COMMENT ON THE QUARK MASSES IN THE SU(5)×U(1) MODEL
DERIVED FROM THE 4D FERMIONIC SUPERSTRING

G.K. Leontaris
CERN -- Geneva

ABSTRACT

In this letter, the problems of the up quark and neutrino mass matrix, as well as the lepton non-conservation are discussed within the $N = 1$ supersymmetric SU(5)×U(1) model derived from the four-dimensional fermionic superstring.
Nowadays, it is widely recognized that superstring theories appear to be the only candidates for a unified theory of all known interactions [1]. Thus, during the last three years much effort has been devoted to the construction of realistic superstring models in four dimensions using either fermionic or bosonic formulations [2,3].

Although string theories are very promising and seem to incorporate all the known particles, the road from the Planck scale down to low-energy physics is not an easy one. Thus it is important to answer the question whether superstring theories can lead to a realistic low-energy gauge group that is consistent with the well-known world of elementary particles and their interactions.

In this paper, we will concentrate our attention on an \( N = 1 \) supersymmetric model based on the group \( SU(5) \times U(1) \) which can be derived from the four-dimensional formulation of superstrings.

This model has been reported [4,5] as an attempt at model building using the fermionic formulation of string theories in four dimensions. Although this model does not seem to be a unified one, like the standard \( SU(5) \), it has some very nice features which make it quite attractive. Indeed, first of all, the fermion fields as well as the Higgs fields needed to break the group down to the standard \( SU(3)_C \times SU(2)_L \times U(1)_Y \) belong only to the \( 5, \overline{5}, 1, 10 \) and \( \overline{10} \) representations of \( SU(5) \). Secondly, the doublet-triplet mass splitting is solved easily in the above model by a missing partner mechanism. Indeed, one can use the 10-dimensional Higgs representation of \( SU(5) \) to give a large mass to the triplet components of 5 and \( \overline{5} \) Higgs fields, while the doublet components remain massless at the first stage of symmetry breaking. The latter are those which will realize the second stage of symmetry breaking from \( SU(3)_C \times SU(2)_L \times U(1)_Y \) down to \( SU(3)_C \times U(1)_{EM} \).

Finally, we notice here the absence of the adjoint representation as well as any higher ones. However, as we will soon realize, one encounters some defects of the above "GUT" model [4] when trying to obtain the necessary Yukawa couplings from the 4D fermionic superstring [5]. To be more precise, the fermionic formulation informs us which Yukawa couplings can be present in the superpotential. In our case, one can have all the necessary Yukawa couplings to give masses to the fermions, but not those which can give the necessary Kobayashi-Maskawa mixing between the first and the other two generations of quarks [5]. In principle, one could generate the necessary Yukawa terms by modifying the boundary conditions which generate the string model, but one can always do it at the price of the appearance of some new extra Higgses.

Here, we will make an attempt to generate the necessary Kobayashi-Maskawa mixing in the above string model. We will see that this can be done assuming one-
A coordinate chart for a manifold \( X \) is a set
\[ \{ (x^1, \ldots, x^n) \} \]
of elements of \( \mathbb{R}^m \) such that the \( C^\infty \) map \( x \mapsto (x^1(x), \ldots, x^n(x)) \) of \( x : U \to \mathbb{R}^n \) is onto, and \( C^\infty \) invertible. Given such a coordinate system, \( \mathcal{O}(X) \) can be described more explicitly in a theory that ties up with tensor analysis (i.e., elements of \( \mathcal{O}(X) \) as \( r \)-fold, skew-symmetric tensors.) Choose an \( r \times t \) coordinate system.

The \( \omega \) and \( \omega' \) are elements of \( \mathcal{O}^r(X) \). Hence, their "differentials" \( d\omega \) are elements of \( \mathcal{O}^{r+1}(X) \). We can then use the exterior multiplication to build up higher order elements.

\[ d\omega = \omega_1 \wedge \cdots \wedge \omega_r \]

The axioms of \( \wedge \):
\[ d\omega + d\omega' = -d(\omega \wedge \omega') \]

The \( \omega \) and \( \omega' \) commute:
\[ \omega \wedge \omega' = \omega' \wedge \omega \]

One proves now that the elements \( u \in \mathcal{O}^r(X) \), \( r=1,2,\ldots \), can be written uniquely in the form:
\[ u = u_1 \cdots u_r \]

with coefficients \( (u_1, \ldots, u_r) \) which are elements of \( \mathcal{O}(X) \), and which depend skew-symmetrically on their indices. This representation, and the following algebraic rules:
\[ d(u_1 \cdots u_r) = du_1 \wedge u_2 + \cdots + (-1)^{r+1} u_1 \cdots u_r \wedge du_r \]
determine the action of \( d \) completely.
and we see that in principle "differential geometry" can be "supersimplified".

The prototype for this is work by H. Cartan [29] that was enormously influential for
the work of my generation, but is not well-known nowadays. The "superspace-
structure" was recently put together by various physicist to, from the mathematician's point of view, an intriguing variant of H. Cartan's ideas. I will now elaborate a bit with some formal definitions.

2. AN ALGEBRAIC GENERALIZATION OF DIFFERENTIAL GEOMETRY. SUPERSPACE AND "QUANTUM

GEOMETRY" AS SUPERGROUPS

**Definition.** A set \( Q \) will be called a differential form algebra if it has
the following properties:

a) \( Q \) is a vector space, say, over the real numbers as field of scalars,

b) \( Q \) is integer-graded, i.e., as a direct sum

\[ Q^0 \oplus Q^1 \oplus Q^2 \oplus \cdots \]

of subspaces labelled by a non-negative integer.

c) \( Q \) is an algebra, i.e., a bilinear map \( Q \times Q \rightarrow Q \) is given. Denote this
product map by

\[ \omega_1 \wedge \omega_2 = \omega_1 \wedge \omega_2 \]

It satisfies:

d) \( Q^n \wedge Q^m \subseteq Q^{n+m} \)

e) it is associative, i.e.,

\[ \omega_1 \wedge (\omega_2 \wedge \omega_3) = (\omega_1 \wedge \omega_2) \wedge \omega_3 \]

for \( \omega_1, \omega_2, \omega_3 \in Q \)

For \( r = 0 \) this means that \( Q^0 \) is a subalgebra. Denote the multiplication in \( Q^0 \) simply as just position \( (f, g) \rightarrow fg \).

f) there is given a linear map \( d : Q \rightarrow Q \) such that \( d(Q^n) \subseteq Q^{n+1} \). Hence:

\[ d(fg) = df \wedge g + f dg \]

for \( f, g \in Q^0 \)

\[ d^2 = 0 \]

i.e.,

\[ d(dw) = 0, \quad \text{for all } w \in Q \]

\[ df(a) = df \wedge a \]

if \( df = 0 \), \( f \in Q^0 \)

\[ d(df) = (-1)^{m+1} \wedge df \]

if \( m \in Q^0 \), \( f \in Q^0 \) and if \( df = 0 \)

b) \( Q \) is generated by \( Q^0 \), and the operations \( \wedge, d \) plus the vector space
structure, i.e., every element \( w \in Q^1 \) can be written as a linear combination of elements of the form

\[ f(\omega_1) \wedge g(\omega_2) \wedge \cdots \omega_k \wedge df 

for \( f, \omega_1, \omega_2, \cdots, \omega_k \in Q \)

Note especially that we do **not** assume that multiplication in \( Q^0 \) --which
determines everything else--is commutative, as it was for standard "differential
geometry". This gives us an additional degree of freedom, which can, in fact,
be exploited to fit in (at least) two additional cases of interest for physics--
superspace and what I call [30] "quantum" (boson) differential geometry. To see
such possibilities, suppose \( f, g \) are two elements of \( Q^0 \) with:

\[ fg \neq gf \]

and \( c \in \mathbb{R} \). (Suppose the algebra structure on \( Q^0 \) has a unit element "1", so that
\( c \) is just a). Apply (3.1)

\[ (df)g + f dg \neq (dg)f + gf = 0 \]

Suppose (as an ansatz) that we want to require that \( df, dg \) be "independent". (3.5) then requires that:

\[ (df)g + fdg = 0 \]

Apply \( d \) to both sides of (3.6) using (3.1) - (3.4):

\[ df \wedge dg \neq 0 \]

Hence:

\[ df, dg \text{ anticommute (commute) if } (-1) \text{ occurs in (3.6).} \]

Thus, we see that \( Q, d, f \) will be determined if we assume that \( Q^0 \) is
generated (as an algebra) by pairs of elements \((f, g)\) satisfying commutation
relations like (3.6).
\[ (x^i) \text{ in } \mathbb{R}^n \text{ are elements of } \mathbb{R}^n, \quad 1 \leq i \leq n, \quad \lambda \in \mathbb{R}, \]

satisfying
\[ \begin{align*}
  x^i x^j &= x^j x^i ; & y^i y^j &= y^j y^i ; & x^i y^j &= y^j x^i.
\end{align*} \tag{1.10} \]

Their differentials must then satisfy the following commutation relations (3.11):
\[ \begin{align*}
  dx^1 \wedge dx^2 &= - dx^2 \wedge dx^1, \\
  dx^1 \wedge dy^a &= - dy^a \wedge dx^1, \\
  dy^a \wedge dy^b &= + 1 dy^b \wedge dy^a.
\end{align*} \tag{3.11} \]

This enables us to write an element \( \omega \in \Omega^0 \) in the following form:
\[ \omega = \sum_{1 \leq i_1, \ldots, i_k \leq k} \frac{d^k \omega_{i_1 \ldots i_k}^X}{i_1! \ldots i_k!} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge \cdots \wedge dx^X \]

\[ \wedge dy^{a_1} \wedge \cdots \wedge dy^{a_k} \wedge \cdots \wedge dy^X, \tag{3.12} \]

with \( X = a \), and the \( i \)'s are elements of \( \mathbb{R}^n \) which depend skew-symmetrically on the indices \( 1 \), symmetrically on the indices \( a \).

Now, one can use the commutation relations (3.10) to write an \( f \in \Omega^0 \) in the form
\[ f = f_1(x^i) a^1 + f_2(x^i) a^2 + \cdots. \tag{3.13} \]

Using this in (3.12) shows that \( \omega \) can be written as sums of the form
\[ \omega_{i_1 \ldots i_k}^X \wedge dy^{a_1} \wedge \cdots \wedge dy^{a_k} \wedge \cdots \wedge dy^X \wedge \cdots \wedge dy^X, \]

\[ \wedge dy^{a_1} \wedge \cdots \wedge dy^{a_k} \wedge \cdots \wedge dy^X, \tag{3.14} \]

where the \( a_i \)'s are differential forms of the usual type in terms of the variables \( x \).

This is the algebraic prototype of the physicist talking about using "superspace" to put together in a "unified" way "fields" of the usual type of various spins.

**Example 2:** Quantum differential geometry.

Suppose that \( \mathbb{C}^n \) is generated (as an associative algebra) by elements that we label
\[ \{ x^i, p_i \}, \quad 1 \leq i, j \leq n, \]

which satisfy the following commutation relations:
\[ \begin{align*}
  x^i x^j &= x^j x^i ; & p_i p_j &= p_j p_i ; \\
  p_i x^j &= x^j p_i = h_i^j.
\end{align*} \tag{3.15} \]

In a constant, which is, of course, to be identified (up to \( 2 \pi \)) with \( i \hbar \) and \( \eta \).

Thus relations (3.15) are essentially the Heisenberg commutation relations. (However, it is important to note that they are not to be considered as "quantum relations," but an commutation relations for an "abstract" associative bracket algebra.)

We can now assume that \( \mathbb{C}^n \) is part of a differential form algebra \( \Omega^0 \),

\[ n > 0. \]

As an Ansatz, let us say that the \( dp_i \) are linearly independent (in the \( \Omega^0 \)-module sense). As we have seen, this condition, plus the commutation relations (3.15), determines the commutation relations between their differentials.

In fact, they take the following form:
\[ \begin{align*}
  dx^i \wedge dx^j &= - dx^j \wedge dx^i, \\
  dx^i \wedge dp_j &= - dp_j \wedge dx^i, \\
  dp_k \wedge dp_j &= - dp_j \wedge dp_k, \\
  x^i dp_j &= dp_j x^i, \\
  etc.
\end{align*} \]

In words, the differentials \( dx^i, dp_j \) satisfy exactly the same commutation relations as they do "classically," i.e., when \( \hbar = 0 \).

The usual quantum-mechanical story may now be obtained by asking for irreducible representations of \( \Omega^0 \) by Hermitian operators. However, it is interesting to note that it is not necessary to do so, even to form the "dynamics." Thus, it is possible to completely decouple (as far as differential-geometric properties go) the structure of quantum mechanics from the need to use a Hilbert space (which always seemed to me to be the weakest link in the mathematical structure of quantum mechanics—which becomes especially acute when field theories enter the picture).

To see this, it is most convenient to follow Cartan's brilliant (but difficult and esoteric) approach to classical mechanics presented in his book, *Lectures on Exterior Differential Forms* [29]. Introduce another variable \( \tau \), which commutes with everything. Let \( \hbar \) be an element of \( \mathbb{C}^n \). Set
\[ \omega = dp_1 \wedge dx^1 + \cdots + dp_n \wedge dx^n \tag{3.11} \]

Write:
\[ db = \eta_1 dx^1 + \eta_n dp_n \tag{3.17} \]

The \( \eta_1, \eta_n \) are defined by this relation. In case \( \hbar = 0 \), they are, as usual, the partial derivatives.

\[ \omega = dp_1 \wedge dx^1 + \cdots + dp_n \wedge dx^n \tag{3.11} \]

Write:
\[ db = \eta_1 dx^1 + \eta_n dp_n \tag{3.17} \]

The \( \eta_1, \eta_n \) are defined by this relation. In case \( \hbar = 0 \), they are, as usual, the partial derivatives.
First of all, let us suppose that $x$ is a point topologically null on $Y$ and $X$. This means that there are coordinate open sets $U_y \subseteq Y$, $U_x \subseteq X$ such that

$$a(U_y) = U_x$$

Let $(x^i)$, $1 \leq i, j \leq n$, be a coordinate system of functions on $U_x$. Let

$$x^i(x^j)$$

be the pull-back into functions on $Y$, i.e.,

$$x^i(x^j(y)) = x^i(y)$$

We now suppose that the following condition is satisfied:

There are additional sets $(y^a)$, $1 \leq a, b \leq n$, of smooth functions on $U_x$ such that the functions

$$(x^i(x^j), y^a)$$

form a coordinate system for $U_x$.

For notational convenience, we leave off the $x^i$, and just label this coordinate system (which is "specially adapted" to the fiber space) as $(x^1, y^a)$.

These coordinates are not, of course, uniquely defined. They can be changed to

$$x^1 = x^1(x)$$
$$y^a = y^a(x, y)$$

The special "triangular" nature of the allowed coordinate changes (4.1) is what is typical of the "fiber space" situation.

However, it is our aim to think as much as possible in coordinate-free terms, and use such coordinates as little as possible. Given a point $x_0 \in X$, the set $Y(x_0)$ = $x^{-1}(x_0)$ of all points of $X$ which map into $x_0$ is called the fiber of the fiber space above the point $x_0$. 

1. FIBER SPACES

Let $X$, $Y$ be manifolds, and let $x: Y \rightarrow X$ be a (smooth) map. I want to say what it means for $x$ to define a "fiber space". There are various technical ways of doing this. I will choose the one which appears closest to the way the theory of fiber spaces is involved in the theory of "gauge fields".

$$h^i = \frac{\partial}{\partial x^i}$$
$$h^i = \frac{\partial}{\partial x^i}$$
The fibers are submanifolds of \( Y \); the \( y^i \) are then coordinates of these fibers. It is this geometric picture—the manifold \( Y \) "fibered" by giving a submanifold at each point—which is the most important. (In fact, there is a more global differential geometric concept—a foliation—which has this property. They are also important, but not at the moment lie outside our domain.)

We want to endow these fibers with geometric structures. One convenient way is to prescribe a group acting on each fiber. One does this by postulating a group \( G \) assigned to each \( x \in X \), and an action

\[
(x, y) \rightarrow g(x, y)
\]

of \( G \) by transformation on \( Y \). (One must also prescribe \( G \) as "varying smoothly with \( x \)" in an appropriate way. We ignore such technical details.) Call this structure \( (G^x) \), a bundle of groups. Giving this action on \( Y \) is called giving the fiber space \( Y \) a bundle of structure groups. Often, each group \( G^x \) is isomorphic to a single group \( G \). One then speaks of the fiber space as having a structure group \( G \).

In physics, one most often (but not always!) encounters fiber spaces which are

\[
Y = X \times F
\]

where \( Y \rightarrow X \) is the Cartesian product map.

Physicists call this a gauge transformation.

**2. CONNECTIONS**

Let \( Y \rightarrow X \) be a fiber space map. Let \( \omega(X) \) and \( \Omega(X) \) denote the algebra of differential forms of \( X \) and \( Y \). Any differential form on \( Y \) which is a linear combination (with coefficients in \( \Omega(X) \)) of forms in \( \Omega(X)(x) \) is said to be vertical. \( \omega^X \) denotes the set of vertical \( r \)-forms.

Focus on \( \omega^X \). In the "adapted" coordinates \( (x^1, x^2) \) described in section 1, they are those which can be written in the form

\[
\omega = \omega(x_1, x_2) dx^1
\]

Notice that there is an uniquely defined "complementary" set of one-forms. In these coordinates \( \omega \) forms such a set, but it is not at all unique. For example, the change of variables (4.2) creates a new complementary set \( \omega_1 \) with

\[
\omega_1 = \omega + \theta \frac{\partial}{\partial x^1}
\]

What we must do is to prescribe a complementary set. In fact, this turns out to be precisely the geometric object called a connection.

**Definition.** A connection for the fiber space \( (Y, X, G) \) is a set \( \mathcal{S} \) of one-forms on \( Y \) such that:

\[
\mathcal{S} \text{ is a } \mathcal{D}^0(Y)\text{-module, i.e., if } \omega, \eta \in \mathcal{S}, \eta \in \mathcal{D}^0(Y), \text{ then}
\]

\[
\omega + \eta, \omega, \eta \in \mathcal{S}
\]

\[
\mathcal{D}^1(Y) \text{ is a direct sum (as an } \mathcal{D}^0(Y)\text{-module)}
\]

of the vertical one-form \( \omega^X \) and \( \mathcal{S} \).

Here is what this means in the adapted coordinates \( (x^1, x^2) \). \( \mathcal{S} \) has a basis (as a \( \mathcal{D}^0(Y)\)-module) consisting of a set of one-forms \( \omega^a \) of the form:

\[
\omega^a = \omega^a - u^a
\]

where the \( u^a \) are vertical one-forms. They are said to be the components of the connection in this adapted coordinate system.

In order to see this relevance to physics, let us see how the components change when the adapted coordinates are changed. For example, suppose that

\[
y^a = f^a(x, x)
\]

where \( x \) are new "fiber" coordinates. Then,

\[
\omega^a = \omega^a - u^a
\]

Insert (5.2) into (5.1):

\[
\omega^a = \omega^a - u^a + f^a - (f^a) c^c - e^a d^i x^i
\]
\[ \omega^a = \partial x^a - \omega^a_b \partial x^b \]  
(5.4)

(5.5) exhibits the change of the components when the fiber coordinates are changed in an \( x \)-dependent way, i.e., as what physicists traditionally call a "gauge transformation".

**Linear Connections and Gauge Transformations**

The ideas developed above are valid for both linear and nonlinear situations. In general, they are simplest to understand in the linear situations.

Consider a connection described by formula (5.1). It is linear (in these coordinates) if the \( \omega^a \) are of the following form:

\[ \omega^a = \partial x^a - \Gamma^a_{bc}(x) \partial x^b \partial x^c \]  
(5.5)

Choosing \( y^a \) as a vector \( X \) (6.1) can be written more conveniently in vectorial form:

\[ y^a = m^a_b(x) y^b \partial x^c \]  
(5.6)

where \( m^a_b(x) \) are a matrix \( (m \times m) \)-valued functions of \( x \).

We can now consider the changes in the "adapted" coordinates \( (x, y) \) which preserve the linear property of the vectorial connection form \( \partial X \). Among these are \( \omega^a \)-collining transformation, sometimes called "linear, local gauge transformations".

\[ X = \partial y = y^a \partial x^a \]  
(5.7)

We can readily find the expression for the connection in the new coordinates:

\[ \frac{dy^a}{dt} = \frac{dx^a}{dt} + \omega^a_{bc} \frac{dx^b}{dt} \partial x^c \]  
(5.8)

The \( \omega^a \) are the bases for the connection form in the "new" coordinates \( (x, y) \).

Notice the typical "non-tensorial", "gauge-like" transformation law:

\[ F^a = \omega^a_b \partial x^b \]  
(5.3)

It is the second term on the right hand side of (5.9) which is "non-tensorial".

6. **Parallel Transport and Curvature**

Let \( y: Y \to X \) continue to be a fiber space. Assume also that \( Y \) has a global adapted coordinate system \( (y^a, x^i) \). Suppose a connection is given, defined by a basis of one-forms:

\[ e^a = \partial y^a = y^a_{i}(y, x) \partial x^i \]  
(6.1)

The \( y^a \) are arbitrary functions of these variables.

Suppose that we give a curve \( t \to x(t), a \leq t \leq b, \) in the base \( X \). Let \( x^i(t) \) be its components in these coordinates. Substitute these functions \( x^i(t) \) into (6.1), obtaining the following one-forms:

\[ e^a = \partial y^a = y^a_{i}(y(t), x(t)) \partial x^i \]  
(6.2)

The variables \( y^a \) are still to be thought of as functions of the parameter variable \( t \) along with the curve.

Now, let us ask for \( y \) as a function of \( t \), so that the forms \( e^a \) become zero. Thus \( y^a(t) \) must then satisfy the following differential equations:

\[ \frac{dy^a}{dt} = \Gamma^a_{bc}(y(t), x(t)) \frac{dx^b}{dt} \partial x^c \]  
(6.3)

The usual existence and uniqueness theorem for ordinary differential equations (assuming, for simplicity, that it can be applied over all of the interval \( a \leq t \leq b \)) implies that there is a unique solution of \( t \to y^a(t) \) (given \( x(t) \)) beginning at a given point \( y^a(0) \). Let us say that a curve satisfying (6.2) is horizontal.

We can summ up as follows:

**Theorem 6.1.** Let \( Y: Y \to X \) be a fiber space map, and let \( \mathcal{F} \) be a collection of one-forms on \( Y \) which defines a connection. Given a curve \( t \to x(t), a \leq t \leq b, \) in \( X \), and a point \( y_0 \in Y \) such that \( y_0^a = x^a(a) \), there is a unique horizontal curve in \( Y, t \to y(t) \) (found locally by solving a time-dependent system of ordinary differential equations) such that:

\[ y(y(t)) = x(t), \quad \text{for all } t. \]
This curve is called a horizontal lift of the base curve \( t \mapsto x(t) \).

Now, fix the curve \( t \mapsto x(t) \) in the base space \( X \). For \( x \in X \), let

\[
P(x) = \tau^{-1}(x) \in \text{fiber above } x.
\]

For each point \( y \) of \( P(x(t)) \), there is then a unique horizontal lift beginning at \( y \) at \( t = a \). Call this lifted curve, and we obtain a point \( y' \) of \( P(x(b)) \) which is a function solely of \( x \) and the base curve. This defines a map

\[
P(x(a)) = f(x(b))
\]

between fibers, called parallel transport along the curve. It is the basic geometric operation of connection theory.

In general, this transformation of fibers will be path dependent. However, let us suppose that it is independent of path. Returning to the local coordinates \((x^i, y^a)\) used above, we see that there will be functions \( y^a(x) \) such that \( t \mapsto x(t) \) is a curve in \( X \), its parallel transport is the curve \( t \mapsto y^a(x(t)) \).

In particular, \( y(x) \) will satisfy the Pfaffian differential equations

\[
dy^a = \frac{\partial}{\partial t} y^a(x, t) \, dt.
\]

In the classical literature, these are called Pfaffian equations. The condition that they can be solved in this way is traditionally known as "complete integrability," i.e., that the relations obtained by applying exterior derivative \( d \) to both sides, and the relation \( d^2 = 0 \), are satisfied identically. In modern differential geometry, this is done (following Cartan [39, 40]) by saying that the forms \( \omega^a \) define a completely integrable Pfaffian system in the sense that there are one-forms \( u^a \), such that

\[
d^u \omega^a = u^a \wedge \omega^b \text{.}
\]

In other words, it can be shown that the existence of one-forms \( u^a \) satisfying (6.5) is equivalent to parallel transport in the fiber space being independent of path.

Let us now work out the conditions for (6.4) using (6.1).

\[
d\omega^a = -\frac{\partial \omega^a}{\partial x^i} \, dx^i
\]

\[
= -\frac{\partial}{\partial x^i} \omega^a \wedge dx^i - \frac{\partial}{\partial x^j} \omega^j \wedge dx^i = \left( \frac{\partial}{\partial x^i} \omega^a \wedge \frac{\partial}{\partial x^j} \omega^j \right) \wedge dx^i
\]

\[
= -\frac{\partial}{\partial x^i} \left( \omega^a \wedge dx^i \right) \wedge dx^i
\]

Now, the forms

\[
\omega^a = \left( \frac{\partial}{\partial x^i} \omega^a \right) \wedge dx^i
\]

can only be written in form (6.5) if they vanish identically, since the one-forms \( \omega^a, dx^i \) are linearly independent. These two forms are called the curvature forms. It can be verified (a tedious calculation) that the curvature transforms tensorially whereas the \( \omega^a \) do not. In fact, the curvature can be defined in a completely coordinate-free way. It then turns out to be a more exotic object that we have considered up to now—a two-differential form on \( X \) with values in the vector bundle over \( X \) whose fibers are the vertical vector fields, i.e., the vector fields on \( Y \) which are tangent to the fibers of \( \tau \).

The curvature becomes a less formidable object in case the connection is linear. To work out this case—which is the basic one needed for physics—
In the matrix notation used earlier:

\[ g = g(x) = H(x) x \, dx^4 \]

Then

\[ \delta g = -\delta H x \, dx^4 - H_4 \delta x \, dx^4 \]

\[ \delta g = -\delta H x \, dx^4 - H_4 \delta x \, dx^4 \]

\[ \varepsilon = \frac{\delta g}{\delta x^4} = \frac{\partial H}{\partial x^4} + H_4 \delta x_4 \]

\[ \varepsilon = \frac{\delta g}{\delta x^4} = \frac{\partial H}{\partial x^4} + H_4 \delta x_4 \]

This is now the object whose vanishing ensures complete integrability, i.e., path-independence (locally) of parallel transport. Of course, the object occurring in (6.7)

\[ \delta_{ij} = \frac{\delta g}{\delta x^4} = \frac{\partial H}{\partial x^4} + H_4 \delta x_4 \]

(6.9)

is very familiar to "gauge" physicists—the curvature tensor of the Yang-Mills field.

The fact that it is the commutator that appears in (6.9) is significant—it indicates that the Lie algebras should get into the game. We will consider the appropriate way of doing this after a short diversion to consider the topic of "Lie algebra-valued one-forms", which is of interest in its own right.

LIE ALGEBRA-VALUED ONE-FORMS AND THE CURVATURE OPERATOR

The analytical part of the theory of connections "with structure group" involves the topic described by the title of this section.

Forget about fiber spaces for the moment, and consider a manifold \( X \).

Denote \( \Omega^r(X) \), \( r = 0, 1, \ldots \), the \( r \)-th degree differential forms. \( \Omega^0(X) \) itself forms a (commutative, associative) algebra.

Let \( G \) be a real Lie algebra. The tensor product

\[ \Omega^r(X) \otimes G \]

is defined as a \( G \)-valued differential form on \( X \). Explicitly, an element

\[ \omega \in \Omega^r(X, G) \]

is a linear combination

\[ \omega = \omega^1 \otimes A_1 + \cdots + \omega^p \otimes A_p \]

(7.1)

with

\[ A_1, \ldots, A_p \in G, \quad \omega^1, \ldots, \omega^p \in \Omega^r(X) \]

Define an operation

\[ \mathcal{D} \colon \Omega^r(X, G) \to \Omega^r(X, G) \]

by means of the following formula:

\[ \mathcal{D} \omega = \omega^1 \otimes A_1 + \cdots + \omega^p \otimes A_p + \frac{1}{2} \sum_{a,b=1}^p \left( \omega^a \mathcal{L}_a \mathcal{L}_b - \mathcal{L}_b \mathcal{L}_a \mathcal{L}_a \right) \]

(7.2)

Note that \( \mathcal{D} \) is a non-linear differential operator (if \( G \) is a non-Abelian Lie algebra) called the curvature operator. (We will see why it is given this "geometric" name in the next section.)

\( \mathcal{D} \) can be put into a more familiar form (to physicists) if the \( A \) are chosen as a basis for the Lie algebra \( G \). Let \( c_{ab}^c \) be the structure constants of the Lie algebra, i.e.,

\[ [A_a, A_b] = c_{ab}^c A_c \]

(7.3)

Suppose that

\[ \omega^a = \mathcal{L}_a \]

(7.4)

Then,

\[ \mathcal{D} \omega^a = \mathcal{L}_a \mathcal{L}_a \mathcal{L}_a + \frac{1}{2} \sum_{b=1}^p \left( \omega^a \mathcal{L}_b \mathcal{L}_b - \mathcal{L}_b \mathcal{L}_a \mathcal{L}_a \right) c_{ab}^c \mathcal{L}_c \]

(7.5)

with

\[ \mathcal{L}_a \mathcal{L}_a \mathcal{L}_a = \frac{1}{2} \mathcal{L}_a \mathcal{L}_a \mathcal{L}_a \]

(7.6)

which is again the familiar "Yang-Mills field" formula (reducing to one half of "Maxwell's equations" if \( G \) is Abelian).

The vanishing of \( \mathcal{D} \) has a special geometric significance. Let \( G \) be a Lie group whose Lie algebra is \( G \). Let \( s^a \) be the left-invariant one-forms on \( G \).
which are "dual" to the bases \( A \) of \( G \). (The \( n^a \) are called the Cartan vectors.) Then, \( \mathcal{D}^g = 0 \), i.e.,

\[
\frac{\partial}{\partial y^a} \frac{1}{2} \frac{\partial}{\partial \phi^a} A \phi^a = 0
\]

implies that there is a map

\( \phi : X \rightarrow G \)

such that

\( \phi^*(\eta^a) = \psi^a \).

We shall see that this has a special importance in the theory of solitons.

3. CONNECTIONS WITH STRUCTURAL GROUPS

Return to the fiber space \( Y \rightarrow X \), with adapted coordinates \( (x^1, y^a) \). Let \( G \) be a Lie group of transformations on the variables \( y \).

**Definition.** A connection for the fiber space has \( G \) for structure group if it is defined in these coordinates by one-forms \( \psi^a \) of the type

\[
\psi^a = \psi(x) = \epsilon(x) dx^a + \frac{1}{2} \frac{\partial}{\partial \phi^a} A \phi^a
\]

with \( \epsilon(x) \in \Omega^1(X, G) \).

Let us see what this means analytically and geometrically. Suppose that \( \phi_a \)

\( \phi^a \in \Omega^1(x, y^a) \)

is a basic for \( \psi^a \), with

\[
\psi^a = \phi^a = \epsilon^a(x) dx^a + \phi^a(x) dx^a
\]

and

\[
\frac{\partial}{\partial \phi^a} A \phi^a = \frac{\partial}{\partial \phi^a} A \phi^a(x) \frac{\partial}{\partial \phi^a} A \phi^a(y) \frac{\partial}{\partial \phi^a} A \phi^a(z)
\]

Then,

\[
\frac{\partial}{\partial \phi^a} A \phi^a = \frac{\partial}{\partial \phi^a} A \phi^a(x) \frac{\partial}{\partial \phi^a} A \phi^a(y) \frac{\partial}{\partial \phi^a} A \phi^a(z)
\]

Suppose that \( t = x(t) \) is a curve in the base space \( X \). Looking at (8.4), we see that the differential equations for the parallel-transported curve are

\[
\frac{\partial}{\partial t} \frac{\partial}{\partial \phi^a} A \phi^a = \frac{\partial}{\partial \phi^a} A \phi^a(x(t)) \frac{\partial}{\partial \phi^a} A \phi^a(y(t))
\]

(8.5)

which are associated with the action of the group \( G \) on \( \phi^a \).

**Theorem:** The connection defined via formula (8.1) by the Lie algebra-valued one-form \( \psi^a \) is flat, i.e., parallel transport is (locally) independent of the path, if and only if the curvature two-form \( \mathcal{D}^g \) is identically zero.

Here is a main result which may be proved by a straightforward computation.

**Theorem:** The connection defined via formula (8.1) by the Lie algebra-valued one-form \( \psi^a \) is flat, i.e., parallel transport is (locally) independent of the path, if and only if the curvature two-form \( \mathcal{D}^g \) is identically zero.

In Part III of E. Cartan's **Collected Works** [22] many examples are described of geometrically interesting connections with various structure groups. For example, \( X = R^n \) if \( Y = R^n \times R^n = T(X) \) (the tangent bundle to \( X \)) then there are four choices of \( G \) of prime geometric interest:

- **a)** \( G = \) group of affine transformations in \( R^n \);
- **b)** \( G = \) group of orthogonal transformations on \( R^n \);
- **c)** \( G = \) group of (nonlinear) projective transformations on \( R^n \), i.e., the space of maps \( R^n \rightarrow R^n \) which map hyperplanes into hyperplanes;
- **d)** \( G = \) group of (nonlinear) conformal maps, i.e., the maps \( R^n \rightarrow R^n \) which preserve angles.

The corresponding connections are associated with affine, Riemannian, projective, and conformal geometry. Geometrically, this whole business is a marvelous realization of Klein's " Erlanger program" for studying "geometry" with a group-theoretic motivation and unification. Up to now, only the Riemannian connections of general relativity have been of interest for physics—one must now ask what the connections that appear in Yang-Mills and soliton theory have to do with geometry! However, I will not pursue these general geometric ideas further in this overview—keep in mind that they are sitting in the background while we develop the analytical machinery needed to understand the physics of " gauge fields" and "solitons".

9. GAUGE TRANSFORMATIONS OF LIE-ALGEBRA VALUED ONE-FORMS

Let \( X \) be a manifold, \( G \) a Lie group, \( \mathfrak{g} \) its Lie algebra. Let \( \mathfrak{g}(X, G) \) be the space of maps \( X \rightarrow G \). If \( X \) is fixed, denote \( \mathfrak{g}(X, G) \) as \( \mathfrak{g} \). (In physics, three choices of \( X \) are used in various contexts: \( X = R, R^3 \) (space), \( R^4 \) (space-time).) Elements of \( \mathfrak{g} \) denote an element of \( G \) by \( g \).
Let $G$ be a group of $m \times n$ matrices. Let $x \to y(x)$ be a map of $X \to G$.

Set:

$$y' = y(x)^{-1} y$$

or

$$y = y(x)y'$$

Then,

$$dy' = (dy)y' + ydy'$$

Hence,

$$\delta = dyy' + gdy' - M_1(x)g dy'dx^i \tag{9.1}$$

$$q^{-1}g = dy' + q^{-1} dy' - q^{-1} M_1(x)g dy'dx^i \tag{9.2}$$

Set:

$$u = M_1(x)y dx^i \tag{9.3}$$

$$u' = q(x)^{-1} M_1(x)g dy'dx^i - q(x)^{-1} M_1(x)dy(x) \tag{9.4}$$

$$\delta = dy - uy \tag{9.5}$$

Then,

$$\delta' = q^{-1} \delta \tag{9.6}$$

$\delta'$ represents the transformed connection under the gauge transformation. (9.6)

show how the Lie algebra-valued one-forms on the base transform under this gauge transformation. Now, (9.4) can be written in a matrix-independent way as:

$$u' = Ad (g(x)(u) - q(x)(\eta)) \tag{9.7}$$

where $\eta$ is the left Lie algebra-valued (Maurer-Cartan) one-form on $G$.

We can now abstract this material out of the connection-theoretic context in which it was derived. We can say that two Lie-algebra valued one-forms $\omega$, $\omega'$ on a manifold $X$ are **gauge-equivalent** if there exists a map $g : X \to G$ such that

$$\omega' = Ad g(x)(\omega) - q(x)(\eta)$$

in the usual way. Particularly notice how $u$ and $u'$ can have quite different singularity structure, the inhomogeneous term, i.e., the second term on the right hand side of (9.8), can possibly cancel a singularity present in $u$. This possibility is particularly important in the study of "instantons" and "monopoles."
We now want to construct "gauge invariants". The obvious one is the curvature two-form described earlier,

$$\Omega^w = du + \frac{1}{2} (ad_u)^w$$ \hspace{1cm} (9.9)

It is readily seen that it transforms tensorially under a gauge transformation, i.e.,

$$\Omega'^w = \text{Ad} g(x)^w (\Omega^w)$$ \hspace{1cm} (9.10)

In particular, the Yang-Mills field equations can be constructed from these invariants. If \( G \) is semi-simple and if \( X \) has a Riemannian metric, there is a "duality" operator

$$\ast : \Omega \rightarrow \Omega$$

on \( G \)-valued two-forms. (\( \ast \) is an \((n, r)\)-form, where \( n = \dim X \).) This is an obvious generalization of the "Hodge" duality operator [41,42]. Then, the covariant derivative with respect to \( u \) can be applied to \( \ast \Omega \)

$$D_u \ast \Omega$$ \hspace{1cm} (9.11)

It is an \((m, 1)\)-form, with values in \( \mathcal{G} \). The vanishing of (9.11) then expresses the "Yang-Mills" equations; they are, by their very construction, "gauge invariant", i.e., \( \ast \Omega \) satisfies

$$D_u \ast \Omega = 0$$

and if \( u' \) is gauge-related to \( u \), then it too satisfies these equations. This isomorphism of the system of nonlinear differential equations (9.12) under the "infinite parameter" group of all gauge transformations is what makes them so important physically, but it also creates substantial mathematical difficulties. A general theory of higher spin "gauge-invariant" equations is presented briefly in [43].

Now, it is possible to construct invariants of gauge transformations whose existence gives a necessary condition that singularity may be "gauged" away. These invariants are called characteristic classes by mathematicians. We will now briefly indicate how they may be defined.

### 15. Characteristic Classes

Suppose \( G \) is a Lie algebra, \( X \) is a manifold, and \( u \) is a Lie-algebra valued one-form on \( X \). Its curvature

$$\Omega = du + \frac{1}{2} \{ [u, u] \} : D_u$$

defines, for each \( x \in X \), a bilinear mapping

$$\Omega(x) : \mathcal{G}_x \times \mathcal{G}_x \rightarrow \mathcal{G}_x$$

which is skew-symmetric. We can then construct a quadrilinear mapping

$$\Omega(x) \ast \Omega(x) : \mathcal{G}_x \times \mathcal{G}_x \times \mathcal{G}_x \times \mathcal{G}_x \rightarrow \mathcal{G}_x$$

by the following formula:

$$\Omega(x) \ast \Omega(x)(v_1, v_2, v_3, v_4) = \Omega(x)(v_1, v_2) \ast \Omega(x)(v_3, v_4)$$ \hspace{1cm} (10.1)

We can then follow this with any linear mapping

$$\alpha : \mathcal{G}_x \ast \mathcal{G}_x \rightarrow \mathbb{R}$$

to obtain a mapping

$$\Omega : \alpha(\mathcal{G}(x) \ast \mathcal{G}(x)) : \mathcal{G}_x \times \mathcal{G}_x \times \mathcal{G}_x \times \mathcal{G}_x \rightarrow \mathbb{R}$$ \hspace{1cm} (10.2)

This formula will define a differential form on \( X \) if \( \alpha \) is symmetric, i.e.,

$$\alpha(A_1, A_2) = \alpha(A_2, A_1)$$ \hspace{1cm} (10.3)

for \( A_1, A_2 \in \mathcal{G} \).

Suppose in addition that \( \alpha \) satisfies the following condition:

$$\alpha(\text{Ad} g(\mathcal{A}_1) \ast \text{Ad} g(\mathcal{A}_2)) = \alpha(\mathcal{A}_1, \mathcal{A}_2)$$ \hspace{1cm} (10.4)

for \( A_1, A_2 \in \mathcal{G} \).

(One says that \( \alpha \) is invariant under the adjoint representation of \( \mathcal{G} \).) One can then prove that \( \Omega \) is a closed differential form, i.e.,

$$d \Omega = 0$$ \hspace{1cm} (10.5)

(this is a consequence of the Bianchi identities) and:

$$\Omega \ast \Omega$$ is invariant under gauge transformations applied to \( u \). \hspace{1cm} (10.6)
11. DIFFERENTIAL EQUATIONS DEFINED BY VANISHING OF THE CURVATURE.

Continuing with $X$ as a manifold, $\mathfrak{g}$ as a Lie algebra,

$$\omega: T(X) \to \mathfrak{g}$$

as a Lie-algebra valued one-form on $X$. The curvature form

$$\Omega = d\omega + \frac{1}{2} [\omega, \omega]$$

is then a $\mathfrak{g}$-valued two-form on $X$. If $\{e_a\}, 1 \leq a, b \leq n$, is a basis for $\mathfrak{g}$, then

$$\Omega = \sum_{a,b} \omega_a \wedge \omega_b$$

where $\omega_a$ are scalar valued two-forms on $X$.

Following E. Cartan [40] one can now define the exterior differential system

on $X$ generated by the $S$-submanifolds

$$\phi: Z \to X$$

are said to be integral submanifolds of this system if the forms $S$ are zero when restricted to the submanifolds, i.e., if

$$\phi^* S = 0$$

for such a submanifold, the Cartan-Maurer operations hold, i.e.,

$$d\phi^* (\omega) + \frac{1}{2} [\phi^*(\omega), \phi^*(\omega)] = 0$$

Then there is a map

$$\sigma: X \to G$$

(whose $G$-Lie group whose Lie algebra is $\mathfrak{g}$) such that:

$$\phi^* (\omega) = \text{pull back under } \phi \text{ of the left-invariant form on } G.$$
to be combined to make "more complicated" cases. These formulas were first found by Zakari for the Sine-Gordon equation and by Wahlquist and Estabrook for Korteweg-de Vries [61].

Finding the prolongations is closely related to the "inverse scattering" structure. Work by Kadomtsev and Savel'ev [67, 68] gives the most convenient (so far) way to write down the prolongations. A glance at these papers should show the reader the unified role that the Lie group SL(2,R) plays in the study of the various equations (Sine-Gordon, Korteweg-de Vries, nonlinear Schrödinger, ...), which in fact are the unique examples with the same properties. (I believe there are other "nice" examples yet to be found associated with other groups—even infinite dimensional ones!)

12. SL(2,R) PROLONGATIONS, PROLONGATION AND BÉNARD TRANSFORMATIONS

Let G be the Lie group SL(2,R) of 2x2 real matrices. Its Lie algebra is the set of 2x2 real matrices of trace zero:

\[ A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11} \end{pmatrix} \]

Let \( \mathfrak{g} \) be a manifold. A \( \mathfrak{g} \)-valued one-form is then a matrix

\[ u = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & -u_{11} \end{pmatrix} \]

of scalar one-forms. Suppose \( X \) is the space of variables \((x,t,u_1,u_2,...)\).

Let us now work in the context of the "theory of nonlinear waves"; \((x,t)\) are independent variables \( x=x(t) \) then stand for derivatives.) Consider \( u \) of the form

\[ u = A \, dx + B \, dt \]

where \( A \) and \( B \) are 2x2 matrices of trace zero consisting of functions of the variables \( u_1, u_2, ... \). Set

\[ \theta : \mathfrak{g} \to \mathfrak{g} \quad \theta = \frac{1}{2} \, \{ [u, u] \} \]

\[ = \frac{1}{2} \, \{ du_1 A \, dx + du_2 A \, dt + \frac{1}{2} \, (Adx + Bdt, Adx + Bdt) \} \]

\[ = \frac{1}{2} \, \{ du_1 A \, dx + du_2 A \, dt + [A, B] \, dx \, A \, dt \} \]

\[ A(u(x,t),...,t) = B(u(x,t),...,t) = [A,B] \, u(x,t),...,t) \]

(12.4)

(12.4) then defines a system of partial differential equations which we call "PDE" for \( u(x,t) \). Requiring that these equations coincide with a given system of equations (e.g., Korteweg-de Vries, Sine-Gordon) then determines differential equations for \( A \) and \( B \). These equations (which sometimes can be solved) play the role in soliton theory analogous to the "Yang-Mills" equations in the theory of gauge fields.

Once such \( A,B \) are found, Bénard transformations can be constructed. Here is the method of Estabrook and Wahlquist [66, 67] for doing this.

Let

\[ \theta = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \]

(12.1)

be a co-vector. Set:

\[ \theta = dy - uy \]

Let \( X \) be the space of variables \( y_1, y_2 \). \( \theta \) are then one-forms on \( X \times Y \), which define a linear connection. \( \theta \) are the curvature forms. Now, look for a map

\[ \beta : X \times Y \to X \]

(12.2)

such that

\[ \beta^*(\theta) \text{ lies in the Grassmann algebra ideal generated by } \theta, \beta. \]

\( \beta \) is called a Bénard map. It enables one to generate new solutions of the underlying partial differential equations for \( u(x,t) \). Namely, if \( u(x,t) \) is one solution of PDE, if \( y(x,t) \) is defined as follows

\[ y_x = A(u(x,t),...,t) \]

\[ y_t = B(u(x,t),...,t) \]

(12.5)

then

\[ \beta(x,t,u(x,t),u_1(x,t),...,y(x,t)) = (x,t,u(x,t),...,t) \]

(12.5) are called the prolongation equations. \( \beta \) then maps each solution of the
we obtain a system (consisting of (2.5) and the PDE's for \(u\)) into a solution \(u'\) of the PDE. Converting formulas for the Krichever maps are given in a paper by Komon and Wadati [68]. In these formulas \(\beta\) is determined by a simple natural map

\[(u,y) \mapsto u'\]  

We shall (engaging in a bit of "abus de langage") call this the "Krichever map", and denote it by \(\beta\).

13. SOLITONS

As in Section 12, suppose a \(\mathbb{C}\)-valued one-form \(u\) is given such that a given set of equations for \(u(x,t)\) are the conditions for

\[du + \frac{1}{2} [du, du] = R = 0\]  

Suppose also that

\[\beta : X \times Y \to X\]  

is given as a Krichever map. Start off with the solution

\[u(x,t) \in \mathbb{C}\]  

The integrability conditions (12.4) are then

\[\{A(0), B(0)\} = 0\]  

The prolongation equations (12.5) take the following form:

\[y'_x = A(0)y\]
\[y'_t = B(0)y\]

Equations (13.2) can be solved explicitly:

\[y(x,t) = \exp(xA(0))(y(0,t)) = \exp(xA(0) + tB(0))(y(0,0))\]

Thus the map \(S : (\mathbb{R}^2) \to \mathbb{C}\) determined by the "rest" solution of the PDE is the map

\[(x,t) \mapsto \exp(xA(0) + tB(0))\]

Notice that its image in \(\mathbb{C}\) is a one-parameter subgroup. Then we can set

\[u'' = \delta(y, u)\]

where \(u = 0\), \(y\) is given by (13.3), \(u'\) is called the one-solution solution of the PDE.

The next prolongation equations are:

\[y'' = A(u',...,u)\]
\[y'' = B(u',...,u)\]

The

\[u'' = \delta(y', u')\]

are called the two-solution solutions.

One continues in this way to define sequences \((0, y, u', u'', y', y'', ..., )\), called the soliton ladder. Now, it looks like these equations for these functions become progressively harder to solve. In fact, in all the cases we know about, there is a "miracle" and they are all solvable in terms of explicit rational functions of the functions of the form (13.3). This is a consequence of the superposition formulas of Hirota type.

14. THE SOLITON LADDER FOR KORTWEG-DE VRIES

To illustrate the generalities described in Sections 12 and 13, let us turn to the Korteweg-de Vries (abbreviated K-dV) equation, which is the main example which has motivated this line of research in nonlinear waves and solitons. I will follow the formulation developed by Komon and Wadati [68] which is very convenient. Set:

\[A = \begin{pmatrix} 1 & u \\ -1 & -1 \end{pmatrix}\]
\[B = \begin{pmatrix} -4u^2 - 2uu_x - u_{xx} - 2ku_x - 4k^2 u - 2u^2 \\ 4k^2 + 2uu_x + 4k^2 + 2uu_x \end{pmatrix}\]

(13.3) \(A\) is a parameter (essentially the "eigenvalue parameter" of the "inverse scattering problem"). The PDE satisfied by \(u\) is K-dV:

\[u_{tt} + 6uu_x + u_{xxx} = 0\]  

The equation (14.1), together with the prolongation equations
form a composite PBK system for \((u,v)\) which is "completely integrable" in the classical sense used, e.g., by E. Cartan [39,40]. (Workers in "nonlinear waves" often use the term "completely integrable" for another property; it is often confusing because the K-dV equation is also "completely integrable" in their sense; and there is no obvious relation between the two concepts. I think it is best to reserve the term for its classical version, especially because the newer use of the term is rather confusing and ill-defined.)

The Backlund map is given by the following formula:

\[
\delta(u,v) = -\left(u + 2\frac{y_1}{y_2} + 4\lambda \frac{y_1}{y_2}\right)
\]

(14.3)

We can start the calculation of the soliton ladder off with the choice \(v = 0\). Then

\[
A(0) = \begin{pmatrix}
\lambda & 0 \\
-1 & -\lambda
\end{pmatrix}
\]

\[
E(0) = \begin{pmatrix}
-4\lambda^2 & 0 \\
4\lambda^2 & 4\lambda^2
\end{pmatrix} = -4\lambda^2 A(0)
\]

\[
Y_x = A(0)y
\]

\[
y_c = -4\lambda^2 A(0)y
\]

\[
u^* = -2\frac{y_1}{y_2} \frac{2}{y_2} + 4\lambda \frac{y_1}{y_2}
\]

(14.10)

Now, equations (14.5) take the form

\[
Y_{1,x} = -4y_1
\]

\[
Y_{2,x} = -y_1 - 4y_2
\]

hence:

\[
\left(\frac{y_1}{y_2}\right)_x = \frac{y_1(\lambda y_1 - y_1 - 2y_1)}{y_2}
\]

(14.2)

\[
= \frac{2\left(y_2^2 + y_1^2\right)}{y_1} = \frac{1}{2} u^*
\]

(14.7)

Hence, the second prolongation equations are given as:

\[
y^*_x = \begin{pmatrix}
\lambda & -2 \frac{y_1}{y_2} \\
0 & 0
\end{pmatrix} y^*
\]

(14.8)

\[
y^*_c = \left(\lambda, 0\right)
\]

(14.9)

(The time-derivatives eventually become so complicated in the K-dV theory that it is difficult to write them out.)

is then the two-soliton solution of K-dV. Approach directly, the equations (14.8) and (14.9) are easy. However, the superposition formula (proved in this case by Wahlquist and Estabrook [41]) enables one to express the general solution of (14.8) in terms of explicit natural functions of the general solution of (14.5) for \(\lambda\) and another value of \(a\), say \(a\). These formulas then show that the two-solitons "decouple" as \(x \rightarrow -\infty\) into a sum of "one-solitons" with shifted phases. This leads to the remarkable and basic physical property of solitons—they interact in a very

(14.4)

(14.5)

No one knows if such properties can persist for nonlinear PDE's in more than one space variable, but I am confident that the basic "Backlund-prolongation" formalism must generalize in some way: if so, it would be very useful in a widespread spectrum of nonlinear physics and engineering problems.

15. EXTERIOR DIFFERENTIAL SYSTEMS, GENERALIZED CONSERVATION LAWS, AND LINKS

Up to now, I have emphasized the more traditional approach to prolongations—Backlund transformations, solitons, etc. It is important to realize that it can be described beautifully also in terms of E. Cartan's theory of exterior differential systems [40, 29, 16], which is, in a sense, the "pure" geometric theory of differential equations. If there is a reasonable extension of the theory to
Let $X$ be a manifold. An exterior differential system is a collection of differential forms, denoted by $\mathcal{E}$, with the following properties:

$$\mathcal{E} + \mathcal{E} \subset \mathcal{E}$$

$$\omega_1 \wedge \omega_2 \in \mathcal{E}$$

for $\omega_2 \in \mathcal{E}$, $\omega_1$ an arbitrary differential form on $X$

$$\omega \in \mathcal{E} \text{ if } 0 \in \mathcal{E}$$

Let $\mathcal{E}$ be such a system. An integral submanifold is a map between manifolds:

$$\phi: Z \rightarrow X$$

such that:

a) The induced map $\phi_\ast: T(Z) \rightarrow T(X)$ on tangent bundles is one-to-one (this is the "submanifold" condition).

b) $\phi^\ast(\omega) = 0$, for all $\omega \in \mathcal{E}$.

Partial differential equations of any type with $n$ independent variables, give rise to exterior differential systems with the property that there is a one-to-one correspondence between their solutions and certain integral submanifolds with $Z = \mathbb{R}^n$.

Typically, there might be certain limiting and degenerate integral submanifolds which do not arise from solutions. In [15] I have explained how these limiting cases can be used to geometrize certain ideas of "singular perturbation theory.

The most extensive and useful discussion of this is in Cartan's book [40]. Courant's book [7] is also very useful as a guide to the classical literature. Cartan usually worked with $\mathcal{E}$'s generated by one-forms. (These are also called Pfaffian systems.) E. Cartan, E. Cartan and Wahlquist have shown [41, 28] that it is also very useful and important to work with $\mathcal{E}$'s generated by two-forms. In fact, the first step in constructing "prolongations" and "Backlund transformations" for a given system of PDE's is to write it in this way.

Let $\mathcal{E}$ be a given exterior differential system. A conservation law is a differential form $\omega$ such that

$$\omega \in \mathcal{E}$$

$$\partial_\lambda \omega \in \mathcal{E}$$

for $\lambda = (\alpha, \beta, \gamma)$, for a description of the relation between this concept and the physicists' notion of "conserved currents".

In general, $\omega$ is a form of arbitrary degree $r$. (For field theories with $n$ independent variables, typically $r = n-1$.) Here is a way to geometrically generalize:

**Definition.** A generalized conservation law for $\mathcal{E}$ of degree $r$ is a set

$$\omega_j \in \mathcal{E}, 1 \leq j \leq m$$

such that

$$\omega_j \wedge \omega_j = 0, \omega_j \wedge \omega_k = 0$$

(15.2)

**Remark.** Here is a geometric meaning of this condition in the case $r = 1$.

Let

$$\phi: Z \rightarrow X$$

be an integral submanifold. (15.2) says that the "pulled-back" Pfaffian system

$$\phi^\ast(\omega_j) = 0$$

(15.3)

is completely integrable, i.e., there are functions $(z^i, \xi_j)$ on $Z$ such that

$$dz^i = \xi_j^i \phi^\ast(\omega_j)$$

(15.4)

In the Estabrook-Wahlquist theory [28] a simplified version of these relations is encountered. Again working with the case $r = 1$ (and we do not really know any interesting examples in higher degrees), suppose that the $\omega_j$ are of the following form

$$\omega_j \equiv - \frac{1}{2} C_{jk} \eta^k$$

(15.5)

where $(g_{ij})$ are the structure constants of a Lie algebra $G$. This means that if $G$ is a Lie group whose Lie algebra is $\mathfrak{g}$, there is a basis $\eta^i$ for the left-invariant ("Cartan-Nurowski") one-forms on $G$ such that

$$\eta^i \wedge \eta^j \wedge \eta^k = 0$$

Then (15.5) means that

$$dz^i = \frac{1}{2} C_{jk} \eta^j \wedge \eta^k \in \mathcal{E}$$

(15.6)

If (15.4) is satisfied, the $\omega_j$ are said to be a generalized conservation law for $\mathcal{E}$ with $G$ as structure group.
Let 
\[ \phi: \mathbb{R}^n \times X \]
be a submanifold of \( M \). Then (15.6) means that
\[ (15.7) \]
\[ d_\phi^*(f^1) = \frac{1}{2} C \delta^k_{i,j} \partial^n_{x_j} \phi^*(f^2) \wedge \phi^*(f^k) = 0. \]
This implies (if \( Z \) is simply connected) that there is a map
\[ \phi: \mathbb{R} \times G \]
such that
\[ \phi^*(f^1) = \phi^*(f^1). \]
(15.8)
Thus, we can assign (with the help of the generalized conservation law) to each integral submanifold \( \phi \) (i.e., to each solution of the underlying PDE) a map
\[ \mathcal{H} \times G. \]

Let \( H \) be a closed subgroup of \( G \). Now, instead of \( G \), consider the map \( \phi \) defined by the map \( \gamma \times G \), i.e., \( \gamma \wedge \partial^n_{x_j} \phi^*(f^2) \wedge \phi^*(f^k) = 0. \)

In the cases which are familiar, \( Z \) is \( \mathbb{R} \), parametrized, say, by variables \( s \) and \( t \). Denote a point of \( G/H \) as \( p \). Thus, \( \gamma \) is a map
\[ \gamma, p : \mathbb{R} \times G. \]
Suppose that:
\[ \lim_{t \to \infty} \gamma^t(p(x,t)) = p. \]
(15.9)
Then, for fixed \( s \), the curve
\[ x = \gamma^t(p(x,t)) \]
goes from \( p_1 \) to \( p_2 \) as \( x = t \). This defines an element of
\[ \mathcal{H}(G/H). \]
the fundamental group of \( G/H \).

Now relation (15.9) is more likely if \( G/H \) is compact. It is easy to guess what \( H \) must be for the usual equations (KdV, sine-Gordon, etc.). There are generalized conservation laws associated with \( SL(2,\mathbb{R}) \). Let \( H \) be the set of matrices of the form
\[ \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \]
where \( a \) and \( b \) are integers. Thus, we can define an integral topological invariant, which is \( \mathbb{Z} \)-valued. It would be interesting to get other Lie groups \( G \) into the game. (I believe there are interesting classes of \( G \)'s which are yet to be discovered associated with at least all the simple Lie groups.) If \( G \) is noncompact and simple, the generalization of what we did for \( G = SL(2,\mathbb{R}) \) is to use the Pontrjagin duality (19.11)
\[ G = \text{KAN}, \]
with \( K \) a maximal compact subgroup of \( G \) and \( \text{AaN} \) subgroups. (\( A \) is abelian, \( N \) nilpotent.) Set:
\[ H = \text{KAN}, \]
where \( H \) is a subgroup of \( K \). Then \( G/H \) is compact; topologically, \( G \) is \( K/H \).

Thus, knowing the topological properties of the compact group, it can readily be arranged that \( G/H \) has topological properties which are potentially useful for generalizations of the "link" idea. However, these interesting possibilities must await further progress in discovering new sort of generalized conservation laws.

16. SUMMARY AND FURTHER COMMENTS

"Modern" differential geometry has a characteristic flavor. It is really not all that "modern", but has its roots solidly in the 19th century, which was really much more of a Golden Age of Geometry than today. The work of Sophus Lie and Elie Cartan is the basis for most of what we do.) The basic objects are manifolds, vector fields, differential forms, fiber spaces, and connections. The mathematical formalism underlies much of science and engineering. (There are only unexploited possibilities of applying these ideas in areas outside of physics—systems and control theory, continuum mechanics, chemistry, biology, even economics.)

In this article I have concentrated on "soliton" theory as that area in which to illustrate the influence of these geometric ideas. (This is also the area in which I am working and which I know best.) However, it should be clear that the mathematics is closely related to that involved in the study of "causes fields", particularly the topics of "instantons" and "monopoles". (I have encountered among physicists an attitude that the theory of solitons is a curiosity of little significance for elementary particle physics. This may, in fact, turn out to be so (self-fulfilling prophecy?), but I am somewhat disturbed that people who express this put-down usually do not understand the full mathematical ramifications and structure of soliton theory, particularly the marvelous techniques (inverse scattering, Bäcklund transformation, Bianchi-style superposition, KdV-breather, soliton prolongations, etc.) that go along with it. What we are doing is developing new
mathematics especially adapted to the geometric structure of certain types of nonlinear partial differential equations—and that mathematics is closely linked to physics. I believe these techniques will turn out to be classics of 20th-century mathematics—comparable historically to elliptic functions, say, in the 19th century—and I want to see as wide an application to physics and engineering. I hope more physicists will invest the time and effort needed to understand the mathematical foundation of these ideas, which are so full of promise.

BIBLIOGRAPHY
