Lie Derivatives on Manifolds

William C. Schulz
Department of Mathematics and Statistics,
Northern Arizona University, Flagstaff, AZ 86011

1. INTRODUCTION

This module gives a brief introduction to Lie derivatives and how they act on various geometric objects. The principal difficulty in taking derivatives of sections of vector bundles is that the there is no cannonical way of comparing values of sections over different points. In general this problem is handled by installing a connection on the manifold, but if one has a tangent vector field then its flow provides another method of identifying the points in nearby fibres, and thus provides a method (which of course depends on the vector field) of taking derivatives of sections in any vector bundle. In this module we develop this theory. [1]

2. TANGENT VECTOR FIELDS

A tangent vector field is simply a section of the tangent bundle. For our purposes here we regard the section as defined on the entire manifold M. We can always arrange this by extending a section defined on an open set U of M by 0, after some appropriate smoothing. However, since we are going to be concerned with the *flow* generated by the tangent vector field we do need it to be defined on all of M.

The objects in the $T(M)$ are defined to be first order linear operators acting on the sheaf of C^∞ functions on the manifold. $X_p(f)$ is a real number (or a complex number for complex manifolds). At each point $p \in M$ we have

$$\begin{aligned} X_p(f+g) &= X_p(f) + X_p(g) \\ X_p(fg) &= X_p(f)g(p) + f(p)X_p(g) \end{aligned}$$

$X_p(f)$ should be thought of as the directional derivative of F in the direction X at P. Thus X inputs a function (at p) and outputs a real (or complex) number. We regard f as being defined on some neighborhood of p (which depends on f). The rules above describe how X_p acts on sums and products.

It is possible to show, although we will not do so here, that $T_p(M)$ is an n-dimensional vector space ($n = \dim M$, the dimension of the Manifold M) with a basis in local coordinates $\{\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^n}\}$. Hence

$$X_p = X^i(p) \frac{\partial}{\partial u^i}\Big|_p$$

[1] 7 Oct 2011

and

$$X_p(f) = X^i(p)\frac{\partial f}{\partial u^i}\Big|_p$$

If we change coordinates from u^i to $\tilde{u}^j$ the the local expression of X changes like a contravariant tensor:

$$\tilde{X}^i\frac{\partial f}{\partial \tilde{u}^i} = \tilde{X}^i\frac{\partial f}{\partial u^j}\frac{\partial u^j}{\partial \tilde{u}^i} = X^j\frac{\partial f}{\partial u^j}$$

so

$$X^j = \tilde{X}^i\frac{\partial u^j}{\partial \tilde{u}^i} \quad \text{and} \quad \tilde{X}^i = X^j\frac{\partial \tilde{u}^i}{\partial u^j}$$

3. THE LIE BRACKET

If we attempt to compose X and Y the results are not encouraging; locally

$$\begin{aligned} Y(f) &= Y^i\frac{\partial f}{\partial u^i} \\ X(Y(f)) &= X^j\frac{\partial}{\partial u^j}\left(Y^i\frac{\partial f}{\partial u^i}\right) \\ &= X^j\frac{\partial Y^i}{\partial u^j}\frac{\partial f}{\partial u^i} + X^jY^i\frac{\partial^2 f}{\partial u^j\partial u^i} \end{aligned}$$

which shows that XY is not a tangent vector since it contains the second derivative of f and is thus not a first order operator. However, we take heart from the observation that the objectionable term is *symmetric* in i and j. Thus if we form

$$Y(X(f)) = Y^j\frac{\partial X^i}{\partial u^j}\frac{\partial f}{\partial u^i} + Y^jX^i\frac{\partial^2 f}{\partial u^j\partial u^i}$$

and subtract, the objectionable terms will drop out and we have

$$X(Y(f)) - Y(X(f)) = \left(X^j\frac{\partial Y^i}{\partial u^j} - Y^j\frac{\partial X^i}{\partial u^j}\right)\frac{\partial f}{\partial u^i}$$

which is a first order operator and hence a tangent vector. We now introduce new notation

$$\begin{aligned} [X, Y] &= X(Y(\cdot)) - Y(X(\cdot)) = \left(X^j\frac{\partial Y^i}{\partial u^j} - Y^j\frac{\partial X^i}{\partial u^j}\right)\frac{\partial}{\partial u^i} \\ [X, Y]^i &= \left(X^j\frac{\partial Y^i}{\partial u^j} - Y^j\frac{\partial X^i}{\partial u^j}\right) \end{aligned}$$

For practise the user may wish to verify that $[X, Y]^i$ transforms properly under coordinate change.

X and Y are said to *commute* if $[X, Y] = 0$. For example $\frac{\partial}{\partial u^i}$ and $\frac{\partial}{\partial u^j}$ commute.

It is obvious that $[X, Y]$ is linear in each variable.

Next we form

$$\begin{aligned}
[X,[Y,Z]](f) &= X((YZ-ZY)(f)) - (YZ-ZY)(X(f)) \\
&= X(Y(Z(f))) - X(Z(Y(f))) - Y(Z(X(f))) + Z(Y(X(f)))
\end{aligned}$$

Let us abbreviate this by surpressing the f and the parentheses and then permute cyclically.

$$\begin{aligned}
[X,[Y,Z]] &= XYZ - XZY - YZX + ZYX \\
[Y,[Z,X]] &= YZX - YXZ - ZXY + XZY \\
[Z,[X,Y]] &= ZXY - ZYX - XYZ + YXZ
\end{aligned}$$

Now if we add the three equations the terms cancel in pairs and we have the important *Jacobi Identity*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

Thus the vector fields in $T(M)$ form an (infinite dimensional) Lie Algebra, which is a vector space with a skew symmetric multiplication $[X,Y]$ linear in each variable satisfying the previous identity. The Jacobi identity substitutes for associativity.

4. BACK AND FORTH WITH DIFFEOMORPHISMS

The user may wonder why this area of mathematics tends to emphasize diffeomorphisms instead of differentiable maps. In fact, the level of generality implied by $\phi : M \to N$ is largely spurious; most of the time $N = M$ and so diffeomorphisms are the natural objects to study. Nevertheless we follow convention in this regard and formulate the results for $\phi : M \to N$

The user will recall that a mapping $\phi : M \to N$ induces a mapping $\phi_* = d\phi$ from $T_p(M)$ to $T_q(N)$ where $q = \phi(p)$. Since a vector in $T_q(N)$ may be identified by its action on a function $f : V \to \mathbb{R}$, V a neighborhood of q in N, we can define ϕ_* by

$$(\phi_* X)_q(f) = X_p(f \circ \phi)$$

since $f \circ \phi : M \to \mathbb{R}$. For notational amusement we define

$$\phi^* f = f \circ \phi$$

and then we have

$$(\phi_* X)(f) = X(\phi^* f)$$

and thus

$$\phi_*(X) = X \circ \phi^*$$

for $X \in T_p(M)$ and $\phi_* : T_p(M) \to T_{\phi(p)}(N)$.

In local coordinates we describe ϕ as follows: p has coodinates $u^1, \ldots, u^n$ and $\phi(p)$ has coodinates $v^1, \ldots, v^n$ and thus $p \to \phi(p)$ is given by

$$v^i(u^1, \ldots, u^n) \qquad i = 1, \ldots, n$$

Then, if

$$X = X^i \frac{\partial}{\partial u^i} \qquad\qquad Y = Y^i \frac{\partial}{\partial v^j}$$

and $Y_{\phi(p)} = \phi_*(X_p)$ we have, (surpressing some p subscripts)

$$Y_{\phi(p)}(f) = Y^j \frac{\partial f}{\partial v^j} = Y^j \frac{\partial f \circ \phi}{\partial u^i} \frac{\partial u^i}{\partial v^j}$$

and

$$Y_{\phi(p)}(f) = X_p(f \circ \phi) = X^i \frac{\partial f \circ \phi}{\partial u^i}$$

from which we see that

$$X^i = Y^j \frac{\partial u^i}{\partial v^j}$$

and symmetrically

$$Y^j = X^i \frac{\partial v^j}{\partial u^i}$$

which resembles the coordinate change rules. This is for the best of reasons; because ϕ is a diffeomorphism a coordinate patch on N becomes, via ϕ, a coordinate patch on M. Thus, locally, a diffeomorphism looks like a coordinate change. This is not particularly helpful in keeping things straight in ones mind, although occasionally technically useful.

Now if ϕ were just a differentiable mapping then it could not be used to map vector fields on M to vector fields on N. The obvious way to do this is as follows: given $q \in N$, select $p \in M$ so that $\phi(p) = q$ and then let $Y_q = \phi_* X_p$. This won't work for two reasons. First, there might be $q \in N$ not in the range of ϕ, and even if q is in the range of ϕ we might have $\phi(p_1) = \phi(p_2) = q$ for $p_1 \neq p_2$ so that Y_q would not be uniquely defined. However, neither of these is a problem for diffeomorphisms so that

Def If X is a vector field on M and $\phi : M \to N$ is a diffeomormorphism then we can define a vector field $Y = \phi_*(X)$ on N by

$$(\phi_* X)_q = X_{\phi^{-1}q} \circ \phi = \phi_*(X_{\phi^{-1}q})$$

If $f : N \to \mathbb{R}$ then (recall $\phi^*(f) = f \circ \phi$)

$$\begin{aligned} (\phi_* X)_q(f) &= \big(X_{\phi^{-1}q} \circ \phi\big)(f) \;=\; X_{\phi^{-1}q}(\phi^* f) \\ &= X_{\phi^{-1}q}(f \circ \phi) \end{aligned}$$

Expressed slightly differently

$$\begin{array}{lrcl} \text{if} & Y &=& \phi_* X \\ \text{then} & Y_q(f) &=& X_{\phi^{-1}q}(f \circ \phi) \end{array}$$

Notice that if $\phi_1 : M_1 \to M_2$ and $\phi_2 : M_2 \to M_3$ are diffeomorphisms then

$$\phi_{2*} \circ \phi_{1*} = (\phi_2 \circ \phi_1)_*$$

This is just an abstract expression of the chain rule, but we can say it really fancy: $*$ is a covariant functor from the category of Manifolds and Diffeomorphisms to the category of vector bundles and isomorphisms.

Now we want to show that the Lie Bracket is preserved under diffeomorphisms. First we note that if $\phi : M \to N$ is a diffeomorphism then it can be regarded locally as a coordinate change. Those persons who verified that the Lie Bracket was invariant under coordinate change when I requested them to do so need not read the following.

Let $\phi : M \to N$ be a diffeomorphism and let $u^1, \ldots, u^n$ be coordinates around $p \in M$ and $v^1, \ldots, v^n$ be coordinates around $q = \phi(p) \in N$. Then we have, for $f : N \to \mathbb{R}$,

$$\begin{aligned} (\phi_* X)_q f &= X_{\phi^{-1}q}(f \circ \phi) \\ &= X^i_{u^i(v^j)} \frac{\partial}{\partial u^i}(f \circ \phi) \\ &= X^i_{u^i(v^j)} \frac{\partial f}{\partial v^i} \frac{\partial v^j}{\partial u^i} \;=\; \tilde{X}^j_{v^k} \frac{\partial f}{\partial v^j} \end{aligned}$$

where we set

$$\tilde{X}^j_{v^k} = X^i_{u^\ell(v^k)} \frac{\partial v^j}{\partial u^i}$$

Now we can calculate the Lie Bracket. We surpress the subscripts on $X, \tilde{X}, Y, \tilde{Y}$ because they are always the same. Then

$$\begin{aligned} [\tilde{X}, \tilde{Y}]^i &= \tilde{X}^j \frac{\partial \tilde{Y}^i}{\partial v^j} - \tilde{Y}^j \frac{\partial \tilde{X}^i}{\partial v^j} \\ &= X^k \frac{\partial v^j}{\partial u^k} \frac{\partial}{\partial u^\ell}\Big(Y^m \frac{\partial v^i}{\partial u^m}\Big) \frac{\partial u^\ell}{\partial v^j} - Y^k \frac{\partial v^j}{\partial u^k} \frac{\partial}{\partial u^\ell}\Big(X^m \frac{\partial v^i}{\partial u^m}\Big) \frac{\partial u^\ell}{\partial v^j} \\ &= X^k \frac{\partial Y^m}{\partial u^\ell} \frac{\partial u^\ell}{\partial v^j} \frac{\partial v^j}{\partial u^k} \frac{\partial v^i}{\partial u^m} + X^k Y^m \frac{\partial v^j}{\partial u^k} \frac{\partial^2 v^i}{\partial u^i \partial u^m} \frac{\partial v^\ell}{\partial u^j} \\ &- Y^k \frac{\partial X^m}{\partial u^\ell} \frac{\partial u^\ell}{\partial v^j} \frac{\partial v^j}{\partial u^k} \frac{\partial v^i}{\partial u^m} - Y^k X^m \frac{\partial v^j}{\partial u^k} \frac{\partial^2 v^i}{\partial u^i \partial u^m} \frac{\partial v^\ell}{\partial u^j} \\ &= X^k \frac{\partial Y^m}{\partial u^\ell} \delta^\ell_k \frac{\partial v^i}{\partial u^m} - Y^k \frac{\partial X^m}{\partial u^\ell} \delta^\ell_k \frac{\partial v^i}{\partial u^m} \\ &+ X^k Y^m \delta^\ell_k \frac{\partial^2 v^i}{\partial u^\ell \partial u^m} - Y^k X^m \delta^\ell_k \frac{\partial^2 v^i}{\partial u^\ell \partial u^m} \\ &= \Big(X^\ell \frac{\partial Y^m}{\partial u^\ell} - Y^\ell \frac{\partial X^m}{\partial u^\ell}\Big) \frac{\partial v^i}{\partial u^m} + X^k Y^m \frac{\partial^2 v^i}{\partial u^k \partial u^m} - Y^k X^m \frac{\partial^2 v^i}{\partial u^k \partial u^m} \\ &= [X, Y]^m \frac{\partial v^i}{\partial u^m} + 0 \\ &= \widetilde{[X, Y]}^i \end{aligned}$$

5. FLOWS

A vector field (section of $T(M)$) gives rise to a flow $\phi_t(p) = \phi(p,t) : M \times I \to M$ where the ϕ_t are diffeomorphisms and the interval I is $(-\epsilon, \epsilon)$, where ϵ may, in general, depend on the value of p. Let $\{u^1, \dots, u^n\}$ be local coordinates and the vector field be given locally by

$$X = X^i \frac{\partial}{\partial u^i}$$

The ϕ_t also have a local expression where $\phi_t(p)$ is given by $\{u^1(t), \dots, u^n(t)\}$ (where p corresponds to $u^i(0) = u^i_0$). Then $\phi(t)$ is determined by differential equations with initial conditions. We denote the $u^i(t)$ with initial conditions $u^i(0) = u^i_0$) by $u^i(u^k_0, t)$. The differential equations are:

$$\frac{du^k(u^i_0, t)}{dt} = X^k(u^j(u^i_0, t)) \qquad\qquad u^k(u^i_0, 0) = u^k_0$$

The u^i_0 function as parameters in these equations. The flowline begins with the point p that has these coordinates and flows out of p to points with the coordinates $u^j(u^i_0, t)$. The Picard-Lindelöf theorem guarantees that the solutions will be as differentiable as the input data and depend as differentiably on the parameters as the X^i do. Solutions will exist for some interval $(-\epsilon, \epsilon)$ where $\epsilon > 0$. However, in general the ϵ will depend on the intial point p. This won't do us any harm. Uniqueness of the solutions guarantees that

$$\phi_{t_1}(\phi_{t_2}(p)) = \phi_{t_1+t_2}(p)$$

as long as $t_1, t_2, t_1 + t_2$ remain within $(-\epsilon, \epsilon)$. If M is compact we can say much more; ϕ_t is defined for all $t \in \mathbb{R}$. However, for Lie purposes the local existence is sufficient, so we will not pursue the matter here.

6. LIE DERIVATIVES

Here is the idea of the Lie Derivative. Given a tangent vector field X on M, the derivative in the X direction of an object is the rate of change along the flowline $\phi_t(p)$ if this makes sense. Unfortunately, it only makes sense for functions. To find derivatives of objects that live in bundles we must change the game slightly, since objects in neighboring fibres cannot be directly compared. Let E be a bundle over M. To compare an object W_q in the fibre E_q over q to an object W_p in the fibre E_p over p we need an isomorphism from E_q to E_p. This is not generally available, but if q is on the flowline out of p determined by the vector field X then we may be able to find such an isomorphism and this suffices to define the Lie Derivative. Indeed if $q = \phi_t(p)$ then $p = \phi_t^{-1}(q) = \phi_{-t}(q)$ and if we can find $(\phi_{-t})_* : \pi^{-1}[q] \to \pi^{-1}[p]$ then we can compare $(\phi_{-t})_* W_q$ with W_p. In fact

$$(\phi_{-t})_* W_{\phi_t(p)} \in \pi^{-1}[p]$$

is, for small t, a *curve* in $\pi^{-1}[p]$ and hence can be differentiated at $t=0$, yielding an object in $T_p(\pi^{-1}[p]) = T_pW_p$. This is the Lie Derivative for the bundle E.

The importance of understanding this construction is easy to underestimate. If we understand it we understand a) where the Lie Derivative lives (in the tangent space to the fibre of the bundle). If the fibre is a vector space it is possible to identify the tangent space with the fibre itself, and this is often done. b) Knowing that we are dealing with a *curve* in the fibre W_p focuses our attention properly when we are doing the technical calculations. c) it prepares us for the idea of connections in vector bundles and principal fibre bundles where the ideas are somewhat similar.

Naturally for different bundles $(\phi_{-t})_*$ will have a different form, but this is just technical stuff; the basic idea is given above.

First we will deal with the Lie Derivative of a function. In this case $(\phi_{-t})_* =$ Identity and $(\phi_{-t})_*f = f$. Hence

$$\begin{aligned}
\pounds_X(f) &= \frac{d}{dt}(f\circ\phi_t)\big|_{t=0} \\
&= \frac{d}{dt}(f(u^i(t)))\big|_{t=0} \\
&= \frac{\partial f}{\partial u^i}\frac{du^i}{dt}\Big|_{t=0} \\
&= \frac{\partial f}{\partial u^i}X^i \\
&= X(f)
\end{aligned}$$

Our next project is the Lie Derivative of a Tangent Vector Field $Y \in T(M)$. X is still the vector field with flow $\phi_t(p)$ with $\frac{d}{dt}\phi_t|_{t=0} = X$. Recall that $\phi_{-t} = \phi_t^{-1}$. The isomorphism is

$$(\phi_{-t})_* : T_{\phi_t(p)} \to T_p(M)$$

where $p = \phi_0(p)$ and $q = \phi_t(p)$. The rest is mere intricate calculation. Let p have coordinates $u^1(0), \ldots, u^n(0)$ and q have coordinates $u^1(t), \ldots, u^n(t)$. For ease of presentation we will use $v^i = u^i(t)$. Then $(\phi_t)_* : T_p(M) \to T_q(M)$ is given in coordinates, with $Y_q = (\phi_t)_* Z_p$, by

$$Y_q^i = \frac{\partial v^i}{\partial u^j} Z_p^j$$

and then $(\phi_{-t})_* = (\phi_t^{-1})_*$ is given by

$$\left(\frac{\partial v^i}{\partial u^j}\right)^{-1} \begin{pmatrix} Y_q^1 \\ \vdots \\ Y_q^n \end{pmatrix} = \begin{pmatrix} Z_p^1 \\ \vdots \\ Z_p^n \end{pmatrix} \in T_p(M)$$

Notice that

$$\left(\frac{\partial v^i}{\partial u^j}\right)\Big|_{t=0} = \left(\frac{\partial u^i(t)}{\partial u^j}\right)\Big|_{t=0} = \text{ Id}$$

We are going to need

$$
\begin{aligned}
\frac{d}{dt}\left(\frac{\partial v^i}{\partial u^j}\right)^{-1}\Big|_{t=0} &= (-1)\left(\frac{\partial v^i}{\partial u^j}\right)^{-2}\frac{d}{dt}\left(\frac{\partial v^i}{\partial u^j}\right)\Big|_{t=0} \\
&= (-1)\left(\frac{\partial v^i}{\partial u^j}\right)^{-2}\left(\frac{\partial}{\partial u^j}\frac{d}{dt}u^i(t)\right)\Big|_{t=0} \\
&= (-1)(\mathrm{Id})\left(\frac{\partial}{\partial u^j}X^i\right) \\
&= -\left(\frac{\partial X^i}{\partial u^j}\right)
\end{aligned}
$$

Now we have, for $Y_q \in T_q(M)$, or more precisely

$$
\begin{aligned}
Y_{\phi_t(p)} &\in T_{\phi_t(p)}(M) \\
\begin{pmatrix} Z_p^1 \\ \vdots \\ Z_p^n \end{pmatrix} &= \left(\frac{\partial v^i}{\partial u^j}\right)^{-1}\begin{pmatrix} Y_{\phi_t(p)}^1 \\ \vdots \\ Y_{\phi_t(p)}^n \end{pmatrix}
\end{aligned}
$$

a curve in the fibre $T_p(M)$ whose derivative at $t = 0$ is the Lie Derivative $\pounds_X(Y) \in T_{p,0}(T_p(M))$. We now compute

$$
\begin{aligned}
\pounds_X(Y) &= \frac{d}{dt}\begin{pmatrix} Z_p^1 \\ \vdots \\ Z_p^n \end{pmatrix}\Bigg|_{t=0} \\
&= \left[\frac{d}{dt}\left(\left(\frac{\partial v^i}{\partial u^j}\right)^{-1}\right)\begin{pmatrix} Y_{\phi_t(p)}^1 \\ \vdots \\ Y_{\phi_t(p)}^n \end{pmatrix} + \left(\frac{\partial v^i}{\partial u^j}\right)^{-1}\frac{d}{dt}\begin{pmatrix} Y_{\phi_t(p)}^1 \\ \vdots \\ Y_{\phi_t(p)}^n \end{pmatrix}\right]_{t=0} \\
&= -\left(\frac{\partial X^i}{\partial u^j}\right)\begin{pmatrix} Y_p^1 \\ \vdots \\ Y_p^n \end{pmatrix} + \left[\left(\frac{\partial v^i}{\partial u^j}\right)^{-1}\frac{\partial}{\partial v^j}\begin{pmatrix} Y_{\phi_t(p)}^1 \\ \vdots \\ Y_{\phi_t(p)}^n \end{pmatrix}\frac{dv^j}{dt}\right]_{t=0} \\
&= -\left(\frac{\partial X^i}{\partial u^j}\right)\begin{pmatrix} Y_p^1 \\ \vdots \\ Y_p^n \end{pmatrix} + \mathrm{Id}\,\frac{\partial}{\partial u^j}\begin{pmatrix} Y_p^1 \\ \vdots \\ Y_p^n \end{pmatrix}X^j
\end{aligned}
$$

If the coordinates of $\pounds_X(Y)$ in the local basis are denoted by $\pounds_X(Y)^i$ this gives

$$
\pounds_X(Y)^i = X^j\frac{\partial}{\partial u^j}Y^i - Y^j\frac{\partial}{\partial u^j}X^i = [X,Y]^i
$$

and we have proved

$$
\pounds_X(Y) = [X,Y]
$$

We note that $\pounds_X(Y)$ has two terms because it is the derivative of $(\phi_{-t})_* Y_{\phi_t(p)}$ both factors of which depend on t, and the origin of the negative sign lies in the inverse: $\phi_{-t} = \phi_t^{-1}$. We also note that the above derivation is highly dependent on the continuity of the second derivatives.

Next we want to derive a formula for the Lie Derivative of the Lie Bracket of two Vector Fields. This is immediate from the Jacobi Identity. Indeed recall

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

from which we get

$$\begin{aligned} [X,[Y,Z]] &= [Y,[X,Z]] - [Z,[X,Y]] \\ \pounds_X([Y,Z]) &= \pounds_Y([X,Z]) - \pounds_Z([X,Y]) \\ \pounds_X([Y,Z]) &= \pounds_Y \pounds_X(Z) - \pounds_Z \pounds_X(Y) \end{aligned}$$

In contrast to the preceding the calculation of the Lie Derivative of a 1-form is somewhat easier as the formulation of $(\phi_{-t})_*$ is a bit easier. It is simply the familiar pullback of differential forms. If $\phi_t(p) = q$ then

$$\phi_t^* : T_q^*(M) \to T_p^*(M)$$

is the pullback and we set, for this bundle, $(\phi_{-t})_* = \phi_t^*$. (Remember, to get a Lie Derivative on a bundle we must have an isomorphism $(\phi_{-t})_*$ from $\pi^{-1}[q]$ to $\pi^{-1}[p]$. We hope to find such an isomorphism which makes sense for the bundle. In the sequel we will also have to worry about whether the Lie Derivatives for the various bundles fit together well. For $T^*(M)$ the above turns out to be a good choice, as well as the only obvious choice.) Let's now recall how ϕ^* works. If $\phi : M \to N$ and $u^1, \ldots, u^n$ are coordinates on M and $v^1, \ldots, v^n$ are coordinates on N and ϕ is given locally by $v^i(u^1, \ldots, u^n)$ and a section ω is given locally on N by $\omega_i dv^i$ then $\tilde{\omega} = \phi^*(\omega)$ is given locally on M by

$$\tilde{\omega}_i(u^j)\, du^i = \omega_j(v^k(u^j)) \frac{\partial v^j}{\partial u^i}\, du^i$$

In our case ϕ_t is expressed by $v^i = v^i(u^j)$, and we know that

$$\frac{dv^k}{dt}\Big|_{t=0} = X^k$$

since

$$\frac{d\phi_t}{dt}\Big|_{t=0} = X$$

Thus

$$\begin{aligned} \pounds_X(\omega) &= \frac{d}{dt}\left[(\phi_{-t})_* \omega(\phi_t(p))\right]_{t=0} \\ &= \frac{d}{dt}\left[(\phi_t)^* \omega(\phi_t(p))\right]_{t=0} \end{aligned}$$

$$
\begin{aligned}
&= \frac{d}{dt}\left[\omega_i \frac{\partial v^i}{\partial u^j} du^j\right]_{t=0} \\
&= \left[\frac{\partial \omega_i}{\partial u^k}\frac{du^k}{dt}\frac{\partial v^i}{\partial u^j}\, du^j + \omega_i\, \frac{d}{dt}\frac{\partial v^i}{\partial u^j}\, du^j\right]_{t=0} \\
&= \frac{\partial \omega_i}{\partial u^k} X^k \delta^i_j\, du^j + \left[\omega_i \frac{\partial}{\partial u^j}\frac{dv^i}{dt}\, du^j\right]_{t=0} \\
&= \frac{\partial \omega_i}{\partial u^k} X^k\, du^i + \omega_i \frac{\partial}{\partial u^j} X^i\, du^j \\
&= \left[\frac{\partial \omega_i}{\partial u^j} X^j + \omega_j \frac{\partial X^j}{\partial u^i}\right]\, du^i \\
&= X(\omega_i)\, du^i + \omega_i\, dX^i
\end{aligned}
$$

This last is an interesting formula and we will find use for it from time to time.

The following calculations are not strictly necessary; they can be derived abstractly from the fact that $(\phi_t)^*$ and d commute. However, the formulas may come in handy and and there is (for me) a certain amusement in the formulas. We first derive

$$\pounds_X(df) = d\pounds_X(f)$$

by computing both sides.

$$
\begin{aligned}
df &= \frac{\partial f}{\partial u^j}\, du^j \\
\pounds_X(df) &= \left[\frac{\partial}{\partial u^i}\frac{\partial f}{\partial u^j} X^j + \frac{\partial f}{\partial u^j}\frac{\partial X^j}{\partial u^i}\right]\, du^i \\
&= \left[\frac{\partial^2 f}{\partial u^j \partial u^i} X^j + \frac{\partial f}{\partial u^j}\frac{\partial X^j}{\partial u^i}\right]\, du^i \\
d\pounds_X(f) &= d(Xf) \,=\, d\left(\frac{\partial f}{\partial u^j} X^j\right) \\
&= \left[\frac{\partial^2 f}{\partial u^i \partial u^j} X^j + \frac{\partial f}{\partial u^j}\frac{\partial X^j}{\partial u^i}\right]\, du^i
\end{aligned}
$$

Next I want

$$
\begin{aligned}
d\pounds_X(u^k) &= d(X(u^k)) \,=\, d\left(X^j \frac{\partial u^k}{\partial u^j}\right) \\
&= d(X^j \delta^k_j) \,=\, dX^k
\end{aligned}
$$

and then

$$
\begin{aligned}
\pounds_X(du^k) &= d\pounds_X(u^k) \\
&= dX^k
\end{aligned}
$$

Now we can verify that Leibniz' rule works here; for $\omega \in \Lambda^1$

$$\pounds_X(\omega) = \pounds_X(\omega_i\, du^i)$$

$$
\begin{aligned}
&= \pounds_X(\omega_i)\,du^i + \omega_i \pounds_X(du^i) \\
&= X(\omega_i)\,du^i + \omega_i\,dX^i
\end{aligned}
$$

which coincides with our previous formula. We can read this in two ways; if we have Leibniz' formula for products then this says the general formula for $\pounds_X(\omega)$ is derivable from the formulas for $\pounds_X(f)$ and $\pounds_X(du^k)$, or we can view it as a justification for Leibniz's formula given our previous work.

We can sweat a little more out of the formula

$$\pounds_X(\omega) = X(\omega_i)\,du^i + \omega_i\,dX^i$$

by applying it to a vector field Y. We get

$$
\begin{aligned}
\pounds_X(\omega)(Y) &= X(\omega_i)\,du^i(Y) + \omega_i\,dX^i(Y) \\
&= X(\omega_i)Y^i + \omega_i Y(X^i) \\
&= X(\omega_i Y^i) - \omega_i X(Y^i) + \omega_i Y(X^i) \\
&= X(\omega(Y)) - \omega_i [X,Y]^i \\
&= X(\omega(Y)) - \omega([X,Y]) \\
&= \pounds_X(\omega(Y)) - \omega([X,Y])
\end{aligned}
$$

This is often written as

$$\pounds_X(\omega(Y)) - \pounds_X(\omega)(Y) = \omega([X,Y])$$

This is a handy formula and we will meet it again in another context.

7. THE $d\omega(X_0,\ldots,X_r)$ FORMULA

In this section we wish to derive a formula which shows how $d\omega$ acts on $T(M)$. Specifically, for ω in $\Lambda^r(M)$, we want the formula

$$
\begin{aligned}
d\omega(X_0, X_1, \cdots X_r) &= \sum_{i=0}^{r} (-1)^i X^i(\omega(X_0, \cdots, X_{i-1}, X_{i+1}, \cdots, X_r) \\
&\quad + \sum_{0\le i<j\le r} (-1)^{i+j} \omega([X_i, X_j], \cdots, X_{i-1}, X_{i+1}, \cdots, X_{j-1}, X_{j+1}, \cdots, X_r)
\end{aligned}
$$

This way of expressing the formula is a little cumbersome so we will make of use of the convention that a hat on a letter means the letter is *omitted*. The formula then reads

$$
\begin{aligned}
d\omega(X_0, X_1, \cdots X_r) &= \sum_{i=0}^{r} (-1)^i X^i(\omega(X_0, \cdots, \hat{X}_i, \cdots, X_r) \\
&\quad + \sum_{0\le i<j\le r} (-1)^{i+j} \omega([X_i, X_j], \cdots, \hat{X}_i, \cdots, \hat{X}_j, \cdots, X_r)
\end{aligned}
$$

This interesting formula explicitly gives the value of $d\omega$ evaluated on vectors. It is even possible to use it to *define* $d\omega$, as Godbillion[2], does instead of determining $d\omega$ by the three rules

$$
\begin{aligned}
df &= \frac{\partial f}{\partial u^i}\, du^i \\
d(\omega^1 \wedge \omega^2) &= d\omega^1 \wedge \omega^2 + (-1)^{\deg(\omega_1)} \omega^1 \wedge d\omega^2 \\
dd\omega &= 0
\end{aligned}
$$

However, we will follow the classical route which requires us to *prove* this formula.

It is worth doing the special case of a one-form first to get the feel for the situation. We have $\omega \in \Lambda^1(M)$ and X_0 and X_1 are two vector fields on M. Then locally

$$
\begin{aligned}
\omega &= \omega_i\, du^i \\
d\omega(X_0, X_1) &= \frac{\partial \omega_i}{\partial u^j}\, du^j \wedge du^i(X_0, X_1) \\
&= \frac{\partial \omega_i}{\partial u^j} \Big(du^j(X_0) du^i(X_1) - du^i(X_0) du^j(X_1) \Big) \\
&= \frac{\partial \omega_i}{\partial u^j} \Big(X_0^j X_1^i - X_0^i X_1^j \Big) \\
&= X_0^j \frac{\partial \omega_i}{\partial u^j} X_1^i - X_1^j \frac{\partial \omega_i}{\partial u^j} X_0^i \\
&= X_0^j \frac{\partial}{\partial u^j} \Big(\omega_i X_1^i \Big) - X_0^j \omega_i \frac{\partial X_1^i}{\partial u^j} \\
&\quad - X_1^j \frac{\partial}{\partial u^j} \Big(\omega_i X_0^i \Big) + X_1^j \omega_i \frac{\partial X_0^i}{\partial u^j} \\
&= X_0(\omega(X_1) - X_1(\omega(X_0) - \omega_i \Big(X_0^j \frac{\partial X_1^i}{\partial u^j} - X_1^j \frac{\partial X_0^i}{\partial u^j} \Big) \\
&= X_0(\omega(X_1) - X_1(\omega(X_0) - \omega_i \Big([X_0, X_1]^i \Big) \\
&= X_0(\omega(X_1) - X_1(\omega(X_0) - \omega \Big([X_0, X_1] \Big)
\end{aligned}
$$

This verifies the formula for the case $r = 1$. Unfortunately for $r > 1$ the situation is not so simple. As Helgason[3] remarks, this formula is essentially trivial but it is a bit tricky to push through in detail. In fact, it's not too easy to find a source where this is done. The steps are, in the big picture, the same as those for $r = 1$ but there are so many things to keep straight that the proof is a bit unwieldy.

Before starting we remind the reader of the formula for diffentiating a determinant. It is, for $r = 3$, simply

$$
\frac{\partial}{\partial u} \begin{vmatrix} f & g & h \\ i & j & k \\ l & m & n \end{vmatrix} = \begin{vmatrix} \frac{\partial f}{\partial u} & g & h \\ \frac{\partial i}{\partial u} & j & k \\ \frac{\partial l}{\partial u} & m & n \end{vmatrix} + \begin{vmatrix} f & \frac{\partial g}{\partial u} & h \\ i & \frac{\partial j}{\partial u} & k \\ l & \frac{\partial m}{\partial u} & n \end{vmatrix} + \begin{vmatrix} f & g & \frac{\partial h}{\partial u} \\ i & j & \frac{\partial k}{\partial u} \\ l & m & \frac{\partial n}{\partial u} \end{vmatrix}
$$

as is immediate from the formula

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} = \sum_{\pi\in\mathcal{S}_3} \operatorname{sgn}(\pi)\, a_1^{\pi(1)} a_2^{\pi(2)} a_3^{\pi(3)}$$

By linearity the preceding formula gives, for a vector field X,

$$X\left(\begin{vmatrix} f & g & h \\ i & j & k \\ l & m & n \end{vmatrix}\right) = \begin{vmatrix} X(f) & g & h \\ X(i) & j & k \\ X(l) & m & n \end{vmatrix} + \begin{vmatrix} f & X(g) & h \\ i & X(j) & k \\ l & X(m) & n \end{vmatrix} + \begin{vmatrix} f & g & X(h) \\ i & j & X(k) \\ l & m & X(n) \end{vmatrix}$$

We also remind the reader that the hat on an expression means that it is omitted, so that for example

$$\begin{vmatrix} f & g & \hat{h} & i \\ j & k & \hat{l} & m \\ n & o & \hat{p} & q \end{vmatrix} = \begin{vmatrix} f & g & i \\ j & k & m \\ n & o & q \end{vmatrix}$$

We are now going to prove the general formula

$$\begin{aligned} d\omega(X_0, X_1, \cdots X_r) &= \sum_{i=0}^{r} (-1)^i X^i(\omega(X_0, \cdots, \hat{X}_i, \cdots, X_r) \\ &\quad + \sum_{0\le i<j\le r} (-1)^{i+j} \omega([X_i, X_j], X_0, \cdots, \hat{X}_i, \cdots, \hat{X}_j, \cdots, X_r) \end{aligned}$$

By linearity, it suffices to prove the formula for

$$\omega = f\, du^{i_1} \wedge \ldots \wedge du^{i_r}$$

We recall that, with $X_k = X_k^j \frac{\partial}{\partial u^j}$, that

$$du^{i_1} \wedge \ldots \wedge du^{i_r}(X_1, \ldots, X_r) = \det |du^{i_j}(X_k)|$$

Thus we are trying to prove that

$$\begin{aligned} d\omega(X_0, \ldots, X_r) &= \sum_{i=0}^{r} (-1)^i X_i\big(f du^{i_1} \wedge \ldots \wedge du^{i_r}(X_0, \ldots, \hat{X}_i, \ldots, X_r)\big) \\ &\quad + \sum_{0\le i<j\le r} (-1)^{i+j} f\, du^{i_1} \wedge \ldots \wedge du^{i_r}([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_r) \end{aligned}$$

This is in some sense straightforward, and there is only one substantive step; everything else is cosmetic rearrangement. I will point out the substantive step when we get there. For ease of reading I will leave out the wedges in the calculation. Here we go....

$$d\omega(X_0, \ldots, X_r) = \sum_{j=1}^{r} \frac{\partial f}{\partial u^j}\, du^j du^{i_1} \ldots du^{i_r}(X_0, X_1, \ldots, X_r)$$

$$= \sum_j \frac{\partial f}{\partial u^j} \begin{vmatrix} du^j(X_0) & du^j(X_1) & \dots & du^j(X_r) \\ du^{i_1}(X_0) & du^{i_1}(X_1) & \dots & du^{i_1}(X_r) \\ \dots & \dots & \dots & \dots \\ du^{i_r}(X_0) & du^{i_r}(X_1) & \dots & du^{i_r}(X_r) \end{vmatrix}$$

$$= \sum_j \frac{\partial f}{\partial u^j} \begin{vmatrix} X_0^j & X_1^j & \dots & X_r^j \\ X_0^{i_1} & X_1^{i_1} & \dots & X_r^{i_1} \\ \dots & \dots & \dots & \dots \\ X_0^{i_r} & X_1^{i_r} & \dots & X_r^{i_r} \end{vmatrix}$$

Now we do an expansion off the i_{th} column to get

$$= \sum_j \sum_{i=0}^r (-1)^i \frac{\partial f}{\partial u^j} X_i^j \begin{vmatrix} X_0^{i_1} & \dots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \\ X_0^{i_r} & \dots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix}$$

$$= \sum_{i=0}^r (-1)^i X_i(f) \begin{vmatrix} X_0^{i_1} & \dots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \\ X_0^{i_r} & \dots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix}$$

Now comes the one substantive step; we use $X(f)g = X(fg) - fX(g)$:

$$= \sum_{i=0}^r (-1)^i X_i \left(f \begin{vmatrix} X_0^{i_1} & \dots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots \\ X_0^{i_r} & \dots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix} \right)$$

$$- \sum_{i=0}^r (-1)^i f X_i \left(\begin{vmatrix} X_0^{i_1} & \dots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots \\ X_0^{i_r} & \dots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix} \right)$$

$$= \sum_{i=0}^r (-1)^i X_i \left(f \begin{vmatrix} du^{i_1}(X_0) & \dots & \widehat{du^{i_1}(X_i)} & \dots & du^{i_1}(X_r) \\ du^{i_2}(X_0) & \dots & \widehat{du^{i_2}(X_i)} & \dots & du^{i_2}(X_r) \\ \dots & \dots & \dots & \dots & \dots \\ du^{i_r}(X_0) & \dots & \widehat{du^{i_r}(X_i)} & \dots & du^{i_r}(X_r) \end{vmatrix} \right)$$

$$- \sum_{i=0}^r \sum_{k \neq i} (-1)^i f \begin{vmatrix} X_0^{i_1} & \dots & X_i(X_k^{i_1}) & \cdots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & X_i(X_k^{i_2}) & \cdots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_0^{i_r} & \dots & X_i(X_k^{i_r}) & \cdots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix}$$

$$= \sum_{i=0}^r (-1)^i X_i \left(f du^{i_1} \cdots du^{i_r}(X_0, \dots, \hat{X}_i, \dots X_r) \right)$$

$$
\begin{aligned}
- & \sum_{i=0}^{r} \sum_{k<i} (-1)^i \, f \begin{vmatrix} X_0^{i_1} & \dots & X_i(X_k^{i_1}) & \cdots & \hat{X}_i^{i_1} & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & X_i(X_k^{i_2}) & \cdots & \hat{X}_i^{i_2} & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_0^{i_r} & \dots & X_i(X_k^{i_r}) & \cdots & \hat{X}_i^{i_r} & \dots & X_r^{i_r} \end{vmatrix} \\
- & \sum_{i=0}^{r} \sum_{k>i} (-1)^i \, f \begin{vmatrix} X_0^{i_1} & \dots & \hat{X}_i^{i_1} & \dots & X_i(X_k^{i_1}) & \dots & X_r^{i_1} \\ X_0^{i_2} & \dots & \hat{X}_i^{i_2} & \dots & X_i(X_k^{i_2}) & \dots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & & & \\ X_0^{i_r} & \dots & \hat{X}_i^{i_r} & \dots & X_i(X_k^{i_r}) & \dots & X_r^{i_r} \end{vmatrix} \\
= & \sum_{i=0}^{r} (-1)^i X_i(\omega(X_0, \dots, \hat{X}_i, \dots, X_r) \\
- & \sum_{i=0}^{r} \sum_{k=0}^{i-1} (-1)^i \, f \, (-1)^k \begin{vmatrix} X_i(X_k^{i_1}) & X_0^{i_1} & \cdots & \hat{X}_k^{i_1} & \cdots & \hat{X}_i^{i_1} & \cdots & X_r^{i_1} \\ X_i(X_k^{i_2}) & X_0^{i_2} & \cdots & \hat{X}_k^{i_2} & \cdots & \hat{X}_i^{i_2} & \cdots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_i(X_k^{i_r}) & X_0^{i_r} & \cdots & \hat{X}_k^{i_r} & \cdots & \hat{X}_i^{i_r} & \cdots & X_r^{i_r} \end{vmatrix} \\
- & \sum_{i=0}^{r} \sum_{k=i+1}^{r} (-1)^i \, f \, (-1)^{k-1} \begin{vmatrix} X_i(X_k^{i_1}) & X_0^{i_1} & \cdots & \hat{X}_i^{i_1} & \cdots & \hat{X}_k^{i_1} & \cdots & X_r^{i_1} \\ X_i(X_k^{i_2}) & X_0^{i_2} & \cdots & \hat{X}_i^{i_2} & \cdots & \hat{X}_k^{i_2} & \cdots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_i(X_k^{i_r}) & X_0^{i_r} & \cdots & \hat{X}_i^{i_r} & \cdots & \hat{X}_k^{i_r} & \cdots & X_r^{i_r} \end{vmatrix} \\
= & \sum_{i=0}^{r} (-1)^i X_i(\omega(X_0, \dots, \hat{X}_i, \dots, X_r) \\
- & \sum_{0 \le i < k \le r} (-1)^{i+k} \, f \begin{vmatrix} X_i(X_k^{i_1}) - X_k(X_i^{i_1}) & X_0^{i_1} & \cdots & \hat{X}_i^{i_1} & \cdots & \hat{X}_k^{i_1} & \cdots & X_r^{i_1} \\ X_i(X_k^{i_2}) - X_k(X_i^{i_2}) & X_0^{i_2} & \cdots & \hat{X}_i^{i_2} & \cdots & \hat{X}_k^{i_2} & \cdots & X_r^{i_2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_i(X_k^{i_r}) - X_k(X_i^{i_r}) & X_0^{i_r} & \cdots & \hat{X}_i^{i_r} & \cdots & \hat{X}_k^{i_r} & \cdots & X_r^{i_r} \end{vmatrix}
\end{aligned}
$$

$$
\begin{aligned}
= & \sum_{i=0}^{r} (-1)^i X_i(\omega(X_0, \dots, \hat{X}_i, \dots, X_r) \\
- & \sum_{0 \le i < k \le r} (-1)^{i+k} \, f \begin{vmatrix} du^{i_1}([X_i, X_k]) & du^{i_1}(X_0) & \cdots & \widehat{du^{i_1}(X_i)} & \cdots & \widehat{du^{i_1}(X_k)} & \cdots & du^{i_1}(X_r) \\ du^{i_2}([X_i, X_k]) & du^{i_2}(X_0) & \cdots & \widehat{du^{i_2}(X_i)} & \cdots & \widehat{du^{i_2}(X_k)} & \cdots & du^{i_2}(X_r) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ du^{i_r}([X_i, X_k]) & du^{i_r}(X_0) & \cdots & \widehat{du^{i_r}(X_i)} & \cdots & \widehat{du^{i_r}(X_k)} & \cdots & du^{i_r}(X_r) \end{vmatrix}
\end{aligned}
$$

$$
\begin{aligned}
= & \sum_{i=0}^{r} (-1)^i X_i(\omega(X_0, \dots, \hat{X}_i, \dots, X_r) \\
- & \sum_{0 \le i < k \le r} (-1)^{i+k} \, f \, du^{i_1} du^{i_2} \cdots du^{i_r} \left([X_i, X_k], X_0, \dots, \hat{X}_i, \dots, \hat{X}_k \dots, X_r\right) \\
= & \sum_{i=0}^{r} (-1)^i X_i(\omega(X_0, \dots, \hat{X}_i, \dots, X_r)
\end{aligned}
$$

$$- \sum_{0 \le i < k \le r} (-1)^{i+k}\, \omega([X_i, X_k], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_k \ldots, X_r)$$

and the proof is complete.

8. SOME ALGEBRAIC IDEAS

The calculation in the previous section might be regarded by some people as unpleasant. In the next section we will develop what we will call the French Method for proving identities. This will give us a less computational mode of attack. However, to do this efficiently we want to bring on board some algebraic ideas which will clarify the structure of the theory.

We suppose that we have an Algebra A (i.e. a vector space with an associative multiplication respecting the vector space operations) which is *GRADED* in the sense that

$$\begin{aligned} A &= \bigoplus_{k=0}^{\infty} A^k \\ A^i \cdot A^j &\subseteq A^{i+j} \end{aligned}$$

Elements of A^i are called *homogeneous of degree i*. A^0 and $\bigoplus_{k=0}^{\infty} A^{2k}$ form natural subalgebras of the graded algebra A. A paradigmatic example of this concept is the Exterior Algebra $\Lambda(M) = \bigoplus_{k=0}^{\infty} \Lambda^k(M)$

In recent years the term *communtative* has come to be applied to graded algebras that satisify the equation

$$x_i x_j = (-1)^{ij}\, x_j x_i \qquad\qquad x_i \in A_i, \quad x_j \in A_j$$

However, to prevent confusion we will refer to this as a *Graded Commuatitive* Algebra. There is a commuatator that goes with this situation given by

$$[x_i, x_j] = x_i x_j - (-1)^{ij}\, x_j x_i$$

and of course the algebra is graded communtative if and only if the commutator vanishes.

The case of interest to us is when the graded algebra A is generated by the elements of degrees 0 and 1. In the case of $A = \Lambda(M)$ this would be the functions and $1-$forms. Abstractly we say

$$A \text{ is generated by } A_0 \oplus A_1$$

Next we introduce the concept of a *Graded Derivation of degree p*. We will temporarily use the letters a and b for graded derivations. A function $a : A \to A$ is a graded derivation if for each r it satisfies

$$\begin{aligned} a : &\quad A_r \to A_{r+p} \\ a(xy) &= a(x)y + (-1)^{rp}\, x a(y) \qquad \text{for } x \in A_r \end{aligned}$$

We also define a bracket $[a,b]$ of two graded derivations of degrees p and q respectively by

$$[a,b] = ab - (-1)^{pq}\, ba$$

The theorem is then

Theorem If a and b are graded derivations of degrees p and q respectively then $[a,b]$ is a graded derivation of degree $p+q$.

Proof The proof is straightforward if a trifle intricate. Recall that

$$\begin{aligned}
a &: A_r \to A_{r+p} \\
b &: A_r \to A_{r+q} \\
a(xy) &= a(x)y + (-1)^{rp} x a(y) \qquad x \in A_r \\
b(xy) &= b(x)y + (-1)^{rq} x b(y) \qquad x \in A_r
\end{aligned}$$

By the definition of bracket for graded derivation we have

$$[a,b] = ab - (-1)^{pq} ba$$

Hence

$$\begin{aligned}
[a,b](xy) &= (ab)(xy) - (-1)^{pq}(ba)(xy) \\
&= a\big[(bx)y + (-1)^{rq}x(by)\big] - (-1)^{pq} b\big[(ax)y + (-1)^{rp}x(ay)\big] \\
&= (abx)y + (-1)^{p(r+q)}(bx)(ay) + (-1)^{rq}(ax)(by) + (-1)^{rp}x(aby) \\
-(-1)^{pq}&\big[(bax)y + (-1)^{q(r+p)}(ax)(by) + (-1)^{rp}(bx)(ay) + (-1)^{rp}(-1)^{rq}x(bay)\big] \\
&= \big((ab - (-1)^{pq}ba)x\big)y + (-1)^{r(p+q)}x\big((ab - (-1)^{pq}ba)y\big) \\
&+\big((-1)^{p(r+q)} - (-1)^{pq+rp}\big)(bx)(ay) + \big((-1)^{rq} - (-1)^{pq+qr+qp}\big)(ax)(by) \\
&= \big(([a,b](x)\big)y + (-1)^{(p+q)r} x[a,b](y)
\end{aligned}$$

because the 3^{rd} and 4^{th} terms in the penultimate equation are 0 due to cancellation of the powers of -1 in the coefficients.

The popular terminology for derivations is as follows.
If p is *even* a graded derivation is called a *derivation*
If p is *odd* a graded derivation is called an *antiderivation*

Let a_1 and a_2 be antiderivations with odd degrees p_1 and p_2
Let b_1 and b_2 be derivations with even degrees q_1 and q_2

Then our result about brackets of graded derivations tell us that

$$[a_1, a_2] = a_1 a_2 + a_2 a_1 \text{is a derivation of degree } p_1 + p_2$$

$$[a_1, a_1] = 2a_1^2 \text{is a derivation of degree } 2p_1$$

$$[a_1, b_1] = a_1 b_1 - b_1 a_1 \text{is an antiderivation of degree } p_1 + q_1$$

$$[b_1, b_2] = b_1 b_2 - b_2 b_1 \text{is a derivation of degree } q_1 + q_2$$

The following theorem is almost obvious.

Theorem Suppose that the graded algebra $A = \bigoplus A_i$ is generated by $A_0 \oplus A_1$ as an algebra and suppose a and b are graded derivations of degree p and suppose that they agree on $A_0 \oplus A_1$ (that is, on scalars and "vectors"). Then they agree on the whole graded algebra $\bigoplus A_i$.

Proof Since the elements of A_i are generated by products of vectors from A_1 and scalars from A_0, the value of a on an element of A_i is determined by the values on A_0 and A_1. Since a and b coincide on A_0 and A_1, they coincide on the entire algebra.

Now the obvious next question is whether one can create a graded derivation on an algebra generated by $A_0 \oplus A_1$ by specifying it it on the elements of A_0 and A_1 and requiring it to satisfy the graded derivation law. The answer, in general, is NO because an element of A_i may be a product of elements of A_0 and A_1 in many different ways, and the different ways may lead to inconsistent values of the graded derivation on that element. Algebraically this means that the graded algebra has *relations*.

But suppose the algebra A is a FREE graded Algebra, which means there are no relations between the elements except those that are forced by being a graded algebra. Then indeed the construction of the derivation is possible. An example of such a free graded algebra is the Exterior Algebra of Forms $\Lambda(M)$ on a manifold M. It is for precisely this reason that the operator d may be uniquely determined by its value on functions f and one forms ω and the law $d(\omega\eta) = d(\omega)\eta + (-1)^{\deg\omega}\omega d(\eta)$ which says precisely that d is a graded derivation of degree 1. This bit of trickery works because $\Lambda(M)$ is free.

9. INTERACTION OF THE LIE DERIVATIVE WITH d AND INNER PRODUCT

In this section we will use the ideas in the previous section to develop what we will call the French Method for proving identities. This will give us a less computational mode of attack.

First however we show that $\mathcal{L}_X$ commutes with d. Recall that if $X \in \Gamma(T(M))$ (i.e. a vector field) and ϕ_t is the corresponding flow then for $\omega \in \Lambda^r(T^*(M)) = \Lambda^r(M)$

$$\mathcal{L}_X\omega = \lim_{t\to 0}\frac{\phi_t^*(\omega) - \omega}{t}$$

Recall that $\phi_t^*(\omega_1 \wedge \omega_2) = \phi_t^*(\omega_1) \wedge \phi_t^*(\omega_2)$. Then

$$\begin{aligned}
\phi_t^*(\omega_1 \wedge \omega_2) - \omega_1 \wedge \omega_2 &= \phi_t^*(\omega_1) \wedge \phi_t^*(\omega_2)\omega_1 \wedge \omega_2 \\
&= \phi_t^*(\omega_1) \wedge \phi_t^*(\omega_2) - \omega_1 \wedge \phi_t^*(\omega_2) \\
&\quad + \omega_1 \wedge \phi_t^*(\omega_2) - \omega_1 \wedge \omega_2 \\
\mathcal{L}_X(\omega_1 \wedge \omega_2) &= \lim_{t\to 0}\frac{\phi_t^*(\omega_1 \wedge \omega_2) - \omega_1 \wedge \omega_2}{t}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{t\to 0}\frac{(\phi_t^*(\omega_1)-\omega_1)\wedge\phi_t^*(\omega_2)+\omega_1\wedge(\phi_t^*(\omega_2)-\omega_2)}{t}\\
&= \pounds_X(\omega_1)\wedge\omega_2+\omega_1\wedge\pounds_X(\omega_2)
\end{aligned}
$$

as we would expect for a reasonable product.

In a similar way, since $\phi_t^*(d\,\omega) = d\,\phi_t^*(\omega)$ we have

$$\pounds_X(d\,\omega) = \lim_{t\to 0}\frac{\phi_t^*(d\omega)-d\,\omega}{t} = \lim_{t\to 0}\frac{d\,\phi_t^*(\omega)-d\,\omega}{t} = d\,\pounds_X(\omega)$$

Note that t is not one of the variables involved in d. We previously proved the formula above for $1-$forms but this proof is valid for $r-$forms. A different proof of this is possible using methods we are about to introduce.

Next notice that $\pounds_X$ is a graded derivation of degree 0 and d is a graded derivation of degree 1. Hence by the theorem in the previous section,

$$[d,\pounds_X] = d\circ\pounds_X-(-1)^{-1\cdot 0}\pounds_X\circ d = d\circ\pounds_X-\pounds_X\circ d$$

is a graded derivation of degree 1. We have just shown it is identically 0. However, note that this fact could be proved by showing that $[d,\pounds_X]$ is identically 0 on functions and $1-$forms. It would then follow that it is 0 on all forms, by the discussion at the end of the previous section. We will call this the French Method.

Clearly there is no fun to be had here with d and $\pounds_X$ because they commute. For real fun we need another graded derivation which does NOT commute with d or $\pounds_X$.

Our first example is simply to use two Lie Derivatives, $\pounds_X$ and $\pounds_Y$. We will show that

$$[\pounds_X,\pounds_Y] = \pounds_{[X,Y]}$$

Since these are graded derivations of degree 0, it suffices to show that they coincide on functions and one forms. Note that we have

$$[\pounds_X,\pounds_Y] = \pounds_X\circ\pounds_Y-(-1)^{0\cdot 0}\pounds_Y\circ\pounds_X = \pounds_X\circ\pounds_Y-\pounds_Y\circ\pounds_X$$

On functions this gives

$$
\begin{aligned}
[\pounds_X,\pounds_Y](f) &= \pounds_X\circ\pounds_Y(f)-\pounds_Y\circ\pounds_X(f) = \pounds_X(Y(f))-\pounds_Y(X(f))\\
&= X(Y(f))-Y(X(f)) = (XY-YX)(f) = [X,Y](f)\\
&= \pounds_{[X,Y]}(f)
\end{aligned}
$$

For one-forms it is perfectly possible to perform a direct attack and fill a page with computations that establish that the two derivations coincide on one forms. However, we would like to do this in a smarter way. The miminal one-form we can work with is du^i and we can use the commutativity of $\pounds_X$ and d effectively here; we don't even need to look at the formula for $\pounds_X$ on one-forms.

$$
\begin{aligned}
(\pounds_X\pounds_Y-\pounds_Y\pounds_X)(du^i) &= d(\pounds_X\pounds_Y-\pounds_Y\pounds_X)(u^i)\\
&= d\,\pounds_{[X,Y]}(u^i)\\
&= \pounds_{[X,Y]}(du^i)
\end{aligned}
$$

Since the two derivations $\pounds_X \pounds_Y - \pounds_Y \pounds_X$ and $\pounds_{[X,Y]}$ agree on functions f and the special one forms du^i, they will agree on all one forms $\omega = f_i du^i$, as required.

Another very important example is the *inner product* of a vector field and a differential form. This is defined by
Def

$$\begin{aligned} X \rfloor f &= 0 \text{ for } f \in \Lambda^0(M) \\ X \rfloor \omega &= \omega(X) \text{ for } \omega \in \Lambda^1(M) \\ X \rfloor \text{ is } & \text{ a } \text{ graded derivation of degree } -1 \end{aligned}$$

This completely defines $X \rfloor$ as we discussed previously, and no further formulae are theoretically necessary. But it is nice to know that for $\omega \in \Lambda^r(M)$

$$X \rfloor \omega(Y_2, Y_3, \ldots, Y_r) = \omega(X, Y_2, Y_3, \ldots, Y_r)$$

However, in order not to break up the derivation we will prove the general formula in appendix 1 of this section. The case $r = 2$ we need below.

We now want to use all our equipment to prove Cartan's Homotopy formula
Theorem For X a vector field on M and $\omega \in \Lambda^r(M)$.

$$\pounds_X \omega = d(X \rfloor \omega) + X \rfloor (d\omega)$$

Since d is a graded derivation of degree 1 and $X \rfloor$ is a graded derivation of degree -1 we know that

$$[d, X \rfloor] = d \circ X \rfloor + X \rfloor \circ d$$

will be a graded derivation of degree 0. Since the only graded derivation of degree 0 that we know is $\pounds_X$ the result is not surprising. Moreover, as we discussed above we can prove the formula by verifying it for $f \in \Lambda^0(M)$ and $\omega \in \Lambda^1(M)$. Since $\pounds_X$ and $[d, X \rfloor]$ coincide on $\Lambda^0(M)$ and $\Lambda^1(M)$, they must coincide on all $\Lambda^r(M)$. We proceed to verify the two cases. for $f \in \Lambda^0(M)$

$$\begin{aligned} d(X \rfloor f) + X \rfloor (d f) &= X \rfloor (d f) \\ &= d f(X) = X(f) = \pounds_X(f) \end{aligned}$$

and for $\omega \in \Lambda^1(M)$, using the formula $(X \rfloor \omega)(Y) = \omega(X, Y)$,

$$\begin{aligned} (d(X \rfloor \omega) + X \rfloor (d\omega))(Y) &= d(\omega(X))Y + d\omega(X, Y) \\ &= Y(\omega(X)) + X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]) \\ &= X(\omega(Y)) - \omega([X, Y]) \\ &= \pounds_X(\omega)(Y) \end{aligned}$$

where we used the formula $\pounds_X(\omega(Y)) - \pounds_X(\omega)(Y) = \omega([X, Y])$

Since we have three graded derivations, d, $X\rfloor$ and $\pounds_X$, we should be able to find formulas using any two of them. We just did the first two, and the pair d, $\pounds_X$ is dull since the two commute, so the remaining possibility is $\pounds_X$ and $X\rfloor$. The commutator of these two should give us a graded derivation of degree -1. Let us see if we can turn it up and simultaneously prove the formula. For $f \in \Lambda^0(M)$

$$\pounds_X(Y\rfloor f) - Y\rfloor \pounds_X(f) = 0$$

No help there. For $\omega \in \Lambda^1(M)$

$$\begin{aligned} \pounds_X(Y\rfloor \omega) - Y\rfloor \pounds_X(\omega) &= \pounds_X(\omega(Y)) - (\pounds_X(\omega))(Y) \\ &= \omega([X,Y]) \\ &= [X,Y]\rfloor \omega \end{aligned}$$

AHA!. We now see that

$$[\pounds_X, Y\rfloor] = \pounds_X \circ Y\rfloor - Y\rfloor \circ \pounds_X = [X,Y]\rfloor$$

which is the formula we were seeking. The formula is true on functions and $1-$forms, and since it is a graded derivation is true on all $r-$forms.

We have now exhausted our supply of graded derivations. However, I wish to present one more formula which has a slightly different feel to it and is not proved in the same way. For this formula we must define the *Outer Product* or *Exterior Product*.

Def Let $\omega \in \Lambda^1(M)$. Then the Outer Product $\epsilon(\omega)$ is defined by

$$\epsilon(\omega)(\eta) = \omega \wedge \eta \qquad \text{for } \eta \in \Lambda^r(M)$$

Note that $\epsilon(\omega)$ is not a derivation. However, it does have an interesting formula which looks a little like a derivation formula. First, recall that $X\rfloor$ is a graded derivation of degree -1 so we have for $\omega \in \Lambda^1(M)$

$$X\rfloor(\omega \wedge \eta) = (X\rfloor \omega) \wedge \eta - \omega \wedge (X\rfloor \eta)$$

Rewriting this with the outer product we have

$$\begin{aligned} X\rfloor(\epsilon(\omega)(\eta)) &= (\omega(X))\eta - \epsilon(\omega)(X\rfloor \eta) \\ X\rfloor(\epsilon(\omega)(\eta)) + \epsilon(\omega)(X\rfloor \eta) &= (\omega(X))\eta \end{aligned}$$

and so we have

$$X\rfloor \circ \epsilon(\omega) + \epsilon(\omega) \circ X\rfloor = \omega(X)$$

where the right hand side is to be interpreted as the operator of scalar multiplication by $\omega(X)$.

Appendix 1 Here we prove the formula for the inner product

$$(X\rfloor \omega)(Y_2, \ldots, Y_r) = \omega(X, Y_2, \ldots, Y_r)$$

for $\omega \in \Lambda^r(M)$. We begin in a leisurely manner by proving the formula for $r = 2$. For this we have

$$\begin{aligned}
X\rfloor(\omega_1 \wedge \omega_2) &= (X\rfloor\omega_1)\omega_2 - \omega_1(X\rfloor\omega_2) \\
&= \omega_1(X)\omega_2 - \omega_2(X)\omega_1 \\
(X\rfloor(\omega_1 \wedge \omega_2))(Y) &= \omega_1(X)\omega_2(Y) - \omega_2(X)\omega_1(Y) \\
&= (\omega_1 \wedge \omega_2)(X, Y)
\end{aligned}$$

For the general case we remind the reader of some formulae. First

$$\omega_1 \wedge \omega_2 \wedge \ldots \wedge \omega_r(Y_1, \ldots, Y_r) = \det \begin{pmatrix} \omega_1(Y_1) & \omega_1(Y_2) & \cdots & \omega_1(Y_r) \\ \omega_2(Y_1) & \omega_2(Y_2) & \cdots & \omega_2(Y_r) \\ \cdots & \cdots & \cdots & \cdots \\ \omega_r(Y_1) & \omega_r(Y_2) & \cdots & \omega_r(Y_r) \end{pmatrix}$$

$$\omega_1 \wedge \omega_2 \wedge \ldots \wedge \omega_r(Y_1, \ldots, Y_r) = \sum_{i=1}^{r}(-1)^{i-1}\omega_1(Y_i)\omega_2 \wedge \ldots \wedge \omega_r(Y_1, \ldots, \hat{Y}_i, \ldots, Y_r)$$

Thus for $\omega \in \Lambda^{r-1}$

$$\omega_1 \wedge \omega(Y_1, \ldots, Y_r) = \sum_{i=1}^{r}(-1)^{i-1}\omega_1(Y_i)\omega(Y_1, \ldots, \hat{Y}_i, \ldots, Y_r)$$

Now we begin our proof by induction. Assume the formula is true for $r - 1$. Then for $\omega_1 \in \Lambda^1(M)$ and $\omega \in \Lambda^r(M)$

$$\begin{aligned}
&Y_1\rfloor(\omega_1 \wedge \omega)(Y_2, \ldots, Y_r) \\
&\quad= (Y_1\rfloor\omega_1)\omega(Y_2, \ldots, Y_r) - (\omega_1 \wedge (Y_1\rfloor\omega))(Y_2, \ldots, Y_r) \\
&\quad= \omega_1(Y_1)\omega(Y_2, \ldots, Y_r) - \sum_{i=2}^{r}(-1)^{i}\omega_1(Y_i)(Y_1\rfloor\omega)(Y_2, \ldots, \hat{Y}_i, \ldots, Y_r)
\end{aligned}$$

Now using the induction assumption we have

$$\begin{aligned}
&Y_1\rfloor(\omega_1 \wedge \omega)(Y_2, \ldots, Y_r) \\
&\quad= \omega_1(Y_1)\omega(Y_2, \ldots, Y_r) - \sum_{i=2}^{r}(-1)^{i}\omega_1(Y_i)\omega(Y_1, \ldots, \hat{Y}_i, \ldots, Y_r) \\
&\quad= \sum_{i=1}^{r}(-1)^{i-1}\omega_1(Y_i)\omega(Y_1, \ldots, \hat{Y}_i, \ldots, Y_r) \\
&\quad= (\omega_1 \wedge \omega)(Y_1, \ldots, Y_r)
\end{aligned}$$

This completes the proof.

References

[1] Berline,N; Getzler, E. & Vergne, M. *Heat Kernals and Dirac Operators* Springer Verlag, Berlin, 1992

[2] Godbillon, Claude ; *Géométrie Différentielle et Méchanique Analytique* Hermann Paris 1969

[3] Helgason, Sigurdur; *Differential Geometry and Symmetric Spaces* American Mathematical Society, Tel Aviv, 2000

[4] Loomis, Lynn H.; *Advanced Calculus* 2^{nd} edition, Jones and Bartlett, Boston, 1990

[5] Morrey, Charles B.; *Multiple Integrals in the Calculus of Variations,* Springer Verlag, Berlin, 1966