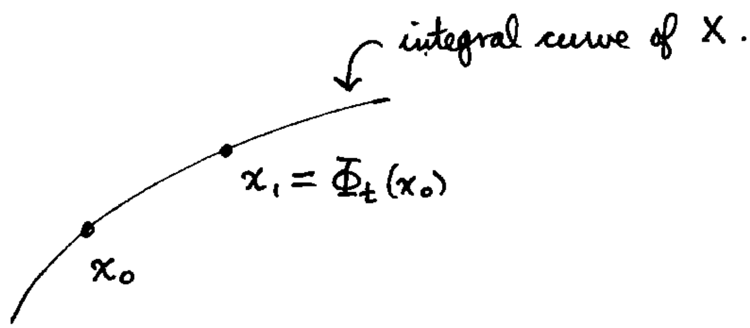
Continue today with the <u>Lie derivative</u>, which is like the convective derivative of ordinary tensor analysis, but generalized to arbitrary manifolds. 3/2/04

<u>Context</u>: Given a manifold M, a vector field $X \in \mathfrak{X}(M)$, with advance map $\Phi_t : M \to M$. Illustrate Lie derivative first with scalar fields, where $\mathcal{L}_X : \mathcal{F}(M) \to \mathcal{F}(M)$ ($\mathcal{L}_X$ is the Lie derivative along vector field X.)

Let $x_0 \in M$ and $x_1 = \Phi_t x_0$. We think of t as small (we will be interested in the limit $t \to 0$). For $f \in \mathcal{F}(M)$, define

$$(\mathcal{L}_X f)(x_0) = \lim_{t \to 0} \frac{1}{t} \left[f(x_1) - f(x_0) \right]$$



It's pretty obvious from this formula that $\mathcal{L}_X f = Xf$, since the vector $X|_{x_0}$ is the small displacement $x_0 \to x_1$ in small time t. ~~But aside~~ Thus, the Lie derivative of a scalar is the obvious generalization of the convective derivative to an arbitrary manifold,

$$\mathcal{L}_X f = Xf = \sum_i X^i \frac{\partial f}{\partial x^i}. \qquad (\text{Think } \vec{v} \cdot \nabla f).$$

But transform eqn. above:

$$(\mathcal{L}_X f)(x_0) = \lim_{t \to 0} \frac{1}{t} \left[f(\Phi_t x_0) - f(x_0) \right]$$

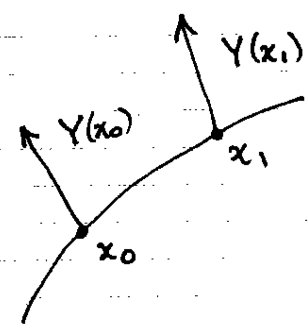
$$= \lim_{t = 0} \frac{1}{t} \left[(\Phi_t^* f)(x_0) - f(x_0) \right]$$

$$= \lim_{t \to 0} \frac{1}{t}\left(\left[\Phi_t^* - 1\right] f\right)(x_0)$$

$$= \left(\left(\left.\frac{d}{dt}\right|_{t=0} \Phi_t^*\right) f\right)(x_0).$$

But recall,
$$\Phi_t^* = e^{tX}$$
when acting on scalars, so we find again, $\left.\frac{d\Phi_t^*}{dt}\right|_{t=0} = X$, $\mathcal{L}_X f = Xf$.

Now generalize to other differential geometric objects, like vector fields. Now we want to define $\mathcal{L}_X$ as an operator: $\mathfrak{X}(M) \to \mathfrak{X}(M)$. Let $Y \in \mathfrak{X}(M)$ (now we have 2 vector fields, X and Y), and we ~~wh~~ wish to define $\mathcal{L}_X Y$. The idea is the same as above, we wish to compare Y at x_1 with Y at x_0 to see how much Y has changed along the integral curves of X. But we cannot just subtract



$Y(x_1) - Y(x_0)$, these vectors belong to two different tangent spaces $T_{x_0}M$ and $T_{x_1}M$ without any natural identification. ~~Not can we as~~ ~~Instead, we~~ However, we can "pull-back" $Y(x_1)$ to a point x_0 using the flow (mapping both base and tip of arrow by Φ_t^{-1}). Note that the pull-back of a vector field is defined in this case because Φ_t is invertible; the pull back is the inverse of the tangent map Φ_{t*}. So, we define

3/2/04

$$\mathcal{L}_X Y = \left(\frac{d}{dt}\bigg|_{t=0} \Phi^{-1}_{t*}\right) Y \text{, or}$$

$$(\mathcal{L}_X Y)(x_0) = \lim_{t\to 0}\frac{1}{t}\left[(\Phi^{-1}_{t*} Y)(x_0) - Y(x_0)\right]$$

In components,

$$(\Phi^{-1}_{t*} Y)^i (x_0) = \frac{\partial x_0^i}{\partial x_1^j} Y^j(x_1).$$

To get x_1 as a fn of x_0, t, we solve the ODE's,

$$\frac{dx^i}{dt} = X^i(x)$$

in power series in t,

$$x_1^i = x_0^i + t X^i(x_0) + \cdots$$

or its inverse,

$$x_0^i = x_1^i - t X^i(x_0) + \cdots$$

So,

$$\frac{\partial x_0^i}{\partial x_1^j} = \delta^i_j - t\, X^i_{,j} + \cdots$$

and,

$$= \left[\delta^i_j - t X^i_{,j}\right] Y^j (x_0 + t X + \cdots)$$

$$= Y^i(x_0) + t\left(X^j Y^i_{,j} - Y^j X^i_{,j}\right).$$

So,

$$\boxed{(\mathcal{L}_X Y)^i = X^j Y^i_{,j} - Y^j X^i_{,j}}$$

This is the Lie derivative of a vector field.

Similarly, you can define the Lie deriv. of a covector field

3/2/04

by
$$\mathcal{L}_X\alpha = \left(\frac{d}{dt}\Big|_{t=0} \Phi_t^*\right)\alpha .$$

If you work it out, you find (in components),

$$\boxed{(\mathcal{L}_X\alpha)_i = X^j \alpha_{i;j} - \alpha_j X^j{}_{;i}}$$

To define $\mathcal{L}_X$ on arbitrary tensors, we develop some general rules. First, $\mathcal{L}_X$ acts on a <u>tensor product</u> of tensors by the <u>Leibnitz rule</u>. An example will illustrate. Consider the tensor product of a covector with a vector (this is a type $(1,1)$ tensor): Define

$$\mathcal{L}_X(\alpha\otimes Y) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\left[(\Phi_\varepsilon^*\alpha)\otimes(\Phi^{-1}_{\varepsilon *} Y)\right],$$

which is the obvious definition. But this is...

$$\boxed{(\mathcal{L}_X\alpha)\otimes Y + \alpha\otimes(\mathcal{L}_X Y).} \quad \boxed{= \mathcal{L}_X(\alpha\otimes Y)}$$

(Leibnitz rule).

The same thing works on contractions. For example, the tensor $\alpha\otimes Y$ has components,

$$(\alpha\otimes Y)_i{}^j = \alpha_i Y^j$$

If we contract (set $i=j$ and sum), we get

$$\alpha_i Y^i = \alpha(Y) = \text{a scalar.}$$

Then we have

$$\mathcal{L}_X[\alpha(Y)] = (\mathcal{L}_X\alpha)(Y) + \alpha(\mathcal{L}_X Y)$$

$$= X(\alpha(Y)).$$

Can use this to calculate $\mathcal{L}_X\alpha$ in components, supposing that we know what $\mathcal{L}_X$ does to scalars and vector fields.

Notice that a scalar multiplied by tensor is a special case of a tensor product:

$$f \otimes T = f\,T \qquad \text{any } T, \quad f \in \mathcal{F}(M).$$

Therefore

$$\mathcal{L}_X (fT) = (\mathcal{L}_X f)T + f(\mathcal{L}_X T) = (Xf)T + f(\mathcal{L}_X T).$$

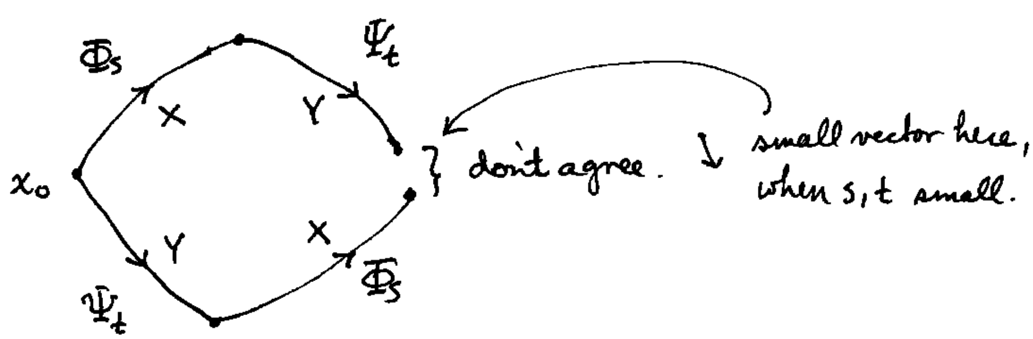
Since an arbitrary tensor field can be written as linear combinations of scalars times tensor products of vector fields and covector fields, the Leibniz rule suffices to compute the Lie derivative of any tensor.

~~Some more rules about $\mathcal{L}_X$. If $f \in \mathcal{F}(M)$, then~~ WRONG ignore

$$\boxed{\mathcal{L}_{fX} = f\,\mathcal{L}_X}$$

~~This is obvious since fX has same integral curves as X, except the t-parameterization is scaled by f. Hence $\left.\frac{d}{dt}\right|_{t=0}$ is scaled by f.~~

The Lie derivative $\mathcal{L}_X Y$ is a special case with a special interpretation. Consider the flows associated with X, Y, call them Φ_s, Ψ_t. These in general do not commute,



3/2/04

When s,t are small, the difference in the endpoints must be a vector. However, since we cannot subtract points, to measure the difference between $\Psi_t \Phi_s x_0$ and $\Phi_s \Psi_t x_0$ we evaluate some scalar $f: M \to \mathbb{R}$ at the 2 points and subtract:

$$f(\Psi_t \Phi_s x_0) - f(\Phi_s \Psi_t x_0)$$

$$= \left((\Psi_t \Phi_s)^* f\right)(x_0) - \left((\Phi_s \Psi_t)^* f\right)(x_0)$$

$$= (\Phi_s^* \Psi_t^* f)(x_0) - (\Psi_t^* \Phi_s^* f)(x_0)$$

$$= \left(\left(\underbrace{\Phi_s^* \Psi_t^* - \Psi_t^* \Phi_s^*}\right) f\right)(x_0).$$

$$= e^{sX} e^{tY} - e^{tY} e^{sX}$$

$$= \left(1 + sX + \frac{s^2}{2} X^2 + \cdots\right)\left(1 + tY + \frac{t^2}{2} Y^2 + \cdots\right) - (X \Leftrightarrow Y)$$

$$= 1 + (sX + tY) + \left(\frac{s^2}{2} X^2 + stXY + \frac{t^2}{2} Y^2\right) + \cdots - (X \Leftrightarrow Y)$$

$$= st\,(XY - YX) + \cdots$$

Thus the small vector is $[X, Y]$ times st.

$$= st\,[X, Y] + \cdots$$

Thus the leading term is the <u>commutator</u> of X and Y (regarded as maps: $\mathcal{F}(M) \to \mathcal{F}(M)$). Thus we have

$$\left.\frac{\partial^2}{\partial s \partial t}\right|_{t=0} \left[f(\Psi_t \Phi_s x_0) - f(\Phi_s \Psi_t x_0) \right] = \left([X, Y] f\right)(x_0),$$

$$\forall f \in \mathcal{F}(M).$$

Now XY is not a vector field (because it is a 2nd order operator), but it turns out that $XY - YX$ <u>is</u> a vector field (all 2nd derivs cancel), and, in fact,

$$\boxed{\mathcal{L}_X Y = [X,Y]}$$

The (important) commutator has the following properties:

a) $[X,Y] = -[Y,X]$

b) $[X,Y]$ linear in X, Y (over $\mathbb{R}$)

c) $[X,[Y,Z]] + \circlearrowright = 0$ (Jacobi).

The set $\mathfrak{X}(M)$ forms a <u>Lie</u> <u>algebra</u>. (of diffeomorphism group).

Some other properties of the commutator:

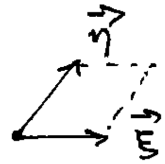
(a) $f_*[X,Y] = [f_*X, f_*Y]$ when $f: M \to N$ is a diffeomorphism

(b) $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$

→ bec. advance maps commute w. diffeomorphisms.

(a) is almost obvious; a diffeomorphism is an isomorphism of differentiable structure, integral curves mapped into integral curves etc.

Now we turn to <u>differential</u> <u>forms</u>. A diff. form of rank r is a <u>completely</u> <u>antisymmetric</u> <u>type $(0,r)$</u> tensor. Why antisymmetric? Because you need these for integrating over oriented, r-dimensional surfaces. Consider an example from 3D vector calculus. Let a small area element be specified by two small vectors $\vec{\xi}$ and $\vec{\eta}$. These might define a small element of a 2D surface.



3/2/04

Then let $\vec{J}$ be a flux vector (of mass, charge, etc., or maybe $\vec{J}=\vec{B}$ = magnetic field). Then the flux through parallelogram is

$$\vec{J}\cdot(\vec{\xi}\times\vec{\eta}).$$

Why $\vec{\xi}\times\vec{\eta}$ and not $\vec{\eta}\times\vec{\xi}$? Because we have to decide which side of the parallelogram is the "outward" oriented side (it's a convention, but the sign of the answer depends on it). So the area element is specified by $\vec{\xi}\times\vec{\eta}$, which is antisymmetric in the two vectors. And the <u>value</u> of the flux is the value of a real-valued <u>linear</u> <u>operator</u> that acts on area elements. It's like a covector (acts on vectors), except that it acts on 2 vectors (effectively area elements). Note, we can write

$$\vec{J}\cdot(\vec{\xi}\times\vec{\eta}) = \tfrac{1}{2}J_{ij}\,(\xi^i\eta^j - \xi^j\eta^i)$$

where $J_{ij} = \varepsilon_{ijk}J^k$. $J_{ij} = -J_{ji}$ are the components of a 2-form.

~~A~~ Special cases of r-forms:

$r=0$ is a <u>scalar</u>, or <u>0-form</u> considered to be antisymmetric in its nonexistent operand.

$r=1$ is a <u>covector</u>, or <u>1-form</u> considered to be antisymmetric in its one operand, $\alpha : \mathfrak{X}(M) \to \mathcal{F}(M)$ (as a field)

$r=2$ is a <u>2-form</u>, an antisymmetric tensor acting on two vector fields,

$$\omega : \mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathcal{F}(M),$$

$$\omega(X,Y) = -\omega(Y,X).$$

3/2/04

Cases $r=0,1,2$

In components: let $x \in M$, x^i = coordinates of x,

let $e_i = \frac{\partial}{\partial x^i}$ = basis vectors of coordinate system.

A scalar (= 0-form) $f: M \to \mathbb{R}$ has ~~no~~ only one component, the value $f(x)$ of f itself.

A covector or 1-form α has components,

$$\alpha_i(x) = \alpha(e_i)\Big|_x .$$

A 2-form ω has components,

$$\omega_{ij}(x) = \omega(e_i, e_j)\Big|_x = -\omega_{ji}(x).$$

The number of independent components of an r-form on an n-dim'l space is

$$\binom{n}{r} = \begin{cases} 1 & r=0 \\ n & r=1 \\ \frac{n(n-1)}{2} & r=2 \\ \vdots & \\ 1 & r=n. \end{cases}$$

Another special case is an n-form, call it ϕ. This is a completely antisymmetric map of n vectors to scalars,

$$\phi: \mathcal{X}(M) \times \dots \times \mathcal{X}(M) \to \mathcal{F}(M)$$

with components

$$\phi_{i_1 \dots i_n}(x) = \text{completely antisymmetric in indices } (i_1, \dots, i_n).$$

$$= \phi(e_{i_1}, \dots, e_{i_n})$$

3/2/04

Thus, a n-form on an n-dimensional ~~sp~~ manifold has components in any given chart that have the form

$$\phi_{i_1 \dots i_n}(x) = \sigma(x) \; \varepsilon_{i_1 \dots i_n},$$

where $\varepsilon_{i_1 \dots i_n}$ is the Levi-Civita symbol, and $\sigma(x)$ is a scalar density. $\sigma(x)$ defines the one and only indep. component of ϕ.

We write the set of all smooth r-forms on M as $\Omega^r(M)$. Thus, $\Omega^0(M) = \mathcal{F}(M)$, $\Omega^1(M) = \mathcal{X}(M)$, etc.

We consider r-forms with $r > n$ to be zero.

How to construct r-forms. One way is to take the exterior product of r 1-forms. The exterior product is an antisymmetrized tensor product. ~~For~~ The exterior product of a r 1-forms is defined as follows.

Let $\alpha^1, \dots, \alpha^r$ be 1-forms $\quad (\alpha^i \in \Omega^1(M))$.

Then $\alpha^1 \wedge \dots \wedge \alpha^r$ is an r-form, defined by its action on r vector fields $X_1, \dots, X_r$ by

$$\underbrace{(\alpha^1 \wedge \dots \wedge \alpha^r)}_{r\text{-form}}(X_1, \dots, X_r) = \sum_{P \in S_r} (-1)^P \alpha^1(X_{P_1}) \alpha^2(X_{P_2}) \dots \alpha^r(X_{P_r})$$

where S_r = set of all permutations P of r objects. More precisely, P is a bijection of the set $\{1, 2, \dots, r\}$ to itself, P_i = value of P acting on i $\quad (1 \le i \le r)$. $(-1)^P$ is the parity of the permutation (+1 if even, −1 if odd).

3/2/04

Can also write this as

$$(\alpha^1 \wedge \ldots \wedge \alpha^r)(X_1, \ldots, X_r) = \begin{vmatrix} \alpha^1(X_1) & \cdots & \alpha^1(X_r) \\ \vdots & & \vdots \\ \alpha^r(X_1) & \cdots & \alpha^r(X_r) \end{vmatrix}$$

Example: Let $\alpha, \beta \in \Omega^1(M)$, $X, Y \in \mathfrak{X}(M)$

$$(\alpha \wedge \beta)(X, Y) = \begin{vmatrix} \alpha(X) & \alpha(Y) \\ \beta(X) & \beta(Y) \end{vmatrix} = \alpha(X)\beta(Y) - \alpha(Y)\beta(X).$$

Properties:

1) $\alpha^1 \wedge \ldots \wedge \alpha^r$ is completely antisymmetric,

$$\alpha^{P_1} \wedge \ldots \wedge \alpha^{P_r} = (-1)^P \alpha^1 \wedge \ldots \wedge \alpha^r.$$

in particular, $\alpha \wedge \beta = -\beta \wedge \alpha$ $\quad (\alpha, \beta \in \Omega^1(M))$.

2) If $\alpha^i = \alpha^j$ for any $i \neq j$, then $\alpha^1 \wedge \ldots \wedge \alpha^r = 0$.

A general r-form is not the exterior product of a set of r 1-forms, but can always be represented as a linear combination of such products. Example: Let A be an antisymmetric, $(0,2)$ tensor,

$$A = A_{\mu\nu}\, dx^\mu \otimes dx^\nu \qquad (A_{\mu\nu} = -A_{\nu\mu})$$

$$= \frac{1}{2} A_{\mu\nu} \left[dx^\mu \otimes dx^\nu - dx^\nu \otimes dx^\mu \right]$$

$$= \frac{1}{2} A_{\mu\nu}\, dx^\mu \wedge dx^\nu.$$

3/2/04

Thus, $dx^\mu \wedge dx^\nu$ $(\mu, \nu = 1, \dots, n)$ is a basis of 2-forms on M.

Similarly, for a general r-form,

$$\overset{\text{r-form}}{\omega} = \frac{1}{r!} \omega_{\mu_1 \dots \mu_r}(x) \underbrace{dx^{\mu_1} \wedge \dots \wedge dx^{\mu_r}}_{\text{basis of r-forms.}}.$$

Here we use summation convention, $\sum_{\mu_1, \dots, \mu_r}$ implied. If you sum only over indices in ascending order,

$$\sum_{\mu_1 < \mu_2 < \dots < \mu_r}$$

you can drop the factor of $\frac{1}{r!}$.

Now generalize the exterior product to arbitrary forms. Let $\alpha \in \Omega^r(M)$, $\beta \in \Omega^s(M)$. Then $\alpha \wedge \beta \in \Omega^{r+s}(M)$, defined by

$$(\alpha \wedge \beta)(X_1, \dots, X_{r+s}) = \frac{1}{r!\, s!} \sum_{P \in S_{r+s}} (-1)^P \cdot \alpha(X_{P_1}, \dots, X_{P_r}) \beta(X_{P_{r+1}}, \dots, X_{P_{r+s}}).$$

Can simplify this.

Properties:

1) $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$ (Associative)

2) $\alpha \wedge \beta = (-1)^{rs} \beta \wedge \alpha$.

Reason: $(-1)^{rs}$ because need rs exchanges to swap order of factors.

Note: 2) implies $\alpha \wedge \alpha = 0$ when r = odd.

Note special case, $r = 0$, $\alpha = 0\text{-form} \equiv f$. Then $f \wedge \beta = f \beta$ (ord. mult.)

3) $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$ where $f: M \to N$.