C

Connections and Curvature

Introduction

In this appendix we present results in differential geometry that serve as a useful background for material in the main body of the book. Material in §1 on connections is somewhat parallel to the study of the natural connection on a Riemannian manifold made in §11 of Chapter 1, but here we also study the curvature of a connection. Material in §2 on second covariant derivatives is connected with material in Chapter 2 on the Laplace operator. Ideas developed in §§3 and 4, on the curvature of Riemannian manifolds and submanifolds, make contact with such material as the existence of complex structures on two-dimensional Riemannian manifolds, established in Chapter 5, and the uniformization theorem for compact Riemann surfaces and other problems involving nonlinear, elliptic PDE, arising from studies of curvature, treated in Chapter 14. Section 5 on the Gauss-Bonnet theorem is useful both for estimates related to the proof of the uniformization theorem and for applications to the Riemann-Roch theorem in Chapter 10. Furthermore, it serves as a transition to more advanced material presented in §§6–8.

In §6 we discuss how constructions involving vector bundles can be derived from constructions on a principal bundle. In the case of ordinary vector fields, tensor fields, and differential forms, one can largely avoid this, but it is a very convenient tool for understanding spinors. The principal bundle picture is used to construct characteristic classes in §7. The material in these two sections is needed in Chapter 10, on the index theory for elliptic operators of Dirac type. In §8 we show how one particular characteristic class, arising from the Pfaffian, figures into the higher-dimensional version of the Gauss-Bonnet formula. The proof given here is geometrical and uses the elements of Morse theory. In Chapter 10 this result is derived as a special case of the Atiyah-Singer index formula.

1. Covariant derivatives and curvature on general vector bundles

Let $E \to M$ be a vector bundle, either real or complex. A *covariant derivative*, or *connection*, on E is a map

$$\nabla_X : C^\infty(M,E) \longrightarrow C^\infty(M,E) \tag{1.1}$$

assigned to each vector field X on M, satisfying the following three conditions:

$$\nabla_X(u+v) = \nabla_X u + \nabla_X v, \tag{1.2}$$

$$\nabla_{(fX+Y)} u = f\nabla_X u + \nabla_Y u, \tag{1.3}$$

$$\nabla_X(fu) = f\nabla_X u + (Xf)u, \tag{1.4}$$

where u, v are sections of E, and f is a smooth scalar function. The examples contained in Chapters 1 and 2 are the Levi-Civita connection on a Riemannian manifold, in which case E is the tangent bundle, and associated connections on tensor bundles, discussed in §2.2.

One general construction of connections is the following. Let F be a vector space, with an inner product; we have the trivial bundle $M \times F$. Let E be a subbundle of this trivial bundle; for each $x \in M$, let P_x be the orthogonal projection of F on $E_x \subset F$. Any $u \in C^\infty(M,E)$ can be regarded as a function from M to F, and for a vector field X, we can apply X componentwise to any function on M with values in F; call this action $u \mapsto D_X u$. Then a connection on M is given by

$$\nabla_X u(x) = P_x D_X u(x). \tag{1.5}$$

If M is imbedded in a Euclidean space $\mathbb{R}^N$, then $T_x M$ is naturally identified with a linear subspace of $\mathbb{R}^N$ for each $x \in M$. In this case it is easy to verify that the connection defined by (1.5) coincides with the Levi-Civita connection, where M is given the metric induced from its imbedding in $\mathbb{R}^N$. Compare with the discussion of submanifolds in §4 below.

Generally, a connection defines the notion of "parallel transport" along a curve γ in M. A section u of E over γ is obtained from $u(\gamma(t_0))$ by parallel transport if it satisfies $\nabla_T u = 0$ on γ, where $T = \dot{\gamma}(t)$.

Formulas for covariant derivatives, involving indices, are produced in terms of a choice of "local frame" for E, that is, a set e_α, $1 \le \alpha \le K$, of sections of E over an open set U which forms a basis of E_x for each $x \in U; K = \dim E_x$. Given such a local frame, a smooth section u of E over U is specified by

$$u = u^\alpha e_\alpha \quad \text{(summation convention)}. \tag{1.6}$$

If $D_j = \partial/\partial x_j$ in a coordinate system on U, we set

$$\nabla_{D_j} u = u^\alpha{}_{;j} e_\alpha = (\partial_j u^\alpha + u^\beta \Gamma^\alpha{}_{\beta j}) e_\alpha, \tag{1.7}$$

the connection coefficients $\Gamma^\alpha{}_{\beta j}$ being defined by

$$\nabla_{D_j} e_\beta = \Gamma^\alpha{}_{\beta j} e_\alpha. \tag{1.8}$$

A vector bundle $E \to M$ may have an inner product on its fibers. In that case, a connection on E is called a *metric connection* provided that

$$X\langle u, v\rangle = \langle \nabla_X u, v\rangle + \langle u, \nabla_X v\rangle, \tag{1.9}$$

for any vector field X and smooth sections u, v of E.

The curvature of a connection is defined by

$$R(X,Y)u = [\nabla_X, \nabla_Y]u - \nabla_{[X,Y]}u, \tag{1.10}$$

where X and Y are vector fields and u is a section of E. It is easy to verify that (1.10) is linear in X, Y, and u, over $C^\infty(M)$. With respect to local coordinates, giving $D_j = \partial/\partial x_j$, and a local frame $\{e_\alpha\}$ on E, as in (1.6), we define the components $R^\alpha{}_{\beta jk}$ of the curvature by

$$R(D_j, D_k)e_\beta = R^\alpha{}_{\beta jk} e_\alpha, \tag{1.11}$$

as usual, using the summation convention. Since D_j and D_k commute, $R(D_j, D_k)e_\beta = [\nabla_{D_j}, \nabla_{D_k}]e_\beta$. Applying the formulas (1.7) and (1.8), we can express the components of R in terms of the connection coefficients. The formula is seen to be

$$R^\alpha{}_{\beta jk} = \partial_j \Gamma^\alpha{}_{\beta k} - \partial_k \Gamma^\alpha{}_{\beta j} + \Gamma^\alpha{}_{\gamma j} \Gamma^\gamma{}_{\beta k} - \Gamma^\alpha{}_{\gamma k} \Gamma^\gamma{}_{\beta j}. \tag{1.12}$$

The formula (1.12) can be written in a shorter form, as follows. Given a choice of local frame $\{e_\alpha : 1 \le \alpha \le K\}$, we can define $K \times K$ matrices $\Gamma_j = (\Gamma^\alpha{}_{\beta j})$ and also $\mathfrak{R}_{jk} = (R^\alpha{}_{\beta jk})$. Then (1.12) is equivalent to

$$\mathfrak{R}_{jk} = \partial_j \Gamma_k - \partial_k \Gamma_j + [\Gamma_j, \Gamma_k]. \tag{1.13}$$

Note that $\mathfrak{R}_{jk}$ is antisymmetric in j and k. Now we can define a "connection 1-form" Γ and a "curvature 2-form" Ω by

$$\Gamma = \sum_j \Gamma_j \, dx_j, \quad \Omega = \frac{1}{2} \sum_{j,k} \mathfrak{R}_{jk} \, dx_j \wedge dx_k. \tag{1.14}$$

Then the formula (1.12) is equivalent to

$$\Omega = d\Gamma + \Gamma \wedge \Gamma. \tag{1.15}$$

The curvature has symmetries, which we record here, for the case of general vector bundles. The Riemann curvature tensor, associated with the Levi-Civita connection, has additional symmetries, which will be described in §3.

Proposition 1.1. *For any connection ∇ on $E \to M$, we have*

$$R(X,Y)u = -R(Y,X)u. \tag{1.16}$$

If ∇ is a metric connection, then

$$\langle R(X,Y)u,v\rangle = -\langle u, R(X,Y)v\rangle. \tag{1.17}$$

Proof. Equation (1.16) is obvious from the definition (1.10); this is equivalent to the antisymmetry of ${R^\alpha}_{\beta jk}$ in j and k noted above. If ∇ is a metric connection, we can use (1.9) to deduce

$$\begin{aligned} 0 &= (XY - YX - [X,Y])\langle u,v\rangle \\ &= \langle R(X,Y)u,v\rangle + \langle u, R(X,Y)v\rangle, \end{aligned}$$

which gives (1.17).

Next we record the following implication of a connection having *zero* curvature. A section u of E is said to be "parallel" if $\nabla_X u = 0$ for all vector fields X.

Proposition 1.2. *If $E \to M$ has a connection ∇ whose curvature is zero, then any $p \in M$ has a neighborhood U on which there is a frame $\{e_\alpha\}$ for E consisting of parallel sections: $\nabla_X e_\alpha = 0$ for all X.*

Proof. If U is a coordinate neighborhood, then e_α is parallel provided $\nabla_j e_\alpha = 0$ for $j = 1, \dots, n = \dim M$. The condition that $R = 0$ is equivalent to the condition that the operators ∇_{D_j} all commute with each other, for $1 \le j \le n$. Consequently, Frobenius's theorem (as expanded in Exercise 5 in §9 of Chapter 1) allows us to solve the system of equations

$$\nabla_{D_j} e_\alpha = 0, \quad j = 1, \dots, n, \tag{1.18}$$

on a neighborhood of p, with e_α prescribed at the point p. If we pick $e_\alpha(p)$, $1 \le \alpha \le K$, to be a basis of E_p, then $e_\alpha(x)$, $1 \le \alpha \le K$, will be linearly independent in E_x for x close to p, so the local frame of parallel sections is constructed.

It is useful to note, in general, several formulas that result from choosing a local frame $\{e_\alpha\}$ by parallel translation along rays through a point $p \in M$, the origin in some coordinate system $(x_1, \dots, x_n)$, so

$$\nabla_{r\partial/\partial r} e_\alpha = 0, \quad 1 \le \alpha \le K. \tag{1.19}$$

This means $\sum x_j \nabla_{D_j} e_\alpha = 0$. Consequently, the connection coefficients (1.8) satisfy

$$x_1 {\Gamma^\alpha}_{\beta 1} + \cdots + x_n {\Gamma^\alpha}_{\beta n} = 0. \tag{1.20}$$

Differentiation with respect to x_j gives

$${\Gamma^\alpha}_{\beta j} = -x_1 \partial_j {\Gamma^\alpha}_{\beta 1} - \cdots - x_n \partial_j {\Gamma^\alpha}_{\beta n}. \tag{1.21}$$

In particular,

$$\Gamma^\alpha{}_{\beta j}(p) = 0. \tag{1.22}$$

Comparison of (1.21) with

$$\Gamma^\alpha{}_{\beta j} = x_1 \partial_1 \Gamma^\alpha{}_{\beta j}(p) + \cdots + x_n \partial_n \Gamma^\alpha{}_{\beta j}(p) + O(|x|^2) \tag{1.23}$$

gives

$$\partial_k \Gamma^\alpha{}_{\beta j} = -\partial_j \Gamma^\alpha{}_{\beta k}, \quad \text{at } p. \tag{1.24}$$

Consequently, the formula (1.12) for curvature becomes

$$R^\alpha{}_{\beta jk} = 2\ \partial_j \Gamma^\alpha{}_{\beta k}, \quad \text{at } p, \tag{1.25}$$

with respect to such a local frame. Note that, near p,

$$R^\alpha{}_{\beta jk} = \partial_j \Gamma^\alpha{}_{\beta k} - \partial_k \Gamma^\alpha{}_{\beta j} + O(|x|^2). \tag{1.26}$$

Given vector bundles $E_j \to M$ with connections ∇^j, there is a natural covariant derivative on the tensor-product bundle $E_1 \otimes E_2 \to M$, defined by the derivation property

$$\nabla_X(u \otimes v) = (\nabla^1_X u) \otimes v + u \otimes (\nabla^2_X v). \tag{1.27}$$

Also, if A is a section of $\operatorname{Hom}(E_1, E_2)$, the formula

$$(\nabla^\#_X A)u = \nabla^2_X (Au) - A(\nabla^1_X v) \tag{1.28}$$

defines a connection on $\operatorname{Hom}(E_1, E_2)$.

Regarding the curvature tensor R as a section of $(\otimes^2 T^*) \otimes \operatorname{End}(E)$ is natural in view of the linearity properties of R given after (1.10). Thus if $E \to M$ has a connection with curvature R, and if M also has a Riemannian metric, yielding a connection on T^*M, then we can consider $\nabla_X R$. The following, known as *Bianchi's identity*, is an important result involving the covariant derivative of R.

Proposition 1.3. *For any connection on $E \to M$, the curvature satisfies*

$$(\nabla_Z R)(X, Y) + (\nabla_X R)(Y, Z) + (\nabla_Y R)(Z, X) = 0, \tag{1.29}$$

or equivalently

$$R^\alpha{}_{\beta ij;k} + R^\alpha{}_{\beta jk;i} + R^\alpha{}_{\beta ki;j} = 0. \tag{1.30}$$

Proof. Pick any $p \in M$. Choose normal coordinates centered at p, and choose a local frame field for E by radial parallel translation, as above. Then, by (1.22) and (1.26),

$$R^\alpha{}_{\beta ij;k} = \partial_k \partial_i \Gamma^\alpha{}_{\beta j} - \partial_k \partial_j \Gamma^\alpha{}_{\beta i}, \quad \text{at } p. \tag{1.31}$$

Cyclically permuting (i, j, k) here and summing clearly give 0, proving the proposition.

Note that we can regard a connection on E as defining an operator

$$\nabla : C^\infty(M,E) \longrightarrow C^\infty(M, T^* \otimes E), \tag{1.32}$$

in view of the linear dependence of ∇_X on X. If M has a Riemannian metric and E a Hermitian metric, it is natural to study the adjoint operator

$$\nabla^* : C^\infty(M, T^* \otimes E) \longrightarrow C^\infty(M,E). \tag{1.33}$$

If u and v are sections of E, ξ a section of T^*, we have

$$\begin{aligned} \big(v, \nabla^*(\xi \otimes u)\big) &= (\nabla v, \xi \otimes u) \\ &= (\nabla_X v, u) \\ &= (v, \nabla_X^* u), \end{aligned} \tag{1.34}$$

where X is the vector field corresponding to ξ via the Riemannian metric. Using the divergence theorem we can establish:

Proposition 1.4. *If E has a metric connection, then*

$$\nabla^*(\xi \otimes u) = \nabla_X^* u = -\nabla_X u - (div\ X)u. \tag{1.35}$$

Proof. The first identity follows from (1.34) and does not require E to have a metric connection. If E does have a metric connection, integrating

$$\langle \nabla_X v, u\rangle = -\langle v, \nabla_X u\rangle + X\langle v, u\rangle$$

and using the identity

$$\int_M Xf\ dV = -\int_M (\mathrm{div}\ X) f\ dV, \quad f \in C_0^\infty(M), \tag{1.36}$$

give the second identity in (1.35) and complete the proof.

Exercises

1. If ∇ and $\widetilde{\nabla}$ are two connections on a vector bundle $E \to M$, show that

$$\nabla_X u = \widetilde{\nabla}_X u + C(X,u), \tag{1.37}$$

   where C is a smooth section of $\mathrm{Hom}(T \otimes E, E) \approx T^* \otimes \mathrm{End}(E)$. Show that conversely, if C is such a section and ∇ a connection, then (1.37) defines $\widetilde{\nabla}$ as a connection.

2. If ∇ and $\widetilde{\nabla}$ are related as in Exercise 1, show that their curvatures R and $\widetilde{R}$ are related by

$$(R - \widetilde{R})(X,Y)u = [C_X, \widetilde{\nabla}_Y]u - [C_Y, \widetilde{\nabla}_X]u - C_{[X,Y]}u + [C_X, C_Y]u, \tag{1.38}$$

   where C_X is the section of $\mathrm{End}(E)$ defined by $C_X u = C(X,u)$.

In Exercises 3–5, let $P(x),\ x \in M$, be a smooth family of projections on a vector space F, with range E_x, forming a vector bundle $E \to M$; E gets a natural connection via (1.5).

3. Let $\gamma : I \to M$ be a smooth curve through $x_0 \in M$. Show that parallel transport of $u(x_0) \in E_{x_0}$ along I is characterized by the following (with $P'(t) = dP(\gamma(t))/dt$):

$$\frac{du}{dt} = P'(t)u.$$

4. If each $P(x)$ is an orthogonal projection of the inner-product space F onto E_x, show that you get a metric connection. (*Hint*: Show that $du/dt \perp u(\gamma(t))$ via $P'P = (I-P)P'$.)
5. In what sense can $\Gamma = -dP\ P = -(I-P)\,dP$ be considered the connection 1-form, as in (1.13)? Show that the curvature form (1.15) is given by

$$\Omega = P\ dP \wedge dP\ P. \tag{1.39}$$

For more on this, see (4.50)–(4.53).

6. Show that the formula

$$d\Omega = \Omega \wedge \Gamma - \Gamma \wedge \Omega \tag{1.40}$$

follows from (1.15). Relate this to the Bianchi identity. Compare with (2.13) in the next section.

7. Let $E \to M$ be a vector bundle with connection ∇, with two local frame fields $\{e_\alpha\}$ and $\{f_\alpha\}$, defined over $U \subset M$. Suppose

$$f_\alpha(x) = g^\beta{}_\alpha(x)e_\beta(x), \quad e_\alpha(x) = h^\beta{}_\alpha(x)f_\beta(x);$$

note that $g^\beta{}_\gamma(x)h^\gamma{}_\alpha(x) = \delta^\beta{}_\alpha$. Let $\Gamma^\alpha{}_{\beta j}$ be the connection coefficients for the frame field $\{e_\alpha\}$, as in (1.7) and (1.8), and let $\widetilde{\Gamma}^\alpha{}_{\beta j}$ be the connection coefficients for the frame field $\{f_\alpha\}$. Show that

$$\widetilde{\Gamma}^\alpha{}_{\beta j} = h^\alpha{}_\mu \Gamma^\mu{}_{\gamma j} g^\gamma{}_\beta + h^\alpha{}_\gamma(\partial_j g^\gamma{}_\beta). \tag{1.41}$$

2. Second covariant derivatives and covariant-exterior derivatives

Let M be a Riemannian manifold, with Levi-Civita connection, and let $E \to M$ be a vector bundle with connection. In §1 we saw that the covariant derivative acting on sections of E yields an operator

$$\nabla : C^\infty(M,E) \longrightarrow C^\infty(M, T^* \otimes E). \tag{2.1}$$

Now on $T^* \otimes E$ we have the product connection, defined by (1.27), yielding

$$\nabla : C^\infty(T^* \otimes E) \longrightarrow C^\infty(M, T^* \otimes T^* \otimes E). \tag{2.2}$$

If we compose (2.1) and (2.2), we get a second-order differential operator called the *Hessian*:

$$\nabla^2 : C^\infty(M,E) \longrightarrow C^\infty(T^* \otimes T^* \otimes E). \tag{2.3}$$

If u is a section of E and X and Y are vector fields, (2.3) defines $\nabla^2_{X,Y}$ as a section of E; using the derivation properties, we have the formula

$$\nabla^2_{X,Y} u = \nabla_X \nabla_Y u - \nabla_{(\nabla_X Y)} u. \tag{2.4}$$

Note that the antisymmetric part is given by the curvature of the connection on E:

$$\nabla^2_{X,Y} u - \nabla^2_{Y,X} u = R(X,Y)u. \tag{2.5}$$

Now the metric tensor on M gives a linear map $T^* \otimes T^* \to \mathbb{R}$, hence a linear bundle map $\gamma : T^* \otimes T^* \otimes E \to E$. We can consider the composition of this with ∇^2 in (2.3):

$$\gamma \circ \nabla^2 : C^\infty(M,E) \longrightarrow C^\infty(M,E). \tag{2.6}$$

We want to compare $\gamma \circ \nabla^2$ and $\nabla^* \nabla$, in the case when E has a Hermitian metric and a metric connection.

Proposition 2.1. *If ∇ is a metric connection on E, then*

$$\nabla^* \nabla = -\gamma \circ \nabla^2 \quad \text{on } C^\infty(M,E). \tag{2.7}$$

Proof. Pick a local orthonormal frame of vector fields $\{e_j\}$, with dual frame $\{v_j\}$. Then, for $u \in C^\infty(M,E)$, $\nabla u = \sum v_j \otimes \nabla_{e_j} u$, so (1.35) implies

$$\nabla^* \nabla u = \sum \bigl[-\nabla_{e_j} \nabla_{e_j} u - (\operatorname{div} e_j) u \bigr]. \tag{2.8}$$

Using (2.4), we have

$$\nabla^* \nabla u = -\sum \nabla^2_{e_j,e_j} u - \sum \bigl[\nabla_{\nabla_{e_j} e_j} u + (\operatorname{div} e_j) \nabla_{e_j} u \bigr]. \tag{2.9}$$

The first term on the right is equal to $-\gamma \circ \nabla^2 u$. Now, given $p \in M$, if we choose the local frame $\{e_j\}$ such that $\nabla_{e_j} e_k = 0$ at p, the rest of the right side vanishes at p. This establishes the identity (2.7).

We next define a "covariant-exterior derivative" operator

$$d^\nabla : C^\infty(M, \Lambda^k T^* \otimes E) \longrightarrow C^\infty(M, \Lambda^{k+1} T^* \otimes E) \tag{2.10}$$

as follows. For $k = 0$, $d^\nabla = \nabla$, given by (2.1), and we require

$$d^\nabla(\beta \wedge u) = (d\beta) \wedge u - \beta \wedge d^\nabla u \tag{2.11}$$

whenever β is a 1-form and u is a section of $\Lambda^k T^* \otimes E$. The operator d^∇ is also called the "gauge exterior derivative." Unlike the case of the ordinary exterior derivative,

$$d^\nabla \circ d^\nabla : C^\infty(M, \Lambda^k T^* \otimes E) \longrightarrow C^\infty(M, \Lambda^{k+2} T^* \otimes E)$$

is not necessarily zero, but rather

$$d^\nabla d^\nabla u = \Omega \wedge u, \tag{2.12}$$

where Ω is the curvature, and we use the antisymmetry (1.16) to regard Ω as a section of $\Lambda^2T^* \otimes \operatorname{End}(E)$, as in (1.15). The verification of (2.12) is a straightforward calculation; (2.5) is in fact the special case of this, for $k = 0$.

The following is an alternative form of Bianchi's identity (1.29):

$$d^\nabla \Omega = 0, \tag{2.13}$$

where the left side is a priori a section of $\Lambda^3T^* \otimes \operatorname{End}(E)$. This can also be deduced from (2.12), the associative law $d^\nabla(d^\nabla d^\nabla) = (d^\nabla d^\nabla)d^\nabla$, and the natural derivation property generalizing (2.11):

$$d^\nabla(A \wedge u) = (d^\nabla A) \wedge u + (-1)^j A \wedge d^\nabla u, \tag{2.14}$$

where u is a section of $\Lambda^kT^* \otimes E$ and A is a section of $\Lambda^jT^* \otimes \operatorname{End}(E)$.

Exercises

1. Let $E \to M$ be a vector bundle with connection ∇, $u \in C^\infty(M, E)$. Fix $p \in M$. Show that if $\nabla u(p) = 0$, then $\nabla^2_{X,Y}u(p)$ is independent of the choice of connection on M.
2. In particular, Exercise 1 applies to the trivial bundle $\mathbb{R} \times M$, with trivial flat connection, for which $\nabla_X u = \langle X, du\rangle = Xu$. Thus, if $u \in C^\infty(M)$ is real-valued and $du(p) = 0$, then $D^2u(p)$ is well defined as a symmetric bilinear form on T_pM. If, in a coordinate system, $X = \sum X_j\, \partial/\partial x_j$, $Y = \sum Y_j\, \partial/\partial x_j$, show that

$$D^2_{X,Y}u(p) = \sum \frac{\partial^2 u}{\partial x_j \partial x_k}(p)\ X_jY_k. \tag{2.15}$$

   Show that this invariance fails if $du(p) \neq 0$.
3. If u is a smooth section of $\Lambda^kT^* \otimes E$, show that

(2.16)
$$\begin{aligned} d^\nabla u(X_0, \dots, X_k) = &\sum_j (-1)^j \nabla_{X_j} u(X_0, \dots, \widehat{X}_j, \dots, X_k) \\ &+ \sum_{j<\ell} (-1)^{j+\ell} u([X_j, X_\ell], X_0, \dots, \widehat{X}_j, \dots, \widehat{X}_\ell, \dots, X_k). \end{aligned}$$

   Compare with formula (13.56) of Chapter 1 and Exercises 2 and 3 in §3 of Chapter 2.
4. Verify the identity (2.12), namely, $d^\nabla d^\nabla u = \Omega \wedge u$.
5. If ∇ and $\widetilde{\nabla}$ are connections on $E \to M$, related by $\nabla_X u = \widetilde{\nabla}_X u + C(X, u)$, $C \in C^\infty(M, T^* \otimes \operatorname{End}(E))$, with curvatures R and $\widetilde{R}$, and curvature forms Ω and $\widetilde{\Omega}$, show that

$$\Omega - \widetilde{\Omega} = d^\nabla C + C \wedge C. \tag{2.17}$$

   Here the wedge product of two sections of $T^* \otimes \operatorname{End}(E)$ is a section of the bundle $\Lambda^2T^* \otimes \operatorname{End}(E)$, produced in a natural fashion, as in (1.15). Show that (2.17) is equivalent to (1.38).

3. The curvature tensor of a Riemannian manifold

The Levi-Civita connection, which was introduced in §11 of Chapter 1, is a metric connection on the tangent bundle TM of a manifold M with a Riemannian metric, uniquely specified among all such connections by the zero-torsion condition

$$\nabla_Y X - \nabla_X Y = [Y, X]. \tag{3.1}$$

We recall the defining formula

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle = {} & X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle \\ & + \langle [X, Y], Z\rangle - \langle [X, Z], Y\rangle - \langle [Y, Z], X\rangle, \end{aligned} \tag{3.2}$$

derived in (11.22) of Chapter 1. Thus, in a local coordinate system with the naturally associated frame field on the tangent bundle, the connection coefficients (1.8) are given by

$$\Gamma^\ell{}_{jk} = \frac{1}{2} g^{\ell\mu} \Big[\frac{\partial g_{j\mu}}{\partial x_k} + \frac{\partial g_{k\mu}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_\mu}\Big]. \tag{3.3}$$

The associated curvature tensor is the Riemann curvature tensor:

$$R(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z. \tag{3.4}$$

In a local coordinate system such as that discussed above, the expression for the Riemann curvature is a special case of (1.12), namely,

$$R^j{}_{k\ell m} = \partial_\ell \Gamma^j{}_{km} - \partial_m \Gamma^j{}_{k\ell} + \Gamma^j{}_{\nu\ell}\Gamma^\nu{}_{km} - \Gamma^j{}_{\nu m}\Gamma^\nu{}_{k\ell}. \tag{3.5}$$

Consequently, we have an expression of the form

$$R^j{}_{k\ell m} = L(g_{\alpha\beta}, \partial_\mu \partial_\nu g_{\gamma\delta}) + Q(g_{\alpha\beta}, \partial_\mu g_{\gamma\delta}), \tag{3.6}$$

where L is linear in the second-order derivatives of $g_{\alpha\beta}(x)$ and Q is quadratic in the first-order derivatives of $g_{\alpha\beta}(x)$, each with coefficients depending on $g_{\alpha\beta}(x)$.

Building on Proposition 1.2, we have the following result on metrics whose Riemannian curvature is zero.

Proposition 3.1. *If (M, g) is a Riemannian manifold whose curvature tensor vanishes, then the metric g is flat; that is, there is a coordinate system about each $p \in M$ in which $g_{jk}(x)$ is constant.*

Proof. It follows from Proposition 1.2 that on a neighborhood U of p there are parallel vector fields $V_{(j)}$, $j = 1, \dots, n = \dim M$, namely, in a given coordinate system

$$\nabla_{D_k} V_{(j)} = 0, \quad 1 \le j, \ k \le n, \tag{3.7}$$

such that $V_{(j)}(p)$ form a basis of T_pM. Let $v_{(j)}$ be the 1-forms associated to $V_{(j)}$ by the metric g, so

$$v_{(j)}(X) = g(X, V_{(j)}), \tag{3.8}$$

for all vector fields X. Hence

$$\nabla_{D_k} v_{(j)} = 0, \quad 1 \le j,\ k \le n. \tag{3.9}$$

We have $v_{(j)} = \sum v^k_{(j)}\, dx_k$, with $v^k_{(j)} = v_{(j)}(D_k) = \langle D_k, v_{(j)}\rangle$. The zero-torsion condition (3.1), in concert with (3.8), gives

$$\partial_\ell \langle v_{(j)}, D_k\rangle - \partial_k \langle v_{(j)}, D_\ell\rangle = \langle v_{(j)}, \nabla_{D_\ell} D_k\rangle - \langle v_{(j)}, \nabla_{D_k} D_\ell\rangle = 0, \tag{3.10}$$

which is equivalent to

$$d\, v_{(j)} = 0, \quad j = 1, \dots, n. \tag{3.11}$$

Hence, locally, there exist functions $x_j,\ j = 1, \dots, n$, such that

$$v_{(j)} = dx_j. \tag{3.12}$$

The functions $(x_1, \dots, x_n)$ give a coordinate system near p. In this coordinate system the inverse of the matrix $\big(g_{jk}(x)\big)$ has entries $g^{jk}(x) = \langle dx_j, dx_k\rangle$. Now, by (1.9),

$$\partial_\ell g^{jk}(x) = \langle \nabla_{D_\ell}\, dx_j, dx_k\rangle + \langle dx_j, \nabla_{D_\ell}\, dx_k\rangle = 0, \tag{3.13}$$

so the proof is complete.

We have seen in Proposition 1.1 that R has the following symmetries:

$$R(X,Y) = -R(Y,X), \tag{3.14}$$

$$\langle R(X,Y)Z, W\rangle = -\langle Z, R(X,Y)W\rangle. \tag{3.15}$$

In other words, in terms of

$$R_{jk\ell m} = \langle R(D_\ell, D_m)D_k, D_j\rangle, \tag{3.16}$$

we have

$$R_{jk\ell m} = -R_{jkm\ell} \tag{3.17}$$

and

$$R_{jk\ell m} = -R_{kj\ell m}. \tag{3.18}$$

The Riemann tensor has additional symmetries:

Proposition 3.2. *The Riemann tensor satisfies*

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0 \tag{3.19}$$

and

$$\langle R(X,Y)Z, W\rangle = \langle R(Z,W)X, Y\rangle, \tag{3.20}$$

or, in index notation,

$$R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk} = 0 \tag{3.21}$$

and

$$R_{ijk\ell} = R_{k\ell ij}. \tag{3.22}$$

Proof. Plugging in the definition of each of the three terms of (3.19), one gets a sum that is seen to cancel out by virtue of the zero-torsion condition (3.1). This gives (3.19) and hence (3.21). The identity (3.22) is an automatic consequence of (3.17), (3.18), and (3.21), by elementary algebraic manipulations, which we leave as an exercise, to complete the proof. Also, (3.22) follows from (3.50) below.

The identity (3.19) is sometimes called *Bianchi's first identity*, with (1.29) then called *Bianchi's second identity.*

There are important contractions of the Riemann tensor. The *Ricci tensor* is defined by

$$\mathrm{Ric}_{jk} = R^i{}_{jik} = g^{\ell m} R_{\ell jmk}, \tag{3.23}$$

where the summation convention is understood. By (3.22), this is symmetric in j, k. We can also raise indices:

$$\mathrm{Ric}^j{}_k = g^{j\ell}\mathrm{Ric}_{\ell k}; \quad \mathrm{Ric}^{jk} = g^{k\ell}\mathrm{Ric}^j{}_\ell. \tag{3.24}$$

Contracting again defines the scalar curvature:

$$S = \mathrm{Ric}^j{}_j. \tag{3.25}$$

As we will see below, the special nature of $R_{ijk\ell}$ for dim $M = 2$ implies

$$\mathrm{Ric}_{jk} = \frac{1}{2} S g_{jk} \ \text{ if } \dim M = 2. \tag{3.26}$$

The Bianchi identity (1.29) yields an important identity for the Ricci tensor. Specializing (1.30) to $\alpha = i$, $\beta = j$ and raising the second index give

$$R^{ij}{}_{ij;k} + R^{ij}{}_{jk;i} + R^{ij}{}_{ki;j} = 0, \tag{3.27}$$

hence, $S_{;k} - \mathrm{Ric}^i{}_{k;i} - \mathrm{Ric}^j{}_{k;j} = 0$, or

$$S_{;k} = 2\ \mathrm{Ric}^j{}_{k;j}. \tag{3.28}$$

This is called the *Ricci identity.* An equivalent form is

$$\mathrm{Ric}^{jk}{}_{;j} = \frac{1}{2}(S\ g^{jk})_{;j}. \tag{3.29}$$

The identity in this form leads us naturally to a tensor known as the *Einstein tensor*:

$$G^{jk} = \mathrm{Ric}^{jk} - \frac{1}{2} S\ g^{jk}. \tag{3.30}$$

The Ricci identity is equivalent to

$$G^{jk}{}_{;j} = 0. \tag{3.31}$$

As shown in Chapter 2, this means the Einstein tensor has zero divergence. This fact plays an important role in Einstein's equation for the gravitational field. Note that by (3.26) the Einstein tensor always vanishes when dim $M = 2$. On the other hand, the identity (3.31) has the following implication when dim $M > 2$.

Proposition 3.3. *If dim $M = n > 2$ and the Ricci tensor is a scalar multiple of the metric tensor, the factor necessarily being $1/n$ times the scalar curvature:*

$$Ric_{jk} = \frac{1}{n} S g_{jk}, \tag{3.32}$$

then S must be a constant.

Proof. Equation (3.32) is equivalent to

$$G^{jk} = \Big(\frac{1}{n} - \frac{1}{2}\Big) S g^{jk}. \tag{3.33}$$

By (3.31) and the fact that the covariant derivative of the metric tensor is 0, we have

$$0 = \Big(\frac{1}{n} - \frac{1}{2}\Big) S_{;k} g^{jk},$$

or $S_{;k} = 0$, which proves the proposition.

We now make some comments on the curvature of Riemannian manifolds M of dimension 2. By (3.17) and (3.18), in this case each component $R_{jk\ell m}$ of the curvature tensor is either 0 or ± the quantity

$$R_{1212} = R_{2121} = gK, \quad g = \det(g_{jk}). \tag{3.34}$$

One calls K the *Gauss curvature* of M when dim $M = 2$.

Suppose we pick normal coordinates centered at $p \in M$, so $g_{jk}(p) = \delta_{jk}$. We see that if dim $M = 2$,

$$\mathrm{Ric}_{jk}(p) = R_{1j1k} + R_{2j2k}.$$

Now, the first term on the right is zero unless $j = k = 2$, and the second term is zero unless $j = k = 1$. Hence, $\mathrm{Ric}_{jk}(p) = K(p)\delta_{jk}$, in normal coordinates, so in arbitrary coordinates

$$\mathrm{Ric}_{jk} = K g_{jk}; \quad \text{hence} \quad K = \frac{1}{2} S \quad \text{if dim } M = 2. \tag{3.35}$$

Explicit formulas for K when M is a surface in $\mathbb{R}^3$ are given by (4.22) and (4.29), in the next section. (See also Exercises 2 and 5–7 below.)

The following is a fundamental calculation of the Gauss curvature of a two-dimensional surface whose metric tensor is expressed in orthogonal coordinates:

$$ds^2 = E(x)\, dx_1^2 + G(x)\, dx_2^2. \tag{3.36}$$

Proposition 3.4. *Suppose* dim $M = 2$ *and the metric is given in coordinates by (3.36). Then the Gauss curvature* $k(x)$ *is given by*

$$k(x) = -\frac{1}{2\sqrt{EG}}\left[\partial_1\Big(\frac{\partial_1 G}{\sqrt{EG}}\Big) + \partial_2\Big(\frac{\partial_2 E}{\sqrt{EG}}\Big)\right]. \tag{3.37}$$

To establish (3.37), one can first compute that

$$\begin{aligned}\Gamma_1 = \big({\Gamma^j}_{k1}\big) &= \frac{1}{2}\begin{pmatrix} E^{-1}\partial_1 E & E^{-1}\partial_2 E \\ -G^{-1}\partial_2 E & G^{-1}\partial_1 G\end{pmatrix},\\ \Gamma_2 = \big({\Gamma^j}_{k2}\big) &= \frac{1}{2}\begin{pmatrix} E^{-1}\partial_2 E & -E^{-1}\partial_1 G \\ G^{-1}\partial_1 G & G^{-1}\partial_2 G\end{pmatrix}.\end{aligned}$$

Then, computing $\mathfrak{R}_{12} = ({R^j}_{k12}) = \partial_1\Gamma_2 - \partial_2\Gamma_1 + \Gamma_1\Gamma_2 - \Gamma_2\Gamma_1$, we have (3.38)

$$\begin{aligned}{R^1}_{212} = &-\frac{1}{2}\partial_1\Big(\frac{\partial_1 G}{E}\Big) - \frac{1}{2}\partial_2\Big(\frac{\partial_2 E}{E}\Big)\\ &+\frac{1}{4}\Big(-\frac{\partial_1 E}{E}\frac{\partial_1 G}{E} + \frac{\partial_2 E}{E}\frac{\partial_2 G}{G}\Big) - \frac{1}{4}\Big(\frac{\partial_2 E}{E}\frac{\partial_2 E}{E} - \frac{\partial_1 G}{E}\frac{\partial_1 G}{G}\Big).\end{aligned}$$

Now $R_{1212} = E\ {R^1}_{212}$ in this case, and (3.34) yields

$$k(x) = \frac{1}{EG}\ R_{1212} = \frac{1}{G}\ {R^1}_{212}. \tag{3.39}$$

If we divide (3.38) by G and then in the resulting formula for $k(x)$ interchange E and G, and ∂_1 and ∂_2, and sum the two formulas for $k(x)$, we get

$$\begin{aligned}k(x) = &-\frac{1}{4}\left[\frac{1}{G}\partial_1\Big(\frac{\partial_1 G}{E}\Big) + \frac{1}{E}\partial_1\Big(\frac{\partial_1 G}{G}\Big)\right]\\ &-\frac{1}{4}\left[\frac{1}{E}\partial_2\Big(\frac{\partial_2 E}{G}\Big) + \frac{1}{G}\partial_2\Big(\frac{\partial_2 E}{E}\Big)\right],\end{aligned}$$

which is easily transformed into (3.37).

If $E = G = e^{2v}$, we obtain a formula for the Gauss curvature of a surface whose metric is a conformal multiple of the flat metric:

Corollary 3.5. *Suppose* dim $M = 2$ *and the metric is given in coordinates by*

$$g_{jk}(x) = e^{2v}\delta_{jk}, \tag{3.40}$$

for a smooth v. Then the Gauss curvature $k(x)$ is given by

$$k(x) = -(\Delta_0 v)e^{-2v}, \tag{3.41}$$

where Δ_0 is the flat Laplacian in these coordinates:

$$\Delta_0 v = \frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2}. \tag{3.42}$$

For an alternative formulation of (3.41), note that the Laplace operator for the metric g_{jk} is given by

$$\Delta f = g^{-1/2}\,\partial_j(g^{jk}g^{1/2}\,\partial_k f),$$

and in the case (3.40), $g^{jk} = e^{-2v}\delta^{jk}$ and $g^{1/2} = e^{2v}$, so we have

$$\Delta f = e^{-2v}\Delta_0 f, \tag{3.43}$$

and hence (3.41) is equivalent to

$$k(x) = -\Delta v. \tag{3.44}$$

The comparison of the Gauss curvature of two surfaces that are conformally equivalent is a source of a number of interesting results. The following generalization of Corollary 3.5 is useful.

Proposition 3.6. *Let M be a two-dimensional manifold with metric g, whose Gauss curvature is $k(x)$. Suppose there is a conformally related metric*

$$g' = e^{2u}g. \tag{3.45}$$

Then the Gauss curvature $K(x)$ of g' is given by

$$K(x) = \big(-\Delta u + k(x)\big)e^{-2u}, \tag{3.46}$$

where Δ is the Laplace operator for the metric g.

Proof. We will use Corollary 3.5 as a tool in this proof. It is shown in Chapter 5, §11, that (M, g) is locally conformally flat, so we can assume without loss of generality that (3.40) holds; hence $k(x)$ is given by (3.41). Then

$$(g')_{jk} = e^{2w}\delta_{jk}, \quad w = u + v, \tag{3.47}$$

and (3.41) gives

$$K(x) = -(\Delta_0 w)e^{-2w} = \big[-(\Delta_0 u)e^{-2v} - (\Delta_0 v)e^{-2v}\big]e^{-2u}. \tag{3.48}$$

By (3.43) we have $(\Delta_0 u)e^{-2v} = \Delta u$, and applying (3.41) for $k(x)$ gives (3.46).

We end this section with a study of $\partial_j\partial_k g_{\ell m}(p_0)$ when one uses a geodesic normal coordinate system centered at p_0. We know from §11 of Chapter 1 that in such a coordinate system, ${\Gamma^\ell}_{jk}(p_0) = 0$ and hence $\partial_j g_{k\ell}(p_0) = 0$. Thus, in such a coordinate system, we have

$$R^j{}_{k\ell m}(p_0) = \partial_\ell \Gamma^j{}_{km}(p_0) - \partial_m \Gamma^j{}_{k\ell}(p_0), \tag{3.49}$$

and hence (3.3) yields

$$R_{jk\ell m}(p_0) = \frac{1}{2}\Bigl(\partial_j\partial_m g_{k\ell} + \partial_k\partial_\ell g_{jm} - \partial_j\partial_\ell g_{km} - \partial_k\partial_m g_{j\ell}\Bigr). \tag{3.50}$$

In light of the complexity of this formula, the following may be somewhat surprising. Namely, as Riemann showed, one has

$$\partial_j\partial_k g_{\ell m}(p_0) = -\frac{1}{3}R_{\ell jmk} - \frac{1}{3}R_{\ell kmj}. \tag{3.51}$$

This is related to the existence of nonobvious symmetries at the center of a geodesic normal coordinate system, such as $\partial_j\partial_k g_{\ell m}(p_0) = \partial_\ell\partial_m g_{jk}(p_0)$. To prove (3.51), by polarization it suffices to establish

$$\partial_j^2 g_{\ell\ell}(p_0) = -\frac{2}{3}R_{\ell j\ell j}, \quad \forall\ j,\ell. \tag{3.52}$$

Proving this is a two-dimensional problem, since (by (3.50)) both sides of the asserted identity in (3.52) are unchanged if M is replaced by the image under Exp_p of the two-dimensional linear span of D_j and D_ℓ. All one needs to show is that if $\dim M = 2$,

$$\partial_1^2 g_{22}(p_0) = -\frac{2}{3}K(p_0) \ \text{ and } \ \partial_1^2 g_{11}(p_0) = 0, \tag{3.53}$$

where $K(p_0)$ is the Gauss curvature of M at p_0. Of these, the second part is trivial, since $g_{11}(x) = 1$ on the horizontal line through p_0. To establish the first part of (3.53), it is convenient to use geodesic polar coordinates, (r, θ), in which

$$ds^2 = dr^2 + G(r,\theta)\, d\theta^2. \tag{3.54}$$

It is not hard to show that $G(r,\theta) = r^2 H(r,\theta)$, with $H(r,\theta) = 1 + O(r^2)$. For the metric (3.54), the formula (3.37) implies that the Gauss curvature is

$$K = -\frac{1}{2G}\partial_r^2 G + \frac{1}{4G^2}(\partial_r G)^2 = -\frac{H_r}{rH} - \frac{H_{rr}}{2H} + \frac{H_r^2}{4H^2}, \tag{3.55}$$

so at the center

$$K(p_0) = -H_{rr} - \frac{1}{2}H_{rr} = -\frac{3}{2}H_{rr}. \tag{3.56}$$

On the other hand, in normal coordinates (x_1, x_2), along the x_1-axis, we have $g_{22}(s,0) = G(s,0)/s^2 = H(s,0)$, so the rest of the identity (3.53) is established.

Exercises

Exercises 1–3 concern the problem of producing two-dimensional surfaces with constant Gauss curvature.

1. For a two-dimensional Riemannian manifold M, take geodesic polar coordinates, so the metric is

$$ds^2 = dr^2 + G(r,\theta)\,d\theta^2.$$

Use the formula (3.55) for the Gauss curvature, to deduce that

$$K = -\frac{\partial_r^2 \sqrt{G}}{\sqrt{G}}.$$

Hence, if $K = -1$, then

$$\partial_r^2 \sqrt{G} = \sqrt{G}.$$

Show that

$$\sqrt{G}(0,\theta) = 0, \quad \partial_r \sqrt{G}(0,\theta) = 1,$$

and deduce that $\sqrt{G}(r,\theta) = \varphi(r)$ is the unique solution to

$$\varphi''(r) - \varphi(r) = 0, \quad \varphi(0) = 0,\ \varphi'(0) = 1.$$

Deduce that

$$G(r,\theta) = \sinh^2 r.$$

Use this computation to deduce that any two surfaces with Gauss curvature -1 are locally isometric.

2. Suppose M is a surface of revolution in $\mathbb{R}^3$, of the form

$$x^2 + y^2 = g(z)^2.$$

If it is parameterized by $x = g(u)\ \cos v,\ y = g(u)\ \sin v,\ z = u$, then

$$ds^2 = \Bigl(1 + g'(u)^2\Bigr)\,du^2 + g(u)^2\,dv^2.$$

Deduce from (3.37) that

$$K = -\frac{g''(u)}{g(u)\Bigl(1 + g'(u)^2\Bigr)^2}.$$

Hence, if $K = -1$,

$$g''(u) = g(u)\Bigl(1 + g'(u)^2\Bigr)^2.$$

Note that a sphere of radius R is given by such a formula with $g(u) = \sqrt{R^2 - u^2}$. Compute K in this case.

2A. Suppose instead that M is a surface of revolution, described in the form

$$z = f\Bigl(\sqrt{x^2 + y^2}\Bigr).$$

If it is parameterized by $x = u\ \cos v,\ y = u\ \sin v,\ z = f(u)$, then

$$ds^2 = \Bigl(1 + f'(u)^2\Bigr)\,du^2 + u^2\,dv^2.$$

Show that

$$K = -\frac{1}{u\sqrt{1+f'(u)^2}} \frac{d}{du}\left(\frac{1}{\sqrt{1+f'(u)^2}}\right) = -\frac{\varphi'(u)}{2u}, \quad \varphi(u) = \frac{1}{1+f'(u)^2}.$$

Thus deduce that

$$K = -1 \Rightarrow \varphi(u) = u^2 + c \Rightarrow f(u) = \int \sqrt{\frac{1}{u^2+c} - 1}\, du.$$

We note that this is an elliptic integral, for most values of c. Show that, for $c = 0$, you get

$$f(u) = \sqrt{1-u^2} - \frac{1}{2}\log\Bigl(1 + \sqrt{1-u^2}\Bigr) + \frac{1}{2}\log\Bigl(1 - \sqrt{1-u^2}\Bigr).$$

3. Suppose M is a region in $\mathbb{R}^2$ whose metric tensor is a conformal multiple of the standard flat metric

$$g_{jk} = E(x)\delta_{jk} = e^{2v}\ \delta_{jk}.$$

Suppose $E = E(r),\ v = v(r)$. Deduce from (3.37) and (3.41) that

$$K = -\frac{1}{2E^2}\Bigl(E''(r) + \frac{1}{r}E'(r)\Bigr) + \frac{1}{2E^3}E'(r)^2 = -\Bigl(v''(r) + \frac{1}{r}v'(r)\Bigr)e^{-2v}.$$

Hence, if $K = -1$,

$$v''(r) + \frac{1}{r}v'(r) = e^{2v}.$$

Compute K when

$$g_{jk} = \frac{4}{(1-r^2)^2}\ \delta_{jk}.$$

4. Show that whenever $g_{jk}(x)$ satisfies $g_{jk}(p_0) = \delta_{jk},\ \partial_\ell g_{jk}(p_0) = 0$, at some point p_0, then (3.50) holds at p_0. If dim $M = 2$, deduce that

$$K(p_0) = -\frac{1}{2}\Bigl(\partial_1^2 g_{22} + \partial_2^2 g_{11} - 2\partial_1\partial_2 g_{12}\Bigr). \tag{3.57}$$

5. Suppose $M \subset \mathbb{R}^3$ is the graph of

$$x_3 = f(x_1, x_2),$$

so, using the natural (x_1, x_2)-coordinates on M,

$$ds^2 = (1 + f_1^2)\, dx_1^2 + 2f_1 f_2\ dx_1\ dx_2 + (1 + f_2^2)\, dx_2^2,$$

where $f_j = \partial_j f$. Show that if $\nabla f(0) = 0$, then Exercise 4 applies, so

$$\nabla f(0) = 0 \Longrightarrow K(0) = f_{11}f_{22} - f_{12}^2. \tag{3.58}$$

Compare the derivation of (4.22) in the next section.

6. If $M \subset \mathbb{R}^3$ is the surface of Exercise 5, then the *Gauss map* $N : M \to S^2$ is given by

$$N\Bigl(x, f(x)\Bigr) = \frac{(-f_1, -f_2, 1)}{\sqrt{1 + f_1^2 + f_2^2}}.$$

Show that if $\nabla f(0) = 0$, then, at $p_0 = \Big(0, f(0)\Big)$, $DN(p_0) : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$DN(p_0) = -\begin{pmatrix} \partial_1^2 f(0) & \partial_1\partial_2 f(0) \\ \partial_2\partial_1 f(0) & \partial_2^2 f(0) \end{pmatrix}. \tag{3.59}$$

Here, $T_{p_0}M$ and $T_{(0,0,1)}S^2$ are both identified with the (x_1, x_2)-plane. Deduce from Exercise 5 that

$$K(p_0) = \det DN(p_0).$$

7. Deduce from Exercise 6 that whenever M is a smooth surface in $\mathbb{R}^3$, with Gauss map $N : M \to S^2$, then, with $DN(x) : T_xM \to T_{N(x)}S^2$,

$$K(x) = \det DN(x), \quad \forall\, x \in M. \tag{3.60}$$

(*Hint*: Given $x \in M$, rotate coordinates so that T_xM is parallel to the (x_1, x_2)-plane.)
This result is Gauss' *Theorema Egregium* for surfaces in $\mathbb{R}^3$. See Theorem 4.4 for a more general formulation; see also (4.35), and Exercises 5, 8, 9, and 14 of §4.

8. Recall from §11 of Chapter 1 that if $\gamma_s(t)$ is a family of curves $\gamma_s : [a, b] \to M$ satisfying $\gamma_s(a) = p,\ \gamma_s(b) = q$, and if $E(s) = \int_a^b \langle T, T\rangle dt, \quad T = \gamma_s'(t)$, then, with $V = (\partial/\partial s)\gamma_s(t)|_{s=0}$, $E'(s) = -2\int_a^b \langle V, \nabla_T T\rangle\, dt$, leading to the stationary condition for E that $\nabla_T T = 0$, which is the geodesic equation. Now suppose $\gamma_{r,s}(t)$ is a two-parameter family of curves, $\gamma_{r,s}(a) = p,\ \gamma_{r,s}(b) = q$. Let $V = (\partial/\partial s)\gamma_{r,s}(t)|_{0,0},\ W = (\partial/\partial r)\gamma_{r,s}(t)|_{0,0}$. Show that

$$\frac{\partial^2}{\partial s \partial r} E(0,0) = 2\int_a^b \Big[\langle R(W,T)V, T\rangle + \langle \nabla_T V, \nabla_T W\rangle - \langle \nabla_W V, \nabla_T T\rangle\Big]\, dt. \tag{3.61}$$

Note that the last term in the integral vanishes if $\gamma_{0,0}$ is a geodesic.

9. If Z is a Killing field, generating an isometry on M (as in Chapter 2, §3), show that

$$Z_{j;k;\ell} = R^m{}_{\ell k j}\ Z_m.$$

(*Hint*: From Killing's equation $Z_{j;k} + Z_{k;j} = 0$, derive $Z_{j;k;\ell} = -Z_{k;\ell;j} - R^m{}_{k\ell j}\ Z_m$. Iterate this process two more times, going through the cyclic permutations of (j, k, ℓ). Use Bianchi's first identity.) Note that the identity desired is equivalent to

$$\nabla^2_{(X,Y)} Z = R(Y, Z)X \quad \text{if } Z \text{ is a Killing field.}$$

10. Derive the following equation of Jacobi for a variation of geodesics. If $\gamma_s(t)$ is a one-parameter family of geodesics, $X = \gamma_s'(t)$, and $W = (\partial/\partial s)\gamma_s$, then

$$\nabla_X \nabla_X\ W = R(X, W)X.$$

(*Hint*: Start with $0 = \nabla_W \nabla_X X$, and use $[X, W] = 0$.)

11. Raising the second index of $R^j{}_{k\ell m}$, you obtain $R^{jk}{}_{\ell m}$, the coordinate expression for $\mathcal{R}$, which can be regarded as a section of $\mathrm{End}(\Lambda^2 T)$. Suppose $M = X \times Y$ with a product Riemannian metric and associated curvatures $\mathcal{R}, \mathcal{R}_X, \mathcal{R}_Y$. Using the splitting

$$\Lambda^2(V \oplus W) = \Lambda^2 V \oplus \Big(\Lambda^1 V \otimes \Lambda^1 W\Big) \oplus \Lambda^2 W,$$

write $\mathcal{R}$ as a 3×3 block matrix. Show that

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathcal{R}_Y \end{pmatrix}.$$

In Exercises 12–14, let X, Y, Z, and so forth, belong to the space $\mathfrak{g}$ of left-invariant vector fields on a Lie group G, assumed to have a bi-invariant Riemannian metric. (Compact Lie groups have these.)

12. Show that any (constant-speed) geodesic γ on G with $\gamma(0) = e$, the identity element, is a subgroup of G, that is, $\gamma(s+t) = \gamma(s)\gamma(t)$. Deduce that $\nabla_X X = 0$ for $X \in \mathfrak{g}$.
(*Hint*: Given $p = \gamma(t_0)$, consider the "reflection" $R_p(g) = pg^{-1}p$, an isometry on G that fixes p and leaves γ invariant, though reversing its direction. From this, one can deduce that $p^2 = \gamma(2t_0)$.)
13. Show that $\nabla_X Y = (1/2)[X, Y]$ for $X, Y \in \mathfrak{g}$. (*Hint*: $0 = \nabla_X X = \nabla_Y Y = \nabla_{(X+Y)}(X+Y)$.)
This identity is called the *Maurer-Cartan* structure equation.
14. Show that

$$R(X,Y)Z = -\frac{1}{4}[[X,Y],Z], \quad \langle R(X,Y)Z, W\rangle = -\frac{1}{4}\Big\langle [X,Y],[Z,W]\Big\rangle.$$

15. If $E \to M$ is a vector bundle with connection $\widetilde{\nabla}$, and $\nabla = \widetilde{\nabla} + C$, as in Exercises 1 and 2 of §1, and M has Levi-Civita connection D, so that $\operatorname{Hom}(T \otimes E, E)$ acquires a connection from D and $\widetilde{\nabla}$, which we'll also denote as $\widetilde{\nabla}$, show that (1.38) is equivalent to

$$(R - \widetilde{R})(X,Y)u = (\widetilde{\nabla}_X C)(Y,u) - (\widetilde{\nabla}_Y C)(X,u) + [C_X, C_Y]u. \tag{3.62}$$

This is a general form of the "Palatini identity."

16. If g is a metric tensor and h a symmetric, second-order tensor field, consider the family of metric tensors $g_\tau = g + \tau h$, for τ close to zero, yielding the Levi-Civita connections

$$\nabla^\tau = \nabla + C(\tau),$$

where $\nabla = \nabla^0$. If $C' = C'(0)$, show that

$$\langle C'(X,Y), Z\rangle = \frac{1}{2}(\nabla_X h)(Y,Z) + \frac{1}{2}(\nabla_Y h)(X,Z) - \frac{1}{2}(\nabla_Z h)(X,Y). \tag{3.63}$$

(*Hint*: Use (3.2).)

17. Let $R(\tau)$ be the Riemann curvature tensor of g_τ, and set $R' = R'(0)$. Show that (3.62) yields

$$R'(X,Y)Z = (\nabla_X C')(Y,Z) - (\nabla_Y C')(X,Z). \tag{3.64}$$

Using (3.63), show that

$$\begin{aligned} 2\langle R'(X,Y)Z, W\rangle = {} & (\nabla^2_{Y,W} h)(X,Z) + (\nabla^2_{X,Z} h)(Y,W) - (\nabla^2_{X,W} h)(Y,Z) \\ & - (\nabla^2_{Y,Z} h)(X,W) + h\Big(R(X,Y)Z, W\Big) + h\Big(R(X,Y)W, Z\Big). \end{aligned} \tag{3.65}$$

(*Hint*: Use the derivation property of the covariant derivative to obtain a formula for $\nabla_X C'$ from (3.63).)

18. Show that

$$
\begin{aligned}
6\langle R(X,Y)Z,W\rangle = {} & \tilde{K}(X+W,Y+Z) - \tilde{K}(Y+W,X+Z) \\
& - \tilde{K}(X,Y+Z) - \tilde{K}(Y,X+W) - \tilde{K}(Z,X+W) \\
& - \tilde{K}(W,Y+Z) + \tilde{K}(X,Y+W) + \tilde{K}(Y,Z+W) \\
& + \tilde{K}(Z,Y+W) + \tilde{K}(W,X+Z) + \tilde{K}(X,Z) \\
& + \tilde{K}(Y,W) - \tilde{K}(X,Y) - \tilde{K}(Y,Z),
\end{aligned}
\tag{3.66}
$$

where

$$\tilde{K}(X,Y) = \langle R(X,Y)Y,X\rangle. \tag{3.67}$$

See (4.34) for an interpretation of the right side of (3.67).

19. Using (3.51), show that, in exponential coordinates centered at p, the function $g = \det(g_{jk})$ satisfies, for $|x|$ small,

$$g(x) = 1 - \frac{1}{3}\sum_{\ell,m} \mathrm{Ric}_{\ell m}(p) x_\ell x_m + O\Big(|x|^3\Big). \tag{3.68}$$

Deduce that if A_{n-1} = area of $S^{n-1} \subset \mathbb{R}^n$ and V_n = volume of unit ball in $\mathbb{R}^n$, then, for r small,

$$V\Big(B_r(p)\Big) = \Big(V_n - \frac{A_{n-1}}{6n(n+2)} S(p) r^2 + O(r^3)\Big) r^n. \tag{3.69}$$

4. Geometry of submanifolds and subbundles

Let M be a Riemannian manifold, of dimension n, and let S be a submanifold, of dimension k, with the induced metric tensor. M has a Levi-Civita connection ∇ and Riemann tensor R. Denote by ∇^0 and R_S the connection and curvature of S, respectively. We aim to relate these objects. The *second fundamental form* is defined by

$$II(X,Y) = \nabla_X Y - \nabla^0_X Y, \tag{4.1}$$

for X and Y tangent to S. Note that II is linear in X and in Y over $C^\infty(S)$. Also, by the zero-torsion condition,

$$II(X,Y) = II(Y,X). \tag{4.2}$$

Proposition 4.1. *$II(X,Y)$ is normal to S at each point.*

Proof. If X, Y and Z are tangent to S, we have

$$\langle \nabla_X Y, Z\rangle - \langle \nabla^0_X Y, Z\rangle = -\langle Y, \nabla_X Z\rangle + X\langle Y,Z\rangle + \langle Y, \nabla^0_X Z\rangle - X\langle Y,Z\rangle,$$

and making the obvious cancellation, we obtain

$$\langle II(X,Y), Z\rangle = -\langle Y, II(X,Z)\rangle. \tag{4.3}$$

Using (4.2), we have

$$\langle II(X,Y),Z\rangle = -\langle Y, II(Z,X)\rangle; \tag{4.4}$$

that is, the trilinear form given by the left side changes sign under a cyclic permutation of its arguments. Since three such permutations produce the original form, the left side of (16.4) must equal its own negative, hence be 0. This proves the proposition.

Denote by $\nu(S)$ the bundle of normal vectors to S, called the normal bundle of S. It follows that II is a section of $\mathrm{Hom}(TS \otimes TS, \nu(S))$.

Corollary 4.2. *For X and Y tangent to S, $\nabla^0_X Y$ is the tangential projection on TS of $\nabla_X Y$.*

Let ξ be normal to S. We have a linear map, called the *Weingarten map*,

$$A_\xi : T_pS \longrightarrow T_pS \tag{4.5}$$

uniquely defined by

$$\langle A_\xi X, Y\rangle = \langle \xi, II(X,Y)\rangle. \tag{4.6}$$

We also define the section A of $\mathrm{Hom}(\nu(S)\otimes TS, TS)$ by

$$A(\xi, X) = A_\xi X. \tag{4.7}$$

We define a connection on $\nu(S)$ as follows; if ξ is a section of $\nu(S)$, set

$$\nabla^1_X \xi = P^\perp \nabla_X \xi,$$

where $P^\perp(x)$ is the orthogonal projection of T_xM onto $\nu_x(S)$. The following identity is called the *Weingarten formula.*

Proposition 4.3. *If ξ is a section of $\nu(S)$,*

$$\nabla^1_X \xi = \nabla_X \xi + A_\xi X. \tag{4.8}$$

Proof. It suffices to show that $\nabla_X\xi + A_\xi X$ is normal to S. In fact, if Y is tangent to S,

$$\begin{aligned}\langle \nabla_X\xi, Y\rangle + \langle A_\xi X, Y\rangle &= X\langle \xi, Y\rangle - \langle \xi, \nabla_X Y\rangle + \langle \xi, II(X,Y)\rangle \\ &= 0 - \langle \xi, \nabla^0_X Y\rangle - \langle \xi, II(X,Y)\rangle + \langle \xi, II(X,Y)\rangle \\ &= 0,\end{aligned}$$

which proves the proposition.

An equivalent statement is that, for X tangent to S, ξ normal to S,

$$\nabla_X \xi = \nabla^1_X \xi - A_\xi X \tag{4.9}$$

is an orthogonal decomposition, into components normal and tangent to S, respectively. Sometimes this is taken as the definition of A_ξ or, equivalently, by (4.6), of the second fundamental form.

In the special case that S is a hypersurface of M (i.e., dim $M =$ dim $S+1$), if $\xi = N$ is a smooth unit normal field to S, we see that, for X tangent to S,

$$\langle \nabla_X N, N\rangle = \frac{1}{2} X\langle N, N\rangle = 0,$$

so $\nabla^1_X N = 0$ in this case, and (4.9) takes the form

$$\nabla_X N = -A_N X, \tag{4.10}$$

the classical form of the Weingarten formula.

We now compare the tensors R and R_S. Let X, Y and Z be tangent to S. Then

$$\begin{aligned} \nabla_X\nabla_Y Z &= \nabla_X\big(\nabla^0_Y Z + II(Y,Z)\big) \\ &= \nabla^0_X\nabla^0_Y Z + II(X, \nabla^0_Y Z) - A_{II(Y,Z)}X + \nabla^1_X II(Y,Z). \end{aligned} \tag{4.11}$$

Reversing X and Y, we have

$$\nabla_Y\nabla_X Z = \nabla^0_Y\nabla^0_X Z + II(Y, \nabla^0_X Z) - A_{II(X,Z)}Y + \nabla^1_Y II(X,Z).$$

Also,

$$\nabla_{[X,Y]}Z = \nabla^0_{[X,Y]}Z + II([X,Y],Z). \tag{4.12}$$

From (4.11) and (4.12) we obtain the important identity

$$\begin{aligned} (R - R_S)(X,Y)Z =& \big\{II(X,\nabla^0_Y Z) - II(Y, \nabla^0_X Z) - II([X,Y],Z) \\ &+ \nabla^1_X II(Y,Z) - \nabla^1_Y II(X,Z)\big\} \\ &- \big\{A_{II(Y,Z)}X - A_{II(X,Z)}Y\big\}. \end{aligned} \tag{4.13}$$

Here, the quantity in the first set of braces $\{\quad\}$ is normal to S, and the quantity in the second pair of braces is tangent to S. The identity (4.13) is called the *Gauss-Codazzi equation.* A restatement of the identity for the tangential components is the following, known as Gauss' *Theorema Egregium.*

Theorem 4.4. *For X, Y, Z and W tangent to S,*

$$\begin{aligned} \langle (R - R_S)(X,Y)Z, W\rangle = &\langle II(Y,W), II(X,Z)\rangle \\ &- \langle II(X,W), II(Y,Z)\rangle. \end{aligned} \tag{4.14}$$

The normal component of the identity (4.13) is specifically Codazzi's equation. It takes a shorter form in case S has codimension 1 in M. In that case, choose a unit normal field N, and let

$$II(X,Y) = \widetilde{II}(X,Y)N; \tag{4.15}$$

$\widetilde{II}$ is a tensor field of type $(0,2)$ on S. Then Codazzi's equation is equivalent to

$$\langle R(X,Y)Z, N\rangle = (\nabla^0_X \widetilde{II})(Y,Z) - (\nabla^0_Y \widetilde{II})(X,Z), \tag{4.16}$$

for X, Y, Z tangent to S, since of course $R_S(X,Y)Z$ is tangent to S.

In the classical case, where S is a hypersurface in flat Euclidean space, $R = 0$, and Codazzi's equation becomes

$$(\nabla^0_Y \widetilde{II})(X,Z) - (\nabla^0_X \widetilde{II})(Y,Z) = 0, \tag{4.17}$$

that is, $\nabla^0 \widetilde{II}$ is a symmetric tensor field of type $(0,3)$. In this case, from the identity $\widetilde{II}_{jk;\ell} = \widetilde{II}_{\ell k;j}$, we deduce $A_j{}^k{}_{;k} = A_k{}^k{}_{;j} = (\mathrm{Tr}\ A)_{;j}$, where $A = A_N$ is the Weingarten map. Equivalently,

$$\mathrm{div}\ A = d(\mathrm{Tr}\ A). \tag{4.18}$$

An application of the Codazzi equation to minimal surfaces can be found in the exercises after §6 of Chapter 14.

It is useful to note the following characterization of the second fundamental form for a hypersurface M in $\mathbb{R}^n$. Translating and rotating coordinates, we can move a specific point $p \in M$ to the origin in $\mathbb{R}^n$ and suppose M is given locally by

$$x_n = f(x'), \quad \nabla f(0) = 0,$$

where $x' = (x_1, \dots, x_{n-1})$. We can then identify the tangent space of M at p with $\mathbb{R}^{n-1}$.

Proposition 4.5. *The second fundamental form of M at p is given by the Hessian of f:*

$$\widetilde{II}(X,Y) = \sum_{j,k=1}^{n-1} \frac{\partial^2 f}{\partial x_j \partial x_k}(0)\ X_j Y_k. \tag{4.19}$$

Proof. From (4.9) we have, for any ξ normal to M,

$$\langle II(X,Y), \xi\rangle = -\langle \nabla_X \xi, Y\rangle, \tag{4.20}$$

where ∇ is the flat connection on $\mathbb{R}^n$. Taking

$$\xi = (-\partial_1 f, \dots, -\partial_{n-1} f, 1) \tag{4.21}$$

gives the desired formula.

If S is a surface in $\mathbb{R}^3$, given locally by $x_3 = f(x_1, x_2)$ with $\nabla f(0) = 0$, then the Gauss curvature of S at the origin is seen by (4.14) and (4.19) to equal

$$\det\Big(\frac{\partial^2 f(0)}{\partial x_j \partial x_k}\Big). \tag{4.22}$$

Consider the example of the unit sphere in $\mathbb{R}^3$, centered at $(0,0,1)$. Then the "south pole" lies at the origin, near which S^2 is given by

$$x_3 = 1 - (1 - x_1^2 - x_2^2)^{1/2}. \tag{4.23}$$

In this case (4.22) implies that the Gauss curvature K is equal to 1 at the south pole. Of course, by symmetry it follows that $K = 1$ everywhere on the unit sphere S^2.

Besides providing a good conception of the second fundamental form of a hypersurface in $\mathbb{R}^n$, Proposition 4.5 leads to useful formulas for computation, one of which we will give in (4.29). First, we give a more invariant reformulation of Proposition 4.5. Suppose the hypersurface M in $\mathbb{R}^n$ is given by

$$u(x) = c, \tag{4.24}$$

with $\nabla u \neq 0$ on M. Then we can use the computation (4.20) with $\xi = \operatorname{grad} u$ to obtain

$$\langle II(X,Y), \operatorname{grad} u\rangle = -(D^2u)(X,Y), \tag{4.25}$$

where D^2u is the Hessian of u; we can think of $(D^2u)(X,Y)$ as $Y\cdot(D^2u)X$, where D^2u is the $n\times n$ matrix of second-order partial derivatives of u. In other words,

$$\widetilde{II}(X,Y) = -|\operatorname{grad} u|^{-1}(D^2u)(X,Y), \tag{4.26}$$

for X, Y tangent to M.

In particular, if M is a two-dimensional surface in $\mathbb{R}^3$ given by (4.24), then the Gauss curvature at $p\in M$ is given by

$$K(p) = |\operatorname{grad} u|^{-2}\det(D^2u)\big|_{T_pM}, \tag{4.27}$$

where $D^2u|_{T_pM}$ denotes the restriction of the quadratic form D^2u to the tangent space T_pM, producing a linear transformation on T_pM via the metric on T_pM. With this calculation we can derive the following formula, extending (4.22).

Proposition 4.6. *If $M\subset\mathbb{R}^3$ is given by*

$$x_3 = f(x_1, x_2), \tag{4.28}$$

then, at $p = (x', f(x'))\in M$, the Gauss curvature is given by

$$K(p) = \big(1 + |\nabla f(x')|^2\big)^{-2}\ \det\Big(\frac{\partial^2 f}{\partial x_j\partial x_k}\Big). \tag{4.29}$$

Proof. We can apply (4.27) with $u(x) = f(x_1, x_2) - x_3$. Note that $|\nabla u|^2 = 1 + |\nabla f(x')|^2$ and

$$D^2 u = \begin{pmatrix} D^2 f & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.30}$$

Noting that a basis of T_pM is given by $(1, 0, \partial_1 f) = v_1,\ (0, 1, \partial_2 f) = v_2$, we readily obtain

$$\det D^2 u\big|_{T_pM} = \frac{\det\big(v_j \cdot (D^2 u) v_k\big)}{\det(v_j \cdot v_k)} = \big(1 + |\nabla f(x')|^2\big)^{-1} \det D^2 f, \tag{4.31}$$

which yields (4.29).

If you apply Proposition 4.6 to the case (4.23) of a hemisphere of unit radius, the calculation that $K = 1$ everywhere is easily verified. The formula (4.29) gives rise to interesting problems in nonlinear PDE, some of which are studied in Chapter 14.

We now define the sectional curvature of a Riemannian manifold M. Given $p \in M$, let Π be a 2-plane in T_pM, $\Sigma = \operatorname{Exp}_p(\Pi)$. The sectional curvature of M at p is

$$K_p(\Pi) = \text{ Gauss curvature of } \Sigma \text{ at } p. \tag{4.32}$$

If U and V form an orthonormal basis of $T_p\Sigma = \Pi$, then by the definition of Gauss curvature,

$$K_p(\Pi) = \langle R_\Sigma(U, V)V, U\rangle. \tag{4.33}$$

We have the following more direct formula for the sectional curvature.

Proposition 4.7. *With U and V as above, R the Riemann tensor of M,*

$$K_p(\Pi) = \langle R(U, V)V, U\rangle. \tag{4.34}$$

Proof. It suffices to show that the second fundamental form of Σ vanishes at p. Since $II(X, Y)$ is symmetric, it suffices to show that $II(X, X) = 0$ for each $X \in T_pM$. So pick a geodesic γ in M such that $\gamma(0) = p,\ \gamma'(0) = X$. Then $\gamma \subset \Sigma$, and γ must also be a geodesic in S, so

$$\nabla_T T = \nabla^0_T T, \quad T = \gamma'(t),$$

which implies $II(X, X) = 0$. This proves (4.34).

Note that if $S \subset M$ has codimension 1, $p \in S$, and $\Pi \subset T_pS$, then, by (4.14),

$$K_p^S(\Pi) - K_p^M(\Pi) = \det \begin{pmatrix} \widetilde{II}(U, U) & \widetilde{II}(U, V) \\ \widetilde{II}(V, U) & \widetilde{II}(V, V) \end{pmatrix}. \tag{4.35}$$

Note how this is a direct generalization of (3.60).

The results above comparing connections and curvatures of a Riemannian manifold and a submanifold are special cases of more general results on subbundles, which arise in a number of interesting situations. Let E be a vector bundle over a manifold M, with an inner product and a metric connection ∇. Let $E_0 \to M$ be a subbundle. For each $x \in M$, let P_x be the orthogonal projection of E_x onto E_{0x}. Set

$$\nabla^0_X u(x) = P_x \nabla_X u(x), \tag{4.36}$$

when u is a section of E_0. Note that, for scalar f,

$$\begin{aligned} \nabla^0_X f u(x) &= P_x\big(f\nabla_X u(x) + (Xf)u\big) \\ &= fP_x\nabla_X u(x) + (Xf)u(x), \end{aligned}$$

provided u is a section of E_0, so $P_x u(x) = u(x)$. This shows that (4.36) defines a connection on E_0. Since $\langle \nabla^0_X u, v\rangle = \langle \nabla_X u, v\rangle$ for sections u, v of E_0, it is clear that ∇^0 is also a metric connection. Similarly, if E_1 is the orthogonal bundle, a subbundle of E, a metric connection on E_1 is given by

$$\nabla^1_X v(x) = (I - P_x)\nabla_X v(x), \tag{4.37}$$

for a section v of E_1.

It is useful to treat ∇^0 and ∇^1 on an equal footing, so we define a new connection $\widetilde{\nabla}$ on E, also a metric connection, by

$$\widetilde{\nabla} = \nabla^0 \oplus \nabla^1. \tag{4.38}$$

Then there is the relation

$$\nabla_X = \widetilde{\nabla}_X + C_X, \tag{4.39}$$

where

$$C_X = \begin{pmatrix} 0 & II^1_X \\ II^0_X & 0 \end{pmatrix}. \tag{4.40}$$

Here, $II^0_X : E_0 \to E_1$ is the second fundamental form of $E_0 \subset E$, and $II^1_X : E_1 \to E_0$ is the second fundamental form of $E_1 \subset E$. We also set $II^j(X,u) = II^j_X u$. In this context, the Weingarten formula has the form

$$C^t_X = -C_X, \quad \text{i.e., } II^1_X = -\big(II^0_X\big)^t. \tag{4.41}$$

Indeed, for any two connections related by (4.39), with $C \in \mathrm{Hom}(TM \otimes E, E)$, if ∇ and $\widetilde{\nabla}$ are both metric connections, the first part of (4.41) holds.

We remark that when γ is a curve in a Riemannian manifold M, and for $p \in \gamma$, $E_p = T_pM$, $E_{0p} = T_p\gamma$, $E_{1p} = \nu(\gamma)$, the normal space, and if ∇ is the Levi-Civita connection on M, then $\widetilde{\nabla}$ is sometimes called the *Fermi-Walker connection* on γ. One also (especially) considers a timelike curve in a Lorentz manifold.

Let us also remark that if we start with metric connections ∇^j on E_j, then form $\widetilde{\nabla}$ on E by (4.38), and then define ∇ on E by (4.39), provided that (4.40) holds, it follows that ∇ is also a metric connection on E, and the connections ∇^j are recovered by (4.36) and (4.37).

In general, for any two connections ∇ and $\widetilde{\nabla}$, related by (4.39) for some $\operatorname{End}(E)$ valued 1-form C, we have the following relation between their curvature tensors R and $\widetilde{R}$, already anticipated in Exercise 2 of §1:

$$(R-\widetilde{R})(X,Y)u = \big\{[C_X,\widetilde{\nabla}_Y]-[C_Y,\widetilde{\nabla}_X]-C_{[X,Y]}\big\}u+[C_X,C_Y]u. \tag{4.42}$$

In case $\widetilde{\nabla}=\nabla^0\oplus\nabla^1$ on $E=E_0\oplus E_1$, and ∇ has the form (4.39), where C_X exchanges E_0 and E_1, it follows that the operator in brackets $\{\ \ \}$ on the right side of (4.42) exchanges sections of E_0 and E_1, while the last operator $[C_X,C_Y]$ leaves invariant the sections of E_0 and E_1. In such a case these two components express respectively the Codazzi identity and Gauss' *Theorema Egregium*.

We will expand these formulas, writing $R(X,Y)\in\operatorname{End}(E_0\oplus E_1)$ in the block matrix form

$$R=\begin{pmatrix} R_{00} & R_{01}\\ R_{10} & R_{11}\end{pmatrix}. \tag{4.43}$$

Then Gauss' equations become

$$\begin{aligned}(R_{00}-R_0)(X,Y)u &= II^1_X II^0_Y u - II^1_Y II^0_X u,\\ (R_{11}-R_1)(X,Y)u &= II^0_X II^1_Y u - II^0_Y II^1_X u,\end{aligned} \tag{4.44}$$

for a section u of E_0 or E_1, respectively. Equivalently, if v is also a section of E_0 or E_1, respectively,

$$\begin{aligned}\langle (R_{00}-R_0)(X,Y)u,v\rangle &= \langle II^0_X u, II^0_Y v\rangle - \langle II^0_Y u, II^0_X v\rangle,\\ \langle (R_{11}-R_1)(X,Y)u,v\rangle &= \langle II^1_X u, II^1_Y v\rangle - \langle II^1_Y u, II^1_X v\rangle.\end{aligned} \tag{4.45}$$

The second part of (4.45) is also called the Ricci equation.

Codazzi's equations become

$$\begin{aligned}R_{10}(X,Y)u &= II^0_X\nabla^0_Y u - II^0_Y\nabla^0_X u - II^0_{[X,Y]}u + \nabla^1_X II^0_Y u - \nabla^1_Y II^0_X u,\\ R_{01}(X,Y)u &= II^1_X\nabla^1_Y u - II^1_Y\nabla^1_X u - II^1_{[X,Y]}u + \nabla^0_X II^1_Y u - \nabla^0_Y II^1_X u,\end{aligned} \tag{4.46}$$

for sections u of E_0 and E_1, respectively. If we take the inner product of the first equation in (4.46) with a section v of E_1, we get

$$\begin{aligned}&\langle R_{10}(X,Y)u,v\rangle\\ &\quad= -\langle \nabla^0_Y u, II^1_X v\rangle + \langle \nabla^0_X u, II^1_Y v\rangle - \langle II^0_{[X,Y]}u,v\rangle\\ &\qquad + \langle II^0_X u, \nabla^1_Y v\rangle - \langle II^0_Y u, \nabla^1_X v\rangle + X\langle II^0_Y u, v\rangle - Y\langle II^0_X u, v\rangle,\end{aligned} \tag{4.47}$$

using the metric property of ∇^0 and ∇^1, and the antisymmetry of (4.40). If we perform a similar calculation for the second part of (4.46), in light of the fact that $R_{10}(X,Y)^t = -R_{01}(X,Y)$, we see that these two parts are equivalent, so we need retain only one of them. Furthermore, we can rewrite the first equation in (4.46) as follows. Form a connection on $\mathrm{Hom}(TM \otimes E_0, E_1)$ via the connections ∇^j on E_j and a Levi-Civita connection ∇^M on TM, via the natural derivation property, that is,

$$(\widetilde{\nabla}_X II^0)(Y,u) = \nabla^1_X II^0_Y u - II^0_Y \nabla^0_X u - II^0(\nabla^M_X Y, u). \tag{4.48}$$

Then (4.46) is equivalent to

$$R_{10}(X,Y)u = (\widetilde{\nabla}_X II^0)(Y,u) - (\widetilde{\nabla}_Y II^0)(X,u). \tag{4.49}$$

One case of interest is when E_1 is the trivial bundle $E_1 = M \times \mathbb{R}$, with one-dimensional fiber. For example, E_1 could be the normal bundle of a codimension-one surface in $\mathbb{R}^n$. In this case, it is clear that both sides of the last half of (4.45) are tautologically zero, so Ricci's equation has no content in this case.

As a parenthetical comment, suppose E is a trivial bundle $E = M \times \mathbb{R}^n$, with complementary subbundles E_j, having metric connections constructed as in (4.36) and (4.37), from the trivial connection D on E, defined by componentwise differentiation, so

$$\nabla^0_X u = PD_X u, \quad \nabla^1_X u = (I-P)D_X u, \tag{4.50}$$

for sections of E_0 and E_1, respectively. There is the following alternative approach to curvature formulas. For $\widetilde{\nabla} = \nabla^0 \oplus \nabla^1$, we have

$$\widetilde{\nabla}_X u = D_X u + (D_X P)(I-2P)u. \tag{4.51}$$

Note that with respect to a choice of basis of $\mathbb{R}^n$ as a global frame field on $M \times \mathbb{R}^n$, we have the connection 1-form (1.13) given by

$$\Gamma = dP(I-2P). \tag{4.52}$$

Since $dP = dP\ P + P\ dP$, we have $dP\ P = (I-P)\,dP$. Thus, writing the connection 1-form as $\Gamma = P\,dP(I-P) - (I-P)\,dP\,P$ casts $\Gamma = -C$ in the form (4.40). We obtain directly from the formula $\Omega = d\Gamma + \Gamma \wedge \Gamma$, derived in (4.15), that the curvature of $\widetilde{\nabla}$ is given by

$$\Omega = dP \wedge dP = P\ dP \wedge dP\ P + (I-P)\ dP \wedge dP\,(I-P), \tag{4.53}$$

the last identity showing the respective curvatures of E_0 and E_1. Compare with Exercise 5 of §1.

Our next goal is to invert the process above. That is, rather than starting with a flat bundle $E = M \times \mathbb{R}^n$ and obtaining connections on subbundles and second fundamental forms, we want to start with bundles $E_j \to M$, $j = 1,2$, with metric connections ∇^j, and proposed second fundamental forms II^j, sections of $\mathrm{Hom}(TM \otimes E_j, E_{j'})$, and then obtain a flat connection ∇ on E via (4.38)–(4.40). Of course, we assume II^0 and II^1 are related by (4.41), so (4.39) makes ∇ a metric connection. Thus, according to equations (4.45) and (4.49), the connection ∇ is flat if and only if, for all sections u, v of E_0,

$$\begin{aligned} &(\widetilde{\nabla}_X II^0)(Y,u) - (\widetilde{\nabla}_Y II^0)(X,u) = 0, \\ &\langle R_0(X,Y)u, v\rangle = \langle II^0_Y u, II^0_X v\rangle - \langle II^0_X u, II^0_Y v\rangle, \end{aligned} \tag{4.54}$$

and, for all sections u, v of E_1,

$$\langle R_1(X,Y)u, v\rangle = \langle II^1_Y u, II^1_X v\rangle - \langle II^1_X u, II^1_Y v\rangle. \tag{4.55}$$

If these conditions are satisfied, then E will have a global frame field of sections $e_1, \dots, e_n$, such that $\nabla e_j = 0$, at least provided M is simply connected. Then, for each $p \in M$, we have an isometric isomorphism

$$J(p) : E_p \longrightarrow \mathbb{R}^n \tag{4.56}$$

by expanding elements of E_p in terms of the basis $\{e_j(p)\}$. Thus $E_0 \subset E$ is carried by $J(p)$ to a family of linear subspaces $J(p)E_{0p} = V_p \subset \mathbb{R}^n$, with orthogonal complements $J(p)E_{1p} = N_p \subset \mathbb{R}^n$.

We now specialize to the case $E_0 = TM$, where M is an m-dimensional Riemannian manifold, with its Levi-Civita connection; E_1 is an auxiliary bundle over M, with metric connection ∇^1. We will assume M is simply connected. The following result is sometimes called the *fundamental theorem of surface theory.*

Theorem 4.8. *Let* II^0 *be a section of* $Hom(TM \otimes TM, E_1)$*, and set* $II^1_X = -(II^0_X)^t$. *Make the symmetry hypothesis*

$$II^0(X,Y) = II^0(Y,X). \tag{4.57}$$

Assume the equations (4.54) and (4.55) are satisfied, producing a trivialization of $E = E_0 \oplus E_1$*, described by (4.56). Then there is an isometric immersion*

$$X : M \longrightarrow \mathbb{R}^n, \tag{4.58}$$

and a natural identification of E_1 *with the normal bundle of* $S = X(M) \subset \mathbb{R}^n$*, such that the second fundamental form of* S *is given by* II^0.

To get this, we will construct the map (4.58) so that

$$DX(p) = J(p)\big|_{TM}, \tag{4.59}$$

for all $p \in M$. To see how to get this, consider one of the n components of J, $J_\nu(p) : E_p \to \mathbb{R}$. In fact, $J_\nu u = \langle e_\nu, u\rangle$. Let $\beta_\nu(p) = J_\nu(p)\big|_{T_pM}$; thus β_ν is a 1-form on M.

Lemma 4.9. *Each* β_ν *is closed, that is,* $d\beta_\nu = 0$.

Proof. For vector fields X and Y on M, we have

$$\begin{aligned} d\beta_\nu(X,Y) &= X \cdot \beta_\nu(Y) - Y \cdot \beta_\nu(X) - \beta_\nu([X,Y]) \\ &= X \cdot \beta_\nu(Y) - Y \cdot \beta_\nu(X) - \beta_\nu(\nabla^0_X Y - \nabla^0_Y X). \end{aligned} \tag{4.60}$$

Using $\nabla_X = \nabla^0_X + II^0_X$ on sections of $E_0 = TM$, we see that this is equal to

$$\begin{aligned} &X \cdot J_\nu(Y) - Y \cdot J_\nu(X) - J_\nu(\nabla_X Y - \nabla_Y X) + J_\nu(II^0_X Y - II^0_Y X) \\ &= (\nabla_X J_\nu)Y - (\nabla_Y J_\nu)X + J_\nu(II^0_X Y - II^0_Y X). \end{aligned}$$

By construction, $\nabla_X J_\nu = 0$, while (4.57) says $II^0_X Y - II^0_Y X = 0$. Thus $d\beta_\nu = 0$.

Consequently, as long as M is simply connected, we can write $\beta_\nu = dx_\nu$ for some functions $x_\nu \in C^\infty(M)$. Let us therefore define the map (4.58) by $X(p) = (x_1(p), \dots, x_\nu(p))$. Thus (4.59) holds, so X is an isometric mapping. Furthermore, it is clear that $J(p)$ maps E_{1p} precisely isometrically onto the normal space $N_p \subset \mathbb{R}^n$ to $S = X(M)$ at $X(p)$, displaying II^0 as the second fundamental form of S. Thus Theorem 4.8 is established.

Let us specialize Theorem 4.8 to the case where dim $M = n - 1$, so the fibers of E_1 are one-dimensional. As mentioned above, the Ricci identity (4.55) has no content in that case. We have the following special case of the fundamental theorem of surface theory.

Proposition 4.10. *Let M be an $(n-1)$-dimensional Riemann manifold; assume M is simply connected. Let there be given a symmetric tensor field $\widetilde{II}$, of type (0,2). Assume the following Gauss-Codazzi equations hold:*

$$\begin{gathered} \langle R^M(X,Y)Z,W\rangle = \widetilde{II}(Y,Z)\widetilde{II}(X,W) - \widetilde{II}(X,Z)\widetilde{II}(Y,W), \\ \big(\nabla^M_X \widetilde{II}\big)(Y,Z) - \big(\nabla^M_Y \widetilde{II}\big)(X,Z) = 0, \end{gathered} \tag{4.61}$$

where ∇^M is the Levi-Civita connection of M and R^M is its Riemann curvature tensor. Then there is an isometric immersion $X : M \to \mathbb{R}^n$ such that the second fundamental form of $S = X(M) \subset \mathbb{R}^n$ is given by $\widetilde{II}$.

Exercises

1. Let $S \subset M$, with respective Levi-Civita connections ∇^0, ∇, respective Riemann tensors R_s, R, and so on, as in the text. Let $\gamma_{s,t} : [a,b] \to S$ be a two-parameter family of curves. One can also regard $\gamma_{s,t} : [a,b] \to M$. Apply the formula (3.52) for the second variation of energy in these two contexts, and compare the results, to produce another proof of Gauss' formula (4.14) for $\langle (R - R_S)(X,Y)Z,W\rangle$ when X,Y,Z,W are all tangent to S.
2. With the Ricci tensor Ric given by (3.23) and the sectional curvature $K_p(\Pi)$ by (4.32), show that, for $X \in T_pM$, of norm 1, if Ξ denotes the orthogonal

complement of X in T_pM, then

$$\operatorname{Ric}(X,X) = \frac{n-1}{\operatorname{vol} S_p(\Xi)} \int\limits_{S_p(\Xi)} K_p(U,X)\; dV(U),$$

where $S_p(\Xi)$ is the unit sphere in Ξ, $n = \dim M$, and $K_p(U,X) = K_p(\Pi)$, where Π is the linear span of U and X. Show that the scalar curvature at p is given by

$$S = \frac{n(n-1)}{\operatorname{vol} G_2} \int\limits_{G_2} K_p(\Pi)\; dV(\Pi),$$

where G_2 is the space of 2-planes in T_pM.

3. Let γ be a curve in $\mathbb{R}^3$, parameterized by arc length. Recall the Frenet apparatus. At $p = \gamma(t)$, $T = \gamma'(t)$ spans $T_p\gamma$, and if the curvature κ of γ is nonzero, unit vectors N and B span the normal bundle $\nu_p(\gamma)$, satisfying the system of ODE

$$\begin{aligned} T' &= \quad \kappa N, \\ N' &= -\kappa T \qquad\quad + \tau B, \\ B' &= \quad -\tau N, \end{aligned} \tag{4.62}$$

and furthermore, $B = T \times N$, $T = N \times B$, $N = B \times T$. Compare with Exercises 4–6 in Chapter 1, §5. Let ∇ denote the standard flat connection on $\mathbb{R}^3$, and ∇^0, ∇^1 the connections induced on $T(\gamma)$ and $\nu(\gamma)$, as in (4.1), (4.8). Show that

$$II(T,T) = \kappa N \tag{4.63}$$

and that

$$\begin{aligned} \nabla^1_T N &= \tau B, \\ \nabla^1_T B &= -\tau N. \end{aligned} \tag{4.64}$$

Compute the right side of the Weingarten formula

$$\nabla_T - \nabla^1_T = -(II_T)^t, \tag{4.65}$$

and show that (4.63)–(4.65) are *equivalent* to (4.62).

4. Let $S \subset \mathbb{R}^3$ be a surface, with connection ∇^S, second fundamental form II^S, and unit normal ν. Let γ be the curve of Exercise 3, and suppose γ is a curve in S. Show that

$$\begin{aligned} II^S(T,T) &= \kappa \langle N, \nu \rangle \nu \\ &= \kappa N - \nabla^S_T T. \end{aligned}$$

If A_ν denotes the Weingarten map of S, as in (4.5), show that

$$A_\nu T = \kappa T - \tau B \quad \text{and} \quad N = \nu,$$

provided γ is a geodesic on S.

5. Use Theorem 4.4 to show that the Gauss curvature K of a surface $S \subset \mathbb{R}^3$ is equal to $\det A_\nu$. Use the symmetry of A_ν to show that each T_pS has an orthonormal basis T_1, T_2 such that $A_\nu T_j = \kappa_j T_j$; hence $K = \kappa_1 \kappa_2$. An eigenvector of A_ν is called a *direction of principal curvature.* Show that

$T \in T_pS$ is a direction of principal curvature if and only if the geodesic through p in direction T has vanishing torsion τ at p.

6. Suppose M has the property that each sectional curvature $K_p(\Pi)$ is equal to K_p, independent of Π. Show that
$$\mathcal{R} = K_p I \quad \text{in } \operatorname{End}(\Lambda^2 T_p),$$
where $\mathcal{R}$ is as in Exercise 4 of §3. Show that K_p is constant, on each connected component of M, if dim $M \geq 3$. (*Hint*: To do the last part, use Proposition 3.3.)
7. Show that the formula (4.42) for $R - \widetilde{R}$ is equivalent to the formula (2.17). (This reiterates Exercise 5 of §2.) Also, relate (4.44) and (4.49) to (3.54).

Let M be a compact, oriented hypersurface in $\mathbb{R}^n$. Let
$$N : M \to S^{n-1}$$
be given by the outward-pointing normal. This is called the *Gauss map*.

8. If $n = 3$, show that $N^*\omega_0 = K\omega$, where ω_0 and ω are the area forms of S^2 and M, respectively, and K is the Gauss curvature of M. Note that the degree of the Gauss map is
$$\operatorname{Deg}(N) = \frac{1}{4\pi} \int\limits_M N^*\omega_0.$$
See §19 of Chapter 1 for basic material on degrees of maps.
9. For general n, show that $N^*\omega_0 = J\omega$, with
$$J = (-1)^{n-1} \det A_N,$$
where ω and ω_0 are the volume forms and $A_N : T_pM \to T_pM$ is the Weingarten map (4.5). Consequently,
$$\operatorname{Deg}(N) = \frac{(-1)^{n-1}}{A_{n-1}} \int\limits_M \Big(\det A_N\Big)\, dV, \tag{4.66}$$
where A_{n-1} is the area of S^{n-1}. (*Hint*: There is a natural identification of T_pM and $T_q(S^{n-1})$ as linear subspaces of $\mathbb{R}^n$, if $q = N(p)$. Show that the Weingarten formula gives
$$DN(p) = -A_N \in \operatorname{End}(T_pM) \approx \mathcal{L}(T_pM, T_qS^{n-1}).) \tag{4.67}$$
10. Let S be a hypersurface in $\mathbb{R}^n$, with second fundamental form $\widetilde{II}$, as in (4.15). Suppose $\widetilde{II}$ is proportional to the metric tensor, $\widetilde{II} = \lambda(x)g$. Show that λ is constant, provided S is connected. (*Hint*: Use the Codazzi equation (4.17), plus the fact that $\nabla^0 g = 0$.)
11. When S is a hypersurface in $\mathbb{R}^n$, a point p, where $\widetilde{II} = \lambda g$, is called an *umbilic point*. If every point on S is umbilic, show that S has constant sectional curvature λ^2.
(*Hint*: Apply Gauss' *Theorema Egregium*, in the form (4.14).)
12. Let $S \subset \mathbb{R}^n$ be a k-dimensional submanifold ($k < n$), with induced metric g and second fundamental form II. Let ξ be a section of the normal bundle

$\nu(S)$. Consider the one-parameter family of maps $S \to \mathbb{R}^n$,

$$\varphi_\tau(x) = x + \tau\xi(x), \quad x \in S,\ \tau \in (-\varepsilon, \varepsilon). \tag{4.68}$$

Let g_τ be the family of Riemannian metrics induced on S. Show that

$$\frac{d}{d\tau} g_\tau(X,Y)\Big|_{\tau=0} = -2\langle \xi, II(X,Y)\rangle. \tag{4.69}$$

More generally, if $S \subset M$ is a submanifold, consider the one-parameter family of submanifolds given by

$$\varphi_\tau(x) = \operatorname{Exp}_x\Bigl(\tau\xi(x)\Bigr), \quad x \in S,\ \tau \in (-\varepsilon, \varepsilon), \tag{4.70}$$

where $\operatorname{Exp}_x : T_xM \to M$ is the exponential map, determined by the Riemannian metric on M. Show that (4.69) holds in this more general case.

13. Let $M_1 \subset M_2 \subset M_3$ be Riemannian manifolds of dimension $n_1 < n_2 < n_3$, with induced metrics. For $j < k$, denote by II^{jk} the second fundamental form of $M_j \subset M_k$ and by A^{jk} the associated Weingarten map. For $x \in M_j$, denote by N_x^{jk} the orthogonal complement of T_xM_j in T_xM_k and by ${}^{jk}\nabla^1$ the natural connection on $N^{jk}(M_j)$. Let X and Y be tangent to M_1, and let ξ be a section of $N^{12}(M_1)$. Show that

$$A_\xi^{12}X = A_\xi^{13}X.$$

Also show that

$${}^{13}\nabla_X^1\xi = {}^{12}\nabla_X^1\xi + II^{23}(X,\xi), \text{ orthogonal decomposition,}$$

and that

$$II^{13}(X,Y) = II^{12}(X,Y) + II^{23}(X,Y), \text{ orthogonal decomposition.}$$

Relate this to Exercises 3–5 when $n_j = j$.

14. If $S \subset M$ has codimension 1 and Weingarten map $A : T_pS \to T_pS$, show that the Gauss equation (4.14) gives

$$\Bigl\langle (R - R_S)(X,Y)Z, W\Bigr\rangle = \Bigl\langle (\Lambda^2 A)(X \wedge Y), Z \wedge W\Bigr\rangle, \quad X, Y, Z, W \in T_pS. \tag{4.71}$$

Show that (with N a unit normal to S) the scalar curvatures of M and S are related by

$$S_M - S_S = -2 \operatorname{Tr} \Lambda^2 A + 2 \operatorname{Ric}_M(N,N). \tag{4.72}$$

5. The Gauss-Bonnet theorem for surfaces

If M is a compact, oriented Riemannian manifold of dimension 2, the Gauss-Bonnet theorem says that

$$\int_M K\, dV = 2\pi\, \chi(M), \tag{5.1}$$

if K is the Gauss curvature of M and $\chi(M)$ its Euler characteristic. There is an associated formula if M has a boundary. There are a number of significant variants of this, involving, for example, the index of a vector field. We present several proofs of the Gauss-Bonnet theorem and some of its variants here.

We begin with an estimate on the effect of parallel translation about a small closed, piecewise smooth curve. This first result holds for a general vector bundle $E \to M$ with connection ∇ and curvature

$$\Omega = \frac{1}{2}\, R^\alpha{}_{\beta jk}\, dx_j \wedge dx_k,$$

with no restriction on $\dim M$.

Proposition 5.1. *Let γ be a closed, piecewise smooth loop on M. Assume it is parameterized by arc length for $0 \le t \le b$, $\gamma(b) = \gamma(0)$. If $u(t)$ is a section of E over γ defined by parallel transport (i.e., $\nabla_T u = 0,\ T = \dot\gamma$), then*

$$u^\alpha(b) - u^\alpha(0) = -\frac{1}{2}\sum_{j,k,\beta} R^\alpha{}_{\beta jk}\Bigl(\int_A dx_j \wedge dx_k\Bigr) u^\beta(0) + O(b^3), \tag{5.2}$$

where A is an oriented 2-surface in M with $\partial A = \gamma$, and the $u^\alpha(t)$ are the components of u with respect to a local frame.

Proof. If we put a coordinate system on a neighborhood of $p = \gamma(0) \in M$ and choose a frame field for E, then parallel transport is defined by

$$\frac{du^\alpha}{dt} = -\Gamma^\alpha{}_{\beta k} u^\beta \frac{dx_k}{dt}. \tag{5.3}$$

As usual, we use the summation convention. Thus

$$u^\alpha(t) = u^\alpha(0) - \int_0^t \Gamma^\alpha{}_{\beta k}(\gamma(s)) u^\beta(s)\, \frac{dx_k}{ds}\, ds. \tag{5.4}$$

We hence have

$$u^\alpha(t) = u^\alpha(0) - \Gamma^\alpha{}_{\beta k}(p) u^\beta(0)(x_k - p_k) + O(t^2). \tag{5.5}$$

We can solve (5.3) up to $O(t^3)$ if we use

$$\Gamma^\alpha{}_{\beta j}(x) = \Gamma^\alpha{}_{\beta j}(p) + (x_k - p_k)\, \partial_k \Gamma^\alpha{}_{\beta j} + O\bigl(|x-p|^2\bigr). \tag{5.6}$$

Hence

$$\begin{aligned} u^\alpha(t) = u^\alpha(0) - \int_0^t &\bigl[\Gamma^\alpha{}_{\beta k}(p) + (x_j - p_j)\, \partial_j \Gamma^\alpha{}_{\beta k}(p)\bigr] \\ &\cdot \bigl[u^\beta(0) - \Gamma^\beta{}_{\gamma\ell}(p) u^\gamma(0)(x_\ell - p_\ell)\bigr]\, \frac{dx_k}{ds}\, ds + O(t^3). \end{aligned} \tag{5.7}$$

If $\gamma(b) = \gamma(0)$, we get

$$\begin{aligned}u^\alpha(b) = u^\alpha(0) &- \int_0^b x_j\,dx_k\big(\partial_j \Gamma^\alpha{}_{\beta k}\big)u^\beta(0)\\ &+ \int_0^b x_j\,dx_k \Gamma^\alpha{}_{\beta k}\Gamma^\beta{}_{\gamma j}u^\gamma(0) + O(b^3),\end{aligned} \tag{5.8}$$

the components of Γ and their first derivatives being evaluated at p. Now Stokes' theorem gives

$$\int\limits_\gamma x_j\,dx_k = \int\limits_A dx_j \wedge dx_k,$$

so

$$u^\alpha(b) - u^\alpha(0) = \Big[-\partial_j\Gamma^\alpha{}_{\beta k} + \Gamma^\alpha{}_{\gamma k}\Gamma^\gamma{}_{\beta j}\Big]\int\limits_A dx_j \wedge dx_k\ u^\beta(0) + O(b^3). \tag{5.9}$$

Recall that the curvature is given by

$$\Omega = d\Gamma + \Gamma \wedge \Gamma,$$

that is,

$$R^\alpha{}_{\beta jk} = \partial_j \Gamma^\alpha{}_{\beta k} - \partial_k \Gamma^\alpha{}_{\beta j} + \Gamma^\alpha{}_{\gamma j}\Gamma^\gamma{}_{\beta k} - \Gamma^\alpha{}_{\gamma k}\Gamma^\gamma{}_{\beta j}. \tag{5.10}$$

Now the right side of (5.10) is the antisymmetrization, with respect to j and k, of the quantity in brackets in (5.9). Since $\int_A dx_j \wedge dx_k$ is antisymmetric in j and k, we get the desired formula (5.2).

In particular, if dim $M = 2$, then we can write the End(E)-valued 2-form Ω as

$$\Omega = \mathcal{R}\mu, \tag{5.11}$$

where μ is the volume form on M and $\mathcal{R}$ is a smooth section of End(E) over M. If E has an inner product and ∇ is a metric connection, then $\mathcal{R}$ is skew-adjoint. If γ is a geodesic triangle that is "fat" in the sense that none of its angles is small, (5.2) implies

$$u(b) - u(0) = -\mathcal{R}u(0)(\text{Area } A) + O\big((\text{Area } A)^{3/2}\big). \tag{5.12}$$

If we specialize further, to oriented two-dimensional M with $E = TM$, possessing the Levi-Civita connection of a Riemannian metric, then we take $J : T_pM \to T_pM$ to be the counterclockwise rotation by 90°, which defines an almost complex structure on M. Up to a scalar this is the unique skew-adjoint operator on T_pM, and, by (3.34),

$$\mathcal{R}u = -K\ Ju, \quad u \in T_pM, \tag{5.13}$$

where K is the Gauss curvature of M at p. Thus, in this case, (5.12) becomes

$$u(b) - u(0) = K\ Ju(0)(\text{Area } A) + O\big((\text{Area } A)^{3/2}\big). \tag{5.14}$$

FIGURE 5.1

On the other hand, if a tangent vector $X_0 \in T_pM$ undergoes parallel transport around a geodesic triangle, the action produced on T_pM is easily seen to be a rotation in T_pM through an angle that depends on the angle defect of the triangle. The argument can be seen by looking at Fig. 5.1. We see that the angle from X_0 to X_3 is

$$(\pi + \alpha) - (2\pi - \beta - \gamma - \xi) - \xi = \alpha + \beta + \gamma - \pi. \tag{5.15}$$

In this case, formula (5.14) implies

$$\alpha + \beta + \gamma - \pi = \int K\ dV + O\big((\text{Area } A)^{3/2}\big). \tag{5.16}$$

We can now use a simple analytical argument to sharpen this up to the following celebrated formula of Gauss.

Theorem 5.2. *If A is a geodesic triangle in M^2, with angles α, β, and γ, then*

$$\alpha + \beta + \gamma - \pi = \int_A K \, dV. \tag{5.17}$$

Proof. Break up the geodesic triangle A into N^2 little geodesic triangles, each of diameter $O(N^{-1})$, area $O(N^{-2})$. Since the angle defects are *additive*, the estimate (5.17) implies

$$\begin{aligned} \alpha + \beta + \gamma - \pi &= \int_A K \, dV + N^2 O\big((N^{-2})^{3/2}\big) \\ &= \int_A K \, dV + O(N^{-1}), \end{aligned} \tag{5.18}$$

and passing to the limit as $N \to \infty$ gives (5.17).

Note that any region that is a contractible geodesic polygon can be divided into geodesic triangles. If a contractible region $\Omega \subset M$ with smooth boundary is approximated by geodesic polygons, a straightforward limit process yields the Gauss-Bonnet formula

$$\int_\Omega K \, dV + \int_{\partial\Omega} \kappa \, ds = 2\pi, \tag{5.19}$$

where κ is the geodesic curvature of $\partial\Omega$. We leave the details to the reader. Another proof will be given at the end of this section.

If M is a compact, oriented two-dimensional manifold without boundary, we can partition M into geodesic triangles. Suppose the triangulation of M so produced has

$$F \text{ faces (triangles)}, \quad E \text{ edges}, \quad V \text{ vertices}. \tag{5.20}$$

If the angles of the jth triangle are α_j, β_j, and γ_j, then summing all the angles clearly produces $2\pi V$. On the other hand, (5.17) applied to the jth triangle, and summed over j, yields

$$\sum_j (\alpha_j + \beta_j + \gamma_j) = \pi F + \int_M K \, dV. \tag{5.21}$$

Hence $\int_M K \, dV = (2V - F)\pi$. Since in this case all the faces are triangles, counting each triangle three times will count each edge twice, so $3F = 2E$. Thus we obtain

$$\int_M K \, dV = 2\pi(V - E + F). \tag{5.22}$$

This is equivalent to (5.1), in view of Euler's formula

$$\chi(M) = V - E + F. \tag{5.23}$$

Figure 5.2

We now derive a variant of (5.1) when M is described in another fashion. Namely, suppose M is diffeomorphic to a sphere with g handles attached. The number g is called the *genus* of the surface. The case $g = 2$ is illustrated in Fig. 5.2. We claim that

$$\int\limits_M K \; dV = 4\pi(1 - g) \tag{5.24}$$

in this case. By virtue of (5.22), this is equivalent to the identity

$$2 - 2g = V - E + F = \chi(M). \tag{5.25}$$

Direct proofs of this are possible, but we will provide a proof of (5.24), based on the fact that

$$\int\limits_M K \; dV = C(M) \tag{5.26}$$

depends only on M, not on the metric imposed. This follows from (5.22), by forgetting the interpretation of the right side. The point we want to make is, given (5.26)—that is, the independence of the choice of metric—we can work out what $C(M)$ is, as follows.

First, choosing the standard metric on S^2, for which $K = 1$ and the area is 4π, we have

$$\int\limits_{S^2} K \; dV = 4\pi. \tag{5.27}$$

Now suppose M is obtained by adding g handles to S^2. Since we can alter the metric on M at will, we can make sure it coincides with the metric of a sphere near a great circle, in a neighborhood of each circle where a handle is attached to the main body A, as illustrated in Fig. 5.2. If we imagine adding two hemispherical caps to each handle H_j, rather than attaching it to A, we turn each H_j into a new sphere, so by (5.27) we have

$$4\pi = \int\limits_{H_j \cup \text{ caps}} K\ dV = \int\limits_{H_j} K\ dV + \int\limits_{\text{caps}} K\ dV. \tag{5.28}$$

Since the caps fit together to form a sphere, we have $\int_{\text{caps}} K\ dV = 4\pi$, so for each j,

$$\int\limits_{H_j} K\ dV = 0, \tag{5.29}$$

provided M has a metric such as described above. Similarly, if we add $2g$ caps to the main body A, we get a new sphere, so

$$4\pi = \int\limits_{A \cup \text{ caps}} K\ dV = \int\limits_{A} K\ dV + 2g(2\pi), \tag{5.30}$$

or

$$\int\limits_{A} K\ dV = 2\pi(2 - 2g). \tag{5.31}$$

Together (5.29) and (5.31) yield (5.24), and we get the identity (5.25) as a bonus.

We now give another perspective on Gauss' formula, directly dealing with the fact that TM can be treated as a complex line bundle, when M is an oriented Riemannian manifold of dimension 2. We will produce a variant of Proposition 5.1 which has no remainder term and which hence produces (5.16) with no remainder, directly, so Theorem 5.2 follows without the additional argument given above. The result is the following; again dim M is unrestricted.

Proposition 5.3. *Let $E \to M$ be a complex line bundle. Let γ be a piecewise smooth, closed loop in M, with $\gamma(0) = \gamma(b) = p$, bounding a surface A. Let ∇ be a connection on E, with curvature Ω. If $u(t)$ is a section of E over γ defined by parallel translation, then*

$$u(b) = \Bigl[\exp\Bigl(-\int\limits_{A} \Omega\Bigr)\Bigr] u(0). \tag{5.32}$$

Proof. Pick a nonvanishing section (hence a frame field) ξ of E over S, assuming S is homeomorphic to a disc. Any section u of E over S is of the form $u = v\xi$ for a complex-valued function v on S. Then parallel transport along $\gamma(t) = (x_1(t), \dots, x_n(t))$ is defined by

$$\frac{dv}{dt} = -\Big(\Gamma_k \frac{dx_k}{dt}\Big)\ v. \tag{5.33}$$

The solution to this single, first-order ODE is

$$v(t) = \Big[\exp\Big(-\int_0^t \Gamma_k(\gamma(s)) \frac{dx_k}{ds}\ ds\Big)\Big]\ v(0). \tag{5.34}$$

Hence

$$v(b) = \Big[\exp\Big(-\int\limits_\gamma \Gamma\Big)\Big] v(0), \tag{5.35}$$

where $\Gamma = \sum \Gamma_k\, dx_k$. The curvature 2-form Ω is given, as a special case of (5.10), by

$$\Omega = d\Gamma, \tag{5.36}$$

and Stokes' theorem gives (5.32), from (5.35), provided A is contractible. The general case follows from cutting A into contractible pieces.

As we have mentioned, Proposition 5.3 can be used in place of Proposition 5.1, in conjunction with the argument involving Fig. 5.1, to prove Theorem 5.2.

Next, we relate $\int_M \Omega$ to the "index" of a section of a complex line bundle $E \to M$, when M is a compact, oriented manifold of dimension 2. Suppose X is a section of E over $M \setminus S$, where S consists of a finite number of points; suppose that X is nowhere vanishing on $M \setminus S$ and that, near each $p_j \in S$, X has the following form. There are a coordinate neighborhood $\mathcal{O}_j$ centered at p_j, with p_j the origin, and a nonvanishing section ξ_j of E near p_j, such that

$$X = v_j \xi_j \ \text{ on } \mathcal{O}_j, \quad v_j : \mathcal{O}_j \setminus p \to \mathbb{C} \setminus 0. \tag{5.37}$$

Taking a small counterclockwise circle γ_j about p_j, $v_j/|v_j| = \omega_j$ maps γ_j to S^1; consider the degree ℓ_j of this map, that is, the winding number of γ_j about S^1. This is the index of X at p_j, and the sum over p_j is the total index of X:

$$\text{Index}(X) = \sum_j \ell_j. \tag{5.38}$$

We will establish the following.

Proposition 5.4. *For any connection on $E \to M$, with curvature form Ω and X as above, we have*

$$\int\limits_M \Omega = -(2\pi i) \cdot\ \textit{Index}(X). \tag{5.39}$$

Proof. You can replace X by a section of $E \setminus 0$ over $M \setminus \{p_j\}$, homotopic to the original, having the form (5.37) with

$$v_j = e^{i\ell_j\theta} + w_j, \tag{5.40}$$

in polar coordinates (r,θ) about p_j, with $w_j \in C^1(\mathcal{O}_j),\ w_j(0) = 0$. Excise small disks $\mathcal{D}_j$ containing p_j; let $\mathcal{D} = \cup\mathcal{D}_j$. Then, by Stokes' theorem,

$$\int\limits_{M\setminus\mathcal{D}} \Omega = -\sum_j \int\limits_{\gamma_j} \Gamma, \tag{5.41}$$

where $\gamma_j = \partial\mathcal{D}_j$ and Γ is the connection 1-form with respect to the section X, so that, with $\nabla_k = \nabla_{D_k},\ D_k = \partial/\partial x_k$ in local coordinates,

$$\nabla_k X = \Gamma_k X. \tag{5.42}$$

Now (5.37) gives (with no summation)

$$\Gamma_k v_j \xi_j = (\partial_k v_j + v_j \widetilde{\Gamma}_{jk})\xi_j \tag{5.43}$$

on $\overline{\mathcal{D}}_j$, where $\widetilde{\Gamma}_{jk}\,dx_k$ is the connection 1-form with respect to the section ξ_j. Hence

$$\Gamma_k = v_j^{-1}\,\partial_k v_j + \widetilde{\Gamma}_{jk}, \tag{5.44}$$

with remainder term $\widetilde{\Gamma}_{jk} \in C^1(\mathcal{O}_j)$. By (5.40), we have

$$\int\limits_{\gamma_j} \Gamma = 2\pi i\ell_j + O(r) \tag{5.45}$$

if each $\mathcal{D}_j$ has radius $\le Cr$. Passing to the limit as the disks $\mathcal{D}_j$ shrink to p_j gives (5.39).

Since the left side of (5.39) is independent of the choice of X, it follows that the index of X depends only on E, not on the choice of such X. In Chapter 10, this formula is applied to a meromorphic section of a complex line bundle, and in conjunction with the Riemann-Roch formula it yields important information on Riemann surface theory.

In case M is a compact, oriented Riemannian 2-manifold, whose tangent bundle can be given the structure of a complex line bundle as noted above, (5.39) is equivalent to

$$\int\limits_M K\ dV = 2\pi\ \mathrm{Index}(X), \tag{5.46}$$

for any smooth vector field X, nonvanishing, on M minus a finite set of points. This verifies the identity

$$\mathrm{Index}(X) = \chi(M) \tag{5.47}$$

in this case.

As a further comment on the Gauss-Bonnet formula for compact surfaces, let us recall from Exercise 8 of §4 that if M is a compact, oriented surface in $\mathbb{R}^3$, with Gauss map $N : M \to S^2$, then

$$\operatorname{Deg}(N) = \frac{1}{4\pi} \int\limits_M N^* \omega_0 = \frac{1}{4\pi} \int\limits_M K \, dV. \tag{5.48}$$

Furthermore, in §20 of Chapter 1, Corollary 20.5 yields an independent proof that, in this case,

$$\operatorname{Deg}(N) = \frac{1}{2} \operatorname{Index}(X), \tag{5.49}$$

for any vector field X on M with a finite number of critical points. Hence (5.48)–(5.49) provide another proof of (5.1), at least for a surface in $\mathbb{R}^3$. This line of reasoning will be extended to the higher-dimensional case of hypersurfaces of $\mathbb{R}^{n+1}$, in the early part of §8, as preparation for establishing the general Chern-Gauss-Bonnet theorem.

To end this section, we provide a direct proof of the formula (5.19), using an argument parallel to the proof of Proposition 5.3. Thus, assuming that M is an oriented surface, we give TM the structure of a complex line bundle, and we pick a nonvanishing section ξ of TM over a neighborhood of $\overline{\mathcal{O}}$. Let $\gamma = \partial \mathcal{O}$ be parameterized by arc length, $T = \gamma'(s),\ 0 \le s \le b$, with $\gamma(b) = \gamma(0)$. The geodesic curvature κ of γ, appearing in (5.19), is given by

$$\nabla_T T = \kappa N, \quad N = JT. \tag{5.50}$$

If we set $T = u\xi$, where $u : \overline{\mathcal{O}} \to \mathbb{C}$, then, parallel to (5.33), we have (5.50) equivalent to

$$\frac{du}{ds} = -\sum \Gamma_k \frac{dx_k}{ds} u + i\kappa u. \tag{5.51}$$

The solution to this single, first-order ODE is (parallel to (5.34))

$$u(t) = \Bigl[\exp\Bigl(i \int_0^t \kappa(s) \, ds - \int_0^t \Gamma_k\bigl(\gamma(s)\bigr) \frac{dx_k}{ds} \, ds \Bigr) \Bigr] \, u(0). \tag{5.52}$$

Hence

$$u(b) = \Bigl[\exp\Bigl(i \int\limits_\gamma \kappa(s) \, ds - \int\limits_{\mathcal{O}} \Omega \Bigr) \Bigr] \, u(0). \tag{5.53}$$

By (5.13), we have

$$\Omega = -iK \, dV, \tag{5.54}$$

and since $u(b) = u(0)$, we have

$$\exp\Bigl(i \int\limits_\gamma \kappa(s) \, ds + i \int\limits_{\mathcal{O}} K \, dV \Bigr) = 1, \tag{5.55}$$

or

$$\int_{\mathcal{O}} K\ dV + \int_{\gamma} \kappa(s)\ ds = 2\pi\nu, \tag{5.56}$$

for some $\nu \in \mathbb{Z}$. Now if $\mathcal{O}$ were a tiny disc in M, it would be clear that $\nu = 1$. Using the contractibility of $\mathcal{O}$ and the fact that the left side of (5.56) cannot jump, we have $\nu = 1$, which proves (5.19).

Exercises

1. Given a triangulation of a compact surface M, within each triangle construct a vector field, vanishing at seven points as illustrated in Fig. 5.3, with the vertices as attractors, the center as a repeller, and the midpoints of each side as saddle points. Fit these together to produce a smooth vector field X on M. Show directly that

$$\operatorname{Index}(X) = V - E + F.$$

FIGURE 5.3

2. Let $L \to M$ be a complex line bundle, and let u and v be sections of L with a finite number of zeros. Show directly that u and v have the same index. (*Hint*: Start with $u = fv$ on $M \setminus Z$, where Z is the union of the zero sets and $f : M \setminus Z \to \mathbb{C} \setminus 0$.)
3. Let M_1 and M_2 be n-dimensional submanifolds of $\mathbb{R}^k$. Suppose a curve γ is contained in the intersection $M_1 \cap M_2$, and assume

$$p = \gamma(s) \Longrightarrow T_pM_1 = T_pM_2.$$

   Show that parallel translations along γ in M_1 and in M_2 coincide. (*Hint*: If $T = \gamma'(s)$ and X is a vector field along γ, tangent to M_1 (hence to M_2), show that $\nabla^{M_1}_T X = \nabla^{M_2}_T X$, using Corollary 4.2.)
4. Let $\mathcal{O}$ be the region in $S^2 \subset \mathbb{R}^3$ consisting of points in S^2 of geodesic distance $< r$ from $p = (0, 0, 1)$, where $r \in (0, \pi)$ is given. Let $\gamma = \partial\mathcal{O}$. Construct

a cone, with vertex at $(0, 0, \sec r)$, tangent to S^2 along γ. Using this and Exercise 3, show that parallel translation over one circuit of γ is given by

$$\text{counterclockwise rotation by } \theta = 2\pi(1 - \cos r).$$

(*Hint*: Flatten out the cone, as in Fig. 5.4. Notice that γ has length $\ell = 2\pi \ \sin r$.)
Compare this calculation with the result of (5.32), which in this context implies

$$u(\ell) = \left[\exp i \int_{\mathcal{O}} K \ dV\right] u(0).$$

FIGURE 5.4

5. Let $\gamma : [a, b] \to \mathbb{R}^3$ be a smooth, closed curve, so $\gamma(a) = \gamma(b)$ and $\gamma'(a) = \gamma'(b)$. Assume γ is parameterized by arc length, so $\gamma'(t) = T(t)$ and $T : [a, b] \to S^2$; hence T is a smooth, closed curve in S^2. Note that the normal space to γ at $p = \gamma(t)$ is naturally identified with the tangent space to S^2 at $T(t) = q$:

$$\nu_p(\gamma) = T_q S^2.$$

a) Show that parallel translation along γ of a section of the normal bundle $\nu(\gamma)$, with respect to the connection described in Exercise 3 of §4, coincides with parallel translation along the curve T of vectors tangent to S^2. (*Hint*: Recall Exercise 3 of §1.)
b) Suppose the curvature κ of γ never vanishes, so the torsion τ is well defined, as in (4.62). Show that parallel translation once around γ acts on $\nu_p(\gamma)$ by multiplication by

$$\exp\left(-i \int_{\gamma} \tau(s) \ ds\right).$$

Here we use the complex structure on $\nu_p(\gamma)$ given by $JN = B, \ JB = -N$. (*Hint*: Use (4.64).) Compare the results of parts a and b.

6. The principal bundle picture

An important tool for understanding vector bundles is the notion of an underlying structure, namely that of a principal bundle. If M is a manifold and G a Lie group, then a principal G-bundle $P \xrightarrow{p} M$ is a locally trivial fibration with a G-action on P such that G acts on each fiber $P_x = p^{-1}(x)$ in a simply transitive fashion. An example is the frame bundle of an oriented Riemannian manifold M, $F(M) \to M$, where $F_x(M)$ consists of the set of ordered oriented orthonormal bases of the tangent space T_x to M at x. If $n = \dim M$, this is a principal $\mathrm{SO}(n)$-bundle.

If $P \to M$ is a principal G-bundle, then associated to each representation π of G on a vector space V is a vector bundle $E \to M$. The set E is a quotient space of the Cartesian product $P \times V$, under the equivalence relation

$$(y, v) \sim (y \cdot g, \pi(g)^{-1} v), \quad g \in G. \tag{6.1}$$

We have written the G-action on P as a right action. One writes $E = P \times_\pi V$. The space of sections of E is naturally isomorphic to a certain subspace of the space of V-valued functions on P:

$$C^\infty(M, E) \approx \{u \in C^\infty(P, V) : u(y \cdot g) = \pi(g)^{-1} u(y),\ g \in G\}. \tag{6.2}$$

We describe how this construction works for the frame bundle $F(M)$ of an oriented Riemannian manifold, which, as mentioned above, is a principal $\mathrm{SO}(n)$-bundle. Thus, a point $y \in F_x(M)$ consists of an n-tuple $(e_1, \dots, e_n)$, forming an ordered, oriented orthonormal basis of $T_x M$. If $g = (g_{jk}) \in \mathrm{SO}(n)$, the G-action is given by

$$(e_1, \dots, e_n) \cdot g = (f_1, \dots, f_n), \quad f_j = \sum_\ell g_{\ell j} e_\ell. \tag{6.3}$$

One can check that $(f_1, \dots, f_n)$ is also an oriented orthonormal basis of $T_x M$ and that $(y \cdot g) \cdot g' = y \cdot (gg')$, for $g, g' \in \mathrm{SO}(n)$. If π is the "standard" representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$, given by matrix multiplication, we claim that there is a natural identification

$$F(M) \times_\pi \mathbb{R}^n \approx TM. \tag{6.4}$$

In fact, if $y = (e_1, \dots, e_n) \in F_x(M)$ and $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, the map (6.4) is defined by

$$(y, v) \mapsto \sum_j v_j e_j \in T_x M. \tag{6.5}$$

We need to show that this is constant on equivalence classes, as defined by (6.1), that is, for any $g \in \mathrm{SO}(n)$,

$$z = y \cdot g = (f_1, \dots, f_n), \quad w = \pi(g)^{-1} v \Longrightarrow \sum w_k f_k = \sum v_j e_j. \tag{6.6}$$

In fact, setting $g^{-1} = h = (h_{jk})$, we see that

$$\sum_k w_k f_k = \sum_{j,k,\ell} h_{kj} v_j g_{\ell k} e_\ell = \sum_{j,\ell} \delta_{\ell j} v_j e_\ell \tag{6.7}$$

since

$$\sum_k g_{\ell k} h_{kj} = \delta_{\ell j}, \tag{6.8}$$

and this implies (6.6).

Connections are naturally described in terms of a geometrical structure on a principal bundle. This should be expected since, as we saw in §1, a connection on a vector bundle can be described in terms of a "connection 1-form" (1.14), depending on a choice of local frame for the vector bundle.

The geometrical structure giving a connection on a principal bundle $P \to M$ is the following. For each $y \in P$, the tangent space $T_y P$ contains the subspace $V_y P$ of vectors tangent to the fiber $p^{-1}(x),\ x = p(y)$. The space $V_y P$, called the "vertical space," is naturally isomorphic to the Lie algebra $\mathfrak{g}$ of G. A connection on P is determined by a choice of complementary subspace, called a "horizontal space":

$$T_y P = V_y P \oplus H_y P, \tag{6.9}$$

with the G-invariance

$$g_*(H_y P) = H_{y\cdot g} P, \tag{6.10}$$

where $g_* : T_y P \to T_{y\cdot g} P$ is the natural derivative map.

Given this structure, a vector field X on M has a uniquely defined "lift" $\widetilde{X}$ to a vector field on P, such that $p_* \widetilde{X}_y = X_x\ (x = p(y))$ and $\widetilde{X}_y \in H_y P$ for each $y \in P$. Furthermore, if E is a vector bundle determined by a representation of G and $u \in C^\infty(M, V)$ corresponds to a section v of E, the V-valued function $\widetilde{X} \cdot u$ corresponds to a section of E, which we denote $\nabla_X v$; ∇ is the covariant derivative on E defined by the connection on P just described. If V has an inner product and π is unitary, E gets a natural metric and ∇ is a metric connection on E.

If the π_j are representations of G on V_j, giving vector bundles $E_j \to M$ associated to a principal bundle $P \to M$ with connection, then $\pi_1 \otimes \pi_2$ is a representation of G on $V_1 \otimes V_2$, and we have a vector bundle $E \to M,\ E = E_1 \otimes E_2$. The prescription above associating a connection to E as well as to E_1 and E_2 agrees with the definition in (1.29) of a connection on a tensor product of two vector bundles. This follows simply from the derivation property of the vector field $\widetilde{X}$, acting as a first-order differential operator on functions on P.

The characterization (6.9)–(6.10) of a connection on a principal bundle $P \to M$ is equivalent to the following, in view of the natural isomorphism $V_y P \approx \mathfrak{g}$. The splitting (6.9) corresponds to a projection of $T_y P$ onto $V_y P$, hence to a linear map $T_yP \to \mathfrak{g}$, which gives the identification $V_yP \approx \mathfrak{g}$ on the linear subspace V_yP of T_yP. This map can be regarded as a $\mathfrak{g}$-valued 1-form ξ on P, called the *connection form*, and the invariance property (6.10) is equivalent to

$$g^*\xi = Ad_{g^{-1}}\xi, \quad g \in G, \tag{6.11}$$

where $g^*\xi$ denotes the pull-back of the form ξ, induced from the G-action on P.

The Levi-Civita connection on an oriented Riemannian manifold gives rise to a connection on the frame bundle $F(M) \to M$ in the following way. Fix $y \in F(M)$, $x = p(y)$. Recall that the point y is an ordered (oriented) orthonormal basis $(e_1, \dots, e_n)$ of the tangent space T_xM. The parallel transport of each e_j along a curve γ through x thus gives a family of orthonormal bases for the tangent space to M at $\gamma(t)$, hence a curve $\gamma^\#$ in $F(M)$ lying over γ. The tangent to $\gamma^\#$ at y belongs to the horizontal space $H_yF(M)$, which in fact consists of all such tangent vectors as the curve γ through x is varied. This construction generalizes to other vector bundles $E \to M$ with connection ∇. One can use the bundle of orthonormal frames for E if ∇ is a metric connection, or the bundle of general frames for a general connection.

Let us restate how a connection on a principal bundle gives rise to connections on associated vector bundles. Given a principal G-bundle $P \to M$, consider a local section σ of P, over $U \subset M$. If we have a representation π of G on V, the associated vector bundle $E \to M$, and a section u of E, then we have $u \circ \sigma : U \to V$, using the identification (6.2). Given a connection on P, with connection 1-form ξ, we can characterize the covariant derivative induced on sections of E by

$$(\nabla_X u) \circ \sigma = \mathcal{L}_X(u \circ \sigma) + \Gamma(X) u \circ \sigma, \tag{6.12}$$

where $\mathcal{L}_X$ acts componentwise on $u \circ \sigma$, and

$$\Gamma(X) = (d\pi)\big(\xi_y(\widehat{X})\big), \quad y = \sigma(x), \quad \widehat{X} = D\sigma(x)X, \tag{6.13}$$

$d\pi$ denoting the derived representation of $\mathfrak{g}$ on V. That (6.12) agrees with $(\mathcal{L}_{\widetilde{X}} u) \circ \sigma$ follows from the chain rule, the fact that $\widehat{X} - \widetilde{X}$ is vertical, and the fact that if $v \in T_yP$ is vertical, then, by (6.2), $\mathcal{L}_v u = -d\pi\big(\xi(v)\big)u$. Note the similarity of (6.12) to (1.7).

Recall the curvature $R(X, Y)$ of a connection ∇ on a vector bundle $E \to M$, defined by the formula (1.10). In case $E = P \times_\pi V$, and ∇u is defined as above, we have (using the identification in (6.2))

$$R(X, Y)u = \mathcal{L}_{[\widetilde{X}, \widetilde{Y}]} u - \mathcal{L}_{\widetilde{[X,Y]}} u. \tag{6.14}$$

Alternatively, using (6.12) and (6.13), we see that the curvature of ∇ is given by

$$\begin{aligned} &R(X,Y)u\circ\sigma \\ &\quad=\Big\{\mathcal{L}_X\Gamma(Y)-\mathcal{L}_Y\Gamma(X)+\big[\Gamma(X),\Gamma(Y)\big]-\Gamma\big([X,Y]\big)\Big\}u\circ\sigma. \end{aligned} \tag{6.15}$$

This is similar to (1.13). Next we want to obtain a formula similar to (but more fundamental than) (1.15).

Fix $y\in P,\ x=p(y)$. It is convenient to calculate (6.15) at x by picking the local section σ to have the property that

$$D\sigma(x):T_xM\longrightarrow H_yP, \tag{6.16}$$

which is easily arranged. Then $\widehat{X}=\widetilde{X}$ at y, so $\Gamma(X)=0$ at y. Hence, at x,

$$\begin{aligned} R(X,Y)u\circ\sigma&=\big\{\mathcal{L}_X\Gamma(Y)-\mathcal{L}_Y\Gamma(X)\big\}u\circ\sigma \\ &=(d\pi)\big\{\widehat{X}\cdot\xi(\widehat{Y})-\widehat{Y}\cdot\xi(\widehat{X})\big\}u\circ\sigma \\ &=(d\pi)\big\{(d\sigma^*\xi)(X,Y)+(\sigma^*\xi)([X,Y])\big\}u\circ\sigma. \end{aligned} \tag{6.17}$$

Of course, $\sigma^*\xi=0$ at x. Thus we see that

$$R(X,Y)u=(d\pi)\big\{(d\xi)(\widetilde{X},\widetilde{Y})\big\}u, \tag{6.18}$$

at y, and hence everywhere on P. In other words,

$$R(X,Y)=(d\pi)\big(\Omega(\widetilde{X},\widetilde{Y})\big), \tag{6.19}$$

where Ω is the $\mathfrak{g}$-valued 2-form on P defined by

$$\Omega(X^\#,Y^\#)=(d\xi)(\varkappa X^\#,\varkappa Y^\#), \tag{6.20}$$

for $X^\#,Y^\#\in T_yP$. Here, $\varkappa$ is the projection of T_yP onto H_yP, with respect to the splitting (6.9). One calls Ω the curvature 2-form of the connection ξ on P.

If V and W are smooth vector fields on P, then

$$(d\xi)(V,W)=V\cdot\xi(W)-W\cdot\xi(V)-\xi\big([V,W]\big). \tag{6.21}$$

In particular, if $V=\widetilde{X},\ W=\widetilde{Y}$ are horizontal vector fields on P, then since $\xi(\widetilde{X})=\xi(\widetilde{Y})=0$, we have

$$(d\xi)(\widetilde{X},\widetilde{Y})=-\xi\big([\widetilde{X},\widetilde{Y}]\big). \tag{6.22}$$

Hence, given $X^\#,Y^\#\in T_yP$, we have

$$\Omega(X^\#,Y^\#)=-\xi\big([\widetilde{X},\widetilde{Y}]\big), \tag{6.23}$$

where $\widetilde{X}$ and $\widetilde{Y}$ are any horizontal vector fields on P such that $\widetilde{X}=\varkappa X^\#$ and $\widetilde{Y}=\varkappa Y^\#$ at $y\in P$. Since ξ annihilates $[\widetilde{X},\widetilde{Y}]$ if and only if it is horizontal, we see that Ω measures the failure of the bundle of horizontal spaces to be involutive.

It follows from Frobenius's theorem that, if $\Omega = 0$ on P, there is an integral manifold $S \subset P$ such that, for each $y \in S$, $T_yS = H_yP$. Each translate $S \cdot g$ is also an integral manifold. We can use this family of integral manifolds to construct local sections $v_1, \dots v_K$ of E ($K = \dim V$), linearly independent at each point, such that $\nabla v_j = 0$ for all j, given that $\Omega = 0$. Thus we recover Proposition 1.2, in this setting.

The following important result is *Cartan's formula* for the curvature 2-form.

Theorem 6.1. *We have*

$$\Omega = d\xi + \frac{1}{2}[\xi, \xi]. \tag{6.24}$$

The bracket $[\xi, \eta]$ of $\mathfrak{g}$-valued 1-forms is defined as follows. Suppose, in local coordinates,

$$\xi = \sum \xi_j \, dx_j, \quad \eta = \sum \eta_k \, dx_k, \quad \xi_j, \ \eta_k \in \mathfrak{g}. \tag{6.25}$$

Then we set

$$[\xi, \eta] = \sum_{j,k} [\xi_j, \eta_k] \, dx_j \wedge dx_k = \sum_{j<k} \bigl([\xi_j, \eta_k] + [\eta_j, \xi_k]\bigr) \, dx_j \wedge dx_k, \tag{6.26}$$

which is a $\mathfrak{g}$-valued 2-form. Equivalently, if U and V are vector fields on P,

$$[\xi, \eta](U, V) = [\xi(U), \eta(V)] + [\eta(U), \xi(V)]. \tag{6.27}$$

In particular,

$$\frac{1}{2}[\xi, \xi](U, V) = [\xi(U), \xi(V)]. \tag{6.28}$$

Note that if π is a representation of G on a vector space V and $d\pi$ the derived representation of $\mathfrak{g}$ on V, if we set $A_j = d\pi(\xi_j)$, then, for

$$d\pi(\xi) = \alpha = \sum A_j \, dx_j, \tag{6.29}$$

we have

$$\alpha \wedge \alpha = \sum_{j,k} A_j A_k \, dx_j \wedge dx_k = \frac{1}{2} \sum_{j,k} (A_j A_k - A_k A_j) \, dx_j \wedge dx_k. \tag{6.30}$$

Hence

$$\alpha \wedge \alpha = \frac{1}{2} (d\pi)[\xi, \xi]. \tag{6.31}$$

Thus we see the parallel between (6.24) and (1.15).

To prove (6.24), one evaluates each side on $(X^\#, Y^\#)$, for $X^\#, Y^\# \in T_yP$. We write $X^\# = \widetilde{X} + X_v$, with $\widetilde{X} \in H_yP$, $X_v \in V_yP$, and similarly write $Y^\# = \widetilde{Y} + Y_v$. It suffices to check the following four cases:

$$\Omega(\widetilde{X}, \widetilde{Y}), \quad \Omega(\widetilde{X}, Y_v), \quad \Omega(X_v, \widetilde{Y}), \quad \Omega(X_v, Y_v). \tag{6.32}$$

Without loss of generality, one can assume that $\widetilde{X}$ and $\widetilde{Y}$ are horizontal lifts of vector fields on M and that $\xi(X_v)$ and $\xi(Y_v)$ are constant $\mathfrak{g}$-valued functions on P. By (6.20) and (6.28), we have

$$\Omega(\widetilde{X},\widetilde{Y}) = (d\xi)(\widetilde{X},\widetilde{Y}), \quad \frac{1}{2}[\xi,\xi](\widetilde{X},\widetilde{Y}) = [\xi(\widetilde{X}),\xi(\widetilde{Y})] = 0, \tag{6.33}$$

so (6.24) holds in this case. Next, clearly

$$\Omega(\widetilde{X},Y_v) = 0, \quad [\xi(\widetilde{X}),\xi(Y_v)] = 0, \tag{6.34}$$

while

$$d\xi(\widetilde{X},Y_v) = \widetilde{X}\cdot\xi(Y_v) - Y_v\cdot\xi(\widetilde{X}) - \xi\big([\widetilde{X},Y_v]\big). \tag{6.35}$$

Now, having arranged that $\xi(Y_v)$ be a constant $\mathfrak{g}$-valued function on P, we have that $\widetilde{X}\cdot\xi(Y_v)=0$. Of course, $Y_v\cdot\xi(\widetilde{X})=0$. Also, $[\widetilde{X},Y_v] = -\mathcal{L}_{Y_v}\widetilde{X}$ is horizontal, by (6.10), so $\xi\big([\widetilde{X},Y_v]\big)=0$. This verifies (6.24) when both sides act on $(\widetilde{X},Y_v)$, and similarly we have (6.24) when both sides act on $(X_v,\widetilde{Y})$. We consider the final case. Clearly,

$$\Omega(X_v,Y_v) = 0, \tag{6.36}$$

while

$$d\xi(X_v,Y_v) = X_v\cdot\xi(Y_v) - Y_v\cdot\xi(X_v) - \xi\big([X_v,Y_v]\big) = -\xi\big([X_v,Y_v]\big) \tag{6.37}$$

and

$$\frac{1}{2}[\xi,\xi](X_v,Y_v) = [\xi(X_v),\xi(Y_v)] = \xi\big([X_v,Y_v]\big), \tag{6.38}$$

so (6.24) is verified in this last case, and Theorem 6.1 is proved.

We next obtain a form of the Bianchi identity that will play an important role in the next section. Compare with (1.40) and (2.13).

Proposition 6.2. *We have*

$$d\Omega = [\Omega,\xi]. \tag{6.39}$$

Here, if $\Omega = \sum \Omega_{jk}\, dx_j\wedge dx_k$ in local coordinates, we set

$$\begin{aligned}[\Omega,\xi] &= \sum_{j,k,\ell} [\Omega_{jk},\xi_\ell]\, dx_j\wedge dx_k\wedge dx_\ell \\ &= -\sum_{j,k,\ell} [\xi_\ell,\Omega_{jk}]\, dx_\ell\wedge dx_j\wedge dx_k = -[\xi,\Omega].\end{aligned} \tag{6.40}$$

To get (6.39), apply d to (6.24), obtaining (since $dd\xi=0$)

$$d\Omega = \frac{1}{2}[d\xi,\xi] - \frac{1}{2}[\xi,d\xi] = [d\xi,\xi], \tag{6.41}$$

which differs from $[\Omega, \xi]$ by $(1/2)[[\xi,\xi],\xi]$. We have

$$[[\xi,\xi],\xi] = \sum_{j,k,\ell} [[\xi_j,\xi_k],\xi_\ell]\ dx_j \wedge dx_k \wedge dx_\ell. \tag{6.42}$$

Now cyclic permutations of (j,k,ℓ) leave $dx_j \wedge dx_k \wedge dx_\ell$ invariant, so we can replace $[[\xi_j,\xi_k],\xi_\ell]$ in (6.42) by the average over cyclic permutations of (j,k,ℓ). However, Jacobi's identity for a Lie algebra is

$$[[\xi_j,\xi_k],\xi_\ell] + [[\xi_k,\xi_\ell],\xi_j] + [[\xi_\ell,\xi_j],\xi_k] = 0,$$

so $[[\xi,\xi],\xi] = 0$, and we have (6.39).

Exercises

1. Let $P \xrightarrow{p} M$ be a principal G-bundle with connection, where M is a Riemannian manifold. Pick an inner product on $\mathfrak{g}$. For $y \in P$, define an inner product on $T_yP = V_yP \oplus H_yP$ so that if $Z \in T_yP$ has decomposition $Z = Z_v + Z_h$, then

$$\|Z\|^2 = \|\xi(Z_v)\|^2 + \|Dp(y)Z_h\|^2.$$

Show that this is a G-invariant Riemannian metric on P.

2. Conversely, if $P \xrightarrow{p} M$ is a principal G-bundle, and if P has a G-invariant Riemannian metric, show that this determines a connection on P, by declaring that, for each $y \in P$, H_yP is the orthogonal complement of V_yP.
3. A choice of section σ of P over an open set $U \subset M$ produces an isomorphism

$$j_\sigma : C^\infty(U,E) \longrightarrow C^\infty(U,V). \tag{6.43}$$

If $\tilde{\sigma}$ is another section, there is a smooth function $g : U \to G$ such that

$$\tilde{\sigma}(x) = \sigma(x) \cdot g(x), \quad \forall\ x \in U. \tag{6.44}$$

Show that

$$j_{\tilde{\sigma}} \circ j_\sigma^{-1} v(x) = \pi\Big(g(x)\Big)^{-1} v(x). \tag{6.45}$$

4. According to (6.12), if $u \in C^\infty(U,E)$ and $v = j_\sigma u,\ \tilde{v} = j_{\tilde{\sigma}} u$, we have

$$(\nabla_X u) \circ \sigma = X \cdot v + \Gamma(X)v, \quad (\nabla_X u) \circ \tilde{\sigma} = X \cdot \tilde{v} + \widetilde{\Gamma}(X)\tilde{v}. \tag{6.46}$$

Show that

$$\widetilde{\Gamma}(X) = \pi\Big(g(x)\Big)^{-1}\Gamma(X)\pi\Big(g(x)\Big) + d\pi\Big(D\lambda_{g(x)}(g(x)) \circ Dg(x)X\Big), \tag{6.47}$$

where $Dg(x)X \in T_{g(x)}G,\ \lambda_g(h) = g^{-1}h,\ D\lambda_g(g) : T_gG \to T_eG \approx \mathfrak{g}$. Compare with (1.41). (*Hint*: Make use of (6.11), plus the identity $(d\pi)(Ad_{g^{-1}}A) = \pi(g)^{-1}d\pi(A)\pi(g)$, $A \in \mathfrak{g}$.)

5. Show that, for X, Y vector fields on M, $\Omega(\widetilde{X},\widetilde{Y})$ satisfies

$$\Omega(\widetilde{X},\widetilde{Y})(y \cdot g) = \mathrm{Ad}(g)^{-1}\Omega(\widetilde{X},\widetilde{Y}). \tag{6.48}$$

Deduce that setting

$$\Omega^b(X,Y) = \Omega(\widetilde{X},\widetilde{Y}) \tag{6.49}$$

defines Ω^b as a section of $\Lambda^2 T^* \otimes (\mathrm{Ad}\ P)$, where Ad P is the vector bundle

$$\mathrm{Ad}\ P = P \times_{\mathrm{Ad}} \mathfrak{g}. \tag{6.50}$$

6. If ξ_0 and ξ_1 are connection 1-forms on $P \to M$, show that $t\xi_1 + (1-t)\xi_0$ is also, for any $t \in \mathbb{R}$. (*Hint*: If P_0 and P_1 are projections, show that $tP_1 + (1-t)P_0$ is also a projection, *provided* that P_0 and P_1 have the same range.)
7. Let ξ_0 and ξ_1 be two connection 1-forms for $P \to M$, and let ∇ be an arbitrary third connection on P. Consider

$$\alpha = \xi_1 - \xi_0. \tag{6.51}$$

If X is a vector field on M and $\widetilde{X}$ the horizontal lift determined by ∇, show that

$$\alpha^b(X) = \alpha(\widetilde{X}) \tag{6.52}$$

defines α^b as an element of $C^\infty(M, \Lambda^1 T^* \otimes \mathrm{Ad}\ P)$. Show that α^b is independent of the choice of ∇.
8. In the setting of Exercise 7, if Ω_j are the curvatures of the connection 1-forms ξ_j, show that

$$\Omega_1 - \Omega_0 = d\alpha + [\alpha, \xi_0] + \frac{1}{2}[\alpha, \alpha]. \tag{6.53}$$

Compare with (2.17) and (3.62). If $d^\nabla \alpha^b$ is the (Ad P)-valued 2-form defined as in §2, via the connection ξ_0, relate $d^\nabla \alpha^b$ to $d\alpha + [\alpha, \xi_0]$.

7. The Chern-Weil construction

Let $P \to M$ be a principal G-bundle, endowed with a connection, as in §6. Let Ω be its curvature form, a $\mathfrak{g}$-valued 2-form on P; equivalently, there is the Ad P-valued 2-form Ω^b on M. The Chern-Weil construction gives closed differential forms on M, whose cohomology classes are independent of the choice of connection on P. These "characteristic classes" are described as follows.

A function $f : \mathfrak{g} \to \mathbb{C}$ is called "invariant" if

$$f(\mathrm{Ad}(g)X) = f(X), \quad X \in \mathfrak{g}, \quad g \in G. \tag{7.1}$$

Denote by $\mathcal{I}_k$ the set of polynomials $p : \mathfrak{g} \to \mathbb{C}$ which are invariant and homogeneous of degree k. If $p \in \mathcal{I}_k$, there is associated an Ad-invariant k-linear function P on $\mathfrak{g}$, called the *polarization* of p, given by

$$P(Y_1, \dots, Y_k) = \frac{1}{k!} \frac{\partial^k}{\partial t_1 \cdots \partial t_k} p(t_1 Y_1 + \cdots + t_k Y_k), \tag{7.2}$$

such that $p(X) = P(X, \dots, X)$. Into the entries of P we can plug copies of Ω, or of Ω^b, to get $2k$-forms

$$p(\Omega) = P(\Omega, \dots, \Omega) \in \Lambda^{2k} P \tag{7.3}$$

and

$$p(\Omega^b) = P(\Omega^b, \dots, \Omega^b) \in \Lambda^{2k} M. \tag{7.4}$$

Note that if $\pi : P \to M$ is the projection, then

$$p(\Omega) = \pi^* p(\Omega^b); \tag{7.5}$$

we say $p(\Omega)$, a form on P, is "basic," namely, the pull-back of a form on M. The following two propositions summarize the major basic results about these forms.

Proposition 7.1. *For any connection ∇ on $P \to M$, $p \in \mathcal{I}_k$, the forms $p(\Omega)$ and $p(\Omega^b)$ are closed. Hence $p(\Omega^b)$ represents a deRham cohomology class*

$$[p(\Omega^b)] \in \mathcal{H}^{2k}(M, \mathbb{C}). \tag{7.6}$$

If $q \in \mathcal{I}_j$, then $pq \in \mathcal{I}_{j+k}$ and $(pq)(\Omega) = p(\Omega) \wedge q(\Omega)$. Furthermore, if $f : N \to M$ is smooth and ∇_f the connection on f^*P pulled back from ∇ on P, which has curvature $\Omega_f = f^*\Omega$, then

$$p(\Omega_f^b) = f^* p(\Omega^b). \tag{7.7}$$

Proposition 7.2. *The cohomology class (7.6) is independent of the connection on P, so it depends only on the bundle.*

The map $\mathcal{I}_* \to \mathcal{H}^{2*}(M, \mathbb{C})$ is called the *Chern-Weil homomorphism.* We first prove that $d\, p(\Omega) = 0$ on P, the rest of Proposition 7.1 being fairly straightforward. If we differentiate with respect to t at $t = 0$ the identity

$$P\big(\mathrm{Ad}(\mathrm{Exp}\, tY)X, \dots, \mathrm{Ad}(\mathrm{Exp}\, tY)X\big) = p(X), \tag{7.8}$$

we get

$$\sum P(X, \dots, [Y, X], \dots, X) = 0. \tag{7.9}$$

Into this we can substitute the curvature form Ω for X and the connection form ξ for Y, to get

$$\sum P(\Omega, \dots, [\xi, \Omega], \dots, \Omega) = 0. \tag{7.10}$$

Now the Bianchi identity $d\Omega = -[\xi, \Omega]$ obtained in (6.8) shows that (7.10) is equivalent to $d\, p(\Omega) = 0$ on P. Since $\pi^* : \Lambda^j M \to \Lambda^j P$ is injective and (7.5) holds, we also have $d\, p(\Omega^b) = 0$ on M, and Proposition 7.1 is proved.

The proof of Proposition 7.2 is conveniently established via the following result, which also has further uses.

Lemma 7.3. *Let ξ_0 and ξ_1 be any $\mathfrak{g}$-valued 1-forms on P (or any manifold). Set $\alpha = \xi_1 - \xi_0$, $\xi_t = \xi_0 + t\alpha$, and $\Omega_t = d\xi_t + (1/2)[\xi_t, \xi_t]$. Given*

$p \in \mathcal{I}_k$, we have

$$p(\Omega_1) - p(\Omega_0) = k\ d\Big[\int_0^1 P(\alpha, \Omega_t, \dots, \Omega_t)\ dt\Big]. \tag{7.11}$$

Proof. Since $(d/dt)\Omega_t = d\alpha + [\xi_t, \alpha]$, we have

$$\frac{d}{dt}\, p(\Omega_t) = k\, P(d\alpha + [\xi_t, \alpha], \Omega_t, \dots, \Omega_t). \tag{7.12}$$

It suffices to prove that the right side of (7.12) equals $k\ dP(\alpha, \Omega_t, \dots, \Omega_t)$. This follows by the "Bianchi" identity $d\Omega_t = -[\xi_t, \Omega_t]$ and the same sort of arguments used in the proof of Proposition 7.1. Instead of (7.8), one starts with

$$P\big(\mathrm{Ad}(\mathrm{Exp}\ tY)Z, \mathrm{Ad}(\mathrm{Exp}\ tY)X, \dots, \mathrm{Ad}(\mathrm{Exp}\ tY)X\big) = P(Z, X, \dots, X).$$

To apply this to Proposition 7.2, let ξ_0 and ξ_1 be the connection forms associated to two connections on $P \to M$, so Ω_0 and Ω_1 are their curvature forms. Note that each ξ_t defines a connection form on P, with curvature form Ω_t. Furthermore, $\alpha = \xi_1 - \xi_0$, acting on $X^\# \in T_yP$, depends only on $\pi_* X^\# \in T_xM$ and gives rise to an Ad P-valued 1-form α^b on M. Thus the right side of (7.11) is the pull-back via π^* of the $(2k-1)$-form

$$k\ d\Big[\int_0^1 P(\alpha^b, \Omega_t^b, \dots, \Omega_t^b)\ dt\Big] \tag{7.13}$$

on M, which yields Proposition 7.2.

We can also apply Lemma 7.3 to $\xi_1 = \xi$, a connection 1-form, and $\xi_0 = 0$. Then $\xi_t = t\xi$; denote $d\xi_t + (1/2)[\xi_t, \xi_t]$ by Φ_t. We have the $(2k-1)$-form on P called the *transgressed form*:

$$Tp(\Omega) = k \int_0^1 P(\xi, \Phi_t, \dots, \Phi_t)\ dt, \tag{7.14}$$

with

$$\Phi_t = t\ d\xi + \frac{1}{2}t^2[\xi, \xi]. \tag{7.15}$$

Then Lemma 7.3 gives

$$P(\Omega) = d\ Tp(\Omega); \tag{7.16}$$

that is, $p(\Omega)$ is an exact form on P, not merely a closed form. On the other hand, as opposed to $p(\Omega)$ itself, $Tp(\Omega)$ is not necessarily a basic form, that is, the pull-back of a form on M. In fact, $p(\Omega^b)$ is not necessarily an exact form on M; typically it determines a nontrivial cohomology class on M. Transgressed forms play an important role in Chern-Weil theory.

The Levi-Civita connection on an oriented Riemannian manifold of dimension 2 can be equated with a connection on the associated principal

S^1-bundle. If we identify S^1 with the unit circle in $\mathbb{C}$, its Lie algebra is naturally identified with $i\mathbb{R}$, and this identification provides an element of $\mathcal{I}_1$, unique up to a constant multiple. This is of course a constant times the product of the Gauss curvature and the volume form, and the invariance of Proposition 7.2 recovers the independence (5.26) of the integrated curvature from the metric used on a Riemannian manifold of dimension 2. More generally, for any complex line bundle L over M, a manifold of any dimension, L can be associated to a principal S^1-bundle, and the Chern-Weil construction produces the class $[\Omega] \in \mathcal{H}^2(M, \mathbb{C})$. The class $c_1(L) = -(1/2\pi i)[\Omega] \in \mathcal{H}^2(M, \mathbb{C})$ is called the first Chern class of the line bundle L. In this case, the connection 1-form on P can be identified with an ordinary (complex-valued) 1-form, and it is precisely the transgressed form (7.14).

Note that if $\dim M = 2$, then (5.39) says that

$$c_1(L)[M] = \text{ Index } X,$$

for any nonvanishing section X of L over $M \setminus \{p_1, \dots, p_K\}$.

For general G, there may be no nontrivial elements of $\mathcal{I}_1$. In fact, if $p : \mathfrak{g} \to \mathbb{R}$ is a nonzero linear form, $V = \ker p$ is a linear subspace of $\mathfrak{g}$ of codimension 1, which is Ad G-invariant if $p \in \mathcal{I}_1$. This means V is an ideal: $[V, \mathfrak{g}] \subset V$. Thus there are no nontrivial elements of $\mathcal{I}_1$ unless $\mathfrak{g}$ has an ideal of codimension 1. In particular, if $\mathfrak{g}$ is semisimple, $\mathcal{I}_1 = 0$.

When G is compact, there are always nontrivial elements of $\mathcal{I}_2$, namely, Ad-invariant quadratic forms on $\mathfrak{g}$. In fact, any bi-invariant metric tensor on G gives a positive-definite element of $\mathcal{I}_2$. Applying the Chern-Weil construction in this case then gives cohomology classes in $\mathcal{H}^4(M, \mathbb{C})$.

One way of obtaining elements of $\mathcal{I}_k$ is the following. Let π be a representation of G on a vector space V_π, and set

$$p_{\pi k}(X) = \text{ Tr } \Lambda^k d\pi(X), \quad X \in \mathfrak{g}, \tag{7.17}$$

where $d\pi(X)$ denotes the representation of $\mathfrak{g}$ on V_π. In connection with this, note that

$$\det\big(\lambda I + d\pi(X)\big) = \sum_{j=0}^{M} \lambda^{M-j} \text{ Tr } \Lambda^j d\pi(X), \quad M = \dim V_\pi. \tag{7.18}$$

If $P \to M$ is a principal $\mathrm{U}(n)$-bundle or $\mathrm{Gl}(n, \mathbb{C})$-bundle, and π the standard representation on $\mathbb{C}^n$, then consider

$$\det\Big(\lambda - \frac{1}{2\pi i}\Omega\Big) = \sum_{k=0}^{n} c_k(\Omega)\lambda^{n-k}. \tag{7.19}$$

The classes $[c_k(\Omega)] \in \mathcal{H}^{2k}(M, \mathbb{C})$ are the Chern classes of P. If $E \to M$ is the associated vector bundle, arising via the standard representation π, we also call this the kth Chern class of E:

$$c_k(E) = [c_k(\Omega)] \in \mathcal{H}^{2k}(M, \mathbb{C}). \tag{7.20}$$

The object

$$c(E) = \sum c_k(E) \in \bigoplus_{k=0}^{n} \mathcal{H}^{2k}(M, \mathbb{C}) \tag{7.21}$$

is called the total Chern class of such a vector bundle.

If $P \to M$ is a principal $\mathrm{O}(n)$-bundle, and π the standard representation on $\mathbb{R}^n$, then consider

$$\det\Big(\lambda - \frac{1}{2\pi}\Omega\Big) = \sum_{k=0}^{n} d_k(\Omega)\lambda^{n-k}. \tag{7.22}$$

The polynomials $d_k(\Omega)$ vanish for k odd, since $\Omega^t = -\Omega$, and one obtains Pontrjagin classes:

$$p_k(\Omega) = d_{2k}(\Omega) \in \mathcal{H}^{4k}(M, \mathbb{R}). \tag{7.23}$$

If $F \to M$ is the associated vector bundle, arising from the standard representation π, then $p_k(F)$ is defined to be (7.23).

Exercises

1. If E and F are complex vector bundles over M, we can form $E \oplus F \to M$. Show that
$$c(E \oplus F) = c(E) \wedge c(F),$$
where $c(E)$ is the total Chern class given by (7.21), that is,
$$c(E) = \det\Big(I - \frac{1}{2\pi i}\Omega\Big) \in \mathcal{H}^{\text{even}}(M, \mathbb{C}),$$
for a curvature 2-form arising from a connection on M.
2. Define the Chern character of a complex vector bundle $E \to M$ as the cohomology class $\mathrm{Ch}(E) \in \mathcal{H}^{\text{even}}(M, \mathbb{C})$ of
$$\mathrm{Ch}(\Omega) = \mathrm{Tr}\ e^{-\Omega/2\pi i},$$
writing $\mathrm{Tr}\ e^{-\Omega/2\pi i} \in \bigoplus_{k\ge 0} \Lambda^{2k} P$ via the power-series expansion of the exponential function. Show that
$$\mathrm{Ch}(E \oplus F) = \mathrm{Ch}(E) + \mathrm{Ch}(F),$$
$$\mathrm{Ch}(E \otimes F) = \mathrm{Ch}(E) \wedge \mathrm{Ch}(F)$$
in $\mathcal{H}^{\text{even}}(M, \mathbb{C})$.
3. If $F \to M$ is a real vector bundle and $E = F \otimes \mathbb{C}$ is its complexification, show that
$$p_j(F) = (-1)^j c_{2j}(E).$$
4. Using $\mathrm{so}(4) \approx \mathrm{so}(3) \oplus \mathrm{so}(3)$, construct two different characteristic classes in $\mathcal{H}^4(M, \mathbb{C})$, when M is a compact, oriented, four-dimensional manifold.

8. The Chern-Gauss-Bonnet theorem

Our goal in this section is to generalize the Gauss-Bonnet formula (5.1), producing a characteristic class derived from the curvature tensor Ω of a Riemannian metric on a compact, oriented manifold M, say $e(\Omega) \in \Lambda^n(M)$, such that

$$\int_M e(\Omega) = \chi(M), \tag{8.1}$$

the right side being the Euler characteristic of M.

A clue to obtaining $e(\Omega)$ comes from the higher-dimensional generalization of the index formula (5.47), namely,

$$\operatorname{Index}(X) = \chi(M), \tag{8.2}$$

valid for any vector field X on M with isolated critical points. The relation between these two when dim $M = 2$ is noted at the end of §5. It arises from the relation between Index(X) and the degree of the Gauss map.

Indeed, let M be a compact, n-dimensional submanifold of $\mathbb{R}^{n+k}$, X a (tangent) vector field on M with a finite number of critical points, and $\overline{\mathcal{T}}$ a small tubular neighborhood of M. By Corollary 20.5 of Chapter 1, we know that if $N : \partial\mathcal{T} \to S^{n+k-1}$ denotes the Gauss map on $\partial\mathcal{T}$, formed by the outward-pointing normals, then

$$\operatorname{Index}(X) = \operatorname{Deg}(N). \tag{8.3}$$

As noted at the end of §5 of this chapter, if M is a surface in $\mathbb{R}^3$, with Gauss map N_M, then $\operatorname{Deg}(N_M) = (1/4\pi)\int_M K\,dV$, where K is the Gauss curvature of M, with its induced metric. If $\mathcal{T}$ is a small tubular neighborhood of M in this case, then $\partial\mathcal{T}$ is diffeomorphic to two oppositely oriented copies of M, with approximately the same metric tensor. The outer component of $\partial\mathcal{T}$ has Gauss map approximately equal to N_M, and the inner component has Gauss map approximately equal to $-N_M$. From this we see that (8.2) and (8.3) imply (8.1) with $e(\Omega) = (1/2\pi)K\,dV$ in this case.

We make a further comment on the relation between (8.2) and (8.3). Note that the right side of (8.3) is independent of the choice of X. Thus, any two vector fields on M with only isolated critical points, have the same index. Suppose M has a triangulation τ into n-simplices. There is a construction of a vector field X_τ, illustrated in Fig. 5.3 for $n = 2$, with the property that X_τ has a critical point at each vertex, of index $+1$, and a critical point in the middle of each j-simplex in τ, of index $(-1)^j$, so that

$$\operatorname{Index}(X_\tau) = \sum_{j=0}^{n} (-1)^j \nu_j(M), \tag{8.4}$$

where $\nu_j(M)$ is the number of j-simplices in the triangulation τ of M. We leave the construction of X_τ in higher dimensions as an exercise.

A proof that any smooth, paracompact manifold M is triangulable is given in [Wh]. There it is shown that if you imbed M smoothly in $\mathbb{R}^N$, produce a fine triangulation of $\mathbb{R}^N$, and then perhaps jiggle the imbedding a bit, the intersection provides a triangulation of M.

Now, in view of the invariance of $\mathrm{Index}(X)$, it follows that the right side of (8.4) is independent of the triangulation of M. Also, if M has a more general cell decomposition, we can form the sum on the right side of (8.4), where ν_j stands for the number of j-dimensional cells in M. Each cell can be divided into simplices in such a way that a triangulation is obtained, and the sum on the right side of (8.4) is unchanged under such a refinement. This alternating sum is one definition of the Euler characteristic, but as we have used another definition in §§8 and 9 of Chapter 5, namely

$$\chi(M) = \sum_{j=0}^{n} (-1)^j \dim \mathcal{H}^j(M), \tag{8.5}$$

we will temporarily denote the right side of (8.4) by $\chi_c(M)$.

Now we tackle the question of representing (8.3) as an integrated curvature, to produce (8.1). We begin with the case when M is a compact hypersurface in $\mathbb{R}^{n+1}$. In that case we have, by (4.66),

$$\mathrm{Deg}(N) = \frac{2}{A_n} \int_M \bigl(\det A_N\bigr)\, dV, \quad \text{for } n \text{ even}, \tag{8.6}$$

where A_n is the area of S^n and $A_N : T_pM \to T_pM$ is the Weingarten map. The factor 2 arises because $\partial\mathcal{T}$ consists of two copies of M. We can express $\det A_N$ directly in terms of the Riemann curvature tensor $R_{jk\ell m}$ of M, using Gauss' *Theorema Egregium.*

In fact, with respect to an oriented orthonormal basis $\{e_j\}$ of T_pM, the matrix of A_N has entries $A_{jk} = \widetilde{II}(e_j, e_k)$, and by (4.14),

$$R_{jk\ell m} = \langle R(e_\ell, e_m)e_k, e_j\rangle = \det \begin{pmatrix} A_{mk} & A_{mj} \\ A_{\ell k} & A_{\ell j} \end{pmatrix}. \tag{8.7}$$

In other words, the curvature tensor captures the action of $\Lambda^2 A_N$ on $\Lambda^2 T_pM$. If $n = 2k$ is even, we can then express $\det A_N$ as a polynomial in the components $R_{jk\ell m}$, using

$$\begin{aligned} (\det A_N) e_1 \wedge \cdots \wedge e_n &= \bigl(\Lambda^n A_N\bigr)(e_1 \wedge \cdots \wedge e_n) \\ &= (Ae_1 \wedge Ae_2) \wedge \cdots \wedge (Ae_{n-1} \wedge Ae_n). \end{aligned} \tag{8.8}$$

Now, by (8.7),

$$Ae_j \wedge Ae_k = \frac{1}{2} \sum R_{\ell m jk}\, e_\ell \wedge e_m. \tag{8.9}$$

Replacing $(1, \dots, n)$ in (8.8) with all its permutations and summing, we obtain

$$\det A_N = \frac{1}{2^{n/2} n!} \sum_{j,k} (\operatorname{sgn} j)(\operatorname{sgn} k) R_{j_1 j_2 k_1 k_2} \cdots R_{j_{n-1} j_n k_{n-1} k_n}, \tag{8.10}$$

where $j = (j_1, \dots, j_n)$ stands for a permutation of $(1, \dots, n)$. The fact that the quantity (8.10), integrated over M, is equal to $(A_n/2)\chi(M)$ when M is a hypersurface in $\mathbb{R}^{n+1}$ was first established by E. Hopf, as a consequence of his result (8.2). The content of the generalized Gauss-Bonnet formula is that for any compact Riemannian manifold M of dimension $n = 2k$, integrating the right side of (8.10) over M gives $(A_n/2)\chi(M)$.

One key point in establishing the general case is to perceive the right side of (8.10) as arising via the Chern-Weil construction from an invariant polynomial on the Lie algebra $\mathfrak{g} = \mathrm{so}(n)$, to produce a characteristic class. Now the curvature 2-form can in this case be considered a section of $\Lambda^2 T^* \otimes \Lambda^2 T^*$, reflecting the natural linear isomorphism $\mathfrak{g} \approx \Lambda^2 T^*$. Furthermore, $\Lambda^* T^* \otimes \Lambda^* T^*$ has a product, satisfying

$$(\alpha_1 \otimes \beta_1) \wedge (\alpha_2 \otimes \beta_2) = (\alpha_1 \wedge \alpha_2) \otimes (\beta_1 \wedge \beta_2). \tag{8.11}$$

If we set

$$\Omega = \frac{1}{4} \sum R_{jk\ell m} (e_j \wedge e_k) \otimes (e_\ell \wedge e_m), \tag{8.12}$$

then we form the k-fold product, $k = n/2$, obtaining

$$\Omega \wedge \cdots \wedge \Omega = 2^{-n} \sum_{j,k} (\operatorname{sgn} j)(\operatorname{sgn} k) R_{j_1 j_2 k_1 k_2} \cdots R_{j_{n-1} j_n k_{n-1} k_n} (\omega \otimes \omega), \tag{8.13}$$

with $\omega = e_1 \wedge \cdots \wedge e_n$. Thus, the right side of (8.10), multiplied by $\omega \otimes \omega$, is equal to $2^{n/2}/n!$ times the right side of (8.13). (Observe the distinction between the product (8.11) and the product on $\mathrm{End}(E) \otimes \Lambda^* T$, used in (7.19) and (7.22), which assigns a different meaning to $\Omega \wedge \cdots \wedge \Omega$.)

Now the Chern-Weil construction produces (8.13), with $\omega \otimes \omega$ replaced by ω, if we use the *Pfaffian*

$$\mathrm{Pf} : \mathrm{so}(n) \longrightarrow \mathbb{R}, \quad n = 2k, \tag{8.14}$$

defined as follows. Let $\xi : \mathrm{so}(n) \to \Lambda^2 \mathbb{R}^n$ be the isomorphism

$$\xi(X) = \frac{1}{2} \sum X_{jk} e_j \wedge e_k, \quad X = (X_{jk}) \in \mathrm{so}(n). \tag{8.15}$$

Then, if $n = 2k$, take a product of k factors of $\xi(X)$ to obtain a multiple of $\omega = e_1 \wedge \cdots \wedge e_n$. Then $\mathrm{Pf}\,(X)$ is uniquely defined by

$$\xi(X) \wedge \cdots \wedge \xi(X) = k!\ \mathrm{Pf}\,(X) \omega. \tag{8.16}$$

Note that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear, then $T^* \xi(X) = \xi(T^t X T)$, so

$$\mathrm{Pf}\,(T^t X T) = (\det T) \mathrm{Pf}\,(X). \tag{8.17}$$

Now any $X \in \mathrm{so}(n)$ can be written as $X = T^t Y T$, where $T \in \mathrm{SO}(n)$, that is, T is an orthogonal matrix of determinant 1, and Y is a sum of 2×2 skew-symmetric blocks, of the form

$$Y_\nu = \begin{pmatrix} 0 & \lambda_\nu \\ -\lambda_\nu & 0 \end{pmatrix}, \quad \lambda_\nu \in \mathbb{R}. \tag{8.18}$$

Thus $\xi(Y) = \lambda_1 e_1 \wedge e_2 + \cdots + \lambda_k e_{n-1} \wedge e_n$, so

$$\mathrm{Pf}\,(Y) = \lambda_1 \cdots \lambda_k. \tag{8.19}$$

Note that $\det Y = (\lambda_1 \cdots \lambda_k)^2$. Hence, by (8.17), we have

$$\mathrm{Pf}\,(X)^2 = \det X, \tag{8.20}$$

when X is a real, skew-symmetric, $n \times n$ matrix, $n = 2k$. When (8.17) is specialized to $T \in \mathrm{SO}(n)$, it implies that Pf is an invariant polynomial, homogeneous of degree k (i.e., $\mathrm{Pf} \in \mathcal{I}_k,\ k = n/2$).

Now, with Ω in (8.12) regarded as a $\mathfrak{g}$-valued 2-form, we have the left side of (8.13) equal to $(1/k!)\mathrm{Pf}\,(\Omega)$. Thus we are on the way toward establishing the generalized Gauss-Bonnet theorem, in the following formulation.

Theorem 8.1. *If M is a compact, oriented Riemannian manifold of dimension $n = 2k$, then*

$$\chi(M) = (2\pi)^{-k} \int\limits_M Pf(\Omega). \tag{8.21}$$

The factor $(2\pi)^{-k}$ arises as follows. From (8.10) and (8.13), it follows that when M is a compact hypersurface in $\mathbb{R}^{n+1}$, the right side of (8.6) is equal to $C_k \int_M \mathrm{Pf}\,(\Omega)$, with

$$C_k = \frac{2^{k+1}}{A_n}\,\frac{k!}{n!}. \tag{8.22}$$

Now the area of the unit sphere is given by

$$A_{2k} = \frac{2\pi^{k+1/2}}{\Gamma(k+\frac{1}{2})} = \frac{2\pi^k}{\left(k-\frac{1}{2}\right)\cdots\left(\frac{1}{2}\right)},$$

as is shown in Appendix A of Chapter 3; substituting this into (8.22) gives $C_k = (2\pi)^{-k}$.

We give a proof of this which extends the proof of (5.24), in which handles are added to a surface. To effect this parallel, we consider how the two sides of (8.21) change when M is altered by a certain type of *surgery*, which we will define in the next paragraph.

First, we mention another ingredient in the proof of Theorem 8.1. Namely, the right side of (8.21) is independent of the choice of metric on M. Since different metrics produce different $\mathrm{SO}(2k)$ frame bundles, this assertion is not a simple consequence of Proposition 7.2. We will postpone the proof of this invariance until near the end of this section.

We now describe the "surgeries" alluded to above. To perform surgery on M_0, a manifold of dimension n, excise a set H_0 diffeomorphic to $S^{\ell-1}\times B^m$, with $m+\ell-1=n$, where $B^m=\{x\in\mathbb{R}^m : |x|<1\}$, obtaining a manifold with boundary X, ∂X being diffeomorphic to $S^{\ell-1}\times S^{m-1}$. Then attach to X a copy of $B^\ell\times S^{m-1}$, sewing them together along their boundaries, both diffeomorphic to $S^{\ell-1}\times S^{m-1}$, to obtain M_1. Symbolically, we write

$$M_0 = X\#H_0,\quad M_1 = X\#H_1. \tag{8.23}$$

We say M_1 is obtained from M_0 by a surgery of type (ℓ,m).

We compare the way each side of (8.21) changes when M changes from M_0 to M_1. We also look at how $\chi_c(M)$, defined to be the right side of (8.4), changes. In fact, this definition easily yields

$$\chi(X\#H_1) = \chi(X\#H_0) - \chi(H_0) + \chi(H_1). \tag{8.24}$$

For notational simplicity, we have dropped the "c" subscript. It is more convenient to produce an identity involving only manifolds without boundary, so note that

$$\begin{aligned}\chi(H_0\#H_0) &= 2\chi(H_0) - \chi(\partial H_0),\\ \chi(H_1\#H_1) &= 2\chi(H_1) - \chi(\partial H_1),\end{aligned} \tag{8.25}$$

and, since $\partial H_0 = \partial H_1$, we have

$$\chi(H_1) - \chi(H_0) = \frac{1}{2}\chi(H_1\#H_1) - \frac{1}{2}\chi(H_0\#H_0), \tag{8.26}$$

hence

$$\chi(M_1) = \chi(M_0) + \frac{1}{2}\chi(H_1\#H_1) - \frac{1}{2}\chi(H_0\#H_0). \tag{8.27}$$

Note that $H_0\#H_0 = S^{\ell-1}\times S^m$, $H_1\#H_1 = S^\ell\times S^{m-1}$. To compute the Euler characteristic of these two spaces, we can use multiplicativity of χ. Note that products of cells in Y_1 and Y_2 give cells in $Y_1\times Y_2$, and

$$\nu_j(Y_1\times Y_2) = \sum_{i+k=j} \nu_i(Y_1)\nu_k(Y_2); \tag{8.28}$$

then from (8.4) it follows that

$$\chi(Y_1\times Y_2) = \sum_{j\ge 0}(-1)^j \sum_{i+k=j}\nu_i(Y_1)\nu_k(Y_2) = \chi(Y_1)\chi(Y_2). \tag{8.29}$$

Using the fairly elementary result that

$$\begin{aligned}\chi(S^j) = 2 &\quad \text{if } j \text{ is even},\\ 0 &\quad \text{if } j \text{ is odd},\end{aligned} \tag{8.30}$$

we have $\chi(H_0\#H_0) - \chi(H_1\#H_1)$ equal to 4 if ℓ is odd and m even, -4 if ℓ is even and m odd, and 0 if ℓ and m have the same parity (which does not arise if dim M is even).

The change in $\chi_c(M)$ just derived in fact coincides with the change in $\chi(M)$, defined by (8.5). This follows from results on deRham cohomology obtained in Chapter 5. In fact (B.8) of Chapter 5 implies (8.24), from which (8.25)–(8.27) follow; (8.52) of Chapter 5 implies (8.29) when Y_j are smooth, compact manifolds; and (8.56) and (8.57) of Chapter 5 imply (8.30).

Thus, for $e(M) = \int_M e(\Omega)$ to change the same way as $\chi(M)$ under a surgery, we need the following properties in addition to "functoriality." We need

$$e(S^j \times S^k) = \begin{matrix} 0 & \text{if } j \text{ or } k \text{ is odd,} \\ 4 & \text{if } j \text{ and } k \text{ are even.} \end{matrix} \tag{8.31}$$

If $e(\Omega)$ is locally defined, we have, upon giving X, H_0, and H_1 coherent orientations,

$$\int\limits_{M_1} e(\Omega) = \int\limits_{M_0} e(\Omega) - \int\limits_{H_0} e(\Omega) + \int\limits_{H_1} e(\Omega), \tag{8.32}$$

parallel to (8.24). Place metrics on M_j that are product metrics on $(-\varepsilon, \varepsilon) \times S^{\ell-1} \times S^{m-1}$ on a small neighborhood of ∂X. If we place a metric on $H_j\#H_j$ which is symmetric with respect to the natural involution, we will have

$$\int\limits_{H_j} e(\Omega) = \frac{1}{2} \int\limits_{H_j\#H_j} e(\Omega), \tag{8.33}$$

provided $e(\Omega)$ has the following property. Given an oriented Riemannian manifold Y, let $Y^{\#}$ be the same manifold with orientation reversed, and let the associated curvature forms be denoted by Ω_Y and $\Omega_{Y^{\#}}$. We require

$$e(\Omega_Y) = -e(\Omega_{Y^{\#}}). \tag{8.34}$$

Now $e(\Omega) = \mathrm{Pf}\,(\Omega/2\pi)$ certainly satisfies (8.34), in view of the dependence on orientation built into (8.16). To see that (8.31) holds in this case, we need only note that $S^\ell \times S^k$ can be smoothly imbedded as a hypersurface in $\mathbb{R}^{\ell+k+1}$. This can be done via imbedding $S^\ell \times I \times B^k$ into $\mathbb{R}^{\ell+k+1}$ and taking its boundary (and smoothing it out). In that case, since $\mathrm{Pf}(\Omega/2\pi)$ is a characteristic class whose integral is independent of the choice of metric, we can use the metric induced from the imbedding. We now have (8.31)–(8.33). Furthermore, for such a hypersurface $M = H_j\#H_j$, we know that the right side of (8.21) is equal to $\chi_c(H_j\#H_j)$, by the argument preceding the statement of Theorem 8.1, and since (8.29) and (8.30) are both valid for both χ and χ_c, we also have this quantity equal to $\chi(H_j\#H_j)$.

It follows that (8.21) holds for any M obtainable from S^n by a finite number of surgeries. With one extra wrinkle we can establish (8.21) for all compact, oriented M. The idea for using this technique is one the author learned from J. Cheeger, who uses a somewhat more sophisticated variant in work on analytic torsion [Ch].

Assume M is connected. Give $M \times \mathbb{R}$ the product Riemannian metric, fix a point $p \in M$, and, with $q = (p, 0) \in M \times \mathbb{R}$, consider on $M \times \mathbb{R}$ the function $f_0(x, t) = \operatorname{dist}((x, t), q)^2$. For R sufficiently large, $f_0^{-1}(R)$ is diffeomorphic to two copies of M, under the map $(x, t) \mapsto x$. For $r > 0$ sufficiently small, $f_0^{-1}(r)$ is diffeomorphic to the sphere S^n.

Our argument will use some basic results of Morse theory. A Morse function $f : Z \to \mathbb{R}$ is a smooth function on a manifold Z all of whose critical points are nondegenerate, that is, if $\nabla f(z) = 0$, then $D^2 f(z)$ is an invertible $\nu \times \nu$ matrix, $\nu = \dim Z$. One also assumes that f takes different values at distinct critical points and that $f^{-1}(K)$ is compact for every compact $K \subset \mathbb{R}$. Now the function f_0 above may not be a Morse function on $Z = M \times \mathbb{R}$, but there will exist a smooth perturbation f of f_0 which is a Morse function. A proof is given in Proposition 4.3 of Appendix B. The new f will share with f_0 the property that $f^{-1}(r)$ is diffeomorphic to S^n and $f^{-1}(R)$ is diffeomorphic to two copies of M. Note that an orientation on M induces an orientation on $M \times \mathbb{R}$, and hence an orientation on any level set $f^{-1}(c)$ that contains no critical points. In particular, $f^{-1}(R)$ is a union of two copies of M with opposite orientations. The following is a basic tool in Morse theory.

Theorem 8.2. *If $c_1 < c_2$ are regular values of a Morse function $f : Z \to \mathbb{R}$ and there is exactly one critical point z_0, with $c_1 < f(z_0) < c_2$, then $M_2 = f^{-1}(c_2)$ is obtained from $M_1 = f^{-1}(c_1)$ by a surgery. In fact, if $D^2 f(z_0)$ has signature (ℓ, m), M_2 is obtained from M_1 by a surgery of type (m, ℓ).*

This is a consequence of the following result, known as the Morse lemma.

Proposition 8.3. *Let f have a nondegenerate critical point at $p \in Z$. Then there is a coordinate system $(x_1, \dots, x_n)$ centered at p in which*

$$f(x) = f(p) + x_1^2 + \cdots + x_\ell^2 - x_{\ell+1}^2 - \cdots - x_{\ell+m}^2 \tag{8.35}$$

near the origin, where $\ell + m = \nu = \dim Z$.

Proof. Suppose that in some coordinate system $D^2 f(p)$ is given by a nondegenerate, symmetric, $\nu \times \nu$ matrix A. It will suffice to produce a coordinate system in which

$$f(x) = f(p) + \frac{1}{2} A x \cdot x, \tag{8.36}$$

near the origin, since going from here to (8.35) is a simple exercise in linear algebra. We will arrange (8.36) by an argument, due to Palais, similar to the proof of Darboux' theorem in Chapter 1, §14.

Begin with any coordinate system centered at p. Let

$$\omega_1 = df, \quad \omega_0 = dg, \quad \text{where } g(x) = \frac{1}{2} Ax \cdot x, \tag{8.37}$$

with $A = D^2 f(0)$ in this coordinate system. Set $\omega_t = t\omega_1 + (1-t)\omega_0$, which vanishes at p for each $t \in [0,1]$. The nondegeneracy hypothesis on A implies that the components of each ω_t have linearly independent gradients at p; hence there exists a smooth, time-dependent vector field X_t (not unique), such that

$$\omega_t \rfloor X_t = g - f, \quad X_t(p) = 0. \tag{8.38}$$

Let $\mathcal{F}_t$ be the flow generated by X_t, with $\mathcal{F}_0 = Id$. Note that $\mathcal{F}_t$ fixes p. It is then an easy computation using (8.38), plus the identity $\mathcal{L}_X \omega = d(\omega \rfloor X) + (d\omega) \rfloor X$, that

$$\frac{d}{dt} \big(\mathcal{F}_t^* \omega_t \big) = 0. \tag{8.39}$$

Hence $\mathcal{F}_1^* \omega_1 = \omega_0$, so $f \circ \mathcal{F}_1 = g$ and the proof of Proposition 8.3 is complete.

From Theorem 8.2, it follows that, given any compact, oriented, connected M, of dimension n, a finite number of surgeries on S^n yields two copies of M, with opposite orientations, say M and $M^{\#}$. Hence (8.21) holds with M replaced by the disjoint union $M \cup M^{\#}$. But, in view of (8.34), both sides of the resulting identity are equal to twice the corresponding sides of (8.21); for χ_c this follows easily from (8.4), and for χ it follows immediately from (8.5). We hence have the Chern-Gauss-Bonnet formula and also the identity $\chi(M) = \chi_c(M)$, modulo the task of showing the invariance of the right side of (8.21) under changes of metric on M.

We turn to the task of demonstrating such invariance. Say g_0 and g_1 are two Riemannian metric tensors on M, with associated SO(n)-bundles $P_0 \to M$, $P_1 \to M$, having curvature forms Ω_0 and Ω_1. We want to show that $\mathrm{Pf}\,(\Omega_1^b) - \mathrm{Pf}\,(\Omega_0^b)$ is exact on M. To do this, consider the family of metrics $g_t = tg_1 + (1-t)g_0$ on M, with associated SO(n)-bundles $P_t \to M$, for $t \in [0,1]$. These bundles fit together to produce a principal SO(n)-bundle $\widetilde{P} \to M \times [0,1]$. We know there exists a connection on this principal bundle. Let $T = \partial/\partial t$ on $M \times [0,1]$, and let $\widetilde{T}$ denote its horizontal lift (with respect to a connection chosen on $\widetilde{P}$). The flow generated by $\widetilde{T}$ commutes with the SO(n)-action on $\widetilde{P}$. Flowing along one unit of time then yields a diffeomprphism $\Phi : P_0 \to P_1$, commuting with the SO(n)-action, hence giving an isomorphism of SO(n)-bundles. Now applying Proposition 7.2 to the original connection on P_0 and to that pulled back from P_1 gives the desired invariance.

Before Chern's work, H. Hopf had established Theorem 8.1 when M is a compact hypersurface in $\mathbb{R}^{2k+1}$. Then C. Allendoerfer [Al] and W. Fenchel [Fen] proved (8.21) for the case when M is isometrically imbedded in $\mathbb{R}^{n+k}$, by relating the integral on the right to the integral over $\partial\mathcal{T}$ of the Gauss curvature of the boundary of a small tubular neighborhood $\mathcal{T}$ of M, and using the known result that $\chi(\partial\mathcal{T}) = 2\chi(M)$. At that time it was not known that every compact Riemannian manifold could be isometrically imbedded in Euclidean space. By other means, Allendoerfer and A.Weil [AW] proved Theorem 8.1, at least for real analytic metrics, via a triangulation and local isometric imbedding. Chern then produced an intrinsic proof of Theorem 8.1 and initiated a new understanding of characteristic classes.

In Chern's original paper [Cher], it is established that $\int_M \mathrm{Pf}\,(\Omega/2\pi)$ is equal to the index of a vector field X on M, by a sophisticated variant of the argument establishing Proposition 5.4, involving a differential form on the unit-sphere bundle of M related to, but more complicated than, the transgressed form (7.14). An exposition of this argument can also be found in [Poo] and in [Wil]. When dim $M = 2$, one can identify the unit-sphere bundle and the frame bundle, and in that case the form coincides with the transgressed form and the argument becomes equivalent to that used to prove Proposition 5.4. An exposition of the proof of Theorem 8.1 using tubes can be found in [Gr].

The Chern-Gauss-Bonnet theorem can also be considered as a special case of an index theorem for an elliptic differential operator. A proof in that spirit is given in Chapter 10, §7. More material on the Pfaffian is also developed there.

We mention a further generalization of the Gauss-Bonnet formula. If $E \to X$ is an SO($2k$)-bundle over a compact manifold X (say of dimension n), with metric connection ∇ and associated curvature Ω, then $\mathrm{Pf}\,(\Omega/2\pi)$ is defined as a ($2k$)-form on X. This gives a class $\mathrm{Pf}\,(E) \in \mathcal{H}^{2k}(X)$, independent of the choice of connection on E, as long as it is a metric connection. There is an extension of Theorem 8.1, describing the cohomology class of $\mathrm{Pf}\,(E)$ in $\mathcal{H}^{2k}(X)$. Treatments of this can be found in [KN] and in [Spi].

Exercises

1. Verify that when Ω is the curvature 2-form arising from the standard metric on S^{2k}, then
$$\int_{S^{2k}} \mathrm{Pf}\left(\frac{\Omega}{2\pi}\right) = 2.$$
2. Generalize Theorem 8.1 to the nonorientable case. (*Hint*: If M is not orientable, look at its orientable double cover $\widetilde{M}$. Use (8.4) to show that $\chi(\widetilde{M}) = 2\chi(M)$.) Using (8.16) as a local identity, define a *measure* $\widetilde{\mathrm{Pf}}(\Omega)$ in the nonorientable case.

3. Let M be a compact, complex manifold of complex dimension n (i.e., real dimension $2n$). Denote by $\mathcal{T}$ its tangent bundle, regarded as a complex vector bundle, with fibers $\mathcal{T}_p$ of complex dimension n. Show that
$$\int\limits_M c_n(\mathcal{T}) = \chi(M),$$
where $c_n(\mathcal{T})$ is the top Chern class, defined by (7.19) and (7.20).
4. If M_j are compact Riemannian manifolds with curvature forms Ω_j and $M_1 \times M_2$ has the product metric, with curvature form Ω, show directly that
$$\pi_1^* \mathrm{Pf}\,(\Omega_1) \wedge \pi_2^* \mathrm{Pf}\,(\Omega_2) = \mathrm{Pf}\,(\Omega),$$
where π_j projects $M_1 \times M_2$ onto M_j. If dim M_j is odd, set $\mathrm{Pf}\,(\Omega_j) = 0$. Use this to reprove (8.31) when $e(\Omega) = \mathrm{Pf}\,(\Omega)$.
5. Show directly that the right sides of (8.2) and (8.3) both vanish when M is a hypersurface of odd dimension in $\mathbb{R}^{n+1}$.

References

[AbM] R.Abraham and J.Marsden, *Foundations of Mechanics*, Benjamin/Cummings, Reading, Mass., 1978.

[Al] C. Allendoerfer, The Euler number of a Riemannian manifold, *Amer. J. Math.* 62(1940), 243–248.

[AW] C. Allendoerfer and A. Weil, The Gauss-Bonnet theorem for Riemannian polyhedra, *Trans. AMS* 53(1943), 101–129.

[Ar] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.

[BGM] M. Berger, P. Gauduchon, and E. Mazet, *Le Spectre d'une Variété Riemannienne*, LNM no.194, Springer-Verlag, New York, 1971.

[Bes] A. Besse, *Einstein Manifolds*, Springer-Verlag, New York, 1987.

[Car] M. doCarmo, *Riemannian Geometry*, Birkhauser, Boston, 1992.

[Ch] J. Cheeger, Analytic torsion and the heat equation, *Ann. Math.* 109(1979), 259–322.

[CE] J. Cheeger and D. Ebin, *Comparison Theorems in Riemannian Geometry*, North Holland, Amsterdam, 1975.

[Cher] S. S. Chern, A simple intrinsic proof of the Gauss-Bonnet formula for closed Riemannian manifolds., *Ann. Math.* 45(1944), 747–752.

[Chv] C. Chevalley, *Theory of Lie Groups*, Princeton Univ. Press, Princeton, N. J., 1946.

[EGH] T. Eguchi, P. Gilkey, and A. Hanson, Gravitation, Gauge Theories, and Differential Geometry, *Physics Reports*, Vol. 66, no. 6 (Dec. 1980).

[Fen] W. Fenchel, On total curvature of Riemannian manifolds, *J. London Math. Soc.* 15(1940), 15–22.

[FU] D. Freed and K. Uhlenbeck, *Instantons and Four-Manifolds*, Springer-Verlag, New York, 1984.

[Gil] P. Gilkey, *The Index Theorem and the Heat Equation*, Publish or Perish, Boston, 1974.

[Gr] A. Gray, *Tubes*, Addison-Wesley, New York, 1990.

[Gu] R. Gunning, *Lectures on Riemann Surfaces*, Princeton Univ. Press, Princeton, N. J., 1967.

[Gun] M. Günther, On the perturbation problem associated to isometric embeddings of Riemannian manifolds, *Ann. Global Anal. Geom.* 7(1989), 69–77.

[Hi] N. Hicks, *Notes on Differential Geometry*, Van Nostrand, New York, 1965.

[Jos] J. Jost, *Nonlinear Methods in Riemannian and Kahlerian Geometry*, Birkhauser, Boston, 1988.

[Kaz] J. Kazdan, Prescribing the curvature of a Riemannian manifold, *CBMS Reg. Conf. Ser. in Math.*, no. 57, AMS, Providence, R. I., 1985.

[KN] S. Kobayashi and N. Nomizu, *Foundations of Differential Geometry*, Interscience, New York, Vol. 1, 1963; Vol. 2, 1969.

[LM] H. B. Lawson and M. L. Michelson, *Spin Geometry*, Princeton Univ. Press, Princeton, N. J., 1989.

[Mil] J. Milnor, *Morse Theory*, Princeton Univ. Press, Princeton, N. J., 1963.

[Mil2] J. Milnor, *Topology from the Differentiable Viewpoint*, Univ. Press of Virginia, Charlottesville, 1965.

[MiS] J. Milnor and J. Stasheff, *Characteristic Classes*, Princeton Univ. Press, Princeton, N. J., 1974.

[Pal] R. Palais, ed., *Seminar on the Atiyah-Singer Index Theorem*, Princeton Univ. Press, Princeton, N. J., 1963.

[Poo] W. Poor, *Differential Geometric Structures*, McGraw-Hill, New York, 1981.

[Spi] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Vols. 1–5, Publish or Perish Press, Berkeley, Calif., 1979.

[Stb] S. Sternberg, *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, N. J., 1964.

[Str] D. Struik, *Lectures on Classical Differential Geometry*, Addison-Wesley, Reading, Mass., 1964.

[Wh] H. Whitney, *Geometric Integration Theory*, Princeton Univ. Press, Princeton, N. J., 1957.

[Wil] T. Wilmore, *Total Curvature in Riemannian Geometry*, Wiley, New York, 1982.