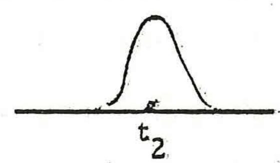
<u>Lemma.</u> If $M: I \longrightarrow \mathbb{R}$ is a smooth function and if

$$\int_{t_o}^{t_1} M\eta \, dt = 0$$

for all smooth functions $\eta: I \longrightarrow \mathbb{R}$ with $\eta(t_o) = \eta(t_1) = 0$, then $M(t) = 0$ for all t with $t_o \leq t \leq t_1$.

<u>Proof.</u> Suppose instead that $M(t) \neq 0$ for some $t = t_2$, say that $M(t_2) > 0$. Then $M(t) > 0$ on some small interval about t_2 and we can choose a "bump" function $b: I \longrightarrow \mathbb{R}$ which is smooth, zero outside the interval, positive inside this interval, and 1 at t_2:



Then choosing $\eta = bM$ as variation, the hypothesis gives

$$\int_{t_o}^{t_1} M\eta \, dt = \int_{t_o}^{t_1} bM^2 \, dt > 0.$$

This contradiction gives $M = 0$, as desired.

The methods used here are those of the Calculus of Variations. The result can be formulated more generally, as follows

Given $h: T.U \times I \longrightarrow \mathbb{R}$, consider paths $c_o: I \longrightarrow U$ which make $\int_{t_o}^{t_1} h\tilde{c} \, dt$ stationary in comparison with other paths c, $c(t_o) = c_o(t_o)$, $c(t_1) = c_o(t_1)$. A necessary condition for this is Euler's equation:

$$\frac{d}{dt}\left(\frac{\partial h}{\partial \dot{q}^i}\right) = \frac{\partial h}{\partial q^i}, \quad i = 1, \ldots, n.$$

In the special case when h is the Lagrangian function L_1, Euler's equations are Lagrange's equations. In more general treatments, the smooth paths used above can be replaced by "piecewise" smooth paths.

§ 8 Bilinear and Quadratic Forms

The kinetic energy $\mathcal{T}$ is usually a quadratic function of the velocities; that is, $\mathcal{T}: T.U \longrightarrow \mathbb{R}$ restricted to the fiber (tangent space) over a point of U is a quadratic function on that tangent space. We now study certain properties of such quadratic functions.

Let V be a finite dimensional vector space. Consider a function $B: V \times V \longrightarrow \mathbb{R}$ $((v, w) \rightsquigarrow B(v, w))$. We define B to be bilinear if $B(v, w)$ is linear in v (with w fixed) and linear in w (with v fixed). We define $Q: V \longrightarrow \mathbb{R}$ to be quadratic when,

1° $Q(-v) = Q(v)$

2° $Q(u+v) - Q(u) - Q(v) \overset{\text{def.}}{=\!=} 2Q^\flat(u, v)$ is bilinear in u and v.

That is, Q determines a symmetric bilinear function $Q^\flat$.

As a consequence of bilinearity, we have

$Q(u+v+w) - Q(u) - Q(v+w) = Q(u+v) - Q(u) - Q(v) + Q(u+w) - Q(u) - Q(w)$.

Letting $u = v = -w$, $Q(u) - Q(u) = Q(2u) - Q(u) - Q(u) + 0 - Q(u) - Q(u)$ and thus $Q(2u) = 4Q(u)$. (We must have $Q(0) = 0$ since $Q^\flat(0, 0) = 0$.)

The assignment $\underset{\text{quadratic}}{Q} \rightsquigarrow Q^\flat \overset{\text{def.}}{=\!=} \frac{1}{2}[Q(u+v) - Q(u) - Q(v)]$ symmetric and bilinear, has an inverse $B \rightsquigarrow B^{\#}$. Define $B^{\#}(u) = B(u, u)$, for B symmetric and bilinear. Clearly, $(Q^\flat)^{\#} = Q$, since

$$(Q^\flat)^\#(u) = Q^\flat(u,u) = \frac{1}{2}[Q(u+u) - Q(u) - Q(u)] = \frac{1}{2}[4Q(u) - 2Q(u)] = Q(u).$$

Conversely, $B^\#$ is quadratic:

1° $B^\#(-u) = B(-u,-u) = B(u,u)$

and

2° $B^\#(u+v) - B^\#(u) - B^\#(v)$ $(\overset{\text{def}}{=\!=}\ 2(B^\#)^\flat(u,v)) = B(u+v,u+v) - B(u,u) - B(v,v)$

$$= B(u,v) + B(v,u) = 2B(u,v).$$

And from this calculation, clearly $(B^\#)^\flat = B$.

Given $Q\colon V \longrightarrow \mathbb{R}$ quadratic and $W \xrightarrow{\varphi} V$ a linear transformation, then $Q\varphi\colon W \longrightarrow \mathbb{R}$ is quadratic. For the proof, note

$$Q = V \xrightarrow{\Delta} V \times V \xrightarrow{B} \mathbb{R}$$
$$v \rightsquigarrow (v,v) \rightsquigarrow B(v,v) = Q(v)$$

where $B = Q^\flat$ is symmetric and bilinear. Check that

$W \times W \xrightarrow{\varphi\times\varphi} V \times V \xrightarrow{B} \mathbb{R}$ is symmetric and bilinear. The rest of the proof is indicated by the commutative diagram:

$$\begin{tikzcd}
W \arrow[r, "\varphi"] \arrow[d, "\Delta"'] & V \arrow[d, "\Delta"'] \arrow[dr, "Q"] & \\
W\times W \arrow[r, "\varphi\times\varphi"] & V\times V \arrow[r, "B"] & \mathbb{R}
\end{tikzcd} .$$

Choosing a basis $e_1,\ldots,e_n$ for V, and letting $v = \sum_{i=1}^{n} q^i e_i$, we have $B(\sum_i q^i e_i, \sum_j q^j e_j) = \sum_{i,j} q^i q^j B(e_i, e_j)$. Defining $g_{ij} = B(e_i, e_j)$, we have $Q(v) = \sum_{i,j=1}^{n} g_{ij} q^i q^j$; we may call $\|g_{ij}\|$ the matrix of Q for the basis e_i.

Consider the functions indicated in the following diagram

$$\begin{tikzcd}
\mathbb{R} & T.U \arrow[l, shift right, "\dot q^1"'] \arrow[l, shift left, "\dot q^n"] \arrow[r, shift left, "J"] \arrow[r, shift right, "L"'] \arrow[d, "\pi"] & \mathbb{R} \\
\mathbb{R} & U \arrow[l, shift right, "q^1"'] \arrow[l, shift left, "q^n"] \arrow[r, "\mathcal{V}"] & \mathbb{R}
\end{tikzcd}$$

The $\dot{q}^i$ and q^i are the usual coordinates and $\mathcal{T}$ is quadratic and positive definite on each fiber, thus $\mathcal{T}$ is a Riemann metric on U. $L \underline{\underline{\mathrm{def}}}\ \mathcal{T} - \mathcal{V}\pi$ is a quadratic plus a constant on each fiber. Let $\mathcal{T} = \sum_{i,j=1}^{n} g_{ij}\dot{q}^i\dot{q}^j$ where $g_{ij}: U \longrightarrow \mathbb{R}$, that is for $u \in U$, $(g_{ij}(u))$ is the positive definite symmetric matrix associated with the quadratic form $\mathcal{T}$ induces on $T_u U$, the fiber over U. But we have already noted that such a Riemann metric gives an isomorphism

$$T_u U \longrightarrow (T_u U)^* = T^u U$$

of the tangent space to its dual, the cotangent space. This isomorphism carries a point with coordinates $(\dot{q}^1, \ldots, \dot{q}^n)$ to the point with coordinate p_i, where

$$p_i = \sum g_{ij}\dot{q}^j = \frac{\partial \mathcal{T}}{\partial \dot{q}^i} .$$

If we apply this isomorphism to each fiber of the tangent bundle, we get a smooth map $\mathcal{L}: T_{\cdot}U \longrightarrow T^{\cdot}U$ called the Legendre transformation, as in the diagram

$$\begin{tikzcd}
\mathbb{R} & T_{\cdot}U \arrow[l, shift right, "\dot{q}^1"'] \arrow[l, shift left, "\dot{q}^n"] \arrow[rr, "\mathcal{L}"] \arrow[dr, "\pi_{\cdot}"'] & & T^{\cdot}U \arrow[r, shift left, "p^1"] \arrow[r, shift right, "p^n"'] \arrow[dl, "\pi^{\cdot}"] & \mathbb{R} \\
& & U \arrow[r, shift left, "q^1"] \arrow[r, shift right, "q^n"'] & \mathbb{R} &
\end{tikzcd}$$

The coordinates for $T_{\cdot}U$ are $q^1\pi_{\cdot}, \ldots, q^n\pi_{\cdot}, \dot{q}_1, \ldots, \dot{q}_n$ and for $T^{\cdot}U$ they are $q^1\pi^{\cdot}, \ldots, q^n\pi^{\cdot}, p_1, \ldots, p_n$. The map $\mathcal{L}$ is then defined by

$$q^i\pi^{\cdot}\mathcal{L} = q^i\pi_{\cdot} \quad , \quad p_i\mathcal{L} = \frac{\partial \mathcal{T}}{\partial \dot{q}^i} .$$

We now consider the properties of this Legendre transformation $\mathcal{L}$

§9 The Legendre Transformation

Suppose U is an open set in $\mathbb{R}^n$ with coordinates $q^1, \ldots, q^n$ and we are given a smooth kinetic energy function $\mathcal{J}$ from $T.U$ to $\mathbb{R}$. We shall assume $\mathcal{J}$ quadratic; i.e., $\mathcal{J} = \frac{1}{2} \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial \dot{q}^i} \dot{q}^i$. Such a function determines a smooth function $\mathcal{L}$ defined by $p_i\mathcal{L} = \frac{\partial \mathcal{J}}{\partial \dot{q}^i}$, $q^i\pi^{\cdot}\mathcal{L} = q^i\pi_{.}$. We set $\hat{\mathcal{J}} = \mathcal{J}\mathcal{L}^{-1}$. From the quadratic assumption on $\mathcal{J}$ we have

$$\begin{tikzcd} \mathbb{R} & T.U \arrow[l, "\mathcal{J}"'] \arrow[r, "\mathcal{L}"] \arrow[dr, "\pi_{.}"'] & T^{\cdot}U \arrow[ll, bend right, "\hat{\mathcal{J}}"'] \arrow[d, "\pi^{\cdot}"] \\ & & U \end{tikzcd}$$

$$2d\mathcal{J} = \sum^{n} d\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right)\dot{q}^i + \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial \dot{q}^i} d\dot{q}^i .$$

Subtract $d\mathcal{J} = \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial q^i\pi_{.}} dq^i\pi_{.} + \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial \dot{q}^i} d\dot{q}^i$ to get

$d\mathcal{J} = \sum_{i=1}^{n} d\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right)\dot{q}^i - \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial q^i\pi_{.}} dq^i\pi_{.}$ Then

$$\begin{aligned} d\hat{\mathcal{J}} &= d(\mathcal{J}\mathcal{L}^{-1}) = (\mathcal{L}^{-1})^* d\mathcal{J} \\ &= \sum_{i=1}^{n} d\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\mathcal{L}^{-1}\right)\dot{q}^i\mathcal{L}^{-1} - \sum_{i=1}^{n} \frac{\partial \mathcal{J}}{\partial q^i\pi_{.}}\mathcal{L}^{-1} d(q^i\pi_{.}\mathcal{L}^{-1}) . \end{aligned}$$

(Recall that if $X \xrightarrow{\phi} Y \xrightarrow{f} \mathbb{R}$ are smooth functions, then f pulls back via ϕ to the function $f\phi$ and we have $d(f\phi) = \phi^* df$, $\phi^*(g df) = (g\phi)\phi^* df$.)

Finally, using the defining equations for $\mathcal{L}$,

$$d\hat{\mathcal{J}} = \sum_{i=1}^{n} (dp_i)\dot{q}^i\mathcal{L}^{-1} - \sum_{i=1}^{n} \left(\frac{\partial \mathcal{J}}{\partial q^i\pi_{.}}\mathcal{L}^{-1}\right) dq^i\pi^{\cdot} .$$

But also $d\hat{\mathcal{J}} = \sum_{i=1}^{n} \frac{\partial \hat{\mathcal{J}}}{\partial p_i} dp_i + \sum_{i=1}^{n} \frac{\partial \hat{\mathcal{J}}}{\partial q^i \pi^\cdot} dq^i \pi^\cdot$, and the coefficients of the differentials are unique. Thus

$$\dot{q}^i \zeta^{-1} = \frac{\partial \hat{\mathcal{J}}}{\partial p_i} \quad \text{and} \quad -\frac{\partial \mathcal{J}}{\partial q^i \pi_\cdot} \zeta^{-1} = \frac{\partial \hat{\mathcal{J}}}{\partial q^i \pi^\cdot},$$

which are the equations of the transferred kinetic energy $\hat{\mathcal{J}}$ in terms of the given one $\mathcal{J}$.

Now if we have a mechanical system with Lagrangian $L = \mathcal{J} - \mathcal{V}$ in which a path c satisfies Lagrange's equations, then these last equations yield Hamilton's equations for c. Let's follow the convention in mechanics of not writing in the maps $\pi_\cdot, \pi^\cdot$ and ζ, so the last equations are

$$\frac{dq^i}{dt} = \frac{\partial \hat{\mathcal{J}}}{\partial p_i} \quad \text{and} \quad -\frac{\partial \mathcal{J}}{\partial q^i} = \frac{\partial \hat{\mathcal{J}}}{\partial q^i}.$$

By definition of ζ

$$\frac{dp_i}{dt} = \frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right) = \frac{d}{dt}\left(\frac{\partial \zeta}{\partial \dot{q}^i}\right)$$

$$= \frac{\partial \zeta}{\partial q^i} = \frac{\partial \mathcal{J}}{\partial q^i} - \frac{\partial \mathcal{V}}{\partial q^i} = -\left(\frac{\partial \hat{\mathcal{J}}}{\partial q^i} + \frac{\partial \mathcal{V}}{\partial q^i}\right).$$

So by setting $\mathcal{H} = \hat{\mathcal{J}} + \mathcal{V}\zeta^{-1}$, we have Hamilton's equations

$$\frac{dq^i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i} \quad \text{and} \quad \frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q^i}.$$

<u>Exercise:</u> Let $n = 1$ and $\mathcal{J} = g\dot{q}^2$, where $U \xrightarrow{g} \mathbb{R}$. Calculate $\hat{\mathcal{J}}$.

We want to understand better how this scheme produced these equations. Notice first that χ maps $T_a U$ into $T^a U$ -- which is to say $(\chi, 1_U)\colon \pi_{\cdot} \longrightarrow \pi^{\cdot}$ is a morphism of prebundles. Thus we may as well look at χ on each fibre of $T_{\cdot}U$ and paste the fibres together where we must. So consider a finite dimensional vector space V (think of this as $T_a U$ for some $a \in U$):

a) At each point $v \in V$, $T_v V \cong V$ $(\tau_v c \rightsquigarrow w)$, where $c(t) = v + tw$ is a curve $I \longrightarrow V$. (Identify V with $T_v V$ by this isomorphism.)

b) At each $v \in V$, $T^v V \cong V^*$ $(d_v f \rightsquigarrow \overline{d_v f})$ where

$$\overline{d_v f}(w) = \frac{d}{dt}(f(v+tw))\Big|_{t=0} \quad \text{and} \quad V \xrightarrow{f} \mathbb{R} .$$

(Again identify V^* and $T^v V$.)

c) $T^{\cdot}V$ can be identified with $V \times V^*$ via

$$(v, d_v f) \rightsquigarrow (v, \overline{d_v f})$$

$$\begin{tikzcd}
T^{\cdot}V \arrow[r] \arrow[d, "\pi^{\cdot}"'] & V \times V^* \arrow[d, "\pi"] \\
V \arrow[r, "="] & V
\end{tikzcd}$$

where $\pi(v, \alpha) = v$.

We have another projection $V \times V^* \longrightarrow V^*$ $(v, \alpha) \rightsquigarrow \alpha$.

d) A 1-form ω on V is a smooth function $\omega\colon T$— $T^{\cdot}V$ such that $\pi \circ \omega = 1_V$. Composition of σ with a 1-form ω produces a smooth function

$$\chi_\omega = \sigma \circ \omega .$$

$$\begin{tikzcd}
V \times V^* \arrow[r, "\sigma"] \arrow[d, shift left, "\pi"] & V \\
V \arrow[u, shift left, "\omega"] \arrow[ur, "\chi_\omega"'] &
\end{tikzcd}$$

e) In particular, suppose $L: V \longrightarrow \mathbb{R}$ is smooth (think of L as a Lagrangian function restricted to one fiber of $T.U$). Then dL is a 1-form and so determines $V \xrightarrow{\mathcal{L}_{dL}} V^*$. To compare this with the example at the beginning of this section, let the potential energy be zero so $L = \mathcal{J}$. The first defining equation for $\mathcal{L}$ says $\mathcal{L}_{d\mathcal{J}} = \sigma \circ d\mathcal{J}$ when written in coordinates. Returning to the general case, the explicit formula for $\mathcal{L}_{dL}$ gives each value $\mathcal{L}_{dL}V$ as a function of w:

$$\mathcal{L}_{dL}V(w) = \frac{d}{dt} L(v+tw)\Big|_{t=0} .$$

f) Let $e_1, \ldots, e_n$ be a basis for V with coordinates $e^1, \ldots, e^n$; then $e^1, \ldots, e^n$ are a dual basis for V^* with coordinates $e_1, \ldots, e_n$. The formula for $\mathcal{L}_{dL} = \mathcal{L}$ in these bases is

$$\mathcal{L}v(w) = \sum_{j=1}^{n} \left(\frac{\partial L}{\partial e^j} \Big|_v \right) e^j w .$$

Apply e_i to both sides and use $e_i e^j = \delta^i_j$ to get $e_i \mathcal{L} v = \frac{\partial L}{\partial e^i}\Big|_v$ or, as functions

$$e_i \mathcal{L} = \frac{\partial L}{\partial e^i} \quad , \quad i = 1, \ldots, n$$

In the mechanical rotation $e^i = \dot{q}^i$ and $e_i = p_i$, since $V = T_a U$ and $V^* = T^a U$ for $a \in U$, U open in $\mathbb{R}^n$, and the result reads

$$p_i \mathcal{L} = \frac{\partial L}{\partial \dot{q}^i}$$

as before.

The Hamiltonian function arises from asking the question: when is $\mathcal{L}$ invertible? (This is probably not the way Hamilton found it.)

By the Inverse Function Theorem (see Abraham p. 10), $\mathcal{L}^{-1}$ exists if and only if the matrix $\left(\frac{\partial^2 L}{\partial e^i \partial e^j}\Big|_v \right)$ is non-singular at every $v \in V$.

g) Suppose $\mathcal{L}$ invertible. Does $\mathcal{L}^{-1}$ come from a smooth function on V^* in the same manner $\mathcal{L}$ came from a smooth L on V? That is, is there $V^* \xrightarrow{H} \mathbb{R}$ such that $\mathcal{L}^{-1} = \pi \circ dH$?

$$\begin{tikzcd} V \times V^* \arrow[r, "\pi"] \arrow[d, shift left, "\sigma"] & V \arrow[dl, "\mathcal{L}^{-1}"] \\ V^* \arrow[u, shift left, "dH"] & \end{tikzcd}$$

From part f) we have

$$dL = \sum_{i=1}^{n} (e_i \mathcal{L}) de^i.$$

Let's try the formula dual to this one:

We want H so that $dH = \sum_{i=1}^{n} (e^i \mathcal{L}^{-1}) de_i$. Use the derivation property of d:

$$\sum_{i=1}^{n} (e^i \mathcal{L}^{-1}) de_i = d\left[\sum_{i=1}^{n} (e^i \mathcal{L}^{-1}) e_i\right] - \sum_{i=1}^{n} e_i d(e^i \mathcal{L}^{-1}).$$

The second term on the right is

$$- \sum_{i=1}^{n} e_i \mathcal{L} \mathcal{L}^{-1} d(e^i \mathcal{L}^{-1}) = - (\mathcal{L}^{-1})^* \sum_{i=1}^{n} (e_i \mathcal{L}) de^i$$

$$= -(\mathcal{L}^{-1})^* dL = -d(L\mathcal{L}^{-1}).$$

Substitute this in our conjectured dH to get

$$dH = d\left[\sum_{i=1}^{n} (e^i \mathcal{L}^{-1}) e_i - L\mathcal{L}^{-1}\right],$$

so H should be

$$H = \sum_{i=1}^{n} (e^i \mathcal{L}^{-1}) e_i - L\mathcal{L}^{-1}.$$

The steps reverse so this is indeed the right formula. On elements,

$$H_y = \mathcal{L}^{-1}_y \cdot y - L\mathcal{L}^{-1}_y \qquad \text{for } y \in V^*$$

usual inner product

and in mechanical notation

$$H = \sum_{i=1}^{n} \dot{q}^i p_i - L.$$

(Notice that we are leaving out $\mathcal{L}$ and $\mathcal{L}^{-1}$ as is customary in mechanics.) If, in particular, $L = \mathcal{J} - \mathcal{V}$ where $\mathcal{V}$ is a function of only the q^i's and $\mathcal{J}$ is quadratic as in our example, then

$$H = \sum_{i=1}^{n} \dot{q}^i \frac{\partial \mathcal{J}}{\partial \dot{q}^i} - L = 2\mathcal{J} - L = \mathcal{J} + \mathcal{V}$$

Given the Lagrangian L, define the <u>Action</u> $A: V \longrightarrow \mathbb{R}$, and <u>Energy</u> $E: V \longrightarrow \mathbb{R}$ by $Av = \langle \mathcal{L}_{dL} v, v \rangle$, $Ev = \langle \mathcal{L}_{dL} v, v \rangle - Lv$. Both are smooth functions.

The last lecture proved most of the following theorem:

<u>Theorem</u> (Legendre): If V is a finite dimensional real vector space of dimension n, then

1) for each $v \in V$, there is a natural isomorphism $T^v V \cong V^*$ (which we consider equality below)

2) for each smooth $L: V \longrightarrow \mathbb{R}$, there is a smooth

$$V \xrightarrow{\mathcal{L}_{dL}} V^* \qquad (v \rightsquigarrow d_v L)$$

In coordinates

$$dL = \sum_{i=1}^{n} (e_i \mathcal{L}_{dL}) de^i$$

with $e_1, \ldots, e_n$ a basis for V. $\mathcal{L}_{dL}$ is called the <u>Legendre transformation for L</u>.

3) the function $\mathcal{L} = \mathcal{L}_{dL}$ is invertible if and only if the matrix $\left(\frac{\partial^2 L}{\partial e^i \partial e^j}\Big|_v\right)$ is non-singular at each $v \in V$. If this is so $\mathcal{L}^{-1}: V^* \longrightarrow V$ is the Legendre transformation for $H: V^* \longrightarrow \mathbb{R}$ defined by

$$Hy = \langle y, \mathcal{L}^{-1} y \rangle - L\mathcal{L}^{-1} y \quad = E\mathcal{L}^{-1} y \, .$$

4) in particular, if L is quadratic, then $\mathcal{L}$ is the isomorphism of V with its dual given by the inner product induced by L and $E = L$ in this case.

[See Sternberg pp. 150-153, Goldstein pp. 215-   , Abraham §17 .]

<u>Corollary 1.</u> If U open in $\mathbb{R}^n$, I open interval $\subset \mathbb{R}$ and $L: T.U \times I \longrightarrow \mathbb{R}$ is smooth, then L on each fibre determines

$$\begin{tikzcd} T.U \times I \arrow[rr, "\mathcal{L}"] \arrow[dr] & & T^{\cdot}U \times I \arrow[dl] \\ & U \times I & \end{tikzcd}$$

and all parts of the theorem hold for this $\mathcal{L}$. (Abraham calls $\mathcal{L}$ the fibre derivative of L.)

<u>Corollary 2.</u> Let c be a path in U, $\tilde{c}$ lifted path in $T.U$. If $\tilde{c}$ satisfies Lagrange's equations for L, then $\mathcal{L}\tilde{c}$ satisfies the canonical differential equations for H (Hamilton's equations):

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i} \, .$$

<u>Exercise</u> Prove Corollary 2.

Forget the projection, so $T^{\cdot}U$ is an open set in $\mathbb{R}^{2n}$ and on it sits a first order differential equation -- Hamilton's equation of Corollary 2. We shall consider changes of coordinates in this 2n-dimensional space which leave this differential equation invariant. To do this we need to consider 2-forms, and in general k-forms.

Chapter II. Tensors and Exterior Forms

§10. Vector Fields

Let U be open in $\mathbb{R}^n$ with coordinates $q^1, \ldots, q^n$. A <u>vector field</u> X on U is a cross-section of the tangent bundle of U; i.e., $\pi_{\bullet} X(a) = a$ for all $a \in U$.

For example, for given coordinates q^i, we can define a vector field D "along the axis q^i" by $D^i(a) = (a, \overline{D}^i a)$, where $\overline{D}^i(a) = \tau_a$ (path along i^{th} coordinate axis) = unit vector in i^{th} direction

$$\begin{tikzcd} T.U \arrow[d, bend left, "\pi"] \\ U \arrow[u, bend left, "X"] \end{tikzcd}$$

in T_aU, for $i = 1, \ldots, n$. This clearly defines a cross-section D^i of the tangent bundle. The set of vector fields on U is an $\mathcal{F}$-module, where $\mathcal{F}$ is the ring of smooth functions $U \longrightarrow \mathbb{R}$. The $\mathcal{F}$-module structure is given by the equations

$$(X_1 + X_2)a = X_1 a + X_2 a$$

$$fX_1(a) = f(a) \cdot X_1 a,$$

for $a \in U$, $f \in \mathcal{F}$ and X_1, X_2 vector fields. The vectors $D^i(a)$, $i = 1, \ldots, n$ form a basis of T_aU, so

$$Xa = \sum_{i=1}^{n} (X_i a) D^i(a);$$

thus $X = \sum_{i=1}^{n} X_i D^i$, where the functions $X_i: U \longrightarrow \mathbb{R}$ are smooth and unique. This says that the vector fields $D^1, \ldots, D^n$ are a basis for the set of vector fields on U as a real vector space.

Each vector field X produces a function called the <u>Lie derivative</u> $\mathcal{F} \xrightarrow{L_X} \mathcal{F}$ with

$$L_X f(a) = \langle d_a f, Xa \rangle = \text{derivative of } f \text{ along } X.$$

Here we need "smooth" to mean C^{∞}, since otherwise $L_X f$ has one lower order of differentiability than f. The function L_X has the properties

(1) L_X is $\mathbb{R}$-linear,

(2) $L_X(f \cdot g) = f \cdot L_X g + g \cdot L_X f$.

Property (1) is a consequence of the linearity of d_a and $\langle\ , Xa\rangle$. For (2)

$$\begin{aligned} L_X(f\cdot g)a &= \langle d_a f\cdot g, Xa\rangle \\ &= \langle f(a)\cdot d_a g + g(a)\cdot d_a f, Xa\rangle \\ &= f(a)\langle d_a g, Xa\rangle + g(a)\langle d_a f, Xa\rangle \\ &= (f\cdot L_X g + g\cdot L_X f)a . \end{aligned}$$

In coordinates, $X = \sum_{i=1}^{n} X_i D^i$ so

$$\begin{aligned} L_X f &= \left\langle \sum_{j=1}^{n} \frac{\partial f}{\partial q^i}\, dq^j \ , \ \sum_{i=1}^{n} X_i D^i \right\rangle \\ &= \sum_{i=1}^{n} X_i \sum_{j=1}^{n} \frac{\partial f}{\partial q^i} \langle dq^j, D^i \rangle \\ &= \sum_{i=1}^{n} X_i \frac{\partial f}{\partial q^i} \ , \quad \text{since } \langle dq^j, D^i\rangle = \delta_i^j . \end{aligned}$$

In particular, if $X = D^i$, then

$$L_{D^i} f = \frac{\partial f}{\partial q^i},$$

so L_{D^i} is sometimes written $\partial/\partial q^i$.

<u>Definition.</u> A derivation on the ring $\mathfrak{F}$ is an $\mathbb{R}$-linear function $\theta: \mathfrak{F} \longrightarrow \mathfrak{F}$ such that

$$\theta(f \cdot g) = f \cdot \theta g + g \cdot \theta f.$$

Each L_X is a derivation on $\mathfrak{F}$; in fact, these are all the derivations on $\mathfrak{F}$.

<u>Theorem.</u> For every derivation θ on $\mathfrak{F}$, there is a unique vector field X such that $\theta = L_X$.

<u>Proof.</u> Take $a \in U$. Since translations are invertible functions which preserve all differentiable structures, we may as well assume $a = 0$. Since U is open, it contains a ball with center at the origin: for any u in that ball define a path c in U by

$$c(t) = tu.$$

The Fundamental Theorem of Calculus gives us the equation

$$\begin{aligned} fc(1) - fc(0) &= \int_0^1 \frac{dfc}{dt}\, dt \\ &= \int_0^1 \sum_{i=1}^n \frac{\partial fc}{\partial q^i}\, q^i(u)\, dt. \end{aligned}$$

Set $h_i(u) = \int_0^1 \frac{\partial fc}{\partial q^i}\, dt$ and notice that $q^i(u)$ is independent of t:

$$f(u) = f(0) + \sum_{i=1}^n h_i(u) q^i(u).$$

Then, by the defining property of θ:

$$\begin{aligned} \theta f(0) &= \sum_{i=1}^n h_i(0)\theta q^i(0) + q^i(0)\theta h_i(0) \\ &= \sum_{i=1}^n \theta q^i(0) \left.\frac{\partial f}{\partial q^i}\right|_0 , \end{aligned}$$

since $q^i(0) = 0$ and $h_i(0) = \left.\dfrac{\partial f}{\partial q^i}\right|_0$. For any $f \in \mathcal{F}$ then we have

$$\theta f = \sum_{i=1}^{n} \theta q^i \cdot \frac{\partial f}{\partial q^i} ,$$

which is exactly $L_X f$ when $X_i = \theta q^i$. The X_i uniquely determine X, so the theorem is proved.

§11. The Tensor Product

This section begins with the material in MacLane and Birkhoff Algebra, Chapter VI, §§4 and 5, and Chapter IX §§7 and 8.

A tensor is sometimes described by symbols with many indices, upper and lower. To really understand tensors, we must understand their relation to the basic vector space V under discussion. Tensors are in fact elements of new vector spaces built up out of V and its dual space by the operation of tensor product.

Given vector spaces V and W, a tensor product of V and W is a vector space, which we will write $V \otimes W$, together with a bilinear function $\otimes : V \times W \longrightarrow V \otimes W$, which have the following property: if $B : V \times W \longrightarrow U$ is any bilinear function, then there is a unique linear map $F : V \otimes W \longrightarrow U$ such that the diagram below commutes:

$$\begin{tikzcd}
 & V \otimes W \arrow[d, "F"] \\
V \times W \arrow[ur] \arrow[r, "B"'] & U
\end{tikzcd}$$

Briefly, we say that $\otimes$ is "universal" among bilinear functions on $V \times W$. If we write the image of the pair $(v, w) \in V \times W$ under the map $\otimes$ as

$v \otimes w$, then the commutativity of the diagram is expressed by the equation $B(v, w) = F(V \otimes W)$. Similarly, the bilinearity of $\otimes$ is equivalent to the equations

(i) $V \otimes (w_1 \times w_2) = v \otimes w_1 \times v \otimes w_2$, (iii) $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$,

(ii) $v \otimes kw = k(v \otimes w)$, (iv) $kv \otimes w = k(v \otimes w)$

for all $v, v_1, v_2 \in V$; $w, w_1, w_2 \in W$, and $k \in \mathbb{R}$.

The universality of $\otimes$ means that the elements $v \otimes w$ generate $V \otimes W$ as a vector space. Thus

$$V \otimes W = \left\{ \sum_{i=1}^{n} (v_i \otimes w_i) k_i \,\middle|\, v_i \in V,\ w_i \in W,\ k_i \text{ scalars and satisfying the relations (i) - (iv) above} \right\}$$

describes $V \otimes W$ in terms of elements but without bases. This decomposition may be used to prove the existence of the tensor product (Algebra, Ch. IX) What in the world is this space $V \otimes W$? Let $V = W = \mathbb{R}^3$, and let e_1, e_2, e_3 be a basis for $\mathbb{R}^3$. A bilinear function $B: \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow U$ satisfies

$$B\left(\sum_{i=1}^{3} x_i e_i , \sum_{j=1}^{3} y_j e_j\right) = \sum_{i,j=1}^{3} x_i y_i \, B(e_i, e_j),$$

so B is determined by the 3×3 matrix $(B(e_i, e_j))$. Let U be the artificial space on the basis $\{e_{ij} \mid i, j = 1, 2, 3\}$ and set $B(e_i, e_j) = e_{ij}$. Then by the universal property of tensor products we have a unique linear h such that $e_{ij} = B(e_i, e_j) = h(e_i \otimes e_j)$

$$\begin{tikzcd}
\mathbb{R}^3 \times \mathbb{R}^3 \arrow[rr, "\otimes"] \arrow[dr, "B"'] & & \mathbb{R}^3 \otimes \mathbb{R}^3 \arrow[dl, dashed, bend left, "h"] \\
& U \arrow[ur, "g"] &
\end{tikzcd}$$

On the other hand, define a linear transformation $g: U \longrightarrow \mathbb{R}^3 \otimes \mathbb{R}^3$ by $g(e_{ij}) = e_i \otimes e_j$. Then U is isomorphic to $\mathbb{R}^3 \otimes \mathbb{R}^3$, since

$$gh(e_i \otimes e_j) = e_i \otimes e_j$$

and

$$hg(e_{ij}) = e_{ij} \;.$$

This analysis works for any pair of finite dimensional vector spaces V and W but not in more general situations. If $\dim V = n$ and $\dim W = n$, it proves that $V \otimes W$ is a finite dimensional vector space of dimension mn.

In forming the tensor product of given spaces, we must say "a tensor product of V and W" instead of "the tensor product," because for all we know there may be many non-isomorphic spaces $V \otimes W$ and maps $\otimes$ enjoying the above properties. These doubts are removed by the following theorem.

Theorem. Any two tensor products of V and W are isomorphic.

Proof. We can show somewhat more. For let $V \otimes W$ and $V \square W$ be two tensor products of V and W, with associated maps $\otimes$ and $\square$ respectively: that is, $\otimes : (v, w) \rightsquigarrow v \otimes w$ and $\square : (v, w) \rightsquigarrow v \square w \in V \square W$. Then we will show that there is a unique isomorphism $\ell : V \otimes W \longrightarrow V \square W$ such that $\ell(v \otimes w) = v \square w$.

First, since $V \otimes W$ is a tensor product of V and W, we may replace U and B in the definition above by $V \square W$ and $\square$, getting a map f making the following diagram commute

$$\begin{tikzcd} V \times W \arrow[rr, "\otimes"] \arrow[dr, "\square"'] & & V \otimes W \arrow[dl, "B"] \\ & V \square W & \end{tikzcd}$$

But $\square$ is also a tensor product of V and W, so we may reverse the roles of $\otimes$ and $\square$ to get the map g in the following commutative diagram

$$\begin{tikzcd} V \times W \arrow[r, "\otimes"] \arrow[dr, "\square"'] & V \otimes W \arrow[d, shift right, "f"'] \\ & V \square W \arrow[u, shift right, "g"'] \end{tikzcd}$$

Thus commutativity means that for all $(v, w) \in V \times W$, we have

$$f(v \otimes w) = v \square w, \qquad g(v \square w) = v \otimes w .$$

In particular, $gf(v \otimes w) = v \otimes w$, $fg(v \square w) = v \square w$. Now use again the fact that $V \otimes W$ is a tensor product, replacing U and B in the definition this time by $V \otimes W$ and $\otimes$. We have just shown that gf and $1_{V \otimes W}$ both make the diagram commute; hence, by theuniqueness assertion in the definition, $1_{V \otimes W} = gf$. Similarly, we show that $fg = 1_{V \square W}$. This means exactly that f is an isomorphism of $V \otimes W \longrightarrow V \square W$ with inverse g.

$$\begin{tikzcd} V \times W \arrow[r, "\otimes"] \arrow[dr, "\otimes"'] & V \otimes W \arrow[d, shift right, bend right, "gf"'] \arrow[d, shift left, bend left, "1_{V \otimes W}"] \\ & V \otimes W \end{tikzcd}$$

We have shown that if tensor products exist, they are unique, "up to isomorphism". But there might not be any tensor products of V and W at all.

<u>Theorem</u> If V and W are vector spaces then they have a tensor product.

Proof. Even though V and W may have infinite dimension, we can still find bases $\{e_i\}$ for V, $i \in I$, and $\{d_j\}$ for W, $j \in J$. The basis elements are indexed not by integers but by members of the (possibly uncountable) sets I and J; that is, to each element i of I there exists a basis element e_i of V. Then every element v of V will be uniquely expressible as a linear combination of some finite subset of the $\{e_i\}$.

We are likewise free to form a new set of symbols b_{ij}, one for each element (i,j) of the cartesian product $I \times J$. The set of all possible finite expressions $\sum_{k=1}^{m} r_k b_{(k)}$, where each $b_{(k)}$ is a b_{ij}-symbol for some (i,j), and each $r_k \in \mathbb{R}$, forms a perfectly good abstract vector space when we define addition and scalar multiplication in the obvious way. We claim that this new space L is a tensor product of V and W.

Define $\Box : V \times W \longrightarrow L$ by $(\sum x^i e_i, \sum y^j d_j) \rightsquigarrow \sum x^i y^i b_{ij}$. Clearly, $\Box$ is bilinear. If now $B: V \times W \longrightarrow U$ is bilinear, we have $B(\sum x^i e_i, \sum y^i d_j) = \sum_i x^i B(e_i, \sum y^j d_j) = \sum_{i,j} x^i y^j B(e_i, d_j)$. Hence if we define $f: L \longrightarrow U$ on the basis elements $\{b_{ij}\}$ of L by the formula $f(b_{ij}) = B(e_i, d_j)$, computation shows that the tensor product diagram commutes, and that this is the only f which will make the diagram commutative. Hence L is a tensor product of V and W.

Denoting the elements b_{ij} by the symbols $e_i \otimes d_j$, we derive as a corollary that if $\{e_i\}$ is a basis of the vector space V and $\{d_j\}$ is a basis of W, then $\{e_i \otimes d_j\}$ is a basis of $V \otimes W$. Notice that not every vector of $V \otimes W$ is of the form $v \otimes w$ for some v in V and w in W. There will usually be sums $\sum_i v_i \otimes w_i$ which cannot be reduced to single terms of the form $v \otimes w$; e.g. $e_1 \otimes d_1 + e_2 \otimes d_2$.

We now define the space of 2-tensors on V as $T_2(V) = V \otimes V$. Since $V \otimes V$ has basis $\{e_i \otimes e_j\}$, we can write any element t of $T_2(V)$ as $t = \sum_{i,j} x^{ij} e_i \otimes e_j$, where the x^{ij} are real numbers. The traditional viewpoint is that the tensor is the array x^{ij}. Of course, the matrix elements depend on which basis of V we pick; we can derive the rule for transforming to the new basis e'_i, where $e^i = \sum a^j_i e'_j$, as follows

$$e_i \otimes e_k = \sum a^j_i e'_j \otimes \sum a^\ell_k e'_\ell = \sum_{j,k} a^j_i a^\ell_k e'_j \otimes e'_\ell$$

Therefore,

$$t = \sum x^{ik} e_i \otimes e_k = \sum_{i,j,k,\ell} x^{ik} a^j_i a^\ell_k e'_j \otimes e'_\ell$$

So

$$x'^{j\ell} = \sum_{i,k} x^{ik} a^j_i a^\ell_k \qquad (1)$$

Thus we have replaced the usual opaque definition that a tensor is an array of coordinates relative to a basis which transforms according to equation (1).

Our new definition helps us see how tensors behave under linear transformations. In the most general case, we have linear maps $g: V \longrightarrow V'$ and $h: W \longrightarrow W'$ giving us a map $g \times h. V \times W \longrightarrow V' \times W'$ defined by sending (v, w) to (gv, hw). The composite $\otimes(g \times h)$ is a bilinear map on $V \times W$, hence by the definition of tensor product, factors uniquely through

$V \otimes W$ as below, giving us a new map $g \otimes h: V \otimes W \longrightarrow V' \otimes W'$. Thus $g(v \otimes w) = gv \otimes hw$. It is also easy to see that if $g': V' \longrightarrow V''$ and $h': W' \longrightarrow W''$, then $g'g \otimes h'h = (g' \otimes h')(g \otimes h)$ (use the uniqueness property in the definition of tensor product).

$$\begin{tikzcd}
V \times W \arrow[rr, "\otimes"] \arrow[dr, "g \times h"'] & & V \otimes W \arrow[dd, dashed, "g \otimes h"] \\
& V' \times W' \arrow[dr, "\otimes"'] & \\
& & V' \otimes W'
\end{tikzcd}$$

A historical note: we have defined elements of the tangent space to be "contravariant" vectors. Actually, modern usage considers tangent vectors to be covariant in nature, since given $\varphi: V \longrightarrow U$, the induced map φ_* on the tangent spaces maps $T.(V) \longrightarrow T.(U)$; if the induced map reversed the arrow, taking $T.(U) \longrightarrow T.(V)$, it would be not "co" but "contra." The reason for the traditional terminology (which we will stick to) is that the coordinate transformations under change of basis, given by (1), do interchange the position of primed and unprimed letters.

Similar arguments to the ones we have been using prove that either $V \otimes (W \otimes U)$ or $(V \otimes W) \otimes U$ is a universal object for trilinear maps from $V \times W \times U$. (cf. MacLane and Birkhoff, Ch.16). Thus we can unambiguously define the tensor product $V \otimes W \otimes U$ to be $V \otimes (W \otimes U)$ and similarly the tensor product of any finite number of vector spaces. The elements of $T_n(V) = V \otimes V \otimes \ldots \otimes V$ (n times) are called n-contravariant tensors, or contravariant tensors of rank n.

<u>Theorem.</u> Let V and W be finite-dimensional vector spaces, $V \otimes W$ a tensor product, and $V^* \square W^*$ a tensor product of the dual spaces V^* and W^*. Then there is an isomorphism $V^* \square W^* \xrightarrow{\cong} (V \otimes W)^*$.

<u>Proof.</u> Recall that V^* was the set of linear maps $f: V \to \mathbb{R}$. If now $(f, g) \in V^* \times W^*$, $f \otimes g: V \otimes W \to \mathbb{R} \otimes \mathbb{R} \cong \mathbb{R}$; therefore $f \otimes g \in (V \otimes W)^*$. (cf. the corollary above). Furthermore, the map $(f, g) \rightsquigarrow f \otimes g \in (V \otimes W)^*$ is bilinear. By the universality of $\square$, there is a map $h: V^* \square W^* \longrightarrow (V \otimes W)^*$ such that $h(f \square g) = f \otimes g$. We wish to show that h is an isomorphism. It suffices to prove that h maps the basis $\{e^i \square d^j\}$ of $V^* \square W^*$ to a basis of $(V \otimes W)^*$. Thus it will be enough to show that $\{e^i \otimes d^j\}$ is the dual basis of $\{e_i \otimes d_j\}$. But by the definition of the tensor product of two maps,

$$(e^i \otimes d^j)(e_k \otimes d_\ell) = (e^i e_k \otimes d^i d_\ell) = (e^i e_k)(d^j d_\ell) \qquad \text{since } k \otimes \ell = k\ell \text{ for } k, \ell \in$$

$$= \delta^i_k \delta^j_\ell = \begin{cases} 1 & \text{if } i = k, j = \ell \\ 0 & \text{otherwise} \end{cases} .$$

Thus $e^i \otimes d^j$ is the dual basis. We can now state the conclusion of the theorem as $(V \otimes W)^* = V^* \otimes W^*$; that is, we identity $(V \otimes W)^*$ with $V^* \otimes W^*$ by the isomorphism here established.

§12. <u>Tensor Algebras and Graded Algebras.</u>

Let us now consider all the vector spaces $T_n(V) = V \otimes \ldots \otimes V$ (n times), for $n = 1, 2, \ldots$. For $n = 0$ we will define $T_0(V) = \mathbb{R}$. Suppose that $a = v_1 \otimes v_2 \otimes \ldots \otimes v_r$ is in $T_r(V)$ and that $b = w_1 \otimes w_2 \ldots \otimes w_s$

is in $T_s(V)$. Since all of the vectors v_i and w_j belong to V, we can form $v_1 \otimes \dots \otimes v_r \otimes w_1 \otimes \dots \otimes w_s \in T_{r+s}(V)$, which we will call the product of a and b. Thus if we refer to the whole collection of spaces $\{T_n(V)\}$ as $T_*(V)$, extending the above multiplication by the distributive law gives a natural multiplicative structure on this set, which has the unusual property that the product of an element of $T_r(V)$ with one of $T_s(V)$ lies in $T_{r+s}(V)$, so that the product is usually in a space different from those of the factors.

We generalize this situation as follows. An algebra A is a vector space which in addition possesses a multiplication; that is, not only can we multiply elements of the space by scalars, but any two elements of the space itself have a product. This multiplication is required to be bilinear and associative, and to have a unit element. Thus, we have the rules

$$(k_1a_1 + k_2a_2)b = k_1(a_1b) + k_2(a_2b) \ , \ b(k_1a_1 + k_2a_2) = k_1(ba_1) + k_2(ba_2)$$

$$a(bc) = (ab)c \ , \ 1a = a1 = a$$

$$(ka)b = k(ab) \ , \ a(kb) = k(ab)$$

for $a, b, a_1, a_2 \in A$ and k, k_1, and k_2 scalars. A graded algebra is a string G of vector spaces, $G_0, G_1, \dots$, with an additional product structure such that if $a \in G_n$ and $b \in G_m$, then $ab \in G_{n+m}$. For each m and n, this product must be a bilinear function from $G_n \times G_m$ to G_{n+m}, must be associative, and must have a unit element $1 \in G_0$. Notice in particular that a graded algebra G is not an algebra, since the sum of any two elements is defined only if they both have the same degree ; that is, if they both lie in the same G_i.* Now we can see that $T_*(V)$ is a graded algebra, with unit element $1 \in T_0(V) = \mathbb{R}$. Call $T_*(V)$ the tensor algebra of V. There are other examples of graded algebras: consider the set G_n of all homogeneous polynomials of degree n in the two letters X and Y. A typical element of G_n is $\sum_{i=0}^{n} a_i X^i Y^{n-i}$, where $a_i \in \mathbb{R}$. G_n is a vector space under the usual addition and scalar multiplication of polynomials, and since the product of a homogeneous polynomial of degree n and one of degree m is a homogeneous polynomial of degree n+m, it is easy to see that the set of all polynomials in S and Y contains the graded algebra G.

§13. Exterior Algebra.

An exterior algebra E is a graded algebra with the property that the square of any element of degree one is zero. Now if a and b are of degree one, so is a+b, hence $0 = (a+b)^2 = 0 + ab + ba + 0$. Thus in any exterior algebra, $ab = -ba$ for any two elements of degree one. What does the rest of an exterior algebra look like? We can get some idea of the answer by considering the simple example of an exterior algebra Λ for which Λ_1 is a two-dimensional vector space. If $\{e^1, e^2\}$ is a basis for Λ_1, then $e^1 e^2 = -e^2 e^1$ will lie in Λ_2. Thus Λ_2 is forced to be at least one-dimensional. (Note: We assume that no other relations besides

---

*If $a \in G_i$ then we say a has degree i.

($a^2 = 0$ for $\deg a = 1$) are satisfied; in particular, $e^1e^2 \neq 0$.) It's also easy to see that any product of elements of Λ_1 which lies in Λ_n (for $n > 2$) must be zero. So we are tempted to form Λ by letting Λ_n be the zero vector space for $n > 2$, with Λ_2 a one-dimensional vector space whose basis is $\{e^{12}\}$, where by definition $e^{12} = e^1e^2$. This does give us an exterior algebra: for example, if $x_1e^1 + x_2e^2$ is any element of Λ_1, the formula

$$(x_1e^1 + x_2e^2)(y_1e^1 + y_2e^2) = (x_1y_2 - x_2y_1)e^1e^2$$

tells us that $(x_1e^1 + x_2e^2)^2 = 0$.

Move now to the case where Λ_1 is three-dimensional, with basis $\{e^1, e^2, e^3\}$. Then e^1e^2, e^2e^3, e^1e^3 all must lie in Λ_2; denote them by e^{12}, e^{23}, e^{13} respectively. Then we can let Λ_2 be three-dimensional, with basis $\{e^{12}, e^{23}, e^{13}\}$ with Λ_3 a one-dimensional space generated by $e^{123} = e^1e^2e^3$, and $\Lambda_n = 0$ for $n > 3$. Multiplication is now always possible since for example $e^2e^{12} = e^2e^1e^2 = -e^2e^2e^1 = 0$ and $e^3e^2e^1 = -e^3e^1e^2 = e^1e^3e^2 = -e^1e^2e^3 = -e^{123}$. Calculation gives the rule

$$\left(\sum_{i=1}^{3} x_ie^i\right)\wedge\left(\sum_{j=1}^{3} x_je^j\right)\wedge\left(\sum_{k=1}^{3} z_ke^k\right) = \det\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} e^{123}$$

where we have used the traditional "wedge" symbol to denote the anti-symmetric multiplication of an exterior algebra.

It is now apparent how to extend the above construction to the n-dimensional case.

<u>Theorem</u>. Given a vector space W of dimension m, there exists a (free) exterior algebra $\Lambda = \Lambda(W)$ with $\Lambda_1 = W$.

<u>Proof</u>. Let $\{e^1, \ldots, e^m\}$ be a basis of W. Let $\Lambda_k(W)$ be a vector space whose basis is the set of symbols $\{e^{i_1 i_2 \ldots i_k}\}$, where $1 \leq i_1 < i_2 < \ldots < i_k \leq m$. Thus $\Lambda_k(W)$ has dimension $\binom{m}{k} = \frac{m!}{k!(m-k)!}$. To write out the rules of multiplication in a simple form, we will use the language of permutations, considering a permutation σ of order n as a one-to-one function from the set of the first n integers to that set. The sign of σ, sometimes written $(-1)^\sigma$, is defined to be $+1$ if σ is the product of an even number of transpositions, and -1 otherwise. Then if $i_1 < \ldots < i_n$, while σ is a permutation, we define $e^{i_{\sigma(1)} i_{\sigma(2)} \ldots i_{\sigma(k)}} = (-1)^\sigma e^{i_1 \ldots i_k}$; it is not hard to see that this agrees with our two- and three-dimensional examples. Now define exterior multiplication by

$$e^{i_1 \ldots i_k} \wedge e^{j_1 \ldots j_\ell} = \begin{cases} 0 & \text{if some } i_s \text{ equals some } j_t \\ e^{i_1 i_2 \ldots i_k j_1 \ldots j_\ell} & \text{otherwise} \end{cases}$$

Finally, let $1 \cdot e^{j_1 j_2 \ldots j_r} = e^{j_1 j_2 \ldots j_r}$, and extend this multiplication linearly to all of $\Lambda(W)$. Now all the properties in the definition of exterior algebra follow: for example

$$(e^{i_1 \ldots i_r} \wedge e^{j_1 \ldots j_s}) \wedge e^{k_1 \ldots k_t} = e^{i_1 \ldots i_r j_1 \ldots j_s k_1 \ldots k_t}$$
$$= e^{i_1 \ldots i_r} \wedge (e^{j_1 \ldots j_s} \wedge e^{k_1 \ldots k_t})$$

is the associative law. Moreover, if $a_i = \sum_{i=1}^{r} c_{ij} e^j$, then the formula

$$a_1 \wedge \cdots \wedge a_r = \det(\cancel{a}_{ij}^{\,c})\cancel{e}^{\,\overline{i_1 \cdots i_r}} .$$

$\det(c_{ij})\, e^{1 \cdots r}$

holds. If $a = \sum x_i e^i$ is any element of degree one,

$$a^2 = \sum x_i x_j e^i \wedge e^j = \sum_{i=j} + \sum_{i<j} + \sum_{i>j} = 0 + \sum_{i<j} x_i x_j e^i \wedge e^j + \sum_{i>j} x_i x_j e^j \wedge e^i = 0,$$

so Λ is an exterior algebra.

This construction is in fact the most general way of obtaining an n-dimensional free exterior algebra; it does not depend on the choice of basis. To prove this, we note first that if E is any exterior algebra, and f is a linear transformation mapping W to E_1, then there are unique linear transformations $f_i : \Lambda_i W \longrightarrow E_i$ such that the collection $\{f_i\}$ forms a morphism of graded algebras; that is, $f_0(1) = 1$, and $f_i(a)f_j(b) = f_{i+j}(ab)$ if a and b are of degree i and j respectively. This is true since we must have $f_k(e^{i_1 \cdots i_k}) = f_k(e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_k}) = f(e^{i_1}) \wedge f(e^{i_2}) \wedge \cdots \wedge f(e^{i_k})$; since f_k is now defined on a basis, it extends uniquely to all of Λ_k.

Our proof that the construction of an exterior algebra on V does not depend on the choice of basis of V will be based on the following very general principles, which we have been using implicitly for some time. Let us define a <u>functor</u> on algebras to vector spaces to be a function $\mathcal{F}$ which to every algebra A assigns a vector space $\mathcal{F}(A)$, and to every map $A \xrightarrow{f} A'$ of algebras assigns a map $\mathcal{F}(f): \mathcal{F}(A) \longrightarrow \mathcal{F}(A')$ of vector spaces (a linear transformation) with the rules that $\mathcal{F}(1_A) = 1_{\mathcal{F}(A)}$, where 1_A is the identity map on A, and $\mathcal{F}(gf) = \mathcal{F}(g)\,\mathcal{F}(f)$ if $A \xrightarrow{f} A' \xrightarrow{g} A''$.

Given $\mathfrak{U}$ and a vector space V, a <u>universal construction</u> for $\mathfrak{U}$ and V is an <u>algebra</u> R and an arrow $u: V \longrightarrow \mathfrak{U}(R)$ (that is, a linear transformation) with the following property: given any arrow $t: V \longrightarrow \mathfrak{U}(A)$, there is a unique map of algebras $f: R \longrightarrow A$ such that $t = \mathfrak{U}(f)u$. That is, $\mathfrak{U}(f)$ makes the following diagram commute and f is the only map which has this property.

$$\begin{tikzcd} V \arrow[r, "u"] \arrow[dr, "t"'] & \mathfrak{U}(R) \arrow[d, "\mathfrak{U}(f)"] \\ & \mathfrak{U}(A) \end{tikzcd} .$$

We saw this situation in the construction of the tensor product (§11 above). In that case we showed that any two universal arrows, i. e. , any two tensor products, were isomorphic. Exactly the same proof works in the general situation:

<u>Theorem.</u> If (R, u) and (R', u') are both universal arrows for $\mathfrak{U}$ and V, then there is a unique isomorphism $h: R \longrightarrow R'$ such that $\mathfrak{U}(h)u = u'$.

To prove this, simply repeat the proof of the above-mentioned theorem about tensor products; since nothing in the argument there depended on whether the objects in question were algebras or vector spaces, this <u>universal property</u> is all we need in the proof.

Now consider the functor $\mathfrak{G}$ from graded algebras to vector spaces, defined by $\mathfrak{G}(G) = G_1$.

Theorem 1. The tensor algebra T_*V is a universal construction for V and $\mathcal{B}$; that is, given a graded algebra G and a map $V \xrightarrow{t} G_1$, there is a unique map $T_*(V) \xrightarrow{r} G$ of algebras making the diagram below commute.

$$\begin{tikzcd} V \arrow[r, "1_V"] \arrow[dr] & \mathcal{B}(T_*(V)) = V \arrow[d, "\mathcal{B}(r) = r_1"] \\ & \mathcal{B}(G) = G_1 \end{tikzcd}$$

Proof. It is easy to see that we are forced to take $r_0 = 1_{\mathbb{R}}$, $r_1 = t$; and to define $r_n(v_1 \otimes \dots \otimes v_n) = (tv_1)(tv_2)\dots(tv_n)$.

Theorem 2. Define $\mathcal{B}$ as before, but mapping graded exterior algebras to vector spaces. Then if W is a finite-dimensional vector space, $\Lambda(W)$ is a universal object for W and $\mathcal{B}$.

Proof. In the same situation as that of Theorem 1, defining $r_n(a_1 \wedge \dots \wedge a_n) = (ta_1) \wedge (ta_2) \wedge \dots \wedge (ta_n)$ does the trick.

Now we can prove our assertion about the construction of the exterior algebra on a given finite dimensional vector space V. Theorem 2 shows that the algebra $\Lambda(W)$ constructed from a particular basis $e_1, \dots, e_n$ is a universal object; likewise, the exterior algebra coming from a different basis $d_1, \dots, d_n$ is universal. Since any two universal objects for a functor are isomorphic, the two exterior algebras we have constructed are really the same. Incidentally, it is possible to define $\Lambda(W)$ directly as an invariant object; for details, see the last chapter of MacLane and Birkhoff.

§14. Alternating Tensors

We now go on to derive a new and very useful way of looking at $\Lambda(W)$ as a special subspace of $T_*(W)$. In §1~~2~~3 we studied elements of S_k, the set of permutations of k letters. Now each $\sigma \in S_k$ can be interpreted as a linear transformation on $T_k(W)$: define $\sigma(w_1 \otimes \ldots \otimes w_k) = w_{\sigma(1)} \otimes \ldots \otimes w_{\sigma(k)}$ and check that this can be extended to a well-defined linear map (this amounts to checking that $\overline{\sigma}$ on $W \times \ldots \times W$ defined by $\overline{\sigma}(w_1 \times \ldots \times w_k) = w_{\sigma(1)} \otimes \ldots \otimes w_{\sigma(k)}$ is a k-linear map).

Definition. A tensor $t \in T_k(W)$ is called alternating if $\sigma(t) = (-1)^{\sigma} t$ for all $\sigma \in S_k$.

Theorem. The set of alternating tensors in $T_*(W)$ can, in a natural way, be made into a graded exterior algebra isomorphic to $\Lambda(W)$.

Proof. The set of alternating k-tensors clearly forms a subspace of $T_k(W)$; furthermore, any tensor t can be symmetrized to yield an alternating tensor. Specifically, let $\overline{A}(t) = \sum_{\sigma \in S_k} (-1)^{\sigma} \sigma t$ if t is any k-tensor, not necessarily alternating. Then

$$\tau \overline{A}(t) = \sum_{\sigma \in S_k} (-1)^{\sigma} \tau\sigma(t) = (-1)^{\tau}(-1)^{\tau} \sum_{\sigma \in S^k} (-1)^{\sigma} \tau\sigma(t) = (-1)^{\tau} \sum_{\sigma \in S^k} (-1)^{\sigma\tau} \tau\sigma(t),$$

since in general $(-1)^{\sigma}(-1)^{\tau} = (-1)^{\sigma\tau}$. But since S_k is a group, as σ runs over all the elements of S_k, $\sigma\tau$ also runs over all the elements of S_k (perhaps in a different order, but each element is counted once and only once). Hence $\tau \overline{A}(t) = (-1)^{\tau} \overline{A}(t)$, for $t \in T_k(W)$, so $\overline{A}(t)$ is alternating, whether or not t is. If t does happen to be alternating, $\overline{A}(t) = (k!)t$, since there are $k!$ elements in S_k; so the mapping A defined by $A(t) = (1/k!)\overline{A}(t)$ has the property that At is always alternating, and that if t is already alternating, then $At = t$.

Now if s and t are any two tensors in $T_*(W)$, define $s \wedge t$ to be $A(s \otimes t)$, where $(s,t) \rightsquigarrow s \otimes t$ is just the usual product in the tensor algebra $T_*(W)$. We assert that the alternating tensors form a graded exterior algebra under $\wedge$. First, given t of degree one, we have $t \wedge t = 0$, since S_2 consists of only two permutations, one of each sign, and hence $A(t \times t) = t \otimes t - t \otimes t = 0$. Next, we check the associative law. The proof of this is divided into steps.

1. $A(\tau t) = (-1)^{\tau} A(t)$ for $t \in T_k(W)$, $\tau \in S_k$.

This follows easily from the definition of A.

2. Given $\sigma \in S_k$, define $\tilde{\sigma} \in S_{m+k}$ by

$$\tilde{\sigma}(i) = i \quad \text{for } 1 \leq \pm m$$

$$\tilde{\sigma}(i+m) = \sigma(i) + m \quad \text{for } m+1 \leq m+i \leq m+k .$$

Then $(-1)^{\sigma} = (-1)^{\tilde{\sigma}}$.

3. If s is an m-tensor and t is a k-tensor, then $s \wedge t = s \wedge (At)$.

<u>Proof.</u>

$$s \wedge At = A(s \otimes At) = \frac{1}{k!} A\Big(\sum_{\sigma \in S_k} s \otimes (-1)^{\sigma} \sigma t\Big) = \frac{1}{k!} A\Big(\sum_{\sigma \in S_k} (-1)^{\sigma} \tilde{\sigma}(s \otimes t)\Big)$$

$$= \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{\sigma} A(\tilde{\sigma}(s \otimes t)) = (\text{by } 1) \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{\sigma} (-1)^{\tilde{\sigma}} A(s \otimes t)$$

$$= A\Big(\frac{1}{k!} \sum_{\sigma} s \otimes t\Big) = A(s \otimes t) = s \wedge t.$$

4. $r \wedge (s \wedge t) = r \wedge A(s \otimes t) = r \wedge (s \otimes t)$ by 3

$$= A(r \otimes (s \otimes t)) = A((r \otimes s) \otimes t) = (r \otimes s) \wedge t$$

$$= A(r \otimes s) \wedge t = (r \wedge s) \wedge t.$$

Similarly, we check that the alternating tensors have the other properties of an exterior algebra.

To finish the proof we must establish theisomorphism between the set of alternating tensors and $T_*(W)$. Let $e_1, \ldots, e_n$ be a basis of W; then a linear map on $\Lambda(W)$ is determined by its values on all basis elements $e_{i_1} \wedge \ldots \wedge e_{i_k}$. Then mapping $e_{i_1} \wedge \ldots \wedge e_{i_k}$ to the alternating tensor $A(e_{i_1} \otimes \ldots \otimes e_{i_k})$ is an isomorphism: it is one-to-one since

$$A(\sum_i a_i (e_{i_1} \otimes \ldots \otimes e_{i_k}) = \sum_i a_i (\tfrac{1}{k!}) \sum_{\sigma \in S_k} (-1)^\sigma e_{\sigma(i_1)} \otimes \ldots \otimes e_{\sigma(i_k)}$$

is a sum of distinct basis elements of $T_k(W)$ with non-zero coefficients, hence non-zero; it is onto since given any alternating tensor $t \in T_k(W)$, write $t = \sum_i a_i e_{i_1} \otimes \ldots \otimes e_{i_k}$; then the element $\sum a_i e_{i_1} \wedge \ldots \wedge e_{i_k}$ of $\Lambda(W)$ is mapped to $A(\sum a_i e_{i_1} \otimes \ldots \otimes e_{i_k}) = At = t$. This completes the proof.

§15. Local Manifolds: Invariant Description.

Eventually we will describe mechanical systems (Lagrange's equations, Hamiltonians, etc.), on smooth manifolds obtained by piecing together open sets of Euclidean spaces $\mathbb{R}^n$. Hitherto, all our treatment has been local, so has been formulated for open sets U in $\mathbb{R}^n$. Each such set $U \subset \mathbb{R}^n$ comes equipped with a natural set of coordinates -- those of $\mathbb{R}^n$. However, the basic constructions such as the tangent bundle $T.U$ or the Lagrange equations for a smooth function $L: T.U \longrightarrow \mathbb{R}$ have in fact been independent of the choice of coordinates. Indeed, all of our previous discussion of such open sets U can be made more clearly invariant if U is replaced by a "local manifold" M in the sense of the following definition. The sense is this: Given coordinates q^i on U, a function $f: U \longrightarrow \mathbb{R}$ is smooth if all the higher partial derivatives $\frac{\partial f}{\partial q^i}$ are continuous. The set $\mathcal{F}$ of all smooth functions is then the same for any allowable choice of coordinates; hence we can give an invariant description in terms of $\mathcal{F}$.

Definition. A local manifold is a set M together with a set $\mathcal{F}$ of functions $f: M \longrightarrow \mathbb{R}$ such that

1) there exist $q^1, \ldots, q^n \in \mathcal{F}$ such that the map

$$m \overset{\varphi}{\rightsquigarrow} (q^1 m, \ldots, q^n m)$$
$$M \xrightarrow{\varphi} \mathbb{R}^n$$

is one-to-one onto an open set $U \subset \mathbb{R}^n$,

2) $f \in \mathcal{F}$ if and only if $f\varphi^{-1}$ is smooth on U.

We leave the reader to carry out the replacement.

§16. The Exterior Bundle

We have seen that given any finite-dimensional vector space W, there exists (uniquely up to isomorphism) the exterior algebra $\Lambda = \Lambda(W)$ of W, i.e., a graded algebra universal for the properties

1. W = elements of degree 1,

2. $w \wedge w = 0$.

We have also seen that

$$\Lambda_k(W) \simeq \text{alternating tensors on } W$$
$$= \{t \in T_k(W) \mid \sigma t = (-1)^\sigma t \quad \text{for all } \sigma \in S_k\}.$$

As we have done in the case of a two-fold tensor product, we can show that

$$T_k(W) \simeq \mathrm{Mult}\,(\overbrace{W^*, \ldots, W^*}^{\text{k-factors}} : \mathbb{R})$$
$$= \text{all multilinear maps } f\colon \overbrace{W^* \times \ldots \times W^*}^{\text{k factors}} \longrightarrow \mathbb{R}.$$

Then an alternating tensor corresponds to an alternating multilinear map

$$a\colon \underbrace{W^* \times \ldots \times W^*}_{\text{k factors}} \longrightarrow \mathbb{R}$$

such that

$$a(v_1, \ldots, v_k) = (-1)^\sigma a(v_{\sigma(1)}, \ldots, v_{\sigma(k)}), \quad \text{for all } \sigma \in X_k,\ v_i \in W^*.$$

Suppose U is an open set in $\mathbb{R}^n$. (or better, a local manifold as defined above). We define the $\underline{k^{\text{th}} \text{ exterior bundle over } U}$ to be

$$\begin{array}{ccc} \Lambda^k U = \Lambda_k(T^\cdot U) & = & \{(a, d) \mid a \in U,\ d \in \Lambda_k(T^a U)\}, \\ \downarrow & & \downarrow \\ U & & a \in U \end{array}$$

i.e., the k^{th} exterior bundle is the bundle whose fibers are the k^{th} exterior algebra of the cotangent spaces of U. If $k = 0$, each $\Lambda_0(T^aU)$ is just $\mathbb{R}$, so $\Lambda^0(U) = U \times \mathbb{R}$ (all fibers isomorphic to the 1-dimensional vector space $\mathbb{R}$.

Λ^kU is a local manifold -- for if U has coordinates $q^1, \ldots, q^n$ and T^aU has basis $e^1, \ldots, e^n$, then Λ^kU has coordinates $q^1, \ldots, q^n$ and $p_{i_1 \ldots i_k}$, where the $p_{i_1 \ldots i_k}$ are the dual basis to the basis $\{e^{i_1} \wedge \ldots \wedge e^{i_k}\}$ for $\Lambda_k(T^aU)$; we can define a function on Λ^kU to be smooth if it is smooth in terms of these coordinates.

For any integers k and m, we may form the "pullback" of Λ^kU and Λ^mU over U

$$\Lambda^kU \underset{U}{\times} \Lambda^mU = \{(x, y) \in \Lambda^kU \times \Lambda^mU \mid \pi(x) = \pi(y)\}$$

$$\begin{tikzcd} & \Lambda^kU \underset{U}{\times} \Lambda^mU \arrow[dl, "\rho_k"'] \arrow[dr, "\rho_m"] & \\ \Lambda^kU \arrow[dr, "\pi"'] & \text{pullback} & \Lambda^mU \arrow[dl, "\pi"] \\ & U & \end{tikzcd}$$

This pullback is itself a bundle over U. Now we can define a multiplication on $\Lambda^kU \underset{U}{\times} \Lambda^mU$ to $\Lambda^{k+m}U$ by means of the multiplication in $\Lambda(T^aU)$ for each $a \in U$:

$$\Lambda^kU \underset{U}{\times} \Lambda^mU \xrightarrow{\wedge} \Lambda^{k+m}U$$

$$((a, d), (a, d')) \rightsquigarrow^{\wedge} (a, d \wedge d') \in \Lambda^{k+m}U .$$