Roberto Percacci

GEOMETRICAL ASPECTS OF NONLINEAR SIGMA MODELS
GEOMETRICAL ASPECTS OF NONLINEAR SIGMA MODELS

Roberto Percacci
Istituto di Fisica Teorica dell'Università di Trieste and
Istituto Nazionale di Fisica Nucleare, Sezione di Trieste

Abstract: Like general relativity and gauge theories, the nonlinear sigma models are examples of field theories with a geometric interpretation; their study involves the use of many ideas from the theories of harmonic maps, Lie groups, symmetric spaces, fibre bundles and differential geometry in general. The present work is intended to give an account of some of these aspects, together with an introduction to the mathematical tools which are necessary to understand them. It is focused essentially on two problems: derive various equivalent forms for the action functional, and give a global definition of nonlinear sigma model as a field theory defined on a nontrivial fibre bundle over spacetime. This broader definition allows the fields to have a "state of twistedness" and thus enlarges considerably the class of topological configurations of the model.
### Summary

#### SECT. 1 PRELIMINARIES

1.1 Plan of the paper  
1.2 Introduction  

#### SECT. 2 A SHORT INTRODUCTION TO FIBRE BUNDLES

2.1 Bundles  
2.2 Groups of transformations  
2.3 Fibre bundles  
2.4 Coordinates in a fibre bundle  
2.5 Principal bundles, associated fibre bundles  
2.6 Prolongation and reduction  
2.7 Example: bundle of bundles  
2.8 Equivariant mappings & cross-sections  
2.9 Connections in a principal bundle  
2.10 Connection form, covariant derivative  
2.11 Connection in an associated fibre bundle  
2.12 Covariant derivative of sections  
2.13 Metric in a principal bundle  

#### SECT. 3 LOCAL DEFINITION OF THE NONLINEAR SIGMA MODELS

3.1 Field configurations  
3.2 The action for harmonic mappings  
3.3 Harmonic sections of a trivial bundle  
3.4 Summary of subsequent results  
3.5 Some facts about Lie groups  
3.6 Homogeneous riemannian manifolds  
3.7 Sigma models with gauge symmetry  
3.8 The principal field  
3.9 Reductive homogeneous space-valued field  
3.10 Some properties of the classical groups  
3.11 The Stiefel manifolds  
3.12 The Grassmann manifolds  
3.13 The canonical connection in the Stiefel bundles  
3.14 The Grassmannian sigma models  

#### SECT. 4 GLOBAL DEFINITION

4.1 Global field configurations  
4.2 Yang-Mills field on a principal bundle  
4.3 Gauge transformations of the sigma field  
4.4 The action for harmonic sections  
4.5 Symmetry breaking  
4.6 The related Higgs model  
4.7 Example  

#### REFERENCES
1.1 In the effort of making this paper fairly self-contained, I have given in sect. 2 all the fundamental definitions and properties of fibre bundles which will be used. The treatment is elementary and proofs have been skipped; the reader is referred to the classical texts of Steenrod\textsuperscript{(1)}, Husemoller\textsuperscript{(2)} and Kobayashi & Nomizu\textsuperscript{(3)} for more details (the latter for the theory of connections).

The third section contains essentially a compilation of the various forms of the action functional for several specific cases, the most general form being that of the action for harmonic maps. In this section we will be dealing only with local properties, i.e., the fields will be defined on a (pseudo)-riemannian space homeomorphic to an open ball in $\mathbb{R}^n$. In the fourth section a global definition of nonlinear sigma model will be given, involving the introduction of Yang-Mills fields; it will be shown that any solution leads to a spontaneous symmetry breaking.

1.2 The prototype physical theory with a geometrical interpretation is without doubt general relativity. In recent years, however (see the work of Wu and Yang\textsuperscript{(4)}) it has become clear that another fundamentally important class of physical theories, gauge theories, also have a geometrical meaning, the gauge potentials being identified with local representatives of a connection in a principal bundle (a brief review of this will be given in § 4.2).

It is not hard to understand the fundamental reason why fibre bundles have to do with gauge theories. In gauge theories one has quantities which are defined only modulo a spacetime-dependent
action of the gauge group. A fibre bundle is a formalization of this idea: roughly speaking, it is a space $E$ which is obtained by glueing a copy of a space $Y$ (the typical fibre) at each point of a space $X$ (the base space) and then forgetting where the origin of $Y$ is. More precisely, if $G$ is a group acting on $Y$, the fibre over $x \in X$ is a subspace of $E$ homeomorphic to $Y$ modulo the action of $G$. This is precisely what happens in gauge theories, where $X$ is spacetime and $Y$ is some "internal space" (for instance, the space of a field on which $G$ acts, or the group $G$ itself).

The scope of the present paper is to give an account of the geometrical aspects of still another class of physical theories: the nonlinear sigma models.

The prototype of these is the so-called $O(3)$ nonlinear sigma model, an $S^2$-valued theory motivated by the fact that the locus of minima of the potential for a certain linear ($\mathbb{R}^3$-valued) sigma model is a sphere: the linear model goes over to the nonlinear one in the limit of a very deep potential. More recently $CP^n$ models and Grassmannian models made their appearance and were thoroughly analyzed. Their interest lies mainly in the fact that they can be formulated in such a way as to mimick a gauge theory, and there is a very deep analogy between these model in two-dimensional spacetime and Yang-Mills theory in four-dimensional spacetime. The ultimate generalization, however, was the recognition that the action functional of all these models is a special case of what mathematicians introduced under the name "energy"; the maps $\varphi: \mathcal{M} \to \mathcal{N}$ that are stationary points for the energy are called harmonic maps. If we now interpret $\mathcal{M}$ as spacetime and $\mathcal{N}$ as space of the fields, we see that nonlinear sigma models are simply $\mathcal{M}$-valued functions over $\mathcal{M}$ and their dynamics is specified requiring $\varphi$ to be harmonic.

In this paper, I will generalize this definition identifying the field configurations with sections of a (possibly nontrivial)
bundle $\Sigma$ with base space $\mathcal{M}$ and fibre $\mathcal{N}$: these generalized models could be called twisted nonlinear sigma models\(^{(20)}\). The possibility of having twisted field configurations depends on the topology of spacetime and on the structure group of the bundle; in fact there is a classification of fibre bundles in homotopy language which, for four dimensional base space, can be turned into a more practical cohomological classification \(^{(21)}\). For instance, if $\mathcal{M}$ is homeomorphic to $\mathbb{R}^4$, then it supports only one bundle with given fibre $\mathcal{N}$, which is the product $\mathcal{M} \times \mathcal{N}$; in this case the earlier definition of a field configuration as a map $\varphi: \mathcal{M} \rightarrow \mathcal{N}$ would be sufficient, but if spacetime had a more complicated topology, the twisted field configurations would represent independent degrees of freedom which could not be ignored.

Another interesting aspect of the generalized definition is that it requires the introduction of a connection in the principal bundle $\Pi$ associated to $\Sigma$, i.e. a Yang-Mills field\(^{(3)}\). We shall see also that in the case when the internal space $\mathcal{N}$ is a homogeneous space $G/H$, a field configuration of the sigma model leads to a spontaneous breakdown of the symmetry from $G$ to $H$ which in our framework shows up as a reduction of the principal $G$-bundle $\Pi$ to a principal $H$-bundle $\Pi'$. It is perhaps interesting to notice here that gravitation also has a structure of this sort: the metric tensor (or better the vierbein) appears as a section of a bundle $\Sigma$ with fibre $GL(4,\mathbb{R})/O(4)$ associated to the bundle of linear frames of spacetime, $L(\mathcal{M})$, and is related to a reduction of $L(\mathcal{M})$ to $O(\mathcal{M})$, the bundle of orthonormal frames\(^{(3)}(17)(21)(22)(30)\). However, the action for general relativity and for the $GL(4,\mathbb{R})/O(4)$ - sigma model

\(^{(3)}\) This Yang-Mills field is not to be confused with the one that appears in the formulation of the sigma models with gauge symmetry (see eq. 3.7.5), which is a composite field and can be eliminated. For instance in the $\mathbb{C}P^n$ models this would be a $U(1)$ gauge field, but the field I am referring to here would be a $U(n)$ gauge field.
are different.

One possible objection to this kind of geometrization (both for the sigma-model and for gauge theories) is that one needs a manifold larger than four-dimensional spacetime to achieve it. There are two possible attitudes in this respect: on one hand, one can think that the fibre bundle is a mere mathematical abstraction; devoid of any physical reality: after all, the gravitational affine connection can be described as a connection in the bundle of linear frames, and usually one does not infer from this that spacetime has 20 dimensions (although, in some sense, some people do). On the other hand one might be tempted to ascribe a real physical significance to the whole bundle space, trying to find some reason why only four dimensions are directly accessible to our everyday experience. Thus one arrives to the Kaluza-Klein idea, that the fibres of the bundle are compact with a characteristic length so small that we can't perceive them; putting in some numbers (essentially Newton's constant and the charge of the electron), the characteristic length turns out to be of the order of Planck's length, $10^{-33}$ cm.

I will not discuss further this amazing idea but rather refer to refs. (24)(25)(26)(27)(28)(29)(30)(31) for the Yang-Mills case and to (15) for a discussion of the bundle $\Sigma$ from this viewpoint.
SECT. 2 A SHORT INTRODUCTION TO FIBRE BUNDLES

2.1 The general reference for paragraphs 2.1 to 2.5 are Steenrod (1) and Husemoller (2). In the former text, fibre bundles are defined ab initio with all their structure, using a coordinate language, while in the latter they are defined, in a coordinate free manner, by restricting progressively a larger class of objects, called bundles. I will follow the second way, but using frequently the coordinate language, which is more intuitive, when it proves more convenient. Other references where the physical content is displayed are (32)(29).

A bundle \( E \) (sometimes denoted \( E \to X \)) is the triple \( (E, \pi, X) \) where \( E \) (the total space) and \( X \) (the base space) are topological spaces and \( \pi \) (the projection) is a map of \( E \) onto \( X \).

If \( x \in X \), the counterimage \( \pi^{-1}(x) \subseteq E \) is called the fibre over \( x \). A cross section of the bundle \( E \to X \) is a map \( s : X \to E \) such that \( \pi \circ s = \text{id}_X \) (the identity of \( X \)).

A bundle \( E' \to X' \) is a subbundle of the bundle \( E \to X \) if \( E' \subseteq E \), \( X' \subseteq X \) and \( \pi' = \pi|_{E'} \) (the restriction of \( \pi \) to \( E' \)). In particular if \( U \) is an open set contained in \( X \), \( \xi|_U = (\pi^{-1}(U), \pi|_{\pi^{-1}(U)}, U) \) is a subbundle of \( \xi \), called the restriction of \( \xi \) to \( U \). The interesting maps between bundles are those that preserve the bundle structure, i.e. send fibres into fibres. This motivates the following definition: a bundle morphism (or a bundle map) of a bundle \( E \to X \) to a bundle \( E' \to X' \) is a couple of maps \( (\omega, \xi) \), \( \omega : E' \to E \), \( \xi : X \to X' \) such that the following diagram commutes:

\[
\begin{array}{ccc}
E & \xrightarrow{\omega} & E' \\
\downarrow{\pi} & & \downarrow{\pi'} \\
X & \xrightarrow{\xi} & X'
\end{array}
\]
Since we have supposed $\pi$ to be surjective, $f$ is completely determined by $U$, but nevertheless we will denote the bundle morphism with the couple $(u, f)$.

$(u, f)$ is a bundle isomorphism if there exists another bundle morphism $(v, f')$ with $v: E' \to E$, $f': X' \to X$ such that $u' \circ v = \text{Id}_E$, $v \circ u' = \text{Id}_{E'}$ (and consequently $f' \circ f = \text{Id}_X$, $f \circ f' = \text{Id}_{X'}$).

An isomorphism of the form $(u, \text{Id}_X)$ is called an $X$-isomorphisms.

Finally, isomorphisms of a bundle to itself are called automorphisms and $X$-isomorphisms of a bundle to itself are called, for reasons that will become clear in §2.9, vertical automorphisms (this term applies only to fibre bundles).

A product bundle is a bundle of the form $\xi = (X \times Y, \pi_x, X)$; a bundle which is $X$-isomorphic to a product bundle is said to be trivial. Notice that a subbundle of a trivial bundle need not be trivial. If $\xi = (E, \pi, X)$ and $f: X' \to X$, one can define the bundle $\xi \times f$ induced from $\xi$ through $f$ in the following way:

$$\xi \times f = (E', \pi', X')$$

where $E' = \{(x', p) \in X' \times E \mid f(x') = \pi(p)\}$ and $\pi': (x', p) \mapsto x'$. In other words, the following diagram must commute

$$
\begin{array}{ccc}
E' & \xrightarrow{f} & E \\
\downarrow{\pi'} & & \downarrow{\pi} \\
X' & \xrightarrow{f} & X \\
\end{array}
\quad \quad
\begin{array}{ccc}
\xi & \xrightarrow{f} & \xi \\
\downarrow{\pi} & & \downarrow{\pi} \\
X & \xrightarrow{f} & X \\
\end{array}
$$

(2.1.2)

### 2.2

Let us recall here some fundamental facts about Lie groups of transformations, which will be used in the following. Given a topological group $G$ and a topological space $Y$, we say that $Y$ is a right $G$-space (resp. left $G$-space) if there is an open map $R: Y \times G \to Y$ (resp. $L: G \times Y \to Y$) such that
\[ R_a \circ R_b = R_{ba} \quad (L_a \circ L_b = L_{ab}) \]
\[ R_e = \text{Id}_Y \quad (L_e = \text{Id}_Y) \] (2.2.1)

where we have defined \( R_a : Y \to Y \) by \( R_a(y) = R(y,a) \) (resp.
\( L_a : Y \to Y \) by \( L_a(y) = L(a,y) \)). In the following
we shall often denote the image of \( y \) under \( R_a \) (\( L_a \)) by \( ya \) (\( ay \)).
It is to be noted that if \( Y \) is a left \( G \)-space, then we can define
on it also a right \( G \)-space structure (and vice-versa) by

\[ L_g(y) = L^{-1}_g(y) \quad (L_g(y) = R^{-1}_g(y)) \] (2.2.2)

The following examples will arise frequently in this work. A group
\( G \) acts transitively and freely on itself, both from the left and
from the right. If \( H \) is a closed subgroup of \( G \) and \( G/H \) is
the space of left cosets of the form \( gH \), then \( G \) acts transi-
tively on \( G/H \) from the left by \( L_a(gH) = (ag)H \) and \( H \) acts tri-
vially on \( G/H \) from the right. Notice that \( H \), being a subgroup
of \( G \), acts on \( G/H \) also from the left, nontrivially. Let \( H_0 \)
be the intersection of all subgroups \( gHg^{-1} \) conjugate to \( H \) in
\( G \); \( H_0 \) is the largest subgroup of \( H \) which is invariant in
\( G \). Using this property, \( h_o g H = g H h_o g H = g H \) for \( h_o \in H_0, g \in G \)
so that \( H_0 \) acts trivially on \( G/H \) also from the left. Conversely
any element \( a \) of \( G \) acting from the left on \( G/H \) as the iden-
tity must belong to \( H_0 : ah = gH \iff ag \in gH \iff a \in gHg^{-1} \).
Therefore, the group $G/H_o$ acts transitively and effectively on $G/H$ from the left.

2.3 A bundle is a fibre bundle with typical fibre $Y$ (for $Y$ a topological space) if

II) the fibre over $x$, $\pi^{-1}(x)$ is homeomorphic to $Y$ for all $x \in X$.

This definition is still too general to be of use in physics, therefore in the course of this paragraph we will progressively restrict the class of objects we are dealing with, imposing more and more structure. First we require that a group acts on the bundle:

a(right) $G$-bundle is a bundle $\xi = (E, \pi, X)$ such that

III) $E$ is a right $G$-space and the action of the group is a vertical automorphism, i.e.

$$\pi \circ R_g = \pi$$

(2.3.1)

$G$ is called the structure group of the bundle $\xi$. If $G$ acts transitively on the fibres, $X$ is homeomorphic to $E \mod G$ (the space of the orbits of $G$ in $E$).

$G$-bundle morphisms are morphisms of the underlying bundle that preserve the group action:

$$E \times G \xrightarrow{u \times \text{id}_G} E' \times G$$

(2.3.2)

or in other words

$$R'_g \circ u = u \circ R_g$$

The last property that we will impose is local triviality:

IV) If $\xi = (E, \pi, X)$ is a fibre bundle with typical fibre $Y$, $\forall x \in X \exists U$ a neighborhood $U$ such that the restriction of $\xi$ to $U$ is $U$-isomorphic to $(U \times Y, \text{pr}_U, U)$.
A local trivialization of the fibre bundle $\xi$ is a homeomorphism $\psi_U : U \times Y \rightarrow \pi^{-1}(U)$ such that $\pi(\psi_U(x,y)) = x$, $\forall y \in Y$.

We have thus defined a topological fibre bundle with typical fibre $Y$ and structure group $G$, as a bundle with the additional properties II), III), IV). Now, in physics one needs to set up differential equations of motion and thus maps have to be not just continuous but also differentiable up to a given order; this motivates the definition of a differentiable fibre bundle with typical fibre $Y$ and structure group $G$, henceforth denoted $\xi = (E, \pi, X; Y, G)$; the definition can be obtained from I), II), III), IV) replacing everywhere topological spaces with differentiable manifolds and homeomorphisms with diffeomorphisms.

Fibre bundle morphisms are bundle morphisms together with a homomorphism between the structure groups (3). In most applications this homomorphism will be the inclusion of a subgroup or the identity; in the latter case (2.3.2) is required to hold.

2.4 Let $\{U_\lambda\}$ be a locally finite open covering of $X$, and $\psi_{U_\lambda} : U_\lambda \times Y \rightarrow \pi^{-1}(U_\lambda)$ be local trivializations of the fibre bundle $\xi$.

Define the maps $\phi_{U_\lambda,x} : Y \rightarrow \pi^{-1}(x)$ by $\phi_{U_\lambda,x}(y) = \psi_{U_\lambda}(x,y)$ and $\pi_{U_\lambda} : \pi^{-1}(x) \rightarrow Y$ with $x \in U_\lambda$ by $\pi_{U_\lambda}(p) = \phi_{U_\lambda,x}^{-1}(p)$, $x = \pi(p)$.

$\pi_{U_\lambda}$ is a local projection onto the fibre, defined in the neighborhood $U_\lambda$. The maps $\pi$ and $\pi_{U_\lambda}$ define the inverse of $\psi_{U_\lambda} : \pi^{-1}(U_\lambda) \rightarrow U_\lambda \times Y$ by $\psi_{U_\lambda}^{-1}(p) = (\pi(p), \pi_{U_\lambda}(p))$. Conditions III) and IV) together imply that the typical fibre itself has to be a $G$-space. The relation between the right action of $G$ on $E$ and the right action of $G$ on $Y$ is

$$R_g(\psi_{U_\lambda}(x,y)) = \psi_{U_\lambda}(x, R_g(y)) \quad (2.4.1)$$

Clearly if $G$ acts trivially, transitively, effectively or freely on $Y$, then so does it also on $E$, and vice-versa.
If $\xi$ is a $G$-bundle, it can be proven that for any pair $(A,B)$, and $x \in U_A \cap U_B$, the map $\Phi_{A,x}^{-1} \circ \Phi_{B,x} = \pi_A \circ \pi_B^{-1} : Y \to Y$ corresponds to the action of an element of $G$, and thus we can define functions $\rho_{AB} : U_A \cap U_B \to G$ by $\rho_{AB}(x) \cdot y = \pi_A \circ \pi_B^{-1}(y)$; these are known as the structure functions of the fibre bundle $\xi$. They have the following properties:

$$
\begin{align*}
\rho_{AB}(x) &= \rho_{AC}(x) \cdot \rho_{CB}(x) & y \in U_A \cap U_B \cap U_C, \\
\rho_{AB}(x) &= (\rho_{AB}(x))^{-1} & y \in U_A \cap U_B, \\
\rho_{AA}(x) &= e & y \in U_A.
\end{align*}
$$

(2.4.2)

The structure functions give a relation between different local trivializations; if $x \in U_A \cap U_B$,

$$
\Psi_A(x,y) = \Psi_B(x,\rho_{AB}(x) \cdot y) (2.4.3)
$$

The structure functions completely fix the topology of the fibre bundle (once the base space and the fibre are given); this is the content of the following Reconstruction theorem:

If $G$ is a transformation group of $Y$, $\{U_a\}$ an open covering of $X$, and $\rho_{AB} : U_A \cap U_B \to G$ is a set of functions satisfying (2.4.2), then there exists a unique bundle with base $X$, fibre $Y$, group $G$ and structure functions $\rho_{AB}$.

The functions $\Psi_A$ can be used to coordinatize the bundle. If $\Psi_A^{-1}_a(x,y) \in U_a \times Y$ and $x^a, y^a$ are the coordinates of the points $x$ and $y$ in some local coordinate system, then we say that $p$ has coordinates $(x^a, y^a)$. While $x^a$ are uniquely defined by the projection $\pi$, $y^a$ are not since the local trivialization is not unique. One way of fixing the trivialization is the following: let $s_A : U_A \to \pi^{-1}(U_A)$ be a local section of $E$ (such sections always exists since $\xi|_{U_A}$ is $U_A$-isomorphic to $U_A \times Y$ and a trivial bundle always admits sections). Then we fix an origin $o$ in $Y$ and require that the coordinates of the section be $(x^a, o)$.
Since is supposed to act transitively on the fibres, any point in can be written as for some and therefore it has coordinates where are the auxiliary functions ( are coordinates in )

\[(g,y)^x = L^x(g^i, y^\beta)\]  

2.5 A principal bundle is a fibre bundle whose typical fibre is the structure group acting on itself by right multiplication. Notice that acts freely on denoting this action, eq. (2.4.1) reads

\[\pi_A \circ R_q = R_q \circ \pi_A\]  

The trivializations of a principal bundle can be defined by means of local sections as in the general case; the origin of eq. (2.4.4) is conveniently taken to be the identity element and in this case the auxiliary functions are just the multiplication table of the group \(G = (g^i)^j = F^i(g^i, g^j)\).

Given a principal bundle as above and a manifold \(Y\) on which \(G\) acts from the left, we can define a fibre bundle \(\xi = (E, \pi, X; Y, G)\) which is said to be associated to \(\pi\) by means of the following (coordinate-free) construction. Define a right action of \(G\) on \(P \times Y\) by \((p, y) \mapsto (p, g^i y)\); the total space of \(\xi\) is taken as the quotient of \(P \times Y\) by this action.

\[E = (P \times Y) \mod G =: P \times \frac{Y}{G}\]  

The natural projection \(\chi: P \times Y \to E\) defined in this way is called a principal map. The projection \(\pi\) is induced from
\[ \chi \text{ and } \pi \text{ by means of the following commutative diagram} \]

\[
\begin{array}{ccc}
P \times Y & \xrightarrow{\chi} & E \\
\downarrow \text{pr}_P & & \downarrow \pi \\
\downarrow & & \downarrow \\
P & \xrightarrow{\pi} & X
\end{array}
\] (2.5.3)

In other words, \( \bar{\pi}(p, y) = \pi(p) \), and its independence from the action of \( G \) on \( P \times Y \) follows from eq. (2.3.1).

The right action \( \bar{R} \) of \( G \) on \( P \) induces an action \( \bar{R} \) of \( G \) on \( E \) by

\[ \bar{R}_g(\chi(p, y)) = \chi(\tilde{R}_g(p), y) \] (2.5.4)

Also, given a covering \( \{U_\lambda\} \) of \( X \), a trivialization \( \psi_\lambda \) of the principal bundle \( \Pi \) induces a trivialization \( \bar{\psi}_\lambda \) of the fibre bundle \( \xi \) through the commutative diagram

\[
\begin{array}{ccc}
\Pi^{-1}(U_\lambda) \times Y & \xleftarrow{\psi_\lambda} & U_\lambda \times G \times Y \\
\downarrow \chi & & \downarrow L \\
\Pi^{-1}(U_\lambda) & \xrightarrow{\bar{\psi}_\lambda} & U_\lambda \times Y
\end{array}
\] (2.5.5)

Here \( L \) stands for the left action of \( G \) on \( Y \). In other words, if \( p \in \Pi^{-1}(U_\lambda) \times Y, x \in U_\lambda \), and \( E \ni w = \chi(p, y) \), then \( w = \bar{\psi}_\lambda(x, g, y) \); this definition is meaningful since if we act on the upper row of the diagram with \( G \) on the right, i.e. we write \( w = \chi(pa, a'y) \) for some \( a \in G \), then \( pa = \psi_\lambda(x, ga) \) and \( w = \bar{\psi}_\lambda(x, ga, a'y) = \bar{\psi}_\lambda(a, a'y) \) independently of \( a \), as it should.

The motivation for the definition of associated bundle is that, as can be easily seen, the structure functions of \( \xi \) and \( \Pi \) are the same, once the covering \( \{U_\lambda\} \) is fixed. Two fibre bundles \( \xi \) and \( \xi' \) having the same base space and structure group are said to be associated if their associated principal bundles are
isomorphic. The relation of being associated is an equivalence relation among fibre bundles and so one can divide the set of all fibre bundles over a space $X$ with structure group $G$ into equivalence classes and in each class, due to the reconstruction theorem, there is a unique (up to $X$-isomorphisms) principal bundle and a unique (up to $X$-isomorphisms) fibre bundle with a given fibre. Therefore, if one wants to classify all fibre bundles over $X$ with structure group $G$, it is sufficient to classify the principal $G$-bundles.

2.6 Let $H$ be a closed subgroup of $G$.

A principal bundle $\mathcal{P}' = (P', \pi', X; H)$ is said to be a reduction of a principal bundle $\mathcal{P} = (P, \pi, X; G)$ if there exists an injective diffeomorphism $\xi: P' \to \xi(P') \subseteq P$ such that $\xi(p'h) = \xi(p')h$, for $p' \in P$, $h \in H$; at the same time, $\mathcal{P}$ is said to be a prolongation of $\mathcal{P}'$. Any principal $H$-bundle admits a prolongation to a principal $G$-bundle; since $H$ acts on $G$ from the left, one has simply to take the fibre bundle with fibre $G$ associated to $\mathcal{P}'$ and regard the structure functions as $G$-valued rather than $H$-valued. On the other hand, a principal $G$-bundle does not in general admit a reduction to a principal $H$-bundle; a necessary and sufficient condition for this to happen will be given by the theorem in § 4.5. We will also see that the reduction of the structure group regards all fibre bundles associated to $\mathcal{P}$.

When $G$ does not act effectively on $Y$, the structure group of the bundle $\xi$ can be reduced to $G/H_0$ (notation of § 2.2) and the principal $G$-bundle $\mathcal{P}$ is reducible to a principal $G/H_0$-bundle. In this case we say that $\xi$ is only weakly associated to $\mathcal{P}$.

Finally, let $Y'$ be a subspace of $Y$ which is invariant
under the action of $\mathcal{G}$, and $\xi'(E', \pi', X; Y, \mathcal{G})$, $\xi = (E, \pi, X; Y, \mathcal{G})$ be associated fibre bundles. Then $E'$ can be regarded as a submanifold of $E$ and

$$R^{\mathcal{G}}(E') \subseteq E'$$

We shall say then that $\xi'$ is a $\mathcal{G}$-invariant subbundle of $\xi'$. (*)

2.7 Let $\mathcal{G}$ be a Lie group and $\mathcal{H} \leq \mathcal{G}$ a closed subgroup, as in § 2.2. The natural map $\mu : \mathcal{G} \rightarrow \mathcal{G}/\mathcal{H}$ assigning to each $g \in \mathcal{G}$ the coset $g\mathcal{H}$ to which it belongs, defines a principal bundle $(\mathcal{G}, \mu, \mathcal{G}/\mathcal{H}; \mathcal{H})$. In the discussion on Stiefel bundles we will need the following generalization. Let $\mathcal{K} \leq \mathcal{H}$ be a closed subgroup of $\mathcal{H}$ and $\mathcal{K}_o$ be the largest subgroup of $\mathcal{K}$ invariant in $\mathcal{H}$. Then

$$\begin{align*}
Y = \mathcal{G}/\mathcal{H} &\xleftarrow{=} \mathcal{G}/\mathcal{K} = Y' \\
(\mathcal{G}, \mu', \mathcal{G}/\mathcal{K}; \mathcal{K}) &\quad (\mathcal{G}, \mu, \mathcal{G}/\mathcal{H}; \mathcal{H}) \quad \text{principal bundles} \\
(\mathcal{G}/\mathcal{K}, \eta, \mathcal{G}/\mathcal{H}; \mathcal{H}/\mathcal{K}/\mathcal{K}_o)
\end{align*}$$

where $\eta : g\mathcal{K} \mapsto g\mathcal{H}$.

We come now to a situation which will arise later in this work and is referred to by Steenrod as a "bundle of bundles". With the notation used previously, let $\mathcal{P} = (P, \pi, X; \mathcal{G})$ be a principal bundle and $\xi = (E, \pi, X; \mathcal{G}/\mathcal{H}, \mathcal{G})$, $\xi' = (E', \pi', X; \mathcal{G}/\mathcal{K}, \mathcal{G})$ the fibre bundles associated to $\mathcal{P}$. If $\{U_\lambda\}$ are coordinate neighborhoods of $\mathcal{P}$ and $\overline{\mathcal{P}}_\lambda, \overline{\mathcal{P}}'_\lambda$ coordinate functions of $\xi$ and $\xi'$ respectively, define the natural map $\nu : E' \rightarrow E$ by

$$\nu(w') = \overline{\mathcal{P}}_\lambda(x, \eta \overline{\mathcal{P}}'_\lambda(w')) \quad w' \in E', \ x = \overline{\pi}'(w') \in U_\lambda$$

(*) This nonstandard terminology is used here in order to make a distinction from the definition of subbundle of § 2.1.
It is easily seen that this definition is independent on the coordinate neighborhood, and that the following diagram commutes:
\[
\begin{array}{c}
E' \\
\downarrow \pi' \\
X \\
\downarrow \pi \\
E
\end{array}
\]

Then, we have the following

**Theorem:** \( \nu \) is a bundle projection, that is \( E' \) can be considered as the total space of a fibre bundle

\[
(E', \nu, E; H/K, H/K_0)
\]  

(2.7.2)

Specializing to the case \( K = \{ e \} \) we obtain the

**Corollary:** the total space of \( \xi \) is \( E = P \mod H \)

Similarly, \( E' = P \mod K \). Thus we can extend the diagram above to

2.8 In the theory of connections and in physical applications it is often convenient to make use of the following

**Proposition:** if \( \xi = (E, \pi, X; Y, G) \) is a fibre bundle associated to \( \Pi = (P, \pi, X; G) \), there is a one-one correspondence between cross-sections of \( \xi \), \( \sigma: X \to E \) and mappings \( \bar{\sigma}: P \to Y \) such that

\[
\bar{\sigma} \circ \tilde{\xi} = L^{-1}_q \circ \bar{\sigma} \quad \forall q \in G
\]  

(2.8.1)

(such mappings are said to be equivariant with respect to \( G \)).
Thus we have a commutative diagram

\[
\begin{array}{ccc}
P \times G & \xrightarrow{\tilde{R}} & P \\
\circ \times \text{id}_G \downarrow & & \downarrow \circ \\
Y \times G & \xrightarrow{\tilde{R}} & Y
\end{array}
\]

(2.8.2)

where \( \tilde{R} \) stands for the right action of \( G \) on \( P \) and \( \tilde{R} \) for the right action of \( G \) on \( Y \) defined in (2.2.2).

If \( \lambda \) is the principal map \( P \times Y \rightarrow E \), define a map

\[ \lambda_p : Y \rightarrow \pi^{-1}(\pi(P)) \subset E \]

by \( \lambda_p(y) = \lambda(p, y) \). Given a section \( \sigma : X \rightarrow E \) a unique equivariant map \( \delta \) is induced by the diagram

\[
\begin{array}{ccc}
P & \xrightarrow{\text{id}_P \times \delta} & P \times Y \\
\pi \downarrow & & \downarrow \lambda \\
X & \xrightarrow{\sigma} & E
\end{array}
\]

(2.8.3)

Thus \( \delta(p) = \lambda_{\pi(p)} \circ \sigma \circ \pi(p) \); the equivariance follows from the fact that \( \lambda_{\pi(p)}(y) = \lambda(p, y) \) and so \( \delta(p)= \lambda_{\pi(p)}(\sigma(\pi(p))) = L^X_{\sigma(\pi(p))}\lambda_p(\sigma(\pi(p))). \)

Conversely given the equivariant map \( \delta : P \rightarrow Y \), a unique cross section of \( E \) is induced by the diagram (2.8.3), which is given by \( \sigma(\pi(p)) = \lambda(p, \delta(p)) \). Using (2.8.1) and \( \lambda(p, x) = \lambda(p, y) \) it is immediate to check that \( \sigma \) does not depend on \( p \) but only on \( X = \pi(p) \).

2.9 Connections in a principal bundle. The general reference for § 2.9 to § 2.12 is Kobayashi & Nomizu (3), vol. I.

Let us consider the space \( T_p(P) \) tangent to a principal bundle \( \Pi = (P, \pi, X; G) \) in \( p \). If \( \{e_i\} \) is a basis for the Lie algebra \( \mathfrak{g} \) of \( G \), the right action of \( G \) on \( P \) defines a set of vector fields in \( P \):

\[
\tilde{e}_i(p) = \frac{d}{dt} \tilde{R}_{\exp(te_i)}(p) \bigg|_{t=0}
\]

(2.9.1)
these are called the fundamental vector fields in $P$. Due to
eq (2.3.1) they are tangent to the fibres and indeed it is clear
that at each point they form a basis for the subspace of $T_p(P)$
which is tangent to the fibre through $p$. Motivated by this, we
define the vertical subspace $V_p \subseteq T_p(P)$ as the kernel of
$t_p: T_p(P) \rightarrow T_{\pi(p)}(X)$ at $p$. Due to the lack of a canonical
projection on the typical fibre $G$, it is clear that there is
no preferred choice of a subspace complementary to $V_p$ in $T_p(P)$.
A connection $\Gamma$ in $P$ is the assignment of a horizontal subspace
$H_p$ at each point such that

\[ a) \quad T_p(p) = V_p \oplus H_p \]

\[ b) \quad \tilde{\gamma}^{(p)}(H_p) = H \quad \forall g \in G \quad (2.9.2) \]

\[ c) \quad H_p \text{ depends differentiably on } p. \]

A connection defines a way to go from one fibre to another
(parallelism). If $\mathbf{u} \in T_x(X)$ and $\mathbf{p} \in \pi^{-1}(x)$, there exists
a unique vector $\mathbf{v} \in T_p(p)$ such that $\mathbf{v} \in H_p$ and $\pi_\mathbf{p}\mathbf{v} = \mathbf{u}$; $\mathbf{v}$
is called the horizontal lift of $\mathbf{u}$. If $c(t)$ is a curve in $X$ with $0 \leq t \leq 1$, $c(0) = x$, $c(1) = x'$ and $p \in \pi^{-1}(x)$, there exists a
unique curve $\tilde{c}(t)$ through $p$ in $P$ with $0 \leq t \leq 1$, $\tilde{c}(0) = p$ such that
the tangent vector $\frac{d}{dt} \tilde{c}(t) \big|_{t=t}$ is horizontal $\forall t$.

The point $p' = \tilde{c}(1)$ is said to be the parallel displacement
of $p$ along $c(t)$ from $\pi^{-1}(x)$ to $\pi^{-1}(x')$. Varying $p$, we
have obtained a map between fibres $\mathcal{T}_c(x, x') : \pi^{-1}(x) \rightarrow \pi^{-1}(x')$ which
is a group isomorphism.

Given the connection $\Gamma$, we can split vectors and forms
over $P$ into their vertical and horizontal parts:

$\mathbf{u} = \text{Ver } \mathbf{u} + \text{Hor } \mathbf{u}$

$\mathbf{u} \in T_p(P)$, $\text{Ver } \mathbf{u} \in V_p$, $\text{Hor } \mathbf{u} \in H_p$

$\text{Ver } \alpha(\mathbf{u}_1, \ldots, \mathbf{u}_q) = \alpha(\text{Ver } \mathbf{u}_1, \ldots, \text{Ver } \mathbf{u}_q)$

$\text{Hor } \alpha(\mathbf{u}_1, \ldots, \mathbf{u}_q) = \alpha(\text{Hor } \mathbf{u}_1, \ldots, \text{Hor } \mathbf{u}_q)$

$\alpha \in \Lambda^q(P)$
Before we go on to define the connection form a digression is in order concerning vector-space-valued differential forms. If \( \mathcal{Y} \) is a vector space, a \( \mathcal{Y} \)-valued \( q \)-form at a point \( p \) in a manifold \( P \) is a totally antisymmetric linear mapping \( \alpha : \bigwedge^q T_p(P) \to \mathcal{Y} \). We shall denote by \( \bigwedge^q (P, \mathcal{Y}) \) the space of \( \mathcal{Y} \)-valued differential \( q \)-forms over \( P \). If \( \{e_i\} \) is a basis for \( \mathcal{Y} \), we may write \( \alpha = \alpha^i e_i \); where \( \alpha^i \) are ordinary real-valued differential forms. \( \mathcal{Y} \)-valued forms have the same properties of real valued forms except that in general there is no exterior product among them; it is possible to define an exterior product if there is a composition law in \( \mathcal{Y} \), as is the case when \( \mathcal{Y} \) is a Lie algebra: in this case the composition law is given by the Lie bracket and
\[
\alpha \wedge \beta = \alpha^i \wedge \beta^j [e_i, e_j] \in \bigwedge^q (P, \mathcal{Y})
\]
Returning now to the case when \( P \) is the total space of the principal bundle \( \Pi \), we shall say that the form \( \alpha \in \bigwedge^q (P, \mathcal{Y}) \) is of type \( q \) if \( \rho : \mathfrak{g} \to \text{GL}(\mathcal{Y}) \) is a representation of \( \mathfrak{g} \) and
\[
\widetilde{R}_q^* \alpha = \rho(q^{-1}) \cdot \alpha
\]
(2.10.1)

Given a connection \( \Gamma \) in \( P \) we can define a unique vertical \( \mathcal{Y} \)-valued one form of type \( \text{Ad} \) on \( P \), called the connection form \( \omega \), by the requirement that it maps a vector \( \xi \in T_p(P) \) to the element of the algebra generating the fundamental vector field that coincides with \( \text{Ver} \xi \) at \( p \). In other words

\begin{align*}
\text{a') } & \quad \omega(e_i) = e_i; \\
\text{b') } & \quad \widetilde{R}_q^* \omega = \text{Ad}(q^{-1}) \omega \quad (2.10.2) \\
\text{c') } & \quad \omega \text{ depends differentiably on } p
\end{align*}
Conversely a one form of this sort defines a horizontal subspace by \( H_p = \ker \omega_p \) at each point. The reader is referred to Kobayashi and Nomizu for the proofs of existence of connection and of the relation between \( \Gamma \) and \( \omega \). Let us now define a basis for vectors and forms in a neighborhood \( \pi^{-1}(U) \). Let \( \{e_\mu\} \) be a local basis in \( U \subset X \) and \( \{e_i\} \) be a basis in \( \mathfrak{g} \) (eventually we can take \( e_i \) the invariant vector fields forming a basis for the algebra); then a basis is induced in \( \pi^{-1}(U) \) consisting of the fundamental vector fields \( \{\tilde{e}_i\} \) and the vectors \( \{\tilde{e}_\mu\} = s_* \{e_\mu\} \) tangent to the constant sections \( s(x) = \psi_\nu(x, q) \). This is the so-called local direct product basis; we will use also another basis which is obtained by taking the horizontal parts of the vectors \( \{\tilde{e}_\mu\} : \)

\[
\tilde{e}_\mu = \text{Hor} \tilde{e}_\mu = \tilde{e}'_\mu - \text{Ver} \tilde{e}_\mu = \tilde{e}'_\mu - B_i^\mu \tilde{e}_i;
\]

where we have defined the connection coefficients:

\[
B_i^\mu(x, q) = \omega^i(\tilde{e}_\mu);
\]

(2.10.4)

When \( \{e_\mu\} \) is a coordinate basis \( \{e_i\} \), it is customary to denote \( \tilde{e}_\mu = D_\mu \). The vectors \( \{\tilde{e}_i, e_\mu\} \) form the so-called horizontal lift basis. Relation (1) implies that the components \( \omega^i \) of the connection form are the duals of the fundamental vector fields \( \tilde{e}_i \):

\[
\omega^i(\tilde{e}_i) = \delta_i^i
\]

Finally, if we take as a basis for horizontal forms the duals of the horizontal lifts of \( \{e_\mu\} \), we can summarize

\[
\langle \tilde{e}_\mu, \tilde{e}_\nu \rangle = \delta^\mu_\nu \quad \langle \tilde{e}_\mu, \tilde{e}_i \rangle = 0 \\
\langle \tilde{e}_i, \tilde{e}_\nu \rangle = 0 \quad \langle \tilde{e}_i, \tilde{e}_j \rangle = \delta^i_j
\]

(2.10.5)
A more detailed description of this and other bases is given in ref. (29). Given a connection $\Gamma$ in $P$ we can define the covariant exterior derivative of a form $\alpha \in \Lambda^j(P,Y)$ as

$$D\alpha = \text{Hor} \; d\alpha$$

(2.10.6)

It can be shown that if $\alpha$ is a $q$-form of type $P$, $d\alpha$ is a $(q+1)$-form of type $P$. In particular

$$\Omega = d\omega$$

(2.10.7)

is a $G$-valued two-form of type $Ad$, called the curvature of $\Gamma$.

2.11 A connection $\Gamma$ in the principal bundle $\Pi = (P,\pi,\chi,G)$ induces a connection $\Gamma_E$ in any fibre bundle $\xi = (E,\pi,\chi,Y,G)$ associated to $\Pi$, that is a splitting of $T_w(E)$, $\forall w \in E$ in vertical and horizontal subspaces. To see this, let $O$ be an arbitrarily chosen point in $Y$, which we will refer to as the origin; since $G$ acts transitively on $Y$, it is always possible to find a point $p \in P$ such that a given point $w \in E$ be the image of $(p,0)$ under the principal map: $w = \chi(p,0)$. Thus, having fixed $O$, we define $\chi_0: P \to E$ by $\chi_0(p) = \chi(p,0)$.

It is clear from the diagram (2.5.3) that $\chi_0$ is a fibre bundle morphism inducing the identity in $\chi$ and with the identity as group homomorphism.

The vertical subspace $V_w \subseteq T_w(E)$ is defined as for the principal bundle as the set of all vectors tangent to the fibre through $w$.

\(^{(\star)}\) In the situation of (2.7.3) $\chi_0: P \to E$ and $\chi': P \to E'$ coincide with the natural maps $\tau$ and $\tau'$.
23

\[ V_w = \ker \pi_* |_w \]

It is easy to see that it is the image of \( V_p \) under \( \pi_0 \). The horizontal subspace \( H_w \subseteq T_w(E) \) is defined as the image of \( H_p \) under \( \pi_0 : \)

\[ H_w = \pi_0_* H_p \]

It is easily seen that this definition does not depend from the choice of \( \sigma \) and that indeed

\[ H_w \cap V_w = T_w(E) \]

Also, a condition similar to b) is satisfied

\[ R_0^* (H_w) = H_{R_0^*(w)} \]

where the action \( R_0^* \) was defined in (2.5.4) and satisfies \( R_0^* \).

As for principal bundles, the horizontal subspaces can be used to define parallelism in the bundle space, i.e. a diffeomorphism among the different fibres in the bundle.

We will denote by \( \overline{T}_c(x,x') : \overline{\pi}^{-1}(x) \rightarrow \overline{\pi}^{-1}(x') \) the parallel displacement along a curve \( c(t), 0 \leq t \leq 1 \) with \( c(0) = x \) and \( c(1) = x' \).

2.12 Unlike sections of principal bundles which, as we shall see, have no direct physical meaning, sections of associated fibre bundles will represent field configurations and thus we will need their derivatives. The derivative of a cross-section is not a well defined object, and we have to define their covariant derivative. Let \( \sigma : X \rightarrow E \) be a section of \( \xi \) and \( \nu \) be a vector tangent to the curve \( C(t) \) in \( C(0) = x \). Varying \( t \), \( \sigma(c(t)) \) is a curve in \( E \) and \( \overline{T}_c(C(t), C(0)) \sigma(c(t)) \) is a curve in the fibre \( \overline{\pi}^{-1}(x) \). We define the covariant derivative of \( \sigma \) at \( x \).
along $\mathfrak{V}$ to be the vertical vector

$$
\mathfrak{V} \sigma |_x = \frac{\partial}{\partial t} \mathcal{F}(c(t), x) \mathcal{V}(c(t)) \bigg|_{t=0} \quad (2.12.1)
$$

If the fibre is a vector space, we can identify the tangent space with the space itself, so the derivative of a section of a vector bundle is another section of the same bundle.

An alternative definition is based on the equivariant map $\mathfrak{V} : \mathcal{P} \to \mathcal{Y}$ associated to the section $\sigma$. The covariant derivative of $\mathfrak{V}$ is defined in analogy to (2.10.6) (which holds if $\mathcal{Y}$ is a vector space)

$$
\mathfrak{V} \sigma = \text{Hor} \mathfrak{V} : H_p \to T_{\mathfrak{V}(p)}(\mathcal{Y}) \quad (2.12.2)
$$

$$
\mathfrak{V} \sigma = \langle \mathfrak{V} \sigma, \mathfrak{V} \sigma \rangle = d\mathfrak{V}(\sigma) \in T_{\mathfrak{V}(p)}(\mathcal{Y}) \quad (2.12.3)
$$

$\mathfrak{V}$ being the horizontal lift of $\mathfrak{V}$. This is related to (2.12.1) by

$$
\mathfrak{V} \sigma = \chi_{\mathfrak{V} \sigma} \mathfrak{V} \sigma \quad (2.12.4)
$$

The proof is the direct generalization of the one of the lemma at p. 115 of vol. I of ref. (3), which was given only for the case when $\mathfrak{V}$ is a $\mathcal{Y}$-valued zero-form on $\mathcal{P}$.

Since transformation properties are often used in physics to characterize certain quantities, we shall defer their discussion until § 4.2, where the physical meaning of the formalism will be clarified.
2.13 In this paragraph we will briefly go through the metric properties of a principal $G$-bundle $\Pi^\pi (P, \pi, X; G)$. Let $\mathfrak{h}$ be a (pseudo)-riemannian structure in $X$ and $\gamma$ an inner product (eventually the negative of the Cartan-Killing form $B$ of eq. (3.5.1)) in the Lie algebra $\mathfrak{g}$ of $G$, and $\omega$ a connection form in $P$. Then we can define a metric $\overline{\gamma}$ in $P$ by

$$\overline{\gamma}_P (\mathfrak{u}, \mathfrak{u}') = g_{\pi(p)} (\pi_* \mathfrak{u}, \pi_* \mathfrak{u}') + \gamma (\omega (\mathfrak{u}), \omega (\mathfrak{u}')) \quad (2.13.1)$$

This is the unique metric in $P$ with the properties that

$$\pi_* : \text{Hor} T_p (P) \to T_{\pi(p)} (X)$$

and

$$\pi_{A*} : \text{Ver} T_p (P) \to T_{\pi_A(p)} (G) \text{ with } x = \pi(p) \in U_A$$

are scalar product-preserving isomorphisms (if we take in $G$ the natural left-invariant metric (3.5.2)) and $\text{Hor} T_p (P), \text{Ver} T_p (P) \forall p \in P$.

Conversely, given a metric $\overline{\gamma}$ and a connection $\Gamma$ in $P$, and imposing these properties, we can generate a metric $g$ in the base space and a metric $\mathfrak{h}$ in the typical fibre by

$$g_x (\mathfrak{u}, \mathfrak{u}') = \overline{\gamma}_P (\mathfrak{u}^*, \mathfrak{u}'^*) \quad \forall \mathfrak{u}, \mathfrak{u}' \in T_x (X), \quad p \in \pi^{-1} (x) \quad (2.13.2)$$

$$\mathfrak{h}_\mathfrak{g} (\mathfrak{u}, \mathfrak{u}') = \overline{\gamma}_P (\mathfrak{U}, \mathfrak{U}') \quad \forall \mathfrak{U}, \mathfrak{U}' \in T_{\mathfrak{g}}$$

where $\mathfrak{u}^*, \mathfrak{u}'^*$ are horizontal lifts of $\mathfrak{u}$ and $\mathfrak{u}'$ and $\mathfrak{U}, \mathfrak{U}'$ are the fundamental vector fields generated by $L_{\mathfrak{u}^*} \mathfrak{u}$, $L_{\mathfrak{u}'^*} \mathfrak{u}'$ through eq. (2.9.1).
SECT. 3 LOCAL DEFINITION OF THE NONLINEAR SIGMA MODELS

3.1 Let $\mathcal{M}$ be a four dimensional manifold endowed with a (pseudo)-riemannian structure $g$, which is to be looked upon as "spacetime" and $\mathcal{N}$ a d-dimensional manifold with a riemannian structure $\mathcal{g}$, which is to be looked upon as the "space of the fields".

In most field theories, $\mathcal{N}$ is further specialized to be a linear vector space: I will call these theories "linear", irrespective of any interaction term in the lagrangian.

In this section, we will deal only with local properties of the fields, therefore we shall restrict our attention to an open set $\mathcal{U} \subseteq \mathcal{M}$, homeomorphic to an open ball in $\mathbb{R}^d$. A field configuration for a nonlinear sigma model is a (smooth) map $\varphi : \mathcal{U} \to \mathcal{N}$. The dynamics of the theory is specified by means of the action (or the energy, if $\mathcal{U}$ is riemannian) $S[\varphi]$, a functional of the fields and their derivatives, from which all properties of the system can in principle be derived. In particular applying the Euler-Lagrange variational principle to the functional $S[\varphi]$ one obtains equations of motion, whose solutions, called classical solutions for the system, are critical points of $S[\varphi]$ in the space of all field configurations. An even more important role is played by the action in the quantum theory of the system.

3.2 Using the map $\varphi$ we can induce on $\mathcal{U}$ the vector bundle $\varphi^* T(\mathcal{N})$, and in this bundle the riemannian structure

(\#) In the current research on nonlinear sigma models, it is often assumed that $\mathcal{M}$ be two-dimensional; the dimensionality of $\mathcal{M}$ will not affect the considerations contained in this paper.
The differential of the map \( \varphi \) can be thought of as a \( \varphi^* T(\mathcal{X}) \)-valued one-form on \( \mathcal{U} \) or as a section of \( T^*(\mathcal{U}) \otimes \varphi^* T(\mathcal{X}) \); this bundle has a natural Riemannian metric which derives from \( g \) on \( T^*\mathcal{M} \) and \( \varphi^* h \) on \( \varphi^* T(\mathcal{X}) \), and which we will indicate \((..)\). In local coordinates \( x^\mu \) (\( \mu = 1, 2, 3, 4 \)) on \( \mathcal{U} \) and \( y^\kappa \) (\( \kappa = 1, 2, \ldots, d \)) on \( \mathcal{X} \);

\[
\begin{align*}
g &= g_{\mu\nu} \, dx^\mu \otimes dx^\nu, \\
h &= h_{\kappa\rho} \, dy^\kappa \otimes dy^\rho, \\
(\varphi^* h)_{\mu\nu} &= \partial_\mu \varphi^* \partial_\nu \varphi^* h_{\kappa\rho}, \\
(d\varphi, d\varphi) &= g_{\mu\nu} \, \partial_\mu \varphi^* \partial_\nu \varphi^* h_{\kappa\rho}(\varphi)
\end{align*}
\]

Let \( \eta = \sqrt{\det g} \, dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^4 \) be the volume element on \( \mathcal{U} \) canonically defined by the Riemannian structure \( g \). Then, we choose the action to be

\[
S_u[\varphi] = \frac{1}{2} \int_\mathcal{U} (d\varphi, d\varphi) \cdot \eta
\]

Requiring that \( \varphi \) extremizes this functional is equivalent to requiring that \( \varphi \) be a harmonic mapping of \( \mathcal{U} \) into \( \mathcal{X} \). Excellent reviews of the theory of harmonic maps are (18) and (19); physical applications are discussed in (17).

3.3 I will repeat now what we did in § 3.1 and 3.2 using a slightly different language: this will serve as a motivation for what we shall do in sect. 4.

(*) In this work I will denote the differential of a map \( \varphi \) either by \( \varphi^* \) or by \( d\varphi \), according to the context.
Consider the trivial bundle $\mathcal{E}_u = (u \times \mathcal{H}, \pi, u)$; a function $\varphi: u \to \mathcal{H}$ defines a section of this bundle $\sigma: u \to u \times \mathcal{H}$ by $x \mapsto \sigma(x) = (x, \varphi(x))$. (\(\sigma\) is sometimes called the graph of $\varphi$.). A section $\sigma$ minimizing $S_u$ (i.e., a section $\sigma$ whose representative $\varphi$ is a harmonic map) will be called a harmonic section of the trivial bundle $\mathcal{E}_u$.

Suppose we can extend the trivial bundle $\mathcal{E}_u$ to the trivial bundle $\Sigma = (\mathcal{M} \times \mathcal{H}, \pi, \mathcal{H})$. Any such bundle admits global cross-sections of the form $\sigma(x) = (x, \varphi(x))$. Thus we can define

$$S[\varphi] = \frac{1}{2} \int_{\mathcal{H}} (d\varphi, d\varphi).$$

In the fourth section we will generalize this to the case when $\Sigma$ is non-trivial.

### 3.4

For the rest of this section, we will look for alternative forms of the action, restricting ourselves to increasingly specialized models.

I will summarize here the results:

- the form (3.2.1) is the most general, being valid for any Riemannian space $\mathcal{H}$; it will be taken as a reference.
- the forms (3.7.4), (3.9.1) holds if $\mathcal{H}$ is a reductive Riemannian homogeneous space.
- the forms (3.14.3-4-5) hold if $\mathcal{H}$ is a Grassmann manifold.

Although constant multiplicative factors in the action are physically unimportant, we will keep track of them in such a way that the various forms all coincide when using the canonical geometrical structures. Before going on with this task, I will devote the following two sections to recalling some properties of Lie groups of transformations.
3.5 The notions in this paragraph can be found in any textbook on differential geometry and Lie groups, such as (35)(36)(37)(38).

Let $G$ be an $n$-dimensional Lie group and $\mathfrak{g}$ its Lie algebra, i.e. the set of all left-invariant vector fields on $G$. By a well-known theorem, $\mathfrak{g}$ is isomorphic to the tangent space to $G$ at the identity, $T_e(G)$; we shall often forget this isomorphism in the following and simply identify the two spaces. The adjoint representation $Ad(G)$, is a representation of $G$ on $\mathfrak{g}$, i.e. a homomorphism of $G$ to $GL(\mathfrak{g})$ (the group of general linear transformations of $\mathfrak{g}$) defined by

$$Ad: \ g \mapsto Ad(g) = (L_gR_g^{-1})_e$$

The kernel of $Ad$ is the center of $G$, denoted $Z$, that is the set of elements of $G$ that commute with all other elements of $G$. The differential of $Ad$ is the adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}$, denoted $ad(\mathfrak{g})$ or simply $\text{ad}(\mathfrak{g})$; if $X, Y \in \mathfrak{g}$

$$\text{ad}(X): Y \mapsto \text{ad}(X)Y = [X,Y]$$

Let $\{e_i\}$ $(i=1,2,\ldots,n)$ be a basis for $\mathfrak{g}$. Then $\text{ad}(e_i) = c_i$ is a basis for the adjoint representation of $\mathfrak{g}$; being elements of $GL(\mathfrak{g})$, the $c_i$'s can be represented by $n \times n$ matrices

$$(C_i)_e^k = c_i^k \quad \text{where} \quad [e_i, e_k] = c_i^k e_k.$$

In the following we will often use normal coordinates in the group manifold. These are defined assigning to a group element $g$ the coordinates of the element of $\mathfrak{g}$ whose exponential is $g$:

$$g = \exp(y^i e_i) \iff y^i = g^i.$$ 

In normal coordinates, one has

$$Ad(\exp(y^i e_i)) = \exp(\text{ad}(y^i e_i)) = \exp(y^i c_i)$$

We can define a canonical (pseudo-) riemannian structure in $G$. 

---

29
in the following way. The Cartan-Killing metric in $\mathfrak{g}$ is defined by

$$\mathcal{B}(X,Y) = \text{Tr} (\text{ad}(X) \circ \text{ad}(Y))$$  \hspace{1cm} (3.5.1)

It is well known that $\mathcal{B}$ is nondegenerate if and only if $\mathfrak{g}$ is semisimple; furthermore, if $\mathfrak{g}$ is semisimple, $\mathcal{B}$ is negative definite if $\mathfrak{g}$ is compact and indefinite if $\mathfrak{g}$ is noncompact (being negative definite on the subalgebra of the maximal compact subgroup and positive definite on the complementary subspace).

If $X,Y \in T_\mathfrak{g}(q)$, we define the canonical left invariant metric as

$$h_\mathfrak{g}(X,Y) = -\mathcal{B}(L_\mathfrak{g}^1 X, L_\mathfrak{g}^1 Y)$$  \hspace{1cm} (3.5.2)

The canonical left invariant (Maurer-Cartan) form on a Lie group $\mathfrak{g}$ is the $\mathfrak{g}$-valued one form $\Theta$ on $\mathfrak{g}$ such that

$$\Theta_\mathfrak{g}(L_\mathfrak{g} X) = X \hspace{1cm} \forall \mathfrak{g} \in \mathfrak{g}$$  \hspace{1cm} (3.5.3)

3.6 Let $\mathcal{X}$ be a space and $\mathcal{G}$ a Lie group acting on $\mathcal{X}$ from the left: $L_\mathfrak{g} : \mathcal{X} \to \mathcal{X}, \forall \mathfrak{g} \in \mathcal{G}$. If $\mathcal{G}$ acts transitively on $\mathcal{X}$ then $\mathcal{X}$ is said to be a homogeneous space. Let $\mathfrak{o}$ be a distinct element of $\mathcal{X}$ and $\mathcal{H}$ be the subgroup of $\mathcal{G}$ leaving $\mathfrak{o}$ fixed; $\mathcal{H}$ is said to be the isotropy subgroup at $\mathfrak{o}$. Then, by a well known theorem (see e.g. (35) p. 111) $\mathcal{X}$ is homeomorphic to $\mathcal{G}/\mathcal{H}$. The distinct point $\mathfrak{o} \in \mathcal{X}$ is the origin of $\mathcal{G}/\mathcal{H}$ i.e. the coset $\mathcal{H}$. The homomorphism of $\mathcal{H}$ into $\text{GL}(T_\mathfrak{o}(\mathcal{G}/\mathcal{H}))$ which assigns to each $\mathfrak{h} \in \mathcal{H}$ its differential at $\mathfrak{o}$
is called the linear isotropy representation. The image of $H$ under this homomorphism, denoted $H^*$, is called the linear isotropy group at $0$. Clearly, if the linear isotropy representation is faithful then $G$ acts effectively on $X$. $G/H$ is said to be a reductive homogeneous space if $G$ admits a decomposition

$$G = H \oplus P \quad H \cap P = \emptyset$$

such that

$$\text{Ad}_G(H)P \subseteq P \quad (3.6.1)$$

where $\text{Ad}_G(H)$ denotes the adjoint representation of $H$ on $G$. $(3.6.1)$ implies

$$[H, P] \subseteq P \quad (3.6.2)$$

and vice versa $(3.6.2)$ implies $(3.6.1)$ if $H$ is connected.

If $G/H$ is reductive, we can identify $P$ with $T_o(G/H)$.

We will restrict our attention to the case when $\text{Ad}_G(H)$ is compact; this surely happens if $H$ itself is compact (vice-versa, $H$ is compact if $\text{Ad}_G(H)$ is compact and faithful, i.e. $H \cap Z = \{e\}$ where $Z$ is the center of $G$).

It is possible to define in $G$ a positive definite inner product $\gamma(\cdot, \cdot)$ which is invariant under $\text{Ad}_G(H)$, such that $P$ is the orthogonal complement of $H$ in $G$ with respect to $\gamma$. As in $(3.5.2)$, $\gamma$ defines a left $G$-invariant and right $H$-invariant metric in $G$, and the restriction of $\gamma$ to $P$ defines a left $G$-invariant metric in $G/H$. If $G$ itself is compact, $\gamma$ can be taken $\text{Ad}_G(G)$ invariant (for example, one can take the negative of the Cartan-Killing form $B$) and thus
define a biinvariant metric in $G$.

A homogeneous riemannian manifold is a homogeneous space with a differential structure and an invariant metric; or in other words, a riemannian manifold with a transitive group of isometries.

3.7 With the notation and terminology introduced in the previous section, let us now consider a nonlinear sigma model with values in some homogeneous space $G/H : \varphi : U \rightarrow G/H$. We can lift $\varphi$ to a map $\tilde{\varphi} : U \rightarrow G$ requiring the following diagram to commute:

$$
\begin{array}{ccc}
G & \xrightarrow{\mu} & G/H \\
\downarrow & & \downarrow \\
U & \xrightarrow{\varphi} & G/H
\end{array}
$$

(3.7.1)

Correspondingly, there is a "tangent" diagram

$$
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{\mathfrak{g}} & T(G) \\
\downarrow & & \downarrow \\
T(U) & \xrightarrow{\mathfrak{d}\varphi} & T(G/H)
\end{array}
$$

(3.7.2)

but $\mathfrak{d}\tilde{\varphi}$ is defined only up to a vertical vector. Thus we need a connection in the bundle $G \rightarrow G/H$. In general a connection always exists, but we will impose one more requirement, namely that it is invariant under left translations of $G$.

Then we have the following (13, p. 103 vol. I)

**Theorem:** a left invariant connection $\omega$ in the bundle $G \rightarrow G/H$ exists if and only if $G/H$ is a reductive homogeneous space. In this case, $\omega$ is given by the $\mathfrak{h}$ -component of the left-invariant Maurer-Cartan form of $G$:
Thus, we can define

the \( H \)-covariant derivative

\[
D\bar{\varphi} = \text{Hor} \, d\bar{\varphi} = d\bar{\varphi} - \text{Ver} \, d\bar{\varphi} = d\bar{\varphi} - L_{\bar{\varphi}^*} \omega(d\bar{\varphi})
\]

(3.7.3)

and (3.7.2) is replaced by

\[
\begin{array}{ccc}
T(\mathcal{U}) & \xrightarrow{d\varphi} & T(\mathcal{G}) \\
\text{Hor} & \downarrow & \mu_* \\
& \downarrow & \\
& T(\mathcal{G}/H)
\end{array}
\]

(3.7.4)

where now \( \mu_* \) is a vector space isomorphism.

Observe now that the metric in the total space of the principal bundle \( \mathcal{Q} \to \mathcal{G}/H \) obtained from \(2.13.1\) is the canonical metric in \( \mathcal{Q} \) if we take the canonical metrics in \( \mathcal{G}/H \) and \( H \) and the invariant connection of the theorem above.

Now since \( D \) is horizontal and \( \mu_* D\bar{\varphi} = \mu_\mu d\bar{\varphi} = d\varphi \)

we have from \(2.13.1\)

\[
\overline{h}_\varphi(D\bar{\varphi},D\bar{\varphi}) = h_\varphi(d\varphi,d\varphi)
\]

and hence the action can be rewritten (13)

\[
S[\varphi] = \frac{1}{\kappa} \int u (D\varphi,D\varphi) \cdot \eta
\]

(3.7.5)

where, explicitly
This formulation of the theory is known as "nonlinear sigma model with gauge symmetry".

3.8 As a preparation for the next paragraph we will consider now a field \( \varphi \) with values in an \( n \)-dimensional connected Lie group \( G \). Let \( g(x) \) denote the adjoint representation of the image of \( x \) under \( \varphi \). We shall show now that the harmonic mapping action can be written (12) (16)

\[
S[\varphi] = -\frac{1}{2} \int_{u} g^{\mu \nu} \text{Tr} (\partial_{\mu} \varphi \cdot \partial_{\nu} \varphi^{-1}) \cdot \eta
\]

which is invariant by right and left multiplications:

\[ S[L_g(\varphi)] = S[\varphi] \quad \forall \ g' \in G, \quad d\varphi = 0 \]

Differentiating \( g^{-1} = e \) we obtain \( d\varphi^{-1} = -g^{-1} d\varphi g^{-1} \), so we can rewrite

\[
S[\varphi] = \frac{1}{2} \int_{u} g^{\mu \nu} \text{Tr} (g' g \cdot g' g^{-1} \cdot \eta)
\]

Let us now choose normal coordinates \( y^i (i = 1, \ldots, n) \) (see § 3.5); then \( g(x) = \exp y^i \eta_i \).

Expanding in Taylor series, we arrive at (39) p. 340

\[
e^{-y^c \partial_c} \quad dy^c \cdot e^{y^c} = dy^c - \frac{1}{2!} [y^c, dy^c] + \frac{1}{3!} [y^c, [y^c, dy^c]] - \ldots
\]

The quantity in parenthesis is precisely the inverse of the left auxiliary function, \( L^{-1}(y)_k \) (see e.g. (38)) so

\[
\tilde{g}^{-1} d\varphi = e^{-y^c \partial_c} \varphi^c = dy^k L^{-1}(y)_k \eta_c
\]

and therefore recalling \( (L^{-1}(\tilde{g}))_k^i X^k = L^{-1}(\tilde{g})_k^i X^k \) and \( (3.5.2) \)
we have

\[ S[\mathbf{q}] = \frac{1}{2} \sum_u g^{u \nu} \partial_{\nu} \mathbf{y}^k \partial \mathbf{y}^\lambda h_{k\lambda} \mathbf{e} \cdot \mathbf{m} \]

Q.E.D.

If \( G \) is a simple group (by which we mean that it has no continuous invariant subgroups) then we can replace the adjoint representation with any other and (3.8.1) gets only multiplied by a constant (see eq. (3.10.4) and the following comments).

3.9 I will show now that the action can be written in a form similar to (3.8.1) also if \( \mathcal{N} \) is a reductive homogeneous space.

Let us introduce a basis in \( \mathcal{G} \) such that the first \( n \) vectors lie in \( \mathcal{P} \) and the remaining \( m \) lie in \( \mathcal{H} \). Since we shall be working with the adjoint representation, we denote

- \( \mathbf{C}_i \), \( i = 1, \ldots, m \) basis for \( \mathcal{G} \)
- \( \tilde{\mathbf{C}}_i = \mathbf{C}_i \), \( i = 1, \ldots, d \) basis for \( \mathcal{P} \)
- \( \mathbf{C}_j' = \mathbf{C}_j - \mathbf{m} \), \( j = 1, \ldots, n-d \) basis for \( \mathcal{H} \)

The condition that \( G/H \) be reductive insures that this basis can be taken orthonormal and that \( \tilde{\mathbf{C}}_i \) is a basis for the vector space \( T_0(G/H) \); translating with \( g \in G \) from the left one can then obtain a basis for the tangent space at any point \( T_{g0}(G/H) \). Then, any point in \( G/H \) can be represented by a matrix \( \hat{\mathbf{y}} = \exp \tilde{\mathbf{y}}' \tilde{\mathbf{C}}_i \) with the sum over \( i \) running from 1 to \( d \), and \( \tilde{\mathbf{y}}' \) are normal coordinates in \( G/H \) centered at 0. Notice that \( \hat{\mathbf{y}} \hat{h} \) with \( h \in H \) would still represent the same point of \( G/H \), but we have so to say "fixed the gauge" (in the language of § 3.7) choosing \( h = e \).

The claim is that the action can be written (12) (16)
\[ S[\gamma] = - \frac{1}{2} \sum_{\mu} g^{\mu\nu} \text{Tr}(\partial_\nu \hat{\gamma} \partial_\mu \hat{\gamma}^{-1}) \cdot \eta \] (3.9.1)

The proof goes exactly as in the previous paragraph, provided one allows the indices to run over \( \mathcal{P} \) only.

The form (3.9.1) is invariant by left \( \mathcal{G} \) multiplication and right \( \mathcal{H} \)-multiplication.

Let us note here that the condition of being reductive homogeneous is not very strong; it includes as special cases all symmetric spaces (see (35) (36) and the second volume of (3)) and all homogeneous spaces \( \mathcal{G}/\mathcal{H} \) such that \( \mathcal{H} \) is either discrete, compact or semisimple and connected (40).

3.10 This and the following three paragraphs contain a discussion on the geometric properties of Stiefel manifolds, Grassmann manifolds and Stiefel bundles; they serve as an introduction to the Grassmanian valued field theories.

Let us use a compact notation taken from ref. (2) which allows to treat simultaneously the case of real, complex or quaternionic spaces. Let \( \mathbb{F} \) be the field of real or complex members, or the algebra of quaternions: \( \mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H} \). If \( x \in \mathbb{F} \), let \( x^* = x \) if \( \mathbb{F} = \mathbb{R} \) and let \( x^* \) denote the conjugate in the usual sense if \( \mathbb{F} = \mathbb{C}, \mathbb{H} \). Let \( \mathbb{F}^n \) denote the vector space of \( n \)-tuples of real, complex or quaternionic numbers and define the inner product in \( \mathbb{F}^n \):

\[ (u,v) = \sum_{i=1}^{n} u_i^* v_i; \quad u,v \in \mathbb{F}^n \] (3.10.1)

Let us denote by \( U_\mathbb{F}(n) \) the group of real, complex or quaternionic matrices which leave this scalar product invariant:
\[
M \in U_F(m) \iff (MU, MV) = (u, v) \quad \forall u, v \in F^m
\]

Clearly

\[
U_F(m) = \begin{cases} 
O(m) & F = \mathbb{R} \\
U(m) & F = \mathbb{C} \\
Sp(m) & F = \mathbb{H}
\end{cases}
\]

Let us also define the unimodular subgroups

\[
SU_F(m) = \begin{cases} 
SO(m) & \text{for } F = \mathbb{R} \\
SU(m) & \text{for } F = \mathbb{C}
\end{cases}
\]

Define the hermitian conjugate of a matrix as usual

\[
(M^*)_{ij} = (M_{ji})^* \quad \text{(in the real case this is just the transpose)}.
\]

If we arrange the components of vectors in columns, then the conjugate \( V^\dagger \) is a row vector; with this convention we may write the inner product simply as \((u, v) = u^\dagger v\). The matrices of \( U_F(m) \) satisfy

\[
M^\dagger M = I_n \\
MM^\dagger = I_n
\]

\[(3.10.2)\]

(here and in the following \( I_n \) denotes the \( n \times n \) unit matrix).

It is also easy to see that the algebra of \( U_F(m) \) (\( SU_F(m) \)), \( U_F(m) \) (\( SU_F(m) \)) consists of the antihermitian (antihermitian and traceless) matrices. Notice that we are not actually dealing with the abstract group but rather with a representation of it on the space \( F^n \); this is the so-called fundamental representation.

The group \( U_F(m) \) (\( SU_F(m) \)) can be embedded in \( F^{n^2} \), to be thought of as the space of all \( n \times n \) matrices with \( F \)-valued elements; the embedding conditions are \((3.10.2)\) and in addition \( \det M = 1 \) for \( SU_F(m) \). An \( n \times n \) matrix \( X \) represents a vector tangent to \( U_F(m) \) (\( SU_F(m) \)) in \( M \) if and only if
\[ M^tX + X^tM = 0 \] for \( U_F(n) \)  
\[ \operatorname{Tr} M^tX = 0 \] for \( SU_F(n) \) \hfill (3.10.3)

To see this, parametrize a curve in \( F_{m^2} \) passing through \( M \) by
\[ c(t) = M + Xt + O(t^2) \] ; the curve is tangent to \( U_F(m) \) if and only if \( c^t(t) c(t) = I_n \) to first order in \( t \) and it is tangent to \( SU_F(n) \) in \( M \) if in addition \( \det c(t) = 1 \) to first order in \( t \) : this implies (3.10.3).

Left translations of vectors are given simply by matrix multiplications: if \( X \in T_M(G) \), \( M'X \in T_{M'(G)} \).
Differentiating (3.10.2), we see that \( \mathrm{d}M \) is \( T_M(G) \)-valued.
Acting on \( X \in T_M(G) \), it gives \( \mathrm{d}M(X) = X \). Therefore, the Maurer-Cartan form of \( U_F(m) \) is \( M^t \mathrm{d}M \), since
\[ M^t \mathrm{d}M(X) = M^tX \in T_M(G) \approx \mathfrak{g} \]
and in addition \( \operatorname{Tr} M^t \mathrm{d}M(X) = \operatorname{Tr} M^tX = 0 \) if \( X \in T_M(SU_F(m)) \).
The standard flat metric in \( F_{m^2} \) induces a metric in \( U_F(m) \) given by
\[ \langle X, Y \rangle_M = \Re \operatorname{Tr} X^t Y \] \quad \( X, Y \in T_M(U_F(m)) \) \hfill (3.10.4)

If \( M = I_n \), this is a metric in the algebra \( U_F(m) \). It does not coincide with the Cartan-Killing metric
\[ \operatorname{Tr} (\mathrm{ad}(X) \circ \mathrm{ad}(Y)) \]
but is proportional to it for simple algebras (the algebras of the groups \( O(m), SU(m) \) and \( Sp(m) \) are simple). The proportionality coefficient \( K \) is the ratio of the so-called indexes of the representations \( K = \text{index (fundamental)/ index (adjoint)} \) \hfill (36).
Also if \( X, Y \in T_M(U_F(m)) \), \( \Re \operatorname{Tr}(L_M^{-1} X)(L_M^{-1} Y) = \Re \operatorname{Tr} X^t Y \) so that the metric (3.10.4) is proportional to the canonical metric in the group manifold (3.5.2), with the same coefficient.
as the scalar products in the algebras:

\[ \langle X, Y \rangle_M = \kappa \nabla_M(X, Y) \quad (3.10.5) \]

3.11 The Stiefel space \( V_k(F^n) \) is defined as the set of all \( k \)-beins in \( F^n \) (a \( k \)-bein is a \( k \)-tuple of orthonormal vectors). We can assemble the column vectors \( w_i, \ldots, w_k \) in a matrix \( Z = (w_i, \ldots, w_k) \) with \( n \) rows and \( k \) columns and of rank \( k \); then the orthonormalization condition is

\[ Z^T Z = I_k \quad \iff \quad (w_i, w_j) = \delta_{ij} \quad (3.11.1) \]

The left action of \( U_F(m) \) on \( V_k(F^n) \) is defined through its action on the individual vectors \( \lambda : U_F(m) \times V_k(F^n) \to V_k(F^n) \) by \( \lambda(M, Z) = \lambda(M, (w_i, \ldots, w_k)) = (Mw_i, \ldots, Mw_k) \).

This action is clearly transitive. Now choose an orthonormal basis in \( F^n \) and let \( Z^0 = (\overline{w}_i, \ldots, \overline{w}_k) \) be a distinct element of \( V_k(F^n) \), for instance the one such that the vectors \( \overline{w}_i \) coincide with the last \( k \) vectors of the basis. It is easily seen that the isotropy group of this element is given by matrices of \( U_F(m) \) forming a subgroup \( U_F(n-k) \):

\[
M = \begin{pmatrix}
M' & 0 \\
0 & I_k
\end{pmatrix}
\]

\[ M^T M' = I_{n-k}, \quad M' M^T = I_{n-k} \]

Therefore the map \( \gamma : \frac{U_F(m)}{U_F(n-k)} \to V_k(F^n) \) defined by \( \gamma(MU_F(n-k)) = \lambda(M, Z^0) \) is bijective. A similar argument leads to a bijective map \( \hat{\gamma} : \frac{SU_F(m)}{SU_F(n-k)} \to V_k(F^n) \).

Since \( \frac{U_F(m)}{U_F(n-k)} \) and \( \frac{SU_F(m)}{SU_F(n-k)} \) have a canonical differential...
structure, we can turn \( V_k(F^n) \) to a differentiable manifold requiring the maps \( \eta \) and \( \eta' \) to be diffeomorphisms.

In the following we will simply identify

\[
V_k(F^n) = \frac{U_F(m)}{U_F(m-k)} = \frac{SU_F(m)}{SU_F(m-k)}
\]

(3.11.2)

Special cases of Stiefel manifolds are

\[
V_n(F^n) = U_F(n) \quad V_n(R^n) = SO(n) \quad V_n(C^n) = SU(n) \\
V_n(R^n) = S_{n}^{\text{even}} \quad V_n(C^n) = S_{2n}^{\text{even}} \quad V_n(H^n) = S_{4n-1}^{\text{even}}
\]

Sometimes an element of \( V_k(F^n) \) is represented by a matrix \( M \in U_F(n) \) (for definiteness, we will suppose that \( k \leq n-k \)):

\[
M = \begin{bmatrix} Z & \bar{Z} \\ n-k & k \end{bmatrix}
\]

(3.11.3)

The submatrix \( Z \) is determined, modulo right multiplications by \( U_F(n-k) \), by conditions (3.10.2) which in this case read

\[
ZZ^T + \bar{Z}\bar{Z}^T = I_n \\
Z^T\bar{Z} = I_k \\
\bar{Z}^T\bar{Z} = I_{n-k} \\
Z^T\bar{Z} = 0
\]

(3.11.4)

Notice that \( \bar{Z} \) represents an element of \( V_{n-k}(F^n) \).

The manifold \( V_k(F^n) \) can be embedded in \( F^{nk} \); then by the same argument as in (3.10.3) a \( m \times k \) matrix \( X \) represents a vector tangent to \( V_k(F^n) \) in \( Z \) if and only if

\[
Z^TX + X^TZ = 0
\]

(3.11.5)

The standard flat metric in \( F^{nk} \) induces in \( V_k(F^n) \) the metric

\[
\langle X, Y \rangle_{Z} = \mathrm{Re} \, \mathrm{Tr} \, X^T Y
\]

(3.11.6)
which is a direct generalization of (3.10.4).

3.12 The Grassmann space $G_k(F^n)$ is defined as the set of all $k$-dimensional linear subspaces of $F^n$ (remember that $F^n$ has to be thought of as a vector space, so we are actually considering only the linear subspaces through the origin).

Similarly, the oriented Grassmann space $SG_k(F^n)$ is defined as the set of all $k$-dimensional oriented linear subspaces of $F^n$.

Clearly, $G_k(F^n)$ is related to $V_k(F^n)$ by means of a fibration, the projection map $\eta$ being the natural one that associates to a given $k$-bein the $k$-dimensional linear subspace that it spans; furthermore rotations of the $k$-bein in that subspace correspond to a group $U_F(k)$ and leave the point in $G_k(F^n)$ fixed. Thus one must have a relation of the type

$$G_k(F^n) = V_k(F^n) \mod U_F(k) \quad (3.12.1)$$

A similar reasoning for the oriented case would lead us to

$$SG_k(F^n) = V_k(F^n) \mod SU_F(k) \quad (3.12.2)$$

To see all this in more detail, let us fix an orthonormal basis in $F^n$ and let $z^0$ be a distinct element of $G_k(F^n)$, for instance the one spanned by the last $k$ vectors of the basis. Let $\lambda : U_F(k) \times G_k(F^n) \to G_k(F^n)$ be the (clearly transitive) action of $U_F(k)$ on the Grassmann space, defined by rotating all the vectors in a linear subspace by means of the same matrix $M : \lambda(M,z) = Mz = \{ v \in F^n \mid v = Mw, \ w \in z \}$.
It is easily seen that the isotropy group of the distinct element \( Z^* \) is a subgroup \( U_F(n-k) \oplus U_F(k) \):

\[
M = \left( \begin{array}{c|c}
M' & \mathbb{I}_{n-k} \\
\hline 
M'' & \mathbb{I}_k 
\end{array} \right)
\]

Therefore the map \( \delta \) defined by

\[
\delta \left( (M(U_F(n-k) \oplus U_F(k))) \right) = \lambda(M, z^*)
\]

leads to a bijective map \( \hat{\delta} \):

\[
\hat{\delta} : \frac{U_F(m)}{U_F(n-k) \oplus U_F(k)} \rightarrow SG_k(F^n)
\]

Since \( \frac{U_F(m)}{U_F(n-k) \oplus U_F(k)} \) and \( \frac{SU_F(m)}{SU_F(n-k) \oplus SU_F(k)} \) have a canonical differential structure, we can turn \( G_k(F^n) \) and \( SG_k(F^n) \) to differentiable manifolds by requiring \( \delta \) and \( \hat{\delta} \) to be diffeomorphisms. Again, we will forget these subtleties and simply identify

\[
G_k(F^n) = \frac{U_F(m)}{U_F(n-k) \oplus U_F(k)} = \frac{SU_F(m)}{SU_F(n-k) \oplus SU_F(k)}
\]

(3.12.3)

where \( S(U_F(n-k) \oplus U_F(k)) = SU_F(m) \cap (U_F(n-k) \oplus U_F(k)) = SU_F(n-k) \oplus \oplus SU_F(k) \oplus U(F) = SU_F(n-k) \oplus U_F(k) = U_F(n-k) \oplus SU_F(k) \)

and for \( F = \mathbb{R} \) or \( \mathbb{C} \)

\[
SG_k(F^n) = \frac{SU_F(m)}{SU_F(n-k) \oplus SU_F(k)}
\]

(3.12.4)

The Grassmann manifolds can be obtained from the corresponding oriented ones by forgetting orientation; in the real case this means taking equivalence classes with respect to the action of the group \( Z^* \), and in the complex case this means taking equivalence classes with respect to the group \( U(0) \):

\[
G_k(R^n) = SG_k(R^n) \mod Z^* \\
G_k(C^n) = SG_k(C^n) \mod U(0)
\]
The dimensions are
\[ \dim G_k(F^n) = q_k(n-k) \quad (q = \dim F) \]
\[ \dim S G_k(R^n) = k(n-k) \]
\[ \dim S G_k(C^n) = 2k(n-k) - 1 \]

Special cases are projective spaces and spheres:
\[ G_1(F^n) = F P^{n-1} \]
\[ S G_1(R^n) = S^{n-1} \]
\[ S G_1(C^n) = S^2n-1 \]

Let us note here that there exists a canonical bijection between
\[ G_k(F^n) \] and \[ G_{k-h}(F^n) \], which associates to any \( k \)-dimensional linear subspace the \((n-k)\)-dimensional subspace orthogonal to it. Thus, we have the following natural projections:
\[ V_k(F^n) \]
\[ \cong \]
\[ V_{n-k}(F^n) \]
\[ (3.12.5) \]

where \( \cong \) means "diffeomorphic to"; we will refer to the fibrations on the right as the "duals" to the fibrations on the left.

A similar diagram holds in the oriented case. Thus we have:
\[ \Sigma_k(F^n) = (V_k(F^n), \eta, G_k(F^n); U F(k)) \]

"Stiefel bundle"
\[ \hat{\Sigma}_k(F^n) = (V_k(F^n), \hat{\eta}, S G_k(F^n); S U F(k)) \]

"oriented Stiefel bundle"

Also, the following relations hold:
\[ (U F(n), \mu, G_k(F^n); U F(n-k) \otimes U F(k)) = \Sigma_k(F^n) \otimes \Sigma_{n-k}(F^n) \]
\[ (S U F(n), \hat{\mu}, S G_k(F^n); S U F(n-k) \otimes S U F(k)) = \hat{\Sigma}_k(F^n) \otimes \hat{\Sigma}_{n-k}(F^n) \]

\( \otimes \) is the "Whitney sum", defined as follows: if \( \xi = (E, \pi, X) \) and \( \xi' = (E', \pi', X') \), the "direct product bundle" is \( \xi \otimes \xi' = (E \oplus E', \pi \oplus \pi', X \oplus X') \)

where \( \pi \oplus \pi'(w,w') = (\pi(w), \pi(w')) \). If \( X = X' \), \( \xi \otimes \xi' = (E \oplus E', \pi, X) = \text{diag } \xi \otimes \xi' \)

i.e. \( E \oplus E' = \{ (w,w') \in E \oplus E' \mid \pi(w) = \pi'(w') \} \) and \( \pi(w,w') = \pi(w) = \pi'(w). \)
3.13. We will now introduce a connection in the Stiefel bundles and use it to induce a metric in the Grassmann manifolds. Given $Z \in \mathcal{V}_k(F^n)$ let $\overline{Z} \in \mathcal{V}_{n-k}(F^n)$ be determined (modulo $U_F(n-k)$) by relations (3.11.4) and define the following $n \times n$ matrices:

$$P(Z) = ZZ^+ \quad \overline{P}(Z) = \overline{Z} \overline{Z}^+ \quad \text{(3.13.1)}$$

We will drop the argument where no confusion can arise. Using (3.11.4) we have the following relations:

$$p^2 = p \quad \overline{p}^2 = \overline{p} \quad p + \overline{p} = I_n \quad p^+ = \overline{p} \quad \text{(3.13.2)}$$

$$\text{Rank } P = k \quad \text{Rank } \overline{P} = n-k \quad \text{Tr } P = k \quad \text{Tr } \overline{P} = n-k$$

Therefore $P$ and $\overline{P}$ are projection operators. In fact, if $\nu \in F^n$ and $Z = (w_1, \ldots, w_k)$, we have

$$P(Z)\nu = \sum_{i=1}^{k} (w_i, \nu) w_i \quad \text{(3.13.3)}$$

so that $P(Z)$ projects on the $k$-dimensional linear subspace spanned by $Z$, while $\overline{P}(Z)$ projects on the orthogonal $(n-k)$-dimensional linear subspace spanned by $\overline{Z}$.

A vector $X \in T_Z(\mathcal{V}_k(F^n))$ will be vertical with respect to the projection $\eta$ if and only if a curve $c(t) = Z + tX + O(t^2)$ through $Z$ in $F^{n_k}$ is tangent to the fibre through $Z$:

$$c(t) \in \mathcal{V}_{n-k} \quad M(t) \in U_F(k) \quad \text{then}$$

$$X = \frac{dc(t)}{dt} \bigg|_{t=0} = ZM \quad \eta M \in U_F(k) \quad \text{(3.13.4)}$$

The operator $P(Z)$ is the vertical projector in $T_Z(\mathcal{V}_k(F^n))$ since $Z^+X$ is antihermitian by (3.11.5) and hence belongs to $U_F(k)$, and $P(Z)X = ZZ^+X \in \text{Ver} T_Z(\mathcal{V}_k(F^n))$ by (3.13.4).
The canonical $U_F(n)$-invariant connection in $\Sigma_k(F^n)$ is defined as the $U_F(k)$-component of the Maurer-Cartan form $M^t dM$ of $U_F(n)$.

Splitting $M$ as in (3.11.3) we obtain

$$A = Z^t dZ $$

(3.13.5)

It is immediately seen that $\text{Im} \overline{P} = \ker A$ so that $\overline{P}(Z)$ is actually the projector onto the horizontal subspace in $T_Z(V_k(F^n))$.

Notice that $P$ and $\overline{P}$ are orthogonal with respect to the inner product (3.11.6).

Similarly, in $T_Z(V_{n-k}(F^n))$, $\overline{P}(Z)$ is the vertical projector and $P(Z)$ the horizontal projector, and the connection form is $\overline{A} = Z^t dZ$.

The connection form in $\Sigma_k(F^n)$ is similarly defined as the $SU_F(k)$ component of the canonical form $M^t dM$ of $SU_F(n)$.

$$\overline{A} = Z^t dZ - \frac{1}{n-k}(\text{Tr}Z^t dZ)I_{n-k}$$

(3.13.6)

As in § 2.13, the metric in $G_k(F^n)$ is derived from the metric (3.11.6) in $V_k(F^n)$ and the connection in the Stiefel bundle.

Let $\tilde{X}$ and $\tilde{Y}$ be vectors tangent to $G_k(F^n)$ at a point $Z$; we will represent them by their horizontal lifts in $V_k(F^n)$, matrices satisfying (3.11.5) with $\eta(Z) = \tilde{Z}$. Next, introduce the matrices $\tilde{X}$ and $\tilde{Y}$ containing $X$ and $Y$ as submatrices and representing vectors tangent to $U_F(n)$ at $M$, where $M$ projects on $Z$ (i.e. in the situation of eq. (3.11.3)):

$$\tilde{X} = \begin{bmatrix} \tilde{X} & X \end{bmatrix} \quad \tilde{Y} = \begin{bmatrix} \tilde{Y} & Y \end{bmatrix}$$

(3.13.7)

Exploiting condition (3.10.3), we obtain

$$\tilde{X}^t \tilde{Z} + \tilde{Z}^t \tilde{X} = 0 \quad \tilde{Y}^t \tilde{Z} + \tilde{Z}^t \tilde{Y} = 0$$

$$X^t Z + Z^t X = 0 \quad Y^t Z + Z^t Y = 0$$

$$X^t \tilde{Z} + \tilde{Z}^t X = 0 \quad Y^t \tilde{Z} + \tilde{Z}^t Y = 0$$

(3.13.8)

so that $\tilde{X}$ and $\tilde{Y}$ have to be vectors tangent to $V_{n-k}(F^n)$ satisfying one additional condition (the third relation in (3.13.8)).
The inner product (3.10.4) in $U_F(m)$ is now

$$\langle \tilde{X}, \tilde{Y} \rangle = \text{Re} \text{Tr} (\tilde{X}^+ \tilde{Y}) + \text{Re} \text{Tr} (\tilde{X}^+ \tilde{Y})$$

Upon using the differential of the fourth relation in (3.11.4) it is easily seen that the contributions to the two addenda of the horizontal parts of the vectors $X, Y$ and $\tilde{X}, \tilde{Y}$ are equal:

$$\text{Re} \text{Tr} (ZZ^+X)(ZZ^+Y) = \text{Re} \text{Tr} (ZZ^+X)(ZZ^+Y)$$

Thus using the isomorphisms $\eta^*: \text{Hor}_z(V_n(F^n)) \rightarrow T_z(G_k(F^n))$ and $\overline{\eta}^*: \text{Hor} T_z(V_n-k(F^n)) \rightarrow T_z(G_k(F^n))$, we can induce in $T_z(G_k(F^n))$ the inner product

$$\langle X, Y \rangle_z = 2 \text{Re} \text{Tr} X^+Y = 2 \text{Re} \text{Tr} \tilde{X}^+Y$$

(3.13.9)

Being directly derived from the metric $\langle ., . \rangle$ in $U_F(n)$, this metric is related to the canonical metric in $G_k(F^n)$ derived from (3.5.2) by the same relation as in (3.10.5):

$$\langle X, Y \rangle_z = k h_z(X, Y)$$

(3.13.10)

3.14. We are now in a position to rewrite the action in a couple of ways which are peculiar of Grassmannian-valued theories. Let $z: \mathcal{J} \rightarrow G_k(F^n)$ and $\overline{z}: \mathcal{J} \rightarrow V_k(F^n)$, $\overline{z}: \mathcal{J} \rightarrow V_{n-k}(F^n)$, $M: \mathcal{J} \rightarrow U_F(m)$ be lifts of $\mathcal{J}$, such that the following diagram commutes

$$\begin{array}{ccc}
U_F(m) & \overset{\eta^*}{\longrightarrow} & U_F(m) \\
\downarrow M & & \downarrow M \\
V_k(F^n) & \overset{z}{\longrightarrow} & V_{n-k}(F^n) \\
\downarrow \eta & & \downarrow \eta \\
G_k(F^n) & \approx & G_{n-k}(F^n)
\end{array}$$

(3.14.1)
This generalizes (3.7.1); the corresponding generalization of (3.7.4) is

\[
\begin{array}{ccc}
\text{Hor } T(U_F(n)) & \mu' & \tilde{\mu}' \\
\downarrow & \downarrow & \downarrow \\
\text{Hor } T(V_{k}(F^n)) & \tilde{\eta} & \tilde{\tilde{\eta}} \\
\downarrow & \downarrow & \downarrow \\
T(G_k(F^n)) & T(G_{n-k}(F^n)) & \\
\end{array}
\]

(3.14.2)

By \( \text{Hor } T(U_F(n)) \) I mean here the horizontal part with respect to the projection \( \mu \) rather than \( \mu' \) or \( \tilde{\mu}' \); in this diagram \( \mu', \tilde{\mu}', \eta, \tilde{\eta} \) are vector space isomorphisms. The connections and related covariant derivatives are

\[
\begin{array}{c}
\mu': \quad \bar{A} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ Z'dZ \end{bmatrix} \\
\tilde{\mu}': \quad A = \begin{bmatrix} 0 \\ 0 \\ Z'dZ \end{bmatrix}
\end{array}
\]

\[
\begin{array}{c}
\eta: \quad A = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\tilde{\eta}: \quad \bar{A} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\mu: \quad \tilde{\bar{A}} = A + \bar{A} = \begin{bmatrix} dZ - ZA \\ d\bar{Z} - \bar{Z}A \end{bmatrix}
\end{array}
\]

The last one coincides with the canonical invariant connection given by the theorem of §3.7.
Eq. (3.7.3) reads now (remembering (3.13.10))

$$S[z] = \frac{1}{2K} \int_u \langle \bar{D}M, D\bar{M} \rangle \cdot \eta \tag{3.14.3}$$

or according to (3.14.9) (14)(15)

$$S[z] = \frac{1}{K} \int_u \langle DZ, D\bar{Z} \rangle \cdot \eta \tag{3.14.4}$$

or the same form with $DZ$ replaced by $D\bar{Z}$; in these formulas we can omit taking the real part since the integrand is already real.

The following relation holds:

$$\text{Tr}(DZ)^\dagger(DZ) = \text{Tr}(dZt d\bar{Z} + A^2)$$

and the same with $D\bar{Z}$ and $\bar{A}$.

Using the differential of the second relation in (3.11.4), one can then turn this into (11)(12)(16)

$$S[z] = \frac{1}{2K} \int_u \text{Tr}(d\bar{P}d\bar{P}) \cdot \eta \tag{3.14.5}$$

and the same with $\bar{P}$.
SECT. 4 GLOBAL DEFINITION

4.1 In § 3.3 we have seen that an \( J \)-valued function \( f \) over \( J \) defines (and is defined by) a unique cross section of a trivial bundle with total space \( J \times J \). We will now generalize our definition of the action functional in such a way that it makes sense also for a nontrivial bundle.

First of all we define a global field configuration for the nonlinear sigma model to be a (smooth) section of the fibre bundle \( \Sigma = (E, \pi, J; J, G) \) with base space \( J \), typical fibre \( J \) and structure group \( G \). The configuration space for the nonlinear sigma model is the space of all sections of all possible bundles \( \Sigma : \Gamma = \oplus \Gamma(E) \).

Now we recall from § 2.7 that the derivative of a cross-section is not a well-defined concept and one has to introduce a connection in the associated principal bundle \( \Pi = (P, \pi, J; G) \) in order to define a covariant derivative.

This connection can be given a priori, however this is a quite unnatural procedure and it is better to introduce a dynamically active Yang-Mills field with gauge group \( G \), minimally coupled to the nonlinear sigma model. Therefore in the next paragraph we shall recall briefly how gauge fields can be described by means of the fibre bundle machinery of sect. 1.

4.2 An extended treatment on the subject of this paragraph can be found in references (33)(34). In the fibre bundle approach to Yang-Mills theory, one is given a priori a principal bundle.
\( \Pi = (P, \pi, \mathcal{H}; G) \), \( G \) being the Yang-Mills group. A local choice of gauge on \( \mathcal{U} \subseteq \mathcal{H} \) is the choice of a trivialization of \( \Pi|_\mathcal{U} \); by the discussion in § 2.4, this is equivalent to the choice of a preferred local cross-section \( s: \mathcal{U} \to P \) such that \( \mathcal{S}(x) = \psi(x, e) \). If \( \omega = \omega^i e_i \) is a connection form in \( P \), the pull-back of \( \omega \) by the preferred section:

\[
\mathcal{A} = s^* \omega \quad \mathcal{A} = A^i \quad \mathcal{A} = e_i \quad (4.2.1)
\]
is a \( \mathcal{G} \)-valued one-form on \( \mathcal{H} \). Observe that this definition, together with the definition of the local direct product basis and eq. (2.10.4), implies that \( A^i = B^i \). A gauge transformation on \( \mathcal{U} \) is a change of trivialization of \( \Pi|_\mathcal{U} \), therefore a change of cross-section. By eq. (2.3.1) and the definition of cross-section, it is clear that any two sections can be joined by an \( x \)-dependent group translation in \( \pi^*(\mathcal{U}) \), so the gauge transformations are the vertical automorphisms of \( \Pi \):

\[
\mathcal{S}(x) = \mathcal{R}_{g(x)}(\mathcal{S}(x)) = \mathcal{S}(x) g(x) \quad (4.2.2)
\]

Accordingly, the coordinates of a fixed point \( p \in \pi^*(\mathcal{X}) \) change by

\[
p = \psi(x, a) = \psi'(x, g^{-1} a) \quad (4.2.3)
\]

If \( \mathcal{U} = \mathcal{U}_A \cap \mathcal{U}_B \) is the intersection of two coordinate neighborhoods of \( \mathcal{H} \), eq. (2.4.3) may be regarded as the effect of a gauge transformation:

\[
\mathcal{S}_B(x) = \psi_B(x, e) \quad \mathcal{S}_A(x) = \psi_A(x, e) = \psi_B(x, g_{AB}) \quad (4.2.4)
\]

\[
\mathcal{S}_B(x) = \mathcal{R}^{-1}_{g_{AB}(x)} \left( \mathcal{S}_A(x) \right) = \mathcal{S}_A(x) g_{AB}(x)
\]
Now it can be seen using condition $b'$ of (2.10.2) and proposition 1.4 in (3), vol. I, p. 11 (the explicit computation can be found in (29)) that $A' = s^* \omega$ is related to $A$ by

$$A' = g^{-1} A g + g^{-1} d g$$  \hspace{1cm} (4.2.5)

We thus recognize that $A$ represents the **Yang-Mills potential**. Similarly, the pull-back of the curvature form (2.10.7) is the $\mathfrak{g}$-valued two-form on $\mathcal{H}$

$$F = s^* \Omega$$  \hspace{1cm} (4.2.6)

that transforms like

$$F' = g^{-1} F g$$  \hspace{1cm} (4.2.7)

and is related to the potential by (see § 2.10)

$$F = dA + [A, A]$$  \hspace{1cm} (4.2.8)

Thus, $F$ represents the **Yang-Mills field strength**. By a well-known theorem (essentially a corollary of the theorem in § 4.5) a principal bundle $\mathcal{H}$ admits a global cross-section (i.e. it is possible to fix the gauge globally throughout $\mathcal{H}$ without singularities) if and only if $\mathcal{H}$ is trivial ($\mathcal{P} = \mathcal{H} \times \mathfrak{g}$). This will always be the case if $\mathcal{H}$ is homeomorphic to $\mathbb{R}^n$ (notice however that here a sort of "effective topology" is relevant where all singularities in $A$ like Dirac strings a.s.o. have been removed and eventually $\mathcal{H}$ has been compactified by imposing some appropriate boundary conditions). When the Yang-Mills field is dynamically active, its action functional is ($\mathcal{F}$ is the inner product in $\mathfrak{g}$):
Thus, while the topology of \( \Pi \) is given a priori, its connection and curvature are determined by the principle of stationary action.

4.3 We shall resume now the discussion on the covariant derivative of a cross-section that was begun in § 2.12. Let \( \sigma : \mathcal{K} \rightarrow \mathcal{E} \) be a (global) cross-section of \( \Sigma \) and \( \bar{\sigma} : \mathcal{P} \rightarrow \mathcal{K} \) be the equivariant mapping associated to \( \sigma \) by (2.8.3). A local representative of \( \bar{\sigma} \) in the gauge defined by a local preferred section \( s : \mathcal{K} \rightarrow \mathcal{U} \rightarrow \mathcal{P} \) is the map

\[
\varphi_u = \bar{\sigma} \circ s : \mathcal{U} \rightarrow \mathcal{K}
\]  

(4.3.1)

We shall now show that \( \varphi_u \) is also a local representative of the section \( \sigma \) in the neighborhood \( \mathcal{U} \). Using diagrams (2.5.4) and (2.8.3) and the following discussions, we have

\[
\sigma(x)|_u = X(p, \bar{\sigma}(p)) = \Psi_u(x, u, (x, s(p)), \bar{\sigma}(p))
\]

where \( \pi_u(p) \) is the local coordinate of \( p \) in \( \mathcal{G} \). Due to the freedom to move \( p \) in the fibre over \( x \), we can choose \( p = s(x) \) in which case \( \pi_u(p) = e \) and by (4.3.1),

\[
\sigma(x)|_u = \Psi_u(x, s(x))
\]

(4.3.2)

which shows indeed that \( \varphi_u \) is a local representative of \( \sigma \) in \( \mathcal{U} \). From (2.8.1) it follows that under the gauge transformation (4.2.2) it transforms as
Let us now derive an explicit form for the covariant derivative of \( \sigma \) in terms of its local representative. Starting from eq. (2.12.2) we have ( \( \mathcal{V} \) is a vector field on \( \mathcal{P} \))

\[
\mathcal{D}_\mathcal{V} \sigma = \mathcal{D}_\mathcal{V} (\sigma) = \mathcal{D}_\sigma (\mathcal{V}) - \mathcal{V} \mathcal{D}_\sigma (\mathcal{V}) = \mathcal{D}_\sigma (\mathcal{V}) - \mathcal{D}_\mathcal{V} (\mathcal{V})
\]

Since \( \mathcal{V} \mathcal{D}_\sigma (\mathcal{V}) \) \( = \mathcal{D}_\mathcal{V} (\mathcal{V}) \) \( \mathcal{D}_\sigma (\mathcal{V}) \)

\[
\mathcal{D}_\sigma (\mathcal{V}) = \frac{d}{dt} \left[ \mathcal{D}_\mathcal{V} (\exp(tw(\mathcal{V})) \mathcal{P}) \right]_{t=0}
\]

The \( \alpha \)-component of this quantity is ( \( \gamma^\alpha \) are coordinates in \( \mathcal{M} \)):

\[
(\mathcal{D}_\sigma (\mathcal{V}))^\alpha = \frac{d}{dt} (\mathcal{L}^\alpha (\exp(-tw(\mathcal{V})), \sigma(\mathcal{V})))_{t=0}
\]

\[
= - \mathcal{L}^\alpha (\sigma(\mathcal{V}), \omega^\alpha (v))
\]

where \( \mathcal{L}^\alpha \) are the auxiliary functions (2.4.5) and

\[
\mathcal{L}^\alpha \left( \gamma^\alpha \right) = \left. \frac{\partial \mathcal{L}^\alpha (g^\gamma, \gamma^\alpha)}{\partial g^\gamma} \right|_{g^{(e)}}
\]

Thus

\[
(\mathcal{D}_\mathcal{V} \sigma)^\alpha \left| \mathcal{P} \right. = \mathcal{D}_\mathcal{V} (\sigma) + \omega^\alpha (v) \mathcal{D}_\mathcal{V} \sigma \left. \right| \mathcal{P}
\]

Now choose \( \mathcal{P} = s(x) \) and \( \mathcal{V} = s_x X \), where \( X \) is a vector field in \( \mathcal{M} \); defining

\[
\mathcal{D}_x \sigma \left| x \right. = \mathcal{D}_{s_x X} \sigma \left. \right| s(x)
\]

and using (4.2.1), (4.3.1) we obtain

\[
(\mathcal{D}_x \sigma)^\alpha \left| x \right. = \mathcal{D}_x (\sigma) \left| x \right. + A^\alpha (x) \left. \mathcal{L}^\alpha (\sigma) \left| x \right. \right.
\]
Choosing $X = e_{\mu}$ and dropping the subscripts $\times$, we have the form

$$\mathcal{D}_{\mu} \varphi_{a} = \partial_{\mu} \varphi_{a} + A_{\mu} L^{i} \left( \varphi_{a} \right) \tag{4.3.8}$$

When $\mathfrak{X}$ is a vector space $\mathfrak{V}$ supporting a representation $\rho$ of $G$, as in §2.10, the functions $L^{i}$ become linear:

$$\mathcal{D}_{\mu} \varphi_{a} = \partial_{\mu} \varphi_{a} + A_{\mu} \left( T_{i} \right)_{c}^{b} \varphi_{b} \tag{4.3.9}$$

where $T_{i} = \frac{d}{dt} \rho(e^{\alpha t}) \mid_{t=0}$ are the representatives of the generators of the algebra, $e_{i}$, in the representation $\rho$, and $(T_{i})^{a}_{b}$ their matrix elements. If we regard $\varphi_{a}$ as the representative of the section $\sigma$, eq. $(4.3.8)$ is the local form of the implicit definition $(2.12.1)$.

The differential of $(2.8.1)$ and property b) in $(2.9.2)$ imply

$$\mathcal{D} \sigma \circ L^{-1}_{g} = L_{g} \circ \mathcal{D} \sigma$$

Therefore under a gauge transformation $(4.2.2)$ we have

$$\mathcal{D}_{x} \varphi_{a} \mid_{x} = \mathcal{D}_{s_{x} \sigma} \mid_{s_{x}(x)} = \mathcal{D} \sigma \left( L^{-1}_{g} \left( s_{x}X \right) \right) \mid_{s_{x}(x)} = L^{-1}_{g} \left( \mathcal{D}_{x} \varphi_{a} \mid_{x} \right) \tag{4.3.10}$$

Comparing with $(4.3.3)$ we see that indeed the covariant derivative has the same transformation properties of the field $\varphi_{a}$ (the derivative $\mathcal{D} \sigma$ as defined in $(2.12.1)$ is obviously formulated in gauge-independent terms). In the case of vector bundles this transformation law can be easily deduced from $(4.2.5)$ and $(4.3.9)$. 
4.4 We will assume now that \( \mathcal{X} \) is a homogeneous space equipped with a left \( G \)- and right \( H \)-invariant metric \( h \) (i.e. a riemannian homogeneous space as defined in § 3.6).

If \( \psi_A, \psi_B \) are local representatives of \( \sigma \) on the open sets \( \mathcal{U}_A, \mathcal{U}_B \), defined by 

\[
\sigma(x) = \psi_A(x, \varphi_A(x)) \quad \forall x \in \mathcal{U}_A
\]

and 

\[
\sigma(x) = \psi_B(x, \varphi_B(x)) \quad \forall x \in \mathcal{U}_B
\]

then (2.4.3) implies 

\[
\psi_A(x) = \varphi_A(x) \varphi_B(x) \quad \forall x \in \mathcal{U}_A \cap \mathcal{U}_B.
\]

The covariant derivative of these local representatives transforms as in eq. (4.3.10), therefore

\[
h(\partial_{\psi_A}, \partial_{\psi_B}) = h(\partial_{\varphi_B}, \partial_{\varphi_B})
\]

If \( \{ \psi_A \} \) is a partition of unity subordinate to the covering \( \{ \mathcal{U}_A \} \), we can then write the action for the sigma model

\[
S[\sigma] = \frac{1}{2} \sum_{\mathcal{U}_A} \int_{\mathcal{U}_A} (\partial_{\psi_A}, \partial_{\psi_B}) \cdot \eta
\]

where explicitly

\[
(\partial_{\psi_A}, \partial_{\psi_B}) = g^{\mu \nu} \partial_{\varphi_A}^\mu \partial_{\varphi_B}^\nu h_{\psi_A \psi_B}
\]

This could be written also in the more compact form

\[
S[\sigma] = \frac{1}{2} \int_{\mathcal{U}_B} (\partial_{\sigma}, \partial_{\sigma}) \cdot \eta
\]

Of course, one has to supplement this with the action for the Yang-Mills fields, which is given by (4.2.9).

A section \( \sigma \) which is a stationary point for \( S \) in the space of all field configurations will be called a harmonic section of \( \Sigma \).
4.5 As was mentioned in the previous section, the physical requirement of existence of smooth, global sections of the fibre bundle $\Sigma$ is a very strong requirement.

One has to apply here the following important theorem (1) p. 71, (3) vol. I p. 57: the bundle $\Sigma = (E, \pi, \mathcal{K}; \eta, \mathcal{H}, \mathcal{G})$ admits a cross section if and only if the associated principal bundle $\Pi = (P, \pi, \mathcal{K}; \eta, \mathcal{H}, \mathcal{G})$ admits a reduction to a principal bundle $\Pi' = (P', \pi', \mathcal{K}; H)$. Moreover, there is a one-one correspondence between sections $\sigma : \mathcal{K} \rightarrow E$ and reduced bundles. I will not give the proof here but will only show how a cross-section $\sigma$ defines the submanifold $P' \subseteq P$. First of all, it follows from the corollary of § 2.7 that $E = P \mod H$. Denote by $\tau$ the canonical projection $P \rightarrow E$; then $P'$ is defined as the inverse image of the section $\sigma$ under $\tau$ (here we identify the bundle $\Pi'$ with its image under the injection $\mathcal{I}$ of § 2.6):

$$P' = \tau^{-1}(\sigma(\mathcal{K})) = \{ p \in P | \tau(p) = \sigma(\pi(p)) \}$$

(4.5.1)

From the commuting diagrams

\[
\begin{array}{ccc}
P & \xrightarrow{\pi} & \mathcal{K} \\
\downarrow{\tau} & & \downarrow{\pi} \\
E & \xrightarrow{\pi} & \mathcal{K}
\end{array}
\quad \quad \quad
\begin{array}{ccc}
P' & \xrightarrow{\pi'} & \mathcal{K} \\
\downarrow{\tau} & & \downarrow{\pi} \\
E & \xrightarrow{\sigma} & \mathcal{K}
\end{array}
\]

(4.5.2)

it is easily seen that $\Pi'$ can also be regarded as the pull-back of the principal bundle $(P, \tau, E)$ on $\mathcal{K}$ through $\sigma$. Indeed, the total space of $\sigma^*(P \rightarrow E)$ is given by

$$\{ (x, p) \in \mathcal{K} \times P | \sigma(x) = \tau(p) \}$$

(4.5.3)
and the projection of $\sigma^*(P \to E)$ maps $(x,p)$ to $x$. But
$\sigma(x) \in \tau(p)$ implies $x = \pi(p)$ and therefore (4.5.3) coincides
with (4.5.1).

Since homotopic maps induce $\mathcal{H}$-isomorphic bundles, the
topology of the reduced bundle $\Pi'$ depends only on the topology
of $\Pi$ and on the homotopy class of the section $\sigma$.

The fibre bundle $\Sigma$, associated to $\Pi$, is also associated
to $\Pi'$ and thus has to be an $\mathcal{H}$-bundle. This is the result of
a theorem (exercise p. 72) which ensures that the fibre bundle with
fibre $G/H$ associated to $\Pi$ is $\mathcal{H}$-isomorphic to the fibre
bundle with fibre $G/H$ associated to $\Pi'$. For the total spaces
we have

$$P \times \mathcal{A} G/H = E \approx E' = P' \times \mathcal{A} G/H$$

The principal map of the second term is related to the principal
map of the first one by

$$X'(p',y) = (p',y) H = (f(p'),y) G = X(f(p'),y)$$

where $f$ is the injection $P' \to P$.

In conclusion, requiring $\Sigma$ to admit sections forces the
gauge group to break down from $G$ to $H$. This is the phenomenon
of spontaneous symmetry breaking: the action functional (4.4.3)
is invariant under $G$ but the solutions are invariant only under
$H$. With regards to the problem of classification, it is clear
that we are not interested in those fibre bundles which do not
admit cross-sections. Thus, for a given space $\mathcal{A} = G/H$ we can
obtain any bundle $\Sigma$ over a spacetime $\mathcal{H}$ admitting sections,
by choosing a principal $H$-bundle over $\mathcal{H}$ (eventually prolonging
it to a $G$-bundle) and taking the associated fibre bundle with
fibre $\mathcal{A}$. The classification of principal $H$-bundles makes
use of suitable topological invariants (the characteristic classes);
it is treated in detail in (21).
The symmetry breaking of the previous section is physically unescapable since there is no solution to the field equations which has the symmetry group $G$ and thus the $G$ symmetry can never manifest itself. It is therefore interesting to study the relation between the nonlinear sigma model and the related linear theory, which admits solutions with symmetry group $G$ (for instance, the constant zero solution).

We shall see now how a linear theory (the Higgs model) has the nonlinear sigma model as a limiting case and then it will appear how the Higgs model is obtained from the nonlinear sigma model by relaxing a constraint. An explicit example will be given in the next paragraph. Let us introduce a vector bundle

$$\alpha = (A, \pi, \kappa; \mathcal{Y}, \rho(u))$$

associated to $\Pi$, having the vector space $\mathcal{Y}$ as typical fibre and the image of $G$ under the representation $\rho : G \rightarrow GL(\mathcal{Y})$ as structure group. A scalar locally $\mathcal{Y}$-valued field over $\kappa$, that is a field whose configurations are the sections of $\alpha$, $\phi : \kappa \rightarrow A$ will be called a Higgs field. All the considerations leading to the action $S[\sigma]$ go through also in this case, so we could write $S[\phi]$ just substituting $\phi$ for $\sigma$ in (4.4.3). However, due to the flatness of $\mathcal{Y} = F^n$ as a manifold, what we would obtain in this case is a free (noninteracting) theory; since an interacting theory (like the nonlinear sigma model) cannot obviously be obtained as a limit of a free theory, we add to the action a potential term describing a (nongeometric) interaction:

$$S[\phi] = \int_\kappa \left[ \frac{1}{2} (\nabla \phi, \nabla \phi) + U(|\phi|) \right] \eta$$  \hspace{1cm} (4.6.1)

(*) Vector bundles are bundles (in general not fibre bundles!) whose fibre have a vector space structure. For our purposes it is sufficient to regard them as fibre bundles with typical fibre $\mathcal{Y} = F^n$, in the notation of § 3.10.
where $\nabla$ denotes the covariant derivative (4.3.9), $(\cdot, \cdot)$ the scalar product deriving from the metric $g$ in $T^*\mathcal{V}$ and an inner product $\langle \cdot, \cdot \rangle$ in $\mathcal{V}$, and $\lvert \phi \rvert$ is the norm of $(\langle \phi, \phi \rangle)^{1/2}$.

Suppose that $U(\lvert \phi \rvert)$ is such that the locus of its minima form a connected submanifold $\mathcal{N} = \mathcal{G}/H$ embedded in $\mathcal{V}$, and that $\mathcal{N}$ is invariant under the action of $\mathcal{G}$ (in other words, $\mathcal{N}$ is an orbit of $\mathcal{G}$ in $\mathcal{V}$). Also, let $\eta$ be the metric in $\mathcal{N}$ induced from the metric in $\mathcal{V}$ by the embedding (the metric in $\mathcal{V}$ regarded as a euclidean space $\mathbb{F}^n$, derives from the inner product $(\cdot, \cdot)$ by transporting vectors at a point to the origin). Without loss of generality, we can shift the potential in such a way that $U(\mathcal{N}) = 0$.

Then $\alpha$ admits an invariant subbundle ($\S$ 2.6) which is (eventually weakly) associated to $\Pi$ and therefore, by the construction theorem, is $\mathcal{N}$-isomorphic to $\Sigma$.

In the following we identify $\mathcal{E}$ with its image under the injection $f: E \to f(E) \subseteq \mathfrak{A}$. It is clear that any section of $\Sigma$ is also a section of $\alpha$, so that any configuration of the sigma model is also a configuration for the related Higgs model, but the converse is not true. Therefore the configuration space $\Gamma(\Sigma)$ is a subspace of $\Gamma(\alpha)$ (observe that $\Gamma(\alpha)$ is a vector space, unlike $\Gamma(\Sigma)$). There is no similar result concerning solutions: for instance if $\phi_0$ is a solution for the sigma model (and thus a configuration of the Higgs model), perturbations of $\phi_0$ in the direction of the normal of $\mathcal{N}$ change the action also to first order and thus $\phi_0$ cannot be a solution for the Higgs model. In fact it can be seen that the only common solutions are those configurations in $\Gamma(\Sigma)$ for which

$$\mathcal{D}\sigma = 0 \quad (4.6.2)$$

(*) This is always possible due to a theorem of Mostow and Palais, see (41)(42).
Coming now to the Yang-Mills field, let $\omega$ be a connection in the principal bundle $\Pi$, describing a gauge field minimally coupled to $\phi$ through the covariant derivative $D$.

As was shown in the preceding paragraph, any section of $\alpha$ which is also a section of $\Sigma$ gives rise to a reduction of the principal $G$-bundle $\Pi$ to a principal $H$-bundle $\Pi'$.

It is known then (Kobayashi & Nomizu vol. I, p. 88) that the restriction of $\omega$ to $P'$ defines an $\mathfrak{h}$-valued connection $\omega'$ in $\Pi'$ if and only if

$$D\phi = 0$$

This condition selects a special class of solutions for the Higgs model (and for the sigma model) which are simply minima of the potential.

In view of what was said before, these are the only solutions of the Higgs model which give a symmetry breaking by means of the mechanism of § 4.5. There are other mechanisms, for instance in the case of the nonabelian monopole; in these cases the field $\phi$ does not belong to $\Gamma(\Sigma)$ and the connection in the reduced bundle is not obtained by restriction.

4.7 In order to make things a bit more concrete, I will discuss in this paragraph the case of the grassmannian model with $U^k = G_k(F^n)$. The same discussion goes through, with little changes for $U^k = S G_k(F^n)$. The associated linear model has $\mathfrak{g} = F^{n_k}$ and $P$ is the direct product of $k$ times the fundamental representation of $U_F(m)$; the components of the Higgs field are arranged in a $m \times k$ matrix $\phi$. The action is

$$S[\phi] = \sum_{\phi} \left[ \frac{1}{2} g^{\mu \nu} \operatorname{Tr} (\partial_\mu \phi)^\dagger (\partial_\nu \phi) + U(\phi) \right]$$

$$U(\phi) = \sum_{\mu} \left( \frac{\lambda}{4} (\phi^* \phi - \alpha^2) - \frac{\lambda}{4}  \right) + \frac{\mu^2}{2} \mu \phi^* \phi + \frac{\mu^4}{4\lambda}$$
Where $\overline{D}_\mu$ is the $U_F(m)$ covariant derivative and $\lambda$ and $\alpha^2$ are constants (multiplied by the unit matrix $I_k$). $\lambda$ is as usual the coupling constant and $\alpha^2 = -\mu^2/\lambda$, where $\mu$ is the mass. If $\mu < 0$ the potential has a local maximum of $\lambda\alpha^2/4 = \mu^2/4\lambda$ for $\phi = 0$ and a locus of minima at $\phi^2 = \alpha^2$; by (3.11.1), this is a Stiefel manifold $V_k(F^n)$ embedded in $F^{n^2}$.

In order to see the transition from the model defined by (4.7.1) to the corresponding nonlinear sigma model, let us introduce $k^2$ auxiliary fields arranged in a $k \times k$ matrix and rewrite the potential as

$$U(\phi', \alpha) = -\text{Tr} \alpha(\phi'^2 - \alpha^2) - \frac{1}{\lambda} \text{Tr} \alpha^2.$$

$\alpha$ is not given a kinetic term. It is immediate to check that, using the Euler-Lagrange equation for $\alpha$, $U(\phi', \alpha)$ reduces to the $U(\phi')$ given in (4.7.1).

The nonlinear sigma model is obtained from the Higgs model letting $\lambda \to \infty$ while keeping $\phi$ fixed; $U(\phi')$ is meaningless in this limit, but $U(\phi', \alpha)$ becomes

$$-\text{Tr} \alpha(\phi'^2 - \alpha^2) \quad (4.7.2)$$

which is no longer a potential, but rather a constraint, since now variation with respect to $\alpha$ yields

$$\phi'^2 = \alpha^2.$$

So far, we have a bundle of Stiefel manifolds. In order to obtain a bundle of Grassmann manifolds we have to enlarge the isotropy group from $U_F(m-k)$ to $U_F(m-k) \otimes U_F(k)$, introducing as in § 3.14 the $U_F(k)$ covariant derivative $\overline{D}$; denoting $\overline{D}$ the $U_F(m) \otimes U_F(k)$ covariant derivative (the first group refers to the structure of the bundle $\Sigma$, the second to that of the Stiefel bundle) we can rewrite the action for the Grassmannian
\[
S[\phi] = \left[\frac{1}{2} g^{\mu \nu} \text{Tr} (\partial_\mu \phi) (\partial_\nu \phi) - \text{Tr} \alpha (\phi^T \phi - \alpha^2)\right] \cdot \eta
\]  
(4.7.3)

\[
\partial_r \phi = \partial_\mu \phi - \phi (\phi^T \partial_\mu \phi) + A_\mu \phi
\]

We have now a bundle of Stiefel bundles. Referring to (2.7.1) we have \( G = U_F(m), H = U_F(m-k) \otimes U_F(k), K = K_0 = U_F(k) \). The requirement that \( \Sigma \) admit sections forces the gauge group to break down to \( U_F(k) \otimes U_F(m-k) \).

Since the Stiefel bundles are universal for the classical groups, it would be interesting at this point to study the universality properties of the bundles (2.7.1) in the present case, in other words to see whether and under what conditions could the analysis of ref. (11) be applied also to the twisted sigma models.
References

22. A. Trautman "Elementary introduction to fibre bundles and gauge fields" Warsaw University (1978).
25. O. Klein; Z. Phys. 27, 895 (1926).