On the other hand, applying $\frac{\partial}{\partial q^i}$ to Hamilton-Jacobi gives the following "derived H-J equation:

$$(2) \qquad 0 = \frac{\partial^2 S}{\partial t\, \partial q^i} + \frac{\partial H}{\partial q^i}(t, dS_t) + \sum \frac{\partial H}{\partial p_j} \frac{\partial^2 S}{\partial q^j \partial q^i} .$$

Now $\frac{\partial^2 S}{\partial q^j \partial q^i} = \frac{\partial^2 S}{\partial q^i \partial q^j}$. Hence subtracting the last two equations gives 2

(with suitable arguments)

$$(3) \qquad \frac{\partial p_i c}{dt} = \frac{\partial H}{\partial q^i} .$$

This is the second half of Hamilton's equations.

In the last equation everything is to be regarded as a function of t. The "suitable arguments" required to make this the case are indicated <u>without</u> ambiguity by the diagram of the functions involved.

$$\begin{tikzcd}
 & I \times T^*C \arrow[r, "H"] & \mathbb{R} \\
I \arrow[ur, "(1,c)"] \arrow[r, "(1,b)"'] & I \times C \arrow[u, "dS"'] &
\end{tikzcd}$$

It should be possible to make a systematic use of such mapping diagrams to indicate which (composite) arguments we intended in equations (such as the Hamilton and Hamilton-Jacobi equations).

Now consider the converse part of the theorem. Take a solution b of the first Hamilton equations; then equations (1) above hold. By hypothesis (1) implies (3); subtracting, (2) holds along b. But by the existence theorem for ordinary differential equations, there is a solution b through each point of $I \times C$; hence (2) holds at any point $(t, q^1, \ldots, q^n)$. But (2) states that

$$\frac{\partial}{\partial q^i}(HJ(S)) = 0 \quad , \quad i = 1, \ldots, n,$$

where $HJ(S)$ denotes the left-hand side of the Hamilton-Jacobi equations.
ie $HJ(S)$ depends only on t, so
Therefore, there is a smooth function $\theta : \mathbb{R} \to \mathbb{R}$ with $HJ(S) = \theta \circ t : I \times C \to \mathbb{R}$. Take a function ψ such that $\frac{d\psi}{dt} = \theta$. Then it is not hard to see that

$$HJ(S - \psi) = 0.$$

This gives the conclusion of the theorem.

44. Transformation to Equilibrium.

Now let Y be an n-dimensional manifold, and suppose $S : \mathbb{R} \times C \times Y \to \mathbb{R}$ is a function such that everywhere

$$\det \left| \frac{\partial^2 S}{\partial q^i \partial y^j} \right| \neq 0$$

where q^i are coordinates for C and the y^j are coordinates for Y.
Let $\Theta : R \times C \times Y \to R \times T^*C$ be given by

$$p_i \Theta = \frac{\partial S}{\partial q^i}$$

$$t \Theta = t$$

$$q^i \Theta = q^i \quad , \quad i = 1, \ldots, n.$$

Thus the assumption on S is equivalent to saying that Θ is regular everywhere. Consider also the mapping $\Psi : \mathbb{R} \times C \times Y \to \mathbb{R} \times T^* Y$ given by

$$t\Psi = t$$

$$y^i\Psi = y^i$$

$$x_i\Psi = \frac{\partial S}{\partial y_i}\ , \quad i = 1, \ldots, n,$$

where the y^i are coordinates on Y and the x_i the corresponding (momentum) coordinates on T^*Y. This map is also regular everywhere. Hence we have the diagram

$$\begin{tikzcd}
R \times T^*C & & \mathbb{R} \times T^*Y \arrow[ll, dotted, "\chi"'] \\
& \mathbb{R} \times C \times Y = N \arrow[ul, "\textcircled{H}"] \arrow[ur, "\Psi"'] &
\end{tikzcd}\quad ,$$

and locally at least there is a map χ from one time dependent phase space to another (compare §26).

<u>Theorem</u>. Take $H: \mathbb{R} \times T^*C \to \mathbb{R}$. For each point $a \in Y$, suppose that $S: N \to \mathbb{R}$ satisfies the H-J partial differential equation for the function H. Let $c: \mathbb{R} \to \mathbb{R} \times T^*Y$ be a curve of the form $c(t) = (t, \text{const.})$. Then the curve χc satisfies the Hamilton equations for H on $\mathbb{R} \times T^*C$.

<u>Proof</u>. It will be more convenient to look at everything in N rather than in $\mathbb{R} \times T^*C$. To do this we pull everything back locally by $\textcircled{H}^{-1}$. Thus we are interested in the curve $\Psi^{-1}c$ and the function $H\textcircled{H}$.

Let x_i denote the coordinates as above. Then $x_i = \frac{\partial S}{\partial y_i}$ on N. Taking $\frac{d}{dt}$ of this, we get

$$0 = \frac{\partial^2 S}{\partial y^i \partial t} + \sum \frac{\partial^2 S}{\partial y^i \partial q^j}\frac{dq^j}{dt} + \sum \frac{\partial^2 S}{\partial y^i \partial y^j}\frac{\partial y^i}{\partial t}\ .$$

The third sum vanishes because $\frac{\partial y^i}{\partial t} = 0$. Since we are assuming

$$\frac{\partial S}{\partial t} + H = 0$$

on N (more precisely, we should write $H\Theta$ instead of H), applying $\frac{\partial}{\partial y^i}$ yields

$$\frac{\partial^2 S}{\partial t \partial y^i} + \sum \frac{\partial H}{\partial p^j} \frac{\partial^2 S}{\partial q^i \partial y^j} = 0,$$

which holds on the curve $\Psi^{-1}c$. Hence, again on the curve,

$$\sum \frac{\partial^2 S}{\partial q^i \partial y^j} \left(\frac{dq^j}{dt} - \frac{\partial H}{\partial p_j}\right) = 0.$$

Since the determinant of $\left(\frac{\partial^2 S}{\partial q^i \partial y^j}\right)$ was assumed to be non-zero, this means that

$$\frac{dq^j}{dt} = \frac{\partial H}{\partial p_j}$$

holds for all j. T his is the first Hamilton equation. By the previous theorem, we get the remaining half of the Hamilton equations.

For fixed t_o, the submanifolds of N of the form $t_o \times C \times Y$ have a symplectic form given by

$$\sum \frac{\partial^2 S}{\partial q^i \partial y^j} dq^i \wedge dy^j .$$

By the theorem proved in § 26 of Part I, the functions Θ and Ψ are symplectic mappings, whence the usual symplectic structures are taken on T^*C and T^*Y.

In the theorem just proved, the trajectories c in $\mathbb{R} \times T^*Y$ are constant in T^*Y. Hence one says that the map χ^{-1} of the theorem transforms the Hamiltonian H "to equilibrium".

45. Characteristics.

The previous results indicate a close relation between the Hamilton-Jacobi equation, a partial differential equation, and Hamilton's equations, a system of ordinary first order differential equations. This is a special case of the theory which relates a first order partial differential equation to its characteristics, which are solutions of a corresponding system of ordinary first order differential equations.

Sources: a) Courant and Hilbert, Methods of Mathematical Physics, II

b) Caratheodory, Calculus of Variations and PDE's of 1st order, Part I.

The case of the arbitrary first order equation will be reached in stages. We first consider the linear case, involving the following functions on a configuration space C:

$$\mathbb{R} \xrightarrow{c} C \xrightarrow[q^1,\dots,q^n]{u} \mathbb{R},$$

$$(1) \qquad \sum_{i=1}^{n} a_i \frac{\partial u}{\partial q^i} = bu + d \quad , \qquad \begin{cases} a_i : C \to \mathbb{R}, \\ b : C \to \mathbb{R}, \\ d : C \to \mathbb{R} \ . \end{cases}$$

Equation (1) for the linear case has a_i, b, d functions of position in C. The a_i determine a vector field $X = \sum_{i=1} \frac{\partial}{\partial q^i}$ on C which appears in a_i

the following coordinate independent form of (1): $L_X u = bu + d$. Call c a characteristic curve of the PDE (1) when

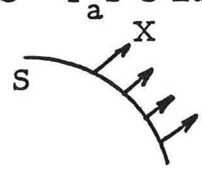
$$(2) \qquad \frac{d(q^i c)}{dt} = a^i \quad , \quad i = 1, \dots, n.$$

Thus the characteristics are the trajectories of the vector field X. In view of the definition of X, this equation can be written as

$$(2') \qquad \frac{duc}{dt} = bu + d.$$

First suppose that $b = d = 0$. Then if the function u satisfies the PDE (1), $L_X u$ it is constant along the characteristics of (1). More generally, for any b and d, the equation states that the values of u along a characteristic are determined by the value at any one (initial) point there. This suggests that we can obtain a solution u by taking initial values along a suitable set S, and then prolong these values by solving (2').

More explicitly, find a submanifold S of dimension n-1 transverse to X (i.e., with $T_aC = T_aS \oplus \mathbb{R}X(a)$ at each point a of S). According



to the basic theorem on the integration of (smooth) vector fields, the trajectories of X through S cover some neighborhood of S, determining on some neighborhood of any $\mathring{S}$ a unique function u which agrees on $\mathring{S}$ with some chosen $u_o : \mathring{S} \to \mathbb{R}$, and which satisfies (2) along characteristics (trajectories of X) on the neighborhood. Here $\mathring{S}$ is an open submanifold of S with compact closure in S. In local coordinates it is immediate that, for smooth u_o, the function u is smooth and satisfies (1). So we have found a local solutuon.

Next we consider a first order P.D.E. in an unknown u, of the form

$$(1) \qquad \sum_{i=1}^{n} a_i(u, q^1, \ldots, q^n) \frac{\partial u}{\partial q^i} = b(u, q^1, \ldots, q^n) \quad , \quad u = u(q^1, \ldots, q^n).$$

This is linear in all the partial derivatives of u, but not in u itself, hence is said to be quasilinear. We interpret the $q^1, \ldots, q^n$ as coordinates in an n-dimensional configuration space C, so that $u: C \to \mathbb{R}$. The equation thus has the form

$$(1') \qquad \sum_{i=1} a_i \frac{\partial u}{\partial q^i} = b, \qquad \begin{tikzcd} \mathbb{R} \times C \arrow[r, bend left, "a_1"] \arrow[r, "\vdots\, a_n"'] \arrow[r, bend right, "b"'] & \mathbb{R}, \end{tikzcd}$$

for given coefficient functions a_i and b.

We plan to reduce this to the previous case for a linear P.D.E. in an unknown $v: \mathbb{R} \times C \longrightarrow \mathbb{R}$ in one more variable, constructing the function u via its graph $\hat{u}: C \xrightarrow{\hat{u}} \mathbb{R} \times C$. Let $r: \mathbb{R} \times C \to \mathbb{R}$ be the projection on the first coordinate. We introduce a function $v: \mathbb{R} \times C \xrightarrow{v} \mathbb{R}$ defined by $v = u - r$. Now $\frac{\partial v}{\partial r} = -1$ and $\frac{\partial v}{\partial q^i} = \frac{\partial u}{\partial q^i}$ (while on hypersurfaces $v = 0$ we will have $a_i(u, q^i) = a_i(r, q)$ and $b(u, q) = b(r, q)$) so that (1) becomes

$$0 = \sum_{i=1}^{n} a_i \frac{\partial u}{\partial q^i} - b = \sum_{i=1}^{n} a_i \frac{\partial v}{\partial q_i} + b \frac{\partial v}{\partial r} = \Big(\sum_{i=1}^{n} a_i \frac{\partial}{\partial q_i} + b \frac{\partial}{\partial r} \Big) v.$$

This is a homogeneous linear P.D.E. in $v: \mathbb{R} \times C \to \mathbb{R}$. Its characteristics are thus given by a suitable vector field $\hat{X}$. Indeed we now define the vector field $\hat{X}$ on $\mathbb{R} \times C$ by $\hat{X} = \sum_{i=1}^{n} a_i \frac{\partial}{\partial q^i} + b \frac{\partial}{\partial r}$, then (on the hypersurface $v = 0$) the equation (1) becomes:

$$0 = \left(\sum a_i \frac{\partial}{\partial q^i} + b \frac{\partial}{\partial r}\right)v = \langle dv, \hat{X} \rangle .$$

So $\hat{X}$ at each point is in the tangent plane to the hypersurface at that point, and the trajectories of $\hat{X}$ remain in the hypersurface. Moreover:

<u>Proposition.</u> For $x_o \in \mathbb{R} \times C$, let $v: \mathbb{R} \times C \to \mathbb{R}$ with

$$\begin{cases} \langle dv, \hat{X} \rangle = 0 \\ \left.\dfrac{\partial v}{\partial r}\right|_{x_o} \neq 0, \; v(x_o) = 0. \end{cases}$$

Then the function $u: N_{x_o} \to \mathbb{R}$ such that $v(q^1, \ldots, q^n, u) = 0$, constructed for a suitable neighborhood $N_{x_o} \subset C$ via the implicit function theorem, is a solution of the P.D.E. (1).

<u>Proof.</u> By the construction of u,

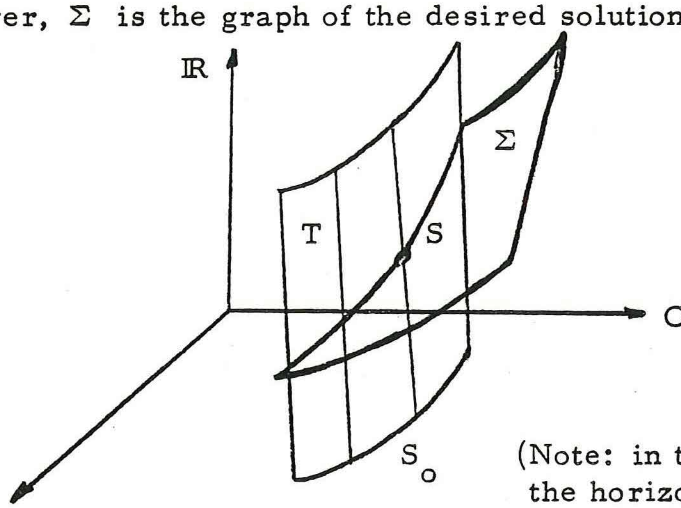
$$0 = \frac{\partial}{\partial q_i}(v \circ \hat{u}) \qquad\qquad C \xrightarrow{\hat{u}} \mathbb{R} \times C \xrightarrow{v} \mathbb{R},$$

$$= \frac{\partial v}{\partial q^i} + \frac{\partial v}{\partial r}\frac{\partial u}{\partial q^i} \quad \text{for each } i, \text{ so that we have}$$

$$0 = \langle dv, \hat{X} \rangle = \sum_{i=1}^{n} a_i \frac{\partial v}{\partial q_i} + b \frac{\partial v}{\partial r} = \sum_{i=1}^{n} -a_i \frac{\partial v}{\partial r}\frac{\partial u}{\partial q_i} + b \frac{\partial v}{\partial r} .$$

Therefore for $\frac{\partial v}{\partial r} \neq 0$, u will be a solution of (1), q.e.d.

Let S_o be a submanifold of C of dimension $n-1$, and let $u_o : S_o \to \mathbb{R}$ be a smooth function. Through S_o in $\mathbb{R} \times C$ pass the vertical hypersurface $T = \{(r, x) \mid r \in \mathbb{R}, \; x \in S_o\}$. Define $v_o : T \to \mathbb{R}$ by $v_o(r, x) = u_o(x) - r$. Suppose that the characteristic field $\hat{X}$ is transverse to T at a point x_o of the $n-1$ dimensional submanifold S on which $v_o = 0$. Then, it is immediate in local coordinates (see the figure below) that the trajectories of $\hat{X}$ through some neighborhood in S of x_o determine an n-dimensional submanifold Σ of $\mathbb{R} \times C$ (locally unique). Moreover, Σ is the graph of the desired solution u. For the



(Note: in the figure, C is the horizontal plane and the $\mathbb{R}$ axis is the vertical.)

fact that $\hat{X}$ is transversal to T at x_o implies that the function $v(x,t) = v_o(x)$ (where $x \in T$ and $\hat{X} = \frac{\partial}{\partial t}$) is well-defined on a neighborhood of x_o in $\mathbb{R} \times C$ and satisfies the conditions of the previous proposition $\left(\frac{\partial v}{\partial r} \Big|_{x_o} = -1 \neq 0 \right)$.

Any point of $\mathbb{R} \times C$ at which $\hat{X}$ is non-vertical lies on the graph of such solutions. Explicitly, the hypersurface $S_o \subset C$ may be described as the locus where some smooth function $f: C \to \mathbb{R}$ is constant (i.e., as a level hypersurface of f). Then the vertical hypersurface T is

$$T = \{(r,y) \mid f(y) = f(y_o)\}$$

for y_o a fixed and y any point of C.

We have proved

Theorem. For smooth functions $a_1, \dots, a_n, b: \mathbb{R} \times C \to \mathbb{R}$ let $S_o \subset C$ be defined by a point $y_o \in C$ and a smooth function $f: C \to \mathbb{R}$ as

$$S_o = \{y \mid y \in C \text{ and } f(y) = f(y_o)\}.$$

If $u_o: S_o \to \mathbb{R}$ is a smooth function satisfying the "transversality" condition

$$\sum a_i(u_o(y_o), y_o) \frac{\partial f}{\partial q^i} \neq 0,$$

then in some neighborhood of y_o there is a unique solution u of the P.D.E. $\sum a_i \frac{\partial u}{\partial q^i} = b$ with values u_o on S_o.

46. The General First Order P.D.E.

Consider an equation

$$(1'') \qquad E(u, q^1, \dots, q^n, \frac{\partial u}{\partial q^1}, \dots, \frac{\partial u}{\partial q^n}) = 0$$

in an unknown function $u: C \to \mathbb{R}$, where $q^1, \dots, q^n$ are local coordinates in the configuration space C. We can regard the "equation" E as a given (smooth) function $E: \mathbb{R} \times T^* C \to \mathbb{R}$. The differential du is a function $C \to T^* C$; we also have $d'u: C \to \mathbb{R} \times T^* C$ given locally as

$$(q^1, \dots, q^n) \to (u, q^1, \dots, q^n, \frac{\partial u}{\partial q^1}, \dots, \frac{\partial u}{\partial q^n}).$$

Thus the equation (1") becomes $E \circ d'u = 0$. If $r: \mathbb{R} \times T^* C \to \mathbb{R}$ is the projection on the first factor, then $\frac{\partial}{\partial q^i}$ applied to (1") yields the i^{th} derived P.D.E.

$$\frac{\partial E}{\partial r} \frac{\partial u}{\partial q^i} + \frac{\partial E}{\partial q^i} + \sum_{j=1}^{n} \frac{\partial E}{\partial p_j} \frac{\partial}{\partial q^i} (\frac{\partial u}{\partial q^j}) = 0, \quad i = 1, \dots, n.$$

Interchanging the order of partial derivatives, this is

$$\text{(2)} \qquad \frac{\partial E}{\partial r}\frac{\partial u}{\partial q^i} + \frac{\partial E}{\partial q^i} + \sum_{j=1}^{n} \frac{\partial E}{\partial p_j}\frac{\partial}{\partial q^j}\left(\frac{\partial u}{\partial q^i}\right) = 0 .$$

By way of motivation, observe that the i^{th} equation of (2) may now be regarded as a quasilinear P.D.E. in the unknown $p_i = \frac{\partial u}{\partial q^i}$. The characteristics of this quasilinear equation are then given by the (n+1)-dimensional vector field (see above):

$$\hat{X} = \sum_j \frac{\partial E}{\partial p_j}\frac{\partial}{\partial q^j} + \left(-\frac{\partial E}{\partial q^i} - \frac{\partial E}{\partial r}p_i\right)\frac{\partial}{\partial r} .$$

The differential equations of these characteristics are then the n+1 equations

$$\frac{dq^j}{dt} = \frac{\partial E}{\partial p_j} , \qquad j = 1, \ldots, n ,$$

$$\frac{dp^i}{dt} = -\frac{\partial E}{\partial q^i} - \frac{\partial E}{\partial r}p_i .$$

As i varies, the first n equations are the same. Note also that this reduces to Hamilton's equations when $\frac{\partial E}{\partial r} = 0$.

Our actual interpretation of (2) will be slightly different, as an equation on $\mathbb{R} \times T^{*}C$ itself, with characteristics in $\mathbb{R} \times T^{*}C$ which are solution curves of:

$$\text{(3)} \qquad \begin{gathered} \frac{dq^j c}{dt} = \frac{\partial E}{\partial p_j} , \quad \frac{dp_i c}{dt} = -\frac{\partial E}{\partial q_i} - \frac{\partial E}{\partial r}p_i , \\ \frac{dr}{dt} = \sum_j p_j \frac{\partial E}{\partial p_j} . \end{gathered}$$

The third set of equations is included since $E = E(r, q, p)$ is constant along trajectories of the vector field

$$(3') \quad X_E = \sum \frac{\partial E}{\partial p_j} \frac{\partial}{\partial p^j} + \sum_j \left(- \frac{\partial E}{\partial q^j} - p_j \frac{\partial E}{\partial r} \right) \frac{\partial}{\partial p_j} + \sum p_j \frac{\partial E}{\partial p_j} \frac{\partial}{\partial r}$$

on $T^{\cdot}C \times \mathbb{R}$. For

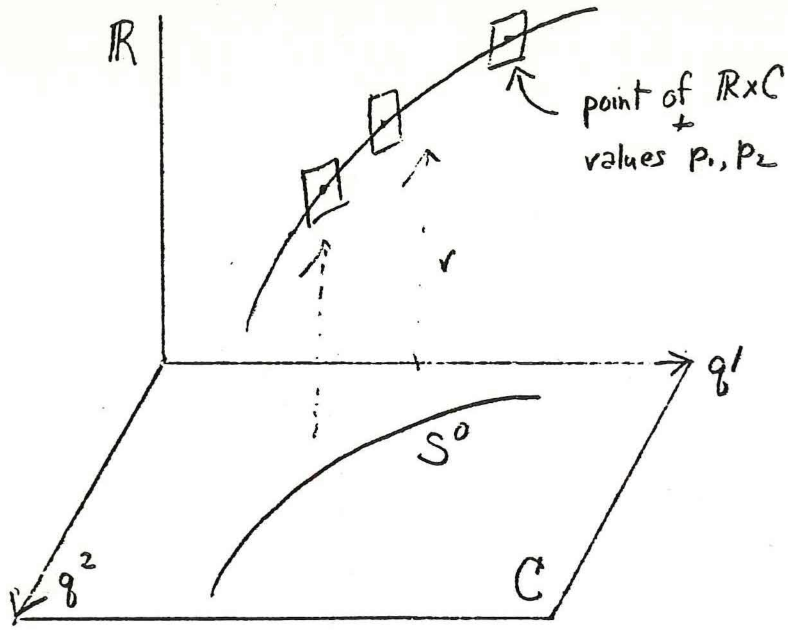
$$L_{X_E} E = \sum_j \frac{\partial E}{\partial p_j} \frac{\partial E}{\partial p^j} + \left(\sum_j - \frac{\partial E}{\partial q^j} \frac{\partial E}{\partial p_j} - \sum p_j \frac{\partial E}{\partial r} \frac{\partial E}{\partial p_j} \right) + \sum p_j \frac{\partial E}{\partial p_j} \frac{\partial E}{\partial r} = 0 .$$

For our previous cases, the last equation disappears, while the middle equations all collapse to the equation for r (relabeled $p_i c$) from the quasilinear case, giving trajectories "parallel" to those of the earlier cases. We state our existence theorem in the form:

Theorem: Given in C a compact submanifold S_o of dimension $n-1$ and initial values u_o of u on S_o such that a certain determinant (which appears as (5) below) does not vanish, then there exists an open set $U \supset S_o$ and a smooth function $u: U \to \mathbb{R}$ which satisfies E on U and agrees on S_o with u_o.

It will be clear from the proof that the conditions on the initial surface could be taken as before, and that the determinant condition corresponds to our previous transversality condition, with no loss of applicability.

Proof. We operate in $\mathbb{R} \times T^{\cdot}C$, where we already have defined in (3) the characteristic curves. In the submanifold T of dimension $2n$ above S_o, we distinguish a surface S which will correspond to $\hat{U} \cap T$. This submanifold (diffeomorphic to S_o) with local coordinates $x^1, \ldots, x^{n-1}$ embedded by $v: S_o \xrightarrow{v} S \subsetneq T$, should have



Proof. In the configuration space C we have local coordinates $q^1, \ldots, q^n$, an initial manifold $S_o \subset C$ of dimension 1 and initial values $u_o : S_o \to \mathbb{R}$. On $\mathbb{R} \times T^*C$ we have $2n+1$ local coordinates $r, q^1, \ldots, q^n, p_1, \ldots, p_n$. We can define a map $v : S_o \to \mathbb{R} \times T^*C$; this amounts to choosing "initial" values of r, q^i and p_j along S_o. Specifically we make $r \circ v = u_o, q^i \circ v = q^i$ and then we choose $p_1, \ldots, p_n$ so that

$$E = 0 \quad , \quad dr - \sum p_i dq^i = 0$$

both along S_o. The last condition on dr may be written in terms of local parameters $x^1, \ldots, x^{n-1}$ on the $(n-1)$-manifold S_o as

$$(4) \qquad 0 = du_o - \sum_{i=1}^{n} p_i dq^i = \sum_k \left(\frac{\partial u}{\partial x^k} - \sum_i p_i \frac{\partial q^i}{\partial x^k} \right) dx^k .$$

Hence $p_1, \ldots, p_n$ are determined uniquely along S_o if

$$(5) \qquad \begin{vmatrix} \frac{\partial E}{\partial p_1} & \frac{\partial E}{\partial p_2} & \cdots & \frac{\partial E}{\partial p_n} \\ \frac{\partial q^1}{\partial x^1} & & \cdots & \frac{\partial q^n}{\partial x^1} \\ \vdots & & & \\ \frac{\partial q^1}{\partial x^{n-1}} & & \cdots & \frac{\partial q^n}{\partial x^{n-1}} \end{vmatrix} \neq 0 \,.$$

This condition may be readily satisfied, since we can assume that the first row is nowhere zero. (This amounts to assuming that the given partial differential equation effectively involves at least one of the partial derivatives $p_i = \frac{\partial u}{\partial q^i}$). Given such a first row, the submanifold S_o can be chosen to make (5) hold; for example, if $\frac{\partial E}{\partial p_n} \neq 0$ we can choose the submanifold S_o given locally by the equation $q^n = 0$, with local coordinates $x^1 = q^1, \ldots, x^{n-1} = q^{n-1}$; then the determinant (5) is simply $(-1)^n \frac{\partial E}{\partial p_n}$. Indeed, the condition (5) is then exactly the condition that S_o be transversal to the projection of X_E.

We now have $v: S_o \to \mathbb{R} \times T^*C$, with image an $(n-1)$-manifold S in $R \times T^*C$; moreover one can show S transversal to the characteristic vector field X_E. Therefore the trajectories of X_E through points of S fill up locally a manifold T of dimension n. Now $E = 0$ holds along S, so by the properties earlier established for characteristics it holds along T. In other words, T gives the graph of functions $u, q^1, \ldots, q^n, p_1, \ldots, p_n$ on C (or on a neighborhood of S_o in C) which satisfy $E(u, q^1, \ldots, q^n, p_1, \ldots, p_n) = 0$.

What remains to be verified is that $p_i = \frac{\partial u}{\partial q_i}$ for $i = 1, \ldots, n$ on this manifold T. If $\nu : T \to \mathbb{R} \times T^*C$ is the inclusion map, this amounts to showing that the induced 1-form $\theta = \nu^*(dr - \sum p_i dq^i)$ is zero on T. We may calculate θ in local coordinates $x^1, \ldots, x^{n-1}$ (on S_o) and t (the parameter along the trajectories of X_E) as

$$\theta = \sum_{k=1}^{n-1} \left(\frac{\partial r}{\partial x^k} - \sum p_i \frac{\partial q^i}{\partial x^k} \right) dx^k + \left(\frac{\partial r}{\partial t} - \sum p_i \frac{\partial q^i}{dt} \right) dt .$$

The last term is zero by the equation (3) for the characteristics. It thus remains to show that

$$D_k = \frac{\partial r}{\partial x^k} - \sum p_i \frac{\partial q^i}{\partial x^k} , \quad k = 1, \ldots, n-1$$

is zero. But $D_k = 0$ on S_o by the choice of the initial values of p_i, while a calculation with the equations (3) shows

$$\frac{\partial D_k}{\partial t} = \frac{\partial E}{\partial x^k} + \frac{\partial E}{\partial r} D_k = \frac{\partial E}{\partial r} D_k .$$

This is a linear first order differential equation for D_k as a function of the parameter t, with initial values zero on S_o. Hence (by the uniqueness of the solutions of such equations) $D_k = 0$, q.e.d..

47. <u>Contact Manifolds.</u> The use of the characteristic vector field X_E for the partial differential equation E raises the following question. For a symplectic manifold any two smooth functions f and g have a Poisson bracket given by

$$\{f, g\} = L_{X_f} g = -L_{X_g} f .$$

On the manifold $N = \mathbb{R} \times T^*C$, each smooth function $F: N \to \mathbb{R}$ determines (by characteristics, as in (3) above) a vector field X_F, and hence two such functions F and G have a "superbracket" defined as $X_F(G)$. We wish to examine the geometric structure producing this operation; it will turn out that this structure depends essentially on the 1-form $dr - \sum p_i dq^i$ used in the calculations to the last theorem.

Another approach is in terms of "elements". An "element" of the space $\mathbb{R} \times C$ is a point of this space plus a (non-vertical) hyperplane through this point; for example, if C has dimension 2 an element is just this: $\boxed{\cdot}$. In coordinates $r, q^1, \ldots, q^n$ any hyperplane through the origin has an equation $a_o r + a_1 q^1, \ldots, + a_n q^n = 0$ for suitable constants a_i; it is non-vertical precisely when $a_o \neq 0$, and in this case we may take $a_o = 1$ and write $r + a_1 q^1 + \ldots a_n q^n = 0$. Thus the hyperplane is determined by $a_1, \ldots, a_n$, which we now write as $p_1, \ldots, p_n$ (in case $n = 3$ they are the direction cosines of the normal to the hyperplane). Thus an element is given by coordinates $r, q^1, \ldots, q^n, p_1, \ldots, p_n$, and so is exactly a point in $\mathbb{R} \times T^*C$.

Take a curve $c: \mathbb{R} \to \mathbb{R} \times T^*C$; it consists of points of $\mathbb{R} \times T^*C$ and so may instead be regarded as a curve in $\mathbb{R} \times C$ consisting of elements there. In the classical treatments, such a curve of elements is called a characteristic strip when the 1-form $\theta = dr - \sum p_i \, dq_i$ is zero along the strip.

These elements are also used in geometrical optics in the space $\mathbb{R} \times C$. Huyghens principle for the propagation of a wave front W gives the new wave front W_t after time t as the envelope of the spherical waves centered at all the points of W. The transformation from W to W_t then does not carry points of the space (say the space $\mathbb{R} \times C$) into points, but elements into elements, so is really a transformation of $\mathbb{R} \times T^*C$ into itself. Since such a transformation of elements carries tangent wave fronts V and W into tangent wave fronts, it is called a "contact transformation".

This pictorial representation of a contact transformation is again connected with first order partial differential equations. One finds (for example, see Luneburg, Mathematical theory of optics) that Maxwell's equations yield wave fronts of the form

$$\psi(x, y, z) - ct = 0 \quad \text{where } \psi \text{ satisfies } \left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 \left(\frac{\partial\psi}{\partial z}\right)^2 - n^2 = 0$$

in the medium whose "index of refraction" is n.

A contact manifold is a manifold N of dimension $2n+1$ with a distinguished one-form θ such that $\theta \wedge (d\theta)^n \neq 0$ everywhere. We also consider submanifolds $\nu : S \to N$ such that $\nu^*\theta = 0$; one dimensional such are called strips. A transformation $(N, \theta) \xrightarrow{h} (N', \theta)$ is called a contact transformation when there is a map $\rho : N \to \mathbb{R}$ for which $h^*\theta' = \rho\theta$.

<u>Theorem.</u> A mapping h is a contact transformation if and only if it takes strips into strips.

<u>Proof.</u> $\Longrightarrow$ Trivial since $R \xrightarrow{c} (N,\theta) \xrightarrow{h} (N',\theta)$.

$\Longleftarrow$ Left to the reader.

Clearly, this notion of a contact transformation is preserved under composition; however, we also needed the characteristics before. Given smooth mappings E and F of N into R, there were X_E defined by (3') and $X_E(F) = [E,F]$, analogous to the Poisson bracket though not quite satisfying the Jacobi identity.

Suppose we are given a mapping $\Phi: T^*(N) \longrightarrow T_*(N)$, with

$$\Phi(dq^i) = -\frac{\partial}{\partial p_i}$$

$$\Phi(dp^i) = \frac{\partial}{\partial q^i} + p_i \frac{\partial}{\partial r}$$

$$\Phi(dr) = -\sum p_i \frac{\partial}{\partial p_i}$$

$$\begin{tikzcd} T_*N \arrow[dr] & & T^*N \arrow[ll, "\Phi"'] \arrow[dl] \\ & N & \end{tikzcd}$$

In particular (by construction) $\Phi(dE) = X_E$. The matrix of Φ will be of the form:

$$\Phi^{\alpha\beta} = \begin{array}{c} \\ q \\ p \\ r \end{array} \begin{array}{c} \begin{array}{ccc} q & p & r \end{array} \\ \begin{pmatrix} 0 & -I & p \\ I & 0 & p \\ 0 & -p & 0 \end{pmatrix} \end{array}$$

where α and β are indices ranging from 1 to 2n+1. It will be of rank 2n. Since $\Phi(dr - \sum p_i dq^i) = -\sum p_i \frac{\partial}{\partial p_i} + \sum p_i \frac{\partial}{\partial p_i} = 0$, the form θ will be determined up to a scalar factor as the kernel of Φ. At any point

Φ is a twice contravariant tensor field on N. We may define such a pair (N, Φ) to be a bracket manifold, where Φ is determined up to multiplication by a smooth $\rho \neq 0$. Given a configuration space C, we may construct $N = T^{*}C \times \mathbb{R}$ and define $\Phi^{\alpha\beta}$ as above, showing that this matrix as defined in local coordinates transforms properly under coordinate changes. In terms of Φ, we have $[E, F] = \langle dF, \Phi(dE) \rangle$. Thus the tensor Φ on N is indeed sufficient to define the bracket operation. For a mapping h to be a bracket transformation of (N, Φ) into (N', Φ'), we require that $h^{*}[E, F] = \rho[h^{*}E, h^{*}F]$ for some $\rho : N \to N'$ with $\rho \neq 0$ on N. We could instead require that h^{*} multiply by a common $\rho \neq 0$ the relations between canonical coordinates:

$$[q^i, q^j] = 0 = [p_i, p_j]$$
$$[q^i, p_j] = \delta^i_j$$
$$[q^i, r] = 0 \qquad [p_i, r] = p_i$$

We suggest that all contact manifolds are bracket manifolds. Note that $d\theta(X_E, X_F) = [E, F]$.

We now develop the suggestion of the previous paragraph. References are Cartan's Lecons sur les integrals invariants (1922) or the article by John Gray in the "Annals of Mathematics" 69(1959), pp. 421-450. Let (N, θ) be a given contact manifold. We will define in terms of the basis form θ a vector field Y_θ and a bracket $[\ \ ]_\theta$: Since the matrix of the 2-form $d\theta$ is of rank 2n, we may define a vector field Y by:

$$L_Y(d\theta) = 0 \quad \text{and} \quad \langle \theta, Y \rangle = 1.$$

For it suffices to do this locally, where we may take θ in the form $\theta = dr - \sum p_i dq^i$ by the Darboux theorem, in which case we have

$$d\theta = \sum dq^i \wedge dp_i \ , \quad Y = \sum Y_i \frac{\partial}{\partial q^i} + \sum \overline{Y}_i \frac{\partial}{\partial p_i} + Z\frac{\partial}{\partial r}$$

and $L_Y d\theta = 0$ if and only if $Y = Z\frac{\partial}{\partial r}$ with $<\theta, Z\frac{\partial}{\partial r}> = Z$.

Equivalent would be an appeal to the fact that the $(2n+1)$ forms on N are spanned at each point by $\theta \wedge (d\theta)^n$, so that any such form, and in particular the form $dE \wedge (d\theta)^n$ may be represented uniquely as $h(r, p, q)\theta\wedge(d\theta)^n$ for some smooth function h. The quantity $h \overset{\text{def}}{\equiv} Y_\theta(E)$ is easily shown to be a derivation, and therefore determines a vector field Y_θ. That this is the vector field of the previous paragraph is verified by evaluation of $<Y_\theta, \theta>$ and $i_{Y_\theta} d\theta$.

We claim that θ also defines $[E,F]$ by:

$$dE \wedge dF \wedge \theta \wedge (d\theta)^{n-1} = [E,F]\theta \wedge d\theta^n .$$

The verification that the function so defined is $[E,F]$ is by direct calculation in canonical coordinates:

$$\theta \wedge (d\theta)^{n-1} = (dr - \sum_i p_i dq^i) \wedge (\sum_i dp_i \wedge dq^i)^{n-1}$$

$$= (dr - \sum_i p_i dq^i) \wedge (n-1)! \sum_i (-1)^{(n-1)(n-2)/2} dp_{\hat{j}} \wedge dq_{\hat{j}}$$

$$= (-1)^{(n-1)(n-2)/2}(n-1)! [\sum dp_{\hat{j}} \wedge dq_{\hat{j}} \wedge dr - \sum (-1)^{n+j-2} p_j dp_{\hat{j}} \wedge dq,$$

where $dp_{\hat{j}}$ is short for the product of all the dp_i with only dp_j omitted from the product, while $dq = dq^1 \wedge \ldots \wedge dq^n$ with no terms omitted.

We have $dE \wedge dF$ computed as:

$$\left(\sum_i \frac{\partial E}{\partial q^i} dq^i + \sum_i \frac{\partial E}{\partial p_i} dp_i + \frac{\partial E}{\partial r} dr\right) \wedge \left(\sum_i \left(\frac{\partial F}{\partial q^i} dq^i + \frac{\partial F}{\partial p_j} dp_j\right) + \frac{\partial F}{\partial r} dr\right)$$

and by direct computation the nonzero terms of $dE \wedge dF \wedge (\theta \wedge (d\theta)^{n-1})$, the only one without products $\theta \wedge \theta$, are:

$$(-1)^{n+\{(n-1)(n-2)/2\}}(n-1)!\sum_i \left(\frac{\partial E}{\partial q^i}\frac{\partial F}{\partial p_i} - \frac{\partial E}{\partial p_i}\frac{\partial F}{\partial q^i} - p_i \frac{\partial E}{\partial p_i}\frac{\partial F}{\partial r} + p_i \frac{\partial E}{\partial r}\frac{\partial F}{\partial p_i}\right) dp \wedge dq \wedge dr$$

$= (-1)^{n(n-1)/2}(n-1)!\,[E,F]\,dp \wedge dq \wedge dr$ according to our former definition. Since $\theta \wedge (d\theta)^n$ is $(-1)^{n(n-1)/2}\, n!\, dp \wedge dq \wedge dr$, our two definitions of $[E,F]$ have been shown to agree (except for $n!$ versus $(n-1)!$).

CHAPTER VII. COVARIANT DIFFERENTIATION

By David Golber

The following material summarizes, in outline form, lectures on covariant differentiation.

I (§48). Riemannian and Pseudo-Riemannian Metrics

A. Definition. A Riemannian metric g on a manifold M assigns in a C^∞ fashion to each point x of M an inner product g_x on the tangent vector space T_xM. A pseudo-Riemannian metric g assigns, in a C^∞ fashion, to each point $x \in M$ a non degenerate symmetric bilinear form g_x on T_xM.

Certain important results, especially III B. 3 below hold for pseudo-Riemannian as well as Riemannian metrics.

B. Local expression: In a coordinate patch on M, we have coordinates $x_1, \ldots, x_n$, and vector fields $\partial/\partial x_1, \ldots, \partial/\partial x_n$. Let us use the abbreviation $\partial/\partial x_i = \partial_i$ for these fields. Suppose M has a Riemannian or pseudo-Riemannian metric g. Then we can define C^∞ functions on the coordinate patch by

$$g_{ij}(x) = g_x(\partial_i(x), \partial_j(x)).$$

If g is Riemannian, then, for each x, the matrix $(g_{ij}(x))$ is positive definite and symmetric. If g is pseudo-Riemannian, then $(g_{ij}(x))$ is non-singular and symmetric.