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

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

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

$$\mathcal{C}yc\ [X, [Y, Z]] = 0. \tag{2.6}$$

Exercises

2.13 Left invariant forms.

Let G be a group and M be a set. A **left action** of G on M consists of a map

$$\phi : G \times M \to M$$

satisfying the conditions

$$\phi(a, \phi(b, m))) = \phi(ab, m)$$

(an associativity law) and

$$\phi(e, m) = m, \quad \forall m \in M$$

where e is the identity element of the group. When there is no risk of confusion we will write am for $\phi(a, m)$.(But in much of the beginning of the following exercises there *will* be a risk of confusion since there will be several different actions of the same group G on the set M). We think of an action as assigning to each element $a \in G$ a transformation, ϕ_a, of M:

$$\phi_a : M \to M, \quad \phi_a : m \mapsto \phi(a, m).$$

So we also use the notation

$$\phi_a m = \phi(a, m).$$

For example, we may take M to be the group G itself and let the action be left multiplication, L, so

$$L(a,m) = am.$$

We will write

$$L_a : G \to G, \quad L_a m = am.$$

We may may also consider the (left) action of right multiplication:

$$R : G \times G \to G, \quad R(a,m) = ma^{-1}.$$

(The inverse is needed to get the order right in $R(a, R(b,m)) = R(ab,m)$.) So we will write

$$R_a : G \to G, \quad R_a m = ma^{-1}.$$

We will be interested in the case that G is a Lie group, which means that G is a manifold and the multiplication map $G \times G \to G$ and the inverse map $G \to G, \quad a \mapsto a^{-1}$ are both smooth maps. Then the differential, $(dL_a)_m$ maps the tangent space to G at m, to the tangent space to G at am:

$$dL_a : TG_m \to TG_{am}$$

and similarly

$$dR_a : TG_m \to TG_{ma}.$$

In particular,

$$dL_{a^{-1}} : TG_a \to TG_e.$$

Let $G = Gl(n)$ be the group of all invertible $n \times n$ matrices. It is an open subset (hence a submanifold) of the n^2 dimensional space $\mathrm{Mat}(n)$ of all $n \times n$ matrices. We can think of the tautological map which sends every $A \in G$ into itself thought of as an element of $\mathrm{Mat}(n)$ as a matrix valued function on G. Put another way, A is a matrix of functions on G, each of the matrix entries A_{ij} of A is a function on G. Hence $dA = (dA_{ij})$ is a matrix of differential forms (or, we may say, a matrix valued differential form). So we may consider

$$A^{-1}dA$$

which is also a matrix valued differential form on G. Let B be a fixed element of G.

1. Show that

$$L_B^*(A^{-1}dA) = A^{-1}dA. \tag{2.7}$$

So each of the entries of $A^{-1}dA$ is *left invariant*.

2. Show that

$$R_B^*(A^{-1}dA) = B(A^{-1}dA)B^{-1}. \tag{2.8}$$

So the entries of $A^{-1}dA$ are not right invariant (in general), but (2.8) shows how they are transformed into one another by right multiplication.

For any two matrix valued differential forms $R = (R_{ij})$ and $S = (S_{ij})$ define their matrix exterior product $R \wedge S$ by the usual formula for matrix product, but with exterior multiplication of the entries instead of ordinary multiplication, so

$$(R \wedge S)_{ik} := \sum_j R_{ij} \wedge S_{jk}.$$

Also, if $R = (R_{ij})$ is a matrix valued differential form, define dR by applying d to each of the entries. So

$$(dR)_{ij} := (dR_{ij}).$$

Finally, if $\psi : X \to Y$ is a smooth map and $R = (R_{ij})$ is a matrix valued form on Y then we define its pullback by pulling back each of the entries:

$$(\psi^* R)_{ij} := (\psi^* R_{ij}).$$

2.14 The Maurer Cartan equations.

3. In elementary calculus we have the formula $d(1/x) = -dx/x^2$. What is the generalization of this formula for the matrix function A^{-1}. In other words, what is the formula for $d(A^{-1})$?

4. Show that if we set $\omega = A^{-1}dA$ then

$$d\omega + \omega \wedge \omega = 0. \tag{2.9}$$

Here is another way of thinking about $A^{-1}dA$: Since $G = Gl(n)$ is an open subset of the vector space $\mathrm{Mat}(n)$, we may identify the tangent space TG_A with the vector space $\mathrm{Mat}(n)$. That is we have an isomorphism between TG_A and $\mathrm{Mat}(n)$. If you think about it for a minute, it is the form dA which effects this isomorphism at every point. On the other hand, left multiplication by A^{-1} is a linear map. Under this identification, the differential of a linear map L looks just like L. So in terms of this identification, $A^{-1}dA$, when evaluated at the tangent space TG_A is just the isomorphism $dL_A^{-1} : TG_A \to TG_I$ where I is the identity matrix.

2.15 Restriction to a subgroup

Let H be a Lie subgroup of G. This means that H is a subgroup of G and it is also a submanifold. In other words we have an embedding

$$\iota : H \to G$$

which is a(n injective) group homomorphism.Let

$$\mathfrak{h} = TH_I$$

denote the tangent space to H at the identity element.

5. Conclude from the preceding discussion that if we now set

$$\omega = \iota^*(A^{-1}dA)$$

then ω takes values in $\mathfrak{h}$. In other words, when we evaluate ω on any tangent vector at any point of H we get a matrix belonging to the subspace $\mathfrak{h}$.

6. Show that on a group, the only transformations which commute with all the right multiplications, $R_b, \quad b \in G$, are the left multiplications, L_a.

For any vector $\xi \in TH_I$, define the vector field X by

$$X(A) = dR_{A^{-1}}\xi.$$

(Recall that $R_{A^{-1}}$ is right multiplication by A and so sends I into A.) For example, if we take H to be the full group $G = Gl(n)$ and identify the tangent space at every point with $\mathrm{Mat}(n)$ then the above definition becomes

$$X(A) = \xi A.$$

By construction, the vector field X is right invariant, i.e. is invariant under all the diffeomorphisms R_B.

7. Conclude that the flow generated by X is left multiplication by a one parameter subgroup. Also conclude that in the case $H = Gl(n)$ the flow generated by X is left multiplication by the one parameter group

$$\exp t\xi.$$

Finally conclude that for a general subgroup H, if $\xi \in \mathfrak{h}$ then all the $\exp t\xi$ lie in H.

8. What is the space h in the case that H is the group of Euclidean motions in three dimensional space, thought of as the set of all four by four matrices of the form

$$\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix}, \quad AA^\dagger = I, \ v \in \mathbf{R}^3?$$

2.16 Frames.

Let V be an n dimensional vector space. Recall that **frame** on V is, by definition, an isomorphism $\mathbf{f} : \mathbf{R}^n \to V$. Giving $\mathbf{f}$ is the same as giving each of the vectors $f_i = \mathbf{f}(\delta_i)$ where the δ_i range over the standard basis of $\mathbf{R}^n$. So giving a frame is the same as giving an ordered basis of V and we will sometimes write

$$\mathbf{f} = (f_1, \ldots, f_n).$$

If $A \in Gl(n)$ then A is an isomorphism of $\mathbf{R}^n$ with itself, so $\mathbf{f} \circ A^{-1}$ is another frame. So we get an action, $R : Gl(n) \times \mathbf{F} \to \mathbf{F}$ where $\mathbf{F} = \mathbf{F}(V)$ denotes the space of all frames:

$$R(A, \mathbf{f}) = \mathbf{f} \circ A^{-1}. \tag{2.10}$$

If $\mathbf{f}$ and $\mathbf{g}$ are two frames, then $\mathbf{g}^{-1} \circ \mathbf{f} = M$ is an isomorphism of $\mathbf{R}^n$ with itself, i.e. a matrix. So given any two frames, $\mathbf{f}$ and $\mathbf{g}$, there is a unique $M \in Gl(n)$ so that $\mathbf{g} = \mathbf{f} \circ M^{-1}$. Once we fix an $\mathbf{f}$, we can use this fact to identify $\mathbf{F}$ with $Gl(n)$, but the identification depends on the choice of $\mathbf{f}$. But in any event the (non-unique) identification shows that $\mathbf{F}$ is a manifold and that (2.10) defines an action of $Gl(n)$ on $\mathbf{F}$. Each of the f_i (the $i-$th basis vector in the frame) can be thought of as a V valued function on $\mathbf{F}$. So we may write

$$df_j = \sum \omega_{ij} f_i \tag{2.11}$$

where the ω_{ij} are ordinary (number valued) linear differential forms on $\mathbf{F}$. We think of this equation as giving the expansion of an infinitesimal change in f_j in terms of the basis $\mathbf{f} = (f_1, \ldots, f_n)$. If we use the "row" representation of $\mathbf{f}$ as above, we can write these equations as

$$d\mathbf{f} = \mathbf{f}\omega \tag{2.12}$$

where $\omega = (\omega_{ij})$.

9. Show that the ω defined by (2.12) satisfies

$$R_B^*\omega = B\omega B^{-1}. \tag{2.13}$$

To see the relation with what went on before, notice that we *could* take $V = \mathbf{R}^n$ itself. Then $\mathbf{f}$ is just an invertible matrix, A and (2.12) becomes our old equation $\omega = A^{-1}dA$. So (2.13) reduces to (2.8).

If we take the exterior derivative of (2.12) we get

$$0 = d(d\mathbf{f}) = d\mathbf{f} \wedge \omega + \mathbf{f} d\omega = \mathbf{f}(\omega \wedge \omega + d\omega)$$

from which we conclude

$$d\omega + \omega \wedge \omega = 0. \tag{2.14}$$

2.17 Euclidean frames.

We specialize to the case where $V = \mathbf{R}^n, n = d + 1$ so that the set of frames becomes identified with the group $Gl(n)$ and restrict to the subgroup, H, of

Euclidean motions which consist of all $n \times m$ matrices of the form

$$\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix}, \quad A \in O(d),\ v \in \mathbf{R}^d.$$

Such a matrix, when applied to a vector

$$\begin{pmatrix} w \\ 1 \end{pmatrix}$$

sends it into the vector

$$\begin{pmatrix} Aw + v \\ 1 \end{pmatrix}$$

and $Aw + v$ is the orthogonal transformation A applied to w followed by the translation by v. The corresponding *Euclidean frames* (consisting of the columns of the elements of H) are thus defined to be the frames of the form

$$f_i = \begin{pmatrix} e_i \\ 0 \end{pmatrix}, \quad i = 1, \dots d,$$

where the e_i form an orthonormal basis of $\mathbf{R}^d$ and

$$f_n = \begin{pmatrix} v \\ 1 \end{pmatrix},$$

where $v \in \mathbf{R}^d$ is an arbitrary vector. The idea is that v represents a choice of origin in d dimensional space and $\mathbf{e} = (e_1, \dots, e_d)$ is an orthonormal basis. We can write this in shorthand notation as

$$\mathbf{f} = \begin{pmatrix} \mathbf{e} & v \\ 0 & 1 \end{pmatrix}.$$

If ι denotes the embedding of H into G, we know from the exercise **5** that

$$\iota^*\omega = \begin{pmatrix} \Omega & \theta \\ 0 & 0 \end{pmatrix},$$

where

$$\Omega_{ij} = -\Omega_{ji}.$$

So the pull back of (2.12) becomes

$$d \begin{pmatrix} \mathbf{e} & v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{e}\Omega & \mathbf{e}\theta \\ 0 & 0 \end{pmatrix} \tag{2.15}$$

or, in more expanded notation,

$$de_j = \sum_i \Omega_{ij} e_i, \quad dv = \sum_i \theta_i e_i.$$

Let (,) denote the Euclidean scalar product. Then we can write

$$\theta_i = (dv, e_i) \tag{2.16}$$

and

$$(de_j, e_i) = \Omega_{ij}.$$

If we set

$$\Theta = -\Omega$$

this becomes

$$(de_i, e_j) = \Theta_{ij}. \tag{2.17}$$

Then (2.14) becomes

$$d\theta = \Theta \wedge \theta, \quad d\Theta = \Theta \wedge \Theta. \tag{2.18}$$

Or, in more expanded notation,

$$d\theta_i = \sum_j \Theta_{ij} \wedge \theta_j, \quad d\Theta_{ik} = \sum_j \Theta_{ij} \wedge \Theta_{jk}. \tag{2.19}$$

Equations (2.16)-(2.18) or (2.19) are known as the **structure equations of Euclidean geometry**.

2.18 Frames adapted to a submanifold.

Let M be a k dimensional submanifold of $\mathbf{R}^d$. This determines a submanifold of the manifold, H, of all Euclidean frames by the following requirements:

i) $v \in M$ and

ii) $e_i \in TM_v$ for $i \leq k$. We will usually write m instead of v to emphasize the first requirement - that the frames be based at points of M. The second requirement says that the first k vectors in the frame based at m be tangent to M (and hence that the last $n-k$ vectors in the frame are normal to M). We will denote this manifold by $\mathcal{O}(M)$. It has dimension

$$k + \frac{k(k-1)}{2} + \frac{(d-k-1)(d-k)}{2}.$$

The first term comes from the point m varying on M, the second is the dimension of the orthogonal group $O(k)$ corresponding to the choices of the first k vectors in the frame, and the third term is dim $O(d-k)$ correspond to the last $(n-k)$ vectors. We have an embedding of $\mathcal{O}(M)$ into H, and hence the forms θ and Θ pull back to $\mathcal{O}(M)$. As we are running out of letters, we will continue to denote these pull backs by the same letters. So the pulled back forms satisfy the same structure equations (2.16)-(2.18) or (2.19) as above, but they are supplemented by

$$\theta_i = 0, \quad \forall i > k. \tag{2.20}$$

2.19 Curves and surfaces - their structure equations.

We will be particularly interested in curves and surfaces in three dimensional Euclidean space. For a curve, C, the manifold of frames is two dimensional, and we have

$$\begin{aligned} dC &= \theta_1 e_1 && (2.21)\\ de_1 &= \Theta_{12}e_2 + \Theta_{13}e_3 && (2.22)\\ de_2 &= \Theta_{21}e_1 + \Theta_{23}e_3 && (2.23)\\ de_3 &= \Theta_{31}e_1 + \Theta_{32}e_2. && (2.24) \end{aligned}$$

One can visualize the manifold of frames as a sort of tube: about each point of the curve there is a circle in the plane normal to the tangent line corresponding the possible choices of e_2.

For the case of a surface the manifold of frames is three dimensional: we can think of it as a union of circles each centered at a point of S and in the plane tangent to S at that point. Then equation (2.21) is replaced by

$$dX = \theta_1 e_1 + \theta_2 e_2 \qquad (2.25)$$

but otherwise the equations are as above, including the structure equations (2.19). These become

$$\begin{aligned} d\theta_1 &= \Theta_{12} \wedge \theta_2 && (2.26)\\ d\theta_2 &= -\Theta_{12} \wedge \theta_1 && (2.27)\\ 0 &= \Theta_{31} \wedge \theta_1 + \Theta_{32} \wedge \theta_2 && (2.28)\\ d\Theta_{12} &= \Theta_{13} \wedge \Theta_{32} && (2.29)\\ d\Theta_{13} &= \Theta_{12} \wedge \Theta_{23} && (2.30)\\ d\Theta_{23} &= \Theta_{21} \wedge \Theta_{13} && (2.31) \end{aligned}$$

Equation (2.29) is known as Gauss' equation, and equations (2.30) and (2.31) are known as the Codazzi-Mainardi equations.

2.20 The sphere as an example.

In computations with local coordinates, we may find it convenient to use a "cross-section" of the manifold of frames, that is a map which assigns to each point of neighborhood on the surface a preferred frame. If we are given a parametrization $m = m(u, v)$ of the surface, one way of choosing such a cross-section is to apply the Gram-Schmidt orthogonalization procedure to the tangent vector fields m_u and m_v, and take into account the chosen orientation.

For example, consider the sphere of radius R. We can parameterize the sphere with the north and south poles (and one longitudinal semi-circle) removed by the $(u,v) \in (0,2\pi) \times (0,\pi)$ by $X = X(u,v)$ where

$$X(u,v) = \begin{pmatrix} R\cos u \sin v \\ R\sin u \sin v \\ R\cos v \end{pmatrix}.$$

Here v denotes the angular distance from the north pole, so the excluded value $v = 0$ corresponds to the north pole and the excluded value $v = \pi$ corresponds to the south pole. Each constant value of v between 0 and π is a circle of latitude with the equator given by $v = \frac{\pi}{2}$. The parameter u describes the longitude from the excluded semi-circle.

In any frame adapted to a surface in $\mathbf{R}^3$, the third vector e_3 is normal to the surface at the base point of the frame. There are two such choices at each base point. In our sphere example let us choose the outward pointing normal, which at the point $m(u,v)$ is

$$e_3(m(u,v)) = \begin{pmatrix} cosu \sin v \\ \sin u \sin v \\ cosv \end{pmatrix}.$$

We will write the left hand side of this equation as $e_3(u,v)$. The coordinates u,v are orthogonal, i.e. X_u and X_v are orthogonal at every point, so the orthonormalization procedure amounts only to normalization: Replace each of these vectors by the unit vectors pointing in the same direction at each point. So we get

$$e_1(u,v) = \begin{pmatrix} -\sin u \\ \cos u \\ 0 \end{pmatrix}, \qquad e_2(u,v) = \begin{pmatrix} \cos u \cos v \\ \sin u \cos v \\ -\sin v \end{pmatrix}.$$

We thus obtain a map ψ from $(0,2\pi) \times (0,\pi)$ to the manifold of frames,

$$\psi(u,v) = (X(u,v), e_1(u,v), e_2,(u,v), e_3(u,v)).$$

Since $X_u \cdot e_1 = R\sin v$ and $X_v \cdot e_2 = R$ we have

$$dX(u,v) = (R\sin v du)e_1(u,v) + (Rdv)e_2(u,v).$$

Thus we see from (2.25) that

$$\psi^*\theta_1 = R\sin v du, \quad \psi^*\theta_2 = Rdv$$

and hence that

$$\psi^*(\theta_1 \wedge \theta_2) = R^2 \sin v du \wedge dv.$$

Now $R^2 \sin v du dv$ is just the area element of the sphere expressed in u,v coordinates. The choice of e_1, e_2 determines an orientation of the tangent space to the sphere at the point $X(u,v)$ and so $\psi^*(\theta_1 \wedge \theta_2)$ is the pull-back of the corresponding oriented area form.

10. Compute $\psi^*\Theta_{12},\ \psi^*\Theta_{13}$, and $\psi^*\Theta_{23}$ and verify that

$$\psi^*(d\Theta_{12}) = -\psi^*(K)\psi^*(\theta_1 \wedge \theta_2)$$

where $K = 1/R^2$ is the curvature of the sphere.

We will generalize this equation to an arbitrary surface in $\mathbf{R}^3$ in section **??**.

2.21 Ribbons

The idea here is to study a curve on a surface, or rather a curve with an "infinitesimal" neighborhood of a surface along it. So let C be a curve and $\mathcal{O}(C)$ its associated two dimensional manifold of frames. We have a projection $\pi : \mathcal{O}(C) \to C$ sending every frame into its origin. By a **ribbon** based on C we mean a section $n : C \to \mathcal{O}(C)$, so n assigns a unique frame to each point of the curve in a smooth way. We will only be considering curves with non-vanishing tangent vector everywhere. With no loss of generality we may assume that we have parametrized the curve by arc length, and the choice of e_1 determines an orientation of the curve, so $\theta = ds$. The choice of e_2 at every point then determines e_3 up to a $\pm$ sign. So a good way to visualize s is to think of a rigid metal ribbon determined by the curve and the vectors e_2 perpendicular to the curve (determined by n) at each point. The forms Θ_{ij} all pull back under n to function multiples of ds:

$$n^*\Theta_{12} = kds, \quad n^*\Theta_{23} = -\tau ds, \quad n^*\Theta_{13} = wds \tag{2.32}$$

where k, τ and w are functions of s. We can write equations (2.21)- (2.24) above as

$$\frac{dC}{ds} = e_1,$$

and

$$\frac{de_1}{ds} = ke_2 + we_3, \quad \frac{de_2}{ds} = -ke_1 - \tau e_3, \quad \frac{de_3}{ds} = -we_1 + \tau e_3. \tag{2.33}$$

For later applications we will sometimes be sloppy and write Θ_{ij} instead of $n^*\Theta_{ij}$ for the pull back to the curve, so along the ribbon we have $\Theta_{12} = kds$ etc. Also it will sometimes be convenient in computations (as opposed to proving theorems) to use parameters other than arc length.

11. Show that two ribbons (defined over the same interval of s values) are congruent (that is there is a Euclidean motion carrying one into the other) if and only if the functions k, τ, and w are the same.

A ribbon is really just a curve in the space, H, of all Euclidean frames, having the property that the base point, that is the v of the frame (v, e_1, e_2, e_3) has non-vanishing derivative. The previous exercise says that two curves, $i : I \to H$ and $j : I \to H$ in H differ by an overall left translation (that is satisfy $j = L_h \circ i$)if and only if the forms $\theta, \Theta_{12}, \Theta_{13}, \Theta_{23}$ pull back to the same forms on I. The