Part III
Bundles
TABLE OF CONTENTS

Lecture 55. Lie Groups and 1-Parameter Subgroups 1

1. Introduction 1
   Problem 90 5
2. 1-Parameter Subgroups 5
   Problem 91 7

Lecture 56. The Lie Algebra of a Lie Group 8

1. Lie Algebra 8
   Problem 92 9
2. The Exponential Map 10
3. Examples of Lie Algebras 11
   Problem 93 13

Lecture 57. Lie Algebras and Subgroups 15

1. Left Invariant Fields Generate Right Translations 15
2. Commutators of Matrices 16
   Problem 94 17
3. Right Invariant Vector Fields 17
4. Subgroups and Subalgebras 18

Lecture 58. Vector Bundles 21

1. Vector Bundles 21
2. Product or Trivial Bundles 25
   Problem 95 26
   Problem 96 28

Lecture 59. Connections in a Vector Bundle 29

1. Connection in a Vector Bundle 29
2. Complex Vector Spaces 32
3. Complex Vector Bundles 34

Lecture 60. The Electromagnetic Connection 37

1. Minimal Coupling 37
2. The Quantum Electron in the Electromagnetic Field 39
   Problem 97-98 42
3. Remark on Weyl's Principal of Gauge Invariance 42a

Lecture 61. Fiber Bundles and Principal Bundles 43

1. Fiber Bundles 43
2. Principal Bundles 44
3. Coset Spaces 48
Problem 99 50
4. Remarks on Coset Spaces 50a
Problem 100 51

Lecture 62. Chern's Proof of Gauss-Bonnet-Pointcaré Theorem 52

1. A Connection in the Frame Bundle of a Surface 52

Lecture 63. Lie Algebra Valued Forms 58

Problem 101 59

Lecture 64. Connections in a Principal Bundle 64

Problem 102-103 69

Lecture 65. The Yang-Mills Field 70

1. The Electromagnetic Connection Revisited; Weyl's Interpretation 70
2. The Heisenberg Nucleon 71
3. The Yang-Mills Nucleon 72
4. Transformation fo A and F under Gauge Transformations 74

Lecture 66. The Groups SO(3) and SU(2) 76

1. SO(3), The Rotation Group of $\mathfrak{R}^3$ 76
2. SU(2) 78
3. Digression on the Topology of SU(2) 80

Lecture 67. Hamilton, Clifford, and Dirac 83

1. Rotations of $\mathfrak{R}^3$ and Pauli Matrices 83
Problem 104 84
2. Clifford Algebras 85
3. The Dirac Program; The Square Root of the D'Alembertian 87

Lecture 68. The Dirac Algebra 89

1. The Lorentz Group 89
2. The Dirac Algebra 93

Lecture 69. The Dirac Operator ∂ in Minkowski Space 95

1. Transformation Properties of the Dirac Matrices 95
2. The Dirac Operator 98

Lecture 70. The Dirac Operator in Curved Space 100

1. The Spinor Bundle 100
2. Connection for an Associated Bundle 101
3. The Spin Conneciton for the Curved M^4 104

Lecture 71. p-Form Sections of a Vector Bundle 107

1. Associated Bundles and Connections 107
2. p-Form Sections of a Vector Bundle 110
Problem 105 114

Lecture 72. Yang-Mills Fields, I 115

1. Invariant Product in *G* 115
Problem 106 115
2. The Yang-Mills Action 118

Lecture 73. Yang-Mills fields, II 121

1. The Yang-Mills Equations 121
2. Yang-Mills Analogy with Electromagnetism 125
Problem 107 126

Lecture 74. Yang-Mills wth Finite Action 127

1. Finite Action 127
2. Chern's Proof Revisited 131
Problem 108 134

Lecture 75. DeRham for Compact Lie Groups 135

1. Bi-invariant p-Forms 135
2. Bi-invariant Riemannian Metrics on Compact Groups 137
Problem 109 138
3. Harmonic Forms in the Bi-invariant Metric 139

Lecture 76. Winding Numbers and Chern Forms 143

1. The Yang-Mills "Winding Number" 143
2. Winding Number in Terms of Field Strength 145
3. The Chern Forms for a U(n) Bundle 147
Problem 110-111 148

Lecture 77. Instanton Number 149

Lecture 78. Homotopy and the Fundamental Group, $\Pi_1(M^n)$ 150

1. Homotopy 150
2. Covering Homtopies 151
3. Some Topology of SU(n) 152
Problem 112 154

4. The Fundamental Group $\Pi_1(M^n)$ 155

Lecture 79. The Higher Homotopy Groups, $\Pi_k(M^n)$ 158

1. $\Pi_k(M^n)$ 158
2. Exact Sequences 161
3. The Exact Homotopy Squence for a Bundle 162

Lecture 80. Some Examples of $\Pi_k(M^n)$, The Hopf Map 165

1. Covering Spheres 165
2. The Concept of a Covering Space 165
3. Back to SU(n) 168
   Problem 113-114 169
4. The Hopf Map and Fibering 169

Lecture 81. Chern Forms as Obstructions 171

1. Covariant Differential of Forms of Type Ad 171
2. The Chern forms C_r are Closed 172
3. $C_2(E)$ as an obstruction "cocycle" 173

Part III Bundles

Lecture 55. Lie Groups and 1-Parameter Subgroups

(1) Introduction. A Lie group is a group G whose pointset is topologically a differentiable manifold and such that the group actions

$$G \times G \longrightarrow G \qquad (g,h) \longmapsto gh$$

$$G \longrightarrow G \qquad g \longmapsto g^{-1}$$

are differentiable

Examples (a) $G = \mathbb{R}$, the additive group of real numbers

(b) $G = \mathbb{R}^+$, the multiplicative group of positive real numbers

(c) $G = Gl(n,\mathbb{R})$, the general linear group of all real $(n \times n)$ matrices g with $\det g \neq 0$.

(d) $G = O(n)$, the orthogonal group.

(e) $G = A(1)$, the "affine group" of the line, i.e those real 2×2 matrices

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \quad \text{with} \quad x > 0.$$

The manifold for $A(1)$ can be considered as the "right half-plane" (x,y), $x > 0$.

A matrix group is a subgroup of $Gl(n,\mathbb{R})$ (or $Gl(n,\mathbb{C})$ if it is complex) that is a submanifold of $Gl(n,\mathbb{R})$, ie of $\mathbb{R}^{n^2}$. While there are Lie groups that cannot be realized as matrix groups they are rare. For our calculations we shall pretend that our group is indeed a matrix group.

As a manifold, a Lie group is very special, for the following reasons. First, a Lie group G always has two families of diffeomorphisms, the left and right translations

$$L_g : G \longrightarrow G \qquad L_g(h) = gh$$

$$R_g : G \longrightarrow G \qquad R_g(h) = hg$$

Then given a tangent vector to G at e, X_e we may left (right) translate X_e to each point of G

$$X_g \equiv (L_g)_* X_e$$

$$\text{or} \quad X_g \equiv (R_g)_* X_e$$

yielding two non vanishing vector fields on all of G. In fact if we take a basis $X_1, \ldots, X_n$ for G_e, ie the tangent space to G at e, then e.g. we can left translate this basis to give n linearly independent vector fields

$$(L_g)_* X_1, \ \ldots, \ (L_g)_* X_n$$

on all of G. In particular, <u>every Lie group is an orientable manifold</u>. Consider, e.g. a closed surface M^2 of genus g sitting in $\mathbb{R}^3$. Of these surfaces, only genus $g = 1$ can have a nonvanishing tangent vector field by Poincare's theorem (see Lecture 28), and in fact the torus T^2 does have a pair of non vanishing vector fields $\partial/\partial\theta$ and $\partial/\partial\varphi$. T^2 is indeed a Lie group, the commutative group of "translations";

$$T^2 \cong S^1 \times S^1 \ , \quad \theta \in S^1, \quad \varphi \in S^1$$

with "multiplication"

$$(\theta_1, \varphi_1) \cdot (\theta_2, \varphi_2) = (\theta_1 + \theta_2, \varphi_1 + \varphi_2)$$

Topologically, the only compact Lie group of dimension 2 is the torus.

<u>Definition</u>: A vector field X on a Lie group G is said to $\left\{\begin{matrix}\text{left}\\ \text{right}\end{matrix}\right\}$ <u>invariant</u> provided it is invariant under $\left\{\begin{matrix}\text{left}\\ \text{right}\end{matrix}\right\}$ translations

$$X_{gh} = (L_g)_* X_h$$

or

$$X_{hg} = (R_g)_* X_h$$

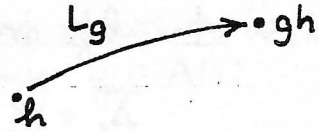
You should convince yourself that, e.g, <u>if we translate X_e at e over the group by</u>

$$X_g \equiv (L_g)_* X_e$$

<u>then the resulting X is indeed left invariant</u>!

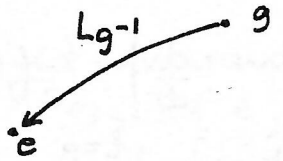
Likewise, e.g, an exterior p-form α^p on G is left invariant provided

$$L_g^* \alpha^p_{gh} = \alpha_h$$



To get a left invariant p-form on G one translates a p-form at e over the entire group

$$\alpha_g \equiv (L_{g^{-1}})^* \alpha_e$$



Example : $G = A(1)$, the affine group of the line

$$A(1) = \left\{ \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}, x > 0 \right\}$$

We may consider A(1) as a submanifold of $\mathbb{R}^4$ or as the right half plane. Since

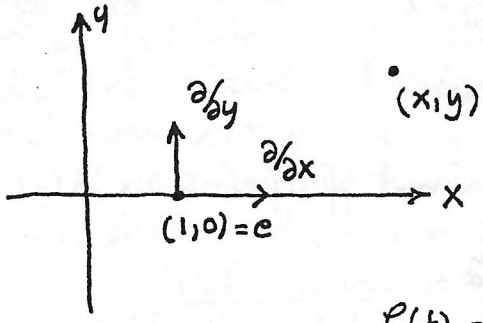
$$\begin{pmatrix} x_1 & y_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 & y_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x_1 x_2 & x_1 y_2 + y_1 \\ 0 & 1 \end{pmatrix}$$

we see that the right half plane is endowed with a rather unusual "multiplication" $(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2, x_1 y_2 + y_1)$. We identify

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \longleftrightarrow (x, y)$$

for the points of the Lie group. For tangent vectors we identify

$$\begin{pmatrix} \dot{x} & \dot{y} \\ 0 & 0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \dot{x} \frac{\partial}{\partial x} + \dot{y} \frac{\partial}{\partial y}$$



Let us left translate the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ at e to the point (x, y). For $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \partial/\partial x$ we consider the curve thru e

$$C(t) = \begin{pmatrix} 1+t & 0 \\ 0 & 1 \end{pmatrix}$$

whose tangent at e is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Then

$$\left(L_{\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}}\right)_* \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{d}{dt}\Big|_{t=0} \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1+t & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$$

and so the left translate of $\partial/\partial x$ to (x,y) is

$$X_1 = x\frac{\partial}{\partial x}.$$

Likewise the left translate of $\frac{\partial}{\partial y}$ to (x,y) is

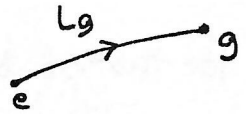
$$\left.\frac{d}{dt}\right|_{t=0} \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}$$

$$X_2 = x\frac{\partial}{\partial y}.$$

Thus a basis for the left invariant vector fields on $A(1)$ is given by the pair

$$X_1 = x\frac{\partial}{\partial x}, \quad X_2 = x\frac{\partial}{\partial y} \tag{55.1}$$

Next note that in any Lie group, if $X_1, \dots, X_n$ is a basis for left invariant vector fields on G and if $\sigma^1, \dots, \sigma^n$ is the dual basis to $X_1, \dots, X_n$, then the $\sigma^1, \dots, \sigma^n$ form a basis for left invariant 1-forms on G. To see that the σ's are indeed left invariant



$$\left[(L_g)^*\sigma^r_{(g)}\right]\left(X_{(e)s}\right)$$

$$= \sigma^r_{(g)}\left[(L_g)_* X_{(e)s}\right] = \sigma^r_{(g)}\left(X_{(g)s}\right) = \delta^r{}_s = \sigma^r_{(e)}\left(X_{(e)s}\right)$$

and so $(L_g)^*\sigma^r_{(g)} = \sigma^r_{(e)}$, as desired!

Thus the dual basis of left invariant pfaffians on $A(1)$ is given by

$$\sigma^1 = \frac{dx}{x} \qquad \sigma^2 = \frac{dy}{x} \tag{55.2}$$

Now, of course, if α^p and β^q are invariant under left translations then so is $d\alpha^p$ and also $\alpha^p \wedge \beta^q$. Thus in $A(1)$

$$\sigma^1 \wedge \sigma^2 = \frac{dx \wedge dy}{x^2} \tag{55.3}$$

is a left invariant area form or "left Haar measure"; for any compact region $\mathcal{U} \subset A(1)$, and for any $g \in A(1)$

$$\iint_{L_g(\mathcal{U})} \frac{dx \wedge dy}{x^2} = \iint_{\mathcal{U}} \frac{dx \wedge dy}{x^2}$$

Problem 90 For the group $A(1)$, find the right invariant vector fields coinciding with $\partial/\partial x$ and $\partial/\partial y$ at e, find the dual right invariant 1-forms, and find the right Haar measure.

(2) 1-Parameter Subgroups. A homomorphism of groups is a function

$$f : G \longrightarrow H$$

such that

$$f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$$

($g_1 \cdot g_2$: product in G; $f(g_1) \cdot f(g_2)$: product in H)

For example, the usual exponential function $f(t) = e^t$ defines a homomorphism

$$\exp : \mathbb{R} \longrightarrow \mathbb{R}^+$$

of the additive group of the reals to the multiplicative group of the positive reals. Note that exp is in fact also a differentiable map. Actually, of course exp is 1:1 (i.e the homomorphism is "injective") and also "onto" ("surjective"). In short exp is an isomorphism of Lie groups. exp is a diffeomorphism with inverse $\ln : \mathbb{R}^+ \longrightarrow \mathbb{R}$.

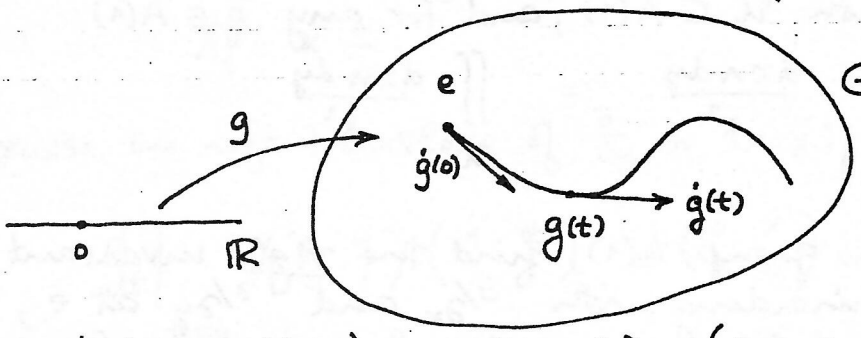
Definition. A 1-parameter subgroup of a Lie group is a differentable homomorphism (in particular, a path)

$$g : \mathbb{R} \longrightarrow G \qquad t \longmapsto g(t) \in G$$

Thus

$$g(s+t) = g(t+s) = g(t) \cdot g(s) = g(s) \cdot g(t)$$

Consider now a 1-parameter subgroup g of a matrix group G



As matrices $g(t+s) = g(t) \cdot g(s)$ (i.e $g_{ij}(t+s) = \sum_k g_{ik}(t)\, g_{kj}(s)$)

Differentiate both sides with respect to s and put $s = 0$

$$\dot{g}(t) = g(t) \cdot \dot{g}(0) \tag{55.4}$$

Since $\dot{g}(0)$ is a constant matrix, the solution to (55.4) is

$$g(t) = g(0)\, e^{t\dot{g}(0)} \tag{55.5}$$

where

$$e^S \equiv \exp S \equiv I + S + \tfrac{1}{2!} S^2 + \tfrac{1}{3!} S^3 + \cdots \tag{55.6}$$

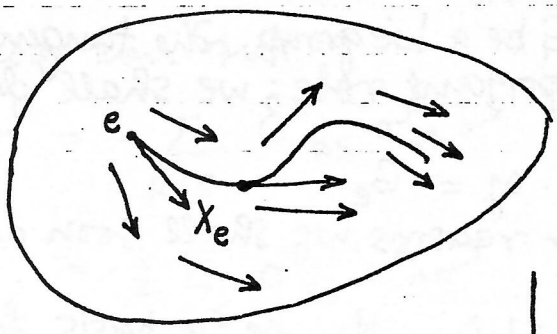
converges for all matrices S. Since $g(0) = e$ for any homomorphism $g : \mathbb{R} \longrightarrow G$, we conclude that

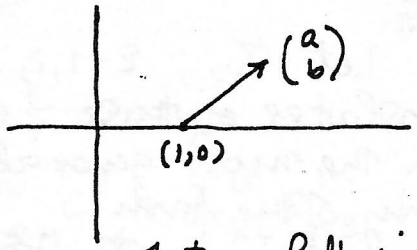
$$\boxed{g(t) = e^{t\dot{g}(0)}} \tag{55.7}$$

is the most general form for a 1-parameter subgroup of a matrix group G. Equation (55.4) tells us how to proceed even if G is not a matrix group, for it really says

$$\dot{g}(t) = (L_{g(t)})_* \dot{g}(0) \tag{55.4'}$$

i.e, the tangent vector X to the 1-parameter subgroup is left translated along the subgroup. In other words, given a tangent vector X_e at $e \in G$, the 1-parameter subgroup of G whose tangent at e is X_e is an integral curve of the vector field X on G resulting from left translating X_e over the entire group





For example, in $A(1)$, to find the 1-parameter subgroup having velocity vector $\begin{pmatrix} a \\ b \end{pmatrix}$ at the identity, we express this vector in terms of the left invariant basis (55.1) at $x=1, y=0$;

$$\begin{pmatrix} a \\ b \end{pmatrix} = a X_1 + b X_2 .$$

Then we need to solve

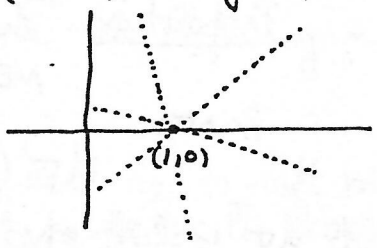
$$\frac{d\vec{x}}{dt} = a X_1 + b X_2 = ax \frac{\partial}{\partial x} + bx \frac{\partial}{\partial y}$$

i.e

$$\begin{cases} \frac{dx}{dt} = ax \\ \frac{dy}{dt} = bx \end{cases} \qquad \text{subject to } x(0)=1,\ y(0)=0 . \tag{55.8}$$

The solutions are clearly straight lines $\frac{dy}{dx} = \left(\frac{b}{a}\right)$, but to see the parameterization we must solve (55.8) to get

$$\begin{cases} x(t) = e^{at} \\ y(t) = \frac{b}{a} e^{at} - \frac{b}{a} . \end{cases}$$



Problem 91 In $Gl(2,\mathbb{R})$ consider $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ as a tangent vector at the identity. Use $J^2 = -I$, $J^3 = -J$, $J^4 = I$ to show

$$e^{\theta J} = \cos\theta\ I + \sin\theta\ J$$

i.e

$$\exp \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

This 1-parameter subgroup of $Gl(2,\mathbb{R})$ is thus the entire subgroup of rotations of the plane, $SO(2)$.

Lecture 56. The Lie Algebra of a Lie Group

(1) Lie Algebra. Let G be a Lie group. The tangent vector space G_e at e plays an important role; we shall denote it by a lower case German $\mathfrak{g}$

$$\mathfrak{g} = G_e$$

and call it (for reasons we shall soon discuss) the Lie algebra of G.

Let X_R, $R=1,2,\dots,N$ be a basis for $\mathfrak{g}$. The left translates of these fields over G will again be called X_R. Then the most general left invariant vector field to G is clearly of the form

$$X = \sum a^R_R X_R = \sum X_R a^R$$

where the a^R are constants.

Let σ^R, $R=1,\dots,N$ be the dual basis of left invariant 1-forms on G; of course they are determined by their values on $\mathfrak{g}$; the most general 1-form on G that is left invariant is of the form $\sum a_R \sigma^R$, and in fact the most general left invariant p-form is of the form

$$\alpha^p = \sum_{R_1<\dots<R_p} a_{R_1\dots R_p}\, \sigma^{R_1}\wedge\dots\wedge\sigma^{R_p}.$$

It is again determined by its value on p-tuples from $\mathfrak{g}$.

Proposition. The Lie bracket $[X_R, X_S]$ of two left invariant vector fields is again left invariant.

Proof.

$$d\sigma^T(X_R, X_S) = -\sigma^T([X_R, X_S]).$$

But $d\sigma^T$ is left invariant, hence the left hand side is constant. Since $\sigma^T([X_R, X_S])$ is constant, $[X_R, X_S]$ must be left invariant. □.

We may then write

$$[X_R, X_S] = X_T C^T_{RS} \qquad C^T_{RS} = -C^T_{SR} \qquad (56.1)$$

for some constants C^T_{RS}; these are the "structure constants" for G, using basis $X_1,\dots,X_N$.

Problem 92. Show then that

$$d\sigma^U = -\sum_{R<S} C^U_{RS}\, \sigma^R \wedge \sigma^S \qquad (56.2)$$

$$= -\tfrac{1}{2} \sum_{R,S} C^U_{RS}\, \sigma^R \wedge \sigma^S$$

and show that $d^2\sigma^U = 0$ gives the relation

$$0 = C^U_{R[S}\, C^R_{LM]} \equiv C^U_{RS} C^R_{LM} + C^U_{RM} C^R_{SL} + C^U_{RL} C^R_{MS} \qquad (56.3)$$

(the notation [S LM] indicates that skew symmetrization with respect to these indices should be performed; this is standard physics notation).

Furthermore show that for any triple of vector fields X_R, X_L, X_M on any manifold, considering them as first order differential operators, we have the "Jacobi identity"

$$[[X_R, X_L], X_M] + [[X_M, X_R], X_L] + [[X_L, X_M], X_R] = 0 \qquad (56.4)$$

and that in the case of a Lie group this yields, from (56.1) again the relation (56.3)

We make $\mathfrak{g} = \mathfrak{G}_e$ into a "Lie Algebra" by defining a product

$$\mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$$

as follows. Let $X \in \mathfrak{g}$, $X_2 \in \mathfrak{g}$. Extend X_1 and X_2 to left invariant fields on all of G. Then define their product to be the Lie product

$$[X_1, X_2],$$

ie the Lie bracket! This product satisfies the two relations

$$[X_1, X_2] = -[X_2, X_1]$$

and the Jacobi identity (56.4). One can see easily from (56.4) that this algebra is not associative.

In the future we shall consistently identify the Lie algebra $\mathfrak{g}$ of G with the vector space of left invariant fields on G!

(2) The Exponential Map

Lemma. For any matrix A

$$\det e^A = e^{\operatorname{tr} A} \tag{56.5}$$

Proof. Linear algebra tells us that by a (complex) change of basis A can be reduced to "superdiagonal" form

$$P^{-1} A P = \begin{bmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_N \end{bmatrix}$$

Then

$$\det e^A = \det (P^{-1} e^A P) = \det P^{-1} \Big(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \Big) P$$

$$= \det \Big(\sum_{k=0}^{\infty} \frac{1}{k!} P^{-1} A^k A \Big)$$

But $(P^{-1} A P)^k = (P^{-1} A P)(P^{-1} A P) \cdots (P^{-1} A P) = P^{-1} A^k P$, and so

$$\det e^A = \det \Big(\sum_{k=0}^{\infty} \frac{1}{k!} (P^{-1} A P)^k \Big) = \det e^{P^{-1} A P}$$

$$= \det e^{\begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix}} = \det \begin{bmatrix} e^{\lambda_1} & & * \\ & \ddots & \\ 0 & & e^{\lambda_N} \end{bmatrix}$$

$$= e^{\lambda_1 + \cdots + \lambda_N} = e^{\operatorname{tr} A}$$

□

Prop. The map $\exp : \mathfrak{g} \to G$ sending $A \to e^A$ is a diffeomorphism of some nbrd of $0 \in \mathfrak{g}$ onto some nbhd of $e \in G$.

Proof. For a matrix group, look at the differential of the exponential map applied to a vector $X \in \mathfrak{g}$

$$\exp_*(X) = \frac{d}{dt} (\exp tX)_{t=0} = \frac{d}{dt} \Big(I + tX + \frac{t^2}{2} X^2 + \cdots \Big) \Big|_{t=0}$$

$$= X.$$

Thus $\exp_* : \mathfrak{g} \longrightarrow \mathfrak{g}$ is the identity and exp. is a local diffeomorphism. □.

(3) Examples of Lie Algebras

(i) $G = Gl(n, \mathbb{R})$. Let $M(n\times n)$ be the vector space of all real $n\times n$ matrices; $M(n\times n) \approx \mathbb{R}^{n^2}$. For $A \in M(n\times n)$

$$\det e^A = e^{\operatorname{tr} A} > 0$$

and so

$$\exp : M(n\times n) \longrightarrow Gl(n, \mathbb{R})$$

Since $\dim M(n\times n) = n^2 = \dim Gl(n, \mathbb{R})$ we see that

$$\underset{\uparrow}{gl(n, \mathbb{R})} = M(n\times n)$$

Lie algebra of $Gl(n, \mathbb{R})$

(ii) $G = SO(n)$. First we need two elementary facts about the exponential of a matrix

$$\text{Since } e^A \cdot e^{-A} = \left(I + A + \frac{A^2}{2!} + \cdots\right)\left(I - A + \frac{A^2}{2!} - \cdots\right) = I$$

we conclude

$$(e^A)^{-1} = e^{-A}.$$

Next, from the power series it is evident that for transposes

$$(e^A)^T = e^{A^T}$$

Now suppose that for some matrix A, $e^A \in SO(n)$. Thus

$$e^{-A} = e^{A^T}$$

Since exp is a local diffeomorphism, exp is 1:1 in a nbhd of $0 \in gl(n, \mathbb{R})$. Thus if A is a small matrix

$$-A = A^T$$

i.e. A is skew symmetric. Let $so(n)$ = the vector space of all real skew symmetric matrices. Then for $A \in so(n)$

$$(e^A)^{-1} = e^{-A} = e^{A^T} = (e^A)^T$$

shows us that

$so(n)$ is the Lie algebra of $SO(n)$

(iii) $G = U(n)$, the group of unitary matrices, $u^{-1} = \bar{u}^T$. Then

$$e^A \in U(n) \text{ means } (\overline{e^A})^T = e^{-A}$$

i.e

$$e^{\bar{A}^T} = e^{-A}.$$

Again we conclude, if A is small, that $\bar{A}^T = -A$. If we let

$u(n) =$ all skew hermitian $n \times n$ complex matrices

then

$u(n)$ is the Lie algebra of $U(n)$.

(iv) $G = SU(n)$, unitary matrices u with $\det u = 1$. To the condition $\bar{A}^T = -A$ we need only add a condition to ensure that

$$1 = \det e^A = e^{\operatorname{tr} A},$$

ie

$SU(n) =$ skew hermitian matrices with trace $= 0$

is the Lie algebra of the special unitary group.

(v) $G = Sl(n, \mathbb{R})$ = all $n \times n$ real matrices g with $\det g = 1$. Clearly for Lie algebra we get

$sl(n, \mathbb{R})$ = all $n \times n$ matrices with trace $= 0$.

(4) Is $\exp : \mathfrak{g} \longrightarrow G$ onto ?

Given a $g \in G$, is there always an $A \in \mathfrak{g}$ such that

$$e^A = g \; ?$$

The answer is yes provided the manifold G is compact and connected, otherwise it need not be so! Consider, for example, the group $Sl(2, \mathbb{R})$ the real special linear group in 2 variables. For $g \in Sl(2, \mathbb{R})$

$$g = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \qquad xw - yz = 1$$

i.e, the coordinates x, y, z, w satisfy the above simple quadratic equation. This locus is not compact since, e.g, x can take on

arbitrarily large values. "We" shall show that

$$\begin{bmatrix} -2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \in Sl(2,\mathbb{R})$$

is never of the form e^A for any $A \in sl(2,\mathbb{R})$, ie for any 2×2 matrix A with trace $=0$.

Problem 93. let A be a 2×2 matrix (real) with trace $=0$. The Cayley-Hamilton theorem says that A satisfies its own characteristic equation

$$A^2 - (\operatorname{tr} A)A + (\det A)I = 0$$

i.e

$$\begin{cases} A^2 = -\Delta I \\ \Delta \equiv \det A \end{cases}$$

Show that

$$e^A = \begin{cases} (\cos\sqrt{\Delta}\,)I + (\frac{1}{\sqrt{\Delta}}\sin\sqrt{\Delta}\,)A & \text{if } \Delta > 0. \\ (\cosh\sqrt{|\Delta|})I + (\frac{1}{\sqrt{|\Delta|}}\sinh\sqrt{|\Delta|})A & \text{if } \Delta < 0 \end{cases}$$

and of course $e^A = I+A$ if $\Delta = 0$. Conclude then that

$$\operatorname{tr} e^A \geq -2.$$

Thus

$$g_0 \equiv \begin{bmatrix} -2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$$

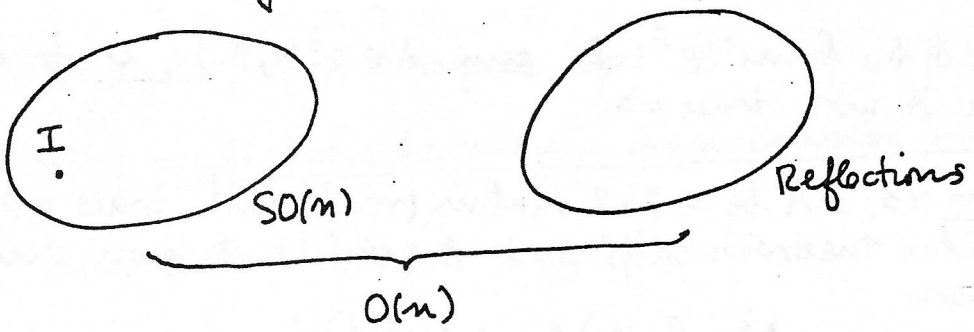
is never of the form e^A, $A \in sl(2,\mathbb{R})$. In particular g_0 does not lie on any 1-parameter subgroup of $Sl(2,\mathbb{R})$!

As mentioned above, since $SO(n)$ is compact and connected (we shall show connectedness later)

$$\exp : so(n) \longrightarrow SO(n)$$

is onto; any othogonal matrix with $\det > 0$ is of the form

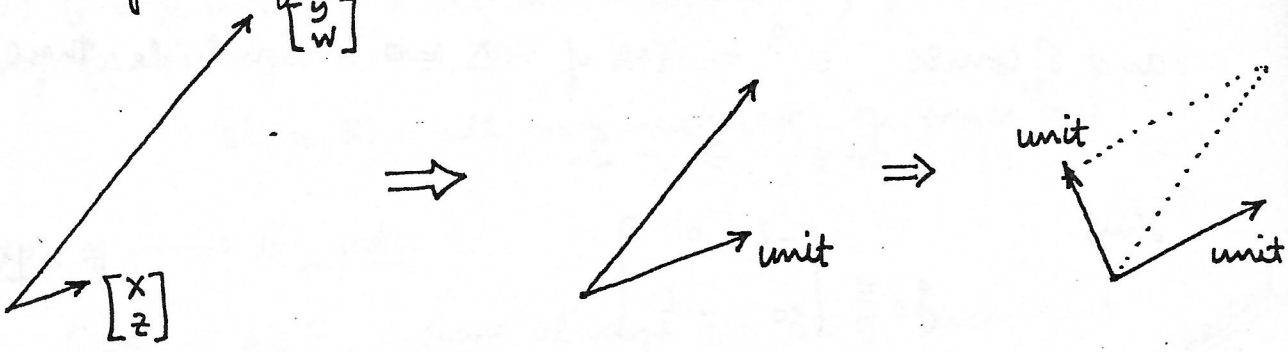
e^A for some skew symmetric matrix A. Now $O(n)$ is not connected, the "reflections", those $g \in O(n)$ with $\det g = -1$, cannot be joined to the identity by a continuous path of orthogonal matrices since $\det g$ cannot continuously change from -1 to $+1$.



Consider now the group $Sl(2, \mathbb{R})$. Since it is defined by $\det g = 1$ it is a 3 dimensional Lie group. We claim that $Sl(2, \mathbb{R})$ is connected. To see this

$$g = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in Sl(2, \mathbb{R})$$

can be pictured as a pair of column vectors $\begin{bmatrix} x \\ z \end{bmatrix}$ and $\begin{bmatrix} y \\ w \end{bmatrix}$ spanning a parallelogram of area $+1$.



By a Gram-Schmidt-like process we can continuously deform this pair into an orthonormal pair, always keeping area $= +1$. This gives us a deformation

$$Sl(2, \mathbb{R}) \longrightarrow SO(2)$$

onto the subgroup of rotation matrices. But $SO(2)$ is topologically a circle, and hence is connected.

Lecture 57. Lie Algebras and Subgroups

(1) **Left Invariant Fields Generate Right Translations.** Let X be a left invariant vector field on the Lie group G. If X_e is the value of X at the identity, then

$$e^{tX_e}$$

is the 1-parameter subgroup generated by X_e, and X_e is said to be the infinitesimal generator of this subgroup. X is tangent to this subgroup at each point of the subgroup.

Now X as a vector field on the manifold G generates a local flow, ie a local 1-parameter group of diffeomorphisms

$$\varphi_t(X) : G \longrightarrow G$$

having the property that the integral curves of X are these stream lines $g \longmapsto \varphi_t(X)g$,

$$\frac{d}{dt}\left(\varphi_t(X)\, g\right)\Big|_{t=0} = X_g.$$

But note that the curve $g \longmapsto g\, e^{tX_e}$ has the property

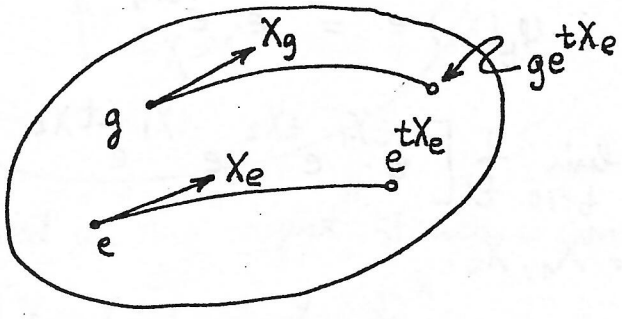
$$\frac{d}{dt}\left(g e^{tX_e}\right)\Big|_{t=0} = g \cdot \frac{d}{dt} e^{tX_e}\Big|_{t=0} = gX_e$$

$$= (L_g)_* X_e = X_g.$$

We conclude from the uniqueness of integral curves that

$$\varphi_t(X)(g) = g\, e^{tX_e} \tag{57.1}$$

i.e **the flow generated by the left invariant field is a 1-parameter group of right translations!**



Since a right invariant vector field Y is then automatically invariant under the flow generated by a left invariant vector field X, we conclude that their bracket vanishes.

$$\left[X^{\text{left invariant}}, Y^{\text{right invariant}} \right] = 0 \qquad (57.2)$$

Of course right invariant vector fields generate left translations, etc, etc.

(2) Commutators of Matrices. If $X_1 \in \mathfrak{g}$ and $X_2 \in \mathfrak{g}$, then the Lie bracket

$$[X_1, X_2] = \mathcal{L}_{X_1} X_2 \in \mathfrak{g}$$

is given by means of the Lie derivative, or as first order differential operators

$$[X_1, X_2] f = X_1 (X_2 f) - X_2 X_1 (f).$$

Now if G is a matrix group, each $X \in \mathfrak{g}$ is itself a matrix (not in G, but in the tangent space $\mathfrak{g}$). For example we have seen that if $G = SO(n)$ then X is a skew symmetric matrix. We claim that $[X_1, X_2]$ is then merely the commutator product of the matrices

$$[X_1, X_2] = X_1 X_2 - X_2 X_1 .$$

To see this, use equation (17.9). We have, at $e = I \in Gl(n, \mathbb{R})$

$$[X_1, X_2] = \lim_{t \to 0} \left[\varphi_{-t}(X_2)\, \varphi_{-t}(X_1)\, \varphi_t(X_2)\, \varphi_t(X_1)\, (e) - e \right] / t^2$$

But since X_1 and X_2 are left invariant their flows are right translations. For example

$$\varphi_t(X_1)\, e = e \cdot e^{tX_1} .$$

Thus

$$[X_1, X_2] = \lim_{t \to 0} \frac{1}{t^2} \left[e^{tX_1} e^{tX_2} e^{-tX_1} e^{-tX_2} - e \right]$$

for the matrices X_1, X_2.

Problem 94. Use (55.6) to show

$$[X_1, X_2] = X_1 X_2 - X_2 X_1$$

as matrices.

This shows, for instance, since $X_1 \in \mathfrak{g}$ & $X_2 \in \mathfrak{g} \Rightarrow [X_1, X_2] \in \mathfrak{g}$, that if X_1 and X_2 are skew symmetric (i.e in $so(n)$), then $[X_1, X_2] = X_1 X_2 - X_2 X_1$ is also skew symmetric, etc. etc.

(3) Right Invariant Vector Fields. All that we have said regarding left invariant vector fields can be redone for right invariant fields. For example, right invariant vector fields generate left translations. We have defined the Lie algebra $\mathfrak{g}$ to be essentially the vector space of left invariant fields, and then

$$[X_i, X_j] = X_k C^k_{ij}.$$

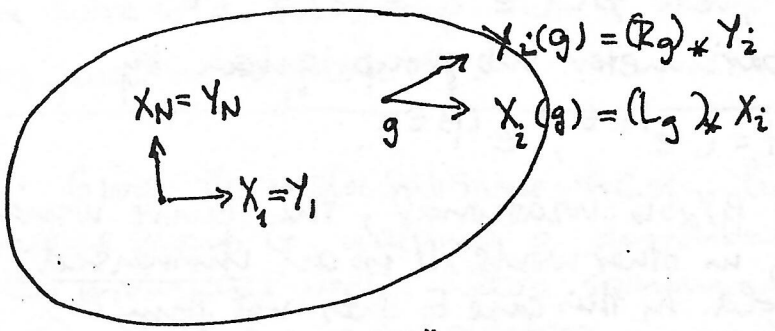
What would happen if we had used right invariant fields instead?

Let $X_1(e), \ldots, X_N(e)$ be a basis for G_e. Extend to left invariant fields $X_i(g)$ on all N of G.

$$X_i(g) = (L_g)_* X_i(e)$$

Now let $Y_1(e), \ldots, Y_N(e)$ coincide with $X_1(e), \ldots, X_N(e)$, and extend $Y_i(g)$ to be right invariant

$$Y_i(g) = (R_g)_* X_i(e) = (R_g)_* Y_i(e).$$



We are interested in the "right" structure constants

$$[Y_i, Y_j] = Y_k D^k_{ij}$$

We shall calculate this in the case of a matrix group, though the result will be true in general! Since the Y_i are right invariant the flow generated by Y_i consists of left translations. Then, repeating the steps going into Problem 14, but for right invariant Y, we see

$$[Y_1, Y_2]_{\text{right}} = Y_2 Y_1 - Y_1 Y_2$$

as matrices. Thus we conclude (since $Y_i = X_i$ at e)

$$[Y_i, Y_j]_{\text{right}} = - Y_k C^k_{ij}$$

ie the right structure constants are merely the negatives of the left structure constants.

By "Lie algebra", we shall always mean the algebra based on left invariant vector fields.

(4) Subgroups and Subalgebras. We shall be interested in subgroups of a Lie group (we have already discussed 1-parameter subgroups)! For example, $SO(n)$

$$SO(n) \subset Gl(n, \mathbb{R})$$

is a subgroup of the general linear group and it is an embedded submanifold (we showed this in §5 of Lecture 1)

The 2-torus, consisting of points

$$(e^{i\theta}, e^{i\varphi}) \in S^1 \times S^1$$

is a 2 dimensional abelian group

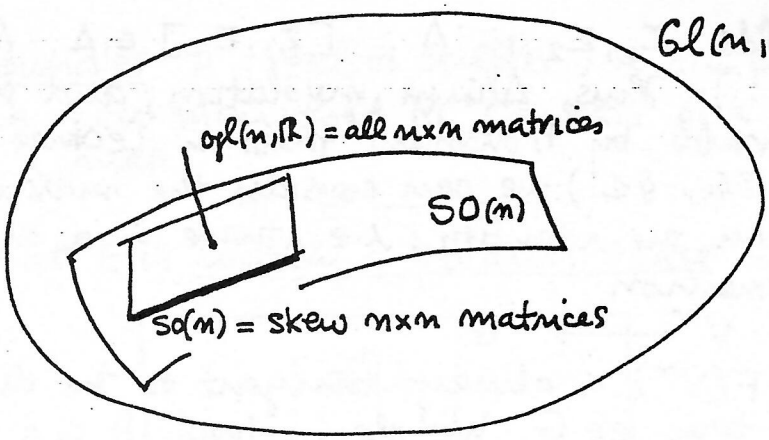
$$(e^{i\theta}, e^{i\varphi}) \cdot (e^{i\theta'}, e^{i\varphi'}) = (e^{i(\theta+\theta')}, e^{i(\varphi+\varphi')}).$$

with a 1-parameter subgroup given by

$$H = (e^{i\alpha t}, e^{i\beta t}).$$

However, if β/α is irrational, this curve winds densely on the torus; in other words H is an immersed, not imbedded submanifold. In this case H does not form a closed subset of T^2

For $H \subset G$ to be a subgroup of G we demand that H itself be a Lie group, that the inclusion map $H \subset G$ be a homomorphism, and the inclusion map is at least an immersion.



The tangent space to $Gl(n,\mathbb{R})$ consists of all $n\times n$ real matrices. The tangent space to $SO(n)$ is the vector subspace of all $n\times n$ real skew symmetric matrices. Let X_1, X_2 be skew matrices. Left translate them over all of $Gl(n,\mathbb{R})$. Since the resulting vector fields X_1, X_2 are tangent to $SO(n)$, so is their Lie bracket,

$$[X_1, X_2] \in so(n)$$

especially at e. This says that not only is $so(n)$ a vector subspace of $gl(n,\mathbb{R})$, it is a <u>subalgebra</u>.

In general, if $H \subset G$ is a subgroup of G, then the Lie algebra of H, $h = H_e$, is a subalgebra of g! The converse of this is also true, and very important.

> <u>Theorem</u>: Let G be a Lie group with Lie algebra g. Let $h \subset g$ be a vector subspace of g that is also a subalgebra
>
> $$[h, h] \subset h.$$
>
> Then there is a subgroup $H \subset G$ whose Lie algebra is the given subalgebra h.

"Proof": Given the vector subspace $h \subset g$, left translate h over the entire group G, yielding a distribution Δ. Let $X_1, \ldots, X_r$ be left invariant vector fields spanning the distribution Δ (everywhere). Since $\underline{h}$ is a <u>subalgebra</u>

$$[X_i, X_j] \in \Delta$$

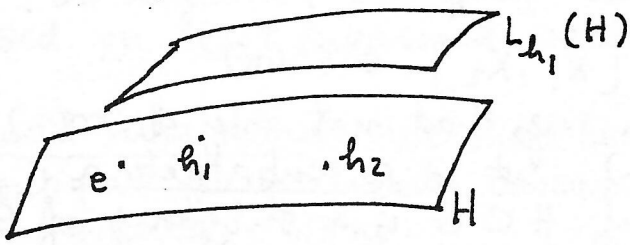
Since $[X_i, X_j]_g = (L_g)_* [X_i, X_j]_e$. It then follows easily that

for any vector fields Z_1, Z_2 in Δ, $[Z_1, Z_2] \in \Delta$ (convince yourself of this!). Thus Δ is in involution, and so is completely integrable by Frobenius' theorem Lecture 25, § 4. Thus (see Lecture 26, § 1) we can consider the maximal leaf passing thru the identity; i.e. there is a manifold V^r and an immersion

$$F : V^r \longrightarrow G$$

such that $H = F(V^r)$ is always tangent to the distribution Δ and H passes thru $e \in G$. We claim that H is a subgroup of G, ie H is closed under multiplication and taking inverses.

Suppose $h_1 \in H$ and $h_2 \in H$. Then by the definition of Δ, L_{h_1} must send the leaf H into another (perhaps distinct) leaf of the distribution Δ



However $L_{h_1}(e) = h_1 \in H$ shows that h_1 is in both leaves H and $L_{h_1}(H)$! But H is a maximal leaf, so $L_{h_1}(H) = H$. In particular

$$h_1 h_2 = L_{h_1}(h_2) \in H$$

and so H is closed under multiplication. A similar argument shows that H is closed under taking inverses. □.

Example. $gl(n, \mathbb{R})$ = all $n \times n$ real matrices. Let X_1 and X_2 be $n \times n$ matrices with trace = 0. Then, as matrices

$$\text{tr}\,[X_1, X_2] = \text{tr}\,(X_1 X_2 - X_2 X_1) = \text{tr}(X_1 X_2) - \text{tr}(X_2 X_1) = 0$$

Shows that the trace-less $n \times n$ matrices form a subalgebra of $gl(n, \mathbb{R})$. The subgroup of $Gl(n, \mathbb{R})$ whose Lie algebra is the set of trace-less matrices is of course $Sl(n, \mathbb{R})$. Note that

$$\exp: sl(n, \mathbb{R}) \longrightarrow Sl(n, \mathbb{R})$$

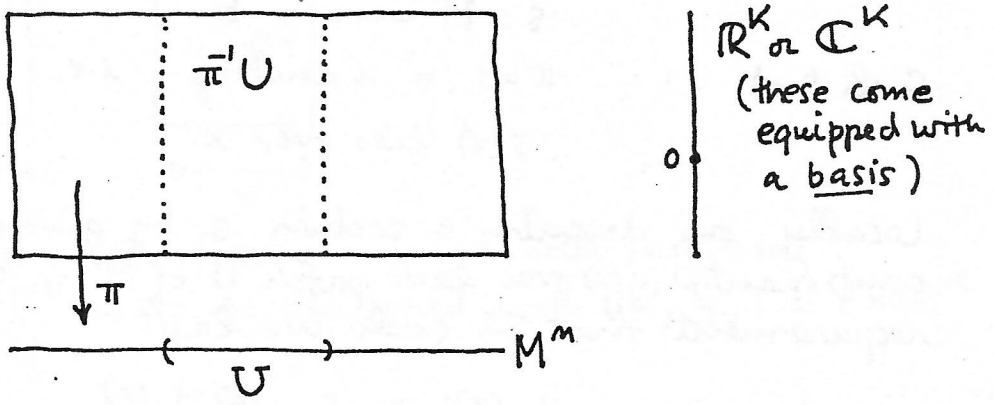
but that this map is not onto (see Problem 93).

Lecture 58 Vector Bundles

(1) **Vector Bundles** A (real or complex) K dimensional vector bundle E over a base manifold M^n consists of a topological space E and a map

$$\pi : E \longrightarrow M$$

such that E is <u>locally a product space</u> in the following sense.



Here is a covering $\{U, V, \dots\}$ of M^n by open subsets $U, \dots$ such that

$U \times \mathbb{R}^K$ (or $\mathbb{C}^K$) is diffeomorphic to the part of the bundle "over" U,

i.e. there are diffeomorphisms (we illustrate the real case)

$$\Phi_U : U \times \mathbb{R}^K \longrightarrow \pi^{-1}(U) \qquad (58.1)$$

$$\Phi_U(x, y_U) \in \pi^{-1}(x).$$

Thus the point $e \in \pi^{-1}(U) \subset E$ can be represented by a pair consisting of a point $\pi(e) = x \in U$, and a K-tuple of real (or complex) numbers $y_U = (y_U^1, \dots, y_U^K)$. We further demand that in an overlap $U \cap V$, the "vector components" of the same point e should be related by a non-singular <u>linear</u> transformation (that will vary differentiably with x)

$$y_V = c_{VU}(x)\, y_U \qquad (58.2)$$

i.e $$y_V^i = \sum_j c_{VU}(x)^i{}_j\, y_U^j$$

Note that each "fiber" $\pi^{-1}(x)$ is a K dimensional vector space and in the name "K dimensional vector bundle", K refers to the dimension of the fiber, not the dimension of E (which is $n+K$ or $n+2K$ in the real or complex cases).

A (cross-) section of E is a differentiable map

$$s : M^n \longrightarrow E$$

Such that $\pi \circ s = \text{identity}$, i.e

$s(x)$ lies over x.

Locally, one describes a section s by giving its vector components $y(x)$ over each patch U of M^n, subject to the requirement that in each overlap

$$y_V(x) = c_{VU}(x)\, y_U(x)$$

Since, in a triple overlap

$$y_W(x) = c_{WV}(x) y_V(x)$$
$$= c_{WV}(x)\, c_{VU}(x)\, y_U$$

we must have

$$c_{WU}(x) = c_{WV}(x)\, c_{VU}(x).$$

The transition functions $\{c_{UV}\}$ satisfy then the two relations

$$\begin{cases} c_{VU}(x) = c_{UV}^{-1}(x) \\ c_{WV}(x)\, c_{VU}(x)\, c_{UW}(x) = 1 \end{cases} \tag{58.3}$$

Conversely, suppose that we are given a manifold M^n with a covering $\{U, V, \dots\}$ and suppose that we are given matrix valued functions in each overlap

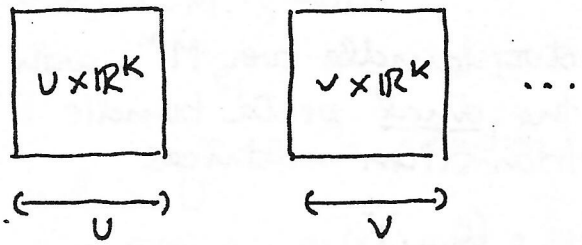
$$c_{VU} : U \cap V \longrightarrow Gl(n, \mathbb{R}) \text{ or } Gl(n, \mathbb{C})$$

that satisfy (58.3). Then we may construct a bundle E over M^n whose transition matrices are these given $\{c_{VU}\}$ as follows (we illustrate the real case)

We take a <u>disjoint</u> collection of manifolds

$$U \times \mathbb{R}^K, \quad V \times \mathbb{R}^K, \ \ldots$$

one for each patch on M^n. These are disjoint even though U and V might not be disjoint



Now we make identifications; we will say that $(x, y_U) \in U \times \mathbb{R}^K$ is to be identified with $(x', y_V) \in V \times \mathbb{R}^K$ if and only if

$$x = x'$$

and

$$y_V = c_{VU}(x)\, y_U$$

It can be shown that the resulting identification space E is indeed a K-dimensional vector bundle over M^n with the given transition matrices $\{c_{VU}\}$.

Let us recall that this really is the procedure we used to construct the <u>tangent bundle</u> TM^n to an M^n (see Lecture 3, § 1). On an overlap of coordinate patches U, V the vector components are related by

$$X^i_V = \sum_j \frac{\partial x^i_V}{\partial x_U{}^j}\, X^j_U$$

and so the transition matrices are

$$c_{VU}(x) = \frac{\partial x_V}{\partial x_U} \tag{58.4}$$

On the other hand, for the <u>cotangent</u> bundle T^*M^n

$$a^V_i = a^U_j \frac{\partial x^j_U}{\partial x^i_V}$$

ie

$$c_{VU}(x) = \left(\frac{\partial x_U}{\partial x_V}\right)^t \tag{58.5}$$

and so

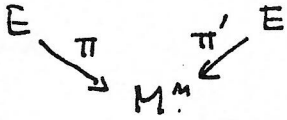
$$c_{VU}(T^*M) = \left[c_{VU}^{-1}(TM)\right]^t$$

$$c_{VU}(T^*M) = c_{UV}(TM)^t \qquad (58.6)$$

If E is any vector bundle over M^n, with transition matrices c_{VU}, then the <u>dual</u> vector bundle E^* is the vector bundle with transition matrices

$$c_{VU}^* = (c_{UV})^t.$$

If E and E' are vector bundles over the same base manifold M^n, then the <u>tensor product</u> bundle $E \otimes E'$ is the vector bundle with transition matrices



$$c_{VU} \otimes c'_{VU}$$

A point in $\pi^{-1}(U)$ has vector components $y_U = (y_U^1, \cdots y_U^K)$, a point in $\pi'^{-1}(U)$ has vector components $z_U = (z_U^1, \cdots, z_U^L)$, and a point in the tensor product bundle has $K \cdot L$ vector components

$$(y_U \otimes z_U)^{ir} = y_U^i z_U^r$$

and by definition

$$c_{VU} \otimes c'_{VU}\,(y_U \otimes z_U) = (c_{VU}\, y_U) \otimes (c'_{VU} z_U)$$

For example, the mixed tensors, once contravariant, once covariant (ie linear transformations on M^n) form the vector bundle

$$T(M) \otimes T^*M.$$

(2) Product or Trivial Bundles. A vector bundle E is locally a product manifold. The diffeomorphisms Φ_U of (58.1) that exhibit the local product structure are frequently obtained by the following procedure which mimicks the "moving frames" of Cartan (Lecture 38).

Let $e = (e_1, \dots e_K)$ be a "frame" of sections of E over the open set $U \subset M^n$; thus

$$e_i : U \longrightarrow E \quad \text{are differentiable}$$

$$\pi \circ e_i = \text{identity} : U \longrightarrow U$$

and at each $x \in U$

$$e_1(x), \dots, e_K(x) \text{ give a } \underline{\text{basis}} \text{ for } \pi^{-1}(x) \quad (\approx \mathbb{R}^K \text{ or } \mathbb{C}^K).$$

Then $\Phi_U : U \times \mathbb{R}^K \longrightarrow \pi^{-1}(U)$ is given by

$$\Phi_U(x, \psi^1, \dots, \psi^K) = \sum_{i=1}^{K} e_i(x)\psi^i$$

where the $\psi = (\psi^1, \dots, \psi^K)$ are constants. The most general cross section over U is of the form

$$\Psi(x) = \sum_{i=1}^{K} e_i(x)\,\psi^i(x)$$

where the $\psi^i(x)$ are differentiable functions. We abreviate this by matrix notation

$$\Psi(x) = e^U(x)\,\psi_U(x),$$

$$e(x) = (e_1^U(x), \dots e_K^U(x)), \quad \psi_U(x) = \begin{bmatrix} \psi_U^1(x) \\ \vdots \\ \psi_U^K(x) \end{bmatrix}$$

Of course if Ψ is a cross section over $U \cap V$ we have, over $U \cap V$

$$\Psi(x) = e^U(x)\,\psi_U(x) = e^V(x)\,\psi_V(x) \tag{58.7}$$

$$\psi_V(x) = c_{VU}(x)\,\psi_U(x)$$

If we can find a frame e over the entire M^n, not just over a subset U then we say that the bundle E is trivial;

In this case

$$\Phi(x, \psi^1, \ldots, \psi^K) = \sum_{i=1}^{K} e_i(x)\,\psi^i$$

gives a diffeomorphism

$$\Phi : M^n \times \mathbb{R}^K \longrightarrow E$$

making E globally a product manifold.

For example, TG, the tangent bundle to a Lie group G is a trivial bundle, $TG \approx G \times \mathbb{R}^N$ for left translating $\mathcal{G}$ (with basis $X_1, \ldots, X_N$) gives N globally independent cross-sections.

(When the tangent bundle TM^n is trivial we say that M^n is a parallelizable manifold).

Note that every vector bundle E possesses a global cross-section, the zero section, defined locally in each U by $\psi^1(x) = 0, \ldots, \psi^K(x) = 0$.

Problem 95. Show that the zero section is a global cross-section.

It is unusual for a non-trivial bundle E to have a global section that is independent of the zero section everywhere. For example, from Lecture 28 we know that among all closed orientable surfaces M^2_g of genus g, only the torus, $g = 1$ possesses a non vanishing section of the tangent bundle, and the torus $T^2 = M^2_{g=1}$ is parallelizable since it has two globally independent sections

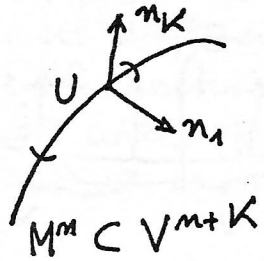
$$\partial/\partial\theta \quad \text{and} \quad \partial/\partial\varphi.$$

On the other hand, the Klein bottle has a single non vanishing section

$$\partial/\partial\theta$$

but does not have two independent sections.

For more examples, consider a Riemannian manifold V^{n+K} and a submanifold $M^n \subset V^{n+K}$. We define the <u>normal bundle</u> $\nu(M^n)$ to M^n in V^{n+K} to consist of those tangent vectors to V^{n+K} that are based at points of M^n and which are orthogonal to M^n.



It should be "clear" that if $U \subset M^n$ is sufficiently small then one can find K fields of normal vectors $n_1, \ldots, n_K$ that are independent, and even orthonormal, along all of U. Then if

$$E = \nu(M^n) \xrightarrow{\pi} M^n$$

denotes the normal bundle

$$\Phi : U \times \mathbb{R}^K \longrightarrow \pi^{-1}(U)$$

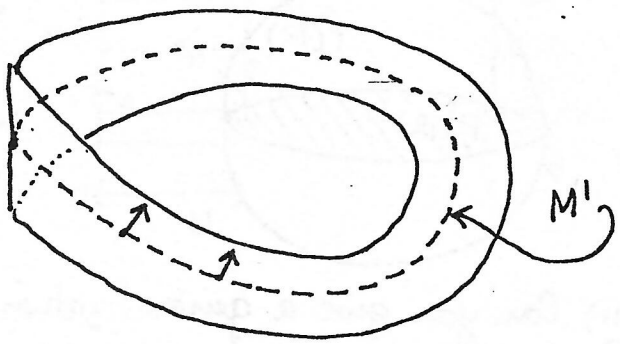
is defined again by

$$\Phi(x, \lambda^1, \ldots, \lambda^K) = \sum_{i=1}^{K} n_i(x)\, \lambda^i$$

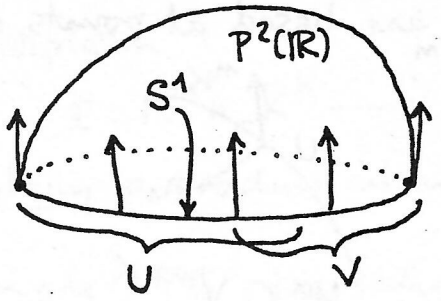
etc., etc.

For example, the normal bundle to $M^2 = S^2 \subset \mathbb{R}^3 = V^3$ is a product, $\nu(S^2) = S^2 \times \mathbb{R}^1$ since we have a global non-vanishing section given by the outward pointing unit normal.

The normal bundle to the central circle M^1 of the Möbius band is <u>not</u> trivial



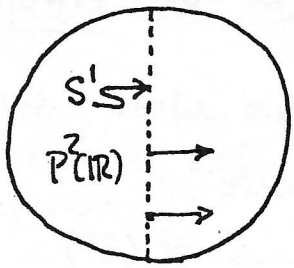
The normal bundle $\nu(S^1)$ to the indicated circle $S^1 \subset P^2(\mathbb{R})$ in the real projective plane is clearly itself an ∞ Möbius band (the lengths of the vectors are not bounded)



For this $S^1 \subset P^2(\mathbb{R})$

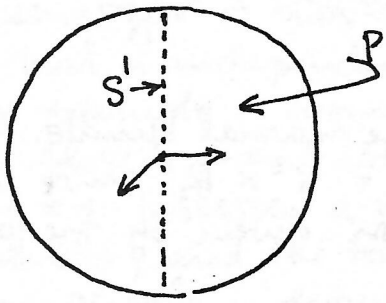
$\nu(S^1)$ is not trivial

Note that if we use the model of $P^2(\mathbb{R})$ consisting of a disc with antipodal points on the boundary identified then S^1 can be deformed into a diameter



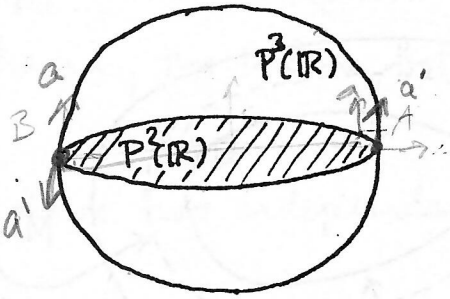
$\nu(S^1)$ not trivial

Problem 96. (i) Is the normal bundle to the circle $S^1 \subset P^3(\mathbb{R})$ trivial?



$P^3(\mathbb{R})$, the solid ball with boundary points identified antipodally.

(ii) Is the normal bundle to $P^2(\mathbb{R})$ in $P^3(\mathbb{R})$ trivial?



(iii) Can you give a generalization of (i) to $S^1 \subset V^3$ for some V^3?

Lecture 59 Connections in a Vector Bundle

(1) **Connection in a Vector Bundle.** Let $E \xrightarrow{\pi} M^n$ be a K dimensional vector bundle (real or complex) over a manifold M^n. Imitating the procedure of Lecture 37, § 2 we wish to have a notion of a connection or covariant differential, an operation that will associate to each section of E an E-valued 1-form on M^n ("E-valued" will replace "vector valued" of Lecture 37). For $U \subset M^n$ let us denote by

$$\mathcal{C}(U; E)$$

the (infinite dimensional) vector space (coefficients in $\mathbb{R}$ or $\mathbb{C}$) of sections $\Psi : U \to E$ of E over U. A connection $\mathcal{D}$ will take a section and give back an E-valued 1-form., ie

$$\mathcal{D} : \mathcal{C}(U; E) \longrightarrow \mathcal{C}(U; E \otimes T^*M). \tag{59.1}$$

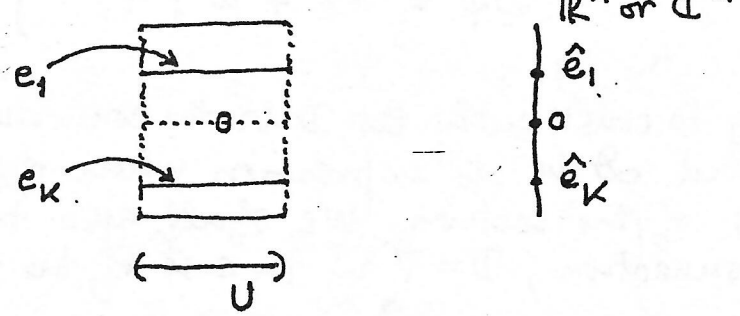
This means the following. Suppose that U is one of our open sets on M^n over which the bundle is trivial, as in (58.1). We then say that Φ_U gives a "local trivialization". Consider the frame of K sections over U

$$e_1, e_2, \cdots, e_K$$

given by this local trivialization; i.e

$$e_i(x) = \Phi_U(x, \hat{e}_1) \tag{59.2}$$

where $\hat{e}_1, \hat{e}_2, \cdots, \hat{e}_K$ is the given standard basis $\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \cdots \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$ of $\mathbb{R}^K$ (or $\mathbb{C}^K$)



Then to give $\mathcal{D}$ on U is to give a matrix $\omega = (\omega^i{}_j)$ of 1-forms on U

$$\left.\begin{aligned} \mathcal{D} e_i &= \sum_{j=1}^{K} e_j \otimes \omega^j{}_i \\ &\text{or} \\ \mathcal{D} e &= e \otimes \omega \\ \omega^j{}_i(x) &= \sum_{\mu=1}^{n} \Gamma^j_{\mu i}\, dx^\mu \end{aligned}\right\} \qquad (59.3)$$

(we will try to use consistently the notation

Greek lower case, $\alpha, \beta, \ldots \mu, \nu, \ldots$ run from $1, \ldots, n$, indices for M^n
Latin lower case, $i, j, \ldots$ run from $1, \ldots, K$, indices for E

Note that the connection coefficients $\Gamma^j_{\mu i}(x)$ have a mixture of M^n and E indices. We are also assuming that (x^μ) are local coordinates in $U \subset M^n$. Since we wish a Leibnitz rule, if Ψ is a section and f a function on U

$$\mathcal{D}(\Psi f) = (\mathcal{D}\Psi) f + \Psi\, df.$$

Thus

$$\mathcal{D}(\Psi) = \mathcal{D}(e_i \psi^i) = \mathcal{D}(e\psi) = (De)\psi + e\, d\psi$$

i.e

$$\left.\begin{aligned} \mathcal{D}\Psi &= \mathcal{D}(e\psi) = e \otimes \mathscr{D}\psi \\ \mathscr{D}\psi &= d\psi + \omega\psi \\ \mathscr{D}\psi^i &= d\psi^i + \omega^i{}_j\, \psi^j \end{aligned}\right\} \qquad (59.4)$$

We shall try to consistently let $\mathcal{D}$ be the connection operator on sections and $\mathscr{D}$ be its expression when operating on components of the section. We shall also try to reserve ∇ for the connection, $\mathcal{D} = \nabla$ in the case of the tangent bundle.

In order that D be <u>globally defined</u> we demand that in an overlap $U \cap V$

$$\psi_V = c_{VU}\,\psi_U \quad \Longrightarrow \quad \mathcal{D}\psi_V = c_{VU}\,\mathcal{D}\psi_U \tag{59.5}$$

Just as in (38.7) we have then;

$$\omega_V = c_{UV}^{-1}\,\omega_U\,c_{UV} + c_{UV}^{-1}\,d\,c_{UV}. \tag{59.6}$$

Note that in our convention $\psi_V = c_{VU}\,\psi_U$ and so

$$e_V = e_U\,c_{VU}^{-1} = e_U\,c_{UV} \tag{59.7}$$

As usual, we define the covariant derivative $D_X\Psi$ of the section Ψ with respect to the tangent vector X on M^m by

$$\begin{aligned} D_X\Psi &\equiv D\Psi(X) \\ &= (e\otimes\mathcal{D}\psi)(X) \\ &= e\cdot[\mathcal{D}\psi(X)] \\ &= e\ (d\psi + \omega\psi)(X) \\ &= e\ (X(\psi) + \omega(X)\psi) \end{aligned} \tag{59.8}$$

and where

$$\omega^i{}_j(X) = \Gamma^i_{\mu j}\,dx^\mu(X) = \Gamma^i_{\mu j}\,X^\mu$$

Then

$$D_X\Psi = e_i\left(\frac{\partial\psi^i}{\partial x^\mu}X^\mu + \Gamma^i_{\mu j}X^\mu\psi^j\right)$$

which we write

$$\left.\begin{aligned} D_X\Psi &= e\,\mathcal{D}_X\psi \\ \mathcal{D}_X\psi &= X^\mu\mathcal{D}_\mu\psi \\ \mathcal{D}_\mu\psi^i &= \frac{\partial\psi^i}{\partial x^\mu} + \Gamma^i_{\mu j}\psi^j \end{aligned}\right\} \tag{59.9}$$

We have, so far, the connection

$$\mathcal{D}: \quad \mathcal{C}(U;E) \longrightarrow \mathcal{C}(U;\, E\otimes \Lambda^1 M)$$

where $\Lambda^1 M = T^*M$ is the cotangent bundle. We can then let $\mathcal{D}$ act on E-valued p-forms on M

$$\mathcal{D}: \quad \mathcal{C}(U;\, E\otimes \Lambda^p M) \longrightarrow \mathcal{C}(U;E\otimes \Lambda^{p+1} M)$$

as usual by (recall (37.6))

$$\mathcal{D}(\Psi \otimes \alpha^p) = \mathcal{D}\Psi \otimes \alpha^p + \Psi \otimes d\alpha^p \tag{59.10}$$

leading, as usual, to <u>curvature</u>

$$\left.\begin{aligned} \mathcal{D}^2(e) &= \mathcal{D}(e\otimes\omega) = e\otimes\vartheta \\ \vartheta &\equiv d\omega + \omega\wedge\omega \\ \vartheta^i{}_j &= \tfrac{1}{2} R^i{}_{j\mu\nu}\, dx^\mu \wedge dx^\nu \end{aligned}\right\} \tag{59.11}$$

Note again the mixture of indices in the curvature tensor

(2) <u>Complex Vector Spaces</u>. Consider the complex plane $\mathbb{C}$ with coordinate $z = x+iy$. $\mathbb{C}$ is of course a vector space of dimension 1 when we allow complex scalars, but $\mathbb{C}$ can also be thought of as a 2 dimensional real vector space

$$z = x+iy \quad \longleftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

and addition of complex numbers corresponds exactly to addition of vectors in $\mathbb{R}^2$. The interesting thing about $\mathbb{C}$ is that it admits an interesting multiplication

$$z_1 z_2 = (x_1+iy_1)(x_2+iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

Of course this can be expressed purely in real terms.

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} * \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1x_2 - y_1y_2 \\ x_1y_2 + x_2y_1 \end{pmatrix}$$

In particular, multiplication in $\mathbb{C}$ by i gets translated in real terms as a linear transformation

$$J : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

whose matrix is simply

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{59.12}$$

with naturally $J^2 = -I$. Likewise, complex K space, $\mathbb{C}^K$, the vector space (with complex coefficients, i.e. scalars) of complex K-tuples

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = \begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_K + iy_K \end{bmatrix}$$

can be considered as $\mathbb{R}^{2K}$

$$z \longleftrightarrow \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \\ x_K \\ y_K \end{bmatrix} = \underset{\sim}{x}$$

and then multiplication by i in $\mathbb{C}^K$, $z \longmapsto iz$, is translated into a linear transformation J whose matrix is

$$J = \begin{bmatrix} 0 & -1 & & & & \\ 1 & 0 & & & 0 & \\ & & 0 & -1 & & \\ & & 1 & 0 & & \\ & 0 & & & \ddots & \\ & & & & 0 & -1 \\ & & & & 1 & 0 \end{bmatrix} \tag{59.13}$$

with again $J^2 = -I$.

Note that $J : \mathbb{R}^{2K} \longrightarrow \mathbb{R}^{2K}$ is also an <u>isometry</u> with respect to the usual Euclidean metric of $\mathbb{R}^{2K}$

$$\langle \underset{\sim}{x}, \underset{\sim}{x}' \rangle = \langle J\underset{\sim}{x}, J\underset{\sim}{x}' \rangle \tag{59.14}$$

Let now E^{2K} be a real, even dimensional vector space with an inner product $\langle , \rangle$. Let

$$J : E^{2K} \longrightarrow E^{2K}$$

be a linear transformation that is an isometry (59.14) and also an <u>involution</u>, i.e

$$J^2 = -I.$$

Then automatically J is skew-symetric

$$\langle J\underset{\sim}{x}, \underset{\sim}{x}' \rangle = \langle J^2\underset{\sim}{x}, J\underset{\sim}{x}' \rangle = -\langle \underset{\sim}{x}, J\underset{\sim}{x}' \rangle$$

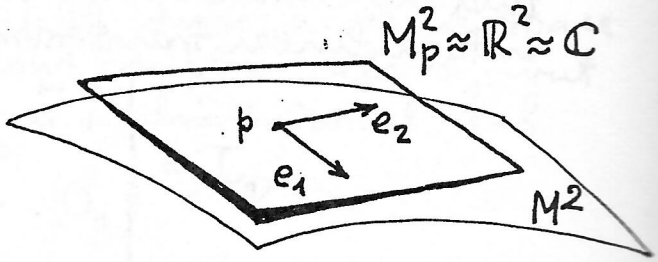
and linear algebra then tells us that an o.n basis for E^{2K} can be chosen so that J has matrix (59.13). Then one can introduce <u>complex coordinates</u> by putting $z_1 = x_1 + iy_1, \cdots, z_K = x_K + iy_K$, and $J: E^{2K} \longrightarrow E^{2K}$ then corresponds simply to multiplication by i.

In particular, $\mathbb{R}^2$ with J as above can be considered a complex 1 dimensional vector space $\mathbb{C}^1 = \mathbb{C}$, which we call the <u>complex line</u>.

(3) <u>Complex Vector Bundles</u>. Let M^2 be a Riemannian <u>surface</u> and we shall assume that M^2 is orientable. We can cover M^2 by patches $U, V, \cdots$, each of which supports an o.n frame $e_U, e_V, \cdots$ of tangent vectors. Suppressing the patch index

$$e_U = (e_1, e_2)$$

is an o.n frame in U.



We shall assume that M^2 has been oriented and that the frames $e_U, e_V, \cdots$ are always positively oriented. It is then clear that each transition matrix for TM is of the form

$$c_{VU}(x) = \begin{bmatrix} \cos\alpha(x) & -\sin\alpha(x) \\ \sin\alpha(x) & \cos\alpha(x) \end{bmatrix} \in SO(2)$$

[<u>An aside</u>: Ordinarily, for the tangent bundle TM^n, $c_{VU}(x) \in Gl(n, \mathbb{R})$; we say that the "structure group" for TM^n is the Lie group $Gl(n, \mathbb{R})$. However a <u>Riemannian</u> manifold allows us to "reduce the structural group" to $O(n)$. Orientability allows us to further reduce to $SO(n)$].

Define $J: M_p^2 \longrightarrow M_p^2$ simply by rotation thru 90° in the positive sense; thus

$$Je_1 = e_2 \qquad Je_2 = -e_1 .$$

It should be clear that J is <u>globally</u> defined; in an overlap $U \cap V$ the action of J using the frame e_V coincides with the action of J using the frame e_U. <u>Thus J allows us to consider each fiber of TM^2 as a complex line</u>! The real vector

$$e_1 \in M_p^2 \approx \mathbb{R}^2$$

can be considered as a complex basis vector

$$e = e_1 \in M_p^2 \approx \mathbb{C}$$

of the complex line M_p^2; then

$$ie = Je_1 = e_2 .$$

In terms of these bases $e^U = e^U_1$, $e^V = e^V_1, \ldots$ the transition matrices $c_{VU}(x)$ become simply the complex <u>numbers</u>

$$c_{VU}(x) = e^{i\alpha(x)}$$

<u>The tangent bundle TM^2 to an orientable Riemannian surface can be considered as a complex line bundle</u>! We can say that we have reduced the structure group from $Gl(2,\mathbb{R})$ to $U(1)$, the unitary group in 1 variable.

The Riemannian connection for M^2 is a connection in the real 2 dimensional tangent bundle. In terms of the o.n frames $e^U, e^V, \ldots$, we have

$$De_i = e_j \otimes \omega^j{}_i = e_j \otimes \omega_{ji}$$

and we know $\omega_{ij} = -\omega_{ji}$ (see Lecture 39)

$$De_1 = e_2 \otimes \omega_{21} \qquad (59.15)$$

A connection matrix in a complex line bundle would be a 1×1 matrix, ie a function $\tilde{\omega}$. We would have in our complex line bundle version of TM, where $e = e_1$

$$\tilde{D}e = e \otimes \tilde{\omega}$$

and since $e_2 = i e_1$ we can rephrase (59.15) as

$$\tilde{D}e = ie_1 \otimes \omega_{21} = e_1 \otimes i\omega_{21}$$

i.e

$$\left.\begin{aligned} \tilde{D}e &= e \otimes \tilde{\omega} \\ \text{where} \quad \tilde{\omega} &\equiv i\omega_{21} = -i\omega_{12} \end{aligned}\right\} \qquad (59.16)$$

Does this really mean that $\tilde{\omega} = -i\omega_{12}$ defines a connection in the complex line bundle? For this to be true we certainly must have that

$$\tilde{D}(i\psi) = i\tilde{D}\psi$$

for any cross section ψ (i.e tangent vector field to M^2). But, e.g.

$$D(ie) = De_2 = e_1 \otimes \omega_{12} = -ie_1 \otimes i\omega_{12}$$

$$= i \cdot (e_1 \otimes -i\omega_{12})$$

$$= i \cdot De$$

as desired!

What is the curvature of the connection for this complex line bundle version of TM^2? It is the single 2-form

$$\tilde{\vartheta} = d\tilde{\omega} + \tilde{\omega} \wedge \tilde{\omega} = d\tilde{\omega}$$

since $\tilde{\omega}$ is now a single 1-form.

$$\tilde{\vartheta} = d\tilde{\omega} = d(-i\omega_{12}) = -i\, d\omega_{12}$$

$$= -i\,\vartheta_{12}$$

and so

$$\boxed{\tilde{\vartheta} = -iK\sigma^1 \wedge \sigma^2} \qquad (59.17)$$

where again $K = R^{12}{}_{12}$ is the Gauss-Riemann curvature of M^2

Lecture 60. The Electromagnetic Connection

(1) Minimal Coupling. Consider a particle moving in $\mathbb{R}^3$ non-relativistically with Hamiltonian H, and with no electromagnetic field present. Hamilton's equations are (see Lecture 20)

$$i_X \Omega = 0$$

i.e

$$\begin{cases} i_{\underset{\sim}{v}} + \frac{dp}{dt} + \frac{\partial}{\partial t} & (\omega - dH \wedge dt) = 0 \\ \omega = \sum_{\nu=0}^{3} dp_\nu \wedge dx^\nu. \end{cases}$$

Written out in full, we have

$$0 = i_X \Omega = \left(\frac{dp_\nu}{dt} + \frac{\partial H}{\partial x^\nu} \right) dx^\nu - \left(\frac{dx^\nu}{dt} - \frac{\partial H}{\partial p_\nu} \right) dp_\nu - \left(\frac{dH}{dt} - \frac{\partial H}{\partial t} \right) dt. \tag{60.1}$$

Suppose now that the particle has a charge e and that an electromagnetic field is introduced. The particle then suffers an additional Lorentz force

$$\frac{dp_\nu}{dt} = -\frac{\partial H}{\partial x^\nu} + e \left(\mathcal{E} - i_v \mathcal{B} \right)_\nu$$

(we have put $\gamma = 1$ since particle is non-relativistic) and the energy of the particle also changes because of the work done by the electromagnetic field at rate $e \underset{\sim}{E} \cdot \underset{\sim}{v}$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + e\, i_v \mathcal{E}$$

Thus, instead of (60.1) we wish to have

$$\left(\frac{dp_\nu}{dt} + \frac{\partial H}{\partial x^\nu} - e \left(\mathcal{E} - i_v \mathcal{B} \right)_\nu \right) dx^\nu - \left(\frac{dx^\nu}{dt} - \frac{\partial H}{\partial p_\nu} \right) dp_\nu - \left(\frac{dH}{dt} - \frac{\partial H}{\partial t} - e\, i_v(\mathcal{E}) \right) dt = 0 \tag{60.2}$$

But note that if $\mathcal{F} = \mathcal{E} \wedge dt + \mathcal{B}$, then

$$e\, i_X \mathcal{F} = i_{V + \frac{dp}{dt} + \frac{\partial}{\partial t}}\, e(\mathcal{E} \wedge dt + \mathcal{B})$$

$$= e\,(i_V \mathcal{E})\, dt + e\, i_V \mathcal{B} - e\,\mathcal{E}$$

$$= -e\,(\mathcal{E} - i_V \mathcal{B}) + e\, i_V(\mathcal{E})\, dt$$

and so (60.2) is equivalent to

$$i_X(\Omega + e\mathcal{F}) = 0 \tag{60.3}$$

In other words, we can take the electromagnetic field into account merely by replacing the symplectic 2-form ω on $T^*M \times \mathbb{R}$ by the 2-form

$$\omega + e\mathcal{F}$$

(actually, instead of $\mathcal{F}$ we are using $\pi^*\mathcal{F}$, where π is projection

$$\pi : T^*M \times \mathbb{R} \longrightarrow M \times \mathbb{R}).$$

Finally, by using the vector potential

$$\mathcal{F} = dA = d(A_0\, dt + \underset{\sim}{A} \cdot d\underset{\sim}{x})$$

$$\Omega + e\mathcal{F} = d(\underset{\sim}{p} \cdot d\underset{\sim}{x} - H\, dt + e\,\underset{\sim}{A}\, d\underset{\sim}{x} + eA_0\, dt)$$

$$= d\left[(\underset{\sim}{p} + e\underset{\sim}{A}) \cdot d\underset{\sim}{x} - (H - eA_0)\, dt\right]$$

we see that we can, alternately keep the original symplectic form ω but redefine momentum and energy by

(canonical momentum)

$$\left.\begin{aligned} \underset{\sim}{p} &\to \underset{\sim}{p} + e\underset{\sim}{A} \\ H &\to H - eA_0 \end{aligned}\right\} \tag{60.4}$$

"minimal coupling".

(2) The Quantum Electron in the Electromagnetic Field. Consider an electron of charge e moving in an electromagnetic field in Minkowski space. If we neglect the spin of the electron it is represented in quantum mechanics by a wave function

$$\psi(x) = \psi(\underset{\sim}{x}, t)$$

a complex valued function. With no electromagnetic field present the operators for momentum and energy are given by

$$\begin{cases} p_i = -i\hbar\, \partial/\partial x^\nu & \nu = 1,2,3. \\ H = i\hbar\, \partial/\partial t = i\hbar\, \partial/\partial x^0 \end{cases}$$

If an electromagnetic field is introduced, minimal coupling tells us to replace these operators by

$$-i\hbar\, \partial/\partial x^\nu + e A_\nu = -i\hbar \left(\partial/\partial x^\nu + \frac{ie}{\hbar} A_\nu \right)$$

$$i\hbar\, \partial/\partial x^0 - e A_0 = i\hbar \left(\partial/\partial x^0 + \frac{ie}{\hbar} A_0 \right)$$

(The A's are multiplication operators)

Schrödingers equation for ψ involve now not the partial derivatives alone but rather the "covariant derivatives"

$$\mathcal{D}_\mu \psi = \frac{\partial \psi}{\partial x^\mu} + \frac{ie}{\hbar} A_\mu \psi = \frac{\partial \psi}{\partial x^\mu} + \Gamma_\mu \psi \qquad (60.5)$$

$\Gamma^i_{\mu j}$ $i,j = 0$ only

Now it is not merely a convenience to introduce covariant derivatives; it might be a necessity, for it might be that that the vector potential A is not globally defined in all of Minkowski space (as we shall see below). In that case we shall apparently have a collection of patches $U, V, \ldots$ each with its own A_μ and consequently with its own version of Schrödingers equation. We can not expect then to find a well defined wave function ψ for the entire Minkowski space. Such a procedure however is expected if we consider ψ to be not a function on M^4

$$\psi : M^4 \longrightarrow \mathbb{C}$$

$M^4 \times \mathbb{C}$ (ψ is cross-section)

but rather a cross section of some complex line bundle E over M^4.

What is the bundle E? To understand this we go back to the vector potential A. We wish that for our unknown bundle

$$\mathcal{D}\psi = \mathcal{D}_\mu \psi \, dx^\mu = d\psi + \frac{ie}{\hbar} A_\mu dx^\mu \psi \qquad (60.6)$$

and so we wish the 1-form

$$\frac{ie}{\hbar} A_\mu dx^\mu = \omega = \frac{ie}{\hbar} A \;, \qquad \Gamma_\mu = \frac{ie}{\hbar} A_\mu \qquad (60.7)$$

to be our connection form. In two patches U, V we have

$$\omega^U = \frac{ie}{\hbar} A^U \qquad \omega^V = \frac{ie}{\hbar} A^V.$$

must be simply connected and regions so closed curves are exact

Now the vector potentials must differ by an exact differential since $\mathcal{F} = dA$ must be well defined (at least in the case when U, V, ..., are simply connected)

$$A^V = A^U + df^{VU} \qquad \text{in } U \cap V$$

where f^{VU} is a real valued function on $U \cap V$. Then

$$\omega^V = \omega^U + d\frac{ie}{\hbar} f^{UV}.$$

But a connection in a complex line bundle transforms, from (59.6) when c_{UV} are 1×1 matrices

$$\omega^V = \omega^U + c_{UV}^{-1} \, d\, c_{UV} = \omega^U + d(\ln c_{UV}) \qquad (60.8)$$

and so we can choose

$$\ln c_{UV}(x) = \frac{ie}{\hbar} f^{UV}$$

or

$$c_{UV}(x) = \exp\left(\frac{ie}{\hbar} f^{UV}\right) \qquad (60.9)$$

Thus the complex line bundle in question, the natural bundle to consider whose cross sections will be our local wave functions, is the line bundle whose transition functions are given by (60.9).

A wave function is then not a single complex valued function ψ but rather a collection of functions ψ_U, ψ_V such that in an overlap

$$\psi_V(x) = c_{VU}(x)\,\psi_U(x) = e^{-\frac{ie}{\hbar} f^{UV}} \psi_U \tag{60.10}$$

This brings us back to the starting point of gauge theories in quantum mechanics, namely H. Weyl's principle of gauge invariance;

> If ψ satisfies the classical Schrödinger equation (which involves the potential A) then
>
> $$e^{-\frac{ie}{\hbar} f}\,\psi$$
>
> satisfies Schrödingers equation where A has been replaced by
>
> $$A + df$$

Note that the transition functions (60.9) for our complex line bundle are functions of absolute value 1; the map

$$\mathbb{C} \longrightarrow \mathbb{C} \text{ given by } \psi \longrightarrow c_{VU}\,\psi$$

is a unitary transformation, i.e the structure group for this complex line bundle is the unitary group $U(1)$. Note that the curvature of the line bundle is the E&M field tensor

$$\boxed{\vartheta = d\omega = \frac{ie}{\hbar} dA = \frac{ie}{\hbar} \mathcal{F}} \tag{60.11}$$

Example; the Monopole. Consider an electron in $\mathbb{R}^3 - 0$ and a magnetic monopole of strength q fixed at the origin. The $\underset{\sim}{B}$ field for this monopole is (see Lecture 21, §4)

$$B^2 = i_{\frac{q}{r^2}\frac{\partial}{\partial r}} \text{vol}^3 = d[(1-\cos\theta)\,d\varphi]$$

ie

$$A'_U = q(1-\cos\theta)\,d\varphi \qquad V = \sqrt{\mathbb{R}^3}$$

in the region $U = \mathbb{R}^3$ – negative z axis. In the region $\mathbb{R}^3$ – positive z axis we can use $\theta' = \pi - \theta$ and $\varphi' = -\varphi$ as coordinates

and get

$$A_V = -q(1+\cos\theta)\,d\varphi$$

Problem 97 Show that the transition function for the monopole bundle is

$$c_{VU}(x) = e^{\frac{2ieq}{\hbar}\varphi}$$

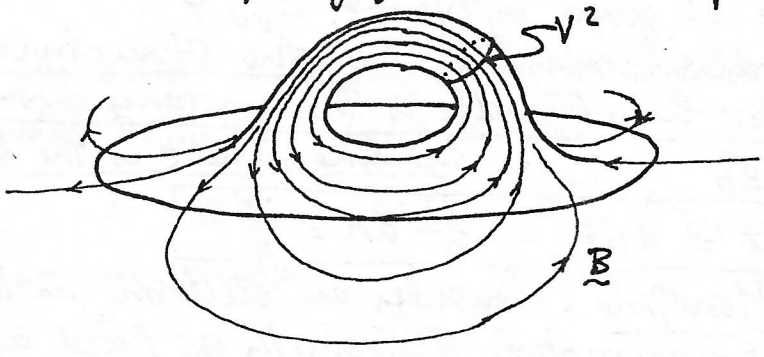
And we can conclude, with Dirac (1931), that in order for $c_{VU}(x)$ to be single valued (as is necessary) in $U \cap V$ we must have

$$\boxed{\frac{2eq}{\hbar} \text{ must be an integer}} \qquad 60.12$$

(this tells us that if a monopole exists, then electric charge e must be quantized!)

We shall see in a future lecture that the monopole bundle is not a product bundle.

Instead of considering an exotic particle (the monopole) in a rather mundane space ($\mathbb{R}^3$) we can get other electromagnetic line bundles of interest by considering 3-dimensional spaces M^3 with exotic topology. In the example of Lecture 23



if $\underset{\sim}{B}$ has a non-zero flux thru the closed orientable V^2 then certainly there can be no global single valued vector potential A. As we shall see this will again give rise to a nontrivial complex line bundle (with connection $\frac{ie}{\hbar c}A$ and curvature $\frac{ie}{\hbar c}\mathcal{F}$).

Problem 98. In the monopole bundle with $\frac{2eq}{\hbar} \neq 0$, the function $\psi_U \equiv 1$ in $U = \mathbb{R}^3$ – neg. z-axis is a cross section over U. Can ψ be extended to be a global cross section in all of $U \cup V = \mathbb{R}^3$ – origin?

(3) Remark on Weyl's Principle of Gauge Invariance. Schrödingers equation, in the presence of an electromagnetic field (or more generally in the presence of a vector potential A (e.g. in the exterior to an infinite solenoid, as in the Bohm-Aharonov effect)) is, in a patch U.

$$i\hbar \mathcal{D}_0 \psi_U = -\frac{\hbar^2}{2m} \mathcal{D}_\mu \mathcal{D}_\mu \psi_U + V \psi_U$$

In an overlapping patch V we have a similar equation for ψ_V, and recall

$$\mathcal{D}_\mu \psi_U = \frac{\partial \psi_U}{\partial x^\mu} + \frac{ie}{\hbar} A^U_\mu \psi \qquad \mu = 0, 1, 2, 3$$

$$\mathcal{D}_\mu \psi_V = \frac{\partial \psi_V}{\partial x^\mu} + \frac{ie}{\hbar} A^V_\mu \psi$$

Since $\psi_V = c_{VU} \psi_U$ and

$$\mathcal{D} \psi_V = \mathcal{D}(c_{VU} \psi_U) = c_{VU} \mathcal{D} \psi_U \quad \text{(see (59.5))}$$

we get immediately

$$i\hbar \mathcal{D}_0 \psi_V \quad \text{=====} \quad -\frac{\hbar^2}{2m} \mathcal{D}_\mu \mathcal{D}_\mu \psi_V + V \psi_V$$

$$\| \qquad\qquad\qquad\qquad\qquad\qquad \|$$

$$c_{VU}\, i\hbar \mathcal{D}_0 \psi_U \quad \text{=====} \quad c_{VU} \left(-\frac{\hbar^2}{2m} \mathcal{D}_\mu \mathcal{D}_\mu \psi_U + V \psi_U\right)$$

which illustrates Weyl's principle. Note that a <u>direct</u> verification is rather horrible since Schrödingers equation, when written out <u>without</u> connections involves

$$\mathcal{D}_\mu \mathcal{D}_\mu \psi = \frac{\partial^2 \psi}{\partial x^{\mu 2}} + \frac{ie}{\hbar}\left(A_\mu \frac{\partial \psi}{\partial x^\mu} + \frac{\partial}{\partial x^\mu}(A_\mu \psi)\right) - \frac{e^2}{\hbar^2} A_\mu A_\mu \psi$$

Note also the Weyl's principle is in no way restricted to Schrödingers equation, as the above calculation indicates.

Note also that the connection $\mathcal{D}$ has <u>nothing</u> to do with a connection for Minkowski space!

Lecture 61 Fiber Bundles and Principal Bundles

(1) Fiber Bundles. This is generalizes the notion of vector bundle; we shall replace the vector space $\mathbb{R}^K$ or $\mathbb{C}^K$ by an arbitrary manifold F called the fiber. For a fiber bundle we need the following. There is again a bundle space E over a manifold M^n (the base space)

$$\pi : E \longrightarrow M^n$$

E shall be locally a product space in the following sense. There should be a covering of M^n by open sets $U, V, \ldots$ such that

$$\pi^{-1}(U) \text{ is diffeomorphic to } U \times F$$

for a given fixed manifold F, the fiber,

$$\Phi_U : U \times F \longrightarrow \pi^{-1}(U)$$

is a diffeomorphism with $\Phi_U(x, y) \in \pi^{-1}(x)$. In an overlap $U \cap V$, a point $e \in \pi^{-1}(U \cap V)$ will have two representations

$$e = \Phi_U(x, y_U) = \Phi_V(x, y_V)$$

and we demand that

$$y_V = c_{VU}(x)\,(y_U)$$

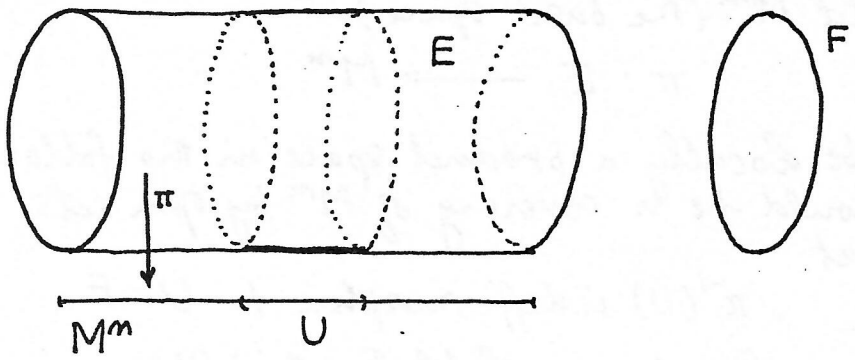
Here $c_{VU}(x)$, for each $x \in U \cap V$ is a diffeomorphism of the fiber F

$$c_{VU} : U \cap V \longrightarrow \text{the group of all diffeomorphisms of the manifold } F$$

We dont ask that $c_{VU}(x) : F \to F$ be linear since F need not be a vector space. For example F may be a K-sphere S^K; then $c_{VU}(x)$ is simply, for each x, a diffeomorphism of S^K onto itself. If each $c_{VU}(x)$ belongs to a specific group G of diffeomorphisms, e.g. if $c_{VU}(x) : S^K \to S^K$ is always a rotation, $G = SO(K)$, then we say that G is the structure group of the bundle. As usual we demand

$$c_{VU}(x) = c_{UV}(x)^{-1}$$

$$c_{UW} \circ c_{WV} \circ c_{VU} = \text{identity on } U \cap V \cap W$$



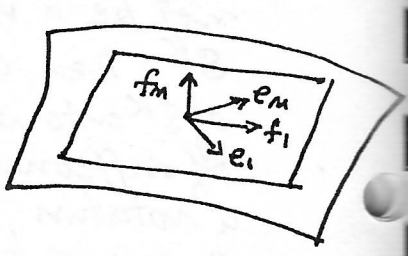
A (cross) section s is again a map $s: M^n \longrightarrow E$ such that $\pi \circ s = \text{identity}$. We can also talk about (cross) sections over a subset $A \subset M^n$.

For example, if TM^n is the tangent bundle of a Riemannian manifold, we may restrict ourselves to considering only vectors of unit length, ie the unit tangent bundle T_0M (see Lecture 6, §1). Here the fiber F is the $(n-1)$ sphere S^{n-1}.

(2) Principal Bundles.

Another example, related again to TM^n, will be of more interest to us. Let M^n be a Riemannian manifold and let FM^n be the collection of all frames $f_1, \dots, f_n$ of o.n vectors at points of M^n. $\pi: FM \longrightarrow M$ assigns to each frame $f = (f_1, \dots, f_n)$ the point of M^n at which the frame is located. What is the fiber $\pi^{-1}(x)$ over x? Let $e = (e_1, \dots, e_n)$ be a given frame at x. Then the most general frame f at x is given by

$$f_i = \sum_j e_j g^j{}_i$$



where the matrix $(g^j{}_i) = g \in O(n)$. We write:

$$f = e\,g \qquad g \in O(n)$$

Thus <u>after a single frame e at x has been chosen</u>, the fiber

$\pi^{-1}(x)$ of all frames at x is in 1:1 correspondence with the group $G = O(n)$ of orthogonal $n \times n$ matrices. The fiber F for the frame bundle $FM = E$ is the Lie group $O(n)$

$$F = O(n).$$

How do we exhibit the local trivialization (product structure). Suppose we can find a differentiable frame $e_U = e_1(x), \ldots, e_n(x)$ field on the open $U \subset M^n$ (e.g. apply Gram-Schmidt to a coordinate frame). Then a general frame on U can be written

$$f = e_U(x)\, g_U(x) \tag{61.1}$$

the local trivialization

$$\Phi_U : U \times O(n) \longrightarrow \pi^{-1}(U)$$

assigns to each $x \in U$ and each $g \in O(n)$ the frame

$$\Phi_U(x, g) = e_U(x)\, g.$$

In an overlap, the same frame (61.1) will have another representation

$$f = e_V(x)\, g_V(x)$$

where

$$e_V(x) = e_U(x)\, c_{UV}(x), \tag{61.2}$$

$c_{UV}(x) \in O(n)$ being the transition matrix for the tangent bundle (this assures us that for a vector $y = e_U y_U = e_V y_V$, then $y_V = c_{VU}\, y_U$). Thus

$$f = e_U(x) g_U(x) = e_V(x) g_V(x) = e_U(x)\, c_{UV}(x)\, g_V(x)$$

gives

$$g_U(x) = c_{UV}(x)\, g_V(x)$$

Thus the diffeomorphism

$$c_{UV}(x) : O(n) \longrightarrow O(n)$$

is simply <u>left translation</u> of $G = O(n)$ by the orthogonal matrix $c_{UV}(x)$!

In a sense, the local "coordinates" of the frame f in U are the matrices $g_U(x) = g_U{}^i{}_j(x)$.

We say that a fiber bundle $\{E, M, \pi, F, G\}$ is a <u>principal</u> bundle, if the fiber F is the same as the group G and if the transition matrices $c_{UV}(x)$ act on G simply by <u>left</u> translation.

Note that with the tangent bundle we could have used general frames instead of o.n. frames. This frame bundle would then have $F = G = O(n, \mathbb{R})$ instead (this makes sense even if there is no Riemannian metric given on M). A Riemannian metric allows us to reduce the structure group to $O(n)$

In any case, we say that <u>the frame bundle FM is the principal fiber bundle associated to the vector bundle TM.</u>

By exactly the same procedure as above, given any <u>vector</u> bundle $E \to M$, we can, by considering frames of sections, get an associated principal bundle to E.

The frame bundle FM has a remarkable property that is not shared by the tangent bundle; the structure group G acts in a natural way as a group of transformations on $E = FM$. Let $g \in G$ be a given matrix and let $f = (f_1, \dots, f_n)$ be a frame at the point $x \in M^m$. Then we can send this frame f into a new frame, written fg by the usual

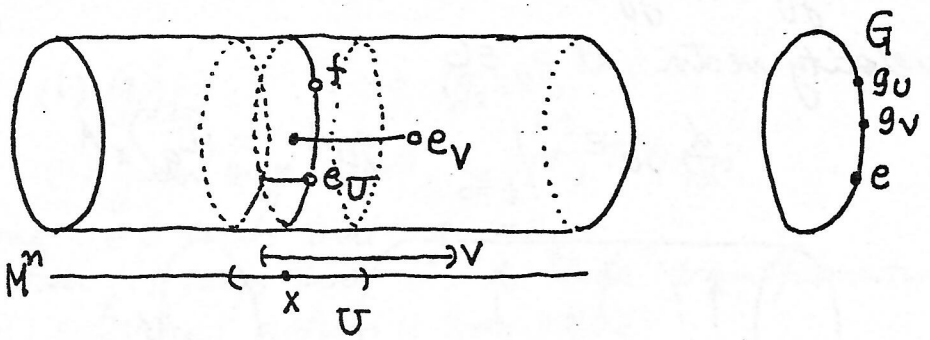
$$(f \cdot g)_i = \sum_j f_j \, g^j{}_i$$

Note that this definition is intrinsic, we have not used the local product structure at all. You should convince yourselves that there is no natural way to let $g \in G$ act on the tangent bundle, one can't define $g(v)$ for v a tangent vector, intrinsically. This property holds for <u>all</u> principal bundles.

<u>Proposition</u>. The structure group G of a <u>principal</u> bundle E acts "on the right" on E without fixed points, and preserving fibers.

$$\pi(fg) = \pi f$$

Proof. Consider $f \in E$, and $g \in G$. Let $\pi(f) \in U$ and let $\Phi_U : U \times G \to \pi^{-1}(U)$ be a local trivialization



Let e_U be the section over U defined by $e_U(x) = \Phi(x, e)$. Then of course we can write

$$f = e_U(x)\, g_U$$

for a unique $g_U \in G$. We then let $g \in G$ act on f by

$$f \cdot g = e_U(x)\,(g_U \cdot g)$$

We claim that this action is actually globally defined; for suppose the same f lies over V. By same procedure

$$f = e_V(x)\, g_V \quad \text{where } e_V(x) = e_U\, c_{UV}.$$

Then note

$$e_V(x)(g_V \cdot g) = e_U(x)\, c_{UV}(x)\, g_V \cdot g$$
$$= e_U(x)\, g_U(x) \cdot g$$

since $g_U = c_{UV}\, g_V$! □

Each right action $f \to f \cdot g$ is a diffeomorphism $E \to E$. Let the 1-parameter subgroup e^{tA}, $A \in \mathfrak{g}$ act. Its velocity vector field on E is then

$$A^*_{(f)} \equiv \frac{d}{dt} f e^{tA}\Big|_{t=0}$$

In terms of the local product structure, $f = e_U g_U$, and then

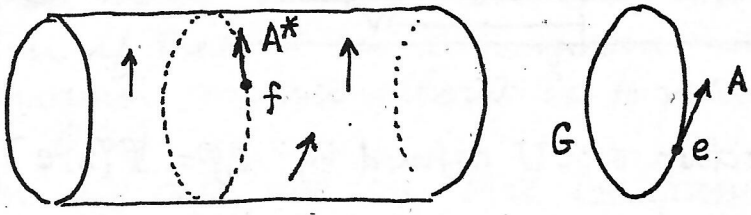
$$A^* = \frac{d}{dt} e_U g_U e^{tA}\Big|_{t=0}$$

the given action $f \to f e^{tA}$ is completely described by the action in G

$$g_0 \longrightarrow g_0 e^{tA}$$

with velocity vector, at $g_0 \in G$

$$\frac{d}{dt} g_0 e^{tA}\Big|_{t=0} = g_0 A = (L_{g_0})_* A$$

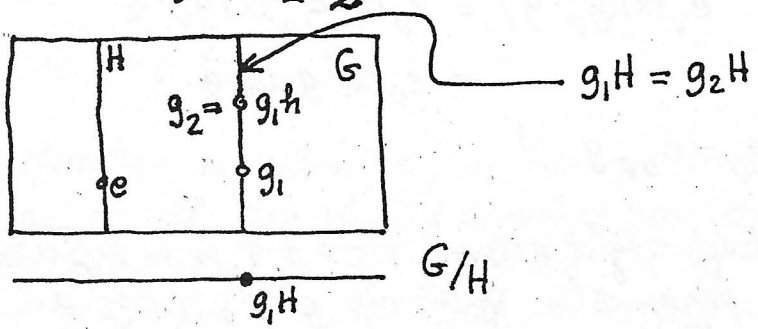


The vector field A^* on E generated by $A \in \mathfrak{g}$ is said to the the "fundamental vector field" associated with A.

(3) Coset Spaces. Let G be a Lie group and let $H \subset G$ be a subgroup. The "left coset space" is the set of equivalence classes of elements of G

$$g_1 \sim g_2 \quad \text{iff} \quad g_2 = g_1 h, \text{ some } h \in H$$

$$\text{i.e} \qquad \text{iff} \quad g_1^{-1} g_2 \in H$$



Thus we identify any and all elements of G that lie on the same left translate of the subgroup H. We denote the equivalence class of $g \in G$ by the notation $[g]$ or else gH.

Many familiar spaces are in fact such coset spaces! Let us say that a group G acts transitively as a group of transformations of a space M provided we have a map

$$G \times M \longrightarrow M$$
$$(g, x) \longmapsto g \cdot x$$

such that

(i) $(g_1 g_2) \cdot x = g_1 \cdot (g_2 x)$

(ii) given any $x_1 \in M$, $x_2 \in M$, there is at least one $g \in G$ such that $g x_1 = x_2$.

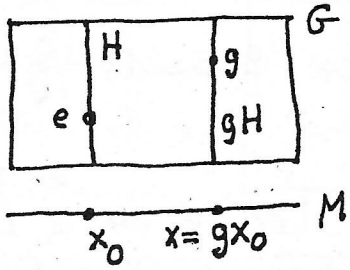
Condition (i) says that G acts as a transformation group, condition (ii) says that action is transitive.

Examples $SO(3)$ acts as a transitive transformation group on the 2-sphere

$$SO(3) \times S^2 \longrightarrow S^2$$

Fundamental Principle. Let G act transitively on the set M. Let $x_0 \in M$ and let $H \subset G$ be the subgroup leaving x_0 fixed (H is called the "stability group" or the "little group" of x_0). Then the points of M are in 1:1 correspondence with the left cosets $\{gH\}$ of G/H

Proof



Let $x \in M$. Since action is transitive $\exists$ at least one g such that $g x_0 = x$. Suppose that $g_2 x_0 = x$ also. Then

$$g^{-1} g_2 x_0 = g^{-1} x = x_0 \quad \text{(why)}$$

and so $g^{-1} g_2 \in H$. Thus to $x \in M$ we may associate the coset gH of G/H. Likewise, to the coset gH we associate the point gx_0

$$gH\, x_0 = g x_0 \quad \text{since } H x_0 = x_0. \qquad \square$$

Example. $SO(3)$ acts transitively on the 2-sphere $M = S^2$.
Let x_0 be the north pole of S^2

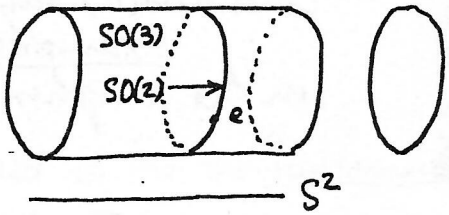
$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The little group of x_0 is clearly the 1-parameter subgroup of rotations about the z axis

$$H = SO(2) \subset SO(3) \qquad SO(2) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for all θ. We conclude that

$$SO(3)/SO(2) \approx S^2$$



Note that in any coset decomposition $G \longrightarrow G/H$ that the subgroup H acts <u>from the right</u> as a group of transformations of G that sends each coset into itself

$h \in H$ sends $g \in G$ into gh;

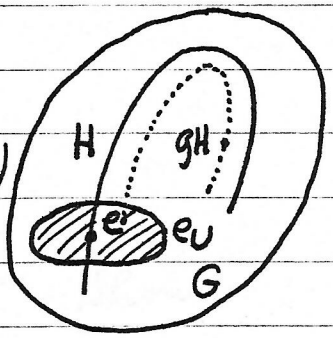
and so h sends gH into gH.

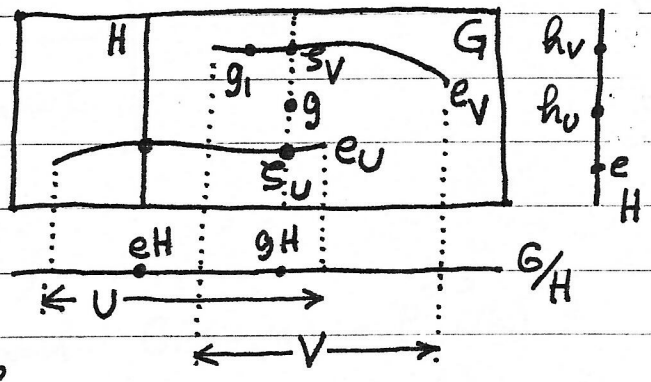
It is a fact that

<u>Theorem</u>. Let G be a Lie group and H a <u>closed</u> subgroup. Then G/H can be made into a manifold of dimension = dim G − dim H. Furthermore G is a principal bundle with structure group H and base space G/H

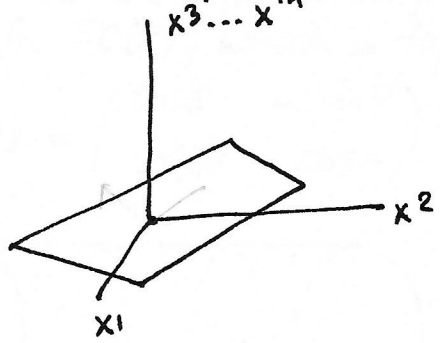
<u>Problem 99</u>. Exhibit the additive group $\mathbb{R}$ of real numbers as a principle bundle over the circle S^1.

(4) Remark on Cosets Spaces. Our theorem claims that G/H can be made into a manifold and $\pi: G \to G/H \;;\; g \to gH$ makes G a principal bundle over G/H with fiber = group = H. To show all this one first needs to find a small piece of submanifold e_U of G that is transverse to H at e (and of complementary dimension to H) and is so small that every left coset gH that is close to H will meet e_U once and only once. (This will not be possible if, e.g. H winds densely in G, but this is forbidden by insisting that H be (topologically) closed in G). The cosets that meet e_U are then in 1:1 correspondence with the points of e_U and the points of G that are "near" H are in 1:1 correspondence with the points of $e_U \times H$. By left translating e_U by the point $g_1 \in G$ we get that the points of G near $g_1 H$ are in 1:1 correspondence with the points of $e_V \times H$, where $e_V = g_1 e_U$. Whenever a coset gH meets e_U, say at the point ξ_U we shall assign to g a "base" point gH and a "fiber" point $h_U \in H$, where h_U is that unique element such that $g = \xi_U h_U$. In this way the point g can be described by the point $(gH, h_U) \in G/H \times H$. If gH also meets the transversal e_V, say at the point ξ_V, then we can also assign to g the point $(gH, h_V) \in G/H \times H$. Certainly $g = \xi_U h_U = \xi_V h_V \Rightarrow h_V = \xi_V^{-1} \xi_U \cdot h_U$ where $c_{VU} \equiv \xi_V^{-1} \xi_U \in H$. This exhibits the transition function for a principal bundle over G/H, as desired.





• **Problem 100.** The real projective $(n-1)$ space $P^{n-1}(\mathbb{R})$ is the space of all un-oriented lines thru the origin of $\mathbb{R}^n$. The Grassmann manifold $Gr(2,n)$ is the space of all un-oriented 2-planes thru the origin of $\mathbb{R}^n$



Let $O(n)$ be the orthogonal group of $\mathbb{R}^n$. (i) Show that

$$Gr(2,n) \simeq \frac{O(n)}{O(2) \times O(n-2)}$$

$$\left[\begin{array}{c|c} O(2) & 0 \\ \hline 0 & O(n-2) \end{array}\right]$$

(ii) Generalize to $Gr(p,n)$

(iii) What does this say about $Gr(1,n) = P^{n-1}(\mathbb{R})$?

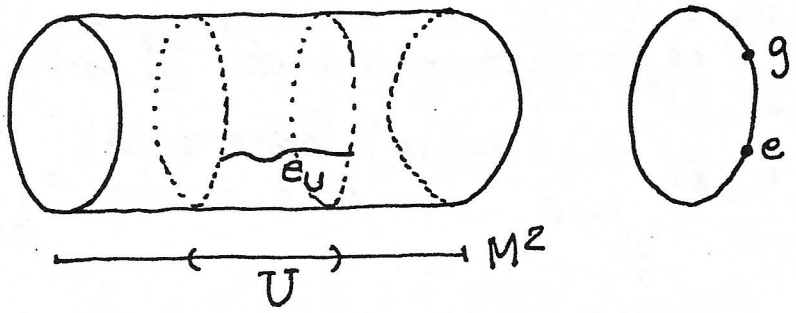
(iv) What is the dimension of $Gr(p,n)$? $\binom{n}{p}$

Lecture 62. Chern's Proof of Gauss-Bonnet-Poincaré Theorem

(1) **A Connection in the Frame Bundle of a Surface.** Let M^2 be an orientable surface with a Riemannian metric. Its tangent bundle TM^2 is an $\mathbb{R}^2$-bundle whose structure group G can be considered $G = SO(2)$. Of course we have the Riemannian (= Levi-Civita) connection in TM^2

Let FM^2 be the associated frame bundle. The fiber and group are both $SO(2)$, which is topologically a circle

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{62.1}$$



While we have defined the notion of a connection in a vector bundle we have not defined a connection in a fiber bundle.

We are now going to define a connection in the principal bundle FM^2.

Let $e_U = (e_1^U, e_2^U)$ be a frame of positively oriented vector fields on U, yielding thereby a cross section of FM^2 over U. From the Riemannian connection in the tangent bundle (= the Riemannian connection "on M^2") we have

$$\nabla e_U = e_U \otimes \omega_U$$

i.e

$$\nabla e_i = e_j \otimes \omega^j{}_i = e_j \otimes \omega_{ji}$$

and recall that $\omega_{ji} = -\omega_{ij}$. This condition of skew symmetry of the connection forms (with respect to o.n frames) arose because we demanded that the metric tensor be covariant constant (see Lecture 38, §4). This is easily equivalent to the demand that parallel displacement sends o.n frames into o.n frames. In other words, let e_U, as above, be a section of our (principal) frame bundle and let $x(t)$ be a curve in M starting at $x(0)$

Parallel translate $e_U(0)$ along $x(t)$, yielding a frame $\tilde{e}(t)$. Since $\tilde{e}(t)$ is assumed o.n, we have, along $x(t)$

$$\tilde{e}(t) = e_U \cdot g(t)$$

for $g(t) \in SO(2)$, i.e. $g(t)$ is in our structure group. But then

$$0 = \frac{\nabla \tilde{e}}{dt} = \frac{\nabla e}{dt} g + e \frac{dg}{dt} = e\, \omega\left(\frac{dx}{dt}\right) g + e \frac{dg}{dt}$$

and so

$$\omega\left(\frac{dx}{dt}\right) = \frac{dg}{dt} g^{-1} \tag{62.2}$$

ie, the value of ω on the tangent vector $\frac{dx}{dt}$ is the matrix $\frac{dg}{dt} g^{-1}$. But from (62.1)

$$\frac{dg}{dt} g^{-1} = \dot{\theta} \begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 0 & -\dot{\theta} \\ \dot{\theta} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \dot{\theta} \tag{62.3}$$

ie $\frac{dg}{dt} g^{-1}$ is in the Lie algebra of $SO(2)$!, and so of course $\omega\left(\frac{dx}{dt}\right)$ is skew symmetric. (It should not surprise us that $\frac{dg}{dt} g^{-1} \in so(2)$ since $\frac{dg}{dt}$ is a tangent vector to $SO(2)$ and we are right translating it back to the identity). <u>The Riemannian condition on our $\mathbb{R}^2$-bundle connection is the demand that the connection 1-forms ω take their values in the Lie algebra of the structure group</u>

$$\omega(X) \in \mathfrak{g} = so(2) \tag{62.4}$$

for every tangent vector X to M!

Now back to the task of <u>defining a "connection" in the frame bundle</u> FM^2. Given our section e_U of the frame bundle, let f be another section over U. Then we have

$$f(x) = e_U g_U(x).$$

Then

$$\nabla f = \nabla(e_U g_U) = e_U \otimes \omega_U g_U + e_U\, dg_U$$

$$= f \otimes \left(g_U^{-1} \omega_U g_U + g_U^{-1} dg_U \right) \qquad (62.5)$$

Now ω_U is a matrix of 1-forms on $U \subset M$, and so can be considered as a matrix of 1-forms on $\pi^{-1}(U)$ (we should write $\pi^* \omega_U$!). In fact, in our case, since $SO(2)$ is commutative

$$g_U^{-1} \omega_U g_U = \omega_U = \begin{pmatrix} 0 & \omega_{12} \\ -\omega_{12} & 0 \end{pmatrix}.$$

Again, in our commutative case $g_U^{-1} dg_U = dg_U g_U^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\theta$, and so

$$\nabla f = f \otimes \begin{pmatrix} 0 & \omega_{12} - d\theta \\ -\omega_{12} + d\theta & 0 \end{pmatrix}$$

Note that

$$\tilde{\omega}_U = g_U^{-1} \omega_U g_U + g_U^{-1} dg_U = \begin{pmatrix} 0 & \omega_{12} - d\theta \\ -\omega_{12} + d\theta & 0 \end{pmatrix} \qquad (62.6)$$

must be considered a matrix of 1-forms on FM, not M, since θ is a coordinate in the fiber, not on M, ie, a frame f at $x \in M$ is described by the local coordinates

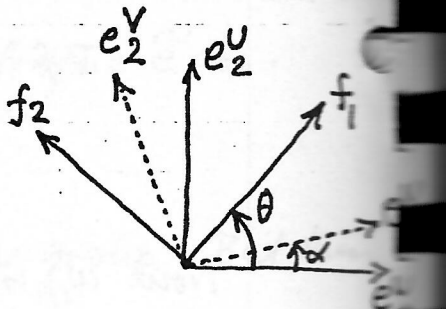
$$x^1_U,\ x^2_U,\ \theta$$

The local connection form $\tilde{\omega}_U$ on $\pi^{-1}(U)$ is locally defined by (62.6) and again has its values in the Lie algebra !

$$\tilde{\omega}_U \left(a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + a^\theta \frac{\partial}{\partial \theta} \right) \in so(2) = \mathfrak{g}$$

But now we notice something remarkable. In (62.5) we know that since ∇f makes sense independent of e_U, it should be that

Proposition: On an overlap, $\tilde{\omega}_U = \tilde{\omega}_V$, and thus the collection $\{\tilde{\omega}_U\}$ defines a $\mathfrak{g}$-valued 1-form on all of FM, $\tilde{\omega}$, the connection form for FM (62.7)



Proof. Let e_V be a section over V

$$e_V = e_U c_{UV}$$

where $c_{UV} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Then the same f has two representations, $f = e_U g_U = e_V g_V$

$$g_U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad g_V = \begin{pmatrix} \cos(\theta-\alpha) & -\sin(\theta-\alpha) \\ \sin(\theta-\alpha) & \cos(\theta-\alpha) \end{pmatrix}$$

Then in our abelian case

$$\tilde{\omega}_V = \omega_V - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d(\theta-\alpha) = \omega_U - c_{UV}^{-1} dc_{UV} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d(\theta-\alpha)$$

But $c_{UV} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow c_{UV}^{-1} dc_{UV} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} d\alpha$, and so

$$\tilde{\omega}_V = \omega_U \quad \text{on } U \cap V$$

□

Thus, although $\{\omega_U\}$ are only locally defined 1-forms on U, $\{\tilde{\omega}_U\}$ define a global 1-form, values in $\mathfrak{g} = so(2)$, on FM. In particular

$$\{\tilde{\omega}_{12}\} = \{\omega_{12} - d\theta\} = \{\pi^*\omega_{12} - d\theta\}.$$

is globally defined on FM! In particular

$$\{\tilde{\vartheta}_{12}\} = \{d\tilde{\omega}_{12}\} = \{\pi^*\vartheta_{12}\}$$

is a globally defined 2-form, values in $\mathfrak{g}$, on all of FM. This is the <u>curvature</u> 2-form on FM. Thus

<u>Proposition</u> The lift $\pi^*\vartheta_{12} = \pi^*(K\sigma^1\wedge\sigma^2)$ of the curvature 2-form on M is <u>globally exact</u> on FM

$$\pi^*(K\sigma^1\wedge\sigma^2) = \{d\omega_{12}\} \qquad (62.8$$

Theorem. Let M^2 be a compact Riemannian surface and let $\tilde{W}$ be a vector field on M^2 having a finite number of zeros at $p_1, p_2, \dots, p_N$. Then

$$\underbrace{\frac{1}{2\pi}\iint_{M^2} K\,dA}_{\text{Gauss-Bonnet}} = \underbrace{\chi(M) = \sum_{k=1}^{N} j_k(W)}_{\text{Poincaré}}$$

Recall that $\chi(M)$ is the Euler characteristic (see Lecture 28) and that the right-hand equality was proven in Lecture 28. The proof we give is due to S.S. Chern and proves directly the equality of the two extreme terms. It has vast generalizations, as we shall see!

Proof (Chern). We prove the Theorem only in the case when M^2 is orientable; the general case follows rather easily from it, as we shall see later on. First remove small discs $D_1, \dots, D_N$ centered at the zeros. Then $\tilde{v} = \tilde{W}/|\tilde{W}| = f_1$ is a well defined unit vector field on $M^2 - \bigcup_k D_k$, the complement of these discs. Complete f_1 to an o.n. frame $f = (f_1, f_2)$ that is positively oriented. We then have a cross section

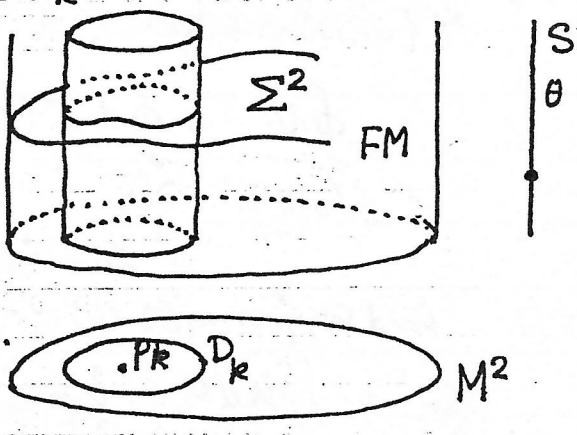
$$f : M^2 - \bigcup_{k=1}^{N} D_k \longrightarrow FM^2$$

of the frame bundle to M^2 over $M^2 - \bigcup D_k$.

Let

$$\Sigma^2 = f(M^2 - \bigcup_{1}^{N} D_k) \subset FM^2$$

be the image of $M - \bigcup D_k$ under this cross section; it is a 2 dimensional submanifold of FM that is diffeomorphic to $M - \bigcup D_k$ since $\pi \circ f = \text{identity}$. Then



$$\iint_{M-\bigcup D_k} K\,dA = \iint_{\pi\Sigma^2} K\,dA = \iint_{\Sigma^2} \pi^*(K\,dA) = \iint_{\Sigma^2} d\tilde{\omega}_{12}$$

$$= \int_{\partial\Sigma^2} \tilde{\omega}_{12} = \int_{\partial\Sigma^2} \pi^*\omega_{12} - d\theta \qquad (62.9)$$

In the above calculation one may use any representation for $\pi^* \omega_{12} - d\theta$, i.e, any of the open sets $U, V, \ldots$ in a local trivialization of the frame bundle, since they fit together to give a global 1-form.

Now let each disc D_k shrink down to their centers p_k. Now in local coordinates x^1_U, x^2_U near p_k we have

$$\omega_{12} = \Gamma^1_{\alpha 2}(x)\, dx^\alpha$$

Thus the part of $\partial \Sigma^2$ that lies over ∂D_k has vanishingly small x extent, and so in the limit as D_k shrunks to p_k

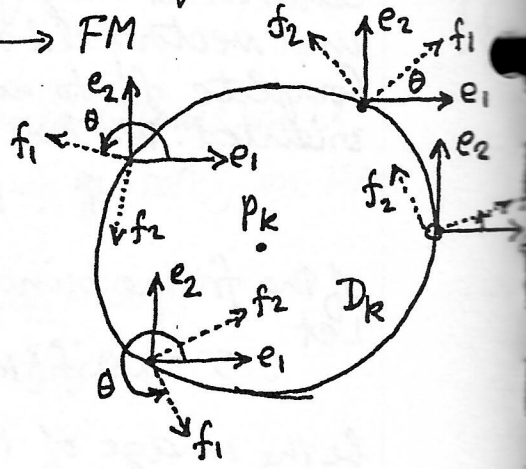
$$\int_{\partial \Sigma} \omega^1_2 \longrightarrow 0.$$

and therefore

$$\lim_{D_k \searrow p_k} \iint_{M - U D_k} K\, dA = - \lim \int_{\partial \Sigma^2} d\theta \qquad (62.10)$$

Consider the part of $\partial \Sigma^2$ that lies over ∂D_k. In a region that includes D_k we may take a cross section $e : D_k \longrightarrow FM$

and then $(f_1\ f_2) = (e_1\ e_2)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.



Note that the part of $\partial \Sigma^2$ that lies over ∂D_k has orientation opposite to ∂D_k. Furthermore

$$\oint_{f(\partial D_k)} d\theta = \oint_{\partial D_k} d \measuredangle (e_1, f_1)$$

$$= 2\pi \cdot \text{index of } v \text{ at } p_k = 2\pi j_k(w)$$

And so, from (62.10)

$$\iint_{M^2} K\, dA = \lim_{D_k \searrow p_k} \iint_{M - U D_k} K\, dA = \sum_{k=1}^{N} \lim - \oint_{-f(\partial D_k)} d\theta$$

$$= \sum_{k=1}^{N} 2\pi j_k(w)$$

□

Corollary: If M^2 is a closed Riemannian surface Then (62.11)

$$\frac{1}{2\pi} \iint_{M^2} K\, dA \text{ is an integer}$$

Lecture 63. Lie Algebra Valued Forms

Let M^n be a manifold, G a Lie group and $\mathfrak{g}$ its Lie algebra. We shall consider vector valued differential forms φ on M^n whose values lie in the fixed vector space $\mathfrak{g}$. We have in mind, e.g., the connection form $\tilde{\omega}$ and the curvature form $\tilde{\Theta}$ in the frame bundle FM^2 to a surface. A more easily understood example is the Maurer-Cartan 1-form Ω on the Lie group G, defined as follows. Let $E_1, \dots, E_N$ be a fixed basis of $\mathfrak{g}$, let $X_1, \dots, X_N$ be the left invariant vector fields on G obtained by left translating the X's over the group, and finally let $\sigma^1, \dots, \sigma^N$ be the dual left invariant 1-forms. Then

$$\Omega = \sum_R E_R \otimes \sigma^R \tag{63.1}$$

takes a tangent vector Y to G at g and left translates Y back to the identity!

Classically this would be written differently. On any manifold, Cartan wrote dp for the vector valued 1-form at $p \in M$ that takes each vector Y at p and gives back the same vector. In coordinates

$$dp = \frac{\partial}{\partial x^i} \otimes dx^i.$$

Then Cartan would write, instead of Ω

$$\Omega = g^{-1} dg.$$

Thus dg takes Y at g, gives Y back at g, and then g^{-1} left translates it back to the identity. We should write

$$\Omega = (L_{g^{-1}})_* \circ dg.$$

For a matrix group, each $E_R \in \mathfrak{g}$ is simply a matrix of a special type and so Ω is a matrix with

left invariant 1-form coefficients

Proposition. In any matrix group

$$\Omega = g^{-1}dg \qquad (63.2)$$

is a matrix each of whose entries is a left invariant 1-form.

Proof. Let $g_0 \in G$ be a fixed matrix. Then

$$\Omega_{g_0 g} = (g_0 g)^{-1} d(g_0 g) = g^{-1} g_0^{-1} g_0 \, dg = g^{-1} dg = \Omega_g .$$

For example, in $G = SO(2)$

$$g = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and then

$$g^{-1}dg = \begin{pmatrix} \cos\theta & +\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -\sin\theta\, d\theta & -\cos\theta\, d\theta \\ \cos\theta\, d\theta & -\sin\theta\, d\theta \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -d\theta \\ d\theta & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \otimes d\theta$$

Similarly in a matrix group G

$$dg\, g^{-1}$$

is a matrix of right invariant 1-forms

Problem 101. Exhibit the left invariant and right invariant 1-forms on the affine group of the line, $g = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$ by means of $g^{-1}dg$ and $dg\, g^{-1}$.

We shall need to develope some machinery for handling $\mathfrak{g}$-valued p-forms on a manifold M^n. The most general $\mathfrak{g}$-valued p-form φ on M^n is of the form

$$\varphi = \sum_{R=1}^{N} E_R \otimes \varphi^R$$

where each φ^R is a p-form on M^n; note that R indicates which E_R is involved, not the degree of the form φ

Since the E_R's are not varying (lying in the fixed vector space $\mathfrak{g}$), it is natural to define

$$d\varphi = d\left(E_R \otimes \varphi^R \right) \equiv E_R \otimes d\varphi^R \qquad (63.3)$$

a $\mathfrak{g}$-valued $(p+1)$ form on M^n.

If φ is a $\mathfrak{g}$-valued p-form and ψ is a $\mathfrak{g}$-valued q form, then, as we shall see, $\varphi \wedge \psi$ cannot always be defined to be a $\mathfrak{g}$ valued $(p+q)$ form; as matrices with form entries $\varphi \wedge \psi$ will be a matrix with form entries but there is no reason to expect that, e.g. if φ and ψ are skew symmetric matrices that $\varphi \wedge \psi$ will also be skew symmetric! Rather than define $\varphi \wedge \psi$ in general (sometimes it will make sense) we shall define a generalized bracket

$$[\varphi, \psi] = \left[E_R \otimes \varphi^R,\ E_S \otimes \psi^S \right] \equiv [E_R, E_S] \otimes \varphi^R \wedge \psi^S \qquad (63.4)$$

For example, consider the Maurer–Cartan 1-form Ω on a Lie group

$$\Omega = E_R \otimes \sigma^R.$$

From the "Maurer-Cartan" equations (56.2)

$$d\Omega = E_R \otimes \left(-\tfrac{1}{2} C^R_{ST}\, \sigma^S \wedge \sigma^T \right)$$

while

$$\begin{aligned} [\Omega, \Omega] &= \left[E_S \otimes \sigma^S,\ E_T \otimes \sigma^T \right] \\ &= [E_S, E_T] \otimes \sigma^S \wedge \sigma^T \\ &= E_R \otimes C^R_{ST}\, \sigma^S \wedge \sigma^T \end{aligned}$$

and so

$$\boxed{d\Omega + \tfrac{1}{2}[\Omega, \Omega] = 0} \qquad (63.5)$$

Maurer-Cartan

Remark: (62.4) defines $[\varphi,\psi]$ by means of a basis but it is not difficult to give an intrinsic definition. If $X_I = X_{i_1},\dots,X_{i_{p+q}}$ are tangent vectors to M^n we could have defined $[\varphi,\psi]$ by

$$[\varphi,\psi](X_I) = \delta^{J,K}_{I}\,[\varphi(X_J),\psi(X_K)] \tag{63-}$$

$$= \delta^{J,K}_{I}\,[E_R\otimes\varphi^R(X_J),\ E_S\otimes\psi^S(X_K)]$$

$$= \delta^{J,K}_{I}\,[E_R,E_S]\ \varphi^R(X_J)\,\psi^S(X_K)$$

$$= [E_R,E_S]\ \varphi^R\wedge\psi^S\,(X_I) \qquad \square$$

We have some immediate consequences of our definitions

$$[\psi,\varphi] = [E_S,E_R]\otimes\psi^S\wedge\varphi^R$$

$$= -[E_R,E_S]\otimes\psi^S\wedge\varphi^R$$

$$= (-1)^{pq+1}\,[E_R,E_S]\otimes\varphi^R\wedge\psi^S$$

$$[\psi,\varphi] = (-1)^{pq+1}\,[\varphi,\psi] \tag{63-7}$$

Also

$$d[\varphi,\psi] = [E_R,E_S]\otimes(d\varphi^R\wedge\psi^S + (-1)^p\varphi^R\wedge d\psi^S)$$

$$d[\varphi,\psi] = [d\varphi,\psi] + (-1)^p\,[\varphi,d\psi] \tag{63-8}$$

Finally, we need to interpret $[\varphi,\psi]$ in the case of matrices when G is a matrix group. For example, the Maurer Cartan 1-form for the affine group of the line is

$$\begin{bmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\otimes\frac{dx}{x} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\otimes\frac{dy}{x}$$

$$= E_1\otimes\frac{dx}{x} + E_2\otimes\frac{dy}{x}$$

What does $[\varphi,\psi]$ say about these matrices with form entries?
In general, when (E_R) are matrices

$$[\varphi,\psi] = [E_R, E_S]\otimes\varphi^R\wedge\psi^S$$
$$= (E_R E_S - E_S E_R)\otimes\varphi^R\wedge\psi^S$$
$$= (E_R\otimes\varphi^R)\wedge(E_S\otimes\psi^S) - E_S E_R\otimes\varphi^R\wedge\psi^S$$

where in $(E_R\otimes\varphi^R)\wedge(E_S\otimes\psi^S)$ we are simply multiplying the matrices but using the exterior product of the entries.
For example $g^{-1}dg\wedge g^{-1}dg =$

$$\begin{bmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{bmatrix}\wedge\begin{bmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{dx\wedge dy}{x^2} \\ 0 & 0 \end{bmatrix} = E_2\otimes\frac{dx\wedge dy}{x^2}$$

as we always have done when discussing $\vartheta = d\omega+\omega\wedge\omega$, etc.
Continuing

$$[\varphi,\psi] = (E_R\otimes\varphi^R)\wedge(E_S\otimes\psi^S) - E_S E_R\otimes(-1)^{pq}\psi^S\wedge\varphi^R$$
$$= (E_R\otimes\varphi^R)\wedge(E_S\otimes\psi^S) - (-1)^{pq}(E_S\otimes\psi^S)\wedge E_R\otimes\varphi^R.$$

$$\boxed{[\varphi,\psi] = \varphi\wedge\psi - (-1)^{pq}\psi\wedge\varphi \quad \text{as matrices}} \tag{63.9}$$

For example, if p is odd $[\varphi,\varphi] = \varphi\wedge\varphi+\varphi\wedge\varphi$, i.e

$$\tfrac{1}{2}[\varphi,\varphi] = \varphi\wedge\varphi \quad \text{if } p \text{ is odd} \tag{63.10}.$$

Consider, for example a Riemannian manifold M^n with its Riemannian connection forms $\omega = (\omega^i_j)$.
We may restrict ourselves to o.n frames, and thus the structure group of the tangent vector bundle is $O(n)$, with Lie algebra

$$\mathfrak{g} = so(n) = \text{skew symmetric } n\times n \text{ matrices.}$$

Just as in lecture 61, since parallel translation of an o.n frame yields an o.n frame, it must be that the connection form ω_{ij} must be an $so(n)$-valued 1-form, i.e

$$\omega_{ij} = -\omega_{ji}.$$

To see this we again let $\tilde{e}(t)$ be a frame that results from parallel translating e_j along a curve $x(t)$ in M^n;

$$\tilde{e}(t) = e_j(x(t))\, g_j(t)$$

where $g_j(t) \in O(n)$ and $g_j(0) = e$. Then, as in (62.2)

$$\omega\left(\frac{dx}{dt}\right) = \frac{dg}{dt}\, g^{-1}$$

is the result of right translating $\frac{dg}{dt}$ at $g \in O(n)$ back to the identity, ie, back into $\mathcal{G} = so(n)$.

Thus ω_{ij} is a skew symmetric matrix of forms, and then of course so is $d\omega_{ij}$. But why is $\omega \wedge \omega$ also skew symmetric? This follows directly from (63.10)

$$\omega \wedge \omega = \tfrac{1}{2}[\omega, \omega]$$

since $[\omega, \omega] \in so(n)$. Thus

$$\Omega = d\omega + \omega \wedge \omega = d\omega + \tfrac{1}{2}[\omega, \omega] \qquad (63.11)$$

is $so(n)$-valued, ie skew symmetric

In Riemannian geometry we also dealt with the 2nd Bianchi identities (44.9)

$$d\Omega + \omega \wedge \Omega - \Omega \wedge \omega = 0.$$

This again makes sense in our Lie algebra $so(n)$ because from (63.9)

$$d\Omega + [\omega, \Omega] = d\Omega + \omega \wedge \Omega - (-1)^{1\cdot 2}\, \Omega \wedge \omega \qquad (63.12)$$

Lecture 64. Connections in a Principal Bundle

While it is possible to define the notion of a connection in any principal bundle we shall only be concerned with the principal bundle associated with some given vector bundle E over M^n (i.e., the bundle of frames of sections of E).

Let E be any vector bundle over M^n with fiber $\mathbb{R}^K$ or $\mathbb{C}^K$, and with structure group G; thus the transition matrices $c_{UV}(x) \in G$ are linear transformations of $\mathbb{R}^K$ or $\mathbb{C}^K$ into itself.

Let $\{\omega_U\}$ be a connection for this bundle. Thus if e_U is a frame of sections of E over U, then

$$De_U = e_U \otimes \omega_U$$

where $\omega_U{}^i_j = \Gamma^i_{\mu j}(x)\, dx^\mu$ is a matrix of 1-forms on U. As in Lecture 63 we assume that if X is any tangent vector to M in U, that $\omega_U(X) \in \mathfrak{g}$. Thus ω_U is a $\mathfrak{g}$-valued 1-form on U; from Lecture 63 it is of the form

$$\omega_U = \sum_{R=1}^{N} E_R \otimes \omega^R$$

where $N = \dim G = \dim \mathfrak{g}$. For example, when E is TM^2, the tangent bundle to a surface with a Riemannian connection, with structure group $G = SO(2)$

$$\omega_U = \begin{bmatrix} 0 & \omega_{12} \\ -\omega_{12} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \omega_{12}$$

and of course $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in so(2)$.

We also have the local curvature forms

$$\Omega_U = d\omega_U + \omega_U \wedge \omega_U$$
$$= d\omega_U + \tfrac{1}{2}[\omega_U, \omega_U]$$

Both forms ω_U and ϑ_U are locally defined on M^n; a change of local frames

$$e_V = e_U\, c_{UV} \tag{64.1}$$

yields

$$\left.\begin{aligned} \omega_V &= c_{UV}^{-1}\,\omega_U\, c_{UV} + c_{UV}^{-1}\, dc_{UV} \\ \text{and}\quad \vartheta_V &= c_{UV}^{-1}\,\vartheta_U\, c_{UV} \end{aligned}\right\} \tag{64.2}$$

(the transformation law for ϑ_V was worked out in §8.26) in the case of the tangent bundle $E = TM^n$).

Consider now the principal bundle P associated to the given vector bundle E, i.e, the bundle of frames of sections of E.

$$\pi : P \longrightarrow M.$$

This is the bundle with fiber $F = G$ (the structure group of E) and where the transition functions are <u>the very same matrices</u> $c_{UV}(x)$ but now acting on the matrices of G by matrix multiplication from the left

$$g \in G \text{ is sent to } c_{UV}(x)\cdot g \in G.$$

We now define connection forms $\tilde{\omega}$ in P (i.e, they will be $\mathfrak{g}$-valued 1-forms defined on P rather than UCM) by analogy with the case of the frame bundle associated to the tangent bundle to a surface (Lecture 62). For a for a point $f \in \pi^{-1}(U)$ (i.e. a frame of sections of E over U)

$$f = e_U g_U\,, \quad g_U \in G$$

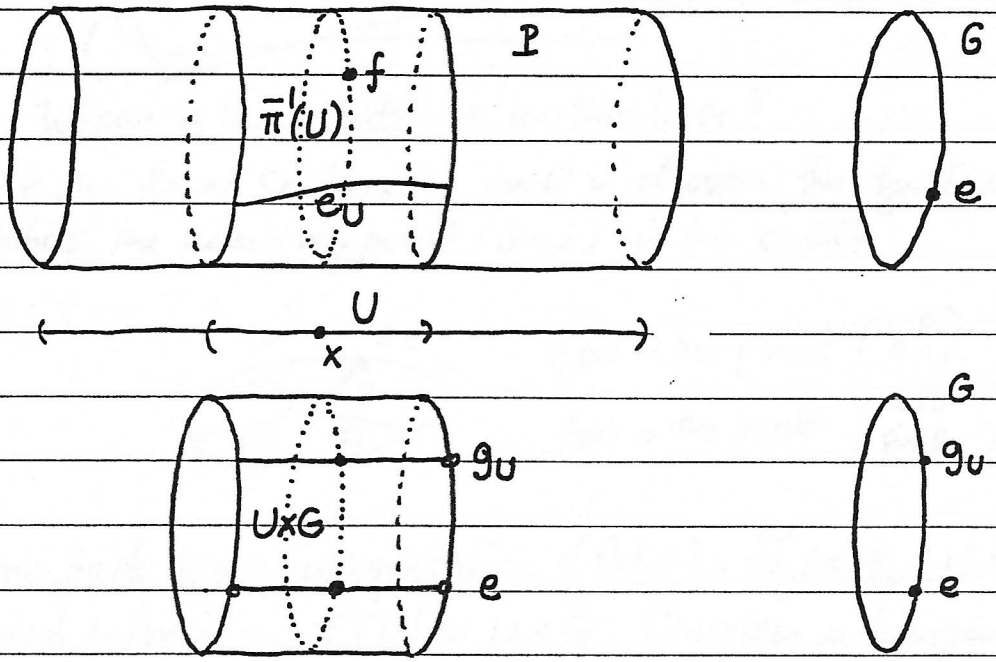
we compute as before

$$Df = f \otimes (g_U^{-1}\,\pi^*\omega_U\, g_U + g_U^{-1} dg_U)$$

and then define $\tilde{\omega}_U(x, g_U)$, ie $\tilde{\omega}_U$ at the point f, by

$$\tilde{\omega}_U(x, g_U) = g_U^{-1} \pi^* \omega_U(x)\, g_U + g_U^{-1} dg_U \qquad (64.3)$$

and we shall usually write ω_U rather than $\pi^* \omega_U$ since ω_U involves only x and dx and x can also be used as partial coordinates in $\pi^{-1}(U)$. (64.3) is interpreted as follows. The local section e_U of P allows us, as usual, to picture $\pi^{-1}(U)$ as $U \times G$



The map

$$\Phi_U : U \times G \longrightarrow \pi^{-1}(U)$$

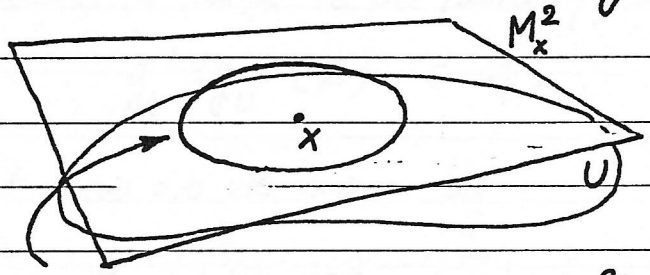
sends $(x, g_U) \in U \times G$ to the frame $e_U(x) \cdot g_U$. In particular the fiber over x, $\pi^{-1}(x)$ is simply a copy of the group G, call it G_x, where

$e \in G$ is sent into the "point" $e_U(x)$

$g_U \in G$ is sent into the "point" $e_U(x) \cdot g_U$

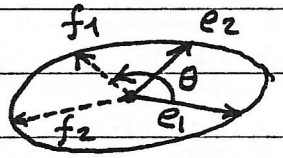
The frame $f = e_U g_U$ is identified with the point $g_U \in G$. Thus the local trivialization of P given by e_U allows us to see a <u>specific copy</u> G_x of the group G over each $x \in U$, whereas before

the local frame e_U was chosen we had a diffeomorphic copy of G over x but we couldn't see the group structure. For example, in the case FM^2 over the oriented surface M^2, each fiber is a copy of $SO(2)$, ie a circle S^1, but we couldn't even see where the identity was



Where is the identity e on this fiber?

but after the local section e_U on U is chosen the first vector e_{U1} exhibits the identity point ($\theta = 0$) in the circle

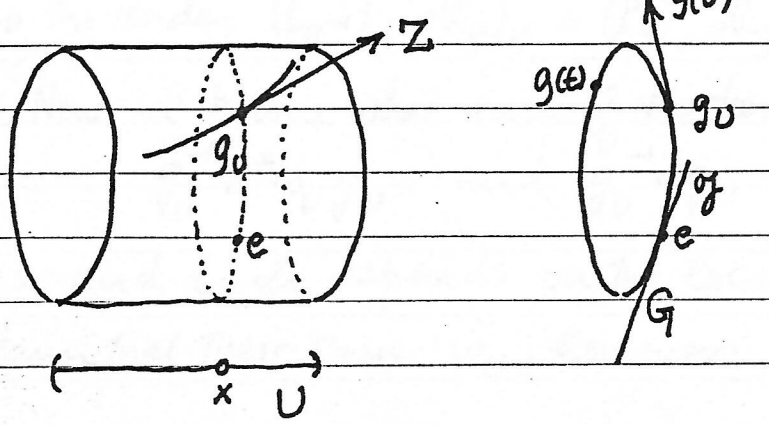


$e_U(x)$ is the point $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$f(x)$ is the point $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Now back to our interpretation of (64.3). $\tilde{\omega}_U(x, g_U)$ is to be a $\mathfrak{g}$-valued 1-form on $\pi^{-1}(U) \approx U \times G$. Consider a tangent vector Z to P at the point $f = e_U(x) g_U$. This is the velocity vector to some curve

$$x = x(t), \quad g = g(t)$$
$$x(0) = x \qquad g(0) = g_U$$



Now by $g_U^{-1} dg_U (Z)$ we simply mean the vector

$$g^{-1}(t)\, \dot{g}(t)\Big|_{t=0}$$

ie we take $\dot{g}(0)$, a tangent vector to G at $g_U \in G$ and we left translate it back to the identity, ie into the Lie algebra $\mathfrak{g}$.

$$g_U^{-1} dg_U (Z) \in \mathfrak{g}.$$

What now do we mean by

$$g_U^{-1} (\pi^* \omega_U)\, g_U (Z) \; ?$$

Since $\omega_U(x) = (\Gamma^i_{\mu j}\, dx^\mu)$ acts only on $\partial/\partial x$

$$g_U^{-1} \pi^* \omega_U\, g_U (Z) = g_U^{-1}\, \omega_U\left(\frac{dx}{dt}\right) g_U.$$

Now $\omega_U\left(\frac{dx}{dt}\right)$ is assumed to be in $\mathfrak{g}$; we then take this vector at e, right translate it to g_U and then left translate it to e again

$$g_U^{-1} \pi^* \omega_U\, g_U (Z) = (L_{g^{-1}})_* \circ (R_g)_*\, \omega_U\left(\frac{dx}{dt}\right) \in \mathfrak{g}$$

(Note that right translations commute with left translations

$$h \nearrow g_1 h \longrightarrow g_1 h g_2 \quad (=)$$
$$h \searrow h g_2 \longrightarrow g_1 h g_2$$

and so the order $(L_{g^{-1}})_* \circ (R_g)_* = (R_g)_* \circ (L_{g^{-1}})_*$ is unimportant)

Now we know what each of the terms

$$g_U^{-1} \pi^* \omega_U\, g_U \quad \text{and} \quad g_U^{-1} dg_U$$

means and each depends on the local section e_U chosen. We claim that their sum $\tilde{\omega}_U$, however, is independent

of the choice of e_U.

Proposition. *On $\pi^{-1}(U \cap V)$ we have*

$$\tilde{\omega}_U = \tilde{\omega}_V \tag{64.4}$$

and so $\{\tilde{\omega}_U\}$ defines a global $\mathfrak{g}$-valued 1-form on the principal bundle P.

Problem 102. Prove this! At a given $f \in \pi^{-1}(U \cap V)$

$$f = e_U g_U = e_V g_V, \quad e_V = e_U c_{UV} \text{ etc.}$$

Show that the transformation rule (59.6) for the connection form $\{\omega_U\}$ for the vector bundle E implies (64.4).

We then define the curvature form $\tilde{\Theta}$ for the principal bundle P to be the global $\mathfrak{g}$-valued 2-form defined locally by

$$\tilde{\Theta}_U \equiv d\tilde{\omega}_U + \tfrac{1}{2}[\tilde{\omega}_U, \tilde{\omega}_U] \tag{64.5}$$

and we know from (63.9) that in terms of matrices with 2-form entries, we can write

$$\tilde{\Theta}_U = d\tilde{\omega}_U + \tilde{\omega}_U \wedge \tilde{\omega}_U \tag{64.6}$$

Of course we also have the curvature forms Θ_U for the original vector bundle E

$$\Theta_U = d\omega_U + \omega_U \wedge \omega_U$$

Problem 103. Show that (64.5) and (64.3) imply that

$$\left.\begin{aligned} \tilde{\Theta}(x, g_U) &= g_U^{-1} \pi^* \Theta_U g_U \\ &= g_U^{-1} \Theta_U g_U \end{aligned}\right\} \tag{64.7}$$

Lecture 65. The Yang-Mills Field

(1) The Electromagnetic Connection Revisited; Weyl's Interpretation.

We shall discuss soon the relativistic wave equation for an electron, the Dirac equation. The Lagrangian for the electron field is some expression involving the 4 component "spinor"

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}.$$

Suppose that we were unaware of the existence of a notion of a field like the electromagnetic field, and so it does not appear in the Lagrangian. The usual Lagrangian is invariant under constant phase transformations

$$\psi \longrightarrow e^{ia}\psi \tag{65.1}$$

where a = constant. Such gauge transformations are called by physicists global, since the same phase transformation is carried out at each space time point. It is also clear that if a ψ satisfies a Schrödinger or Dirac equation so will $e^{ia}\psi$. For the Dirac equation this symmetry of the Lagrangian implies, via the Noether procedure, conservation of charge.

Suppose now that one had reason to believe that local gauge transformations

$$\psi \longrightarrow e^{-\frac{ie}{\hbar}f}\psi \tag{65.2}$$

f a function, should also be a symmetry of the "correct" Lagrangian. Then, just as in Lecture 60, one would be forced to switch over to a "covariant derivative"

$$\partial/\partial x^\mu \longmapsto \partial/\partial x^\mu + \frac{ie}{\hbar}A_\mu$$

where A_μ is some <u>new field</u> (with the appropriate transformation rule

$$A_\mu \longrightarrow A_\mu + \partial_\mu f),$$

that must be introduced; our insistence on the symmetry (65.2) would force us to introduce an "electromagnetic field". The Lagrangian would also have to be modified to include "kinetic" terms for the new field A introduced. An immediate candidate involving first derivatives of this new field would be the square of the curvature ν. Since

$$\nu = d\omega = ie\, dA = ie\mathcal{F}$$

we could add a constant multiple of

$$\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}. \qquad (65.3)$$

Quantization of the field A (ie 2nd quantization) would then yield particles, <u>photons</u>; exchange of photons between electrons would then "explain" the force between electrons!

(2) <u>The Heisenberg Nucleon</u>. Heisenberg postulated that if we disregard electromagnetic effects, then the proton p and the neutron n behave very similarly during nuclear "strong" interactions. He then considered p and n as two aspects of a single particle, the <u>nucleon</u>, represented by a 2 component wavefunction.

$$\psi = \begin{pmatrix} p \\ n \end{pmatrix} \qquad (65.4)$$

(Each p and n is really a 4 component spinor; we gloss over this). Thus a particle that is definitely a

proton is given by $\begin{pmatrix} p \\ 0 \end{pmatrix}$, a neutron by $\begin{pmatrix} 0 \\ n \end{pmatrix}$, while $\begin{pmatrix} p \\ n \end{pmatrix}$ is a mixture.

Heisenberg felt that the Lagrangian for the nucleon should admit the global symmetry

$$\psi \to u\psi = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} p \\ n \end{pmatrix} \qquad (65.5)$$

where $u \in U(2)$ is a constant unitary matrix. Since

$$\begin{pmatrix} e^{i\alpha} p \\ e^{i\alpha} n \end{pmatrix} \text{ and } \begin{pmatrix} p \\ n \end{pmatrix}$$

represent the same mixture, Heisenberg eliminated consideration of this special symmetry by restricting u to have determinant $= 1$; the symmetry group consists then of global (ie constant) matrices $u \in SU(2)$, the special unitary group.

(3) The Yang-Mills Nucleon. Yang and Mills (1954) postulated that the correct Lagrangian for a nucleon will admit $SU(2)$ as a local symmetry group. Thus while originally ψ could be considered as a $\mathbb{C}^2$ valued function on Minkowski space, we should now consider ψ as a cross section of a $\mathbb{C}^2$ vector bundle over M^4 with a covariant derivative

$$\frac{\partial}{\partial x^\mu} + \omega_\mu$$

where $\omega_\mu = \omega\left(\frac{\partial}{\partial x^\mu}\right)$ and ω is an $su(2)$ valued connection form. Now $su(2)$ consists of skew-hermitian matrices with trace zero; since Physicists prefer hermitian matrices, a basis for $su(2)$ is given by the matrices

$$(i\sigma_R) \qquad R = 1, 2, 3$$

where σ_R are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{65.6}$$

Thus each ω_μ is of the form

$$\omega_\mu = iq\,\sigma_R A^R_\mu \equiv iq\,\vec{\sigma}\cdot\vec{A}_\mu . \tag{65.7}$$

Here q is some coupling constant (similar to charge for the electron field), and we have had to introduce 3 new fields A^1_μ, A^2_μ, A^3_μ, the "Yang-Mills" fields, to mediate the force between nucleons. Each A^R is a 4-vector field in M^4.

$$\begin{cases} \partial_\mu \psi^K = \dfrac{\partial \psi^K}{\partial x^\mu} + iq\,\vec{\sigma}\cdot\vec{A}_\mu\,\psi^K \qquad K=1,2. \\ \vec{\sigma}\cdot\vec{A}_\mu = A^1_\mu \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + A^2_\mu \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + A^3_\mu \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv A_\mu \end{cases} \tag{65.8}$$

One then has to introduce "kinetic" terms for the fields A_μ into the new Lagrangian, terms involving derivatives of the fields A_μ. The only object that comes to mind, immediately involving derivatives of ω is the curvature of the connection

$$\begin{aligned} \vartheta &= d\omega + \tfrac{1}{2}[\omega,\omega] \\ &= d\omega + \omega\wedge\omega \end{aligned}$$

as matrices of forms. Then

$$\vartheta_{\mu\nu} = \vartheta(\partial_\mu,\partial_\nu) = d\omega(\partial_\mu,\partial_\nu) + \omega\wedge\omega\,(\partial_\mu,\partial_\nu)$$

$$\left.\begin{aligned} \vartheta_{\mu\nu} &= \partial_\mu\omega_\nu - \partial_\nu\omega_\mu + \omega_\mu\omega_\nu - \omega_\nu\omega_\mu \\ &= \partial_\mu\omega_\nu - \partial_\nu\omega_\mu + [\omega_\mu,\omega_\nu] \end{aligned}\right\} \tag{65.9}$$

Using the matrix $A_\mu = \vec{\tau}\cdot\vec{A}_\mu$, and (65.7)

$$\vartheta_{\mu\nu} = ig(\partial_\mu A_\nu - \partial_\nu A_\mu) - g^2[A_\mu, A_\nu]$$

$$= ig\, F_{\mu\nu} \qquad (65.10)$$

where

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu) + g[A_\mu, A_\nu] \qquad (65.11)$$

is again the <u>field strength</u> (a <u>hermitian</u> 2×2 matrix valued 2-form). One adds to the Lagrangian a term proportional to

$$F_{\mu\nu}F^{\mu\nu}$$

just as in E&M. After 2nd quantization the fields A^1, A^2, A^3 yield 3 particles that mediate the nuclear force.

(4) <u>Transformation of A and F under Gauge Transformations</u>.

As usual, under changes of frames in the vector bundle, ie gauge transformations (local)

$$\psi \longrightarrow u\psi \qquad u \in SU(2)$$

we have

$$\omega' = u^{-1}\omega u + u^{-1}du$$

or

$$\omega'_\mu = u^{-1}\omega_\mu u + u^{-1}\partial_\mu u.$$

Putting $\omega_\mu = igA_\mu$

$$A'_\mu = u^{-1}A_\mu u - \frac{i}{g}u^{-1}\partial_\mu u \qquad (65.12)$$

This brings up a minor point. Many people call A a "connection" but this is not technically correct. ω is the connection, and a multiple ($\neq 1$) of a connection is <u>not</u> a connection, as is evidenced in (65.12) by the

coefficient $-\frac{i}{q}$ before $u^{-1}\partial_\mu u$!

To consider infinitesimal gauge transformations (as physicists do), write

$$u(\lambda, x) = e^{\lambda\eta(x)} \tag{65.13}$$

where $\eta(x)$ is a family of skew hermitian matrices. Put

$$\omega^u \equiv u^{-1}\omega u + u^{-1}du$$

and consider

$$\delta\omega \equiv \omega^u - \omega = u^{-1}\omega u + u^{-1}du - \omega$$
$$= e^{-\lambda\eta}\omega e^{\lambda\eta} - \omega + e^{-\lambda\eta}d(e^{\lambda\eta})$$
$$= (1-\lambda\eta+\cdots)\omega(1+\lambda\eta+\cdots) - \omega + (1-\lambda\eta-\cdots)de^{\lambda\eta}.$$

Now $d(e^{\lambda\eta}) = d(1+\lambda\eta+\cdots) = \lambda\, d\eta$, and so putting $\lambda\eta = \delta\eta$, and neglecting $\lambda^2, \lambda^3, \ldots$

$$\delta\omega \cong d(\delta\eta) + [\omega, \delta\eta], \tag{65.14}$$

ie

$$\delta A = [A, \delta\eta] - \frac{i}{q}d(\delta\eta) \tag{65.15}$$

Similarly, putting (see (64.2))

$$\vartheta^u = u^{-1}\vartheta u$$

we get $$\delta F \cong [F, \delta\eta] \tag{65.16}$$

Lecture 66. The Groups SO(3) and SU(2)

(1) SO(3), The Rotation Group of $\mathbb{R}^3$. Of course $\mathfrak{so} = so(3)$ consists of all skew symmetric 3×3 matrices. Rotations about the z axis form a 1-parameter subgroup

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \exp \theta E_3$$

for some $E_3 \in so(3)$. Then

$$E_3 = \frac{d}{d\theta} \exp \theta E_3 \Big|_{\theta=0} = \frac{d}{d\theta} R_z(\theta) \Big|_{\theta=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and likewise for E_1, E_2. Use as basis for $so(3)$

$$E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{66.1}$$

For Lie algebra, we compute

$$[E_1, E_2] = E_3, \quad [E_2, E_3] = E_1, \quad [E_3, E_1] = E_2$$

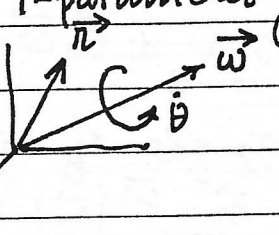
i.e.

$$[E_i, E_j] = \varepsilon_{ijk} E_k \tag{66.2}$$

and so these are the structure constants

$$c^k_{ij} = \varepsilon_{kij}.$$

Consider now a 1-parameter group of rotations about an axis, $\|\vec{\omega}\| = \dot{\theta}$



then we know that

$$\frac{d\vec{n}}{dt}\Big|_{t=0} = \vec{\omega} \times \vec{n}(0)$$

On the other hand this one parameter group is of the form

$$R(t) = e^{tS}$$

for some skew symmetric matrix S, and so

$\vec{r}(t) = R(t)\,\vec{r}(0) = e^{tS}\,\vec{r}(0)$. Then

$$\left.\frac{d\vec{r}}{dt}\right|_{t=0} = S\,\vec{r}(0)$$

and we conclude $S(\vec{r}) = \vec{\omega} \times \vec{r}$. Now as usual we may associate with the (pseudo) vector $\vec{\omega} \in \mathbb{R}^3$ a 2-form ω^2, ie a skew symmetric matrix ω_{ij}. Then

$$\left.\begin{aligned} S(r)_i &= S_{ij}\, x_j \\ \text{while} \quad (\vec{\omega} \times \vec{r})_i &= (2_{\vec{r}}\,\omega^2)_i = x_j\, \omega_{ji} \end{aligned}\right\}.$$

We conclude that $S_{ij} = -\omega_{ij}$,

$$\begin{cases} R(t) = e^{-t\omega} \\ \omega = (\omega_{ij}) \text{ is the skew matrix assoc. to } \vec{\omega} \end{cases} \tag{66.3}$$

Thus $\vec{\omega} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \leftrightarrow -E_1$, $\vec{\omega} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow -E_2$, $\vec{\omega} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow -E_3$

In other words ω, as a skew symmetric matrix has an expansion

$$\omega = -\sum E_i\, \omega^i \equiv -\vec{E} \cdot \vec{\omega}$$

and then

$$R(t) = e^{t\,\vec{E}\cdot\vec{\omega}}.$$

In terms of the angle of rotation, $\theta = t\,\dot{\theta}$, and the unit vector $\vec{n} = \vec{\omega}/|\vec{\omega}| = \vec{\omega}/\dot{\theta}$

$$\boxed{R(\theta) = e^{\theta\,\vec{E}\cdot\vec{n}}} \tag{66.4}$$

(2) SU(2). The Lie algebra su(2) consists of skew hermitian 2×2 matrices with trace 0. If we call su(2) $\mathfrak{g}$, then $i\mathfrak{g}$ is the space of hermitian 2×2 matrices with trace zero. A basis for $i\mathfrak{g}$ is given by the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{66.5}$$

Note that the Lie algebra is given by

$$[\sigma_i, \sigma_j] = 2\varepsilon_{ijk}\,\sigma_k, \tag{66.6}$$

the same as for SO(3)*. It is a theorem of Lie group theory that SO(3) and SU(2) must locally be the same (locally "isomorphic"). We shall now exhibit a specific homomorphism, a classical one,

$$R : SU(2) \longrightarrow SO(3)$$

of SU(2) onto all of SO(3). We start out by looking more carefully at the Lie algebra $\mathfrak{g}$ of SU(2). Of course $\mathfrak{g}$ and $i\mathfrak{g}$ are 3 dimensional vector spaces, $i\mathfrak{g}$ having a specific basis given by (66.5). Define a map

$$\left.\begin{aligned} &\sim : \mathbb{R}^3 \longrightarrow i\mathfrak{g} \\ &\text{by} \\ &\vec{x} \in \mathbb{R}^3 \longmapsto \vec{x}\cdot\vec{\sigma} = x^\mu \sigma_\mu = \begin{bmatrix} z & x-iy \\ x+iy & -z \end{bmatrix} \end{aligned}\right\} \tag{66.7}$$

This map is clearly onto the set of all trace-less hermitian matrices, and it has an inverse because we can solve for $\vec{x}$

$$x = \tfrac{1}{2}\,\mathrm{tr}(\underset{\sim}{x}\sigma_1) \qquad y = \tfrac{1}{2}\,\mathrm{tr}(\underset{\sim}{x}\sigma_2) \qquad z = \tfrac{1}{2}\,\mathrm{tr}(\underset{\sim}{x}\sigma_3) \tag{66.8}$$

---

* after using as new basis $\sigma_i/2$ instead.

Note that under the map $\sim$

$$\begin{pmatrix}1\\0\end{pmatrix} \longrightarrow \sigma_1 \qquad \begin{pmatrix}0\\1\end{pmatrix} \longrightarrow \sigma_2 \qquad \begin{pmatrix}0\\0\end{pmatrix} \longrightarrow \sigma_3$$

and so

$$\tfrac{1}{2}\mathrm{tr}(\sigma_1\sigma_1) = \tfrac{1}{2}\mathrm{tr}(\sigma_2\sigma_2) = \tfrac{1}{2}\mathrm{tr}(\sigma_3\sigma_3) = 1$$

while

$$\tfrac{1}{2}\mathrm{tr}(\sigma_\mu\sigma_\nu) = 0, \quad \mu \neq \nu.$$

Thus, if we define a scalar product $\langle , \rangle$ in $i\mathfrak{sl}$ by

$$\langle \underset{\sim}{X}, \underset{\sim}{X} \rangle \equiv \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{X}\underset{\sim}{X}) \qquad (66.9)$$

then $\sigma_1, \sigma_2, \sigma_3$ form an o.n basis and

$$\sim : \mathbb{R}^3 \longrightarrow i\mathfrak{sl}$$

is an isometry of vector spaces; $\vec{x} \longmapsto \underset{\sim}{X} = x^\mu\sigma_\mu = \vec{x}\cdot\vec{\sigma}$

Now $SU(2)$ operates on $\mathfrak{sl} = su(2)$ by the "adjoint" action

$u \in SU(2)$ sends $X \in su(2)$ into uXu^{-1}

$$X \longmapsto uXu^{-1}$$

(of course $(L_u)_* \circ (R_{u^{-1}})_* : \mathfrak{sl} \longrightarrow \mathfrak{sl}$). Thus we have an action

$$\underset{\sim}{X} \in i\mathfrak{sl} \xrightarrow{Ad(u)} u\underset{\sim}{X}u^{-1} \in i\mathfrak{sl}.$$

$$Ad : SU(2) \longrightarrow Gl(i\mathfrak{sl})$$

where $Gl(i\mathfrak{sl})$ are the nonsingular 3×3 matrices acting on $i\mathfrak{sl}$.

For each 2×2 $u \in SU(2)$ we are associating a 3×3 $Ad(u) : i\mathfrak{sl} \longrightarrow i\mathfrak{sl}$, ie $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$. Note that

$$Ad(u_1u_2)(\underset{\sim}{X}) = (u_1u_2)\,\underset{\sim}{X}\,(u_1u_2)^{-1}$$
$$= Ad(u_1) \circ Ad(u_2)(\underset{\sim}{X})$$

and so Ad is a homomorphism of $SU(2)$ into $Gl(3,\mathbb{R})$. We say that Ad is a representation of $SU(2)$ by 3×3 real matrices (recall that $\sim$ allows us to identify $\mathbb{R}^3$ with $i\mathfrak{sl}$).

Note

$$\langle \mathrm{Ad}(u)\underset{\sim}{X}, \mathrm{Ad}(u)\underset{\sim}{Y}\rangle = \tfrac{1}{2}\mathrm{tr}\left(\mathrm{Ad}(u)\underset{\sim}{X}\ \mathrm{Ad}(u)\underset{\sim}{Y}\right)$$

$$= \tfrac{1}{2}\mathrm{tr}\left(u\underset{\sim}{X}u^{-1}u\underset{\sim}{Y}u^{-1}\right) = \tfrac{1}{2}\mathrm{tr}\left(\underset{\sim}{X}\underset{\sim}{Y}\right)$$

$$= \langle \underset{\sim}{X}, \underset{\sim}{Y}\rangle$$

and so each $\mathrm{Ad}(u)$ is an isometry of $\mathbb{R}^3$, ie

$\mathrm{Ad}(u)$ is an orthogonal 3×3 matrix for each $u \in SU(2)$.

$$\mathrm{Ad} : SU(2) \longrightarrow O(3).$$

and so $\det \mathrm{Ad}(u) = \pm 1$.

(3) Digression on the Topology of $SU(2)$. The usual representation of $SU(2)$ is as 2×2 matrices

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

We shall show that topologically $SU(2)$ is simply S^3, the 3 dimensional sphere! To see this, note that that S^3 can be represented as the set of unit vectors in $\mathbb{C}^2 \approx \mathbb{R}^4$

$$S^3 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \;\middle|\; |z_1|^2 + |z_2|^2 = 1 \right\}.$$

Note that $SU(2) : S^3 \longrightarrow S^3$; this is the meaning of being unitary. Further note that $SU(2)$ acts transitively on S^3, for $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in S^3$ can be sent into the generic point $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in S^3$ simply by

$$u = \begin{pmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{pmatrix} \in SU(2)$$

From lecture 61, § 3, we then know that topologically

$$S^3 = SU(2)/H$$

where H is the stability subgroup for the point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. But

But $u\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$ for $u \in SU(2)$ only if $u = \begin{pmatrix}1&0\\0&1\end{pmatrix} = I$. Thus $H = I$ and

$$S^3 \sim SU(2)$$

topologically.

In particular $SU(2)$ is connected; and since $\det Ad(u) = \pm 1$ must be continuous we conclude $\det Ad(u) = 1$, and so

$$Ad : SU(2) \longrightarrow SO(3).$$

This homomorphism however is not an isomorphism, i.e. it is not 1:1, for clearly $Ad(u) = Ad(-u)$. In fact Ad is a 2:1 homomorphism. This should not surprise us since $SO(3)$ is topologically $P^3(\mathbb{R})$ which is covered twice by the sphere S^3!

Let now $u(t) = e^{itA}$, A hermitian 2×2, be a 1-parameter subgroup of $SU(2)$; then we have a 1-parameter subgroup of rotations of $\mathbb{R}^3$ given by

$$Ad\, u(t)\, x = e^{itA}\, \underset{\sim}{x}\, e^{-itA}.$$

For velocity vector at $x \in \mathbb{R}^3$ we have (recall $\mathbb{R}^3 \sim i\circ f$)

$$\frac{d}{dt} Ad\, u(t)\, x \Big|_{t=0} = \frac{d}{dt} e^{itA}\, \underset{\sim}{x}\, e^{-itA} \Big|_{t=0}$$

$$= i\,[A, \underset{\sim}{x}\,]$$

$$= i\,[A^j \sigma_j, x^k \sigma_k] = i A^j x^k\, [\sigma_j, \sigma_k]$$

$$= (\text{from (66.6)}) \quad = -2 A^j x^k\, \varepsilon_{jk\ell}\, \sigma_\ell$$

$$= -2\,(A \times r)^\ell \sigma_\ell$$

From our discussion preceding (66.3) we see that this 1-parameter group of rotations of $\mathbb{R}^3$ has an angular velocity vector $\vec{\omega} = -2\vec{A}$, and

$$\mathrm{Ad}(e^{itA})\,x = e^{2t\vec{E}\cdot\vec{A}}\,x$$

$\therefore$ the 1-parameter subgroup $u = e^{itA}$, A hermitian gets sent by Ad into the 1-parameter group of rotations of $\mathfrak{so} = \mathbb{R}^3$ $e^{2t\vec{E}\cdot\vec{A}}$

for example, the vector $A = \begin{pmatrix}0\\0\\1\end{pmatrix}$, ie $A = \sigma_3 = \begin{bmatrix}1&0\\0&-1\end{bmatrix}$ generates the 1-parameter subgroup

$$\exp i\theta\begin{bmatrix}1&0\\0&-1\end{bmatrix} = \begin{bmatrix}e^{i\theta}&0\\0&e^{-i\theta}\end{bmatrix}$$

of $SU(2)$, and this corresponds, under Ad, to the 1-parameter group of rotations of $\mathbb{R}^3 = i\sigma$

$$\exp 2\theta E_3 = \exp\begin{bmatrix}0&-2\theta&0\\2\theta&0&0\\0&0&0\end{bmatrix} = \begin{bmatrix}\cos 2\theta&-\sin 2\theta&0\\\sin 2\theta&\cos 2\theta&0\\0&0&1\end{bmatrix}.$$

A final remark: we have $\mathrm{Ad}: SU(2) \longrightarrow SO(3)$ and we have indicated that this map is 2:1; in fact $\ker \mathrm{Ad} = \{u \in SU(2) \mid \mathrm{Ad}\,u = I \in SO(3)\}$ is, as we said the subgroup $\{\pm I\}$ of $SU(2)$. Is Ad onto?, ie is every rotation of $\mathbb{R}^3$ of the form $\mathrm{Ad}\,u$ for some $u \in SU(2)$? We have shown that at the identity $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ of $SU(2)$

$$(\mathrm{Ad})_*: \quad , su(2) \longrightarrow so(3)$$

is in fact

$$\mathrm{Ad}_*(A) = 2A.$$

Ad_* is 1:1 at $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ of $SU(2)$. The same is true at $\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$ since $\mathrm{Ad}(u) = \mathrm{Ad}(-u)$. Thus Ad_* is onto at the inverse image of $\begin{pmatrix}1&0\\0&1\end{pmatrix} \in SO(3)$, and by our correspondence $i\sigma \sim \mathbb{R}^3$ orientations are preserved under Ad. Hence Brouwer degree of $\mathrm{Ad}: SU(2) \to SO(3)$ is $+2$. In particular Ad is <u>onto</u>, for if $g \in SO(3)$ had no preimage then g would be a regular point (see Lecture 3, § 1), implying the degree is 0, a contradiction. We conclude $SO(3) \approx SU(2)/\{\pm I\}$.

Lecture 67. Hamilton, Clifford, and Dirac

(1) Rotations of $\mathbb{R}^3$ and Pauli Matrices. We saw in the last lecture that there is a representation of $SU(2)$ as a group of rotations of $\mathbb{R}^3$

$$e^{i\theta \vec{A}\cdot\vec{\sigma}} \xrightarrow{Ad} e^{-2\theta \vec{A}\cdot\vec{E}} \qquad (67.1)$$

and $Ad : SU(2) \longrightarrow SO(3)$ is a 2:1 covering. We can also say, roughly, that to the rotation of $\mathbb{R}^3$ about the axis given by the unit vector $\vec{A}$ thru an angle θ one can associate two 2×2 special unitary matrices $e^{-i\frac{\theta}{2}\vec{A}\cdot\vec{\sigma}}$ and $e^{-i(\frac{\theta}{2}+\pi)\vec{A}\cdot\vec{\sigma}}$. In other words $SO(3)$ not only has the usual representation as 3×3 matrices operating on $\mathbb{R}^3$, it also has a double valued representation as 2×2 special unitary matrices acting on $\mathbb{C}^2$! $SO(3)$ thus normally acts on $\mathbb{R}^3$, ie vectors, but it can also act in a double valued way on $\mathbb{C}^2$; these $\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$ on which $SO(3)$ acts are called spinors. Mathematicians do not like double valued anythings; they prefer to say that $SU(2)$ furnishes naturally a spinor representation of the 2-fold cover of $SO(3)$; when $SU(2)$ is thought of as the 2-fold cover of $SO(3)$ it is called the spin group $Spin(3)$. Spinors form a representation of $Spin(3)$. The topological reason that $SO(3)$ can have a "2 fold cover" is that $SO(3)$ is not simply connected. The reader can now understand the remarks of Lecture 1, §2e!

The relation (67.1) is of great practical importance because the Pauli matrices σ_i are much easier to deal with than the 3×3 matrices E_i. For example

we will consider the problem of composing two rotations $R_1 = e^{\theta \vec{A}\cdot\vec{E}}$ and $R_2 = e^{\varphi \vec{B}\cdot\vec{E}}$ in $\mathbb{R}^3$. But first note explicitly the properties of the Pauli matrices.

$$\sigma_1\sigma_2 = i\sigma_3, \quad \sigma_2\sigma_3 = i\sigma_1, \quad \sigma_3\sigma_1 = i\sigma_2. \tag{67.2}$$

from which one immediately gets the commutation relations

$$[\sigma_i, \sigma_j] = 2i\,\varepsilon_{ijk}\,\sigma_k. \tag{67.3}$$

We shall also need the anticommutation relations

$$\{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} I \tag{67.4}$$

where I is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Problem 104. (i) Show that (67.4) holds.

(ii) Using (67.3) and (67.4), show that if $\vec{A}$ and $\vec{B}$ are two unit vectors in $\mathbb{R}^3$, then

$$(\vec{A}\cdot\vec{\sigma})(\vec{B}\cdot\vec{\sigma}) = (\vec{A}\cdot\vec{B})I + i(\vec{A}\times\vec{B})\cdot\vec{\sigma} \tag{67.5}$$

and in particular $(\vec{A}\cdot\vec{\sigma})^2 = I$

(iii) Show that

$$\boxed{e^{-\frac{\theta}{2}\vec{A}\cdot\vec{\sigma}} = (\cos\tfrac{\theta}{2})I - (i\sin\tfrac{\theta}{2})(\vec{A}\cdot\vec{\sigma})} \tag{67.6}$$

(iv) Let B be a second unit vector. Show

$$R_1R_2 \equiv e^{-i\frac{\theta}{2}(\vec{A}\cdot\vec{\sigma})}\, e^{-i\frac{\varphi}{2}(\vec{B}\cdot\vec{\sigma})} =$$

$$\left[\cos\tfrac{\theta}{2}\cos\tfrac{\varphi}{2} - (\sin\tfrac{\theta}{2}\sin\tfrac{\varphi}{2})(A\cdot B)\right] I$$

$$- i\left[\sin\tfrac{\theta}{2}\cos\tfrac{\varphi}{2}(\vec{A}\cdot\vec{\sigma}) + \cos\tfrac{\theta}{2}\sin\tfrac{\varphi}{2}(\vec{B}\cdot\vec{\sigma}) + \sin\tfrac{\theta}{2}\sin\tfrac{\varphi}{2}(A\times B)\cdot\sigma\right]$$

(iv) Let R_1 = rotation about z axis of 90°, R_2 = rotation about y axis of 90°. Use (iii) to describe the rotation R_1R_2.

01 Clifford Algebras

The expression (67.6) was known essentially to Hamilton in terms of quaternions rather than Pauli matrices (this was the motivation for Hamilton's invention of quaternions; for more discussion see Misner, Thorne & Wheeler, Chapter 41).

Let us isolate some of the properties of the Pauli matrices that will be necessary for important generalizations.

First note that σ_1, σ_2 and σ_3 generate a 3 dimensional linear space V^3 under addition and multiplication by real scalars. In this vector space there is a scalar product $\langle , \rangle$ and

$$\langle \sigma_i, \sigma_j \rangle = g_{ij} = \delta_{ij}$$

(see (66.9)). Furthermore there is a multiplication (matrix multiplication) in V^3 but V^3 is not closed under this multiplication, e.g. $\sigma_1\sigma_2 = i\sigma_3$ and $i\sigma_3$ is not in V^3 since i is not real. If we put

$$e_1 = \sigma_1 \qquad e_2 = \sigma_2 \qquad e_3 = \sigma_3$$

$$e_4 = e_1 e_2 \qquad e_5 = e_2 e_3 \qquad e_6 = e_3 e_1$$

$$e_7 = e_1 e_2 e_3$$

$$e_8 = e_7 \cdot e_7$$

then the resulting vector space (real coefficients) is 8 dimensional and it is closed under (matrix) multiplication (note that $e_7 = i\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $e_8 = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$) It thus forms an algebra (a linear space with an associative product). Furthermore

$$e_i e_j + e_j e_i = 2 g_{ij} I \qquad i,j = 1,2,3$$

This 8 dimensional vector space is called the

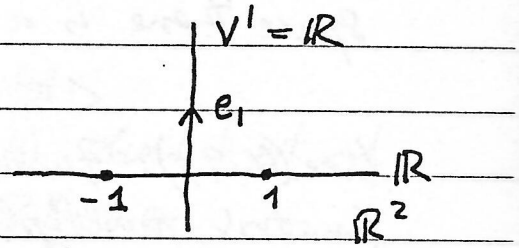
Clifford Algebra $\mathcal{C}(V^3)$ generated by the Pauli matrices, the "Pauli algebra"

In general, if $\mathcal{C}$ is an associative algebra generated by a subspace V^n and if V^n has a scalar product $\langle , \rangle$ and a basis $e_1, \dots, e_m$ such that

$$\underbrace{e_i e_j + e_j e_i}_{\text{products in } \mathcal{C}} = 2g_{ij} \cdot 1 = 2\langle e_i, e_j \rangle 1 \qquad (67.7)$$

Then we say that $\mathcal{C} = \mathcal{C}(V^n)$ is the Clifford Algebra generated by V^n.

Ex. 1. $V^1 = \mathbb{R} = \mathcal{C}(V^1)$ is the Clifford algebra of real numbers, $e_1 = 1$, $\langle e_1, e_1 \rangle = 1$



Ex. 2. Let $\mathcal{C} = \mathbb{R}^2$

Let $e_1 = (0,1)$. Define scalar product in V^1 (the y axis) by $\langle e_1, e_1 \rangle = -1$. Define multiplication in $\mathcal{C}$ by putting

$$e_1^2 = (-1, 0)$$

and then

$$(a + be_1)(c + de_1) = ac + (bc + ad)e_1 + bd\, e_1^2$$

$$= (ac - bd) + (bc + ad)e_1$$

The resulting algebra $\approx \mathbb{C}$. We could say in this case that $\mathbb{C}$ is the Clifford algebra generated by the real line $\mathbb{R}$ with scalar product $\langle e_1, e_1 \rangle = -1$.

Ex. 3. Let V^2 be a 2 dimensional real vector space with basis e_1, e_2 and scalar product $\langle e_i, e_j \rangle = -\delta_{ij}$. We want to define a product such that $e_1^2 = -1$, $e_2^2 = -1$ and $e_1 e_2 = -e_2 e_1$, in order to satisfy (67.7). We then add to V^2 another copy

of the real line $\mathbb{R}$ to take care of scalars like -1 and we add another one dimensional vector space with basis e_3 and put $e_1^2 = -1 = e_2^2$

$$e_1 e_2 = -e_2 e_1 = e_3$$

With Hamilton we call these

$$e_1 = i, \quad e_2 = j, \quad e_3 = k$$

and then $\mathcal{C}(V^2)$ is the algebra of quaternions

$$q = a_0 + a_1 i + a_2 j + a_3 k.$$

Of course automatically

$$k^2 = kk = (e_1 e_2)(e_1 e_2) = e_1 (e_2 e_1) e_2$$
$$= -e_1 (e_1 e_2) e_2 = -(e_1 e_1)(e_2 e_2) = -1$$

Ex. 4 Let V^3 be a 3 dimensional vector space with, this time, a positive definite scalar product

$$\langle e_i, e_j \rangle = \delta_{ij}$$

We therefore want a product such that

$$e_1^2 = e_2^2 = e_3^2 = 1$$

$$e_i e_j + e_j e_i = 0 \qquad i \neq j$$

We have just talked about this algebra; it is the Pauli algebra with basis

$$e_1 = \sigma_1, \quad e_2 = \sigma_2, \quad e_3 = \sigma_3.$$

Do we need more Clifford algebras?

(3) The Dirac Program; the Square Root of the D'Alembertian.

The Klein-Gordon equation in Minkowski space M_0^4 for a particle of mass m is

$$\left.\begin{aligned} \Box^2 \psi &= -m^2 \psi \\ \text{ie} \quad g^{ij} \partial_i \partial_j \psi &= -m^2 \psi \end{aligned}\right\} \tag{67.8}$$

where

$$(g_{ij}) = (\eta_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{67.9}$$

η_{ij} is the Minkowski metric ; note that we have changed our sign convention for η_{ij}. The present sign convention is used e.g. in Bleckers' book "Gauge Theory and Variational Principles" which is a good reference for material in this lecture.

With Dirac we shall try to construct a first order differential operator

$$D = \alpha^i \partial_i$$

such that

$$\Box^2 \psi = D(D\psi)$$

i.e. we shall try to construct a "square root" of the d'Alembertain $\Box^2$. Then we can solve the Klein Gordon equation (which, by the way was first proposed as a relativistic wave equation by Schrödinger) by solving the first order equation

$$D\psi = \alpha^j \partial_j \psi = -im\psi \qquad (67.10)$$

for then

$$\Box^2 \psi = D(D\psi) = D(-im\psi) = -im\, D\psi$$
$$= -m^2 \psi$$

as desired!

Back then to

$$\Box^2 = D \cdot D = (\alpha^i \partial_i)(\alpha^j \partial_j)$$
$$= \alpha^i \alpha^j \partial_i \partial_j = \frac{1}{2}(\alpha^i \alpha^j + \alpha^j \alpha^i)\partial_i \partial_j$$

and so we require

$$\alpha^i \alpha^j + \alpha^j \alpha^i = 2g^{ij} = 2\eta^{ij}$$

i.e. the α's must be the 4 generators of a Clifford algebra ; (they certainly cannot be scalars). Dirac also wanted, of course, some sort of "covariance" under Lorentz transformations. We do all this next time.

Lecture 68. The Dirac Algebra

① How go to $\Delta x \to 0$
② How decompose into nodes

(1) The Lorentz Group. Our principal reference is again Bleecker's book. With the Minkowski metric

$$(\eta_{ij}) = \begin{pmatrix} 1 & & & \\ & -1 & & 0 \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$$

we define the Lorentz group L to be the linear isometries of M^4_0

$$L = \left\{ 4\times 4 \text{ matrices } B : M^4_0 \to M^4_0 \;\middle|\; \langle Bx, By\rangle = \langle x, y\rangle \right\}.$$

If we write the scalar product, as usual, as

$$\langle x, y\rangle = x^T \eta\, y$$

then

$$\langle Bx, By\rangle = (Bx)^T \eta\, By = x^T B^T \eta\, By$$

requires then

$$B^T \eta B = \eta \tag{68.1}$$

Since $\det \eta = -1$ we see $(\det B)^2 = 1$, $\det B = \pm 1$ since B is a real matrix. Since $1 = \langle e_0, e_0\rangle = \langle Be_0, Be_0\rangle$ and

$$B(e_0) = \begin{pmatrix} B^0_0 \\ B^1_0 \\ B^2_0 \\ B^3_0 \end{pmatrix}$$

we see that $(B^0_0)^2 \geq 1$. Bleecker shows that L breaks up into 4 connected components

$$L^{\uparrow}_{+} : \det B > 0 \;\&\; B^0_0 \geq 1$$
$$L^{\uparrow}_{-} : \det B < 0 \;\&\; B^0_0 \geq 1$$
$$L^{\downarrow}_{+} : \det B > 0 \;\&\; B^0_0 \leq -1$$
$$L^{\downarrow}_{-} : \det B < 0 \;\&\; B^0_0 \leq -1$$

where $L^{\uparrow}_{+}$ is the component holding the identity

Look at the Lie algebra of L. Write $B = e^{tS}$. Then

$$(e^{tS})^T \eta\, e^{tS} = \eta$$

and differentiating with respect to t and putting $t = 0$

$$S^T\eta + \eta S = 0$$

Since $\eta^T = \eta$, this says

$$(\eta S)^T = -(\eta S) \tag{68.2}$$

which says that when we lower the upper index of S the resulting covariant 2nd order tensor is <u>skew symmetric</u>

$$S_{ij} = \eta_{ik} S^k{}_j = -S_{ji}$$

Thus $\dim L = 4\cdot 3/2 = 6$.

Just as $SO(3)$ is covered twice by $SU(2)$, we shall indicate why $L_+^\uparrow$ is covered twice by $Sl(2,\mathbb{C})$, the complex 2×2 matrices with determinant 1 (which of course is also 6 dimensional).

Let $H(2,\mathbb{C})$ be the space of 2×2 hermitian matrices $A = A^* = \bar{A}^T$, with <u>no</u> requirement on the trace. This space is 4 dimensional and we shall now identify $H(2,\mathbb{C})$ with $\mathbb{R}^4$. Let us augment the Pauli matrices by adding in the identity matrix

$$\left.\begin{aligned} \tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \tau_1 = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \tau_2 = \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \tau_3 = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}\right\} \tag{68.3}$$

These matrices give us a basis for $H(2,\mathbb{C})$. Define a map $\sim : \mathbb{R}^4 \longrightarrow H(2,\mathbb{C})$ by

$$x \in \mathbb{R}^4 \longmapsto x\tau = \underset{\sim}{x} = x^0\tau_0 + x^1\tau_1 + x^2\tau_2 + x^3\tau_3 \tag{68.4}$$

$$\underset{\sim}{x} = \begin{pmatrix} x_0 + z & x_1 - iy \\ x + iy & x_0 - z \end{pmatrix}$$

We can write $\underset{\sim}{x} = x^0 + \vec{x}\cdot\vec{\sigma}$, and we can solve for x

$$x^i = \frac{1}{2}\operatorname{tr}(\underset{\sim}{x}\tau^i) \tag{68.5}$$

We also have the easily verified

$$\det \underset{\sim}{x} = \langle x, x \rangle \tag{68.6}$$

We shall also have need of another identification of $\mathbb{R}^4$ with $H(2,\mathbb{C})$, namely

$$\tilde{x} = x^0\tau_0 - (x^1\tau_1 + x^2\tau_2 + x^3\tau_3) \tag{68.7}$$

Thus

$$\underset{\sim}{x} = x^0\tau_0 + \vec{x}\cdot\vec{\sigma}$$

$$\tilde{x} = x^0\tau_0 - \vec{x}\cdot\vec{\sigma} \quad (\text{"space-inversion"})$$

Then

$$\tilde{x} = \begin{pmatrix} x^0 - z & -x + iy \\ -x - iy & x^0 + z \end{pmatrix}$$

and one computes

$$\begin{cases} \det \tilde{x} = \langle x, x \rangle \\ \underset{\sim}{x}\,\tilde{x} = \tilde{x}\,\underset{\sim}{x} = \langle x, x \rangle I \end{cases} \tag{68.8}$$

Theorem: There is a homomorphism

$$\Lambda : SL(2,\mathbb{C}) \longrightarrow L_+^{\uparrow}$$

given by (when $A \in SL(2,\mathbb{C})$)

$$\underline{\Lambda(A)(x)} = A\,\underset{\sim}{x}\,A^* = A\,\underset{\sim}{x}\,\bar{A}^T$$

Furthermore Λ is onto and is a 2-fold covering with $\Lambda^{-1}(I) = \pm I$.

Note that Λ is similar to what we used for $SU(2) \longrightarrow SO(3)$; in fact when $A \in SU(2) \subset SL(2,\mathbb{C})$, $A^* = A^{-1}$!

We refer to Bleecker's book for a proof of this theorem. We shall merely indicate one or two aspects.

To see that Λ is a homomorphism, note that

$$\langle \Lambda(A)x, \Lambda(A)x\rangle = \det \underline{\Lambda(A)x} = \det A\underset{\sim}{x}A^*$$

$$= \det A \det\underset{\sim}{x} \det A^* = \det\underset{\sim}{x} = \langle x, x\rangle$$

and so $\Lambda(A)$ is a Lorentz transformation, $\det\Lambda(A) = \pm 1$. To see that $\det\Lambda(A) = +1$ we need only know that $Sl(2,\mathbb{C})$ is connected. But

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}$$

can be thought of as a pair of complex vectors $\begin{pmatrix} a_{11}\\ a_{21}\end{pmatrix}$ & $\begin{pmatrix} a_{12}\\ a_{22}\end{pmatrix}$ spanning an "area" $\det A = 1$. By our usual Gram-Schmidt-like process used for $Sl(2,\mathbb{R})$ (See Lecture 5b, § 4) we may deform $Sl(2,\mathbb{C})$ into the subgroup $SU(2)$, which is connected (in fact it is topologically S^3). Hence $\det\Lambda(A) = +1$. Since $A = I$ yields a Lorentz transformation with $B^0{}_0 \geq 1$, connectedness also shows $B^0{}_0 \geq 1$ for all $A \in Sl(2,\mathbb{C})$, ie Λ maps into $L_+^\uparrow$. It is shown then that the image $\Lambda(Sl(2,\mathbb{C}))$ is an open subgroup of the connected group $L_+^\uparrow$ and thus, rather clearly it is all of $L_+^\uparrow$.

Just as we used $SU(2) \longrightarrow SO(3)$ to yield a double valued spinor representation of $SO(3)$, so the above $Sl(2,\mathbb{C}) \longrightarrow L_+^\uparrow$ allows us to represent $L_+^\uparrow$ as a group of 2×2 complex matrices acting on $\mathbb{C}^2$, and again this representation is double valued. This representation, which is basically just the usual representation of $Sl(2,\mathbb{C})$ is a double valued spinor representation of the Lorentz group identity group component $L_+^\uparrow$. It is better to consider instead, that it is the usual representation of $Sl(2,\mathbb{C})$. As a

representation of $Sl(2,\mathbb{C})$ or as a 2 valued representation of $L_+^\uparrow$ it is written $\mathcal{D}^{(\frac{1}{2},0)}$

(2) The Dirac Algebra. We have seen in Lecture 67 that the Pauli matrices generate a Clifford algebra

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} 1.$$

There is a rather standard procedure leading from a Clifford algebra with n generators to a new Clifford algebra with $(n+1)$ generators. We are only interested in going from the Pauli algebra to the next higher one. There is a complication due to the fact that the Pauli algebra uses the metric δ_{ij} (3×3) while for relativistic purposes the "Dirac algebra" must use the Minkowski metric η_{ij} (4×4). We proceed as follows

Note that in the case of the Pauli algebra, the map $\sim : \mathbb{R} \longrightarrow 2\times 2$ matrices ; $\vec{x} \longmapsto \vec{x}\cdot\vec{\sigma}$, can be thought of as a map $\sigma : \mathbb{R}^3 \longrightarrow gl(2,\mathbb{C})$

$$\sigma(\vec{x}) = \underset{\sim}{x} = \vec{x}\cdot\vec{\sigma}.$$

For example

$$\sigma\begin{pmatrix}1\\0\\0\end{pmatrix} = \sigma_1 \qquad \sigma\begin{pmatrix}0\\1\\0\end{pmatrix} = \sigma_2 \qquad \sigma\begin{pmatrix}0\\0\\1\end{pmatrix} = \sigma_3.$$

We now define a map $\gamma : \mathbb{R}^4 \longrightarrow gl(4,\mathbb{C})$ of $\mathbb{R}^4$ into all 4×4 complex matrices by

$$\gamma(x) = \begin{bmatrix} 0 & \underset{\sim}{x} \\ \underset{\sim}{\tilde{x}} & 0 \end{bmatrix}. \qquad (68.9)$$

The meaning of this will be discussed in the next lecture.

In particular

$$\gamma_1 \equiv \gamma(e_1) = \begin{bmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{bmatrix} = \left[\begin{array}{cc:cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \hdashline 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array}\right]$$

$$\gamma_2 \equiv \gamma(e_2) = \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix}$$

$$\gamma_3 \equiv \gamma(e_3) = \begin{bmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\gamma_0 \equiv \gamma(e_0) = \begin{bmatrix} 0 & \tau_0 \\ \tau_0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(68.10)

the famous Dirac matrices.

The matrices $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ generate a Clifford algebra. In fact we have

Theorem For all $x \in \mathbb{R}^4$, $y \in \mathbb{R}^4$

$$\gamma(x)\gamma(y) + \gamma(y)\gamma(x) = 2\langle x, y\rangle I \qquad (68.11)$$

where $\langle\ ,\ \rangle$ is the Minkowski metric.

Proof. Both sides of (68.11) are bilinear, symmetric functions of x and y and hence ("polarization") it is sufficient to verify (68.11) when $y = x$. But

$$\gamma(x)\gamma(x) = \begin{bmatrix} 0 & \underset{\sim}{x} \\ \tilde{x} & 0 \end{bmatrix}\begin{bmatrix} 0 & \underset{\sim}{x} \\ \tilde{x} & 0 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{x}\tilde{x} & 0 \\ 0 & \tilde{x}\underset{\sim}{x} \end{bmatrix}$$

$$= \begin{bmatrix} \langle x, x\rangle I & 0 \\ 0 & \langle x, x\rangle I \end{bmatrix} \quad \text{(from (68.8))}$$

$$= \langle x, x\rangle \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \quad \text{as desired.}$$

Lecture 69. The Dirac Operator $\not{\partial}$ in Minkowski Space

(1) Transformation Properties of the Dirac Matrices. Recall that our goal was to factor the d'Alembertian $\Box^2\psi = \not{D}(\not{D}\psi)$

$$\not{D}\psi = \alpha^\mu \partial_\mu = \alpha^\mu \partial/\partial x^\mu$$

in Minkowski space M_0^4 (with metric $\eta = \text{diag}(1,-1,-1,-1)$. The α's then had to satisfy the Clifford anti-commutation relations

$$\{\alpha^\mu, \alpha^\nu\} = \alpha^\mu\alpha^\nu + \alpha^\nu\alpha^\mu = 2\eta^{\mu\nu}\cdot 1.$$

Clearly the α's cannot be numbers

. The α's generate a Clifford algebra and in the last lecture we derived a candidate for α, namely $\alpha^\mu = \gamma_\mu$. In order, however, for $\not{D}$ to be a well defined differential operator we must investigate the behavior of the γ's under a Lorentz transformation of M_0^4.

First what do we mean by γ? γ is a frame dependent matrix defined on tangent vectors $\vec{X}$ to M_0^4; if $e_\mu = \partial/\partial x^\mu$ is a Lorentzian frame at a point, and if $\vec{X}$ is a vector at this point, then $\vec{X} = e_\mu X^\mu = eX$ and we have defined

$$\gamma(\vec{X}) = \begin{bmatrix} 0 & \underset{\sim}{X} \\ \tilde{X} & 0 \end{bmatrix}$$

the answer clearly depending on the frame used since the matrices $\tau_0, \tau_1, \tau_2, \tau_3$ are fixed for all time, independent of anything.

Suppose we are given two Lorentzian frames $e = \partial/\partial x$ and $e' = \partial/\partial x'$ at the same point. Then $\vec{X} = eX = e'X'$

$$X' = \frac{\partial x'}{\partial x} X = \Lambda X$$

$$e' = e\frac{\partial x}{\partial x'}$$

The Lorentz transformation Λ will correspond, under

$$\Lambda : SL(2,\mathbb{C}) \longrightarrow L_+^\uparrow$$

to two matrices $\pm A \in Sl(2,\mathbb{C})$; pick one of them. By definition

$$\underset{\sim}{\Lambda(A)X} = AXA^* \tag{69.1}$$

Lemma The 2×2 matrix associated to $\Lambda(A)X$ under $\sim$ is

$$\widetilde{\Lambda(A)X} = A^{*-1}\,\tilde{X}\,A^{-1} \tag{69.1}$$

Proof. Recall from (68.8) that $\underset{\sim}{X}\tilde{X} = \tilde{X}\underset{\sim}{X} = \langle X, X\rangle I$, and so

$$\tilde{X} = \langle X, X\rangle \underset{\sim}{X}^{-1}$$

at least in the case that X is not a null vector, since $\det \tilde{X} = \langle x, x\rangle$. But if we have proven (69.1) when $\langle X, X\rangle \neq 0$ it will follow by continuity for all X. Thus assume $\langle X, X\rangle \neq 0$. Then

$$\begin{aligned}\widetilde{\Lambda(A)X} &= \langle \Lambda(A)X, \Lambda(A)X\rangle \left[\underset{\sim}{\Lambda(A)X}\right]^{-1} \\ &= \langle X, X\rangle \left[A\underset{\sim}{X}A^*\right]^{-1} = \langle X, X\rangle A^{*-1}\underset{\sim}{X}^{-1}A^{-1} \\ &= A^{*-1}\langle X, X\rangle \underset{\sim}{X}^{-1}A^{-1} = A^{*-1}\tilde{X}A^{-1} \qquad \square.\end{aligned}$$

Theorem. Let $\rho : Sl(2,\mathbb{C}) \longrightarrow Gl(4,\mathbb{C})$ be the representation of $Sl(2,\mathbb{C})$ by 4×4 matrices defined by

$$\rho(A) = \begin{bmatrix} A & 0 \\ 0 & A^{*-1} \end{bmatrix} \qquad A \in Sl(2,\mathbb{C}) \tag{69.2}$$

Then the frame dependent Dirac matrices transform as

$$\gamma'(\Lambda(A)X) = \rho(A)\,\gamma(X)\,\rho^{-1}(A) \tag{69.3}$$

(Note: $\Lambda(A)X$ and X are the components of the same vector $\vec{x}$ but given in the two Lorentz coordinate systems x' and x)

Proof.

$$\gamma'(\Lambda(A)X) = \begin{bmatrix} 0 & \underset{\sim}{\Lambda(A)X} \\ \widetilde{\Lambda(A)X} & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & A\underset{\sim}{X}A^* \\ A^{*-1}\tilde{X}A^{-1} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} A & 0 \\ 0 & A^{*-1} \end{bmatrix} \begin{bmatrix} 0 & \underset{\sim}{X} \\ \tilde{X} & 0 \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & A^* \end{bmatrix}$$

$$= \rho(A)\,\gamma(X)\,\rho(A^{-1}) = \rho(A)\,\gamma(X)\,\rho^{-1}(A) \qquad \square.$$

Interpretation. The matrix $\gamma(\vec{X}) : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ depends very much on the Lorentzian frame used in M_o^4. However, if to each Lorentz transformation Λ of M_o^4 we also make a change of frame in the 4 dimensional complex vector space $\mathbb{C}^4 = \mathbf{V}$ given by the change of basis matrix $\rho(A)$, then we see from (69.3) that $\gamma(\vec{X})$ defines a linear transformation $\gamma(\vec{X}) : \mathbf{V} \longrightarrow \mathbf{V}$ that is independent of the Lorentzian frame. This is because the matrix of a linear transformation changes under a change of frame precisely by $\gamma \longrightarrow \rho\gamma\rho^{-1}$! $\gamma'\psi' = (\rho\gamma\rho^{-1})(\rho\psi) = \rho(\gamma\psi) = (\gamma\psi)'$

As mentioned before, the usual representation of $SL(2,\mathbb{C})$ by 2×2 matrices, when thought of as a 2 valued representation of $L_+^\uparrow$ is called the spinor representation of $L_+^\uparrow$, written $\mathcal{D}^{(\frac{1}{2},0)}$. The representation using A^{*-1} instead of A is called $\mathcal{D}^{(0,\frac{1}{2})}$. Thus for $\gamma(\vec{X})$ to be a well defined linear transformation

$$\psi \in \mathbb{C}^4 \longrightarrow \gamma(\vec{X})\psi \in \mathbb{C}^4$$

we must have ψ be a "4 component spinor"; it must transform via

$$\psi \longrightarrow \begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix}\psi$$

for every Lorentz transformation Λ of $\mathbb{R}^4$. This representation

ρ of $sl(2,\mathbb{C})$ is written $\mathcal{D}^{(\frac{1}{2},\frac{1}{2})}$; it is the direct sum of $\mathcal{D}^{(\frac{1}{2},0)}$ and $\mathcal{D}^{(0,\frac{1}{2})}$

A Lorentz transformation $\Lambda : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ must always be accompanied by a change of basis $\rho(A) : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$ (69.4) in spinor space!

Then $\gamma(\vec{X})$ acts on 4 component spinors.

(2) The Dirac Operator. Consider M^4_0 with a given Lorentzian coordinate system x. A "wave function" ψ will then be a function on M^4_0 taking its values in $\mathbb{C}^4$

$$\psi(x) = \begin{pmatrix} \psi^0 \\ \psi^1 \\ \psi^2 \\ \psi^3 \end{pmatrix}$$

We define the Dirac operator $\not{\partial}$ sending wave functions into wave functions by

$$\boxed{\not{\partial}\psi = \gamma_\mu \eta^{\mu\nu} \frac{\partial \psi}{\partial x^\nu}} \qquad (69.5)$$

i.e

$$\boxed{\not{\partial} = \gamma^\nu \partial_\nu}$$

Recall that

$$\gamma_\mu = \gamma\left(\frac{\partial}{\partial x^\mu}\right) = \begin{bmatrix} 0 & \tau_\mu \\ \pm\tau_\mu & 0 \end{bmatrix}.$$

This defines $\not{\partial}$ using the Lorentzian coordinate system x; what happens if we consider the same definition using a Lorentzian system $x' = \Lambda x$; then we must put

$$\frac{\partial}{\partial x'} = \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} \quad \text{and} \quad \psi' = \rho(A)\psi.$$

We have

$$\not{\partial}'\psi' = \gamma'_\mu \eta'^{\mu\nu} \frac{\partial \psi'}{\partial x'^\nu} = \gamma'\left(\frac{\partial}{\partial x'^\mu}\right) \eta'^{\mu\nu} \rho(A) \frac{\partial \psi}{\partial x'^\nu}$$

$$= \gamma'\left(\frac{\partial}{\partial x'^\mu}\right) \eta'^{\mu\nu} \rho(A) \frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial \psi}{\partial x^\alpha}.$$

Now the vector $\partial/\partial x'^\mu$ has components in the original frame given from

$$\partial/\partial x'^\mu = \frac{\partial x^\beta}{\partial x'^\mu} \frac{\partial}{\partial x^\beta}$$

The transformation rule (69.3) then says

$$\gamma'\left(\frac{\partial}{\partial x'^\mu}\right) = \rho(A)\, \gamma\left(\frac{\partial x^\beta}{\partial x'^\mu} \frac{\partial}{\partial x^\beta}\right) \rho^{-1}(A)$$

$$= \rho(A) \frac{\partial x^\beta}{\partial x'^\mu} \gamma\left(\frac{\partial}{\partial x^\beta}\right) \rho^{-1}(A)$$

$$= \rho(A)\, \gamma_\beta\, \rho^{-1}(A) \frac{\partial x^\beta}{\partial x'^\mu}.$$

Then

$$\not{\partial}'\psi' = \rho(A)\, \gamma_\beta\, \rho^{-1}(A) \underbrace{\frac{\partial x^\beta}{\partial x'^\mu} \eta'^{\mu\nu} \frac{\partial x^\alpha}{\partial x'^\nu}} \rho(A) \frac{\partial \psi}{\partial x^\alpha}$$

$$= \rho(A)\, \gamma_\beta\, \eta^{\beta\alpha} \frac{\partial \psi}{\partial x^\alpha} = \rho(A)\, \gamma^\alpha \frac{\partial \psi}{\partial x^\alpha}$$

$$= \rho(A) \cdot \not{\partial}\psi .$$

Thus $\not{\partial}$ is a well defined first order differential operator on 4 component spinors of type $\mathcal{D}^{(\frac{1}{2},\frac{1}{2})}$!

Now

$$\gamma^\alpha = \eta^{\alpha\beta}\gamma_\beta$$

and so, from (68.10)

$$\gamma^0 = \begin{bmatrix} 0 & -I \\ -I & 0 \end{bmatrix}, \quad \gamma^1 = \begin{bmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{bmatrix}, \quad \gamma^2 = \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{bmatrix}, \quad \gamma^3 = \begin{bmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{bmatrix}$$

and finally $\not{\partial} = \begin{bmatrix} 0 & -I\partial_t + \sigma_1\partial_x + \sigma_2\partial_y + \sigma_3\partial_z \\ -I\partial_t - \sigma_1\partial_x - \sigma_2\partial_y - \sigma_3\partial_z & 0 \end{bmatrix}$

Lecture 70. The Dirac Operator in Curved Space

(1) The Spinor Bundle. Consider now a pseudo-Riemannian 4 manifold M^4, rather than Minkowski space M_0^4. We suppose that there are patches $\{U, V, \ldots\}$ on M^4 and "orthonormal" frame fields $e_U, e_V, \ldots$, on each. Thus

$$\langle e^U_\mu, e^U_\nu \rangle = \eta_{\mu\nu}$$

and in an overlap

$$e_V(x) = e_U(x)\, c_{UV}(x)$$

where $c_{UV} : U \cap V \longrightarrow L_+^\uparrow$. In the case of M_0^4 there was no problem associating a $c'_{UV}(x) \in Sl(2,\mathbb{C})$ with each $c_{UV}(x) \in L_+^\uparrow$ since $c_{UV}(x)$ in fact was a constant (a global Lorentz transformation) but now we have variable $c_{UV}(x)$; some how, in order to construct the vector $\mathbb{C}^4$ bundle of 4 component spinors, we must pick, for each $c_{UV}(x) \in L_+^\uparrow$ one of the two possible $c'_{UV}(x) \in Sl(2,\mathbb{C})$ covering $c_{UV}(x)$ and this must be done in an unambiguous and differentiable manner in order that $\psi_V = \rho(c'_{UV}(x))\psi_U$ make sense. If we cannot do this, i.e. if we cannot "lift the structure group $L_+^\uparrow$ to the group $Sl(2,\mathbb{C})$" then there will be no spinor bundle over M^4 whose cross sections can serve as wave functions ψ for the Dirac electron! We then say that that the pseudo-Riemannian M^4 does not admit a spin structure! We shall consider the question of the existence of a spin structure in Lecture , but for the present we shall assume that such a spin structure exists on our M^4!

This means that to our manifold M^4 we may associate to the $\mathbb{R}^4$ bundle TM^4, the tangent bundle to M^4 with transition functions $c_{UV} : U \cap V \longrightarrow L_+^\uparrow$, anew

$\mathbb{C}^4$-vector bundle SM^4 with transition functions

$$\rho_{UV} : U \cap V \longrightarrow Gl(4, \mathbb{C})$$

defined by (See (69.2))

$$\rho_{UV}(x) = \rho(c'_{UV}(x)) = \begin{bmatrix} c'_{UV}(x) & 0 \\ 0 & [c'_{UV}(x)]^{*-1} \end{bmatrix}$$

The Dirac operator $\partial\!\!\!/ = \gamma^\mu \partial_\mu$ constructed for Minkowski space M^4_0 will not work in this new situation; in our proof that

$$\partial\!\!\!/'\psi' = \rho(A)\partial\!\!\!/\psi$$

for M^4_0 we used that fact that the matrices $A \in Sl(2,\mathbb{C})$ were constant (global Lorentz transformations). Our $A = c'_{UV}(x)$ are not constant; we shall have to replace $\partial_\mu\psi = \partial\psi/\partial x^\mu$ by some sort of covariant derivative $\mathcal{D}_\mu\psi$. Before considering our specific problem we shall look at the general situation involved with "associated bundles".

(2) Connection for an Associated Bundle. Let E be an $\mathbb{R}^N$ or $\mathbb{C}^N$ vector bundle over a manifold M^n with transition functions $c_{UV}(x) \in G$ taking values in a structure Lie group G. Now let V ($= \mathbb{R}^K$ or $\mathbb{C}^K$) be any vector space (with usual basis) and suppose that there is some representation ρ of G as a group of linear transformations of V unto itself

$$\rho : G \longrightarrow Gl(V) = Gl(\mathbb{R}^K \text{ or } \mathbb{C}^K)$$

then we can construct a new bundle over M^n, with fiber V and transition functions

$$\rho_{UV} : U \cap V \longrightarrow Gl(V)$$

by

$$\rho_{UV}(x) = \rho\left(c_{UV}(x)\right)$$

i.e

$$\rho_{UV} = \rho \circ c_{UV} \qquad (70.1)$$

$$\rho_{UV}(x) : V \longrightarrow V.$$

This new bundle is said to be <u>the bundle associated to the original bundle E via the representation</u> ρ.

Let $\{\omega_U\}$ be a connection for the original vector bundle E; thus for each tangent vector dx/dt to M^n $\omega_U(dx/dt) \in \mathfrak{g}$. As such $\omega_U(dx/dt)$ is a matrix acting on the original fiber $\mathbb{R}^N$ or $\mathbb{C}^N$. A connection $\{\Omega_U\}$ for the associated bundle will consist of matrices with $\Omega_U(dx/dt)$ acting now on V ($\approx \mathbb{R}^K$ or $\mathbb{C}^K$) instead of $\mathbb{R}^N$ or $\mathbb{C}^N$. But $\rho : G \longrightarrow Gl(V)$ is a differentiable map of Lie groups and $\omega_U(dx/dt) \in \mathfrak{g} = G_e$. We define

$$\Omega_U\left(\frac{dx}{dt}\right) \equiv \rho_*\left(\omega_U\left(\frac{dx}{dt}\right)\right)$$

which of course is a tangent vector to $Gl(V)$ at the identity $e \in Gl(V)$, ie $\Omega_U(dx/dt) \in \mathfrak{gl}(V)$. We abbreviate all this by

$$\Omega_U = \rho_* \omega_U \qquad (70.2)$$

<u>Proposition</u> $\{\Omega_U\}$ defines a connection for the bundle associated to E via the representation ρ.

Proof. We must show

$$\Omega_V\left(\frac{dx}{dt}\right) = \rho_{UV}^{-1}\, \Omega_U\left(\frac{dx}{dt}\right)\rho_{UV} + \rho_{UV}^{-1}\, d\rho_{UV}.$$

Now

$$\Omega_V\left(\frac{dx}{dt}\right) = \rho_*\left[\omega_V\left(\frac{dx}{dt}\right)\right] = \rho_*\left[c_{UV}^{-1}\, \omega_U\left(\frac{dx}{dt}\right) c_{UV} + c_{UV}^{-1}\, dc_{UV}\left(\frac{dx}{dt}\right)\right].$$

Look at the two pieces on the right hand side.

In this expression $\frac{dx}{dt} = \frac{dx}{dt}\big|_{t=0}$ is a given tangent vector to M^n at a given point $x \in U \cap V$. Abbreviate

$$\omega_U = \omega_U\left(\frac{dx}{dt}\big|_{t=0}\right).$$

Now $\omega_U \in \mathfrak{g}$ is the tangent vector at $e \in G$ to the 1 parameter subgroup $e^{t\omega_U}$. By the geometric meaning of the differential ρ_* of a map ρ

$$\rho_*\left(c_{UV}^{-1}\,\omega_U\,c_{UV}\right) = \frac{d}{dt}\,\rho\left[c_{UV}^{-1}\,e^{t\omega_U}\,c_{UV}\right]\Big|_{t=0}$$

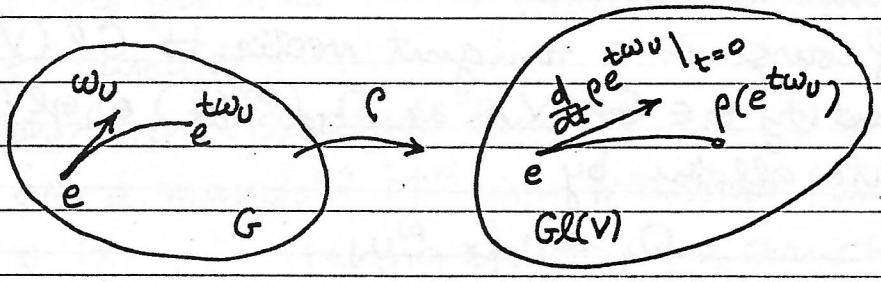
$$= \frac{d}{dt}\left[\rho(c_{UV}^{-1})\,\rho(e^{t\omega_U})\,\rho(c_{UV})\right]\Big|_{t=0}$$

(since ρ is a homomorphism)

$$= \rho^{-1}(c_{UV})\,\frac{d}{dt}\,\rho(e^{t\omega_U})\Big|_{t=0}\,\rho(c_{UV})$$

(since we are at a specific point x, i.e. c_{UV} is constant)

$$= \rho^{-1}(c_{UV})\,\rho_*(\omega_U)\,\rho(c_{UV})$$



and so

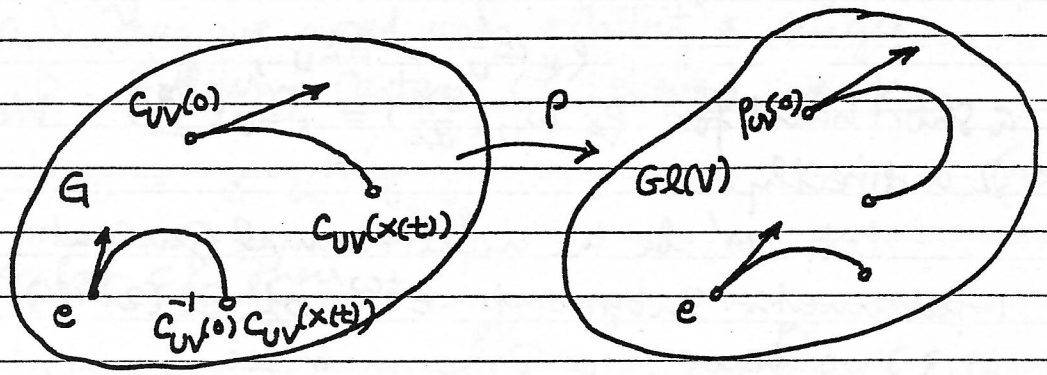
$$\rho_*\left(c_{UV}^{-1}\,\omega_U\,c_{UV}\right) = \rho_{UV}^{-1}\,\Omega_U\,\rho_{UV} \qquad (70.3)$$

We now must consider the 2nd term

$$\rho_*\left(c_{UV}^{-1}\,dc_{UV}\left(\frac{dx}{dt}\right)\right)$$

Consider a curve $x = x(t)$ on M^n with the given $\frac{dx}{dt}$ at $t=0$. We then have a curve in the Lie group G

$$c_{UV}^{-1}(0)\; c_{UV}(x(t))$$



$$\rho_*\left(c_{UV}^{-1}\, dc_{UV}\left(\frac{dx}{dt}\right)\right) = \frac{d}{dt}\,\rho\left[c_{UV}^{-1}(0)\, c_{UV}(x(t))\right]_{t=0}$$

$$= \rho_{UV}^{-1}(0)\cdot \frac{d}{dt}\,\rho_{UV}(x(t))\Big|_{t=0}$$

$$= \rho_{UV}^{-1}\; d\rho_{UV}\left(\frac{dx}{dt}\right)\Big|_{t=0}$$

and so

$$\rho_*\left[c_{UV}^{-1}\, dc_{UV}\left(\frac{dx}{dt}\right)\right] = \rho_{UV}^{-1}\, d\rho_{UV}\left(\frac{dx}{dt}\right). \qquad (70.4)$$

(10.3) & (70.4) prove our proposition.

(3) <u>The Spin Connection for the Curved M^4</u>. Let M^4 be a pseudo-Riemannian manifold with connection (for the tangent bundle TM^4) $\{\omega_U\}$. Thus $\omega_U(dx/dt) \in \mathcal{L}_4$, the Lie algebra of the Lorentz group $L_+^\uparrow$ (which is of course also the Lie algebra of the full Lorentz group L. We are assuming that M^4 has a spin structure and we want a connection for the associated $\mathbb{C}^4$ bundle (the 4 component spinor bundle) via the representation $\rho: Sl(2,\mathbb{C}) \longrightarrow Gl(4,\mathbb{C})$. Since $\Lambda: Sl(2,\mathbb{C}) \longrightarrow L_+^\uparrow$ is a 2:1 covering, to $\omega_U(dx/dt) \in \mathcal{L}_4$ there is a unique $\omega'(\frac{dx}{dt}) \in sl(2,\mathbb{C})$ such that $\Lambda_* \omega' = \omega$ (there are two vectors "above ω" but only one of them ...

is located at $e \in Sl(2,\mathbb{C})$. We need Ω such that

$$\rho_* \omega'_U = \Omega_U, \tag{70.5}$$

a shorthand for $\rho_* \omega'_U(\frac{dx}{dt}) = \Omega_U(\frac{dx}{dt})$. We shall find Ω indirectly.

Let ω' be the infinitesimal generator of the 1-parameter subgroup $e^{t\omega'}$ of $Sl(2,\mathbb{C})$. Then from (69.3) we have, for Dirac matrices

$$\gamma(\Lambda(e^{t\omega'})X) = \rho(e^{t\omega'})\,\gamma(X)\,\rho(e^{-t\omega'}) \tag{70.6}$$

where $X \in M^4_x$ is a fixed vector in the tangent space to M^4 at the point $x \in M^4$. Differential with respect to t and put $t=0$

$$\gamma((\Lambda_* \omega')X) = \rho_*(\omega')\gamma(X) - \gamma(X)\rho_*(\omega')$$

ie

$$\gamma(\omega(X)) = \Omega\,\gamma(X) - \gamma(X)\,\Omega \tag{70.7}$$

Now let $X = e^{\nu}{}_{\mu}$ be the μ^{th} basis vector of M^4_x

$$\gamma(\omega(e_\mu)) = \Omega\gamma_\mu - \gamma_\mu\Omega$$

or

$$\gamma(e_\nu\, \omega^\nu{}_\mu) = \Omega\gamma_\mu - \gamma_\mu\Omega$$

where again $\omega^\nu{}_\mu = \omega^\nu{}_\mu(\frac{dx}{dt})$, $\frac{dx}{dt}$ a fixed vector at x. Thus

$$\gamma_\nu\, \omega^\nu{}_\mu = \Omega\gamma_\mu - \gamma_\mu\Omega \tag{70.8}$$

is to be solved for $\Omega = \Omega(dx/dt)$.

This is a Clifford algebra equation and it can be shown that the solution, if it exists, is unique. (Two solutions would give us $(\Omega_1 - \Omega_2)\gamma_\mu = -\gamma_\mu(\Omega_1 - \Omega_2)$, and so $\Omega_1 - \Omega_2$ would commute with the entire Clifford

algebra and it can then be shown that the only possible $\Omega_1 - \Omega_2$ is 0). Thus we need only exhibit a single solution Ω. Following Cartan (I believe), we try

$$\Omega = c\,\omega^{\mu\nu}\gamma_\mu\gamma_\nu \tag{70.9}$$

for some scalar c. Then

$$\frac{1}{c}(\Omega\gamma_\alpha - \gamma_\alpha\Omega) = \omega^{\mu\nu}\gamma_\mu\gamma_\nu\gamma_\alpha - \omega^{\mu\nu}\gamma_\alpha\gamma_\mu\gamma_\nu$$

$$= \omega^{\mu\nu}\gamma_\mu(-\gamma_\alpha\gamma_\nu + 2g_{\nu\alpha}I) - \omega^{\mu\nu}\gamma_\alpha\gamma_\mu\gamma_\nu$$

$$= -\omega^{\mu\nu}\gamma_\mu\gamma_\alpha\gamma_\nu + 2\omega^{\mu\nu}\gamma_\mu g_{\nu\alpha}I - \omega^{\mu\nu}\gamma_\alpha\gamma_\mu\gamma_\nu$$

$$= \omega^{\mu\nu}\gamma_\alpha\gamma_\mu\gamma_\nu - 2\omega^{\mu\nu}\gamma_\nu g_{\mu\alpha} + 2\omega^{\mu\nu}\gamma_\mu g_{\nu\alpha} - \omega^{\mu\nu}\gamma_\alpha\gamma_\mu\gamma_\nu$$

$$= -2\omega^{\mu\nu}\gamma_\nu g_{\mu\alpha} + 2\omega^{\mu\nu}\gamma_\mu g_{\nu\alpha}$$

But $\omega^{\mu\nu}$ is skew in μ,ν (see (68.2)). Then

$$\frac{1}{c}(\Omega\gamma_\alpha - \gamma_\alpha\Omega) = 2(\gamma_\nu\omega^{\nu\mu}g_{\mu\alpha} + \gamma_\mu\omega^{\mu\nu}g_{\nu\alpha})$$

$$= 2(\gamma_\nu\omega^{\nu}{}_{\alpha} + \gamma_\mu\omega^{\mu}{}_{\alpha}) = 4\gamma_\nu\omega^{\nu}{}_{\alpha}.$$

But (70.8) says $\Omega\gamma_\alpha - \gamma_\alpha\Omega = \gamma_\nu\omega^{\nu}{}_{\alpha}$, and so $4c = 1$.

Spin Connection

$$\boxed{\Omega = \frac{1}{4}\omega^{\mu\nu}\gamma_\mu\gamma_\nu = \frac{1}{4}\omega_{\mu\nu}\gamma^\mu\gamma^\nu = \frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu]} \tag{70.10}$$

For intrinsic (covariant) derivative and Dirac operators

$$\boxed{\frac{D\psi}{dt} = \frac{d\psi}{dt} + \frac{1}{4}\omega_{\mu\nu}\left(\frac{dx}{dt}\right)\gamma^\mu\gamma^\nu \qquad \not{D}\psi \equiv \gamma^\alpha D_{e_\alpha}\psi = \gamma^\alpha\left[e_\alpha(\psi) + \frac{1}{4}\Gamma^\mu_{\alpha\nu}\gamma_\mu\gamma^\nu\right]} \tag{70.11}$$

Lecture 71. p Form Sections of a Vector Bundle

(1) Associated Bundles and Connections. Let us review these ideas. For concreteness we shall assume that our vector bundles are complex (as they are in quantum physics).

Let $E \xrightarrow{\pi} M^n$ be a $\mathbb{C}^K$ vector bundle over M^n. In overlaps of patches on M^n we have transition functions

$$c_{UV} : U \cap V \longrightarrow G \subset Gl(\mathbb{C}^K)$$

each $c_{UV}(x)$ is a $K \times K$ matrix. Thus in $\pi^{-1}(U \cap V)$ we are identifying

$$(x, y_V) \sim (x, y_U) \quad \text{iff} \quad y_V = c_{VU}(x)\, y_U .$$

A connection assigns a $\mathfrak{g}$-valued 1-form ω_U to each U and we have covariant differential

$$D y_U = d y_U + \omega_U y_U \tag{71.1}$$

And in an overlap

$$D y_V = c_{VU}\, D y_U \tag{71.2}$$

Now suppose that $\rho : G \longrightarrow Gl(\mathbb{C}^N)$ is some representation of our structure group G, thus each $\rho(g)$ is an $N \times N$ matrix operating on the new fiber $\mathbb{C}^N$ and ρ is a homomorphism

$$\rho(g_1 g_2) = \rho(g_1)\rho(g_2).$$

We then define a new vector bundle $E_\rho \xrightarrow{\pi} M^n$ associated to E, with fiber $\mathbb{C}^N$, by making identifications in $U \times \mathbb{C}^N$ and $V \times \mathbb{C}^N$ by

$$(x, \psi_V) \sim (x, \psi_U) \quad \text{iff} \quad \psi_V(x) = \rho(c_{UV}(x))\, \psi_U$$

and we define an associated connection for E_ρ by using as connection form in U

(71.3)

$\rho_* \omega_U$.

This is the $\mathfrak{gl}(\mathbb{C}^N)$-valued 1-form on $U \subset M$ defined by

$$(\rho_* \omega_\cdot)\left(\frac{dx}{dt}\right) = \rho_*\left(\omega\left(\frac{dx}{dt}\right)\right);$$

in particular

$$\rho_* \omega_\nu = \rho_*(\omega(e_\nu)).$$

Recall that $\rho: G \longrightarrow GL(\mathbb{C}^N)$ and so $\rho_*: \mathfrak{g} \to \mathfrak{gl}(\mathbb{C}^N)$; for any $\omega_\nu \in \mathfrak{g}$, e.g.

$$\rho_* \omega_\nu = \frac{d}{dt} \rho(e^{t\omega_\nu})\Big|_{t=0}.$$

We then define the covariant differential for our associated bundle by

$$D\psi = d\psi + (\rho_* \omega)\psi \tag{71.4}$$

and automatically

$$D\psi_V = \rho(c_{UV}) \cdot D\psi_U. \tag{71.5}$$

For covariant derivative, we have

$$D_\mu \psi = \partial_\mu \psi + (\rho_* \omega_\mu)\psi \tag{71.6}$$

If we do not suppress the indices on ψ

$$D_\mu \psi^R = \partial_\mu \psi^R + (\rho_* \omega_\mu)^R{}_S \psi^S \tag{71.7}$$

Example. The associated ad bundle. For any vector bundle E with structure group G we can consider the adjoint action of G on its own Lie algebra $\mathfrak{g}$; for $Y \in \mathfrak{g}$

$$\rho(g): \mathfrak{g} \longrightarrow \mathfrak{g} \quad \text{by} \quad \rho(g)(Y) = g\, Y\, g^{-1} \tag{71.}$$

This representation is written

$$\rho(g) = Ad(g)$$

When ρ_* is usually written ad. For $\mathcal{X} \in \mathcal{G}$

$$\mathrm{ad}(\mathcal{X})\, \mathcal{Y} = (\rho_* \mathcal{X})(\mathcal{Y}) = \frac{d}{dt} \mathrm{Ad}(e^{t\mathcal{X}})\mathcal{Y}\Big|_{t=0}$$

$$= \frac{d}{dt}\left(e^{t\mathcal{X}} \mathcal{Y}\, e^{-t\mathcal{X}}\right)\Big|_{t=0}$$

$$= \mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X} = [\mathcal{X}, \mathcal{Y}].$$

Thus if our original vector bundle has transition matrices $c_{UV}(x) \in G$, then we can construct the associated Ad-bundle with transition functions given by

$$\rho(c_{UV}(x))\, \mathcal{Y} = c_{UV}(x)\, \mathcal{Y}\, c_{UV}^{-1}(x) \qquad (71.9)$$

and associated connection forms $\rho_* \omega_\mu$

$$\rho_* \omega_\mu\ (\mathcal{Y}) = [\omega_\mu, \mathcal{Y}]. \qquad (71.10)$$

and covariant differential

$$D\psi = d\psi + [\omega, \psi] \qquad (71.11)$$

and covariant derivative

$$D_\mu \psi = \partial_\mu \psi + [\omega_\mu, \psi]. \qquad (71.12)$$

(2) p-Form Sections of a Vector Bundle. A p-form section φ of a vector bundle E over M^m is a collection of p-forms $\{\varphi_U\}$, φ_U defined on $U \subset M^m$ and having values in the fixed fiber $\mathbb{C}^K$ or $\mathbb{R}^K$ of the bundle E, such that in an overlap $U \cap V$

$$\varphi_V = c_{VU}(x)\, \varphi_U$$

ie
$$\varphi_V(\partial_{\alpha_1}, \cdots, \partial_{\alpha_p}) = c_{VU}(x)\, \varphi_U(\partial_{\alpha_1}, \cdots, \partial_{\alpha_p})$$

Each φ_U is a column of p-forms

$$\varphi_U(x) = (\varphi_U^R(x)) = \begin{pmatrix} \varphi_U^1(x) \\ \vdots \\ \varphi_U^K(x) \end{pmatrix}$$

$$\varphi^R = \varphi^R_{\underline{A}}\, dx^A = \varphi^R_{\underline{\alpha_1 \cdots \alpha_p}}(x)\, dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_p}$$

We define the "covariant exterior differential" of φ, $D\varphi$, to be the collection of forms (where $B = (\beta_1, \cdots, \beta_{p+1})$)

$$(D\varphi_U)_B = \delta^{\alpha \underline{A}}_B \left(\partial_\alpha \varphi_A + \omega_\alpha(\varphi_A)\right) \qquad (7.13)$$

ie
$$(D\varphi_U)^R_B = \delta^{\alpha \underline{A}}_B \left(\partial_\alpha \varphi^R_A + \omega_\alpha{}^R{}_S\, \varphi^S{}_A\right)$$

(recall that $\omega_\alpha = \omega(\partial_\alpha)$ is a matrix $\omega_\alpha{}^R{}_S \in \mathfrak{g}$).

Note that this simply says

$$(D\varphi)^R = d\varphi^R + \omega^R{}_S \wedge \varphi^S \qquad (7.14)$$

It follows, as usual, that $D\varphi$ is a $(p+1)$ form section of E, ie $D\varphi$ is "covariant"

$$D\varphi_V = c_{VU}\, D\varphi_U$$

We discussed in §1 the Ad bundle associated to the vector bundle E by means of the adjoint representation $\mathrm{Ad}\, g : \mathfrak{g} \to \mathfrak{g}$ by $\mathrm{Ad}\, g\,(y) = g\, y\, g^{-1}$. It has transition functions $\mathrm{Ad}\,(c_{UV}(x))$, i.e

$$\mathrm{Ad}\, c_{VU}(x)\ (y)_U \;=\; c_{VU}(x)\, y\, c_{VU}^{-1}(x)$$

Now we know that the local curvature forms $\vartheta_U = d\omega_U + \frac{1}{2}[\omega_U, \omega_U]$ of the vector bundle E are $\mathfrak{g}$ valued 2-forms that transform by

$$\vartheta_V = c_{UV}^{-1}\, \vartheta_U\, c_{UV} = c_{VU}\, \vartheta_U\, c_{VU}^{-1}$$

Thus the collection of local forms $\{\vartheta_U\}$ fit together to give a global 2-form section of the Ad bundle!

Let us compute the covariant exterior differential of this curvature ϑ by our above procedure. Let $A = (\alpha_1, \alpha_2)$, $B = (\beta_1, \beta_2, \beta_3)$. $\vartheta_{\alpha_1\alpha_2} = \vartheta_A \in \mathfrak{g}$. Then

$$\begin{aligned}
(D\vartheta)_B &= \delta^{\mu A}_{\ B} \left(\partial_\mu \vartheta_A + (\mathrm{Ad})_* \omega_\mu\ (\vartheta_A) \right) \\
&= (\text{see (71.10)}) \\
&= \delta^{\mu A}_{\ B} \left(\partial_\mu \vartheta_A + [\omega_\mu, \vartheta_A] \right) \\
&= (\text{see (63.6)}) \\
&= (d\vartheta)_B + [\omega, \vartheta]_B
\end{aligned}$$

$$\text{i.e.} \quad \left.\begin{aligned} D\vartheta &= d\vartheta + [\omega, \vartheta] \\ &= d\vartheta + \omega\wedge\vartheta - (-1)^2 \vartheta\wedge\omega \end{aligned}\right\} \qquad (71.15)$$

and since $\vartheta = d\omega + \frac{1}{2}[\omega,\omega] = d\omega + \omega\wedge\omega$, we know

$$\boxed{\text{Bianchi Identity, } D\vartheta = 0} \qquad (71.16)$$

In general, for any p-form section of the Ad bundle

$$\boxed{\begin{aligned} DF^p &= dF^p + [\omega, F^p] \\ &= dF^p + \omega \wedge F^p - (-1)^p F^p \wedge \omega \end{aligned}} \tag{71.16}$$

In the more usual physics notation, for a 1-form F^1

$$(DF^1)_{\alpha\beta} = \partial_\alpha F_\beta - \partial_\beta F_\alpha + [\omega_\alpha, F_\beta] - [\omega_\beta, F_\alpha] \tag{71.17}$$

and for a 2-form

$$\left.\begin{aligned} (DF^2)_{\alpha\beta\gamma} &= \delta^{\mu\rho\sigma}_{\alpha\beta\gamma} \left(\partial_\mu F_{\rho\sigma} + [\omega_\mu, F_{\rho\sigma}] \right) \\ &= \partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} \\ &\quad + [\omega_\alpha, F_{\beta\gamma}] + [\omega_\gamma, F_{\alpha\beta}] + [\omega_\beta, F_{\gamma\alpha}] \end{aligned}\right\} \tag{71}$$

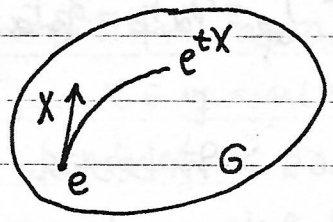
A final note; $\vartheta = d\omega + \omega \wedge \omega = d\omega + \frac{1}{2}[\omega, \omega]$ looks something like the covariant differential of the connection ω, but this is completely wrong; ω does not transform as a 1-form section of the Ad bundle (due to the term $c_{UV}^{-1} d c_{UV}$)

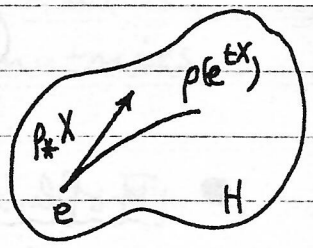
Insert for top of page 109

More Remarks; The Adjoint Group. Let $\rho: G \longrightarrow H$ be a (differentiable) homomorphism of Lie groups.

What is $\rho_*: \mathfrak{g} \longrightarrow \mathfrak{h}$?

For $X \in \mathfrak{g}$ consider the 1-parameter subgroup e^{tX} and its image $\rho(e^{tX})$. Then, as usual $\rho_*(X) = \frac{d}{dt}(\rho e^{tX})|_{t=0}$. Now $\rho(e^{tX})$ is a 1-parameter subgroup of H and so it has an infinitesimal generator and it clearly is $\rho_* X$





$$\rho(e^{tX}) = e^{t\rho_* X} \tag{11.8a}$$

Consider the adjoint group $Ad(G)$ of G. This is the group of linear transformations of $\mathfrak{g} \longrightarrow \mathfrak{g}$ given by $Ad(g)(Y) = gYg^{-1}$. Note, for example that when G is abelian (say the n-torus T^n) then $Ad(G)$ just consists of the identity, $Ad(T^n) = \{e\}$. For a second example, we have seen in Lecture 66, §2, that $Ad(SU(2)) = SO(3)$.

Look at $\rho = Ad: G \longrightarrow Ad(G)$. Then, for $X \in \mathfrak{g}$, $\rho_* X$ is the infinitesimal generator of the subgroup $Ad\, e^{tX}$. This subgroup is a 1-parameter group of transformations of $\mathfrak{g}$ and so its infinitesimal generator $\rho_* X$ is also a linear transformation of $\mathfrak{g} \longrightarrow \mathfrak{g}$. What is it? For $Y \in \mathfrak{g}$

$$\rho_*(X)(Y) = \frac{d}{dt}(e^{tX} Y e^{-tX})|_{t=0} = [X, Y] \tag{11.8b}$$

The linear transformation $Y \longrightarrow [X,Y]$ is named "$ad\, X$"

$$Ad\, e^{tX} = e^{t\, ad\, X} \qquad ad\, X = Ad_*\, X \tag{11.8c}$$

We can even write

$$\begin{aligned} Ad\, e^{tX}(Y) &= \left(I + t\, ad\, X + \frac{t^2}{2!} ad\, X\, ad\, X + \dots\right) Y \\ &= Y + t[X,Y] + \frac{t^2}{2!}[X,[X,Y]] + \dots \end{aligned} \tag{11.8d}$$

Correction. Equation 71.17 should read

$$(DF')_{\alpha\beta} = \partial_\alpha F_\beta - \partial_\beta F_\alpha + [\omega_\alpha, F_\beta] - [\omega_\beta, F_\alpha] \qquad (71.17)$$

- Problem 105. In Lecture 63 we considered brackets of $\mathfrak{g}$-valued exterior forms. There is generalized Jacobi identity satisfied. It is (you needn't prove this), for $\alpha^p, \beta^q, \gamma^r$

$$(-1)^{pr}[[\alpha,\beta],\gamma] + (-1)^{qr}[[\beta,\alpha],\beta] + (-1)^{pq}[[\beta,\gamma],\alpha] = 0 \qquad (71.19)$$

Use this and other properties of d and $[,]$ to give a direct proof of the Bianchi identity for curvature. That is, use the Lie algebra version

$$\vartheta = d\omega + \tfrac{1}{2}[\omega,\omega]$$

to show $D\vartheta = d\vartheta + [\omega, \vartheta] = 0$

without interpreting them as matrices (ie, don't use $d\omega + \omega \wedge \omega$, etc.).

Lecture 72. Yang-Mills Fields, I

(1) Invariant scalar Product in $\mathfrak{g}$. Let $G = U(n)$, the group of $n \times n$ unitary matrices, $\bar{g}^T = g^{-1}$. Then $\mathfrak{g} = u(n)$ is the vector space of skew hermitian $n \times n$ matrices

$$\bar{X}^T = -X, \quad X \in \mathfrak{g} = u(n).$$

Define a (real) scalar product $\langle , \rangle$ in $\mathfrak{g}$ by

$$\langle X, Y \rangle_{\mathfrak{g}} \equiv -\operatorname{tr} XY = -X_{ij} Y_{ji} \tag{72.1}$$

- Problem 106 Show that this form on $u(n)$ is ① real, ② symmetric and ③ positive definite, and so does define a scalar product in $u(n)$

Note: The Lie algebra $\mathfrak{g}$ of any Lie group G will be considered as a real vector space

Note also that this scalar product is invariant under the adjoint action of $G = U(n)$ on $\mathfrak{g}$;

$$u \in U(n) \Rightarrow \langle uXu^{-1}, uYu^{-1} \rangle_{\mathfrak{g}}$$
$$= -\operatorname{tr}(uXu^{-1}uYu^{-1}) = -\operatorname{tr} XY = \langle X, Y \rangle_{\mathfrak{g}}$$

The group $U(n)$ is a compact group. To see this we consider $U(n)$ as the subset of $\mathbb{C}^{n^2}$ satisfying $u\bar{u}^T = I$, ie

$$\sum_j u_{ij} \bar{u}_{kj} = \delta_{ik}$$

ie, in particular

$$\sum_{j,k} u_{kj} \bar{u}_{kj} = \sum_k \delta_{kk} = n.$$

(but not all such)

Thus $U(n)$ consists of points of $\mathbb{C}^{n^2}$ that lie on the sphere of radius $\sqrt{n}$ (recall u_{kj} are coordinates for $\mathbb{C}^{n^2}$). Thus $U(n)$ is a bounded subset of $\mathbb{C}^{n^2}$. It is also clear that a limit of a sequence of unitary matrices is again unitary; thus $U(n)$ is a closed, bounded subset of $\mathbb{C}^{n^2}$ and so is compact.

Consider now a Lie group G acting as a group of linear transformations on some $\mathbb{C}^n$ (as matrices). For example $SO(n)$ acts on $\mathbb{R}^n$ so of course it can be considered as acting on $\mathbb{C}^n$ as well. We claim that we can introduce some hermitian scalar product $\langle , \rangle$ in $\mathbb{C}^n$ so that this scalar product is invariant under the group action of G on $\mathbb{C}^n$

$$\langle gX, gY \rangle = \langle X, Y \rangle \quad \text{all } g \in G. \tag{72.2}$$

We do this by taking any hermitian scalar product $\ll , \gg$ in $\mathbb{C}^n$ and "averaging over the group G", as follows. Let ω be a right invariant Haar measure on G (Lecture 55, § 1); thus if $\tau^1, \dots, \tau^N$ ($N = \dim G$) is a basis for the right invariant 1-forms on G

$$\omega = \tau^1 \wedge \tau^2 \wedge \cdots \wedge \tau^N.$$

We let ω_g be ω at the point $g \in G$. We may also "normalize" ω so that

$$\int_G \omega = 1$$

(since G is compact, $\therefore$ bounded, $\int_G \omega$ is finite.)

We then define a new scalar product in $\mathbb{C}^n$ by

$$\langle X, Y \rangle \equiv \int_G \ll g'X, g'Y \gg \omega_{g'} \tag{72.3}$$

$X \in \mathbb{C}^n$, $Y \in \mathbb{C}^n$. This is of the form

$$\langle X, Y \rangle = \int_G f(g') \omega_{g'}$$

We claim that this new $\langle , \rangle$ is invariant under G. Note

$$\langle gX, gY\rangle = \int_G \langle\langle g'gX, g'gY\rangle\rangle\, \omega_{g'}$$

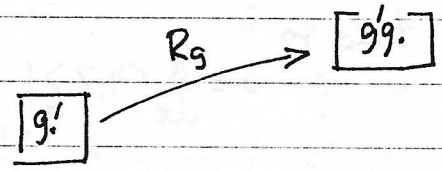
$$= \int_G f(g'g)\, \omega_{g'}$$

Formally, we proceed as follows; the right invariance of ω is formally written $\omega_{g'} = \omega_{g'g}$, all $g \in G$. Then

$$\langle gX, gY\rangle = \int_G f(g'g)\omega_{g'} = \int_G f(g'g)\omega_{g'g}$$

$$= \int_G f(g'')\omega_{g''} = \langle X, Y\rangle.$$

In more precise terms, right invariance of ω says



$$R_g^*\, \omega_{g'g} = \omega_{g'}$$

and so, putting $g'' = g'g$

$$\langle gX, gY\rangle = \int_G f(g'g)\, \omega_{g'} = \int_G (f \circ R_g)(g')\, R_g^*\, \omega_{g'g}$$

$$= \int_G R_g^* \{ f(g'')\, \omega_{g''} \}$$

$$= \int_{R_g(G)=G} f(g'')\, \omega_{g''} = \langle X, Y\rangle$$

and (72.2) is proven.

Now since $\langle\ ,\ \rangle$ is a hermitian scalar product (72.2) says that each $g \in G$ is a unitary transformation of $\mathbb{C}^n$ in this new metric. In this way we may always assume that a compact group G of matrices is a subgroup of the unitary group!

Let now G be a compact group of matrices, a subgroup of $U(n)$. We have seen that there is a scalar product in $u(n)$

$$\langle X, Y \rangle_{\mathfrak{g}} = -\operatorname{tr} XY$$

that is invariant under the adjoint action of $U(n)$. Clearly then $\langle\ ,\ \rangle$ defines a scalar product in $\mathfrak{g} \subset u(n)$ that is invariant under $\operatorname{Ad} G$.

Then, for X, Y, Z in $\mathfrak{g}$

$$\langle e^{tX} Y e^{-tX},\ e^{tX} Z e^{-tX} \rangle_{\mathfrak{g}} = \langle Y, Z \rangle_{\mathfrak{g}}$$

and so

$$0 = \frac{d}{dt} \langle Y, Z \rangle_{\mathfrak{g}} \Big|_{t=0} = \langle [X, Y], Z \rangle_{\mathfrak{g}} + \langle Y, [X, Z] \rangle_{\mathfrak{g}}$$

$$\langle \operatorname{ad} X(Y), Z \rangle_{\mathfrak{g}} + \langle Y, \operatorname{ad} X(Z) \rangle_{\mathfrak{g}} = 0 \qquad (72.4)$$

(2) <u>The Yang-Mills Action</u> Let $E \longrightarrow M^n$ be a vector bundle with <u>compact</u> group G as structure group. We are mainly concerned with $M^n = M^4$ being space-time and in the original Yang-Mills Nucleon $G = SU(2)$, see Lecture 65. If ω is a connection in E, then

$$\omega = iq A \qquad (72.5)$$

where A is called the "<u>gauge field</u>" or "<u>gauge potential</u>" and q is a coupling constant (generalized "charge"). Since $G \subset U(n)$, $\omega_\mu = \omega(\partial_\mu)$ is skew hermitian and A_μ takes its values in the hermitian matrices. For curvature (See (65.11))

$$\left.\begin{aligned} \vartheta &= d\omega + \tfrac{1}{2}[\omega, \omega] = iq F \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + iq [A_\mu, A_\nu] \end{aligned}\right\} \qquad (72.6)$$

and F is called the "<u>field strength</u>". It also is hermitian.

In all that follows we should use the physical fields A and F, and the physics student should do this; however the mathematics is slightly simpler if we avoid the q's and the i's by using instead the connection ω and the curvature ϑ. We can put in iq later!

We define the <u>action functional</u> for the Y-M field by

$$S = \frac{1}{4}\int_M -\mathrm{tr}\,(\vartheta_{\mu\nu}\vartheta^{\mu\nu})\,\mathrm{vol}^n \qquad (72.7)$$

Note that each $\vartheta_{\mu\nu} = R^S{}_{T\mu\nu}$ is a skew-hermitian matrix, i.e

$$\vartheta_{\mu\nu} \in \mathcal{G}$$

and $-\mathrm{tr}\,(\vartheta_{\mu\nu}\vartheta^{\mu\nu})$ is the scalar product in $\mathcal{G}$ of these matrices. The indices in $\vartheta^{\mu\nu}$ have been raised <u>by the metric tensor of</u> M^n, i.e. (g_{ij}). Note that we could write

$$S = -\frac{1}{4}\int_{M^n} \mathrm{tr}\,(\vartheta\wedge *\vartheta) = -\frac{1}{4}\int_{M^n} \langle\vartheta,\vartheta\rangle\,\mathrm{vol}^n \qquad (72.8)$$

where we are thinking of ϑ as a matrix valued 2-form and of course $*$ is the Hodge operator (see (48.4)). We shall be concerned with S as a functional of the connection ω (i.e. the gauge potential A) and we shall consider the integrals in (72.7), (72.8) as defining global (Hilbert-space) scalar products

$$S(\omega) = \tfrac{1}{2}(\vartheta,\vartheta). \qquad (72.9)$$

How does S depend on the connection? Take a 1-parameter family of connections $\omega = \omega(\varepsilon)$ with "velocity" $\dot\omega = \dot\omega(0)$

For first variation (remember, metric in M^m is fixed)

$$\dot{S}(\omega) = \frac{d}{d\varepsilon} S[\omega(\varepsilon)]\Big|_{\varepsilon=0} = \tfrac{1}{2}(\vartheta, \vartheta)^{\cdot}$$

$$= (\dot{\vartheta}, , \vartheta) = \left((d\omega + \tfrac{1}{2}[\omega,\omega])^{\cdot}, \vartheta \right)$$

$$= \left(d\dot{\omega} + \tfrac{1}{2}[\dot{\omega},\omega] + \tfrac{1}{2}[\omega,\dot{\omega}], \vartheta \right)$$

$$= \left(d\dot{\omega} + [\omega,\dot{\omega}], \vartheta \right) \qquad (72.10)$$

since ω and $\dot{\omega}$ are 1-forms. Now if ω and ω' are connections, their difference is a $\mathfrak{g}$-valued 1-form $\Delta\omega$ that transforms as

$$(\Delta\omega)_V = c_{VU}\,(\Delta\omega)\,c_{VU}^{-1}$$

and so is a 1-form section of the Ad bundle associated to the G bundle E (see lecture 71, §1). Likewise $\dot{\omega}$ is a 1-form section of the Ad bundle. Of course ϑ is a 2-form section of this same bundle. But then

$$d\dot{\omega} + [\omega,\dot{\omega}] = D\dot{\omega}$$

is the covariant exterior differential of $\dot{\omega}$ (see (71.11)) and (72.10) gives

$$\left.\begin{aligned} \dot{S}(\omega) &= (D\dot{\omega}, \vartheta) \\ &= (\dot{\omega}, D^*\vartheta) \end{aligned}\right\} \qquad (72.11)$$

where D^* is the Hilbert-space adjoint of D.

As usual we demand that $\dot{S}(\omega) = 0$ for all variations $\dot{\omega}$. This leads, as usual, to the equation

$$\boxed{\begin{aligned} &D^*\vartheta = 0 \quad \text{(Yang-Mills eqn).} \\ &D\vartheta = 0 \quad \text{(Bianchi identity)} \end{aligned}} \qquad (72.12)$$

Also

The coordinate expressions for these appear in the next lecture.

Insert, p.117

As pointed out by Lee Lin, the "precise proof" of $\langle gX, gY\rangle = \langle X, Y\rangle$ is incorrect. Jack Smith gave the following

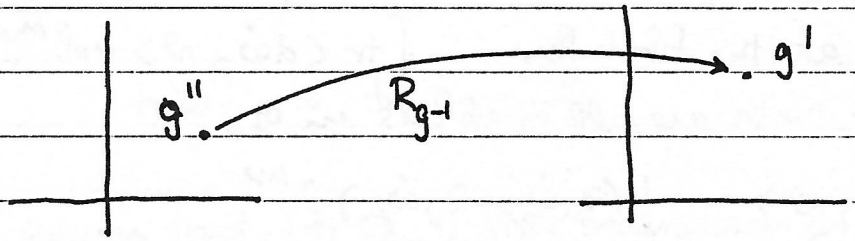
$$\langle X, Y\rangle = \int_G \langle\langle g'X, g'Y\rangle\rangle\, \omega_{g'}$$

$$= \int_G F(g')\, \omega_{g'}$$

$$\langle gX, gY\rangle = \int_G \langle\langle g'gX, g'gY\rangle\rangle\, \omega_{g'}$$

$$= \int_G F(g'g)\, \omega_{g'} = \int H(g')\, \omega_{g'}$$

where $H(g') = F(g'g)$



Put $g' = g''g^{-1} = R_{g^{-1}}(g'')$, $g'' = g'g$. Then

$$\int_G H(g')\, \omega_{g'} = \int_{R_{g^{-1}}(G)} H(g')\, \omega_{g'}$$

$$= \int_G R_{g^{-1}}^{*} \{ H(g')\omega_{g'} \}$$

$$= \int_G H(g''g^{-1})\, R_{g^{-1}}^{*}\, \omega_{g'}$$

$$= \int_G H(g''g^{-1})\, \omega_{g''} = \int_G F(g''g^{-1}g)\, \omega_{g''}$$

$$= \int_G F(g'')\, \omega_{g''} = \langle X, Y\rangle.$$

Lecture 73. Yang-Mills Fields, II

(1) The Yang-Mills Equations. We have seen that the Yang-Mills curvature ϑ ($= iqF$) must satisfy

$$D^* \vartheta = 0$$

where D^* is the Hilbert space adjoint of the covariant differential D for the Ad bundle. We must compute the coordinate expression for D^* (we did the case of ordinary exterior forms last quarter; see (48.11). D^* must satisfy

$$\int \mathrm{tr} \langle d\dot{\omega} + [\omega, \dot{\omega}], \vartheta \rangle \, \mathrm{vol} = \int \mathrm{tr} \langle \dot{\omega}, D^* \vartheta \rangle \, \mathrm{vol} \qquad (73.1)$$

for all 1-form sections $\dot{\omega}$ and all 2 form sections ϑ of the Ad bundle. $\langle \, , \rangle$ is the Riemannian pointwise scalar product

Look at the first term $\int \mathrm{tr} \langle d\dot{\omega}, \vartheta \rangle \mathrm{vol}^M$. Note (remember these are all matrices in $\mathfrak{g}$)

$$\langle d\dot{\omega}, \vartheta \rangle = \tfrac{1}{2} (\partial_\mu \dot{\omega}_\nu - \partial_\nu \dot{\omega}_\mu) \vartheta^{\mu\nu}$$

$$= \tfrac{1}{2} (\dot{\omega}_{\nu|\mu} - \dot{\omega}_{\mu|\nu}) \vartheta^{\mu\nu}$$

where $\dot{\omega}_{\nu|\mu} = \partial_\mu \dot{\omega}_\nu - \dot{\omega}_\alpha \Gamma^\alpha_{\mu\nu}$ is the covariant derivative in the Riemannian metric of M^m, not the bundle covariant derivative. With all indices showing

$$(\dot{\omega}^R_{\nu s})_{|\mu} = \partial_\mu \dot{\omega}^R_{\nu s} - \dot{\omega}^R_{\alpha s} \Gamma^\alpha_{\mu\nu},$$

nothing happens to the $\mathfrak{g}$ indices.

Then, as in getting (48.11)

$$\langle d\dot{\omega}, \vartheta \rangle = \tfrac{1}{2} \{ (\dot{\omega}_\nu \vartheta^{\mu\nu})_{|\mu} - (\dot{\omega}_\mu \vartheta^{\mu\nu})_{|\nu} - \dot{\omega}_\nu \vartheta^{\mu\nu}{}_{|\mu} + \dot{\omega}_\mu \vartheta^{\mu\nu}{}_{|\nu} \} \qquad (73.2)$$

Now $\dot{\omega}$ transforms by $\dot{\omega}_V = c_{VU} \dot{\omega}_U c^{-1}_{VU}$, and so if X is a tangent vector to M in $U \cap V$

$$\dot{\omega}_V(X) = c_{VU} \dot{\omega}_U(X) c^{-1}_{VU} \qquad (73.3)$$

and

$$\mathrm{tr}\,\dot{\omega}_V(X) = \mathrm{tr}\,\dot{\omega}_U(X)$$

and so

$$\mathrm{tr}\dot{\omega} = \{\mathrm{tr}\dot{\omega}_U\}$$

is a well defined, ordinary, global exterior 1-form on M^n.

Note that (73.3) says

$$\dot{\omega}_V{}^\alpha = \dot{\omega}_V\left(\frac{\partial}{\partial x_V^\alpha}\right) = c_{VU}\,\dot{\omega}_U\left(\frac{\partial}{\partial x_V^\alpha}\right)c_{VU}^{-1}$$

$$= c_{VU}\,\dot{\omega}_U\left(\frac{\partial}{\partial x_U^\beta}\right)c_{VU}^{-1}\,\frac{\partial x_U^\beta}{\partial x_V^\alpha}$$

or

$$\dot{\omega}_V{}^\alpha = c_{VU}\,\dot{\omega}_U{}^\beta\,c_{VU}^{-1}\,\frac{\partial x_U^\beta}{\partial x_V^\alpha} \tag{73.4}$$

and then

$$\mathrm{tr}\,\dot{\omega}_V{}^\alpha = \mathrm{tr}\,\dot{\omega}_U{}^\beta\,\frac{\partial x_U^\beta}{\partial x_V^\alpha}$$

showing again that $(\mathrm{tr}\,\dot{\omega}_\alpha)$ does transform as an ordinary covector on M^n.

Likewise, for the curvature form, $\vartheta_V{}^{\mu\nu} = c_{VU}\,\vartheta_U{}^{\alpha\beta}\,c_{VU}^{-1}\,\frac{\partial x_U^\alpha}{\partial x_V^\mu}\,\frac{\partial x_U^\beta}{\partial x_V^\nu}$

$$(\mathrm{tr}\,\vartheta_V{}^{\mu\nu}) = (\mathrm{tr}\,\vartheta_U{}^{\alpha\beta})\,\frac{\partial x_U^\alpha}{\partial x_V^\mu}\,\frac{\partial x_U^\beta}{\partial x_V^\nu}$$

shows that $(\mathrm{tr}\,\vartheta^{\mu\nu})$ defines an ordinary skew symmetric twice covariant. Furthermore, then

$$\mathrm{tr}\left(\dot{\omega}_{V\nu}\,\vartheta_V^{\mu\nu}\right) = \mathrm{tr}\left(\dot{\omega}_{U\alpha}\,\vartheta_U^{\beta\alpha}\right)\cdot\frac{\partial x_V^\mu}{\partial x_U^\beta}$$

defines a global contravariant vector $\mathrm{tr}\,\dot{\omega}_\nu\vartheta^{\mu\nu}$ etc, etc.

Now back to (73.2). Take trace of both sides. The first term

$$\tfrac{1}{2}\,\mathrm{tr}\left(\dot{\omega}_\nu\vartheta^{\mu\nu}\right)_{|\mu}$$

is the divergence of a globally defined vector field; likewise

for the 2nd term. Integrate over M^n and apply the divergence theorem to the first two terms. These terms will vanish under various conditions

(i) M^n is compact without boundary

(ii) M^n is noncompact, without boundary, and $\dot{\omega}$ vanishes outside a compact subset

(iii) $\dot{\omega}$ vanishes on ∂M

Assuming one of the above we get

$$\int_{M^n} \mathrm{tr} \langle d\dot{\omega}, \vartheta \rangle \, \mathrm{vol}^n = \frac{1}{2} \int_{M^n} \mathrm{tr} \{ -\dot{\omega}_\nu \vartheta^{\mu\nu}{}_{|\mu} + \dot{\omega}_\mu \vartheta^{\mu\nu}{}_{|\nu} \} \, \mathrm{vol}^n$$

or

$$\int_{M^n} \mathrm{tr} \langle d\dot{\omega}, \vartheta \rangle \, \mathrm{vol}^n = \int_{M^n} -\mathrm{tr} \, \dot{\omega}_\nu \vartheta^{\mu\nu}{}_{|\mu} \, \mathrm{vol}^n \qquad (73.5)$$

Now for the second term in (73.1), $\mathrm{tr} \langle [\omega, \dot{\omega}], \vartheta \rangle$.

Now $[\omega, \dot{\omega}] = [\dot{\omega}, \omega]$, and (using (63.9))

$$\langle [\dot{\omega}, \omega], \vartheta \rangle = \frac{1}{2} [\dot{\omega}, \omega]_{\mu\nu} \vartheta^{\mu\nu}$$

$$= \frac{1}{2} \{ \dot{\omega} \wedge \omega + \omega \wedge \dot{\omega} \}_{\mu\nu} \vartheta^{\mu\nu}$$

$$= \frac{1}{2} \{ \dot{\omega}_\mu \omega_\nu - \dot{\omega}_\nu \omega_\mu + \omega_\mu \dot{\omega}_\nu - \omega_\nu \dot{\omega}_\mu \} \vartheta^{\mu\nu}$$

$$= \frac{1}{2} \{ [\dot{\omega}_\mu, \omega_\nu] - [\dot{\omega}_\nu, \omega_\mu] \} \vartheta^{\mu\nu}$$

$$= [\dot{\omega}_\mu, \omega_\nu] \vartheta^{\mu\nu}.$$

Then

$$\int_M \mathrm{tr} \langle [\omega, \dot{\omega}], \vartheta \rangle \, \mathrm{vol}^n = \int_M \mathrm{tr} \, [\dot{\omega}_\mu, \omega_\nu] \vartheta^{\mu\nu} \, \mathrm{vol}^n$$

But, using the skew symmetry of ad in the of scalar product (72.4),

$$\int_M \mathrm{tr} \langle [\omega, \dot{\omega}], \vartheta \rangle \mathrm{vol}^n = -\int_M \mathrm{tr}\,[\omega_\nu, \dot{\omega}_\mu]\, \vartheta^{\mu\nu}\, \mathrm{vol}^n$$

$$= \int_{M^n} \mathrm{tr}\ \dot{\omega}_\mu\ [\omega_\nu, \vartheta^{\mu\nu}]\ \mathrm{vol}^n = -\int_{M^n} \mathrm{tr}\, \dot{\omega}_\mu [\omega_\nu, \vartheta^{\nu\mu}]\, \mathrm{vol}^n$$

(72.4)

Combining with (73.5), we have

$$\int_M \mathrm{tr} \langle D\dot{\omega}, \vartheta \rangle \mathrm{vol}^n = \int_M \mathrm{tr} \langle d\dot{\omega} + [\omega, \dot{\omega}], \vartheta \rangle \mathrm{vol}^n$$

$$= \int_M -\mathrm{tr} \left(\dot{\omega}_\nu \left\{ \vartheta^{\mu\nu}{}_{|\mu} + [\omega_\mu, \vartheta^{\mu\nu}] \right\} \right) \mathrm{vol}^n$$

$$= \int_M \mathrm{tr} \langle \dot{\omega}, D^*\vartheta \rangle \mathrm{vol}^n$$

provided we put

$$\boxed{(D^*\vartheta)^\nu = -\left\{ \vartheta^{\mu\nu}{}_{|\mu} + [\omega_\mu, \vartheta^{\mu\nu}] \right\}} \qquad (73.6)$$

In Summary, we have the Yang Mills equations

$$\left.\begin{aligned} D^*\vartheta &= 0 \\ \vartheta^{\mu\nu}{}_{|\mu} + [\omega_\mu, \vartheta^{\mu\nu}] &= 0 \end{aligned}\right\} \qquad (73.7.)$$

which supplement the Bianchi identities

$$D\vartheta = 0$$

We now must put in the physical fields A and F and we shall then compare these equations with those of electromagnetism.

(2) Yang-Mills Analogy with Electromagnetism. If we now put $\omega = igA$ and $\vartheta = igF$, as we have seen ((65.11))

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig\,[A_\mu, A_\nu]$$

generalizes the case in E&M. The Yang-Mills action is, except for a constant, in the case when no sources are present

$$S(A) \sim \int_M \mathrm{tr}\, F_{\mu\nu} F^{\mu\nu}\, \mathrm{vol}^n \tag{73.8}$$

$$= \int_M \mathrm{tr}\,(\partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu])(\partial^\mu A^\nu - \partial^\nu A^\mu + ig[A^\mu, A^\nu])\,\mathrm{vol}^n$$

While the E&M action has only the quadratic terms $(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)$ we see that Y-M also has cubic and quartic terms in the A's. The Y-M equations $D^*F = 0$ and the Bianchi identities $DF = 0$, when written out, become (see (71.18)

$$\left.\begin{aligned} &\underline{\text{Y-M:}} \quad F^{\mu\nu}{}_{|\mu} + ig[A_\mu, F^{\mu\nu}] = 0 \\ &\underline{\text{Bianchi:}}\ 0 = \partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} + \\ &\qquad + ig\{[A_\alpha, F_{\beta\gamma}] + [A_\gamma, F_{\alpha\beta}] + [A_\beta, F_{\gamma\alpha}]\} \end{aligned}\right\} \tag{73.9}$$

It is amusing to compare these with Maxwell's equations, even in M_0^4, Minkowski space. We write, for example $G = SU(n)$

$$F^{0i} = E^i, \quad F^{12} = B^3, \quad i = 1,2,3,$$

even though $\vec{E}$ is now a vector of hermitian matrices, etc. Look e.g. at Y-M for $\nu = 0$. We have

$$\left.\begin{aligned} &F^{i0}{}_{|i} + ig[A_i, F^{i0}] = 0 \\ &\mathrm{div}\,\vec{E} + ig(\vec{A}\cdot\vec{E} - \vec{E}\cdot\vec{A}) = 0 \end{aligned}\right\} \text{"Gauss' eqn"} \tag{73.10}$$

We see here that even though we started out without sources $-ig(\vec{A}\cdot\vec{E} - \vec{E}\cdot\vec{A})$ serves the role of a "charge density."

Thus the Y-M field $\vec{E}$ and the gauge potential $\vec{A}$ combine to act as a source for the Y-M field, very much unlike Gauss' law in E&M, and it is the non-abelian nature of the structure group G, ie $\vec{A}\cdot\vec{E} - \vec{E}\cdot\vec{A} \neq 0$ that is responsible for this.

Looking at Y-M again, this time considering only a spatial index $\nu = i$

$$F^{\mu i}{}_{|\mu} + ig\,[A_\mu, F^{\mu i}] = 0$$

ie

$$F^{0i}{}_{|0} + F^{ji}{}_{|j} + ig\,[A_0, F^{0i}] + ig\,[A_j, F^{ji}] = 0 \qquad (73.11)$$

$$\frac{\partial \vec{E}}{\partial t} - \text{curl}\,\vec{B} + ig\,(A_0\vec{E} - \vec{E}A_0) - ig\,(\vec{A}\times\vec{B} + \vec{B}\times\vec{A}) = 0)$$

Ampere-Maxwell

- Problem 107. Show that Bianchi (73.9) yields non abelian versions of Faraday's law and absence of monopoles.

Lecture 74. Yang-Mills with Finite Action

(1) Finite Action. Consider a quantum particle interacting with a Yang-Mills field. This particle is described by a "wave function" ψ, a cross section of a complex $\mathbb{C}^N$ vector bundle over M_0^4, Minkowski space. We assume that the structure group for this bundle is $SU(n)$; thus $G = SU(n)$ operates on $\mathbb{C}^N$ via some representation. For our purposes it is sufficient to consider the standard representation of $SU(n)$ on $\mathbb{C}^n$. This bundle has a connection $\omega = igA$ and a curvature $\vartheta = igF$, where A and F are hermitian matrix valued local forms on M_0^4. On $U \subset M_0^4$ we have a frame of sections $\Psi_U = e_U = (e_1^U, \dots, e_n^U)$, ω_U and ϑ_U. In an overlap

$$e_V = e_U c_{UV} \qquad c_{UV}(x) \in SU(n).$$

For this lecture we shall be more concerned with the background Y-M field than with the particle. The action for this Y-M field alone is essentially

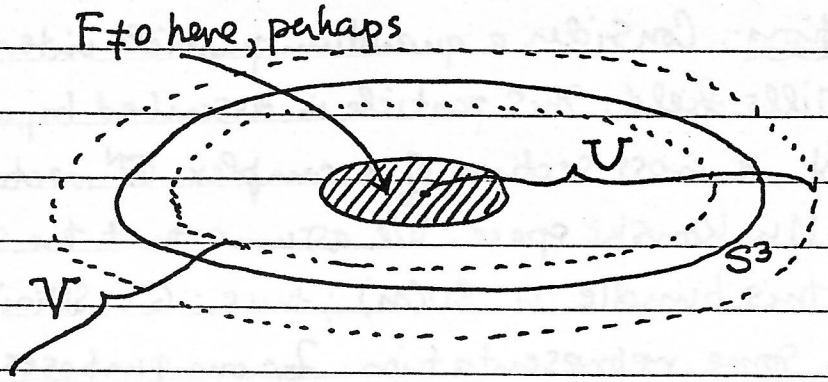
$$\int_{M_0^4} -\operatorname{tr} F_{\mu\nu} F^{\mu\nu} * 1 \sim \int_{M_0^4} \operatorname{tr}(E^2 - B^2) * 1.$$

For consideration in a Feynman integral one considers the "Euclidean" rather than Minkowski action, ie one uses the positive definite metric instead of $-dt^2 + d\vec{x} \cdot d\vec{x}$ and we shall be concerned with YM fields having finite Euclidean action

$t \to it$ transforms Minkowski to Euclidean

$$\left| \operatorname{tr} \int_{M^4 = \mathbb{R}^4} (E^2 + B^2) * 1 \right| < \infty$$

For purposes of visualization we shall in fact assume that the field strength F vanishes outside some 3-sphere S^3 centered at the origin of $\mathbb{R}^4$!



Let U be a coordinate patch holding this S^3 and its interior and let V be a coordinate patch holding S^3, extending to ∞ in $\mathbb{R}^4$, and such that $F=0$ in V. V is exterior to some sphere.

Now in V we have $\vartheta_V = 0$. We claim that we can make a change of frame in V so that in the new frame $\omega_V = 0$. This should not be a complete surprise; it is a global version of Riemann's theorem on curvature 0. To see it, let ω' be the original connection form for V. We wish $g: V \to SU(n)$ so that

$$\omega_V \equiv g^{-1}\omega' g + g^{-1} dg = 0$$

ie

$$dg + \omega' g = 0. \tag{74.1}$$

Can we solve this 1-form system for $g = g^r{}_s(x)$?

$$\begin{aligned} d(dg + \omega' g) &= d\omega' g - \omega' \wedge dg \\ &\equiv d\omega' g - \omega' \wedge (-\omega' g) = (d\omega' + \omega' \wedge \omega') g \\ &= \vartheta' g = 0, \text{ in } U. \end{aligned}$$

By Frobenius we may locally solve (74.1) for g, subject to any initial $g_0 = g(p)$ at $p \in V$.

Suppose we have two solutions g_1 and g_2 in two overlapping patches. Then $dg_1 = -\omega' g_1$ and $dg_2 = -\omega' g_2$, so

$$\begin{aligned} d(g_1^{-1} g_2) &= -g_1^{-1} dg_1 g_1^{-1} g_2 + g_1^{-1} dg_2 \\ &= g_1^{-1}\omega' g_1 g_1^{-1} g_2 - g_1^{-1}\omega' g_2 = 0, \end{aligned}$$

and so two overlapping solutions are always related by a constant $SU(n)$ matrix k; $g_2(x) = k\, g_1(x)$. This means that if we take a closed path $\mathcal{C}$ in V, starting and ending at p, if we choose $g_0 = g(p)$ and continue the solution along this path, then when we return to p g_0 has become $k g_0$, for some $k \in SU(n)$. But V is the region exterior to some sphere and so is simply connected. By the usual Cauchy argument familiar from complex variable theory (contract the curve to p) it must be that $k = e$. Thus a global solution $g : V \longrightarrow SU(n)$ to (74.1) exists and

$$\omega_V = 0$$

when we use the new frame of sections $e_V = e_U' g$. □.

Note that the original connection ω' is of the form

$$\omega' = - dg\, g^{-1} \tag{74.2}$$

and is said to be "pure gauge"

Since $\omega_V = 0$ in $U \cap V$, and in particular on S^3

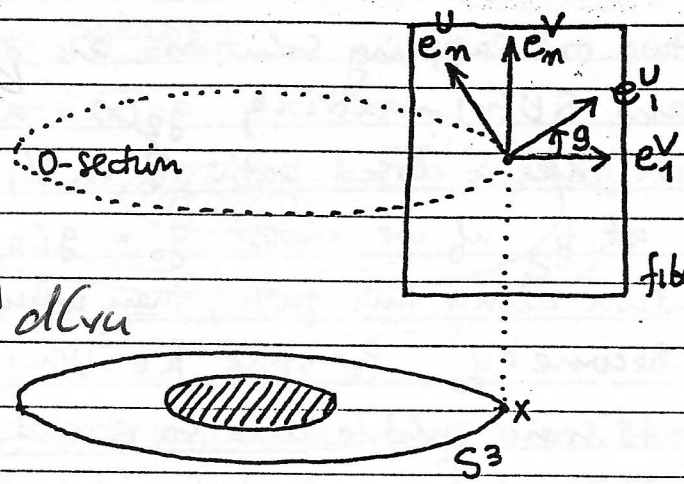
$$\omega_U = c_{VU}^{-1} \omega_V c_{VU} + c_{VU}^{-1} dc_{VU} = c_{VU}^{-1} dc_{VU}$$

which we again write in the simplified form (put $c_{VU} = g$)

$$\omega_U = g^{-1} dg \quad \text{on } S^3 \tag{74.3}$$

where $g : U \cap V \longrightarrow SU(n)$ are new matrices, not those of (74.1).

We then have the following picture. Consider the part of our wave function bundle that lies over the sphere S^3



$$\omega_U = C_{VU}^{-1} \omega_V C_{VU} + C_{VU}^{-1} dC_{VU}$$

$\omega_V \to 0$

let $C_{VU} = g$

Over the sphere S^3 we have two frames given, the "flat" frame e_V, and the frame e_U over U. The flat frame e_V consists of sections $e^V_1, \ldots, e^V_n$ each of which is covariant constant

$$\nabla e^V_R = e_S \, \omega^S_{V\,R} = 0,$$

ie these sections are parallel displaced along S^3. We are comparing the U-frames e_U with these parallel displaced frames e_V

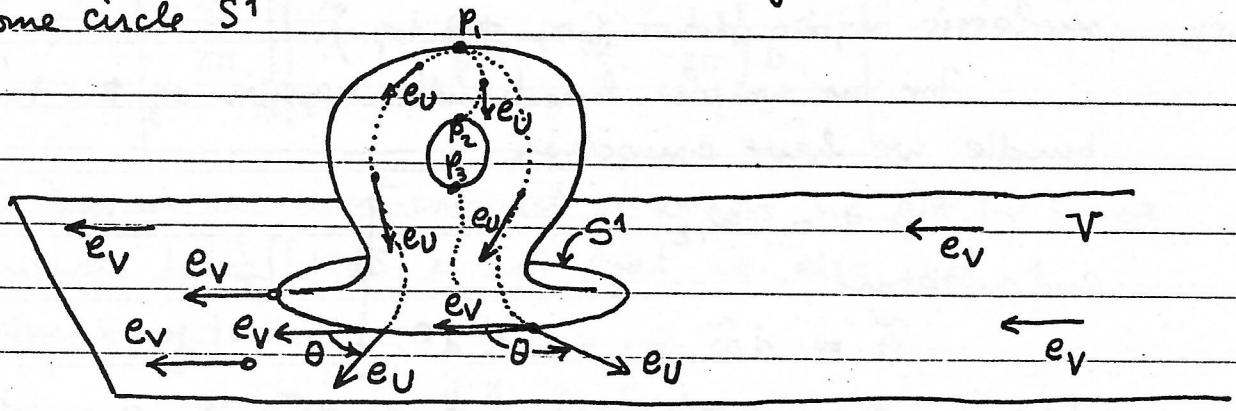
$$e_U = e_V C_{VU} = e_V g \tag{14.4}$$

and consequently we have a mapping

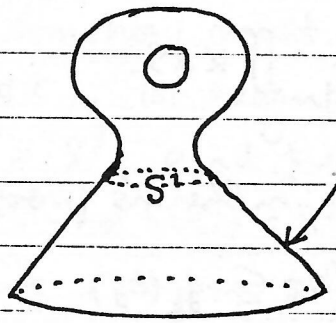
$$g : S^3 \longrightarrow SU(n). \tag{14.5}$$

This situation is similar to that encountered in Chern's proof of the Poincaré index theorem (see the figure after eqn (62.10)). Let us go back and reconsider Chern's proof in the light of our Y-M field with finite action. This will suggest, as it did to Chern, how Poincaré's theorem can be generalized to the YM case, and more!

(2) Chern's Proof Revisited Consider, instead of a closed M^2 as in Lecture 62, a curved version M^2 of the plane R^2, but such that the curvature vanishes in the region V exterior to some circle S^1



The bundle we are considering is the tangent bundle TM^2 to M^2 but considered as a complex $\mathbb{C}$ line bundle over the oriented M^2. By using "orthonormal" frames e_U, e_V, we may consider the group of this bundle to be $U(1)$. We have indicated the "flat" covariant constant frame e_V in the exterior region (Note: Unlike the case where M^n has dimension $n \geq 3$, the region V is <u>not</u> simply connected. One cannot always find a global flat frame in the region V, e.g



M^2 is flat in this V region; it is a cone. But a parallel displaced e_V will <u>not</u> return to itself after traversing S^1 (See Lect. 35, § 2).

We have chosen an example where e_V does globally exist, as in the Y-M case!)

In the curved region U we have indicated a cross section e_U that has singularities at three points p_1, p_2, p_3 (the field e_U is essentially the velocity flow (normalized) of molassis oozing down from the top)

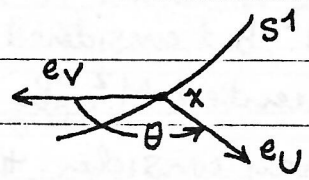
For our complex line bundle version of the tangent bundle we have connection

$$\left.\begin{aligned} \tilde{\omega} &= -i\,\omega_{12} \\ \text{and curvature} \\ \tilde{\vartheta} &= d\tilde{\omega} = -i\,K\,dS \end{aligned}\right\} \qquad (74.6)$$

where K is the Gauss curvature and dS is the area element, (see (59.16) and (59.17)).

On the circle S^1 we have

$$e_U = e_V\, e^{i\theta}$$



and so $g(x) = e^{i\theta}$

$$\left.\begin{aligned} \text{and} \\ \tilde{\omega}_U &= g^{-1}dg = e^{-i\theta}\cdot i\, e^{i\theta} d\theta = i\,d\theta \end{aligned}\right\} \qquad (74.7)$$

In the situation of Poincaré's theorem, Chern considered $\iint K\,dS$ integrated over a closed M^2. If we integrate

$$\iint_{M^2} K\,dS = \iint_U K\,dS$$

since $K=0$ outside U. In Chern's proof

$$\iint K\,dS = 2\pi \sum_p j_p(e_U)$$

while in our non-closed M^2, since $K\,dS = d\omega_{12}$, we get

$$\iint_U K\,dS = 2\pi \sum_p j_p(e_U) + \oint_{\partial U = S^1} \omega_{12}^U \qquad (74.$$

since $\partial U = S^1$. We can then write

$$\frac{i}{2\pi}\iint_{M^2}\tilde{\vartheta} = \sum_p j_p(e_U) + \frac{i}{2\pi}\oint_{S^1}\tilde{\omega}_U$$

rotation # of e_U about e_V

or

$$\boxed{\frac{i}{2\pi}\iint_{M^2}\tilde{\vartheta} = \sum_p j_p(e_U) - \frac{1}{2\pi}\oint_{S^1} d\theta} \qquad (74.9)$$

Thus in our non-compact situation we get the same answer for $\frac{1}{2\pi}\iint K\,dS$ except that we also get a boundary term

$$-\frac{1}{2\pi}\oint d\measuredangle(e_V, e_U)$$

telling us <u>how many times the given cross section</u> e_U <u>rotates around the "flat" cross section</u> e_V !

$$\frac{1}{2\pi}\iint K\,dS = \sum j_p(e_U) - \frac{1}{2\pi}\oint d\measuredangle(e_V, e_U) \qquad (74.10)$$

Note that this last "rotation number" $\frac{1}{2\pi}\oint d\measuredangle(e_V, e_U)$ is exactly the <u>degree of the map</u>

$$g: S^1 \longrightarrow S^1, \quad x \longrightarrow g(x) = e^{i\theta}$$

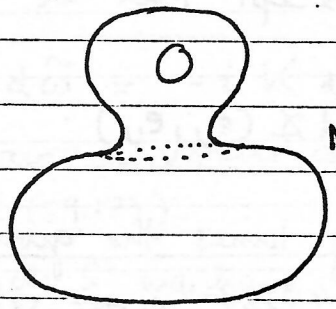
This last object, somehow must correspond to the Y-M situation in (74.5). For example, in the case $n=2$, ie $SU(2) = G = S^3$, and the map (74.5)

$$g: S^3 \longrightarrow SU(2) = S^3$$

indeed <u>does</u> have a degree. Can this Y-M degree be evaluated by looking at curvature N for the Y-M bundle? The answer, given long before YM by Chern ≈ 1945 is Yes! To understand Chern's result however

we will need to know more about the topology of Lie groups and something about Chern's "characteristic classes". We proceed with this program in the next Lecture.

Problem 108 Verify (14.10) in the case of our specific example drawn on the top of p. 131. To evaluate $\iint K dS$ I suggest that you relate this to $\iint K dS$ over the closed surface M'



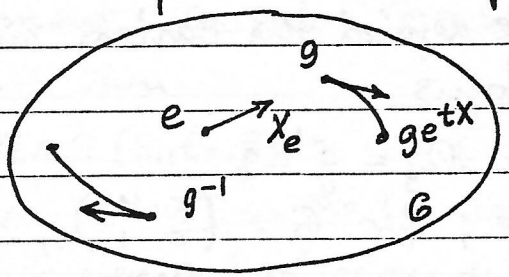
$B_R = \varepsilon_{ijk} F^{ik}$

Lecture 75. DeRham for Compact Lie Groups

(1) Bi-invariant p-forms. Let $inv : G \longrightarrow G$ be the inverse map, $inv(g) = g^{-1}$.

Proposition. If X is a left invariant vector field on G then $(inv)_* X$ is the right invariant vector field that coincides with $-X$ at e.

Proof. The left invariant vector field X generates right translations. Its integral curve thru g is $g \longmapsto g e^{tX_e}$, with velocity vector at g given by $\frac{d}{dt} g e^{tX_e}\Big|_{t=0} = g X_e$



Under the inverse map this curve is sent into the curve $g^{-1} \longrightarrow e^{-tX_e} g^{-1}$, with velocity vector at g^{-1} given by $\frac{d}{dt}(e^{-tX_e} g^{-1})\Big|_{t=0} = -X_e g^{-1}$ □.

Definition. A form or vector field is bi-invariant if it is invariant under both left and right translations.

For example, for $G = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$ $x > 0$, dx/x is bi-invariant.

Theorem If α^p is a bi-invariant p-form, then α^p is closed

$$d\alpha^p = 0$$

Proof. Let $\sigma^1, \dots, \sigma^n$ be a basis for the left invariant 1-forms on G, and let $\tau^1, \dots, \tau^n$ be the basis for right invariant 1-forms such that $\tau^i = \sigma^i$ at e. We know (Lecture 57, § 3) that the right structure constants are then negatives of the left, ie

$$d\sigma^i = -\frac{1}{2} C^i_{jk} \sigma^j \wedge \sigma^k$$

$$d\tau^i = \frac{1}{2} C^i_{jk} \tau^j \wedge \tau^k$$

Let α^p be bi-invariant. Then

$$\alpha^p = a_{\underrightarrow{i_1 \cdots i_p}} \sigma^{i_1} \wedge \cdots \wedge \sigma^{i_p}$$

where a_I are constants. Since α^p is also right invariant

$$\alpha^p = a_{\underrightarrow{i_1 \cdots i_p}} \tau^{i_1} \wedge \cdots \wedge \tau^{i_p}$$

with the same constants. Compute $d\alpha^p$ at e from both expressions. □

We have defined the Cartan matrix of left invariant 1-forms

$$\Omega_g = g^{-1} dg.$$

For example, for $G = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$,

$$\Omega_g = \begin{pmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{pmatrix}.$$

We can also consider exterior powers $\Omega^2 = \Omega \wedge \Omega$, $\Omega^3, \ldots$

For example

$$\begin{pmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{pmatrix} \wedge \begin{pmatrix} \frac{dx}{x} & \frac{dy}{x} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{dx \wedge dy}{x^2} \\ 0 & 0 \end{pmatrix}$$

We define the <u>Cartan p-forms</u> $\Omega_1, \Omega_2, \ldots \Omega_{n = \dim G}$ by

$$\boxed{\Omega_p \equiv \operatorname{tr} \underbrace{\Omega \wedge \cdots \wedge \Omega}_{p} = \operatorname{tr} g^{-1} dg \wedge \cdots \wedge g^{-1} dg} \qquad (75.1).$$

These are of course <u>left invariant</u> p forms on G. For $\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$ $\Omega_1 = \frac{dx}{x}$, $\Omega_2 = 0$

<u>Proposition</u> The Cartan p-forms Ω_p are <u>bi-invariant</u>, and hence closed. Also $\Omega_{2p} \equiv 0$

Proof. For $k \in G$

$$\operatorname{tr}\left[(gk)^{-1} d(gk) \wedge (gk)^{-1} d(gk) \wedge \cdots \wedge (gk)^{-1} d(gk)\right]$$

$$= \operatorname{tr}\left[k^{-1} \left(g^{-1} dg \wedge \cdots \wedge g^{-1} dg\right) k\right] = \Omega_p$$

and so Ω_p is right invariant, hence bi-invariant.

Next note $\Omega_2 = \mathrm{tr}\,\Omega\wedge\Omega = \Omega_{ij}\wedge\Omega_{ji} = -\Omega_{ji}\wedge\Omega_{ij} = -\Omega_2$.
Similarly $\Omega_{2p} = 0$, all p. □

The Cartan 3-form plays an especially important role. Noting that $\Omega(X) = X$ for any $X \in \mathfrak{g} = G_e$ see page 58

$$\Omega\wedge\Omega\ (X,Y) = \Omega(X)\Omega(Y) - \Omega(Y)\Omega(X)$$
$$= [X,Y].$$
$$(\Omega\wedge\Omega)\wedge\Omega\ (X,Y,Z) = [X,Y]Z + [Z,X]Y + [Y,Z]X \qquad (75.2)$$

Pause for a moment; consider $G = SO(3)$, and let E_1, E_2, E_3 be the basis (66.1). Then

$$\Omega\wedge\Omega\wedge\Omega\,(E_1,E_2,E_3) = (E_1)^2 + (E_2)^2 + (E_3)^2 \qquad (75.3)$$

and this matrix is not in the Lie algebra $so(3)$! It is called a "Casimir" element. (See our discussion preceding (66.4)).

Back to our general result (75.2)

$$\Omega_3(X,Y,Z) = \mathrm{tr}\left([X,Y]Z + [Z,X]Y + [Y,Z]X\right)$$
$$= 3\,\mathrm{tr}[X,Y]Z \qquad (75.4)$$

and if we use our Lie algebra scalar product $\langle X,Y\rangle_{\mathfrak{g}} = -\mathrm{tr}XY$

$$\boxed{\Omega_3(X,Y,Z) = -3\langle [X,Y], Z\rangle_{\mathfrak{g}}} \qquad (75.5)$$

(2) Bi-invariant Riemannian Metrics on Compact Groups. Let

$$\langle\ ,\ \rangle_{\mathfrak{g}}$$

be a scalar product in $\mathfrak{g}$ that is invariant under the adjoint action of G. For example, if $G = U(n), SU(n), O(n), SO(n)$ we can put $\langle X,Y\rangle_{\mathfrak{g}} = -\mathrm{tr}XY$. For any compact group G we can take any real scalar product $\ll\ ,\gg$ in $\mathfrak{g}$ and average

$$\langle X,Y\rangle_{\mathfrak{g}} \equiv \int_G \ll g'Xg'^{-1}, g'Yg'^{-1}\gg \omega_{g'}$$

using the right Haar measure $\omega_{g'}$ (check that this works)
We then define a Riemannian metric for G by

$$\langle X_g, Y_g \rangle \equiv \langle (l_{g^{-1}})_* X_g, (l_{g^{-1}})_* Y_g \rangle_{0_f} = \langle g^{-1} X_g, g^{-1} Y_g \rangle_{0_f}$$

and it is clear that this metric on G is <u>left</u> invariant by construction. We claim that it is also <u>right</u> invariant, hence

$$\langle\ ,\ \rangle \text{ is bi-invariant.}$$

This is clear; $\langle X_e g^{-1}, Y_e g^{-1} \rangle = \langle g X_e g^{-1}, g Y_e g^{-1} \rangle = \langle X_e, Y_e \rangle$ □.
We have shown

> On every compact Lie group there is a bi-invariant Riemannian metric (75.6)

- <u>Problem 109</u>. Show that for any scalar product $\langle\langle\ ,\ \rangle\rangle_{0_f}$ invariant under the adjoint action of G

$$\langle\langle [X, Y], Z \rangle\rangle$$

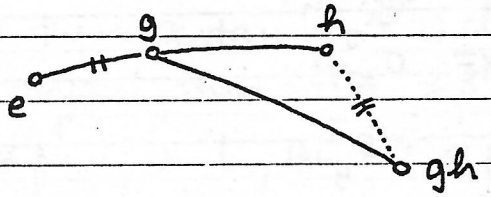
is an exterior 3-form (i.e skew in X, Y, Z). This is except for a factor, the Cartan 3-form.

> <u>Theorem</u>. If G carries a bi-invariant metric, then the geodesics thru e are the 1-parameter subgroups (75.7)

Proof. We know that if $\mathcal{C}$ is a 1-parameter subgroup and if $g \in \mathcal{C}$, then $g\mathcal{C}$ is pointwise the same curve, ie 1-parameter subgroups are invariant under left translation by elements of the subgroup. We shall show that geodesics γ thru e have this same property, and since their velocity vectors shall then be left invariant in the above sense they must coincide with the 1-parameter subgroups.

The sketch we give below is due (as far as I know) to H. Samelson. We shall use the fact (unproven by us) that if two points in a Riemannian manifold are sufficiently close, then there is a unique curve of minimal length joining them, and it is an unbroken geodesic.

Consider a small geodesic segment $\widehat{e,g,h}$ thru $e \in G$



and suppose that the left translate of this, $\widehat{g,gh}$ does not lie along the same curve. For lengths, $\ell(\widehat{e,g,h}) = \ell(\widehat{g,gh})$ by left invariance of the metric. But by right invariance $\ell(\widehat{h,gh}) = \ell(\widehat{e,g})$. Thus

$$\ell(\widehat{g,h}) + \ell(\widehat{h,gh}) = \ell(\widehat{g,h}) + \ell(\widehat{e,g}) = \ell(\widehat{e,g,h})$$
$$= \ell(\widehat{g,gh})$$

This would show that the small geodesic $\widehat{g,gh}$ joining g to gh would have the same length as the broken geodesic $(\widehat{g,h}) + (\widehat{h,gh})$, a contradiction. □.

It is a theorem of Hilbert that any pair of points on a compact Riemannian manifold can be joined by a (minimizing) geodesic. We conclude

> Every point $g \in G$ of a compact Lie group G can be joined to e by a 1-parameter subgroup. (7.5.8)

(3) Harmonic Forms in the Bi-invariant Metric.

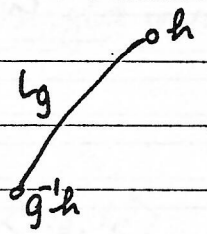
Let $\sigma^1, \ldots, \sigma^n$ be a basis for the left invariant 1-forms on G, and consider a bi-invariant metric for G; we know such exist if G is compact. Use the orientation for G

such that $\sigma^1 \wedge \cdots \wedge \sigma^n$ is positive. We know, for multi-inde
$I = (i_1, \ldots, i_p)$

$$* \sigma^I = \sigma^J$$

where J is the correctly ordered complementary index to I. Then

$$L_g^* (* \sigma_h^I) = L_g^* (\sigma_h^J)$$
$$= \sigma_{g^{-1}h}^J$$
$$= * \sigma_{g^{-1}h}^I$$
$$= * L_g^* \sigma_h^I$$



and so $L_g^* \circ * = * \circ L_g^*$. Likewise $R_g^* \circ * = * \circ R_g^*$

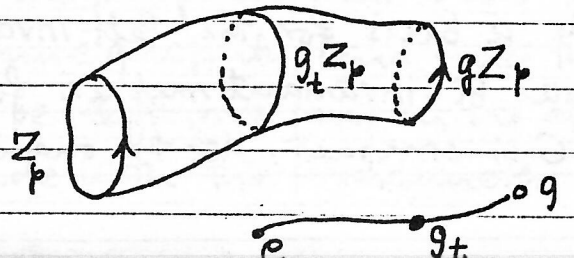
Suppose now that h^p is a harmonic p-form on G (in the bi-invariant metric)

$$dh^p = 0 = d * h^p$$

Then $d L_g^* h^p = 0$ and $d * L_g^* h^p = d L_g^* * h^p = 0$, and so the left (and likewise right) translations of h^p are also harmonic. We claim that if G is connected, then $L_g^* h^p = h^p = R_g^* h^p$, ie

Harmonic forms in a bi-invariant metric are bi-invariant.

To show e.g. that $L_g^* h^p = h^p$ we need only show that $L_g^* h^p$ and h^p have the same periods (see (53.4)). But if Z_p is a p-cycle on G, note that gZ_p is homologous to Z_p; for we can join e to g by a curve g_t and then $\{g_t Z_p\}$ defines a deformation of Z_p into gZ



But then

$$\int_{Z_p} L_g^* h^p = \int_{gZ_p \sim Z_p} h^p = \int_{Z_p} h^p, \quad \text{as desired.} \qquad \square.$$

Conversely, we claim that bi-invariant p-forms are harmonic in the bi-invariant metric. If β^p is bi-invariant then certainly β^p is closed $d\beta^p = 0$. Also, if β^p is bi-invariant, then so is $*\beta^p$

$$L_g^* * \beta^p = * L_g \beta^p = * \beta^p \quad \text{etc.}$$

But if $*\beta^p$ is bi-invariant it is closed, $d*\beta^p = 0$, and so β^p is harmonic

> In a bi-invariant metric on G the bi-invariant forms coincide with the harmonic forms (75.9)

> Corollary: On a compact Lie group the left Haar measure is also a right Haar measure (75.10)

Proof (the hard way). $\sigma^1 \wedge \cdots \wedge \sigma^n$ is the left Haar measure; of course it is closed. But $*(\sigma^1 \wedge \cdots \wedge \sigma^n)$ = the constant function 1, and so $d * (\sigma^1 \wedge \cdots \wedge \sigma^n) = 0$. $\therefore \sigma^1 \wedge \cdots \wedge \sigma^n$ is harmonic and so bi-invariant.

> Theorem. If G is a compact non-abelian Lie group, then the third Betti number of G does not vanish
>
> $$b_3(G) = \dim H_3(G; \mathbb{R}) \neq 0$$ (75.11)

Proof. The Cartan 3-form

$$\Omega_3(X,Y,Z) = -3\langle [X,Y],Z\rangle$$

is bi-invariant, hence harmonic. We need only show that Ω_3 does not vanish identically. But if

$$\langle [X,Y],Z\rangle = 0$$

for all Z, then $[X,Y]=0$ for all X,Y, and so the Lie algebra is abelian, $XY=YX$ all $X\in\mathcal{G}$, $Y\in\mathcal{G}$. But it follows easily from the power series that if X and Y commute

$$e^X e^Y = e^{X+Y} = e^Y e^X$$

and so G must then be abelian. □

In conclusion, we have shown that the Cartan bi-invariant 3-form

$$\boxed{\operatorname{tr}\, g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg} \qquad (15.12)$$

defines, via deRham, a non trivial 3 dimensional "cohomology class", at least if G is not abelian.

Lecture 76. Winding Numbers and Chern Forms

(1) *The Yang-Mills "Winding Number"* Recall that in Lect. 74 we were comparing, on a distant 3-sphere $S^3 \subset \mathbb{R}^4$, the frame e_U with the covariant constant frame e_V, (74.4) & (74.5).

$$e_U(x) = e_V(x)\, g_{VU}(x)$$

$$g_{VU} : S^3 \longrightarrow SU(n)$$

The bi-invariant Cartan 3-form on $SU(n)$

$$\Omega_3 = \mathrm{tr}\, g^{-1}dg \wedge g^{-1}dg \wedge g^{-1}dg$$

defines a non-trivial de Rham "cohomology" class on $SU(n)$ (as it does on any compact non-abelian Lie group) and we can form the number

$$\int_{S^3} g_{VU}^{*}(\Omega_3) = \int_{g_{VU}(S^3) \subset SU(n)} \Omega_3 . \qquad (76.1)$$

While g_{VU} depends strongly on the Y-M connection $\omega_U = ig A_U$ in the curved region U it is clear that the number (76.1) is a homotopy invariant of A_U; a deformation of the connection A_U will leave (76.1) unchanged. For our purposes it will be important to normalize the form Ω_3. For this we shall consider the $SU(2)$ subgroup of $SU(n)$

$$SU(2) = \left(\begin{array}{c|c} SU(2) & 0 \\ \hline 0 & 1 \end{array}\right) \subset SU(n)$$

The Cartan 3-form Ω_3 of $SU(n)$ restricts to Ω_3 for $SU(2)$, and we shall use as a normalization constant

$$\int_{SU(2)} \Omega_3 .$$

Let us compute this number.

Ω_3 is a bi-invariant 3-form on SU(2) and so is the volume form vol^3; it is then clear that Ω_3 is some constant multiple of vol^3. Letting $\sigma_x, \sigma_y, \sigma_z$ be the Pauli matrices, $i\sigma_x, i\sigma_y, i\sigma_z$ give a basis for su(2), and

$$\Omega_3(i\sigma_x, i\sigma_y, i\sigma_z) = 3\,\text{tr}\left([i\sigma_x, i\sigma_y]\, i\sigma_z\right) = 12.$$

and $$\text{vol}^3(i\sigma_x, i\sigma_y, i\sigma_z) = \|i\sigma_x\|\,\|i\sigma_y\|\,\|i\sigma_z\| = (\sqrt{2})^3$$

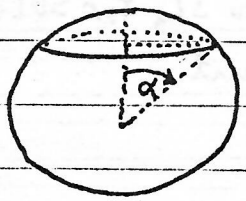
and so

$$\Omega_3 = \frac{12}{(\sqrt{2})^3}\,\text{vol}^3 \tag{76.2}$$

for SU(2). What is the volume of SU(2), $\int_{SU(2)} \text{vol}^3$?

The bi-invariant metric on SU(2) based on $\langle X, Y\rangle_{\text{of}} = -\text{tr}\,XY$ must be some constant multiple of the Euclidean metric on $S^3 = SU(2)$,

. The 1-parameter subgroup $e^{i\sigma_z t}$ describes a great circle on S^3 since it is a geodesic on SU(2), and here $0 \le t \le 2\pi$. The length of this great circle, in the $\langle\,,\rangle_{\text{of}}$ metric is $2\pi\,\|i\sigma_z\| = 2\pi\sqrt{2}$. Thus SU(2) carries the same metric as the Euclidean sphere S^3 of radius $\sqrt{2}$. The volume of the unit $S^3 \subset \mathbb{R}^4$ is



$$\int_0^\pi 4\pi \sin^2\alpha\, d\alpha = 2\pi^2$$

and so $$\int_{SU(2)} \text{vol}^3 = 2\pi^2 \cdot (\sqrt{2})^3.$$

We conclude from (76.2)

$$\int_{SU(2)} \Omega_3 = 24\,\pi^2 \tag{76.3}$$

Finally, we define the "winding number" of the Y-M field "at infinity", to be

$$\left.\begin{aligned} \frac{1}{24\pi^2}\int_{S^3} g_{VU}^{*}\,\Omega_3 \;&=\; \frac{1}{24\pi^2}\int_{g_{VU}(S^3)} \Omega_3 \\ &= \frac{1}{24\pi^2}\int_{g_{VU}(S^3)} \mathrm{tr}\; g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg \end{aligned}\right\} \qquad (76.4).$$

(2) Winding Number in terms of Field Strength. Chern's expression (74.10) in the U(1) case suggests an expression for this winding number in terms of a 4-tuple integral involving curvature (we are assuming that the YM potential ω_U is globally defined in U, there are no singularities and $j(e_U)=0$)

Note $\vartheta = d\omega + \omega\wedge\omega$, and so

$$\vartheta\wedge\vartheta = (d\omega+\omega\wedge\omega)\wedge(d\omega+\omega\wedge\omega)$$

$$= d\omega\wedge d\omega + d\omega\wedge\omega\wedge\omega + \omega\wedge\omega\wedge d\omega + \omega\wedge\omega\wedge\omega\wedge\omega$$

But $\mathrm{tr}(\omega\wedge\omega\wedge d\omega) = \mathrm{tr}(d\omega\wedge\omega\wedge\omega)$ and $\mathrm{tr}(\omega\wedge\omega\wedge\omega\wedge\omega)=0$

and so

$$\mathrm{tr}(\vartheta\wedge\vartheta) = \mathrm{tr}(d\omega\wedge d\omega) + 2\,\mathrm{tr}(d\omega\wedge\omega\wedge\omega)$$

$$= d\,\mathrm{tr}(\omega\wedge d\omega) + 2\,\mathrm{tr}(d\omega\wedge\omega\wedge\omega).$$

Also

$$d(\omega\wedge\omega\wedge\omega) = d\omega\wedge\omega\wedge\omega - \omega\wedge d\omega\wedge\omega + \omega\wedge\omega\wedge d\omega$$

$$\mathrm{tr}\; d(\omega\wedge\omega\wedge\omega) = 3\,\mathrm{tr}(d\omega\wedge\omega\wedge\omega)$$

$$\therefore \quad \mathrm{tr}(\vartheta\wedge\vartheta) = d\,\mathrm{tr}(\omega\wedge d\omega) + \frac{2}{3}\,d\,\mathrm{tr}(\omega\wedge\omega\wedge\omega)$$

$$\mathrm{tr}(\vartheta\wedge\vartheta) = d\left[\mathrm{tr}\left(\omega\wedge d\omega + \frac{2}{3}\,\omega\wedge\omega\wedge\omega\right)\right] \qquad (76.5)$$

In our Y-M case

$$\int_U \mathrm{tr}\,\vartheta\wedge\vartheta = \int_{\partial U=S^3} \mathrm{tr}\,(\omega\wedge d\omega + \tfrac{2}{3}\,\omega\wedge\omega\wedge\omega).$$

But since $\vartheta=0$ on S^3, $\omega\wedge d\omega = \omega\wedge\vartheta - \omega\wedge\omega\wedge\omega = -\omega\wedge\omega\wedge\omega$ and

$$\int_U \mathrm{tr}\,\vartheta\wedge\vartheta = -\frac{1}{3}\int_{S^3} \mathrm{tr}\,\omega\wedge\omega\wedge\omega.$$

But $\omega_U = g^{-1}dg$ on S^3 (see (14.3)). We have finally

$$\boxed{\text{winding number on } S^3 = \frac{1}{24\pi^2}\int_{S^3}\mathrm{tr}(\omega_U\wedge\omega_U\wedge\omega_U) = -\frac{1}{8\pi^2}\int_{\mathbb{R}^4}\mathrm{tr}\,\vartheta\wedge\vartheta} \qquad (76.6)$$

Some comments.

(1) $\mathrm{tr}\,\vartheta\wedge\vartheta$ is not the Lagrangian; the Lagrangian was essentially $\mathrm{tr}\,\vartheta\wedge *\vartheta$.

$$(F\wedge F) = (F\wedge F)_{0123}\,dt\wedge dx\wedge dy\wedge dz = \left(\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right)dt\wedge dx\wedge dy\wedge dz \qquad (76.7)$$

while

$$(F\wedge *F) = \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu}\,dt\wedge dx\wedge dy\wedge dz$$

(2) In physics books they say that $\mathrm{tr}\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ is a divergence. This is their way of stating (76.5). One should be careful; it is globally exact in our Y-M situation. However it will not be globally exact in topologically more interesting situations. The statement that $\mathrm{tr}\,F\wedge F$ is locally exact cannot be surprising since it is a (closed) 4-form on an M^4. The topological significance of $\mathrm{tr}\,F\wedge F$, generalizin

Poincaré's theorem for closed surfaces

$$\frac{1}{2\pi}\int_{M^2} K\,dS = \chi(M) \tag{76.8}$$

was understood by Chern. We turn to these matters.

(3) The Chern Forms for a $U(n)$ Bundle. Let E be a complex $\mathbb{C}^k$ vector bundle with structure group $U(k)$ and base manifold M^n. We have the local curvature forms ($u(k)$ valued 2-forms) ϑ_U.

For any matrix (with real number entries) K we can expand (K is a $k\times k$ matrix)

$$\det(I+K) = 1 + \operatorname{tr}K + \operatorname{tr}\hat{K}^2 + \operatorname{tr}\hat{K}^3 + \cdots + \operatorname{tr}\hat{K}^k$$

and where, if $\lambda_1,\dots,\lambda_k$ are the eigenvalues of K,

$$\left.\begin{aligned} \operatorname{tr}K &= \lambda_1+\lambda_2+\cdots+\lambda_k \\ \operatorname{tr}\hat{K}^2 &= \sum_{i<j}\lambda_i\lambda_j = \lambda_1\lambda_2+\lambda_1\lambda_3+\cdots+\lambda_{k-1}\lambda_k \\ \operatorname{tr}\hat{K}^3 &= \sum_{i<j<k}\lambda_i\lambda_j\lambda_k \\ &\vdots \\ \operatorname{tr}\hat{K}^k &= \lambda_1\lambda_2\cdots\lambda_k = \det K \end{aligned}\right\} \tag{76.9}$$

are the elementary symmetric functions of the eigenvalues of K.

Consider now formally the result of substituting the curvature 2-form matrix $\vartheta = \vartheta_U$ for K; in fact we consider

$$\det\left(I+\frac{i}{2\pi}\vartheta\right) = \begin{vmatrix} 1+\frac{i}{2\pi}\vartheta_{11} & \frac{i}{2\pi}\vartheta_{12} & \cdots & \frac{i}{2\pi}\vartheta_{1k} \\ \frac{i}{2\pi}\vartheta_{21} & 1+\frac{i}{2\pi}\vartheta_{22} & \cdots & \frac{i}{2\pi}\vartheta_{2k} \\ \vdots & & & \\ \frac{i}{2\pi}\vartheta_{k1} & \frac{i}{2\pi}\vartheta_{k2} & \cdots & 1+\frac{i}{2\pi}\vartheta_{kk} \end{vmatrix}$$

where we expand out the determinant in the usual way; since ϑ_{rs} is a 2-form there is no problem with ordering. The answer

is a well define collection of forms

$$\left.\begin{aligned} \det\left(I+\frac{i}{2\pi}\vartheta\right) &= 1+\frac{i}{2\pi}\operatorname{tr}\vartheta+\cdots \\ &\equiv 1+c_1(E)+c_2(E)+\cdots+c_k(E) \end{aligned}\right\} \quad (76.10)$$

where $c_r(E)$ is a $2r$-form on $U\subset M^n$, the $\underline{r^{th}\text{ Chern form}}$.
Note that in an overlap $U\cap V$

$$\det\left(I+\tfrac{i}{2\pi}\vartheta_V\right)=\det\left(I+\tfrac{i}{2\pi}c_{VU}\vartheta_U c_{VU}^{-1}\right)=\det c_{VU}\left(I+\tfrac{i}{2\pi}\vartheta_U\right)c_{VU}^{-1}$$
$$=\det\left(I+\tfrac{i}{2\pi}\vartheta_U\right)$$

and so each c_r is in fact a globally defined $2r$-form on M^n!
The form c_1 is familiar

$$c_1=\frac{i}{2\pi}\operatorname{tr}\vartheta$$

For c_2, note that $\sum_{r<s}\lambda_r\lambda_s=\frac{1}{2}\left[\left(\sum_j\lambda_j\right)^2-\sum_j\lambda_j^2\right]$ (76.11)

$$c_2=-\frac{1}{4\pi^2}\cdot\frac{1}{2}\left[(\operatorname{tr}\vartheta)\wedge(\operatorname{tr}\vartheta)-\operatorname{tr}(\vartheta\wedge\vartheta)\right] \quad (76.\ )$$

Note that in the case of a $U(1)$ bundle, $c_1=\frac{i}{2\pi}\vartheta$. When E is the complex line bundle, the tangent bundle to a surface M^2, $c_1=\frac{1}{2\pi}K\cdot dS$ (see (59.17)). In c_2, note that the term

$$\frac{1}{8\pi^2}\operatorname{tr}(\vartheta\wedge\vartheta)$$

is exactly what occurs in the Y.M winding number (76.6).

Note that in general, from Bianchi

$$dc_1=\frac{i}{2\pi}d\operatorname{tr}\vartheta=\frac{i}{2\pi}\operatorname{tr}(d\vartheta)=\frac{i}{2\pi}\operatorname{tr}\left[\vartheta\wedge\omega-\omega\wedge\vartheta\right]=0.$$

● Problem 110. Show directly (without using (76.5)) that

$$dc_2=0$$

● Problem 111. In (76.11) we expressed $\sum\lambda_r\lambda_s$ in terms of $\operatorname{tr}K$ and $\operatorname{tr}(K^2)$.
Express $\sum_{r<s<t}\lambda_r\lambda_s\lambda_t$ in terms of $\operatorname{tr}K$, $\operatorname{tr}(K^2)$, $\operatorname{tr}(K^3)$

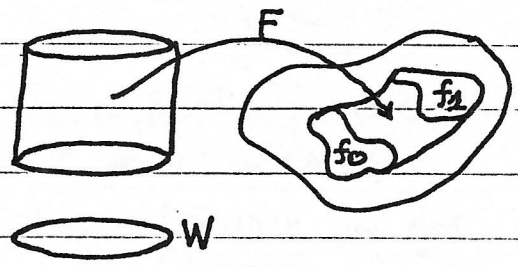
Lecture 77 Instanton Number

This lecture was not written up. It considered selfdual and anti-selfdual connections as yielding solutions to the Yang-Mills equations.

In Lecture 76 we defined the Chern classes for a complex vector bundle and remarked that c_1 and c_2 are closed forms (of degree 2 and 4 respectively). In fact all the Chern forms c_r are closed, as we shall see, and consequently yield cohomology classes of the base manifold M^n via de Rham's theorem. To discuss these cohomology classes we shall need to know more about the topology of Lie groups, and in particular their "homotopy groups."

(1) Homotopy. If $f_1 : W \to M^n$ and $f_2 : W \to M^n$ are two maps of a space W into M^n, we say that f_0 and f_2 are homotopic if there is a map $F : W \times I \longrightarrow M^n$ of the "cylinder" $W \times [0,1]$ into M^n such that

$$F(w,0) = f_0(w)$$
$$F(w,1) = f_2(w)$$

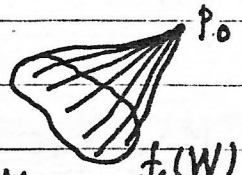


Each of the maps F_t

$$F_t(w) = F(w,t)$$

is homotopic to the original F_0.

If f_2 maps all of W to a single point $p_0 \in M^n$ we say that f_0 is homotopic to the constant map



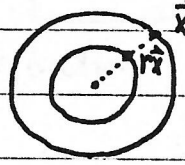
We shall be especially interested in the situation where $W = S^k$ is a k-sphere, $k = 0, 1, 2, \ldots$, even when $k > n = \dim M$. The k-sphere is of course the boundary of the $k+1$ ball D^{k+1}. Our main concern centers around the following simple but very important observation

Proposition $f: S^k \to M^n$ is homotopic to a constant map if and only if f can be extended to a map $\tilde{F}: D^{k+1} \to M^n$ (78.1).

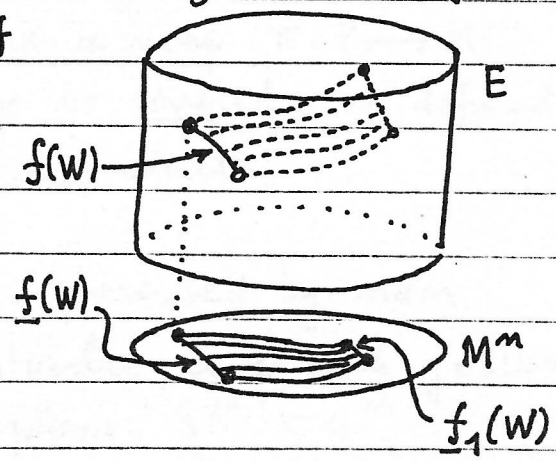
Proof. Suppose $\tilde{F}: D^{k+1} \to M^n$ extends $f: S^k \to M^n$; thus $\tilde{F}(\vec{x}) = f(\vec{x})$ for $\|\vec{x}\| = 1$. Define $F: S^k \times I \to M^n$ by $F(\vec{x}, r) = \tilde{F}((1-r)\vec{x})$, for $\vec{x} \in S^k$, $0 \le r \le 1$. Then $F(\vec{x}, 0) = \tilde{F}(\vec{x}) = f(\vec{x})$ and $F(\vec{x}, 1) = \tilde{F}(\vec{0})$ shows that f is homotopic to the constant map $\vec{x} \mapsto \tilde{F}(\vec{0})$.

Suppose, on the other hand, that $f = f_0$ is homotopic to the constant map $f_1(\vec{x}) = p_0 \in M^n$; then we have $F: S^k \times I \to M^n$ with $F(\vec{x}, 0) = f(\vec{x})$, $F(\vec{x}, 1) = f_1(\vec{x}) = p_0$. Define an extension $\tilde{F}: D^{k+1} \to M^n$ by

$$\tilde{F}(r\vec{x}) = F(\vec{x}, 1-r)$$



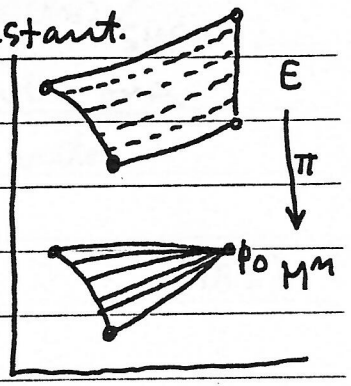
(2) Covering Homotopies. Let $E \xrightarrow{\pi} M^n$ be a vector bundle, with connection ω. Let $f: W \to E$ be a map of a space W into the bundle space E; then of course we get a map $\underline{f}: W \to M^n$ by $\underline{f} = \pi \circ f$. Suppose now that we have a homotopy of $\underline{f}$, $\underline{F}$, to a new map $\underline{f}_1: W \to M^n$. We claim that we can "cover" this homotopy by a homotopy of the original map $f: W \to E$, ie a map $F: W \times I \to E$ such that $\pi F(w,t) = \underline{F}(w,t)$. To do this we consider a fixed $w \in W$ and look at the curve $t \to \underline{F}(w,t)$ in M^n. Over this curve in M^n there is a unique curve over it in E representing.



"parallel displacement" along $F(w,t)$, starting at $f(w)$. That is, we look at the unique curve in E that lies over $F(w,t)$, starts at $f(w)$, and is tangent to the n-plane distribution in E defined locally by

$$d\psi^i + \omega^i{}_j(F(w,t))\,\psi^j = 0. \qquad (78.2)$$

(Note that if f is homotopic to a constant, it is not necessarily true that f will be homotopic to a constant. It need only be true that all points of $F(w,1)$ will lie on the same fiber $\pi^{-1}(p_0)$ over p_0).



What we have said for a vector bundle also holds for a principal fiber bundle $P \xrightarrow{\pi} M^n$ over M^n. Instead of (78.2) we consider the distribution

$$\tilde{\omega} = g^{-1}\omega g + g^{-1}dg = 0$$

defined in P. It turns out that one can cover homotopies in any fiber bundle, without any use of a connection. In fact one can define the notion of a fiber space over M^n; this is a space P and a map $\pi: P \to M$ such that homotopies can always be covered, as defined above. Such spaces need not be local products!

(3) Some Topology of $SU(n)$. $SU(n)$ is represented by $n \times n$ matrices acting on $\mathbb{C}^n$. In particular, since each $g \in SU(n)$ is unitary $SU(n)$ sends the unit sphere $S^{2n-1} \subset \mathbb{C}^n$

$$S^{2n-1} = \{ z \in \mathbb{C}^n \mid |z_1|^2 + \cdots + |z_n|^2 = 1 \}$$

into itself. It is clear that $SU(n)$ acts transitively on S^{2n-1}, for the point $(1,0,\cdots,0)$ can be sent into the point $(z_1,\cdots,z_n) \in S^{2n-1}$ simply by writing down some $g \in SU(n)$

having $(z_1, \dots, z_n)^t$ as first column. The stability subgroup ("little group") of the point $(1, 0, \dots, 0)$ is clearly the subgroup.

$$SU(n-1) = \begin{bmatrix} 1 & 0 \dots 0 \\ 0 & \\ \vdots & \quad SU(n-1) \\ 0 & \end{bmatrix}$$

and so we have a diffeomorphism $S^{2n-1} \approx SU(n)/SU(n-1)$ and in fact SU(n) is a principal SU(n-1) bundle over S^{2n-1} (see Lecture 61). In mathematics one writes

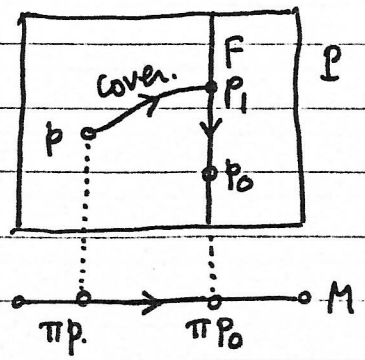
$$\begin{array}{ccc} SU(n-1) & \longrightarrow & SU(n) \\ & & \downarrow \pi \\ & & S^{2n-1} \end{array} \qquad (78.3)$$

Proposition: SU(n) is connected and simply connected. (78.4)

We first show SU(n) is connected. First we note a much more general fact;

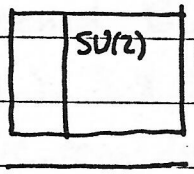
If $F \to P \xrightarrow{\pi} M$ is a fiber bundle, and if the base space M and the fiber F are connected, then P is connected. (78.5)

To see this, let p and p_0 be points in P. Project them down into M. Since M is connected there is a curve in M joining πp to πp_0. This curve can be considered as a homotopy from the constant map of a point w into πp to the constant map of the point w to πp_0. Cover this homotopy by a path from p to the fiber over πp_0,



ie from ϕ to a point p_1 in the fiber holding p_0. Since F is $\pi^{-1}(p_0)$ connected, we can find a curve in F joining this point p_1 to p_0. This finishes the general statement.

Now $SU(1)$ is a single point; no problem. $SU(2)$ is the 3-sphere S^3 (Lecture 66) and so is connected. $SU(3)$ is a principal $SU(2)$ bundle over S^5 and S^5 is connected

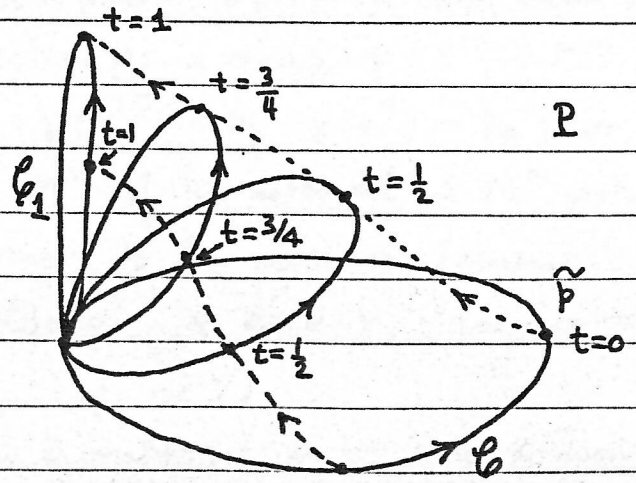


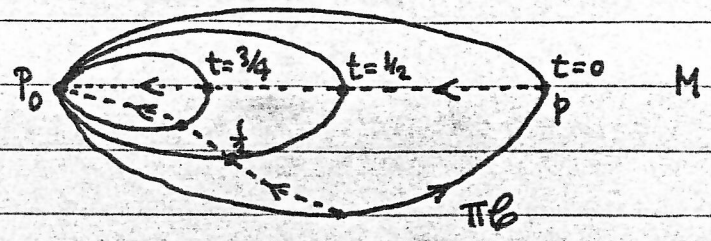
Thus $SU(3)$ is connected. $SU(4)$ is a principal $SU(3)$ bundle over S^7, and so is connected. Etc. We have shown that $SU(n)$ is connected.

Problem 112. Show that $SO(n)$ is connected

We now show that $SU(n)$ is simply connected, ie any closed curve on $SU(n)$ can be contracted to a point. First, again note a more general fact.

(78.6) Let $F \to P \downarrow \pi M$ be a fiber bundle with M simply connected and fiber F simply connected. Then P is simply connected.





Proof. Let $\mathcal{C}$ be a closed curve in P. Project it down to $\pi\mathcal{C}$ in M^m. Since M^m is simply connected, $\pi\mathcal{C}$ can be contracted to a point $p_0 \in M^m$. Each of the deformation curves joining points p of $\pi\mathcal{C}$ to p_0 can be covered by a deformation curve in P joining $\tilde{p} \in \pi^{-1}(p)$ to an endpoint lying in the fiber $\pi^{-1}(p_0)$. The resulting endpoints form a closed curve $\mathcal{C}_1$ in $\pi^{-1}(p_0)$. We have deformed $\mathcal{C}$ into a closed curve lying entirely in the fiber $\pi^{-1}(p_0)$. But this fiber is simply connected, hence this closed curve can be deformed to a point in $\pi^{-1}(p_0)$. □.

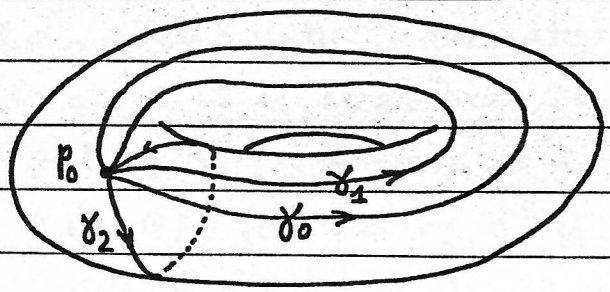
In the case of $SU(n)$, note $SU(1)$ is a point; no problem. $SU(2)$ is a 3-sphere, which is simply connected (it is $\mathbb{R}^3$ with a single point "at infinity" thrown in. $SU(3)$ is a principal $SU(2)$ bundle over S^5. Since $SU(2)$ and S^5 are simply connected, so is $SU(3)$. $SU(4)$ is (etc. etc.). □

(4) The Fundamental Group $\pi_1(M^m)$. Consider all closed curves on an M^m that start and end at a given point $p_0 \in M^m$ (called the "base point" of M^m). Each such curve can be considered either as a map $\gamma: S^1 \to M^m$ of the circle in M^m (i.e $\gamma(e^{i\theta})$ with $\gamma(e^0) = p_0$) or as a map $\gamma: [0,1] \to M^m$ of the interval in M^m with $\gamma(0) = p_0 = \gamma(1)$; the latter seems to be most convenient. We shall identify two such "loops", γ_0 and γ_1, saying they are <u>homotopic</u>

$$\gamma_0 \sim \gamma_1$$

provided γ_1 is homotopic to γ_0 via a homotopy $F: [0,1] \times [0,1] \to M^m$ that preserves the base point, ie $F = F(\theta, t)$ with

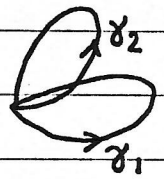
$$F(0,t) = p_0 = F(1,t)$$



$\gamma_1 \sim \gamma_0$ on T^2
but $\gamma_2 \not\sim \gamma_0$.

Given any two loops γ_1 and γ_2 on M^n we may compose them and form a new loop $\gamma_2\gamma_1$, defined by

$$\gamma_2\gamma_1(t) = \begin{cases} \gamma_1(2t) & \text{for } 0 \le t \le \frac{1}{2} \\ \gamma_2(2t-1) & \text{for } \frac{1}{2} \le t \le 1 \end{cases}$$

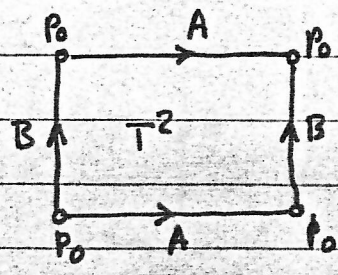


One can show that if $\gamma_1' \sim \gamma_1$ and $\gamma_2' \sim \gamma_2$, then $\gamma_2'\gamma_1' \sim \gamma_2\gamma_1$. The homotopy classes of loops on M^n form a group under "multiplication" $(\gamma_1, \gamma_2) \to \gamma_1\gamma_2$. This group is Poincaré's fundamental group, written $\pi_1(M^n, p_0)$. We shall usually suppress the point p_0 since if M^n is connected, the groups $\pi_1(M^n, p_0)$ are all isomorphic, all $p_0 \in M^n$. We write $\pi_1(M^n)$.

Consider loops on the circle $M^1 = S^1$. These are homotopy classes of maps $\gamma: S^1 \to S^1$. Certainly if $\gamma_1: S^1 \to S^1$ and $\gamma_2: S^1 \to S^1$ are <u>homotopic</u> they will have the same Brouwer degree. It turns out that mappings of the same Brouwer degree are homotopic. Thus the homotopy class of a map $\gamma: S^1 \to S^1$ is characterized by an integer, the degree. We write

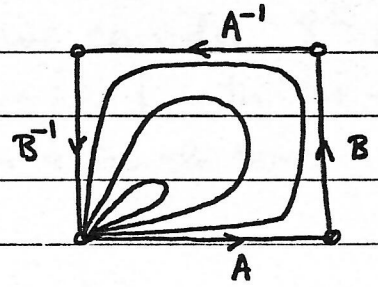
$$\pi_1(S^1) \approx \mathbb{Z}.$$

It turns out that the fundamental group of the 2-torus



is an abelian group with 2 generators, A and B. It is abelian because it is clear that the closed loop A followed by B followed by A^{-1} (ie A traversed in the opposite direction, $A^{-1}(\theta) = A(1-\theta)$) followed by B^{-1} is contractible to p_0, ie

$$B^{-1}A^{-1}BA \sim 1$$

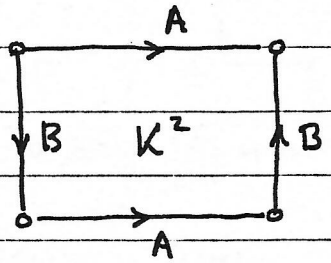


(the identity element $1 \in \pi_1(M^n)$ is the homotopy class of the loops contractible to the point p_0).

Thus

$$BA \sim AB.$$

For the Klein Bottle



$$BA^{-1}BA \sim 1.$$

and so

$$BA = AB^{-1}$$

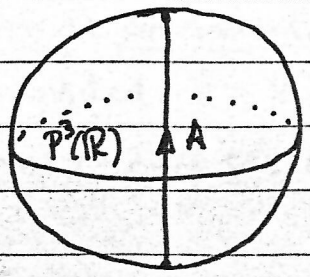
The fundamental group $\pi_1(K^2)$ of the Klein bottle is the <u>non-commutative group</u> with two generators A and B satisfying the single relation $BA = AB^{-1}$.

The rotation group SO(2) is topologically a circle S^1 and so SO(2) is not simply connected

$$\pi_1 SO(2) \approx \mathbb{Z}.$$

The rotation group SO(3) is topologically $P^3(\mathbb{R})$ and so is not simply connected

$$\pi_1(SO(3)) \approx \mathbb{Z}_2 A$$

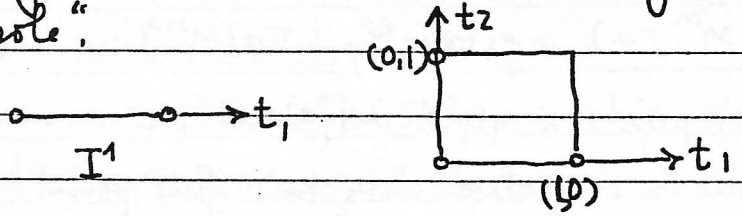


$A^2 = A \cdot A$ <u>is</u> contractible.

Lecture 78. The Higher Homotopy Groups $\pi_k(M^n)$

(1) $\pi_k(M^n)$. We shall consider maps $f: S^k \to M^n$ of a k-sphere S^k into M^n. We shall always ask that a distinguished point on S^k, the "north pole", be sent to a distinguished "base point", usually written $*$, in M.

For technical reasons we consider S^k to be the unit k-cube $I^k = [0,1] \times \cdots \times [0,1]$ with the entire boundary $\dot{I}^k$ identified with a single point, the "north pole".



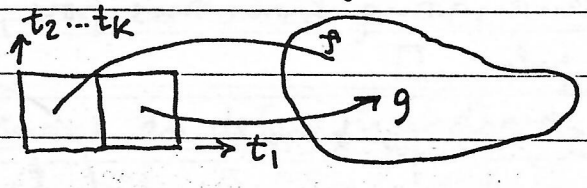
Then our $f: S^k \to M^n$ is a map $f: I^k \to M^n$ such that $f(\dot{I}^k) = * \in M^n$. To say f_0 is homotopic to f_1, $f_0 \sim f_1$ is to say there is a map

$F: S^k \times I \to M^n$ such that $F(y,0) = f_0(y)$ and $F(y,1) = f_1(y)$

$F(\text{northpole}, t) = *$, all $t \in [0,1]$.

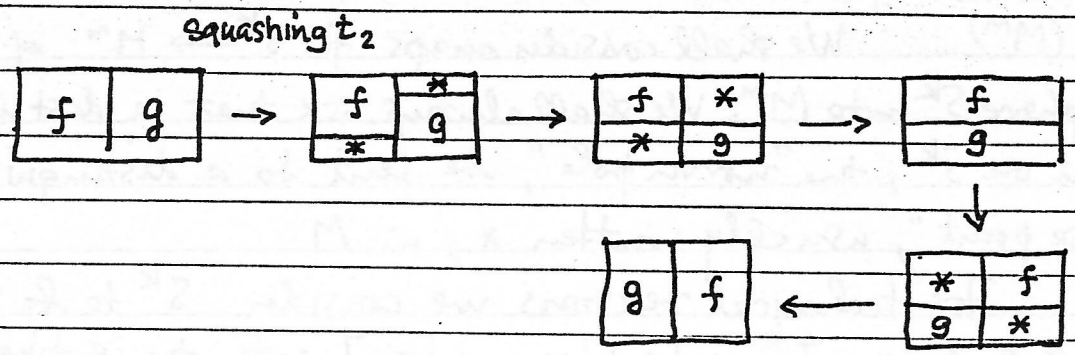
We compose two maps $f: S^k \to M^n$ and $g: S^k \to M^n$ as we did loops; this time it is called $f+g$

$$(f+g)(t_1 \cdots t_k) = \begin{cases} f(2t_1, t_2, \ldots t_k) & 0 \le t_1 \le \frac{1}{2} \\ g(2t_1 - 1, t_2, \ldots t_k) & \frac{1}{2} \le t_1 \le 1. \end{cases} \tag{78.1}$$



briefly $\boxed{f \mid g}$.

Again, two maps f and g are identified if they are homotopic. We use the additive notation $f+g$ since these mappings form a commutative group under our composition, provided $k \ge 2$. (we know $\pi_1(M^n)$ need not be commutative). We see commutativity from the following homotopies

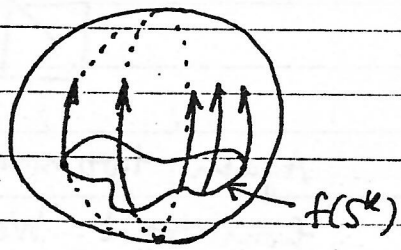


The resulting abelian group (for $k \geq 2$) is written

$$\pi_k(M^n, *) \quad \text{or briefly} \quad \pi_k(M^n)$$

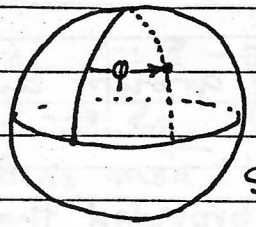
Different basepoints * yield different but isomorphic groups. We shall not concern ourselves with this fine point.

Consider the case $M^n = S^n$. $\pi_k(S^n)$ consists of homotopy classes of maps of a k-sphere into an n-sphere. Consider the case $k < n$. It seems evident (but is actually false!) that $f(S^k)$ can not cover all of S^n if $k < n$. This can be proved if f is differentiable; see Sard's theorem. It is a fact that a continuous f is always homotopic to a differentiable f (by approximations); thus we may assume that $f(S^k)$ does not cover all S^n. If the south pole of S^n is not covered, we can push the image $f(S^k)$ along great circles emanating from this pole, collapsing $f(S^k)$ at the north pole. □



We conclude $\pi_k(S^n) = 0$ for $k < n$.

For $k = n$ we know that homotopic maps of S^n into itself have the same Brouwer degree. A theorem of Heinz Hopf shows conversely that maps $S^n \to S^n$ of the same degree are in fact homotopic. Thus homotopy classes of maps $S^n \to S^n$ are characterized by an integer, the degree. It is easy to construct a map $S^n \to S^n$ of any given

degree N; 

send $\varphi \longrightarrow N\varphi$

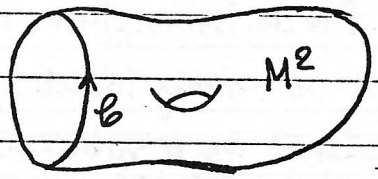
We conclude

$$\pi_k(S^n) = \begin{cases} 0 & \text{if } k<n \\ \mathbb{Z} & \text{if } k=n. \end{cases} \qquad (78.2)$$

Two remarks (1) One can also define $\pi_0(M)$; we shall try to avoid using it (for lack of time).

(2) (78.2) says nothing of $\pi_k(S^n)$ for $k>n$. Hopf made the surprising discovery that $\pi_3(S^2) \approx \mathbb{Z}$. We shall discuss this later!

Warning. Do not confuse homotopy with homology. It is true that each $f: S^k \longrightarrow M^n$ does define a singular k-cycle of M^n (after choice of orientation of S^k). It is also true that if f is homotopic to a constant map then $f(S^k)$ is homologous to zero but the converse is not true



The curve C is the image of a circle. It bounds in M^2 but it cannot be contracted to a point!

Homotopic maps do define homologous cycles; this is the "deformation lemma" of Lecture 51. Homology deals with general cycles, in particular general submanifolds, while homotopy groups deal only with images of spheres.

(2) Exact Sequences. A sequence of groups and homomorphisms

$$\longrightarrow G_j \xrightarrow{f_j} G_{j+1} \xrightarrow{f_{j+1}} G_{j+2} \longrightarrow \cdots$$

is said to be exact at G_{j+1} provided that the kernel of f_{j+1} (the subgroup of G_{j+1} sent into the identity of G_{j+2}) coincides with the image $f_j(G_j) \subset G_{j+1}$ of G_j under f_j. If the sequence is exact at each G_j we say the sequence is exact. 0 indicates the group with just the identity elt.

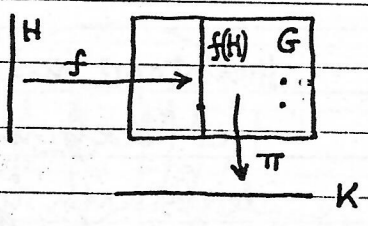
For example, if $0 \xrightarrow{f} H \xrightarrow{g} G$ is exact at H, then $\ker(g) = \text{im}(f) = 0$. Since the kernel of g is 0, g must be 1:1. If $H \xrightarrow{r} G \xrightarrow{s} 0$ is exact at G, then $\text{im}(r) = \ker s =$ all of G, ie $r: H \longrightarrow G$ is onto. If $0 \longrightarrow H \xrightarrow{f} G \longrightarrow 0$ is exact, meaning exact at all the interior groups, then f must be both 1:1 and onto, ie f is an isomorphism.

$$\left.\begin{array}{ll} 0 \longrightarrow H \xrightarrow{f} G & f \text{ is 1:1 or } \underline{\text{injective}} \\ H \xrightarrow{f} G \longrightarrow 0 & f \text{ is onto or } \underline{\text{surjective}} \\ 0 \longrightarrow H \xrightarrow{f} G \longrightarrow 0 & f \text{ is an isomorphism or } \underline{\text{bijective}} \end{array}\right\} \quad 78.3$$

In $0 \longrightarrow H \xrightarrow{f} G$, since f is 1:1 we may identify H with its image f(H), in other words H may be considered as a subgroup of G.

Consider the exact sequence of abelian groups

$$0 \longrightarrow H \xrightarrow{f} G \xrightarrow{\pi} K \longrightarrow 0$$



The kernel of π consists of the subgroup f(H). Thus K consists of equivalence classes of elements of G, $g \sim g'$ if and only if $g - g' \in f(H)$. That is, $K \approx G/f(H)$, or, since f is 1:1, we write $K = G/H$

$$\left.\begin{array}{c} 0 \longrightarrow H \longrightarrow G \longrightarrow K \longrightarrow 0 \\ \text{says } K = G/H \end{array}\right\} \quad (78.4)$$

For example, if $f: \mathbb{Z} \to \mathbb{Z}$ sends $n \in \mathbb{Z} \to 2n \in \mathbb{Z}$ and if $\pi: \mathbb{Z} \to \mathbb{Z}_2$ sends each integer into its equivalence class mod 2, then

$$0 \to \mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}_2 \to 0$$

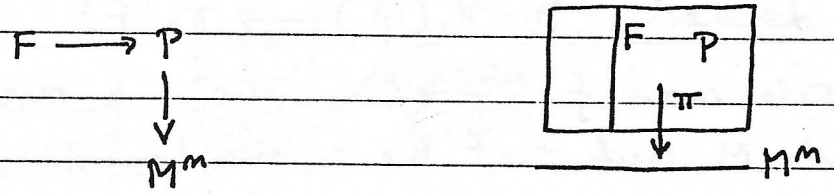
is exact.

If $C_k(M)$ are k-chains on M, $Z_k(M)$ the k-cycles, and $B_k(M)$ the k-boundaries, and $H_k(M)$ the homology classes

$$\left.\begin{aligned} &0 \to Z_k(M) \to C_k(M) \xrightarrow{\partial} B_{k-1}(M) \to 0 \\ &0 \to B_k(M) \to Z_k(M) \xrightarrow{\pi} H_k(M) \to 0 \end{aligned}\right\} \quad (78.5)$$

are both exact sequences. The two unnamed maps $Z_k \to C_k$ and $B_k \to Z_k$ are simply inclusions.

(3) The Exact Homotopy Sequence for a Bundle. For simplicity only, we shall consider a fiber bundle P with connected fiber F

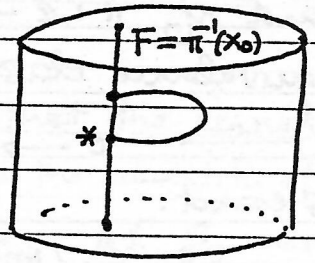


Theorem: If the fiber F is connected, we have the exact sequence of homotopy groups

$$\cdots \to \pi_k(F) \xrightarrow{i_*} \pi_k(P) \xrightarrow{\pi_*} \pi_k(M) \xrightarrow{\partial} \pi_{k-1}(F) \to \cdots$$
$$\cdots \to \pi_2(F) \xrightarrow{i_*} \pi_2(P) \xrightarrow{\pi_*} \pi_2(M) \xrightarrow{\partial} \pi_1(F) \xrightarrow{i_*} \pi_1(P) \xrightarrow{\pi_*} \pi_1(M) \to 0 \quad (78.6)$$

i_* is the homomorphism coming from inclusion $F \xrightarrow{i} P$, π_* comes from the projection $\pi: P \to M$. We must explain the remaining homomorphism $\partial: \pi_k(M) \to \pi_{k-1}(F)$. We illustrate this for the case $k=2$, though our explanation is general. In fact, let us consider $k=1$ first.

First we must indicate base points in F, P, & M.
Take for base point in M some $x_0 \in M$.
Let F be the specific fiber $\pi^{-1}(x_0)$.
Let $*$ be an arbitrary (fixed) point in F. We shall use $*$ as base point for both F and P.

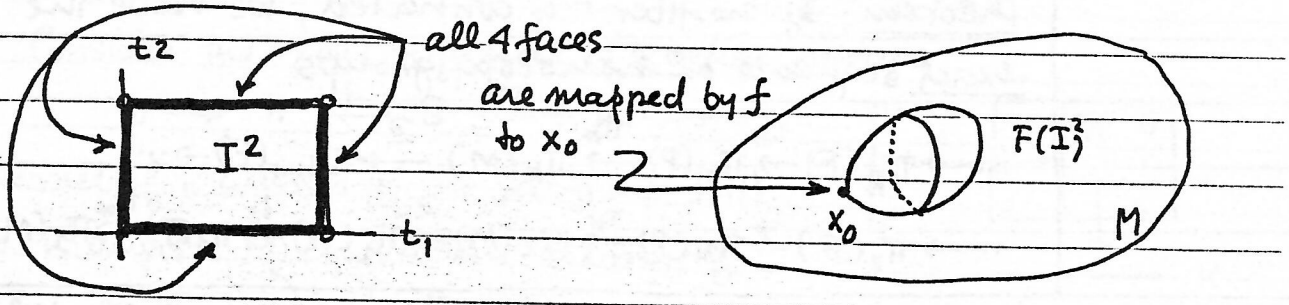


Let $f: S^1 \longrightarrow M$ be a map

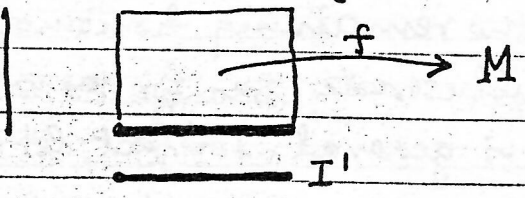


of a 1-sphere into M, based at $x_0 \in M$. We may consider this as a homotopy of the constant map of a point into M. This loop can be covered (via the covering homotopy theorem) by the image of $[0,1]$ starting at $*$ and ending at some other point of F. This pair of points clearly can be considered as the map of a 0-sphere into F, since a zero sphere is the boundary of the interval $\overset{0}{\circ}\!\!-\!\!-\!\!\overset{1}{\circ}$, ie, S^0 is 2 pts. This 0-sphere in F is by definition $\partial(f)$, ie we have defined $\partial : \pi_1(M) \longrightarrow \pi_0(F)$.

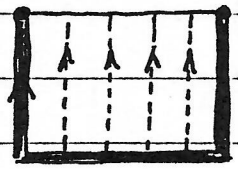
Consider now $f: S^2 \longrightarrow M$. This is a map of a square I^2 into M such that the entire boundary $\dot{I}^2$, ie all 4 faces, are mapped into $x_0 \in M$



Now this map $f: I^2 \longrightarrow M$ can be considered as a homotopy of the map f restricted to the face I^1 defined by $t_2 = 0$



f restricted to this face is of course the constant map $I^1 \to x_0 \in M$. The base point $*$ of P lies over the point x_0. By the covering homotopy property, the homotopy f can be covered by a homotopy of the constant map $I^1 \to * \in P$. Under f, the two sides and the bottom of the square are mapped constantly into x_0 and so when we consider the vertical deformation lines on $f(I^2)$, the lines of the two sides will be covered by constant maps into the point $*$. However, when covering the top edge,

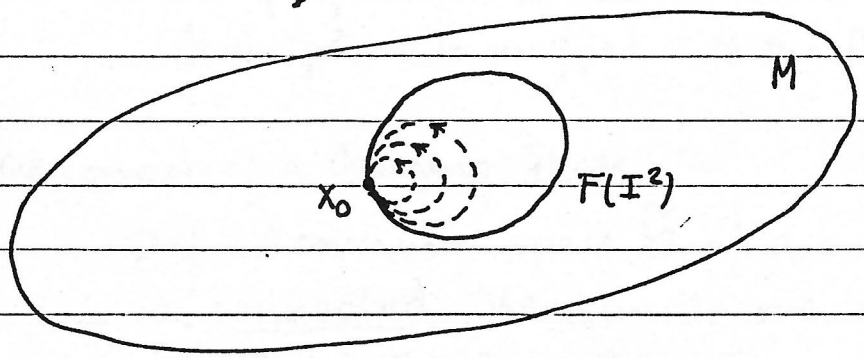


we only know that this edge will be covered by a closed loop lying in F



This loop in F is by definition $\partial(f)$ in

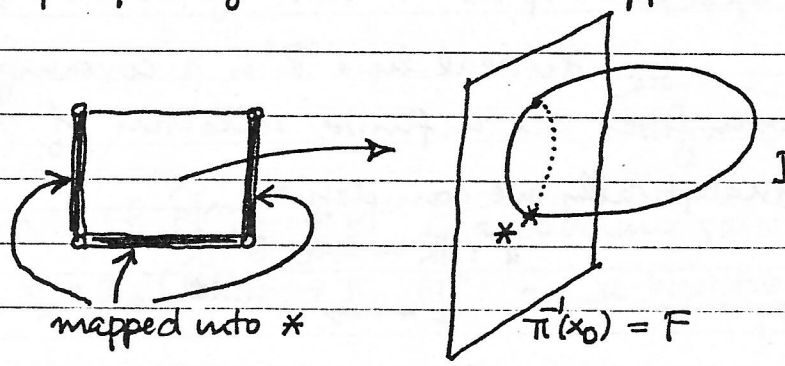
$$\partial : \pi_2(M) \to \pi_1(F).$$



We shall not prove exactness of the sequence, though some of the steps are almost immediate.

Lecture 80. Some Examples of $\pi_k(M^n)$. The Hopf Map.

(1) **Covering Spheres.** In our description of the homomorphism $\partial : \pi_k(M^n, x_0) \longrightarrow \pi_{k-1}(P, *)$ in the exact homotopy sequence of a bundle we have actually demonstrated the following fact which is of independent interest. We had considered the map $f: I^2 \longrightarrow M$ of a $k=2$ sphere into M^n (recall that $\dot{I}^2$ was mapped into x_0) and we had covered this by a map of I^2 into P in which all sides except the top $t_2 = 1$ were mapped into the base point $*$, and the remaining top edge $t_2 = 1$ was mapped into the fiber $\pi^{-1}(x_0)$



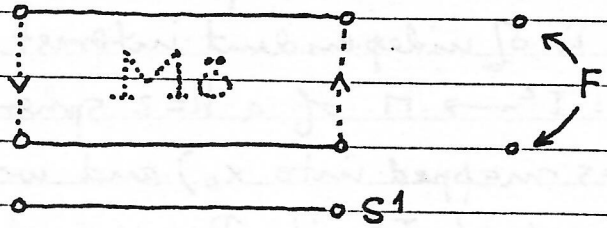
It should be clear that we can consider this as a map of a 2 disc into P in which the boundary circle ($t_2 = 1$) is mapped into F.

> **Proposition.** Any map of a k sphere into M^n (with base point x_0) can be covered by a map of k-disc into P in which the boundary $(k-1)$ sphere is mapped into the fiber $\pi^{-1}(x_0)$ (80.1)

(2) **The Concept of a Covering Space.**

> **Def.** A covering space $\tilde{M}$ of a manifold M^n is a connected fiber bundle over base space M^n $\tilde{M} \xrightarrow{\pi} M^n$ whose fiber F consists of a discrete set of points

Ex. The edge of a (finite) Möbius band is a covering space of the circle S^1 with fiber consisting of 2 points



Note that $\tilde{S}^1$ is again topologically just S^1.

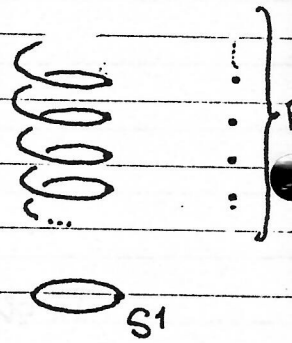
Ex. The n-sphere S^n is a covering space of the real projective n space $P^n(\mathbb{R})$ again with fiber 2 points

Ex. The real line $\mathbb{R}^1$ is a covering space of the circle S^1 with fiber an infinite collection of points.

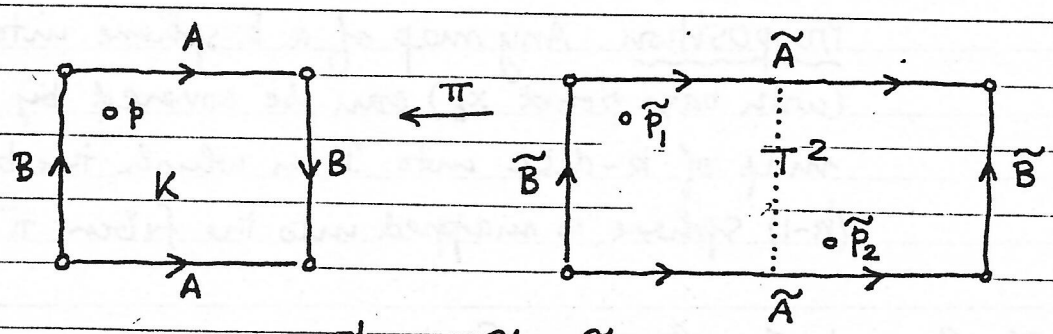
Analytically we can define

$$\pi : \mathbb{R} \longrightarrow S^1$$
$$t \longmapsto e^{2\pi i t}$$

by



Ex. The 2-torus T^2 is 2-fold covering space of the Klein bottle



$$\pi^{-1}(p) = \tilde{p}_1 \cup \tilde{p}_2$$

Ex. Real n-space $\mathbb{R}^n$ is an ∞-fold covering of the n-torus T^n, $\pi : \mathbb{R}^n \longrightarrow T^n = S^1 \times \cdots \times S^1$

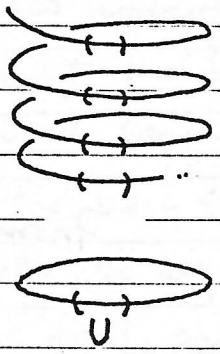
$$(t_1, \ldots, t_n) \longmapsto (e^{2\pi i t_1}, \ldots, e^{2\pi i t_n}).$$

Of course the notion of covering space can be given without talking about fiber bundles

Def'. $\tilde{M} \xrightarrow{\pi} M^n$ is a covering manifold of the manifold M^n if $\tilde{M}$ is connected and if each $x \in M^n$ has a nbhd U such that $\pi^{-1}(U)$ consists of disjoint open subsets $U_1, U_2, \ldots$ of $\tilde{M}$ such that each

$$\pi : U_\alpha \longrightarrow U$$

is a diffeomorphism.



Theorem. If $\tilde{M}^n \xrightarrow{\pi} M^n$ is a covering space then
$\pi_* : \pi_k(\tilde{M}^n) \xrightarrow{\simeq} \pi_k(M^n)$ is a bijection for $k \geq 2$ (80.2)

Proof. First show $\pi_* : \pi_k(\tilde{M}^n) \longrightarrow \pi_k(M^n)$ is 1:1.

Suppose that $f(I^k)$ is a k-sphere at in $\tilde{M}^n$ that when projected down $\pi \circ f(I^k)$ is homotopic to a constant We can cover this homotopy to yield a homotopy of $f(I^k)$ into the fiber $\pi^{-1}(x_0)$. This would yield a map of I^k into a discrete set of points. Since I^k is connected we conclude I^k would be mapped into the single point $*$. Thus if $\pi \circ f(I^k)$ is contractible to a point x_0 then $f(I^k)$ is contractible to the point $*$, and so is trivial. $\therefore \pi_*$ is 1:1.

We must show that π_* is onto. Let $f(I^k)$ be a k-sphere based at $x_0 \in M^n$. By proposition (80.1) this k-sphere can be covered by a k disc on $\tilde{M}$ whose boundary k-1 sphere lies in $\pi^{-1}(x_0)$. Since

$k-1 \geq 1$ this sphere is connected and again must collapse to the point $*$. Thus the k-disc is actually the map of a k-sphere into $\tilde{M}^n$, and π_* maps this k-sphere onto the given $f(I^k)$. We have shown π_* is onto. □

Note that we did not use $k \geq 2$ in our 1:1 argument.

Corollary. For a covering

$$\pi_* : \pi_1(\tilde{M}^n) \longrightarrow \pi_1(M^n) \tag{80.3}$$

is 1:1, ie injective

From our previous examples we see e.g

$$\left.\begin{aligned} \pi_k S^n &= \pi_k P^n(\mathbb{R}) \\ \pi_k T^n &= \pi_k \mathbb{R}^n = 0 \\ \pi_k(\text{Klein}) &= \pi_k T = 0 \end{aligned}\right\} \text{ for } k \geq 2 \tag{80.4}$$

(3) Back to $SU(n)$. We know that $SU(n)$ is connected and simply connected, $\pi_1 SU(n) = 0$. We now show

$$\pi_2 SU(n) = 0 \tag{80.5}$$

Proof. From the fibering $SU(n-1) \longrightarrow SU(n) \downarrow S^{2n-1}$

we have the exact homotopy sequence

$$\cdots \longrightarrow \pi_2 S^{2n-1} \longrightarrow \pi_2 SU(n-1) \longrightarrow \pi_2 SU(n) \longrightarrow \pi_2 S^{2n-1} \longrightarrow \cdots$$

For $n \geq 3$ this yields

$$0 \longrightarrow \pi_2 SU(n-1) \longrightarrow \pi_2 SU(n) \longrightarrow 0$$

and so $\pi_2 SU(n-1) \cong \pi_2 SU(n)$ for $n \geq 3$,

$$\therefore \quad \pi_2 SU(n) = \pi_2 SU(2) = 0 \qquad n \geq 3$$ □

Fact. E. Cartan has shown that

$$\boxed{\pi_2(G) = 0} \qquad (80.6)$$

for any Lie Group. This is much harder

- Problem 113 Show that $\boxed{\pi_3 SU(n) = \pi_3 SU(2) = \mathbb{Z},\ n \geq 2}$ (80.7)

- Problem 114. We have stated the exact homotopy sequence for a fiber bundle (19.6) in the case that the fiber F is connected. When the fiber is not connected (as in a covering) the only difference is that in the very last term $\pi_*: \pi_1(P) \longrightarrow \pi_1(M)$ need not be onto, so we do not necessarily have $\pi_1(M) \longrightarrow 0$. Accept this fact, and then show that theorem (80.2) is an immediate consequence of this exact sequence.

(4) The Hopf Map and Fibering. The whole starting point for Hurewicz's invention of the higher homotopy groups $\pi_k(M^n)$ must have been Heinz Hopf's discovery of an essential map of a 3-sphere onto a 2-sphere, ie a map $\pi: S^3 \longrightarrow S^2$ that was not homotopic to a constant map. With our machinery we can easily exhibit such a map.

From the fibering

$$\begin{array}{ccc} S^1 = SO(2) & \longrightarrow & SO(3) \\ & & \downarrow \pi \\ & & S^2 \end{array}$$

we see immediately that $P^3(\mathbb{R})$ is a bundle over S^2 with

circle as fiber. Then $\mathbb{Z} \approx \pi_3(S^3) \approx \pi_3 P^3(\mathbb{R}) \approx \pi_3(S^2)$ from the exact sequence, $\therefore$

$$\boxed{\pi_3(S^2) = \mathbb{Z}} \tag{80.8}$$

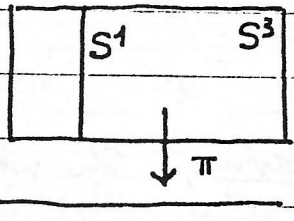
In fact, if we look at the 2-fold covering of $SO(3)$, ie $SU(2) \longrightarrow SO(3)$, we have the following picture

$$\begin{array}{ccc} S^1 & \longrightarrow & SU(2) = S^3 \\ & & \downarrow \pi \\ & & S^2 \end{array}$$

S^1 is the subgroup $\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$

ie S^3 is a fiber bundle over S^2 with (non-intersecting) circles as fibers

This is the Hopf fibration



and

$$\pi_* : \pi_3(S^3) \xrightarrow{\approx} \pi_3(S^2)$$

is a bijection. This says in particular that the Hopf map $\pi : S^3 \longrightarrow S^2$ is not homotopic to the constant map; it is an "essential" map.

There are many other ways of picturing the Hopf map, but time is fleeting.

Lecture 81. Chern Forms as Obstructions

(1) Covariant Differential of Forms of Type Ad.

Proposition. Let B^p be a p-form section of the Ad-bundle. Then

$$d(\mathrm{tr}\, B) = \mathrm{tr}\, DB \tag{81.1}$$

Proof.

$$DB = dB + [\omega, B] = dB + \omega \wedge B - (-1)^p B \wedge \omega$$

and so

$$\mathrm{tr}\, DB = \mathrm{tr}\, dB + 0 \qquad \square.$$

Proposition.

$$\boxed{D^* B = \pm * D * B} \tag{81.2}$$

Proof. Let C be a $(p-1)$ form section of the Ad bundle with small support. Then (see (72.7))

$$(DC, B) = -\frac{1}{2}\int \mathrm{tr}\, DC \wedge * B = -\frac{1}{2}\int \mathrm{tr}\, D(C \wedge * B) \pm \frac{1}{2}\int C \wedge D * B$$

$$= -\frac{1}{2}\int d\, \mathrm{tr}(C \wedge * B) \pm \frac{1}{2}\int \mathrm{tr}(C \wedge D * B).$$

By Stokes theorem and the small support of C the first integral will vanish. We conclude

$$(DC, B) = \pm \frac{1}{2}\mathrm{tr} \int C \wedge **D*B = \pm (C, *D*B) \qquad \square.$$

Corollary. If a Yang Mills field strength $F = (ig)^{-1}\nu$ is either self dual, $F = *F$, or anti self dual, $-F = *F$, then F satisfies the (pure) Yang Mills equations

$$D^* F = \pm * D * F = 0$$

since F automatically satisfies Bianchi $DF = 0$.

(2) <u>The Chern Forms c_r are Closed</u>. In Lecture 7b we defined the Chern forms for a complex vector bundle E over M by

$$\left.\begin{aligned} \det\left(I + \frac{i}{2\pi}\vartheta\right) &= 1 + c_1(E) + c_2(E) + \cdots \\ &= 1 + \frac{i}{2\pi}\operatorname{tr}\vartheta - \frac{1}{8\pi^2}\left[(\operatorname{tr}\vartheta)\wedge(\operatorname{tr}\vartheta) - \operatorname{tr}(\vartheta\wedge\vartheta)\right] + \cdots \end{aligned}\right\} \quad (81.3)$$

$c_r(E)$ is a $2r$ form defined on the base manifold M^n, and we have verified that c_1 and c_2 are closed. In fact all c_r are closed. We shall not give a complete proof. Instead, let us accept the fact (following from "Newton's identities" for symmetric functions) that each c_r can be expressed as a polynomial (using $\wedge$ product) in the forms $\operatorname{tr}\vartheta$, $\operatorname{tr}(\vartheta\wedge\vartheta)$, $\cdots$. We have exhibited c_1 and c_2 as such in (81.3), and you have verified this also in the case c_3 (problem III). But then $dc_r = 0$ will follow then from $d\operatorname{tr}(\vartheta) = \operatorname{tr} D\vartheta = 0$. For example

$$d\operatorname{tr}(\vartheta\wedge\vartheta) = \operatorname{tr} D(\vartheta\wedge\vartheta) = \operatorname{tr}\left(D\vartheta\wedge\vartheta + \vartheta\wedge D\vartheta\right) = 0,$$

from Bianchi.

We are dealing with $U(K)$ bundles, and so $i\vartheta$ is a Hermitian matrix with 2-form entries. It turns out that each c_r is a real form on M^n.

By de Rham's theorem each c_r defines a real cohomology class on M^n; for any real $2r$-cycle Z_{2r} on M^n (basically a compact, oriented $2r$ dimensional submanifold of M^n) we associate the "period"

$$\int_{Z_{2r}} c_r$$

a real number that is the same for all Z'_{2r} that are

homologous to Z_{2r}. $\int c_r$ is a linear form on $H_{2r}(M;\mathbb{R})$

In fact it turns out that if we consider integer cycles, ie $Z_{2r} \in H_{2r}(M^m;\mathbb{Z})$ then $\int_{Z_{2r}} c_r$ is always an integer.

For example, when $M = M^2$ and E is the complex line bundle version of the tangent bundle to M^2, M^2 compact

$$\int_{M^2} c_1 = \int_{M^2} \frac{i}{2\pi}\tilde{\vartheta} = \frac{1}{2\pi}\int_{M^2} K\,dS = \chi(M^2)$$

is the Euler characteristic of M^2. This integer, by Poincarés theorem

$$\chi(M^2) = \sum j_p(v)$$

is the sum of the indices of any vector field v defined on M whose only zeros (if any) are isolated. We can say that $\int_M c_1$ represents the "obstruction" to constructing a non-vanishing cross section of the tangent bundle to M; certainly if $\int_M c_1 \neq 0$ there is no chance of constructing a non-vanishing crosssection to TM over all of M.

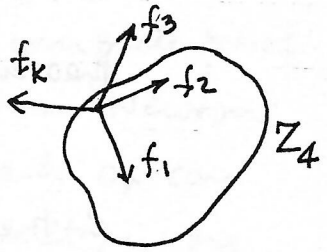
It turns out that all of the Chern cforms, c_r, have such an interpretation, representing obstructions to constructing cross sections of various kinds, associated with the given bundle E. We shall not do this, but we will consider in detail the case of $c_2(E)$ for a vector bundle with structure group $SU(k)$.

(3) $c_2(E)$ as an obstruction "cocycle". Let $E \longrightarrow M^m$ be a vector bundle over M^m with structure group $SU(k)$. We wish to evaluate

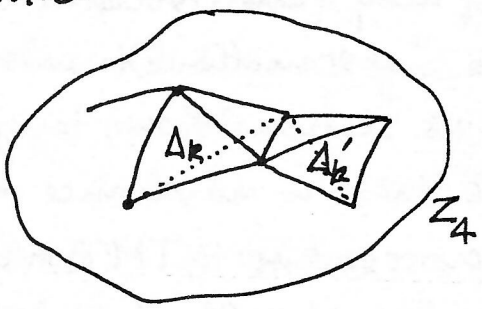
$$\int_{Z_4} c_2$$

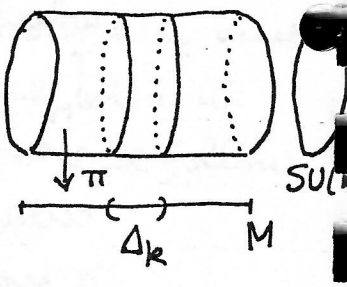
where Z_4 is an integral 4-cycle on M^n, ie it represents an element of $H_4(M^n; \mathbb{Z})$. For simplicity we shall take for Z_4 a compact oriented 4 dimensional submanifold of M^n.

We consider the problem of constructing a <u>frame</u> of k linearly independent cross sections of E over points of the submanifold Z_4. Since $SU(k)$ is the structure group, we can say that we shall try to construct a cross section of the principal $SU(k)$ bundle associated to E, over the subset Z_4.



We triangulate Z_4 into simplexes, each of which is so small that the part of our principal bundle over each Δ_k is trivial

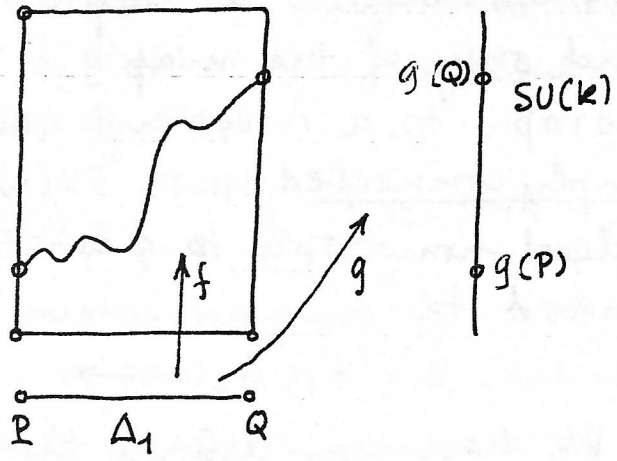




Now we begin to construct our cross section. Over each 0-simplex (vertex) Δ_0 we pick arbitrarily a group element $f(\Delta_0)$ (remember, $\pi^{-1}(\Delta_4) \approx \Delta_4 \times SU(k)$ is trivial).

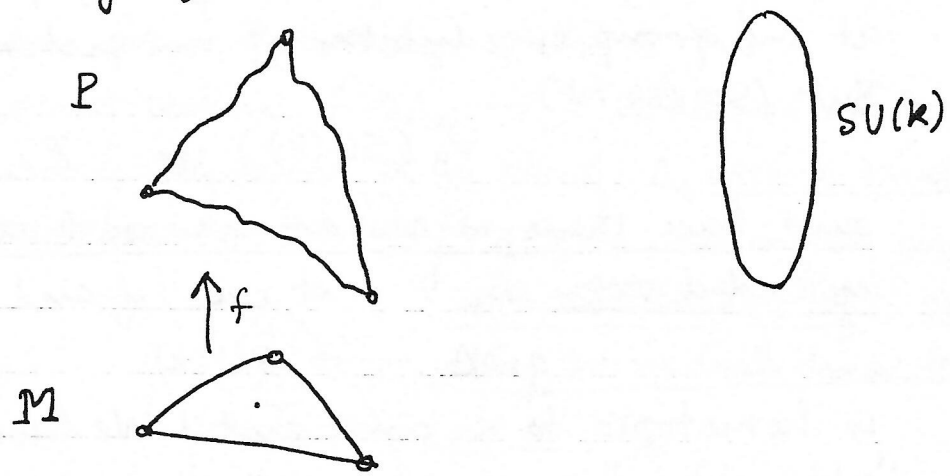
We have thus constructed a cross section over the "0-skeleton" of Z_4.

We now wish to extend this given cross section to a cross section defined over each 1-simplex.



We have, f defined on the boundary of each 1-simplex Δ_1 Can we extend f to all of Δ_1? Since SU(k) is connected the answer is certainly, yes!, there are many ways to extend over Δ_1. Choose one for each Δ_1.

We now wish to extend our cross section over each 2-simplex Δ_2. We have already defined f on the boundary of each 2 simplex



Again, since $\pi^{-1}(\Delta_4) = \Delta_k \times SU(k)$ we effectively have a map

$$g : \partial\Delta_2 \longrightarrow SU(k)$$

simply by looking at the $G = SU(k)$ coordinates of the cross section f. $\partial\Delta_2$ is topologically a circle S^1 We know that a map of a circle $= \partial\Delta_2$ into SU(k)

can be extended to a map of all of Δ_2 into $SU(k)$ if and only if the map $g : \partial\Delta_2 \longrightarrow SU(k)$ is homotopic to a constant (See (78.1)). But $SU(k)$ is <u>simply connected</u>, $\pi_1 SU(k) = 0$, and so $g : \partial\Delta_2 \to SU$ is indeed homotopic to a constant and f can be extended to f

$$f : \Delta_2 \longrightarrow \pi^{-1}(\Delta_4) \subset P.$$

We have then defined the cross section to each $\partial\Delta_3$. Can we extend to Δ_3? $\partial\Delta_3$ is a 2-sphere and we again have a map $g : S^2 = \partial\Delta_3 \longrightarrow SU(k)$. But the 2nd homotopy group $\pi_2(SU(k)) = 0$ (see (80.5)) and so f can be extended to

$$f : \Delta_3 \longrightarrow \pi^{-1}(\Delta_4) \subset P.$$

We have defined f on each $\partial\Delta_4$ in Z_4; can we extend to each Δ_4? $\partial\Delta_4$ is a 3-sphere. By looking at the group coordinates of our section, $g : \partial\Delta_4 = S^3 \longrightarrow SU$. But (See (80.7))

$$\pi_3(SU(k)) \approx \mathbb{Z} \qquad (81.$$

and <u>thus there is no assurance that f can be extended over Δ_4</u>!; it can if and only if

$$g : \partial\Delta_4 \longrightarrow SU(k)$$

is homotopic to a constant! We have met an "<u>obstruction</u>".

In all of the above there has been no use made of the cycle Z_4. For any Δ_4 in a triangulation of M^n (subject to $\pi^{-1}(\Delta_4)$ is a product $\Delta_4 \times SU(k)$) we have associated, via (81.3) an "integer". Such an assignment is basically a 4-"cochain".

We need to explain this integer more fully. We know from Prob. III $\pi_3 SU(k) = \mathbb{Z}$ is the result of 2 facts. First $SU(2)$ is a subgroup of $SU(k)$

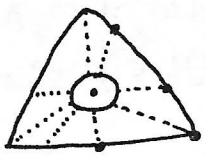
$$SU(2) = \begin{pmatrix} SU(2) & 0 \\ 0 & 1 \ddots 1 \end{pmatrix} \subset SU(k) \qquad (81.4)$$

and the exact homotopy sequences for $SU(2) \subset SU(3) \subset \cdots \subset SU(k)$ give $\pi_3 SU(k) = \pi_3 SU(2) = \pi_3 S^3 = \mathbb{Z}$. That is $SU(2) \subset SU(k)$ is a 3-sphere that realizes the generator of the group $\pi_3 SU(k)$, ie any 3-sphere mapped into $SU(k)$ can be deformed so as to have its image lie on the $SU(2)$ subgroup!

Then for our partial cross section $g: \partial\Delta_4 \longrightarrow SU(k)$ this map of $S^3 = \partial\Delta_4$ can be deformed to give a map $\partial\Delta_4 \longrightarrow SU(2) \subset SU(k)$. This map $S^3 \longrightarrow S^3$ has a degree, and this degree is the integer we associate to the 4-simplex Δ_4.

We propose now to evaluate this integer by integrating the 2nd Chern form c_2 over Δ_4! Why is this?

The cross section $f: \partial\Delta_4 \longrightarrow P$ is defined on $\partial\Delta_4$; but we can immediately extend it to all of Δ_4 with a small 4 ball removed about the barycenter x_0 of Δ_4. We merely make the $SU(k)$ coordinates of f constant along the radial lines leading out to $\partial\Delta_4$.



Let (ε-ball) be the ball of radius ε centered at the barycenter $x_0 \in \Delta_4$. We are going to compute

$$\int_{\Delta_4} c_2 = \lim_{\varepsilon \to 0} \int_{\Delta_4 - (\varepsilon\,\text{ball})} c_2$$

We can be very brief since the procedure mimicks that given in Lecture 62 for Chern's proof of the Poincare theorem.

Let $\Sigma_4 = f(\Delta_4 - \varepsilon\text{ball})$ be the graph of our local section. Then

$$\int_{\Delta_4 - \varepsilon\text{ball}} c_2 = \int_{\pi(\Sigma_4)} c_2 = \int_{\Sigma_4} \pi^* c_2$$

Now

$$c_2 = -\frac{1}{8\pi^2}\left[\operatorname{tr}\vartheta \wedge \operatorname{tr}\vartheta - \operatorname{tr}\vartheta\wedge\vartheta\right] = \frac{1}{8\pi^2}\operatorname{tr}(\vartheta\wedge\vartheta) \tag{81.5}$$

since $c_1 = \frac{i}{2\pi}\operatorname{tr}\vartheta = 0$ for an $SU(k)$ connection (traceless skew hermitian ω and ϑ); <u>the first Chern form vanishes for a $SU(k)$ bundle!</u>

Now let e_U be the frame on the open $U \subset M^4$ holding Δ_4 for which $\pi^{-1}(\Delta_4)$ is a product; $\pi^{-1}U \approx U \times SU(k)$. $\vartheta = \vartheta_U$ on U.

$$c_2 = \frac{1}{8\pi^2}\operatorname{tr}\vartheta_U \wedge \vartheta_U$$

$$\pi^* c_2 = \frac{1}{8\pi^2}\operatorname{tr}\pi^*\vartheta_U \wedge \pi^*\vartheta_U$$

$$= \frac{1}{8\pi^2}\operatorname{tr}\left\{g^{-1}\pi^*\vartheta_U\, g \wedge g^{-1}\pi^*\vartheta_U\, g\right\}$$

$$\pi^* c_2 = \frac{1}{8\pi^2}\operatorname{tr}(\tilde{\vartheta}_U \wedge \tilde{\vartheta}_U) \tag{81.6}$$

(see 64.7). Now $\tilde{\vartheta}_U = d\tilde{\omega}_U + \tilde{\omega}_U \wedge \tilde{\omega}_U$, (64.6), and the same calculation that gave (76.5) shows

$$\pi^* c_2 = \frac{1}{8\pi^2}\operatorname{tr}(\tilde{\vartheta}\wedge\tilde{\vartheta}) = d\,\frac{1}{8\pi^2}\operatorname{tr}\left\{\tilde{\omega}\wedge d\tilde{\omega} + \frac{2}{3}\tilde{\omega}\wedge\tilde{\omega}\wedge\tilde{\omega}\right\} \tag{81.}$$

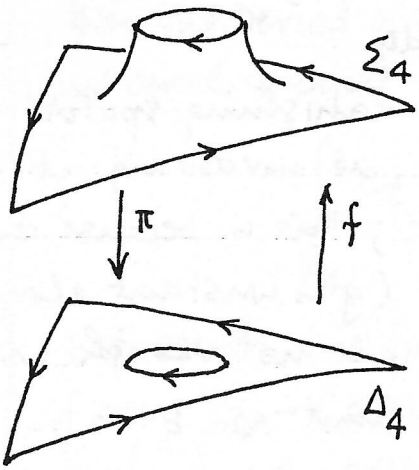
the 3-form

$$\operatorname{tr}\left\{\tilde{\omega}\wedge d\tilde{\omega} + \frac{2}{3}\,\tilde{\omega}\wedge\tilde{\omega}\wedge\tilde{\omega}\right\}$$

is the "<u>Chern-Simons</u>" 3-form.

$$\int_{\Sigma_4} \pi^* c_2 = \frac{1}{8\pi^2}\int_{\partial\Sigma_4} \operatorname{tr}\left\{\tilde{\omega}_U\wedge d\tilde{\omega}_U + \frac{2}{3}\,\tilde{\omega}_U\wedge\tilde{\omega}_U\wedge\tilde{\omega}_U\right\} \tag{81.8}$$

Now $\partial\Sigma_4$ consists of 2 pieces



← This is a bad picture. The section should have constant height along radial lines; this is the way we designed the extension

(i) The first part consists of $f(\partial\Delta_4)$, the part of the section over the edge of Δ_4. If we integrate c_2 over all of the Δ_4 of the given 4-cycle Z_4 it is clear that all the integrals over these $f(\partial\Delta_4)$ pieces will cancel, as usual.

(ii) The second part consists of the part of the section over the boundary of the ε-ball but with orientation opposite to the usual boundary of this ball.

From (81.8)

$$\int_{U\Sigma_4} \pi^* c_2 = -\frac{1}{8\pi^2} \sum_{\varepsilon\text{balls}} \int_{\partial(\varepsilon\text{-balls})} \mathrm{tr}\left\{ \tilde{\omega}_U \wedge d\tilde{\omega}_U + \frac{2}{3} \tilde{\omega}_U \wedge \tilde{\omega}_U \wedge \tilde{\omega}_U \right\} \quad (81.9$$

Now

$$\tilde{\omega}_U = g^{-1} \pi^* \omega_U \, g + g^{-1} dg$$

and so our 3 fold integrals will involve terms containing $g^{-1}dg$ and $\omega_U = \Gamma(x)\,dx$. We are going to let the radius ε of the ε-balls tend to 0. The integrals we will

get will involve

$$dx\wedge dg\wedge dg \quad \text{or} \quad dx\wedge dx\wedge dg \quad \text{or} \quad dx\wedge dx\wedge dx \quad \text{or} \quad dg\wedge dg\wedge dg$$

Since the ε balls will have vanishing spatial extent, any integral in the first line, ie involving at least one dx will vanish in the limit; this is because the g integrals will be independent of ε (g is constant along radial lines). The only integrals that will not vanish will involve $dg\wedge dg\wedge dg$. Thus, in the limit as $\varepsilon\to 0$ (81.9) will give

$$\int_{Z_4}\pi^* c_2 = \lim_{\varepsilon\to 0} -\frac{1}{8\pi^2}\sum_{\varepsilon\text{balls}}\int_{\partial(\varepsilon\text{ balls})} \mathrm{tr}\left\{\bar g^{-1}dg\wedge d(\bar g^{-1}dg)+\frac{2}{3}\bar g^{-1}dg\wedge\bar g^{-1}dg\wedge\bar g^{-1}dg\right.$$

$$= \frac{1}{24\pi^2}\sum\int_{\partial(\varepsilon\text{balls})} \mathrm{tr}\ \bar g^{-1}dg\wedge\bar g^{-1}dg\wedge\bar g^{-1}dg$$

Now each map $g:\partial(\varepsilon\text{-ball})\longrightarrow SU(k)$ that arises from the cross section $f:\partial\Delta_4\to P$, is homotopic to a map $\hat g:\partial(\varepsilon\text{ ball})\longrightarrow$ the $SU(2)$ subgroup of $SU(k)$, as was explained following (81.4). Now from (76.3)

$$\frac{1}{24\pi^2}\int_{SU(2)}\mathrm{tr}\ g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg = 1$$

and so

$$\frac{1}{24\pi^2}\int_{\partial(\varepsilon\text{ ball})}\mathrm{tr}\,g^{-1}dg\wedge g^{-1}dg\wedge g^{-1}dg$$

is indeed the Brouwer degree of the map $:\partial(\varepsilon\text{-ball})\xrightarrow{\hat g} SU(2)$!

In conclusion, we have shown the following

CHERN'S

Theorem. Let P be a principal $SU(k)$ bundle over M^4. Then the period of the 2nd Chern form c_2 over an integral 4-cycle Z_4

$$\int_{Z_4} c_2$$

represents the following. There always exists a cross section f over $Z_4 - (\text{a finite union } \bigcup_\alpha p_\alpha \text{ of points})$. About each p_α we construct a small 3-sphere S^3_α and map it into $SU(k)$ via the section f. Each image S^3_α in $SU(k)$ is deformable so as to lie on the $SU(2)$ subgroup of $SU(k)$. Let $j_\alpha(f)$ denote the number of times the image of S^3_α covers $SU(2)$, i.e

$$j_\alpha(f) = \text{the Brouwer degree of } S^3_\alpha \longrightarrow SU(2).$$

Then

$$\int_{Z_4} c_2 = \sum_\alpha j_\alpha(f).$$

In particular $\sum_\alpha j_\alpha(f)$ is <u>independent of the section f chosen</u>! It is also immediate that

$$\int_{Z_4} c_2$$

is always an <u>integer</u>. It is also clear that

$$\int_{Z_4} c_2 = \frac{i}{8\pi^2} \int_{Z_4} \operatorname{tr} \vartheta \wedge \vartheta$$

is in fact <u>independent of the $SU(k)$ connection</u> used in the bundle!