Acta Physica Austriaca, Suppl. XXIII, 433–476 (1981)

A BRIEF INTRODUCTION TO THE GEOMETRY OF GAUGE FIELDS+

by

M. E. MAYER
Dept. of Physics, Univ. of California
Irvine, California, 92717, USA

and

A. TRAUTMAN
Inst. of Theoretical Physics, Warsaw Univ.
Hoża 69, 00-681 Warsaw, Poland

1. INTRODUCTION

In view of the common background required for the understanding of the lectures of both authors, and in order to avoid unnecessary duplications, we have decided to present jointly this brief introduction to the language and properties of fiber bundles. By now the advantages of the fiber-bundle formulation of gauge field theories have led to a widespread acceptance of this language, and a number of reviews of the subject have appeared or are in course of publication. These, together with a number of standard textbooks are listed in the references to this

+Lectures given at the XX. Internationale Universitätswochen für Kernphysik, Schladming, Austria, February 17 - 26, 1981.

introduction. Nevertheless, we felt that it would be convenient for the reader of these proceedings to have at his disposal a summary of the basic facts. We also tried to clarify a number of concepts and propose an acceptable terminology wherever a standard has not been established in the literature. This refers, in particular, to the terms gauge transformation , pure gauge transformation, and the related (infinite-dimensional) groups as well as to the concepts of extension, prolongation, restriction, and reduction of bundles, which are used with slightly varying meaning in different texts.

In the oral presentation most of the general background material was presented by Andrzej Trautman, and the material related to reduction and symmetry of connections was given in Meinhard Mayer's lectures. Little, if anything, in this introduction is original. The actual text has been written in California by the first author and slightly revised by the second during his stay in France after the Schladming meeting.

No detailed proofs are given here, but wherever possible illustrations and examples are used to make the concepts plausible to physicists. Many proofs are straightforward and can be carried out by introducting local coordinates and bases. However, we recommend to the reader who wants to become familiar with the spirit of modern, coordinate-free, differential geometry to try to stay away from bases and indices as much as possible.

2. EXAMPLES OF BUNDLES. FUNDAMENTAL DEFINITIONS

We assume that the reader is familiar with such fundamental notions as differential manifold (including the concept of charts, atlases, diffeomorphism, etc.) and with the calculus of exterior differential forms, as well as with the fundamental concepts related to Lie groups and Lie algebras. In this section we list a few examples, illustrated by pictures, which will provide the intuitive background for understanding the more formal definitions and statements of the remainder of this lecture.

Of course, the most important examples of bundle structures appearing in contemporary physics are furnished by abelian and nonabelian gauge theories, and by the (pseudo-) Riemannian manifolds of general relativity. In the most familiar abelian gauge theory - electrodynamics of a charged field - the field ψ defined on Minkowski space and with values in some complex vector space (for simplicity, consider a complex scalar, or a Dirac or Pauli spinor) is subjected to the "point-dependent phase transformation", $\psi(x) \to \exp\{i\alpha(x)\}\psi(x)$, and in order to reestablish invariance of the equations of motion, one replaces in them (or in the Lagrangian) the ordinary space-time derivatives ∂_μ by the "covariant" derivative $\nabla_\mu = \partial_\mu + iA_\mu$ (we set $e = c = 1$) where the new field A_μ subject to the "gauge transformation of the second kind" $iA_\mu \to i(A_\mu + \partial_\mu\alpha) = iA_\mu + g(x)^{-1}\partial_\mu g(x)$, where $g(x) = \exp(i\alpha(x))$ is a smooth function on spacetime with values in the group U(1). We see here the appearance of a function on spacetime with values in a Lie group, or more generally, a copy of the group U(1) attached to each point of spacetime - the trivial principal bundle $M \times U(1)$. A simple analysis of magnetic monopoles shows that this picture is not adequate, and that in some situations a more intimate mixture of spacetime and the gauge group U(1) becomes necessary, where a product representation is valid only locally.

Similarly, if one considers a field which transforms under a representation of a nonabelian compact Lie group G (e.g., G = SU(2), in the original work of Yang and Mills), one is led to a structure in which a copy of G is attached to each point of space-time, and locally, in a neighborhood

U of a point, this can be represented as the product $U \times G$. The place of the electromagnetic vector potential is taken by the Lie-algebra-valued (matrix-valued) one-form $A = \sum A^a_\mu(x)\, e_a dx^\mu$, where (e_a) is a basis of the Lie algebra $\mathcal{G}$ of G (in the case of SU(2), $e_a = i\sigma_a$, $a = 1,2,3$, σ_a are the Pauli matrices), and as the field $\phi(x)$ is subjected to the "local gauge transformation" $\phi(x) \to g(x)\phi(x)$ the equations of motion are preserved if A is subjected to the affine transformation $A(x) \to g(x)^{-1}A(x)g(x) + g(x)^{-1}dg(x)$ provided ordinary differentials are replaced by "covariant differentials" $d\phi \to D\phi = d\phi + [A,\phi]$. These differentials do not commute and their "commutator" is related to the Yang-Mills field strength two-form

$$F = dA + \frac{1}{2}[A,A] \quad . \tag{2.1}$$

This two-form is subject to the "Bianchi identity" (just as the electromagnetic field strength two-form satisfies $dF = 0$)

$$DF = dF + [A,F] = 0 \quad . \tag{2.2}$$

In distinction from the abelian case, F is not gauge-independent, but transforms as

$$F(x) \to g(x)^{-1}F(x)g(x) \quad ,$$

and, if one wishes to write down the inhomogeneous Yang-Mills equation, one may generalize the Maxwell equations to

$$D{*}F(x) = -{*}J(x) \quad . \tag{2.3}$$

Here $*$ means the Hodge-duality operator which associates to a p-form in n-space the dual (n-p)-form (provided the manifold is oriented and has a Riemannian metric). In particular, in Minkowski space $*$ can be defined through its action on the basis coordinate forms:

$$*dx^{0} = dx^{1} \wedge dx^{2} \wedge dx^{3} \text{ (cycl.perm.)},$$

$$*(dx^{1} \wedge dx^{2}) = dx^{0} \wedge dx^{3} \text{ , etc..} \tag{2.4}$$

Alternatively, $*F$ can be thought of as the two-form with components given by $*F_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}$. $*J$ denotes the current three-form of the matter field which satisfies the "covariant divergence" equation $D{*J} = 0$.

In order to see that the concept of "a group labeled by a point in a manifold". i.e., a principal bundle, or a vector space labeled by a point in a manifold, i.e., a vector bundle, appear naturally in geometry, and that matrix-valued forms, such as A and F, are usually associated with such objects, we consider some more elementary geometric examples.

In looking at these examples it is important to remember that the concept of fiber bundle generalizes the notion of direct (or cartesian) product of two spaces, and that the concept of section generalizes the graph of a function. Thus, the simplest example (and the one easiest to picture) is the cartesian product of two sets X (the domain space of the function $f: X \to Y$) and Y (the range space). The cartesian product $X \times Y$ can be viewed as formed by copies of Y (fibers) attached to each point of X, called the base space. In the usual treatment of cartesian products X and Y are treated on an equal footing. In fiber-bundle theory the base space X plays the role of a label space, whereas the typical fiber Y is quite distinct (and may be of a different nature, e.g., a group G, a vector space V, a homogeneous (coset) space G/H, such as a sphere, etc.). A distinguished role is also played by the projection $\pi : X \times Y \to X : (x,y) \to x$. Another distinction to be kept in mind is the fact that in a cartesian product we automatically identify points in different fibers "which lie on the same horizontal". In a fiber bundle there is no such automatic identification; it must be introduced as extra structure, by defining a section (or a basis of sections, in a vector bundle), i.e., picking a distinguished point or basis in each fiber. As we shall see this cannot always be done (if it can, and the bundle is principal, then it may be identified with a product, and is called trivial). Some of these notions are illustrated in Fig.1. The reader should keep in mind that most of our illustrations are for products (i.e., trivial bundles, since the nontrivial cases are difficult to draw; the reader might think of the Möbius band or the Klein bottle as examples of nontrivial bundles).

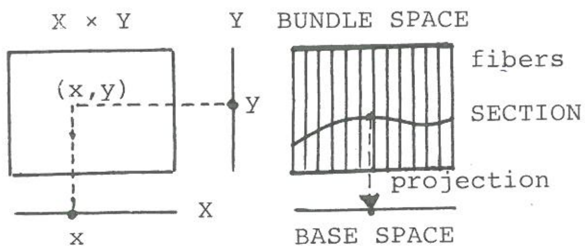

Fig. 1

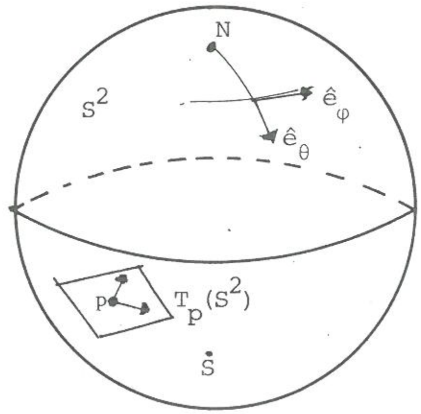

Fig. 2

An example which illustrates nontrivial bundles and some of the general concepts to be defined later is the tangent bundle of the sphere S^2 and the associated frame bundle of dyads on an oriented sphere (Fig. 2). Consider the unit sphere S^2 in R^3 as the base space X of our bundle. This sphere is a differentiable manifold, and an atlas of charts consists, e.g., of two open sets obtained by removing caps around the north and south poles, and the mappings of the remaining portions of the sphere onto R^2 realized by stereographic projections from these two poles. To each of the points on S^2 we attach the two-dimensional tangent space spanned, e.g., by the two tangent vectors $\hat{e}_\theta$, $\hat{e}_\phi$. We denote this tangent space at the point p by $T_p(S^2)$. The union of all these two-dimensional vector spaces, as the point p ranges over the sphere, is the tangent bundle TS^2. It is a four-dimensional manifold, since each point in it is labelled by the two coordinates of p and the two coordinates of the tangent vector at p. Moreover, since the "sphere cannot be combed", we cannot represent this bundle as a cartesian product of S^2 and R^2 globally (although in any chart of S^2 this is possible), and therefore we are dealing with a vector bundle on S^2 which is nontrivial, but locally trivial. Associated with this vector bundle is a bundle - the frame bundle - the fiber of which is isomorphic to a group. Indeed, consider the bundle of oriented dyads (pairs of unit vectors tangent to S^2 at each point). In an embedding into R^3 it is clear that the total space of this bundle is isomorphic to SO(3), since any dyad can be taken into any other dyad by an orientation-preserving rotation of R^3. Locally, such a rotation can be viewed as consisting of the two parameters labelling the point p on the sphere and an angle ψ which takes a "standard dyad" (e.g., the east-north dyad) into the given dyad, i.e., locally, SO(3) decomposes into the product of an open set U in S^2 and a copy of the group SO(2) - the typical fiber, the projection π being the smooth map associating the point p to the frame at that point.

The bundle $SO(3) \to S^2$ is an example of a principal bundle, in which the typical fiber is a Lie group G (just as for a gauge theory), which also acts on the bundle space on the right (since we like to write frame transformations as right multiplication by matrices, reserving the left multiplication for actions of the group on vector components), and the tangent bundle is a vector bundle associated with this principal bundle by the fundamental representation of SO(2) by rotation matrices in two dimensions.

A generalization of this example to four dimensions is the bundle of tetrads (Vierbeins) often used in general relativity. Here the typical fiber is the Lorentz group which takes the standard tetrad $e^o_a(x)$ at a given point into an arbitrary tetrad $e_b(x)$. In distinction from the previous example, if the base space is homeomorphic to R^4, such a bundle will be trivial, i.e., will admit a product representation $M \times SO(1,3)$.

The principal bundles of dyads, tetrads, and, more generally, frame bundles (which may be of higher differential order than one) are more special or "richer" than the principal bundles occurring in gauge theories of the Yang-Mills type. The elements of the former bundles are "concrete": they can be defined and easily visualized in terms of geometrical constructions referring only to the base space. In other words, frame bundles are soldered to the base. This important concept distinguishes the theory of gravitation among all gauge theories; it is discussed in some detail in the lectures by the second author.

To illustrate the concept of a nontrivial principal bundle and of its local triviality, consider a last example (Fig. 3), where the base-space is the circle S^1 and the fiber is the discrete group of all integers Z. We easily obtain two bundles: the trivial product $S^1 \times Z$ (left) which is just the union of a countable set of circles "stacked"

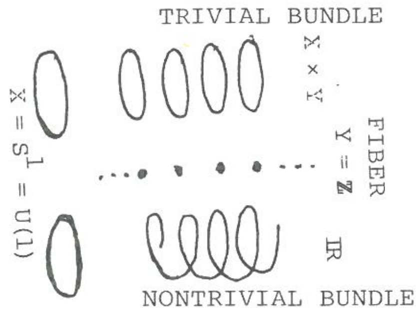

Fig. 3

over S^1, and the nontrivial helix (right) homeomorphic to R with projection map $\pi : R \rightarrow S^1$ given by $\pi(t) = \exp(2\pi it)$. It is obvious that acts on the bundle space by "translation" $(t,n) \rightarrow t + n$, where $t \in R$, $n \in Z$. The attentive reader will recognize in this example an essential ingredient of the Riemann surface of the logarithm. The nontriviality of this bundle reflects the impossibility of defining the logarithm as a smooth function on the pointed complex plane $C - \{0\}$.

We are now ready to summarize the precise definitions to be used in the sequel. Further examples will be discussed in Section 5.

A gauge theory consists of various mathematical structures associated with a <u>principal bundle</u> $P(M,G) = (P,M,G,\pi)$, where M, the <u>base space,</u> is spacetime (Minkowski space, its imaginary-time version R^4, or the compactification S^4 of the latter, or one of the (pseudo)-Riemannian manifolds of general relativity), G is the <u>structure group</u> of the bundle (the gauge group, in the physics literature), P is a smooth manifold which locally, i.e., over a covering of M by open sets U_i, has a product structure: $P|U_i \simeq U_i \times G$, and π is the projection map

π : P $\rightarrow$ M, a smooth surjection of P onto M, such that the inverse image $\pi^{-1}(x) = P_x$, the <u>fiber</u> over $x \in M$, is diffeomorphic to G. The bundle is called <u>trivial</u> if it is isomorphic to M $\times$ G (globally). The above mentioned isomorphism of the restriction of P to each open set U_i to a product is called a local <u>trivialization.</u> The group G acts on P to the right in such a manner that the equivalence induced by this action is the same as that induced by the projection π, i.e. all points in the same orbit of G project onto the same point in M.

A <u>section</u> (or cross section) of a principal bundle is a (smooth) mapping s : M $\rightarrow$ P such that $\pi(s(x)) = x$. Global sections may not exist in principal bundles; in fact if such a section exists, the bundle is trivial, and vice versa. However, in view of the local triviality, local sections, i.e., sections over properly chosen open sets U_i always exist and can be used to describe the local trivialization. Indeed, let us consider the diffeomorphism between $P|_U$ and $U \times G$ given by $y \rightarrow (\pi(y), \phi(y)), y \in P$, $\phi(y) \in G$, such that for any $g \in G$, $\phi(yg) = \phi(y)\cdot g$. Then the local section s_U is defined by $s_U(x) = y\cdot\phi(y)^{-1}$ for y in $\pi^{-1}(x)$, obviously independent of the point y, because one can be taken into another by the action of G. The local section corresponds to an identification of the identity in the group with the submanifold of $\pi^{-1}(U)$ corresponding to U. It is sometimes convenient to describe the principal bundle in terms of charts, and the change of coordinates from one chart to another by means of <u>transition functions</u> (G-valued cocycles), i.e., maps $g_{UV} : U \cap V \rightarrow G$ satisfying the cocycle identity on $U \cap V \cap W$: $g_{UW}(x) = g_{UV}(x)\cdot g_{VW}(x)$. A change of local trivialization subjects the sections to a <u>local gauge transformation</u> $s'_U(x) = s_U(x)\cdot g_U(x)$ where g_U is a G-valued function on U, and the transition functions to $g'_{UV}=g_U^{-1}g_{UV}g_V$.

The principal bundle P(M,G) is the kinematic background of the gauge theory, since it specifies both the spacetime manifold M and the gauge group G. It will carry some of the dynamics of the field theory, since the connections and curvatures (see Section 3) will be identified as gauge potentials and gauge fields, and the metric structure of the base space may be identified with the gravitational field.

However, in order to accommodate the various matter fields describing particles, we have to associate to the bundle P(M,G) various vector bundles (or bundles with homogeneous spaces G/H) as fibers), in which the particle fields can be considered as sections. Alternatively, we may consider the particle fields as smooth maps from the bundle space P into a vector space V on which G acts (on the left) by means of a representation r, and which are equivariant under this action. More precisely, let $\phi : P \to V$ be such a map. Then equivariance means that for any $g \in G$ and $p \in P$

$$\phi(p \cdot g) = r(g^{-1})\phi(p) \quad . \tag{2.5}$$

In other words, the field ϕ defined on the bundle space P with values in the vector space V in fact depends only on the projection $x = \pi(p)$, i.e., is a field defined on space-time in the usual sense. When composed with, or pulled-back by, a section s_U, the field will be $\phi_U = s_U^{*}\phi = \phi \circ s_U$, where ϕ_U is a V-valued function of $x \in U$, and if the trivialization is changed by a local gauge transformation , $s_U(x) \to$ $\to s_U(x) g_U(x) = s'_U(x)$, the field will be subjected to the "gauge transformation of the second kind" $\phi'_U(x) = r(g_U(x)^{-1})\phi_U(x)$. The same method works if V is replaced by an arbitrary manifold F on which G acts on the left.

Alternatively, one can first define a fiber bundle with typical fiber F (or V, in the case of vector bundles) associated with P(M,G) by the representation (action) r of G in the following manner: consider the cartesian product

$P \times F$. The group G acts naturally on the right on this product in the following way. Let $p \in P$, $f \in F$, $g \in G$. Then $(p,f)\cdot g = (p\cdot g,\ r(g^{-1})f)$. The orbit space of this action, i.e., the set of equivalence classes under the equivalence $(p,f) \sim (p\cdot g, r(g^{-1})f)$ is denoted by $P \times {}_G F$ and is called the fiber bundle associated with P by the action of G by r on the typical fiber F. It is denoted by $E(M,F,G)$ and has a natural projection $\pi_E : E \to M$, obtained by factoring the composition of the projection of $P \times F$ on P with the bundle map $P \to M$ through the quotient map $P \times F \to E$. If $F = V$ is a vector space we obtain a vector bundle, and if $F = G/H$ is a coset space we obtain a bundle with homogeneous spaces G/H as fibers.

It is an easy exercise to show that a section of $E(M,F,G)$ defines an equivariant function on P with values in F, and vice versa. Therefore, matter fields can also be regarded as sections of the associated (vector) bundle E, which is convenient in some constructions.

Before going on to the definition of connections, curvature and holonomy, we recall several general constructions involving fiber bundles which will be used in the physical applications.

3. MORPHISMS OF BUNDLES. FIBERED PRODUCTS AND PULLBACKS. EXTENSION AND RESTRICTION OF THE STRUCTURE GROUP

We limit our definitions to the case of principal bundles, although some are valid for more general bundles, and all transpose easily to associated fiber bundles. We list only the most important definitions, referring the reader for more details to the literature.

Let $P(M,G,\pi)$ and $P'(M',G',\pi')$ be two principal bundles. A morphism m of P into P' is a pair $m = (u,h)$, where $u : P \to P'$ is a C^∞-map and $h : G \to G'$ is a Lie-group homomorphism such that $u(p \cdot g) = u(p) \cdot h(g)$ for all p in P and g in G. It is clear that u takes fibers of P into fibers of P' (remembering that a fiber in a principal bundle is the orbit under the right group action), and therefore induces a C^{∞}-map $v : M \to M'$ such that the diagram

$$\begin{array}{ccc} P & \xrightarrow{\;u\;} & P' \\ \downarrow{\scriptstyle\pi} & & \downarrow{\scriptstyle\pi'} \\ M & \xrightarrow{\;v\;} & M' \end{array}$$

is commutative. This can be taken as the definition of a morphism of bundles which are not necessarily principal. If h is an isomorphism (we then identify G with G') and v is a diffeomorphism then the map u is also a diffeomorphism, and the morphism m is called a bundle isomorphism. In particular, an isomorphism of P onto itself is called a <u>bundle automorphism.</u> The group of all bundle automorphisms of a given principal bundle is an infinite-dimensional group, denoted by Aut P. Usually, in gauge theories, we are interested in bundle automorphisms which reduce to the identity map on the base space, i.e., for which $v = Id_M$. We call such bundle automorphisms <u>vertical</u> or <u>based.</u> Automorphisms which leave some absolute elements invariant will be called <u>gauge transformations.</u> A gauge transformation which is vertical is called a <u>pure gauge transformation</u> (for a more detailed discussion of this distinction which is important in general-relativistic contexts, cf. Section 6).

If we consider a local trivialization of the bundle P over the open set U of M, defined by a section s_U (the reader may think of the local section s_U as a way of identifying the group identity e_x over each point x of the base M), then it is easy to see that the pure gauge transformation $u : P \to P$ is implemented by a function

$g_U : U \to G$ which is smooth, so that the section s_U is taken into the section $s'_U(x) = s_U(x) g_U(x)$ (i.e., it reduces to a gauge transformation in the sense employed by physicists). Thus, the group of pure gauge transformations $\mathcal{G}_o$ may be viewed as the infinite-dimensional group of all smooth G-valued functions on M (cf. Section 6).

Another concept which is useful and where the group structure of the principal bundle does not intervene is the concept of fibered product and pullback.

The <u>fibered product</u> of two bundles (here it does not matter whether we are dealing with principal bundles, vector bundles or other fiber bundles) (T_1, M, P_1), (T_2, M, P_2), over the same base manifold M, with total spaces T_i, and projections P_i is defined as the bundle

$$(T_1 \times_M T_2, M, p_1 \times_M p_2) \tag{3.1}$$

which has as total space the submanifold of $T_1 \times T_2$ consisting of all pairs (t_1, t_2) such that $p_1(t_1) = p_2(t_2)$ (i.e., points in the fibers project onto the same point of the base), and as projection, the restriction of the product-projection to that subspace. The local triviality of the fibered product is easily established. This concept is also useful for maps.

An important special case of the fibered product is obtained if one considers instead of one of the bundles above a (smooth) mapping of a manifold M' into the base manifold M of a bundle. On obtains by the same construction the <u>pullback</u> or <u>induced bundle</u> over M with the same fiber. More precisely, let $\lambda = (E, M, p)$ be a bundle over M with fiber F (a group G in the case of principal bundles, a vector space in the case of vector bundles), and let $f : M' \to M$ be a (smooth) map of M' into M. Then the triple $f^*(\lambda) = (M' \times_M E, M', p')$ is a bundle, called the pullback of λ by f (or the induced bundle, or reciprocal image bundle) and denoted by $f^*(\lambda) = M' \times_M \lambda$. Here p' is the restriction of the projection onto the first factor of $M' \times E$ to the fibered product $M' \times_M E$. In particular, if M' is a submanifold and f is the injection, the pullback is the same as the induced bundle on the submanifold.

The fiber of the pullback bundle is the same as the fiber of the original bundle, but that the base space has been replaced by M'. In particular, this will be one way of defining gauge theories over extensions of the usual space-time manifold. A section $s : M \to E$ of the bundle ξ is taken into a section $s' : M' \to f^*(\xi)$ of the pullback bundle by means of the relation $s'(x') = (x', s(f(x')))$, where x' is a point in M'. This also defines a map $f' : M' \times_M E \to E$ such that $f' \circ s' = s \circ f$. The map f' is the restriction of the second projection of the product $M' \times E$ to the fibered product.

In physical and geometrical applications of fiber bundles it is often necessary to change the structure group of the bundle. Thus, if one deals with a gauge theory with symmetry breaking the original Lagrangian is defined on a principal bundle P with structure group G, and the "vacuum" (or classical critical field) may have the lower symmetry group H, a closed subgroup of G. There arises the question of constructing a principal bundle Q with structure group H, and the relation between P and Q. Similarly, in general-relativistic contexts, one considers the frame bundle with structure group GL(4,R), whereas physics often imposes either restrictions of GL(4,R) to one of its subgroups, such as SO(1,3), if the metric is to be contemplated on the same footing as a Higgs field, or extensions to larger groups, such as the affine group $GL(4,R) \times R^4$ (semidirect product). In general, one can consider four distinct operations related to changes of the structure group. The terminology employed in the literature in this context is not uniform (and there are sometimes subtle differences in definitions). Here we adopt the terminology proposed by one of the authors (A.T., 1976) which is at variance with that used in some of the literature and in some of the lectures by the other author (although the distinction plays no important role in the Yang-Mills context).

Consider two principal bundles $\xi = (Q, M, H, p)$ and $\eta = (P, M, G, \pi)$ over the same base space M, and a morphism $m = (f,v)$ of Q into P, where f is a smooth map of the bundle space Q into the bundle space P, v is a diffeomorphism of M onto itself (not necessarily the identity), and the corresponding homomorphism of the structure groups is denoted, as before, by $h : H \to G$.

If both f and h are <u>injective immersions</u> (i.e., one-to-one into, and such that the tangent map is injective at each point of the respective manifold), then ξ is called a <u>restriction</u> of η relative to the morphism $m = (f,v,h)$, or, simpler, Q is a <u>restriction</u> of P to the subgroup H of G, and η is called an <u>extension</u> of ξ relative to the morphism m, or simpler, P is called an <u>extension</u> of Q to the structure group G.

If both f and h are <u>surjective submersions</u> (i.e., their tangent maps are surjective at each point), then η is called a <u>reduction</u> of ξ relative to m, or simpler, P is a reduction of Q to the structure group G, and ξ is called a <u>prolongation</u> of η relative to m, or Q is a prolongation of P to the group H.

In particular, we shall assume in the sequel that $v = Id_M$ and that H is a closed Lie subgroup of G, such that the coset space G/H is a differentiable manifold. If $f : Q \to f(Q) \subset P$ is a diffeomorphism of Q onto a closed submanifold of P, such that $f(q \cdot h) = f(q) \cdot h$ for all $h \in H$, $q \in Q$, $f(q) \in P$, then we are in the restriction-extension situation, which we discuss in more detail.

Note. The terminology here follows that of Trautman (1976) and is closest to Bourbaki and Dieudonné, whereas Kobayasi-Nomizu (as well as Mayer) call our restriction reduction. In the Yang-Mills context there is usually no risk of confusion, but in the discussion of G-structures and spin structures some care is indicated.

There is no difficulty in obtaining an extension of the bundle Q with structure group H to a larger group G. Indeed, construct the associated bundle $Q \times_H G$ with fiber G based on the left action of H on G. We now let G act on the right on this space (consisting of H-orbits) by $p \cdot g = ((q,g')H) \cdot g = (q,g' \cdot g)H$ with $p \in P$, $q \in Q$, $g,g' \in G$ (recall that $Q \times_H G$ consists of equivalence classes of pairs $(q,g) \sim (q \cdot h, h^{-1} \cdot g)$, where we have not written out the homomorphism of H into G which is understood), and the action fibers this manifold over M, resulting in the bundle P. The morphism $f : Q \to P$ is given by $f(q) = (q,e)H$, where e is the identity in G (and H). It is easy to see that the projection of the associated bundle is the projection in P, and that the local triviality of Q induces local triviality of P.

On the other hand, restriction of P to a subgroup H of G is not always possible. This can be easily seen if one notices that a necessary and sufficient condition for a restriction to exist is that there should be a covering of M by open sets such that the corresponding transition functions take values in the subgroup H of G. Globally, one obtains a more interesting condition:

The bundle $P(M,G)$ has a restriction to the group H iff the associated bundle $E(M,G/H,G,P) = P \times_G G/H = P/H$ admits a cross section $\sigma : M \to E$.

It is easy to see that the orbit space of P under the action of the subgroup H of G, P/H, can be identified with the associated bundle E. Denoting by γ the canonical projection of G onto G/H, we can set for $p \in P, \delta(p) = p\cdot\gamma(e)$, where e is the identity of G. The mapping $\delta: P \to E$ is a projection for the new principal bundle (P, H, E, δ) over the larger base $E = P \times_G G/H$ which is canonically identified with the orbit space P/H (this is illustrated in Fig. 4, in the middle).

Let now $\sigma : M \to E$ denote a section of E and $\sigma^* : (P, H, E, \delta) = (Q, H, M, \rho)$ the pullback (induced bundle) of this map. It is obvious (cf. Fig. 4, right) that this is now a principal bundle with structure group H over M, and its extension to G is isomorphic to the original bundle P. Two different sections σ_1 and σ_2 of E will define isomorphic restrictions iff they are mapped into each other by a pure gauge transformation (G-M-automorphism) of P. Otherwise different sections of E determine different (nonisomorphic) restrictions.

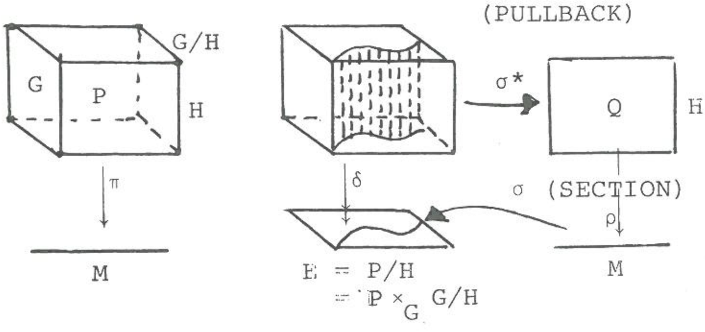

Fig. 4

4. CONNECTIONS, CURVATURE, AND HOLONOMY

We have seen in Sec. 2 that gauge potentials and fields are described by differential forms with values in the Lie algebra of the gauge group and act on fields through covariant differentiation. We have also made it plausible that fields should be considered as equivariant functions on a principal bundle with values in a vector space V (or a homogeneous space G/H) on which the gauge group G acts on the left. In this section we discuss the concept of a connection and its curvature on a principal bundle and show how these concepts are related to the familiar Yang-Mills potential and its field-strength (considered as matrix-valued one- and two-forms, respectively, on spacetime) and to the Levi-Civita connection (represented by the familiar Christoffel symbols) and the Riemann-Christoffel curvature tensor in the four-dimensional pseudo-Riemannian manifolds describing spacetime in general relativity.

The modern concept of a connection in a principal fiber bundle, and the associated covariant differentiations in associated bundles, has evolved during the first half of this century through the work of many geometers. There exist several equivalent definitions of a connection. We restrict our attention to one which is most useful in the context of gauge theories.

The fibered structure of a principal bundle, in which each fiber is isomorphic to the Lie group G, the structure group of the bundle, suggests that the tangent space $T_p(P)$ to the bundle space P at the point p contains a distinguished subspace of <u>vertical vectors,</u> Ver_pP, vectors which are tangent to the fiber P_x over a point x of the base M. As is well known, the tangent vectors to a Lie group G at a point g form a vector space $T_g(G)$ which is linearly isomorphic to $T_e(G) = G$, the Lie algebra of G. Thus, there exists an isomorphism between vertical vectors in T(P), the tangent bundle of P, and elements of the Lie algebra G, allowing us to identify these (sometimes it is convenient to introduce a field A^* of vertical vectors corresponding to an element A in G; A^* is called the fundamental vector field, and the right action of G on P intertwines with the adjoint action on G, i.e., $R_{g^*} A^*$ is the fundamental vector field corresponding to $Ad(g^{-1})A)$. Since locally the bundle space P is isomorphic to the product between a subset of M and the group G, the tangent space $T_p(P)$ is isomorphic to a direct sum of $Ver_p(P)$ and a vector space $Hor_p(P)$ which must be isomorphic to $T_{\pi(p)}(M)$. Unfortunately, since the bundle P has no intrinsic "orthogonality" structure, there is no canonical way of identifying the horizontal subspace, and defining a connection in P means <u>defining</u> at each point p such a horizontal subspace $Hor_p(P)$, smoothly over p, and equivariantly under the action of G, i.e., $R_{g^*} Hor_p(P) = Hor_{p \cdot g}(P)$ for $g \in G$. Giving such a distribution of horizontal subspaces allows us to compare tangent vectors in different fibers, i.e., leads to a notion of <u>parallel transport.</u>

This definition of a connection in terms of horizontal subspaces of T(P) is easy to explain, but hard to compute with. Therefore it is more convenient to introduce a dual definition, in terms of differential forms on P, which pick out the vertical component of tangent vectors and vanish on horizontal vectors (we remind the reader that this is the only way of defining a direct sum decomposition in a vector space without metric). More precisely, a <u>connection</u> on P is defined (globally) by a one-form ω on P with values in the Lie-algebra G such that for a vector $X \in T_p(P)$ we have

$$\omega(X) = A \in G \text{ where } A^*(p) = Ver_p X \ , \qquad (4.1)$$

i.e., the value of the one-form on the vector X equals the element of the Lie algebra which is isomorphic to the vertical part of X (the part tangent to the fiber). The horizontal subspace $Hor_p(P)$ of $T_p(P)$ is then the kernel of the one-form ω, i.e. $\omega(Hor_p(P)) = 0$. The invariance of the horizontal space under the right action of G implies that for any vector $X \in T_p(P)$

$$\omega(R_{g*}X) = Ad(g^{-1})\,\omega(X), \quad \text{or} \quad R_g^{*}\omega = Ad(g^{-1})\,\omega . \tag{4.2}$$

We remind the reader that R_g means the right action of g on a manifold, R_{g*} denotes the tangent (derivative) linear map of this action on vector fields, whereas R_g^* denotes the pullback of this action to differential forms. $Ad(g^{-1})\omega = g^{-1}\omega g$ if both g and ω are interpreted as matrices.

The existence of a connection establishes an isomorphism between $Hor_p(P)$ and $T_x(M)$, where $x = \pi p$, and thus the connection defines a lift to T(P) of any vector field X on M, denoted by X^* and called the horizontal lift of X. Similarly, any curve in the base space M can be lifted into a horizontal curve in P, i.e., a curve which has the horizontal lifts of the tangent vectors as tangents. This allows one to establish a correspondence between points in different fibers along a curve, correspondence which is called parallel transport of fibers. Indeed, starting from a point p_o in the bundle, the horizontal lift of a curve $\gamma = x_t$, $0 \leq t \leq 1$, (assumed smooth, or piecewise smooth and continuous) in M defines a curve γ^* in P, with end point p_1. As p_o varies over the fiber P_{x_o}, p_1 will vary over the fiber P_{x_1}, and the horizontal curves establish an isomorphism $\gamma : P_{x_o} \to P_{x_1}$ between fibers, obviously commuting with the right action of G on fibers.

The reader familiar with the theory of Lie groups will have noticed a certain similarity between the connection form ω and the left-invariant 1-form θ on a Lie group; in particular, the left-invariant form θ on G transforms according to (4.2) under the right action of G on itself. Moreover, the form θ satisfies the Maurer-Cartan structure equation

$$d\theta + \frac{1}{2}[\theta,\theta] = 0 \, . \qquad (4.3)$$

The left-hand side of this equation measures the deviation obtained by parallel transport around an infinitesimal parallelogram (spanned, e.g., by two vectors X, Y tangent to G), i.e., it shows that a Lie group is flat with respect to this parallel transport. Since the bundle consists of fibers isomorphic to G, it makes sense to calculate an expression of the form (4.3) for the connection form, expression which will measure by how much a connection "differs from being a Maurer-Cartan form". This leads to the definition of the curvature two-form on P:

$$\Omega \underset{\text{def}}{=} d\omega + \frac{1}{2}[\omega,\omega] = D\omega \, . \qquad (4.4)$$

Here D denotes covariant exterior differentiation of vector-valued forms defined by

$$D\alpha = \text{hor } d\alpha \qquad (4.5)$$

where, for any k-form β on P, one has

$$\text{hor } \beta\,(u_1,\ldots,u_k) = \beta(\text{hor } u_1,\ldots,\text{hor } u_k)$$

for any vectors $u_1,\ldots,u_k \in T_pP$. In particular, since Ω is a horizontal form, the Bianchi identity

$$D\Omega = d\Omega + [\omega,\Omega] = 0 \qquad (4.6)$$

follows easily. Both the structure equation (4.4) and the Bianchi identity can be verified by evaluating the right- and left-hand sides on horizontal and vertical vectors. Under the right action of G the two-form Ω transforms according to the adjoint representation of G on G:

$$R_g^*\Omega = \mathrm{Ad}(g^{-1})\Omega \ . \tag{4.7}$$

Although there is a superficial similarity between these concepts and the Yang-Mills fields introduced in Section 2, the most striking difference is the fact that the latter are defined on spacetime, i.e., on the base space of the bundle. A connection one-form ω and its curvature two-form Ω can be pulled down to the base space M of the principal bundle P locally, i.e., over an open set U in M, where a section σ_U defines a local trivialization of the bundle. More precisely, let $\{U_i\}$ be an open covering of M and let $\phi_i : \pi^{-1}(U_i) \to U_i \times G$ be the local trivializations, $\psi_{ij} = \phi_i\phi_j^{-1}$, and $\sigma_i(x) = \phi_i^{-1}(x,e)$ the corresponding local sections. Let θ denote the left-invariant (Maurer-Cartan) one-form on G. Then the transition functions ψ_{ij} (which are G-valued on the intersection $U_i \cap U_j$) define a G-valued one-form θ_{ij} on $U_i \cap U_j$ by pullback:

$$\theta_{ij} = \psi_{ij}^*\theta \quad . \tag{4.8}$$

In each open set U_i the section σ_i pulls back the connection form ω to U_i thus defining a family of G-valued one-forms on M

$$A_i = \sigma_i^*\omega \tag{4.9}$$

which are subject to the following "overlap conditions" in $U_i \cap U_j$:

$$A_j = Ad(\psi_{ij}^{-1})A_i + \theta_{ij} \quad . \tag{4.10}$$

In particular, ψ_{ij} may be thought of as the transition function describing the transition from one local trivialization to another $\sigma_i = \sigma_j \cdot \psi_{ij}$; then the forms θ_{ij} are obviously given by the expression $\theta_{ij} = \psi_{ij}^{-1} d\psi_{ij}$, and leaving out the indices i, j, if we operate in a fixed open set, we obtain the law of gauge transformation:

$$A \to A' = Ad(\psi^{-1})A + \psi^{-1}d\psi \quad , \tag{4.11}$$

which agrees with the form given in Sec. 2 for the Yang-Mills potential.

Similarly, the curvature two-form Ω pulls back to a family of G-valued two-forms defined on the open covering $\{U_i\}$ of M:

$$F_i = \sigma_i^* \Omega \quad , \tag{4.12}$$

which under a change of trivialization (or in the overlap of two open sets) transforms under the adjoint action of the transition function:

$$F_j = Ad(\psi_{ij}^{-1})F_i \quad . \tag{4.13}$$

It is easy to verify that the structure equation (4.4) and the Bianchi identity (4.6) "pull down" to the local forms F_i, A_i in the form:

$$F_i = dA_i + \frac{1}{2}[A_i, A_i] \quad , \tag{4.14}$$

$$DF_i = dF_i + [A_i, F_i] = 0 \quad , \tag{4.15}$$

coinciding, respectively, with the definition of the (local)

Yang-Mills field strength and the "homogeneous" Yang-Mills equation (2.1), (2.2).

It should be noted, that if the principal bundle P admits a global section, i.e., is trivializable (which is the case, e.g., if the base space is all of Minkowski space, or any space homeomorphic to an R^n), then the open cover consists of one set, and the forms A, F are globally defined. Conversely, it can be shown that if P is non-trivial, then A and F cannot be globally defined; the reader should keep in mind the example of the Dirac monopole, where A cannot be globally defined, but a connection form can be defined in terms of two pull-backs, to two overlapping sets on S^2, with a gauge transformation on the overlap, or the situation encountered in the Bohm-Aharonov effect.

Finally, it should be noted that parallel transport and covariant differentiation are defined in an obvious manner in any associated vector bundle.

The curvature two-form measures the "nonintegrability" of the connection (or parallel transport) locally, i.e., for transport around an infinitesimal parallelogram in the base space. Parallel transport around an arbitrary loop in the base space leads to the concept of <u>holonomy group</u> of a connection in P.

We consider continuous curves in M which are piece-wise differentiable, and call closed curves starting and ending at a point x <u>loops</u>. The loops based at a point x form a group under the obvious composition, with the zero loop playing the role of identity, and the loop with opposite orientation playing the role of inverse. We denote the group of loops based on the point x in M by L_x, and the subgroup of contractible loops by L_x^o (we assume that M is a connected manifold). Parallel transport of fibers associated with a connection ω in the principal bundle P(M,G) leads to a representation of the group L_x by automorphisms of the fiber P_x, called the holonomy group of P with reference point x, and denoted by H(x). If we restrict the loops to contractible loops, we obtain a subgroup of H(x), denoted by $H^o(x)$ and called the restricted holonomy group at x. Both these groups can be realized as subgroups of the structure group G of P, since the automorphism of the fiber P_x associated to parallel transport around a loop γ can also be realized by the right action of a group element g, with the obvious composition property for successive transport around two loops, the inverse, or the trivial loop, provided one chooses a fixed point p in the fiber above x, where the parallel transport starts. Thus the choice of p and the group of loops L_x (or contractible loops L_x^o) determines a subgroup H(p) (or $H^o(p)$) of G, called holonomy group (or restricted holonomy group) at $p \in P$. Another way of defining H(p) is as that subgroup of elements $h \in G$ such that p and $p \cdot h$ can be joined by a horizontal curve. It can be shown that the holonomy groups H(p) and $H^o(p)$ are actually Lie subgroups of G, that $H(p)/H^o(p)$ is discrete (countable), and that the holonomy groups based at different points x or p are conjugate of each other.

Let P(M,G) be a principal bundle (with M, as always, connected, paracompact) with a connection ω, and p_o an arbitrary point of P. Then the set of points in P which can be joined to p_o by horizontal curves coincides with the restriction of P(M,G) to the subgroup $H(p_o)$ of G; this set is denoted by $P(p_o)$ and is called the holonomy bundle at p_o. Moreover, the connection G is reducible to a connection in $P(p_o)$, i.e., the Lie algebra G of G admits a decomposition into a direct sum G = H + M, where H is the Lie algebra of $H(p_o)$, and the H-component of ω restricted to $P(p_o)$ is a connection form on $P(p_o)$. In other words, the horizontal subspace of $T_q(P)$ is tangent to

$P(p_o)$ for every $q \in P(p_o)$. It is clear that the points in P lie either in the same holonomy bundle, or else their holonomy bundles are disjoint, i.e., the bundle space P is decomposed into a disjoint union of its holonomy bundles, and these are taken into each other by those elements of G which are not in $H(p_o)$. For this reason we may always consider that a gauge group which survives a symmetry breaking is a holonomy group of the bundle (or by abuse of language, a holonomy group of the vacuum).

It is clear that, since the curvature Ω of the connection measures the "infinitesimal holonomy", there must be a relation between curvature and holonomy. There are various theorems establishing such relations, the most important one being the theorem of Ambrose and Singer. This theorem states that in a principal bundle $P(M,G)$ over a connected manifold M, with connection ω and curvature Ω, the Lie algebra of the holonomy group $H(p)$ with reference point p is the subspace of G spanned by all elements of the form $\Omega_v(X,Y)$ where v is a point in the holonomy bundle $P(p)$ and X,Y are arbitrary horizontal vectors at v.

Let $E(M,V,P,G)$ be a vector bundle associated with P and $s : M \to E$ a section of E, which can be represented by an equivariant function $f : P \to V$. Then parallel transport of sections can be defined by the parallel transport in P. In particular, each element of the holonomy group at p is represented by a linear transformation on V, and in terms of the vector-valued function f one can represent this action by a matrix defined on the horizontal lifts of loops in M, acting on the function f, or alternatively, introducing local bases of sections, as product-integrals of matrices around loops (or path-ordered exponentials).

5. MORE EXAMPLES. UNIVERSAL BUNDLES AND UNIVERSAL CONNECTIONS

Before proceeding to a discussion of symmetries of connections we give in this section several examples of principal bundles which seem to be less familiar to the physics community. Some of these examples (Stiefel manifolds fibered over Grassmann manifolds) have a "universal" character, meaning that any bundle with an orthogonal or unitary structure group can be obtained from such a bundle as a pullback of a mapping of the base space into the appropriate manifold (more precisely, a homotopy class of such mappings, since homotopically equivalent maps lead to isomorphic bundles). We also mention briefly the existence of "universal connections", a construction which will certainly find many physical applications in the near future.

One of the simplest examples of principal bundles can be obtained if one considers a Lie group G with two nested closed subgroups H, K ,

$$K \subset H \subset G \,, \tag{5.1}$$

where K is an invariant (normal) subgroup of H, and such that G/K and G/H are differentiable manifolds. Then it is easy to see that the projection

$$\pi : G/K \to G/H \tag{5.2}$$

yields a principal bundle over G/H with structure group H/K. The projection is the mapping which associates to each left coset of K in G the left coset of H in G which contains it (if K is not a normal subgroup, (5.2) is still a bundle, with fibers the homogeneous space H/K).

In this case, if G is semisimple and the pair G, H is reductive, i.e., $G = H \dot{+} M$ (direct sum) and $[H,M] \subset M$, then the canonical left-invariant form θ on G has a H-component which projects to an H/K connection on G/K (this result is a special case of a theorem by Narasimhan and Ramanan, which was used in a theorem by Harnad on invariant gauge fields, cf. also Nowakowski and Trautman, 1978).

This construction becomes more transparent if we specialize the groups G, H, K, to be orthogonal, unitary or symplectic groups, which act, respectively, on R^n, C^n, or H^n (H denotes the quaternions). For simplicity, we discuss only the real case, but a simple change of notation (replacing R by C or H, SO(n) by U(n), or Sp(n)) yields the results in the complex or quaternionic cases. It should be noted that we could have started from the general linear groups, but there are theorems stating that the bundles defined by the maximal compact subgroups suffice, i.e., the restriction of GL(n,R) to O(n) (or SL(n) to SO(n)) is always possible.

The real Stiefel manifold $S_{n,k}(R)$ (we omit the R in the sequel) is defined as the manifold of all orthonormal k-frames ($0 \leqq k \leqq n$) in R^n, i.e., the set of all linear isometric mappings of R^k into R^n, defined by the set of $n \times k$ matrices with orthonormal rows. To see that this is a manifold, embed the matrices in R^{nk} and verify that the gradients of the orthonormality conditions are nonzero and mutually orthogonal, thus defining a compact manifold without boundary of dimension $nk - k(k+1)/2$. $S_{n,n}$ is clearly identical to the manifold O(n), and $S_{n,0}$ is a point, whereas $S_{n,1}$ is S^{n-1} and $S_{n,2}$ is the submanifold of TS^{n-1} consisting of unit tangent vectors. It is somewhat harder to show that SO(n) is diffeomorphic to $S_{n,n-1}$ (the diffeomorphism adds to the $n \times (n-1)$ matrix in $S_{n,n-1}$ a column which makes it into an SO(n)-matrix, i.e., completes the orthonormal n-1-frame to an n-frame).

The subgroup of $O(n)$ which stabilizes the orthonormal k-frame is isomorphic to $O(n-k)$, hence $S_{n,k}$ can be interpreted as the coset space $S_{n,k} = O(n)/O(n-k)$. We could have started from oriented orthonormal frames, and then we would have obtained $S_{n,k} = SO(n)/SO(n-k)$. (The reader should be warned that the complex Stiefel manifold $S_{n,k}(C) = U(n)/U(n-k)$ is a real $2nk - k^2$ dimensional compact manifold without boundary which is not a complex manifold in the technical sense; the quaternionic Stiefel manifold $S_{n,k}(H) = Sp(n)/Sp(n-k)$ has real dimension $4nk-(2k^2-k)$ and is also compact without boundary).

The real <u>Grassmann manifold</u> $G_{n,k}(R)$ $(0 \leqq k \leqq n)$ is the set of all k-planes (k-dimensional subspaces) through the origin of R^n. By considering an orthonormal frame in such a subspace, and its complementary (n-k)-frame one can see that the Grassmann manifold admits charts which map its open sets into the space $R^{k(n-k)}$ of real $k \times (n-k)$ matrices. In fact, it is clear that $G_{n,k}$ is a homogeneous space: each k-plane can be considered as a coset in the group $O(n)$ with respect to the subgroup $O(k) \times O(n-k)$,

$$G_{n,k}(R) = O(n)/(O(k) \times O(n-k)) \ .$$

If we replace k-planes by oriented k-planes in the above definition, we obtain the oriented Grassmann manifolds $G^+_{n,k}(R) = SO(n)/(SO(k) \times SO(n-k)) = O(n)/(O(k) \times SO(n-k))$. Important special cases are: $G_{n,0} = G_{n,n} =$ point; $G_{n,1}(R) = RP^{n-1}$, the n-1 dimensional real projective space through the origin (which is conveniently parametrized by means of the homogeneous coordinates $(x_1/x_j, \ldots, x_{j-1}/x_j, x_{j+1}/x_j \ldots x_n/x_j))$.

The generalizations to the complex and quaternionic cases are obvious:

$$G_{n,k}(C) = U(n)/(U(k) \times U(n-k)), G_{n,k}(H) = Sp(n)/(Sp(k) \times Sp(n-k)). \tag{5.3}$$

The construction given at the beginning of this section now can be applied to nested groups $O(n) \supset O(k) \times O(n-k) \supset O(k)$ or their unitary or symplectic counterparts. This leads to a particularly simple principal bundle with structure group $O(k)$ (respectively, $U(k)$, $Sp(k)$):

$$\xi : S_{n,k} \xrightarrow{\pi} G_{n,k} , \qquad (5.4)$$

where the projection associates to each k-dimensional frame the k-plane spanned by it. Rewriting this in the form

$$O(n)/O(n-k) \xrightarrow{\pi} O(n)/(O(k) \times O(n-k)) \qquad (5.5)$$

it is clear that the fiber of this bundle is $O(k)$, and that $O(k)$ acts on the homogeneous space on the right by acting on $O(n)$ (remember that $S_{n,k}$ consists of left cosets, and that $O(k)$ is isomorphic to the factor group $O(k) \times O(n-k)/O(n-k)$). The principal bundle ξ is in a certain sense universal, in that for a given $O(k)$ and large enough n any $O(k)$-principal bundle over a compact manifold can be obtained as a pullback of a mapping of that manifold into $G_{n,k}$ (up to homotopy, which leads to an isomorphic bundle). The same construction holds in the complex and quaternionic cases.

The bundle (5.4) also has a canonical connection - called a universal connection .(The existence of universal connections was proved by Narasimhan and Ramanan, and a recent improvement was found by Roger Schlafly; since they involve embeddings in spaces of high dimension, they may be of use in the discussion of the large N limit of SU(N) gauge theories, and may have other physical applications.) Starting from the n × n matrix defined by an orthonormal frame in R^n, an element of $O(n)$ which we denote by X, we have the left-invariant Maurer-Cartan form on $O(n)$:

$$\theta = X^{-1}\, dX = -{}^t\theta \quad . \tag{5.6}$$

If we restrict the skew-symmetric matrix θ to the $k \times k$-matrix corresponding to the k-frames of the Stiefel manifold $S_{n,k}$, we obtain a connection on the bundle (5.4). Whereas the connection θ is flat, on account of the Maurer-Cartan structure equation (4.3), its restriction ω to the first k rows and columns will have a nonvanishing curvature, as can be seen explicitly by rewriting (4.3) and (5.6) in a basis adapted to $S_{n,k}$ (replacing X^{-1} by ${}^tX$).

The Narasimhan-Ramanan theorem considers an arbitrary compact Lie group G embedded in O(k). Then the canonical connection on $S_{n,k}$ can be pulled back into a connection on the principal bundle with structure group G:

$$O(n)/O(n-k) \to O(n)/(G \times O(n-k)) \quad . \tag{5.7}$$

This yields a canonical connection on this canonical G-bundle. Now for an arbitrary manifold M of dimension $\dim M \leqq m$, and for $n \geqq \frac{1}{2}((k+m)^2 + 7(k+m) + 10)$ any principal G-bundle with connection over M may be obtained as the pullback of the canonical G-bundle (5.7) and its canonical connection by some smooth map $f : M \to O(n)/(G \times O(n-k))$. Schlafly has extended this theorem to the case when the base space is a compact Riemannian manifold with an isometric action of a compact Lie group H, which also acts on the principal bundle space P and preserves the connection (see next section for the meaning of the last statement). He has shown that in this situation there always exists a representation $r : H \to O(n)$ and an H-map $f{:}M \to O(n)/G \times O(n-k)$, such that the bundle is isomorphic to the bundle induced by the pair of maps f, r.

We now turn to the promised discussion of symmetries an invariance properties of connections on principal bundles.

6. INVARIANCE AND SYMMETRIES OF CONNECTIONS

In spite of the fact that conditions for the invariance of a connection have been discussed in the mathematical literature over twenty years ago, and Wang's theorem can be found in textbooks, physicists rediscovered them only in 1978-79. This section contains a brief survey of this topic, which has been discussed from a more physical point of view by Jackiw in last year's Schladming lectures.

The problem is quite simple when viewed globally, on the principal bundle; complications arise only when one tries to express the invariance conditions for the connection forms on local trivializations of P.

Before discussing connections we summarize the definitions of gauge transformations to be used. An isomorphism of a principal bundle onto itself is called an automorphism of the bundle. Such an automorphism consists of a pair of diffeomorphisms (u,v) of P and M such that $\pi \circ u = v \circ \pi$ (Eq.(3.1)), and $u(p\cdot g) = u(p)\cdot g$ for all $p \in P$, $g \in G$. An automorphism is called vertical if $v = Id_M$. If we denote the group of all automorphisms (an infinite-dimensional group) by Aut P, the subgroup of all vertical automorphisms $Aut_M P$ is a normal subgroup, the quotient being the group of all diffeomorphisms of M onto itself, i.e., we have the exact sequence of homomorphisms:

$$I \to Aut_M P \xrightarrow{i} Aut\ P \xrightarrow{j} Diff\ M \to I , \qquad (6.1)$$

where i is the canonical injection and $v = j(u)$. If $u \in Aut_M P$, its action is in the fiber and therefore can be implemented by an element U(p) of G such that for any p in P and g in G

$$u(p) = p\cdot U(p),\ U(p\cdot g) = g^{-1}U(p)g \quad . \qquad (6.2)$$

Thus, there is a natural isomorphism of $Aut_M P$ onto the multiplicative group of (smooth) maps $U : P \to G$, subject to the equivariance condition (6.2), or equivalently, to sections of the associated bundle $P \times_{AdG} G$ with fibers G, but the right action replaced by the adjoint action.

The group Aut P (as well as $Aut_M P$) acts on (local) sections of P in the following manner: if $s : V \to P$ (V an open subset of M), then its transform is $s' = u \circ s \circ v^{-1}$. If $u \in Aut_M P$, the subset V of M is left invariant and the section is subject to what a physicist would call a gauge transformation:

$$s'(x) = s(x) \cdot U(s(x)), \quad x \in V \subset M \quad . \tag{6.3}$$

If one deals only with Yang-Mills fields over a flat spacetime (or a Euclidean, compact version thereof) one is thus entitled to identify $Aut_M P$ with the group of gauge transformations (this is the definition adopted by Atiyah, Singer, and many other mathematicians). However, in theories involving gravity, or other structures on spacetime, it is convenient to introduce a further differentiation.

<u>Definition.</u> The <u>gauge group</u> of a theory in which the bundle has some <u>absolute elements,</u> such as the metric tensor of special relativity, or some other structure element of P or M, is the subgroup $\mathcal{G}$ of Aut P such that the diffeomorphism v and the projection preserve the absolute elements of M. The group of <u>pure gauge transformations</u> consists of the vertical automorphisms in $\mathcal{G}$; this group will be denoted by $\mathcal{G}_0 = \mathcal{G} \cap Aut_M P$, it is a normal subgroup of $\mathcal{G}$, and the quotient $\mathcal{G}/\mathcal{G}_0$ in the exact sequence

$$I \to \mathcal{G}_0 \xrightarrow{i} \mathcal{G} \xrightarrow{j} \mathcal{G}/\mathcal{G}_0 \to I \tag{6.4}$$

is the subgroup of Diff M leaving the absolute elements invariant (e.g., if M is Minkowski space, $\mathcal{G}/\mathcal{G}_o$ is the Poincaré group; this corresponds to the necessity of sometimes combining a gauge transformation with a change of Lorentz frame in some calculations).

Invariance of connections under automorphisms of the bundle P is simply expressed as the fact that the pullback of the connection form ω on P by the mapping $u \in \text{Aut } P$, $\omega' = u^*\omega$ is again a connection form on P. If u is a vertical automorphism (in particular, a pure gauge transformation), then

$$\omega' = \text{Ad}(U^{-1}(p))\omega + U^{-1}(p)dU(p) \quad , \tag{6.5}$$

where U(p) is the map defined in Eq.(6.2). We see that the form ω is subject to the usual gauge transformation of a gauge potential (albeit, on P rather than on M). The curvature form Ω' of the pullback $u^*\omega$ is given by the adjoint action of U(p) on the original curvature form:

$$\Omega' = \text{Ad}(U^{-1}(p))\Omega \quad . \tag{6.6}$$

The equations (6.5), (6.6) can easily be pulled down to the forms A, F on the base space given by a locally trivializing section s. Here one can either pull ω back to M by the transformed section, or pull ω' back by the original section, obtaining the usual gauge transformation formulas for A and F:

$$A' = \text{Ad}(S^{-1})A + S^{-1}dS \ , \qquad F' = \text{Ad}(S^{-1})F \ , \tag{6.7}$$

where $S = U \circ s$.

Among the automorphisms of the principal bundle P with a connection ω and the associated bundles carrying the particle fields, symmetries are distinguished by the fact that they preserve the connection ω and the absolute elements of the theory (e.g., they preserve the action, or they modify the Lagrangian density by a divergence). In particular, a symmetry of a gauge theory is a gauge transformation (in the wider sense defined above) which leaves the connection form ω invariant (in addition to the other absolute elements):

$$u^*\omega = \omega; \; u^*\Omega = \Omega \quad ; \tag{6.8}$$

since a nonabelian gauge theory is not completely determined by the curvature, it is not sufficient to require invariance only of the curvature form.

When this condition is pulled back by a local trivialization to the base space, it will usually be formulated as the requirement that the one-form A be unchanged up to a pure gauge transformation, or in other words, a gauge field is invariant under a symmetry, if the symmetry transformation can be compensated by a gauge transformation of the locally trivializing section (this is the formulation given by Bergmann and Flaherty, Trautman, Jackiw, and other authors).

To write the invariance condition (6.8) for the physical fields A, F, we consider first a one-parameter group $u_t : R \to \mathrm{Aut}\, P$ of automorphisms of P. Let Y denote the corresponding vector field on P, and X the projection of Y onto M:

$$X = \pi_* Y \quad . \tag{6.9}$$

The vector field X generates a one-parameter group $v_t =$ $= j(u_t)$ of transformations on M. Let ω be a u_t-invariant connection on P,

$$u_t^*\omega = \omega, \quad u_t^*\Omega = \Omega \quad . \tag{6.10}$$

For an arbitrary point p_o in P the groups u_t, v_t define curves in P, M, respectively:

$$p_t = u_t(p_o) \ , \quad x_t = v_t(\pi p_o) = \pi(p_t) \ . \tag{6.11}$$

The connection defines a horizontal lift of x_t which we denote by h_t. Then it is obvious that $p_t = h_t g_t$ for a suitable element g_t of G, and g_t is a one-parameter Lie subgroup of G, generated by the Lie algebra element $T = \omega_{p_o}(Y)$. The invariance of the connection and its curvature on P can be expressed infinitesimally as the vanishing of their Lie derivatives with respect to Y:

$$L_Y\omega = 0 \ , \qquad L_Y\Omega = 0 \ . \tag{6.12}$$

(Recall that for forms the Lie derivative is defined by $L_Y = d \circ Y \rfloor + Y \rfloor \circ d$, where denotes the interior product of Y with the differential form following the sign.)

The expressions (6.12) for the invariance of connections are identical to the usual conditions for the invariance of fields encountered in physics, but hidden behind the simple form is the gauge freedom inherent in the theory, particularly if one works in terms of the pullbacks A, F, to the base space. If we denote the value of the one-form ω_p (at the point p in P) on the vector field Y at p by $Z = \omega_p(Y)$, we obtain an equivariant map of P into the Lie algebra $Z : P \to G$, $Z \circ R_g = \mathrm{Ad}(g^{-1}) \circ Z$. Its covariant exterior differential

$$DZ = dZ + [\omega, Z] \tag{6.13}$$

is a horizontal one-form (with values in G) of type Ad, and the definition of the Lie derivative and Eq.(6.14) yield the detailed form of the invariance condition:

$$L_Y\omega = Y \lrcorner \Omega + DZ = 0, \quad L_Y\Omega = D(Y \lrcorner \Omega) + [\Omega, Z] = 0 \qquad (6.14)$$

(Trautman, 1979). If we use a local section s to pull back the connection and curvature to the gauge potential A, and the field strength F on M, the vector field Y is to be replaced by the generator X of the transformations in M, and the Lie-algebra-valued function on P, Z, defines a function on M, $\Phi = Z \circ s : M \to G$. Then the invariance conditions for A and F under the symmetry induced on M by the vector field X (such a vector field always has a horizontal lift under the given connection; adding an arbitrary vertical vector field of the type of Z to it, will give a field on P) can be written in the form

$$X \lrcorner F + D\Phi = 0 \qquad (6.15)$$

where $D\Phi = d\Phi + [A,\Phi]$, and

$$D(X \lrcorner F) + [F,\Phi] = 0 . \qquad (6.16)$$

In terms of the potential one-form A the invariance condition can be rewritten as $L_X A = DW(X)$, where $W(X)$ differs from Φ by the zero-form $-X \lrcorner A$. The right-hand side of the last equation has the infinitesimal form of a gauge transformation, and under a change of chart (gauge transformation) with transition functions g_{ij} the function W is subject to the transformation

$$W_j = \mathrm{Ad}(g_{ij}^{-1})\, W_i + g_{ij}^{-1}\, X \lrcorner dg_{ij} \quad . \qquad (6.17)$$

If X_1 and X_2 denote two vector fields on M inducing symmetries of the connection A, then consistency requires that

$$2F(X_1,X_2) = \Phi([X_1,X_2]) - [\Phi(X_1),\Phi(X_2)] \quad , \qquad (6.18)$$

where the left-hand side denotes the value of the two-form

F on the two vector fields X_1, X_2, and the right-hand side expresses the dependence of the G-valued 0-form on the vector field X_i (and implicitly, on the trivializing section s). The infinitesimal forms of the invariance conditions have been independently discovered by Forgacs and Manton, Harnad, Shnider and Vinet, and Jackiw (cf. the bibliography to Mayer's contribution for references), and the usefulness of Eqs.(6.15), (6.18) (with a difference in sign) has been discussed in Jackiw's 1980 Schladming lectures.

To end this section we give, for the convenience of the reader, a brief statement of Wang's theorems on invariant connections, in a notation which is close to the one used by Kobayashi and Nomizu, where the detailed proofs can be found.

Consider, as before, a principal bundle P(M,G), with a connection ω which is invariant with respect to a group of automorphisms K of P(M,G), assumed to be a connected Lie group with fiber-transitive action, i.e., for any two fibers there is an element of K which maps one into the other, hence K acts transitively on the base space M. We denote by u_o a reference point in P, chosen once and for all, and by x_o its projection in M, $x_o = \pi(u_o)$. Furthermore we denote by J the isotropy subgroup of K at x_o, i.e., the subgroup of all transformations in K which leave x_o invariant (it is clear that M can then be viewed as the homogeneous space K/J). We denote the Lie algebras of the groups G, K, J by $\mathfrak{g}$, $\mathfrak{k}$, $\mathfrak{j}$, respectively and, when it exists, the subspace of $\mathfrak{k}$ complementary to $\mathfrak{j}$ by $\mathfrak{m}$: $\mathfrak{k} = \mathfrak{j} \dotplus \mathfrak{m}$ (direct sum). Then we define a linear mapping $\Lambda : \mathfrak{k} \to \mathfrak{g}$ by $\Lambda(X) = \omega_{u_o}(\tilde{X})$, where $X \in \mathfrak{k}$ and $\tilde{X}$ is the vector field on P induced by X, which has the properties

(i) $\Lambda(X) = \lambda_*(X)$ for $X \in \mathfrak{j}$; here λ_* is the homomorphism $\lambda_* : \mathfrak{j} \to \mathfrak{g}$ defined as the differential of the homomorphism

$\lambda : J \to G$, which assigns an element $g \in G$ taking the point u_o into the same point as the left action of $j \in J$: $ju_o = u_o g$: $g = \lambda(j)$;

(ii) for $j \in J$ and $X \in \kappa$, $\Lambda(Ad(j)(X)) = Ad(\lambda(j))(\Lambda(X))$, where $Ad(j)$ is the adjoint action of J on κ and $Ad(\lambda(j))$ is that of G on g. The geometric meaning of these homomorphisms should be clear from our discussion of the lifting of the horizontal projection of any one-parameter group of automorphisms given by Eq.(6.11) and the discussion following it. Note that u_o denotes our previous p_o (and <u>not</u> the value of the automorphism at $t = 0$), and the vertical action $\lambda(j)$ is the same as the previous g_t.

It is easy to verify, by using the definition of curvature (the structure equation), that the curvature form Ω satisfies the condition (from which Eq.(6.18) follows by pullback to M):

$$2\Omega_{u_o}(\tilde{X},\tilde{Y}) = [\Lambda(X),\Lambda(Y)] - \Lambda([X,Y]), \text{ for } X, Y \in \kappa . \quad (6.19)$$

What Wang's theorem asserts is the existence of a bijection between the set of K-invariant connections in P and the set of linear mappings $\Lambda : \kappa \to g$ satisfying the conditions listed above, bijection which is given by

$$\Lambda(X) = \omega_{u_o}(X), \text{ for } X \in \kappa \quad . \quad (6.20)$$

The proof is straightforward and can be found, e.g., in Kobayashi and Nomizu (p.107, with the same notations as here).

It also follows immediately that a K-invariant connection is flat (i.e., has vanishing curvature) iff $\Lambda : \kappa \to g$ is a Lie algebra homomorphism (since then the right-hand side of Eq.(23) vanishes, and hence so does the left-hand side).

Moreover, if in addition the Lie subalgebra $\mathfrak{j}$ admits a complementary subspace $\mathfrak{m}$ in $\mathfrak{k}$ such that Ad(J)($\mathfrak{m}$) = ($\mathfrak{m}$), then there is a bijection between the set of K-invariant connections in P and the set of linear mappings $\Lambda_{\mathfrak{m}} : \mathfrak{m} \to \mathfrak{g}$, such that for $X \in \mathfrak{m}$, $j \in J$ we have $\Lambda_{\mathfrak{m}}(Ad(j)(X)) = Ad(\lambda(j))(\Lambda_{\mathfrak{m}}(X))$, with the bijection given in terms of the Λ defined above by $\Lambda(X) = \lambda(X)$ if $X \in \mathfrak{j}$, and $\Lambda(X) = \Lambda_{\mathfrak{m}}(X)$ if $X \in \mathfrak{m}$. The curvature form of the K-invariant connection defined by the linear mapping $\Lambda_{\mathfrak{m}}$ satisfies the following condition:

$$2\Omega_{u_o}(X,Y) = [\Lambda_{\mathfrak{m}}(X),\Lambda_{\mathfrak{m}}(Y)] - \Lambda_{\mathfrak{m}}([X,Y]_{\mathfrak{m}}) - \lambda([X,Y]_{\mathfrak{j}}), \qquad X,Y \in \mathfrak{m},$$

where the subscripts on the brackets denote components in the corresponding subspaces of the algebra $\mathfrak{k}$ where the bracket is originally defined. If $\Lambda_{\mathfrak{m}} = 0$ then the corresponding invariant connection is called the <u>canonical invariant connection</u> with respect to the decomposition $\mathfrak{k} = \mathfrak{j} \dot{+} \mathfrak{m}$. Physically, this corresponds to choosing the gauge functions Z and the connection A in eqs.(6.13) - (6.18) so that the components of Φ in the subspace $\mathfrak{m}$, corresponding to the given decomposition, should vanish.

It is to be noted that the existence of a complementary subspace $\mathfrak{m}$ invariant under the adjoint action of J is equivalent to the reductivity of the homogeneous space K/J = M, a rather restrictive condition on the base space M.

Finally, it should be noted that the Lie algebra of the holonomy group of a K-invariant connection at u_o is defined by a sum of iterated brackets of $\Lambda(\mathfrak{k})$ with the subspace $\mathfrak{m}_o$ of $\mathfrak{g}$ spanned by the right-hand side of eq. (6.19) (for details we refer the reader again to Kobayashi-Nomizu, p.110-111).

J. Nowakowski, A. Trautman, Natural Connections on Stiefel Bundles are Sourceless Gauge Fields, J. Math. Phys. 19 (1978) 1100.

R. Schlafly, Universal Connections, Inventiones math. 59 (1980) 59-65.

V.A. Rokhlin, D.B. Fuks, Nachal'nyi kurs topologii (Introductory course in topology), Nauka, 1977.

Printed in Austria