$$(f^*\alpha)_p(X) = \alpha_{f(p)}(f_*X) \ , \quad \forall \ p \in M$$

10/2/08

where the subscript indicates the point at which the field is evaluated. Thus 1-forms on N get pulled back into 1-forms on M.

A simpler example of the pull-back is for scalar fields. Let $\phi \in \mathcal{F}(N)$, $\phi: N \to \mathbb{R}$ be a scalar field. Then we define the pull back $\psi = f^*\phi \in \mathcal{F}(M)$ by

$$\psi(p) = (f^*\phi)(p) = \phi(f(p)) = (\phi \circ f)(p),$$

that is, $f^*\phi = \phi \circ f$.

Return to the pull-back of covectors, and put it into coordinate language. Let x^i, y^i be coordinates on M and N as above, and let $\beta = f^*\alpha$, so that

$$\alpha = \sum_{i=1}^{n} \alpha_i \, dy^i|_q$$

$$\beta = \sum_{i=1}^{m} \beta_i \, dx^i|_p$$

Then

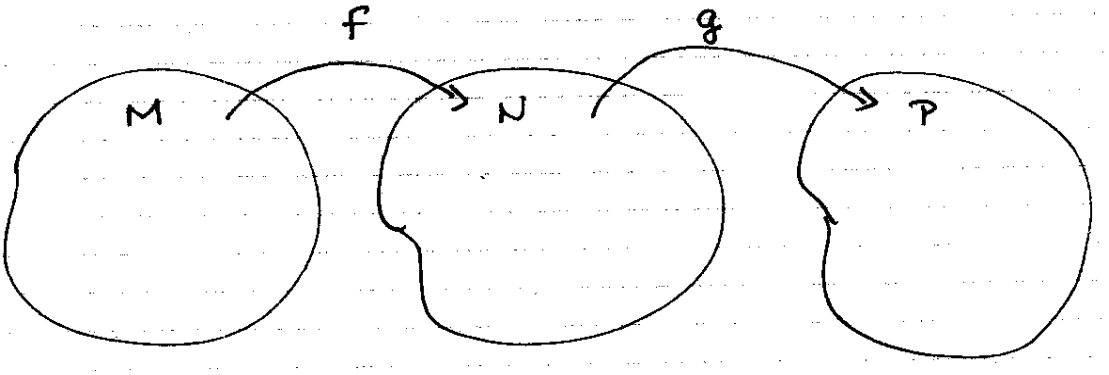
$$\beta_i = \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i} \alpha_j .$$

Notice that f^{-1} need not be defined in order to define f_* and f^*. In particular, M and N need not have the same dimensionality. But if f^{-1} does exist (then $\dim M = \dim N$), then we can "push forward" covectors from M to N (by f^{-1*}) and

(continued page after next)

10/2/08

Behavior of f^*, f_* under compositions. Let $f: M \to N$, $g: N \to P$



Then

$$g \circ f : M \to P,$$

and

$$(g \circ f)_* = g_* \circ f_* .$$

This is fairly obvious, you just map the small displacement vector (in M) under a succession of 2 linear maps, first by f_*, then g_*, to get $(g \circ f)_*$.

As for pull-backs, the rule is

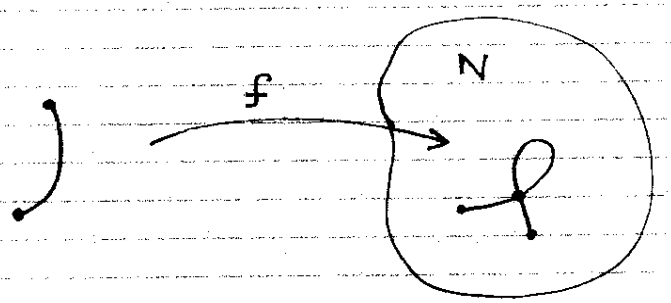
$$(g \circ f)^* = f^* \circ g^* \quad \text{(in reverse order).}$$

"pull-back" vectors from N to M (by f_*^{-1}).

Now consider the mapping of one manifold of a certain dimensionality into one of higher dimensionality, $f: M \to N$, $\dim M \leq \dim N$. Then f is called an <u>immersion</u> if f_* is of maximal rank,

$$\text{rank } f_*: T_pM \to T_{f(p)}N = \dim M.$$

This means that each little piece of M, which looks like $\mathbb{R}^m$, gets mapped into a subset of N that also looks like $\mathbb{R}^m$. This means that f_* is injective (the image of M under f is locally m-dimensional ($m = \dim M$)). However, an immersion does not preclude self-intersections:

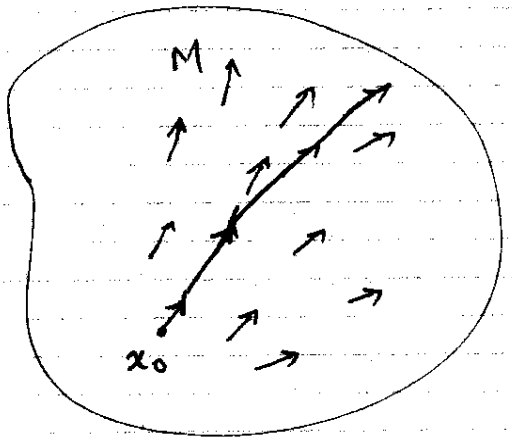


To exclude self-intersections, we can demand that f itself be an injection. ~~This means~~ Then f is called an <u>imbedding</u> (because im f "looks like" M).

Now we consider <u>ordinary differential equations</u> (ODE's) and <u>flows</u>. Begin with an intuitive picture of a <u>vector field</u>, as a small displacement (each understood to be taking place in some elapsed parameter Δt), attached to each point of M:

10/2/08

picture of $X \in \mathfrak{X}(M)$



Idea: Each point gets up and moves a small amount in time Δt.

If you just follow these arrows, starting with some initial point x_0, you trace out a curve called the integral curve of X. By following in integral curve for time t, starting at x_0, you get a final point x described by a function

at least sometimes it has this domain

$$\Phi : M \times \mathbb{R} \to M,$$

$$x = \Phi(x_0, t),$$

where Φ is called the advance map.

To make this more precise, express the vector field X in some chart:

$$X = \sum_i X^i(x) \frac{\partial}{\partial x^i}$$

x = a point
x^i = its coordinates
X = a vector
X^i = its components.

This is an operator which when acting on scalars f gives a number interpreted as df/dt. In particular, letting it act on the coordinates themselves gives a set of ODEs:

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

Thus, a vector field on a manifold is a generalization of a system of ODE's on $\mathbb{R}^n$. Standard theorems on ODE's say that the system above has a unique solution $x^i(t)$ satisfying $x^i(0) = x_0^i$ for t in some interval

$[0,T]$, if the functions $X^i(x)$ are smooth. This is the (important) 10/2/08 <u>uniqueness theorem</u> for ODE's (really, existence and uniqueness). However, even if the vector field $X^i(x)$ is smooth, the solution may not exist for all t (for example, it may run off to infinity in finite t). (For an example of this, consider $\dot{x} = x^2$, $x_0 = 1$ $(x \in \mathbb{R})$, for which $x \to \infty$ as $t \to 1$.) <u>For simplicity</u>, we will <u>assume</u> that this does not happen, i.e., that solutions $x^i(t)$ <u>exist for all time</u>, for any x_0^i. Then we can speak of the "general solution functions" $\Phi^i(t, x_0)$ that give $x^i(t)$, assuming $x^i(0) = x_0^i$. These ~~solution~~ functions satisfy, [we write $x^i(t) = \Phi^i(t, x_0)$, indicating explicitly the dependence on i.c.'s]

1) $$\Phi^i(0, x_0) = x_0^i$$

2) $$\frac{\partial \Phi^i}{\partial t}(t, x_0) = X^i(\Phi(t, x_0)).$$

These are just the initial conditions and ODE's expressed in terms of Φ^i. [We would normally write them,

1) $$x^i(0) = x_0^i$$

2) $$\frac{dx^i}{dt} = X^i(x(t)) \quad .$$].

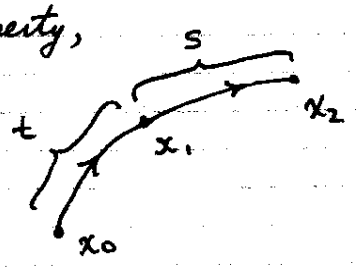
All of the above is in one chart. By mapping a solution $x^i(t)$ in the given chart back onto M, we get a segment of an integral curve. Just before we run off one chart we can switch to another, thereby continuing the integral curve. The result is that we define a map $\Phi : \mathbb{R} \times M \to M$ or maps $\Phi_t : M \to M$ (a different notation), such that

10/2/08

$$x = \Phi(t, x_0) = \Phi_t(x_0)$$

is the point on the integral curve starting at x_0 at $t=0$, reached after time t.

The advance map satisfies an important property,



$$\Phi_s \Phi_t = \Phi_{s+t}$$

$$\text{or} \quad \Phi(s, \Phi(t, x_0)) = \Phi(s+t, x_0).$$

(the <u>composition property</u>). This is intuitive: if you start at x_0, follow the integral curve for elapsed time t, reaching x_1, then treat x_1 as initial conditions and follow the integral curve for elapsed time s, you must get the same thing as starting at x_0 and following the integral curve for time $s+t$.

We will prove this working in a single chart, ignoring the complications that result when we must switch charts. We want to show that

$$\Phi^i(s, \Phi(t, x_0)) = \Phi^i(s+t, x_0).$$

Let $x_1^i = \Phi^i(t, x_0)$, $\quad \xi^i(s) = \Phi^i(s, x_1)$, $\quad \eta^i(s) = \Phi^i(s+t, x_0)$.

We need to show that $\xi^i(s) = \eta^i(s)$. First, at $s=0$, we have

$$\xi^i(0) = \Phi^i(0, x_1) = x_1^i$$
$$\eta^i(0) = \Phi^i(t, x_0) = x_1^i.$$

Next, we have

$$\frac{d\xi^i}{ds} = \frac{\partial \Phi^i}{\partial s}(s, x_1) = X^i(\Phi(s, x_1)) = X^i(\xi(s)),$$

10/2/08

and

$$\frac{d\eta^i}{ds} = \frac{\partial \Phi^i}{\partial s}(s+t, x_0) = \frac{\partial \Phi^i}{\partial (s+t)}(s+t, x_0)$$

$$= X^i(\Phi(s+t, x_0)) = X^i(\eta(s)).$$

Thus, both $\xi^i(s)$ and $\eta^i(s)$ satisfy the same ODE's and the same initial conditions, so by the uniqueness theorem they must be equal. QED.

By the composition property, $\Phi_{-t} \circ \Phi_t = \mathrm{id}_M$, so Φ_t is a <u>diffeomorphism</u> : $M \to M$. In fact, the set

$$\{\Phi_t \mid t \in \mathbb{R}\}$$

constitutes a <u>one-parameter group of diffeomorphisms</u> of M onto itself. It is an <u>action</u> of the group $\mathbb{R}$ (meaning t) on M. This group is sometimes called the <u>flow</u>.

Now about the <u>exponential</u> <u>notation</u> for the flow. This is a way of connecting Φ_t with the vector field X. The notation used in many books is

$$\boxed{\Phi_t = e^{tX}.}$$

Taken ~~as written~~ literally, this has no meaning (we must assign a meaning to it). First note that a vector field X is a mapping : $\mathcal{F}(M) \to \mathcal{F}(M)$, namely, $f \mapsto \sum_i X^i \frac{\partial f}{\partial x^i}$. Therefore $X^2 = X \cdot X = X \circ X$ has a meaning as a map: $\mathcal{F}(M) \to \mathcal{F}(M)$. The higher powers of a vector field are not vector fields (they are not 1st order partial differential operators, they are higher order diff. ops.), but they are perfectly good maps : $\mathcal{F}(M) \to \mathcal{F}(M)$. [Notice that a vector <u>at a point</u> is a map: $\mathcal{F}(M) \to \mathbb{R}$, so a power of it has no meaning.]

10/2/08

So, an exponential series like

$$e^{tX} = 1 + tX + \frac{t^2}{2!}X^2 + \cdots$$

has meaning as a map: $\mathcal{F}(M) \to \mathcal{F}(M)$, at least if we ignore convergence questions (which we will). The 1 above means id_M. On the other hand, Φ_t is a map: $M \to M$, not: $\mathcal{F}(M) \to \mathcal{F}(M)$. This is why $\Phi_t = e^{tX}$ has no meaning as it stands. But let us apply the exponential series to a function f and see what we get. We evaluate the function at x_0, which is an initial condition. ~~(where t instead of x_0)~~

$$e^{tX} f = f + tXf + \frac{t^2}{2} X^2 f + \cdots$$

$$= f + t \sum_i X^i \frac{\partial}{\partial x^i} f + \frac{t^2}{2} \sum_i X^i \frac{\partial}{\partial x^i} \sum_j X^j \frac{\partial}{\partial x^j} f + \cdots$$

$$= f + t \text{ "}\frac{df}{dt}\text{"} + \frac{t^2}{2} \text{ "}\frac{d^2 f}{dt^2}\text{"} + \cdots$$

where we put the t-derivatives in quotes because what is actually meant is

$$\text{"}\frac{df}{dt}\text{"} = \frac{d}{dt}(f \circ c)$$

where c is the integral curve passing through x at $t=0$. Thus, if the series converges, it does so to "$f(t)$", which means $f(x(t))$, where $x(t)$ is the integral curve, $x(t) = \Phi_t(x_0)$. So,

means same as $c(t)$

$$(e^{tX} f)(x_0) = f(\Phi_t(x_0)) = (\Phi_t^* f)(x_0).$$

This is true for all x_0, so we have

$$e^{tX} f = \Phi_t^* f.$$

10/2/08

And this is true for all f, so we have

$$\Phi_t^* = e^{tX}. \qquad \text{(A)}$$

The usual formula in books is meaningful if we put a $*$ on Φ_t (turning it into a pull-back), and interpret both sides as maps : $\mathcal{F}(M) \to \mathcal{F}(M)$.

On the other hand, we can regard e^{tX} as a formal notation for Φ_t, without trying to interpret it as a power series. This notation has the virtue of making some of the properties of the advance map obvious:

$$\Phi_s \Phi_t = e^{sX} e^{tX} = e^{(s+t)X} = \Phi_{s+t}$$

$$\text{and } \Phi_0 = e^{0X} = 1 = \mathrm{id}_M.$$

Restoring the star $*$, we can differentiate $\Phi_t^* = e^{tX}$ formally, to get

$$\frac{d}{dt}\Phi_t^* = X e^{tX} = e^{tX} X,$$

which implies

$$\frac{d}{dt}\Phi_t^* = X \Phi_t^* = \Phi_t^* X, \qquad \text{(B)}$$

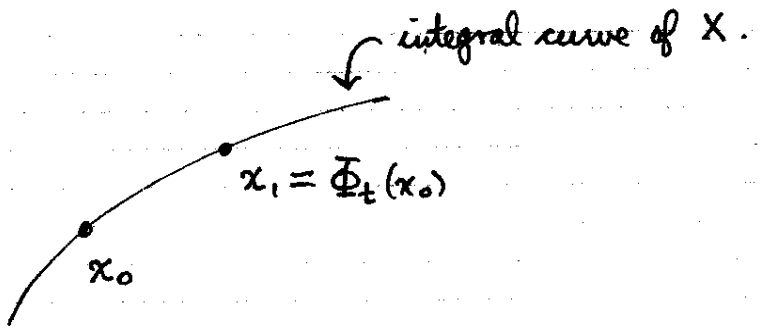
an equation that is perfectly meaningful as operators : $\mathcal{F}(M) \to \mathcal{F}(M)$. We have not proved it (because of questions of convergence of series), but in fact this result is true, and can be proved by other means.

Altogether, we have derived relationships between a vector field X and the 1-parameter group of diffeomorphisms $\{\Phi_t\}$ that it generates, in an intrinsic notation (not tied to a coordinate system). These are Eqs. (A) and (B) above.

Continue today with the <u>Lie derivative</u>, which is like the convective derivative of ordinary tensor analysis, but generalized to arbitrary manifolds. 10/2/08

<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. ~~[illegible]~~ Thus, the Lie derivative of a scalar is the obvious generalization of the convective derivative to an arbitrary manifold,

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

But transform eqn. above:

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

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

10/2/08

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

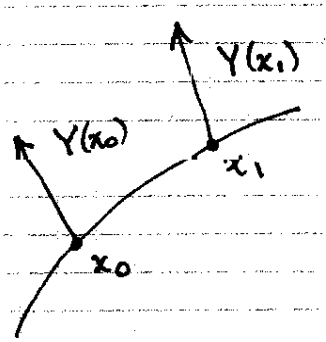
$$= \left(\left(\frac{d}{dt}\Big|_{t=0} \Phi_t^*\right) f\right)(x_0).$$

But recall,

$$\Phi_t^* = e^{tX}$$

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

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 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. However, we can "pull-back" $Y(x_1)$ to 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_{tX}. So, we define

10/2/08

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

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

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

10/2/08

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

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

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

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

(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. 10/2/08

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

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

Therefore

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

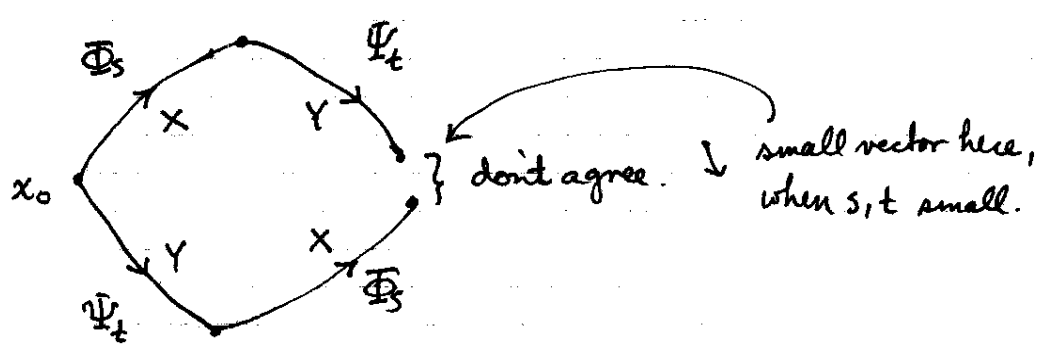
Since an arbitrary tensor field can be written as linear combinations of scalars times tensor products of vector fields and covector fields, the Leibnitz rule suffices to compute the Lie derivative of any tensor.

~~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-parametrization is scaled by f. Hence $\frac{d}{dt}\big|_{t=0}$ 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,



10/2/08

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:

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

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

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

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

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

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

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

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

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

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

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

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

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