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)

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:

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

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) $\quad x(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^{3n}$ $\quad$ 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 $\frac{v}{\|v\|}$!

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 $G(X) = 0$

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

<u>Example</u>: 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 <u>still</u> equivalent to:

$$\delta_M \int_{t_1}^{t_2} K(\frac{dX}{dt})dt = 0$$

$\left\{ \begin{array}{ll} \delta_M : & \text{only vary} \\ & \text{w.r.t. paths} \\ & \text{in } M. \end{array} \right.$

<u>OR</u>

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)

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 kinematic 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 different 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 dynamical 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, not 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.

<u>Definition of smooth k-dimension vector bundle over M</u>.

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

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

$\theta \subseteq M \quad s : \theta \to E \quad$ smooth is a <u>section</u> 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

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

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 <u>local gauges</u> for E over θ. <u>Basic defining axiom for smooth vector bundle is that each $p \in \theta$ has a neighborhood θ for which there is a local gauge $F : \theta \times \mathbb{R}^k \simeq E|\theta$</u>.

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 $p \in M$ $\quad f|(E_1)_p$ 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 <u>group of gauge transformations</u> 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 $0 \to E_1 \to E_2 \to E_2/E_1 \to 0$