimensional plane.

on the earth's surface, which is homeomorphic to a sphere, one can appreciate the usefulness in having charts, which are flat subsets of the plane. The fact that the plane and the sphere are not topologically equivalent necessitates the use of more than one chart to cover the globe; one needs in fact a complete atlas.

One indeed extends the terminology of the above example to a special class of manifolds. One defines a system $\sum$ of differentiable coordinates in an n-manifold by giving an open covering $\{V_j\}_{j \in J}$, and for each $j \in J$, a homeomorphic map $\psi_j: E_j \rightarrow V_j$ for some open set E_j of R^n . These homeomorphisms must satisfy the property that the map

$$\psi_j^{-1} \psi_i: \psi_i^{-1}(V_i \cap V_j) \rightarrow \psi_j^{-1}(V_i \cap V_j)$$

be differentiable for $i, j \in J$. The pair (V_j, ψ_j) is called a chart, and the family of charts $\{V_j, \psi_j\}_{j \in J}$ is called the atlas. Finally the system $\sum$ is of class r if each $\psi_j^{-1} \psi_i$ has continuous derivatives of order r . If this is true for every r , $\sum$ is of class ∞ . If the maps $\psi_j^{-1} \psi_i$ are analytic, then $\sum$ is of class ω .

The n -manifold X is in fact a construction out of patches $\{E_j\}_{j \in J}$ of R^n which are glued together by the $\{\psi_j\}_{j \in J}$. To see this let us proceed in the general way we have used before with bundles, and define a patch as

$$\bar{E}_1 = E_1 \times \{1\}$$

which is an open subset of $R^n \times J$, where the topology on J is such that every subset of J is an open set. We next take the union over all patches

$$\bar{E} = \bigcup_{i \in J} \bar{E}_i$$

which is a disjoint union of open sets. Now over each open set E_1 , we have a homeomorphism $\psi_1: E_1 \rightarrow V_1$. We use these homeomorphisms to define a continuous function $\psi: E \rightarrow X$ by

$$\psi(a, i) = \psi_1(a) \text{ whenever } a \in E_1.$$

Since E_1 are disjoint, ψ is well defined. To show ψ is continuous, observe for any open set $U \subset X$, $\psi^{-1}(U) = \bigcup_{i \in J} (\psi^{-1}(U) \cap E_i)$. So if each $\psi^{-1}(U) \cap E_1$ is open, then $\psi^{-1}(U)$ is open and hence ψ is continuous. On E_1 , $\psi^{-1}(U) \cap E_1 = \psi_1^{-1}(U \cap V_1)$, and since V_1 is open, so is $U \cap V_1$ and hence $\psi_1^{-1}(U \cap V_1)$ by the continuity of ψ_1 .

It is also clear that since $\psi(E_j) = V_j$ and the $\{V_j\}_{j \in J}$ cover X , then ψ maps E onto X . But ψ will not in general be a 1-1 map. The points of E which ψ identifies makes up the space X . Let R be the equivalence relation defined by ψ :

$$(a, i) \equiv (b, j) \pmod{R} \text{ iff } \psi(a, i) = \psi(b, j).$$

Let q be the natural projection of $E \rightarrow E/R$ defined by $q(a, i) = [a, i]$, the set of all $(b, j) \in E$ such that $\psi(a, i) = \psi(b, j)$. As usual, the topology on E/R is such that q is continuous. Finally, there is a natural map ψ_0 from $E/R \rightarrow X$ defined by $\psi_0[a, i] = \psi(a, i)$; the function is unique since the elements of $[a, i]$ have the same image under ψ . The function is continuous since for any open $U \subset X$, $q^{-1}\psi_0^{-1}(U) = \psi^{-1}(U)$ is open, and by the quotient topology on E/R this implies $\psi_0^{-1}(U)$ is also open. Next, ψ_0 is a 1-1 map, since if $\psi_0[a, i] = \psi_0[b, j]$ then $\psi(a, i) = \psi(b, j)$ so $(a, i) \equiv (b, j) \pmod{R}$, and so $[a, i] = [b, j]$. Finally, since ψ is onto, ψ_0 is also onto. We have now shown the general result that a map $\psi: E \rightarrow X$ has a canonical decomposition $E \xrightarrow{q} E/R \xrightarrow{\psi_0} X$ where ψ_0 is a 1-1 continuous function

of E/R onto X .

To show E/R and X are homeomorphic, we have to show ψ_0^{-1} is continuous.

Let W be any open set of E/R , then $(\psi_0^{-1})^{-1}(W) = \psi_0(W) = \psi(q^{-1}(W))$
 $= \psi\left(\bigcup_{i \in J} (q^{-1}(W) \cap E_i)\right) = \bigcup_{i \in J} \psi_i(q^{-1}(W) \cap E_i)$. Since q is continuous
 and E_i is open, then $q^{-1}(W) \cap E_i$ is open. Because each ψ_i is a homeomorphism, $\psi_i(q^{-1}(W) \cap E_i)$ is also open; the union of open sets is open implying
 $\bigcup_{i \in J} \psi_i(q^{-1}(W) \cap E_i)$ is open. Thus $\psi_0^{-1}(W)$ is open for any $W \subset E/R$,
 thereby showing ψ_0^{-1} is continuous. Thus the n -manifold X is topologically
 the same as the patches E_i pasted together by ψ , namely E/R .

If we examine the pasting operation in more detail, we see $(a, i) \equiv (b, j)$
 (mod R) is true iff $\psi_i(a) = \psi_j(b) \in V_i \cap V_j$ and $a \in E_i, b \in E_j$. Or in
 other words, $b = \psi_j^{-1}\psi_i(a)$; in particular $\psi_j^{-1}\psi_i$ maps all points of $\psi_i^{-1}(V_i \cap V_j)$
 onto $\psi_j^{-1}(V_i \cap V_j)$. The kinds of n -manifolds are distinguished by noting
 whether for all $i, j \in J$, $\psi_j^{-1}\psi_i$ has continuous derivatives of order r , for
 some r . We obviously want to call two systems of coordinates Σ and Σ'
 r -equivalent if the composite atlas $\{V_j, \psi_j\}_{j \in J} \cup \{V'_k, \psi'_k\}_{k \in J'}$ forms
 a system of coordinates of class r . This defines an equivalence relation;
 a differentiable n -manifold of class r is an n -manifold X together with
 an r -equivalence class of coordinates of class r .

B. Coordinate Transformations

Up to now we have said nothing about bundles, but have rather summarized
 the description of a differentiable manifold. Our ordinary space-time of re-
 lativity is such a differentiable manifold. Given such a manifold, in physics
 we construct over this manifold various tensor fields. We now turn to examine
 such constructions.

Since X is not necessarily Euclidean, an important consideration is how
 the coordinates transform in going from one patch to another. Such a trans-
 formation is effected by $\psi_j^{-1}\psi_i$; over each $V_j \cap V_i$, $\psi_j^{-1}\psi_i$ transforms
 $\psi_i^{-1}(V_i \cap V_j)$ onto $\psi_j^{-1}(V_i \cap V_j)$. Specifically, let $x \in V_j \cap V_i$, $\psi_i^{-1}(x) = b_i$

and $\psi_j^{-1}(x) = b_j$. Then the coordinates at b_1 are related to the coordinates at b_j by a linear transformation $a_{ji}(x)$ which is the Jacobian

matrix of $\psi_j^{-1}\psi_1$:

$$db_j = d\psi_j^{-1}(x)$$

$$= d(\psi_j^{-1}\psi_1(b_1))$$

$$\begin{aligned} (db_j)^\alpha &= \sum_{\beta=1}^n \frac{\partial}{\partial b_1^\beta} (\psi_j^{-1}\psi_1(b_1))^\alpha (db_1)^\beta \\ &= \sum_{\beta=1}^n (a_{ji}(x))_\beta^\alpha (db_1)^\beta \end{aligned}$$

where the Euclidean vectors $b_j = (b_j^1, \dots, b_j^n)$ have components b_j^α , $\alpha = 1, \dots, n$. The matrices $a_{ji}(x)$ for each $j, i \in J$ and $x \in X$ are elements of L_n , non-singular $n \times n$ matrices. The set L_n forms a group, and is a subset of points of R^{n^2} if the elements of each matrix of L_n are considered as coordinates of a n^2 -dimensional vector. To each point A in L_n all points sufficiently close to A are likewise in L_n . This means L_n is a neighborhood of each of its points and hence L_n is an open subset of R^{n^2} . Since L_n is an open set of R^{n^2} , it is necessarily an analytic manifold;

the ω -equivalence class of L_n defines an analytic differentiable manifold. These properties of L_n will be used to define the group structure of tensor bundles.

One other property of the $a_{ji}(x)$ will be needed if we are to identify the $a_{ji}(x)$ as elements of a group of a bundle, namely the transitive property of Jacobean matrices

$$a_{jk}(x) a_{ki}(x) = a_{ji}(x),$$

$$x \in V_i \cap V_j \cap V_k.$$

To work this out, we write out explicitly the matrix multiplications

$$\sum_{\beta=1}^n (a_{jk}(x))_{\beta}^{\alpha} (a_{ki}(x))_{\gamma}^{\beta} = (a_{ji}(x))_{\gamma}^{\alpha}.$$

Then in terms of the Jacobians, this is

$$\begin{aligned} \sum_{\beta=1}^n \frac{\partial}{\partial b_k^{\beta}} (\psi_j^{-1} \psi_k(b_k))^{\alpha} \frac{\partial}{\partial b_i^{\gamma}} (\psi_k^{-1} \psi_i(b_i))^{\beta} \\ = \frac{\partial}{\partial b_i^{\gamma}} (\psi_j^{-1} \psi_i(b_i))^{\alpha}. \end{aligned}$$

Because of these transitive relations, $a_{ii}(x) = e$, the identity matrix, and $a_{ij}(x)^{-1} = a_{ji}(x)$. This shows again that $a_{ij}(x) \in L_n$, and indeed gives a continuous map $a_{ij}: V_i \cap V_j \rightarrow L_n$.

C. Tensor Bundles

We now have all the elements with which to define a bundle structure on X . We call X the base, $\{V_j\}_{j \in J}$ the coordinate neighborhoods, L_n the group and $g_{ij}(x) = a_{ij}(x)$ the coordinate transformation functions. The fibre γ is the space of real n -dimensional vectors on which L_n acts. The bundle structure theorem can now be invoked to define the remaining elements of the coordinate

bundle, namely the bundle space B and the coordinate function ϕ_j . The bundle so defined is the tangent bundle of X .

Suppose now there exists a (global) cross section $f: X \rightarrow B$, namely a continuous map satisfying $pf(x) = x$ for every x in X . Since each fibre $p^{-1}(x)$ over x in X is homeomorphic to the same fibre Y (the space of real n -dimensional vectors), the value $f(x)$ can be thought of as a specific element of Y , namely a vector. Thus the cross section f gives a continuous vector valued field defined over all of X . The converse is true: when there is a vector field over X then there exists a global cross section. Note that there will always be local cross sections, but there can be cases where global cross sections do not exist. This is a restatement in bundle language that not every manifold admits a vector field continuously defined over the whole manifold. One application of bundle theory is to the existence question of various tensor fields.

Just as there are other tensors over X besides vectors, there are also other bundle structures associated with X . We know that X is the natural base, and $\{V_j\}_{j \in J}$ the natural family of neighborhoods. Now let Y be any fibre space, and G a group acting effectively on Y . If there exists a continuous homomorphism $h: L_n \rightarrow G$ then $g_{j1}(x) = h(a_{j1}(x))$ defines a continuous map of $V_j \cap V_1 \rightarrow G$. We take g_{j1} to be the coordinate transformation functions, and call the bundle constructed from the bundle structure theorem a tensor bundle of type h . The tangent bundle was a special case of $G = L_n$ and h the identity.

Another example is to take $G = L_n$, Y the space of n -dimensional vectors, and h the mapping taking each matrix of L_n into the transpose of its inverse:

$$h(A) = (A^{-1})^T.$$

For a vector $y_\beta, \beta=1, \dots, n,$

$$(h(A)y)_\beta = \sum_Y (A^{-1})^\gamma_\beta y_Y,$$

which shows y_β transforms covariantly; For vectors $y^\beta,$

$$(Ay)^\beta = \sum_Y A^\beta_Y y^Y$$

so y^β transforms contravariantly. Thus the tangent bundle is the bundle of contravariant vectors, and for $h(A) = (A^{-1})^T$ we get the bundle of covariant vectors.

As a final class of examples, G can be the 1-dimensional linear group (1-1 non-singular matrices), and $h: L_n \rightarrow G$ defined by $h(A) = (\det A)^N$ for some positive integer N . The space of 1-dimensional vectors Y is the set R of real numbers. The bundle space one constructs is called the bundle of relative scalars of weight N .

Under certain circumstances, we can show that the bundle space itself is a differential manifold. Under these circumstances, tensor bundles can also be constructed over B . The argument we give uses extensively the idea that a function $f: X \rightarrow X'$ is of class r if it is a mapping of one differentiable manifold of class $\geq r$ into another of class $\geq r$, and if there are representative systems of neighborhoods Σ and Σ' of X and X' respectively such that $\forall j \in J, j' \in J'$

$$\psi_{j'}^{-1} \circ f \circ \psi_j: \psi_j^{-1}(V_j \cap f^{-1}(V_{j'})) \rightarrow E'_{j'}$$

has continuous derivatives to the order r . If this is true for one pair Σ, Σ' , it will be true for all pairs. If there is a third differentiable manifold X'' of class $\geq r$ and $g: X' \rightarrow X''$ is of class r , then $gf: X \rightarrow X''$

is of class r . Let Σ' and Σ'' be the pair of neighborhoods such that

$$\psi_j''^{-1} \circ \psi_j' : \psi_j'^{-1} (V_j' \cap g^{-1} (V_j'')) \rightarrow E_j''$$

has continuous derivatives to order r . Then

$$\psi_j''^{-1} \circ \psi_j' \circ \psi_j'^{-1} \circ f \circ \psi_j = \psi_j''^{-1} \circ g \circ \psi_j :$$

$$\psi_j^{-1} (V_j \cap f^{-1} (V_j' \cap g^{-1} (V_j''))) \rightarrow E_j''$$

has continuous derivatives to order r , for every set V_j' , so the argument holds also over the domain $\psi_j^{-1} (V_j \cap f^{-1} g^{-1} (V_j''))$.

We conclude by showing an example of when the bundle space B of a tensor bundle of type h is itself a differentiable manifold. The example is provided by the case in which G is a Lie Group. Let X as usual be the base (which is assumed to be a differentiable manifold of type r), Y the fibre, B the bundle space from the bundle construction theorem, and ϕ_j the coordinate functions. Assume Y is a differentiable manifold of type $\geq r-1$, and let $\{V_k'', \psi_k''\}_{k \in K}$ be the atlas. The homeomorphic sets in R^n are $E_k'' [\psi_k'' : E_k'' \rightarrow V_k'']$. Now the atlas for B is constructed as follows. Let $\psi_{jk}' :$

$E_j \times E_k'' \rightarrow B$ be defined by $\psi_{jk}'(u, v) = \phi_j(\psi_j(u), \psi_k''(v))$. The open sets in B are $V_{jk}' = \phi_j(V_j \times V_k'')$, which form a covering of B . One has to check that $\psi_{\ell m}^{-1} \circ \psi_{jk}' : \psi_{jk}'^{-1}(V_{jk}' \cap V_{\ell m}') \rightarrow \psi_{\ell m}^{-1}(V_{jk}' \cap V_{\ell m}')$ has continuous derivatives to some order. Pick a point $z \in V_{jk}' \cap V_{\ell m}'$ so there exists an $x \in V_j$ and $y \in V_k''$ such that $z = \phi_j(x, y)$; similarly there exists an $\bar{x} \in V_\ell$ and $\bar{y} \in V_m''$ such that $z = \phi_\ell(\bar{x}, \bar{y})$. Since $z = \phi_j(x, y) = \phi_\ell(\bar{x}, \bar{y})$ we have $x = \bar{x} \in V_j \cap V_\ell$ and $\bar{y} = g_{\ell j}(x) \cdot y$.

Now the points x and y are the images of $u \in E_j$ and $v \in E_k''$

respectively; similarly the points $\bar{x}$ and $\bar{y}$ are the images of $\bar{u} \in E_\ell$ and $\bar{v} \in E_m''$ respectively. We see that

$$\bar{u} = \psi_\ell^{-1} \psi_j(u)$$

$$\bar{v} = \psi_m''^{-1} g_{\ell j}(x) \cdot \psi_k''(v)$$

gives the operation $\psi_{\ell m}'^{-1} \psi_{jk}'$ since

$$(\bar{u}, \bar{v}) = \psi_{\ell m}'^{-1} \psi_{jk}'(u, v)$$

The first component $\psi_\ell^{-1} \psi_j(u)$ has derivatives up to order r . The second component $\psi_m''^{-1} (g_{\ell j}(x) \cdot \psi_k''(v))$ has the form $\psi_m''^{-1} f \psi_k''(v)$ where f is a composition of several maps each between differentiable manifolds. The map f is composed of $(V_\ell \cap V_j) \times Y \xrightarrow{a_{\ell j} \times 1} L_n \times Y \xrightarrow{h \times 1} G \times Y \xrightarrow{\cdot} Y$. The product of two differentiable manifolds of degree $r-1$ is again of degree $r-1$. The map $a_{\ell j}$ is of class $r-1$, so the product map $a_{\ell j} \times 1$ $[(a_{\ell j} \times 1)(x, y) = (a_{\ell j}(x), y)]$ is of class r . The Lie Group G is a differentiable manifold of type ω , and a homomorphism $h: L_n \rightarrow G$ is analytic: thus $h \times 1$ is analytic. The group action " $\cdot$ " takes $G \times Y \rightarrow Y$ and we assume that it is at least of class $r-1$. The combined maps f then are of class $r-1$ and f maps

$$(V_\ell \cap V_j) \times Y \rightarrow Y:$$

$$f(x, y) = h a_{\ell j}(x) \cdot y.$$

This proves $\psi_{\ell m}'^{-1} \psi_{jk}'$ has derivatives to order $r-1$.

The thorough study of manifolds and tensor bundles is somewhat orthogonal to the plan of these lectures. For those curious to learn more about applications of fibre bundles to manifolds, consult [Steenrod]. Below I list a few of the results he obtains about tensors using tensor bundles for base spaces X

which are differentiable manifolds.

1. Covariant and contravariant bundles are equivalent.
2. A differentiable manifold always admits a Riemannian metric.
3. Scalar bundles of even weight are equivalent to product bundles.
4. Any bundle over X with group L_n is equivalent in L_n to a bundle with group O_n (the orthogonal group in n -dimensions).

To learn more about manifolds one can study Steenrod, as well as consult texts on Differential Geometry (e.g. [Kobayashi and Nomizu], or the introductory chapter of [Hawking and Ellis]).

Bibliography

S. W. Hawking and G. F. R. Ellis, The Large Scale Structure of Space Time (Cambridge University Press, Cambridge 1973).

S. Kobayashi and K. Nomizu, Foundations of Differential Geometry, vol.1 (Interscience, New York, 1963).

N. Steenrod, The Topology of Fibre Bundles (Princeton University Press, Princeton, New Jersey 1951).

VI. COVERING HOMOTOPY THEOREMS

A. Bundle Homotopy

We have discussed the notion of homotopy equivalence of maps. (A long discussion is also given in chapter VII.) It is natural to extend this notion to fibre bundle structures. Recall that maps $\bar{h}_0$ and $\bar{h}_1$ from X to X' are homotopic if $\exists \bar{h}: X \times I \rightarrow X'$ such that $\bar{h}$ is continuous, $I = [0,1]$, $\bar{h}(x,0) = \bar{h}_0(x)$ and $\bar{h}(x,1) = \bar{h}_1(x)$. The extension will be to bundle maps h_0 and h_1 of B to B' . We must deform the bundle map h_0 into h_1 whenever the two maps are "bundle homotopic". Thus we must make clear the idea of "deform", as applied to bundle maps.

To start, one defines a bundle $B \times I$ which is the product of a coordinate bundle B with the unit interval $[0,1]$. Specifically, $B \times I$ has bundle space $B \times I$, base space $X \times I$, projection $q: q(b,t) = (p(b),t)$, fibre Y , group G , coordinate neighborhoods $V_j \times I$ and coordinate functions $\psi_j((x,t), y) = (\phi_j(x,y),t)$ all defined in terms of the elements of $B = (B, X, p, Y, G, V_j, \phi_j)$. Two maps $h_0, h_1: B \rightarrow B'$ are bundle homotopic if there exists a bundle homotopy $h: B \times I \rightarrow B'$ such that h on $B \times \{0\}$ agrees with h_0 and h on $B \times \{1\}$ agrees with h_1 .

This definition has several implications. For example, $h: B \times I \rightarrow B'$ is a homotopy of $h_0: B \rightarrow B'$ into $h_1: B \rightarrow B'$. This homotopy is fibre preserving so that $q^{-1}(x,t)$ is mapped homeomorphically onto some $p'^{-1}(x')$. In particular there is an induced map $\bar{h}: X \times I \rightarrow X'$ such that $p'h = \bar{h}q$, and $\bar{h}$ is a homotopy of the induced

FIG. 1 "Axes" of a coordinate bundle.

maps $\bar{h}_0$ and $\bar{h}_1$ of h_0 and h_1 respectively. By definition of a bundle map, the coordinate bundle $\bar{h}$ induces over $X \times I$ is strictly equivalent to $B \times I$. This induced bundle has the same bundle space, base space, projection, fibre and group as $B \times I$. It has the coordinate neighborhoods $\{\bar{h}^{-1}(V'_j)\}$ and coordinate functions $\rho_j((x,t), y) = h_{(x,t)}^{-1} \phi'_j(\bar{h}(x,t), y)$. Strict equivalence requires

$$\bar{g}_{jk}(x,t) = \phi'_{j,\bar{h}(x,t)} h_{(x,t)}^{-1} \psi_{k,(x,t)}$$

$$(x,t) \in \bar{h}^{-1}(V'_j) \cap (V_k \times I)$$

to coincide with an operation of G , and

$$\bar{g}_{jk} : \bar{h}^{-1}(V'_j) \cap (V_k \times I) \rightarrow G$$

is continuous.

This reminder of the general definitions emphasizes that bundle homotopic maps are those which are deformed into each

other by an operation which is continuous, and defines a bundle map at each stage of the deformation.

B. Covering Homotopy Theorem for Product Bundle

The simplest bundle structure to test these definitions on is the product bundle B' : $B' = X' \times Y$, $G = \{e\}$. Any bundle map h_0 from (an arbitrary bundle) B to a product bundle B' induces a map $\bar{h}_0$ on the base spaces. A question of some importance later is to ask whether there is a (bundle) homotopy h of h_0 , whenever one is given a homotopy $\bar{h}$ of $\bar{h}_0$. In our simple example, we can in fact construct such a bundle homotopy. The coordinate function p' is the identity map, so define

$$h(b,t) = (\bar{h}(x,t), p'_j h_0(b))$$

where $p'_j(b')$ is the projection on the fibre Y . This definition uses only h_0 and the given homotopy $\bar{h}$ of the induced map $\bar{h}_0$ of the base spaces. This mapping preserves fibres since all points $\Psi_k((x,t), y) = (\phi_k(x,y), t)$ in the fibre of (x,t) are mapped into $(\bar{h}(x,t), p'_j h_0 \phi_k(x,y))$, which is a point in the fibre of $\bar{h}(x,t)$. Furthermore, for fixed (x,t) , the mapping h is a homeomorphism since for fixed x , $y \mapsto p'_j h_0 \phi_k(x,y)$ is a homeomorphism of Y onto Y .

The mapping $h: B \times I \rightarrow B'$ defined above for a trivial bundle B' has another interesting property. Whenever for each $b \in B$, $\bar{h}(p(b), t)$ is constant for t in some closed interval of I , then since $p'_j h_0(b)$ is independent of t we see $h(b, t)$ is constant. This leads to a general definition: a bundle homotopy h is stationary with the induced homotopy $\bar{h}$ if whenever for each $b \in B$ and each interval $[t_1, t_2]$, $\bar{h}(p(b), t)$ is constant for all $t \in [t_1, t_2]$, then $h(b, t)$ is constant for all $t \in [t_1, t_2]$. Steenrod also says

in this case that h covers $\bar{h}$.

C. Covering Homotopy Theorem for X Compact

For product bundles B' we have a covering homotopy theorem, that whenever there is a bundle map $h_0: B \rightarrow B'$ and a homotopy $\bar{h}$ of the induced map $\bar{h}_0$, then there is a bundle homotopy h which covers $\bar{h}$. This theorem is in fact true under more general circumstances. Following Steenrod, we investigate first the case that B' is an arbitrary bundle, and B is a bundle such that its base X is compact (open coverings have a finite sub-covering, and X is Hausdorff). We are given $h_0: B \rightarrow B'$ a bundle map, and $\bar{h}: X \times I \rightarrow X'$ a homotopy of $\bar{h}_0: X \rightarrow X'$ which is induced by h_0 ($p'h_0 = \bar{h}_0 p$).

The construction of h depends upon the above observations concerning trivial bundles, and the property of coordinate bundles being locally a product.

1. Choosing the right cover

The open cover $\{\bar{h}^{-1}(V'_j)\}$ of $X \times I$ consists of open sets such as $\bar{h}^{-1}(V'_j)$, which by the definition of the product topology on $X \times I$, is the union of open sets $U_{\alpha j} \times I_{\beta j}$; $U_{\alpha j}$ is open in X and $I_{\beta j}$ is open in I . The base $X \times I$ is compact since X is compact (by assumption), I is compact, and one of the elementary theorems is that products of compact spaces are compact. The open covering $\{U_{\alpha j} \times I_{\beta j}\}$ over all α, β, j has therefore a finite subcover. From the sets $\{U_{\alpha_m j_m} \times I_{\beta_m j_m}\}$ of the subcover, one can refine $\{I_{\beta_m j_m}\}$ to a finite open cover $I_1, \dots, I_r$ of I such that I_1 intersects only I_2 , I_2 intersects only I_1 and $I_3, \dots$, I_{r-1} intersects I_{r-2} and I_1 , and I_r intersects only I_{r-1} .

Fig. 2 The solid 3 rectangles illustrate a given covering, the dotted line shows a refinement so that $X \times I$ is covered by $U_\lambda \times I_v$, $\lambda, v = 1, 2$.

The above figure illustrates the refinement process, and suggests that for our problem the cover $\{\bar{h}^{-1}(V'_j)\}$ can be refined to a finite open cover $\{U_\lambda \times I_v\}$ with I_v as above, $v = 1, \dots, r$ and $\{U_\lambda\}$ an open cover of X .

Because of the choice made for $\{I_v\}$, we can choose a set $t_v \in I_v \cap I_{v+1}$, $v = 1, \dots, r-1$ which satisfy

$$0 = t_0 < t_1 < \dots < t_{r-1} < t_r = 1.$$

[Of course $I_v \cap I_{v+1}$ cannot be an empty set since if it were, $[0, 1]$ would be the disjoint union of open sets, and hence disconnected; however, $[0, 1]$ is a connected interval.] The picture which results is shown in the figure below.

Fig. 1. Decomposition of $X \times I$.

Now for the bundle $B \times [0, t_0]$, the construction of a covering homotopy is trivial since we take h to be h_0 . Steenrod's proof proceeds by induction. Assume we have constructed a covering homotopy for $B \times [0, t_v]$, we must extend this homotopy to $B \times [t_v, t_{v+1}]$, which is the part of $B \times I$ over the base $X \times [t_v, t_{v+1}]$. Referring to the above figure, we assume the theorem for $B \times I$ over $X \times [0, t_v]$ and have to construct a covering homotopy only in the strip $X \times [t_v, t_{v+1}]$.

2. Digression on separation properties for compact X

The next part of the construction involves breaking the strip $X \times [t_v, t_{v+1}]$ into convenient parts. To do this we make use of some elementary theorems of topology, which we shall first present. Recall that a compact space X is defined to be Hausdorff (distinct points of X have disjoint neighborhoods), and every open covering of X has a finite subcovering. From [Bourbaki] we find:

a. Compact is regular

Thm: "Every compact space is regular".

Def: "A topological space X is regular if it is Hausdorff and if for any closed subset F of X and any point $x \notin F$ there is a neighborhood of x and a neighborhood of F which do not intersect."

Fig. 4 F is a closed subset of X , x is not in F , U is a neighborhood of x , V is a neighborhood of F and $U \cap V = \emptyset$.

The above figure illustrates the property regular spaces must have. Rephrasing the definition, for exterior points x of a closed set F , points sufficiently close to x are disjoint from points sufficiently close to F . We need one other theorem from the standard texts:

b. Compact is normal

Thm: Every compact space is normal.

Def: X is normal if it is Hausdorff and if for any two disjoint closed subsets of X , A and B , there exists disjoint open sets U , V such that $A \subset U$ and $B \subset V$.

Fig. 5 A and B are closed subsets of X , A is in U , B is in V for disjoint open sets U and V .

c. Urysohn's theorem

We note that Hausdorff, regular, and normal spaces characterize three different attributes of spaces we are used to dealing with (like the real line, or line intervals). These attributes have sufficiently important consequences, and arise frequently enough so that they have been given special names by the mathematicians. A discussion of these properties can be found in the appendix to this lecture, as well as in standard texts such as [Bourbaki], and [Pontyagin]. We quote the last result (proved in the Appendix) which we will need, which is Urysohn's Theorem: A Hausdorff space X is normal iff for any two disjoint closed subsets A and B of X , there exists a continuous mapping of X into $[0,1]$ which is equal to 0 at every point of A and 1 at every point of B .

3. Completing the choice of cover for 2-fold induction step

Let us now return to Steenrod's proof. We are assuming there is a covering homotopy h of $\bar{h}$ defined for $B \times [0, t_v]$. We want to extend this to a covering homotopy h of $\bar{h}$ over $B \times [t_v, t_{v+1}]$. Let

us consider the strip $X \times [t_v, t_{v+1}]$. Now each x lies in some U_λ . Since X is regular, $X \setminus U_\lambda = F$ is closed and $x \notin F$, then there exists disjoint neighborhoods of F and of x . Take the neighborhoods both to be open; call W the neighborhood around x , and A the neighborhood around F . Taking complements of F and A , $x \in W \subset X \setminus A \subset U_\lambda$. Since $X \setminus A$ is closed, the closure $\bar{W}$ is also in U_λ . Therefore, X being regular implies that for each x , and for the open set U_λ that contains x , there exists a set W such that $x \in W \subset \bar{W} \subset U_\lambda$.

Now X is also a normal space, and $\bar{W}$ and $F = X \setminus U_\lambda$ are two disjoint closed subsets of X . Therefore, there exists disjoint open sets W' and W'' such that $\bar{W} \subset W'$ and $F \subset W''$. Taking complements, we have $\bar{W} \subset W' \subset X \setminus W'' \subset U_\lambda$. Since $X \setminus W''$ is closed, the closure $\bar{W}' \subset U_\lambda$. We now have the result we want, namely for every $x \in X$, there is a pair of neighborhoods (W, W') such that

$$x \in W \subset \bar{W} \subset W' \subset \bar{W}' \subset U_\lambda.$$

The set of all such sets W of the pair covers X , so there exists a finite subcovering. Label the corresponding pairs (W_α, W'_α) , $\alpha = 1, \dots, s$.

Now according to Urysohn's theorem, there is a map $u_\alpha: X \rightarrow [t_v, t_{v+1}]$ such that $u_\alpha(\bar{W}_\alpha) = t_{v+1}$ and $u_\alpha(X \setminus W'_\alpha) = t_v$ since $\bar{W}_\alpha$ and $X \setminus W'_\alpha$ are disjoint closed subsets of a normal space X . This one uses to define an ordered sequence of functions $\tau_\alpha: X \rightarrow [t_v, t_{v+1}]$. Define $\tau_0(x) = t_v$, $\tau_1(x) = u_1(x)$, $\tau_2(x) = \text{Max}(u_1(x), u_2(x))$, ..., $\tau_{s-1}(x) = \text{Max}(u_1(x), \dots, u_{s-1}(x))$ and $\tau_s(x) = t_{v+1}$. See the figure below. By definition, for each $x \in X$ we have

$$\tau_0(x) \leq \tau_1(x) \leq \dots \leq \tau_s(x) .$$

Fig. 6 The construction of the τ functions.

The reason for introducing these functions is that they provide a systematic way of breaking the strip $X \times [t_v, t_{v+1}]$ into a finite number of pieces in order to construct the covering homotopy h . Let X_α be the set of points (x, t) in the strip such that $t_v \leq t \leq \tau_\alpha(x)$. In the figure, X_α consists of the points bounded by $\tau_\alpha(x)$ and t_v . Let S_α be the part of $S \times I$ over X_α . Then S_α form a nested sequence of bundles $S \times [t_v] = S_0 \subset S_1 \subset \dots \subset S_s = S \times [t_v, t_{v+1}]$. By definition, h is defined in S_0 . We extend h inductively to each of the S_α . So suppose we have constructed

an h covering $\bar{h}$ over $B_{\alpha-1}$. If we extend h over B_α we will have proved by induction that h is extended over $B \times [t_v, t_{v+1}] = B_s$.

4. Induction step

Now the difference between X_α and $X_{\alpha-1}$ consists of those points in X_α not in $X_{\alpha-1}$. Since we certainly don't add in X_α points outside of W'_α , we have $X_\alpha \setminus X_{\alpha-1} \subset W'_\alpha \times [t_v, t_{v+1}]$. By the construction of the pair (W_α, W'_α) , $W'_\alpha \subset U_\lambda$ for some U_λ ; also by construction $[t_v, t_{v+1}] \subset I_{v+1}$ and $U_\lambda \times I_{v+1} \subset \bar{h}^{-1}(V'_j)$ for some j . We are now in a position, finally, to extend h . We do this in a very similar way to the extension for a trivial bundle:

$$h(b, t) = \phi_j^* (\bar{h}(x, t), p_j^* h(b, \tau_{\alpha-1}(x)))$$

$$x = p(b)$$

$$(x, t) \in X_\alpha \setminus X_{\alpha-1}$$

This defines a continuous function of (b, t) . The closure of $X_\alpha \setminus X_{\alpha-1}$ is $X_\alpha \setminus \overset{\circ}{X}_{\alpha-1}$, where $\overset{\circ}{X}_{\alpha-1}$ is the interior of $X_{\alpha-1}$ and by definition X_α is closed. The interior of $X_{\alpha-1}$ is defined by the condition $t_v < t < \tau_{\alpha-1}(x)$. So the closure of $X_\alpha \setminus X_{\alpha-1}$ consists of the points $\tau_{\alpha-1}(x) \leq t \leq \tau_\alpha(x)$. The function defined above is defined in this closure, and when $t = \tau_{\alpha-1}(x)$, the definition of h agrees with $h(b, \tau_{\alpha-1}(x))$ which was assumed to be given on $B_{\alpha-1}$. The function h has therefore been extended from the bundle space $B_{\alpha-1}$ to B_α . The projection $p^* h(b, t) = \bar{h}q(b, t)$ so that h maps fibres into fibres, and the induced map coincides with $\bar{h}$. The variable t does not occur in the y -coordinate of ϕ_j^* , so if $\bar{h}(x, t) =$

$\bar{h}(x, t')$ for $\tau_{\alpha-1}(x) \leq t < t' \leq \tau_{\alpha}(x)$, then $h(b, t) = h(b, t')$. In other words h is stationary with $\bar{h}$.

Finally, we have to verify that the h so defined is a bundle map; it remains to show that the induced bundle by $\bar{h}$ on the closure of $B_{\alpha} \setminus B_{\alpha-1}$ is strictly equivalent to the given bundle structure. To do this we must consider any point $(x, t) \in \overline{B_{\alpha} \setminus B_{\alpha-1}}$, and suppose $x \in V_k$. Using the notation $\tau = \tau_{\alpha-1}(x)$, we must show

$$\bar{g}_{jk}(x, t) = p_j^! h(x, t) \psi_k(x, t)$$

coincides with an operation of G , and $\bar{g}_{jk}$ is a continuous map into G . We compute the effect of $\bar{g}_{jk}(x, t)$ on y :

$$\begin{aligned} \bar{g}_{jk}(x, t) \cdot y &= p_j^! h(x, t) \psi_k((x, t), y) \\ &= p_j^! h(\phi_k(x, y), t) \\ &= p_j^! \phi_j^! (\bar{h}(x, t), p_j^! h(\phi_k(x, y), \tau)) \\ &= p_j^! h(\phi_k(x, y), \tau) \\ &= p_j^! h(x, \tau) \psi_k((x, \tau), y) \\ &= \bar{g}_{jk}(x, \tau) \cdot y \\ &= \bar{g}_{jk}(x, \tau_{\alpha-1}(x)) \cdot y \end{aligned}$$

By assumption, h is a bundle map over $B_{\alpha-1}$ so $\bar{g}_{jk}(x, \tau)$ is a continuous map into G of $V_k \times I \cap (\text{boundary of } X_{\alpha} \setminus X_{\alpha-1})$. The points $(x, \tau_{\alpha-1}(x))$ are in the boundary of $X_{\alpha} \setminus X_{\alpha-1}$. Thus $\bar{g}_{jk}(x, t)$ is a

continuous map into G of $V_k \times I \cap (\overline{X_\alpha \setminus X_{\alpha-1}})$. This proves h is a bundle map.

Now we have extended a covering homotopy h of $\bar{h}$ from $B_{\alpha-1}$ to B_α . This proves by induction that h can be extended to $B \times [t_v, t_{v+1}]$, and so again by induction it shows that h can be extended to all of $B \times I$. The theorem is now proved: For a bundle map $h_0: B \rightarrow B'$, and a homotopy $\bar{h}$ of $\bar{h}_0: X \rightarrow X'$, when X is compact there is a covering homotopy h of $\bar{h}$.

D. Covering Homotopy Theorem for Non-Compact X

For some purposes, one needs the covering homotopy theorem for non-compact base spaces. There are two generalizations on the market, one of which is given in [Steenrod] and will be reproduced here, the other is in the literature [See W. Huebsch]. As before, we assume as given two coordinate bundles B and B' , a bundle map $h_0: B \rightarrow B'$, and a homotopy $\bar{h}: X \times I \rightarrow X'$ of the induced map $\bar{h}_0: X \rightarrow X'$. In this proof, one imposes the restriction on X that it be a normal space, locally compact and have the property that every open covering of X has a countable open subcover. [Huebsch assumes only that X is paracompact.]

The result we have to demonstrate is the existence of a (bundle) homotopy h of h_0 which covers $\bar{h}$. That is, $h: B \times I \rightarrow B'$ is a bundle map; $h(b, 0) = h_0(b)$ for all $b \in B$; and for each $b \in B$ and each interval $[t_1, t_2] \subset I$ where $\bar{h}(x, t)$ is constant for $t \in [t_1, t_2]$ and $x = p(b)$, then $h(b, t)$ is likewise constant. The idea of the proof is to make use of the theorem already proved for a

compact base space X by breaking our given space X up into a sequence of compact subspaces.

1. Choosing the right cover

Our first task is to see that a locally compact space X with the property that open covers have countable subcovers, is the union of open sets $\{W_n\}$ such that the closure $\bar{W}_n$ is compact and $\bar{W}_n \subset W_{n+1}$ for each integer n . By definition, a locally compact space X is Hausdorff and is one in which every point $x \in X$ has a compact neighborhood V . Now if V is compact it is necessarily a closed subspace of the Hausdorff space X . For suppose $z \notin V$, then each $y \in V$ is distinct from z and there exist two families of open sets W_y and U_y such that $y \in W_y$, $z \in U_y$ and $W_y \cap U_y = \emptyset$ for each $y \in V$. The sets $\{W_y\}$ give V an open covering, so there exists a finite subcovering $W_1 \dots W_m$. Form $W = W_1 \cup \dots \cup W_m$, and with the corresponding $U_1, \dots, U_m$ form $U = U_1 \cap \dots \cap U_m$. By construction U and W are disjoint open sets, $z \in U$ and $V \subset W$. This shows that every point $z \in X \setminus V$ has an open neighborhood $U \subset X \setminus V$, which implies $X \setminus V$ is open; thus V is a closed set.

Now for each point $x \in X$, we have a closed, compact neighborhood V of x ; let U_x be the interior of V (the largest open set contained in V) so that $\bar{U}_x$ is thus closed and compact (since $\bar{U}_x = V$ by the definitions of interior and closure). The family $\{U_x\}_{x \in X}$ provides an open covering of X ; we now invoke the property of X that the open covering $\{U_x\}$ have a countable open subcovering $U_1, U_2, \dots$. By construction each $\bar{U}_n$ is compact. We build a new countable open covering as follows. Define $W_1 = U_1$; since $\bar{W}_1$ is compact, and the $\{U_n\}$ cover $\bar{W}_1$ (they in fact cover all of X), there are a finite number of open sets $U_{k_1}, \dots, U_{k_m}$ which cover $\bar{W}_1$.

Define $W_2 = (\bigcup_{k=1}^m U_{k_1}) \cup (U_1 \cup U_2)$ and note W_2 is open, $W_1 \subset W_2$ and $\bar{W}_2 \subset (\bigcup_{k=1}^m \bar{U}_{k_1}) \cup \bar{U}_1 \cup \bar{U}_2$. Since a finite union of closed (compact) sets is closed (compact), $\bar{W}_2$ is a closed subset of a compact space, hence W_2 is itself compact. One now proceeds by induction; given sets $W_1 \subset \bar{W}_1 \subset W_2 \subset \bar{W}_2 \dots \subset W_n \subset \bar{W}_n$ where $\bar{W}_i, i=1, \dots, n$ are compact a similar argument to the one above shows the existence of a set W_{n+1} such that $\bar{W}_n \subset W_{n+1}$ and W_{n+1} is compact. Moreover, the above construction suggests putting $W_{n+1} = (\bigcup_{m=1}^{n+1} U_m) \cup (\bigcup_{j=1}^r U_{k_j})$, a finite union of open sets. This definition insures W_{n+1} is compact, and that $\bigcup_{n=1}^{\infty} W_n$ covers the whole space. We have now verified that the property of a locally compact (and hence Hausdorff) space in which open covers have countable subcovers is that there exist a countable open covering $\{W_n\}$ of X such that the $\{\bar{W}_n\}$ are compact and $\bar{W}_n \subset W_{n+1}$.

2. Completing the choice of cover for induction step

Now to prove the covering homotopy theorem we want to use the expanding sequence of open sets W_n to break up X into successively larger compact sets, for on these sets the theorem has already been proved. This is done using the property that X is normal; since $\bar{W}_n$ and $X \setminus W_{n+1}$ are disjoint closed sets of X , by Urysohn's theorem there exists a continuous function $\tau_n: X \rightarrow [0, 1]$

Fig. 7 The τ_n function.

such that $\tau_n(X \setminus W_{n+1}) = 0$ and $\tau_n(W_n) = 1$. This defines a countable family of continuous functions $\tau_1, \tau_2, \dots$. One defines also a function τ_0 by $\tau_0(x) = 0, \forall x \in X$. Since the sequence of sets W_n is expanding, $\tau_n(x) \leq \tau_{n+1}(x), \forall x \in X, n = 0, 1, \dots$. Also since the $\{W_n\}$ cover X , for each x there is some n such that $x \in W_n$, hence for this n , $\tau_n(x) = 1$. One uses these τ_n functions to define on $X \times I$ an expanding family of sets X_n defined by points (x, t) where $0 \leq t \leq \tau_n(x)$. Over each of these sets X_n , one denotes by B_n the part of the bundle space whose projection gives X_n ; the corresponding coordinate bundle is called $\mathcal{B}_n$. The $\{X_n\}$ cover $X \times I$, so $\{B_n\}$ cover $B \times I$ and hence $\{\mathcal{B}_n\}$ cover $\mathcal{B} \times I$.

3. Induction step

There is now sufficient machinery to construct inductively a covering homotopy h of h_0 . The bundle $\mathcal{B}_0$ consists of $\mathcal{B} \times \{0\}$ and so the covering homotopy in this bundle is the given bundle map h_0 . Now suppose we have constructed a bundle homotopy h of h_0 covering $\bar{h}$ in the bundle $\mathcal{B}_{n-1}$. If we can extend this to points of $\mathcal{B}_n$, by induction we will have shown that h exists in every $\mathcal{B}_n$.

$n=0,1,\dots$. We then have to show this defines a covering homotopy over all of $B \times I$.

The set A

The points (x,t) of $X \times I$ which are in X_n but not X_{n-1} satisfy $\tau_{n-1} < t \leq \tau_n(x)$. So we start by considering points of X which satisfy $\tau_{n-1}(x) < \tau_n(x)$. Let us define the closed set A as the intersection of two closed sets, $\bar{W}_{n+1}$ and $X \setminus W_{n-1}$; A is a closed subset of a compact space $\bar{W}_{n+1}$, hence A is compact. For each $x \in X$, we saw generally that $\tau_{n-1}(x) \leq \tau_n(x)$. To show all points of x such that $\tau_{n-1}(x) < \tau_n(x)$ lie in A , let us break X up into three disjoint regions, the closed set $\bar{W}_{n-1}$, the open set $W_{n+1} \setminus \bar{W}_{n-1} = (X \setminus \bar{W}_{n-1}) \cap W_{n+1}$ and the closed set $X \setminus W_{n+1}$.

Fig.8 Regions of X .

We use the fact that $\bar{W}_{n-1} \subset W_n \subset \bar{W}_n \subset W_{n+1}$. Note that X is the union of $X \setminus W_{n+1}$ and W_{n+1} , and W_{n+1} is the union of $\bar{W}_{n-1}$ and $W_{n+1} \setminus \bar{W}_{n-1}$. For each $x \in \bar{W}_{n-1}$, since $\bar{W}_{n-1} \subset \bar{W}_n$, $\tau_{n-1}(x) = \tau_n(x) = 1$. For each $x \in X \setminus W_{n+1}$, since $X \setminus W_{n+1} \subset X \setminus W_n$, $\tau_{n-1}(x) = \tau_n(x) = 0$. For each $x \in W_{n+1} \setminus \bar{W}_{n-1} \subset \bar{W}_{n+1} \setminus W_{n-1} = A$, $\tau_{n-1}(x) \leq \tau_n(x)$. Thus all points $x \in X$ such that $\tau_{n-1}(x) < \tau_n(x)$ are in A .

Now we want to apply the covering homotopy theorem to the bundle lying over A . The theorem assumes the closed interval on the real line is $[0,1]$, yet for A , the natural interval is $[\tau_{n-1}(x), \tau_n(x)]$. For each $x \in A$, we can map $[0,1] \rightarrow [\tau_{n-1}(x), \tau_n(x)]$ by the linear map

$$\lambda_x(s) = s\tau_n(x) + (1-s)\tau_{n-1}(x).$$

Inverting this equation, we get a linear map $[\tau_{n-1}(x), \tau_n(x)] \rightarrow [0,1]$ defined by

$$\lambda_x^{-1}(t) = \frac{t - \tau_{n-1}(x)}{\tau_n(x) - \tau_{n-1}(x)}$$

for every $x \in A$ such that $\tau_{n-1}(x) < \tau_n(x)$.

Applying covering theorems to bundle over A

We are assuming that there exists a covering homotopy h of h_0 defined on B_{n-1} . So for each $b \in B$, $p(b) \in A$, $t \in [0,1]$, h is defined so long as $0 \leq t \leq \tau_{n-1}(p(b))$. In particular, h defines a bundle map h'_0 for the bundle lying over A :

$$h'_0(b) = h(b, \tau_{n-1}(p(b)))$$

$$p(b) \in A$$

The induced map $\bar{h}'_0: A \rightarrow X'$ is

$$\bar{h}'_0(x) = \bar{h}(x, \tau_{n-1}(x))$$

$$x \in A$$

The linear map $\lambda_x(s)$ can be used to define a homotopy $\bar{h}': A \times I \rightarrow X'$ of $\bar{h}'_0$, namely

$$\bar{h}'(x, s) = \bar{h}(x, \lambda_x(s))$$

$$x \in A$$

$$s \in [0, 1]$$

Since A is compact, the covering homotopy theorem provides a bundle homotopy h' of h'_0 which covers $\bar{h}'$. Next, we would like to use $h'(b, s)$ to define a homotopy h on points (b, t) where $x = p(b) \in A$ and $\tau_{n-1}(x) \leq t \leq \tau_n(x)$.

The homotopy h

For all points (b, t) such that $\tau_{n-1}(x) < t \leq \tau_n(x)$, $x = p(b)$, (and from above, x must be in A), we can use $\lambda_x^{-1}(t)$ to define the function

$$h(b, t) = h'(b, \lambda_x^{-1}(t)),$$

since nowhere will λ_x^{-1} be singular. The function h is also defined for $t = \tau_{n-1}(x)$, having the value $h'(b, 0) = h'_0(b) = h(b, \tau_{n-1}(x))$, thereby providing an extension of h into the region $\tau_{n-1}(x) < t \leq \tau_n(x)$, that is points of B_n not in B_{n-1} . Remember that points $0 \leq t \leq \tau_n(x)$ define points (x, t) in X_n , which are covered by B_n . This extension of h provides a continuous well defined function except at points such as (b, t) where $t = \tau_{n-1}(x) = \tau_n(x)$, $x = p(b)$, since $\lambda_x^{-1}(t)$ is not defined there (q.v. definition of $\lambda_x^{-1}(t)$).

Points where λ_x^{-1} is singular

For each x , whenever $\tau_{n-1}(x) = \tau_n(x)$, by its definition $\lambda_x(s)$ is constant with value $\tau_{n-1}(x)$; therefore $\bar{h}'(x, s) = \bar{h}(x, \tau_{n-1}(x))$ is stationary. Since h' covers $\bar{h}'$, $h'(b, s)$ is also stationary. We want to prove $h(b, t)$ can be defined at (b, t) in such a way that h is continuous at this point. One chooses the value of $h(b, t)$ to be the constant given by $h'(b, s)$. We now check continuity at (b, t) .

Let N' be any neighborhood of $h(b, t)$, which is to say N' is a neighborhood of $h'(b, s)$. Every point of $b \times I$ maps into the point $h'(b, s)$, so $h'^{-1}(N')$ contains $b \times I$. Now b (that is, the subset $\{b\}$) is a compact subset of B since obviously every covering of b has a subcovering consisting of a single open set. Since I is compact, so therefore is the product $b \times I$. For each $s \in I$, h' is continuous at (b, s) so there exists an open neighborhood $U_s \times V_s$ such that $h'(U_s \times V_s) \subset N'$. These neighborhoods $U_s \times V_s$ cover $b \times I$, so there exists a finite subcover $U_1 \times V_1, \dots, U_m \times V_m$ covering $b \times I$. It therefore follows that $N = \bigcap_{k=1}^m U_k$ is a neighborhood of b , $\bigcup_{k=1}^m V_k = I$, $N \times I$ is a neighborhood of $b \times I$, and $h'(N \times I) \subset N'$. Now take some $b_1 \in N$, $x_1 = p(b_1)$, and any t_1 such that $\tau_{n-1}(x_1) < t_1 \leq \tau_n(x_1)$. Consequently $\lambda_{x_1}^{-1}(t_1)$ is well defined, $(b_1, \lambda_{x_1}^{-1}(t_1)) \in N \times I$, and $h'(b_1, \lambda_{x_1}^{-1}(t_1)) = h(b_1, t_1)$ is in N' . To summarize then, for a point (b, t) , and any neighborhood N' of $h(b, t)$ in which $t = \tau_{n-1}(x) = \tau_n(x)$, there exists a neighborhood of points about (b, t) whose image under h is contained in N' . These considerations show $h(b, t)$ is continuous for $x = p(b)$ in A and $\tau_{n-1}(x) \leq t \leq \tau_n(x)$.

h is a bundle map

We need now to check that h is a bundle map. First note that $p'h(b,t) = p'h'(b, \lambda_x^{-1}(t)) = \bar{h}'(p(b), \lambda_x^{-1}(t)) = \bar{h}(p(b), t)$. This shows h preserves fibres. To check h is stationary, suppose b is fixed, $x = p(b)$, and some interval in I is fixed so that $\bar{h}(x, t)$ is constant in this interval. This implies $\bar{h}'(x, \lambda_x^{-1}(t))$ is constant and hence $h'(b, \lambda_x^{-1}(t))$ is constant since h' covers $\bar{h}'$. By definition then, $h(b, t) = h'(b, \lambda_x^{-1}(t))$ is also constant. To see that h is a bundle map, we have to verify as usual that the bundle over A is strictly equivalent to the bundle induced by $\bar{h}^{-1}$. Let $x \in V_j$, $\tau_{n-1}(x) \leq t \leq \tau_n(x)$ and $\bar{h}(x, t) \in V'_k$. We have to verify the properties of

$$\bar{g}_{kj}(x, t) = \phi_{k,x}'^{-1, h}(x, t) \psi_j(x, t) .$$

Evaluating $\bar{g}_{kj}(x, t)$ on y , we see

$$\begin{aligned} \bar{g}_{kj}(x, t) \cdot y &= \phi_{k,x}'^{-1, h}(x, t) \psi_j((x, t), y) \\ &= \phi_{k,x}'^{-1, h}(\phi_j(x, y), t) \\ &= \phi_{k,x}'^{-1, h'}(\phi_j(x, y), \lambda_x^{-1}(t)) \\ &= \phi_{k,x}'^{-1, h'}(x, s) \psi_j(x, s) \cdot y \\ &= \bar{g}_{kj}'(x, s) \cdot y \\ s &= \lambda_x^{-1}(t) , \end{aligned}$$

so that

$$\bar{g}_{kj}(x,t) = \bar{g}'_{kj}(x, \lambda_x^{-1}(t)) .$$

Here $\bar{g}'_{kj}$ is the function associated with the bundle map h' . One is supposed to note that $\bar{g}_{kj}(x,t)$ acts like an element of G and is necessarily continuous at all points (x,t) except where $\tau_{n-1}(x) = t = \tau_n(x)$. Just as one argued before, continuity at this point can also be established.

One has therefore extended h to a covering homotopy of h_0 over all points (b,t) such that $x = p(b) \in A$ and $\tau_{n-1}(x) \leq t \leq \tau_n(x)$. To see this is an extension of h to B_n , consider points $(b,t) \in B_n \setminus B_{n-1}$. These points satisfy $\tau_{n-1}(x) < t \leq \tau_n(x)$ for $x = p(b)$. Now we showed that whenever $\tau_{n-1}(x) < \tau_n(x)$ then $x \in A$. Thus there is an extension of h to B_n . By induction this is true for every n , $n = 0, 1, 2, \dots$.

4. Loose ends

Therefore by the induction principle in each bundle $\mathcal{B}_n$ there exists a covering homotopy of h_0 which is an extension of the covering homotopy of h_0 on $\mathcal{B}_{n-1}$. The next and last point to settle is whether these covering homotopies provide a covering homotopy of h_0 on $\mathcal{B} \times I$. The first observation which we need

which will lead to the desired result is that B_{n-1} lies in the interior of B_n . We prove this observation as follows.

For this part of the proof, we lose none of the generality by slightly redefining our sets X_n to be the points (x,t) such that $0 \leq t \leq \tau_n(x)$ and $x \in W_{n+1}$ for each $n = 0, 1, 2, \dots$. We see in the

figure that X_0 lies in the "square" $W_1 \times I$. This obviously gen-

FIG. 9 $X_0 = W_1 \times 0$ and X_1 is the shaded region.

eralizes to the statement that X_{n-1} lies in the "square" $W_n \times I$. If we show the "square" is open then we will have shown $X_{n-1} \subset^\circ X_n$ for each $n=1,2,\dots$. But note W_n is open in X and I is open in I , so by the product topology $W_n \times I$ is open in $X \times I$. So we have the result that $X_{n-1} \subset^\circ X_n$, and our modified sets X_n still cover $X \times I$.

Given the result for X_n , we consider now the bundle spaces B_n which result from taking the inverse image of X_n using the bundle projection $q: B_n = q^{-1}(X_n)$. Now if $X_{n-1} \subset^\circ X_n$ then $B_{n-1} \subset q^{-1}(X_n)$. Moreover $q^{-1}(X_n)$ is an open set of B_n , hence $q^{-1}(X_n) \subset^\circ B_n$. Thus $B_{n-1} \subset^\circ B_n$ is obtained for $n=1,2,\dots$.

Now the sets B_n cover $B \times I$. For each set, $B_n \subset^\circ B_{n+1}$ (B_{n+1} is the interior of B_{n+1}). Thus $\{B_n\}$ also covers $B \times I$. Moreover,

h is defined on each B_n and is continuous there. Let V' be any closed set of B' ; we want to show $V = h^{-1}(V')$ is closed and thus h is continuous on $B \times I$. Now for each set B_n , since h is continuous on this set, $V \cap B_n$ is closed with respect to B_n . Another way to put it is that $B_n \setminus (V \cap B_n) = C \cap B_n$ is open in B_n , where $C = (B \times I) \setminus V$.

Fig. 10 Sketch demonstrates $B_n \setminus (V \cap B_n) = C \cap B_n$.

By definition this means $C \cap B_n$ is the intersection of some open set U of $B \times I$ and B_n . Another open set is $U \cap \overset{\circ}{B}_n$, and this set intersected with $\overset{\circ}{B}_n$ is the same as $C \cap \overset{\circ}{B}_n$ showing that $C \cap \overset{\circ}{B}_n = U \cap \overset{\circ}{B}_n$ is open in $\overset{\circ}{B}_n$. But by its construction $U \cap \overset{\circ}{B}_n$ is necessarily an open set in $B \times I$. This is true for each n , so using the fact that $\{\overset{\circ}{B}_n\}$ cover $B \times I$ we have that $C = \bigcup_{n=1}^{\infty} C \cap \overset{\circ}{B}_n$ is an open set. We defined C to be $(B \times I) \setminus V$, so this implies V is indeed a closed set.

We thus have shown that h is extended continuously over all of $B \times I$. It is fibre preserving and a stationary map with respect to $\bar{h}$ by its construction. To demonstrate it is a bundle map we note $\bar{g}_{kj}(x, t)$ is defined over all of $V = (V_j \times I) \cap \bar{h}^{-1}(V_k)$ and is by construction an element of G . Its continuity is argued in the same way as for h , by noting that the interiors of $V \cap X_n$ cover V , and $\bar{g}_{kj}$ is defined and continuous on each $V \cap X_n$. The previous argument showed in such an instance that $\bar{g}_{kj}$ is therefore continuous over all of V . This completes the proof of the theorem.

E. Appendix on Digression on Separation Properties

1. Hausdorff

We have had occasion to mention the Hausdorff property: spaces X in which every pair of distinct points have disjoint neighborhoods. In the figure, points sufficiently close to x

Fig. 11 Hausdorff property.

(i.e. the points of U) are disjoint from points sufficiently close to y (i.e. the points of V). In addition to the Hausdorff property, spaces which we will be interested in have a number of other, rather similar properties.

2. Regular

Consider a closed set F and a point x not in F . If points sufficiently close to x are disjoint from

Fig. 12 Regular property.

points sufficiently close to F for any closed set F and any point x not in F , then the space is called regular [we will in addition require X to be Hausdorff when we call X regular, following [Bourbaki's] notations].

3. Normal

Finally a normal space X is one which is Hausdorff and for every pair of disjoint closed sets A and B in X , points sufficiently close to A are disjoint from points sufficiently close to B . These separation properties are not equivalent; although a normal space is regular and a regular space is Hausdorff, there are Hausdorff spaces which are not regular and regular spaces not normal.

Fig. 13 Normal property

4. Proof that compact is Hausdorff, regular and normal

There are common situations in which X has all three properties, for example when X is compact. We shall demonstrate how this comes about. First, in our definition of compact, we require X to be Hausdorff, and for every open covering of X to have a finite subcovering. We wish to show first X is regular, then that X is normal. For the proof we follow [Pontryagin].

Let X be compact, F any closed set in X and x a point in X which is not in F . For each point $y \in F$, since X is assumed Hausdorff, there exist disjoint open neighborhoods U_y of x and V_y of y . Now the set of all V_y forms an open covering of F . Moreover, if $X \setminus F$ is added to this covering, we have an open covering of X . There is thus a finite subcovering of X , and in particular a finite subcovering of F by sets of $\{V_y\}$ (showing generally that closed subsets of compact spaces are compact).

Call the sets of the subcovering $V_{y_1}, \dots, V_{y_n}$ and the corresponding open neighborhoods of x , $U_{y_1}, \dots, U_{y_n}$. The union $\bigcup_{k=1}^n V_{y_k} = V$ is open and contains F ; the intersection $\bigcap_{k=1}^n U_{y_k} = U$ is open and contains x . Any point in common with both U and V would have to be in every U_{y_k} and some V_{y_k} , so in particular it would be in U_{y_k} and V_{y_k} ; these sets have no points in common by construction, so U and V are disjoint. This shows X is regular.

Now we show X is normal. Let A and B be two disjoint closed sets. For each $x \in A$, the regularity property implies the existence of disjoint open sets U_x and V_x such that $x \in U_x$ and $B \subset V_x$. Since $\{U_x\}$ is an open covering of A , and A is compact (since it is closed), there is a finite open subcovering $U_{x_1}, \dots, U_{x_m}$ of A with the corresponding disjoint sets $V_{x_1}, \dots, V_{x_m}$ containing B . So denoting $U = \bigcup_{k=1}^m U_{x_k}$ and $V = \bigcap_{k=1}^m V_{x_k}$ we have disjoint open sets such that $A \subset U$ and $B \subset V$. This shows X is normal.

5. Proof that metric spaces are normal

Another class of spaces which are normal are metric spaces. (Again, the argument is from [Pontryagin].) These spaces have defined on them a distance function d , which serves to define

the open sets. For each x , the sets U_x such that $d(x,y) < \epsilon$ for each real number ϵ are the open sets containing x . Let A and B be two disjoint closed sets, and $x \in A$ and $y \in B$ two distinct points. Call U_x a sphere whose center is at x and whose radius is $\frac{1}{2}d(x,B)$, where the distance from a point x to a closed subset B is the least of the distances $d(x,z)$ over all points $z \in B$. Call V_y the sphere whose center is at y and whose radius is $\frac{1}{2}d(y,A)$. Let U

Fig. 14 Sets U_x and V_y

be the union of all sets U_x (which are an open covering of A) and V be the union of all sets V_y (an open covering of B). Necessarily $A \subset U$ and $B \subset V$ and U and V are open sets. We need to show U and V are disjoint. Let $z \in U \cap V$, then there must be some U_{x_0} and V_{y_0} such that $z \in U_{x_0} \cap V_{y_0}$. The points x of U_{x_0} satisfy $d(x, B) < \frac{1}{2}d(x_0, B)$, and the points y of V_{y_0} satisfy $d(y, A) < \frac{1}{2}d(y_0, A)$.

$$R = \frac{1}{2}d(A,B)$$

Fig.15 Minimal separation of U_{x_0} and V_{y_0} .

Even if x_0 were as close to B as possible, and y_0 were as close to A as possible, the distance between x_0 and y_0 would be $d(A,B)$, and U_{x_0} would consist of points less than $\frac{1}{2}d(A,B)$ in distance to x_0 . V_{y_0} would consist of points less than $\frac{1}{2}d(A,B)$ in distance to y_0 . Thus these two sets can have no points in common. Therefore, if X is a metric space, X is normal.

6. Proof of Urysohn's theorem

The last point concerning normal spaces is Urysohn's theorem. For a normal space X , given any two disjoint closed subspaces A and B , there exists a continuous map $f: X \rightarrow [0,1]$ which is equal to 0 on A and 1 on B . The proof follows [Bourbaki, vol.2].

Fig.16 Urysohn's function f which has $f(A) = 0$ and $f(B) = 1$.

One first notices that a Hausdorff space X is normal iff for any closed set A and any open neighborhood V of A , there exists an open neighborhood W of A such that $\bar{W} \subset V$. For suppose X is normal, A is closed, V is open and $A \subset V$. Then A is disjoint from $X \setminus V$ a closed set. Thus there exist two open disjoint sets U and W such that $X \setminus V \subset U$ and $A \subset W$. Since U and W are disjoint, $W \subset X \setminus U$, also $X \setminus U \subset V$ so $A \subset W \subset X \setminus U \subset V$. Since $X \setminus U$ is closed, the closure of W is also contained in $X \setminus U$, which implies $A \subset W \subset \bar{W} \subset V$. This was what we set out to prove. Conversely suppose for every closed set A , any open neighborhood V of A there exist an open set W such that $A \subset W \subset \bar{W} \subset V$. Let A and B be any two disjoint closed sets. It follows that A is in $X \setminus B$, and $X \setminus B$ is an open neighborhood of A . There exists then an open set U such that $A \subset U \subset \bar{U} \subset X \setminus B$. We have $A \subset U$, $B \subset X \setminus \bar{U}$ and U and $X \setminus \bar{U}$ are disjoint open sets. Hence X is normal.

We now proceed to the construction of a continuous function $f: X \rightarrow [0,1]$ such that $f(A) = 0$ and $f(B) = 1$, when we are given two disjoint closed sets A and B . Now suppose first that we had some continuous function $g: X \rightarrow [-1,1]$ which is equal to -1 on A and $+1$ on B . We will show this defines a characteristic set of open sets. Let $U(t) = g^{-1}[-1, t)$, the inverse image of an open set (in $[-1,1]$) for each $t \in [0,1]$. This is a family of sets satisfying three properties: (i) $A \subset U(0)$, since $g^{-1}(-1) = A$; (ii) $B \subset X \setminus U(1)$, since if $x \in U(1) \cap B$ then $g(x) = 1$ (for $x \in B$), and $g(x) < 1$ since $x \in U(1)$ and $U(1)$ are points whose image under g are in $[-1,1)$, thus $U(1) \cap B = \emptyset$ or $B \subset X \setminus U(1)$; and (iii) for each pair of real numbers t, t' such that $0 \leq t < t' \leq 1$, then $\bar{U}(t) \subset U(t')$ since

the points of $U(t)$ are certainly in the closed set $g^{-1}([-1, t])$ all of whose points are in $g^{-1}([-1, t'])$. The useful point is that if we are given such a family of sets $U(t)$ satisfying these three properties, then there is a continuous function $f: X \rightarrow [0, 1]$ satisfying $f(A) = 0$ and $f(B) = 1$.

Define such a function f as follows. For every point $x \in X$ put $f(x) = 1$ whenever $x \in X \setminus U(1)$, and if $x \in U(1)$ let $f(x)$ be the greatest lower bound of values of t such that $x \in U(t)$. Since t ranges between 0 and 1, $f(x)$ ranges between 0 and 1, and $f(A) = 0$ and $f(B) = 1$. To show f is continuous at a point a , consider every open set $(a - \epsilon, a + \epsilon)$ where ϵ is any positive number. The points of X which map into this interval are in $f^{-1}[0, a + \epsilon) = U(a + \epsilon)$ as well as in the points of X not in the closure of $f^{-1}[0, a - \epsilon)$, $X \setminus \bar{U}(a - \epsilon)$; thus $f^{-1}(a - \epsilon, a + \epsilon) = U(a + \epsilon) \cap (X \setminus \bar{U}(a - \epsilon))$, which is an open set. Since any open set will be the union of sets of type $(a - \epsilon, a + \epsilon)$ for some $0 \leq a \leq 1$ and some positive ϵ , we have the result that the inverse image of any open set is open, hence f is continuous. Thus to prove Urysohn's theorem, we must construct a family of open sets $\{U(t) \mid t \in [0, 1]\}$ satisfying the above three properties.

To start the construction, take $U(1) = X \setminus B$. Now A is a closed set contained in an open set $U(1)$, so that the normality of X implies the existence of an open set $U(0)$ such that $A \subset U(0) \subset \bar{U}(0) \subset U(1)$. The idea is that given $U(0)$ and $U(1)$, we will construct a set $U(\frac{1}{2})$ so that $U(0)$, $U(\frac{1}{2})$ and $U(1)$ satisfy the three properties. Then given these three sets, we can pick sets $U(\frac{1}{4})$ and $U(\frac{3}{4})$ so the new collection satisfies all three properties. This process continues, so we proceed by induction and suppose

we are given open sets $U(0), U\left(\frac{1}{2^n}\right), \dots, U\left(\frac{2^n-1}{2^n}\right), U(1)$ such that $\bar{U}(U) \subset U\left(\frac{1}{2^n}\right), \bar{U}\left(\frac{1}{2^n}\right) \subset U\left(\frac{2}{2^n}\right), \dots, \bar{U}\left(\frac{2^n-1}{2^n}\right) \subset U(1)$. Thus for each $k, 0 \leq k \leq 2^n-1$, there is a closed set $\bar{U}\left(\frac{k}{2^n}\right)$ contained in an open set $U\left(\frac{k+1}{2^n}\right)$ so by the normality of X there exists an open set $V\left(\frac{k}{2^n} + \frac{1}{2^{n+1}}\right)$ such that

$$\bar{U}\left(\frac{k}{2^n}\right) \subset V\left(\frac{k}{2^n} + \frac{1}{2^{n+1}}\right) \subset \bar{V}\left(\frac{k}{2^n} + \frac{1}{2^{n+1}}\right) \subset U\left(\frac{k+1}{2^n}\right).$$

This proves by induction that for every (dyadic) number r of the form $k/2^n$ for any n and any k such that $0 \leq k \leq 2^n-1$, there exists an open set $U(r)$. By construction the sets $\{U(r)\}$ satisfy the properties $A \subset U(0)$, $B \subset X \setminus U(1)$ and for two dyadic numbers r and r' such that $0 \leq r < r' \leq 1$ then

$$\bar{U}(r) \subset U(r').$$

We now extend this family to one over all real numbers in $[0,1]$.

Define the open set $U(t) = \bigcup_{r \leq t} U(r)$ where the union is over any dyadic number less than or equal to t . If t is a dyadic number r_0 , $U(t)$ is the set $U(r_0)$ defined already. Thus $A \subset U(0)$ and $B \subset X \setminus U(1)$ for this family as well. Now suppose we are given two real numbers t and t' such that $0 \leq t < t' \leq 1$. Then there always exist dyadic numbers r and r' such that $t \leq r < r' \leq t'$. Now since $U(t) = \bigcup_{r_0 \leq t} U(r_0)$, then for each $r_0 \leq t \leq r$, $U(r_0) \subset U(r)$ and hence $U(t) \subset U(r)$ and thus $\bar{U}(t) \subset \bar{U}(r)$. Also $\bar{U}(r) \subset U(r')$ since $r < r'$. Finally because $r' \leq t'$, $U(r')$ is one of the open sets in $U(t')$ hence $U(r') \subset U(t')$. Putting these statements together we have

$$\bar{U}(t) \subset \bar{U}(r) \subset U(r') \subset U(t'),$$

so that $\bar{U}(t) \subset U(t')$. Thus all three properties for $|U(t)|_{t \in [0,1]}$ are established and Urysohn's theorem is proved.

We note that the converse of Urysohn's theorem is also true. For a Hausdorff space X , if for all disjoint closed sets A and B there is a continuous map $f: X \rightarrow [0,1]$ such that $f(A) = 0$ and $f(B) = 1$, then X is normal. The sets $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$ are disjoint, open and $A \subset U = f^{-1}[0, \frac{1}{2})$ and $B \subset V = f^{-1}(\frac{1}{2}, 1]$. Since f is continuous, U and V are open; any point $x \in U \cap V$ must have its image $f(x)$ in both $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$. This is impossible so $U \cap V = \emptyset$. Since this was true for any disjoint closed sets A and B , we have proved X is normal.

Bibliography

N. Bourbaki, op. cit. Vol. I and II.

W. Huebsch, *Annals of Math*, 61, 555 (1955).

L. S. Pontryagin, op. cit.

N. Steenrod, op. cit.

VII. THE HOMOTOPY GROUPS

To understand homotopy group, one needs only the rudiments of topology; to calculate such groups for spaces as simple as spheres, more advanced reasoning is required. As an introduction to this subject we will present the descriptive aspects of homotopy groups, and summarize their "elementary" properties. The review given here is based primarily on [Steenrod's] chapter 15.

One of the first trivial obstacles in learning about homotopy seems to be how to say "homotopy". Webster's Third New International Dictionary says ho·mot'o·py, accent on the second syllable. The related word, homotopic, is pronounced homo·top'ic, accent on the third syllable. It is our hope that this disposes with the only serious problem of learning the subject.

A. Homotopies

The idea of homotopy originates in the study of curves in arbitrary topological spaces. By a curve, one means a map (continuous function) f from the unit interval $I = [0,1]$ into the space X under consideration. The curve is said to start at $f(0) = x_0$ and end at $f(1) = x_1$. A homotopy relates the curve f to another curve g by giving a precise meaning to the idea that f can be deformed continuously into g . Specifically, a homotopy h is a continuous map $h: I \times I \rightarrow X$ such that $h(t,0) = f(t)$ and $h(t,1) = g(t)$. More generally, any pair of maps f and g from an arbitrary space

Z to X are connected by a homotopy $h: Z \times I \rightarrow X$ if h is continuous, $h(z,0) = f(z)$ and $h(z,1) = g(z)$. In what follows, Z will be typically some unit cube I^n in n -dimensional Euclidean space.

The concept of homotopy introduces a natural equivalence between the set of all maps of a space Z into a space X . Two maps $f: Z \rightarrow X$ and $g: Z \rightarrow X$ are homotopic iff there exists a homotopy connecting f and g . One writes this relation as $f \sim g$, and shows by elementary arguments that this relation is an equivalence relation.

[1) Identity: $f \sim f$ since the homotopy $h(z,t) = f(z)$ connects f with itself continuously. 2) Symmetry: if $f \sim g$ then $\exists$ a continuous $h: Z \times I \rightarrow X$ such that $h(z,0) = f(z)$ and $h(z,1) = g(z)$. Define $\tilde{h}(z,t) = h(z,1-t)$. Thus $\tilde{h}$ is a homotopy connecting g to f and so $g \sim f$. 3) Transitivity: if $f_0 \sim f_1$ and $f_1 \sim f_2$ for continuous maps f_0, f_1 and f_2 , then $\exists$ homotopies $h_1: Z \times I \rightarrow X$ and $h_2: Z \times I \rightarrow X$ connecting f_0 to f_1 , and f_1 to f_2 respectively. Define the function

$$h(z,t) = \begin{cases} h_1(z,2t) & 0 \leq t \leq \frac{1}{2} \\ h_2(z,2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases}$$

which is a continuous map from $Z \times I$ into X , since at $t = 1/2$, both definitions give $h(z,1/2) = f_1(z)$. The map h is a homotopy of f_0 into f_2 since $h(z,0) = f_0(z)$ and $h(z,1) = f_2(z)$. We thus prove $f_0 \sim f_2$ showing the homotopy relationship is transitive.]

Because homotopy defines an equivalence relation between maps of Z into X , the set of all maps of Z into X breaks up into disjoint equivalence classes, called homotopy classes. For certain reference spaces Z , the set of homotopy classes is given a special name and admits a group structure. The simplest example is the

fundamental group $\pi_1(X, x_0)$, which is the set of homotopy classes of closed curves beginning at x_0 . Note a closed curve beginning at $x_0 \in X$ is a map $f: I \rightarrow X$ which begins and ends at x_0 , $f(0) = f(1) = x_0$. The map f is a representative map of the set $[f]$ of all closed curves at x_0 homotopic to f . We show below that the set $\pi_1(X, x_0)$ of all such homotopy classes admits a group structure. Before doing this, we introduce other sets of homotopy classes for a given space X .

Unit 3-cube, I^3 . A point $t \in I^3$ has coordinates $t = (t_1, t_2, t_3)$, $t_i \in [0, 1]$, $i = 1, 2, 3$. The 2-face I^2 is defined by $t_3 = 0$, J^2 is the union of the remaining faces, and the boundary $\dot{I}^3 = I^2 \cup J^2$.

Consider a unit n -cube I^n , which is the n -fold product of the closed interval $[0, 1] = I$. Consider further all maps of I^n into the topological space X , which we want to learn about. Such maps generalize the idea of curves in X , to n -dimensional surfaces in X . With curves we restricted attention to closed curves. We have a similar option with surfaces: we can demand that the boundary $\dot{I}^n$ of the unit n -cube be mapped into a single point $x_0 \in X$. [The boundary $\dot{I}_n$ consists of all points of I^n not in the interior

of I^n (which is the largest open set contained in I^n and is identified with the n -fold product of the open interval $(0,1)$.)] The set of homotopy classes of closed surfaces in X is denoted $\pi_n(X, x_0)$, and is called the n^{th} homotopy group of X . [We will show the group property later.]

For $n > 1$ it will turn out to be useful to consider a generalization of $\pi_n(X, x_0)$ by requiring in part that the boundary $\dot{I}^n$ be mapped into a subspace A of X . More specifically, one demands that one face, I^{n-1} of the n -cube, be mapped into A , while the union of the remaining faces J^{n-1} be mapped into a single point $x_0 \in A$. The notation one uses is to let $t = (t_1, \dots, t_n)$ be the coordinates of a point in I^n . The face I^{n-1} is defined by convention to be the points $(t_1, \dots, t_{n-1}, 0)$. The boundary $\dot{I}^n$ is the set of points t when some $t_i = 0$ or 1 , $i = 1, \dots, n$, and is also the union of the faces, $I^{n-1} \cup J^{n-1} = \dot{I}^n$. The face I^{n-1} is a unit $n-1$ cube whose boundary consists of points of I^{n-1} which meet one of the other faces, $\dot{I}^{n-1} = I^{n-1} \cap J^{n-1}$. The idea that some map f takes I^n into X , I^{n-1} into A , and J^{n-1} into x_0 is summarized symbolically by

$$f: (I^n, I^{n-1}, J^{n-1}) \rightarrow (X, A, x_0).$$

The set of all such maps is denoted by $F^n(X, A, x_0)$, and the homotopy classes of this set are denoted $\pi_n(X, A, x_0)$, and called the n^{th} relative homotopy group of $X \bmod A$. One has to demonstrate that such sets show a group structure, which they do for $n \geq 2$ when A is larger than the single point x_0 . When $A = \{x_0\}$ then $\pi_n(X, x_0, x_0) = \pi_n(X, x_0)$, showing that the n^{th} homotopy groups are

a special case. The set of maps $f: (I^n, \dot{I}^n) \rightarrow (X, x_0)$ is denoted by $F^n(X, x_0)$, and one has to show that $\pi_n(X, x_0)$ admits a group structure for $n \geq 1$.

F. Group structure of homotopy classes

Let us consider the set $\pi_n(X, A, x_0)$ of homotopy classes of $F^n(X, A, x_0)$ [the set of maps $f: (I^n, \dot{I}^{n-1}, J^{n-1}) \rightarrow (X, A, x_0)$] where $x_0 \in A \subset X$ and X is some arbitrary given topological space. Whenever $n=1$ we assume $A = \{x_0\}$. A product is defined on $\pi_n(X, A, x_0)$ as follows. For any pair of maps $f_1, f_2 \in F^n(X, A, x_0)$ define the product

$$f_1 * f_2(t_1, \dots, t_n) = \begin{cases} f_1(2t_1, t_2, \dots, t_n) & 0 \leq t_1 \leq \frac{1}{2} \\ f_2(2t_1 - 1, t_2, \dots, t_n) & \frac{1}{2} \leq t_1 \leq 1 \end{cases}.$$

As $t_1 = \frac{1}{2}$, both $(2t_1, \dots, t_n)$ and $(2t_1 - 1, \dots, t_n)$ lie on a face in J^{n-1} , hence their image is x_0 under f_1 and f_2 respectively. This insures $f_1 * f_2$ is well defined, continuous and itself an element of $F^n(X, A, x_0)$.

Now suppose f'_1, f'_2 are also in $F^n(X, A, x_0)$ such that $f'_1 \sim f_1$ and $f'_2 \sim f_2$. We show that $f'_1 * f'_2 \sim f_1 * f_2$. Let h_k be the homotopy of f'_k and f_k , $k=1,2$. Defining $h = h_1 * h_2$ in the obvious way, one constructs a homotopy between $f'_1 * f'_2$ and $f_1 * f_2$. Thus for the homotopy classes $[f_k]$ of which f_k is representative, $k=1,2$, one can define the product $[f_1] * [f_2] = [f_1 * f_2]$. This shows the existence of a binary operation on $\pi_n(X, A, x_0)$. We next show that this product satisfies the group axioms of associativity, existence of identity, and existence of inverses.

Associativity. For any three elements of $\pi_n(X, A, x_0)$ one

show: $(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3)$ where $f_i \in F^n(X, A, x_0)$ is a representative of the homotopy class $[f_i]$, $i = 1, 2, 3$. We need to demonstrate only that $(f_1 * f_2) * f_3 \sim f_1 * (f_2 * f_3)$. To prove this, as well as several later assertions, one uses the notion of stretching or contracting the interval $[0, 1]$ of the first coordinate t_1 . Let $u(a, b, t_1) = (t_1 - a)/(b - a)$ for $a < b$ be the map from $[a, b]$ onto $[0, 1]$. This function is also continuous in the parameters a and b , as long as $a < b$. We consider

$$F_{ab}(f_1, f_2, f_3) = \begin{aligned} & f_1(u(0, a, t_1), t_2, \dots, t_n) & t_1 \in [0, a] \\ & f_2(u(a, b, t_1), t_2, \dots, t_n) & t_1 \in [a, b] \\ & f_3(u(b, 1, t_1), t_2, \dots, t_n) & t_1 \in [b, 1] \end{aligned}$$

which corresponds to $(f_1 * f_2) * f_3$ if $a = \frac{1}{4}$ and $b = \frac{1}{2}$. Varying a, b continuously (as a function of a single parameter on $[0, 1]$) defines a homotopy of this product; we can thus stretch a to become $\frac{1}{2}$ and b to become $\frac{3}{4}$ in which case F is $f_1 * (f_2 * f_3)$ and thus homotopic to $(f_1 * f_2) * f_3$.

Identity. Define the map $e \in F^n(X, A, x_0)$ to be the constant map $e(t) = x_0$, $t \in [0, 1]$. Now for any other map $f \in F^n(X, A, x_0)$, by definition

$$e * f(t) = \begin{cases} x_0 & t_1 \in [0, \frac{1}{2}] \\ f(u(\frac{1}{2}, 1, t_1), \dots, t_n) & t_1 \in [\frac{1}{2}, 1] \end{cases}$$

By replacing the intervals above by $[0, \frac{1-s}{2}]$ and $[\frac{1-s}{2}, 1]$ one obtains a homotopy h of $e * f$:

$$h(t, s) = \begin{cases} x_0 & t_1 \in [0, \frac{1-s}{2}] \\ f(u(\frac{1-s}{2}, 1, t_1), \dots, t_n) & t_1 \in [\frac{1-s}{2}, 1] \end{cases}.$$

This shows $e * f$ is homotopic to $h(t, 1) = f(t)$, (where as usual $t = (t_1, \dots, t_n)$). The homotopy class of e thus acts as the identity of $\pi_n(X, A, x_0)$.

For later arguments it is useful to note that any map $f \in F^n(X, A, x_0)$ whose image $f(I^n) \subset A$, is necessarily in the identity homotopy class $[e]$. We have to show $f \sim e$. The homotopy we seek can be built in two stages. First we find the homotopy h connecting the identity map of I^n into I^n with the map of I^n into the face $t_n = 1$ (a subset of J^{n-1}):

$$h(t, \tau) = (t_1, \dots, t_{n-1}, (1-\tau)t_n + \tau).$$

Next we consider $f(h(t, \tau))$ and note that it is a continuous map of $I^n \times I \rightarrow X$; whenever $t \in I^{n-1}$ then $f(h(t, \tau)) \in A$, and whenever $t \in J^{n-1}$ then $f(h(t, \tau)) = x_0$. This shows $f(h(t, \tau))$ is a homotopy of f in $F^n(X, A, x_0)$; since $f(h(t, 1)) = x_0 = e(t)$ we have that $f \sim e$.

Inverse. Consider any $f \in F^n(X, A, x_0)$; define $\bar{f}(t) = f(1-t_1, t_2, \dots, t_n)$. If $\bar{f} * f$ is homotopic to a constant then $[\bar{f}]$ is the group inverse to $[f]$. For simplicity in this argument we suppress all components of t except t_1 . The product $\bar{f} * f$ is

$$\bar{f} * f(t_1) = \begin{cases} f(1-2t_1) & 0 \leq t_1 \leq \frac{1}{2} \\ f(2t_1-1) & \frac{1}{2} \leq t_1 \leq 1 \end{cases}.$$

Consider the domain $I^n \times I$ for a homotopy h of $\bar{f} * f$. In the fig-

ure we plot the (t_1, τ) subplane. Define h so

that at $\tau = 0$, h has the value of $\bar{f} * f$; namely when $t_1 = \bar{t}_1$ or $1 - \bar{t}_1$, then $h(t_1, 0) = f(1 - 2\bar{t}_1)$. These two points are points of an "arch" in the above plane, denoted by the dotted line. Extend the definition of h on this dotted line to remain equal to $f(1 - 2\bar{t}_1)$. By changing $\bar{t}_1$ from 0 to $1/2$, one covers the entire $I \times I$ square; in the process one defines a continuous function h which is a homotopy of $\bar{f} * f$ in $F^n(X, A, x_0)$. For $\tau = 1$, h takes on the value $f(1)$; since $t_1 = 1$ is a face in J^{n-1} , we have that $h(t, 1) = x_0 = e(t)$. This shows $\bar{f} * f \sim e$, hence $[\bar{f}]$ is the inverse of $[f]$.

We thus obtain the important result that $\pi_n(X, A, x_0)$ admits a group structure, called the n dimensional relative homotopy group of X mod A with base point x_0 . It is defined for $n \geq 2$, and also for $n = 1$, when $A = \{x_0\}$ in which case $\pi_1(X, x_0)$ is the fundamental group of X with base point x_0 . The fundamental group is not in general commutative; however a general result is that for $n \geq 2$, $\pi_n(X, x_0)$ is commutative, and for $n \geq 3$, $\pi_n(X, A, x_0)$ is commutative. For this reason Steenrod uses the additive notation $f_1 + f_2$ for the product $f_1 * f_2$.

The demonstration that the relative homotopy groups are

commutative can be understood by studying what happens when $n = 3$. The idea of the proof is to note the homeomorphism between the 3-cube and the cylinder. Suppose f_1 and f_2 are two maps of $F^3(X, A, x_0)$.

Then, for the product $f_1 * f_2$, f_1 has the domain $0 \leq t_1 \leq 1/2$ of the cube, and f_2 the domain $1/2 \leq t_1 \leq 1$. Since the cube and cylinder are homeomorphic, f_1 can be thought of as having the domain of half the cylinder as shown above, and f_2 the other half. Now suppose one rotates the cylinder about its axis by θ , inducing of course a rotation of the cube. Denoting the rotation of a point t in I^3 by $R(\theta)t$, we see that for any $f \in F^3(X, A, x_0)$, $f(R(\pi\tau)t)$ is a homotopy of f in $F^3(X, A, x_0)$ as τ varies from 0 to 1. [During the rotation the image of J^2 stays in x_0 .] Call $f^r(t) = f(R(\pi)t)$, the map induced by rotating the cylinder 180° ; because of the homotopy, $f^r \sim f$. This rotation interchanges the two halves of the cylinder, so that $(f_1 * f_2)^r = f_2^r * f_1^r$. Using the homotopy equivalences we have $f_1 * f_2 \sim (f_1 * f_2)^r = f_2^r * f_1^r \sim f_2 * f_1$. This completes the proof when $n = 3$ that multiplication is commutative.

For $n > 3$, the proof is identical since only the coordinates t_1, t_2 and t_3 are involved. Thus the relative homotopy groups are commutative. For $n \geq 3$, the homotopy groups $\pi_n(X, x_0)$ are special cases of the relative homotopy groups, hence $\pi_n(X, x_0)$ is commutative. With a modified argument one can show $\pi_2(X, x_0)$ is also commutative. During the rotation,

the image of i^2 stays in x_0 , so as before $f_1 * f_2 \sim (f_1 * f_2)^r \sim f_2^r * f_1^r \sim f_2 * f_1$. The argument will not work for $\pi_2(X, A, x_0)$ since during the rotation, the image of J^1 will not in general stay in x_0 .

C. Homomorphisms between homotopy groups

What we have introduced at this point is the assignment to each pair (X, A) of a topological space X and a subspace $A \subset X$, groups $\pi_n(X, A, x_0)$ [$x_0 \in A$ is a base point]. If one also assigns to each map $h: (X, A, x_0) \rightarrow (Y, B, y_0)$, a group homomorphism $h_*: \pi_n(X, A, x_0) \rightarrow \pi_n(Y, B, y_0)$, one can then use algebra to understand topological problems. To this end consider a map $h: X \rightarrow Y$ such that $A \subset X$ is carried into $B \subset Y$ and $x_0 \in A$ has as image $y_0 \in B$. This is denoted briefly by $h: (X, A, x_0) \rightarrow (Y, B, y_0)$. Let f be any map of $F^n(X, A, x_0)$; then the composition hf is a map in $F^n(Y, B, y_0)$. Let H be a homotopy of f into some other map $f' \in F^n(X, A, x_0)$. Then the composition hH is a homotopy of hf into hf' in $F^n(Y, B, y_0)$. Passing to the homotopy classes, h gives a well defined map taking $[f]$ into $[hf]$.

This map is denoted by

$$h_* : \pi_n(X, A, x_0) \rightarrow \pi_n(Y, B, y_0) .$$

To show h_* is a group homomorphism we have only to show it preserves group products, $h_*([f] * [f']) = [hf] * [hf']$. This follows from consideration of the representative elements of the homotopy classes: $h(f_1 * f_2) = (hf_1) * (hf_2)$. The map h_* is called the homomorphism induced by h.

Before turning to the description of the elementary properties of the homotopy groups, we need one more tool, the boundary operator:

$$\partial_* : \pi_n(X, A, x_0) \rightarrow \pi_{n-1}(A, x_0) .$$

Consider any $f \in F^n(X, A, x_0)$. This map takes I^n into X , I^{n-1} into A and J^{n-1} into x_0 . In particular if we restrict attention to the behavior of f on I^{n-1} we have a map which takes I^{n-1} into A . Furthermore, since the boundary $I^{n-1} = I^{n-1} \cap J^{n-1} \subset J^{n-1}$, we see that f takes I^{n-1} into x_0 . Thus to each $f \in F^n(X, A, x_0)$ there is a unique $\partial f \in F^{n-1}(A, x_0)$ defined by setting $\partial f = f$ on I^{n-1} .

Now suppose $f_0 \sim f_1$ are both maps in $F^n(X, A, x_0)$. These maps restricted to I^{n-1} are ∂f_0 and ∂f_1 and are both in $F^{n-1}(A, x_0)$. Let h denote the homotopy connecting f_0 and f_1 . By restricting h to $I^{n-1} \times I$, we obtain a homotopy of ∂f_0 and ∂f_1 . Passing to the homotopy classes, ∂ defines a function ∂_* from $\pi_n(X, A, x_0)$ to $\pi_{n-1}(A, x_0)$. To show ∂_* is a homomorphism, one notes $\partial(f * f') = (\partial f) * (\partial f')$ for any $f, f' \in F^n(X, A, x_0)$, and then passes to the homotopy classes. The homomorphism ∂_* is called the boundary homo-

morphism, and is ordinarily denoted by ∂ .

D. Basic Properties

The rest of this lecture is devoted to showing the basic properties of π_n , ∂ and h_* [where $h: (X, A, x_0) \rightarrow (Y, B, y_0)$]. Maps which play a special role are $i: (A, x_0) \rightarrow (X, x_0)$ which is the inclusion map $i(x) = x$, $\forall x \in A$, and $j: (X, x_0, x_0) \rightarrow (X, A, x_0)$ which is the inclusion map $j(x) = x$, $\forall x \in X$. Note i and j differ only in their domains and ranges. The infinite sequence of maps

$$\begin{aligned} \cdots \xrightarrow{\partial} \pi_n(A, x_0) \xrightarrow{i_*} \pi_n(X, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0) \xrightarrow{\partial} \\ \pi_{n-1}(A, x_0) \xrightarrow{i_*} \cdots \xrightarrow{j_*} \pi_2(X, A, x_0) \xrightarrow{\partial} \pi_1(A, x_0) \xrightarrow{i_*} \pi_1(X, x_0) \end{aligned}$$

is called the homotopy sequence of (X, A, x_0) .

1. Identity

If h is the identity map of (X, A, x_0) then h_* is the identity homomorphism of $\pi_n(X, A, x_0)$.

Proof: For each $f \in F^n(X, A, x_0)$, the composition $hf = f$. Passing to the homotopy classes we see $h_*[f] = [f]$ for each $[f] \in \pi_n(X, A, x_0)$ which proves the theorem.

2. Products

If $h: (X, A, x_0) \rightarrow (Y, B, y_0)$ and $k: (Y, B, y_0) \rightarrow (Z, C, z_0)$ then $(kh)_* = k_*h_*$ for each dimension of (relative) homotopy groups.

Proof: Let f be any element of $F^n(X, A, x_0)$. Now hf is a representative of $h_*[f]$, and $k(hf)$ is therefore a representative of $k_*(h_*[f])$, which by the associativity $(kh)f = k(hf)$ shows $k_*(h_*[f])$ must also equal $(kh)_*[f]$. Since this holds for all $[f] \in \pi_n(X, A, x_0)$ we obtain the desired result $(kh)_* = k_*h_*$.

3. Commutativity of ∂

If $h: (X, A, x_0) \rightarrow (Y, B, y_0)$ and $h|A: (A, x_0) \rightarrow (B, y_0)$ where $h|A = h$ for all $x \in A$, then $\partial h_* = (h|A)_* \partial$, that is the diagram

$$\begin{array}{ccc}
 \pi_n(X, A, x_0) & \xrightarrow{h_*} & \pi_n(Y, B, y_0) \\
 \downarrow \partial & & \downarrow \partial \\
 \pi_{n-1}(A, x_0) & \xrightarrow{(h|A)_*} & \pi_{n-1}(B, y_0)
 \end{array}$$

is commutative.

Proof: Let $[f]$ be any element of $\pi_n(X, A, x_0)$ with f a representative map of the homotopy class $[f]$. Now the map $hf \in F^n(Y, B, y_0)$ is representative of $h_*[f]$, and $\partial(hf) \in F^{n-1}(B, y_0)$ is representative of $\partial(h_*[f]) = (\partial h_*)[f]$. Similarly we see $\partial f \in F^{n-1}(A, x_0)$ is representative of $\partial[f]$, and $(h|A)(\partial f) \in F^{n-1}(B, y_0)$ is representative of $(h|A)_*(\partial[f]) = [(h|A)_* \partial][f]$. Since on the subset I^{n-1} , $(h|A)(\partial f) = hf$ and $\partial(hf) = hf$ we obtain $[(h|A)_* \partial][f] = (\partial h_*)[f]$ for every $[f] \in \pi_n(X, A, x_0)$; this shows $(h|A)_* \partial = \partial h_*$.

4. Exactness

The homotopy sequence is exact. (A sequence $A \xrightarrow{p} B \xrightarrow{q} C$ of groups and homomorphisms is exact if the image of p is the kernel of q : $\text{Im}(p) = \ker(q)$. The image of p is the subgroup $p(A)$ of B . The kernel of q is the subgroup of B which is mapped by q into the identity of C . A simple example is for a map $p: \mathbb{Z} \rightarrow \mathbb{Z}$ ($\mathbb{Z}$ is the integers) defined by $p(n) = 2n$, and a map $q: \mathbb{Z} \rightarrow \mathbb{Z}_2$ ($\mathbb{Z}_2$ is the integers mod-

ulo 2) defined by $\varphi(n) \equiv n \pmod{2}$. The kernels of p and q are

$$\ker(p) = \{0\}$$

$$\ker(q) = \{0, \pm 2, \pm 4, \dots\}$$

and the images of p and q are

$$\operatorname{im}(p) = \{0, \pm 2, \pm 4, \dots\}$$

$$\operatorname{im}(q) = \{0, 1\}.$$

Thus by inspection we see $\operatorname{im}(p) = \ker(q)$ and $\ker(p) = \operatorname{im}(q)$ with $\ker(p) \neq \operatorname{im}(q)$. In other words the sequence $\mathbb{Z} \xrightarrow{p} \mathbb{Z} \xrightarrow{q} \mathbb{Z}$ is exact whereas $\mathbb{Z} \xrightarrow{q} \mathbb{Z} \xrightarrow{p} \mathbb{Z}$ is not exact. An infinite sequence $\dots \xrightarrow{p_1} A_i \xrightarrow{p_{i+1}} A_{i+1} \xrightarrow{p_{i+2}} A_{i+2} \xrightarrow{p_{i+3}} \dots$ is exact if $\operatorname{im}(p_i) = \ker(p_{i+1})$ for each i . One generally proves exactness of $A \xrightarrow{p} B \xrightarrow{q} C$ by showing i) $\operatorname{im}(p) \subset \ker(q)$, which is $qp(A) = e$ the identity of C ; and ii) $\ker(q) \subset \operatorname{im}(p)$, which is that whenever $b \in \ker(q)$, there exists an $a \in A$ such that $p(a) = b$. The proof below of exactness is taken from [Hu]).

Proof: Consider the homotopy sequence of (X, A, x_0) ; it is the infinite sequence of groups and homomorphisms

$$\begin{aligned} \dots \xrightarrow{\partial} \pi_n(A, x_0) &\xrightarrow{j_*} \pi_n(X, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0) \xrightarrow{\partial} \pi_{n-1}(A, x_0) \xrightarrow{j_*} \\ \dots \xrightarrow{j_*} \pi_2(X, A, x_0) &\xrightarrow{\partial} \pi_1(A, x_0) \xrightarrow{j_*} \pi_1(X, x_0) \end{aligned}$$

which terminates at $\pi_1(X, x_0)$. There are three cases to prove

$$(1) \quad \operatorname{im}(j_*) = \ker(\partial) \quad \text{at } \pi_n(X, A, x_0)$$

$$(11) \quad \text{im}(i_*) = \ker(j_*) \text{ at } \pi_n(X, x_0)$$

$$(111) \quad \text{im}(\partial) = \ker(i_*) \text{ at } \pi_n(A, x_0) .$$

Case (1): $\text{im}(j_*) = \ker(\partial)$ for

$$\pi_n(X, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0) \xrightarrow{\partial} \pi_{n-1}(A, x_0) .$$

To show $\text{im} \subset \ker$, for each $[f] \in \pi_n(X, x_0)$ let $f \in F^n(X, x_0)$ be a representative map. Every such map is also a map of $(I^{n-1}, I^{n-1}, j^{n-1}) \rightarrow (X, A, x_0)$, so $F^n(X, x_0) \subset F^n(X, A, x_0)$; thus $f \in F^n(X, A, x_0)$ and so f represents also $j_*[f]$ (recall $j(x) = x$, $\forall x \in X$ is the inclusion map $(X, x_0, x_0) \rightarrow (X, A, x_0)$). Now if we restrict f to I^{n-1} , by definition of $f \in F^n(X, x_0)$, $f|I^{n-1}$ is the constant map. Thus $f|I^{n-1} = \partial f$ is in the identity class $[e] \in \pi_n(X, A, x_0)$ for every $f \in F^n(X, x_0)$, showing $\partial j_*(\pi_n(X, x_0)) = [e]$ or in other words $\text{im}(j_*) \subset \ker(\partial)$.

To show $\ker \subset \text{im}$, for each $[f] \in \ker(\partial)$, let $f \in F^n(X, A, x_0)$ be the representative; therefore ∂f is homotopic to the constant map, e . Let the homotopy of ∂f and e be $h: (I^{n-1} \times I, I^{n-1} \times I) \rightarrow (A, x_0)$, $h(t, 0) = \partial f(t)$ and $h(t, 1) = e(t)$, $t \in I^{n-1}$. Extend h to the set $I^n \times 0$ by defining $h(t, 0) = f(t)$ where now $t \in I^n$. Extend h to the set $J^{n-1} \times I$ by defining $h(t, \tau) = f(t) = x_0$ when $t \in J^{n-1}$, $\tau \in I$.

Since $I^n = I^{n-1} \cup J^{n-1}$ we have defined $h(t, \tau)$ on the set $E = (I^n \times 0) \cup (I^n \times 1)$ which in the figure is the base of the cube and the side faces. The set E defines an n cell which lies on the boundary of the $n+1$ cell $I^n \times I$. Given h defined on E , the idea now is to extend h to all of the cube $I^n \times I$.

Let $\bar{F}: E \rightarrow E$ be the identity map $\bar{F}(t, \tau) = (t, \tau)$ whenever $(t, \tau) \in E$. By its definition E is a closed subset of $I^n \times I$. The problem can be formulated by asking can a map $\bar{F}$ defined on a closed subset A (which is here E) of a normal space Z (which is $I^n \times I$) into a space Y (which is again E) be extendable to the whole space Z ? When the answer is yes, Y is called solid. The definition we used for normal spaces Z before was that any two disjoint closed subsets of Z have disjoint open neighborhoods. We then showed Urysohn's theorem that for two disjoint closed subsets $A, B \subset Z$ there exists a continuous function of Z into $[0, 1]$ which is 0 on A and 1 on B . Another consequence of Urysohn's theorem (which can be found in [Bourbaki], as well as in the appendix to chapter VI) is that for any closed subset $A \subset Z$ and any map $\bar{F}: A \rightarrow [0, 1]$, there exists an extension of $\bar{F}$ mapping Z into $[0, 1]$. Thus by definition the closed interval I is solid. One can then show that products of I are likewise solid, and hence n -cells are solid. By these arguments one deduces that to $\bar{F}: E \rightarrow E$ there exists an extension $r: I^n \times I \rightarrow E$ such that r is the identity on E . Such a map is called a retraction of $I^n \times I$ into E .

Given the retraction r of $I^n \times I$ into E , we form the composition hr which is a homotopy in $F^n(X, A, x_0)$ of f into some map f' . Now $hr: J^{n-1} \times I \rightarrow x_0$ since here the domain is in E and on this part of E , h was defined to give x_0 . Next, $hr: I^{n-1} \times I \rightarrow h(I^{n-1} \times I)$ since the domain here is also in E ; furthermore, since $h: I^{n-1} \times I \rightarrow$

A is the homotopy of f with the identity map, $h_r(t, 1) = x_0$. Thus $h_r(t, 1) = f'(t)$ maps $I^{n-1} \cup J^{n-1}$ into x_0 , so that $f' \in F^n(X, x_0, x_0)$. Furthermore f' represents the element $[f'] \in \pi_n(X, x_0)$, whose image $j_*[f'] = [f]$, and by assumption $\partial[f] = [e]$. These considerations therefore show $\ker(\partial) \subset \text{im}(j_*)$. Combining this result with $\text{im}(j_*) \subset \ker(\partial)$ we prove part (i) that $\text{im}(j_*) = \ker(\partial)$.

Case (ii): $\text{im}(i_*) = \ker(j_*)$ for

$$\pi_n(A, x_0) \xrightarrow{i_*} \pi_n(X, x_0) \xrightarrow{j_*} \pi_n(X, A, x_0).$$

To show $\text{im}(i_*) \subset \ker(j_*)$, for each $[f] \in \pi_n(A, x_0)$ let $f \in F^n(A, x_0)$ be a representative map of the homotopy class. Since $F^n(A, x_0) \subset F^n(X, x_0)$, then f also represents $i_*[f] \in \pi_n(X, x_0)$. Moreover $f \in F^n(X, A, x_0)$ since $F^n(A, x_0) \subset F^n(X, A, x_0)$ and is homotopic to the constant map in $F^n(X, A, x_0)$. Thus f represents $j_*i_*[f]$ as does the constant map. This shows $j_*i_*[f] = [e]$, and hence $\text{im}(i_*) \subset \ker(j_*)$.

To show $\ker(j_*) \subset \text{im}(i_*)$, let $f \in F^n(X, x_0)$ represent $[f] \in \ker(j_*)$ so that $j_*[f] = [e]$. The map $jf \in F^n(X, A, x_0)$ is by assumption homotopic to the constant map, so let h be the homotopy in $F^n(X, A, x_0)$ connecting jf to e . Use h to define a map

$$g(t_1, \dots, t_n, \tau) = \begin{cases} h(t_1, \dots, t_{n-1}, 0, 2t_n) & 0 \leq 2t_n \leq \tau \\ h(t_1, \dots, t_{n-1}, \frac{2t_n - \tau}{2 - \tau}, \tau) & \tau \leq 2t_n \leq 2 \end{cases}$$

and define $g_\tau(t) = g(t_1, \dots, t_n, \tau)$. Now whenever $t \in J^{n-1}$ then,

except possibly when $t_n = 1$, the domain of h in the definition of g is in $J^{n-1} \times I$ so $g_\tau(J^{n-1}) = x_0$. When $t_n = 1$ then $g_\tau(t) = h(t_1, \dots, t_{n-1}, 1, \tau) = x_0$. Therefore $g_\tau(J^{n-1}) = x_0$. For $\tau = 0$, $g_0(t) = h(t, 0) = f(t)$. Finally when $t \in I^{n-1}$, then $t_n = 0$ and using the definition of g , $g_\tau(t) = h(t_1, \dots, t_{n-1}, 0, 0) = f(t)$; when $t \in I^{n-1}$ then $f(t) = x_0$, and so $g_\tau(I^{n-1}) = x_0$. One thus shows that there is a homotopy in $F^n(X, x_0)$ of f to the map $g_1(t)$. Now when $0 \leq t_n \leq 1/2$, then $g_1(t) = h(t_1, \dots, t_{n-1}, 0, 2t_n) \in A$; when $1/2 \leq t_n \leq 1$ then $g_1(t) = h(t_1, \dots, t_{n-1}, 2t_n - 1, 1) = x_0$. As a consequence, $g_1(I^n) \subset A$, showing f is homotopic in $F^n(X, x_0)$ to the map g_1 which in turn is also an element of $F^n(A, x_0)$. The homotopy class or $i_*[g_1]$ has as representative $g_1 \sim f$ in $F^n(X, x_0)$, showing $i_*[g_1] = [f]$. We have proved for any $[f] \in \ker(j_*)$ that there exists a $[g_1]$ such that $i_*[g_1] = [f]$; in other words $\ker(j_*) \subset \text{im}(i_*)$. This concludes the proof that $\ker(j_*) = \text{im}(i_*)$.

Case (iii): $\text{im}(\partial) = \ker(i_*)$ for

$$\pi_{n+1}(X, A, x_0) \xrightarrow{\partial} \pi_n(A, x_0) \xrightarrow{i_*} \pi_n(X, x_0).$$

To show $\text{im}(\partial) \subset \ker(i_*)$, let f be representative of $[f] \in \pi_{n+1}(X, A, x_0)$. The boundary ∂f is in $F^n(A, x_0)$ and represents $\partial[f]$. Since $F^n(A, x_0) \subset F^n(X, x_0)$, ∂f also represents $i_*\partial[f]$. Define the homotopy $h(t_1, \dots, t_n, \tau) = f(t_1, \dots, t_n, \tau)$, and the corresponding functions $h_\tau(t_1, \dots, t_n) = h(t_1, \dots, t_n, \tau)$. Since $f \in F^{n+1}(X, A, x_0)$, then $h_\tau(I^n) \subset X$ and $h_\tau(I^n) \subset f(J^n) = x_0$. Thus h is a homotopy of h_0 into h_1 in $F^n(X, x_0)$. Now $h_0(t_1, \dots, t_n) = f(t_1, \dots, t_n, t_{n+1}=0) = \partial f(t_1, \dots, t_n)$, and $h_1(t_1, \dots, t_n) = f(t_1, \dots, t_n, t_{n+1}=1) \subset f(J^n) = x_0$. Therefore ∂f

is homotopic to the constant map in $F^n(X, x_0)$, showing $i_* \partial[f] = [e]$, and hence $\text{im}(\partial) \subset \ker(i_*)$.

To show $\ker(i_*) \subset \text{im}(\partial)$, let $f \in F^n(A, x_0)$ represent any $[f]$ in $\pi_n(A, x_0)$ such that $i_*[f] = [e]$. Since $F^n(A, x_0) \subset F^n(X, x_0)$, f also represents $i_*[f]$ and so is homotopic to the constant map in $F^n(X, x_0)$. Let the homotopy be denoted $h(t, \tau)$. Now define a function $g(t_1, \dots, t_n, t_{n+1}) = h(t_1, \dots, t_n, t_{n+1})$, from which we see $g: (I^{n+1}, I^n, J^n) \rightarrow (X, A, x_0)$ since for $t_{n+1} = 0$, $g(t, 0) = h(t, 0) = f(t) \in A$; for $t_{n+1} = 1$, $g(t, 1) = h(t, 1) = x_0$; and for $i \neq n+1$, $t_i = 0$ or 1 then $g(t, t_{n+1}) = x_0$. Therefore $g \in F^n(X, A, x_0)$ and represents $[g]$ such that $\partial[g]$ is represented by $g(t, 0) = h(t, 0) = f(t)$. This shows $\partial[g] = [f]$ proving $\ker(i_*) \subset \text{im}(\partial)$. This concludes the proof that $\ker(i_*) = \text{im}(\partial)$, and the last step in showing that the homotopy sequence of (X, A, x_0) is exact.

5. Homotopy

If the maps h and k of $(X, A, x_0) \rightarrow (Y, B, y_0)$ are connected by a homotopy which maps $A \times I \rightarrow B$ and $x_0 \times I \rightarrow y_0$ then h_* and k_* coincide for each n .

Proof: For any $[f] \in \pi_n(X, A, x_0)$ let $f \in F^n(X, A, x_0)$ be a representative map. Then hf and kf represent $h_*[f]$ and $k_*[f]$ respectively. Let H be the homotopy connecting k and h , $H: (X \times I, A \times I, x_0 \times I) \rightarrow (Y, B, y_0)$. Define $\tilde{H}(t, \tau) = H(f(t), \tau)$, and $\tilde{H}_\tau(t) = \tilde{H}(t, \tau)$. We see that $\tilde{H}_0(t) = H(f(t), 0) = h(f(t)) = hf(t)$, and $\tilde{H}_1(t) = H(f(t), 1) = k(f(t)) = kf(t)$. Moreover, for each τ , $\tilde{H}_\tau \in F^n(Y, B, y_0)$ showing that $\tilde{H}$ is a homotopy of hf and kf in $F^n(Y, B, y_0)$, implying $h_*[f] = k_*[f]$. Since $[f]$ was arbitrary, $h_* = k_*$.

6. Trivial Space

If X consists of a single point x_0 , then for each n , $\pi_n(X, x_0)$ consists only of the identity element.

Proof: Since every map $f \in F^n(X, x_0)$ is the constant map, $[f] = [e]$ exhausts the elements of $\pi_n(x_0, x_0)$.

7. Fundamental theorem

Let B be a bundle over X , $A \subset X$, $B_0 = p^{-1}(A)$, $b_0 \in B_0$ and $x_0 = p(b_0)$. Then

$$p_* : \pi_n(B, B_0, b_0) \rightarrow \pi_n(X, A, x_0) \quad n \geq 2$$

is an isomorphism. [A group isomorphism is a 1-1 onto homomorphism. One shows this by proving $\ker(p_*)$ consists only of the identity element, and $\text{im}(p_*) = \pi_n(X, A, x_0)$].

Proof: To prove $\ker(p_*)$ consists only of the identity, one proceeds as follows. Let $[f_0] \in \pi_n(B, B_0, b_0)$ and f_0 be a representative map of $[f_0]$; further suppose $p_*[f_0] = [e]$, which means pf_0 is homotopic to the constant map. Let $\bar{h}$ be the homotopy of pf_0 to the constant map in $F^n(X, A, x_0)$. Now we have a map f_0 of I^n into B , a map pf_0 of I^n into X and a homotopy $\bar{h}$ of $\bar{h}_0 = pf_0$ into the

constant map. Now $\bar{h}_0$ induces on I^n a bundle structure $\bar{h}_0^{-1}(B) = B'$ such that I^n is the base space, $\{\bar{h}_0^{-1}(v_j)\}$ are the coordinate neighborhoods (where $\{v_j\}$ are the coordinate neighborhoods of X) and $g_{ij}(\bar{h}_0(t))$ are the coordinate transformations. Keeping the same fibre as for B , the bundle construction theorem gives a unique (up to bundle equivalence) bundle structure over I^n . We call the induced bundle space B' .

Now corresponding to $pf_0 = \bar{h}_0$ is induced a bundle map $h_0: B' \rightarrow B$ which covers $\bar{h}_0$. We have all the elements needed to apply the covering homotopy theorem, since the base of $\bar{h}_0^{-1}(B)$ is I^n which is a compact space. There is thus a homotopy h of h_0 which covers $\bar{h}$. We would like to use this to show the existence of a homotopy $f: I^n \times I \rightarrow B$ such that $pf = \bar{h}$ and f is stationary with $\bar{h}$. This conclusion, which Steenrod calls the Second Covering Homotopy Theorem, rests on the assumption of a space X' ($=I^n$) with suitable properties such as compactness, a map $f_0: X' \rightarrow B$, its induced map $\bar{h}_0 = pf_0$, and a homotopy of $\bar{h}_0$. We now prove this theorem.

Second Covering Homotopy Theorem

One does this by characterizing the induced bundle space B' of $\bar{h}_0^{-1}(B)$ as the set of points (t, b) of $I^n \times B$ such that $\bar{h}_0(t) = p(b)$. Let the natural projections $I^n \times B \rightarrow I^n$ and $I^n \times B \rightarrow B$ be denoted by p' and h_0 , respectively. The mapping diagram

$$\begin{array}{ccc} B' & \xrightarrow{h_0} & B \\ p' \downarrow & & \downarrow p \\ I^n & \xrightarrow{\bar{h}_0} & X \end{array}$$

is necessarily commutative: $ph_0(t, b) = p(b) = \bar{h}_0 p'(t, b)$. Define $V_j' = \bar{h}_0^{-1}(V_j)$ as the coordinate neighborhoods, and set $\phi_j'(t, y) = (t, \phi_j(\bar{h}_0(t), y))$ where $y \in Y$ (the fibre), and ϕ_j is the coordinate function of B over V_j . Since 1) $p'\phi_j'(t, y) = t$ and 2) one can define $p_j'(t, b) = p_j(b)$ whenever $\bar{h}_0(t) = p(b) \in V_j$, so that $p_j'\phi_j'(t, y) = y$, then ϕ_j' maps $V_j' \times Y$ onto $p'^{-1}(V_j) \cap B'$. One also can show $g_{jk}'(t) \cdot y = p_j'\phi_k'(t, y) = p_j\phi_k(\bar{h}_0(t), y) = g_{jk}(\bar{h}_0(t)) \cdot y$, implying the bundle one has defined here is equivalent to that given by the bundle structure theorem for $\bar{h}^{-1}(B)$.

Using the above definition of the induced bundle $\bar{h}_0^{-1}(B)$, one can define a cross section $\phi: I^n \rightarrow B'$ by $\phi(t) = (t, f_0(t))$ since by definition $\bar{h}_0(t) = pf_0(t)$, showing $(t, f_0(t))$ is always in B' . The map $h_0\phi$ takes t into the point $h_0(t, f_0(t)) = f_0(t)$ of B by its definition so that $ph_0(t, f_0(t)) = \bar{h}_0(t)$. If one now defines $f(t, \tau) = h(\phi(t), \tau)$ using the covering homotopy h of $\bar{h}$, then f is a homotopy

of f_0 which covers $\bar{h}$, the homotopy of pf_0 ; necessarily $pf = \bar{h}$, and f is stationary with $\bar{h}$. This concludes the proof of the second covering homotopy theorem.

$$\begin{array}{ccc}
 B' \times I & \xrightarrow{h} & B \\
 \uparrow \phi \times 1 & \nearrow f & \downarrow p \\
 I^n \times I & \xrightarrow{\bar{h}} & X
 \end{array}$$

Now we recall that $\bar{h}$ was the homotopy connecting $pf_0 = \bar{h}_0$ with the constant map in $F^n(X, A, x_0)$. One would like to show that the covering homotopy f is in $F^n(B, B_0, b_0)$. To this end, consider $f(J^{n-1} \times I)$; since $\bar{h}(J^{n-1} \times I) = x_0$, $\bar{h}$ is stationary over this domain, hence $f(J^{n-1} \times I) = f_0(J^{n-1}) = b_0$. Next since $\bar{h}(I^{n-1} \times I) \subset A$, then $f(I^{n-1} \times I)$ has its image under p in A , hence $f(I^{n-1} \times I) \subset p^{-1}(A) = B_0$. Thus it is true that f is a homotopy in $F^n(B, B_0, b_0)$, of f_0 into the map $f_1: (I^n, I^{n-1}, J^{n-1}) \rightarrow (Y_0, Y_0, b_0)$. We see the last statement by noting $pf_1(I^n \times 1) = \bar{h}(I^n \times 1) = x_0$, so $f_1(I^n \times 1) \subset p^{-1}(x_0) = Y_0$; similarly $f_1(I^{n-1} \times 1) \subset Y_0$.

$$\ker(p_*) = 0$$

Now remember we want to show $\ker(p_*)$ has only the identity element. We would show this if $f_1 \in F^n(B, B_0, b_0)$ is homotopic to the constant map. Define the map

$$k(t_1, \dots, t_n, \tau) = (t_1, \dots, t_{n-1}, (1-\tau)t_n + \tau)$$

of $I^n \times I$ into I^n . For $\tau=0$ it maps I^n into I^n as the identity; for $\tau=1$ it maps I^n into the face $t_n=1$. Thus k is a homotopy of the cube I^n into the face $t_n=1$. The face $t_n=1$ is contained in J^{n-1} , and k maps J^{n-1} into itself. One now defines $k'(t, \tau) = f_1(k(t, \tau))$. We see that $k'(t, 0) = f_1(t)$ and $k'(t, 1) = f_1(J^{n-1}) = b_0$. In other words k' is a homotopy of f_1 into the constant map in $F^n(B, B_0, b_0)$.

To sum up, one started with an element $[f_0] \in \pi_n(B, B_0, b_0)$, a representative $f_0 \in F^n(B, B_0, b_0)$, and it was assumed that $p_*[f_0] = [e]$, so $pf_0 \sim e$. One found that $f_0 \sim f_1$, and further that $f_1 \sim e$. Putting these together, one concludes $f_0 \sim e$ in $F^n(B, B_0, b_0)$, or $[f_0] = [e]$. In other words $\ker(p_*)$ contains only the element $[e]$.

$$\text{im}(p_*) = \pi_n(X, A, x_0)$$

We have proved now half the result; the second half of the result requires showing $\text{im}(p_*) = \pi_n(X, A, x_0)$. So let $[\bar{h}_0]$ be any element of $\pi_n(X, A, x_0)$, and $\bar{h}_0$ a representative of $[\bar{h}_0]$. Use the map

$$k(t_1, \dots, t_n, \tau) = (t_1, \dots, t_{n-1}, (1-\tau)t_n + \tau)$$

to define $\bar{h}(t, \tau) = \bar{h}_0(k(t, \tau))$. Note $\bar{h}(t, 0) = \bar{h}_0(t)$, $\bar{h}_1(t) = \bar{h}(t, 1) = \bar{h}_0(J^{n-1}) = x_0$, and $\bar{h}(J^{n-1} \times I) = x_0$, $\bar{h}(I^{n-1} \times I) \subset X$. So $\bar{h}$ is a homotopy of $\bar{h}_0$ into the constant map; however since $\bar{h}(I^{n-1} \times I)$ is not necessarily in A , then $\bar{h}$ is not in general a homotopy in $F^n(X, A, x_0)$. Let $h_1: I_n \rightarrow B$ be $h_1(t) = b_0$; then $ph_1 = \bar{h}_1 \sim \bar{h}_0$. Using the second

covering homotopy theorem, there exists a homotopy h of h_1 which covers $\bar{h}$, and which is stationary with $\bar{h}$. The homotopy h satisfies $ph = \bar{h}$ so $h(t, 1) = h_1(t)$, and this is homotopic to a map $h_0(t) = h(t, 0)$ with $ph_0(t) = \bar{h}_0(t)$. Because $\bar{h}(J^{n-1} \times I) = x_0$, we have that $h(J^{n-1} \times I) = h_0(t) = b_0$. Now $ph_0(I^{n-1}) = \bar{h}_0(I^{n-1}) \subset A$, so that $h_0(I^{n-1}) \subset B_0$. We therefore deduce that $h_0 \in F^n(B, B_0, b_0)$, and h_0 represents $[h_0]$ whose image $p_*[h_0] = [\bar{h}_0]$. This shows for each $[\bar{h}_0]$ there exists an $[h_0]$ whose image under p_* gives back $[\bar{h}_0]$: thus p_* is onto.

E. Exact Sequence for Bundles

A consequence of the last result when $A = \{x_0\}$ is that $\pi_n(B, Y_0, y_0) \approx \pi_n(X, x_0)$ (" $\approx$ " means "is isomorphic to") under p_* where $Y_0 = p^{-1}(x_0)$. Now consider a segment of the homotopy sequence of (B, Y_0, b_0) :

$$\pi_n(Y_0, b_0) \xrightarrow{1_*} \pi_n(B, b_0) \xrightarrow{j_*} \pi_n(B, Y_0, b_0) \xrightarrow{\partial} \pi_{n-1}(Y_0, b_0).$$

By using the isomorphism p_* we can effectively replace $\pi_n(B, Y_0, b_0)$ by $\pi_n(X, x_0)$. Define $i: (Y_0, b_0) \rightarrow (B, b_0)$ (inclusion map), $j: (B, b_0, b_0) \rightarrow (B, Y_0, b_0)$ and $\Delta = \partial p_*^{-1}: \pi_n(X, x_0) \rightarrow \pi_{n-1}(Y_0, b_0)$. Then deleting the base point to make the notation simpler, one obtains the homotopy sequence of the bundle B:

$$\begin{aligned} \dots \rightarrow \pi_n(Y_0) &\xrightarrow{1_*} \pi_n(B) \xrightarrow{p_*} \pi_n(X) \xrightarrow{\Delta} \pi_{n-1}(Y_0) \rightarrow \\ \dots \rightarrow \pi_2(X) &\xrightarrow{\Delta} \pi_1(Y_0) \xrightarrow{1_*} \pi_1(B) \xrightarrow{p_*} \pi_1(X) \end{aligned}$$

The useful result for later calculations is that the homotopy

sequence of a bundle is exact; this follows since all we have done is replace $\pi_n(B, Y_0)$ in the homotopy sequence of (B, Y_0, b_0) by $\pi_n(X)$, an isomorphic group. The other thing that has been done is to add the term $\pi_1(X)$; this means exactness has still to be proved for the link

$$\pi_1(Y_0) \xrightarrow{i_*} \pi_1(B) \xrightarrow{p_*} \pi_1(X)$$

To show exactness we show $\text{im}(i_*) = \ker(p_*)$, which requires demonstrating $\text{im}(i_*) \subset \ker(p_*)$ and $\ker(p_*) \subset \text{im}(i_*)$. The first is equivalent to showing $p_* i_*$ is the trivial map, i.e. projects $\pi_1(Y_0)$ into the identity of $\pi_1(X)$. Any closed curve in Y_0 is projected by p into the point x_0 , so necessarily $p_* i_*$ is trivial.

To show $\ker(p_*) \subset \text{im}(i_*)$, take a closed curve C in B with base at b_0 , and whose projection pC is homotopic to a constant keeping its end points fixed. Thus C is a representative map of $\ker(p_*)$.

By arguments now familiar to us, one constructs a covering homotopy of C of $I \times I \rightarrow B$ which contracts C into a curve C' lying in Y_0 , since pC is contracted to the constant curve at x_0 . The curve C' is a representative of $\pi_1(Y_0)$ whose image under i_* gives $[C]$; this shows $\ker(p_*) \subset \text{im}(i_*)$. This completes the proof that the homotopy sequence of a bundle is exact.

Bibliography

N. Bourbaki, op. cit.

S.-T. Hu, Homotopy Theory (Academic Press, N. Y., 1959).

N. Steenrod, op. cit.

VIII. THE HOMOLOGY GROUPS

We are interested in finding the possible fibre bundle structures over a given base space X having a specified (internal symmetry) group G . In a previous lecture we formulated [Wu and Yang's] description of Dirac monopoles in this language. The problem of characterizing the monopole solutions was reduced to that of finding all the bundle structures over a sphere. The classification problem of finding all bundle structures with a given base X and given group G is also a central problem of fibre bundle theory. For the monopole example, X is a 2-sphere, S^2 . Whenever $X = S^n$, it can be shown that the classification problem reduces to the calculation of $\pi_{n-1}(G)$ [Steenrod, Sec. 18]. There is no general algorithm for computing the resultant homotopy groups. One powerful tool which mathematicians have is homology theory. One can assign to any topological space a set of commutative groups, called homology groups, which are topological invariants. These groups are effectively computable for most of the spaces which occur in physics. Certain of these groups are isomorphic to certain of the homotopy groups. In a roundabout fashion one can therefore obtain information about the homotopy groups.

This lecture is meant as a brief introduction to the literature on homology theory. We will also indicate the relationship which exists between homology and homotopy theory. For the intuitive notions of homology, the elementary textbook by [John Hocking and Gail Young] is recommended. [Samuel Eilenberg and Norman Steenrod] give one of the best books for serious study by the mathematically minded. The relevant statements connecting homology to homotopy theory can be found in [Steenrod's] book on fibre bundles on which we have based past lectures; one should also consult the article on algebraic

topology given in the Encyclopedia Britannica. All of these sources have been used as a basis for the present chapter.

A. Simplexes and mod 2 Homology Groups

The idea of this lecture is to show by doing. Homology theory has many elements, no one of which is difficult, and each of which is constructive (in the sense one can compute). The proof that one obtains topological invariants, and that one can carry out the constructions for any arbitrary topological space is difficult and we defer this point to the experts. In all cases we assume the result of the constructions will lead to topological invariants, and allow ourselves the freedom from the outset to choose the simplest homeomorphic representative of a given space. For example we will prefer the triangle to a square or circle or simple closed curve, the tetrahedron to a cube or ball. So to begin the discussion we consider as primary objects for study, certain geometric shapes called simplexes.

1. Simplexes

Let $A = \{v_0, \dots, v_k\}$ be $k+1$ vertices in n -dimensional Euclidean space E^n , $n \geq k$. Suppose further that as vectors, the v_i are linearly independent. Then the k -simplex is the set of all points of the hyperplane H^k containing the vertices in A in which the barycentric coordinates $f_j \geq 0$, where each point $x \in H^k$ is represented by $x = \sum_{i=0}^k f_i v_i$ with $\sum_{i=0}^k f_i = 1$. The 0-simplex consists of a single vertex, v_0 . The 1-simplex

consists of the two points v_0 and v_1 and the line joining them. The 2-simplex

consists of the 3 vertices v_0 , v_1 , v_2 ; the lines v_0v_1 , v_1v_2 , v_2v_0 ; and the points in the triangle. The 3-simplex by similar arguments is the points in the tetrahedron with vertices

v_0 , v_1 and v_2 , as well as the boundary faces, and edges.

The simplex can be considered a topological space using the subspace topology for a subset of R^n . The open simplex is the subspace of a simplex with $f_i > 0$, $i=0,1,\dots,k$. The face of a simplex has some $f_i = 0$. The generic symbol for a k -simplex is s^k , and a specific k -simplex with vertices $v_0, \dots, v_k$ is denoted $\langle v_0, \dots, v_k \rangle$.

2. Complexes

Using simplexes as basic building blocks, more complex spaces can be constructed. The rules for joining two simplexes s^m and s^n , $m \leq n$ are that either

$$(1) \quad s^m \cap s^n = \emptyset$$

$$(2) \quad s^m \cap s^n = s^k, \quad k \leq m$$

and s^k is a "face" of both s^m
and s^n .

If two simplexes are joined according to these rules, they are said to be properly joined. A simplicial complex K is a finite collection of properly joined geometric simplexes, such that if s^n is any simplex of K then every face of s^n also belongs to K . A large class of topological spaces, including differentiable manifolds, are homeomorphic to the points of some simplicial complex. When this is true the space is triangulable.

4. Chain Groups

The simplest homology theory is "mod 2 homology theory," and is based on using the integers mod 2. The basic idea of the general constructions are contained already in this simple case. As our first example consider a complex K (unless otherwise stated, complex means simplicial complex) consisting of a 2-simplex (triangle) $\langle v_0 v_1 v_2 \rangle$ and all its faces:

$$K = \{ \langle v_0 v_1 v_2 \rangle, \langle v_1 v_2 \rangle, \langle v_2 v_0 \rangle, \langle v_0 v_1 \rangle, \langle v_0 \rangle, \langle v_1 \rangle, \langle v_2 \rangle \} .$$

Form the group $C_2(K, J_2)$ of 2-chains of the complex K relative to the integers mod 2 ($J_2 = \{0, 1\}$) by the two elements $\{1v_0 v_1 v_2, 0v_0 v_1 v_2\}$. Group addition is commutative and defined by

$$1v_0 v_1 v_2 + 1v_0 v_1 v_2 = 0v_0 v_1 v_2$$

$$0v_0 v_1 v_2 + 1v_0 v_1 v_2 = 1v_0 v_1 v_2$$

$$0v_0 v_1 v_2 + 0v_0 v_1 v_2 = 0v_0 v_1 v_2 .$$

In other words, the symbols " $v_0 v_1 v_2$ " are treated as a dummy label " x ," and group addition is a question only of the coefficient. A graphic way to write these elements is

and so addition is

where $l \equiv m+n \pmod{2}$.

The group of 1-chains $C_1(K, J_2)$ of K relative to the group J_2 is determined by the set of generators $\{gv_0v_1, gv_1v_2, gv_2v_0; g=0,1\}$. The group of 0-chains $C_0(K, J_2)$ of K relative to the groups J_2 is determined by the set of generators $\{v_0, v_1, v_2; g=0,1\}$. In both cases addition is commutative and defined in the same manner as for $C_2(K, J_2)$.

4. Boundary homomorphism ∂

Given an element $gv_0v_1v_2 \in C_2(K, J_2)$ one can obtain an element of $C_1(K, J_2)$ which in some sense represents the boundary of $v_0v_1v_2$. Define the boundary homomorphism $\partial: C_2(K, J_2) \rightarrow C_1(K, J_2)$ by

$$\partial gv_0v_1v_2 = gv_1v_2 + gv_2v_0 + gv_0v_1$$

for $g = 0, 1$. This operation does indeed preserve sums, so that it is a homomorphism. If $g=1$, then the three terms above on the right side of the equation are the boundary of the triangle. If the boundary operator is applied to gv_0v_1 one

obtains

$$\partial g v_0 v_1 = g v_1 + g v_0$$

and similarly for any simplex $[\partial g v_0 \dots v_k = g v_1 \dots v_k + \dots + g v_0 v_1 \dots \hat{v}_i \dots v_k + \dots + g v_0 \dots v_{k-1}]$, where $\hat{v}_i$ indicates that this vertex is missing, $i=0, \dots, k$. Applying ∂ twice to $g v_0 v_1 v_2$ gives

$$\partial^2 g v_0 v_1 v_2 = (g+g) (v_1 + v_2 + v_0) = 0,$$

since $g+g \equiv 0 \pmod{2}$.

5. p -boundary and p -cycle groups

These considerations show

$$\partial: C_2(K, J_2) \rightarrow C_1(K, J_2)$$

is a homomorphism and that $\partial C_2(K, J_2)$ is in the kernel of the map

$\partial: C_1(K, J_2) \rightarrow C_0(K, J_2)$. The definitions obviously extend to any p -simplex in K for any complex K . One defines a group $C_p(K, J_2)$ of p -chains of K relative to the group J_2 , and using the boundary homomorphism twice, one finds $\partial^2 = 0$. For the general case one ordinarily adopts the notation

$$B_p(K, J_2) = \partial C_{p+1}(K, J_2)$$

$$Z_p(K, J_2) = \ker(\partial: C_p(K, J_2) \rightarrow C_{p-1}(K, J_2))$$

Each p -chain in $B_p(K, J_2)$ is called a p -boundary, and B_p is the group of p -boundaries. Each p -chain in Z_p is called a p -dimensional cycle and is mapped into 0 by ∂ . Another way to state these facts is to note

$B_p(K, J_p) \subset Z_p(K, J_p)$ and $Z_p(K, J_p) \subset C_p(K, J_p)$. The first follows since $\partial^2 C_{p+1} = 0$,

and so $\partial C_{p+1} \subset \ker \partial$.

6. p -homology groups

Using this general notation, we return to our simple example of K being a 2-simplex and all its faces. There are no 2-boundaries since there are no 3-simplexes in K and hence no 3-chains: $B_2(K, J_2) = 0$. Only $0 \in C_2(K, J_2)$ is a 2-dimensional cycle since $\partial g v_0 v_1 v_2 = g(v_0 v_1 + v_1 v_2 + v_2 v_0) = 0$ implies $g \equiv 0 \pmod{2}$: $Z_2(K, J_2) = 0$. For $p=1$, $B_1(K, J_2) = \{0, v_0 v_1 + v_1 v_2 + v_2 v_0\}$ has two elements. Graphically B_1 consists of the triangle

with 1's along the boundary or 0's along the boundary of the 2-simplex. The 1-dimensional cycles Z_1 consist of either 0 or the element $v_1 v_2 + v_2 v_0 + v_0 v_1$. One checks that ∂ on either element gives zero. Finally we see for $p=0$, the generators of $B_0(K, J_2) = \{0, v_0 + v_1, v_1 + v_2, v_2 + v_0\}$ and $Z_0(B, J_2) = C_0(K, J_2)$. The last statement is obtained by defining $\partial v_i = 0$, $i = 0, 1, 2$.

For the triangle, a summary of groups of K (up to isomorphism)

p	C_p	Z_p	B_p	Z_p/B_p
0	$J_2 \oplus J_2 \oplus J_2$	$J_2 \oplus J_2 \oplus J_2$	$J_2 \oplus J_2$	J_2
1	$J_2 \oplus J_2 \oplus J_2$	J_2	J_2	0
2	J_2	0	0	0
3	0	0	0	0

Using group isomorphism, the above table summarizes our simple computations. The notation $J_2 \oplus J_2$ indicates a commutative group with two generators, etc. One notices that Z_p and B_p are not always equal. It is generally true that both are commutative groups, and $B_p \subset Z_p$. From this it follows that one can construct a group from the factor set Z_p/B_p . [As usual one constructs equivalence classes: $z_1 \equiv z_2 \pmod{B_p}$ iff $z_1 - z_2 \in B_p$. A typical element of Z_p/B_p is also written as $z_1 \oplus B_p$ for any $z_1 \in Z_p$. The addition of two such elements is defined as $(z_1 \oplus B_p) + (z_2 \oplus B_p) = (z_1 + z_2) \oplus B_p$. This operation defines a commutative group structure on Z_p/B_p]. The commutative group formed on Z_p/B_p is called the mod 2 homology group of dimension p. For the complex K under consideration, the only non-trivial mod 2 homology group is $H_0(K, J_2) \approx J_2$ (read " $\approx$ " as "is isomorphic to").

The amazing thing about the above construction is that the mod 2 homology groups are topological invariants of the triangle from which the complex K was constructed, even though the chain groups, boundary groups and group of cycles are not invariants. For example, we could have split our given triangle up into several triangles, and so the chain

groups of each dimension $p = 0, 1, 2$ would be correspondingly larger. One can verify nevertheless that the homology groups obtained are unchanged.

7. Examples

The general method of calculating homology groups of a topological space X is to triangulate it, namely find a simplicial complex

which is homeomorphic to X . This may not always be possible, though it is in most interesting cases. The above procedure leads one to the evaluation of homology groups. This procedure will be illustrated below by additional examples. A word of warning though: spaces with the same homology groups are not necessarily topologically equivalent. A cube in any number of dimensions is homologically trivial, i.e. the only non zero homology group is H_0 which is isomorphic to J_2 .

As our next example, consider a Möbius strip. A convenient picture is to take a rectangle and identify one set of opposite edges. One then breaks the figure into triangles as shown below.

In this triangulation, the Möbius strip has 6 2-simplexes, 12 1-simplexes and 6 0-simplexes; there are no simplexes of dimension 3 or greater. Let K be the complex formed by all the simplexes of the Möbius strip. All the groups C_3, Z_3, B_3, H_3 are 0. It then follows that $B_2(K, J_2) = 0$, since $B_2 = \partial C_3$. The chain group $C_2(K, J_2)$ of order 2 is not zero. A typical element is

$$\alpha = \varepsilon_1 v_0 v_1 v_3 + \varepsilon_2 v_0 v_2 v_3 + \varepsilon_3 v_2 v_3 v_4 + \\ + \varepsilon_4 v_2 v_4 v_5 + \varepsilon_5 v_4 v_0 v_5 + \varepsilon_6 v_5 v_1 v_0, \quad \varepsilon_k = 0, 1 \text{ for } k=1, \dots, 6$$

Graphically we can think of α as

To determine the $\ker(\partial) = Z_2(K, J_2)$, we determine the values of g_k for which $\partial\alpha = 0$. This gives rise to the equation

$$\partial\alpha = 0$$

$$= (g_1 + g_6) v_0 v_1 + g_1 v_1 v_3 + (g_1 + g_2) v_3 v_0$$

$$+ g_2 v_0 v_2 + (g_2 + g_3) v_2 v_3 + g_3 v_3 v_4$$

$$+ (g_3 + g_4) v_2 v_4 + g_4 v_2 v_5 + (g_4 + g_5) v_4 v_5$$

$$+ g_5 v_4 v_0 + (g_5 + g_6) v_0 v_5 + g_6 v_5 v_1$$

which can be satisfied only if each $g_k = 0, k=1, \dots, 6$. It then follows that $Z_2(K, J_2) = 0$ and so $H_2(K, J_2) = 0$.

Now we consider the groups of dimension 1. The 1-chains have the general element α which we write graphically as

Explicitly, we have

$$\alpha = g_1 v_1 v_3 + g_2 v_2 v_4 + g_3 v_3 v_5$$

$$+ g_4 v_4 v_6 + g_5 v_5 v_0 + g_6 v_6 v_1$$

$$+ \bar{g}_{16} v_0 v_1 + \bar{g}_{12} v_3 v_0 + \bar{g}_{23} v_5 v_3$$

$$+ \bar{g}_{34} v_2 v_4 + \bar{g}_{45} v_4 v_5 + \bar{g}_{56} v_6 v_5$$

The 1-boundaries, elements of $B_1(K, J_2)$, are the images of $C_2(K, J_2)$ by

∂ . Taking a general $\beta \in C_2(K, J_2)$, we find $\alpha = 0$ when

$$\bar{g}_{1j} \equiv g_1 + g_j \pmod{2}, \quad 1, j=1, \dots, 6. \quad \text{In other words } Z_1(K, J_2) \text{ has 6}$$

generators. To obtain $Z_1(K, J_2)$ we take a general $\alpha \in C_1(K, J_2)$ and require

$$\partial\alpha = 0:$$

$$\partial\alpha = 0$$

$$= (g_2 + g_5 + \bar{g}_{16} + \bar{g}_{12} + \bar{g}_{56})v_0 + (g_1 + g_6 + \bar{g}_{16})v_1$$

$$+ (g_2 + g_4 + \bar{g}_{23} + \bar{g}_{34})v_2 + (g_1 + g_3 + \bar{g}_{12} + \bar{g}_{23})v_3$$

$$+ (g_3 + g_5 + \bar{g}_{34} + \bar{g}_{45})v_4 + (g_4 + g_6 + \bar{g}_{45} + \bar{g}_{56})v_5$$

The solution of these equation is

$$\bar{E}_{16} = g_1 + g_6$$

$$\bar{E}_{23} = g_2 + g_3 + (\bar{E}_{12} + g_1 + g_6) = g_1 + g_3 + \bar{E}_{12}$$

$$\bar{E}_{34} = g_3 + g_4 + (\bar{E}_{12} + g_1 + g_6)$$

$$\bar{E}_{45} = g_4 + g_5 + (\bar{E}_{12} + g_1 + g_6)$$

$$\bar{E}_{56} = g_5 + g_6 + (\bar{E}_{12} + g_1 + g_6)$$

$$\bar{E}_{12} = g_1 + g_2 + (\bar{E}_{12} + g_1 + g_6)$$

Therefore a general element of $\pi_1(K, J_2)$ is

$$\begin{aligned} \alpha = & g_1 (v_0 v_1 + v_1 v_3 + v_3 v_2 + v_2 v_4 + v_4 v_5 + v_5 v_0) \\ & + g_2 (v_0 v_2 + v_2 v_4 + v_4 v_5 + v_5 v_0) \\ & + g_3 (v_2 v_3 + v_3 v_4 + v_4 v_2) + g_4 (v_5 v_2 + v_2 v_4 + v_4 v_5) \\ & + g_5 (v_4 v_0 + v_0 v_5 + v_5 v_4) + g_6 (v_5 v_1 + v_1 v_0 + v_0 v_5) \\ & + \bar{E}_{12} (v_3 v_0 + v_0 v_5 + v_5 v_4 + v_4 v_2 + v_2 v_3) \end{aligned}$$

Note that each term is literally a cycle around the Möbius strip. There are basically two kinds of cycles: those which are closed and bound a triangle and those which cut the Möbius strip in half and do not form a closed boundary. Those cycles which form a boundary in the above non-technical sense

generate $H_1(K, J_2)$, and those which do not generate $H_1(K, J_2)$. More precisely, an element of $Z_1(K, J_2)$ for the Möbius strip differs from an element of $H_1(K, J_2)$ by the addition of the 1-cycle $v = \overline{v_0 v_1} + \overline{v_1 v_2} + \overline{v_2 v_3} + \overline{v_3 v_4} + \overline{v_4 v_5} + \overline{v_5 v_0}$. The figure below illustrates this cycle, which is clearly not a boundary which has an inside and an outside. The factor group $H_1(K, J_2)/Z_1(K, J_2) = H_1(K, J_2)$ is generated by the single element $[\overline{v_0 v_1}]$.

The table below summarizes the various groups of the Möbius strip. The only difficult part in obtaining the 0-dimensional groups is

P	C_p	Z_p	B_p	H_p
0	$\#6J_2$	$\#6J_2$	$\#5J_2$	J_2
1	$\#12J_2$	$\#7J_2$	$\#6J_2$	J_2
2	$\#6J_2$	0	0	0

to find $B_0(K, J_2) = \partial C_1(K, J_2)$. One takes a general element α of C_1 and computes $\partial\alpha$:

$$\begin{aligned}
\partial \alpha &= (\bar{g}_{16} + g_1 + g_6)(v_0 + v_1) \\
&+ [(\bar{g}_{16} + g_1 + g_6) + (\bar{g}_{12} + g_1 + g_2)] [v_0 + v_5] \\
&+ [(\bar{g}_{16} + g_1 + g_6) + (\bar{g}_{12} + g_1 + g_2)] [v_2 + v_3] \\
&+ [(\bar{g}_{16} + g_1 + g_6) + (\bar{g}_{12} + g_1 + g_2)] [v_2 + v_4] \\
&+ [(\bar{g}_{16} + g_1 + g_6) + (\bar{g}_{12} + g_1 + g_2)] [v_4 + v_5]
\end{aligned}$$

where we have grouped terms to emphasize the independent elements. There are 5 independent generators of $B_0(K, J_2)$, so $B_0(K, J_2) \approx \mathbb{Z}_2^5 = J_2 \oplus J_2 \oplus J_2 \oplus J_2 \oplus J_2$.

We see from the last example that the Möbius strip has a non-zero mod 2 homology group $H_1(K, J_2) \approx J_2$. We also see some reasons for the terminology of p-boundary and p-cycle. We could continue in like fashion to construct mod 2 homology groups for lots of other examples. We would find that spaces such as the cylinder and Klein bottle which are clearly quite different have the same mod 2 homology groups. This leads one to look for more distinguishing structures; in particular one generalizes the coefficient group from the group of integers mod 2 to an arbitrary abelian (i.e. commutative) group G . For most purposes G can be thought of as the group of integers $\mathbb{Z}$. To avoid confusion with the symbol $\mathbb{Z}_p$ for p-boundaries, we will continue to use the notation G for the coefficient group, though we will mainly consider G to be the group of integers.

B. Oriented Simplexes and Homology groups with coefficient group G

1. Oriented simplex and complex

To extend the construction of chain groups C_p and boundary homomorphisms

to arbitrary coefficient groups G , one introduces the idea of an oriented simplex and an oriented complex. Let

$$\sigma^p = \langle v_0 \dots v_p \rangle$$

be a p -simplex. Consider a fixed ordering of the letters $v_0 < v_1 < \dots < v_p$. Let π be a permutation of the integers $0, 1, \dots, p$. One says that

$$\langle v_{\pi(0)} \dots v_{\pi(p)} \rangle = +\sigma^p$$

is a positively oriented simplex if π is an even permutation, and

$$\langle v_{\pi(0)} \dots v_{\pi(p)} \rangle = -\sigma^p$$

is a negatively oriented simplex if π is an odd permutation. In diagrams we sometimes indicate the orientation by arrows:

$$\langle v_0 v_1 \rangle = +\sigma^1$$

$$\langle v_0 v_1 v_2 \rangle = +\sigma^2$$

We will also write an oriented simplex $\pm \sigma^2$ as $\pm v_0 v_1$, $\pm \sigma^3$ as $\pm v_0 v_1 v_2$, etc.

The boundary operator on $\langle v_0 v_1 \rangle = +\sigma^1$ gives 0-simplexes, and can be defined to give a natural ordering,

$$\partial(+v_0 v_1) = +v_0 - v_1$$

since v_0 is the tail of the arrow, and v_1 is the head.

For a 2-simplex, the boundary of $+v_0 v_1 v_2$ is defined to be

$$\partial(+v_0 v_1 v_2) = +v_1 v_2 - v_0 v_2 + v_0 v_1$$

If we perform ∂ again, we get 0:

$$\partial^2(+v_0 v_1 v_2) = 0v_0 + 0v_1 + 0v_2 = 0$$

The orientations of the boundary are again natural ones. The two-simplex has an orientation indicated by the arrows,

which shows that the 1-simplexes $+v_1 v_2$ and $+v_0 v_1$ agree with that orientation, and $+v_0 v_2$ has the opposite orientation. Hence the minus sign for the $v_0 v_2$ segment.

Now a more complicated example is provided by $+v_0 v_1 v_2 v_3$, an oriented

3-simplex.

The natural boundary operation motivated by the previous example is

$$\begin{aligned}\partial(+v_0v_1v_2v_3) &= +v_1v_2v_3 - v_0v_2v_3 \\ &\quad + v_0v_1v_3 - v_0v_1v_2\end{aligned}$$

an alternating sum. This guarantees that $\partial^2(+v_0v_1v_2v_3) = 0$. Consider $+v_1v_2v_3$, the positively oriented 2-simplex. What is the orientation of $\langle v_0v_2v_3 \rangle$ which is adjacent to it? They have one face in common, and in agreement with our previous examples we have chosen $\langle v_0v_2v_3 \rangle$ to have the opposite orientation, $\langle v_0v_2v_3 \rangle = -v_0v_2v_3$ as a result of the boundary operator.

2. Boundary operator ∂

We now state the general rule for the boundary operator on a positively oriented p -simplex:

$$\partial(+v_0 \dots v_p) = \sum_{j=0}^p (-)^j v_0 \dots \hat{v}_j \dots v_p$$

where $\hat{v}_j$ indicates this vertex is missing from the simplex. The operation of the boundary operator twice gives

$$\partial^2(+v_0 \dots v_p) = \sum_{j,k=0}^p (-)^{j+k} v_0 \dots \hat{v}_j \dots \hat{v}_k \dots v_p$$

$$= 0$$

The definition of ∂ insures that two $(p-1)$ -simplexes that have one face in common will have an opposite orientation. One will also find in the literature the definition of incidence numbers which expresses the same idea. One defines the incidence number $[\sigma^m, \sigma^{m-1}]$ for two simplexes of an oriented simplicial complex K (i.e. a complex such that each simplex in K has an orientation) by

$$[\sigma^m, \sigma^{m-1}] = 0 \quad \sigma^{m-1} \text{ is not a face of } \sigma^m \text{ in } K$$

$$[\sigma^m, \sigma^{m-1}] = \pm 1 \quad \sigma^{m-1} \text{ is a face of } \sigma^m \text{ in } K.$$

One chooses $+1$ if the orientation of σ^{m-1} is the natural one induced by σ^m , and -1 for the other possibility. One then defines the boundary operator as

$$\partial \sigma^m = \sum_{\sigma^{m-1}} [\sigma^m, \sigma^{m-1}] \sigma^{m-1}$$

where the sum goes over all the $(n-1)$ oriented simplexes in the complex K . When the oriented complex K is an oriented p -simplex and all its faces, and when we use the natural orderings to write each σ^{m-1} , we recover the first definition of ∂ . The above formula generalizes ∂ to an arbitrary oriented complex K .

3. Chain group

We can now proceed to define for any oriented complex K , a p -dimensional chain with coefficients in G . To be precise we define a p -chain as a function c_p from the oriented p -simplexes in K with values in G :

$$c_p(\sigma^p) = g.$$

The function satisfies

$$c_p(-\sigma^p) = -c_p(+\sigma^p) .$$

The addition of two chains is defined as

$$(c_p + c'_p)(\pm\sigma^p) = c_p(\pm\sigma^p) + c'_p(\pm\sigma^p) .$$

This addition on the set of all p -chains defines a group $C_p(K, G)$ which is necessarily commutative. We write the generators of the group the way we did for mod 2 homology theory:

$$gv_0 \dots v_p \in C_p(K, G) .$$

This generator is defined as the function satisfying

$$gv_0 \dots v_p (\pm v_0 \dots v_p) = \pm g$$

$$gv_0 \dots v_p (\pm\sigma^p) = 0 \text{ if } \begin{cases} \pm\sigma^p \neq \pm v_0 \dots v_p & \text{and} \\ \pm\sigma^p \in K \end{cases} .$$

Any element α of $C_p(K, G)$ can be written as

$$\alpha = \sum_{i_0 \dots i_p} g_{i_0 \dots i_p} v_{i_0} \dots v_{i_p}$$

where $g_{i_0 \dots i_p} \in G$. The boundary operator on $C_p(K, G)$ is defined by its action on each of the generators

$$\partial(gv_0 \dots v_p) = \sum_{j=0}^p (-)^j gv_0 \dots \hat{v}_j \dots v_p .$$

It is an immediate consequence that

$$\partial^2 g v_0 \dots v_p = 0.$$

4. p -boundary, p -cycle and p -homology groups

The procedure now parallels that for mod 2 homology theory. We have for each dimension p , chain groups $C_p(K, G)$ and boundary homomorphisms

$$\partial: C_p(K, G) \rightarrow C_{p-1}(K, G).$$

As before, the image $\partial C_{p+1}(K, G) = B_p(K, G)$ is the set of p -boundaries; the set of $C_p(K, G)$ whose image is 0 is the set $Z_p(K, G)$ of p -cycles. Necessarily $B_p(K, G) \subset Z_p(K, G)$, and because Z_p is commutative, the factor set $Z_p(K, G)/B_p(K, G)$ has a group structure which is commutative, which one calls the homology group $H_p(K, G)$ of order p of K over G . The homology groups are topological invariants. We shall illustrate the preceding discussion with two examples.

5. Example

As our first example we choose the Möbius strip again. The calculation

proceeds identically to the mod 2 theory, except we have to keep track of signs. There are 6 oriented 2-simplexes, 12 oriented 1-simplexes and 6

oriented 0-simplexes. The chain groups $C_p(K, G) = 0$ for $p \geq 3$, so $B_p(K, G) = Z_p(K, G) = H_p(K, G) = 0$ also for $p \geq 3$. Since $C_3(K, G) = 0$ it follows that $B_2(K, G) = \partial C_3(K, G) = 0$. Now let us compute the 2-cycles, and the 1-boundaries. An arbitrary $\alpha \in C_2(K, G)$ has the form

$$\begin{aligned} \alpha = & g_1 v_0 v_1 v_3 + g_2 v_0 v_2 v_3 + g_3 v_2 v_3 v_4 \\ & + g_4 v_2 v_4 v_5 + g_5 v_0 v_4 v_5 + g_6 v_0 v_1 v_5 \end{aligned}$$

By means of the boundary homomorphism ∂ ,

$$\begin{aligned} \partial \alpha = & g_1 (v_0 v_1 + v_1 v_3 + v_3 v_0) + g_2 (v_3 v_0 + v_0 v_2 + v_2 v_3) \\ & + g_3 (v_2 v_3 + v_3 v_4 + v_4 v_2) + g_4 (v_2 v_4 + v_4 v_5 + v_5 v_2) \\ & + g_5 (v_4 v_5 + v_5 v_0 + v_0 v_4) + g_6 (v_0 v_1 + v_1 v_5 + v_5 v_0) \end{aligned}$$

This shows that $B_1(K, G)$ has six independent generators, each of which generates a subgroup isomorphic to G : $B_1(K, G) \approx 6G$. It also shows that $\partial \alpha = 0$ has only the solution $g_1 = \dots = g_6 = 0$, showing $Z_2(K, G) = 0$.

Now the 1-chains $\alpha \in C_1(K, G)$ have the form

$$\begin{aligned} \alpha = & g_1 v_1 v_3 + g_2 v_0 v_2 + g_3 v_3 v_4 + g_4 v_5 v_2 \\ & + g_5 v_0 v_4 + g_6 v_1 v_5 + \bar{g}_{16} v_0 v_1 + \bar{g}_{12} v_3 v_0 \\ & + \bar{g}_{23} v_2 v_3 + \bar{g}_{34} v_2 v_4 + \bar{g}_{45} v_4 v_5 + \bar{g}_{56} v_5 v_0, \end{aligned}$$

so that $C_1(K, G) \approx 12G$. The 0-boundaries are the images

$$\begin{aligned}
\partial\alpha &= (\bar{g}_{16}-g_1-g_6)(v_0-v_1) \\
&+ [(\bar{g}_{56}-g_5-g_6) + (g_{12}-g_1-g_2)][v_5-v_0] \\
&+ [(\bar{g}_{45}-g_4-g_5) + (g_{12}-g_1-g_2)][v_4-v_5] \\
&+ [(\bar{g}_{34}+g_3-g_4) + (g_{12}-g_1-g_2)][v_2-v_4] \\
&+ [(\bar{g}_{23}-g_2-g_3) - (g_{12}-g_1-g_2)][v_2-v_3]
\end{aligned}$$

which is a group with 5 generators each giving a subgroup isomorphic to G . Thus $B_0(K,G) = \oplus 5G$. When $\partial\alpha = 0$ we can read off the constraints for the g 's:

$$\bar{g}_{16} = g_1 + g_6$$

$$\bar{g}_{56} = g_5 + g_6 - (\bar{g}_{12}-g_1-g_2)$$

$$\bar{g}_{45} = g_4 + g_5 - (\bar{g}_{12}-g_1-g_2)$$

$$\bar{g}_{34} = -g_3 + g_4 - (\bar{g}_{12}-g_1-g_2)$$

$$\bar{g}_{23} = g_2 + g_3 + (\bar{g}_{12}-g_1-g_2)$$

The general element of $Z_1(K,G)$ has the form

$$\alpha = \bar{\alpha} + (\bar{g}_{12}-g_1-g_2)(v_2v_3+v_3v_0+v_0v_5+v_5v_4+v_4v_2)$$

where

$$\begin{aligned}\bar{\alpha} = & g_1 (v_0 v_1 + v_1 v_3 + v_3 v_0) + g_2 (v_3 v_0 + v_0 v_2 + v_2 v_3) \\ & + g_3 (v_2 v_3 + v_3 v_4 + v_4 v_2) + g_4 (v_2 v_4 + v_4 v_5 + v_5 v_2) \\ & + g_5 (v_4 v_5 + v_5 v_0 + v_0 v_4) + g_6 (v_0 v_1 + v_1 v_5 + v_5 v_0)\end{aligned}$$

is (because of its form) an element of $B_1(K, G)$. We conclude that $Z_1(K, G) \approx \oplus 7G$ and the factor group $H_1(K, G) \approx G$. Similar considerations can be used to show $C_0(K, G) \approx \oplus 6G$, $Z_0(K, G) \approx \oplus 6G$. Coupling these results with $B_0(K, G) \approx \oplus 5G$ one obtains $H_0(K, G) \approx G$. The table below summarizes the non zero results.

Groups for Möbius strip (up to isomorphism)

p	C_p	Z_p	B_p	H_p
0	$\oplus 6G$	$\oplus 6G$	$\oplus 5G$	G
1	$\oplus 12G$	$\oplus 7G$	$\oplus 6G$	G
2	$\oplus 6G$	0	0	0

6. Euler and Betti numbers

For the Möbius strip, since G could have been arbitrary, we recover the mod 2 homology theory by taking $G = J_2$. Another observation that turns out to be a general result is that $H_p(K, G)$ will contain the direct sum of G β_p times, $\beta_p > 0$ a unique integer. This integer is the Betti number of order p . From above we see $\beta_0 = \beta_1 = 1$ and $\beta_p = 0$, $p \geq 2$. The Euler number is related to these integers. The Euler number of the complex K is

$$\chi(K) = \sum_{i=0}^{\infty} (-1)^i \alpha_i$$

where α_1 is the number of 1-simplexes in the complex K . The sum is finite for finite dimensional complexes. It can be shown that the Euler number is related to the Betti numbers by

$$\chi(K) = \sum_{m=0}^{\infty} (-1)^m \beta_m.$$

Since the homology groups are asserted to be topological invariants, β_p are also topological invariants, and so is $\chi(K)$. For the Möbius strip

$$\chi(K) = \sum (-1)^k \alpha_k = 6 - 12 + 6 = 0 \text{ and } \sum (-1)^k \beta_k = 1 - 1 = 0.$$

7. Another example; torsion numbers

As a final example of computing homology groups, let us consider the Klein bottle. A triangulation is given below. The top and bottom of

the rectangle are identified to make a cylinder; the left and right are identified with a twist so as to make a Klein bottle. Without the twist one would have a torus. There are $\alpha_2 = 18$ 2-simplexes, $\alpha_1 = 27$ 1-simplexes, and $\alpha_0 = 9$ 0-simplexes; $\alpha_p = 0$ for $p \geq 3$. The complex K is the set of all simplexes in the above figure. The Euler number $\chi(K) = 9 - 27 + 18 = 0$,

showing that $\beta_0 - \beta_1 + \beta_2 = 0$. If it is true that $\beta_0 = 1$, then we will have $\beta_1 - \beta_2 = 1$. The group $C_2(K, G)$ has 27 generators, $C_2(K, G) \approx \#27G$, there are no 2-boundaries $B_2(K, G) = \partial C_3(K, G) = 0$, and it's not difficult to verify that $Z_2(K, G) = 0$. Thus we find $H_2(K, G) = 0$ and so $\beta_2 = 0$. Thus if β_3 were equal to 1, we would expect $\beta_1 = 1$. We will construct $H_1(K, G)$ explicitly to verify this expectation and to learn that the Betti number is not sufficient to characterize the homology group. Indeed we show that $H_1(K, G) = G \oplus J_2$, the direct sum of G and the group of integers mod 2, for any group G . The number θ_1 of finitely generated subgroups of $H_p(K, G)$ of fixed order is called a torsion number; it is an invariant of H_p .

Now the 1-boundaries are generated by closed loops around each of the 2-simplexes. This can be verified by taking any element in $C_2(K, G)$,

where we use the figure (vertices v_i labeled by i , and elements $+h_0 v_0 v_1 v_6$ labeled by the triangle with h_0 in the center, etc.) to represent the general group element. The orientation of each triangle is taken to be the natural one, so for example the triangle $\langle v_2 v_6 v_5 \rangle$ contributes $+h_{16} v_2 v_5 v_6$. The 1-boundaries $\bar{\alpha}$ are 1-cycles such that $\bar{\alpha} = \partial \alpha$ for some $\alpha \in C_2(K, G)$. A gen-

eral 1-cycle α' must satisfy $\partial\alpha' = 0$, and will have the form

where the orientation of each 1-simplex is always from small number to large.

The requirement $\partial\alpha' = 0$ leads to 9 constraint equations

$$\epsilon_1 = -\epsilon_3 - \epsilon_8 - \epsilon_{12} - \epsilon_{13} - \epsilon_{17} - \epsilon_{18} - \epsilon_{19} - \epsilon_{22} - \epsilon_{27}$$

$$\begin{aligned} \epsilon_2 = & -\epsilon_3 + \epsilon_6 - \epsilon_8 - \epsilon_{11} - \epsilon_{12} - \epsilon_{15} - \epsilon_{17} + \epsilon_{21} \\ & - \epsilon_{24} + \epsilon_{26} - \epsilon_{20} \end{aligned}$$

$$\epsilon_4 = \epsilon_{10} - \epsilon_{18} - \epsilon_{21} - \epsilon_{26} - \epsilon_{27}$$

$$\begin{aligned} \epsilon_5 = & -\epsilon_6 - \epsilon_{10} + \epsilon_{11} - \epsilon_{13} + \epsilon_{15} \\ & - \epsilon_{19} + \epsilon_{20} - \epsilon_{22} - \epsilon_{23} \end{aligned}$$

$$\varepsilon_7 = -\varepsilon_8 - \varepsilon_{11} - \varepsilon_{12} - \varepsilon_{15} - \varepsilon_{17} - \varepsilon_{20}$$

$$+ \varepsilon_{21} - \varepsilon_{24} - \varepsilon_{25}$$

$$\varepsilon_9 = \varepsilon_{12} + \varepsilon_{13} + \varepsilon_{17} + \varepsilon_{18} + \varepsilon_{19}$$

$$\varepsilon_{14} = -\varepsilon_{13} + \varepsilon_{19} - \varepsilon_{20} + \varepsilon_{22} + \varepsilon_{23}$$

$$\varepsilon_{16} = \varepsilon_{17} + \varepsilon_{20} - \varepsilon_{21} + \varepsilon_{24} + \varepsilon_{25}$$

Thus a 1-cycle has the general form

$$\begin{aligned} \alpha' = & \varepsilon_3 (v_0 v_2 + v_2 v_1 + v_1 v_0) + \varepsilon_6 (v_2 v_7 + v_7 v_1 + v_1 v_2) \\ & + \varepsilon_8 (v_0 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_0) + \varepsilon_{10} (v_6 v_7 + v_7 v_1 + v_1 v_6) \\ & + \varepsilon_{11} (v_7 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_7) \\ & + \varepsilon_{12} (v_3 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_0 + v_0 v_3) \\ & + \varepsilon_{13} (v_3 v_7 + v_7 v_1 + v_1 v_0 + v_0 v_3) \\ & + \varepsilon_{15} (v_4 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_7 + v_7 v_4) \\ & + \varepsilon_{17} (v_3 v_5 + v_5 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_0 + v_0 v_3) \\ & + \varepsilon_{18} (v_3 v_6 + v_6 v_1 + v_1 v_0 + v_0 v_3) \\ & + \varepsilon_{19} (v_3 v_4 + v_4 v_7 + v_7 v_1 + v_1 v_0 + v_0 v_3) \end{aligned}$$

$$\begin{aligned}
& + g_{20} (v_4 v_5 + v_5 v_8 + v_8 v_2 + v_2 v_1 + v_1 v_7 + v_7 v_4) \\
& + g_{21} (v_5 v_6 + v_6 v_1 + v_1 v_2 + v_2 v_8 + v_8 v_5) \\
& + g_{22} (v_0 v_4 + v_4 v_7 + v_7 v_1 + v_1 v_0) \\
& + g_{23} (v_1 v_4 + v_4 v_7 + v_7 v_1) \\
& + g_{24} (v_1 v_5 + v_5 v_8 + v_8 v_2 + v_2 v_1) \\
& + g_{25} (v_2 v_5 + v_5 v_8 + v_8 v_2) + g_{26} (v_2 v_6 + v_6 v_1 + v_1 v_2) \\
& + g_{27} (v_0 v_6 + v_6 v_1 + v_1 v_0) .
\end{aligned}$$

One can see explicitly that each term traces out a closed cycle through the complex. If we impose the constraint that $\alpha' = \bar{\alpha}$ is a 1-boundary so that $\bar{\alpha} = \partial\alpha$, then we have one additional relationship among the g 's,

$$g_{26} + g_{21} - g_3 - g_8 - g_{12} - g_{17} = 0.$$

Moreover, comparing $\bar{\alpha}$ with $\partial\alpha$ we get 27 equations in 18 unknowns. This shows that in general $B_1(K, G)$ has 18 generators, and $B_1(K, G) \approx \oplus 18 G$.

Solving these equations, we obtain the h 's in terms of the g 's:

$$2h_{17} = g_{13} + g_{18} + g_{19} + g_{21} + g_{22} + g_{23} + g_{24} + g_{25} + 2g_{26}$$

$$h_1 = h_{17} + g_{10} - g_{18} - g_{21} - g_{26}$$

$$h_2 = -h_{17} + g_6 - g_{11} - g_{15} - g_{20} - g_{24} - g_{25} - g_{26}$$

$$h_3 = -h_{17} + g_{11} + g_{15} + g_{20} + g_{24} + g_{25} + g_{26}$$

$$h_4 = -h_{17} + g_3$$

$$h_5 = h_{17} + g_{12} + g_{17} - g_{21} - g_{26}$$

$$h_6 = -h_{17} + g_8 + g_{21} + g_{26}$$

$$h_7 = -h_{17} + g_{19} + g_{22} + g_{23} + g_{24} + g_{25} + g_{26}$$

$$h_8 = h_{17} - g_{15} - g_{20} - g_{24} - g_{25} - g_{26}$$

$$h_9 = -h_{17} + g_{20} + g_{24} + g_{25} + g_{26}$$

$$h_{10} = -h_{17} - g_{17} + g_{21} + g_{26}$$

$$h_{11} = -h_{17} + g_{21} + g_{26}$$

$$h_{12} = h_{17} - g_{22} - g_{23} - g_{24} - g_{25} - g_{26}$$

$$h_{13} = -h_{17} + g_{23} + g_{24} + g_{25} + g_{26}$$

$$h_{14} = h_{17} - g_{24} - g_{25} - g_{26}$$

$$h_{15} = -h_{17} + g_{25} + g_{26}$$

$$h_{16} = h_{17} - g_{26}$$

$$h_0 = -h_{17} - g_{27}$$

$$0 = g_{26} + g_{21} - g_3 - g_8 - g_{12} - g_{17} \quad .$$

Note h_{17} which appears is given in terms of g 's whenever the right hand side of the first equation is an even integer. We thus gain criteria for when a general 1-cycle α' is a 1-boundary. An arbitrary element α' of $Z_1(K, G)$ has the form

$$\alpha' = \alpha'' + \gamma$$

where

$$\gamma = (g_{26}g_3g_8g_{12}g_{17}g_{21})(v_2v_6+v_6v_1+v_1v_2) .$$

$$\alpha'' = \alpha' - \gamma$$

The term γ is not an element of $B_1(K, G)$ except in the trivial case that

$$g = g_{26}g_3g_8g_{12}g_{17}g_{21}$$

is zero. It represents a cycle which does not break the Klein bottle into an inside and an outside. The term α'' will be an element of $B_1(K, G)$ as long as the equations for h 's in terms of g 's have a solution. This they will have whenever $2h_{17}$ is a multiple $2g'$, for some $g' \in G$. This will not be true in general. However, it is always true that $2\alpha'' \in B_1(K, G)$. To summarize, $Z_1(K, G)$ consists of elements of $B_1(K, G)$ plus the subgroup generated by $g(v_2v_6+v_6v_1+v_1v_2)$ plus the subgroup generated by an α'' which satisfies $2\alpha'' \in B_1(K, G)$. This is equivalent to saying $Z_1(K, G)/B_1(K, G) = H_1(K, G) \approx G \oplus J_2$. This shows $H_1(K, G)$ is characterized by more than just the Betti number. In addition it is characterized by one torsion number $\theta = 1$ meaning there is one finitely generated subgroup of order 2. The summary for the Klein bottle is given below.

p	C_p	Z_p	B_p	H_p
0	$\otimes 9G$	$\otimes 9G$	$\otimes 8G$	G
1	$\otimes 27G$	$\otimes 19G \otimes J_2$	$\otimes 18G$	$G \otimes J_2$
2	$\otimes 18G$	0	0	0

C. Relative Homology Theory

1. Simplicial maps

So far we have concentrated on the homology groups of a single space. We turn now to the situation where we have two topological spaces, and a mapping connecting them. Consistent with our previous discussions we suppose both spaces have been triangulated, and we are given two complexes, K_1 and K_2 . It turns out that we can always replace any continuous map of K_1 and K_2 by a homotopic map ϕ which is (continuous and) simplicial: Let $\langle v_0 \dots v_p \rangle$ be any simplex of K_1 ; ϕ is a map of the vertices of K_1 into the vertices of K_2 . Specifically, $\langle \phi(v_0) \dots \phi(v_p) \rangle$ is a simplex of K_2 . Now not all the vertices $\phi(v_i)$ may be distinct, so $\langle \phi(v_0) \dots \phi(v_p) \rangle$ may not be a p-simplex.

Given a simplicial map $\phi: K_1 \rightarrow K_2$, there is an induced map of the chain groups

$$\phi_p: C_p(K_1, G) \rightarrow C_p(K_2, G) .$$

For each simplex $v_0 \dots v_p$ one defines

$$\phi_p(gv_0 \dots v_p) = \begin{cases} 0 & \text{if } \phi(v_i) = \phi(v_j), i \neq j \\ g\phi(v_0) \dots \phi(v_p) & \text{if } \phi(v_i) \neq \phi(v_j), i \neq j . \end{cases}$$

The rule for when ϕ_p gives zero can also be imposed by requiring $gA_1 \dots A_m = 0$ whenever two vertices $A_i = A_j$, $i \neq j$. With the above definition ϕ_p can be extended linearly to all p -chains.

2. Commutativity with ∂

It is not difficult to check that ϕ_p is a homomorphism. It has also another characteristic property, ϕ_p commutes with the boundary homomorphism

$$\partial \phi_p = \phi_{p-1} \partial.$$

This can be checked by explicit calculation on each generator:

$$\partial \phi_p(gv_0 \dots v_p) = \sum_m (-)^m g \phi(v_0) \dots \phi(v_m) \dots \phi(v_p)$$

$$\phi_{p-1} \partial(gv_0 \dots v_p) = \sum_m (-)^m g \phi(v_0) \dots \phi(v_m) \dots \phi(v_p).$$

3. ϕ_p induces homomorphisms between homology

We can also show ϕ_p induces a homomorphism between the homology groups. Suppose $z \in Z_p(K_1, G)$, thus $\partial z = 0$, and $\phi_{p-1} \partial z = 0$, and so $\partial \phi_p(z) = 0$. This shows $\phi_p(z) \in Z_p(K_2, G)$. Suppose $b \in B_p(K_1, G)$, then there exists an $\alpha \in C_{p+1}(K_1, G)$ with $\partial \alpha = b$. Now $\phi_p(b) = \phi_p \partial \alpha = \partial \phi_{p+1}(\alpha)$, showing $\phi_p(b) \in B_p(K_2, G)$. An element in $H_p(K_1, G) = Z_p(K_1, G)/B_p(K_1, G)$ is an equivalence class of elements $[z]$ such that $z \equiv z' \pmod{B}$ iff $z - z' \in B_p(K_1, G)$. From our previous arguments we see that $\phi_p(z), \phi_p(z') \in Z_p(K_2, G)$ and that $\phi_p(z) - \phi_p(z') \in B_p(K_2, G)$. This shows $\phi_p(z) \equiv \phi_p(z') \pmod{B}$: to each class $[z]$ there corresponds a unique class $[\phi_p(z)] \in H_p(K_2, G)$. This correspondence is indeed a homomorphism, and is denoted by ϕ_{p*} :

$$\phi_{p*} : H_p(K_1, G) \rightarrow H_p(K_2, G).$$

By more advanced reasoning it can be shown that ϕ_{p*} depends only on the homotopy class of ϕ . Thus homology theory provides for spaces and maps, groups and homomorphisms. The groups are topological invariants and the maps are homotopic invariants.

4. Pairs of spaces: chain, boundary, cycle and homology groups

Another aspect of homology theory we haven't touched is the possibility of constructing homology groups for a pair of spaces (X, A) where A is a subspace of X . As usual we will be content for now with a complex K and a subcomplex L . Define two simplicial maps (here $\emptyset$ is the empty set)

$$i: (L, \emptyset) \rightarrow (K, \emptyset)$$

$$j: (K, \emptyset) \rightarrow (K, L)$$

as the natural injections. In other words, i acts as the identity on L ; its image is considered an element of K rather than an element of the subset L of K . Similarly j acts as the identity on K ; it certainly takes the empty set into L , and takes K into K as indicated by $j: (K, \emptyset) \rightarrow (K, L)$.

Now let us consider the chain groups for the pair (K, L) . Denote by $C_p(L)$ and $C_p(K)$ the chain groups when the group G is the set of integers. All the generators of $C_p(L)$ are contained in $C_p(K)$; the map i_p induced by i gives a natural isomorphism between $C_p(L)$ and a subgroup of $C_p(K)$. We will use the same notation $C_p(L)$ for this subgroup of $C_p(K)$. We define the group of relative p-chains of K mod L as $C_p(K, L) = C_p(K)/C_p(L)$. These are equivalence classes $[c]$ defined as $c \equiv c' \pmod{C_p(L)}$ iff $c - c' \in C_p(L)$. A boundary homomorphism can be defined on $C_p(K, L)$ by

$$\partial[c] = [\partial c],$$

and so one can define the group $Z_p(K, L)$ of relative p-cycles of $K \bmod L$ as the set of elements $[c]$ such that $\partial[c] = 0$, and the group $B_p(K, L)$ of relative p-boundaries of $K \bmod L$ as $\partial C_{p+1}(K, L)$. One shows $\partial^2 = 0$, so $E_p(K, L) \subset Z_p(K, L)$. Both subgroups are commutative, so the factor set $H_p(K, L) = Z_p(K, L) / B_p(K, L)$ is a commutative group, the relative p-homology group of $K \bmod L$.

5. Exact sequences

We now show how to construct an exact homology sequence. Based on paragraph 3, the injections i and j induce homomorphisms on the homology groups. The injection $i: (L, \emptyset) \rightarrow (K, \emptyset)$ induces the map $i_p: C_p(L) \rightarrow C_p(K)$, which in turn induces the homomorphism

$$i_*: H_p(L) \rightarrow H_p(K).$$

The injection $j: (K, \emptyset) \rightarrow (K, L)$ induces a homomorphism $j_p: C_p(K) \rightarrow C_p(K, L)$, namely for $c \in C_p(K)$, $j_p(c) = [c]$ the equivalence class of $c \bmod C_p(L)$. One then has the homomorphism

$$j_*: (H_p(K) + H_p(K, L))$$

induced by j .

Using paragraph 4, one can show that ∂ induces a homomorphism on the relative homology groups. Let $[z] \in H_p(K, L)$ and let $[z] \in Z_p(K, L)$ be a relative p-cycle which is in the equivalence class $[z] \bmod B_p(K, L)$, and let $z \in C_p(K)$ be a representative chain of the equivalence class $[z]$. Since $[\partial z] = \partial[z] = 0$, ∂z is a $p-1$ chain of L , and must also be a cycle: $\partial z \in Z_{p-1}(L)$. Consideration of the elements of $Z_{p-1}(L)$ equivalent to $\partial z \bmod B_{p-1}(L)$ will show that ∂ determines a unique element $[\partial z]'$ of $H_{p-1}(L)$. Define

$$\partial_* [[z]] = [\partial z]'.$$

One can show that this map is a homomorphism $\partial_*: H_p(K, L) \rightarrow H_{p-1}(L)$.

Putting together these results, one constructs a sequence of groups and homomorphisms:

$$\begin{array}{ccccccc} & 1_* & & \partial_* & & \partial_* & & 1_* \\ \dots & \rightarrow & H_p(K) & \xrightarrow{\partial_*} & H_p(K, L) & \xrightarrow{\partial_*} & H_{p-1}(L) & \xrightarrow{\partial_*} & H_{p-1}(K) & \rightarrow \dots \end{array}$$

One can show that this sequence, called the homology sequence is exact: that is, the image of each homomorphism in the sequence is equal to the kernel of the next homomorphism in the sequence.

Another result one can obtain concerns the effect of removing from L and K a subcomplex M . One shows that the injection of $K \setminus M$ into K induces an isomorphism of $H_p(K \setminus M, L \setminus M)$ onto $H_p(K, L)$ for each dimension p . This property is called the excision property.

D. Singular Homology Theory

In the previous paragraphs we have shown in detail how to compute homology groups for complexes; we have introduced the notion of relative homology groups, and indicated how homomorphisms between homology groups are related to simplicial maps. We have also stated a few of the more important properties of homology theory. We have avoided discussing homology theory for general spaces and continuous maps because it is not known whether every space can be triangulated, and whether continuous map for general spaces can be replaced by simplicial maps. One can nevertheless obtain a satisfactory theory. We will indicate briefly how this is done.

1. Axioms

The first observation is that the homology theory can be axiomatized. It is possible to pick out properties which are easily demonstrated for triangulable spaces and simplicial maps, which indeed hold more generally. We summarize here the axioms for any homology theory. The elements of the axioms will be pairs of topological spaces (X, A) with $A \subset X$, and maps

$$f : (X, A) \rightarrow (Y, B)$$

which means $f: X \rightarrow Y$ is continuous, and the restriction $f|_A$ of f to A defines a continuous map $f|_A: A \rightarrow B$. To each pair of spaces (X, A) and each integer p is associated an abelian group $H_p(X, A)$ called the relative homology group of X mod A of dimension p .

To each map $f: (X, A) \rightarrow (Y, B)$ and integer p is assigned a homomorphism $f_*: H_p(X, A) \rightarrow H_p(Y, B)$ between homology groups. There is defined for each p and each pair of spaces (X, A) , a function

$$\partial : H_p(X, A) \rightarrow H_{p-1}(A) .$$

Finally, one uses the notation

$$i : A \rightarrow X$$

for the inclusion map $i(x) = x$, $x \in A$; and

$$j : X \rightarrow (X, A)$$

for the inclusion map of $(X, \emptyset)$ into (X, A) , $j(x) = x$, $x \in X$. The axioms are

1. If $f = \text{identity}$ then $f_* = \text{identity}$
2. $(gf)_* = f_* g_*$
3. $\partial f_* = (f|_A)_* \partial$
4. (Exactness Axiom) The homology sequence is exact:

$$\cdots \xrightarrow{\partial} H_p(A) \xrightarrow{i_*} H_p(X) \xrightarrow{j_*} H_p(X, A) \xrightarrow{\partial} H_{p-1}(A) \xrightarrow{i_*} \cdots$$

5. (Homotopy Axiom) If f_0 and f_1 are homotopic maps of $(X, A) \rightarrow (Y, B)$, then for each p , f_{0*} and f_{1*} of $H_p(X, A)$ into $H_p(Y, B)$ coincide
6. (Excision Axiom) If U is an open subset of X whose closure $\bar{U}$ is contained in the interior of A , then the inclusion map $(X \setminus U, A \setminus U) \rightarrow (X, A)$ induces isomorphisms $H_p(X \setminus U, A \setminus U) \cong H_p(X, A)$ for each integer p .
7. (Dimension Axiom) If P is a space with a single point, then $H_p(P) = 0$ for all $p \neq 0$.

2. Singular p -simplexes and chain groups

Given these axioms, one is not expected to know whether or not there exist a set of groups and homomorphisms (other than all $H_p = 0$) satisfying them. There may be (and in fact is) several theories satisfying the same set of axioms. The one called Singular Homology Theory is found as follows. Consider the set of maps from a unit p -simplex Δ_p into X . The unit p -simplex is the set of points $(x_0, \dots, x_p)$ in Euclidean space $\mathbb{R}^{p+1}$ with $\sum_{i=0}^p x_i = 1$, $0 \leq x_i \leq 1$, $i = 0, \dots, p$. Each map T is called a singular p -simplex. One uses these maps to generate an abelian group as follows.

Consider the formal set of objects gT where T is any singular p -simplex and g is any integer. Consider the set of objects generated by the formal sums $\sum_T g_T T$. The addition of two such objects is defined by taking for each T , $g_T T + g'_T T = (g_T + g'_T)T$. The set of objects $\sum_T g_T T$ for all sets $\{g_T\}$ of integers defines a group, called the group of singular p -chains of X over the integers, $C_p(X)$. It is similar to the p -chain group we discussed before, but it need not be finitely generated. It may not be calculable.

2. Boundary, cycle and homology groups

One defines on $C_p(X)$ a boundary homomorphism by considering the "faces" T^i of T . Let $e_p^i: \Delta_{p-1} \rightarrow \Delta_p$ be the simplicial map (i.e. one which transforms vertices into vertices) defined by the requirement that it take Δ_{p-1} onto the $p-1$ face of Δ_p which does not contain the i^{th} vertex. Specifically, let the vertices of Δ_p be $d_0, \dots, d_p$; d_i is the point $(0, \dots, 1, 0, \dots, 0)$ having $x_i = 1$ in $\mathbb{R}^{p+1}$. The function e_p^i is defined by

$$e_p^i(d_j) = d_j \quad j < i$$

$$= d_{j+1} \quad i \leq j$$

The i^{th} face of a singular p -simplex is defined as $T^{(i)} = T e_p^i: \Delta_{p-1} \rightarrow X$. The boundary homomorphism is defined by

$$\partial_p gT = \sum_{i=0}^p (-1)^i g T^{(i)}.$$

This gives a homomorphism $\partial_p: C_p(X) \rightarrow C_{p-1}(X)$, satisfying $\partial_{p-1} \partial_p = 0$.

4. Relative chain, boundary, cycle and homology groups

Given the chain group and boundary homomorphism one can proceed to the relative homology groups. For a pair of subspaces (X, A) , one has both the chain group $C_p(X)$ and $C_p(A)$ for each p . The generators of $C_p(A)$ are generators of $C_p(X)$; the factor set $C_p(X)/C_p(A) = C_p(X, A)$ is a group of relative singular p -chains of X mod A . The boundary operator ∂_p induces a boundary operator $\bar{\partial}_p: C_p(X, A) \rightarrow C_{p-1}(X, A)$, which also satisfies $\bar{\partial}_{p-1}\bar{\partial}_p = 0$. Then one can define the group $Z_p(X, A)$ of relative p -cycles of X mod A by elements $z \in C_p(X, A)$ such that $\bar{\partial}z = 0$. Also the group of relative p -boundaries $B_p(X, A)$ is defined as $\bar{\partial}_{p+1}C_{p+1}(X, A)$. The relative p -homology group $H_p(X, A)$ of X mod A is defined as $Z_p(X, A)/B_p(X, A)$ since $B_p \subset Z_p$. All of these definitions are just a repeat of what we did before for triangulable spaces.

5. Relation to the axioms

Formally one now has for each pair of spaces (X, A) and each p , an abelian group $H_p(X, A)$. One can also construct a homomorphism

$$\partial_{\#}: H_p(X, A) \rightarrow H_{p-1}(A)$$

as was done for triangulable spaces. Before checking the axioms one has to show that for each continuous map f there is an induced map $f_{\#}$ on the homology groups.

To find the induced map, one constructs first f_p , defined on $C_p(X)$ by

$$f_p(T) = fT.$$

This gives a map $A_p \rightarrow B_p$ when $f: X \rightarrow Y$. Thus $f_p: C_p(X) \rightarrow C_p(Y)$, and one can pass to relative singular p -chain groups to obtain an $\bar{f}_p: C_p(X, A) \rightarrow C_p(Y, B)$, and finally by considering $\bar{f}_p$ on $Z_p(X, A)$ and equivalence classes mod $B_p(X, A)$ one obtains a map $f_{\#}: H_p(X, A) \rightarrow H_p(Y, B)$.

It is straightforward, but lengthy to show that the groups $H_p(X, A)$ and maps f_* , ∂_* satisfy the axioms. One demonstrates there is a homology theory for arbitrary spaces and maps. One has not demonstrated that such groups are always calculable. For spaces (X, A) which are triangulable, one can prove that the homology groups of the singular theory are equal (up to isomorphism) to the groups as calculated using ordinary simplexes and complexes. In other words we can calculate using the simple rules.

One can also use the axioms directly to obtain theorems concerning homology groups and homomorphisms. This is often more efficient than direct computation. One can show in this way for example that for an n -sphere S^n , $H_n(S^n) = \mathbb{Z}$, the group of integers.

E. Connection between Homology and Homotopy

The close connection between homology and homotopy is evident if one compares the homology axioms to the seven properties satisfied by homotopy groups and their homomorphisms. Homotopy groups don't satisfy the excision axiom, but satisfy instead a property involving fibre bundles.

The useful result for our purposes is the existence of a homomorphism

$$\phi: \pi_n(X, A, x_0) \rightarrow H_n(X, A)$$

between relative homotopy groups and relative homology groups. This is true for each n . When A is an arcwise connected subspace of X and both X and A are simply connected, then for n defined as the largest integer such that $\pi_1(X, A, x_0) = 0$ for $2 \leq n$,

$$\phi: \pi_n(X, A, x_0) \rightarrow H_n(X, A)$$

is an isomorphism. This theorem is due to Hurewicz, and the proof can be found in [Hu]. The result can be stated as "the first non-zero homology group and the first non-zero homotopy group have the

same dimension and are isomorphic under ϕ ." [see Steenrod, pp. 78-80.]

Applied to the n -sphere S^n it shows that $\pi_n(S^n) \approx H_n(S^n) \approx \mathbb{Z}$ the set of integers. For an n -cell E^n and its boundary S^{n-1} , the theorem also shows $\pi_n(E^n, S^{n-1}, x_0) \approx H_n(E^n, S^{n-1}) \approx \mathbb{Z}$. These are the only results that Steenrod uses from homology theory in his Part II on homotopy theory of bundles.

Bibliography

S. Eilenberg and N. Steenrod, Foundations of Algebraic Topology
(Princeton University Press, Princeton, N.J., 1952).

J. Hocking and G. Young, op. cit.

S.-T. Hu, op. cit.

N. Steenrod, op. cit.

T. T. Wu and C. N. Yang, op. cit.