THE GEOMETRIZATION OF PHYSICS

RICHARD S. PALAIS

June-July, 1981

LECTURE NOTES IN MATHEMATICS

INSTITUTE OF MATHEMATICS

NATIONAL TSING HUA UNIVERSITY

HSINCHU, TAIWAN, R.O.C.

**With Support from National Science Council
Republic of China**

THE GEOMETRIZATION OF PHYSICS

RICHARD S. PALAIS*

LECTURE NOTES FROM A COURSE AT
NATIONAL TSING HUA UNIVERSITY
HSINCHU, TAIWAN
JUNE-JULY, 1981

*RESEARCH SUPPORTED IN PART BY:
THE NATIONAL SCIENCE FOUNDATION (USA)
AND THE NATIONAL RESEARCH COUNCIL (ROC)

U.S. and Foreign Copyright reserved by the author.

Reproduction for purposes

of Tsing Hua University or the

National Research Council permitted.

Acknowledgement

I would first like to express my appreciation to the National Research Council for inviting me to Taiwan to give these lectures, and for their financial support.

To all the good friends I made or got to know better while in Taiwan, my thanks for your help and hospitality that made my stay here so pleasantly memorable. I would like especially to thank Roan Shi-Shih, Lue Huei-Shyong, and Hsiang Wu-Chong for their help in arranging my stay in Taiwan.

My wife, Terng Chuu-Lian was not only a careful critic of my lectures, but also carried out some of the most difficult calculations for me and showed me how to simplify others.

The mathematicians and physicists whose work I have used are of course too numerous to mention, but I would like to thank David Bleecker particularly for permitting me to see and use an early manuscript version of his forth coming book, "Gauge Theory and Variational Principles".

Finally I would like to thank Miss Chu Min-Whi for her careful work in typing these notes and Mr. Chang Jen-Tseh for helping me with the proofreading.

Preface

In the Winter of 1981 I was honored by an invitation, from the National Science Council of the Republic of China, to visit National Tsing Hua University in Hsinchu, Taiwan and to give a six week course of lectures on the general subject of "gauge field theory". My initial expectation was that I would be speaking to a rather small group of advanced mathematics students and faculty. To my surprise I found myself the first day of the course facing a large and heterogeneous group consisting of undergraduates as well as faculty and graduate students, physicists as well as mathematicians, and in addition to those from Tsing Hua a sizable group from Taipei, many of whom continued to make the trip of more than an hour to Hsinchu twice a week for the next six weeks. Needless to say I was flattered by this interest in my course, but beyond that I was stimulated to prepare my lectures with greater care than usual, to add some additional foundational material, and also I was encouraged to prepare written notes which were then typed up for the participants. This then is the result of these efforts.

I should point out that there is basically little that is new in what follows, except perhaps a point of view and style. My goal was to develop carefully the mathematical tools necessary to understand the "classical" (as opposed to "quantum") aspects of gauge fields, and then to present the essentials, as I saw them, of the physics.

A gauge field, mathematically speaking, is "just a connection". It is now certain that two of the most important "forces" of physics, gravity and electromagnetism are gauge fields, and there is a rapidly growing segment of the theoretical physics community that believes not only that the same is true for the "rest" of the fundamental forces of physics (the weak and strong nuclear forces, which seem to manifest themselves only in the quantum mechanical domain) but moreover that all these forces are really just manifestations of a single basic "unified" gauge field. The major goal of these notes is to develop, in sufficient detail to be convincing, an observation that basically goes back to Kuluza and Klein in the early 1920's that not only can gauge fields of the "Yang-Mills" type be unified with the remarkable successful Einstein model of gravitation in a beautiful, simple, and natural manner, but also that when this unification is made they, like gravitational field.

disappear as forces and are described by pure geometry, in the sense that particles simply move along geodesics of an appropriate Riemannian geometry.

Contents

Course outline.

a) Outline of smooth vector bundle theory.

b) Connections and curvature tensors (alias gauge potential and gauge fields).

c) Characteristic classes and the Chern-Weil homomorphism.

d) The principal bundle formalism and the gauge transformation group.

e) Lagrangian field theories.

f) Symmetry principles and conservation laws.

g) Gauge fields and minimal coupling.

h) Electromagnetism as a gauge field theory.

i) Yang-Mills fields and Utiyama's theorem.

j) General relativity as a Lagrangian field theory.

k) Coupling gravitation to Yang-Mills fileds (generalized Kaluza-Klein theories).

l) Spontaneous symmetry breaking (Higg's Mechanism).

m) Self-dual fields, instantons, vortices, monopoles.

References.

a) Gravitation, Gauge Theories, and Differential Geometry; Eguchi, Gilkey, Hanson, Physics Reports vol. 66, No. 6, Dec. 1980.

b) Intro. to the fiber bundle approach to Gauge theories, M. Mayer; Springer Lecture Notes in Physics, vol. 67, 1977.

c) Gauge Theory and Variational Principles, D Bleecker (manuscript for book to appear early 1982).

d) Gauge Natural Bundles and Generalized Gauge Theories, D. Eck, Memoiks of the AMS (to appear Fall 1981).

[Each of the above has extensive further bibliographies].

Some Motivational Remarks:

<u>The Geometrization of Physics in the 20$^{\text{th}}$ Century</u>.

Suppose we have n particles with masses $m_1, \ldots, m_n$ which at time t are at $\vec{x}_1(t), \ldots, \vec{x}_n(t) \in \mathbb{R}^3$. How do they move? According to Newton there are functions $\vec{f}_i(\vec{x}_1, \ldots, \vec{x}_n) \in \mathbb{R}^3$ (f_i is force acting on i^{th} particle) such that

$$m_i \frac{d^2\vec{x}_i}{dt^2} = \vec{f}_i$$

---

(1) $x(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^{3n}$ fictitious particle in $\mathbb{R}^{3n}$

$$F = (\frac{1}{m_1}\vec{f}_1, \ldots, \frac{1}{m_n}\vec{f}_n)$$

$$\frac{d^2x}{dt^2} = F$$

Introduce high dimensional space into mathematics

"Free" particle (non-interacting system)

$$F = 0$$

$$\frac{d^2x}{dt^2} = 0 \qquad x = x^0 + tv$$

$$x_i = x_i^0 + tv_i$$

Note: image only depends on $\dfrac{v}{\|v\|}$!

Particle moves in straight line (geometric).

(2)

$$\delta \int_{t_1}^{t_2} K(\frac{dX}{dt})dt = 0$$ Lagrang's Principle of Least Action

$$[K(\frac{dX}{dt}) = \frac{1}{2} \sum_i m_i \left\| \frac{dx_i}{dt} \right\|^2]$$ Riemann metric.

Extremals are geodesics parametrized proportionally to arc length. (pure geometry!)

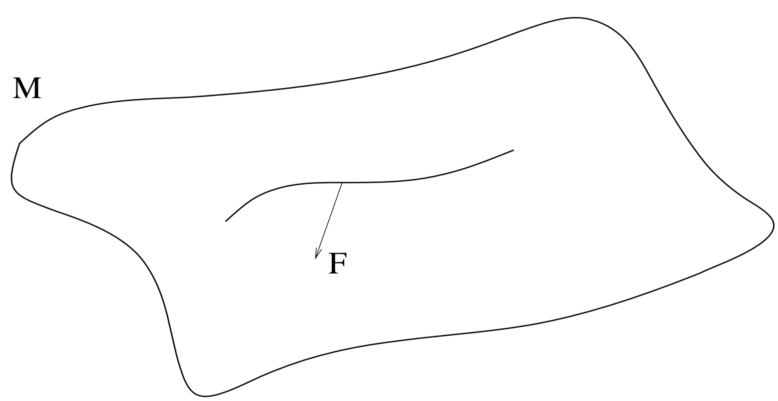
"Constraint Forces" only.

(3) $M \subseteq \mathbb{R}^{3n}$ given by $G(X) = 0$

$$G : \mathbb{R}^{3n} \to \mathbb{R}^k \qquad G(X) = (G_1(X), \ldots, G_k(X))$$
$$G_j(X) = G_j(x_1, \ldots, x_n)$$

Example: Rigid Body $\|x_i - x_j\|^2 = d_{ij} \quad i, j = 1, \ldots, n \quad (k = n(n+1)/2)$

Force F normal to M



K defines an induced Riemannian metric on M. Newton's equations still equivalent to:

$$\delta_M \int_{t_1}^{t_2} K(\frac{dX}{dt})dt = 0$$ $\left\{ \delta_M : \text{ only vary w.r.t. paths in } M. \right.$

OR

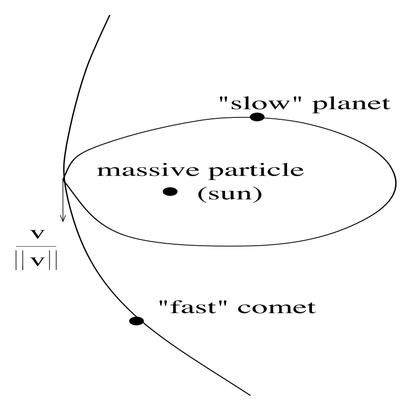
Path of particle is a geodesic on M parametrized proportionally to arc length (Introduces manifolds and Riemannian geometry into physics and mathematics!).

General Case: $F = \nabla V$ (conservation of energy).

$$\begin{aligned} L(X, \frac{dX}{dt}) &= k(\frac{dX}{dt}) - V(X) \\ \delta_M \int_{t_1}^{t_2} L dt &= 0 \end{aligned}$$

(possibly with constraint forces too)

Kinematical dilemma



Can these be geodesics (in the constraint manifold M) w.r.t. some Riemannian metric?

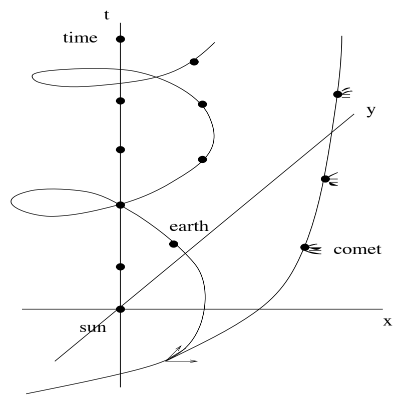
Geodesic image is determined by the direction of any tangent vector. A slow particle and a fast part with same initial direction in gravitational field of massive particle have different pathsin space.

Nevertheless it has been possible to get rid of forces and bring back geometry — in the sense of making particle path geodesics — by "expanding" our ideas of "space" and "time". Each fundamental force took a new effort.

Before 1930 the known forces were gravitation and electromagnetism.

Since then two more fundamental forces of nature have been recognized — the "weak" and "strong" nuclear forces. These are very short range forces — only significant when particles are within 10^{-18} cm. of each other, so they cannot be "felt" like gravity and electromagnetism which have infinite range of action.

The first force to be "geometrized" in this sense was gravitation, by Einstein in 1916. The "trick" was to make time another coordinate and consider a (pseudo) Riemannian structure in space-time $\mathbb{R}^3 \times \mathbb{R} = \mathbb{R}^4$. It is easy to see how this gets rid of the <u>kinematic</u> dilemma:



If we parametrize a path by its length function then a slow and fast particle with the same initial direction $\vec{r} = (\frac{dx_1}{ds}, \frac{dx_2}{ds}, \frac{dx_3}{ds})$ have <u>different</u> initial directions in space time $(\frac{dx_1}{ds}\ \frac{dx_2}{ds}\ \frac{dx_3}{ds}\ \frac{dt}{ds})$, since $\frac{dt}{ds} = \frac{1}{v}$ is just the reciprocal of the velocity. Of course there is still the (much more difficult) problem of finding the correct <u>dynamical</u> law, i.e. finding the physical law which determines the metric giving geodesics which model gravitational motion.

The quickest way to guess the correct dynamical law is to compare Newton's law $\frac{d^2x_i}{dt^2} = -\frac{\partial v}{\partial x_i}$ with the equations for a geodesic $\frac{d^2x_\alpha}{ds^2} + \Gamma^\alpha_{\beta\gamma} \frac{dx_\beta}{ds} \frac{dx_\gamma}{ds}$. Then assuming static weak gravitational fields and particle speeds small compared with the speed of light, a very easy calculation shows that if $ds^2 = g_{\alpha\beta} dx_\alpha dx_\beta$ is approximately $dx_4^2 - (dx_1^2 + dx_2^2 + dx_3^2)$ $(x_4 = t)$ then $g_{44} \sim 1 + 2V$. Now Newton's law of gravitation is essentially equivalent to:

$$\Delta V = 0$$

or $$\delta \int |\nabla_V|^2 dv = 0$$

(where variation have compact support).

So we expect a second order PDE for the metric tensor which is the Euler-Lagrange equations of a Lagrangian variational principle $\delta \int L\, dv = 0$. Where L is some scalar function of the metric tensor and its derivatives. A classical invariant these argument shows that the only such scalar with reasonable invariance properties with respect to coordinate transformation (acting on the metric) is the scalar curvature — and this choice in fact leads to Einstein's gravitational field equations for empty space [cf. A. Einstein's "The Meaning of Relativity" for details of the above computation].

What about electromagnetism?

Given by two force fields $\vec{E}$ and $\vec{B}$.

The force on a particle of electric charge q moving with velocity $\vec{v}$ is:

$$q(\vec{E} + \vec{v} \times \vec{B}) \qquad \text{(Lorentz force)}$$

If in 4-dimensional space-time we define a 2-norm $F = \sum_{\alpha<\beta} F_{\alpha\beta} dx_\alpha \wedge dx_\beta$ (i.e. a skew 2-tensor, the Faraday tensor) by

$$F = E_i dx_i \wedge dx_4 + \frac{1}{2} B_i e_{ijk} dx_j \wedge dx_k$$

so

$$F = \begin{pmatrix} 0 & B_3 & -B_2 & E_1 \\ -B_3 & 0 & B_1 & E_2 \\ B_2 & -B_1 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix}$$

then the 4-force on the particle is:

$$qF_{\alpha\beta} v^\beta$$

Now the (empty-space) Maxwell equation become in this notation

$$dF = 0 \quad \text{and} \quad d({}^*F) = 0$$

where:

$${}^*F = B_i dx_i dx_4 + \frac{1}{2} E_i e_{ijk} dx_j dx_k.$$

The equation $dF = 0$ is of course equivalent by Poincaré's lemma to $F = dA$ for a 1-form $A = \sum_\alpha A_\alpha dx_\alpha$ (the 4-vector potential), while the equation $d({}^*F) = 0$ says that A is "harmonic", i.e. a solution of Lagrangian variational problem

$$\delta \int \|dA\|^2 dv = 0.$$

Now is there some natural way to look at the paths of particles moving under the Lorentz force as geodesics in some Riemannian geometry? In the late 1920's Kaluza and Klein gave a beautiful extention to Einstein's theory that provided a positive answer to this question. On the 5-dimensional space $p = \mathbb{R}^4 \times S^1$ (on which S^1 acts by $e^{i\theta}(p, e^{i\phi}) = (p, e^{i(\theta+\phi)})$, consider metrics γ which are invariant under this S^1 action. What the Kaluza-Klein theory showed was:

1) Such metrics γ correspond 1-1 with pairs (g, A) where g is a metric and A a 1-form on $\mathbb{R}^4$.

2) If the metric γ on P is an Einstein metric, i.e. satisfies the Einstein variational principle

$$\delta \int R(\gamma) dv^5 = 0$$

then

a) the corresponding A is harmonic (so $F = dA$ satisfies the Maxwell equations)

b) the geodesics of γ project exactly onto the paths of charged particles in $\mathbb{R}^4$ under the Faraday tensor $F = dA$.

c) the metric g on $\mathbb{R}^4$ satisfies Einstein's field equations, <u>not</u> for "empty-space", but better yet for the correct "energy momentum tensor" of the electromagnetic field F!

What has caused so much excitement in the last ten years is the realization that the two short range "nuclear forces" can also be understood in the same mathematical framework. One must replace the abelian compact Lie group S^1 by a more general compact simple group G and also generalize the product bundle $\mathbb{R}^4 \times G$ by a more general principal bundle. The reason that the force is now short range (or equivalently why the analogues of photons have mass) depend on a very interesting mathematical phenomenon called "spontaneous symmetry breaking" or "the Higg's mechanism" which we will discuss in the course.

Actually we have left out an extremely important aspect of physics-quantization. Our whole discussion so far has been at the classical level. In the course I will only deal with this "pre-quantum" part of physics.

LECTURE 2. 6/5/81 10AM–12AM RM 301 TSING HUA

REVIEW OF SMOOTH VECTOR BUNDLES

M a smooth (C^∞), paracompact, n-dimensional manifold. $C^\infty(M, W)$ smooth maps of M to W. Here we sketch concepts and notations for theory of smooth vector bundles over M. Details in written notes, extra lectures.

Definition of smooth k-dimension vector bundle over M.

E a smooth manifold, $\pi : E \to M$ smooth

$E = \bigcup\limits_p E_p \quad E_p = \pi^{-1}(p) \quad$ a k-dimensional real v-s.

$\theta \subseteq M \quad s : \theta \to E \quad$ smooth is a section if $s(p) \in E_p$

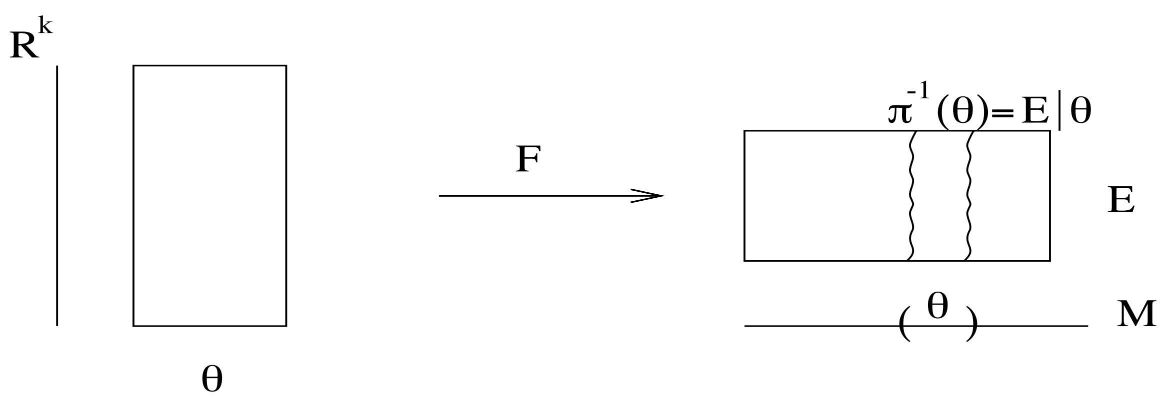
all $p \in \theta$

$\Gamma(E|\theta)$ = all sections of E over θ

$s = s_1, \ldots, s_k \in \Gamma(E|\theta)$ is called a local basis of sections for E over θ if the map

$$\begin{array}{rccl} F^S : & \theta \times \mathbb{R}^k & \to & E|\theta \simeq \pi^{-1}(\theta) \\ & (p, \alpha) & \to & \alpha_1 s_1(p) + \cdots + \alpha_k s_k(p) \end{array}$$

is a diffeo $F^S : I \times \mathbb{R}^k \simeq E|\theta$



Note F^S_p is linear $\{p\} \times R^k \simeq E_p$. Conversely given $F : \theta \times R^k \simeq E|\theta$ diffeo such that for each $p \in \theta \quad F_p = F/\{p\} \times \mathbb{R}^k$ maps $\mathbb{R}^k$ linearly onto E_p, F arises as above [with $s_i(p) = F_p(e_i)$]. These maps $F : \theta \times R^k \simeq E/\theta$ play a central role in what follows. They are called local gauges for E over θ. Basic defining axiom for smooth vector bundle is that each $p \in \theta$ has a neighborhood θ for which there is a local gauge $F : \theta \times \mathbb{R}^k \simeq E|\theta$.

Whenever we are interested in a "local" question about E we can always choose a local gauge and pretend $E|\theta$ is $\theta \times R^k$ — in particular a section of E over θ becomes a map $s : \theta \to \mathbb{R}^k$. Gauge transition functions: Suppose $F^k : \theta_i \times R^k \to E|\theta_i \quad i = 1, 2$ are two local gauges. Then for each $p \in \theta_1 \cap \theta_2$ we have two isomorphisms $F_p^i : \mathbb{R}^k \simeq E_p$, hence there is a unique $g(p) = (F_p^1)^{-1} \circ F_p^2 \in GL(k)$ is easily seen to be smooth and is called the gauge transition map from the local gauge F_1 to the local gauge F_2. It is characterized by:

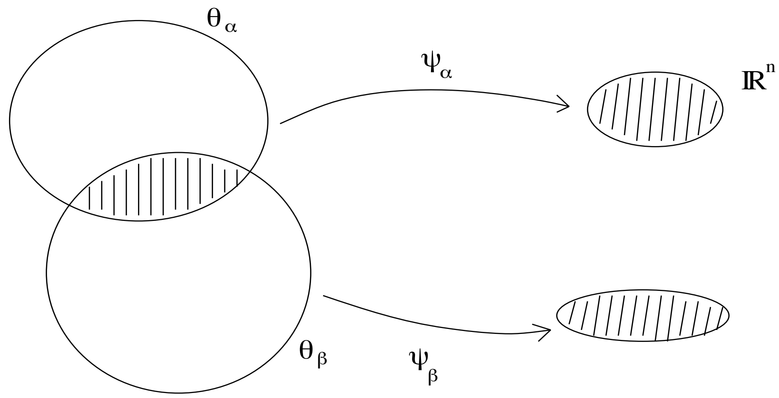
$$F_2 = F_1 g \text{ in } (\theta_1 \cap \theta_2) \times R^k \quad (\text{where } F_1 g(p, \alpha) = F_1(p, g(p)\alpha)).$$

Cocycle Condition If $F_i : \theta_i \times \mathbb{R}^k \simeq E|\theta_i$ are three local gauges and $g_{ij} : \theta_i \cap \theta_j \to GL(k)$ is the gauge transition function from F_j to F_i then in $\theta_1 \cap \theta_2 \cap \theta_3$ the following "cocycle condition" is satisfied: $g_{13} = g_{12} \circ g_{23}$.

Definition: A G-bundle structure for E, where G is a closed subgroup of $GL(k)$, is a collection of local gauges $F_i : \theta_i \times \mathbb{R}^k \to E|\theta_i$ for E such that the $\{\theta_i\}$ cover M and for all i, j the gauge transition function g_{ij} for F_j to F_i has its image in G.

Examples and Remarks: If S is some kind of "structure" for the vector space $\mathbb{R}^k$ which is invariant under the group G, then given a G-structure for E we can put the same kind of structure on each E_p smoothly by carrying S over by any of the isomorphism $(F_i)_p : \mathbb{R}^k \simeq E_p$ with $p \in \theta_i$ (since S is G invariant there is no contradiction). Conversely, if G is actually the group of all symmetries of S then a structure of type S put smoothly on the E_p gives a G-structure for E.

SO: An $O(k)$-structure is the same as a "Riemannian structure" for E, a $GL(m, \mathbb{C})$-structure ($k = 2m$) is the same as complex vector bundle structure, a $U(m)$ structure is the same as a complex-structure together with a hermitian inner product, etc.



Example: $\{\phi_\alpha : \theta_\alpha \to \mathbb{R}^n\}$ the charts defining the differentiable structure of M $\psi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1}$ $g_{\alpha\beta} = D\psi_{\alpha\beta} \circ \phi_\beta$

Maximal G-structures. Every G-structure for E is included in a unique maximal G-structure. [Will always assume maximality] G-bundle Atlas: An (indexed) open cover $\{\theta_\alpha\}_{\alpha\in A}$ of M together with smooth maps $g_{\alpha\beta} : \theta_\alpha \cap \theta_\beta \to G$ satisfying the cocycle condition (again, can be embedded in a unique maximal such atlas).

A G-vector bundle gives a G-bundle atlas. Conversely:

Theorem: If $\{\theta_\alpha,\, g_{\alpha\beta}\}$ is any G-bundle atlas then there is a G-vector bundles having the $g_{\alpha\beta}$ as transition functions.

How unique is this?

If E is a G-cover bundle over M and $p \in M$ then a G-frame for E at p is a linear isomorphism $f : \mathbb{R}^k \simeq E_p$ for some gauge $F : \theta \times \mathbb{R}^k \simeq E|\theta$ of the G-bundle structure for E with $p \in \theta$. Given one such G-frame f_0 then $f = f_0 \circ g$ is also a G-frame for every $g \in G$ and in fact the map $g \to f_0 \circ g$ is a bijection of G with the set of all G-frames for E at p].

Given vector bundles E_1 and E_2 over M a vector bundle morphism between them is a smooth map $f : E_1 \to E_2$ such that for all $p \in M$ $f|(E_1)_p$ is a linear map $f_p : (E_1)_p \to (E_2)_p$. If in addition each f_p is bijective (in which case $f^{-1} : E_2 \to E_1$ is also a vector bundle morphism) then f is called an equivalence of E_1 with E_2. If E_1 and E_2 are both G-vector bundle and f_p maps G-frames of E_1 at p to G-frames of E_2 at p then f is called an equivalence of G-vector bundles.

Theorem. Two G-vector bundles over M are equivalent (as G vector bundles) if and only if they have the same (maximal) G-bundle atlas; hence there is a bijective correspondence between maximal G-bundle atlases and equivalence classes of G-bundles.

If E is a smooth vector bundle over M then $Aut(E)$ will denote the group of automorphisms (i.e. self-equivalences of E as a vector bundle) and if E is a G-vector bundle than $Aut_G(E)$ denotes the sub-group of G-vector bundle equivalences of E with itself. $Aut_G(E)$ is also called the group of gauge transformations of E.

CONSTRUCTION METHOD FOR VECTOR BUNDLES

1) "Gluing". Given (G) vector bundles E_1 over θ_1, E_2 over θ_2 with $\theta_1 \quad \theta_2 = M$ and a G equivalence $E_1|(\theta_1 \quad \theta_2) \overset{\psi}{\simeq} E_2(\theta_1 \quad \theta_2)$ get a bundle E over M with equivalence $\psi_1 : E|\theta_1 \simeq E_1$ and $\psi_2 : E|\theta_2 \simeq E_2$ such that in $\theta_1 \quad \theta_2$ $\psi_2 \cdot \psi_1^{-1} = \psi$.

2) "Pull-back". Given a smooth vector bundles $E \xrightarrow{\pi} M$ and a smooth map $f : N \to M$ get a smooth vector bundles f^*E over N; $f^*E = \{(n, e) \in N \times E | f(n) = \pi e\}$ with projection $\tilde{\pi}(n, e) = n$, so $(f^*E)_n = E_{f(n)}$. A G-structure also pulls back.

3) "Smooth functors". Consider a "functor" (like direct sum or tensor product) which to each r-tuple of vector spaces $v_1, \ldots, v_r$ associate a vector space $F(v_1, \ldots, v_r)$ and to isomorphism $T_1 : v_1 \to w_1, \ldots, T_r : v_r \to w_r$ associate on isomorphism $F(T_1, \ldots, T_r)$ of $F(v_1, \ldots, v_r)$ with $F(w_1, \ldots, w_r)$ and assume $GL(v_1) \times \cdots \times GL(v_r) \xrightarrow{F} GL(F(v_1, \ldots, v_r))$ is smooth. Then given smooth vector bundles $E_1, \ldots, E_r$ over M we can form a smooth vector bundle

$F(E_1, \ldots, E_r)$ over M whose fiber at p is $F((E_1)_p, \ldots, (E_r)_p)$. In particular in this way we get $E_1 \oplus \cdots \oplus E_r$, $E_1 \otimes \cdots \otimes E_r$, $L(E_1, E_2)$, $\Lambda^p(E)$, $\Lambda^p(E, F)$.

4) Sub-bundles and Quotient bundles
   E_1 is said to be a sub-bundle of E_2 if $E_1 \subseteq E_2$ and the inclusion map is a vector bundle morphism. Can always choose local basis for E_2 such that initial element are a local base for E_1. It follows that there is a well defined smooth bundle structure for the quotient E_2/E_1 so that $0 \to E_1 \to E_2 \to E_2/E_1 \to 0$ is a sequence of bundle morphism. Moreover, using a Riemannian structure for E_2 we can find a second sub-bundle $E_1' = E_1^{\perp}$ in E_2 such that $E_2 = E_1 \oplus E_1'$ and then the projection of E_2 on E_2/E_1 maps E_1' isomorphically onto E_2/E_1.

LINEAR DIFFERENTIAL OPERATORS

$\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{Z}^+)^n$ multi-index

$|\alpha| = \alpha_1 + \cdots + \alpha_n$

$D^\alpha = \partial_\alpha = \dfrac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$

$D^\alpha : C^\infty(\mathbb{R}^n, \mathbb{R}^k) \to C^\infty(\mathbb{R}^n, \mathbb{R}^k)$

Definition: A linear map $L : C^\infty(\mathbb{R}^n, \mathbb{R}^k) \to C^\infty(\mathbb{R}^n, \mathbb{R}^\ell)$ is called an r-th order linear differential operator (with smooth coefficients) if it is of the form

$$(Lf)(x) = \sum_{|\alpha| \le r} a_\alpha(x) D^\alpha f(x)$$

where the a_α are smooth maps of $\mathbb{R}^n$ into the space $L(\mathbb{R}^k, \mathbb{R}^\ell)$ of linear maps of $\mathbb{R}^k$ into $\mathbb{R}^\ell$ (i.e. $k \times \ell$ matrices).

Easy exercises: $L : C^\infty(\mathbb{R}^n, \mathbb{R}^k) \to C^\infty(\mathbb{R}^n, \mathbb{R}^\ell)$ is an r^{th}-order linear differential operator if and only if whenever a smooth $g : \mathbb{R}^n \to \mathbb{R}$ vanishes at $p \in \mathbb{R}^n$ then for any $f \in C^\infty(\mathbb{R}^n, \mathbb{R}^k)$ $L(g^{r+1}f) = 0$ (use induction and the product rule for differentiation).

Definition. Let E_1 and E_2 be smooth vector bundles over M and let $L : \Gamma(E_1) \to \Gamma(E_2)$ be a linear map. We call L an r^{th}-order linear differential operator if whenever $g \in C^\infty(M, \mathbb{R})$ vanishes at $p \in M$ then for any section $s \in \Gamma(E_1)$, $L(g^{r+1}s)(p) = 0$. The set of all such L is clearly a vector space which we denote by $\mathrm{Diff}^r(E_1, E_2)$.

Remark. If we choose local gauges $F_i : \theta \times \mathbb{R}^{k_i} \simeq E_i|\theta$ then sections of E_i, restricted to θ get represented by elements of $C^\infty(\theta, \mathbb{R}^{k_i})$. If θ is small enough to be inside the domain of a local coordinate system $x_1, \ldots, x_n$ for M then any $L \in \mathrm{Diff}^r(E_1, E_2)$ has a local representation $\sum\limits_{|\alpha| \le r} a_\alpha D^\alpha$ as above.

DIFFERENTIAL FORMS WITH VALUES IN A VECTOR BUNDLE

If V is a vector space $\Lambda^p(V)$ denote all skew-symmetric p-linear maps of V into $\mathbb{R}$. If W is a second vector space then $\Lambda^p(V) \otimes W$ is canonically identified with all skew-symmetric p-linear maps of V into W. [If $\lambda \in \Lambda^p(V)$ and $\omega \in W$ then $\lambda \otimes \omega$ is the alternating p-linear map $(v_1, \ldots, v_p) \mapsto \lambda(v_1, \ldots, v_p)w$].

If E is a smooth bundle over M then the bundle $\Lambda^p(T(M)) \otimes E$ play a very important role, and so the notation will be shortened to $\Lambda^p(M) \otimes E$. This is called the "bundle of p-forms on M with values in E", and a smooth section $\omega \in \Gamma(\Lambda^p(M) \otimes E)$ is called a smooth p-form on M with values in E. Note that for each $q \in M$ $\omega_q \in \Lambda^p(\mathrm{TM}_q) \otimes E_q$ is an alternating p-linear map of TM_q into E_q, so that if $x_1, \ldots, x_p$ are p vector fields on M then

$$q \to \omega_q((x_1)_q, \ldots, (x_p)_q)$$

is a smooth section $\omega(x_1, \ldots, x_p)$ of E which is skew in the x_i.

Given $\omega_i \in \Lambda(\Gamma^{p_i}(M) \otimes E_i)$ $i = 1, 2$ we define their wedge product $\omega_1 \tilde{\Lambda} \omega_2$ in

$\Gamma(\Lambda^{p_1+p_2}(M) \otimes (E_1 \otimes E_2))$ by:

$$\begin{aligned}
&\omega_1 \tilde{\wedge} \omega_2(v_1, \ldots, v_{p_1+p_2}) \\
= &\frac{p_1! \, p_2!}{(p_1+p_2)!} \sum_{\sigma \in s_{p_1+p_2}} \varepsilon(\sigma) \omega_1(v_{\sigma(1)}, \ldots, v_{\sigma(p_1)}) \otimes \omega_2(v_{\sigma(p_1+1)}, \ldots, v_{\sigma(p_1+p_2)})
\end{aligned}$$

so in particular if ω_1 and ω_2 are one forms with value in E_1 and E_2 then

$$\omega_1 \tilde{\wedge} \omega_2(v_1, v_2) = \frac{1}{2}(\omega_1(v_1) \otimes \omega_2(v_2) - \omega_1(v_2) \otimes \omega_2(v_1))$$

In case $E_1 = E_2 = E$ is a bundle of algebras — i.e. we have a vector bundle morphism $E \otimes E \to E$ then we can define a wedge product $\omega_1 \wedge \omega_2$ which is a $p_1 + p_2$ form on M with values in E again. In particular for the case of two one-forms again we have

$$\omega_1 \wedge \omega_2 = \frac{1}{2}(\omega_1 \otimes \omega_2 - \omega_2 \otimes \omega_1)$$

where $\omega_1 \otimes \omega_2(v_1, v_2) = \omega_1(v_1)\omega_2(v_2)$.

In case E is a bundle of anti-commutative algebras (e.g. Lie algebras)

$$\omega_2 \otimes \omega_1 = -\omega_1 \otimes \omega_2$$

so we have

$$\omega_1 \wedge \omega_2 = \omega_1 \otimes \omega_2$$

for the wedge product of two 1-forms with values in a bundle E of anti-commutative algebras. In particular letting

$$\begin{aligned}
\omega &= \omega_1 = \omega_2 \\
\omega \wedge \omega &= \omega \otimes \omega
\end{aligned}$$

for a 1-form with values in a Lie algebra bundle E. If the product in E is designated, as usual, by the bracket $[\,,\,]$ then it is customary to write $[\omega, \omega]$ instead of $\omega \otimes \omega$. Note that

$$[\omega, \omega](v_1, v_2) = [\omega(v_1), \omega(v_2)]$$

In particular if E is a bundle of matrix Lie algebras, whose [,] means commutation then

$$\omega \wedge \omega(v_1, v_2) = \omega(v_1)\omega(v_2) - \omega(v_2)\omega(v_1).$$

Another situation in which we have a canonical pairing of bundles: $E_1 \otimes E_2 \to E_3$ is when $E_1 = E$, $E_2 = E^*$, the dual bundle and $E_3 = \mathbb{R}_M = M \times R$. Similarly if E is a Riemannian vector bundle and $E_1 = E_2 = E$ we have such a pairing into $\mathbb{R}_M$. Thus in either case if ω_i is a p_i-form with values in E_i then we have a wedge product $\omega_1 \wedge \omega_2$ which is an ordinary <u>real</u> valued $p_1 + p_2$-form.

THE HODGE *-OPERATOR

Let M have a Riemannian structure. The inner product on TM_q induces one on each $\Lambda^p(M) = \Lambda^p(\mathrm{TM}_q)$: (characterized by)

$$< v_1 \wedge \cdots v_p, \omega_1 \wedge \cdots \omega_p >= \sum_{\pi \in s_p} \varepsilon(\pi) < v_{\pi(1)}, \omega_1 > \cdots < v_{\pi(p)}, \omega_p >$$

If $e_1, \ldots, e_p$ is any orthogonal basis for TM_q then the $\binom{n}{p}$ elements $e_{j_1} \wedge \cdots \wedge e_{j_p}$ (where $1 \leq j_1 < \cdots < j_p \leq n$) is an orthonormal basis for $\Lambda^p(M)_q$. In particular $\Lambda^n(M)_p$ is 1-dim. and has two elements of norm 1. If we can choose $\mu \in \Gamma(\Lambda^n(M))$ with $\|\mu_q\| = 1$ M is called <u>orientable</u>. The only possible chooses are $\pm\mu$ and a choice of one of them (call it μ) is called an orientation for M and μ is called the Riemannian volume element.

Now fix p and consider the bilinear map $\lambda, \mu \mapsto \lambda \wedge v$ of $\Lambda^p(M)_p \times \Lambda^{n-p}(M)_p \to \Lambda^n(M_p)$. Since μ is a basis for $\Lambda^n(M)_p$ there is a bilinear form $B_p : \Lambda^p \times \Lambda^{n-p} \to \mathbb{R}$.

$$\lambda \wedge v = B_p(\lambda, v)\mu$$

<u>We shall now prove the easy but very important fact that B_p is non-degenerate and therefore that it uniquely determines an isomorphism $^* : \Lambda^p(M) \simeq \Lambda^{n-p}(M)$.</u> <u>Such that</u>:

$$\lambda \Lambda^* v =< \lambda, v > \mu$$

Given $I = (i_1, \ldots, i_p)$ with $1 \le i_1 < \cdots < i_p \le n$ let $e_I = e_{i_1} \wedge \cdots \wedge e_{i_1}$ for any orthonormal basis $e_1, \ldots, e_n$ of TM_q and let $I^C = (j_1, \ldots, j_{n-p})$ be the complementary set in $(1, 2, \ldots, n)$ in increasing order. Let $\tau(I)$ be the parity of the permutation

$$\begin{pmatrix} 1, 2, \ldots, p, p+1, \ldots, n \\ i_1, i_2, \ldots, i_p, j_1, \ldots, j_{n-p} \end{pmatrix}$$

Then clearly $e_I \wedge e_{I^C} = \tau(I)\mu$ while for any other $(n-p)$-element subset J of $(1, 2, \ldots, n)$ in increasing order $e_I \wedge e_J = 0$. Thus clearly as I ranges over all p-element subsets of $(1, 2, \ldots, n)$ in increasing order $\{e_I\}$ and $\{\tau(I)e_{I^C}\}$ are bases for $\Lambda^p(M)_q$ and $\Lambda^{n-p}(M)_q$ dual w.r.t. B_p.

This proves the non-degeneracy of B_p and the existence of $^* = {}^*_p$, and also that $^*e_I = \tau(I)e_{I^C}$. It follows that $^*{}_{n-p} \circ {}^*{}_p = (-1)^{p(n-p)}$. In particular if n is a multiple of 4 and $p = n/2$ then $^{*2}_p = 1$ so in this case $\Lambda^p(M)_q$ is the direct sum of the $+1$ and -1 eigenspaces of $^*{}_p$ (called self-dual and anti-self dual elements of $\Lambda^p(M)_q$.) Generalized Hodge *-operator.

Now suppose E is a Riemannian vector bundle over M. Then as remarked above the pairing $E \otimes E \to \mathbb{R}_n$ given by the inner product on E induces a pairing $\lambda, v \mapsto \lambda \wedge v$ from $(\Lambda^p(M)_q \otimes E_q) \otimes \Lambda^{n-p}(M)_q \otimes E_q \to \mathbb{R}$ so exactly as above we get a bilinear form B_p so that $\lambda \wedge v = B_p(\lambda, v)\mu_q$. And also just as above we see that B_p is non-degenerate by showing that if $e_1, \ldots, e_n$ is an orthonormal basis for TM_q and $u_1, \ldots, u_k$ an orthonormal basis for E_1 then $\{e_I \otimes u_j\}$ and $\{e_{I^C} \otimes u_j\}$ are dual bases w.r.t. B_p. It follows that:

<u>For $p = 1, \ldots, n$ there is an isomorphism $^*p : \Lambda^p(M) \otimes E \simeq \Lambda^{n-p}(n) \otimes E$ characterized by</u>:

$$\lambda \wedge {}^*v = <\lambda, v> \mu$$

moreover if $e_1, \ldots, e_n$ is an orthonormal basis for TM_q and $u_1, \ldots, u_k$ is an orthonormal base for E_q then $^*(e_I \otimes u_j) = \tau(I)e_{I^C} \otimes u_j$.

The subspace of $\Gamma(E)$ consisting of sections having compact support (disjoint form ∂M if $M \neq \emptyset$) will be denoted by $\Gamma_C(E)$. If $s_1, s_2 \in \Gamma_C(E)$ then

$q \rightarrow < s_1(q), s_2(q) >$ is a well-defined smooth function (when E has a Riemannian structure) and clearly this function has compact support. We denote it by $< s_1, s_2 >$ and define a pre-hilbert space structure with inner product $((s_1, s_2))$ on $\Gamma_C(E)$ by

$$((s_1, s_2)) = \int_M < s_1, s_2 > \mu$$

Then we note that, by the definition of the Hodge *-operator, given $\lambda, \nu \in \Gamma_C(\Lambda^p(M) \otimes E)$

$$((\lambda, \nu)) = \int_M \lambda \Lambda^* \nu$$

THE EXTERIOR DERIVATIVE

Let $E = M \times V$ be a product bundle. Then sections of $\Lambda^p(M) \otimes E$ are just p-forms on M with values in the fixed vector space V. In this case (and this case only) we have a natural first order differential operator

$$d : \Gamma(\Lambda^p(M) \otimes E) \rightarrow \Gamma(\Lambda^{p+1}(M) \otimes E)$$

called the exterior derivative. If $\omega \in \Lambda^p(M) \otimes E$ and $x_1, \ldots, x_{p+1}$ are $p+1$ smooth vector fields on M

$$\begin{aligned} dw(x_1, \ldots, x_{p+1}) \;\; = \;\; & \sum_{i=1}^{p+1} (-1)^{i+1} x_i w(x_1, \ldots, \hat{x}_i, \ldots, x_{p+1}) + \\ & \sum_{1 \leq i < j \leq p+1} (-1)^{i+j} w([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_{p+1}) \end{aligned}$$

[It needs a little calculation to show that the value of $dw(x_1, \ldots, x_{p+1})$ at a point q depends only on the values of the vector fields x_i at q, and not also, as it might at first seem also on their derivative.

We recall a few of the important properties of d.

1) d is linear

2) $d(w_1 \tilde{\wedge} w_2) = (dw_1)\tilde{\wedge} w_2 + (-1)^{p_1} w_1 \tilde{\wedge} d(w_2)$ for $w_i \in \Lambda^{p_i}(M) \otimes E$

3) $d^2 = 0$

4) If $w \in \Gamma_C(\Lambda^{n-1}(M) \otimes E)$ then $\int_M dw = 0$ (This is a special case of Stoke's Theorem)

ADJOINT DIFFERENTIAL OPERATORS

In the following E and F are Riemannian vector bundles over a Riemannian manifold M. Recall $\Gamma_C(E)$ and $\Gamma_C(F)$ are prehilbert space. If $L : \Gamma(E) \to \Gamma(F)$ is in $\mathrm{Diff}^r(E, F)$ then it maps $\Gamma_C(E)$ into $\Gamma_C(F)$. A linear map $L^* : \Gamma(F) \to \Gamma(E)$ in $\mathrm{Diff}^r(F, E)$ is called a (formal) adjoint for L if for all $s_1 \in \Gamma_C(E)$ and $s_2 \in \Gamma_C(F)$

$$((Ls_1, s_2)) = ((s_1, L^* s_2)).$$

It is clear that if such an L^* exists it is unique. It is easy to show formal adjoints exist locally (just integrate by parts) and by uniqueness these local formal adjoints fit together to give a global formal adjoint. That is we have the theorem that any $L \in \mathrm{Diff}^r(E, F)$ has a unique formal adjoint $L^* \in \mathrm{Diff}^r(F, E)$. We now compute explicitly the formal adjoint of $d = d_p : \Gamma(\Lambda^p(M) \otimes E) \to \Gamma(\Lambda^{p+1}(M) \otimes E)$ $(E = M \times V)$ which is denoted by

$$\delta = \delta_{p+1} : \Gamma(\Lambda^{p+1}(M) \otimes E) \to \Gamma(\Lambda^p(M) \otimes E).$$

Let $\lambda \in \Gamma_C(\Lambda^p(M) \otimes E)$ and let $\nu \in \Gamma_C(\lambda^{p+1}(M) \otimes E)$. Then $\lambda \Lambda^* \nu$ is (because of the pairing $E \otimes E \to \mathbb{R}$ given by the Riemannian structure) a real valued $p + (n - (p+1)) = n - 1$ form with compact support, and hence by Stokes theorem

$$0 = \int_M d(\lambda \Lambda^*{}_{p+1} \nu).$$

Now

$$d(\lambda \Lambda^* \nu) = d\lambda \Lambda^*{}_{p+1} \nu + (-1)^p \lambda \Lambda d({}^*{}_{p+1} \nu)$$

so recalling that

$$\begin{aligned} {}^*{}_p{}^*{}_{n-p} &= (-1)^{p(n-p)} \\ d(\lambda\Lambda^*\nu) &= d\lambda\Lambda^*\nu + (-1)^{p+p(n-p)}\lambda\Lambda^*{}_p{}^*{}_{n-p}d^*{}_{p+1} \end{aligned}$$

so integrating over M and recalling the formula for $((\,,\,))$ on forms

$$\begin{aligned} 0 &= \int_M d(\lambda\Lambda^*\nu) \\ &= \int_M d\lambda\Lambda^*\nu + (-1)^{p+p(n-p)}\int_M \lambda\Lambda^*_p({}^*_{n-p}d^*_{p+1})\nu \\ &= ((d\lambda,\,\nu)) - ((\lambda,\,\delta\nu)) \end{aligned}$$

where

$$\delta = \delta_{p+1} = -(-1)^{p(n-p+1)*}{}_{n-p}d^*{}_{p+1}$$

Since $d_{p+1}\circ d_p = 0$ it follows easily that $\delta_{p+1}\circ\delta_{p+2} = 0$

The Laplacian $\Delta = \Delta_p = d_{p-1}\delta_p + \delta_{p+1}d_p$ is a second order linear differential operator $\Delta_p \in \mathrm{Diff}^2(\Lambda^p(M)\otimes E,\,\Lambda^p(M)\otimes E)$.

The kernel of Δ_p is called the space of harmonic p-forms with values in $E = M\times V$ (or values in V).

Theorem. If $w\in\Gamma_C(\Lambda^p(M)\otimes E)$ then w is harmonic if and only if $dw = 0$ and $\delta w = 0$.

Proof.

$$\begin{aligned} 0 &= ((\Delta w,\,w)) \\ &= ((d\delta w + \delta d)w,\,w)) \\ &= ((d\delta w,\,w)) + ((\delta dw,\,w)) \\ &= ((\delta w,\,\delta w)) + ((dw,\,dw)). \end{aligned}$$

So both δw and dw must be zero.

Exercise: Let $\mathcal{H}_p$ denote the space of harmonic p-norms in $\Gamma_C(\Lambda^p(M)\otimes E)$ show that $\mathcal{H}_p$, $\mathrm{im}(d_{p-1})$, and $\mathrm{im}(\delta_{p+1})$ are mutually orthogonal in $\Gamma_C(\Lambda^p(M)\otimes E)$.

HODGE DECOMPOSITION THEOREM

Let M be a closed (i.e. compact, without boundary) smooth manifold and let $E = M \times V$ be a smooth Riemannian bundle over M. Then $\Gamma(\Lambda^p(M) \otimes E)$ is the orthogonal direct sum

$$\mathcal{H}_p \oplus \mathrm{im}(d_{p-1}) \oplus \mathrm{im}(\delta_{p+1})$$

Corollary. If $w \in \Gamma(\Lambda^p(M) \otimes E)$ is closed (i.e. $dw = 0$) then there is a unique harmonic form $h \in \Gamma(\Lambda^p(M) \otimes E)$ which differs from w by an exact form (= something in image of d_{p-1}). That is very de Rham cohomology class contains a unique harmonic representative.

CONNECTIONS ON VECTOR BUNDLES

Notation: In what follows E denotes a k-dimensional smooth vector bundle over a smooth n-dimensional manifold M. G will denote a Lie subgroup of the group $GL(k)$ of non-singular linear transformations of $\mathbb{R}^k$ (identified where convenient with $k \times k$ matrices). The Lie algebra of G is denoted by $\mathcal{G}$ and is identified with the linear transformations A of $\mathbb{R}^k$ such that $\exp(tA)$ is a one-parameter subgroups of G. The Lie bracket $[A, B]$ of two elements of $\mathcal{G}$ is given by $AB - BA$. We assume that E is a G-vector bundle, i.e. has a specified G-bundle structure. (This is no loss of generality, since we can of course always assume $G = GL(k)$). We let $\{F_i\}$ denote the collection of local gauges $F_i : \theta_i \times \mathbb{R}^k \simeq E/\theta_i$ for E defining the G-structure and we let $g_{ij} : \theta_i \cap \theta_j \to G$ the corresponding transition function. Also we shall use $F : \theta \times \mathbb{R}^k \to E/\theta$ to represent a typical local gauge for E and $g : \theta \to G$ to denote a typical gauge transformation, and $s^1, \ldots, s^k$ a typical local base of sections of E ($s^i = F(e^i)$).

Definition. A connection on a smooth vector bundle E over M is a linear map

$$\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$$

such that given $f \in C^\infty(M, \mathbb{R})$ and $s \in \Gamma(E)$

$$\nabla(fs) = f\nabla s + df \otimes s$$

Exercise: $\nabla \in \mathrm{Diff}^1(E, T^*M \otimes E)$

Exercise: If $E = M \times V$ so $\Gamma(E) = C^\infty(M, V)$ then $d : \Gamma(E) \to \Gamma(T^*M \otimes E)$ is a connection for E. (Flat connection)

Exercise: If $f : E_1 \simeq E_2$ is a smooth vector bundle isomorphism then f induces a bijection between connections on E_1 and E_2.

Exercise: Put the last two exercises together to show how a gauge $F : \theta \times \mathbb{R}^k \simeq E/\theta$ defines a connection ∇^F for E/θ. (called flat connection defined by F)

Exercise: If $\{\theta_\alpha\}$ is a locally finite open cover of M, $\{\phi_\alpha\}$ a smooth partition of unity with $\mathrm{supp}(\phi_\alpha) \subseteq \theta_\alpha$, ∇^α a connection for E/θ_α, then $\sum\limits_\alpha \phi_\alpha \nabla^\alpha = \nabla$ is a connection for E. (The preceding two exercises prove connections always exist).

Covariant Derivatives.

Given a connection ∇ for E and $s \in \Gamma(E)$, the value of ∇s at $p \in M$ is an element of $T^*M_p \otimes E_p$, i.e. a linear map of TM_p into E_p. Its value at $X \in \mathrm{TM}_p$ is denoted by $\nabla_x s$ and called the covariant derivative of s in the direction X. (For the flat connection $\nabla = d$ on $E = M \times V$, if $s = f \in C^\infty(M, V)$ then $\nabla_x s = df(X) =$ the directional derivative of f in the direction X). If $X \in \Gamma(\mathrm{TM})$ is a smooth vector field on M then $\nabla_x s$ is a smooth section of E. Thus for each $X \in \Gamma(\mathrm{TM})$ we have a map

$$\nabla_x : \Gamma(E) \to \Gamma(E)$$

(called convariant differentiation w.r.t. X). Clearly:

(1) ∇_x is linear and in fact in $\text{Diff}^1(E, E)$.

(2) The map $X \to \nabla_x$ of $\Gamma(\text{TM})$ into $\text{Diff}^1(E, E)$ is linear. Moreover if $f \in C^\infty(M, \mathbb{R})$ then $\nabla_{fx} = f\nabla_x$.

(3) If $s \in \Gamma(E)$, $f \in C^\infty(M, \mathbb{R})$, $X \in \Gamma(\text{TM})$

$$\nabla_x(fs) = (Xf)s + f\nabla_x s$$

Exercise: Check the above and show that conversely given a map $X \to \nabla_x$ form $\Gamma(\text{TM})$ into $\text{Diff}^1(E, E)$ satisfying the above it defines a connection.

CURVATURE OF A CONNECTION

Suppose E is trivial and let ∇ be the flat connection comming from some gauge $E \simeq M \times V$. If we don't know this gauge is there someway we can detect that ∇ is flat?

Let $f \in \Gamma(E) \simeq C^\infty(M, V)$ and let $X, Y \in \Gamma(E)$. Then $\nabla_x f = Xf$ so $\nabla_x(\nabla_y f) = X(Yf)$, hence if we write $[\nabla_x, \nabla_y]$ for the commutator $\nabla_x\nabla_y - \nabla_y\nabla_x$ of the operators ∇_x, ∇_y in $\text{Diff}^1(E, E)$ then we see $[\nabla_x, \nabla_y]f = X(Yf) - Y(Xf) = [X, Y]f = \nabla_{[X,Y]}f$. In other words:

$$\nabla_{[X,Y]} = [\nabla_x, \nabla_y]$$

or $X \to \nabla_x$ is a Lie algebra homeomorphism of $\Gamma(\text{TM})$ into $\text{Diff}^1(E, E)$. Now in general this will <u>not</u> be so if ∇ is not flat, so it is suggested that with a connection we study the map $\Omega : \Gamma(\text{TM}) \times \Gamma(\text{TM}) \to \text{Diff}^1(E, E)$

$$\Omega(X, Y) = [\nabla_x, \nabla_y] - \nabla_{[X,Y]}$$

which measure the amount by which $X \to \nabla_x$ fails to be a Lie algebra homeomorphism.

Theorem. $\Omega(X, Y)$ is in $\mathrm{Diff}^0(E, E)$; i.e. for each $p \in M$ there is a linear map $\Omega(X, Y)_p : E_p \to E_p$ such that if $s \in \Gamma(E)$ then $(\Omega(X, Y)s)(p) = \Omega(X, Y)_p s(p)$.

Proof. Since $\nabla_y(fs) = (Yf)s + f\nabla_y s$ we get $\nabla_x\nabla_y(fs) = x(yf)s + (yf)\nabla_x s + (xf)\nabla_y s + f\nabla_x\nabla_y s$ and interchanging x and y and subtracting, then subtracting $\nabla_{[x,y]}(fs) = ([x, y]f)s + f_{[x,y]}s$ we get finally:

$$([\nabla_x, \nabla_y] - \nabla_{[x,y]})(fs) = f([\nabla_x, \nabla_y] - \nabla_{[x,y]})s$$

from which at follows that if f vanishes at $p \in M$ then $([\nabla_x, \nabla_y] - \nabla_{[x,y]})(fs)$ vanishes at p, so $[\nabla_x, \nabla_y] - \nabla_{[x,y]} \in \mathrm{Diff}^0(E, E)$. $\square$

Theorem. There is a two from Ω on M with values in $L(E, E)$ (i.e. a section of $\Lambda^2(M) \otimes L(E, E)$) such that for any $x, y \in \Gamma(\mathrm{TM})$

$$\Omega_p(x_p, y_p) = \Omega(x, y)_p.$$

Proof. What we must show is that $\Omega(x, y)_p$ depends only in the value of x and y at p. Since Ω is clearly skew symmetric it will suffice to show that if x is fixed then $y \to \Omega(x, y)$ is an operator of order zero, or equivalently that $\Omega(x, fy) = f(\Omega(x, y))$ if $f \in C^\infty(M, \mathbb{R})$. Now recalling that $\nabla_{fy}s = f\nabla_y s$ we see that

$$\nabla_x\nabla_{fy}s = (xf)\nabla_y s + f\nabla_x\nabla_y s.$$

On the other hand $\nabla_{fy}\nabla_x s = f\nabla_y\nabla_x s$ so $[\nabla_x, \nabla_{fy}] = f[\nabla_x, \nabla_y] + (xf)\nabla_y$. On the other hand since

$$[x, fy] = f[x, y] + (xf)f$$

we have

$$\begin{aligned} \nabla_{[x,yf]} &= \nabla_{f[x,y]} + \nabla_{(xf)y} \\ &= f\nabla_{[x,y]} + (xf)\nabla_y. \end{aligned}$$

Thus

$$\Omega(x, fy) = [\nabla_x, \nabla_{fy}] - \nabla_{[x,fy]}$$

$$
\begin{aligned}
&= f[\nabla_x, \nabla_y] - f\nabla_{[x,y]} \\
&= f\Omega(x, y).
\end{aligned}
$$

$\square$

This two form Ω with values in $L(E, E)$ is called the curvature form of the connection ∇. (If there are several connection under consideration we shall write Ω^∇). At this point we know only the vanishing of Ω is a necessary condition for there to exist local gauges $\theta \times \mathbb{R}^k = E/\theta$ with respect to which ∇ is the flat connection, d. Later we shall see that this condition is also sufficient.

Torsion. Suppose we have a connection ∇ on $E = \mathrm{TM}$. In this case we can define another differential invariant of ∇, its torsion $\tau \in \Gamma(\Lambda^2(M) \otimes \mathrm{TM})$. Given $x, y \in \Gamma(\mathrm{TM})$ define $\tau(x, y) \in \Gamma(\mathrm{TM})$ by

$$\tau(x, y) = \nabla_x y - \nabla_y x - [x, y].$$

Exercise. Show that there is a two-form τ on M with values in TM such that if $x, y \in \Gamma(\mathrm{TM})$ then $\tau_p(x_p, y_p) = \tau(x, y)_p$. [Hint: it is enough to show that if $x, y \in \Gamma(\mathrm{TM})$ then $\tau(x, fy) = f\tau(x, y)$ for all $f \in C^\infty(M, \mathbb{R})$].

Exercise. Let $\phi : \theta \to \mathbb{R}^n$ be a chart for M. Then ϕ induces a gauge F^ϕ : $\theta \times \mathbb{R}^n \simeq \mathrm{TM}/\theta$ for the tangent bundle of M (namely $F(p, v) = D\phi_p^{-1}(v)$). Show that for the flat connection on TM/θ defined by such a gauge not only the curvature but also the torsion is zero.

Remark. Later we shall see that, conversely, if ∇ is a connection for TM and if both Ω^∇ and τ^∇ are zero then for each $p \in M$ there is a chart ϕ at p with respect to which ∇ is locally the flat connection coming from F^ϕ.

STRUCTURE OF THE SPACE $\mathcal{C}(E)$ OF ALL CONNECTION ON E

Let $\mathcal{C}(E)$ denote the set of all connection on E and denote by $\Delta(E)$ the space of all smooth one-form on M with values in $L(E,\, E)$: $\Delta(E) = \Gamma(\Lambda^1(M) \otimes L(E,\, E))$

Definition. If $w \in \Delta(E)$ and $s \in \Gamma(E)$ we define $w \otimes s$ in $\Gamma(\Lambda^1(M) \otimes E)$ by

$$(w \otimes s)(x) = w(x)s(p)$$

for $x \in \mathrm{TM}_p$.

Exercise. Show that $s \to w \otimes s$ is in $\mathrm{Diff}^0(E,\, T^*M \otimes E)$ and in fact $w \mapsto (s \to w \otimes s)$ is a linear isomorphism of $\Delta(E)$ with $\mathrm{Diff}^0(E,\, T^*M \otimes E)$.

The next theorem says that $\mathcal{C}(E)$ is an "affine" subspace of $\mathrm{Diff}^1(E, T^*M \otimes E)$ and in fact that if $\nabla^0 \in \mathcal{C}(E)$ then we have a canonical isomorphism

$$\mathcal{C}(E) \simeq \nabla^0 + \mathrm{Diff}^0(E,\, T^*M \otimes E)$$

of $\mathcal{C}(E)$ with the translate by ∇^0 of the subspace $\mathrm{Diff}^0(E,\, T^*M \otimes E) = \Delta(E)$ of $\mathrm{Diff}^1(E,\, T^*M \otimes E)$.

Theorem. If $\nabla^0 \in \mathcal{C}(E)$ and for each $w \in \Delta(E)$ we define $\nabla^w : \Gamma(E) \to \Gamma(T^*M \otimes E)$ by $\nabla^w s = \nabla^0 s + w \otimes s$, then $\nabla^w \in \mathcal{C}(E)$ and the map $w \mapsto \nabla^w$ is a bijective map $\Delta(E) \simeq \mathcal{C}(E)$.

Proof. It is trivial to verify that $\Delta^w \in \mathcal{C}(E)$. If $\nabla^1 \in (E)$ then since $\nabla^i(f_s) = f\nabla^i s + df \otimes s$ for $i = 0, 1$ it follows that $(\nabla^1 - \nabla^0)(fs) = f(\nabla^1 - \nabla^0)s$ so $\nabla^1 - \nabla^0 \in \mathrm{Diff}^0(E, T^*M \otimes E)$ and hence is of the form $s \mapsto w \otimes s$ for some $w \in \Delta(E)$. This shows $w \to \nabla^w$ is surjective and injectivity is trivial.

We shall call $\Delta(E) = \Gamma(\Lambda^1(M) \otimes L(E,\, E))$ the space of connection forms for E. Note that a connection form w does not by itself define a connection ∇^w, but only relative to another connection ∇^0. Thus $\nabla(E)$ is the space of "differences" of connection.

Connections in a Trivial Bundle

Let E be the trivial bundle in $M \times \mathbb{R}^k$. Then $\Gamma(E) = C^\infty(M, \mathbb{R}^k)$ and $\Gamma(\Lambda^1(M) \otimes E) = \Gamma(\Lambda^1(M) \otimes \mathbb{R}^k) =$ space of $\mathbb{R}^k$ valued one forms, so as remarked earlier we have a natural "origin" in this case for the space $\mathcal{C}(E)$ of connection on E, namely the "flat" connection $d : C^\infty(M, \mathbb{R}^k) \to \Gamma(\Lambda^1(M) \otimes \mathbb{R}^k)$ i.e. the usual differential of a vector valued function. Now $\Delta(E) = \Gamma(\Lambda^1(E) \otimes L(E, E)) = \Gamma(\Lambda^1(M) \otimes L(\mathbb{R}^k, \mathbb{R}^k)) =$ the space of $(k \times k)$-matrix valued 1-forms on M, so the theorem of the preceding section says that in this case there is a bijective correspondence $w \mapsto \nabla^w = d + w$ from $(k \times k)$-matrix valued one-form w on M to connections ∇^w for E, given by

$$\nabla^w f = df + wf.$$

To be more explicit, let $e^\alpha = e^1, \ldots, e^k$ be the standard base for $\mathbb{R}^k$ and $t^{\alpha\beta}$ the standard base for $L(\mathbb{R}^k, \mathbb{R}^k)(t^{\alpha\beta}(e^\gamma) = \delta_{\beta\gamma} e^\alpha)$. Then $w = \sum w_{\alpha\beta} \otimes t^{\alpha\beta}$ where $w_{\alpha\beta}$ are uniquely determined ordinary (i.e. real valued) one forms on M; $w_{\alpha\beta} \in \Gamma(\Lambda^1(M))$. Also $f = \sum_\alpha f_\alpha e^\alpha$ where $f_\alpha \in C^\infty(M, \mathbb{R})$ are real valued smooth function in M. Then, using summation convertion:

$$(\nabla^w f)_\alpha = df_\alpha + w_{\alpha\beta} f_\beta$$

which means that if $x \in \mathrm{TM}_p$ then

$$\begin{aligned}(\nabla^w_x f)_\alpha(p) &= df_\alpha(x) + w_{\alpha\beta}(x) f_\beta(p) \\ &= x f_\alpha + w_{\alpha\beta}(x) f_\beta(p).\end{aligned}$$

It is easy to see how to "calculate" the forms $w_{\alpha\beta}$ given $\nabla = \nabla^w$. If we take $f = e^\gamma$ then $f_\beta = \delta_{\beta\gamma}$ and in particular $df_\alpha = 0$ and $(\nabla^w f)_\alpha = w_{\alpha\beta}\delta_{\alpha\beta} = w_{\alpha\gamma}$.

Thus

$$w_{\alpha\beta} = (\nabla^w e^\beta)_\alpha$$

or

$$\nabla^w e^\beta = \sum_\beta w_{\alpha\beta} \otimes e^\alpha$$

or finally:

$$\nabla^w_x e^\beta = \sum_\beta w_{\alpha\beta}(x) e^\alpha$$

or in words:

Theorem. There is a bijective correspondence $w \to \nabla^w$ between $k \times k$ matrices $w = (w_{\alpha\beta})$ of one forms on M and connections ∇^w on the product bundle $E = M \times \mathbb{R}^k$. ∇^w is determined from w by: $(\nabla^w_x f)(p) = xf + w(x)f(p)$ for $x \in \mathrm{TM}_p$ and $f \in \Gamma(E) = C^\infty(M, \mathbb{R}^k)$. Conversely $\nabla \in \mathcal{C}(E)$ determines w by the following algorithm: let $e^1, \ldots, e^k$ be the standard "constant" sections of E, then for $x \in \mathrm{TM}_p$ $w_{\alpha\beta}(x)$ is the coefficient of e^α when $\nabla_x e$ is expanded in this basis:

$$\nabla_x e^\beta = \sum_\alpha w_{\alpha\beta}(x) e^\alpha(p).$$

Since $\Lambda^2(M) \otimes L(E, E) = \Lambda^2(M) \otimes L(\mathbb{R}^k, \mathbb{R}^k)$, the curvature two-form $\Omega \in \Gamma(\Lambda^2(M) \otimes L(E, E))$ of a connection $\nabla = \nabla^w$ in $\mathcal{C}(E)$ is a $(k \times k)$-matrix $\Omega_{\alpha\beta}$ of 1-form on M:

$$\Omega = \sum_{\alpha, \beta} \Omega_{\alpha\beta} e^{\alpha\beta}$$

or

$$\Omega(x, y) e^\beta = \sum_{\alpha, \beta} \Omega_{\alpha\beta}(x, y) e^\alpha$$

Recall $\Omega(x, y)e^\beta = (\nabla_x \nabla_y - \nabla_y \nabla_x - \nabla_{[x,y]})e^\beta$. Now

$$\begin{aligned} \nabla_y e^\beta &= w_{\alpha\beta}(y) e^\alpha \\ \nabla_x \nabla_y e^\beta &= (x w_{\alpha\beta}(y)) e^\alpha + w_{\gamma\beta}(y) \nabla_x e^\gamma \\ &= ((x w_{\alpha\beta}(y) + w_{\alpha\gamma}(x) w_{\gamma\beta}(y)) e^\alpha \end{aligned}$$

and we get easily

$$\begin{aligned}\Omega(x,\,y)e^\beta &= ((xw_{\alpha\beta}(y)) - (yw_{\alpha\beta}(x)) - w_{\alpha\beta}([x,\,y])e^\alpha \\ &\quad +(w_{\alpha\gamma}(x)w_{\gamma\beta}(y) - w_{\alpha\gamma}(y)w_{\gamma\beta}(x))e^\alpha \\ &= (dw_{\alpha\beta} + (w\wedge w)_{\alpha\beta})(x,\,y)e^\alpha\end{aligned}$$

Thus $\Omega_{\alpha\beta} = dw_{\alpha\beta} + (w\wedge w)_{\alpha\beta}$.

Theorem. If w is a $(k\times k)$-matrix of one-forms on M and $\nabla = \nabla^w = d+w$ is the corresponding connection on the product bundle $E = M\times\mathbb{R}^k$, then the curvature form Ω^w of ∇ is the $(k,\,k)$-matrix of two forms on M given by $\Omega^w = dw + w\wedge w$.

Bianchi Inequality: The curvature matrix of two-forms Ω^w satisfies:

$$d\Omega^w + w\wedge\Omega^w - \Omega\wedge w = 0$$

Proof.

$$\begin{aligned}d\Omega &= d(dw + w\wedge w) \\ &= d(dw) + dw\wedge w - w\wedge dw \\ &= (\Omega - w\wedge w)\wedge w - w\wedge(\Omega + w\wedge w) \\ &= \Omega\wedge w - w\wedge\Omega.\end{aligned}$$

$\square$

Christoffel Symbols: If $x_1,\,\ldots,\,x_n$ is a local coordinate system in $\theta\subseteq M$ then in θ $w_{\alpha\beta} = \sum_{i=1}^{n}\Gamma^\alpha_{i\beta}dx_i$. The $\Gamma^\alpha_{i\beta}$ are Christoffel symbols for the connection ∇^w w.r.t. these coordinates.

REPRESENTATION OF A CONNECTION W.R.T. A LOCAL BASE

It is essentially trivial to connect the above description of connection on a product bundle to a description (locally) of connections on an arbitrary bundle E with respect to a local base $(s^1, \ldots, s^k)$ for E/θ. Namely:

Theorem. There is a bijective correspondence $w \to \nabla^w$ between $(k \times k)$-matrices $w = w_{\alpha\beta}$ of 1-forms on θ and connections ∇^w for E/θ. If $s = \sum f_\alpha s^\alpha$ and $x \in \mathrm{TM}_p$, $p \in \theta$ then $\nabla^w_x s = (x f_\alpha + w_{\alpha\beta}(x) f_\beta(p)) s^\alpha(p)$. Conversely given ∇ we get w such that $\nabla = \nabla^w$ by expanding $\nabla_x s^\beta$ in terms of the s^α, i.e. $\nabla_x s^\beta = \sum_\alpha w_{\alpha\beta}(x) s^\alpha$.

If ${}^*s^\alpha$ is the dual section basis for E^* so ${}^*s^\alpha \otimes s^\beta$ is the local base for $L(E, E) = E^* \otimes E$ over θ then the curvature form Ω^w for ∇^w can be written in θ as

$$\Omega = \sum_{\alpha, \beta} \Omega^w_{\alpha\beta} {}^*s^\alpha \otimes s^\alpha$$

where the $(k \times k)$-matrix of two forms in θ is determined by

$$\Omega(x, y) s^\beta = \sum_{\alpha, \beta} \Omega^\omega_{\alpha\beta}(x, y) s^\alpha.$$

These forms $\Omega^w_{\alpha\beta}$ can be calculated directly form $w = w_{\alpha\beta}$ by

$$\Omega^w = dw + w \wedge w.$$

and satisfy

$$d\Omega^w + w \wedge \Omega^w - \Omega^w \wedge w = 0.$$

Change of gauge.

Suppose we have two local bases of sections for E/θ, say $s^1, \ldots, s^k$ and $\tilde{s}^1, \ldots, \tilde{s}^k$ and let $g : \theta \to GL(k)$ be the gauge transition map from one gauge to the other, i.e.:

$$s = \sum g_{\alpha\beta} s^\alpha.$$

Let ∇ be a connection for E. Then restricted to sections of E/θ, ∇ defines a connection on E/θ which is of the form ∇^w w.r.t. the basis $s^1, \ldots, s^k$, and ∇^w

w.r.t. the basis $\tilde{s}^1, \ldots, \tilde{s}^k$. What is the relation between the $(k \times k)$-matrices of 1-forms w and $\tilde{w}$?

Letting $g^{-1} = g^{-1}_{\alpha\beta}$ denote the matrix inverse to g, so that

$$\begin{aligned}
s^\lambda &= g^{-1}_{\alpha\beta}\ \tilde{s}^\alpha \\
\nabla_x \tilde{s}^\beta &= \nabla_x (g_{\lambda\beta} s^\lambda) \\
&= dg_{\lambda\beta}(x) s^\lambda + g_{\alpha\beta} \nabla_x s^\alpha \\
&= (dg_{\lambda\beta}(x) + g_{\alpha\beta} w_{\lambda\alpha}(x)) s^\lambda \\
&= (g^{-1}_{\alpha\lambda} dg_{\lambda\beta}(x) + g^{-1}_{\alpha\lambda} w_{\lambda\alpha}(x) g_{\alpha\beta}) s^{-\alpha}
\end{aligned}$$

from which we see that:

$$\tilde{w} = g^{-1} dg + g^{-1} w g$$

i.e.

$$\tilde{w}_{\alpha\beta}(x) = g^{-1}_{\alpha\lambda} dg_{\lambda\beta}(x) + g^{-1}_{\alpha\lambda} w_{\lambda\alpha}(x) g_{\alpha\beta}.$$

Exercise: Show that $\tilde{\Omega} = g^{-1}\Omega g$ i.e. that $\tilde{\Omega}_{\alpha\beta}(x, y) = g^{-1}_{\alpha\lambda}\Omega_{\lambda\gamma}(x, y) g_{\gamma\beta}$.

CONSTRUCTING NEW CONNECTION FROM OLD ONES

Proposition. Let $\{\theta_\alpha\}$ be an indexed covering of M by open sets. Suppose ∇^α is a connection on E/θ_α such that ∇^α and ∇^β agree in $E/(\theta_\alpha \cap \theta_\beta)$. Then there is a unique connection ∇ on E such that ∇ restrict to ∇^α in E/θ_α.

Proof. Trivial.

Theorem. If ∇ is any connection on E there is a unique connection ∇^* on the dual bundle E^* such that if $\sigma \in \Gamma(E^*)$ and $s \in \Gamma(E)$ then

$$x(\sigma(s)) = \nabla^*_x \sigma(s) + \sigma(\nabla_x s)$$

for any $x \in$ TM.

Proof. Consider the indexed collection of open sets $\{\theta_s\}$ of M such that the index s is a basis of local section $s^1, \ldots, s^k$ for E/θ_s. By the preceding proposition it will suffice to show that for each such θ_s there is a unique connection $\overset{s}{\nabla}{}^*$ for E^*/θ_s satisfying the given property, for then by uniqueness $\overset{s}{\nabla}{}^*$ and $\overset{s'}{\nabla}{}^*$ must agree on $E^*/(\theta_i \cap \theta_{s'})$. Let $\nabla_x s^\beta = w_{\alpha\beta}(x)s^\alpha$ and let ∇^* be any connection on E^*/θ_s. Let $\sigma^1, \ldots, \sigma^k$ be the bases for E^*/θ_s dual to $s^1, \ldots, s^k$ and let $\nabla^{*\beta}_\sigma = w^*_{\alpha\beta}(x)\sigma^\alpha$. Since $\sigma^\beta(s^\lambda) = \delta_{\beta\lambda}$ if ∇^* satisfies the given condition

$$0 = x(\sigma^\beta(s^\lambda)) = w^*_{\alpha\beta}(x)\sigma^\alpha(s^\lambda) + \sigma^\beta(w_{\alpha\lambda}(x)s^\alpha) = w^*_{\alpha\beta}(x) + w_{\beta\lambda}(x)$$

and thus w^* (and hence ∇^*) is uniquely determined by w to be $w^*_{\lambda\beta}(x) = -w_{\beta\lambda}(x)$. An easy computation left as an exercise shows that with this choice of w^*, ∇^* does indeed satisfies the required condition. $\square$

Theorem. If ∇^i is a connection in a bundle E^i over M, $i = 1, 2$ then there is a unique connection $\nabla = \nabla' \otimes 1 + 1 \otimes \nabla^2$ on $E^1 \otimes E^2$ such that if $s^i \in \Gamma(E^i)$ and $x \in \mathrm{TM}$ then

$$\nabla_x(s^1 \otimes s^2) = (\nabla_x s^1) \otimes s^2 + s^1 \otimes (\nabla_x s^2).$$

Proof. Exercise. [Hint: follow the pattern of the preceding theorem. Consider open sets θ of M for which there exist local bases $s^1, \ldots, s^{k_1}$ for E^1 over θ and $\sigma^1, \ldots, \sigma^{k_2}$ for E^2 over θ. Show that the matrix of 1-form w for ∇ relative to the basis $s^i \otimes s^j$ for $E^1 \otimes E^2$ over θ can be chosen in one and only one way to give ∇ the required property.]

Remark: It follows that given any connection on E there are connections on each of the bundles $\overset{r}{\otimes}\, E^* \otimes \overset{s}{\otimes}\, E$ which satisfy the usual "product formula" and "commute with contractions". Moreover these connections are uniquely determined by this property.

Theorem: If $E_1, \ldots, E_r$ are smooth vector bundles over M then the natural map $(\nabla_1, \ldots, \nabla_r) \mapsto \nabla_1 \oplus \cdots \oplus \nabla_r$ is a bijection $\mathcal{C}(E_1) \times \cdots \times \mathcal{C}(E_r) \simeq \mathcal{C}(E_1 \oplus$

$\cdots \oplus E_r)$.

Proof. Trivial.

From the smooth bundle $E \xrightarrow{\pi} M$ and a smooth map $\phi : N \to M$ we can form the pull back bundle $\phi^*(E)$ over N

$$\phi^*(E) = \{(x,\, v) \in N \times E \,|\, \phi(v) = \pi(v)\} \quad \phi^*(E)_x = E_{\phi(x)}.$$

There is a canonical linear map

$$\begin{aligned} \phi^* : \qquad & \Gamma(E) \to \Gamma(\phi^*(E)) \\ & s \mapsto s \circ \phi \end{aligned}$$

and if $s^1, \ldots, s^k$ is a local base for E over θ then $\phi^*(s^1), \ldots, \phi^*(s^k)$ is a local base for $\phi^*(E)$ over $\phi^{-1}(\theta)$.

Theorem. Given $\nabla \in \mathcal{C}(E)$ there is a uniquely determined connection ∇^ϕ for $\phi^*(E)$ such that if $s \in \Gamma(E)$, $y \in TN$ and $s = D\phi(y)$ then $\nabla^\phi_y(\phi^* s) = \phi^*(\nabla_x s)$.

Proof. Exercise. [Hint: given a local base $s^1, \ldots, s^k$ for E over θ for which the matrix of 1-forms of ∇ is $w_{\alpha\beta}$, show that the matrix of one-forms of ∇^ϕ must be $\phi^*(w)_{\alpha\beta})$ w.r.t. $\phi^*(s^1), \ldots, \phi^*(s^k)$ for the defining condition to hold and that in fact it does hold with this choice.

Special case 1. Let N be a smooth submanifold of M and $i : N \to M$ the inclusion map. Then $i^*(E) : E/N$ and $i^* : \Gamma(E) \to \Gamma(E/N)$ is $s \to s/N$. If $\nabla \in \Gamma(E)$ write $\nabla^{/N} = \nabla^i$. Suppose $x \in TN_p \subseteq \mathrm{TM}_p$. Then for $s \in \Gamma(E)$

$$\nabla_x s = \nabla_x^{/N}(s/N)$$

In particular if s and $\tilde{s}$ are two section of E with the same restriction to N then $\nabla_x s = \nabla_x \tilde{s}$ for all $x \in TN$ [We can see this directly w.r.t. a local trivialization (gauge) $\nabla_x s = x(s) + w(x)s$].

Special case 2. $N = I = [a, b]$. $\sigma = \phi : I \to M$. $\nabla^\sigma \in \mathcal{C}(\sigma^*(E))$. A section $\tilde{s}$ of $\sigma^*(E)$ is a map $t \to \tilde{s}(t) \in E_{\sigma(t)}$. We write $\frac{D\tilde{s}}{dt} = \frac{D^\nabla \tilde{s}}{dt} = \nabla^\sigma_{\frac{\partial}{\partial t}}(\tilde{s})$. Called the covariant derivative of s along σ. With respect to a local base $s^1, \ldots, s^k$ for E with connection forms $w_{\alpha\beta}$ suppose $\tilde{s}(t) = v_\alpha(t) s^\alpha(\sigma(t))$. Then the components $\frac{Dv_\alpha}{dt}$ of $\frac{D\tilde{s}}{dt}$ are given by

$$\frac{Dv_\alpha}{dt} = \frac{dv_\alpha}{dt} + w_{\alpha\beta}(\sigma'(t))v_\beta(t).$$

If $x_1, \ldots, x_n$ are local coordinates in M and we put $w_{\alpha\beta} = \Gamma^i_{\alpha\beta} dx_i$ and $x_i(\sigma(t)) = \sigma_i(t)$ then

$$\frac{Dv_\alpha}{dt} = \frac{dv_\alpha}{dt} + \Gamma^\alpha_{i\beta}(\sigma(t))\frac{d\sigma_i}{dt} v_\beta(t).$$

Note that this is a linear first order ordinary differential operator with smooth coefficients in the vector function $(v_1(t), \ldots, v_k(t)) \in \mathbb{R}^k$.

Parallel Translation

We have already noted that a connection ∇ on E is "flat" (i.e. locally equivalent to d in some gauge) iff its curvature Ω is zero. Since Ω is a two form it automatically is zero if the base space M is one-dimensional, so if $\sigma : [a, b] \to M$ is as above then for connection ∇ on any E over M the pull back connection ∇^σ on $\sigma^*(E)$ is flat. parallel translation is a very powerful and useful tool that is an explitic way of describing this flatness of ∇^σ.

Definition. The kernel of the linear map $\frac{D}{dt} = \nabla^\sigma_{\sigma'} : \Gamma(\sigma^* E) \to \Gamma(\sigma^* E)$ is a linear subspace $P(\sigma)$ of $\Gamma(\sigma^* E)$ called the space of parallel (or covariant constant) vector fields along σ.

Theorem. For each $t \in I$ the map $s \mapsto s(t)$ is a linear isomorphism of $P(\sigma)$ with $E_{\sigma(t)}$.

Proof. An immediate consequence of the form of $\frac{D}{dt}$ is local coordinates (see above) and the standard elementary theory of linear ODE.

Definition. For $t_1,\, t_2 \in I$ we define a linear operator $P_\sigma(t_2,\, t_1) :\, E_{\sigma(t_1)} \to E_{\sigma(t_2)}$ (called parallel translation along σ from t_1 to t_2) by $P_\sigma(t_2,\, t_1)v = s(t_2)$, where s is the unique element of $P(\sigma)$ with $s(t_1) = v$.

Properties:

1) $P_\sigma(t,\, t) =$ identity map of $E_{\sigma(t)}$

2) $P_\sigma(t_3,\, t_2)P_\sigma(t_2,\, t_1) = P_\sigma(t_3,\, t_1)$

3) $P_\sigma(t_1,\, t_2) = P_\sigma(t_2,\, t_1)^{-1}$

4) If $v \in E_{\sigma(t_0)}$ then $t \mapsto P_\sigma(t,\, t_0)v$ is the unique $s \in P(\sigma)$ with $s(t_0) = v$.

Exercise: Show that ∇ can be recovered from parallel translation as follows: given $s \in \Gamma(E)$ and $x \in$ TM let $\sigma :\, [0,\, 1] \to M$ be any smooth curve with $\sigma'(0) = x$ and define a smooth curve $\tilde{s}$ in $E_{\sigma(0)}$ by $\tilde{s}(t) = P_\sigma(0,\, t)s(\sigma(t))$. Then

$$\nabla_x s = \left.\frac{d}{dt}\right|_{t=0} \tilde{s}(t).$$

Remark: This shows that in some sense we shall not try to make precise here covariant differentiation is the infinite-simal form of parallel translation. One should regard parallel translation as the basic geometric concept and the operator ∇ as a convenient computational description of it.

Remark. Let $s^1,\, \ldots,\, s^k$, and $\tilde{s}_1,\, \ldots,\, \tilde{s}_k$ be two local bases for σ^*E and w and $\tilde{w}$ their respective connection forms relative to ∇^σ. If $g :\, I \to GL(k)$ is the gauge transition function from s to $\tilde{s}$ (i.e. $\tilde{s}^\alpha = g_{\alpha\beta}s^\alpha$) we know that $w = \tilde{w} + g^{-1}dg$. Now suppose $v_1,\, \ldots,\, v_k$ is a basis for $E_{\sigma(t_0)}$ and we define $\tilde{s}^\alpha(t) = P_\sigma(t,\, t_0)v_\alpha$, so that the $\tilde{s}^\alpha$ are covariant constant, and hence $\nabla^\sigma \tilde{s}^\alpha = 0$ so $\tilde{w} = 0$. Thus ∇^σ looks like d in the basis $\tilde{s}^\alpha$ and for the arbitrary basis $s^1,\, \ldots,\, s^k$ we see that its connection form w has the form $g^{-1}dg$.

Holonomy. Given $p \in M$ let Λ_p denote the semi-group of all smooth closed loops $\sigma : [0, 1] \to M$ with $\sigma(0) = \sigma(1) = p$. For $\sigma \in \Lambda_p$ let $P_\sigma = P_\sigma(1, 0) : E_p \simeq E_p$ denote parallel translation around σ. It is clear that $\sigma \to P_\sigma$ is a homomorphism of Λ_p into $GL(E_p)$. Its image is called the holonomy group of ∇ (at p; conjugation by $P_\gamma(1, 0)$ where γ is a smooth path from p to q clearly is an isomorphism of this group onto the holonomy group of ∇ at q).

Remark. It is a (non-trivial) fact that the holonomy group at p is a (not necessarily closed) Lie subgroup of $GL(E_p)$. By a Theorem of Ambrose and Singer its Lie algebra is the linear span of the image of the curvature form Ω_p in $L(E_p, E_p)$.

Exercise. If ∇ is flat show that $P_\sigma = id$ if σ is homotopic to the constant loop at p, so that in this case $\sigma \mapsto P_\sigma$ induces a homomorphism $P : \pi_1(M) \to GL(E_p)$. This latter homomorphism need not be trivial. Let $M = \mathbb{C}^* = \mathbb{C} - \{0\}$ and consider the connection $\nabla^c = d + w^c$ on $M \times \mathbb{C} = M \times \mathbb{R}^2$ defined by the connection 1-form w^c on M with values in $\mathbb{C} = L_{\mathbb{C}}(\mathbb{C}, \mathbb{C}) \subseteq L(\mathbb{R}^2, \mathbb{R}^2)$ $w^c = c\frac{dz}{z} = cd(\log z)$.

ADMISSIBLE CONNECTIONS ON G-BUNDLES

We first recall some of the features of a G-bundle E over M. At each $p \in M$

there is a special class of admissible frames $e^1, \ldots, e^k$ for the fiber E_p. Given one such frame every other admissible frame at p, $\tilde{e}^1, \ldots, \tilde{e}^k$ is uniquely of the form $\tilde{e} = g_{\alpha\beta} e^\alpha$ where $g = (g_{\alpha\beta}) \in G$ and conversely every $g \in G$ determines an admissible frame in this way — so once some admissible frame is picked, all admissible frames at p correspond bijectively with the group G itself. A linear map $T : E_p \to E_q$ is called a G-map if it maps admissible frames at p to admissible frames at q — or equivalently if given admissible frames $e_p^1, \ldots, e_p^k$ at p and $e_q^1, \ldots, e_q^k$ at q the matrix of T relative to these frames lies in G. In particular the group of G-maps of E_p with itself is denoted by $\mathrm{Aut}(E_p)$ and called the group of G-automorphism of the fiber E_p. It is clearly a subgroup of $GL(E_p)$ isomorphic to G, and its Lie algebra is thus a subspace of $L(E_p, E_p) = L(E, E)_p$ denoted by $L_G(E_p, E_p)$ and isomorphic to $\mathcal{G}$. ($T \in L_G(E_p, E_p)$ iff $\exp(tT)$ is a G map of E_p for all t).

Definition. A connection ∇ for E is called admissible if for each smooth path $\sigma : [0, 1] \to M$ parallel translation along σ is a G map of $E_{\sigma(0)}$ to $E_{\sigma(1)}$.

Theorem. A NASC that a connection ∇ for E be admissible is that the matrix w of connection one-forms w.r.t. admissible bases have values in the Lie algebra of G.

Proof. Exercise.

Corollary. If ∇ is admissible then its curvature two form Ω has values in the subbundle $L_G(E, E)$ of $L(E, E)$.

QUASI-CANONICAL GAUGES

Let $x_1, \ldots, x_n$ be a convex coordinate system for M in θ with $p_0 \in \theta$ the origin. Given a basis $v^1, \ldots, v^k$ for E_{p_0} we get a local basis $s^1, \ldots, s^k$ for E over θ by letting $s^i(p)$ be the parallel translate of v^i along the ray $\sigma(t)$ joining p_0 to p (i.e. $x_i(\sigma(t)) = tx_i(p)$). If ∇ is an admissible connection and $v^1, \ldots, v^k$ an admissible basis at p_0 then $s^1, \ldots, s^k$ is an admissible local basis called a quasi canonical gauge for E over θ.

Exercise: Show that the s^i really are smooth (Hint: solutions of ODE depending on parameters are smooth in the parameters as well as the initial conditions) and also show that the connection forms for ∇ relative $s^1, \ldots, s^k$ all vanish at p_0.

THE GAUGE EXTERIOR DERIVATIVE

Given a connection ∇ for a vector bundle E we define linear maps

$$D_p^\nabla = D_p : \Gamma(\Lambda^p(M) \otimes E) \to \Gamma(\Lambda^{p+1}(M) \otimes E)$$

called gauge exterior derivative by:

$$\begin{aligned} D_p \cdot w(x_1, \ldots, x_{p+1}) &= \sum_{i=1}^{p+1} (-1)^{i+1} \nabla_{x_i} w(x_1, \ldots, \hat{x}_i, \ldots, x_{p+1}) \\ &+ \sum_{1 \le i < j \le p+1} (-1)^{i+j} w([x_i, x_j], x_1, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_{p+1}) \end{aligned}$$

for $x_1, \ldots, x_{p+1}$ smooth vector fields on M. Of course one must show that this really defines $D_i w$ as a $(p+1)$-form, i.e. that the value of the above expression at a point q depends only on the values of the x_i at q. (Equivalently, since skew-symmetry is clear it suffices to check that $D_p w(f x_1, \ldots, x_{p+1} = f D_p w(x_1, \ldots, x_{p+1})$ for f a smooth real valued function on M).

Exercise: Check this. (Hint: This can be considered well known for the flat connection $\nabla = d$ — and relative to a local trivialization ∇ anyway looks like $d + w$).

Remark: Note that $D_0 = \nabla$.

Theorem. If $s \in \Gamma(E)$ then

$$(D_1 D_0 s)(X, Y) = \Omega(X, Y) s.$$

Proof.

$$\begin{aligned} & D_1(D_0 s)(X, Y) \\ = \; & \nabla_X D_0 s(Y) - \nabla_Y D_0 s(X) - D_0 s([X, Y]) \\ = \; & \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X,Y]} s. \end{aligned}$$

Corollary: $D_p : \Gamma(\Lambda^p(M) \otimes E) \to \Gamma(\Lambda^{p+1}(M) \otimes E)$ is a complex iff $\Omega = 0$.

Theorem. If ∇^i is a connection in E_i, $i = 1, 2$, and ∇ is the corresponding connection on $E^1 \otimes E^2$ then if w_i is a p_i-form on M with values in E_i

$$D^{\nabla}_{p_1+p_2} w_1 \tilde{\wedge} w_2 = D^{\nabla_1}_{p_1} w_1 \tilde{\wedge} w_2 + (-1)^{p_1} w_1 \tilde{\wedge} D^{\nabla^2}_{p_2} w_2.$$

Proof. Exercise. (Hint. Check equality at some point p by choosing a quasi-canonical gauge at p).

Corollary. If w_1 is a real valued p-form and w_2 an E valued q-form then

$$D(w_1 \wedge w_2) = dw_1 \wedge w_2 + (-1)^p w_1 \wedge Dw_2.$$

Corollary. If λ is a real values p-form and s is a section of E

$$D(\lambda \otimes s) = d\lambda \wedge s + (-1)^p \lambda \wedge \nabla s.$$

Theorem. Let λ be a p-form with values in E and let $s^1, \ldots, s^k$ be a local basis for E. Then writing $\lambda = \sum_{\alpha} \lambda_\alpha s^\alpha$ where the λ_α are real valued p-forms

$$D\lambda = (d\lambda_\alpha + w_{\alpha\beta} \wedge \lambda_\beta) s^\alpha$$

where w is the matrix of connection forms for ∇ relative to the s^α. Thus the formula for gauge covariant derivative relative to a local base may be written

$$D\lambda = d\lambda + w \wedge \lambda$$

Proof. Exercise.

Exercise. Use this to get a direct proof that

$$\begin{aligned} D_1 D_0 \lambda &= \Omega\lambda \\ &= \sum (\Omega_{\alpha\beta} \lambda_\beta) s^\alpha \end{aligned}$$

where

$$\Omega = dw + w \wedge w.$$

Theorem. Given a connection ∇ on E let $\tilde{\nabla}$ be the induced connection of $L(E, E) = E^* \otimes E$. If w is the matrix of connection 1-forms for ∇ relative to a local basis $s^1, \ldots, s^k$ and α is a p-form on M with values in $L(E, E)$ (given locally by the matrix $\alpha_{\alpha\beta}$ of real valued p-forms, where $\alpha = \alpha_{\alpha\beta} \overset{*}{s}{}^{\alpha} \otimes s^{\beta}$) then

$$D^{\tilde{\nabla}}\alpha = d\alpha + w \wedge \alpha - (-1)^p \alpha \wedge w$$

Proof. Exercise.

Corollary. The Bianchi identity $d\Omega + w\wedge\Omega - \Omega\wedge w$ is equivalent to the statement $D\Omega = 0$ (or more explicitly $\nabla^{\tilde{\nabla}}\Omega^{\nabla} = 0$).

Proof. Take $p = 2$ and $\alpha = \Omega$.

Exercise: Given a connection ∇ on E and $\gamma \in \Delta(E) = \Gamma(\Lambda^1(M) \otimes L(E, E))$ let $\nabla^\gamma = \nabla + \gamma$ and let $D = D^{\nabla}$. Show that the curvature Ω^γ of ∇^γ is related to the curvature Ω of Ω by

$$\Omega^\gamma = \Omega + D\gamma + \gamma \wedge \gamma$$

(Remark: Note how this generalizes the formula $\Omega = dw + w \wedge w$ which is the special case $\nabla = d$ (so $D = d$) and $\gamma = w$).

CONNECTION ON TM

In this section ∇^T is a connection on TM, τ its torsion, $x_1, \ldots, x_n$ a local coordinate system for M in θ, $X^\alpha = \frac{\partial}{\partial x_\alpha}$ the corresponding natural basis for TM over θ, $w_{\alpha\beta}$ the connection form for ∇^T relative to this basis ($\nabla^T X = w_{\alpha\beta} X^\alpha$) and $\Gamma^\alpha_{\gamma\beta}$ the Christoffel symbols ($w_{\alpha\beta} = \Gamma^\alpha_{\gamma\beta} dx_\gamma$). Also, g will be a pseudo-Riemannian metric for M and $g_{\alpha\beta}$ its components with respect to the coordinate system $x_1, \ldots, x_n$, i.e. $g_{\alpha\beta}$ is the real valued function $g(X^\alpha, X^\beta)$ defined in θ. We note that the torsion τ is a section of $\Lambda^2(M) \otimes \mathrm{TM} \subseteq T^*M \otimes T^*M \otimes \mathrm{TM} \simeq T^*M \otimes L(\mathrm{TM}, \mathrm{TM}) = \Delta(\mathrm{TM})$, hence we can define another connection $\tilde{\nabla}^T$ (called the torsionless part of ∇^T) by

$$\tilde{\nabla}^T = \nabla^T - \frac{1}{2}\tau$$

or more explicitly

$$\tilde{\nabla}^T_x y = \nabla^T_x y - \frac{1}{2}\tau(x,\, y)$$

Exercise. Check that $\tilde{\nabla}^T$ has in fact zero torsion.

The remaining exercises work further details of this situation.

Exercise: Write $\tau = \tau^\alpha_{\gamma\beta} dx_\gamma \otimes dx_\beta \otimes x^\alpha$ and show that

$$\tau^\alpha_{\gamma\beta} = \Gamma^\alpha_{\gamma\beta} - \Gamma^\alpha_{\beta\gamma}$$

so $\tilde{\nabla}^T$ has Christoffel symbols $\tilde{\Gamma}^\alpha_{\gamma\beta}$ given by the "symmetric part" of $\Gamma^\alpha_{\gamma\beta}$ w.r.t. its lower indices

$$\tilde{\Gamma}^\alpha_{\gamma\beta} = \frac{1}{2}(\Gamma^\alpha_{\gamma\beta} + \Gamma^\alpha_{\beta\gamma})$$

and ∇ has zero torsion iff $\Gamma^\alpha_{\gamma\beta}$ is symmetric in its lower indices.

Recall that a curve σ in M is called a geodesic for ∇^T if its tangent vector field σ', considered as a section of $\sigma^*(\mathrm{TM})$. is parallel along σ. Show that if σ lies in θ and $x_\alpha(\sigma(t)) = \sigma_\alpha(t)$ then the condition for this is

$$\frac{d^2\sigma_\alpha}{dt^2} + \Gamma^\alpha_{\gamma\beta}(\sigma(t))\frac{d\sigma}{dt}\frac{d\sigma_\beta}{dt} = 0$$

[Note this depends only on the symmetric part of $\Gamma^\alpha_{\gamma\beta}$ w.r.t. its lower indices, so ∇^T and $\tilde{\nabla}^T$ have the same geodesics].

Exercise. Let $p_0 \in \theta$. Show there is a neighborhood U of 0 in $\mathbb{R}^n$ such that if $v \in U$ then there is a unique geodesics $\sigma_v : [0,\, 1] \to M$ with $\sigma_v(0) = p_0$ at $\sigma'_v(0) = \sum\limits_\alpha v_\alpha x^\alpha(p_0)$. Define "geodesic coordinates" at p_0 by the map $\phi : U \to M$ given by $\phi(v) = \sigma_v(1)$. Prove that $D\phi_0(v) = \sum\limits_\alpha v^\alpha x^\alpha(p_0)$ so $D\phi_0$ is linear isomorphism of $\mathbb{R}^n$ onto TM_{p_0} and hence by the inverse function theorem ϕ is in fact a local coordinate system at p_0. Show that in these coordinate the Christoffel symbols of $\tilde{\nabla}^T$ vanish at p_0. [Hint: show that for small real s, $\sigma_{sv}(t) = \sigma_v(st)$].

Now let ∇ be a connection in any vector bundle E over M. Let $x^1, \ldots, x^n$ be the coordinate basis vectors for TM with respect to a geodesic coordinate system at p_0 as above and let $s^1, \ldots, s^k$ be a quasi-canonical gauge for E at p_0 constructed from these coordinates — i.e. the s^α are parallel along the geodesic rays emanating from p_0. Then if (and only if) ∇^T is torsion free, so $\nabla^T = \tilde{\nabla}^T$, connection forms for ∇^T and ∇ w.r.t. $x^1, \ldots, x^n$ and $s^1, \ldots, s^k$ respectively all vanish at p_0. Using this prove:

Exercise: Let $\hat{\nabla}$ be the connection $\hat{\nabla} : \Gamma(\otimes^p T^*M \otimes E) \to \Gamma(\otimes^{p+1} T^*M \otimes E)$ coming from ∇^T on TM and ∇ on E. If w is a p-form on M with values in E then $D_p w$ coincides with $\mathrm{Alt}(\hat{\nabla} w)$, the skew-symmetrized $\hat{\nabla} w$, provided ∇^T has zero torsion.

Exercise: If we denote still by ∇^T the connection on $T^*M \otimes T^*M$ induced by ∇^T recall that g is admissible for the $0(n)$ structure defined by g iff $\nabla^T g = 0$. Show this is equivalent to the condition

$$\frac{\partial g_{\alpha\beta}}{\partial x_\gamma} = g_{\beta\gamma} \Gamma^\lambda_{\alpha\gamma} + g_{\alpha\lambda} \Gamma^\lambda_{\gamma\beta}.$$

Show that if ∇^T has torsion zero then this can be solved uniquely in θ for the $\Gamma^\lambda_{\alpha\gamma}$ in terms of the $g_{\alpha\beta}$ and their first partials.

YANG-MILLS FIELDS

We assume as given a smooth, Riemannian, n-dimensional manifold M and a k-dimensional smooth vector bundle E over M with structure group a compact subgroup G of $O(k)$ with Lie algebra $\mathcal{G}$. We define for each $\xi \in \mathcal{G}$ a linear map $\mathrm{Ad}(\xi) : \mathcal{G} \to \mathcal{G}$ by $\eta \to [\xi, \eta] = \xi\eta - \eta\xi$. We assume that the "Killing form" $< \xi, \eta >= -\mathrm{tr}(\mathrm{Ad}(\xi)\mathrm{Ad}(\eta))$ is positive definite — which is equivalent to the assumption that G is semi-simple. By the Jacobi identity each $\mathrm{Ad}(\xi)$ is skew-adjoint w.r.t. this inner product. Since

$$\begin{aligned}\exp(t\mathrm{Ad}(\xi))\eta &= (\exp t\xi)\eta(\exp t\xi)^{-1} \\ &= \mathrm{ad}(\exp t\xi)\eta\end{aligned}$$

this means that the action of G on $\mathcal{G}$ by inner automorphisms is orthogonal w.r.t. the Killing form.

Recall that $L_G(E, E) \subseteq L(E, E)$ is the vector bundle whose fiber at p is the Lie algebra of the group of G-automorphisms of E_p. A gauge $\theta \times \mathbb{R}^k \simeq E|\,\theta$ induces a gauge $\theta \times \mathcal{G} \simeq L_G(E, E)|\,\theta$ and the gluing together under different gauge is by ad( ). So $L_G(E, E)$ has a canonical inner product which is preserved by any connection coming from a G-connection in E. Thus there is a Hodge *-operator naturally defined for forms with values in $L_G(E, E)$ by means of the pseudo Riemann structure in TM and this Killing Riemannian structure for $L_G(E, E)$.

Now in particular if ∇ is a G-connection in E then its curvature $\Omega = \Omega^\nabla$ is a two form with values in $L_G(E, E)$ so $^*\Omega$ is an $n-2$ form with values in $L_G(E, E)$ and $\delta\Omega = {}^*D^*\Omega$ is a 1-form with values in $L_G(E, E)$. The ("free") Yang-Mills equation for ∇ are

$$\begin{cases} \delta\Omega = 0 \\ D\Omega = 0 \end{cases}$$

Note that the second equation is actually an identity (i.e. automatically satisfied) namely the Bianchi identity.

Let us define the action of a connection ∇ on M to be

$$A(\nabla) = \frac{1}{2}\int_M \|\Omega^\nabla\|^2\mu = \frac{1}{2}\int_M \Omega^\nabla \wedge {}^*\Omega^\nabla$$

Now let $\gamma \in \Delta_G(E) = \Gamma(\Lambda^1(M) \otimes L_G(E, E))$ so $\nabla^\gamma = \nabla + \gamma$ is another G connection on E. We recall that

$$\Omega^{\nabla\gamma} = \Omega^\nabla + D^\nabla\gamma + \gamma \wedge \gamma.$$

Thus

$$A(\nabla + t\gamma) = A(\nabla) + t\int \Omega^\nabla \wedge {}^*D\gamma + t^2\int \cdots$$

and

$$\left.\frac{d}{dt}\right|_{t=0} A(\nabla + t\gamma) = ((\Omega^{\nabla},\, D^{\nabla}_{\gamma})) = \pm(({}^*D^{\nabla *}\Omega^{\nabla},\, \gamma))$$

Thus the Euler-Lagrange condition that ∇ be an extremal of A is just the non-trivial Yang-Mills equation $\delta\Omega = 0$.

We note that the Riemannian structure of M enters only very indirectly (via the Hodge *-operator on two-forms) into the Yang-Mills equation. Now if we change the metric g on M to a conformally equivalent metric $\tilde{g} = c^2 g$ it is immediate from the definition that the new $*$-operator on k-forms is related to the old by $\tilde{*}_k = c^{2k-n}{*}_k$. In particular if n is even and $k = n/2$ then we see $^*_k$ is invariant under conformal change of metric on M. Thus

Theorem. If $\dim(M) = 4$ then the Yang-Mills equations for connection on G-bundles over M is invariant under conformal change of metric on M. In particular since $\mathbb{R}^4$ is conformal to $S^4 - \{p\}$ under stereographic projection, there is a natural bijective correspondence between Yang-Mills fields on $\mathbb{R}^4$ and Yang-Mills fields on $S^4 - p$.

If Ω is a Yang-Mills field on S^4 then of course by the compactness of S^4 its action

$$\int_{\mathbb{R}^4} \Omega \wedge {}^*\Omega = \int_{S^4-\{p\}} \Omega \wedge {}^*\Omega = \int_{S^4} \Omega \wedge {}^*\Omega$$

is finite. By a remarkable theorem of K. Uhlenbeck, conversely any Yang-Mills field of finite action on $\mathbb{R}^4$ extends to a smooth Yang-Mills field on S^4. Thus the finite action Yang-Mills fields on $\mathbb{R}^4$ can be identified with all Yang-Mills fields in S^4.

Now also when $n = 4$ we recall that $({}^*{}_2)^2 = 1$ and hence

$$\Lambda^2(M) \otimes F \quad (F = L_G(E,\, E))$$

splits as a direct-sum into the sub-bundles $(\Lambda^2(M) \otimes F)_{\pm}$ of ± 1 eigenspaces of ${}^*{}_2$. In particular any curvature form Ω of a connection in M splits into the sum of its projection $\Omega_\pm$ on these two bundles

$$\Omega = \Omega_+ + \Omega_- \quad {}^*(\Omega_\pm) = \pm\Omega_\pm.$$

Now in general a two-form γ is called self-dual (anti-self dual) if ${}^*\gamma = \gamma$ (${}^*\gamma = -\gamma$) and a connection is called self dual (anti-self dual) if its curvature is self-dual (anti-self dual).

Theorem. If $\dim(M) = 4$ then self-dual and anti-self dual connections are automatically solutions of the Yang-Mills equation.

Proof. Since ${}^*\Omega = \pm\Omega$ the Bianchi identity $D\Omega = 0$ implies $D({}^*\Omega) = 0$.

Definition. An instanton (anti-instanton) is a self-dual (anti-self dual) connection on $\mathbb{R}^4$ with finite action, or equivalently a self-dual (anti-self dual) connection on S^4.

It is an important open question whether there exist any Yang-Mills fields in S^4 which are not self dual or anti-self dual.

Theorem. If E any smooth vector bundle over S^4 then the quantity

$$C = C(E) = \int_{S^4} \Omega \wedge \Omega$$

where Ω is the curvature form of a connection ∇ on E is a constant (in fact 8π times an integer) called the 2^{nd} Chern number of E, depending only on E and not on ∇. If we write $a = \int_{S^4} \Omega \wedge {}^*\Omega$ for the Yang-Mills action of Ω and $a_+ = \int \Omega_+ \wedge {}^*\Omega_+$, $a_- = \int \Omega_- \wedge {}^*\Omega_-$ for the lengths of Ω_+ and Ω_-, then

$$a = a_+ + a_- \quad \text{and} \quad c = a_+ - a_-$$

so

$$a = c + 2a_- = -c + 2a_+.$$

Thus if $c \geq 0$ then an instanton (i.e. $a_- = 0$) is an absolute minimum of the action and if $c \leq 0$ then an anti-instanton (i.e. $a_+ = 0$) is an absolute minimum of the action.

Proof. The fact that c is independent of the connection on E follows from our discussion of characteristic classes and numbers below. Now:

$$\begin{aligned} c &= \int(\Omega_+ + \Omega_-)\wedge {}^*(\Omega_+ - \Omega_-) \\ &= \int \Omega_+ \wedge {}^*\Omega_+ - \int \Omega_- \wedge {}^*\Omega_- \\ &= a_+ - a_- \end{aligned}$$

$$\begin{aligned} a &= \int(\Omega_+ + \Omega_-)\wedge {}^*(\Omega_+ + \Omega_-) \\ &= \int \Omega_+ \wedge {}^*\Omega_+ + \int \Omega_- \wedge {}^*\Omega_- \\ &= a_+ + a_-. \end{aligned}$$

□

TOPOLOGY OF VECTOR BUNDLES

In this section we will review some of the basic facts about the topology of vector bundles in preparation for a discussion of characteristic classes.

We denote by $\mathrm{Vect}_G(M)$ the set of equivalence classes of G-vector bundles over M. Given $f : N \to M$ we have an induced map, given by the "pull-back" construction:

$$f^* : \mathrm{Vect}_G(M) \to \mathrm{Vect}_G(N).$$

If E is smooth G-bundle over N we will denote by $E \times I$ the bundle we get over $N \times I$ by pulling back E under the natural projection of $N \times I \to N$ (so the fiber of $(E \times I)$ at $(x,\, t)$ is the fiber of E at x).

Lemma. Any smooth G-bundle $\tilde{E}$ over $N \times I$ is equivalent to one of the form $E \times I$.

Proof. Let $E = {i_0}^*(\tilde{E})$ where $i_0 : N \to N \times I$ is $x \to (x,\, 0)$ so $E_x = \tilde{E}_{(x,\,0)}$ at hence $(E \times I)_{(x,\,t)} = \tilde{E}_{(x,\,0)}$. Choose an admissible connection for $\tilde{E}$ and define an equivalence $\phi : E \times I \simeq \tilde{E}$ by letting $\phi_{(x,t)} : (E \times I)_{(x,t)} \simeq \tilde{E}_{(x,t)}$ be parallel translation of $\tilde{E}_{(x,0)}$ along $\tau \to (x, \tau)$.

Theorem. If $f_0, f_1 : N \to M$ are homotopic then $f_0^* : \mathrm{Vect}_G(M) \to \mathrm{Vect}_G(N)$ and $f_1^* : \mathrm{Vect}_G(M) \to \mathrm{Vect}_G(N)$ are equal.

Proof. Let $F : N \times I \to M$ be a homotopy of f_0 with f_1 at let $i_t : N \to N \times I$ be $x \to (x, t)$. Let $\tilde{E} = F^*(E')$ for some $E' \in \mathrm{Vect}_G(M)$. By the Lemma $\tilde{E} \simeq E \times I$ for some bundle E over N, hence since $f_t = F \circ i_t$ $(t = 0, 1)$ $f_t^*(E') = {i_t}^* F^*(E') \simeq {i_t}^*(E \times I) = E$.

Corollary. If M is a contractible space then every smooth G bundle over M is trivial.

Proof. Let f_0 denote the identity map of M and let $f_1 : M \to M$ be a constant map to some point p. If $E \in \mathrm{Vect}_G(M)$ then $f_0^*(E) = E$ and $f_1^*(E) = E_p \times M$. Since M is contractible f_0 and f_1 are homotopic and $E \simeq E_p \times M$.

Corollary. If E is a G-bundle over S^n then E is equivalent to a bundle formed by taking trivial bundles E^+ and E^- over the (closed) upper and lower hemispheres D^n_+ and D^n_- and gluing them along the equator $S^{n-1} = D^n_+ \cap D^n_-$ by a map $g : s^{n-1} \to G$. Only the homotopy class of g matters in determining the equivalence class of E, so that we have a map $\mathrm{Vect}_G(s^n) \to \pi_{n-1}(G)$ which is in fact an isomorphism.

Proof. Since $D^n_\pm$ are contractible $E|D^n_\pm \simeq D^n \times E_p$, so we can in fact reconstruct E by gluing. It is easy to see that a homotopy between gluing maps g_0 and g_1 of $s^{n-1} \to G$ defines an equivalence of the glued bundles and vice versa.

Remark. It is well-known that for a simple compact group G $\pi_3(G) = \mathbb{Z}$, so in particular $\mathrm{Vect}_G(S^4) \simeq \mathbb{Z}$. In fact $E \mapsto \frac{1}{8\pi} C(E) = \frac{1}{8\pi} C(E) = \frac{1}{8\pi} \int \Omega \wedge \Omega$ $(= 2^{\text{nd}}$ Chern number of $E)$ Where Ω is the curvature of any connection on E gives this map.

The preceding corollary is actually a special case of a much more general fact, the <u>bundle</u> <u>classification</u> <u>theorem</u>. It turns out that for any Lie group G and positive integer N we can construct a space $B_G = B_G^N$ (the classifying space of G for spaces of dimension $< N$) and a smooth G vector bundle $\xi_G = \xi_G^N$ over B_G (the "universal" bundle) so that if M is any smooth manifold of dimension $< N$ and E is any smooth G vector bundle over M then $E = f^*\xi_G$ for some smooth map $f : M \to B_G$; in fact the map $f \to f^*\xi_G$ is a bijective correspondence between the set $[M, B_G]$ of homotopy class of M into B_G and the set $\mathrm{Vect}_G(M)$ of equivalence classes of smooth G-vector bundles over M. This is actually not as hard as it might seem and we shall sketch the proof below for the special case of $GL(k)$ (assuming basic transversality theory).

<u>Notation</u>.

$$Q^r = \{T \in L(\mathbb{R}^k, \mathbb{R}^\ell) \,|\, \mathrm{rank}(T) = r\}$$

<u>Proposition</u>. Q^r is a submanifold of $L(\mathbb{R}^k, \mathbb{R}^\ell)$ of dimension $k\ell - (k-r)(\ell - r)$, hence of condimension $(k-r)(\ell-r)$. Thus if $\dim(M) < \ell - k + 1$, i.e. if $\dim(M) \leq (\ell - k)$, then any smooth map of M into $L(\mathbb{R}^k, \mathbb{R}^\ell)$ which is transversal to $Q^0, \ldots, Q^{k-1}$ will in fact have rank k everywhere.

<u>Proof</u>. We have an action $(g, h)T = gTh^{-1}$ of $GL(\ell) \times GL(k)$ on $L(\mathbb{R}^k, \mathbb{R}^\ell)$ and Q^r is just the orbit of $P_r =$ projection of $\mathbb{R}^k$ into $\mathbb{R}^r \to \mathbb{R}^\ell$. The isotropy group of P_r is the set of (g, h) such that $gP_r = P_r h$. If $e_1, \ldots, e_r, \ldots, e_k, \ldots, e_\ell$ is usual basis then this means $ge_i = P_r he_i$ $i = 1, \ldots, r$ and $P_r he_i = 0$ $i = r+1, \ldots, k$. Thus (g, h) is determined by:

(a) $h|\mathbb{R}^r \in L(\mathbb{R}^r, \mathbb{R}^k)$

(b) $P_r^\perp h|\mathbb{R}^{k-r} \in L(\mathbb{R}^{k-r}, \mathbb{R}^{k-r})$ $\quad (P_r h|\mathbb{R}^{k-r} = 0)$

(c) $g|\mathbb{R}^{\ell-r} \in L(\mathbb{R}^{\ell-r}, \mathbb{R}^\ell)$ $\quad (g|\mathbb{R}^r$ det. by $h)$.

$rk + (k-r)(k-r) + \ell(\ell-r) = (\ell^2 + k^2) - (k\ell - (k-r)(\ell-r))$ is the dimension of the isotropy group of Q^r hence:

$$\dim(Q^r) \quad = \quad \dim(GL(\ell) \times GL(k)/\text{isotropy group})$$

$$= \; k\ell - (k-r)(\ell - r).$$

□

Extension of Equivalence Theorem. If M is a (compact) smooth manifold and $\dim(M) \leq (\ell - k)$ then any smooth k-dimensional vector bundle over M is equivalent to a smooth subbundle of the product bundle $\mathbb{R}^\ell_M = M \times \mathbb{R}^\ell$, and in fact if N is a closed smooth submanifold of M then an equivalence of $E|N$ with a smooth sub-bundle of $\mathbb{R}^\ell_N$ can extended to an equivalence of E with a sub-bundle of $\mathbb{R}^\ell_M$.

Proof. An equivalence of E with a sub-bundle of $\mathbb{R}^\ell_M$ is the same as a section ψ of $L(E, \mathbb{R}^\ell_M)$ such that $\psi_x : E_x \to \mathbb{R}^\ell$ has rank k for all x. In terms of local trivialization of E, ψ is just a map of M into $L(\mathbb{R}^k, \mathbb{R}^\ell)$ and we want this map to miss the submanifolds Q^r, $r = 0, 1, \ldots, k-1$. Then Thom transversality theorem and preceding proposition now complete the proof. □

$$\begin{aligned}
G &= GL(k) \\
B_G &= G(k, \ell) \\
&= k - \text{dim. linear subspace of } \mathbb{R}^\ell \\
&= \{T \in L(\mathbb{R}^k, \mathbb{R}^\ell) \,|\, T^2 = T,\, T^* = T,\, \mathrm{tr}(T) = k\} \\
\xi^G &= \{(P, v) \in G(k, \ell) \times \mathbb{R} \,|\, P_v = v \quad \text{i.e. } v \in \mathrm{im}(P)\}
\end{aligned}$$

This is a vector bundle over B_G whose fiber at p is image of p.

Bundle Classification Theorem

If $\dim(M) < (\ell - k)$ then any smooth G-bundle E over M is equivalent to $f_0^*\xi^G$ for some smooth $f_0 : M \to B_G$. If also E is equivalent to $f_1^*\xi^G$ then f_0 and f_1 are homotopic.

Proof. By preceding theorem we can find an isomorphism ψ^0 of E with a k-dimensional sub-bundle of $\mathbb{R}^\ell_M$. Then if $f_0 : M \to B_G$ is the map $x \mapsto \mathrm{im}(\psi^0_x)$, ψ^0 is an equivalence of E with $f_0^* \xi^G$. Now suppose $\psi^1 : E \simeq f_1^* \xi^G$ and regard $\psi^0 \cup \psi^1$ as an equivalence over $N = M \times \{0\} \cup M \times \{1\}$ of $(E \times I)|N$ with a sub-bundle of $\mathbb{R}^\ell_N$. The above extension theorem says (since $\dim(M \times I) = \dim M + 1 \leq (\ell - k)$) that this can be extended to an isomorphism ϕ of $E \times I$ with a k-dimensional sub-bundle of $\mathbb{R}^\ell_{M \times I}$. Then $F : M \times I \to B_G$. $(x, t) \to \mathrm{im}(\phi_{(x,t)})$ is a homotopy of f_0 with f_1. $\square$

CHARACTERISTIC CLASSES AND NUMBERS

The theory of "characteristic classes" is one of the most remarkable (and mysterious looking) parts of bundle theory. Recall that given a smooth map $f : N \to M$ we have induced maps $f^* : \mathrm{Vect}_G(M) \to \mathrm{Vect}_G(N)$ and also $f^* : H^*(M) \to H^*(N)$, where $H^*(M)$ denotes the de Rham cohomology ring of M. Both of these f^*'s depend only on the homotopy class of $f : N \to M$. Informally speaking a characteristic class C (for G-bundles) is a kind of snapshot of bundle theory in cohomology theory. It associates to each G-bundle E over M a cohomology class $C(E)$ in $H^*(M)$ in a "natural" way — where natural means commuting with f^*.

Definition. A characteristic class for G-bundle is a function c which associates to each smooth G vector bundle E over a smooth manifold M an element $c(E) \in H^*(M)$, such that if $f : N \to M$ is a smooth map then $c(f^*E) = f^*c(E)$. [In the language of category theory this is more elegant: c is a natural transformation from the functor Vect_G to the functor H^*, both considered as contravariant functors in the category of smooth manifolds]. We denote by $\mathrm{Char}(G)$ the set of all characteristic classes of G bundles.

Remark. Since $f^* : H^*(M) \to H^*(N)$ is always a ring homomorphism it is clear that $\mathrm{Char}(G)$ has a natural ring structure.

Exercise. Let ξ_G be a universal G bundle over the G-classifying space B_G. Show that the map $c \mapsto c(\xi_G)$ is a ring isomorphism of $\mathrm{Char}(G)$ with $H^*(B_G)$; i.e. every characteristic class arises as follows: Choose an element $c \in H^*(B_G)$ and given $E \in \mathrm{Vect}_G(M)$ let $f : M \to B_G$ be its classifying map (i.e. $E \simeq f^*\xi_G$). Then define $c(E) = f^*(c)$.

Unfortunately this description of characteristic classes, pretty as it seems, is not very practical for their actual calculation. The Chern-Weil theory, which we discuss below for the particular case of $G = GL(k, c)$ on the other hand seems much more complicated to describe, but is ideal for calculation.

First let us define characteristic numbers. If $w \in H^*(M)$ with M closed, then $\int_M w$ denote the integral over M of the $\dim(M)$-dimensional component of w. The numbers $\int_M c(E)$, where $c \in \mathrm{Char}(G)$, are called the characteristic numbers of a G-bundle E over M. They are clearly invariant, i.e. equal for equivalent bundles, so they provide a method of telling bundles apart. [In particular if a bundle is trivial it is induced by a constant map, so all its characteristic classes and numbers are zero. Thus a non-zero characteristic number is a test for non-triviality]. In fact in good cases there are enough characteristic numbers to characterize bundles, hence their name.

THE CHERN-WEIL HOMOMORPHISM

Let $X = X_{\alpha\beta}$ $1 \leq \alpha, \beta \leq k$ be a $k \times k$ matrix of indeterminates and consider the polynomial ring $\mathbb{C}[X]$. If $P = P(X)$ is in $\mathbb{C}[X]$ then given any $k \times k$ matrix of elements $r = r_{\alpha\beta}$ from a commutative ring R we can substitute r for X in P and get an element $P(r) \in R$.

In particular if V is a k-dimensional vector space over $\mathbb{C}$ and $T \in L(V, V)$ then given any basis $e_1, \ldots, e_k$ for V, $Te_\beta = T_{\alpha\beta}e_\alpha$ defines the matrix $T_{\alpha\beta}$ of T relative to this basis, and substituting $T_{\alpha\beta}$ for $X_{\alpha\beta}$ in P gives a complex number $P(T_{\alpha\beta})$.

Given $g = g_{\alpha\beta} \in GL(k, \mathbb{C})$ define a matrix $\tilde{X} = \mathrm{ad}(g)X$ of linear polynomials in $\mathbb{C}[X]$ by

$$(\mathrm{ad}(g)X)_{\alpha\beta} = g^{-1}_{\alpha\lambda}X_{\lambda\gamma}g_{\gamma\beta}$$

If we substitute $(\mathrm{ad}(g)X)_{\alpha\beta}$ for $X_{\alpha\beta}$ in the polynomial $P \in \mathbb{C}[X]$ we get a new element $p^g = \mathrm{ad}(g)P$ of $\mathbb{C}[X] : p^g(X) = P(g^{-1}Xg)$. Clearly $\mathrm{ad}(g) : \mathbb{C}[X] \to C[X]$ is a ring automorphism of $\mathbb{C}[X]$ and $g \to \mathrm{ad}(g)$ is a homomorphism of $GL(k, \mathbb{C})$ into the group of ring automorphisms of $\mathbb{C}[X]$.

Definition. The subring of $\mathbb{C}[X]$ consisting of all $P \in \mathbb{C}[X]$ such that $P(X) = P^g(X) = P(g^{-1}Xg)$ for all $g \in GL(k, \mathbb{C})$ is called the ring of (adjoint) invariant polynomials and is denoted by $\mathbb{C}^G[X]$.

What is the significance of $\mathbb{C}^G[X]$?

Theorem. Let $P(X_{\alpha\beta}) \in \mathbb{C}[X_{\alpha\beta}]$. A NASC that p be invariant is the following: given any linear endomorphism $T : V \to V$ of a k-dimensional complex vector space V, the value $P(T_{\alpha\beta}) \in \mathbb{C}$ of P on the matrix $T_{\alpha\beta}$ of T w.r.t. a basis $e_1, \ldots, e_k$ for V does not depend on the choice of $e_1, \ldots, e_k$ but only on T and hence gives a well defined element $P(T) \in \mathbb{C}$.

Proof. If $\tilde{e}_\beta = g_{\alpha\beta}e_\alpha$ is any other basis for V then $g \in GL(k, \mathbb{C})$ and if $\tilde{T}_{\alpha\beta}$ is the matrix of T relative to the basis $\tilde{e}_\alpha$ then $\tilde{T}_{\alpha\beta} = g^{-1}_{\alpha\gamma}T_{\gamma\lambda}g_{\lambda\beta}$, from which the theorem is immediate. $\square$

Examples of Invariant Polynomials:

1) $Tr(X) = \sum_{\alpha} X_{\alpha\alpha} = X_{11} + \cdots X_{kk}$
2) $Tr(X^2) = \sum_{\alpha\beta} X_{\alpha\beta}X_{\beta\alpha}$
3) $Tr(X^m)$

4) $\det(X) = \sum_{\sigma \in S_k} \varepsilon(\sigma) X_{1\sigma(1)} X_{2\sigma(2)} \cdots X_{k\sigma(k)}$

Let t be a new indeterminate. If I is the $k \times k$ identity matrix then $tI + X = t\delta_{\alpha\beta} + X_{\alpha\beta}$ is a $k \times k$ matrix of elements in the ring $\mathbb{C}[X][t]$ so we can substitute it in det and set another element of the latter ring $\det(tI + X) = t^k + c_1(X)t^{k-1} + \cdots + c_k(X)$ where $c_1, \ldots, c_k \in \mathbb{C}[X]$ and clearly c_i is homogeneous of degree i. Since $g^{-1}(tI + X)g = tI + gXg^{-1}$ it follows easily from the invariance of det that $c_1, \ldots, c_k$ are also invariant.

Remark. Let T be any endomorphism of a k-dimensional vector space V. Choose a basis $e_1, \ldots, e_n$ for V so that T is in Jordan canonical form and let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of T. Then we see easily

1) $Tr(T) = \lambda_1 + \cdots + \lambda_k$

2) $Tr(T^2) = \lambda_1^2 + \cdots + \lambda_k^2$

3) $Tr(T^m) = \lambda_1^m + \cdots + \lambda_k^m$

4) $\det(T) = \lambda_1 \lambda_2 \cdots \lambda_k$

Also $\det(tI + T) = (t + \lambda_1) \cdots (t + \lambda_k)$ from which it follows that

5) $c_m(T) = \sigma_m(\lambda_1, \ldots, \lambda_k)$

where σ_m is the m^{th} "elementary symmetric function of the λ_i.

$$\sigma_m(\lambda_1, \ldots, \lambda_k) = \sum_{1 \leq i_1 < \cdots < i_m \leq k} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k}$$

This illustrates a general fact.

If $P \in \mathbb{C}^G[x]$ then there is a uniquely determined symmetric polynomial of k-variables $\hat{P}(\Lambda_1, \ldots, \Lambda_k)$ such that if $T : V \to V$ is as above then $P(T) = P(\lambda_1, \ldots, \lambda_k)$. We see this by checking an the open, dense set of diagonalizable T. Now recall the fundamental fact that the symmetric polynomials in $\Lambda_1, \ldots, \Lambda_k$ form a polynomial ring in the k generators $\sigma_1, \ldots, \sigma_k$ or in the k power sums $p_1, \ldots, p_k$; $p_m(\Lambda_1, \ldots, \Lambda_k) = \Lambda_1^m + \cdots \Lambda_k^m$. From this we see the structure of $\mathbb{C}^G[X]$.

Theorem. The ring $\mathbb{C}^G[X]$ of adjoint invariant polynomials in the $k \times k$ matrix of indeterminates $X_{\alpha\beta}$ is a polynomial ring with k generators: $\mathbb{C}^G[X] = C[y_1, \ldots, y_k]$. For the y_k we can take either $y_m(X) = \mathrm{tr}(X^m)$ or else $y_m(X) = c_m(X)$ where the c_m are defined by $\det(tI + X) = \sum_{m=0}^{k} c_m(X)t^k$.

Now let E be a complex smooth vector bundle over M (i.e. a real vector bundle with structure group $G = GL(k, \mathbb{C}) - GL(2k, \mathbb{R})$), and let ∇ be an admissible connection for E. Let $s^1, \ldots, s^k$ be an admissible (complex) basis of smooth sections of E over θ, with connection forms and curvature forms $W_{\alpha\beta}$ and $\Omega_{\alpha\beta}$. Now the $\Omega_{\alpha\beta}$ are two-forms (complex valued) in θ hence belong to the commutative ring of even dimensional differential forms in θ and if $P \in C[X]$ then $P(\Omega_{\alpha\beta})$ is a well defined complex valued differential form in θ. If $\tilde{\Omega}_{\alpha\beta}$ are the curvature forms w.r.t. basis $\tilde{s}^\beta = g_{\alpha\beta}s^\alpha$ for $E|\theta$ (where $g : \theta \to GL(k, \mathbb{C})$ is a smooth map) then $\tilde{\Omega} = g^{-1}\Omega g$ so $P(\tilde{\Omega})_x = P(\tilde{\Omega}_x) = P(g^{-1}(x)\Omega_x g(x)) = p^{g(x)}(\Omega_x)$. Thus if $P \in \mathbb{C}^G[X]$, so $P^g = P$ for all $g \in GL(k, \mathbb{C})$ the $P(\tilde{\Omega})_x = P(\Omega_x) = P(\Omega)_x$ for all $x \in \theta$, and hence $P(\Omega)$ is a well defined complex valued form in θ, depending only on ∇ and not in the choice of $s^1, \ldots, s^k$. It follows of course that $P(\Omega)$ is globally defined in all of M.

Chern-Weil Homomorphism Theorem: ($G = GL(k, \mathbb{C})$). Given a complex vector bundle E over M and a compatible connection ∇, for any $P \in \mathbb{C}^G[X]$ define the complex-valued differential form $P(\Omega^\nabla)$ on M as above. Then

1) $P(\Omega^\nabla)$ is a closed form on M.

2) If $\tilde{\nabla}$ is any other connection on M then $P(\Omega^{\tilde{\nabla}})$ differs from $P(\Omega^\nabla)$ by an exact form; hence $P(\Omega^\nabla)$ defines an element $P(E)$ of the de Rham cohomology $H^*(M)$ of M, depending only on the bundle E and not on ∇.

3) For each fixed $P \in \mathbb{C}^G[X]$ the map $E \to P(E)$ thus defined is a characteristic class; that is an element of $\mathrm{Char}(GL(k, \mathbb{C}))$. [In fact if $f : N \to M$ is a smooth map and ∇^f is the connection on f^*E pulled back from ∇ on E, then since

$\Omega^{\nabla^f} = f^*\Omega^{\nabla}$, $P(\Omega^{\nabla^f}) = P(f^*\Omega^{\nabla}) = f^*P(\Omega^{\nabla})$.]

4) The map $\mathbb{C}^G[X] \to \mathrm{Char}(GL(k, \mathbb{C}))$ (which associates to P this map $E \to P(E)$) is a ring homomorphism (the Chern-Weil homomorphism).

5) In fact it is a ring isomorphism.

Proof. Parts 3) and 4) are completely trivial and it is also evident that the Chern-Weil homomorphism is injective. That it is surjective we shall not try to prove here since it requires knowing about getting the de Rham cohomology of a homogeneous space from invariant forms, which would take us too far afield. Thus it remains to prove parts 1) and 2). We can assume that P is homogeneous of degree m. Let $\tilde{P}$ be the "polarization" or m^{th} differential of P, i.e. the symmetric m-linear form on matrices such that

$$P(X) = \tilde{P}(X, \ldots, X)$$

[Explicitly if $y_1, \ldots, y_m$ are $k \times k$ matrices then $\tilde{P}(y_1, \ldots, y_m)$ equals

$$\left.\frac{\partial^m}{\partial t_1 \cdots \partial t_m}\right|_{t=0} P(t_1 y_1 + \cdots + t_m y_m)$$

]. If we differentiate with respect to t at $t = 0$ the identity

$$\tilde{P}(\mathrm{ad}(\exp ty)X, \ldots, \mathrm{ad}(\exp ty)X) \equiv P(X)$$

we get the identity

$$\sum \tilde{P}(X, \ldots, yX - Xy, X, \ldots, X) = 0$$

If we substitute in this from the exterior algebra of complex forms in θ, $y = w_{\alpha\beta}$ and $X = \Omega_{\alpha\beta}$

$$\sum \tilde{P}(\Omega, \ldots, w \wedge \Omega - \Omega \wedge w, \Omega, \ldots, \Omega) = 0.$$

On the other hand from the multi-linearity of $\tilde{P}$ and $P(\Omega) = \tilde{P}(\Omega, \ldots, \Omega)$ we get $d(P(\Omega)) = \sum \tilde{P}(\Omega, \ldots, d\Omega, \Omega, \ldots, \Omega)$. Adding these equations and recalling the

Bianchi identity $d\Omega + w\wedge\Omega - \Omega\wedge w = 0$ gives $d(P(\Omega)) = 0$ as desired, proving 1). We can derive 2) easily from 1) by a clever trick of Milnor's. Given two connections ∇^0 and ∇^1 on E use $M \times I \to M$ to pull them up to connection $\tilde{\nabla}^0$ and $\tilde{\nabla}^1$ on $E \times I$ and let $\tilde{\nabla} = \phi\tilde{\nabla}^1 + (1-\phi)\tilde{\nabla}^0$ where $\phi :\ M \times I \to \mathbb{R}$ is the projection $M \times I \to I \subseteq \mathbb{R}$. If ε_i is the inclusion of M into $M \times I$, $x \to (x, i)$ for $i = 0, 1$ then $\varepsilon_i^*(E \times I) = E$ and ε_i pulls back $\tilde{\nabla}$ to ∇^i, so as explained in the statement of 3)

$$P(\Omega^{\nabla^i}) = \varepsilon_i^* P(\Omega^{\tilde{\nabla}}).$$

Thus $P(\Omega^{\nabla^1})$ and $P(\Omega^{\nabla^2})$ are pull backs of the <u>same</u> closed (by(1)) form $P(\Omega^{\tilde{\nabla}})$ on $M \times I$ under two <u>different</u> maps $\varepsilon_i^* :\ H^*(M \times I) \to H^*(M)$. Since ε_0 and ε_1 are clearly homotopic maps of M into $M \times I$, $\varepsilon_0^* = \varepsilon_1^*$ which means that $\varepsilon_0^*(\Omega^{\tilde{\nabla}})$ and $\varepsilon_1^*(\Omega^{\tilde{\nabla}})$ are cohomologous in M.

The characteristic class $(\frac{i}{2\pi})^m C_m(E)$ is called the m^{th} Chern class of E. Since the C_m are polynomial generators for $\mathbb{C}^G[X]$ $m = 1, 2, \ldots, k$ it follows that any characteristic class is uniquely a polynomial in the Chern classes.

For a real vector bundle we can proceed just as above replacing $\mathbb{C}$ by $\mathbb{R}$ and defining invariant polynomials $E_m(X)$ by $\det(tI + \frac{1}{2\pi}X) = \sum_m E_m(X)t^{k-m}$ and the Pontryagin classes $P_m(E)$ are defined by $p_m(E) = E_{2m}(\Omega)$ where Ω is the curvature of a connection for E. These are related to the Chern classes of the complexified bundle $E_{\mathbb{C}} = E \otimes \mathbb{C}$ by

$$p_m(E) = (-1)^m C_{2m}(E_{\mathbb{C}})$$

PRINCIPAL BUNDLES

Let G be a compact Lie group and let $\mathcal{G}$ denote its Lie algebra, identified with TG_e, the tangent space at e. If $X \in \mathcal{G}$ then $\exp(tX)$ denotes the unique one parameter subgroup of G tangent to X at e. The automorphism $\mathrm{ad}(g)$ of G $(x \to g^{-1}xg)$ has as its differential at e the automorphism $\mathrm{Ad}(g) :\ \mathcal{G} \to \mathcal{G}$ of the

Lie algebra $\mathcal{G}$. Clearly $\mathrm{ad}(g)(\exp tX) = \exp(t\mathrm{Ad}(g)X)$.

Let P be a smooth manifold on which G acts as a group of diffeomorphisms. Then for each $p \in P$ we have a linear map I_p of $\mathcal{G}$ into TP_p defined by letting $I_p(X)$ be the tangent to $(\exp tX)_p$ at $t = 0$. Thus $I(X)$, $(p \mapsto I_p(X))$, is the smooth vector field on P generating the one-parameter group $\exp tX$ of diffeomorphism of P and in particular $\underline{I_p(X) = 0 \text{ if and only if } \exp tX \text{ fixes } P}$.

Exercise:

a) $\mathrm{im}(I_p) =$ tangent space to the orbit GP at p.

b) $I_{gp}(X) = \mathrm{Dg}(I_p(\mathrm{Ad}(g)X))$

Let γ be a Riemannian metric for P. We call γ G-invariant if $g^*(\gamma) = \gamma$ for all $g \in \mathcal{G}$, i.e. if G is included in the group of isometries of γ. We can define a new metric $\bar{\gamma}$ in P (called then result of averaging γ over G) by $\bar{\gamma} = \int_G g^*(\gamma) d\mu(g)$ where μ is normalized Haar measure in G.

Exercises: Show that $\bar{\gamma}$ is an invariant metric for P. Hence invariant metrics always exists.

Definition. P is a $\underline{G\text{-principal bundle}}$ if G acts freely on P, i.e. for each $p \in P$ the isotropy group $G_p = \{y \in G \mid gp = p\}$ is the identity subgroup of G. We define the base space M of P to be the orbit space P/G with the quotient topology. We write $\pi : P \to M$ for the quotient map, so the topology of M is characterized by the condition that π is continuous and open.

Definition. Let θ be open in M. A local section for P over θ is a smooth submanifold $\sum$ of M such that $\sum$ is transversal to orbits and $\sum$ intersects orbit in θ in exactly one point.

Exercise. Given $p \in P$ there is a local section $\sum$ for P containing p. (Hint: With respect to an invariant metric exponentiate an ε-ball in the space of tangent vector to P at p normal to the orbit Gp).

Theorem. There is a unique differentiable structure on M characterized by either of the following:

a) The projection $\pi : P \to M$ is a smooth submersion.

b) If $\sum$ is a section for P over θ then $\pi | \sum$ is a diffeomorphism of $\sum$ onto θ.

Proof. Uniqueness even locally of a differentiable structure for M satisfying a) is easy. And b) gives existence.

Since $\pi : P \to M$ is a submersion, $\mathrm{Ker}(D\pi)$ is a smooth sub-bundle of TP called the vertical sub bundle. Its fiber at p is denoted by V_p.

Exercises:

a) V_p is the tangent space to the orbit of G thru p.

b) If Σ is a local section for P and $p \in \Sigma$ then $\mathrm{T}\Sigma_p$ is a linear complement to V_p in TP_p.

c) For $p \in P$ the map $I_p : \mathcal{G} \to \mathrm{TP}_p$ is an isomorphism of $\mathcal{G}$ onto V_p.

Definition. For each $p \in P$ define $\tilde{w}_p : V_p \to \mathcal{G}$ to be I_p^{-1}.

Exercise. Show that $\tilde{w}$ is a smooth $\mathcal{G}$ valued form on the vertical bundle V.

Definition. For each $g \in \mathcal{G}$ we define a $\mathcal{G}$ valued form $g^*\tilde{w}$ on V by $(g^*\tilde{w})_p = \tilde{w}_{g^{-1}p} \circ D_{g^{-1}}$ (Note $D_{g^{-1}}$ maps V_p onto $V_{g^{-1}p}$!).

Exercise.

$$g^*\tilde{w} = \mathrm{Ad}(g) \circ \tilde{w}$$

Definition. A connection-form on P is a smooth $\mathcal{G}$ valued one-form w on P satisfying

$$g^* w = \mathrm{Ad}(g) w$$

and agreeing with $\tilde{w}$ on the vertical sub bundle.

Definition. A connection on P is a smooth G-invariant sub bundle H of TP complementary to V.

Remark. H is called the horizontal subbundle. The two conditions mean that $H_{gp} = D_g(H_p)$ and $\mathrm{TP}_p = H_p \oplus V_p$ at all $p \in P$. We will write $\hat{H}_p$ and $\hat{V}_p$ to denote the projection operator of TP_p on the subspaces H_p and V_p w.r.t. this direct sum decomposition. Clearly:

$$\hat{H}_{gp} \circ D_g = D_g \circ \hat{H}_p$$
$$\hat{V}_{gp} \circ D_g = D_g \circ \hat{V}_p$$

Theorem. If H is a connection for P then for each $p \in P$ $D\pi_p : \mathrm{TP}_p \to \mathrm{TM}_{\pi(p)}$ restricts to a linear isomorphism $h_p : H_p \simeq \mathrm{TM}_{\pi(p)}$. Moreover if $g \in \mathcal{G}$ then

$$h_{gp} \circ Dg_p = h_p$$

(where $D_g : \mathrm{TP}_p \to \mathrm{TP}_{gp}$).

Proof. Since $\pi : P \to M$ is a submersion and H_p is complementary to $\mathrm{Ker}(D\pi_p)$ it is clear that h_p maps H_p isomorphically onto $\mathrm{TM}_{\pi(p)}$. the rest is an easy exercise.

Definition. If H is a connection for P we define the connection one form w^H corresponding to H by $w^H = \tilde{w} \circ \hat{V}$, i.e. $w_p^H : \mathrm{TP}_p \to \mathcal{G}$ is the composition of the vertical projection $\mathrm{TP}_p \to V_p$ and $\tilde{w}_p : V_p \to \mathcal{G}$.

Exercise. Check that w^H really is a connection form on P at that H can be recovered from w^H by $H_p = Ker(W_p^H)$.

Theorem. The map $H \to w^H$ is a bijective correspondence between all connections on P and all connection forms on P.

Proof. Exercise.

Remark. Note that if we identify the vertical space V_p at each point with $\mathcal{G}$ (via the isomorphism $\tilde{w}_p$) then the connection form w of a connection H is just:

$$w_p = \text{projection of } \mathrm{TP}_p \text{ on } V_p \text{ along } H_p$$

This is the good geometric way to think of a connection and its associated connection form. The basic geometric object is the G invariant sub bundle H of TP complementary to V, and w is what we use to explicitly describe H for calculational purpose.

Henceforth we will regard a section s of P over θ as a smooth map $s : \theta \to P$ such that $\pi(s(x)) = x$, (i.e. for each orbit $x \in \theta$ $s(x)$ is an element of x). If $\tilde{s}$ is a second section over θ then there is a unique map $g : \theta \to G$ such that $\tilde{s} = gs$ (i.e. $\tilde{s}(x) = g(x)s(x)$ for all $x \in \theta$). We call g the transition function between the sections $\tilde{s}$ and s.

For each section $s : \theta \to P$ and connection form w on P define a $\mathcal{G}$ valued one-form w^s on θ by $w^s = s^*(w)$ (i.e. $w^s(X) = w(\mathrm{Ds}(X))$ for $X \in \mathrm{TM}_p$, $p \in \theta$).

Exercise. If s and $\tilde{s}$ are two section for P over θ, $\tilde{s} = gs$, and if w is a connection 1-form in P then show that

$$w^{\tilde{s}} = \mathrm{Ad}(g) \circ w^s + g^{-1} Dg$$

or more explicitly:

$$w_x^{\tilde{s}} = \mathrm{Ad}(g(x)) w_x^s + g(x)^{-1} Dg_x$$

(the term $g(X)^{-1}Dg_x$ means the following: since $g : \theta \to G$, Dg_x maps TM_p into $TG_{g(x)}$; then $g(x)^{-1}Dg_x(x)$ is the vector in $\mathcal{G} = TG_e$ obtained by left translation $Dg_x(x)$ to e, i.e. $g(x)^{-1}$ really means $D\lambda$ where $\lambda : G \to G$ is $\gamma \to g(x)^{-1}\gamma$).

Exercise. Conversely show that if for each section s of P over θ we have a $\mathcal{G}$ valued one-form w^s in θ and if there satisfy the above transformation law then there is unique connection form w on P such that $w^s = s^*w$.

Definition. If w is a connection form on the principal bundle P we define its curvature Ω to be the $\mathcal{G}$ valued two form on P, $\Omega = dw + w \wedge w$.

Exercise. If H is the connection corresponding to w (i.e. $H = \operatorname{Ker} w$) and $\hat{H}_p =$ projection of TP_p on H_p along V_p show that $\Omega = Dw$, where by definition $Dw(u, v) = dw(\hat{H}u, \hat{H}v)$.

Remark. This shows $\Omega(X, Y) = 0$ if either X or Y is vertical. Thus we may think of Ω_p as a two form on $\mathrm{TM}_{\pi(p)}$ with values in $\mathcal{G}$.

THE PRINCIPAL FAME BUNDLE OF A VECTOR BUNDLE

Now suppose E is a G-vector bundle over M. We will show how to construct a principal G-bundle $P(E)$ with orbit space (canonically diffeomorphic to) M, called the principal frame bundle of E, such that connections in E and connections in $P(E)$ are "really the same" thing — i.e. correspond naturally.

The fiber of $P(E)$ at $x \in M$ is the set $P(E)_x$ of all admissible frames for E_x and so $P(E)$ is just the union of these fibers, and the projection $\pi : P(E) \to M$ maps $P(E)_x$ to x. If $e = (e^1, \ldots, e^k)$ is in $P(E)_x$ and $g = g_{\alpha\beta} \in G\_GL(k)$ then $ge = \tilde{e} = (\tilde{e}^1, \ldots, \tilde{e}^k)$ where $\tilde{e}^\beta = g_{\alpha\beta}e_\alpha$, so clearly G acts freely on $P(E)$ with the fibers $P(E)_x$ as orbits. If $s = (s^1, \ldots, s^k)$ is a local base of section of E over θ then we get a bijection map $\theta \times G \simeq P(E)|\theta = \pi^{-1}(\theta)$ by $(x, g) \to gs(x)$. We make $P(E)$ into a smooth manifold by requiring these to be diffeomorphisms.

Exercise. Check that the action of G on $P(E)$ is smooth and that the smooth local sections of $P(E)$ are just the admissible local bases $s = (s^1, \ldots, s^k)$ of E.

Theorem. Given a admissible connection ∇ for E the collection w^s of local connection forms defined by

$$\nabla s^\beta = w_{\alpha\beta} \otimes s^\alpha$$

defined a unique connection form w in $P(E)$ such that $w^s = s^* w$, and hence a unique connection H in $P(E)$ such that $w = \tilde{w} \circ \hat{V}$. This map $\nabla \to H$ is in fact a bijective correspondence between connections in E and connections in $P(E)$.

Proof. Exercise.

INVARIANT METRICS IN PRINCIPAL BUNDLES

Let $\pi : P \to M$ be a principal G bundle, G a compact group and let α be an adjoint invariant inner product on $\mathcal{G}$ the Lie algebra of G. (We can always get such an α by averaging over the group; if G is simple α must be a constant multiple of the killing form). By using the isomorphism $\tilde{w}_p : V_p \simeq \mathcal{G}$ we get a G-invariant Riemannian metric we shall also call α on the vertical bundle V such that each of the maps $\tilde{w}_p$ is isometric.

By an invariant metric for P we shall mean a Riemannian metric g for P which is invariant under the action of G and restricts to α on V.

Theorem. Let $\tilde{g}$ be an invariant metric for P and let H be the sub bundle of TP orthogonally complementing V with respect to $\tilde{g}$. Then H is a connection for P and moreover there is a unique metric g for M such that for each $p \in P$, $D\pi$ maps H_p isometrically onto $\mathrm{TM}_{\pi(p)}$. This map $\tilde{g} \to (H, g)$ is a bijective correspondence between all invariant metric in P and pairs (H, g) consisting of a connection for P and metric for M. If w^H is the connection form for H then we can recover $\tilde{g}$ from H and g by

$$\tilde{g} = \pi^* g + \alpha \circ w^H.$$

Proof. Trivial.

Corollary: If g is a fixed metric on M there is a bijective correspondence between connections H for P and invariant metrics $\tilde{g}$ for P such that $\pi | H_p : H_p \simeq \mathrm{TM}_{\pi(p)}$ is an isometry for all $p \in P$ (namely $\tilde{g} = \pi^* g + \alpha \circ w^H$).

It turns out, not surprisingly perhaps, that there are some remarkable relations between the geometry of (M, g) and that of $(P\, \tilde{g})$, involving of course the connection. Moreover these relationships are at the very heart of the Kaluza-Klein unification of gravitation and Yang-Mills fields. We will study them with some care now.

MATHEMATICAL BACKGROUND OF KALUZA-KLEIN THEORIES

In this section M will as usual denote an n-dimensional Riemannian (or pseudo-Riemannian) manifold. We let $v_1, \ldots, v_n$ be a local o.n. frame field in M and $\theta_1, \ldots, \theta_n$ the dual coframe (In general indices i, j, k have the range 1 to n).

G will denote a p-dimensional compact Lie group with Lie algebra $\mathcal{G} = TG_e$ having an adjoint invariant inner product α. We let $e_{n+1}, \ldots, e_{n+p}$ denote an o.n. basis for $\mathcal{G}$ and λ_α the dual basis for $\mathcal{G}^*$. (In general indices α, β, γ have the range $n+1$ to $n+p$). We let $C^\gamma_{\alpha\beta}$ be the structure constants for $\mathcal{G}$ relative to the e_α

$$[e_\alpha, e_\beta] = C^\gamma_{\alpha\beta} e_\gamma.$$

Of course $C^\gamma_{\alpha\beta} = -C^\gamma_{\beta\alpha}$. Now for α fixed $C^\gamma_{\alpha\beta}$ is the matrix of the skew adjoint operator $\mathrm{Ad}(e_\alpha) : \mathcal{G} \to \mathcal{G}$ w.r.t. the o.n. basis e_α, hence also $C^\gamma_{\alpha\beta} = -C^\beta_{\alpha\gamma}$.

We let P be a principal G-bundle over M with connection H having connection form w. We recall that we have a canonical invariant metric on P such that $D\pi$

maps H_p isometrically onto $M_{\pi(p)}$. We shall write θ_i also for the forms $\pi^*(\theta_i)$ in P and we write $\theta_\alpha = \lambda_\alpha \circ w$. Then letting the indices A, B, C have the range 1 to $n+p$ it is clear θ_A is an o.n. coframe field in P. We let v_A be the dual frame field. [Clearly the v_α are vertical and agree with the $e_\alpha \in \mathcal{G}$ under the natural identification, and the v_i are horizontal and project onto the v_i in M].

In what follows we shall use these frames fields to first compute the Levi-Civita connection in P. With this in hand we can compute the Riemannian curvature of P in terms of the Riemannian curvature of M, the Riemannian curvature of G (in the bi-invariant metric defined by α) and the curvature of the connection on P. Also we shall get a formula for the "acceleration" or curvature of the projection $\bar{\lambda}$ on M of a geodesic λ in P. The calculations are complicated, but routine so we will give the main steps and leave details to the reader as exercises. First however we list the main important results.

Notation: Let Ω be the curvature two-form of the connection form w (a $\mathcal{G}$ valued two form on P) and define real valued two-forms Ω_α by $\Omega = \Omega_\alpha e_\alpha$ and functions F^α_{ij} (skew in i, j) by $\Omega_\alpha = \frac{1}{2} F^\alpha_{ij} \theta_i \wedge \theta_j$. we let θ_{AB} denote the connection forms of the Levi-Civita connection for P, relative to the frame field v_A, i.e. $\theta_{AB} = -\theta_{BA}$ and

$$\nabla v_B = \theta_{AB} v_A$$

(or equivalently $d\theta_B = \theta_{AB} \wedge \theta_A$). Finally we let $\bar{\theta}_{ij}$ denote the Levi-Civita connection forms on M relative to θ_i, and we also write $\bar{\theta}_{ij}$ for $\pi^* \bar{\theta}_{ij}$, these same forms pulled up to P.

Theorem 1:

$$\begin{cases} \theta_{ij} &= \bar{\theta}_{ij} - \frac{1}{2} F^\alpha_{ij} \theta_\alpha \\ \theta_{\alpha i} &= \frac{1}{2} F^\alpha_{ij} \theta_j \\ \theta_{\alpha\beta} &= -\frac{1}{2} C^\alpha_{\gamma\beta} \theta_\gamma \end{cases}$$

Notation. Let $\gamma : I \to P$ be a geodesic in P and let $\bar{\gamma} = \pi \circ \gamma : I \to M$ be its projection in M. We define a function $q : I \to \mathcal{G}$ called the specific charge of γ by $q(t) = w(\gamma'(t))$ (i.e. $q(t)$ is the vertical component of the velocity of γ).

We define a linear functional $\check{f}(\gamma(t))$ on $\mathrm{TM}_{\gamma(t)}$ by $\check{f}(\gamma(t))(w) = \Omega_{\gamma(t)}(w,\, \bar{\gamma}'(t)) \cdot q$ (where $\cdot$ means inner product in $\mathcal{G}$ and we recall Ω_p can be viewed as a two form on $\mathrm{TM}_{\pi(p)}$) $\check{f}$ is called the Lorentz co-force and the dual element $\check{f}(\gamma(t)) \in \mathrm{TM}_{\gamma(t)}$ is called the Lorentz force.

Remark: If we write:

$$\gamma'(t) = u_i v_i + q_\alpha v_\alpha$$

then $q = q_\alpha v_\alpha$ and $\bar{\gamma}'(t) = u_i v_i$. Recalling $\Omega = \frac{1}{2} F^\alpha_{ij} \theta_i \wedge \theta_j e_\alpha$, we have $\Omega(w,\, \bar{\gamma}'(t)) = \frac{1}{2} F^\alpha_{ij} w_i u_j e_\alpha$ so $\Omega(w,\, \bar{\gamma}'(t)) \cdot q = \frac{1}{2} q_\alpha F^\alpha_{ij} w_i u_j$ thus we see that the Lorentz force is given by $\check{f}(\gamma'(t)) = \frac{1}{2} q_\alpha F^\alpha_{ij} u_j v_i$.

Theorem 2. The specific charge q is a constant. Moreover the "acceleration" $\frac{D\bar{\gamma}'}{dt}$ of the projection $\bar{\gamma}$ of γ is just the Lorentz force.

Remark. Note that in the notation of the above remark this says that the q_α are constant and

$$\frac{D}{dt}(\frac{du_i}{dt}) = \frac{1}{2} q_\alpha F^\alpha_{ij} u_j.$$

Before stating the final result we shall need, we recall some terminology and notation concerning curvature in Riemannian manifolds.

The curvature ${}^M\Omega$ of (the Levi-Civita connection for) M is a two form on M with values in the bundle of linear maps of TM to TM; so if $x,\, y \in \mathrm{TM}_p$ then ${}^M\Omega(x,\, y) : \mathrm{TM}_p \to \mathrm{TM}_p$ is a linear map. The Riemannian tensor of M, ${}^M\mathrm{Riem}$, is the section of $\otimes^4 T^*M$ given by

$${}^M\mathrm{Riem}(x,\, y,\, z,\, w) = < {}^M\Omega(x,\, z)y,\, w > .$$

It's component with respect to a frame v_i are denoted by

$${}^M R_{ijk\ell} = {}^M\mathrm{Riem}(v_i,\, v_j,\, v_k,\, v_\ell)$$

The Ricci tensor ${}^M\mathrm{Ric}$ is the symmetric bilinear form on M given by

$${}^M\mathrm{Ric}(x,\, y) = \mathrm{trace}(z \to {}^M\Omega(x,\, z)y)$$

so its components ${}^M R_{ij}$ are given by

$$ {}^M R_{ij} = \sum_k {}^M R_{ikjk}. $$

Finally, the scalar curvature ${}^M R$ of M is the scalar function

$$ {}^M R = \operatorname{trace} {}^M \mathrm{Ric} = \sum_{i,k} {}^M R_{ikik}. $$

Of course P also has a scalar curvature function ${}^P R$ which by the invariance of the metric is constant on orbits of G so is well defined smooth function on M. Also the adjoint invariant metric α on $\mathcal{G} = TG_e$ defines by translation a bi-invariant metric on G, so G has a scalar curvature ${}^G R$ which of course is a constant. Finally since the curvature Ω of the connection form w is a two form on M with values in $\mathcal{G}$ (and both TM_p and $\mathcal{G}$ have inner products), Ω has a well defined length $\|\Omega\|$ which is a scalar function on M

$$ \|\Omega\|^2 = \sum_{\alpha, i, j} (F^\alpha_{ij})^2. $$

Theorem 3.

$$ {}^P R = {}^M R - \frac{1}{2}\|\Omega\|^2 + {}^G R. $$

Proof of Theorem 1:

1) $\theta_{AB} + \theta_{BA} = 0$

2) $d\theta_A = \theta_{BA} \wedge \theta_B$

are the structure equation. Lifting the structural equation $d\theta_i = \bar{\theta}_{ji} \wedge \theta_j$ on M to P and comparing with the corresponding structural equation on P gives

3) $d\theta_i = \bar{\theta}_{ji} \wedge \theta_j = \theta_{ji} \wedge \theta_j + \theta_{\alpha i} \wedge \theta_\alpha$ on the other hand $w = \theta_\alpha e_\alpha$, and so $d\theta_\alpha e_\alpha = dw = \Omega - [w,\, w] = \frac{1}{2} F^\alpha_{ij} \theta_i \wedge \theta_j e_\alpha - \frac{1}{2}\theta_\beta \wedge \theta_\gamma [e_\beta,\, e_\gamma] = (\frac{1}{2} F^\alpha_{ij}\theta_i \wedge \theta_j - \frac{1}{2} C^\alpha_{\beta\gamma}\theta_\beta \wedge \theta_\gamma) e_\alpha$

which with the structural equation $d\theta_\alpha = \theta_{i\alpha} \wedge \theta_i + \theta_{\beta\alpha} \wedge \theta_\beta$ gives

4) $\theta_{i\alpha} \wedge \theta_i + \theta_{\beta\alpha} \wedge \theta_\alpha = \frac{1}{2} F^\alpha_{ij}\theta_i \wedge \theta_j - \frac{1}{2} C^\alpha_{\beta\gamma}\theta_\beta \wedge \theta_\gamma$

We can solve 1), 2), 3), 4) by comparing coefficient; making use of $C^{\alpha}_{\beta\gamma} = -C^{\gamma}_{\beta\alpha} = C^{\gamma}_{\alpha\beta} = -C^{\beta}_{\alpha\gamma}$. We see easily that the values given in Theorem 1 for θ_{AB} solve these equation, and by Cartan's lemma these are the unique solution.

Proof of Theorem 3:

Recall that ${}^PR$ is by definition ${}^PR_{ABAB}$ where ${}^PR_{ABCD}$ is defined by $d\theta_{AB} + \theta_{AF} \wedge \theta_{FB} = {}^P\Omega_{AB} = \frac{1}{2}{}^PR_{ABCD}\theta_C \wedge \theta_D$. Thus it is clearly only a matter of straight forward computation, given Theorem 1, to compute the ${}^PR_{ABCD}$ in terms of the ${}^MR_{ijk\ell}$, the F^{α}_{ij}, and the $C^{\alpha}_{\beta\gamma}$. The trick is to recognize certain terms in the sum ${}^PR = {}^PR_{ABAB}$ as $\|\Omega\|^2 = F^{\alpha}_{ij}F^{\alpha}_{ij}$ and ${}^GR$. The first is easy; for the second, the formula ${}^GR = {}^GR_{\alpha\beta\alpha\beta} = \frac{1}{4}C^{\gamma}_{\alpha\beta}C^{\gamma}_{\alpha\beta}$ follows easily from section 21 of Milnor's "Morse Theory". We leave the details as an exercise, with the hint that since we only need component of ${}^PR_{ABCD}$ with $C = A$ and $D = B$ some effort can be saved.

Proof of Theorem 2.

To prove that for a geodesic $\gamma(t)$ in P the specific charge $q(t) = w(\gamma'(t))$ is a constant it will suffice (since $w(x) = \sum_{\alpha} < x, e_\alpha > e_\alpha$) to show that the inner product $< \gamma'(t), e_\alpha >$ is constant. Now the $e_{\alpha'}$ considered as vector fields on P generate the one parameter groups $\exp(te_\alpha) \in G$ of isometries of P, hence they are Killing vector fields. Thus the constancy of $< \gamma'(t), v_\alpha >$ is a special case of the following fact (itself a special case of the E. Noether Conservation Law Theorem).

Proposition. If X is a Killing vector field on a Riemannian manifold N and σ is a geodesic of N then the inner product of X with $\sigma'(t)$ is independent of t.

Proof. If $\lambda_s : I \to N$ $|s| < \varepsilon$ is a smooth family of curves in N recall that the first variation formula says

$$\begin{aligned} & \left.\frac{d}{ds}\right|_{s=0} \int_a \frac{1}{2}\|\lambda'(t)\|^2 dt \\ = & \int_a^b < \frac{D}{dt}\lambda_0'(t),\, \eta(t) > dt + < \lambda_0'(t),\, \eta(t) >]_a^b \end{aligned}$$

where $\eta(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \lambda_s(t)$ is the variation vector field along λ_0. In particular if $\lambda_0 = \sigma$ so that $\frac{D}{dt}(\lambda') = 0$, then $\frac{d}{ds}\Big|_{s=0} \int_a^b \|\lambda_s(t)\|^2 dt =< \sigma'(t),\, \eta(t) >\Big|_a^b$. Now let ϕ_s be the one parameter group of isometries of N generated by X and put $\lambda_s(t) = \phi_s(\sigma(t))$. Since the ϕ_s are isometries it is clear that $\|\lambda_s'(t)\|^2 = \|\sigma'(t)\|^2$ hence $\left.\frac{d}{ds}\right|_{s=0} \int_a^b \|\lambda_s'(t)\|^2 = 0$. On the other hand since X generates ϕ_s

$$\left.\frac{d}{ds}\right|_{s=0} \lambda_s(t) = X_{\sigma(t)}$$

so our first variation formula says $< \sigma'(a),\, X_{\sigma(a)} >=< \sigma'(b),\, X_{\sigma(b)} >$. Since $[a,\, b]$ can be any subinterval of the domain of σ this says $< \sigma'(t),\, X_{\sigma(t)} >$ is constant. $\square$

Now let $\gamma'(t) = u_A(t)v_A(\gamma(t))$ so $\bar{\gamma}'(t) = u_i(t)v_i(\bar{\gamma}(t))$. From the definition of covariant derivative on P and M we have

$$\begin{aligned} u_{AB}\theta_B &= du_A - u_B\theta_{BA} \\ \bar{u}_{ij}\bar{\theta}_j &= du_i - u_j\bar{\theta}_{ji} \end{aligned}$$

comparing coefficient when $A = i$ gives

$$u_{ij} = \bar{u}_{ij} - \frac{1}{2}u_\alpha f_{ij}^\alpha.$$

Now the definition of $\frac{D\bar{\gamma}'}{dt}$ is $\bar{u}_{ij}v_i\theta_j(\bar{\gamma}')$. We leave the rest of the computation as an exercise.

GENERAL RELATIVITY

We use atomic clocks to measure time and radar to measure distance: the distance from P to Q is half the time for a light signal from P to reflect at Q

and return to P, so that automatically the speed of light is 1. We thus have coordinates $(t,\, x,\, y,\, z) = (x_0,\, x_1,\, x_2,\, x_3)$ in the space-time of event $\mathbb{R}^4$. It turn out that the metric $d^2 = dt^2 - dx^2 - dy^2 - dz^2$ has an invariant physical meaning: the length of a curve is the time interval that would be measured by a clock travelling along that curve. (Note that a moving particle is described by its "world line $(t,\, x(t),\, y(t),\, z(t))$.

According to Newton a gravitational field is described by a scalar "potential" ϕ on $\mathbb{R}^3$. If a particle under no other force moves in this field it will satisfy:

$$\frac{d^2 x_i}{dt^2} = \frac{\partial \phi}{\partial x_i} \quad i = 1,\, 2,\, 3$$

Along our particle world line $(t,\, x_1(t),\, x_2(t),\, x_3(t))$ we have

$$(\frac{d\tau}{dt})^2 = 1 - v^2 \quad v^2 = \sum_{i=1}^{3} (\frac{dx_i}{dt})^2$$

so if $v \ll 1$ (i.e. velocity is a small fraction of the speed of light) $d\tau \sim dt$ and to good approximation

$$\frac{d^2 x_i}{d\tau^2} = \frac{\partial \phi}{\partial x_i} \quad i = 1,\, 2,\, 3.$$

Can we find a metric $d\tau^2$ approximating the above flat one so that its geodesics

$$\frac{d^2 x_\alpha}{d\tau^2} = \Gamma^\alpha_{\beta\gamma} \frac{dx_\beta}{d\tau} \frac{dx_\gamma}{d\tau}$$

will be the particle paths in the gravitational field described by ϕ (to a closed approximation)? Consider case of a gravitational field generated by a massive object stationary at the spatial origin (world line $(t,\, 0,\, 0,\, 0)$. In this case $\phi = -\frac{GM}{r}$ $(r = x_1^2 + x_2^2 + x_3^2)$. For the metric $d\tau^2$ it is natural to be invariant under $SO(0)$ and the time translation group $\mathbb{R}$. It is not hard to see any such metric can be put into the form

$$d\tau^2 = e^{2A(r)} dt^2 - e^{2B(r)} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2).$$

Let us try

$$B = 0, \quad e^{2A(r)} = (1 + \alpha(r))$$
$$d\tau^2 = (1 + \alpha(r)) dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2$$

Geodesic equation is

$$
\begin{aligned}
\frac{d^2x_\alpha}{d\tau^2} &= -\Gamma^\alpha_{\beta\gamma}\frac{dx_\beta}{d\tau}\frac{dx_\gamma}{d\tau} \\
\frac{d^2x_i}{dt^2} &= -\Gamma^i_{\infty}\{\frac{dx_0}{d\tau}\sim 1 \ \frac{dx_i}{d\tau}\sim 0\} \\
\Gamma^k_{ij} &= \frac{1}{2}g^{k\ell}(\frac{\partial g_{\ell j}}{\partial x_i}+\frac{\partial g_{\ell i}}{\partial x_i}-\frac{\partial g_{ij}}{\partial x_\ell}) \\
\Gamma^k_{00} &= \frac{1}{2}g^{k\ell}(0+0-\frac{\partial g_{00}}{\partial x_k} \\
&= +\frac{1}{2}\frac{\partial g_{00}}{\partial x_k} \\
&= -\frac{1}{2}\frac{\partial \alpha_0}{\partial x_k} \\
\frac{d^2x_i}{d\tau^2} &= \frac{1}{2}\frac{\partial \alpha_0}{\partial x_i}
\end{aligned}
$$

Comparing with Newton's equations $\alpha = 2\phi = -\frac{2GM}{r}$ $d\tau^2 = (1-\frac{2GM}{r})dt^2 - dx^2 - dy^2 - dz^2$. Will have geodesics which very well approximate particle world lines in the gravitational field with potential $-\frac{GM}{r}$.

Now consider the case of a general gravitational potential ϕ and recall that ϕ is always harmonic — i.e. satisfied the "field equations". $\frac{\partial^2\phi}{\partial x_i\partial x_i} = 0$. Suppose the solutions of Newton's $\frac{d^2x_i}{d\tau^2} = -\frac{\partial\phi}{\partial x_i}$ are geodesics of some metric $d\tau^2$. Let us take a family $x^s_\alpha(t)$ of geodesics with $x^0_\alpha(t) = x_\alpha(t)$ and $\frac{\partial}{\partial s}\Big|_{s=0} x^s_\alpha(t) = \eta_\alpha$. Then η_α will be a Jacobi-field along x_α, i.e.

$$\frac{D}{d\tau}(\frac{dx_\alpha}{d\tau}) = (R^\alpha_{\beta\gamma\delta}\frac{dx_\beta}{d\tau}\frac{dx_\gamma}{d\tau})\eta_\delta$$

On the other hand, taking $\frac{\partial}{\partial s}\Big|_{s=0}$ of

$$\frac{d^2x^s_\alpha(t)}{d\tau^2} = \frac{\partial\phi(x^s(t))}{\partial x_i}$$

give

$$\frac{d^2\eta_i}{d\tau^2} = \frac{\partial^2\phi}{\partial x_i\partial x_j}\eta_j$$

so comparing suggest

$$(R^\alpha_{\beta\gamma\delta}\frac{dx_\beta}{d\tau}\frac{dx_\gamma}{d\tau}) \sim \frac{\partial^2\phi}{\partial x_i\partial x_j}$$

Now $0 = \Delta\phi = \frac{\partial^2\phi}{\partial x_i \partial x_i}$ so we expect $R^{\alpha}_{\beta\gamma\alpha}\frac{dx_\beta}{d\tau}\frac{dx_\gamma}{d\tau} = R_{\beta\gamma}\frac{dx_\beta}{d\tau}\frac{dx_\gamma}{d\tau} = 0$. But $R_{\beta\gamma}$ is symmetric and $\frac{dx_\beta}{d\tau}$ is arbitrary so this implies $R_{\beta\gamma} = 0$. We take these as our (empty space) field equations for the metric tensor $d\tau^2 = g_{\alpha\beta}dx_\alpha dx_\beta$. Actually for reasons we shall see soon these equations are usually written differently. Let $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta}$ where as usual $R = g^{\alpha\beta}R_{\alpha\beta}$ is the scalar curvature. Then $g^{\alpha\beta}G_{\alpha\beta} = R - \frac{1}{2}RS^{\alpha}_{\alpha} = R - \frac{n}{2}R$ so $G_{\alpha\beta} = 0 \Rightarrow R = 0$ (if $n \neq 2$) and hence $R_{\alpha\beta} = 0$ and the converse is clear. Thus our field equations are equivalent to $G_{\mu\nu} = 0$. We shall now see that these are the Euler-Lagrange equation of a very simple and natural variational problem.

Let M be a smooth manifold and let θ be a relatively compact open set in M. Let g be a (pseudo) Riemannian metric δg an arbitrary symmetric two tensor with support in θ at $g_\varepsilon = g + \varepsilon\delta g$ (so for small ε, g_ε is also a Riemannian metric for M. Note $\delta_g = \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} f(g_\varepsilon)$. For any function f of metric we shall write similarly $\delta f = \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} f(g_\varepsilon)$.

$$\begin{aligned}
\mu &= \text{Riemannian measure} = \sqrt{g}\,dx_1\cdots dx_n \\
\text{Riem} &= \text{Riemannian tensor} = R^i_{jk\ell} \\
\text{Ric} &= \text{Ricci tensor} = R^i_{jki} \\
R &= \text{scalar curvature} = g^{ij}R_{ij} \\
G &= \text{Einstein tensor} = \text{Ric} - \frac{1}{2}Rg
\end{aligned}$$

We put ε's on these quantities to denote their values w.r.t. $g + \varepsilon\delta g$. Consider the functional $\int_\theta R\mu$.

Theorem. $\delta\int_\theta R\mu = \int_\theta G_{\mu\nu}\delta g^{\mu\nu}\mu$. Hence the NASC that a metric be extremal for $\delta\int_\theta R\mu$ (for all θ and all compact variations in θ) is that $G = 0$.

Lemma. Given a vector field v in M define $\operatorname{div}(v)$ to be the scalar function given by $d(i_v\mu) = \operatorname{div}(v)\mu$. Then in local coordinates

$$\operatorname{div}(v) = \frac{1}{\sqrt{g}}(\sqrt{g}v^\alpha)_{,\alpha} = v^\alpha_{;\alpha}$$

Proof. $\mu = \sqrt{g} dx_1 \wedge \cdots \wedge dx_n$

so

$$\begin{aligned} i_v \mu &= \sum_{\alpha=1}^{n} (-1)^{i+1} \sqrt{g} v^i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n \\ d(i_v \mu) &= \sum_{\alpha=1}^{n} (\sqrt{g} v^\alpha)_{\prime\alpha} dx_1 \wedge \cdots \wedge dx_n \\ &= \sum_{\alpha=1}^{n} \frac{1}{\sqrt{g}} (\sqrt{g} v^\alpha)_{\prime\alpha} \mu \end{aligned}$$

To see $\frac{1}{\sqrt{g}}(\sqrt{g}v^\alpha)_{\prime\alpha} = v^\alpha_{;\alpha}$ at some point, use geodesic coordinate at that point and recall that in these coordinates $\frac{\partial g_{ij}}{\partial x_k} = 0$ at that point. □

Remark. If v has compact support $\int \operatorname{div}(v)\mu = 0$

$$R^\mu_{\nu\rho\sigma} = \Gamma^\mu_{\nu\sigma,\rho} - \Gamma^\mu_{\nu\rho,\sigma} + \Gamma^\mu_{\tau\rho}\Gamma^\tau_{\nu\sigma} + \Gamma^\mu_{\tau\sigma}\Gamma^\tau_{\nu\rho}$$

Lemma. $\delta\Gamma^\mu_{\nu\rho}$ is a tensor field (section of $T^*M \otimes T^*M \otimes \mathrm{TM}$) and

$$\delta R^\mu_{\nu\rho\sigma} = \delta\Gamma^\mu_{\nu\sigma;\rho} - \delta\Gamma^\mu_{\nu\rho;\sigma}, \text{ hence}$$

$$\delta R_{\nu\rho} = \delta\Gamma^\mu_{\nu\mu;\rho} - \delta\Gamma^\mu_{\nu\rho;\mu}$$

$$g^{\nu\rho}\delta R_{\nu\rho} = (g^{\nu\rho}\delta\Gamma^\mu_{\nu\rho})_{;\rho} - (g^{\nu\rho}\delta\Gamma^\mu_{\nu\rho})_{;\mu}$$

(Palatini identities)

Proof. That $\delta\Gamma^\mu_{\nu\rho} = \frac{\partial}{\partial\varepsilon}\Gamma^\mu_{\nu\rho}(g(\varepsilon))$ is a tensor field is a corollary of the fact that the difference of two connection is. The first identity is then clearly true at a point by choosing geodesic coordinates at that point. □

Corollary. $\int_\theta g^{\mu\nu}\delta R_{\mu\nu}\mu = 0$

Proof. $g^{\mu\nu}\delta R_{\mu\nu}$ is a divergence

Lemma. $\delta\mu = -\frac{1}{2} g_{ij}\delta g^{ij}\mu$

Proof. Since $g_{ij}g^{ij} = n$, $\delta g_{ij} g^{ij} = -g_{ij}\delta g^{ij}$. Also $\frac{\partial g}{\partial g_{ij}} = g\, g^{ij}$ by Cramer's rule, so $\frac{\partial\sqrt{g}}{\partial g_{ij}} = \frac{1}{2}\sqrt{g}g^{ij}$ and so

$$\delta\sqrt{g} = \frac{\partial\sqrt{g}}{\partial g_{ij}}\delta g_{ij} \quad = \quad (\frac{1}{2} g^{ij}\delta g_{ij})\sqrt{g}$$

$$= (-\frac{1}{2}g_{ij}\delta g^{ij})\sqrt{g}$$

since $\mu = \sqrt{g}dx_1 \wedge \cdots \wedge dx_n$, lemma follows. We can now easily prove the theorem:

$$R\mu = g^{ij}R_{ij}\mu$$

so

$$\begin{aligned}
\delta(R\mu) &= \delta g^{ij}R_{ij}\mu + g^{ij}\delta R_{ij}\mu + R\delta\mu \\
&= (R_{ij} - \frac{1}{2}Rg_{ij})\delta g^{ij}\mu + (g^{ij}\delta R_{ij})\mu \\
\delta\int_\theta R\mu &= \int_\theta \delta(R\mu) \\
&= \int_\theta G_{ij}\delta g^{ij}\mu
\end{aligned}$$

This completes the proof of the theorem. □

SCHWARZCHILD SOLUTION

Let's go back to our static, SO(3) symmetric metric

$$\begin{aligned}
d\tau^2 &= e^{2A(r)}dt^2 - e^{2B(r)}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \\
&= w_1^2 + w_2^2 + w_3^2 + w_4^2
\end{aligned}$$

$$\begin{cases} w_i = a_i du_i \quad \text{(no sum)} \\ u_1 = t \quad u_2 = r \quad u_3 = \theta \quad u_4 = \phi \\ a_1 = e^{A(r)} \quad a_2 = ie^{B(r)} \quad a_3 = ir \quad w_y = ir\sin\theta \end{cases}$$

$$w_{ij} = \frac{(a_i)_j}{a_j}du_i - \frac{(a_j)_i}{a_i}du_j$$

$$\begin{aligned}
dw_{ij} + w_{ik}w_{kj} &= \Omega_{ij} \\
&= \frac{1}{2}R_{ijk\ell}w_k w_\ell
\end{aligned}$$

<u>Exercise</u>. Prove the following:

$$w_{12} = -iA'e^{A-B}dt$$

$$\begin{aligned}
w_{13} &= w_{14} = 0 \\
w_{23} &= -e^{-B} d\theta \\
w_{24} &= -\sin\theta e^{-B} d\phi \\
w_{34} &= -\cos\theta d\phi
\end{aligned}$$

$$\begin{aligned}
R_{1212} &= (A'' + A'^2 - A'B')e^{-2B} \\
R_{1313} &= \frac{A'}{r} e^{-2B} \\
R_{2323} &= -\frac{B'}{r} e^{-2B} = R_{2424} \\
R_{3434} &= \frac{e^{-2B} - 1}{r^2}
\end{aligned}$$

$$\begin{aligned}
R &= 2[e^{-2B}(A' + A'^2 - A'B' + 2\frac{A' - B'}{r} + \frac{1}{r^2}) - \frac{1}{r^2}] \\
R\mu &= Rr^2 \sin^2\theta dr d\theta d\phi dt \\
&= ((1 - 2rB')e^{A-B} - e^{A+B}) dr (\sin\theta d\theta d\phi dt) \\
&\quad + (r^2 A' e^{A-B})' dr (\sin\theta d\theta d\phi dt)
\end{aligned}$$

[Note this second term is a divergence, hence it can be ignored in computing the Euler-Lagrange equations. If we take for our region θ over which we vary $\int R\mu$ a rectangular box with respect to these coordinates then the integration w.r.t. θ, ϕ, t gives a constant multiplier and we are left having to extremalize

$$\int_{r_1}^{r_2} L(A', B', A, B, r) dr$$

where $L = (1 - 2rB')e^{A-B} - e^{A+B}$. [One must justify only extremalizing w.r.t. variations of the metric which also have spherical symmetry. On this point see "The principle of Symmetric Criticality", Comm. in Math. Physic, Dec. 1979]. The above is a standard 1-variable Calculus of variations problem which gives Euler-Lagrange equations:

$$0 = \frac{\partial L}{\partial A} - \frac{\partial}{\partial r}(\frac{\partial L}{\partial A'})$$

$$\begin{aligned} &= \frac{\partial L}{\partial B} - \frac{\partial}{\partial r}(\frac{\partial L}{\partial B'}) \\ &= (1-2rB')e^{A-B} - e^{A+B} \\ &= (1+2rA')e^{A-B} - e^{A+B} \end{aligned}$$

so $A' + B' = 0$ or $B = -A + k$, and we can take $k = 0$ (since another choice just rescales t) then we have

$$1 = (1+2rA')e^{2A} = (re^{2A})'$$

so

$$\begin{aligned} re^{2A} &= r - 2Gm \\ e^{2A} &= 1 - \frac{2Gm}{r};\ e^{2B} = (1-\frac{2Gm}{r})^{-1} \\ d\tau^2 &= (1-\frac{2Gm}{r})dt^2 - \frac{dr^2}{1-\frac{2Gm}{r}} + r^2(d\theta^2 + \sin^2\phi d\phi^2) \end{aligned}$$

and this metrics gives geodesics which describe the motion of particles in a central gravitational field in better agreement with experiment than the Newtonian theory!

Exercise: Let ϕ_t be the one parameter group of diffeomorphisms of M generated by a smooth vector field X, g a Riemannian metric in M, and show that $(\frac{\partial}{\partial t}\Big|_{t=0} \phi_t^*(g))_{ij} = X_{i;j} + X_{j;i}$ (where ; means covariant derivative with respect to the Riemannian connection) [Hint: Let $g = g_{ij}dx_i \otimes dx_j$ where the x_i are geodesic coordinates at some point p and prove equality at p].

Remark: The Einstein tensor G^{ij} of any Riemannian metric always satisfied the differential identity $G^{ij}_{;j} = 0$. This can be obtained by contracting the Bianchi identities, but there is a more interesting proof. Let ϕ_t and X be as above where X say has compact support contained in the relatively compact open set θ of M. Note that clearly $\phi_t^*(R(g)\mu_g) = R(\phi_t^*(g))\mu_{\phi_t^*(g)}$ and so

$$\int_\theta R(g)\mu_g = \int_{\phi_t(\theta)} R(g)\mu_g = \int_\theta \phi_t^*(R(g)\mu_g) = \int_\theta R(\phi_t^*(g))\mu_{\phi_t^*(g)}$$

so

$$0 = \frac{d}{dt}\bigg|_{t=0} \int_\theta R(g)\mu_g = \int G_{ij}\delta g^{ij}\mu = -\int G^{ij}\delta g_{ij}\mu$$

where by the exercise $\delta g_{ij} = X_{i;j} + X_{j;i}$. Hence, since G^{ij} is symmetric

$$\begin{aligned} 0 &= \int G^{ij}X_{i;j} \\ &= \int (G^{ij}X_i)_{;j^\mu} - \int G^{ij}_{;j}X_i\mu. \end{aligned}$$

Now $(G^{ij}X_i)_{;j}$ is the divergence of the vector field $G^{ij}x_i$, so the first termsor vanishes. Hence $G^{ij}_{;j}$ is orthogonal to all covector fields with compact support and so must vanish.

THE STRESS-ENERGY TENSOR

The special relativity there is an extremely important symmetric tensor, usually denoted $T_{\alpha\beta}$, which describes the distribution of mass (or energy), momentum, and "stress" in space-time. More specifically T^{00} represents the mass-energy density, T^{0i} represents the i^{th} component of momentum density and T^{ij} represents the i, j component of stress [roughly, the rate of flow of the i^{th} component of momentum across a unit area of surface orthogonal to the x^j-direction]. For example consider a perfect fluid with density ρ_0 and world velocity v^α (i.e. if the world line of a fluid particle is given by $x^\alpha(\tau)$ then $v^\alpha = \frac{dx_\alpha}{d\tau}$ along this world line); then $T^{\alpha\beta} = \rho_0 v^\alpha v^\beta$. In terms of $T^{\alpha\beta}$ the basic conservation laws of physics (conservation of mass-energy, momentum, and angular momentum) take the simple unified form $T^{\alpha\beta}_{;\beta} = 0$. When it was recognized that the electromagnetic field

$$F_{\alpha\beta} = \begin{pmatrix} 0 & B_3 & -B_2 & E_1 \\ -B_3 & 0 & B_1 & E_2 \\ B_2 & -B_1 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix}$$

interacted with matter, so it could take or give energy and momentum, it was realized that if the conservation laws were to be preserved then as well as the matter stress-energy tensor $T_M^{\alpha\beta}$ there had to be an electromagnetic stress-energy tensor $T_{EM}^{\alpha\beta}$ associated to $F_{\alpha\beta}$ and the total stress energy tensor would be the sum of these two. From Maxwell's equations one can deduce that the appropriate expression (Maxwell stress energy tensor) is

$$T_{EM}^{\alpha\beta} = \frac{1}{4} g^{\alpha\beta} \|F\|^2 - g^{\alpha\mu} g^{\gamma\nu} g^{\beta\lambda} F_{\mu\nu} F_{\lambda\gamma}$$

where g denotes the Minkowski metric. (We will see how such a terrible expression arises naturally latter, for now just accept it). Explicitly we get:

$$\begin{aligned} T_{EM}^{00} &= \frac{1}{2}(\|E\|^2 + \|B\|^2) = \text{energy density} \\ T_{EM}^{0i} &= (E \times B)_i = \text{Poynting momentum vector} \\ T_{EM}^{ij} &= \frac{1}{2}\delta_{ij}(\|E\|^2 + \|B\|^2) - (E_i E_j + B_i B_j) \\ &= \text{Maxwell stress tensor.} \end{aligned}$$

If there is a distribution of matter with density ρ_0 then the Newtonian gravitational potential ϕ satisfies Poisson's equation

$$\Delta\phi = 4\pi\rho$$

we know for weak, static, gravitational fields to be described by a metric tensor $g_{\alpha\beta}$ we should have $g_{00} = (1 - 2\phi)$ and calculation gives

$$G^{00} = \Delta g_{00} = -2\Delta\phi.$$

Thus since $T^{00} = \rho_0$, Poisson's equation becomes

$$G^{00} = -8\pi T^{00}.$$

Now we also know $G^{\mu\nu}_{;\nu} = 0$ identically for geometric reasons, while $T^{\mu\nu}_{;\nu}$ expresses the basic conservation laws of physics. Where space is empty (i.e. $T^{\mu\nu} = 0$) we know $G^{\mu\nu} = 0$ are very good field equations.

The evidence is overwhelming that the correct field equations in the presence of matter are $G^{\mu\nu} = 8\pi T^{\mu\nu}$!

Let E be a smooth G bundle over an n-dimensional smooth manifold M. Eventually $n = 4$ and M is "space-time". A section ψ of E will be called a particle field, or simply a field. In the physical theory it "represents" (in a sense I will not attempt to explain) the fundamental particles of the theory. The dynamics of the theory is determined by a Lagrangian, $\hat{L}$, which is a (non-linear) first order differential operator from sections of E to n-forms on M;

$$\hat{L} : \Gamma(E) \to \Gamma(\Lambda^n(M)).$$

Recall that to say $\hat{L}$ is a first order operator means that it is of the form $\hat{L}(\psi) = L(j_1(\psi))$ where $j_1(\psi)$ is the 1-jet of ψ and $L : J^1(E) \to \Lambda^n(M)$ is a smooth map taking $J^1(E)_x$ to $\Lambda^n(M)_x$. If we choose a chart ϕ for M and an admissible local basis $s = (s^1, \dots, s^k)$ for E then w.r.t. the coordinates $x = x_1, \dots, x_n$ determined by ϕ and the components ψ^α of ψ w.r.t. s ($\psi = \psi^\alpha s^\alpha$) $(j_1\psi)_x$ is given by $(\psi_i^\alpha(x), \psi^\alpha(x))$ where $\psi_i^\alpha(x) = \partial_i\psi^\alpha(x) = \partial\psi^\alpha/\partial x_i$. Thus $\hat{L}(\psi) = L(j_1\psi)$ is given locally by

$$\hat{L}(\psi) = L^{s,\phi}(\psi_i^\alpha, \psi^\alpha, x)dx_1 \cdots dx_n$$

(we often omit the s, ϕ). The "field equations" which determines what are the physically admissible fields ψ are determined by the variational principle

$$\delta \int \hat{L}(\psi) = 0.$$

What this means explicitly is the following: given an open, relatively compact set θ in M define the action functional $A_\theta : \Gamma(E) \to \mathbb{R}$ by $A_\theta(\psi) = \int_\theta \hat{L}(\psi)$. Given any field $\delta\phi$ with support in θ define $\delta A_\theta(\psi, \delta\phi) = \frac{d}{dt}\Big|_{t=0} A(\psi + t\delta\phi)$. Then ψ is called an extremal of the variational principle $\delta \int \hat{L} = 0$ if for all θ and $\delta\phi$ $\delta A_\theta(\psi, \delta\phi) = 0$. If θ is included in the domain of the chart ϕ then the usual easy calculation gives:

$$\delta A_\theta(\psi, \delta\phi) = \int_\theta (\frac{\partial L}{\partial \psi^\alpha} - \frac{\partial}{\partial x_i}(\frac{\partial L}{\partial \psi_i^\alpha}))\delta\phi^\alpha dx_1 \cdots dx_n$$

where $L = L^{s,\phi}$ and we are using summation convention. Thus a NASC for ψ to be an extremal is that it satisfies the second order system of PDE (Euler-Lagrange equations):

$$\frac{\partial}{\partial x_i}(\frac{\partial L}{\partial \psi_i^\alpha}) = \frac{\partial L}{\partial \psi^\alpha}$$

[Hopefully, with a reasonable choice of L, in the physical case $M = \mathbb{R}^4$ these equations are "causal", i.e. uniquely determined by their Cauchy data, the ψ^α and $\frac{\partial \psi^\alpha}{\partial t}$ restricted to the Cauchy surface $t = 0$]. The obvious, important question is how to choose $\hat{L}$. To be specific, let $M = \mathbb{R}^4$ with its Minkowski-Lorentz metric

$$d\tau^2 = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2$$

and let $s = (s^1, \ldots, s^k)$ be a global admissible gauge for E. Then there are some obvious group invariance condition to impose on L. The basic idea is that physical symmetries should be reflected in symmetries of $\hat{L}$ (or L). For example physics is presumably the same in its fundamental laws everywhere in the universe, so L should be invariant under translation, i.e.

1) $L(\psi_i^\alpha, \psi^\alpha, x) = L(\psi_i^\alpha, \psi^\alpha)$ (no explicit x-dependence). Similarly if we orient our coordinates differently in space by a rotation, or if our origin of coordinates is in motion with uniform velocity relative the coordinates x_i these new coordinates should be as good as the old ones. What this means mathematically is that if $\gamma = \gamma_{ij}$ is a matrix in the group $O(1, 3)$ of Lorentz transformations (the linear transformations of $\mathbb{R}^4$ preserving $d\tau^2$) and if $\tilde{x}$ is a coordinate system for $\mathbb{R}^4$ related to the coordinates x by:

$$x_j = \gamma_{ij}\tilde{x}_i$$

then physics (and hence L!) should look the same relative $\tilde{x}$ as relative to x. Now by the chain rule:

$$\frac{\partial}{\partial \tilde{x}_i} = \frac{\partial x_j}{\partial \tilde{x}_i}\frac{\partial}{\partial x_j} = \gamma_{ij}\frac{\partial}{\partial x_j}$$

so

2) $L(\gamma_{ij}\psi_j^\alpha, \psi^\alpha) = L(\psi_i^\alpha, \psi^\alpha)$ for $\gamma \in O(1, 3)$. (Lorentz Invariance).

The next invariance principle is less obvious. Suppose $g = g_{\alpha\beta}$ is an element of our "gauge group" G. Then ("mathematically") the gauge $\tilde{s}$ related to s by $s^\beta = g_{\alpha\beta}\tilde{s}^\alpha$ is just as good as the gauge s. Since $\psi^\beta s^\beta = (g_{\alpha\beta}\psi^\beta)\tilde{s}^\alpha$ the component of ψ relative to $\tilde{s}$ are $\tilde{\psi}^\alpha = g_{\alpha\beta}\psi^\beta$. Thus it seems ("mathematically") reasonable to demand

3) $L(g_{\alpha\beta}\psi_i^\beta,\, g_{\alpha\beta}\psi^\beta) = L(\psi_i^\alpha,\, \psi^\alpha)$ for $g_{\alpha\beta} \in G$ ("global" gauge invariance). But is this physically reasonable? The indifference of physics to translations and Lorentz transformations is clear, but what is the physical meaning of a gauge rotation in the fibers of our bundle E? Well, think of it this way, G is <u>chosen</u> as the maximal group of symmetries of the physics in the sense of satisfying 3).

Of course we could also demand more generally that we have "local" gauge invariance i.e. if $g_{\alpha\beta} : \mathbb{R}^4 \to G$ is any smooth map we could consider the gauge $\tilde{s} = (\tilde{s}^1(x),\, \ldots,\, \tilde{s}^k(x))$ related to s by $s^\beta = g_{\alpha\beta}(x)s^\alpha(x)$; so again $\tilde{\psi}^\alpha(x) = g_{\alpha\beta}(x)\psi^\beta(x)$, but now since the $g_{\alpha\beta}$ are not constant

$$\tilde{\psi}_i^\alpha(x) = g_{\alpha\beta}(x)\psi_i^\beta(x) + \frac{\partial g_{\alpha\beta}}{\partial x_i}\psi^\beta(x)$$

and the analogue to 3) would be 3′)

$$L(g_{\alpha\beta}\psi_i^\beta + \frac{\partial g_{\alpha\beta}}{\partial x_i}\psi^\beta,\, \psi^\alpha) = L(\psi_{,i}^\alpha,\, \psi^\alpha)$$

for all smooth maps $g_{\alpha\beta} : \mathbb{R}^4 \to G$.

But this would be essentially impossible to satisfy with any L depending non-trivially on the $\psi_{,i}^\alpha$. We recognize here the old problem that the old problem that the "gradient" or "differential" operator d does not transform linearly w.r.t. non-constant gauge transformations — so does not make good sense in a non-trivial bundle. Nevertheless it is possible to make good sense out of local gauge invariance by a process the physicists call "minimal replacement" — and which not surprisingly involves the use of connections. However before considering this idea let us stop to give explicit examples of field theories.

Let us require that our field equations while not necessarily linear, be linear in the derivatives of the fields. This is easily seen to be equivalent to requiring that L

be a quadratic polynomial in the ψ_i^α and ψ^β, plus a function of the ψ^α:

$$L = A_{\alpha\beta}^{ij}\psi_i^\alpha\psi_j^\beta + B_{\alpha\beta}^i\psi_i^\alpha\psi^\beta + c_\alpha^i\psi_i^\alpha + v(\psi).$$

We can omit $c_\alpha^i\psi_i^\alpha = (c_\alpha^i\psi^\alpha)_i$ since it is a divergence. Similarly since $\psi_i^\alpha\psi^\beta + \psi^\alpha\psi_i^\beta = (\psi^\alpha\psi^\beta)i$ we can assume $B_{\alpha\beta}^i$ is skew in α, β and write L in the form

$$L = A_{\alpha\beta}^{ij}\psi_i^\alpha\psi_j^\beta + B_{\alpha\beta}^i(\psi_i^\alpha\psi^\beta - \psi^\alpha\psi_i^\beta) + v(\psi)$$

Inclusion of the second term leads to Dirac type terms in the field equations. To simplify the discussion we will suppose $B_{\alpha\beta}^i = 0$.

Now Lorentz invariance very easily gives $A_{\alpha\beta}^{ij} = c_{\alpha\beta}\eta_{ij}$ where $\eta_{ij}dx_idx_j$ is a quadratic form invariant under $O(1, 3)$. But since the Lorentz group, $O(1, 3)$ acts irreducibly on $\mathbb{R}^4$ it follows that $\eta_{ij}dx_idx_j$ is a multiple of $d\tau^2$, i.e. we can assume $n_{00} = +1$, $n_{ii} = -1$ $i = 1, 2, 3$ and $\eta_{ij} = 0$ for $i \neq j$. Similarly global gauge invariance gives just as easily that $C_{\alpha\beta}\psi^\alpha\psi^\beta$ is a quadratic form invariance under the gauge group G and that the smooth function V in the fiber $\mathbb{R}^k$ of E is invariant under the action of G, (i.e. constant on the orbits of G).

Thus we can write our Lagrangian in the form

$$L = \frac{1}{2}(\|\partial_0\psi\|^2 - \|\partial_1\psi\|^2 - \|\partial_2\psi\|^2 - \|\partial_3\psi\|^2) + V(\psi)$$

where $\| \ \|^2$ is a G-invariant quadratic norm (i.e. a Riemannian structure for the bundle E). Assuming the s^α are chosen orthonormal $\|\partial_i\psi\|^2 = \sum_\alpha (\partial_i\psi^\alpha)^2$, and the Euler-Lagrange equations are

$$\Box\psi^\alpha = \frac{\partial V}{\partial\psi^\alpha} \quad \alpha = 1, \ldots, k$$

where $\Box = \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}$ is the D'Alambertian or wave-operator.

As an example consider the case of linear field equations, which implies that $V = \frac{1}{2}M_{\alpha\beta}\psi^\alpha\psi^\beta$ is a quadratic form invariant under G. By a gauge rotation $g \in G$ we can assume V is diagonal in the basis s^α, so $V = \frac{1}{2}\sum_\alpha m_\alpha^2(\psi_\alpha)^2$ and the field equations are

$$\Box\psi^\alpha = m_\alpha\psi^\alpha \quad \alpha = 1, \ldots, k$$

Note these are k uncoupled equations (Klein-Gordon equations) (Remark: Clearly the set of ψ^α corresponding to a fixed value m of m_α span a G invariant subspace — so if G acts irreducibly — which is essentially the definition of a "unified" field theory, then the m must all be equal).

Now, and this is an important point, the parameters m_α are according to the standard interpretation of this model in physics measurable quantities related to masses of particles that should appear in certain experiments.

Let us go back to the more general case:

$$\Box^2 \psi^\alpha = \frac{\partial V}{\partial \psi^\alpha}$$

We assume V has a minimum value, and since adding a constant to V is harmless we can assume this minimum value is zero. We define Vac$= V^{-1}(0)$ to be the set of "vacuum" field configuration — i.e. a vacuum field ψ is a constant field (in the gauge s^α) such that $V(\psi) = 0$. Since at a minimum of V $\frac{\partial V}{\partial \psi^\alpha} = 0$, every vacuum field is a solution of the fields equations.

The physicists view of the world is that Nature "picks" a particular vacuum or "equilibrium" solution ψ_0 and then the state ψ of the system is of the form $\psi = \psi_0 + \phi$ when ϕ is small. By Taylor's theorem

$$V(\psi) = V(\psi_0) + (\frac{\partial V}{\partial \psi^\alpha})_{\psi_0} \psi^\alpha + \frac{1}{2} M_{\alpha\beta} \psi^\alpha \psi^\beta$$

plus higher order terms in ψ^α, where

$$V(\psi_0) = (\frac{\partial V}{\partial \psi^\alpha})_{\psi_0} = 0$$

and

$$M_{\alpha\beta} = (\frac{\partial^2 V}{\partial \psi^\alpha \partial \psi^\beta})_{\psi_0}$$

is the Hessian of V at ψ_0. Thus if we take ψ_0 as a new origin of our vector space, i.e. think of the $\phi = \psi^\alpha - \psi_0^\alpha$ as our fields, then as long as the ϕ^α are small the theory with potential V should be approximated by the above Klein-Gordon theory with mass matrix M the Hessian of V at ψ_0.

The principal direction $s^1, \ldots, s^k$ of this Hessian are called the "bosons" of the theory and the corresponding eigenvalues m_α $\alpha = 1, \ldots, k$ the "masses" of the theory (for the particular choice of the vacuum ψ_0). Now if $\tilde{w}$ denotes the orbit of G thru ψ_0 then V is constant on $\tilde{w}$, hence by a well known elementary argument the tangent space to $\tilde{w}$ at ϕ_0 is in the null space of the Hessian. We can choose $s^1, \ldots, s^r$ spanning the tangent space to $\tilde{w}$ at ψ_0 and then $m_1 = \cdots = m_r = 0$. [These r massless bosons are called the "Goldstone bosons" of the theory (after Goldstone who pointed out their existence, and physicists usually call this existence theorem "Goldstone's Theorem"). Massless particles should be easy to create and observe and the lack of experimental evidence of their existence caused problems for early versions of the so-called spontaneous symmety breaking field theories which we shall discuss later. These problems are overcome in an interesting and subtle way by the technique we will describe next for making field theories "locally" gauge invariant.

MINIMAL REPLACEMENT

As we remarked earlier, the search for Lagrangians $L(\psi_i^\alpha, \psi^\alpha)$ which are invariant under "local" gauge transformations of the form $\psi^\alpha \to g_{\alpha\beta}\psi^\beta$ where $g_{\alpha\beta} : M \to G$ is a possibly non-constant smooth map leads to a dead end. Nevertheless we can make our Lagrangian formally invariant under such transformations by the elementary expedient of replacing the ordinary coordinate derivative $\psi_i^\alpha = \partial_i \psi^\alpha$ by $\nabla_i^w \psi^\alpha$ where ∇^w is an admissible connection on E and ∇_i^w means the covariant derivative w.r.t. ∇^w in the direction $\frac{\partial}{\partial x_i}$. We can think of ∇^0 as being just the flat connection d with respect to some choice of gauge and $\nabla^w = \nabla^0 + w$ where as usual w is a G-connection form on M, i.e. a $\mathcal{G}$ valued one-form on M. At first glance this seems to be a notational swindle; aren't we just absorbing the offending term in the transformation law for ∂_i under gauge transformations into the w? Yes and no! If we simply made the choice of ∇^w a part of the given of the theory it would indeed be just such a meaningless notational trick. But "minimal replacement", as this process is called, is a more subtle idea by far. The important idea is <u>not to make any a priori choice of</u> ∇^w, but rather let the connection become a "dynamical variable" of the theory itself — on a logical par with the particle fields ψ, and like them determined by field equations coming from a variational principle.

Let $w_{\alpha\beta'}$ as usual be determined by $\nabla^w s^\beta = w_{\alpha\beta} s^\alpha$, and define the gauge potentials, or Christoffel symbols:

$$A^\alpha_{i\beta} = w_{\alpha\beta}(\partial_i)$$

so that

$$\nabla^w_i \psi^\alpha = \partial_i \psi^\alpha + A^\alpha_{i\beta}\psi^\beta$$

so that if our old particle Lagrangian was $\hat{L}_p(\psi) = L_p(\partial_i \psi^\alpha_1 \psi^\alpha)\mu$ then after minimal replacement it becomes

$$\hat{L}_p(\psi,\, \nabla^w) = L_p(\partial_i \psi^\alpha + A^\alpha_{i\beta}\psi^\beta,\, \psi^\alpha)\mu.$$

Now we can define the variational derivative of $\hat{L}_p$ with respect to w, which is usually called the "current" and denoted by J. It is a three form on M with values in $\mathcal{G}$, depending on ψ and w, $J(\psi,\, w)$, defined by

$$\left.\frac{d}{dt}\right|_{t=0} \hat{L}_p(\psi,\, \nabla^{w+t\delta w}) = J(\psi,\, \nabla^w) \wedge \delta w$$

so that if δw has support in a relatively compact set θ then

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} \int_\theta \hat{L}_p(\psi,\, \nabla^{w+t\delta w}) &= \int J(\psi,\, \nabla^w) \wedge \delta w \\ &= \int \delta w \wedge J \\ &= ((\delta w,\, {}^*J)) \end{aligned}$$

so that clearly the component of ${}^*J$, the dual current 1-form are given by:

$${}^*J^\alpha_{i\beta} = \frac{\partial L_p}{\partial A^\alpha_{i\beta}}.$$

For example for a Klein-Gordon Lagrangian

$$L_p = \frac{1}{2}\eta^{ij}(\partial_i\psi^\alpha)(\partial_j\psi^\alpha) + V(\psi)$$

we get after replacement:

$$L_p = \frac{1}{2}\eta^{ij}(\partial_i\psi^\alpha)(\partial_j\psi^\alpha) + \eta^{ij}A^\alpha_{i\beta}\psi^\beta\partial_j\psi^\alpha + \frac{1}{2}\eta^{ij}A^\alpha_{i\beta}A^\alpha_{j\gamma}\psi^\beta\psi^\gamma$$

and the current 1-form ${}^*J$ has the components

$${}^*J^\alpha_{i\beta} = \eta^{ij}(\nabla_j\psi^\alpha)\psi^\beta.$$

Thus if we tried to use $\hat{L}_p$ as our complete Lagrangian and to determine w (or A) by extremalizing the corresponding action w.r.t. w we would get for the connection the algebraic "field equations"

$$0 = \frac{\delta L}{\delta w} = {}^*J$$

or for our special case:

$$\psi^\beta(\partial_i\psi^\alpha + A^\alpha_{i\gamma}\psi^\gamma) = 0.$$

As long as one of the ψ^α doesn't vanish (say $\psi^0 \neq 0$) we can solve this by:

$$A^\alpha_{i\beta} = \begin{cases} 0 & \beta \neq 0 \\ -(\partial_i\psi^\alpha)/\psi^0 & \beta = 0. \end{cases}$$

A very easy calculation shows $F^\alpha_{ij\beta} = \partial_j A^\alpha_{i\beta} - \partial_i A^\alpha_{j\beta} - [A^\alpha_{i\gamma},\, A^\gamma_{j\beta}] = 0$ so in fact the connection is flat!

But this gives an unphysical theory and we get not only a more symmetrical (between ψ and ∇^w) theory mathematically, but a good physical theory by adding to $\hat{L}_p$ a connection Lagrangian $\hat{L}_{\mathcal{C}}$ depending on the one-jet of w

$$\hat{L}_{\mathcal{C}}(\nabla^w) = L_{\mathcal{C}}(A^\alpha_{i\beta,j},\, A^\alpha_{i\beta})\mu$$

(where $A^\alpha_{i\beta,j} = \partial_j A^\alpha_{i\beta}$). Of course we want $L_{\mathcal{C}}$ like L_p to be not only translation and Lorentz invariant, but also invariant under gauge transformations $g : M \to G$. After all, it was this kind of invariance that led us to minimal replacement in the first place.

We shall now explain a simple and natural method for constructing such translation-Lorentz-gauge invariant Lagrangians $L_{\mathcal{C}}$. In the next section we shall prove the remarkable (but very easy!) fact — Utiyama's Lemma — which says this method in fact is the only way to produce such Lagrangian.

Let $\mathbb{R}^{1,3}$ denote $\mathbb{R}^4$ considered as a representation space of the Lorentz group $O(1, 3)$ and let $\mathcal{G}$ as usual denote the Lie algebra of G, considered as representation space of G under ad. Then a two form on $M = \mathbb{R}^4$ with values in $\mathcal{G}$ can we considered a map of M into the representation space $(\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}) \otimes \mathcal{G}$ of $O(1, 3) \times G$. Now given a connection ∇^w for E and a gauge, the matrix Ω^w_{ij} of curvature two-forms is just such a map

$$F : x \mapsto F^{\alpha}_{ij\beta}(x) = \partial_j A^{\alpha}_{i\beta} - \partial_i A^{\alpha}_{j\beta} + [A^{\alpha}_{i\gamma}, A^{\gamma}_{j\beta}]$$

which moreover depends on the 1-jet $(A^{\alpha}_{i\beta,j}, A^{\alpha}_{i\beta})$ of the connection. If we make a Lorentz transformation γ on M and a gauge transformation $g : M \to G$ then this curvature (or field strength) map is transformed to

$$x \mapsto (\gamma \otimes \mathrm{ad}(g(x)))F(x) = \gamma_{ik}\gamma_{jk} g^{-1}_{\alpha\gamma}(x) F^{\lambda}_{k\ell\mu}(x) g_{\mu\beta}(x).$$

Thus if $\Lambda : (\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}) \otimes \mathcal{G} \to \mathbb{R}$ is smooth function invariant under the action of $O(1, 3) \times G$, then $L_{\mathcal{C}} = \Lambda(F)$ will give us a first order Lagrangian with the desired invariance properties for connections. (And as remarked above, Utiyama's Lemma says there are no others).

As for the case $\hat{L}_p$ let us restrict ourselves to the case of field equations linear in the highest (i.e. second order) derivatives of the connection. This is easily seen to be equivalent to assuming that Λ is a quadratic form on $(\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}) \otimes \mathcal{G}$, of course invariant under the action of $O(1, 3) \times G$, that is Λ is of the form $Q^1 \otimes Q^2$ where Q^1 is an $O(1, 3)$ invariant quadratic form on $\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}$ and Q^2 is an ad invariant form on $\mathcal{G}$. To simplify the discussion assume G is simple, so that G acts irreducibly on $\mathcal{G}$ under ad, and Q^2 is uniquely (up to a positive multiplicative constant) determined to be the Killing form:

$$Q^2(x) = -\mathrm{tr}(A(X)^2).$$

If q denotes the $O(1,\,3)$ invariant quadratic form $\eta_{ij}v_iv_j$ on $\mathbb{R}^{1,3}$, then for Q^1 we can take $q \wedge q$ and this gives for Λ the Yang-Mills Lagrangian

$$\hat{L}_{\mathcal{C}} = \hat{L}_{\mathrm{YM}} = \frac{1}{4}\|\Omega\|^2\mu = \frac{1}{4}\Omega \wedge {}^*\Omega$$

or in component form

$$L_{\mathrm{YM}} = \frac{1}{4}\eta_{ik}\eta_{j\ell}F^{\alpha}_{ik\beta}F^{\beta}_{j\ell\alpha}$$

[But wait, is this all? If we knew $O(1,\,3)$ acted irreducibly on $\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}$ then $q \wedge q$ would be the <u>only</u> $O(1,\,3)$ invariant form on $\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}$. Now, quite generally, if V is a vector space with non-singular quadratic form q, then the Lie algebra (V) of the orthogonal group $O(V)$ of V is canonically isomorphic to $V \wedge V$ under the usual identification between skew-adjoint (w.r.t. q) linear endomorphisms of V and skew bilinear forms on V. Thus $V \wedge V$ is irreducible if and only if the adjoint action of $O(V)$ on its Lie algebra is irreducible — i.e. if and only if $O(V)$ is simple (By the way, this argument shows $q \wedge q$ is just the Killing form of (V)). Now it is well known that $O(V)$ <u>is</u> simple except when $\dim(V) = 2$ (when it is abelian) and $\dim(V) = 4$ — the case of interest to us. When $\dim(v) = 4$ the orthogonal group is the product of two normal subgroups isomorphic to orthogonal groups of three dimensional spaces. It follows that there is a self adjoint (w.r.t. $q \wedge q$) map $\tau : \mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3} \to \mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}$ <u>not</u> a multiple of the identity and which commutes with the action of $O(1,\,3)$, such that $(u,\,v) \to q \wedge q(u,\,\tau v)$ together with $q \wedge q$ span the $O^{1,3}$-invariant bilinear forms on $\mathbb{R}^{1,3} \wedge \mathbb{R}^{1,3}$. Clearly τ is just the Hodge *-operator:

$$\tau = *_2 : \Lambda^2(\mathbb{R}^{1,3}) \to \Lambda^{4-2}(\mathbb{R}^{1,3})$$

(and conversely, the existence of $*_2$ shows orthogonal groups in four dimensions are not simple!), so the corresponding Lagrangian in just $\Omega \wedge {}^*(\tau\Omega) = \Omega \wedge \Omega$. But $\Omega \wedge \Omega$ is just the second chern form of E and in particular it is a <u>closed</u> two form and so integrates to zero; i.e. adding $\Omega\wedge\Omega$ to the Yang-Mills Lagrangian $\hat{L}_{\mathrm{YM}}$ would not change the action integral $\int \hat{L}_{\mathrm{YM}}$. Thus, finally (and modulo Utiyama's Lemma below) we see that when G is simple the unique quadratic, first order, translation-Lorentz-gauge invariant Lagrangian $\hat{L}_{\mathcal{C}}$ for connection is up to a scalar multiple the Yang-Mills Lagrangian:

$$\hat{L}_{\mathrm{YM}}(\nabla^w) = -\frac{1}{2}\Omega\wedge{}^*\Omega.$$

UTIYAMA'S LEMMA

As usual we write $A^\alpha_{i\beta}$ for the Christoffel symbols (or gauge potentials) of a connection in some gauge and $A^\alpha_{i\beta,j}$ for their derivatives $\partial_j A^\alpha_{i\beta}$. Then coordinate functions for the space of 1-jets of connections are $(a^\alpha_{i\beta j},\, a^\beta_{i\beta})$ and if $F(a^\alpha_{i\beta j},\, a^\alpha_{i\beta})$ is a function on this space of 1-jets we get, for a particular connection <u>and</u> choice of gauge a function on the base space M by $x \to F(A^\alpha_{i\beta,j}(x),\, A^\alpha_{i\beta}(x))$. We are looking for functions $F(a^\alpha_{i\beta j},\, a^\alpha_{i\beta})$ such that this function on M depends <u>only</u> on the connection w and <u>not</u> on the choice of gauge. Let us make the linear non-singular change of coordinates in the 1-jet space

$$\begin{aligned}\bar{a}^\alpha_{i\beta j} &= \frac{1}{2}(a^\alpha_{i\beta j} + a^\alpha_{j\beta i}) \qquad i \le j\\ \hat{a}_{i\beta j} &= \frac{1}{2}(a^\alpha_{i\beta j} - a^\alpha_{j\beta i}) \qquad i < j\end{aligned}$$

i.e. replace the $a^\alpha_{i\beta j}$ by their symmetric and anti symmetric part relative to $i,\, j$. Then in these new coordinates the function F will become a function

$$\tilde{F}(\hat{a}^\alpha_{i\beta j},\, \bar{a}^\alpha_{i\beta j},\, a^\alpha_{i\beta}) = F(\hat{a}^\alpha_{i\beta j},\, \bar{a}^\alpha_{i\beta j},\, a^\alpha_{i\beta})$$

and the function on the base will be $\tilde{F}(\frac{1}{2}(A^\alpha_{i\beta,j} - A^\alpha_{j\beta,i}),\, \frac{1}{2}(A^\alpha_{i\beta,j} + A^\alpha_{j\beta,i}),\, A^\alpha_{i\beta})$ which again does not depend on which gauge we use. Utiyama's Lemma says we can find a function of just the $\hat{a}^\alpha_{i\beta j}$

$$F^*(\hat{a}^\alpha_{i\beta j})$$

such that the function on M is given by $x \to F^*(F^{\alpha}_{i\beta j}(x))$ where $F^{\alpha}_{i\beta j}$ are the field strengthes:

$$F^{\alpha}_{i\beta j} = A^{\alpha}_{i\beta,j} - A^{\alpha}_{ij,\beta} + [A^{\alpha}_{i\gamma},\, A^{\gamma}_{j\beta}]$$

The function F^* is in fact just given by:

$$F^*(\hat{a}^{\alpha}_{i\beta j}) = \tilde{F}(2\hat{a}^{\alpha}_{i\beta j},\, 0,\, 0).$$

To prove that this F^* works it will suffice (because of the gauge invariance of $\tilde{F}$) to show that given an arbitrary point p of M we can choose a gauge such that in this gauge we have $\underline{\text{at } p}$:

$$A^{\alpha}_{i\beta}(p) = 0$$
$$A^{\alpha}_{i\beta,j}(p) + A^{\alpha}_{j\beta,i}(p) = 0$$

and note that this automatically implies that at p we also have:

$$F^{\alpha}_{ij\beta}(p) = A^{\alpha}_{i\beta,j}(p) - A^{\alpha}_{j\beta,i}(p)$$

The gauge that does this is of course just the quasi-canonical gauge at p; i.e. choose p as our coordinate origin in $\mathbb{R}^4$, pick any frame $s^1(p),\, \ldots,\, s^k(p)$ for E_p and define $s^{\alpha}(x)$ for any point x in $\mathbb{R}^4$ by parallel translating $s^{\alpha}(p)$ along the ray $t \to p + t(x-p)$ $(0 \leq t \leq 1)$ from p to x.

If $x(t) = (x_0(t),\, \ldots,\, x_3(t))$ is any smooth curve in M and $v(t) = \sum_{\alpha} v^{\alpha}(t) s^{\alpha}(x(t))$ is a vector field along $x(t)$, recall that the covariant derivative $\frac{Dv}{dt} = \sum_{\alpha} (\frac{Dv}{dt})^{\alpha} s^{\alpha}(x(t))$ of $v(t)$ along $x(t)$ is given by:

$$(\frac{Dv}{dt})^{\alpha} = \frac{dv^{\alpha}}{dt} + A^{\alpha}_{i\beta}(x(t)) v^{\beta}(t) \frac{dx_i}{dt}$$

Now take for $x(t)$ the ray $x(t) = t(\lambda_0,\, \lambda_1,\, \lambda_2,\, \lambda_3)$ and $v_{(\gamma)}(t) = s^{\gamma}(x(t))$, so $x_i(t) = t\lambda_i$ and $v^{\alpha}_{(\gamma)}(t) = \delta^{\alpha}_{\gamma}$. Then since $V_{(\gamma)}$ is parallel along $x(t)$

$$0 = (\frac{Dv_{(\gamma)}}{dt})^{\alpha} = A^{\alpha}_{i\gamma}(t\lambda_0, t\lambda_1,\, t\lambda_2,\, t\lambda_3)\lambda_i$$

In particular (taking $t = 0$ and noting λ_i is arbitrary) we get $A^{\alpha}_{i\gamma}(p) = 0$ which is part of what we need. On the other hand differentiating with respect to t and setting $t = 0$ gives

$$A^{\alpha}_{i\gamma,j}(p)\lambda_i\lambda_j = 0$$

and again, since λ_i, λ_j are arbitrary, this implies

$$A^{\alpha}_{i\gamma,j}(p) + A^{\alpha}_{j\gamma,i}(p) = 0$$

and our proof of Utiyama's lemma is complete. $\square$

GENERALIZED MAXWELL EQUATIONS

Our total Lagrangian is now

$$\hat{L}(\psi,\, \nabla^w) = \hat{L}(\psi, \nabla^w) + \hat{L}_{\text{YM}}(\nabla^w)$$

where

$$\hat{L}_{\text{YM}} = -\frac{1}{2}\Omega \wedge {}^*\Omega.$$

Our field equations for both the particle field ψ and the gauge field ∇^w are obtained by extremaling the action integrals $\int_\theta \hat{L}(\psi,\, \nabla)$ (where θ is a relatively compact open set of M and the variation $\delta\psi$ and δw have support in θ). Now we have defined the current three-form J by

$$\delta_w \int \hat{L}_p = \int \delta w \wedge J = ((\delta w,\, {}^*J))$$

and long ago we computed that

$$\delta_w \int \hat{L}_{\text{YM}} = ((\delta w,\, -{}^*D^{w*}\Omega^w))$$

so the "inhomogeneous" Yang-Mills field equations for the connection form w is just $0 = \delta_w \int \hat{L}$ or

$$\boxed{D^{w*}\Omega = J}$$

(of course we also have the trivial homogeneous equations $D\Omega = 0$, the Bianchi identity). As we pointed out earlier when $G = SO(2)$, the "abelian" case, these equations are completely equivalent to Maxwell's equations, when we make the identification ${}^*J = (\rho, j_1, j_2, j_3)$ where ρ is the change density and $\vec{j} = (j_1, j_2, j_3)$ the current density and of course $\Omega = F_{ij} dx_i \wedge dx_j$ is identified with the electric field $\vec{E}$ at magnetic field $\vec{B}$ as described before.

These equations are now to be considered as part of a coupled system of equations, the other part of the system being the particular field equations:

$$\frac{\delta \hat{L}_p}{\delta \psi} = 0.$$

[Note that $J = \frac{\delta \hat{L}_p}{\delta w}$ will involve the particle fields and their first derivatives explicitly, and similarly $\frac{\delta \hat{L}_p}{\delta \psi}$ will involve the gauge field and its derivative explicitly, so we must really look at these equations as a coupled system — <u>not</u> as two independent systems, one to determine the connection ∇^w and the other to determine ψ].

COUPLING TO GRAVITY

There is a final step in completing our mathematical model, namely coupling our particle field ψ and gauge field ∇^w to the gravitational field g. Recall that g was interpreted as the metric tensor of our space-time. [That is, $d\tau^2 = g_{\mu\nu} dx_\mu dx_\nu$, where the integral of $d\tau$ along a world line of a particle represent atomic clock time of a clock moving with the particle. Also paths of particles not acted upon by forces (other than gravity) are to be geodesics in this metric. And finally, in geodesic coordinates near a point we expect the metric to be very well approximated over substantial regions by the Lorentz-Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu}$].

The first step is then to replace the metric tensor η_{ij} in $\hat{L} = \hat{L}_p + \hat{L}_{\mathrm{YM}}$ by a metric tensor g_{ij} which now becomes a dynamical variable of our theory, on a par with ψ and ∇^w

$$\hat{L}(\psi, \nabla^w, g) = \hat{L}_p(\psi, \nabla^w, g) + \hat{L}_{\mathrm{YM}}(\nabla^w, g)$$

where

$$\begin{aligned}\hat{L}_{\text{YM}} &= -\frac{1}{2}\|\Omega^w\|^2\mu_g \\ &= -\frac{1}{2}\Omega^w \wedge {}^*\Omega^w \\ &= -\frac{1}{2}(g^{ik}g^{i\ell}F^{\alpha}_{k\ell\beta}F^{\alpha}_{ij\beta})\sqrt{g}dx_1\cdots dx_n\end{aligned}$$

and for a Klein-Gordon type theory $\hat{L}_p$ would have the form

$$\hat{L}_p = (\frac{1}{2}g^{ij}\nabla_i\psi^\alpha\nabla_j\psi^\alpha + v(\psi))\sqrt{g}dx_1\cdots dx_n$$

[Note the analogy between this process and minimal replacement. Just as we replaced the flat connection d by a connection ∇^w to be determined by a variational principle, so now we replace the flat metric η_{ij} by a connection q_{ij} to be determined by a variational principle. This is in fact the mathematical embodiment of Einstein's principle of equivalence or general covariance and was used by him long before minimal replacement].

The next step (in analogy to defining the current J as $\frac{\delta\hat{L}}{\delta w}$) is to define stress-energy tensors $T^{ij} = T^{ij}_p + T^{ij}_{\text{YM}}$ by $T^{ij} = \frac{\delta_L}{\delta g_{ij}}$, $T^{ij}_p = \frac{\widehat{\delta'_p}}{\delta g_{ij}}$, and $T^{ij}_M = \frac{\delta\hat{L}_{\text{YM}}}{\delta g_{ij}}$, [Remark: since L involves the g_{ij} algebraically, that is does <u>not</u> depend on the derivatives of the g_{ij}, no integration by parts is required and $\frac{\delta\hat{L}}{\partial g_{ij}}$ is essentially the same as $\frac{\partial\hat{L}}{\partial g_{ij}}$].

<u>Exercise</u>. Compute T^{ij}_{YM} explicitly and show that for the electromagnetic case ($G = SO(2)$, so $\hat{L}_{\text{YM}} = \frac{1}{2}g^{ik}g^{j\ell}F_{k\ell}F_{ij}\sqrt{g}dx_1\cdots dx_n$) that this leads to the stress-energy tensor T^{ij}_{EM} described in our earlier discussion of stress energy tensors.

Finally, to get our complete field theory we must add to the particle Lagrangian $\hat{L}_p$ and connection or Yang-Mills Lagrangian $\hat{L}_{\text{YM}}$ a gravitational Lagrangian $\hat{L}_G$, depending <u>only</u> on the metric tensor g:

$$\hat{L} = \hat{L}_p(\psi,\, \nabla^w,\, g) + \hat{L}_{\text{YM}}(\nabla^w,\, g) + \hat{L}_G(g).$$

From our earlier discussion we know the "correct" choice for $\hat{L}_G$ is the Einstein-Hilbert Lagrangian

$$\hat{L}_G(g) = -\frac{1}{8\pi} R(g)\mu_g$$

where $R(g)$ is the scalar curvature function of g. The complete set of coupled field equations are now:

1) $\frac{\delta \hat{L}_p}{\delta \psi} = 0$ (particle field equations)

2) $D^{w*}\Omega^w = J$ (Yang-Mills equations)

3) $G = -8\pi T$ (Einstein equations)

Note that while 1) and 2) look like our earlier equations, formally, now the "unknown" metric g rather than the flat metric η_{ij} must be used in interpreting these equations. (Of course 1), 2), and 3) are respectively the consequences of extremalizing the action $\int \hat{L}$ with respect to ψ, ∇^w, and g).

THE KALUZA-KLEIN UNIFICATION

We can now complete our discussion of the Kaluza-Klein unification of Yang-Mills fields with gravity. Let P denote the principal frame bundle of E and consider the space m_p of <u>invariant</u> Riemannian metrics on P. We recall that if m_M denotes all Riemannian metric on M and $\mathcal{C}(E)$ all connections in E then we have a canonical isomorphism $m_p \simeq m_M \times \mathcal{C}(E)$, say $P_g \to ({}^M g, \nabla^w)$. We define the Kaluza-Klein Lagrangian for metric ${}^P g$ to be just the Einstein-Hilbert Lagrangian $R(g)\mu_g$ restricted to the invariant metrics:

$$\hat{L}_{k-k}({}^P g) = R({}^P g)\mu_{(P_g)}$$

Now since ${}^P g$ is invariant it is clear that $R({}^P g)$ is constant on fibers, and in fact we earlier computed that

$$R({}^P g) = R({}^M g) - \frac{1}{2}\|\Omega^w\|^2 + {}^G R$$

where $R({}^Mg)$ is the scalar curvature of the metric ${}^Mg$ (at the projected point) and ${}^GR$ is the constant scalar curvature of the group G. It is also clear that

$$\mu^P g = \pi^*(\mu_{M_g}) \wedge \mu_G$$

where μ_G is the measure on the group. Thus by "Fubini's Theorem":

$$\int_P \hat{L}_{k-k}({}^Pg) d\mu_{P_g} = \mathrm{vol}(G)(\int_M \hat{L}_G({}^Mg) + \int_M \hat{L}_{\mathrm{YM}}(\nabla^w) + {}^GR)$$

It follows easily that ${}^Pg$ is a critical point of the LHS, i.e. a solution of the empty-space Einstein equations on P

$$G = 0,$$

if and only if ${}^Mg$ and ∇^w extremalize the RHS — which we know is the same as saying that ∇^w (or Ω) is Yang-Mills and that ${}^Mg$ satisfies the Einstein field equations:

$$G = -T_{\mathrm{YM}}$$

where T is the Yang-Mills stress-energy tensor of the connection ∇^w. [There is a slightly subtle point; on the LHS we should extremalize with respect to <u>all</u> variations of ${}^Pg$, not just invariant variations. But since the functional $\int_P \hat{L}_{k-k}$ itself is clearly invariant under the action of G on metrics, it really is enough to only vary with respect to invariant metrics. For a discussion of this point see "The Principle of Symmetric Criticality", R. Palais, Comm. in Math. Phys., Dec. 1979].

Kaluza-Klein Theorem

An invariant metric on a principal bundle P satisfies the empty-space Einstein field equations, $G = 0$, if and only if the "horizontal" sub-bundle of TP (orthogonal to the vertical sub-bundle), considered as a connection, satisfies the Yang-Mills equations and the metric on the base (obtained by projecting the metric on P) satisfies the full Einstein-equations

$$G = T,$$

where T is the Yang-Mills stress energy tensor, computed for the above connection. Moreover in this case the paths of particles in the base, moving under the generalized Lorentz force, are exactly the projection on the base of geodesics on P.

Proof. Everything is immediate, either from our earlier discussion or the remarks preceding the statement of the theorem.

THE DISAPPEARING GOLDSTONE BOSONS

There is a very nice bonus consequence of making our field theory ("locally") gauge invariant. By a process that physicists often refer to as the "Higgs mechanism" the unwanted massless goldstone bosons can be made to "disappear" — or rather it turns out that they are gauge artifacts and that in an appropriate gauge (that depends on which particle field we are considering) they vanish identically locally.

The mathematics behind this is precisely the "Slice Theorem" of transformation group theory, which we now explain in the special case we need. Let H be a closed subgroup or our gauge group G (it will be the isotropy group of the vacuum, what physicists call the "unbroken group"). We let $\mathcal{H}$ denote the Lie algebra of H. Denoting the dimension of G/H by d we choose an orthonormal basis $e_1, \ldots, e_f$ for $\mathcal{H}$ with $e_1, \ldots, e_d$ in $\mathcal{H}^\perp$ and $e_{d+1}, \ldots, e_f$ in $\mathcal{H}$.

Clearly every element of G suitably close to the identity can be written uniquely in the form $\exp(X)h$ with X near zero in $\mathcal{H}^\perp$ and h near the identity in H, and more generally it follows from the inverse function theorem that if G acts smoothly in a space X and $x_0 \in X$ has isotropy group H then $x \mapsto \exp(X)x_0$ maps a neighborhood of zero in $\mathcal{H}^\perp$ diffeomorphically onto a neighborhood of x_0 in its orbit Gx_0.

Now suppose G acts orthogonally on a vector space V and let $v_0 \in V$ have isotropy group H and orbit Ω. Consider the map $(x, v) \to \exp(X)(v_0 + \nu)$ of $\mathcal{H}^\perp \times T\Omega_{v_0}^\perp$ into V. Clearly $(0, 0) \to v_0$, and the differential at $(0, 0)$ is bijective, so by the inverse function theorem we have

<u>Theorem</u>. There is an $\varepsilon > 0$ and a neighborhood U of v_0 in V such that each $u \in U$ can be written uniquely in the form $\exp(X)(v_0 + \nu)$ with $X \in \mathcal{H}^\perp$ and $\nu \in T\Omega_{v_0}^\perp$ having norms $< \varepsilon$.

<u>Corollary</u>. Given a smooth map $\psi : M \to U$ (U as above) there is a unique map $X : M \to \mathcal{H}^\perp$ with $\|X(x)\| < \varepsilon$, such that if we define

$$\phi : M \to V$$

by

$$\exp(X(x))^{-1}\psi(x) = v_0 + \phi(x)$$

then ϕ maps M into $T\Omega_{v_0}^\perp$.

Now let us return to our Klein-Gordon type of field theory <u>after</u> minimal replacement. The total Lagrangian is now

$$\hat{L} = \frac{1}{2}\eta^{ij}\nabla_i\psi \cdot \nabla_j\psi + V(\psi) - \frac{1}{4}\|\Omega\|^2$$

where

$$\begin{aligned} \nabla_i\psi &= \partial_i\psi + A_i^\alpha e_{\alpha\psi} \\ \|\Omega\|^2 &= \eta^{ik}\eta^{j\ell}F_{ij}^\alpha F_{k\ell}^\alpha \\ F_{ij}^\alpha &= \partial_i A_j^\alpha - \partial_j A_i^\alpha + A_i^\beta A_j^\gamma C^{\alpha\beta\gamma} \end{aligned}$$

where $C^{\alpha\beta\gamma}$ are the structure constants of $\mathcal{H}$ in the basis e_α; i.e.

$$[e_\beta, e_\gamma] = C^{\alpha\beta\gamma}e_\alpha$$

[Note by the way that if we consider just the pure Yang-Mills Lagrangian $\frac{1}{4}\|\Omega\|^2$, the potential terms are all cubic or quartic in the field variable A_i^α, so the Hessian at the unique minimum $A_i^\alpha \equiv 0$ is zero — i.e. all the masses of a pure Yang-Mills field are zero].

Now, assuming as before that V has minimum zero (so the vacuum fields of this theory are $A_i^\alpha \equiv 0$ and $\psi \equiv v_0$ where $V(v_0) = 0$) we pick such a vacuum (i.e. make a choice of v_0) and let H be the isotropy group of v_0 under the action of the gauge group G. Since $\mathcal{H}$ is the kernel of the map $X \to Xv_0$ of $\mathcal{H}$ into V, this map is bijective on $\mathcal{H}^\perp$; thus $k(X, Y) =< Xv_0, Yv_0 >$ is a positive definite symmetric bilinear form on $\mathcal{H}^\perp$ and we can assume that the basis $e_1, \ldots, e_d$ of $\mathcal{H}^\perp$ is chosen not only orthogonal with respect to the Killing form, but also orthogonal with respect to k and that

$$< e_\alpha v_0, e_\alpha v_0 >= M_\alpha^2 > 0$$

That is $e_\alpha v_0 = M_\alpha u_\alpha$ where u_α $\alpha = 1, \ldots, d$ is an orthonormal basis for $T\Omega_{v_0}$, the tangent space at v_0 of the orbit Ω of v_0 under G. We let $u_{d+1}, \ldots, u_k$ be an orthonormal basis for $T\Omega_{v_0}^\perp$ consisting of eigenvectors for the Hessian of V at v_0, say with eigenvalues $m_\alpha^2 \geq 0$. Putting $\psi = v_0 + \phi$, to second order in

$$\phi = \sum_{\alpha=1}^{k} \phi^\alpha u_\alpha$$

we have

$$V(\psi) = V(v_0 + \psi) = \frac{1}{2} \sum_{\alpha=d+1}^{k} m_\alpha^2 (\phi^\alpha)^2.$$

Now

$$\begin{aligned} \nabla_i \psi &= \partial_i \phi + A_i^\alpha e_\alpha (v_0 + \phi) \\ &= \partial_i \phi + \sum_{\alpha=1}^{\alpha} M_\alpha A_i^\alpha u_\alpha + \sum_{\alpha=1}^{f} A_i^\alpha e_\alpha \phi \end{aligned}$$

(where we have used $e_\alpha v_0 = M_\alpha u_\alpha$). Thus to second order in the A_i^α and the shifted fields ϕ we have

$$\begin{aligned}\hat{L} &= \frac{1}{2}\eta^{ij}(\partial_i\phi)\cdot(\partial_j\phi) + \frac{1}{2}\sum_{\alpha=j+1}^{k} m_\alpha(\phi^\alpha)^2 \\ &+\frac{1}{2}\sum_{\alpha=1}^{d} M_\alpha^2\eta^{ij}A_i^\alpha A_j^\alpha + \sum_{\alpha=1}^{d} M_\alpha\eta^{ij}A_i^\alpha(\partial_j\phi\cdot u_\alpha)\end{aligned}$$

The last term is peculiar and not easy to interpret in a Klein-Gordon analogy. But now we perform our magic! According to the corollary above, by making a gauge transformation on our $\psi(x) = (v_0 + \phi(x))$ (in fact a unique gauge transformation of the form

$$\exp(X(x))^{-1}\psi(x)$$

$X : M \to \mathcal{H}^\perp$) we can insure that in the new gauge $\phi(x)$ is orthogonal to $T\Omega_{v_0}$ {right in front of your eyes the "Goldstone bosons" of ψ, i.e. the component of ϕ tangent to Ω at v_0, have been made to disappear or, oh well, been gauged away}.

Since in this gauge $\phi : M \to T\Omega_{v_0}^\perp$, also $\partial_j\phi \in T\Omega_{v_0}^\perp$. Since the u_α, $\alpha \leq d$ lie in $T\Omega_{v_0}$ they are orthogonal to the $\partial_j\phi$, so the $\partial_j\phi\cdot u_\alpha = 0$ and the offensive last term in $\hat{L}$ goes away. What has happened is truly remarkable. Not only have the d troublesome massless scalar fields ϕ^α $1 \leq \alpha \leq d$ "disappeared" from the theory. They have been replaced by an equal number of massive vector fields A_i^α (recall $M_\alpha^2 = < e_\alpha v_0,\, e_\alpha v_0 >$ is definitely positive).

Of course we still have the $f - d = \dim(H)$ massless vector fields A_i^α $d+1 \leq \alpha \leq f$ in our theory, so unless $H = S^1$, (giving us the electromagnetic or "photon" vector field) we had better have some good explanation of why those "other" massless vector fields aren't observed.

In fact, in the current favorite "electromagnetic-weak force" unification of Weinberg-Salam H is S^1. Moreover particle accelerator energies are approaching the level (about 75 GEV) where the "massive vector bosons" should be observed.

If they are not ....

I would like to thank Professor Lee Yee-Yen of the Tsing-Hua Physics department for helping me to understand the Higgs-Kibble mechanism, and also for helping to keep me "physically honest" by sitting in on my lecture and politely pointing out my mis-statements.