C. How to get Riemannian or pseudo-Riemannian metrics.

(1) On one coordinate patch or on a Euclidean space, we have a system of vector fields $\partial_1, \ldots, \partial_n$ valid everywhere. It suffices to define the functions $g_{ij}(x)$. We could, for example, set $g_{ij}(x) = \delta_{ij}$, making the vector fields $\{\partial_i\}$ orthonormal at each point.

(2) On a manifold, we can use (1) to construct metrics on coordinate patches. Then we can use "partitions of unity" to combine these metrics into a metric on the whole manifold. This is the method usually used to show that any manifold has a metric.

(3) The usual way in which Riemannian metrics arise in practice is as follows: Suppose N is a space (often a Euclidean space) which already has a Riemannian metric h. Suppose we have a manifold M and a C^∞ function $f: M \to N$ which is an immersion (that is, the Jacobian matrix of f is non-singular at every point of M). Then we define a metric $f^*(h) = g$ on M by letting $g_x(X, Y) = h_{f(x)}(f_*(X), f_*(Y))$ for $X, Y \in T_\cdot M$

In local coordinates, this goes as follows: Let $x_1, \ldots, x_m$ be local coordinates on M, and $y_1, \ldots, y_n$ be local coordinates on N. Then f is given by $y_i = f_i(x_1, \ldots, x_m)$ $(i = 1, \ldots, n)$. Let h be given by $h_{ij}(y)$. Then $g = f^*(h)$ is given by

$$\begin{aligned} g_{ij}(x) &= \sum_{k,\ell=1}^{n} h_{k,\ell}(f(x)) \frac{\partial f_k}{\partial x_i} \cdot \frac{\partial f_\ell}{\partial x_j} \\ &= \sum_{k,\ell=1}^{n} h_{k,\ell}(f(x)) \frac{\partial y_k}{\partial x_i} \cdot \frac{\partial y_\ell}{\partial x_j} . \end{aligned}$$

$f^*(h)$ is called the <u>pullback of h by f.</u>

II (§49). Covariant Differentiation

A. Motivation. We want some sort of directional derivative on a manifold. Will L_X do? No! Why?

(1) A directional derivative in the direction X should depend only on the value of X at the point in question. But $L_{f\cdot X}Y = f\cdot L_X Y - (Y\cdot f)X$, showing that $L_X Y$ depends on how X is changing at the point in question (note the term $(Y\cdot f)X$).

(2) We will be interested in Newton's laws, and therefore in acceleration as we move along a curve; i. e., the derivative of the velocity vector in the direction of the velocity vector. But the Lie derivative $L_X X$ is always zero. Thus we cannot use L_X to discuss acceleration.

B. The abstract covariant derivative.

(1) Definition: An (affine) connection ∇ on M is a rule which assigns to two smooth vector fields X and Y on M another smooth vector field $\nabla_X Y$ on M, called the covariant derivative of Y in the direction X (with respect to ∇), obeying

(a) $\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y$ and $\nabla_{fX} Y = f\cdot\nabla_X Y$,

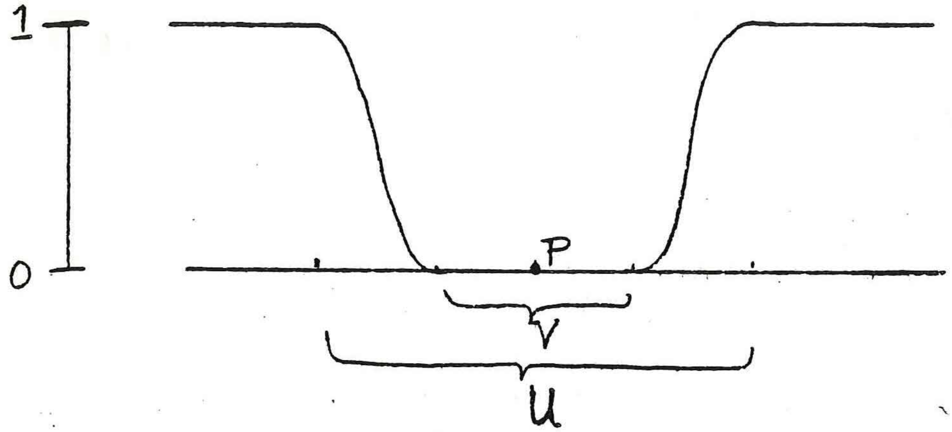
(b) $\nabla_X(Y_1+Y_2) = \nabla_X Y_1 + \nabla_X Y_2$ and $\nabla_X(fY) = (X\cdot f)Y + f\cdot\nabla_X Y$

for $f \in \mathcal{F}(M)$, X_i, Y_i vector fields on M.

(2) If $\nabla_X Y$ is to fulfill our expectations of what a directional derivative ought to be, then the following proposition should hold:

<u>Proposition.</u> For any point p on M, $(\nabla_X Y)_p$ depends only on X_p and on the behavior of Y in a neighborhood of p (actually on the "germ" of Y at p).

<u>Proof.</u> If $Y = Y'$ in a neighborhood U of p, then we take a "bump" function f which is one outside of U and zero on some neighborhood $V \subseteq U$ of p.



Then $Y - Y' = f\cdot(Y - Y')$, so that

$$\begin{aligned}(\nabla_X(Y - Y'))_p &= [\nabla_X(f\cdot(Y - Y'))]_p \\ &= (X\cdot f)_p(Y - Y')_p + f(p)\nabla_X(Y - Y')_p \\ &= 0\cdot(Y - Y')_p + 0\cdot\nabla_X(Y - Y')_p = 0,\end{aligned}$$

so $(\nabla_X Y)_p = (\nabla_X Y')_p$.

If $X_p = X'_p$, then we can write $X - X' = \sum f_i P_i$, where $f \in \mathcal{F}(M)$ and P_i are vector fields on M, with $f_i(p) = 0$. (Details left to the reader.) Then

$$(\nabla_X Y)_p - (\nabla_{X'} Y)_p = (\nabla_{X-X'} Y)_p = (\nabla_{\sum f_i P_i} Y)_p$$

$$= \sum f_i(p)(\nabla_{P_i} Y)_p = 0 , \quad \text{Q.E.D.}$$

(3) By the proposition above, $\nabla_X Y$ is well-defined at a point, even if X and Y are defined only in a neighborhood of that point (rather than on the whole manifold). Thus the following definition makes sense:

For $\{x_1, \ldots, x_n\}$ a local coordinate system on M, $\partial_i = \partial/\partial x_i$ as before, we define n^3 smooth functions $\Gamma_{ij}^k(x)$ $(i, j, k = 1, \ldots, n)$ on the coordinate patch by

$$\Delta_{\partial_i}(\partial_j)_x = \sum_k \Gamma_{ij}^k(x)\, \partial_k(x).$$

The functions Γ_{ij}^k are called the <u>Christoffel symbols</u> of the connection. We can calculate that for $X = \sum a_i(x)\partial_i$, $Y = \sum b_j(x)\partial_j$,

$$\nabla_X Y = \sum_i a_i \left[\sum_j \frac{\partial b_j}{\partial x_i} \partial_j + \sum_{j,k} b_j \Gamma_{ij}^k \partial_k \right].$$

(4) Examples:

(a) In $\mathbb{R}^n$, with the usual coordinate system, let Γ_{ij}^k be identically zero. Then we get the usual directional derivative of vector fields.

(b) If M is embedded in N (particularly $\mathbb{R}^n$), and if N has a connection ∇^N and a Riemannian metric, then we can use these to define a connection on M as follows: For $p \in M$ and X, Y vector fields defined on M in a neighborhood of p, extend X and Y to vector fields defined on N in a neighborhood of p. Define $\overset{M}{\nabla}$ by setting $(\overset{M}{\nabla}_X Y)_p = \mathrm{proj}_M((\overset{N}{\nabla}_X Y)_p)$, where proj_M denotes the perpendicular projection of T_pN onto T_pM. It is easy to verify that $\overset{M}{\nabla}$ satisfies the definition of a connection. $(\overset{M}{\nabla}_X Y)_p$ is independent of the extensions of X and Y to N, and ...)

C. Covariant derivative of a vector field along a curve.

(1) Define a vector field X along a curve $c: I \to M$ to be a map X such that

$$\begin{tikzcd} & TM \arrow[d] \\ I \arrow[ur, "X"] \arrow[r, "c"'] & M \end{tikzcd}$$

commutes.

Note difficulties involved in extending X to M when c crosses itself, has cusps, stationary points, etc. An example of a vector field along a curve c is the velocity $\dot{c}$.

(2) For X a vector field along c, define the <u>covariant derivative of X along</u> c, $\nabla_{\dot{c}} X$, by

(a) where $\dot{c}(t) \neq 0$, extend X to a neighborhood of $c(t)$ in M, and let the covariant derivative along the curve just be the ordinary covariant derivative in M, $\nabla_{\dot{c}}(X)$. We show that the result is independent of the extension by showing that

$$\nabla_{\dot{c}}(Y) = \sum_i (\dot{c} \cdot y_i)\partial_i + \sum_{i,j,k} \dot{c}_i y_j \Gamma^k_{ij} \partial_k ;$$

where $Y = \sum y_i \partial_i$.

(b) Where $\dot{c} = 0$, let $\nabla_{\dot{c}}(Y) = 0$.

D. Parallel translation

(1) For M a manifold with connection ∇; c a smooth curve in M and X a vector field along c, we say that X is parallel along c if $\nabla_{\dot{c}} X = 0$ holds everywhere on c.

In local coordinates $x_1, \ldots, x_n$: let c be given by $c_1(t), \ldots, c_n(t)$; let X be given by $X(t) = \sum X_i(t)\partial_i(c(t))$. Then the equation $\nabla_{\dot{c}} X = 0$ is equivalent to

$$\frac{dX_i}{dt}(t) + \sum_{k,\ell} \Gamma^i_{k\ell}(c(t)) \frac{dc_k}{dt} \cdot X_\ell(t) = 0, \qquad (i = 1, \ldots, n).$$

This is a system of n linear differential equations in n variables. For an initial value t_o and an arbitrarily chosen vector $X(t_o)$ in $T_{c(t_o)}M$, there is a unique vector field $X(t)$ along c which coincides with $X(t_o)$ at $c(t_o)$. The value of this vector field at $c(t_1)$ is said to be the parallel translation of $X(t_o)$ along c to $c(t_1)$.

(2) Note that the parallel translation along c from $c(a)$ to $c(b)$ gives an invertible linear map of $T_{c(a)}M$ to $T_{c(b)}M$. This linear map depends very heavily on c (unless the "curvature" of the connection is zero).

(3) Relation of parallel translation and ∇.

Proposition. Let $X \in T_pM$, Y be a vector field defined in some neighborhood of p. Take any curve c such that $\dot{c}(0) = X$. Then

$$(\nabla_X Y)_p = \lim_{t \to \infty} \frac{(\|^o_{c,t}\ Y(c(t))) - Y(p)}{t}$$

(where $\|^o_{c,t}$ denotes the parallel translation along c from $c(t)$ to $c(0) = p$.)

Proof. Let $\{Z_1, \ldots, Z_n\}$ be a basis of T_pM. Extend Z_i by parallel translation to a vector field along c. Thus, for each t, $\{Z_1(t), \ldots, Z_n(t)\}$ is a basis for $T_{c(t)}M$.

Write $Y(c(t)) = \sum y_i(t) Z_i(t)$. As parallel translation is linear and the Z_i's are parallel along c, we get that

$$\|^{o}_{c,t}\, Y(c(t)) = \sum y_i(t) Z_i(0).$$

Taking the difference and the limit, we find that the right hand side of our conclusion becomes

$$\sum_i \left(\lim_{t\to 0} \frac{y_i(t) - y_i(0)}{t}\right) Z_i(0) = \sum_i (\dot{c} \cdot y_i)_0 Z_i(0) .$$

But, as $\nabla_{\dot{c}} Z_i = 0$, this equals the left hand side of our conclusion:

$$\begin{aligned} \nabla_X Y &= \nabla_{\dot{c}}\left(\sum y_i(t) Z_i(t)\right) \\ &= \sum (\dot{c} \cdot y_i) \cdot Z_i(t) + \sum y_i(\nabla_{\dot{c}} Z_i) \\ &= \sum (\dot{c} \cdot y_i) \cdot Z_i(t) + 0 \quad , \qquad \text{Q.E.D.} \end{aligned}$$

(4) Note the similarity of the above proposition to the proposition giving the Lie derivative L_X in terms of the flow of X. As a parallel to Willmore's theorem, we have:

Theorem. We can extend ∇_X to a unique linear map of the various tensor bundles

$$\nabla_X : T_s^{\,r}(M) \to T_s^{\,r}(M)$$

such that

(1) $\nabla_X f = X \cdot f$ for $f \in \mathfrak{F}(M)$,

(2) For Y a vector field on M, $\nabla_X Y$ is the given covariant derivative.

(3) $\nabla_X \delta = 0$, where $\delta = \sum_i e^i \otimes e_i$.

(4) ∇_X is a derivation of the tensor algebra:

$$\nabla_X(\tau \otimes \tau') = (\nabla_X \tau) \otimes \tau' + \tau \otimes (\nabla_X \tau')$$

Further, we can also extend the notion of parallel translation along c to

$$\|_{c,a}^{b} : T_s^{\,r}(M)_{c(a)} \longrightarrow T_s^{\,r}(M)_{c(b)}$$

and, for any tensor field τ, $\nabla_X \tau$ is given by a limit, as in the previous proposition.

Example: Using (3) and (4), we can find

$$\nabla_{\partial/\partial x_i}(dx^j) = -\sum_k \Gamma^j_{ik}\, dx^k .$$

III (§50) Nice Covariant Derivatives.

A. Torsion

(1) Symmetry. Suppose $S: \mathbb{R}^2 \to M$ is a smooth map. (Call S a "parametrized surface".) Then we get two vector fields on S,

$$S_x = \frac{\partial S}{\partial x} \quad \text{and} \quad S_y = \frac{\partial S}{\partial y} ,$$

$$\begin{tikzcd} & TM \arrow[d] \\ \mathbb{R}^2 \arrow[ur, "S_x, S_y"] \arrow[r, "S"] & M \end{tikzcd}$$

We can form $\nabla_{S_x} S_y$ and $\nabla_{S_y} S_x$. (These correspond to covariant derivatives along the curves $S(t, y_o)$ and $S(x_o, t)$ respectively.) In general, it is not true that

$$\nabla_{S_x} S_y = \nabla_{S_y} S_x \, .$$

It this condition is satisfied for all parametrized surfaces S, then we say that the connection ∇ is <u>torison free</u> or <u>symmetric</u>. (Note: this condition does <u>not</u> correspond to the property $\partial^2/\partial x \partial y = \partial^2/\partial y \partial x$ in ordinary Euclidean space. <u>That</u> property corresponds to the "curvature" of the connection being zero.)

(2) The torsion tensor.

For vector fields X, Y on M, define

$$\mathrm{Tor}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Show

(a) Tor is $\mathcal{F}(M)$-linear in X and Y,

(b) Follows from (a) that $\mathrm{Tor}(X, Y)_p$ depends only on X_p and Y_p, and bilinearly on these.

(c) (b) means that Tor is a tensor, the <u>torsion tensor</u> of the connection ∇. (Actually, more properly speaking, the torsion tensor is the tensor τ of type $\binom{2}{1}$ given by

$$\tau(X, Y, \omega) \equiv \omega(\mathrm{Tor}(X, Y))$$

for X, Y vector fields and ω a 1-form.)

(d) Calculate local form: If we let

$$\mathrm{Tor}(\partial_i, \partial_j) = \sum_k \mathrm{Tor}_{ij}^{\;k} \, \partial_k$$

then we find

$$\mathrm{Tor}_{ij}^{\;k} = \Gamma_{ij}^{\;k} - \Gamma_{ji}^{\;k} .$$

(3) Relation of Tor and symmetry.

Theorem. $\mathrm{Tor} \equiv 0$ if and only if $\nabla_{S_x} S_y = \nabla_{S_y} S_x$ for all parametrized surfaces.

Proof. ($\Longleftarrow$). For any two vectors X and Y at p, we can choose S so that $(S_x)_p = X$ and $(S_y)_p = Y$. It is easy to calculate that, because S_x and S_y both come from S, $[S_x, S_y] = 0$. (The calculation reduces to $\frac{\partial^2}{\partial y \partial x} = \frac{\partial^2}{\partial y \partial x}$ on $\mathbb{R}^2$.) Then

$$\nabla_{S_x} S_y = \nabla_{S_y} S_y \text{ implies } \nabla_X Y - \nabla_Y X = 0 \text{ and } [X, Y] = 0,$$

so $\mathrm{Tor}(X, Y) = 0$ for all X, Y.

($\Longrightarrow$). $\mathrm{Tor}(S_x, S_y) = 0$. But, again $[S_x, S_y] = 0$. Q.E.D.

B. Invariance of g under parallel translation.

(1) Definition. If M has both a connection ∇ and a Riemannian or pseudo-Riemannian metric $g(X, Y) = (X, Y)$, then it will be nice if parallel translation preserves inner products; i.e., whenever $X(t)$ and $Y(t)$ are parallel along c, then $(X(t), Y(t))$ is independent of t.

(2). Proposition. (1) above holds if and only if the following condition holds: if A and B are vector fields on M, then

$$X\cdot(A,B) = (\nabla_X A, B) + (A, \nabla_X B).$$

Proof. ($\Longrightarrow$) Take c a curve with $\dot{c}(0) = X$. Take an orthonormal basis $Y_1, \ldots, Y_n$ at $c(0)$. Extend these by parallel translation along c. By our hypothesis (1) above, the vectors $Y_1(t), \ldots, Y_n(t)$ form an orthonormal basis in $T_{c(t)}M$ for each t.

We can write

$$A(c(t)) = \sum f_i(t) Y_i(t),$$

$$B(c(t)) = \sum g_i(t) Y_i(t).$$

Then

$$(A(c(t)), B(c(t))) = \sum f_i(t)\cdot g_i(t),$$

and

$$\begin{aligned} X\cdot(A,B) &= \frac{d}{dt}(A(c(t)), B(c(t))) = \frac{d}{dt}\Big(\sum_i f_i(t)\cdot g_i(t)\Big) \\ &= \sum_i [(X\cdot f_i) g_i + f_i (X\cdot g_i)] \\ &= (\nabla_X A, B) + (A, \nabla_X B). \end{aligned}$$

($\Longleftarrow$) is even easier. If A and B are parallel along c, then $\nabla_{\dot{c}} A = \nabla_{\dot{c}} B = 0$, so the derivative of (A,B) along c is

$$\begin{aligned} \dot{c}\cdot(A,B) &= (\nabla_{\dot{c}} A, B) + (A, \nabla_{\dot{c}} B). \\ &= 0 + 0 = 0. \end{aligned}$$

Therefore (A,B) is constant along c. Q.E.D.

Note: If we regard (,) as a tensor $g \in T_o^2(M)$, then the condition that parallel translation preserve inner products is equivalent to $\nabla_X g = 0$ for all vector fields X on M. Here, ∇_X is as described in the theorem at the top of page 96.

Note: The theorem above also holds for pseudo-Riemannian metrics. The modification of the proof is left to the reader.

(3) Main theorem (Holds for pseudo-Riemannian metrics).

Theorem. Given M with a pseudo-Riemannian metric (,), there is a unique connection ∇ on M satisfying

(1) $\mathrm{Tor} = 0$

(2) parallel translation preserves inner products.

Proof. Uniqueness: we have from (2) that

$$X\cdot(Y, Z) = (\nabla_X Y, Z) + (Y, \nabla_X Z).$$

Using (1), this becomes

$$X\cdot(Y, Z) = (\nabla_X Y, Z) + (Y, \nabla_Z X) + (Y, [X, Z]).$$

Cyclically permuting X, Y and Z, we get two other equations. Solving for $(\nabla_X Y, Z)$ and eliminating the terms involving $\nabla_Y Z$ and $\nabla_Z X$ (using the symmetry of (,)) we get

$$2(\nabla_X Y, Z) = X\cdot(Y, Z) + Y\cdot(Z, X) - Z\cdot(X, Y) - (Y, [X, Z]) - (Z, [Y, X]) + (X, [Z, Y]).$$

As (,) is nonsingular, this shows that $\nabla_X Y$ is determined.

Conversely, if we define $\nabla_X Y$ by using this formula, then we find that condition (1) and condition (2) of the theorem are satisfied. Q.E.D.

(4) Local form of the above result

In local coordinates, using the fact that $[\partial_i, \partial_j] = 0$ we get

$$2\Gamma_{ij}^{k} = \sum_{\ell} [\partial_i (g_{j\ell}) + \partial_j (g_{i\ell}) - \partial_\ell (g_{ij})] g^{\ell k} ,$$

where $\partial_i = \dfrac{\partial}{\partial x_i}$, and $(g^{\ell k})$ is the inverse matrix of (g_{ij}).

C. Example. Suppose N (especially $\mathbb{R}^n$) is a manifold with a metric g and the unique corresponding covariant derivative ∇^N. Let M be embedded in N. M inherits a metric h (see I. C. 3) and a connection ∇^M (see II. B. 4). Claim that ∇^M is the unique connection on M corresponding to the metric h.

Proof. (1) Tor is zero: if S is a surface in M, then it is a surface in N. Then $\nabla^N_{S_x} S_y = \nabla^N_{S_y} S_x$, so their projections $\nabla^M_{S_x} S_y$ and $\nabla^M_{S_y} S_x$ into M are equal.

(2) Show

$$X \cdot (Y, Z) = (\nabla^M_X Y, Z) + (Y, \nabla^M_X Z) \quad \text{for } X, Y, Z \text{ tangent fields}$$

to M. But the left hand side is independent of whether we look in M or in N. The equation holds in N. $\nabla^N_X Y$ differs from $\nabla^M_X Y$ by a vector perpendicular to M, so $(\nabla^M_X Y, Z) = (\nabla^N_X Y, Z)$, and so on. Q.E.D.

IV. (§51) Lagrange's Equations.

Suppose N particles move in space, subject to certain constraints. For each allowed configuration of the N particles, we get a point in 3N-space. We assume that the set of allowed configurations is a submanifold M of $\mathbb{R}^{3N}$ of dimension n and that arbitrary motions on the submanifolds are possible. This is what it means for the constraints to be "holomonic".

Put the metric on $\mathbb{R}^{3N}$ given by

$$h = \sum_{i=1}^{N} m_i(dx_i^2 + dy_i^2 + dz_i^2),$$

where m_i is the mass and (x_i, y_i, z_i) the coordinates of the i^{th} particle.

Let ω denote the 1-form on $\mathbb{R}^{3N}$ given by

$$\omega = \sum_{i=1}^{N} F_{ix} dx^i + F_{iy} dy^i + F_{iz}\, dz^i \;,$$

where F_{ix} is the force on the i^{th} particle in the x-direction, etc.

Now, we have the inclusion $M \overset{f}{\hookrightarrow} \mathbb{R}^{3N}$. Let $g = f^*(\omega)$, and let ∇ be the unique nice connection on M associated with g. Now, g produces an isomorphism of T_*M with T^*M. Let X_ω be the vector field corresponding to ω_Q under this isomorphism. Then the equations of motion may be expressed for a path c in M as

$$\nabla_{\dot{c}}\, \dot{c} = X_\omega$$

That is, given an initial position $c(t_o)$ and an initial velocity $\dot{c}(t_o)$ the system follows the unique path c(t) satisfying this equation for these initial conditions.

for each k, such that $ds(\omega) + sd(\omega) = \omega$ for every form ω. If then we have an ω with $d\omega = 0$, it will follow that $\omega = d(s\omega)$, showing that ω is exact. We will also let V denote the tangent space (at any point) of U

A k-form ω may be regarded as a smooth map from U to $\Lambda_k(V^*)$, the space of alternating k-tensors on V. Thus for each $u \in U$, ω_u is an alternating k-tensor: $v_1, \ldots, v_k \in V$ implies that $\omega_u(v_1, \ldots, v_k) \in \mathbb{R}$. Write $\omega(u, v_1, \ldots, v_k) = \omega_u(v_1, \ldots, v_k)$; then ω is a function smooth in the first argument, and linear and alternating in the last k arguments.

Suppose f is a smooth real-valued function on U. We define a new function $Df\colon U \times V \longrightarrow \mathbb{R}$ by letting $Df(u,v) = \langle d_u f, v \rangle$; that is, $Df(u,v) = \left.\dfrac{d(f \circ \tilde{v})}{dt}\right|_{t=0}$ where v is the path defined by $\tilde{v}(t) = u+tv$. Hence Df is nothing more than the directional derivative of f in the direction v at the point u. Now if f happens to be a function of other variables as well, we can still form Df by ignoring those other variables as we take the derivative, and then putting them back: thus if

$f = f(u, w_1, \ldots, w_r)$,

$$Df(u, v, w_1, \ldots, w_r) = (d/dt)\, f(u+tv, w_1, \ldots, w_r)\Big|_{t=0} .$$

Notice that Df is a linear function of v; if also f happens to be a linear function (in u), $Df(u,v) = f(v)$.

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; $\omega = \sum w_i dq^i$, 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

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

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

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

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

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

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

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

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

Proof. First look at unconstrained motion in $\mathbb{R}^{3N}$. It is easy to see that the metric h gives us the usual connection $\nabla^R(\Gamma^k_{ij} \equiv 0)$ on $\mathbb{R}^{3N}$. Now specify the path c by giving coordinates c_{jx}, c_{jy}, c_{jz} for the j^{th} particle, $j = 1, \ldots, n$. The definition of the covariant derivative along c then gives

$$\nabla^R_{\dot{c}}\dot{c} = \sum_j \left(\frac{d^2 c_{jx}}{dt^2} \frac{\partial}{\partial x_j} + \frac{d^2 c_{jy}}{dt^2} \frac{\partial}{\partial y_j} + \frac{d^2 c_{jz}}{dt^2} \frac{\partial}{\partial z_j} \right)$$

Further, if X^R_ω is the vector field which corresponds to the form ω by way of h then

$$X^R_\omega = \sum_j \left(\frac{F_j x}{m_j} \frac{\partial}{\partial x_j} + \frac{F_j y}{\partial y_j} + \frac{F_j z}{m_j} \frac{\partial}{\partial z_j} \right).$$

Therefore the statement

$$\nabla^R_{\dot{c}}\, \dot{c} = X^R_\omega$$

is equivalent to Newton's equations in $\mathbb{R}^N$.

Let us decompose this vector identity into components parallel an perpendicular to $M \subsetneq R^N$.

$$(\nabla^R_{\dot{c}}\dot{c})_{||} + (\nabla^R_{\dot{c}}\dot{c})_{\perp} = (X^R_\omega)_{||} + (X^R_\omega)_{\perp} .$$

We know (III, C.) that $(\nabla^R_{\dot{c}}\dot{c})_{||} = \nabla_{\dot{c}}\dot{c}$ in M. Therefore Newton's equations are equivalent to:

$$\nabla_{\dot{c}}\dot{c} = (X^R_\omega)_{||}$$

$$(\nabla^R_{\dot{c}}\dot{c})_{\perp} = (X^R_\omega)_{\perp} .$$

The statement that the motion is constrained to M says that the second equation must balance. Therefore, the first equation is our equation of motion. We will have proven our result if we show that $(X^R_\omega)_{||} = X_\omega$, as defined at the start. As these are both tangent fields to M, it is enough to show that

$$g((X^R_\omega)_{||}, Y) = g(X_\omega, Y) \qquad \text{for } Y \in T_* M.$$

But

$$\begin{aligned} g((X^R_\omega)_{||}, Y) &= h(X^R_\omega, Y) \quad \text{as } Y \in T.M \\ &= \omega(Y) \quad \text{by definition of } X^R_\omega \\ &= \langle \omega, f_* Y \rangle \\ &= \langle f^* \omega, Y \rangle \\ &= g(X_\omega, Y) \text{ by definition of } X_\omega. \quad \text{Q.E.D.} \end{aligned}$$

We make several comments on the material above.

(1) The equation $\nabla_{\dot{c}} \dot{c} = X_\omega$ is the "same" as Lagrange's equations. To see that, take $g(-, \frac{\partial}{\partial x_j})$ of both sides. Then

$$g(X_\omega, \frac{\partial}{\partial x_j}) = \omega(\frac{\partial}{\partial x_j}) = Q_j ,$$

the generalized force in the j^{th} direction. To analyze the term $g(\nabla_{\dot{c}} \dot{c}, \frac{\partial}{\partial x_j})$, let Z be a vector field extending $\dot{c}$,

$$g(\nabla_Z Z, \frac{\partial}{\partial x_j}) = Z \cdot g(\frac{\partial}{\partial q^j}, Z) + g(Z, [\frac{\partial}{\partial q^j}, Z]) - \frac{1}{2} \frac{\partial}{\partial q^j} \cdot g(Z, Z),$$

(using the relation of the Theorem III. B. 3).

Claim that this expression is $\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^j}) - \frac{\partial T}{\partial q^j}$. But, for $v_o \in M_*$ $T(v_o) = \frac{1}{2} g(v_o, v_o)$. Thus $\left.\frac{\partial T}{\partial \dot{q}^j}\right|_{v_o} = g(v_o, \frac{\partial}{\partial q^j})$, and, as $Z = \dot{c}$, $\frac{d}{dt}(\frac{\partial T}{\partial \dot{q}^j}) = Z \cdot g(Z, \frac{\partial}{\partial q^j})$. To show that

$$g(Z, [\frac{\partial}{\partial q^j}, Z]) - \frac{1}{2}\frac{\partial}{\partial q^j} \cdot g(Z, Z) = -\frac{\partial T}{\partial q^j} : \quad \text{Let } Z = \sum_{\ell} z^{\ell} \frac{\partial}{\partial x^{\ell}} .$$

Then

$$[\frac{\partial}{\partial q^j}, Z] = \sum_m \frac{\partial z^m}{\partial q^j} \cdot \frac{\partial}{\partial q^m} .$$

Let g be given by g_{ij} . Then

$$g(Z, [\frac{\partial}{\partial q^j}, Z]) = \sum_{k,\ell} g_{k\ell}\, z^k \frac{\partial z^{\ell}}{\partial q^j}$$

and

$$\frac{1}{2}\frac{\partial}{\partial q^j}(g(Z, Z)) = \frac{1}{2}\frac{\partial}{\partial q^j}\left(\sum_{k,\ell} g_{k\ell}\, z^k z^{\ell}\right)$$

$$= \frac{1}{2}\sum_{k,\ell} \frac{\partial g_{k\ell}}{\partial q^j} z^k z^{\ell} + \sum_{k,\ell} g_{k\ell} z^k \frac{\partial z^{\ell}}{\partial q^j} ,$$

and therefore

$$g(Z, [\frac{\partial}{\partial q^j}, Z]) - \frac{1}{2}\frac{\partial}{\partial q^j} g(Z, Z) = -\frac{1}{2}\sum_{k,\ell} \frac{\partial g_{k\ell}}{\partial q^j} z^k z^{\ell}$$

$$= (-\frac{\partial T}{\partial q^j})(Z),$$

(where $T(q^1, \ldots, q^n; \dot{q}^1, \ldots, \dot{q}^n) = \frac{1}{2}\sum_{k,\ell} g_{k\ell}(q^1, \ldots, q^n) \cdot \dot{q}^k \cdot \dot{q}^{\ell}$) .

Thus we have shown that we do indeed have Lagrange's equations here.

(By being more sophisticated, we could probably have done this with less involvement in the coordinates. But note that our final result involves the coordinates, so we cannot avoid them entirely.)

(2) Note: X_ω does not depend on forces which are perpendicular (relative to h) to M. These "forces of constraint" may therefore be ignored in

setting up the equations $\nabla_{\dot{c}}\dot{c} = X_\omega$ for solution. This is the whole advantage of the method.

(3) If the forces parallel to M are zero, $\nabla_{\dot{c}}\dot{c} = 0$ says that the path is a "geodesic". For a single particle moving on a surface (e. g. a marble moving (without gravity) on a cone) this is particularly reasonable, as it just says that the acceleration is perpendicular to the surface. (Recall III. C.)

(4) Note finally that if we take the inner product of $\nabla_{\dot{c}}\dot{c} = X_\omega$ with $\dot{c}$ and apply III. B. 3 again, we get

$$\dot{c}\cdot[\tfrac{1}{2}g(\dot{c}, \dot{c})] = \omega(\dot{c})$$

which simply says that the rate of change of the kinetic energy T $(= \frac{1}{2} g(\dot{c}, \dot{c}))$ is given by the work-form applied to $\dot{c}$, as it should be.

SUPPLEMENT - EULER'S EQUATIONS

By Raphael Zahler

Frequently in classical mechanics it happens that the configuration space of the dynamical system in question has the structure of a Lie group. This means that it is a differentiable manifold with an additional "multiplication" operation related to the structure of the manifold. The points of the manifold are then thought of as motions of the system; the product xy stands for the motion resulting from the combined effect of the motion x followed by the motion y. For example, the group of all rotations of an asymmetrical three-dimensional body which leaves a particular point fixed is the familiar Lie group $SO(3)$. The reader may consult Helgason, Differential Geometry and Symmetric Spaces, (Academic Press) for a full treatment of the mathematical theory of Lie groups; here we will briefly outline some important facts. For any fixed element g of the Lie group G, multiplication on the left by g gives a map L_g of G into itself called "left translation by g". The induced map on tangent spaces, L_{g*}, maps $T_e(G)$ to $T_g(G)$, where e is the identity element of the group G. In this way the structure of the vector space $T_e(G)$ is closely related to the overall structure of G. $T_e(G)$ is called the Lie algebra of G. There is a function, called the exponential map, which takes vectors of $T_e(G)$ to points of G; if $\exp X$ is the point corresponding to the vector X, then $\exp(t_1+t_2)X = (\exp t_1X)(\exp t_2X)$; in particular, $\exp 0 = e$.

Suppose now that the kinetic-energy metric on our Lie group is left-invariant; that is,

$$(X,Y)_g = (L_{h*}X, L_{h*}Y)_{hg}$$

for all $g,h \in G$, $X,Y \in T_g(G)$. (The subscript "g" in $(X,Y)_g$ denotes that the metric is being applied to tangent vectors at the point g.) It is then a fact that the geodesics of the metric, which represent the motion of the system in time, can be described near e by $c(t) = \exp X(t)$, where $X(t) \in T_e(G)$, all t. We will investigate the behavior of these trajectories close to e.

Consider a geodesic $\exp \gamma(t)$, and let its tangent vector at time t be $\dot{\gamma}(t)$. Then the kinetic energy as a function of time is

$$T(t) = \frac{1}{2}(\dot{\gamma}, \dot{\gamma})_{\exp \gamma} = \frac{1}{2}(L_{\exp(-\gamma)*}\dot{\gamma}, L_{\exp(-\gamma)*}\dot{\gamma})_e ,$$

or, if we write $\xi(t) = L_{\exp(-\gamma)*}\dot{\gamma}$,

$$T(t) = \frac{1}{2}(\xi(t), \xi)t)).$$

Let us assume that there is no potential energy. Then we will be able to show that ξ satisfies Euler's equation: $\dot{\xi} = -B(\xi, \xi)$, where the function $B: T_e(G) \times T_e(G) \to T_e(G)$ is defined uniquely by

$$([X, Y], Z) = (B(Z, X), Y) \quad , \quad \text{all } Y.$$

First of all, since everything involved is invariant under left translation, it will suffice to consider the case $\gamma(0) = 0$; any other geodesic will be a left translation of one of these. Next, it may be proved using a "Taylor expansion" technique that

$$L_{\exp X*} Y = Y - \frac{1}{2}[X, Y] + O(|X|^2),$$

where the symbols X and Y on the right-hand side are understood in terms of a special "canonical" coordinate system $\{x_1, \ldots, x_n\}$ in a neighborhood of e by which we identify the vectors of the various different tangent spaces near e. Let us plug this into our formula for the Lagrangian:

$$L = T = \frac{1}{2}(\xi, \xi)_e = \frac{1}{2}(L_{\exp(-\gamma)*}\dot{\gamma}, L_{\exp(-\gamma)*}\dot{\gamma})_e$$

$$= \frac{1}{2}(\dot{\gamma} - \frac{1}{2}[-\gamma, \dot{\gamma}] + O(|\gamma|^2), \dot{\gamma} - \frac{1}{2}[\gamma, \dot{\gamma}] + O(|\gamma|^2))_e$$

$$= \frac{1}{2}((\dot{\gamma}, \dot{\gamma})_e - 2(\dot{\gamma}, \frac{1}{2}[-\gamma, \dot{\gamma}])_e + O(|\gamma|^2))$$

$$= \frac{1}{2}(\dot{\gamma}, \dot{\gamma})_e - \frac{1}{2}(\dot{\gamma}, [\dot{\gamma}, \gamma])_e + O(|\gamma|^2)$$

$$= \frac{1}{2}(\dot{\gamma}, \dot{\gamma})_e - \frac{1}{2}(B(\dot{\gamma}, \dot{\gamma}), \gamma)_e + (|\gamma|^2).$$

We now invoke the Euler-Lagrange equations, which must be satisfied by any geodesic: in terms of the canonical coordinates,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_r} = \frac{\partial L}{\partial x_r} \qquad r = 1, \ldots, n.$$

Writing $L = L(t, \gamma, \xi) = L(t, x_1, \ldots, x_n, \dot{x}_1, \ldots, \dot{x}_n)$, we get $L = \frac{1}{2}(\xi, \xi)_e = \frac{1}{2}\sum g_{ii}\dot{x}_i^2$, where we assume that the matrix of constants $\{g_{ij}\}$ representing the metric at e has been diagonalized. Hence

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_r} = g_{rr}\frac{d\dot{x}_r}{dt} \qquad \text{(neglecting the "0" term)}$$

Next: $\dfrac{\partial L}{\partial x_r} = \dfrac{\partial}{\partial x_r}\left(\frac{1}{2}(\dot{\gamma}, \dot{\gamma}) - \frac{1}{2}(B(\dot{\gamma}, \dot{\gamma}), \gamma)\right).$

If we write $B(\dot{\gamma}, \dot{\gamma}) = \sum b_i \dfrac{\partial}{\partial x_i}$ then $\dfrac{\partial \dot{x}_r}{dt} = -b_r$, and the Euler-Lagrange equations now imply that Euler's equation is satisfied near the origin by the trajectories of the dynamical system, when the configuration space happens to form a Lie group with left-invariant metric. (Note: this derivation is due to Arnold (Comptes Rendues, v. 260 (May 31, 1965), p. 5668); a derivation independent of Lie-group theory is found in Loomis and Sternberg, <u>Advanced Calculus</u> (Addison-Wesley), p. 541 ff.)

Let us apply this to the case of rigid-body motion. Considering only rotations leaving a point fixed, a moment's reflection shows that the rotation group $SO(3)$ is actually the configuration space of our system; just take a fixed reference position of the body and consider any other position as a rotation from the reference position. The tangent space to the Lie group $SO(3)$ at the origin can be identified with the vector space of all skew-symmetric three-by-three matrices; these are usually referred to in physics texts as "infinitesimal rotations", and this is the Lie algebra we must work with.

Let $F(t)$ be a curve in configuration space; then a particle at a point p of Euclidean space moves in the trajectory $(F(t))(p)$. A physically sensible Riemannian metric in this case is the <u>inertia tensor</u> $(A, B) = \int m(A_p, B_p)dp$ where A and B are skew-symmetric matrices. In general, this is a left- but not right-invariant metric. In terms of a basis $\{e_i\}$ of $\mathbb{R}^3$, we have

$$\begin{aligned}(A, B) &= \int m(A(\sum r_i e_i), B(\sum r_j e_j))\, dp \\ &= \sum_{i,j=1}^{3} (Ae_i, Be_j) \int m r_i r_j\, dp \\ &= \sum_{i,j=1}^{3} I_{ij}(Ae_i, Be_j).\end{aligned}$$

I, the coordinatized version of the inertia tensor, may be diagonalized (principal axis theorem); picking an obvious basis E_{12}, E_{13}, E_{23} of the Lie algebra gives

$$(E_{ij}, E_{ij}) = \begin{cases} I_i + I_j & i \neq j \\ 0 & \text{otherwise} \end{cases}.$$

It is now possible to substitute in Euler s equation as we derived it above, to obtain

$$\frac{da_{12}}{dt}(I_1 + I_2) - (I_2 - I_1)a_{13}a_{23} = 0,$$

and two similar equations obtained by cyclic permutations of the indices. This is the form of Euler's equations without potential usually found in physics texts; it may be used to solve problems like those involving the spinning top.

A similar situation occurs in the physics of fluid flow. If we have a domain filled with a uniform incompressible ideal fluid, the group of volume-preserving diffeomorphisms of this domain forms configuration space, and, in certain conditions, is a Lie group. Euler's equation, in the form in which we have derived it, now yields

$$\frac{d}{dt}(\mathrm{curl}\ \vec{\xi}) = \mathrm{curl}(\vec{\xi} \times \mathrm{curl}\ \vec{\xi})$$

where ξ is now interpreted as the velocity vector field of the fluid. This is known as Euler's equation for fluid mechanics.