If ω is a k-form, redefine the (k+1)-form $d\omega$ by

$$(d\omega)(u, v_o, \ldots, v_k) = \sum_{\ell=0}^{k} (-1)^{\ell} (D\omega)(u, v_\ell, v_o, v_1, \ldots, \hat{v}_\ell, \ldots, v_k) .$$

(Here the $\wedge$ over v_ℓ means that v_ℓ is omitted.) We claim this $d\omega$ is the same as the $d\omega$ defined previously. This is checked by showing that this $d\omega$ is linear and alternating in the $v_o, \ldots, v_k$, and has the same values on the basis elements of $V \times V \times \ldots \times V$ as the old $d\omega$. The linearity is clear, given our comments regarding the operator D; $d\omega$ is alternating since computation shows that it vanishes when any two successive arguments are equal. Suppose now ω is a one-form ; $\omega = \sum_i w_i dq^i$, where $\{q^i\}$ are coordinates on M and $\{e_i\}$ are the corresponding basis elements of $V \cong T_u(M)$. Then $w_i(u) = \omega(u, e_i)$. By our old definition

$$d\omega = \sum_{i<j} (\frac{\partial w_j}{\partial q^i} - \frac{\partial w_i}{\partial q^j}) dq^i \wedge dq^j = \sum_{i<j} dw(u, e_i, e_j) dq^i \wedge dq^j .$$

To prove that the two definitions coincide for one-forms it will thus suffice to show that $d\omega(u, e_i, e_j)$ is the same as in the new definition. But in the new definition

$$d\omega(u, e_i, e_j) = D\omega(u, e_i, e_j) - D\omega(u, e_j, e_i),$$

and

$$Df(u, e_i) = \partial f / \partial q_i .$$

Hence

$$d\omega(u, e_i, e_j) = \frac{\partial \omega(u, e_j)}{\partial q^i} - \frac{\partial \omega(u, e_j)}{\partial q^j} = \frac{\partial \omega_j}{\partial q^i} - \frac{\partial \omega_i}{\partial q^j} ,$$

which is what we were trying to prove. Similar techniques show that the two definitions are the same for general k-forms.

We are now ready to define the map s which makes a $(p-1)$-form out of every p-form. If ω is a k-form, let

$$(s\omega)(u; v_1, \ldots, v_{k-1}) = \int_0^1 t^{k-1} \omega(tu; u, v_1, \ldots, v_{k-1}) \, dt.$$

Here we consider the open set U as part of the vector space $V = \mathbf{R}^n$, which has also been identified with $T_p(U)$. Thus on the right-hand side of the equation, the second argument, $u \in U$, is viewed as a vector of V. But since U is an open ball, tu, the first argument, is in U for all $t \neq 1$. It is now easy to check that $s\omega$ is a $(k-1)$-form -- linear, alternating, and smooth as a function of u.

We now take a k-form ω and show, at last, that $ds(\omega) + sd(\omega) = \omega$. First,

$$D(s\omega)(u, v, v_1, \ldots, v_{k-1}) = \int_0^1 D\,[t^{k-1} \omega(tu, v, u, v_1, \ldots, v_{k-1})] dt$$

(since all functions involved
are smooth and bounded)
$$= \int_0^1 t^k D\omega(tu, v, u, v_1, \ldots, v_{k-1}) \, dt$$

$$+ \int_0^1 t^{k-1} \omega(tu, v, v_1, \ldots, v_{k-1}) \, dt.$$

The latter term appears as it does since ω is linear in the third variable, and it was proved that if f is linear, $Df(u, v) = f(v)$. Now

$$d(s\omega)(u, v_1, \ldots, v_k) = \sum_{\ell=1}^{k} (-1)^{\ell-1} D(s\omega)(u, v_\ell, v_1, \ldots, \hat{v}_\ell, \ldots, v_k)$$

$$= \sum_{\ell=1}^{k} (-1)^{\ell-1} \left[\int_0^1 t^k D\omega(tu, v_\ell, u, v_1, \ldots, \hat{v}_\ell, \ldots, v_k)\, dt \right.$$

$$\left. + \int_0^1 t^{k-1} \omega(tu, v_\ell, v_1, \ldots \hat{v}_\ell, \ldots, v_k)\, dt \right] ,$$

and

$$s(d\omega)(u, v_1, \ldots, v_k) = \int_0^1 t^k d\omega(tu, u, v_1, \ldots, v_k)\, dt$$

$$= \int_0^1 \sum_{\ell=1}^{k} (-1)^\ell t^k D\omega(tu, v_\ell, u, v_1, \ldots, \hat{v}_\ell, \ldots, v_k)\, dt$$

$$+ \int_0^1 t^k D\omega(tu, u, v_1, \ldots, v_k)\, dt .$$

When we add $d(s\omega)$ and $s(d\omega)$, the first terms of each expression cancel; also,

$$\sum_{\ell=1}^{k} (-1)^{\ell-1} \int_0^1 t^{k-1} \omega(tu, v_\ell, v_1, \ldots, \hat{v}_\ell, \ldots, v_k)\, dt$$

$$= \sum_{k=1}^{\ell} (-1)^{\ell-1} \int_0^1 (-1)^{\ell-1} t^{k-1} \omega(tu, v_1, \ldots, v_k)\, dt \qquad \text{since } \omega \text{ is alternating}$$

$$= k \int_0^1 t^{k-1} \omega(tu, v_1, \ldots, v_k)\, dt .$$

Hence

$$(s\, d\omega + d s\, \omega)(u, v_1, \ldots, v_k) = \int_0^1 \left[t^k D\omega(tu, u, v_1, \ldots, v_k) + k t^{k-1} \omega(tu, v_1, \ldots, v_k) \right] dt$$

$$= \int_0^1 \frac{d}{dt} \left[t^k \omega(tu, v_1, \ldots, v_k) \right] dt$$

(since $D\omega(tu, u, \ldots)$ is just the directional derivative in direction u

of $\omega(tu, v_1, \ldots)$);

$$= 1^k \omega(u, v_1, \ldots, v_k) - 0 = \omega(u, v_1, \ldots, v_k). \quad \text{Q. E. D.}$$

§24. The Lie Derivative

Let X be a vector field on U. There is an operation on the

ring $\mathcal{F}$ of smooth functions on U defined by

$$L_X(f) = Df(u, X(u)) = (X_u)(f).$$

This operaton L_X is a derivation, since each $X_u \in T_u(U)$ is a

derivation. L_X is called the Lie derivative. We have shown that a

vector field is determined by the way it acts on the functions of $\mathcal{F}$;

this means that knowing the operator L_X determines X.

Now it is easy to check that if θ and ψ are derivations of $\mathcal{F}$,

then so is $\theta\psi - \psi\theta$. Call this new derivation $[\theta, \psi]$. Then $[L_X, L_Y]$ is

a derivation, so to it there is associated a unique vector field. This

vector field is called the Lie bracket of X and Y and is written $[X, Y]$.

In coordinates $\{q^1, \ldots, q^n\}$, let $X = \sum x^i \frac{\partial}{\partial q^i}$, $Y = \sum y^i \frac{\partial}{\partial q^i}$

$x^i, y^i \in \mathcal{F}$. Then $L_X(f) = \sum x^i \frac{\partial f}{\partial q^i}$ for $f \in \mathcal{F}$, so

$$L_{[X, Y]} = \sum_i (\sum_j x^j \frac{\partial y^i}{\partial q^j} - y^j \frac{\partial x^i}{\partial q^j}) \frac{\partial}{\partial q^i}.$$ In fact, it is easy to check

without using coordinates that $[\,,\,]$ is linear in each argument and satis-

fies the relations

$[X, X] = 0$, $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$ (Jacobi identity).

We will now show that the Lie bracket provides a natural way of extending the Lie-derivative operation to apply to all tensor fields. There is an <u>identity tensor</u> $\delta \in \mathcal{J}^1_1(V)$; δ corresponds to the identity map on V^* under the series of identifications

$$\mathcal{J}^1_1(V) = V \otimes V^* \cong \mathrm{Hom}(V^* \otimes V, \mathbb{R}) \cong \mathrm{Hom}(V^*, \mathrm{Hom}(V, \mathbb{R})) \cong \mathrm{Hom}(V^*, V^*).$$

In coordinates $\{e_j\}$ we find $\delta = \sum e_i \otimes e^i$.

<u>Theorem</u>. Given any vector field X on U, there is a unique linear map L_X, where $L_X: \mathcal{J}^r_s(U) \longrightarrow \mathcal{J}^r_s(U)$ for all non-negative integers r and s, with the properties

1. $L_X f = <df, X>$ for any smooth function f on U,

2. $L_X Y = [X, Y]$ for any vector field Y,

3. $L_X \delta = 0$,

4. L_X is a derivation; that is,

$$L_X(\tau \otimes \tau') = (L_X \tau) \otimes \tau' + \tau \otimes (L_X \tau'),$$

for any tensor fields τ and τ'.

<u>Proof</u>. We put coordinates $q^1, \ldots, q^n$ on U and show that any operator L_X satisfying the given conditions must satisfy

(5) $$L_X(dq^k) = \sum_i \frac{\partial x^k}{\partial q^i} dq^i \qquad \text{for all } k.$$

Write $X = \sum x^i \frac{\partial}{\partial q^i}$; then we have shown that

$$L_X Y = \sum_{i,j} (X^j \frac{\partial Y^i}{\partial q^j} - Y^j \frac{\partial X^i}{\partial q^j}) \frac{\partial}{\partial q^i} \, , \quad \text{if } Y = \sum Y^i \frac{\partial}{\partial q^i} \, .$$

In particular, $L_X(\frac{\partial}{\partial q^k}) = - \sum_i \frac{\partial X^i}{\partial q^k} \frac{\partial}{\partial q^i}$.

Now $\delta = \sum_k \frac{\partial}{\partial q^k} \otimes dq^k$. So

$$0 = L_X \delta = \sum_k L_X(\frac{\partial}{\partial q^k}) \otimes dq^k + \frac{\partial}{\partial q^k} \otimes L_X dq^k$$

$$= -\sum_k \sum_i \frac{\partial X^i}{\partial q^k} (\frac{\partial}{\partial q^i} \otimes dq^k) + \sum_k \frac{\partial}{\partial q^k} \otimes L_X dq^k .$$

If we write $L_X(dq^k) = \sum c_k^i dq^i$, it is clear that we are forced to take $c_k^i = \partial X^k / \partial q^i$. But it is clear from (4) that once L_X is defined on functions, vector fields, and co-vector fields, it extends uniquely to all tensor fields. So we merely define L_X by (1), (2), and (5), and check that (3) and (4) are satisfied. Notice that this is really an invariant proof, since we have shown that <u>any</u> extension of the Lie derivative satisfying 1-4 must, when expressed in coordinates, agree with the operator we've defined.

Corollary 1. $L_X(df) = d(L_X f)$.

For example,

$$d(L_X q^k) = d(\sum X^i \frac{\partial q^k}{\partial q^i}) = dX^k = \sum_i \frac{\partial X^k}{\partial q^i} dq^i = L_X(dq^k).$$

Corollary 2. If V is an open subset of U, and $|_V$ denotes the restriction to V, then

$$(L_X \tau)|_V = L_{(X|_V)}(\tau|_V).$$

Corollary 3. L_X maps $\Omega_k(U)$ into $\Omega_k(U)$.

Proof. Think of an exterior form as an alternating tensor; recall that a tensor τ is alternating if and only if $A\tau = \tau$. Hence, we must show that if τ is alternating, $A(L_X\tau) = L_X\tau$. In fact, $A(L_X\tau) = L_X(A\tau)$, since A is a sum of permutation operators, and it is easy to see that L_X commutes with permutations.

Corollary 4. $L_X(\omega \wedge \eta) = L_X\omega \wedge \eta + \omega \wedge L_X\eta$, if ω and η are exterior forms.

Proof.
$$L_X(\omega \wedge \eta) = L_X(A(\omega \otimes \eta)) = AL_X(\omega \otimes \eta)$$
$$= A(L_X\omega \otimes \eta + \omega \otimes L_X\eta)$$
$$= L_X\omega \wedge \eta + \omega \wedge L_X\eta .$$

Corollary 5. $d(L_X\omega) = L_X(d\omega)$ if ω is a k-form.

Proof. Write $\omega = \sum f dq^1 \wedge dq^2 \wedge \ldots \wedge dq^k$. Then
$$d\omega = \sum df \wedge dq^1 \wedge \ldots \wedge dq^k, \quad \text{while}$$
$$L_X\omega = \sum L_X f \, dq^1 \wedge \ldots \, dq^k + \sum_{i=1}^{k} f \, dq^1 \wedge \ldots \wedge L_X dq^i \wedge \ldots \wedge dq^k.$$
Computation now proves the equality, with the aid of the preceding corollaries.

§25 Transportation along Trajectories

This purely formal proof of the properties of the Lie derivative does not really shed much light on the geometrical meaning of this operator. Actually, as we are about to show, the Lie derivative can be interpreted in a way very much like the ordinary definition of a derivative: the limit of a difference quotient.

Recall that every smooth vector field X on a local manifold M has integral curves c passing through every point of M. An integral curve is one whose tangent at every point p is the same as the tangent vector which is the value of the vector field X at p; symbolically, if $p = c(t_0)$,

$$dc/dt(t_0) = X_{c(t_0)} \, , \quad \text{or} \quad dc/dt = X \circ c.$$

The standard existence and uniqueness theorem for ordinary differential equations, when applied to this equation expressed in coordinates, guarantees the existence of at least one integral curve through each point of M (although each curve may be defined only on a small interval on the real line); furthermore, two integral curves passing through the same point must agree wherever they are both defined.

Now change the point of view slightly, and consider the motion of M which takes each point p to the point t units along the integral curve passing through p. For a fixed t for which all the integral curves are defined, this would describe a map from M to M. In these terms, the existence and uniqueness theorem may be stated:

If X is a smooth vector field on a local manifold M, there is a unique trajectory of X through each point, and for each point p of M there is a neighborhood U of p and an interval $I \subset \mathbb{R}$, together with a function $F: I \times U \longrightarrow M$, such that $F(0, u) = u$ (initial conditions) and for fixed u, the map taking t to $F(t, u)$ is a trajectory of X.

Write F_t for the map of M to M defined by $F_t(u) = F(t, u)$. In the case where M and U happen to be all of $\mathbb{R}^n$ and $I = \mathbb{R}$, we must have $F_t F_{t'}(u) = F_{t+t'}(u)$ for all t and t'. Indeed. $t \rightsquigarrow F_t(F_{t'}(u))$ is a trajectory starting at $F_{t'}(u)$, but so is $t \rightsquigarrow F_{t+t'}(u)$; by uniqueness, they must be equal. In particular,

$$\left.\begin{aligned} F_t(F_{-t}(u)) &= u \\ F_{-t}(F_t(u)) &= u \end{aligned}\right\} \text{ for each } u \in M,$$

so each F_t has a smooth inverse map. In this case we say that F_t is a diffeomorphism. Then $t \rightsquigarrow F_t$ is a map of the additive group of real numbers into the group of diffeomorphisms of $\mathbb{R}^n$. In general, of course, the map F will be defined only on subsets I and U of $\mathbb{R}$ and $\mathbb{R}^n$; it is not hard to see, however, that restricting further to an interval $I' \in I$ and an open set $U' \in U$ gives us a map $F' = F|_{I' \times U'}$ for which F'_t has the smooth inverse F'_{-t} for each $t \in I'$. (Cf. Abraham pp. 39-40). Such a system F' will be called a _flow box_ of X.

Suppose now that B is a functor from vector spaces to vector spaces: for example, $B(V) = V^*$, or $B(V) = V \times V$, or $B(V) = V \wedge \ldots \wedge V$. Associated to such a B is a bundle $B(M)$ over M, whose fiber over p is $B(T_p M)$. Thus if $B(V) = V$, $B(M) = T.M$; if $B(V) = V \wedge V$, $B(M)$ is the bundle of all 2-forms on M; and so on. Furthermore, if $f: M \longrightarrow N$, is a smooth map, then there is a map

$$B(f): B(M) \longrightarrow B(N) \qquad \text{if } B \text{ is covariant};$$

$$B(f): B(N) \longrightarrow B(M) \qquad \text{if } B \text{ is contravariant}.$$

For instance, if $B(V) = V^*$, $B(f) = f^*: T^{\cdot}(N) \longrightarrow T^{\cdot}(M)$.

To define the Lie derivative of a general field T with respect to X we wish to do the following: given a point p, move along the trajectory of X through p for a time t; at this point, find the value of T, and now move T back along the trajectory to get a tensor at p. This pulled-back tensor will not generally be the same as the value of the tensor field at p; but we can form the difference quotient.
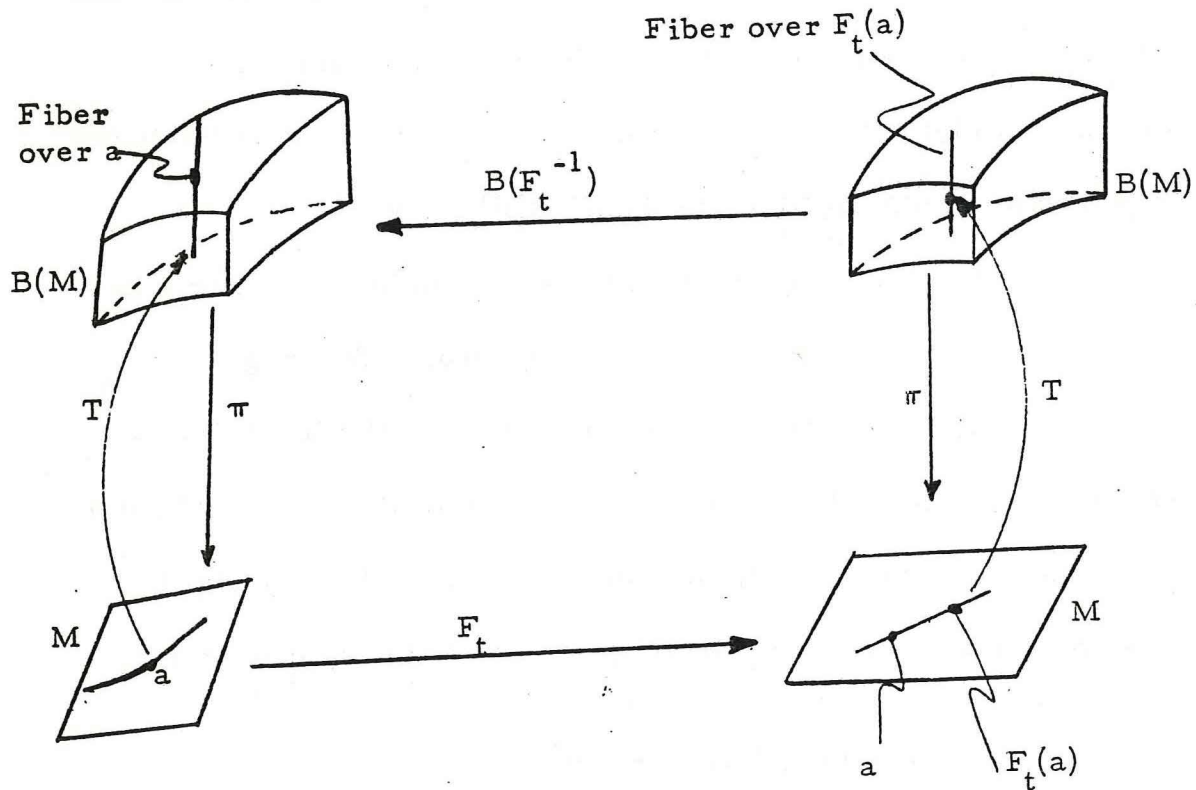
<u>Definition.</u> If X is a smooth vector field on M and T is a co- or contra-variant field on M, define the operator K_X by

$$K_X(T) = \begin{cases} d/dt(B(F_t^{-1}) \circ T \circ F_t)\big|_{t=0} & \text{if } T \text{ is covariant} \\ d/dt(B(F_t) \circ T \circ F_t)\big|_{t=0} & \text{if } T \text{ is contravariant,} \end{cases}$$

where F is a flow box for X.

Notice that this definition does not depend on the choice of F, by the uniqueness of flow boxes.

<u>Theorem.</u> K_X is the same as the Lie derivative L_X.

<u>Proof.</u> By the Theorem of §24 characterizing L_X, it will suffice to show that K_X is a derivation, $K_X \delta = 0$, and that K_X agrees with L_X on functions and vector fields. In fact, examination of the proof of that theorem makes it clear that we can show $K_X = L_X$ on covector fields instead of on vector fields, and the result will still follow.

First, $K_X \delta = 0$ since δ may be expressed as an identity matrix invariant under F_t and $B(F_t)$; its derivative is zero. Showing that K_X is a derivation will involve working with functors $B, B',$ and B'', and a covariant, bilinear, natural transformation $\square : B \times B' \longrightarrow B''$. Then if $T: M \longrightarrow B(M)$, $T': M \longrightarrow B'(M)$, we can define $T \times T'$, mapping M into the pullback bundle $B(M) \times_M B'(M)$; $T \square T'$ is the composite map

$$M \xrightarrow{\;T \times T'\;} B(M) \times_M B'(M) \xrightarrow{\;\square_M\;} B''(M)$$

To show that $K_X(T \square T') = K_X T \square T' + T \square K_X T'$, we compute

$$B''(F_t^{-1}) \circ T \square T' \circ F_t = [B(F_t^{-1})T \square B'(F_t^{-1})T'] \circ F_t$$

$$= [B(F_t^{-1})T]F_t \square [B'(F_t^{-1})T']F_t \; .$$

Now in general, suppose $\sigma(t), \sigma'(t)$ are maps of an interval I to V and V', respectively; $\sigma(t) \square \sigma'(t) \in V \square V' = V''$. Since $\square$ is bilinear, we can write

$$e_i''(\sigma(t) \square \sigma'(t)) = \sum_{j,k} c_{jk}^i \, e_j(\sigma(t)) e_k'(\sigma'(t))$$

for some constants c_{jk}^i where $\{e_i\}$ are coordinates on V, and so on. Now the ordinary Leibnitz rule for the derivative of a product of real-valued functions applies, and it follows that K_X is a derivation.

If f is a real-valued function on M, the fiber over every point is $\mathbb{R}$, and in this case $B(F_t^{-1})$ is always the identity map. Thus

$$K_X(f) = \frac{d}{dt}(f(F_t(a)) = \sum_i \frac{\partial f}{\partial q^i} \frac{\partial F_t^i}{dt} = \sum x^i \frac{\partial f}{\partial q^i} = L_X f \; ,$$

where $q^i F_t = F_t^i$, and $X = \sum_i x^i \frac{\partial}{\partial q^i}$

Finally, to show that K_X agrees with L_X on covector fields, we can use the properties already proved to find $K_X(\sum f_i dq^i)$, once we know $K_X(dg)$ for any function g; hence it will be enough to prove $K_X(dg) = L_X(dg)$. Now in this case $B(V) = V^*$, so $B(F_t) = F_t^*$. In general, however, we defined the pullback of a form ω by a map φ as the form $\varphi^* \omega$ given by

$$(\varphi^* \omega)(a) = \varphi^*(\omega_{\varphi(a)}) = (\varphi^* \circ \omega \circ \varphi)(a) \; .$$

Thus

$$K_X(dg) = \frac{d}{dt}(B(F_t) \circ dg \circ F_t) = \frac{d}{dt}(F_t^* \circ dg \circ F_t)$$

$$= \frac{d}{dt}(F_t^*(dg)) = \frac{d}{dt}(d(g \circ F_t))$$

$$= d(\frac{d}{dt}(g \circ F_t)) = d(L_X g) = L_X(dg)$$

by previous part of the theorem.

This completes the proof of the theorem.

Note. In the above proof we considered a k-form as a cross-section of a suitable (exterior) bundle. This means in particular that we should consider a 0-form (= a smooth function) as a cross-section; we will show that it is a cross-section of the trivial bundle. Explictly, take that functor B which sends each vector space V to the one-dimensional space $\mathbb{R}$ and each linear transformation $f: V \longrightarrow V'$ to $1: \mathbb{R} \longrightarrow \mathbb{R}$. (Thus B is a "constant" functor). If B is used to construct a fiber bundle over M it gives the bundle

$$M \times \mathbb{R} \quad (m, k)$$

$$\downarrow \qquad \wr$$

$$M \qquad m$$

A cross-section here is clearly just a smooth function $M \longrightarrow \mathbb{R}$.

§ 26. Canonical Transformations described by generating functions

A function $F(q^1, \ldots, q^n, P_1, \ldots, P_n)$ of 2n variables will yield a canonical transformation. We first describe informally how this arises. Suppose that the quantities $q^1, \ldots, q^n, P_1, \ldots, P_n$ are coordinates on some local manifold U and that the matrix

$$\left\| \ \frac{\partial^2 F}{\partial q' \partial P_j} \ \right\|$$

is everywhere non-singular. Define 2n more quantities p_i and Q^i (= smooth functions) on U by the equations

$$q^i = q^i \ , \ i = 1, \ldots, n \qquad Q^i = \frac{\partial F}{\partial P_i}$$

$$p_i = \frac{\partial F}{\partial q^i} \ , \ i = 1, \ldots, n \qquad P_i = P_i \ , \ i = 1, \ldots, n \ .$$

The assumption on the matrix above, plus the standard implicit function theorem, tells us that the p_i, q^i or the P_i, Q^i may also serve as coordinates on U. In particular, there is then a transformation from the $p_i q^i$ to the $P_i Q^i$ coordinates. <u>This</u> is the transformation "generated" by the given function F. To show that it is indeed a canonical transformation we calculate the differential

$$d\left(\sum P_i Q^i - F \right) = \sum dP_i Q^i + \sum P_i dQ^i - \sum \frac{\partial F}{\partial q^i} \, dq^i - \sum \frac{\partial F}{\partial P_i} \, dP^i \ .$$

Inserting the values chosen for p_i and Q^i above gives

$$d\left(\sum P_i Q^i - F \right) = \sum P_i dQ^i - \sum p_i dq^i \ .$$

Taking the differential once more gives

$$\sum dP_i \wedge dQ^i = \sum dp_i \wedge dq^i \ .$$

so the indicated transformation is indeed canonical.

Similar transformations may be generated from functions G of other sets of variables, say $G(Q^1, \ldots, Q^n, p_1, \ldots, p_n)$. The formalism may be found in Goldstein; we turn now to a more conceptual explanation.

Theorem. Let M be a $2n$-dimensional local manifold with coordinates $\{\begin{smallmatrix} P_i \\ q^i \end{smallmatrix}\}$, $M \xrightarrow{F} \mathbb{R}$ a smooth function and $\det \left\| \dfrac{\partial^2 F}{\partial P_i \partial q^i} \right\| \neq 0$ everywhere. Then $\omega = -\sum\limits_{i,j} \dfrac{\partial^2 F}{\partial q^i \partial P_j} dq^i \wedge dP_j$ is a closed 2-form with $\underbrace{\omega \wedge \ldots \wedge \omega}_{n} \neq 0$. Thus M is symplectic.

Proof. We must first show that $d\omega = 0$. But

$$d\omega = -\sum \frac{\partial^3 F}{\partial q^k \partial q^i \partial P_j} \underbrace{dq^k \wedge dq^i \wedge dP_j}_{\text{alternating}} - \sum \frac{\partial^3 F}{\partial q^i \partial P_k \partial P_j} dP_k \wedge dq^i \wedge dP_j = 0$$

Next we must show the n-fold exterior product $\omega \wedge \ldots \wedge \omega \neq 0$. Write

$$\omega = \sum a_{ij} dq^i \wedge dP_j .$$

In the n-fold product many terms (iterated factors) drop out; there remain the following terms, for all permutations σ and τ of the symmetric group on n letters:

$$\sum \pm \left(\prod a_{\sigma_i \tau_j} \right) dq^i \wedge \ldots \wedge dq^n \wedge dP_j \wedge \ldots \wedge dP_n$$

so one gets the determinant $n!$ times and

$$\omega \wedge \ldots \wedge \omega = n! \det \| a_{ij} \| \neq 0 .$$

Thus (M, ω) is a symplectic manifold, as required.

Now the definitions

$$\left. \begin{array}{l} P_i = \dfrac{\partial F}{\partial q^i} \\ q^i = q^i \end{array} \right\} \quad i = 1, \ldots, n$$

give 2n coordinates; since $\omega = \sum dp_i \wedge dq^i$ in these coordinates, they

are canonical coordinates. Indeed,

$$\sum dp_i \wedge dq^i = \sum d(\frac{\partial F}{\partial q^i}) \wedge dq^i = \underbrace{\sum \frac{\partial^2 F}{\partial q^i \partial q^i} \wedge dq^i}_{0} + \sum \frac{\partial^2 F}{\partial q^i \partial p_j} \wedge dp^j \wedge dq^i.$$

This is exactly the 2-form ω defined above.

Similarly, the definitions

$$\left. \begin{array}{l} P_i = P_i \\ \\ Q_i = \dfrac{\partial F}{\partial P_i} \end{array} \right\} \quad i = 1, \ldots, n$$

give 2n coordinates which are canonical coordinates, since the 2-form may be written

$$\omega = \sum dP_i \wedge dQ^i.$$

Note the advantage of using differential forms. Specifying any closed non-degenerate 2-form makes U symplectic -- no matter what the coordinates. Here there are three possible coordinate systems

$$q^1, \ldots, q^n, p_1, \ldots, p_n \qquad \omega = \sum dp_i \wedge dq^i$$

$$q^1, \ldots, q^n, P_1, \ldots, P_n \qquad \omega = - \sum \frac{\partial^2 F}{\partial q^i \partial P_j} dq^i \wedge dP_j$$

$$Q^1, \ldots, Q^n, P_1, \ldots, P_n \qquad \omega = - \sum dP_i \wedge dQ^i.$$

The first and third systems are symplectic, but the transformation is generated by going through the intermediate non-symplectic coordinate system.

Such transformations may now be used to simplify a given Hamiltonian function H (a smooth function on U, given in terms of the

coordinates q^i, p_i). If we relabel the generating function F as W, then $p_i = \partial W / \partial q^i$, and

$$H(q^1, \ldots, q^n, p_1, \ldots, p_n) = H(q^1, \ldots, q^n; \frac{\partial W}{\partial q^1}, \ldots, \frac{\partial W}{\partial q^n}) .$$

We propose to choose new canonical coordinates Q^i, P_i so that H will become simply P_1, the first P-coordinate. Writing α_1 for P_1, and considering α_1 as a constant (a parameter) this yields the equation

$$H(q^1, \ldots, q^n, \frac{\partial W}{\partial q^1}, \ldots, \frac{\partial W}{\partial q^n}) = \alpha_1 .$$

This is a first order partial differential equation for the function W. It is called the <u>Hamilton-Jacobi</u> partial differential equation. If we find a solution W depending on $n-1$ additional "constants of integration"

$$W = W(q^1, \ldots, q^n; \alpha_1, \ldots, \alpha_n)$$

satisfying the condition

$$\det \left\| \frac{\partial^2 W}{\partial q^i \partial \alpha_j} \right\| \neq 0 , \quad i, j = 2, \ldots, n$$

we can then prove that

$$\det \left\| \frac{\partial^2 W}{\partial q^i \partial \alpha_j} \right\| \neq 0 , \quad i, j = 1, \ldots, n .$$

Then W, with the α's replaced by P's, may serve as the generating function of a canonical transformation. In the new canonical coordinates the Hamiltonian function is just P_1. Therefore Hamilton's equations become

$$\dot{Q}_i = \delta_{i1} ,$$

$$\dot{P}_i = 0 , \quad i = 1, \ldots, n$$

and may be immediately integrated as

$$Q_i = t\delta_{i1} + \beta_i \ , \quad P_i = \alpha_i \ , \quad i = 1, \ldots, n \ .$$

From the transformation equations

$$P_i = \frac{\partial W}{\partial q^i} \ , \quad Q^j = \frac{\partial W}{\partial \alpha_j} = t\delta_{j1} + \beta_j$$

one may then solve for the original coordinates p_i and q_i as functions

of t.

§27. The Top

As an example, we consider the rigid motion of a heavy top, using

the "Euler angles" $\theta, \psi,$ and ϕ as the parametrization of the rotation

group in 3-space, with the axis of the top along the z-axis (see figure). Let

$\omega_x, \omega_y,$ and ω_z be the angular velocities about the x, y, and z axes. The

kinetic energy of the top is then

$$T = \frac{1}{2} [I_1 \omega_x^2 + I_2 \omega_y^2 + I_3 \omega_z^2] \ ,$$

where I_1, I_2, I_3 are the diagonal terms of the moment of inertia tensor.

(The axes are chosen so that the non-diagonal elements are zero.)

Choose the top symmetric with respect to the x- and y-axes so $I_1 = I_2$.

One may calculate from the definition of the Euler angles that

$$\omega_z = \dot{\psi} + \dot{\phi} \cos\theta \ ,$$

$$\omega_y = \dot{\phi} \sin\theta \sin\psi + \dot{\theta} \cos\psi \ ,$$

$$\omega_x = \dot{\phi} \sin\theta \cos\psi - \dot{\theta} \sin\psi \ .$$

Therefore

$$T = \frac{1}{2}[I_1(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + I_3(\dot{\psi} + \dot{\phi}\cos\theta)^2] \, .$$

Since the momentum coordinates p_i are defined as $p_i = \dfrac{\partial T}{\partial \dot{q}^i}$, we have

$$p_1 = I_1\dot{\theta} \, ,$$

$$p_2 = I_3(\dot{\psi} + \dot{\phi}\cos\theta),$$

$$p_3 = I_3(\dot{\psi}\cos\theta) + (I_1\sin^2\theta + I_3\cos^2\phi)\dot{\phi}$$

from which we obtain

$$H = \frac{1}{2}\left[\frac{p_1^2}{I_1} + \frac{p_2^2}{I_3} + \frac{1}{I_1}\left(\frac{p_2\cos q^1 - p_3}{\sin q^1}\right)^2\right] + Mg\ell\cos q^1$$

where ℓ is the distance along z from the origin to the center of mass
(see figure).

Substituting $p_i = \dfrac{\partial W}{\partial q^i}$ gives the Hamilton-Jacobi partial

differential equation for W. Since H also is independent of q^2 and q^3,

we know that $p_2 = \alpha_2$ and $p_3 = \alpha_3$, α_2 and α_3 constants are solutions of

Hamilton's equations. Hence we "separate the variables" in W, writing

$$W = W_1(q^1) + W_2(q^2) + W_3(q^3) \, ,$$

where we may take

$$W_2(q^2) = \alpha_2 q_2 + \text{constant,}$$

$$W_3(q^3) = \alpha_3 q^3 + \text{constant,}$$

since $\dfrac{\partial W}{\partial q^i} = p_i = \alpha_i$ for $i = 2, 3$.

The remaining equation for $\dfrac{dW}{dq^1}$ is

$$\left(\frac{dW}{dq^1}\right)^2 = I_1\left[2\alpha_1 - 2Mg\ell \cos q^1 - \frac{\alpha_2^2}{I_3} - \frac{1}{I_1}\left(\frac{\alpha_2 \cos q^1 - \alpha_3}{\sin q^1}\right)^2\right].$$

Set $u = \cos q^1$; the equation becomes

$$(1 - u^2)\left(\frac{dW}{dq^1}\right)^2 = I_1\left[\left(2\alpha_1 - \frac{\alpha_2^2}{I_3} - 2Mg\ell u\right)(1 - u^2) - \frac{1}{I_1}(\alpha_2 u - \alpha_3)^2\right]$$

$$= F(u, \alpha_1) = (1 - u^2)(h - ku) - \left(\frac{\alpha_2}{I_1} u - \frac{\alpha_3}{I_1}\right)^2 ,$$

where h and k are defined accordingly so

$$W = \int \sqrt{\frac{F(u, \alpha_1)}{1 - u^2}} \; dq^1 .$$

But we want $Q^1 = \dfrac{\partial W}{\partial \alpha_1}$ so $\dot{Q}^1 = \dfrac{\partial H}{\partial p_1} = 1$ and $Q_1 = t + \beta$, so

$$Q^1 = \frac{\partial W}{\partial \alpha_1} = \int \frac{dq_1(1 - u^2)}{\sqrt{1 - u^2}\sqrt{F(u, \alpha_1)}} = \int \frac{du}{\sqrt{F(u, \alpha_1)}} ,$$

so

$$\beta_1 + t = \int \frac{du}{\sqrt{F(u, \alpha_1)}} , \quad \text{where} \quad u = \cos q^1 = \cos \theta.$$

Since F is a cubic polynomial in u, this is an elliptic integral; the detailed explicit solution is given in Klein and Sommerfeld (4-volumes) on the gyroscope.

We consider now the qualitative properties of the solution. Since $F(u) = ku^3 + \ldots$, with $k > 0$, we have $F(\infty) = +\infty$, $F(-\infty) = -\infty$, and $F(u) = 0$ has three roots. The roots u_1 and u_2 between -1 and $+1$ (see graphs) are the only ones of physical interest since $-1 \leq u \leq 1$ is the

only part which is physically possible. Now

$$\frac{dt}{du} = \frac{1}{\sqrt{F(u)}} \qquad \text{implies} \quad \left(\frac{du}{dt}\right)^2 = F(u)$$

so the zeroes of $F(u)$ are where $\frac{du}{dt}$ is zero.

If we take the positive square root of F, we get a solution for u increasing from u_1 to u_2 as time goes from 0 to, say, A. This may be extended by reflection (negative root of F) to give a solution decreasing from u_2 to u_1. Continuing, as in the figure, we have a solution for all t
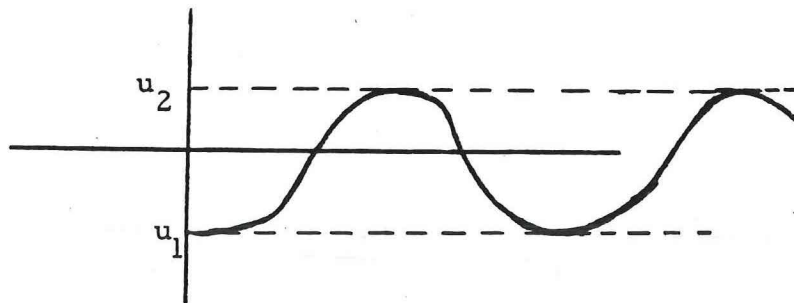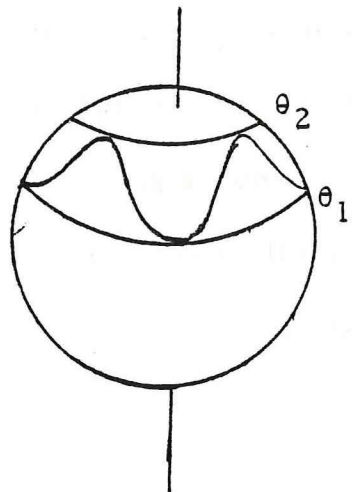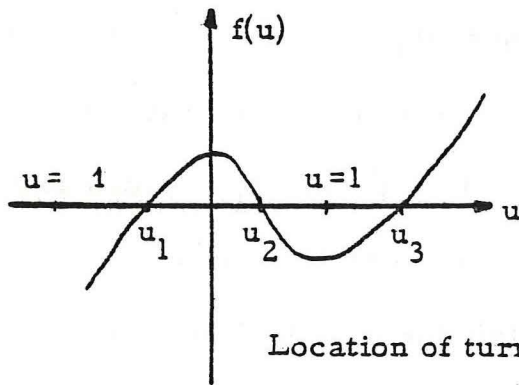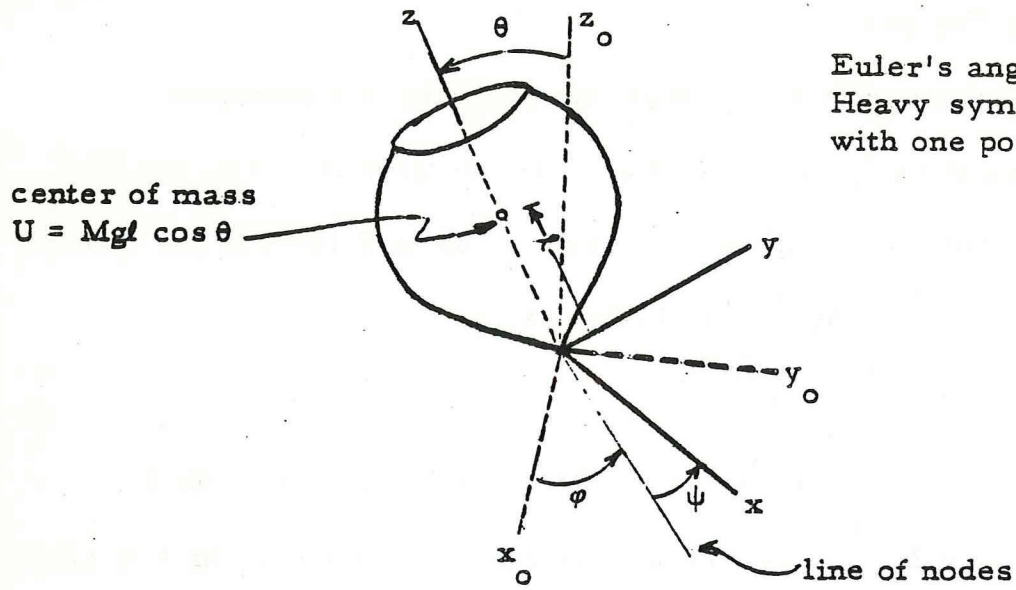

of the general form indicated, so that $u(t)$ is periodic with period $2A$.

This solution may be pictured in terms of the angles (θ, φ) which give the point where the axis of the top pierces the unit sphere. For example, $p_2 = \alpha_2$ and $p_3 = \alpha_3$ allows us to solve the equations above for $\dot{\varphi}$ as

$$\dot{\varphi} = \frac{\alpha_3 - \alpha_2 \cos\theta}{I_1 \sin^2\theta} = \frac{\alpha_3 - \alpha_2 u}{I_1(1 - u^2)} .$$

Here $\dot{\varphi}$ is the speed of <u>precession</u> of the axis of the top around the vertical, while $\dot{\theta}$ is the speed of rotation. For further qualitative discussion, see Goldstein, <u>Classical Mechanics</u>, pp. 164-175 or W.F. Osgood, <u>Mechanics</u>.

Euler's angles and the
Heavy symmetric top
with one point fixed

center of mass
$U = Mgl \cos\theta$

line of nodes

Location of turning angles

One possible Locus of the figure axis on
the unit sphere

§28. Darboux's Theorem.

The typical example of a symplectic manifold is a cotangent bundle of configuration space. In such a bundle we have the usual position and momentum coordinates q^i and p_i, and the basic 2-form of the manifold is given in terms of these coordinates as

$$\omega = \sum dp^i \wedge dq^i .$$

The definition of a symplectic manifold (M, ω) was apparently much more general: Any 2n-dimensional manifold with a 2-form ω which is closed $(d\omega = 0)$ and non-degenerate $(\omega \wedge \ldots \wedge \omega$, to n factors, nowhere 0).

The added generality is illusory: Darboux's theorem asserts that at each point a of such a manifold there is always an open set U containing a and coordinates q^i and p_i -- good in this neighborhood U -- for which ω has the special form above.

One proof of this theorem is done by systematically exploiting the correspondence between vector fields and forms which is given by the basic form ω. This is the proof given in Abraham, pp. 92-94. Another proof (see Sternberg p. 137) depends on a more general theorem of Frobenius on the integration of differential systems. We refer to these texts for details.
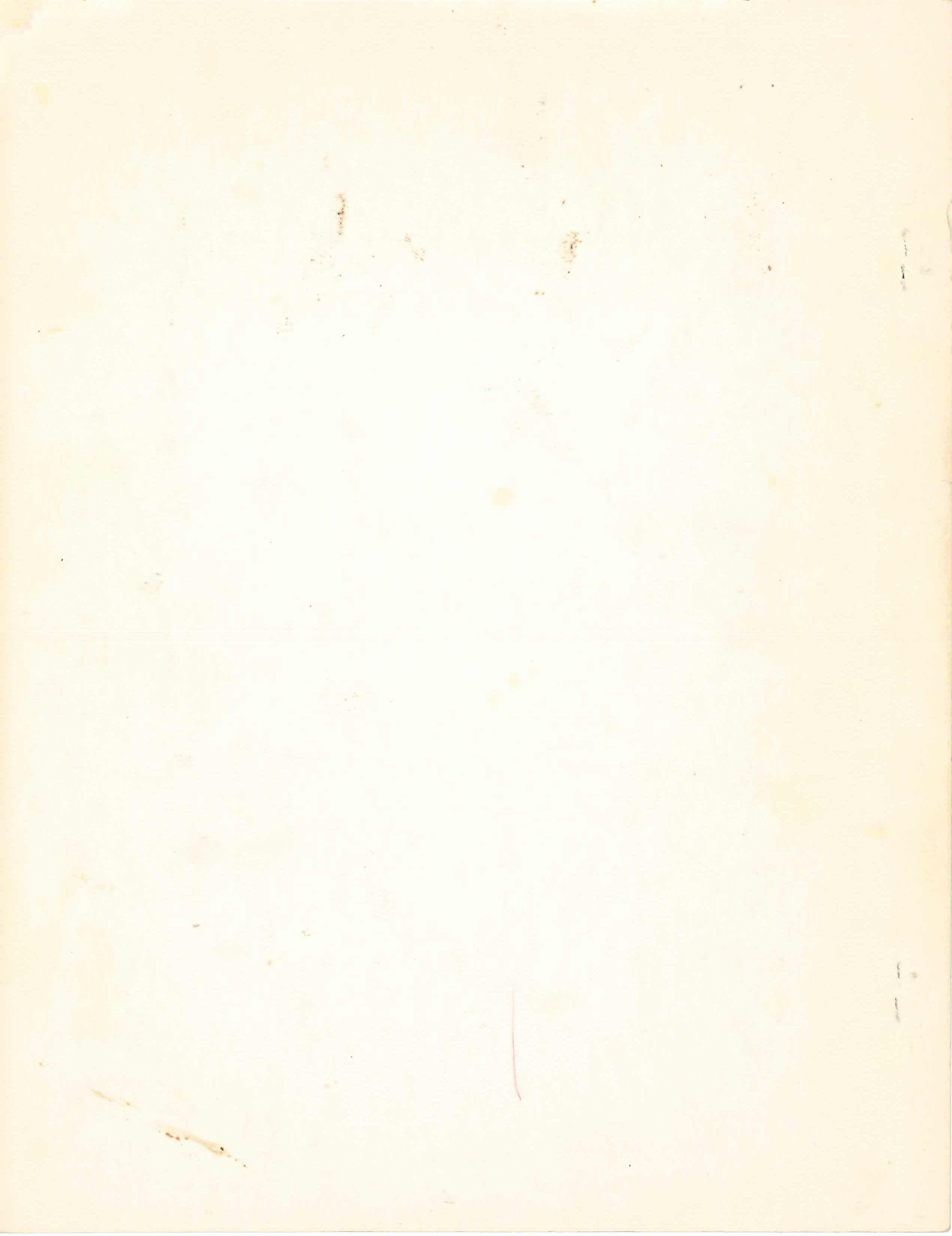
Corrections to Geometrical Mechanics, Part I, Saunders Mac Lane

(1-4 means 4 lines from foot of the page)

p. 10, line 2 (display) gij after $\sum$ should be g_{1j}

p. 10, line 7, 1-5, 1-1: All V^1 should be l.c. v^1.

p. 11, line 6 (display): $V_1 \longrightarrow F_1$

p. 11, line 10 (display): (see below)

p. 18: Proof of theorem incomplete because addition of tangent vectors is not explicitly defined. Definition should be by map $\rho : T_a U \longrightarrow (T^a U)^\#$ given in line 1 via

$$\rho(\tau_a c + \tau_a c') = \rho\tau_a c + \rho t_a c' \text{ (addition in } T^a U)^\#$$

This requires proof that ρ is onto. Use local coordinates $q^1, \ldots, q^n$

$$\theta : T_a U \longrightarrow \mathbb{R}^n \text{ by } \theta\tau_a c = (\frac{\partial c}{\partial q^1}, \ldots, \frac{\partial c}{\partial q^n})$$

$$\varphi : T^a U \longrightarrow (\mathbb{R}^n)^\# \text{ by } \varphi df = (\frac{\partial f}{\partial q^1}, \ldots, \frac{\partial f}{\partial q^n})$$

$$\varphi^\# : \mathbb{R}^n \longrightarrow (T^a U)^\#$$

and thus show by calculation that the diagram

$$
\begin{array}{ccc}
T_a U & \xrightarrow{\;\;\rho\;\;} & (T^a U)^\# \\
& \theta \searrow \quad \nearrow \varphi^\# & \\
& \mathbb{R}^n &
\end{array}
$$

with θ, $\varphi^\#$ both isomorphisms, commute ($\varphi^\#\theta = \rho$)
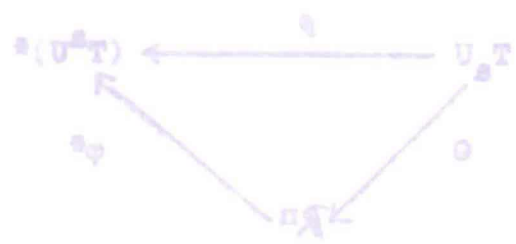
p. 20, 1-1,-2 $\hat{f}$ should be $k : V \longrightarrow \mathbb{R}$

p. 11, line 10 (display): ∂p_1 should be ∂q^1 (1st equation)