Symplectic Geometry
Lecture 5

Review of differential calculus.
The generalized Weil identity.
The Moser trick.

The purpose of today's lecture is to give a rapid review of the basics of the calculus of differential forms on manifolds. We will give two proofs of Weil's formula for the Lie derivative of a differential form: the first of an algebraic nature and then a more general geometric formulation with a "functorial" proof that we learned from Bott. We then apply this formula to the "Moser trick" and give several applications of this method.

Superalgebras.

A (commutative associative) **superalgebra** is a vector space

$$A = A_{even} \oplus A_{odd}$$

with a given direct sum decomposition into even and odd pieces, and a map

$$A \times A \to A$$

which is bilinear, satisfies the associative law for multiplication, and

$$\begin{aligned}
A_{even} \times A_{even} &\to A_{even} \\
A_{even} \times A_{odd} &\to A_{odd} \\
A_{odd} \times A_{even} &\to A_{odd} \\
A_{odd} \times A_{odd} &\to A_{even} \\
\omega \cdot \sigma &= \sigma \cdot \omega \text{ if either } \omega \text{ or } \sigma \text{ are even,} \\
\omega \cdot \sigma &= -\sigma \cdot \omega \text{ if both } \omega \text{ and } \sigma \text{ are odd.}
\end{aligned}$$

$$
\begin{aligned}
A_{even} \times A_{even} &\rightarrow A_{even} \\
A_{even} \times A_{odd} &\rightarrow A_{odd} \\
A_{odd} \times A_{even} &\rightarrow A_{odd} \\
A_{odd} \times A_{odd} &\rightarrow A_{even} \\
\omega \cdot \sigma &= \sigma \cdot \omega \text{ if either } \omega \text{ or } \sigma \text{ are even,} \\
\omega \cdot \sigma &= -\sigma \cdot \omega \text{ if both } \omega \text{ and } \sigma \text{ are odd.}
\end{aligned}
$$

We write these last two conditions as

$$\omega \cdot \sigma = (-1)^{\mathrm{deg}_\sigma \mathrm{deg}_\omega} \sigma \cdot \omega.$$

Here $\deg \tau = 0$ if τ is even, and $\deg \tau = 1 \pmod 2$ if τ is odd.

Differential forms.

A **linear differential form** on a manifold, M, is a rule which assigns to each $p \in M$ a linear function on TM_p. So a linear differential form, ω, assigns to each p an element of TM_p^*. We will, as usual, only consider linear differential forms which are smooth.

The superalgebra $\Omega(M)$ is the superalgebra generated by smooth functions on M (taken as even) and by the linear differential forms, taken as odd.

Multiplication of differential forms is usually denoted by $\wedge$. The number of differential factors is called the *degree* of the form. So functions have degree zero, linear differential forms have degree one.

In terms of local coordinates, the most general *linear* differential form has an expression as $a_1 dx_1 + \cdots + a_n dx_n$ (where the a_i are functions). Expressions of the form

$$a_{12} dx_1 \wedge dx_2 + a_{13} dx_1 \wedge dx_3 + \cdots + a_{n-1,n} dx_{n-1} \wedge dx_n$$

have degree two (and are even). Notice that the multiplication rules require

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

and, in particular, $dx_i \wedge dx_i = 0$. So the most general sum of products of two linear differential forms is a differential form of degree two, and can be brought to the above form, locally, after collections of coefficients. Similarly, the most general differential form of degree $k \leq n$ on an n dimensional manifold is a sum, locally, with function coefficients, of expressions of the form

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \quad i_1 < \cdots < i_k.$$

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

and, in particular, $dx_i \wedge dx_i = 0$. So the most general sum of products of two linear differential forms is a differential form of degree two, and can be brought to the above form, locally, after collections of coefficients. Similarly, the most general differential form of degree $k \leq n$ on an n dimensional manifold is a sum, locally, with function coefficients, of expressions of the form

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \quad i_1 < \cdots < i_k.$$

There are $\left(\begin{array}{c} n \\ k \end{array} \right)$ such expressions, and they are all even, if k is even, and odd if k is odd.

The *d* operator.

There is a linear operator d acting on differential forms called *exterior* differentiation, which is completely determined by the following rules: It satisfies Leibniz' rule in the "super" form

$$d(\omega \cdot \sigma) = (d\omega) \cdot \sigma + (-1)^{\deg \omega} \omega \cdot (d\sigma).$$

On functions it is given by

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

and, finally,

$$d(dx_i) = 0.$$

Since functions and the dx_i generate, this determines d completely.

For example, on linear differential forms

$$\omega = a_1 dx_1 + \cdots a_n dx_n$$

we have

$$\begin{aligned} d\omega &= da_1 \wedge dx_1 + \cdots + da_n \wedge dx_n \\ &= \left(\frac{\partial a_1}{\partial x_1} dx_1 + \cdots \frac{\partial a_1}{\partial x_n} dx_n \right) \wedge dx_1 + \cdots \\ & \qquad \left(\frac{\partial a_n}{\partial x_1} dx_1 + \cdots + \frac{\partial a_n}{\partial x_n} dx_n \right) \wedge dx_n \end{aligned}$$

$$= \left(\frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2} \right) dx_1 \wedge dx_2 + \cdots + \left(\frac{\partial a_n}{\partial x_{n-1}} - \frac{\partial a_{n-1}}{\partial x_n} \right) dx_{n-1} \wedge dx_n$$

In particular, equality of mixed derivatives shows that $d^2 f = 0$.

Rules to remember.

In particular, equality of mixed derivatives shows that $d^2 f = 0$, and hence that $d^2\omega = 0$ for any differential form. Hence the rules to remember about d are:

$$
\begin{aligned}
d(\omega \cdot \sigma) &= (d\omega) \cdot \sigma + (-1)^{\deg_\omega} \, \omega \cdot (d\sigma) \\
d^2 &= 0 \\
df &= \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n .
\end{aligned}
$$

Derivations.

A linear operator $\ell : A \to A$ is called an *odd derivation* if, like d, it satisfies

$$\ell : A_{even} \to A_{odd}, \quad \ell : A_{odd} \to A_{even}$$

and

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + (-1)^{\deg \omega} \, \omega \cdot \ell\sigma.$$

A linear map $\ell : A \to A$,

$$\ell : A_{even} \to A_{even}, \;\; \ell : A_{odd} \to A_{odd}$$

satisfying

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + \omega \cdot (\ell\sigma)$$

is called an *even derivation*. So the Leibniz rule for derivations, even or odd, is

$$\ell(\omega \cdot \sigma) = (\ell\omega) \cdot \sigma + (-1)^{\deg \ell \deg \omega} \, \omega \cdot \ell\sigma.$$

Knowing the action of a derivation on a set of generators of a superalgebra determines it completely. For example, the equations

$$d(x_i) = dx_i, \quad d(dx_i) = 0 \quad \forall i$$

implies that

$$dp = \frac{\partial p}{\partial x_1} dx_1 + \cdots + \frac{\partial p}{\partial x_n} dx_n$$

for any polynomial, and hence determines the value of d on any differential form with polynomial coefficients. The local formula we gave for df where f is any differentiable function, was just the natural extension (by continuity, if you like) of the above formula for polynomials.

The sum of two even derivations is an even derivation, and the sum of two odd derivations is an odd derivation.

Commutator.

The composition of two derivations will not, in general, be a derivation, but an instructive computation from the definitions shows that the *commutator*

$$[\ell_1, \ell_2] := \ell_1 \circ \ell_2 - (-1)^{\deg_{\ell_1} \deg_{\ell_2}} \ \ell_2 \circ \ell_1$$

is again a derivation which is even if both are even or both are odd, and odd if one is even and the other odd.

Derivations and multiplications.

A derivation followed by a multiplication is again a derivation: specifically, let ℓ be a derivation (even or odd) and let τ be an even or odd element of A. Consider the map

$$\omega \mapsto \tau\ell\omega.$$

We have

$$\begin{aligned}\tau\ell(\omega\sigma) &= (\tau\ell\omega)\cdot\sigma + (-1)^{\deg_\ell \deg_\omega}\tau\omega\cdot\ell\sigma \\ &= (\tau\ell\omega)\cdot\sigma + (-1)^{(\deg_\ell + \deg_\tau)\deg_\omega}\omega\cdot(\tau\ell\sigma)\end{aligned}$$

so $\omega \mapsto \tau\ell\omega$ is a derivation whose degree is

$$\deg\tau + \deg\ell.$$

Pullback.

Let $\phi : M \to N$ be a smooth map. Then the pullback map ϕ^* is a linear map that sends differential forms on N to differential forms on M and satisfies

$$\begin{aligned} \phi^*(\omega \wedge \sigma) &= \phi^*\omega \wedge \phi^*\sigma \\ \phi^* d\omega &= d\phi^*\omega \\ (\phi^* f) &= f \circ \phi. \end{aligned}$$

The first two equations imply that ϕ^* is completely determined by what it does on functions. The last equation says that on functions, ϕ^* is given by "substitution": In terms of local coordinates on M and on N ϕ is given by

$$\begin{aligned} \phi(x^1, \ldots, x^m) &= (y^1, \ldots, y^n) \\ y^i &= \phi^i(x^1, \ldots, x^m) \quad i = 1, \ldots, n \end{aligned}$$

In terms of local coordinates on M and on N ϕ is given by

$$\begin{aligned} \phi(x^1, \ldots, x^m) &= (y^1, \ldots, y^n) \\ y^i &= \phi^i(x^1, \ldots, x^m) \quad i = 1, \ldots, n \end{aligned}$$

where the ϕ_i are smooth functions. The local expression for the pullback of a function $f(y^1, \ldots, y^n)$ is to substitute ϕ^i for the y^is as into the expression for f so as to obtain a function of the x's.

It is important to observe that the pull back on differential forms is defined for any smooth map, not merely for diffeomorphisms. This is the great advantage of the calculus of differential forms.

The chain rule.

Suppose that $\psi : N \to P$ is a smooth map so that the composition

$$\phi \circ \psi : M \to P$$

is again smooth. Then the *chain rule* says

$$(\phi \circ \psi)^* = \psi^* \circ \phi^*.$$

On functions this is essentially a tautology - it is the associativity of composition: $f \circ (\phi \circ \psi) = (f \circ \phi) \circ \psi$. But since pull-back is completely determined by what it does on functions, the chain rule applies to differential forms of any degree.

Lie derivative.

Let ϕ_t be a one parameter group of transformations of M. If ω is a differential form, we get a family of differential forms, $\phi_t^*\omega$ depending differentiably on t, and so we can take the derivative at $t = 0$:

$$\frac{d}{dt}\left(\phi_t^*\omega\right)_{|t=0} = \lim_{t=0} \frac{1}{t}\left[\phi_t^*\omega - \omega\right].$$

Since $\phi_t^*(\omega \wedge \sigma) = \phi_t^*\omega \wedge \phi_t^*\sigma$ it follows from the Leibniz argument that

$$\ell_\phi : \ \omega \mapsto \frac{d}{dt}\left(\phi_t^*\omega\right)_{|t=0}$$

is an even derivation. We want a formula for this derivation.

$$\ell_\phi : \ \omega \mapsto \frac{d}{dt}\left(\phi_t^*\omega\right)_{|t=0}$$

is an even derivation. We want a formula for this derivation.

Notice that since $\phi_t^* d = d\phi_t^*$ for all t, it follows by differentiation that

$$\ell_\phi d = d\ell_\phi$$

and hence the formula for ℓ_ϕ is completely determined by how it acts on functions.

Let X be the vector field generating ϕ_t. Recall that the geometrical significance of this vector field is as follows: If we fix a point x, then

$$t \mapsto \phi_t(x)$$

is a curve which passes through the point x at $t = 0$. The tangent to this curve at $t = 0$ is the vector $X(x)$. In terms of local coordinates, X has coordinates $X = (X^1, \dots, X^n)$ where $X^i(x)$ is the derivative of $\phi^i(t, x^1, \dots, x^n)$ with respect to t at $t = 0$. The chain rule then gives, for any function f,

$$\begin{aligned} \ell_\phi f &= \frac{d}{dt} f(\phi^1(t, x^1, \dots, x^n), \dots, \phi_n(t, x^1, \dots, x^n))_{|t=0} \\ &= X^1 \frac{\partial f}{\partial x_1} + \cdots + X^n \frac{\partial f}{\partial x_n}. \end{aligned}$$

Vector fields as differential operators.

The chain rule then gives, for any function f,

$$\begin{aligned}\ell_\phi f &= \frac{d}{dt} f(\phi^1(t, x^1, \ldots, x^n), \ldots, \phi_n(t, x^1, \ldots, x^n))_{|t=0} \\ &= X^1 \frac{\partial f}{\partial x_1} + \cdots + X^n \frac{\partial f}{\partial x_n}.\end{aligned}$$

For this reason we use the notation

$$X = X^1 \frac{\partial}{\partial x_1} + \cdots + X^n \frac{\partial}{\partial x_n}$$

so that the differential operator

$$f \mapsto Xf$$

gives the action of ℓ_ϕ on functions.

As we mentioned, this action of ℓ_ϕ on functions determines it completely. In particular, ℓ_ϕ depends only on the vector field X, so we may write

$$\ell_\phi = D_X$$

where D_X is the even derivation determined by

$$D_X f = Xf, \quad D_X d = dD_X.$$

But we want a more explicit formula for D_X. For this it is useful to introduce an odd derivation associated to X called the *interior product* and denoted by $i(X)$. It is defined as follows: First consider the case where

$$X = \frac{\partial}{\partial x_j}$$

First consider the case where

$$X = \frac{\partial}{\partial x_j}$$

and define its interior product by

$$i\left(\frac{\partial}{\partial x_j}\right) f = 0$$

for all functions while

$$i\left(\frac{\partial}{\partial x_j}\right) dx_k = 0, \quad k \neq j$$

and

$$i\left(\frac{\partial}{\partial x_j}\right) dx_j = 1.$$

$$i\left(\frac{\partial}{\partial x_j}\right) f = 0 \qquad i\left(\frac{\partial}{\partial x_j}\right) dx_k = 0, \quad k \neq j \qquad i\left(\frac{\partial}{\partial x_j}\right) dx_j = 1.$$

The fact that it is a derivation then gives an easy rule for calculating $i(\partial/\partial x_j)$ when applied to any differential form: Write the differential form as

$$\omega + dx_j \wedge \sigma$$

where the expressions for ω and σ do not involve dx_j. Then

$$i\left(\frac{\partial}{\partial x_j}\right)[\omega + dx_j \wedge \sigma] = \sigma.$$

The interior product in general.

The operator

$$X^j i\left(\frac{\partial}{\partial x_j}\right)$$

which means first apply $i(\partial/\partial x_j)$ and then multiply by the function X^j is again an odd derivation, and so we can make the definition

$$i(X) := X^1 i\left(\frac{\partial}{\partial x_1}\right) + \cdots + X^n i\left(\frac{\partial}{\partial x_n}\right). \qquad (1)$$

It is easy to check that this does not depend on the local coordinate system used.

Notice that we can write

$$Xf = i(X)df.$$

In particular we have

$$\begin{aligned} D_X dx_j &= dD_X x_j \\ &= dX_j \\ &= di(X)dx_j. \end{aligned}$$

We can combine these two formulas as follows: Since $i(X)f = 0$ for any function f we have

$$D_X f = di(X)f + i(X)df.$$

Since $ddx_j = 0$ we have

$$D_X dx_j = di(X)dx_j + i(X)ddx_j.$$

Weil's formula.

Hence

$$D_X = di(X) + i(X)d = [d, i(X)] \tag{2}$$

when applied to functions or to the forms dx_j. But the right hand side of the preceding equation is an even derivation, being the commutator of two odd derivations. So if the left and right hand side agree on functions and on the differential forms dx_j they agree everywhere. This equation, (2), known as *Weil's formula*, is a basic formula in differential calculus.

Exterior forms as multilinear functions.

We can use the interior product to consider differential forms of degree k as $k-$multilinear functions on the tangent space at each point. To illustrate, let σ be a differential form of degree two. Then for any vector field, X, $i(X)\sigma$ is a linear differential form, and hence can be evaluated on any vector field, Y to produce a function. So we define

$$\sigma(X,Y) := [i(X)\sigma](Y).$$

$$\sigma(X,Y) := [i(X)\sigma](Y).$$

We can use this to express exterior derivative in terms of ordinary derivative and Lie bracket: If θ is a linear differential form, we have

$$\begin{aligned}
d\theta(X,Y) &= [i(X)d\theta](Y) \\
i(X)d\theta &= D_X\theta - d(i(X)\theta) \\
d(i(X)\theta)(Y) &= Y[\theta(X)] \\
[D_X\theta](Y) &= D_X[\theta(Y)] - \theta(D_X(Y)) \\
&= X[\theta(Y)] - \theta([X,Y])
\end{aligned}$$

where we have introduced the notation $D_XY =: [X,Y]$ which is legitimate since on functions we have

$$(D_XY)f = D_X(Yf) - YD_Xf = X(Yf) - Y(Xf)$$

so D_XY as an operator on functions is exactly the commutator

of X and Y.

An expression for the exterior derivative.

$$
\begin{aligned}
d\theta(X,Y) &= [i(X)d\theta](Y) \\
i(X)d\theta &= L_X\theta - d(i(X)\theta) \\
d(i(X)\theta)(Y) &= Y[\theta(X)] \\
[D_X\theta](Y) &= D_X[\theta(Y)] - \theta(D_X(Y)) \\
&= X[\theta(Y)] - \theta([X,Y])
\end{aligned}
$$

$$
\begin{aligned}
d\theta(X,Y) &= [i(X)d\theta](Y) \\
i(X)d\theta &= D_X\theta - d(i(X)\theta) \\
d(i(X)\theta)(Y) &= Y[\theta(X)] \\
[D_X\theta](Y) &= D_X[\theta(Y)] - \theta(D_X(Y)) \\
&= X[\theta(Y)] - \theta([X,Y])
\end{aligned}
$$

Putting the previous pieces together gives

$$
d\theta(X,Y) = X\theta(Y) - Y\theta(X) - \theta([X,Y]), \tag{3}
$$

with similar expressions for differential forms of higher degree.

Integration.

Let

$$\omega = f dx_1 \wedge \cdots \wedge dx_n$$

be a form of degree n on $\mathbf{R}^n$. (Recall that the most general differential form of degree n is an expression of this type.) Then its integral is defined by

$$\int_M \omega := \int_M f dx_1 \cdots dx_n$$

where M is any (measurable) subset.

$$\int_M \omega := \int_M f dx_1 \cdots dx_n$$

where M is any (measurable) subset. This, of course is subject to the condition that the right hand side converges if M is unbounded. There is a lot of hidden subtlety built into this definition having to do with the notion of orientation. But for the moment this is a good working definition.

The *change of variables formula* says that if $\phi : M \to \mathbf{R}^n$ is a smooth differentiable map which is one to one whose Jacobian determinant is everywhere positive, then

$$\int_M \phi^* \omega = \int_{\phi(M)} \omega.$$

Stokes theorem.

Let U be a region in $\mathbf{R}^n$ with a chosen orientation and smooth boundary. We then orient the boundary according to the rule that an outward pointing normal vector, together with the a positive frame on the boundary give a positive frame in $\mathbf{R}^n$. If σ is an $(n-1)-$form, then

$$\int_{\partial U} \sigma = \int_U d\sigma.$$

$$\int_{\partial U} \sigma = \int_U d\sigma.$$

A manifold is called *orientable* if we can choose an atlas consisting of charts such that the Jacobian of the transition maps $\phi_\alpha \circ \phi_\beta^{-1}$ is always positive. Such a choice of an atlas is called an orientation. (Not all manifolds are orientable.) If we have chosen an orientation, then relative to the charts of our orientation, the transition laws for an $n-$form (where $n = \dim M$) and for a density are the same. In other words, given an orientation, we can identify densities with $n-$forms and $n-$form with densities. Thus we may integrate $n-$forms. The change of variables formula then holds for orientation preserving diffeomorphisms as does Stokes theorem.

Lie derivatives of vector fields.

Let Y be a vector field and ϕ_t a one parameter group of transformations whose "infinitesimal generator" is some other vector field X. We can consider the "pulled back" vector field $\phi_t^* Y$ defined by

$$\phi_t^* Y(x) = d\phi_{-t}\{Y(\phi_t x)\}.$$

In words, we evaluate the vector field Y at the point $\phi_t(x)$, obtaining a tangent vector at $\phi_t(x)$, and then apply the differential of the (inverse) map ϕ_{-t} to obtain a tangent vector at x.

If we differentiate the one parameter family of vector fields $\phi_t^* Y$ with respect to t and set $t = 0$ we get a vector field which we denote by $D_X Y$:

$$D_X Y := \frac{d}{dt} \phi_t^* Y_{|t=0}.$$

If ω is a linear differential form, then we may compute $i(Y)\omega$ which is a function whose value at any point is obtained by evaluating the linear function $\omega(x)$ on the tangent vector $Y(x)$. Thus

$$i(\phi_t^*Y)\phi_t^*\omega(x) = \langle (d(\phi_t)_x)^*\omega(\phi_t x), d\phi_{-t}Y(\phi_t x)\rangle = \{i(Y)\omega\}(\phi_t x).$$

In other words,

$$\phi_t^*\{i(Y)\omega\} = i(\phi_t^*Y)\phi_t^*\omega.$$

We have verified this when ω is a differential form of degree one. It is trivially true when ω is a differential form of degree zero, i.e. a function, since then both sides are zero. But then, by the derivation property, we conclude that it is true for forms of all degrees. We may rewrite the result in shorthand form as

$$\phi_t^* \circ i(Y) = i(\phi_t^*Y) \circ \phi_t^*.$$

$$D_X Y := \frac{d}{dt}\phi_t^* Y_{|t=0}.$$

Since $\phi_t^* d = d\phi_t^*$ we conclude from Weil's formula that

$$\phi_t^* \circ D_Y = D_{\phi_t^* Y} \circ \phi_t^*.$$

Until now the subscript t was superfluous, the formulas being true for any fixed diffeomorphism. Now we differentiate the preceding equations with respect to t and set $t = 0$. We obtain,using Leibniz's rule,

$$D_X \circ i(Y) = i(D_X Y) + i(Y) \circ D_X$$

and

$$D_X \circ D_Y = D_{D_X Y} + D_Y \circ D_X.$$

$$D_X \circ i(Y) = i(D_X Y) + i(Y) \circ D_X$$

and

$$D_X \circ D_Y = D_{D_X Y} + D_Y \circ D_X.$$

This last equation says that Lie derivative (on forms) with respect to the vector field $D_X Y$ is just the commutator of D_X with D_Y:

$$D_{D_X Y} = [D_X, D_Y].$$

For this reason we write

$$[X, Y] := D_X Y$$

and call it the Lie bracket (or commutator) of the two vector fields X and Y. The equation for interior product can then be written as

$$i([X, Y]) = [D_X, i(Y)].$$

The Lie bracket is antisymmetric in X and Y. We may multiply Y by a function g to obtain a new vector field gY. Form the definitions we have

$$\phi_t^*(gY) = (\phi_t^* g)\phi_t^* Y.$$

Differentiating at $t = 0$ and using Leibniz's rule we get

$$[X, gY] = (Xg)Y + g[X, Y] \tag{4}$$

where we use the alternative notation Xg for $D_X g$. The antisymmetry then implies that for any differentiable function f we have

$$[fX, Y] = -(Yf)X + f[X, Y]. \tag{5}$$

From both this equation and from Weil's formula (applied to differential forms of degree greater than zero) we see that the Lie derivative with respect to X at a point x depends on more than the value of the vector field X at x.

Jacobi's identity.

From the fact that $[X, Y]$ acts as the commutator of X and Y it follows that for any three vector fields X, Y and Z we have

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

This is known as **Jacobi's identity**. We can also derive it from the fact that $[Y, Z]$ is a natural operation and hence for any one parameter group ϕ_t of diffeomorphisms we have

$$\phi_t^*([Y,Z]) = [\phi_t^* Y, \phi_t^* Z].$$

If X is the infinitesimal generator of ϕ_t then differentiating the preceding equation with respect to t at $t = 0$ gives

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

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

In other words, X acts as a derivation of the "mutliplication" given by Lie bracket. This is just Jacobi's identity when we use the antisymmetry of the bracket. In the future we we will have occasion to take cyclic sums such as those which arise on the left of Jacobi's identity. So if F is a function of three vector fields (or of three elements of any set) with values in some vector space (for example in the space of vector fields) we will define the cyclic sum $\mathcal{C}yc\ F$ by

$$\mathcal{C}yc\ F(X,Y,Z) := F(X,Y,Z) + F(Y,Z,X) + F(Z,X,Y).$$

With this definition Jacobi's identity becomes

$$\mathcal{C}yc\ [X,[Y,Z]] = 0. \qquad (6)$$

A general version of Weil's formula.

Let W and Z be differentiable manifolds, let I denote an interval on the real line containing the origin, and let

$$\phi : W \times I \to Z$$

be a smooth map. We let $\phi_t : W \to Z$ be defined by

$$\phi_t(w) := \phi(w, t).$$

We think of ϕ_t as a one parameter family of maps from W to Z. We let ξ_t denote the tangent vector field along ϕ_t. In more detail:

$$\xi_t : W \to TZ$$

is defined by letting $\xi_t(w)$ be the tangent vector to the curve $s \mapsto \phi(w, s)$ at $s = t$.

If σ is a differential form on Z of degree $k+1$, we let the expression $\phi_t^* i(\xi_t)\sigma$ denote the differential form on W of degree k whose value at tangent vectors $\eta_1, \ldots, \eta_k$ at $w \in W$ is given by

$$\phi_t^* i(\xi_t)\sigma(\eta_1, \ldots, \eta_k) := (i(\xi_t)(w))\sigma)(d(\phi_t)_w \eta_1, \ldots, d(\phi_t)_w \eta_k). \tag{7}$$

It is only the combined expression $\phi_t^* i(\xi_t)\sigma$ which will have any sense in general: since ξ_t is not a vector field on Z, the expression $i(\xi_t)\sigma$ will not make sense as a stand alone object (in general).

Let σ_t be a smooth one-parameter family of differential forms on Z. The

$$\phi_t^* \sigma_t$$

is a smooth one parameter family of forms on W, which we can then differentiate with respect to t. The general form of Weil's formula is:

$$\frac{d}{dt}\phi_t^* \sigma_t = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* i(\xi_t) d\sigma + d\phi_t^* i(\xi_t)\sigma. \tag{8}$$

Before proving the formula, let us note that it is functorial in the following sense: Suppose that that $F : X \to W$ and $G : Z \to Y$ are smooth maps, and that τ_t is a smooth family of differential forms on Y. Suppose that $\sigma_t = G^*\tau_t$ for all t. We can consider the maps

$$\psi_t : X \to Y, \quad \psi_t := G \circ \phi_t \circ F$$

and then the smooth one parameter familiy of differential forms

$$\psi_t^*\tau_t$$

on X. The tangent vector field ζ_t along ψ_t is given by

$$\zeta_t(x) = dG_{\phi_t(F(x))}\left(\xi_t(F(x))\right).$$

So

$$\psi_t^* i(\zeta_t)\tau_t = F^*\left(\phi_t^* i(\xi_t) G^*\tau_t\right).$$

Therefore, if we know that (8) is true for ϕ_t and σ_t, we can conclude that the analogous formula is true for ψ_t and τ_t.

Consider the special case of (8) where we take the one parameter family of maps

$$f_t : W \times I \to W \times I, \qquad f_t(w, s) = (w, s + t).$$

Let

$$G : W \times I \to Z$$

be the map ϕ, and let

$$F : W \to W \times I$$

be the map

$$F(w) = (w, 0).$$

Then

$$(G \circ f_t \circ F)(w) = \phi_t(w).$$

Thus the functoriality of the formula (8) shows that we only have to prove it for the special case $\phi_t = f_t : W \times I \to W \times I$ as given above!

In this case, it is clear that the vector field ξ_t along ψ_t is just the constant vector field $\frac{\partial}{\partial s}$ evaluated at $(x, s+t)$. The most general differential (t-dependent) on $W \times I$ can be written as

$$ds \wedge a + b$$

where a and b are differential forms on W. (In terms of local coordinates $s, x^1, \ldots, x^n$ these forms a and b are sums of terms that have the expression

$$c dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

where c is a function of s, t and x.) To show the full dependence on the variables we will write

$$\sigma_t = ds \wedge a(x, s, t) dx + b(x, s, t) dx.$$

With this notation it is clear that

$$\phi_t^* \sigma_t = ds \wedge a(x, s+t, t) dx + b(x, s+t, t) dx$$

$$\phi_t^*\sigma_t = ds \wedge a(x, s+t, t)dx + b(x, s+t, t)dx$$

and therefore

$$\frac{d\phi_t^*\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx$$
$$+ds \wedge \frac{\partial a}{\partial t}(x, s+t, t)dx + \frac{\partial b}{\partial t}(x, s+t, t)dx.$$

So

$$\frac{d\phi_t^*\sigma_t}{dt} - \phi_t^*\frac{d\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

$$\frac{d\phi_t^*\sigma_t}{dt} - \phi_t^*\frac{d\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

$$i\left(\frac{\partial}{\partial s}\right)\sigma_t = a dx$$

Therefore

$$d\phi_t^* i(\xi_t)\sigma_t = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + d_W(a(x, s+t, t)dx).$$

$$d\sigma_t = -ds \wedge d_W(adx) + \frac{\partial b}{\partial s} ds \wedge dx + d_W bdx$$

so

$$i\left(\frac{\partial}{\partial s}\right) d\sigma_t = -d_W(adx) + \frac{\partial b}{\partial s}dx$$

and therefore

$$\phi_t^* i(\xi_t) d\sigma_t = -d_W a(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

$$\frac{d\phi_t^*\sigma_t}{dt} - \phi_t^*\frac{d\sigma_t}{dt} = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

$$d\phi_t^* i(\xi_t)\sigma_t = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + d_W(a(x, s+t, t)dx).$$

$$\phi_t^* i(\xi_t)d\sigma_t = -d_W a(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx.$$

So

$$d\phi_t^* i(\xi_t)\sigma_t + \phi_t^* i(\xi_t)d\sigma_t = ds \wedge \frac{\partial a}{\partial s}(x, s+t, t)dx + \frac{\partial b}{\partial s}(x, s+t, t)dx$$

$$= \frac{d\phi_t^*\sigma_t}{dt} - \phi_t^*\frac{d\sigma_t}{dt}$$

proving (8).

A special case of (8) is the following. Suppose that $W = Z = M$ and ϕ_t is a family of diffeomorphisms $f_t : M \to M$. Then ξ_t is given by

$$\xi_t(p) = v_t(f_t(p))$$

where v_t is the vector field

$$v_t(f(p)) = \frac{d}{dt} f_t(p).$$

In this case $i(v_t)\sigma_t$ makes sense, and so we can write (8) as

$$\frac{d\phi_t^* \sigma_t}{dt} = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* D_{v_t} \sigma_t. \tag{9}$$

The Moser trick.

Let M be a differentiable manifold and let ω_0 and ω_1 be smooth k-forms on M. Let us examine the following question: does there exist a diffeomorphism $f : M \to M$ such that $f^*\omega_1 = \omega_0$?

Moser answers this kind of question by making it harder! Let $\omega_t, \quad 0 \leq t \leq 1$ be a family of k-forms with $\omega_t = \omega_0$ at $t = 0$ and $\omega_t = \omega_1$ at $t = 1$. We look for a one parameter family of diffeomorphisms

$$f_t : M \to M, \quad 0 \leq t \leq 1$$

such that

$$f_t^*\omega_t = \omega_0 \tag{10}$$

and

$$f_0 = \mathrm{id}\,.$$

$$\frac{d\phi_t^* \sigma_t}{dt} = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* D_{v_t} \sigma_t. \tag{9}$$

$$f_t^* \omega_t = \omega_0 \tag{10}$$

Let us differentiate (10) with respect to t and apply (9). We obtain

$$f_t^* \dot{\omega}_t + f_t^* D_{v_t} \omega_t = 0$$

where we have written $\dot{\omega}_t$ for $\frac{d\omega_t}{dt}$. Since f_t is required to be a diffeomorphism, this becomes the requirement that

$$D_{v_t} \omega_t = -\dot{\omega}_t. \tag{11}$$

$$D_{v_t}\omega_t = -\dot{\omega}_t. \tag{11}$$

Moser's method is to use "geometry" to solve this equation for v_t if possible. Once we have found v_t, solve the equations

$$\frac{d}{dt}f_t(p) = v_t(f_t(p)), \qquad f_0(p) = p \tag{12}$$

for f_t. Notice that for p fixed and $\gamma(t) = f_t(p)$ this is a system of ordinary differential equations

$$\frac{d}{dt}\gamma(t) = v_t(\gamma(t)), \quad \gamma(0) = p.$$

The standard existence theorems for ordinary differential equations guarantees the existence of of a solution depending smoothly on p at least for $|t| < \epsilon$. One then must make some additional hypotheses that guarantee existence for all time (or at least up to $t = 1$). Two such additional hypotheses might be

- M is compact, or
- C is a closed subset of M on which $v_t \equiv 0$. Then for $p \in C$ the solution for all time is $f_t(p) = p$. Hence for p close to C solutions will exist for a long time. Under this condition

there will exist a neighborhood U of C and a family of diffeomorphisms

$$f_t : U \to M$$

defined for $0 \leq t \leq 1$ such

$$f_0 = \text{id}, \qquad f_{t|C} = \text{id}\, \forall t$$

and (10) is satisfied.

Volume forms.

Let M be a compact oriented connected n-dimensional manifold. Let ω_0 and ω_1 be nowhere vanishing n-forms with the same volume:

$$\int_M \omega_0 = \int_M \omega_1.$$

Moser's theorem asserts that under these conditions there exists a diffeomorphism $f : M \to M$ such that

$$f^* \omega_1 = \omega_0.$$

Moser invented his method for the proof of this theorem.

The first step is to choose the ω_t. Let

$$\omega_t := (1 - t)\omega_0 + t\omega_1.$$

The first step is to choose the ω_t. Let

$$\omega_t := (1-t)\omega_0 + t\omega_1.$$

Since both ω_0 and ω_1 are nowhere vanishing, and since they yield the same integral (and since M is connected), we know that at every point they are either both positive or both negative relative to the orientation. So ω_t is nowhere vanishing. Clearly $\omega_t = \omega_0$ at $t = 0$ and $\omega_t = \omega_1$ at $t = 1$. Since $d\omega_t = 0$ as ω_t is an n-from on an n-dimensional manifold,

$$D_{v_t} = di(v_t)\omega_t$$

by Weil's formula. Also

$$\dot{\omega}_t = \omega_1 - \omega_0.$$

$$D_{v_t}\omega_t = -\dot{\omega}_t. \tag{11}$$

$$\dot{\omega}_t = \omega_1 - \omega_0.$$

Since $\int_M \omega_0 = \int_M \omega_1$ we know that

$$\omega_0 - \omega_1 = d\nu$$

for some $(n-1)$-form ν. Thus (11) becomes

$$di(v_t)\omega_t = d\nu.$$

We will certainly have solved this equation if we solve the harder equation

$$i(v_t)\omega_t = \nu.$$

But this equation has a unique solution since ω_t is no-where vanishing. QED

The classical Morse lemma.

Let $M = \mathbb{R}^n$ and $\phi_i \in C^\infty(\mathbb{R}^n)$, $i = 0.1$. Suppose that 0 is a non-degenerate critical point for both ϕ_0 and ϕ_1, suppose that $\phi_0(0) = \phi_1(0) = 0$ and that they have the same Hessian at 0, i.e. suppose that

$$\left(d^2\phi_0\right)(0) = \left(d^2\phi_1\right)(0).$$

The Morse lemma asserts that there exist neighborhoods U_0 and U_1 of 0 in $\mathbb{R}^n$ and a diffeomorphism

$$f : U_0 \to U_1, \qquad f(0) = 0$$

such that

$$f^*\phi_1 = \phi_0.$$

Proof. Set

$$\phi_t := (1-t)\phi_0 + t\phi_1.$$

The Moser trick tells us to look for a vector field v_t with

$$v_t(0) = 0, \quad \forall\ t$$

and

$$D_{v_t}\phi_t = -\dot{\phi}_t = \phi_0 - \phi_1.$$

The function ϕ_t has a non-degenerate critical point at zero with the same Hessian as ϕ_0 and ϕ_1 and vanishes at 0. Thus for each fixed t, the functions

$$\frac{\partial \phi_t}{\partial x^i}$$

form a system of coordinates about the origin.

If we expand v_t in terms of the standard coordinates

$$v_t = \sum_j v_j(x,t) \frac{\partial}{\partial x^j}$$

then the condition $v_j(0,t) = 0$ implies that we must be able to write

$$v_j(x,t) = \sum_i v_{ij}(x,t) \frac{\partial \phi_t}{\partial x^i}.$$

for some smooth functions v_{ij}. Thus

$$D_{v_t}\phi_t = \sum_{ij} v_{ij}(x,t) \frac{\partial \phi_t}{\partial x^i} \frac{\partial \phi_t}{\partial x^j}.$$

Similarly, since $-\dot{\phi}_t$ vanishes at the origin together with its first derivatives, we can write

$$-\dot{\phi}_t = \sum_{ij} h_{ij} \frac{\partial \phi_t}{\partial x^i} \frac{\partial \phi_t}{\partial x^j}$$

where the h_{ij} are smooth functions. So the Moser equation $D_{v_t} \phi_t = -\dot{\phi}_t$ is satisfied if we set

$$v_{ij}(x, t) = h_{ij}(x, t).$$

Notice that our method of proof shows that if the ϕ_i depend smoothly on some paramters lying in a compact manifold S then the diffeomorphism f can be chosen so as to depend smoothly on $s \in S$.

In differential topology books the classical Morse lemma is usually stated as follows:

Theorem 1 *Let M be a manifold and $\phi : M \to \mathbb{R}$ be a smooth function. Suppose that $p \in M$ is a non-degenerate critical point of ϕ and that the signature of $d^2\phi_p$ is $(k, n-k)$. Then there exists a system of coordinates $(U, x_1, \ldots, x_n)$ centered at p such that in this coordinate system*

$$\phi = c + \sum_{i=1}^{k} x_i^2 - \sum_{i=k+1}^{n} x_i^2.$$

Proof. Choose any coordinate system $(W, y_1, \ldots y_n)$ centered about p and apply the previous result to

$$\phi_1 = \phi - c$$

and

$$\phi_0 = \sum h_{ij} y_i y_j$$

where

$$h_{ij} = \frac{\partial^2 \phi}{\partial y_i \partial y_i}(0).$$

This gives a change of coordinates in terms of which $\phi - c$ has become a non-degenerate quadratic form. Now apply Sylvester's theorem in linear algebra which says that a linear change of variables can bring such a non-degenerate quadratic form to the desired diagonal form.