Riemannian Geometry

Lectures given during the fall of 1990

Hörmander, Lars

RIEMANNIAN GEOMETRY

Lars Hörmander

Department of Mathematics

Box 118, S-221 00 Lund, Sweden

RIEMANNIAN GEOMETRY

Lectures given during the fall of 1990 by

Lars Hörmander

Department of Mathematics

Box 118, S-221 00 Lund, Sweden

Lars Hörmander
Department of Mathematics
University of Lund
Box 118
S-221 00 Lund, Sweden
e-mail lvh@maths.lth.se (internet)

The notes have been typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX.

PREFACE

The inspiration for the theory of partial differential equations has always come from two main sources, physics and geometry. The interaction between all three areas has become intensified in recent years. The solution of the index problem by Atiyah and Singer in the early 60's forced the people working on differential equations to improve their knowledge of differential geometry. This was very useful in the subsequent development of microlocal analysis which mainly involved symplectic geometry, a topic which had previously mainly been cultivated in connection with ordinary differential equations. Riemannian geometry is central in the recent development of gauge theories, which rely on a mixture of geometry, physics and partial differential equations. There has also been a great deal of recent activity in the general theory of relativity, that is, in pseudo-Riemannian geometry — for example, the proof of the positive mass conjecture and global existence theorems for the vacuum Einstein equations. The solution of some purely geometric problems, such as the Yamabe problem and the isometric imbedding problem, have also enriched the theory of non-linear partial differential equations. The fundamental open problems in the theory of overdetermined systems of linear differential equations also require a strong background in geometry even to understand the present state of affairs. All this should be sufficient reason for an analyst to study geometry seriously.

In a half semester course it is only possible to present a brief outline of the most classical Riemannian geometry with a few glimpses of more recent developments. However, if enough interest is manifested, I plan to continue for one or several semesters more in order to be able to approach the research front, and these lectures should then be a convenient platform to build on. One could continue in many different directions. One possibility is to discuss pseudo-Riemannian manifolds with Lorentz signature (general relativity theory). Another is to discuss relations to the theory of functions of several complex variables; the presence of a complex structure gives a much richer structure.

Lund in December 1990

Lars Hörmander

Contents of Lectures on Riemannian Geometry

CHAPTER I

CURVES IN A EUCLIDEAN SPACE

Summary. In Section 1.1 we shall just recall the definition of the curvature of a curve in a Euclidean vector space. The special two dimensional case where the curvature can be given a sign is studied in somewhat greater depth in Section 1.2. We define the torsion of a curve and prove the Frenet formulas in Section 1.3. This leads to a discussion of moving frames in Section 1.4 which is a preparation for the Riemannian geometry in Chapter IV.

1.1. Curvature of a curve. Let V be a finite dimensional vector space over $\mathbf{R}$ with a symmetric scalar product (x, y) defined for $x, y \in V$ and the norm $\|x\| = (x, x)^{\frac{1}{2}}$. By a C^k curve in V we mean a C^k map

$$I \ni t \mapsto x(t) \in V \tag{1.1.1}$$

where I is an interval $\subset \mathbf{R}$, $k \geq 1$, and $x'(t) \neq 0$ for every t. (Self-intersections are allowed.) If J is another interval on $\mathbf{R}$ and $\varphi : J \to I$ is a C^k bijection with $\varphi' > 0$, then we regard the curve defined by $x \circ \varphi$ as the same curve as (1.1.1) with just a change of parametrization. The arc length $s(t)$ on (1.1.1) is defined up to an additive constant by

$$ds(t)/dt = \|x'(t)\|.$$

If we define $X(s(t)) = x(t)$ then X is also in C^k and $\|X'(s)\| = 1$. If $\tilde{s}$ is another parameter with this property, then $d\tilde{s}/ds = 1$ so $\tilde{s} - s$ is a constant.

Differentiation of the equation $1 = \|X'(s)\|^2 = (X'(s), X'(s))$ with respect to s gives

$$2(X''(s), X'(s)) = 0.$$

Thus the second derivative $X''(s)$ is orthogonal to the *tangent* $X'(s)$ of the curve. To interpret it geometrically we consider a circle with two perpendicular radii y, z of length R and center x_0; it is parametrized by

$$x(s) = x_0 + y\cos(s/R) + z\sin(s/R), \text{ thus } x'(s) = (-y/R)\sin(s/R) + (z/R)\cos(s/R).$$

Since $(y, y) = (z, z) = R^2$ and $(y, z) = 0$, we confirm using the second equation that $\|x'(s)\|^2 = 1$. The second derivative

$$x''(s) = -(y/R^2)\cos(s/R) - (z/R^2)\sin(s/R) = (x_0 - x(s))/R^2$$

is directed toward the center of the circle and $\|x''(s)\| = 1/R$.

DEFINITION 1.1.1. If $s \mapsto x(s) \in V$ is a C^2 curve parametrized by the arc length, then $x'(s)$ is the unit tangent vector at $x(s)$. If $x''(s) \neq 0$, then $n = x''(s)/\|x''(s)\|$ is called the principal normal of the curve at $x(s)$, $1/\|x''(s)\|$ is called the radius of curvature, and the circle with center at $x(s) + x''(s)/\|x''(s)\|^2$ lying in the plane spanned by the vectors $x'(s)$ and $x''(s)$ at $x(s)$ is called the osculating circle. One calls $\kappa = \|x''(s)\|$ the curvature at $x(s)$ even if $\kappa = 0$; otherwise $x''(s) = \kappa n$.

From the discussion above it is clear that when the curvature is not 0 then the osculating circle is the only one which has a tangency of higher order with the curve at $x(s)$.

Now suppose that we have a curve $t \mapsto x(t)$ with $x'(t) \neq 0$ which may not have the arc length as parameter. If we write $x(t) = X(s(t))$ as above, we obtain

$$x'(t) = X'(s)ds/dt, \qquad x''(t) = X''(s)(ds/dt)^2 + X'(s)d^2s/dt^2.$$

Since $X''(s)$ is orthogonal to $X'(s)$ we have

$$X''(s)(ds/dt)^2 = x''(t) - (x''(t), X'(s))X'(s)$$

or equivalently,

$$X''(s) = x''(t)/\|x'(t)\|^2 - x'(t)(x''(t), x'(t))/\|x'(t)\|^4. \tag{1.1.2}$$

EXERCISE 1.1.1. Calculate the curvature of a C^2 curve $t \mapsto x(t)$ with arbitrary parametrization.

EXERCISE 1.1.2. Show that the tangential component of $x''(t)$ is equal to d^2s/dt^2 times the unit tangent vector, and that the normal component is $(ds/dt)^2\kappa$ times the principal normal.

1.2. The two dimensional case. For curves in the plane $\mathbf{R}^2$ we can attach a sign to the curvature (which depends on the orientation of the plane). If $s \mapsto x(s)$ is a curve with the arc length s as parameter and if $x''(s) \neq 0$, then we change the direction of the normal $n(s)$ so that $x'(s)$ and $n(s)$ are positively oriented, that is, $\det(x'(s), n(s)) > 0$ and define the curvature $\kappa(s)$ so that $x''(s) = \kappa(s)n(s)$ as before. Thus $\kappa > 0$ means that moving along the curve one sees the curve to the left of the tangent.

Consider now a closed simple C^2 curve

$$\mathbf{R}/L\mathbf{Z} \ni s \mapsto x(s) \in \mathbf{R}^2$$

where s is still the arc length and $x(s) = x(s')$ if and only if $s - s' \in L\mathbf{Z}$. We can write

$$x'(s) = (\cos\theta(s), \sin\theta(s)), \quad 0 \leq s \leq L,$$

where θ is a C^1 function uniquely determined up to an integer multiple of 2π. Since

$$x''(s) = (-\sin\theta, \cos\theta)d\theta/ds = \kappa n = \kappa(-\sin\theta, \cos\theta),$$

we have $\kappa = d\theta/ds$. The integral

$$\int_0^L \kappa\, ds = \int_0^L d\theta = \theta(L) - \theta(0) \tag{1.2.1}$$

is obviously a multiple of 2π.

THEOREM 1.2.1. *The integral* (1.2.1) *of the curvature of a simple closed curve is* $\pm 2\pi$ *with the positive sign if the curve lies entirely to the left of some tangent when one moves around it in the positive direction.*

PROOF. The normalized chord direction

$$x(t,s) = (x(t) - x(s))/\|x(t) - x(s)\|, \quad 0 \le s < t \le L$$

becomes a continuous function in the closed triangle $T = \{(t,s); 0 \le s \le t \le L\}$ if we define $x(s,s) = x'(s)$. Since T is simply connected we can find a continuous function $\varphi(t,s)$ in T such that $\varphi(s,s) = \theta(s)$, $0 \le s \le L$, and

$$x(t,s) = (\cos\varphi(t,s), \sin\varphi(t,s)), \quad (t,s) \in T.$$

We want to find

$$\theta(L) - \theta(0) = \varphi(L,L) - \varphi(L,0) + \varphi(L,0) - \varphi(0,0) = 2(\varphi(L,0) - \varphi(0,0)),$$

where the last equality follows from the fact that $x(L,s) = -x(s,0)$. Choose the origin of the parametrization so that $x_2(0) = \min_s x_2(s)$. Then $\varphi(s,0)$ only varies between 0 and π. If $x_1'(0) > 0$ then $\varphi(0,0) = 0$ and $\varphi(L,0) = \pi$ so the integral of the curvature becomes 2π. If $x_1'(0) < 0$ the sign changes, which proves the theorem.

EXERCISE 1.2.1. Show that when $x_1'(0) = 1$ in the preceding argument, then the variation $d(y)$ of the argument of $(x(s) - y)/\|x(s) - y\|$ for $0 \le s \le L$ is either 0 or 2π for all points y in the complement Ω of the curve, and that $\Omega_e = \{y \in \Omega; d(y) = 0\}$ is an open connected unbounded set while $\Omega_i = \{y \in \Omega; d(y) = 2\pi\}$ is an open connected bounded set, the interior of the curve. (Jordan's curve theorem.) Hint: Examine first a neighborhood of $x(0)$, then a neighborhood of the curve.

EXERCISE 1.2.2. Show that if $\Gamma : [0,T] \ni s \mapsto x(s) \in \mathbf{R}^3$ is a simple closed C^2 curve parametrized by the arc length, and $N \in S^2$, then $s \mapsto x(s) - N(x(s), N)$ is a C^2 curve in the orthogonal plane of N for almost all N. Write an expression for its curvature and conclude that if $\kappa(s)$ is the curvature of Γ, then $\int_0^T \kappa(s)\,ds \ge 2\pi$. (See also Milnor [1].)

EXERCISE 1.2.3. Show that if

$$\mathbf{R}/T\mathbf{Z} \ni t \mapsto x(t) \in \mathbf{R}^2$$

is a simple closed C^2 curve, then

$$L = \int_0^T |\det(x'(t), x''(t))|^{1/3}\,dt$$

is independent of the parametrization. It is called the affine length. Show that L^3/A is invariant under affine transformations if A is the area of the interior of the curve, calculate the quotient for an ellipse. (By a theorem of Blaschke this is the least upper bound for all convex curves; see Burago and Zalgaller [1], p. 7.)

1.3. The Frenet formulas. We shall now return to the study of a C^k curve (1.1.1) taking the arc length s as parameter. If the curvature is not equal to 0 at $x(s_0)$, then the principal normal $n(s)$ is defined for s near s_0. If $k \geq 3$ we can take its derivative $n'(s)$. Since $(n(s), n(s)) = 1$ we obtain $2(n'(s), n(s)) = 0$, and since $(n(s), x'(s)) = 0$ we have

$$(n'(s), x'(s)) + (n(s), x''(s)) = 0, \quad \text{that is, } (n'(s), x'(s)) = -\kappa(s).$$

Hence $n'(s) + \kappa(s)x'(s)$ is orthogonal to the plane spanned by $x'(s)$ and $n(s)$. The length $\tau(s)$ of this vector is called the *torsion* of the curve at $x(s)$. If $\tau(s) \neq 0$, then normalization gives a unit vector $b(s)$, called the *binormal* of the curve at $x(s)$, and we have

$$n'(s) = -\kappa(s)x'(s) + \tau(s)b(s).$$

The procedure can be continued by differentiation of b with respect to s. Differentiation of the equations

$$(b(s), x'(s)) = 0, \quad (b(s), n(s)) = 0, \quad (b(s), b(s)) = 1$$

gives

$$\begin{gathered}(b'(s), x'(s)) = -(b(s), x''(s)) = 0,\\ (b'(s), n(s)) = -(b(s), n'(s)) = -\tau(s), \quad (b'(s), b(s)) = 0.\end{gathered}$$

If $\dim V = 3$ the vectors $x'(s)$, $n(s)$ and $b(s)$ form a complete orthonormal system, so we get the third of the *Frenet formulas*

$$\begin{aligned} x''(s) &= & \kappa(s)n(s) & \\ n'(s) &= -\kappa(s)x'(s) & & + \tau(s)b(s) \\ b'(s) &= & - \tau(s)n(s). & \end{aligned}$$

However, the procedure becomes more illuminating in the higher dimensional case if we use somewhat different notation.

Consider a C^{k+1} curve (1.1.1) with arbitrary parametrization, such that for all $t \in I$ the derivatives $x'(t), \dots, x^{(k)}(t)$ are linearly independent but $x'(t), \dots,$ $x^{(k+1)}(t)$ are linearly dependent. Clearly $k \leq \dim V$. For any given sufficiently differentiable curve this is true if I is replaced by any interval in the dense open subset where the rank of $x'(t), x''(t), \dots$ is locally maximal. The hypothesis is independent of the choice of parametrization, and so is the linear span $E_j(t)$ of $x'(t), \dots, x^{(j)}(t)$ when $j \leq k$. In fact, if $\tilde{t}$ is another parameter then

$$\frac{d^j x}{d\tilde{t}^j} - \left(\frac{dt}{d\tilde{t}}\right)^j \frac{d^j x}{dt^j} \in E_{j-1}$$

so this follows inductively. Application of the Gram-Schmidt orthogonalization procedure to the sequence $x'(t), \dots, x^{(k)}(t)$ gives orthonormal vectors $e_1(t), \dots, e_k(t)$:

$$e_1(t) = x'(t)/c_1(t), \dots, e_j(t) = \big(x^{(j)}(t) - \sum_{i<j}(x^{(j)}(t), e_i(t))e_i(t)\big)/c_j(t),$$

where $c_j(t) > 0$ is chosen so that $\|e_j(t)\| = 1$. These vectors at $x(t)$ do not depend on the choice of parametrization; $e_1(t)$ is the unit tangent vector, $e_2(t)$ is the principal normal, $e_3(t)$ is the binormal, and so on.

Assume now for the sake of simplicity that the parameter t is the arc length s. By the construction and the hypothesis that $E_{k+1}(s) = E_k(s)$ we have

$$e_i'(s) = \sum_{j=1}^{k} \omega_{ij}(s)e_j(s), \quad i \le k,$$

where $\omega_{ij}(s) = (e_i'(s), e_j(s)) = 0$ if $j > i+1$. If we differentiate the equations

$$(e_i(s), e_j(s)) = \delta_{ij}$$

expressing the orthonormality of $e_1(s), \dots, e_k(s)$, we obtain

$$0 = (e_i'(s), e_j(s)) + (e_i(s), e_j'(s)) = \omega_{ij}(s) + \omega_{ji}(s),$$

so the matrix $(\omega_{ij}(s))$ is skew symmetric. Thus $\omega_{ii}(s) = 0$, and since $\omega_{ij}(s) = 0$ when $j > i+1$, this is also true when $i > j+1$. If we set $\kappa_j(s) = \omega_{j,j+1}(s)$, $1 \le j < k$, it follows that the matrix (ω_{ij}) has the special form for $k = 4$, say:

$$\begin{pmatrix} 0 & \kappa_1 & 0 & 0 \\ -\kappa_1 & 0 & \kappa_2 & 0 \\ 0 & -\kappa_2 & 0 & \kappa_3 \\ 0 & 0 & -\kappa_3 & 0 \end{pmatrix}.$$

Here $\kappa_1(s)$ is of course the curvature, and κ_2 is the torsion. The Frenet formulas can now be written

$$\begin{aligned} &e_1'(s) = \kappa_1(s)e_2(s), \quad e_j'(s) = -\kappa_{j-1}(s)e_{j-1}(s) + \kappa_j(s)e_{j+1}(s),\ 1 < j < k, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad e_k'(s) = -\kappa_{k-1}(s)e_{k-1}(s). \end{aligned} \tag{1.3.1}$$

This system of differential equations has a unique solution with given initial data. If these are orthonormal then the solution remains orthonormal, and the solution must belong to the space $E_k(s_0)$, for there is a solution with $e_1(s), \dots, e_k(s)$ contained in this space which is equal to $e_1(s_0), \dots, e_k(s_0)$ at s_0, so $E_k(s)$ is independent of s and the curve lies in an affine subspace of dimension k.

EXERCISE 1.3.1. Let $k \ge 3$ and determine the Taylor expansion of $x(s)$ with error $o(s^3)$ at $0 \in I$ in terms of $x(0)$, $e_1(0)$, $e_2(0)$, $e_3(0)$ and the Taylor expansions of κ_1 and κ_2 at 0.

1.4. Moving frames. In the preceding section we studied orthonormal vectors depending on a parameter which were adapted to a given curve. However, some of the arguments are of a much more general nature so we shall take them up once more.

Let V again be a Euclidean space of dimension $n < \infty$, and denote by $F(V)$ the set of all orthonormal frames $e_1, \dots, e_n \in V$. Since $F(V) \subset V^n$ it makes sense to say that a curve $I \ni t \mapsto f(t) \in F(V)$ is differentiable. Writing $f(t) = (e_1(t), \dots, e_n(t))$ we have by definition of $F(V)$

$$(e_i(t), e_j(t)) = \delta_{ij}$$

and conclude as in Section 1.3 by differentiating that

$$de_i/dt = \sum_{j=1}^{n} \omega_{ij}(t) e_j(t), \quad \text{where } \omega_{ij}(t) = -\omega_{ji}(t). \tag{1.4.1}$$

This can be interpreted as follows. If $e = (e_1, \dots, e_n) \in F(V)$ and $f = (f_1, \dots, f_n) \in F(V)$, then

$$e_i = \sum_{j=1}^{n} O_{ij} f_j, \quad \text{where } \sum_{k=1}^{n} O_{ik} O_{jk} = \delta_{ij}. \tag{1.4.2}$$

Thus $O\,{}^tO = I$, where I is the identity matrix, or equivalently ${}^tO\,O = I$. Such matrices are called *orthogonal*, and one denotes by $\mathbf{O}(n)$ the set of orthogonal $n \times n$ matrices, which is a group under matrix multiplication. When (1.4.2) is valid we write $e = Of$. With any fixed e^0 we can now write $e(t) = O(t)e^0$ where $O(t)$ is differentiable with values in $\mathbf{O}(n)$, and we have $de(t)/dt = (dO(t)/dt)e^0$. In particular, if $t = 0$ and we take $e^0 = e(0)$, then $O(0) = I$, and (1.4.1) means that

(1.4.3) $O(t)$ differentiable at 0 with values in $\mathbf{O}(n)$, $O(0) = I$
$\implies$ $O'(0)$ is skew symmetric.

The proof is obvious: the equation $O(t)\,{}^tO(t) = I$ implies

$$O'(0)\,{}^tO(0) + O(0)\,{}^tO'(0) = 0, \text{ that is, } O'(0) + {}^tO'(0) = 0.$$

This means that infinitesimal rotations are defined by skew symmetric matrices. Every such infinitesimal rotation ω gives rise to a one parameter group of rotations $O(t)$, that is a function $\mathbf{R} \ni t \mapsto O(t) \in \mathbf{O}(n)$ such that

$$O(t)O(s) = O(t+s), \quad O'(0) = \omega. \tag{1.4.4}$$

The first condition implies $O(0) = I$, and differentiation with respect to s gives $O'(t) = O(t)\omega$ when $s = 0$, hence

$$O(t) = e^{t\omega} = \sum_0^\infty (t\omega)^j / j!.$$

The properties (1.4.4) follow at once, and since ${}^tO(t) = O(-t)$ it follows that $O(t) \in \mathbf{O}(n)$.

Let us consider the examples of lowest dimension. If $n = 2$ then

$$\omega = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}, \quad e^{t\omega} = \begin{pmatrix} \cos(t\theta) & \sin(t\theta) \\ -\sin(t\theta) & \cos(t\theta) \end{pmatrix},$$

so $e^{t\omega}$ means rotation by the angle $t\theta$ in the negative direction. If $n = 3$ then

$$\omega = \begin{pmatrix} 0 & \theta_3 & -\theta_2 \\ -\theta_3 & 0 & \theta_1 \\ \theta_2 & -\theta_1 & 0 \end{pmatrix}.$$

Noting that $\omega\theta = 0$ it is easy to see that $e^{t\omega}$ means rotation by the angle $t|\theta|$ around θ.

Exercise 1.4.1. Show that if

$$\omega = \begin{pmatrix} 0 & \theta_3 & -\theta_2 \\ -\theta_3 & 0 & \theta_1 \\ \theta_2 & -\theta_1 & 0 \end{pmatrix},$$

then $e^{t\omega} = A + B\cos(t|\theta|) + C\sin(t|\theta|)$ and determine the 3×3 matrices A, B, C. Use this to find the curve in $\mathbf{R}^3$ with constant curvature and torsion which passes through the origin with tangent, principal normal and binormal along the positive x_1, x_2 and x_3 axes.

Exercise 1.4.2. Show that if n is odd and S is a skew symmetric $n \times n$ matrix, then the equation $S\theta = 0$ has at least one solution $\theta \neq 0$, and show that $e^{tS}\theta = \theta$. Show for any n that if S is a skew symmetric $n \times n$ matrix $\neq 0$, then there is a two dimensional plane $W \subset \mathbf{R}^n$ such that W and its orthogonal complement $W^\perp$ are left invariant by S and therefore by e^{tS}. Describe the structure of S and e^{tS} geometrically in general. Hint: Use that S/i is hermitian symmetric in $\mathbf{C}^n$.

Exercise 1.4.3. Show that if $O \in \mathbf{O}(n)$, then $\mathbf{R}^n$ is an orthogonal direct sum $W_+ \oplus W_- \oplus W$ such that $\pm O$ is the identity on $W_\pm$ and W is the orthogonal direct sum of twodimensional subspaces where O is a rotation. Conclude that the component of the identity in $\mathbf{O}(n)$ is the subgroup $\mathbf{SO}(n)$ of elements with determinant 1, and that every element in $\mathbf{SO}(n)$ is of the form e^S for some skew symmetric S. Hint: Use that O is unitary in $\mathbf{C}^n$.

We end this section with a few exercises giving some information on Lie groups and Lie algebras which will be useful later on. Let $\mathbf{M}(V)$ be the set of linear transformations in the finite dimensional vector space V, and let $\mathbf{GL}(V)$ be the subset of invertible ones. Then $\mathbf{GL}(V)$ is a group under multiplication; the identity is denoted by I. With some Euclidean norm in V we use the corresponding operator norm in $\mathbf{M}(V)$. When $V = \mathbf{R}^n$ we write $\mathbf{M}(n)$ and $\mathbf{GL}(n)$ instead of $\mathbf{M}(\mathbf{R}^n)$ and $\mathbf{GL}(\mathbf{R}^n)$.

EXERCISE 1.4.4. Show that

(1) $e^X = \sum_0^\infty X^n/n!$ converges in $\mathbf{M}(n)$ if $X \in \mathbf{M}(n)$, and that $e^X \in \mathbf{GL}(n)$, $e^{-X}e^X = I$.
(2) $\log(I+X) = \sum_1^\infty (-1)^{n-1}X^n/n$ converges in $\mathbf{M}(n)$ if $X \in \mathbf{M}(n)$ and $\|X\| < 1$, and $e^{\log(I+X)} = I + X$ then.
(3) $\log e^X = X$ if $\|X\| < \log 2$.
(4) if $\mathbf{R} \ni t \mapsto A(t) \in \mathbf{GL}(n)$ is a continuous one parameter subgroup, then

$$A(t)\int_0^\varepsilon A(s)\,ds = \int_t^{t+\varepsilon} A(s)\,ds;$$

deduce that $A : \mathbf{R} \to \mathbf{M}(n)$ is differentiable and that $A(t) = e^{tX}$ where $X = A'(0) \in \mathbf{M}(n)$.
(5) if $X, Y \in \mathbf{M}(n)$, then

$$e^{X+Y} = \lim_{\nu\to\infty}(e^{X/\nu}e^{Y/\nu})^\nu.$$

EXERCISE 1.4.5. Let $\mathbf{G}$ be a closed subgroup of $\mathbf{GL}(n)$, and

$$\mathfrak{g} = \{X \in \mathbf{M}(n); e^{tX} \in \mathbf{G}, t \in \mathbf{R}\}.$$

Show that

(1) $\mathfrak{g}$ is a linear subspace of $\mathbf{M}(n)$.
(2) if $g \in \mathbf{G}$ and $X \in \mathfrak{g}$, then

$$ge^{tX}g^{-1} = e^{tY}, \quad t \in \mathbf{R},$$

where $Y = gXg^{-1} \in \mathfrak{g}$. One writes $Y = \mathrm{Ad}(g)X$ also (the *adjoint representation*).
(3) if $Y, X \in \mathfrak{g}$, then $t \mapsto \mathrm{Ad}(e^{tY})$ is a one parameter subgroup of $\mathbf{GL}(\mathfrak{g})$, hence $\mathrm{Ad}(e^{tY}) = e^{t\,\mathrm{ad}(Y)}$ where $\mathrm{ad}(Y) \in \mathbf{M}(\mathfrak{g})$ (the infinitesimal adjoint representation); show that $\mathrm{ad}(Y)X = [Y, X] = YX - XY$ and deduce that $\mathfrak{g}$ is a Lie algebra, that is, closed also under commutators.
(4) show that $\{e^X; X \in \mathfrak{g}, \|X\| < 1\}$ is a neighborhood of the identity in $\mathbf{G}$. Hint: Assume that this is false and take a sequence $X_k \to 0$ in $\mathbf{M}(n)$ such that $e^{X_k} \in \mathbf{G}$. Write $X_k = Y_k + Z_k$ where $Z_k \in \mathfrak{g}$ and $0 \neq Y_k \in \mathfrak{g}^\perp$ (the orthogonal complement with respect to some Euclidean norm in $\mathbf{M}(n)$). Note that

$$\mathbf{G} \ni e^{X_k}e^{-Z_k} = e^{Y_k+Z_k}e^{-Z_k} = e^{W_k}, \text{ where } W_k = Y_k + O(|Y_k||Z_k|).$$

Pass to a subsequence such that $W_k/\|W_k\| \to W$ and get a contradiction by proving that $W \in \mathfrak{g} \cap \mathfrak{g}^\perp$.

EXERCISE 1.4.6. Show that if $U \in \mathbf{SU}(n)$, the group of unitary $n \times n$ matrices with determinant 1, and $h \in \mathbf{H}_0(n)$, the space of $n \times n$ hermitian symmetric matrices with trace 0, then $UhU^* \in \mathbf{H}_0(n)$, which gives a homomorphism $\mathbf{SU}(n) \to \mathbf{O}(n^2 - 1)$. Describe the one parameter subgroups of $\mathbf{SU}(n)$, and prove that $\mathbf{SU}(2) \to \mathbf{SO}(3)$ is surjective, and in fact a double cover. Show that $\mathbf{SU}(2)$ is simply connected by proving that

$$\mathbf{SU}(2) \ni U \mapsto (\operatorname{Re} U_{11}, \operatorname{Im} U_{11}, \operatorname{Re} U_{21}, \operatorname{Im} U_{21}) \in S^3$$

is a bijection on the unit sphere $S^3 \subset \mathbf{R}^4$, and write down the inverse.

CHAPTER II

CURVATURE OF SUBMANIFOLDS OF A EUCLIDEAN SPACE

Summary. In Section 2.1 we introduce the notions of first and second fundamental forms for a submanifold of a Euclidean space. After introducing the Christoffel symbols and geodesic curvature, we define the Riemann curvature tensor and connect it with the second fundamental form through the Gauss equations. Section 2.2 is devoted to the special case of a hypersurface, and in particular the Gauss map. Basic algebraic (symmetry) properties of the curvature tensor are given in Section 2.3. The discussion includes the first Bianchi identity and the decomposition of the curvature tensor with special emphasis on the four dimensional case.

2.1. Curves on a submanifold of a Euclidean space. Let M be a C^μ submanifold of dimension n in a finite dimensional vector space V of dimension N. This means that to every $X_0 \in M$ there is a neighborhood U in V and

(1) a map $F \in C^\mu(U, \mathbf{R}^{N-n})$ such that F' is surjective and $M \cap U = \{X \in U; F(X) = 0\}$;
(2) there is a neighborhood ω of 0 in $\mathbf{R}^n$ and a map $f \in C^\mu(\omega, V)$ with $f(0) = X_0$ and injective differential, such that $f(\omega) = M \cap U$.

These conditions are equivalent. In fact, given F as in (1) we can choose $G \in C^\mu(U, \mathbf{R}^n)$ such that (F', G') is of rank N at X_0 and $G(X_0) = 0$. By the implicit function theorem there is an inverse $\Phi \in C^\mu$ defined in a neighborhood of 0 in $\mathbf{R}^N$. Since $X = \Phi(F(X), G(X))$ when X is in a neighborhood of X_0, it follows that a neighborhood of X_0 in M is parametrized by $X = \Phi(0, y)$ where $y = G(X)$ is in a neighborhood of 0 in $\mathbf{R}^n$. Conversely, $(F(\Phi(0, y)), G(\Phi(0, y))) = (0, y)$, so $\Phi(0, y) \in M$. Choosing G as the projection on a suitable n dimensional coordinate plane, we obtain a representation (2) where $N - n$ coordinates are C^μ functions of the other n.

On the other hand, given f as in (2) we can choose $g \in C^\mu(\mathbf{R}^{N-n}, V)$ with $g(0) = 0$ so that $(x, y) \mapsto f(x) + g(y)$ has bijective differential at $(0, 0)$. Let $\varphi = (\varphi_1, \varphi_2)$ be an inverse from a neighborhood of X_0 to a neighborhood of 0 in $\mathbf{R}^n \times \mathbf{R}^{N-n}$. Then $\varphi(f(x)) = (x, 0)$ for x in a neighborhood of 0 in $\mathbf{R}^n$, and $X = f(\varphi_1(X)) + g(\varphi_2(X)) = f(\varphi_1(X))$ if $\varphi_2(X) = 0$, so (1) is valid if $F = \varphi_2$.

If $I \ni t \mapsto X(t)$ is a C^ν curve for some $\nu \le \mu$ which is contained in M, and if f is a parametrization of M near $X(t_0) = X_0$ as in (2) above, then we can write $X(t) = f(x(t))$ where $x(t) \in \omega$, if t is close to t_0. The local coordinates $x(t)$ are also C^ν functions of t, for f has a C^μ left inverse φ by the inverse function theorem, and $x(t) = \varphi(X(t))$. Thus it is equivalent to say that $X(t)$ is a C^ν function with values in V or that the local coordinates $x(t)$ are C^ν functions of t. (In the same way we see that a change of local coordinates is made by a C^μ map.) This will be the only available definition for abstract manifolds in Chapter III. We shall consistently use upper case letters for points in V and lower case letters for points in the parameter space. It is the latter which will survive in Chapter III.

If $\nu \geq 1$ then the tangent of our curve has the direction

$$\frac{dX}{dt} = \frac{df(x(t))}{dt} = \sum_1^n \frac{dx_j}{dt} f_j(x(t)); \quad f_j(x) = \partial f(x)/\partial x_j = \partial_j f(x). \tag{2.1.1}$$

Thus the tangent vectors of all C^1 curves passing through $f(x)$ span the plane

$$T_{f(x)}M = f'(x)\mathbf{R}^n,$$

called the *tangent plane* of M at $f(x)$. It is of course independent of the choice of f, but $f'(x)$ identifies it with $\mathbf{R}^n$ in a manner which is not invariant. From now on we assume that V is a Euclidean vector space. The arc length s on the curve is then given by

$$(ds/dt)^2 = \sum_{j,k=1}^n g_{jk}(x(t))\tau_j\tau_k; \quad \tau_j = dx_j/dt, \quad g_{jk}(x) = (f_j(x), f_k(x)). \tag{2.1.2}$$

One calls the quadratic form (2.1.2) the *first fundamental form* of M. It is a quadratic form depending on $x \in \omega$, which can invariantly be regarded as a quadratic form in $T_{f(x)}M$, for $\mathbf{R}^n \ni \tau \mapsto \sum_1^n \tau_j f_j(x) = X \in T_{f(x)}M$ is a bijection, and the form is equal to the square of the norm of the tangent vector X in V.

Assuming from now on that $\mu \geq 2$ and that we have a C^2 curve, we compute the second derivative:

$$\frac{d^2 f(x(t))}{dt^2} = \sum_{j=1}^n \frac{d^2 x_j}{dt^2} f_j(x(t)) + \sum_{j,k=1}^n \frac{dx_j}{dt}\frac{dx_k}{dt} f_{jk}(x(t)), \tag{2.1.3}$$

where $f_{jk}(x) = \partial_j\partial_k f(x)$. Second derivatives of x just occur in the first sum, and the coefficients there are tangent vectors. If we project on the *normal plane* $N_{f(x)}M$ of $T_{f(x)}M$, we can eliminate this sum and obtain:

Theorem 2.1.1 (Meusnier). *Let $h_{jk}(x)$, $j,k = 1,\dots,n$, be the orthogonal projection of $f_{jk}(x) = \partial_j\partial_k f(x)$ in the normal plane $N_{f(x)}M$ of M at $f(x)$. For every C^2 curve on M with the unit tangent vector $f'(x)\tau$ at $f(x)$, the sum*

$$\sum_{j,k=1}^n h_{jk}(x)\tau_j\tau_k \tag{2.1.4}$$

is then equal to the curvature times the orthogonal projection of the principal normal in $N_{f(x)}M$.

We can regard (2.1.4) as a quadratic form H in $T_{f(x)}M$ with values in $N_{f(x)}M$; by Theorem 2.1.1 it is then independent of the choice of parametrization. One calls (2.1.4) the *second fundamental form* of M. Classically it was defined for hypersurfaces (in $\mathbf{R}^3$). In that case, an orientation of M identifies $N_{f(x)}M$ with $\mathbf{R}$, so one can then regard H as a real valued quadratic form. We shall sometimes introduce an orthonormal basis in $N_{f(x)}M$ and will then be able to write H as an $N-n$ tuple of scalar quadratic forms.

We shall now consider the tangential components of (2.1.3). Since $f_1(x), \dots, f_n(x)$ form a basis in $T_{f(x)}M$, we can write

$$f_{ik}(x) = \sum_{l=1}^{n} \Gamma_{ik}{}^{l}(x) f_l(x) + h_{ik}(x), \tag{2.1.5}$$

for the normal component is $h_{ik}(x)$ by definition. To calculate the coefficients $\Gamma_{ik}{}^{l}$ we take the scalar product with f_j which gives

$$\sum_{l=1}^{n} \Gamma_{ik}{}^{l} g_{lj} = (f_{ik}, f_j). \tag{2.1.6}$$

By a miracle one can compute the right-hand side by means of the coefficients of the first fundamental form and their derivatives, for we have

$$\begin{aligned} \partial_k g_{ij} &= (f_{ik}, f_j) + (f_i, f_{jk}) \\ \partial_i g_{jk} &= (f_{ji}, f_k) + (f_j, f_{ki}), \\ \partial_j g_{ki} &= (f_{kj}, f_i) + (f_k, f_{ij}). \end{aligned}$$

The desired quantity occurs in the first two equations so we add them and subtract the third, which gives

$$\Gamma_{ikj} = (f_{ik}, f_j) = \tfrac{1}{2}(\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}), \tag{2.1.7}$$

where the first equality is a new definition. Note the symmetry in the indices i, k, and that

$$\partial_k g_{ij} = \Gamma_{ikj} + \Gamma_{jki}. \tag{2.1.7$'$}$$

If (g^{ij}) denotes the inverse of the matrix (g_{ij}) we obtain from (2.1.6) and (2.1.7)

$$\Gamma_{ik}{}^{l} = \sum_{j=1}^{n} g^{lj}(f_{ik}, f_j) = \sum_{j=1}^{n} g^{lj}\Gamma_{ikj}. \tag{2.1.8}$$

Definition 2.1.2. The functions Γ_{ikj} and $\Gamma_{ik}{}^{j}$ are called Christoffel symbols of the first and second kind; they are determined by the first fundamental form.

Remark. The classical notation was $[ij, k]$ and $\left\{ \begin{matrix} l \\ ij \end{matrix} \right\}$ instead of Γ_{ijk} and $\Gamma_{ij}{}^{l}$.

Exercise 2.1.1. Prove that if $g = \det(g_{ij})$, then $\partial_l g = 2g \sum_i \Gamma_{il}{}^{i}$, that is, $\partial_l \sqrt{g} = \sqrt{g} \sum_i \Gamma_{il}{}^{i}$.

Summing up the preceding discussion, we can now write (2.1.3) in the form

$$\frac{d^2 f(x(t))}{dt^2} = \sum_{j=1}^{n} \Big(\frac{d^2 x_j}{dt^2} + \sum_{i,k=1}^{n} \Gamma_{ik}{}^{j} \frac{dx_i}{dt}\frac{dx_k}{dt}\Big) f_j(x) + \sum_{i,k=1}^{n} \frac{dx_i}{dt}\frac{dx_k}{dt} h_{ik}(x). \tag{2.1.9}$$

We discussed the normal component in Theorem 2.1.1, and we can now examine the tangential component:

THEOREM 2.1.3. *Let $s \to f(x(s))$ be a C^2 curve in M, parametrized by the arc length, that is, $\sum g_{jk}(x)\, dx_j/ds\, dx_k/ds = 1$. Then the curvature times the orthogonal projection of the principal normal in $T_{f(x)}M$ is equal to*

$$\sum_{j=1}^{n} \Big(\frac{d^2 x_j}{ds^2} + \sum_{i,k} \Gamma_{ik}{}^{j} \frac{dx_i}{ds} \frac{dx_k}{ds} \Big) f_j(x). \tag{2.1.10}$$

Thus the vector (2.1.10) does not depend on the choice of parametrization f. The length is called the *geodesic curvature* of the curve at $f(x)$, and the direction is called the *geodesic principal normal direction.*

In (2.1.10) we have just computed the derivative of the tangent vector along the curve. However, the argument is much more general. Suppose that we have a C^1 vector field v defined in a neighborhood of the curve and that v is *everywhere tangent* to M. Then we can write the vector at $f(x)$ as

$$v(x) = \sum_{1}^{n} v_j(x) f_j(x),$$

and we obtain

$$\frac{dv(x(t))}{dt} = \sum_{j,k=1}^{n} ((\partial_k v_j(x)) f_j(x) + v_j(x) f_{jk}(x)) x'_k.$$

By (2.1.5) it follows that

$$\frac{dv(x(t))}{dt} = \sum_{j,k=1}^{n} (\partial_k v_j + \sum_{i=1}^{n} \Gamma_{ik}{}^{j} v_i) x'_k f_j(x) \quad \mod N_{f(x)}. \tag{2.1.11}$$

This is called the *covariant derivative* of the vector field along the curve. Note that one does not need to know the embedding function in order to compute (2.1.11); it suffices to know the first fundamental form. (The right-hand side of (2.1.11) is also well defined if v is just given on the curve.) This aspect will be discussed systematically in Chapter III.

Now we pass to determining the tangential components of the derivatives of a C^1 *normal* vector field $n(x)$ at $f(x)$. We can write

$$\partial_i n(x) = \sum_{k=1}^{n} n_{ik}(x) f_k(x) \quad \mod N_{f(x)}.$$

Since differentiation of the equation $(n, f_j) = 0$ gives

$$(\partial_i n, f_j) + (n, f_{ij}) = 0, \quad \text{that is,} \ \sum_{k=1}^{n} n_{ik} g_{kj} = -(h_{ij}, n)$$

it follows that

$$n_{ik} = -\sum_{j=1}^{n}(h_{ij}, n)g^{kj}. \tag{2.1.12}$$

This means that $y \mapsto \sum y_i \partial_i n(x) \bmod N_{f(x)}$ regarded as a linear transformation in $T_{f(x)}$ is precisely the linear transformation defined by the scalar product $(H, -n)$ of the second fundamental form and $-n$, using the identification of quadratic forms and linear transformations given by the first fundamental form.

It is now easy to find the tangential component of $f_{ijk} = \partial_i\partial_j\partial_k f$. (The following formulas should be understood in the sense of distribution theory if M is only in C^2, but one can also assume that $M \in C^3$ and use approximation to extend the final formulas where no third derivatives occur to the C^2 case.) Since $f_{ij} = \sum \Gamma_{ij}{}^l f_l + h_{ij}$ we obtain

$$f_{kij} = \partial_k f_{ij} = \sum_{m=1}^{n}(\partial_k \Gamma_{ij}{}^m + \sum_{l=1}^{n}\Gamma_{ij}{}^l\Gamma_{lk}{}^m - \sum_{l=1}^{n}(h_{ij}, h_{kl})g^{lm})f_m \mod N_{f(x)}.$$

Since $f_{kij} - f_{jik} = 0$, we have proved

THEOREM 2.1.4 (THE GAUSS EQUATIONS). *The first and the second fundamental form are related by the equations*

$$\begin{aligned} R^m{}_{ijk} &= \partial_j\Gamma_{ik}{}^m - \partial_k\Gamma_{ij}{}^m + \sum_{l=1}^{n}(\Gamma_{ik}{}^l\Gamma_{lj}{}^m - \Gamma_{ij}{}^l\Gamma_{lk}{}^m) \\ &= \sum_{l=1}^{n}\big((h_{ik}, h_{jl}) - (h_{ij}, h_{kl})\big)g^{lm}; \end{aligned} \tag{2.1.13}$$

here the first equality is a definition.

The second equality in (2.1.13) suggests that we should introduce

$$R_{lijk} = \sum_{m=1}^{n} g_{lm}R^m{}_{ijk} \tag{2.1.14}$$

for then it takes the simple form

$$R_{lijk} = (h_{ik}, h_{jl}) - (h_{ij}, h_{kl}). \tag{2.1.15}$$

To rewrite the first equality in (2.1.13) we note that in view of (2.1.7)′

$$\sum_{m=1}^{n} g_{lm}\partial_j\Gamma_{ik}{}^m + \sum_{s=1}^{n}\Gamma_{ik}{}^s\Gamma_{sjl} = \partial_j\sum_{m=1}^{n} g_{lm}\Gamma_{ik}{}^m - \sum_{m=1}^{n}\Gamma_{ik}{}^m(\Gamma_{ljm} + \Gamma_{mjl}) + \sum_{m=1}^{n}\Gamma_{ik}{}^m\Gamma_{mjl}.$$

Using the cancellation between the last two sums we obtain

$$R_{lijk} = \partial_j \Gamma_{ikl} - \partial_k \Gamma_{ijl} + \sum_{m=1}^{n} (\Gamma_{ij}{}^m \Gamma_{lkm} - \Gamma_{ik}{}^m \Gamma_{ljm}). \tag{2.1.13$'$}$$

Again we note that the expression (2.1.13)$'$ shows that R_{ijkl} can be computed from the first fundamental form alone while (2.1.15) only involves the second fundamental form. From (2.1.15) we also see that the corresponding 4-linear form

$$R(t^1, t^2, t^3, t^4) = \sum_{i,j,k,l=1}^{n} R_{ijkl} t_i^1 t_j^2 t_k^3 t_l^4; \quad t^1, \dots, t^4 \in \mathbf{R}^n,$$

is antisymmetric in the pair t^1, t^2 and in the pair t^3, t^4 but symmetric for exchange of the pairs. Considered as a 4-linear form in $f'(x)t^1, \dots, f'(x)t^4 \in T_{f(x)}M$ the form R is independent of the choice of parametrization since this is true for the symmetric bilinear map $H : T_{f(x)}M \times T_{f(x)}M \to N_{f(x)}M$ defined by $\sum h_{ij} t_i^1 t_j^2$. In fact, for the corresponding form $\widetilde{R}$ we have by (2.1.15)

$$\widetilde{R}(t^1, t^2, t^3, t^4) = \big(H(t^1, t^3), H(t^2, t^4)\big) - \big(H(t^1, t^4), H(t^2, t^3)\big). \tag{2.1.15$'$}$$

Definition 2.1.5. The 4-linear form $\widetilde{R}$ on TM is called the Riemann curvature tensor of M.

The curvature tensor has symmetries in addition to those already mentioned. We shall discuss them in Section 2.3. However, already here we introduce the *Ricci tensor* which is obtained by contraction of the Riemann curvature tensor,

$$R_{jl} = \sum_{i,k=1}^{n} g^{ik} R_{ijkl} = \sum_{k=1}^{n} R^k{}_{jkl} = \sum_{k=1}^{n} R^k{}_{lkj}. \tag{2.1.16}$$

The corresponding bilinear form

$$\sum R_{jl} t_j^1 t_l^2$$

is symmetric and invariantly defined on the tangent space of M, for it is the trace of the linear transformation corresponding to the bilinear form $(t, s) \mapsto R(t, t^1, s, t^2)$ and the first fundamental form. Taking the trace once more one obtains the *scalar curvature*

$$S = \sum g^{jl} R_{jl}. \tag{2.1.17}$$

Exercise 2.1.2. Show that with the notation in Exercise 2.1.1

$$R_{ik} = \sum_{j=1}^{n} \partial_j \Gamma_{ik}{}^j - \tfrac{1}{2} \partial_i \partial_k \log g + \tfrac{1}{2} \sum_{l=1}^{n} \Gamma_{ik}{}^l \partial_l \log g - \sum_{l,j=1}^{n} \Gamma_{ij}{}^l \Gamma_{lk}{}^j.$$

2.2. The curvature of a hypersurface. To clarify the geometric meaning of the curvature tensor we shall devote this section to a discussion of hypersurfaces in $\mathbf{R}^N$. (The case $N = 3$ is the classical one of course.) Thus we assume now that $\dim V = N = n+1$. Let $n(x)$ denote one of the unit normals of M at $f(x)$. Then we can write

$$h_{ij}(x) = \tilde{h}_{ij}(x)n(x)$$

where $\tilde{h}_{ij}$ is now a scalar. It is clear that this scalar quadratic form corresponds to an invariantly defined quadratic form on the tangent space of M. The quotient $\sum \tilde{h}_{ij}t_it_j / \sum g_{ij}t_it_j$ is by Meusnier's theorem the curvature at $f(x)$ of a curve in M with tangent vector $f'(x)t$ and the principal normal in the direction of the normal $n(x)$ of the surface, that is, with vanishing geodesic curvature. To understand the quotient it is natural to diagonalize the forms simultaneously, thus introduce new coordinates $s_1, \dots, s_n$ such that

$$\sum_{i,j=1}^n \tilde{h}_{ij}t_it_j = \sum_{i=1}^n K_i s_i^2, \quad \sum_{i,j=1}^n g_{ij}t_it_j = \sum_{i=1}^n s_i^2.$$

This is always possible since $\sum g_{ij}t_it_j$ is positive definite. The eigenvalues K_i are the solutions of the equation

$$\det(\tilde{h}_{ij} - \lambda g_{ij}) = 0. \tag{2.2.1}$$

In particular, the symmetric functions are given by the coefficients in this equation,

$$\sum_{i=1}^n K_i = \sum_{i,j=1}^n \tilde{h}_{ij}g^{ij}, \quad \prod_{i=1}^n K_i = \det(\tilde{h}_{ij})/\det(g_{ij}). \tag{2.2.2}$$

Definition 2.2.1. For an oriented hypersurface the eigenvectors of the second fundamental form with respect to the first fundamental form, both regarded as quadratic forms on the tangent space T_xM of M at x, are called principal curvature directions at x. The corresponding eigenvalues are called principal curvatures. Their product and sum, given by (2.2.2), are called the total (or Gauss) curvature and the mean curvature respectively.

Exercise 2.2.1. Write down explicitly the formulas for the mean curvature and the total curvature of a surface of the form $x_{n+1} = \varphi(x_1, \dots, x_n)$ in $\mathbf{R}^{n+1}$, where $\varphi \in C^2$.

Since $(n, n) = 1$ we have $(n, \partial_i n) = 0$, so (2.1.12) gives

$$\partial_i n = -\sum_{j,k=1}^n \tilde{h}_{ij}g^{kj}f_k,$$

or equivalently

$$(\sum_{i=1}^n t_i\partial_i n, \sum_{k=1}^n s_kf_k) = -\sum_{i,j=1}^n \tilde{h}_{ij}t_is_j.$$

Thus the differential of the *Gauss map*

$$\gamma : M \ni f(x) \to n(x) \in S^n \tag{2.2.3}$$

is the linear transformation in $T_{f(x)}M$ corresponding to minus the second fundamental form and the first fundamental form:

THEOREM 2.2.2. *If γ is the Gauss map $M \to S^n$ where M is an oriented hypersurface and S^n is the unit sphere, then the differential $\gamma' : T_xM \to T_{\gamma(x)}S^n$, $x \in M$, can be regarded as a map $T_xM \to T_xM$ since T_xM and $T_{\gamma(x)}S^n$ are parallel. The principal curvature directions are then eigenvectors of $-\gamma'$ with the principal curvatures as eigenvalues.*

The last statement is a classical theorem of Olinde Rodrigues. In particular we see that γ is a local diffeomorphism precisely when the total curvature is not 0.

EXERCISE 2.2.2. Let M_0 be an open subset of M which is mapped diffeomorphically into S^n by γ, and let f be a continuous function with support in the range $\gamma(M_0) \subset S^n$. Show that

$$\int f\, dS = \int (f \circ \gamma)|K|\, dM$$

where dS and dM are the Euclidean volume elements of S^n and of M.

We shall now compute the Riemann curvature tensor of a hypersurface. If $e_1, \dots, e_n$ is an orthonormal system of principal curvature directions at a point $x \in M$, then

$$\tilde{h}(t) = \sum_{j=1}^{n} K_j (t, e_j)^2, \quad t \in T_x(M),$$

where K_j are the corresponding principal curvatures. For $t^1, \dots, t^4 \in T_x(M)$ we have

$$\begin{aligned}(t^1, e_i)(t^3, e_i)(t^2, e_j)(t^4, e_j) &- (t^1, e_i)(t^4, e_i)(t^2, e_j)(t^3, e_j) \\ + (t^1, e_j)(t^3, e_j)(t^2, e_i)(t^4, e_i) &- (t^1, e_j)(t^4, e_j)(t^2, e_i)(t^3, e_i) \\ &= \begin{vmatrix}(t^1, e_i) & (t^1, e_j) \\ (t^2, e_i) & (t^2, e_j)\end{vmatrix} \begin{vmatrix}(t^3, e_i) & (t^3, e_j) \\ (t^4, e_i) & (t^4, e_j)\end{vmatrix}.\end{aligned}$$

If we multiply by $\frac{1}{2}K_iK_j$ and sum, it follows that

$$\widetilde{R}(t^1, t^2, t^3, t^4) = \tfrac{1}{2} \sum_{i,j=1}^{n} K_iK_j \begin{vmatrix}(t^1, e_i) & (t^1, e_j) \\ (t^2, e_i) & (t^2, e_j)\end{vmatrix} \begin{vmatrix}(t^3, e_i) & (t^3, e_j) \\ (t^4, e_i) & (t^4, e_j)\end{vmatrix}. \tag{2.2.4}$$

We can of course equally well sum for $i < j$, omitting the factor 1/2. In the particular case where $n = 2$ we only have one term then and conclude that $\widetilde{R}(t^1, t^2, t^3, t^4)$ is the product of the total curvature $K = K_1K_2$ and the areas of the parallelograms spanned by t^1, t^2 and by t^3, t^4. Since R is determined by the first fundamental form, we have in particular proved:

Theorem 2.2.3 (Teorema egregium of Gauss). *For a surface in* $\mathbf{R}^3$ *the total curvature is determined by the first fundamental form; it equals* $R(t^1, t^2, t^1, t^2)$ *if* t^1, t^2 *are two orthogonal unit tangent vectors.*

Note that the mean curvature is not determined by the first fundamental form, for a plane and a cylinder have the same fundamental form with suitable coordinates but the mean curvatures are different. In Chapter III we shall return to the historical background of Gauss' discovery.

To compute the Ricci tensor, we take $i \neq j$ in (2.2.4), let $t^1 = t^3 = e_k$ and sum over k. The only contributions $\neq 0$ occur when $k = i$ or $k = j$, so the quadratic form corresponding to the symmetric Ricci tensor is

$$t \mapsto \sum_{i \neq j} K_i K_j (t, e_j)^2. \tag{2.2.5}$$

When $n = 2$ it is the Gauss curvature times the first fundamental form. For any n it has diagonal form, and the diagonal elements are $\sum_{j; j \neq i} K_i K_j$, for $i = 1, \dots, n$. This is the sum of the Gauss curvatures of the sections with three dimensional planes containing n, e_i and e_j for some $j \neq i$. The scalar curvature is

$$S = \sum_{i \neq j} K_i K_j = (\sum_1^n K_i)^2 - \sum K_i^2, \tag{2.2.6}$$

which is twice the Gauss curvature when $n = 2$.

2.3. Algebraic properties of the curvature tensor. Recall that the curvature tensor R_{ijkl} defined by (2.1.15) for a submanifold of $\mathbf{R}^N$ is antisymmetric in the pair ij and in the pair kl but symmetric for exchange of these pairs. This implies that

$$R(t^1, t^2, t^1, t^2)/(g(t^1, t^1) g(t^2, t^2) - g(t^1, t^2)^2)$$

only depends on the two plane spanned by the tangent vectors t^1 and t^2; the denominator is the square of the area of the parallelogram spanned by them. Because of Theorem 2.2.3 one calls this quotient the *sectional curvature* for the two plane in $T_x M$.

We shall now determine if there are additionals restrictions on the tensors which may occur at a point. This is a simple problem in linear algebra. Let $\sum h_{jk} x_j x_k$ be a quadratic form in $\mathbf{R}^n$ with coefficients $h_{jk} = h_{kj}$ in $\mathbf{R}^r$ where $r = N - n$, and define as in (2.1.15)

$$R_{ijkl} = (h_{ik}, h_{jl}) - (h_{il}, h_{jk}). \tag{2.3.1}$$

Since we put no condition on $r = N - n$ and the different coordinates in $\mathbf{R}^r$ give additive contributions, the set $\mathcal{T}$ of tensors R_{ijkl} which can occur is the linear space generated by (2.2.4) where we take only K_1 and K_2 different from 0. Let E be the vector space of all R_{ijkl}, $i, j, k, l = 1, \dots, n$ with the obvious symmetries

$$R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}. \tag{2.3.2}$$

E is a Euclidean space with the norm square $\sum R_{ijkl}^2$, so every linear form on E can be written

$$R \mapsto \sum R_{ijkl} S_{ijkl}$$

where $S \in E$. If the linear form vanishes on the term in (2.2.4) with $i = 1$, $j = 2$, then we have for the 4-linear form defined by S

$$S(e_1, e_2, e_1, e_2) = 0, \tag{2.3.3}$$

for (2.3.2) implies that the four terms in $\tilde{R}$ give the same contribution. Since

$$S(e_1, e_2 + \lambda e_1, e_1, e_2 + \lambda e_1) = S(e_1, e_2, e_1, e_2), \quad \lambda \in \mathbf{R},$$

we must have (2.3.3) for arbitrary $e_1, e_2 \in \mathbf{R}^n$, not necessarily orthogonal. Polarization gives

$$S(e_1, e_2, e_1, e_4) = 0, \quad e_1, e_2, e_4 \in \mathbf{R}^n, \tag{2.3.3$'$}$$

for this is a symmetric bilinear form in e_2 and e_4. Trying to polarize again we just obtain

$$S(e_1, e_2, e_3, e_4) + S(e_3, e_2, e_1, e_4) = 0,$$

or if we use the symmetries (2.3.2)

$$S(e_1, e_2, e_3, e_4) = S(e_1, e_4, e_2, e_3).$$

Thus $S(e_1, e_2, e_3, e_4)$ is invariant under circular permutations of e_2, e_3, e_4. By (2.3.2) an exchange of e_1 and e_2 or e_3 and e_4 changes the sign, so

$$S(e_1, e_2, e_3, e_4) = \operatorname{sgn} \pi S(e_{\pi(1)}, e_{\pi(2)}, e_{\pi(3)}, e_{\pi(4)})$$

if π is any permutation of $1, 2, 3, 4$. This means that S is an alternating 4-linear form (hence equal to 0 if $n \leq 3$). Conversely, every alternating form satisfies (2.3.3), so $\mathcal{T}$ is the orthogonal space in E of the alternating forms. These are spanned by the exterior products

$$(e_1, \dots, e_4) \mapsto \det(t^i, e_j)_{i,j=1}^4 = \sum_\pi \operatorname{sgn} \pi \prod_1^4 (t^{\pi(i)}, e_i)$$

where the sum is taken over all permutations of $1, \dots, 4$ and $t^1, \dots, t^4 \in \mathbf{R}^n$. Thus $R \in \mathcal{T}$ if and only if $R \in E$ and

$$\sum_\pi \operatorname{sgn} \pi R(t^{\pi(1)}, \dots, t^{\pi(4)}) = 0.$$

Because of the symmetries (2.3.2) we can rearrange the arguments so that t^1 stands first and the permutation is positive, so the condition is equivalent to

$$R(t^1, t^2, t^3, t^4) + R(t^1, t^4, t^2, t^3) + R(t^1, t^3, t^4, t^2) = 0. \tag{2.3.4}$$

This is called the *first Bianchi identity.* It is often written $R_{i[jkl]} = 0$ where the bracket is read as summation over all circular permutations. The symmetry conditions are not independent of each other. The condition on R_{ijlk} in (2.3.2) is obviously a consequence of the others, and the condition on R_{klij} follows from the first two equalities in (2.3.2) and the Bianchi identity. In fact, they give

$$\begin{aligned}
&R(t^1,t^2,t^3,t^4)\\
&\quad = -R(t^1,t^4,t^2,t^3) - R(t^1,t^3,t^4,t^2) = R(t^4,t^1,t^2,t^3) + R(t^3,t^1,t^4,t^2)\\
&\quad = -R(t^4,t^3,t^1,t^2) - R(t^4,t^2,t^3,t^1) - R(t^3,t^2,t^1,t^4) - R(t^3,t^4,t^2,t^1)\\
&\qquad = 2R(t^3,t^4,t^1,t^2) + R(t^2,t^4,t^3,t^1) + R(t^2,t^3,t^1,t^4)\\
&\qquad = 2R(t^3,t^4,t^1,t^2) - R(t^2,t^1,t^4,t^3) = 2R(t^3,t^4,t^1,t^2) - R(t^1,t^2,t^3,t^4)
\end{aligned}$$

which proves the remaining symmetry in (2.3.3).

For a submanifold M of $\mathbf{R}^{n+r}$ with dimension n the Gauss curvature at $X \in M$ of the intersection with a plane of dimension $2+r$ containing the normal $N_X M$ and two tangent vectors t^1, t^2 is by Theorem 2.2.3 equal to $R(t^1,t^2,t^1,t^2)$ divided by the square of the area of the parallelogram spanned by t^1 and t^2. Now a tensor $R \in \mathcal{T}$ is uniquely determined by such curvatures, for

$$R \in \mathcal{T}, \quad R(t^1,t^2,t^1,t^2) = 0, \text{ if } t^1,t^2 \in \mathbf{R}^n \implies R = 0.$$

In fact, the condition on R here is the condition (2.3.3), which means that R is orthogonal to $\mathcal{T}$, hence equal to 0.

A symmetric bilinear form in ν variables has $\nu(\nu+1)/2$ independent coefficients whereas an antisymmetric bilinear form has $\binom{\nu}{2} = \nu(\nu-1)/2$ independent coefficients. Since we can interpret E as a space of symmetric bilinear forms on the space $\mathbf{R}^n \wedge \mathbf{R}^n$ which has dimension $\nu = n(n-1)/2$, it is clear that $\dim E = \nu(\nu+1)/2$. The orthogonal space of $\mathcal{T}$ in E has dimension $\binom{n}{4}$, so it follows that

$$\begin{aligned}
\dim \mathcal{T} &= \nu(\nu+1)/2 - \nu(n-2)(n-3)/12\\
&= \nu(3n(n-1)+6-n^2+5n-6)/12 = n^2(n^2-1)/12.
\end{aligned}$$

We sum up the results proved so far in a theorem:

THEOREM 2.3.1. *A 4-linear form R in $\mathbf{R}^n$ can occur as the Riemann curvature tensor for some n dimensional submanifold of a Euclidean space if and only if it has the symmetry properties* (2.3.2) *and satisfies the Bianchi identity* (2.3.4). *Such tensors form a linear space $\mathcal{T}$ of dimension $n^2(n^2-1)/12$. When $R \in \mathcal{T}$ then R is uniquely determined by $R(t^1,t^2,t^1,t^2)$ for $t^1,t^2 \in \mathbf{R}^n$.*

It is clear that one can find r depending only on n such that for every $R \in \mathcal{T}$ there is a solution h of (2.3.1) with values in $\mathbf{R}^r$. The dimension of the space of such quadratic forms h is $rn(n+1)/2$, so it follows from the Morse-Sard theorem that we must have $rn(n+1)/2 \ge n^2(n^2-1)/12$, that is,

$$r \ge n(n-1)/6. \tag{2.3.5}$$

EXERCISE 2.3.1. Use the fact that (2.3.1) is invariant if h is replaced by Oh where $O \in \mathbf{O}(r)$ to show that (2.3.5) can be strengthened to

$$n(n+1)s - s(s-1) \geq n^2(n^2-1)/6 \quad \text{for some } s \in \{1, 2, \dots, r\},$$

which means that $r \geq \big((1 - 1/\sqrt{3})(n^2+1) + (1-\sqrt{3})n + O(1/n)\big)/2$.

Classical results to be discussed in Chapter III will prove that one can always take $r = n(n-1)/2$. Using representation theory Berger, Bryant and Griffiths [1] have proved that one can always take $r = \binom{n-1}{2} + 2$, but the best value of r does not seem to be known. Such a value is an obvious lower bound for the codimension with which a general Riemannian manifold can be locally C^2 embedded. In Chapter III we shall see that a local embedding with high regularity is usually not possible with codimension lower than $\binom{n}{2}$.

The full linear group $\mathbf{GL}(n)$ acts on $\mathcal{T}$ by

$$(gR)(t^1, \dots, t^4) = R(g^{-1}t^1, \dots, g^{-1}t^4), \quad g \in \mathbf{GL}(n),\ R \in \mathcal{T},\ t^j \in \mathbf{R}^n.$$

There is no invariant subspace for this operation (see Berger, Bryant and Griffiths [1]). However, in the context where we encountered the Riemann curvature tensor only the operation of $\mathbf{O}(n)$ is natural, since it corresponds to changing orthonormal basis in the tangent space. At the end of Section 2.1 we also defined the Ricci tensor, which belongs to the space $S^2(\mathbf{R}^n)$ of symmetric bilinear forms in $\mathbf{R}^n$. The passage from the Riemann tensor to the Ricci tensor commutes with the operation of $\mathbf{O}(n)$, so the kernel $\mathcal{W}$ consisting of all $R \in \mathcal{T}$ such that

$$\sum_{i=1}^{n} R_{ijil} = 0, \quad j, l = 1, \dots, n, \tag{2.3.6}$$

is invariant under $\mathbf{O}(n)$. (In view of the symmetries (2.3.2) the contraction of R with respect to any pair of indices is equal to 0 if $R \in \mathcal{W}$.) Since the Euclidean norm we introduced in E above is invariant under $\mathbf{O}(n)$, the orthogonal complement of $\mathcal{W}$ in $\mathcal{T}$ is also invariant under $\mathbf{O}(n)$. We shall determine the decomposition of a general $R \in \mathcal{T}$ in components belonging to these spaces, but first we shall make some elementary remarks on how $\mathbf{O}(n)$ acts on $S^2(\mathbf{R}^n)$.

The metric scalar product $(\cdot, \cdot)$ in $\mathbf{R}^n$ is of course invariant under $\mathbf{O}(n)$, and so is its orthogonal space consisting of symmetric forms which are traceless (with respect to the Euclidean form). If a $\mathbf{O}(n)$ invariant subspace of $S^2(\mathbf{R}^n)$ contains a form which is not proportional to the metric form, then it contains a form $\sum \lambda_j x_j y_j$ with $\lambda_1 \neq \lambda_2$. Hence it also contains the form with λ_1 and λ_2 interchanged and therefore the form $x_1y_1 - x_2y_2$ and so the form $x_jy_j - x_ky_k$ for arbitrary j and k. The linear hull contains all diagonal forms with zero trace, so the only $\mathbf{O}(n)$ invariant subspaces of $S^2(\mathbf{R}^n)$ are the multiples of the metric form and the forms with zero trace. If $n \geq 3$, as we assume now, it follows at once from (2.2.5) that both the metric form and other forms can occur as Ricci tensors, so the map $\mathcal{T}(\mathbf{R}^n) \to S^2(\mathbf{R}^n)$ corresponding to passage from the Riemann tensor to the Ricci tensor is surjective. (It is easy to show that any symmetric tensor can in fact occur as Ricci tensor for a manifold with codimension $r = 2$.)

From the surjectivity just proved it follows that the orthogonal space of $\mathcal{W}$ in $\mathcal{T}$ is of dimension $n(n+1)/2$. If $R \in \mathcal{T}$ and $S = \sum R_{ijij}$ is the scalar curvature while $B_{ij} = R_{ij} - S\delta_{ij}/n$ is the *traceless Ricci tensor*, then we claim that

$$(2.3.7) \quad \begin{aligned} R_{ijkl} &= W_{ijkl} \\ &+ (B_{ik}\delta_{jl} - B_{il}\delta_{jk} + B_{jl}\delta_{ik} - B_{jk}\delta_{il})/(n-2) + S(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})/n(n-1) \end{aligned}$$

decomposes R in an element in $\mathcal{W}$ and two orthogonal ones, one which has traceless Ricci tensor and one which has Ricci tensor proportional to the metric tensor. To verify this it suffices to show that the last two terms are in $\mathcal{T}$, for they are obviously orthogonal to $\mathcal{W}$ and the contractions with respect to i, k are

$$(0 - B_{jl} + nB_{jl} - B_{jl})/(n-2) = B_{jl} \text{ resp. } S(n\delta_{jl} - \delta_{jl})/n(n-1) = S\delta_{jl}/n.$$

The symmetries (2.3.2) are obviously valid, and the Bianchi condition is then equivalent to orthogonality to all antisymmetric 4-linear forms, which is equally obvious since every term is symmetric in some pair of indices.

DEFINITION 2.3.2. The component $W \in \mathcal{W}$ of $R \in \mathcal{T}$ defined by (2.3.7) is called the Weyl tensor or the conformal curvature tensor. We have

$$(2.3.7)' \quad \begin{aligned} R_{ijkl} &= W_{ijkl} \\ &+ (R_{ik}\delta_{jl} - R_{il}\delta_{jk} + R_{jl}\delta_{ik} - R_{jk}\delta_{il})/(n-2) - S(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})/(n-1)(n-2) \end{aligned}$$

The reason for the terminology will be clear in Chapter III, but some motivation will be provided already by the following

EXERCISE 2.3.2. Calculate the Riemann curvature tensor of a sphere of radius R in $\mathbf{R}^{n+1}$ and decompose it according to (2.3.7).

We have proved that $\mathbf{O}(n)$ acts on $\mathcal{T} \ominus \mathcal{W}$ with just two invariant subspaces, displayed in the decomposition (2.3.7), but the proof that $\mathcal{W}$ has no $\mathbf{O}(n)$ invariant subspace requires a bit of invariant theory. We must therefore refer the reader to the proof given in Berger, Gauduchon and Mazet [1, pp. 76–78].

EXERCISE 2.3.3. Prove that

$$\sum_{i,j,k,l} R_{ijkl} R_{iklj} = -\tfrac{1}{2} \sum_{i,j,k,l} {R_{ijkl}}^2.$$

When $n = 3$ then $n(n+1)/2 = 6 = n^2(n^2-1)/12$ so there is no Weyl tensor. When $n > 4$ it is known that $\mathcal{W}$ has no $\mathbf{SO}(n)$ invariant subspace (see references in Besse [1, p. 49]). However, this is not true when $n = 4$, and since further decomposition of $\mathcal{W}$ is then relevant in Yang-Mills theory, we shall discuss it briefly. (For details and references we refer to Atiyah, Hitchin and Singer [1], Besse [1].) Since the 4-linear forms R with the symmetries (2.3.2) can be identified with symmetric bilinear forms on $\wedge^2(\mathbf{R}^n)$, we shall first discuss how $\mathbf{SO}(n)$ acts on $\wedge^2(\mathbf{R}^n)$.

LEMMA 2.3.3. *$\wedge^2(\mathbf{R}^n)$ has no proper invariant subspace under the action of $\mathbf{SO}(n)$ if $n \neq 4$, but $\wedge^2(\mathbf{R}^4) = \wedge_+ \oplus \wedge_-$ where $\wedge_\pm$ are irreducible $\mathbf{SO}(4)$ invariant subspaces of dimension 3.*

PROOF. If S is a skew symmetric $n \times n$ matrix, then $\mathbf{R}^n = V_0 \oplus V_1 \oplus \cdots \oplus V_j$ where $SV_0 = 0$ and V_i is of dimension 2 and $SV_i = V_i$ when $i = 1, \dots, j$. (See Exercise 1.4.2.) This means that every element in $\wedge^2(\mathbf{R}^n)$ can be transformed by the action of $\mathbf{SO}(n)$ to the form

$$c_1 e_1 \wedge e_2 + \cdots + c_j e_{2j-1} \wedge e_{2j}, \tag{2.3.8}$$

where $e_1, \dots, e_n$ is the standard basis in $\mathbf{R}^n$ and $2j \leq n$. Suppose now that $\wedge$ is some $\mathbf{SO}(n)$ invariant subspace $\neq \{0\}$ of $\wedge^2(\mathbf{R}^n)$, and let (2.3.8) be an element in $\wedge$ with $c_1 \dots c_j \neq 0$. If $j = 1$, then $\wedge$ will contain every element of the form (2.3.8) so $\wedge = \wedge^2(\mathbf{R}^n)$ then. If $2j < n$, then $e_1 \to -e_1$, $e_n \to -e_n$ defines an element in $\mathbf{SO}(n)$ which just replaces c_1 by $-c_1$, so $e_1 \wedge e_2 \in \wedge$ and $\wedge = \wedge^2(\mathbf{R}^n)$ then. Assume now that $2j = n$. Then $e_1 \leftrightarrow e_3$, $e_2 \leftrightarrow e_4$ defines an element in $\mathbf{SO}(n)$ which interchanges c_1 and c_2; by subtraction we find that $e_1 \wedge e_2 - e_3 \wedge e_4 \in \wedge$ unless $c_1 = c_2$. The lemma is now proved if $n > 4$ unless $n = 2j$ and $c_1 = c_2 = \cdots = c_j$. Since $e_1 \to e_3 \to e_5 \to e_1$ defines an element in $\mathbf{SO}(n)$ we can then conclude by subtraction that

$$(e_1 - e_3) \wedge e_2 + (e_3 - e_5) \wedge e_4 + (e_5 - e_1) \wedge e_6 \in \wedge.$$

Since $e_1 - e_3 + e_3 - e_5 + e_5 - e_1 = 0$, the rank of this two vector is just 4 and the lemma follows also in this case.

We are now just left with the case $n = 4$, and then we know that $\wedge$ must contain either $e_1 \wedge e_2 + e_3 \wedge e_4$ or $e_1 \wedge e_2 - e_3 \wedge e_4 = e_1 \wedge e_2 + e_4 \wedge e_3$. In the first case $\wedge$ must contain the space $\wedge_+$ spanned by $e_i \wedge e_j + e_k \wedge e_l$ where i, j, k, l is an even permutation of $1, 2, 3, 4$, and in the other case $\wedge$ contains the space $\wedge_-$ defined using the odd permutations. If $\varphi_\pm \in \wedge_\pm$ then

$$\varphi_+ \wedge \varphi_+ = (\varphi_+, \varphi_+)\omega, \quad \varphi_- \wedge \varphi_- = -(\varphi_-, \varphi_-)\omega, \quad \varphi_+ \wedge \varphi_- = 0, \quad (\varphi_+, \varphi_-) = 0,$$

where $\omega = e_1 \wedge e_2 \wedge e_3 \wedge e_4$ and $(\cdot, \cdot)$ is the Euclidean scalar product of two vectors. Thus $W_\pm$ are the eigenspaces with eigenvalues ± 1 of the quadratic form $\varphi \wedge \varphi / \omega$ with respect to (φ, φ), which proves the invariance under $\mathbf{SO}(n)$, since ω is invariant.

EXERCISE 2.3.4. Let $n = 4$ and take

$$(e_1 \wedge e_2 \pm e_3 \wedge e_4)/\sqrt{2}, \quad (e_1 \wedge e_3 \pm e_4 \wedge e_2)/\sqrt{2}, \quad (e_1 \wedge e_4 \pm e_2 \wedge e_3)/\sqrt{2}$$

as basis in $\wedge_\pm$. With this basis the $\mathbf{SO}(4)$ action on $\wedge^2(\mathbf{R}^4)$ gives a homomorphism $\mathbf{SO}(4) \to \mathbf{SO}(3) \times \mathbf{SO}(3)$. Show that the corresponding homomorphism of Lie algebras $\mathfrak{so}(4) \to \mathfrak{so}(3) \oplus \mathfrak{so}(3)$ is an isomorphism and determine it explicitly in terms of the corresponding 4×4 and 3×3 skew symmetric matrices.

REMARK. The decomposition when $n = 4$ will perhaps be more evident when we have introduced the $*$ operator on forms later on.

Still with $n = 4$ we now interpret the space E of 4-linear forms on $\mathbf{R}^4$ satisfying (2.3.2) as the space of symmetric bilinear forms on $\wedge^2(\mathbf{R}^4) = \wedge_+ \oplus \wedge_-$. This space has a $\mathbf{SO}(4)$ invariant block matrix decomposition

$$S^2(\wedge_+) \oplus S^2(\wedge_-) \oplus \mathrm{Hom}(\wedge_+, \wedge_-),$$

with spaces of dimensions 6, 6 and 9. However, since we have an $\mathbf{SO}(3)$ action on the first two spaces they decompose further into the multiples $I(\wedge_\pm)$ of the metric form and the tracless forms $S_0^2(\wedge_\pm)$, so we have in fact an $\mathbf{SO}(4)$ invariant decomposition

$$S_0^2(\wedge_+) \oplus I(\wedge_+) \oplus S_0^2(\wedge_-) \oplus I(\wedge_-) \oplus \mathrm{Hom}(\wedge_+, \wedge_-),$$

where the spaces now have dimensions 5, 1, 5, 1, 9. Since the space of Weyl tensors $\mathcal{W}$ has dimension 10, and the space of traceless Ricci tensors has dimension 9, it is now clear that $\mathrm{Hom}(\wedge_+, \wedge_-)$ is precisely the space of traceless Ricci tensors while $\mathcal{W} = \mathcal{W}_+ \oplus \mathcal{W}_-$ with $\mathcal{W}_\pm = S_0^2(\wedge_\pm)$. The purely antisymmetric tensor and the pure scalar curvature part $S(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})/12$ are found in the two invariant one dimensional spaces; the antisymmetric tensors are generated by the differences of the two metric tensors whereas the scalar curvature term corresponds to their sum. Thus a general curvature tensor in $\mathcal{T}$ is represented in block matrix form by

$$\begin{pmatrix} W_+ + SI/12 & B \\ {}^tB & W_- + SI/12 \end{pmatrix}$$

where B corresponds to the traceless Ricci tensor and the traceless symmetric $W_\pm$ together define the Weyl tensor. We leave the details of the verification for the energetic reader.

The $\mathbf{SO}(n)$ invariant classification of the curvature tensors gives rise to important classes of Riemannian manifolds:

Definition 2.3.4. A Riemannian manifold is called an Einstein manifold if the traceless Ricci tensor vanishes, and it is called conformally flat if the Weyl tensor vanishes. An oriented manifold of dimension 4 is called self-dual if the Weyl tensor is in $\mathcal{W}_+$ and anti-self-dual if it is in $\mathcal{W}_-$.

CHAPTER III

ABSTRACT RIEMANNIAN MANIFOLDS

Summary. In Section 3.1 we extend the intrinsic results proved in Chapter II to abstract Riemannian manifolds, taking the notion of covariant differentiation as the main tool. After introducing geodesic coordinates we also discuss various geometrical interpretations of curvature and prove the classical Gauss-Bonnet theorem. Section 3.2 is devoted to embedding theorems which show that abstract Riemannian manifolds are not really more general than the submanifolds of a Euclidean space studied in Chapter II. In Section 3.3 we show that a Riemannian manifold with vanishing curvature tensor is flat, that is, locally isomorphic to $\mathbf{R}^n$; more generally we also discuss manifolds with constant curvature. Section 3.4 is devoted to the transformation rules for the curvature under conformal changes of metric. For dimensions ≥ 4 it is proved that vanishing of the Weyl tensor is necessary and sufficient for the existence of a flat conformal metric. The Yamabe problem to find conformal metrics of constant scalar curvature is also mentioned briefly, but we leave the study of it for another chapter.

3.1. Covariant derivatives and curvature. In Chapter II we studied n dimensional C^k submanifolds M of some $\mathbf{R}^N$. Locally they were parametrized by maps $\omega \ni x \mapsto \kappa(x) \in M$ where ω is an open set in $\mathbf{R}^n$ and $\kappa \in C^k$ is injective with injective differential. If $\tilde{\omega} \ni \tilde{x} \mapsto \tilde{\kappa}(\tilde{x}) \in M$ is another local parametrization, it follows that the map

$$x \mapsto \tilde{x} = \tilde{\kappa}^{-1} \circ \kappa(x)$$

is defined in the open set $\kappa^{-1}(\kappa(\omega) \cap \tilde{\kappa}(\tilde{\omega})) \subset \omega$ which is in C^k since $\tilde{\kappa}$ has a C^k left inverse.

An abstract C^k manifold M of dimension n is by definition a Hausdorff topological space provided with a family of homeomorphisms, $\kappa : \omega \to f(\omega) \subset M$ with open range, such that $\tilde{\kappa}^{-1} \circ \kappa \in C^k(\kappa^{-1}(\kappa(\omega) \cap \tilde{\kappa}(\tilde{\omega}))$ if $\tilde{\kappa}$ is another member of the family. We shall always assume that M has a countable dense subset. Then it is well known that M can be embedded as a submanifold of $\mathbf{R}^{2n+1}$, so the difference between submanifolds of $\mathbf{R}^N$ and abstract manifolds is more one of principle and attitude than of substance.

The tangent bundle $T(M)$ was defined for a submanifold of $\mathbf{R}^N$ as the set of all (x, t) with $x \in M$ and t in the tangent space $T_x(M)$ of M in $\mathbf{R}^N$. It has a natural generalization for an abstract manifold. One just takes $T_x(M)$ to be the n dimensional vector space of first order differential operators at x which annihilate constants. In a local coordinate patch $\omega \subset \mathbf{R}^n$ with local coordinates x, the first order operators can be written $\sum_1^n t_j \partial/\partial x^j$, which identifies $T(M)$ with $\omega \times \mathbf{R}^n$ over ω. (From now on we shall use superscripts for the coordinates to conform with the conventions of tensor calculus.) If one switches to coordinates $\tilde{x}$ as above, then (x, t) is identified with $(\tilde{x}, \tilde{t})$ if $\kappa(x) = \tilde{\kappa}(\tilde{x})$ and $(\tilde{\kappa}^{-1}\kappa)'t = \tilde{t}$, which means that $\tilde{t} = (\partial \tilde{x}/\partial x)t$.

When studying submanifolds of $\mathbf{R}^N$ we took from the Euclidean metric in $\mathbf{R}^N$ the first fundamental form which we wrote

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}(x) dx^i dx^j \tag{3.1.1}$$

in terms of local coordinates. A manifold is said to be Riemannian if it is provided with such a form:

Definition 3.1.1. A C^k manifold, where $k \geq 1$, is said to be Riemannian if there is given in each fiber T_xM of TM a positive definite quadratic form such that in local coordinates it has the form (3.1.1) with $g_{ij} \in C^{k-1}$.

Shifting to other local coordinates $\tilde{x}$ gives

$$\sum \tilde{g}_{ij}(\tilde{x}) d\tilde{x}^i d\tilde{x}^j = \sum \tilde{g}_{ij}(\tilde{x})(\partial \tilde{x}^i/\partial x^\nu) dx^\nu (\partial \tilde{x}^j/\partial x^\mu) dx^\mu,$$

or in matrix notation $(g_{ij}(x)) = {}^t(\partial\tilde{x}/\partial x)(\tilde{g}_{ij}(\tilde{x}))(\partial\tilde{x}/\partial x)$, so it is sufficient to assume that the coefficients in (3.1.1) are in C^{k-1} for a set of coordinate patches covering M. We note in passing that

$$\det(g_{ij}(x)) = (\det \partial\tilde{x}/\partial x)^2 \det(\tilde{g}_{ij}(\tilde{x})),$$

which means that $\sqrt{\det(g_{ij}(x))}\, dx$, where dx is Lebesgue measure in $\mathbf{R}^n$, is an invariant definition of a positive measure (positive C^{k-1} density) on M, often denoted by $dvol(x)$.

We shall now show that the results in Chapter II which are intrinsic in the sense that they can be expressed in terms of the first fundamental form alone remain valid for abstract Riemannian manifolds. This could be done by proving that every abstract Riemannian manifold can at least locally or to a high degree of approximation be embedded isometrically in $\mathbf{R}^N$ for some N. We shall see in Section 3.2 that this is possible. However, we shall now give a direct approach which only relies on Chapter II for motivation and not for any proofs. This means that we shall establish invariance under changes of coordinates even when we know from Chapter II that this is true for submanifolds of $\mathbf{R}^N$. It would be tedious to mention precise differentiability assumptions, so for the time being we assume that M is a C^∞ Riemannian manifold.

Let X and Y be two smooth vector fields, in local coordinates represented by

$$X = \sum_1^n X^j \partial/\partial x^j, \quad Y = \sum_1^n Y^j \partial/\partial x^j.$$

(2.1.11) suggests that another vectorfield, the *covariant derivative* of Y in the direction X, should be invariantly defined by

$$\nabla_X Y = \sum_{j=1}^{n} (XY^j + \sum_{i,k=1}^{n} \Gamma_{ik}{}^j X^k Y^i) \partial/\partial x^j, \tag{3.1.2}$$

where $\Gamma_{ik}{}^j$ is defined in terms of the metric by the last expression in (2.1.8), (2.1.7). We could verify by direct computation that this definition is indeed independent of the choice of local coordinates, but instead we shall prove that in a local coordinate patch the linear differential operator defined by (3.1.2) has the properties

$$\nabla_X Y - \nabla_Y X = [X, Y], \tag{3.1.3}$$

$$(\nabla_Z X, Y) + (X, \nabla_Z Y) = Z(X, Y), \tag{3.1.4}$$

where Z is a third vector field. From (3.1.3) it follows that $\nabla_X Y$ is linear in X as well as in Y. We shall also prove that ∇_X is uniquely determined by (3.1.3), (3.1.4), which implies the invariance. The proof of (3.1.3), where $[X, Y] = XY - YX$ denotes the commutator of the vector fields, follows immediately from the fact that $\Gamma_{ik}{}^j$, defined by (2.1.7), (2.1.8), is symmetric in i, k. To prove (3.1.4) we write $Z = \sum Z^j \partial/\partial x^j$ and obtain

$$\begin{aligned}(\nabla_Z X, Y) = \sum_{j,l=1}^n (ZX^j + \sum_{i,k=1}^n \Gamma_{ik}{}^j Z^k X^i) g_{jl} Y^l \\ = \sum_{j,l=1}^n g_{jl}(ZX^j)Y^l + \sum_{i,k,l=1}^n Z^k \Gamma_{ikl} X^i Y^l.\end{aligned}$$

If we exchange X and Y and add using (2.1.7)$'$, it follows that (3.1.4) holds. To prove the uniqueness means proving that if the right hand sides of (3.1.3), (3.1.4) are replaced by 0, then $\nabla_X Y$ must be 0 for arbitrary X and Y. This follows since

$$\begin{aligned}(\nabla_X Y, Z) = -(Y, \nabla_X Z) = -(Y, \nabla_Z X) \\ = (\nabla_Z Y, X) = (\nabla_Y Z, X) = -(Z, \nabla_Y X) = -(Z, \nabla_X Y)\end{aligned}$$

where we have alternated using the homogeneous forms of (3.1.3), (3.1.4). This implies $\nabla_X Y = 0$ as claimed. Thus we have proved:

Theorem 3.1.2. *For smooth vector fields X, Y on M the vector field defined in local coordinates by* (3.1.2) *is invariantly defined and the covariant differentiation ∇ is characterized by* (3.1.3), (3.1.4).

Note that $\nabla_X Y(x)$ is defined even if Y is only defined on a curve with tangent X at x. Taking for X and Y the unit tangent of a curve we can define the *geodesic curvature and geodesic principal normal* as the norm and the direction of $\nabla_X Y$. This generalizes the definition given after Theorem 2.1.3 for submanifolds of $\mathbf{R}^N$.

In the Euclidean case the differential operators ∇_X and ∇_Y act componentwise as the scalar operators X and Y, so their commutator is $\nabla_{[X,Y]}$. We shall now examine to what extent this is true in the Riemannian case, so we form with three vector fields X, Y, Z

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

If X is replaced by the product φX with some smooth function φ, then S is just multiplied by φ, for since $[\varphi X, Y] = \varphi[X, Y] - (Y\varphi)X$ and

$$\nabla_{\varphi X} = \varphi\nabla_X, \quad \nabla_Y(\varphi W) = \varphi\nabla_Y W + (Y\varphi)W, \tag{3.1.5}$$

the other terms $-(Y\varphi)\nabla_X Z + (Y\varphi)\nabla_X Z$ will cancel. Thus S contains no derivatives of X, hence no derivatives of Y by the skew symmetry. This is also true for Z since

$$\begin{aligned}&\nabla_X(\varphi\nabla_Y Z + (Y\varphi)Z) - \nabla_Y(\varphi\nabla_X Z + (X\varphi)Z)\\ &= \varphi(\nabla_X\nabla_Y - \nabla_Y\nabla_X)Z + (X\varphi)(\nabla_Y Z - \nabla_Y Z) + (Y\varphi)(\nabla_X Z - \nabla_X Z) + ([X, Y]\varphi)Z.\end{aligned}$$

We can therefore calculate S using local coordinates and assuming that the components of X, Y, Z are constants, hence $[X, Y] = 0$. Then

$$\nabla_X\nabla_Y Z = \nabla_X \sum_{i,j,l} \Gamma_{jl}{}^i Y^l Z^j \partial/\partial x^i = \sum_{ijkl}\Big(\partial_k \Gamma_{jl}{}^i + \sum_\nu \Gamma_{k\nu}{}^i\Gamma_{jl}{}^\nu\Big)X^k Y^l Z^j \partial/\partial x^i.$$

If we subtract the analogous formula for $\nabla_Y\nabla_X Z$ and recall the first definition (2.1.13) of the Riemann curvature tensor, it follows that

$$(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})Z = \sum_{ijkl} R^i{}_{jkl} X^k Y^l Z^j \partial/\partial x^i. \tag{3.1.6}$$

In particular, this proves that the definition of the Riemann curvature tensor given by (2.1.13) is coordinate independent and does define a tensor of type $1, 3$. (2.1.13)$'$ follows as before.

If $[0, a] \ni t \mapsto \gamma(t) \in M$ is a (piecewise) C^1 curve and $X_0 \in T_{\gamma(0)}M$, then there is a unique vector field $X(t) \in T_{\gamma(t)}M$, $t \in [0, a]$, along the curve with $X(0) = X_0$ and $\nabla_{\gamma'(t)}X(t) = 0$. In fact, in local coordinates this is the Cauchy problem for a linear system of differential equations

$$\frac{dX^j(t)}{dt} + \sum_{i,k}\Gamma_{ik}{}^j(\gamma(t))\frac{d\gamma^k(t)}{dt}X^i(t) = 0, \quad j = 1, \dots, n.$$

One calls $X(t)$ the *parallel translation* of X_0 along the curve.

Exercise 3.1.1. Show that in the local coordinates the parallel translation of X_0 from x to $x + \varepsilon e_k$, where e_k is the kth basis vector in $\mathbf{R}^n$, is

$$X_0 + (A(\varepsilon) + \tfrac{1}{2}A(\varepsilon)^2 + O(\varepsilon^3))X_0, \quad \text{as } \varepsilon \to 0; \quad A(\varepsilon)_{ji} = -\varepsilon\Gamma_{ik}{}^j(x + \tfrac{1}{2}\varepsilon e_k).$$

Use this to find the limit as $\varepsilon \to 0$ of $(\widetilde{X}_0(\varepsilon) - X_0)/\varepsilon^2$ if $\widetilde{X}_0$ is the parallel translation of X_0 around the square from x to $x + \varepsilon e_k$ to $x + \varepsilon(e_k + e_l)$ to $x + \varepsilon e_l$ to x.

Covariant differentiation can be extended to arbitrary tensors. Let first ω be a one form, that is, a section of the cotangent bundle $T^*(M)$. To ω corresponds a vector field Y, for the metric tensor identifies $T^*_x(M)$ and $T_x(M)$ for every x; in local coordinates

$$Y^k = \sum_1^n g^{kj}\omega_j, \quad \omega_j = \sum_1^n g_{jk}Y^k, \text{ if } \omega = \sum \omega_j dx_j,\ Y = \sum Y^k \partial/\partial x^k.$$

Sometimes the notation $Y = \omega^\sharp$ and $\omega = Y^\flat$ is used for the "musical" isomorphisms $T^*_x \to T_x \to T^*_x$ raising and lowering indices. With that notation we now have an obvious invariant definition of covariant differentiation of one forms ω,

$$\nabla_X \omega = (\nabla_X \omega^\sharp)^\flat.$$

By (3.1.4) we obtain for any vector field Y

$$X(\omega^\sharp, Y) = (\nabla_X \omega^\sharp, Y) + (\omega^\sharp, \nabla_X Y),$$

or if we use the notation $\langle \cdot, \cdot \rangle$ for the duality between T^* and T,

$$X\langle \omega, Y\rangle = \langle \nabla_X \omega, Y\rangle + \langle \omega, \nabla_X Y\rangle. \tag{3.1.7}$$

This could also have been taken as a definition of $\nabla_X \omega$. In local coordinates (3.1.7) means that

$$\begin{aligned}\langle \nabla_X \omega, Y\rangle = X\sum_{j=1}^n \omega_j Y^j - \sum_{j=1}^n \omega_j (XY^j + \sum_{k,l=1}^n \Gamma_{kl}{}^j X^k Y^l)\\ = \sum_{j=1}^n (X\omega_j)Y^j - \sum_{j,k,l=1}^n \omega_l \Gamma^l_{kj} X^k Y^j\end{aligned}$$

so we obtain

$$(\nabla_X \omega)_j = X\omega_j - \sum_{k,l=1}^n \Gamma_{kj}{}^l X^k \omega_l. \tag{3.1.2$'$}$$

Exercise 3.1.2. Let $\omega = \sum \omega_j dx^j$ be a one form. Show that for vector fields X, Y we have

$$(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})\omega = -\sum R^j{}_{ikl} X^k Y^l \omega_j dx^i.$$

Exercise 3.1.3. Show that for tensor fields f of type k, l, that is, sections of

$$\underbrace{T(M) \otimes \cdots \otimes T(M)}_{k \text{ times}} \otimes \underbrace{T^*(M) \otimes \cdots \otimes T^*(M)}_{l \text{ times}}$$

and a vector field X the formula in local coordinates

$$(3.1.2)'' \quad (\nabla_X f)^{i_1\dots i_k}_{j_1\dots j_l} = X f^{i_1\dots i_k}_{j_1\dots j_l} + \sum_{\nu=1}^{k}\sum_{\mu,j=1}^{n} \Gamma_{\mu j}{}^{i_\nu} f^{i_1\dots i_{\nu-1}\mu i_{\nu+1}\dots i_k}_{j_1\dots j_l} X^j \\ - \sum_{\nu=1}^{l}\sum_{\mu,j=1}^{n} \Gamma_{j_\nu j}{}^{\mu} f^{i_1\dots i_k}_{j_1\dots j_{\nu-1}\mu j_{\nu+1}\dots j_l} X^j$$

defines invariantly a tensor field $\nabla_X f$ of type k, l, and that this extension of ∇_X is uniquely determined by the product rule

$$\nabla_X(Y_1 \otimes \cdots \otimes Y_k \otimes \omega_1 \otimes \cdots \otimes \omega_l) \\ = (\nabla_X Y_1) \otimes Y_2 \otimes \cdots \otimes \omega_l + \cdots + Y_1 \otimes \cdots \otimes \omega_{l-1} \otimes \nabla_X \omega_l,$$

where $Y_1, \dots, Y_k$ are vector fields and $\omega_1, \dots, \omega_l$ are one forms. Prove that the product rule also holds for arbitrary tensor products.

To round off the definition of ∇_X, we define $\nabla_X f = Xf$ if f is just a function on M. Recall that *contraction* of a tensor means in local coordinates that one of the upper indices is put equal to one of the lower indices followed by summation over that index. In particular, the contraction of $Y \otimes \omega$ where Y is a vector and ω a one form is the scalar product $\langle \omega, Y \rangle$ discussed above. For decomposable tensors as in the preceding exercise, the contraction also means precisely taking scalar product of one of the vector factors and one of the one form factors. Since (3.1.7) shows that for $Y \otimes \omega$ it does not matter if one first applies ∇_X and then contracts or if one first contracts and then applies ∇_X, it follows that *covariant differentiation commutes with contraction.*

Let f again by a tensor field of type k, l. Since $\nabla_X f$ at every point is just a linear function of X, not depending on its derivatives, we can regard $\nabla_X f$ as the contraction of X and a tensor field ∇f of type $k, l+1$,

$$(3.1.2)''' \quad (\nabla f)^{i_1\dots i_k}_{j_1\dots j_l, j} = \partial_j f^{i_1\dots i_k}_{j_1\dots j_l} + \sum_{\nu=1}^{k}\sum_{\mu=1}^{n} \Gamma_{\mu j}{}^{i_\nu} f^{i_1\dots i_{\nu-1}\mu i_{\nu+1}\dots i_k}_{j_1\dots j_l} \\ - \sum_{\nu=1}^{l}\sum_{\mu=1}^{n} \Gamma_{j_\nu j}{}^{\mu} f^{i_1\dots i_k}_{j_1\dots j_{\nu-1}\mu j_{\nu+1}\dots j_l}$$

Note the notation $, j$ for the added index. Also note that $\nabla_X f$ is defined at x even if f is only defined on a curve passing through x with tangent X. A tensor f defined along a curve $t \mapsto x(t) \in M$ can therefore be differentiated with respect to t as $\nabla_{x'(t)} f$.

If X and Y are two vector fields and g is the metric tensor, of type $0, 2$, then $(X, Y) = \sum g_{ij} X^i Y^j$ is the contraction of $g \otimes X \otimes Y$ with respect to both pairs of indices. Thus $\nabla(X, Y)$ is the contraction of

$$(\nabla g) \otimes X \otimes Y + g \otimes (\nabla X) \otimes Y + g \otimes X \otimes \nabla Y.$$

On the other hand, by (3.1.4) it is the contraction of the last two terms only. Hence it follows that

$$\nabla g = 0. \tag{3.1.8}$$

Exercise 3.1.4. Prove (3.1.8) directly from the definitions. Show also that $\nabla g^{\sharp\sharp} = 0$. Prove that if X, Y, Z are vector fields then $\nabla_X \nabla_Y Z$ is the sum of the contraction of $(\nabla^2 Z) \otimes Y \otimes X$ and that of $(\nabla Z) \otimes \nabla_X Y$. Deduce that in local coordinates

$$Z^i{}_{,kj} - Z^i{}_{,jk} = \sum_l R^i{}_{ljk} Z^l.$$

Exercise 3.1.5. Let f be a tensor field of type k, l, and compute the antisymmetric part of $\nabla\nabla f$, that is, in local coordinates,

$$f^{i_1 \dots i_k}_{j_1 \dots j_l, rs} - f^{i_1 \dots i_k}_{j_1 \dots j_l, sr}.$$

Exercise 3.1.6. Let V be a linear subspace of the tensor product $\mathbf{R}^n \otimes \cdots \otimes \mathbf{R}^n$ with $k + l$ factors, such that V is invariant under the action of the orthogonal group $\mathbf{O}(n)$ on the tensor product. Show that if M is a Riemannian manifold of dimension n and f is a tensor of type k, l at $x \in M$, then the components $f^{i_1 \dots i_k}_{i_{k+1} \dots i_{k+l}}$ with respect to any orthonormal basis for $T_x M$ are in V if this is true for one orthonormal basis. Such tensors therefore form a vector subbundle $\mathcal{V}$ of the bundle of tensors of type k, l. Prove that if f is a section of $\mathcal{V}$ and X is a vector field, then $\nabla_X f$ is a section of $\mathcal{V}$.

We shall now discuss the notion of curvature from a somewhat different view point which is close to its origin. Surveyors and cartographers perceived quite early the need for solving spherical triangles (on the earth) which are too large for application of Euclidean trigonometry but too small to make it numerically convenient to use spherical trigonometry. (A spherical triangle is bounded by arcs of great circles.) It was known at least since the 16th centrury that the sum of the angles in a spherical triangle is equal to $\pi + S/R^2$ where S is the area and R is the radius. Legendre observed that when the corner angles are not close to 0 or π, then the "spherical excess" S/R^2 is approximately equally divided as an angular excess of $S/3R^2$ at each corner, compared to a Euclidean triangle with the same sides. The practical rule is then to apply Euclidean trigonometry with this correction. In his fundamental paper on surface theory Gauss [1] devoted much attention to Legendre's theorem and determined the next order of approximation (which gives a larger excess opposite a shorter side), and for general surfaces he derived an analogue of Legendre's theorem with $S/3R^2$ replaced by $SK/3$. Here K is what we know as Gauss' curvature, which occurs in this context naturally as an intrinsic property of the surface. Riemann [1] followed the path initiated by Gauss in his extension to dimensions > 2. We shall now follow these papers in principle but using modern notation of course.

The first point in Gauss [1] was to define an analogue of the great circle arcs on the sphere. These are the shortest paths between any two of its points which have no antipodal points between them. To define an analogue for Riemannian manifolds, *geodesic* arcs, we assume given two points x_0, x_1 in the same coordinate patch ω and look for a smooth curve $[0,1] \ni t \mapsto x(t)$ in ω with $x(0) = x_0$, $x(1) = x_1$, such that

$$s = \int_0^1 ds = \int_0^1 (\sum_{j,k=1}^n g_{jk}(x) dx^j/dt dx^k/dt)^{\frac{1}{2}}\, dt$$

is minimized. By Schwarz' inequality we have

$$s^2 \leq \int_0^1 \sum_{j,k=1}^n g_{jk}(x) dx^j/dt dx^k/dt\, dt$$

with equality if and only if ds/dt is constant, that is, the parameter t is proportional to the arc length. This can always be achieved, so the minimum problem is equivalent to minimizing

$$I = \int_0^1 \sum_{j,k=1}^n g_{jk}(x) dx^j/dt dx^k/dt\, dt \tag{3.1.9}$$

for all smooth curves from x_0 to x_1. This has the advantage that one can expect a unique solution if the points are sufficiently close. If the minimum is attained, then the Euler equations

$$2\frac{d}{dt}\sum_{k=1}^n g_{jk}(x)\frac{dx^k}{dt} = \sum_{i,k=1}^n \partial_j g_{ik}(x)\frac{dx^i}{dt}\frac{dx^k}{dt},\ j = 1,\dots,n, \tag{3.1.10}$$

must be valid. (Replace $x(t)$ by $x(t) + \varepsilon y(t)$ where $y(0) = y(1) = 0$, put the derivative with respect to ε equal to 0 when $\varepsilon = 0$, and integrate by parts.) Carrying out the differentiation in the left-hand side we obtain the equations

$$2\sum_{k=1}^n g_{jk}(x)\frac{d^2x^k}{dt^2} + \sum_{i,k=1}^n (2\partial_i g_{jk}(x) - \partial_j g_{ik}(x))\frac{dx^i}{dt}\frac{dx^k}{dt} = 0, \quad j = 1,\dots,n.$$

The second sum does not change if we replace the parenthesis by

$$\partial_i g_{jk}(x) + \partial_k g_{ji}(x) - \partial_j g_{ik}(x) = 2\Gamma_{ikj}$$

so the differential equations (3.1.10) can be written

$$\frac{d^2x^k}{dt^2} + \sum_{i,j=1}^n {\Gamma_{ij}}^k \frac{dx^i}{dt}\frac{dx^j}{dt} = 0. \tag{3.1.10$'$}$$

These equations mean that the covariant derivative of the tangent vector dx/dt along the curve is equal to 0, that is, that the geodesic curvature vanishes (cf. Theorem

2.1.3). The derivation as Euler equations of an invariant variational problem shows that the equations are invariant under a change of variables, and we know that also from the invariance of the covariant derivative.

Without deciding yet if a solution of the minimum problem for I really exists, we can define:

Definition 3.1.3. An integral curve of the differential equations (3.1.10) (or equivalently (3.1.10)$'$) is called a geodesic curve.

By the basic existence theorems for ordinary differential equations the equations (3.1.10) have a unique solution for small t with prescribed initial data $x = x_0, dx/dt = v \in T_{x_0}$ when $t = 0$. The solution $x(x_0, t, v)$ is a C^∞ function and depends on tv rather than on t and v, because (3.1.10) is independent of t and homogeneous in dt. This means that $x(x_0, t, v) = X(x_0, tv)$ where X is a C^∞ function from a neighborhood of the zero section in $T(M)$ to M with $X(x_0, 0) = x_0$ and $\partial_v X(x_0, 0)$ equal to the identity. By the inverse function theorem we can therefore introduce X as local coordinates in a neighborhood of x_0. If we do that for fixed x_0, and denote the metric in the new coordinates by $\sum G_{jk}(X) dX^j dX^k$, the fact that $t \mapsto tX$ is a solution of (3.1.10) for every X means that

$$2\frac{\partial}{\partial t}\sum G_{jk}(tX)X^k = \sum(\partial_j G_{ik})(tX)X^i X^k. \tag{3.1.10$''$}$$

If we multiply by X^j and sum, it follows that $\sum G_{jk}(tX)X^j X^k$ is independent of t, thus $\sum(G_{jk}(tX) - G_{jk}(0))X^j X^k = 0$. (This also follows at once by covariant differentiation.) The derivative with respect to X^j is

$$2\sum(G_{jk}(tX) - G_{jk}(0))X^k + t\sum(\partial_j G_{ik})(tX)X^i X^k$$

which must therefore vanish. Combined with (3.1.10)$''$ this gives

$$2S_j + 2t\partial S_j/\partial t = 0, \qquad S_j = \sum(G_{jk}(tX) - G_{jk}(0))X^k.$$

Thus tS_j is independent of t, hence equal to 0, so we obtain

$$\sum_1^n G_{jk}(X)X^k = \sum_1^n G_{jk}(0)X^k, \qquad j = 1, \dots, n. \tag{3.1.11}$$

Conversely, if we have a coordinate system such that (3.1.11) is valid, then reversing the preceding argument shows that the straight lines through the origin (with respect to the parameters) are geodesics. We could make a linear change of coordinates to make $G_{jk}(0)$ equal to the identity matrix. However, with pseudo-Riemannian geometry in view, where g_{jk} will be non-singular but not positive definite, we prefer to leave $G_{jk}(0)$ arbitrary. The coordinate systems now obtained are called *geodesic coordinates*; they are uniquely determined up to a linear transformation (up to an orthogonal one if one has insisted on the Euclidean normal form).

From Cauchy-Schwarz' inequality and (3.1.11) it follows that

$$\Big(\sum_{j,k=1}^n G_{jk}(X)dX^jdX^k\Big)^{\frac12}$$
$$\geq \sum_{j,k=1}^n G_{jk}(0)X^jdX^k/\Big(\sum_{j,k=1}^n G_{jk}(0)X^jX^k\Big)^{\frac12} = d\Big(\sum_{j,k=1}^n G_{jk}(0)X^jX^k\Big)^{\frac12},$$

and this implies that when X is in the domain of the geodesic coordinates, then the shortest path to the origin is indeed the geodesic ray. Thus we have proved that geodesics do give the shortest path between any two of its points which are sufficiently close. (The importance of the latter restriction was already clear for the sphere.)

So far in this chapter we have only discussed local properties of Riemannian manifolds. A Riemannian manifold is a metric space with the distance $s(x, y)$ between $x, y \in M$ defined as the infimum of the lengths of differentiable curves from x to y. We add a theorem of a global nature:

Theorem 3.1.4 (Hopf-Rinow). *Let M be a connected Riemannian manifold. Then the following properties are equivalent:*

(1) *M is a complete metric space.*
(2) *Every geodesic in M can be extended indefinitely in both directions.*
(3) *There exists a point $x \in M$ such that all geodesics starting at x can be extended indefinitely in both directions.*
(4) *Every closed bounded subset of M is compact.*

They imply that any two points $x, y \in M$ can be joined by a geodesic of length $s(x, y)$.

Proof. (1) $\Longrightarrow$ (2): Let $\mathbf{R} \supset (a, b) \ni s \mapsto x(s)$ be a geodesic with maximal interval of definition and the arc length as parameter. Since $s(x(s_1), x(s_2)) \leq |s_1 - s_2|$, the sequence $x(b - 1/k)$ is a Cauchy sequence if $b < \infty$. By the local existence theorem for ordinary differential equations applied in a neighborhood of the limit, the geodesic with initial data $x(b-1/k), x'(b-1/k)$ can be extended for an interval δ independent of k for large k, which shows that (a, b) is not a maximal interval of definition and proves (2).

The implications (4) $\Longrightarrow$ (1) and (2) $\Longrightarrow$ (3) are trivial. Assume now that (3) is valid. This means that we have a globally defined geodesic exponential map $\gamma : T_xM \ni v \mapsto X(1) \in M$, giving the value at time 1 of the geodesic with initial data x, v. Since γ is continuous and $B(r) = \{v \in T_xM; |v|_x \leq r\}$ is compact, it follows that $\gamma B(r)$ is compact. We know already that $\gamma B(r)$ is equal to $\widehat{B}(r) = \{y \in M; s(x, y) \leq r\}$ for small r, and it is obvious that $\gamma B(r) \subset \widehat{B}(r)$ for every $r > 0$. If we prove that equality holds for every r, then (4) will be proved.

Let r_1 be the supremum of all r such that $\gamma B(\varrho) = \widehat{B}(\varrho)$ when $\varrho < r$. If $r_1 = \infty$ our claim is true so assume that $r_1 < \infty$. Then $\gamma B(r_1) = \widehat{B}(r_1)$, for $\widehat{B}(r_1)$ is the closure of

$$\cup_{r<r_1}\widehat{B}(r_1) = \gamma \cup_{r<r_1} B(r) \subset \gamma B(r_1),$$

and $\gamma B(r_1)$ is compact and therefore closed. Now we claim that $\gamma B(r) = \widehat{B}(r)$ for $r < r_1 + \delta$ if δ is small enough. If $y \in \widehat{B}(r) \setminus \widehat{B}(r_1)$, we can for large ν find a path of length $< s(x,y) + 1/\nu$ from x to y. If z_ν is the last point in $\widehat{B}(r_1) = \gamma B(r_1)$, then $s(x, z_\nu) = r_1$ and $s(z_\nu, y) + r_1 \le s(x,y) + 1/\nu$. If z is a limit point of the sequence z_ν in the compact set $\gamma B(r_1)$, it follows that $s(z,y) \le s(x,y) - r_1 < \delta$. In view of the compactness of $\gamma B(r)$ we can choose δ independently of z so small that this implies that there is a geodesic from z to y of length $s(x,y) - r_1$. We also have a geodesic of length r_1 from x to y. If these did not fit together to one geodesic, we could get a path shorter than $s(x,y)$ from x to y by smoothing out the broken geodesic near z, which is a contradiction completing the proof of Theorem 3.1.4.

Variations of geodesics are controlled by the *Jacobi differential equations*:

Theorem 3.1.5. *Let $\gamma \in C^3(\omega, M)$ where ω is an open set in $\mathbf{R}^2$, and assume that $t \mapsto \gamma(t,s)$ is for fixed s a geodesic with parameter proportional to the arc length, when $(t,s) \in \omega$. Let $T = \gamma_* \partial/\partial t$ be the tangent of the geodesic and let $X = \gamma_* \partial/\partial s$ be a vector field along the geodesic describing the direction in which it moves. Then X satisfies the Jacobi differential equation $\nabla_T \nabla_T X = R(T,X)T$, where in local coordinates*

$$R(T,X)Z = \sum R^i{}_{jkl} Z^j T^k X^l \partial/\partial x^i.$$

Proof. If γ' has rank 2 then the range is locally a two dimensional surface where the vector fields T and X are defined and commute, since $\partial/\partial t$ and $\partial/\partial s$ commute. In this surface we have $\nabla_T T = 0$ for $t \mapsto \gamma(t,s)$ is a geodesic. Since $[T,X] = 0$ we have $\nabla_T X - \nabla_X T = 0$, so (3.1.6) gives

$$\nabla_T \nabla_T X = \nabla_T \nabla_X T = [\nabla_T, \nabla_X]T + \nabla_X \nabla_T T = \nabla_{[T,X]} T + R(T,X)T = R(T,X)T,$$

which proves the Jacobi differential equation in the closure of the open subset of ω where γ' has rank 2. In the complement in ω we have locally $\gamma(t,s) = \gamma_0(\psi(t,s))$ where $\gamma_0(t) = \gamma(t, s_0)$ for a suitable s_0. The geodesic equation means that $\psi(t,s)$ is linear in in t for fixed s. Hence $X = aT$ where $a = (\partial\psi/\partial s)/(\partial\psi/\partial t)$ is linear in t, and

$$\nabla_T X = (\nabla_T a)\nabla T, \quad \nabla_T \nabla_T X = (\nabla_T \nabla_T a)T = 0,$$

which completes the proof.

Exercise 3.1.7. Show that if M_0 is a submanifold of the Riemannian manifold M of codimension ν, then one can at every point in M_0 choose local coordinates $x^1, \dots, x^n$ in M such that M_0 is defined by $x^1 = \cdots = x^\nu = 0$ and

$$\sum_{k=1}^{\nu} g_{jk}(x) x^k = \begin{cases} x^j, & \text{if } 1 \le j \le \nu \\ 0, & \text{if } j > \nu. \end{cases}$$

Hint: This means precisely that the rays $t \mapsto (tx^1, \dots, tx^\nu, x^{\nu+1}, \dots, x^n)$ are geodesics orthogonal to M_0. (Note that when $\nu = 1$ the condition means that $g_{1k}(x) = \delta_{1k}$.)

Exercise 3.1.8. Show that if M_0 is a geodesic curve in a Riemannian manifold M, then one can in a neighborhood of any compact interval on M_0 find coordinates $x = (x^1, \dots, x^n)$ such that M_0 is defined by $x' = (x^1, \dots, x^{n-1}) = 0$ and

$$g_{jk}(x) = \delta_{jk} + G_{jk}(x', x^n) + O(|x'|^3),$$

where $G_{jk}(x', x^n)$ is a quadratic form in x' depending on x^n and

$$\sum_1^{n-1} G_{jk}(x', x^n)x^k = 0, \; j = 1, \dots, n.$$

(Hint: Use Exercise 3.1.6 and make the vector fields $\partial/\partial x^j$ parallel along M_0.) Conclude that if the differential equations

$$2d^2X^j/dx^{n2} = \partial G_{nn}(X', x^n)/\partial x^j, \quad j = 1, \dots, n-1$$

have a solution $X'(x^n) \not\equiv 0$ vanishing when $x^n = a$ and $x^n = b$, then an interval on M_0 containing $[a, b]$ strictly cannot minimize the distance between its endpoints. Show that the Jacobi differential equation along M_0 for a vector field $(X'(x^n), X^n(x^n))$ means that X^n is a linear function of x^n and that $X'(x^n)$ satisfies the preceding differential equations.

It follows from (3.1.11) that if we expand G in a Taylor series,

$$G = G^0 + G^1 + G^2 + \dots$$

where G^j is homogeneous of degree j in X, then $G^1 = 0$, for the equations $\sum G^1_{jk}(X)X^k = 0$ imply that if $G^1_{jk}(X) = \sum G_{jkl}X^l$, then G_{jkl} is symmetric in the first two indices and antisymmetric in the last two. The permutations

$$jkl \to jlk \to ljk \to lkj \to klj \to kjl \to jkl$$

must therefore change the sign although we get back the same elements, which proves the claim.

The first interesting term is therefore G^2. We shall write

$$G^2(X; Y) = \langle G^2(X)Y, Y\rangle,$$

which is a symmetric quadratic form in X as well as in Y and has the fundamental property

$$\partial G^2(X; Y)/\partial Y = 0, \quad \text{if } Y = X. \tag{3.1.12}$$

In particular, $G^2(X; X) = 0$. Since the dimension of the space of quadratic forms in n variables is $n(n + 1)/2$ and that of cubic forms is $n(n + 1)(n + 2)/6$, it is easily seen that the space of forms satisfying (3.1.12) is of dimension

$$(n(n + 1)/2)^2 - n^2(n + 1)(n + 2)/6 = n^2(n^2 - 1)/12.$$

We have seen in Theorem 2.3.1 that this is precisely the dimension of the space $\mathcal{T}$ of curvature tensors. We can polarize G^2 to a 4-linear form $G^2(X_1, X_2; Y_1, Y_2)$ which is symmetric in X_1, X_2 as well as in Y_1, Y_2 and has the property

$$G^2(X; Y) = G^2(X, X; Y, Y).$$

Then (3.1.12) yields

$$G^2(X_1, X_2; X_3, X_4) + G^2(X_3, X_1; X_2, X_4) + G^2(X_2, X_3; X_1, X_4) = 0. \tag{3.1.12$'$}$$

From (3.1.12)$'$ we easily obtain $G^2(X, X; Y, Y) = -2G^2(X, Y; X, Y)$, hence the symmetry

$$G^2(X; Y) = G^2(Y; X). \tag{3.1.13}$$

Since $G'_{jk}(0) = 0$, the corresponding Christoffel symbols vanish at the origin. (Note that the Christoffel symbols are not tensors, for otherwise we could not make them equal to 0 at a point by changing coordinates. The fact that we have raised and lowered an index in the same way as for tensors might suggest otherwise, but it is just for linear transformations that they behave as tensors.) Thus the formulas (2.1.13)$'$ for the curvature tensor simplify at 0 to

$$R_{lijk} = \tfrac{1}{2}(\partial_j \partial_i G_{kl} + \partial_k \partial_l G_{ij} - \partial_j \partial_l G_{ik} - \partial_i \partial_k G_{jl})$$

after cancellation of two terms $\partial_j \partial_k G_{il}$. This means that for the corresponding 4-linear forms we have

$$\begin{aligned} R(X_1, X_2; X_3, X_4) = G^2(X_1, X_4; X_2, X_3) &+ G^2(X_2, X_3; X_1, X_4) \\ - G^2(X_2, X_4; X_1, X_3) &- G^2(X_1, X_3; X_2, X_4). \end{aligned} \tag{3.1.14}$$

Here we have only used so far that $G^1 = 0$. For geodesic coordinates we have the symmetry (3.1.13), which simplifies (3.1.14) to

$$R(X_1, X_2; X_3, X_4) = 2G^2(X_1, X_4; X_2, X_3) - 2G^2(X_1, X_3; X_2, X_4). \tag{3.1.14$'$}$$

Using (3.1.12)$'$ we obtain on the other hand

$$\begin{aligned} -6G^2(X_1, X_2; X_3, X_4) &= R(X_1, X_4; X_2, X_3) + R(X_1, X_3; X_2, X_4), \\ -3G^2(X, X; Y, Y) &= R(X, Y; X, Y). \end{aligned} \tag{3.1.15}$$

We have a one to one correspondence between the two 4-linear forms $G^2(X_1, \dots, X_4)$ and $R(X_1, \dots, X_4)$ with the properties:

symmetry for G^2 when $X_1 \leftrightarrow X_2$ or $X_3 \leftrightarrow X_4$ or $(X_1, X_2) \leftrightarrow (X_3, X_4)$; for R when $(X_1, X_2) \leftrightarrow (X_3, X_4)$;

antisymmetry for R when $X_1 \leftrightarrow X_2$ or $X_3 \leftrightarrow X_4$;

circular antisymmetry for both G^2 and R as described in (3.1.12)$'$. For R this is again the first Bianchi identity.

Exercise 3.1.9. Show that at the center of a geodesic coordinate system we have

$$\Gamma_{ijk}(x) = \tfrac{1}{3}\sum_l (R_{ikjl}(0) + R_{iljk}(0))x^l + O(|x|^2).$$

What is the analogue for ${\Gamma_{ij}}^k$?

Exercise 3.1.10. Show that at the center of a geodesic coordinate system we have

$$\sum_{i,j,k,l} {R_{ijkl}}^2 = 12 \sum_{i,j,k,l} {G_{ijkl}}^2.$$

The preceding observations were expressed in words by Riemann [1, p. 279]: "Führt man diese Grössen ein, so wird für unendlich kleine Werthe von x das Quadrat des Linienelements $= \sum dx^2$, das Glied der Nächsten Ordnung in demselben aber gleich einem homogenen Ausdruck zweiten Grades der $n(n-1)/2$ Grössen $(x_1dx_2 - x_2dx_1)$, $(x_1dx_3 - x_3dx_1)$, $\dots$, also eine unendlich kleine Grösse der vierten Dimension, so dass man eine endliche Grösse erhält wenn man sie durch das Quadrat des unendlich kleinen Dreiecks dividirt, in dessen Eckpunkten die Werthe der Veränderlichen sind $(0, 0, 0, \dots)$, $(x_1, x_2, x_3, \dots)$, $(dx_1, dx_2, dx_3, \dots)$. Diese Grösse behält denselben Werth, so lange die Grössen x und dx in denselben binären Linearformen enthalten sind, oder so lange die beiden kürzesten Linien von den Werthen 0 bis zu den Werthen x und von den Werthen 0 bis zu den Werthen dx in demselben Flächenelement bleiben, und hängt also nur von Ort und Richtung desselben ab. Sie wird offenbar $= 0$, wenn die dargestellte Manningfaltigkeit eben, d.h. das Quadrat des Linienelements auf $\sum dx^2$ reducirbar ist, und kann daher als das Mass der in diesem Punkte in dieser Flächenrichtung stattfindenden Abweichung der Mannigfaltigkeit von der Ebenheit angesehen werden. Multiplicirt mit $-3/4$ wird sie der Grösse gleich, welche Herr Geheimer Hofrath Gauss das Krümmungsmass einer Fläche genannt hat."

The factor 4 here comes from the fact that Riemann divided by the square of the area of a triangle and not the corresponding parallelogram. The factor -3 is the same as in (3.1.15) above, and we shall now show that it is also closely connected to the denominator 3 in Legendre's theorem.

Still with geodesic coordinates we shall consider the geodesic triangle with corners at 0, εY, εZ when ε is small. For the sides from 0 the lengths are $\varepsilon|Y|$ and $\varepsilon|Z|$ where $|X|^2 = \sum G_{jk}(0)X^jX^k$; we denote the corresponding scalar product by $(\cdot,\cdot)$. The third side is not as easy to determine since we do not know the geodesic. However, it is clear that the square of its length is equal to ε^2 times the square of the length of the geodesic from Y to Z for the metric

$$\sum G_{jk}(\varepsilon X)dX^jdX^k = |dX|^2 + \varepsilon^2 G^2(X; dX) + O(\varepsilon^3).$$

This geodesic must differ from the straight line segment $[0,1] \ni t \mapsto Y + t(Z-Y)$ by $O(\varepsilon^2)$, and since the straight line is a geodesic for the metric $|dX|^2$, we obtain for the geodesic distance εa between εY and εZ

$$a^2 = |Y-Z|^2 + \varepsilon^2 \int_0^1 G^2(Y + t(Z-Y); Z-Y)\,dt + O(\varepsilon^3)$$
$$= |Y-Z|^2 + \varepsilon^2 G^2(Y;Z) + O(\varepsilon^3),$$

for $G^2(Y+t(Z-Y); Z-Y) = G^2(Y; Z-Y) = G^2(Y;Z)$. The Riemannian angle α_r at 0 is defined by $(Y,Z) = |Y||Z|\cos\alpha_r$, and the angle α_e opposite εa in the Euclidean triangle with sides $\varepsilon|Y|$, $\varepsilon|Z|$, εa is given by the cosine theorem

$$a^2 = |Y|^2 + |Z|^2 - 2|Y||Z|\cos\alpha_e, \quad \text{hence}$$
$$2|Y||Z|(\cos\alpha_r - \cos\alpha_e) = \varepsilon^2 G^2(Y;Z) + O(\varepsilon^3).$$

Thus $\delta = \alpha_r - \alpha_e = O(\varepsilon^2)$ and

$$\delta = -\frac{G^2(Y;Z)}{(|Y||Z|\sin\alpha_r)^2}\frac{|\varepsilon Y||\varepsilon Z|\sin\alpha_r}{2} + O(\varepsilon^3).$$

Since the first factor is $-1/3$ times the sectional curvature in the YZ plane (cf. (3.1.15) and Theorem 2.2.3), and the second factor is the area of the geodesic triangle $+O(\varepsilon^3)$, we have proved Legendre's theorem. Note that the total angle excess is equal to the sectional curvature times the area $+O(\varepsilon^3)$.

Exercise 3.1.11. Assuming that the earth is a sphere with circumference 40000 km, and that T is an equilateral geodesic triangle on the earth with the base equal to 100 km and angles 60° at the base, estimate how many seconds of arc the third angle differs from 60°.

In the two dimensional case considered by Gauss we can make a subdivision of the geodesic triangle T by geodesics joining the midpoints of each side. Repeating this subdivision and noting that angle excess is additive when we subdivide, we conclude in the limit that

$$\alpha + \beta + \gamma = \pi + \iint_T K\,dS \tag{3.1.16}$$

where α, β, γ are the angles at the corners, K is the Gauss curvature, and dS is the Riemannian area measure. This is Gauss' part of the Gauss-Bonnet theorem; it was extended by Bonnet to more general regions. This we shall now do with a different more analytical proof which will introduce some ideas which will be important later on.

Let us assume that we have a Riemannian metric in a simply connected open set $\omega \subset \mathbf{R}^2$. We shall generalize Theorem 1.2.1 by calculating the integral of the signed geodesic curvature κ_g over a simple closed curve $s \mapsto x(s) \in \omega$, $0 \le s \le L$. We assume that s is the arc length, that is,

$$\sum_{j,k=1}^{2} g_{jk}(x(s)) \frac{dx^j(s)}{ds} \frac{dx^k(s)}{ds} = 1,$$

and we denote by $e(s)$ the unit tangent vector $e(s) = x'(s)$. The covariant derivative of e along the curve has the components

$$(\nabla_e e)^j = \frac{de^j}{ds} + \sum_{i,k=1}^{2} \Gamma_{ik}{}^j(x(s)) e^i(s) \frac{dx^k}{ds},$$

but we shall avoid explicit calculations using this expression. Let n be the unit vector orthogonal to e such that e, n is positively oriented. Then the signed geodesic curvature κ_g is given by

$$\kappa_g = (\nabla_e e, n),$$

and our task is to calculate the integral

$$\int_0^L \kappa_g \, ds. \tag{3.1.17}$$

This is the integral of the differential form

$$\kappa = \sum_{j,l=1}^{2} g_{jl}(de^j + \sum_{k,i=1}^{2} \Gamma_{ik}{}^j(x) e^i dx^k) n^l \tag{3.1.18}$$

in the unit circle bundle

$$S(T\omega) = \{(x, w) \in T(\omega); \sum_{j,k=1}^{2} g_{jk}(x) w^j w^k = 1\},$$

the integral being taken along the curve $\gamma : s \mapsto (x(s), e(s))$. (Recall that n is uniquely determined by e.) Here $e(s) = x'(s)$, but we shall simplify the problem by generalizing it, so we allow any closed C^1 curve γ in $S(T\omega)$ now.

We shall first examine how $\int_\gamma \kappa$ depends on the choice of the unit vector e. Any other choice $\tilde{e}$ can be written

$$\tilde{e} = e \cos\theta + n \sin\theta, \quad \text{thus } \tilde{n} = -e \sin\theta + n \cos\theta,$$

where θ is uniquely determined modulo 2π, so $d\theta$ is uniquely determined. We obtain

$$\nabla\tilde{e} = (\nabla e)\cos\theta + (\nabla n)\sin\theta + (-e\sin\theta + n\cos\theta)d\theta,$$

and since $(e,e)=(n,n)=1$, $(e,n)=0$, it follows from (3.1.4) that

$$(\nabla e, e) = 0, \quad (\nabla n, n) = 0, \quad (\nabla e, n) + (e, \nabla n) = 0.$$

Hence

$$(\nabla \tilde{e}, \tilde{n}) = (\nabla e, n) + d\theta,$$

so $\int_{\tilde{\gamma}} \kappa = \int_\gamma \kappa + \int_0^L d\theta$ if $\tilde{\gamma}$ is the curve $s \mapsto (x(s), \tilde{e}(s))$.

If v and $\tilde{v}$ are two arbitrary vector fields with no zeros defined along the curve $s \mapsto x(s)$, then we can normalize them with respect to the Riemannian metric to unit vector fields e and $\tilde{e}$ and introduce the variation $\int_0^L d\theta$ of the angle from e to $\tilde{e}$ along the curve. This is an integer multiple of 2π which must be independent of the Riemannian metric. In fact, if we replace g_{jk} by $\lambda g_{jk} + (1-\lambda)\delta_{jk}$, $0 \le \lambda \le 1$, then the angle variation depends continuously on λ so it must have the same value when $\lambda = 0$ as when $\lambda = 1$. If for example one of the vector fields is tangent to the curve and the other is $\partial/\partial x^1$, then it follows from Theorem 1.2.1 that the angle variation between them is $\pm 2\pi$.

We shall now determine the integral $\int_\gamma \kappa$ when e is a unit vector field defined in the whole of ω, for example the normalization of the vector field $\partial/\partial x^1$. The advantage of this is that we can then pull the form κ back by the map $x \mapsto (x, e(x))$ to a form in ω with integral over the curve $s \mapsto x(s)$ equal to $\int_\gamma \kappa$. (It also follows that γ is a boundary in $S(T\omega)$.) The interior ω' of the curve is $\Subset \omega$ since ω was assumed simply connected, and we shall calculate the integral of the differential form over $\partial\omega'$ using Stokes' formula. If $\varepsilon_i = \partial/\partial x^i$ and we write $\nabla_i = \nabla_{\varepsilon_i}$ for the sake of brevity, the differential form to integrate is

$$\sum_{i=1}^{2} (\nabla_i e, n)\, dx^i.$$

The integral over $\partial\omega'$ is equal to the integral over ω' of

$$\nabla_1(\nabla_2 e, n) - \nabla_2(\nabla_1 e, n) = ((\nabla_1\nabla_2 - \nabla_2\nabla_1)e, n) + (\nabla_2 e, \nabla_1 n) - (\nabla_1 e, \nabla_2 n). \tag{3.1.19}$$

Here we have used (3.1.4). The last two terms are equal to 0 since $\nabla_j e$ has the direction of n and $\nabla_k n$ has the direction of e. By (3.1.6) the first term, on the right is equal to $R(n, e, \varepsilon_1, \varepsilon_2) = -R(e, n, \varepsilon_1, \varepsilon_2)$. The vectors e, n are positively oriented and span a parallelogram with area 1 while $\varepsilon_1, \varepsilon_2$ span one of area $\sqrt{g}$ where $g = \det(g_{jk})$. In view of Theorem 2.2.3 we therefore conclude that (3.1.19) is equal to $-K\sqrt{g}$, where K is the total (Gaussian) curvature, and we have proved:

Theorem 3.1.6 (Gauss-Bonnet). *Let $\omega \subset \mathbf{R}^2$ be a simply connected open set with Riemannian metric, and let $\omega' \Subset \omega$ be simply connected with C^2 boundary. Then*

$$\int_{\partial\omega'} \kappa_g\, ds + \int_{\omega'} K\, dS = 2\pi, \tag{3.1.20}$$

where K is the total curvature, κ_g is the geodesic curvature of $\partial\omega'$ with the orientation making ω' lie to the left of $\partial\omega'$, ds is the arc length of $\partial\omega'$, and dS is the Riemannian area element $\sqrt{g}\,dx$ in ω'.

In the proof we have actually proved

$$d\kappa = \pi^*(-K\sqrt{g}dx^1 \wedge dx^2) \tag{3.1.21}$$

where π is the projection $S(T\omega) \to \omega$ and κ is defined by (3.1.18). In fact, since exterior differentiation commutes with pullbacks, we have just proved that the two sides of (3.1.21) are equal when pulled back by any local section $x \mapsto (x, e(x))$ of $S(T\omega)$. This implies (3.1.21), for if u is a differential form in a fiber space over a manifold M and $s^*u = 0$ for every local section of M, then $u = 0$ if the degree of u does not exceed $\dim M$. (Verify this as an exercise.) Note that κ cannot be obtained by lifting a form from ω; in fact, we have shown that the restriction of κ to any fiber of $S(T\omega)$ is equal to the natural one form on the oriented circle.

In Theorem 3.1.6 we assumed that $\partial\omega'$ was in C^2, so (3.1.20) does not quite cover (3.1.16). However, it is easy to extend the theorem to the case where $\partial\omega'$ is only piecewise C^2. To get a closed curve in $S(T\omega)$ we must then add at each corner x a circular arc in the fiber connecting the incoming tangent to the outgoing tangent; alternatively we can approximate ω' by domains with the corners rounded off. This gives

$$\int_{\partial\omega'} \kappa_g\, ds + \sum \alpha_j + \int_{\omega'} K\, dS = 2\pi, \tag{3.1.20$'$}$$

where α_j denote the exterior angles at the corners and κ_g is integrated only over the smooth part of $\partial\omega'$. For a geodesic triangle, the three angles are $\pi - \alpha_j$, so the sum is

$$3\pi - \sum \alpha_j = \pi + \int_{\omega'} K\, dS.$$

Thus (3.1.20)$'$ contains (3.1.16). It is also easy to obtain (3.1.20) from (3.1.16) applied to a triangulation of a polygonal approximation of ω'. We leave this also as an exercise.

We shall now discuss the case of a compact oriented Riemannian manifold M of dimension 2 (without boundary). Suppose that M is decomposed by geodesic arcs into a finite number ν_2 of geodesic polygons ω_j. Denote the interior angles of ω_j by β_{jk}. Then (3.1.20)$'$ yields

$$\sum_k (\pi - \beta_{jk}) + \int_{\omega_j} K\, dS = 2\pi.$$

The number of terms in the sum is equal to the number of sides of ω_j. If ν_0 and ν_1 denote the total number of corners and sides occurring in some ω_j, we get by adding

$$\int_M K\, dS + 2\pi\nu_1 - 2\pi\nu_0 = 2\pi\nu_2,$$

for there will be altogether $2\nu_1$ terms π in the left-hand side. Thus

$$\frac{1}{2\pi}\int_M K\,dS = \nu_2 - \nu_1 + \nu_0, \tag{3.1.22}$$

so the right-hand side is independent of how the decomposition of M is made; it is the *Euler characteristic* of M.

We can approach the calculation of the integral in (3.1.22) in another way starting from a function f on M which has only non-degenerate critical points. Then the vector field $F = (df)^\sharp$, usually denoted $\operatorname{grad} f$, has only finitely many zeros. Let

$$\Gamma = \{(x, F(x)/\|F(x)\|); F(x) \neq 0\} \subset S(TM).$$

This is a manifold whose boundary consists of the circles over the zeros of F, that is, the critical points of f. For at such a point we can introduce geodesic coordinates diagonalizing the quadratic terms in f, that is,

$$g_{jk}(x) = \delta_{jk} + O(|x|^2), \quad f(x) = f(0) + (f_1(x^1)^2 + f_2(x^2)^2)/2 + O(|x|^3).$$

Then we have $F(x) = (f_1x^1, f_2x^2) + O(|x|^2)$ and

$$F(x)/\|F(x)\| = (f_1x^1, f_2x^2)/\sqrt{(f_1x^1)^2 + (f_2x^2)^2} + O(|x|).$$

When x winds around a small circle $|x| = \varepsilon$ then $F(x)/\|F(x)\|$ winds around the unit circle in the same or opposite direction depending on the sign of f_1f_2. The boundary of Γ is the limit of the image of the negatively oriented circle of radius ε as $\varepsilon \to 0$, so it consists of the circles in $S(TM)$ with orientation opposite to the sign of the Hessian of f at the critical point. If we integrate (3.1.21) over Γ it follows now from Stokes' formula that

$$-\int K\,dS = \int_\Gamma d\kappa = \int_{\partial\Gamma}\kappa = -2\pi\sum\varepsilon_j$$

where ε_j is the sign of the Hessian of f at the critical point. Thus we have

Theorem 3.1.7. *If M is a compact two dimensional oriented Riemannian manifold with total curvature K and no boundary, then $\int_M K\,dS/2\pi$ is for every real valued function on M with only non-degenerate critical points equal to the sum of the signs of the Hessian of f at the critical points. The integral is also equal to the Euler characteristic.*

Instead of the vector field $F = (df)^\sharp$ we could have used here any vector field with only non-degenerate fixed points (that is, zeros). We shall later on give an extension due to Chern of the preceding arguments which is applicable to any oriented manifold of even dimension. The problem is to find the appropriate differential forms in $S(TM)$. To do so we need a systematic approach to Riemannian geometry using differential forms; the proof of Theorem 3.1.6 was meant to motivate the need for that.

We shall now supplement Theorem 3.1.7 by studying the integral $\int_M K\,dS$ for a compact oriented hypersurface $M \subset \mathbf{R}^{n+1}$ of any dimension n. We denote by γ the

Gauss map (2.2.3); the direction of the normal n is chosen so that a positive system of tangent vectors followed by n is a positive system in $\mathbf{R}^{n+1}$. We choose the orientation of S^n so that with this definition the normal at $x \in S^n$ is $-x$. Set $\check{\gamma} = -\gamma$, which is then the identity map if $M = S^n$. The degree D of the map $\check{\gamma}$ is then defined by

$$\int_M \check{\gamma}^* u = D \int_{S^n} u$$

where u is an arbitrary n-form on S^n. We choose for u the volume form on S^n, which means that $u(t_1, \dots, t_n)$ is the n-dimensional volume with sign of the parallelepiped spanned by the tangent vectors $t_1, \dots, t_n$ at a point on S^n. By Theorem 2.2.2

$$(\check{\gamma}^* u)(t_1, \dots, t_n) = u(\check{\gamma}' t_1, \dots, \check{\gamma}' t_n) = \big(\prod K_j\big) u(t_1, \dots, t_n).$$

Now $t_1, \dots, t_n$ have as tangent vectors of M the same orientation as they have as tangent vectors of S^n at $-n$, since the normal is n there. Hence $\check{\gamma}^* u$ equals K times the volume form of M, and we obtain

$$\int_M K\, dS = D \int_{S^n} dS, \tag{3.1.23}$$

where dS denotes the area element in M and that in S^n in the two integrals.

When $n = 2$ it follows in view of Theorem 3.1.7 that the degree of the mapping $-\gamma$ is equal to half the Euler characteristic. Later on we shall extend this to arbitrary n such that $(n-2)/4$ is an integer.

We shall finally prove an important result on the covariant derivative of the curvature tensor. For a geodesic system of coordinates the Christoffel symbols are $O(|x|)$, so the non-linear terms in $(2.1.13)'$ are $O(|x|^2)$. Hence we obtain

$$R_{ijkl} = \tfrac{1}{2}(\partial_j \partial_k g_{il} + \partial_i \partial_l g_{jk} - \partial_i \partial_k g_{jl} - \partial_j \partial_l g_{ik}) + O(|x|^2),$$

and it follows that

$$R_{ijkl,m} = \tfrac{1}{2}(\partial_j \partial_m \partial_k g_{il} + \partial_i \partial_l \partial_m g_{jk} - \partial_i \partial_m \partial_k g_{jl} - \partial_j \partial_l \partial_m g_{ik})$$

at the center of the geodesic coordinates. Note that the first and the last term are equal apart from the sign and a circular permutation of the indices klm, and that this is also true for the middle terms. Hence we obtain the *second Bianchi identity*

$$R_{ijkl,m} + R_{ijlm,k} + R_{ijmk,l} = 0, \tag{3.1.24}$$

or more briefly $R_{ij[kl,m]} = 0$. Since $R_{ijkl,m}$ is a tensor, this is of course true for any system of coordinates. Contraction of (3.1.24) with respect to the indices ik and the indices jl, that is, multiplication by $g^{ik} g^{jl}$ and summation gives, since contraction and the musical isomorphisms commute with covariant differentiation

$$S_{,m} - \sum g^{ik} R_{im,k} - \sum g^{jl} R_{jm,l} = 0,$$

that is,

$$\partial_k S = 2\sum g^{ij} R_{ik,j}. \tag{3.1.25}$$

Here S is the scalar curvature and R_{ij} is the Ricci tensor.

By Definition 2.3.4 a Riemannian manifold M is said to be an Einstein manifold if $R_{ij} = fg_{ij}$ for some function f on M. This implies that $S = nf$, where n is the dimension of M. Since $R_{ik,j} = g_{ik}\partial_j f$ in view of (3.1.8), using (3.1.25) gives $n\partial_k f = 2\partial_k f$, $k = 1, \dots, n$. If $n > 2$ and M is connected, it follows that f is a constant, so we have proved:

Theorem 3.1.8. *If M is a connected Einstein manifold of dimension > 2, then the Ricci tensor is a constant multiple of the metric tensor.*

3.2. Local isometric embedding. Let us assume given a smooth Riemannian metric in a neighborhood of 0 in $\mathbf{R}^n$. We want to find a map $x \mapsto f(x) \in \mathbf{R}^N$ defined in a neighborhood of 0 so that the given metric is the same as the metric introduced on the embedded manifold by the Euclidean metric, that is, $|df(x)|^2 = \sum g_{jk}(x)dx^j dx^k$, or explicitly

$$(\partial_j f(x), \partial_k f(x)) = g_{jk}(x), \quad 1 \le j \le k \le n. \tag{3.2.1}$$

Altogether there are $n(n+1)/2$ equations, so it follows from general theorems discussed at the end of this section that smooth solutions do not exist in general unless $N \ge n(n+1)/2$. We shall now discuss a classical theorem of Janet and Cartan (see Jacobowitz [1] and references there) which shows that a local real analytic solution always exists when $N = n(n+1)/2$ if g_{jk} are real analytic. The idea of the proof is to argue by induction with respect to n, and extend a local isometric embedding of $\mathbf{R}^{n-1} \times \{0\}$ by solving a Cauchy problem. To do so one must cut down the number of equations to solve:

Lemma 3.2.1. *The equations* (3.2.1) *are valid in a ball with center at* 0 *if and only if they hold for $1 \le j \le k < n$ when $x_n = 0$ and in addition we have the equations*

$$\begin{aligned} &(\partial_n^2 f, \partial_j f) = \Gamma_{nnj}, \quad 1 \le j \le n, \\ &(\partial_n^2 f, \partial_i\partial_j f) = \partial_n \Gamma_{ijn} + (\partial_i\partial_n f, \partial_j\partial_n f) - \tfrac12 \partial_i\partial_j g_{nn}, \quad 1 \le i \le j < n, \end{aligned} \tag{3.2.2}$$

and the boundary conditions

$$\begin{aligned} (\partial_n f, \partial_j f) &= g_{nj}, \quad 1 \le j \le n,\ x_n = 0, \\ (\partial_n f, \partial_i\partial_j f) &= \Gamma_{ijn}, \quad 1 \le i \le j < n,\ x_n = 0. \end{aligned} \tag{3.2.3}$$

Proof. The first n equations (3.2.2) and the second set of initial conditions (3.2.3) follow from the definition (2.1.7) of the Christoffel symbols. To prove the second set of equations (3.2.2) we note that

$$\begin{aligned} \partial_n \Gamma_{ijn} &= \partial_n(\partial_n f, \partial_i\partial_j f) = (\partial_n^2 f, \partial_i\partial_j f) + (\partial_n f, \partial_n\partial_i\partial_j f), \\ \tfrac12 \partial_i\partial_j g_{nn} &= \tfrac12 \partial_i\partial_j(\partial_n f, \partial_n f) = (\partial_n f, \partial_n\partial_i\partial_j f) + (\partial_i\partial_n f, \partial_j\partial_n f). \end{aligned} \tag{3.2.4}$$

Subtraction yields the remaining equations (3.2.2).

Now assume that we have a solution of (3.2.2), (3.2.3) such that (3.2.1) is valid when $x_n = 0$ for $1 \le j \le k < n$. The first equation (3.2.2) with $j = n$ shows that

$$\partial_n(\partial_n f, \partial_n f) = 2\Gamma_{nnn} = \partial_n g_{nn},$$

so using the first boundary condition (3.2.3) with $j = n$ we obtain $(\partial_n f, \partial_n f) = g_{nn}$ also when $x_n \neq 0$. From the first equation (3.2.2) with $j < n$ we now obtain

$$\begin{aligned}\partial_n(\partial_n f, \partial_j f) &= \Gamma_{nnj} + (\partial_n f, \partial_j \partial_n f)\\ &= \Gamma_{nnj} + \tfrac{1}{2}\partial_j(\partial_n f, \partial_n f) = \Gamma_{nnj} + \tfrac{1}{2}\partial_j g_{nn} = \partial_n g_{nj},\end{aligned}$$

and using the boundary condition we conclude that $(\partial_n f, \partial_j f) = g_{nj}$. Now we have for $1 \le i \le j < n$

$$\begin{aligned}\partial_n(\partial_i f, \partial_j f) &= (\partial_i \partial_n f, \partial_j f) + (\partial_i f, \partial_j \partial_n f)\\ &= \partial_i(\partial_n f, \partial_j f) + \partial_j(\partial_i f, \partial_n f) - 2(\partial_n f, \partial_i\partial_j f) = \partial_i g_{jn} + \partial_j g_{in} - 2(\partial_n f, \partial_i \partial_j f).\end{aligned}$$

When $x_n = 0$ this is equal to $\partial_i g_{jn} + \partial_j g_{in} - 2\Gamma_{ijn} = \partial_n g_{ij}$ by the second set of initial conditions (3.2.3). Differentiating again we obtain

$$\partial_n^2(\partial_i f, \partial_j f) = \partial_n \partial_i g_{jn} + \partial_n \partial_j g_{in} - 2\partial_n \Gamma_{ijn} = \partial_n^2 g_{ij},$$

where we have used (3.2.4). In view of the initial conditions it follows that $(\partial_i f, \partial_j f) = g_{ij}$ for $1 \le i \le j < n$, which proves the lemma.

As already mentioned we shall start from a local isometric embedding of $\mathbf{R}^{n-1}$, and we must choose it so that the equations (3.2.3) with $j < n$ can be solved. This requires a stronger condition on the embedding, which we formulate with $\mathbf{R}^{n-1}$ replaced by $\mathbf{R}^n$ in the following definition:

Definition 3.2.2. A C^2 map $x \mapsto f(x)$ from a neighborhood of 0 in $\mathbf{R}^n$ to $\mathbf{R}^N$ is said to be *free* at 0 if the derivatives

$$\partial_i f(0), \partial_i \partial_j f(0), \quad 1 \le i \le j \le n$$

are linearly independent. Their linear hull is then called the *osculating space* at 0.

It is clear that f is free at every point in a neighborhood of 0 if f is free at 0. If a free map exists then $N \ge n + n(n+1)/2$, the dimension of the osculating space. Note that when n is replaced by $n-1$, this condition becomes $N \ge (n-1) + (n-1)n/2 = n(n+1)/2 - 1$.

Theorem 3.2.3 (Janet-Cartan). *If $\sum g_{jk}(x) dx^j dx^k$ is a real analytic Riemannian metric in a neighborhood of* 0 *in $\mathbf{R}^n$ then there is a local real analytic isometric embedding in $\mathbf{R}^N$ with $N = n(n+1)/2$. It can be chosen free if $N = n(n+3)/2$.*

Proof. When $n = 1$ the statement is obvious. We may therefore assume that $n > 1$ and that the theorem has already been proved for dimensions smaller than n. Without restriction we may assume that the coordinates are geodesic or at least that

$$g_{jk}(x) - \delta_{jk} = O(|x|^2), \quad \text{as } x \to 0.$$

Using the inductive hypothesis we choose a free real analytic map f_0 from a neighborhood of 0 in $\mathbf{R}^{n-1}$ to $\mathbf{R}^{N-1}$, where $N = n(n+1)/2$, such that for x' in a neighborhood of 0

$$(\partial_j f_0(x'), \partial_k f_0(x')) = g_{jk}(x', 0), \quad 1 \le j \le k < n.$$

We want to find f satisfying (3.2.2) so that $f(x', 0) = (f_0(x'), 0)$ and (3.2.3) is valid. The equations

$$\begin{aligned}(\partial_n f(x', 0), \partial_j f_0(x')) &= g_{nj}(x', 0), \quad 1 \le j < n,\\ (\partial_n f(x', 0), \partial_i \partial_j f_0(x')) &= \Gamma_{ijn}(x', 0), \quad 1 \le i \le j < n,\end{aligned}$$

determine uniquely the component $v(x')$ of $\partial_n f(x', 0)$ in $\mathbf{R}^{N-1}$, because f_0 is free, and $v(0) = 0$ since the right-hand sides of these equations vanish at 0. The only remaining equation (3.2.3) can now be written

$$(\partial_n f_N(x', 0))^2 = g_{nn}(x', 0) - \|v(x')\|^2$$

where $\| \cdot \|$ denotes the Euclidean norm. Since the right-hand side is positive when $x' = 0$ we have a unique analytic positive solution. Summing up, we have well defined analytic boundary conditions

$$f(x', 0) = (f_0(x'), 0), \quad \partial_n f(x', 0) = (v(x'), \sqrt{g_{nn}(x', 0) - \|v(x')\|^2}). \tag{3.2.5}$$

The last coordinate of $\partial_n f(x', 0)$ is not 0, so $\partial_j f(0)$, $1 \le j \le n$, and $\partial_i \partial_j f(0)$, $1 \le i \le j < n$, form a basis for $\mathbf{R}^N$. The equations (3.2.2) can therefore be solved for $\partial_n^2 f$,

$$\partial_n^2 f = \Phi(x, \{\partial_j f\}_{j \le n}, \{\partial_i \partial_j f\}_{i<n, j\le n}), \tag{3.2.6}$$

where Φ is analytic in a neighborhood of the initial data at 0. This is a Kovalevsky system so it has a unique analytic solution with the data (3.2.5) in a neighborhood of the origin.

It remains to show that we can get a free embedding in a space of dimension $N + n$. To do so, we first construct an embedding in $\mathbf{R}^{N+n-1}$ by changing (3.2.5) to

$$\begin{aligned} f(x', 0) &= (f_0(x'), \underbrace{0, \dots, 0}_{n \text{ times}}),\\ \partial_n f(x', 0) &= (v(x'), \sqrt{g_{nn}(x', 0) - \|v(x')\|^2 - |x'|^2}, x'). \end{aligned} \tag{3.2.5$'$}$$

The equations (3.2.2) still have a unique analytic solution (3.2.6) if we require that $\partial_n^2 f$ shall lie in the linear hull of $\partial_j f$, $1 \le j \le n$ and $\partial_i \partial_j f$, $1 \le i \le j < n$. Then it is clear that all derivatives $\partial_j f$ and $\partial_i\partial_j f$ with $i \le j$ are linearly independent at 0 with the exception of $\partial_n^2 f$, and since we have an isometric embedding in $\mathbf{R}^{N+n-1}$ we could not hope for more. If f is such an embedding for the metric $\sum g_{jk}dx^j dx^k - 4(x^n dx^n)^2$, then the embedding $(f,(x^n)^2)$ in $\mathbf{R}^{N+n}$ will be free.

If the metric is not real analytic, the proof breaks down for there is no reason then why the Cauchy problem should be solvable. Of course one can still use the completely elementary formal part of the Cauchy-Kovalevsky theorem and conclude that if $g_{jk} \in C^\infty$ then f can be chosen so that (3.2.1) holds with an error vanishing of infinite order at 0. Even for $n = 2$ it is not known whether the first part of Theorem 3.2.3 is valid in the C^∞ case when the Gauss curvature has a non-simple zero at the origin. (See e.g. Jacobowitz [2].) However, the second part of Theorem 3.2.3 remains true as will now be shown following Günther [1].

Assume now just that $g_{ij} \in C^{2+\varrho}$ for some $\varrho \in (0,1)$, that is, that $\partial^\alpha g_{ij}$ is Hölder continuous of order ϱ when $|\alpha| \le 2$. Replacing g_{ij} by the second order Taylor polynomial we obtain from Theorem 3.2.3 a free embedding f with values in $\mathbf{R}^N$, $N = n(n+3)/2$, defined in a neighborhood of 0 such that

$$\partial^\alpha(g_{ij}(x) - (\partial_i f(x), \partial_j f(x)) = O(|x|^{\varrho+2-|\alpha|}) \quad \text{as } x \to 0.$$

Choose $\varphi \in C_0^\infty(\mathbf{R}^n)$ such that $\varphi(x) = 1$ when $|x| \le 1$ and $\varphi(x) = 0$ when $|x| \ge 2$, and set for small ε

$$h_{ij}(x) = \varphi(x/\varepsilon)\big(g_{ij}(x) - (\partial_i f(x), \partial_j f(x))\big).$$

Then $|x| \le 2\varepsilon$ if $x \in \operatorname{supp} h_{ij}$, $g_{ij}(x) = h_{ij}(x) + (\partial_i f(x), \partial_j f(x))$ if $|x| < \varepsilon$, and

$$\sup|\partial^\alpha h_{ij}(x)| = O(\varepsilon^{\varrho+2-|\alpha|}), \quad \text{if } |\alpha| \le 2. \tag{3.2.7}$$

Let $\Omega \subset \mathbf{R}^n$ be a ball with center at 0 so small that f is free in $\overline{\Omega}$. It suffices to show that for small ε there is a map $u \in C^{2+\varrho}$ from $\overline{\Omega}$ to $\mathbf{R}^N$ such that $|d(f+u)|^2 = |df|^2 + \sum h_{ij}dx^i dx^j$, that is,

$$(\partial_i u, \partial_j f) + (\partial_j u, \partial_i f) + (\partial_i u, \partial_j u) = h_{ij}. \tag{3.2.8}$$

We choose ε so small that h_{ij} vanishes on $\partial\Omega$. If u vanishes on $\partial\Omega$ then (3.2.8) is equivalent to the equation obtained by letting $\Delta = \sum \partial_k^2$ act on both sides, which gives

$$\begin{aligned}\Delta h_{ij} &= \Delta\big((\partial_i u, \partial_j f) + (\partial_j u, \partial_i f)\big) + \partial_i(\Delta u, \partial_j u) + \partial_j(\Delta u, \partial_i u) + T_{ij}(u)\\ &= \partial_i\big((\Delta u, \partial_j u) + \Delta(\partial_j f, u)\big) + \partial_j\big((\Delta u, \partial_i u) + \Delta(\partial_i f, u)\big) - 2\Delta(\partial_i\partial_j f, u) + T_{ij}(u),\end{aligned}$$

where $T_{ij}(u)$ is quadratic in the second derivatives of u. We can simplify the equations by adding just n equations which annihilate the first two terms in all these equations and conclude that (3.2.8) is valid if

$$\begin{aligned}\Delta(\partial_i f, u) &= -(\partial_i u, \Delta u), \quad 1 \le i \le n,\\ \Delta(\partial_i\partial_j f, u) &= \tfrac{1}{2}(T_{ij}(u) - \Delta h_{ij}), \quad 1 \le i \le j \le n,\end{aligned}$$

in Ω and $u = 0$ on $\partial\Omega$. Let G be the operator solving the Dirichlet problem for the Laplacian in Ω, thus $\Delta G\psi = \psi$ in Ω and $G\psi = 0$ on $\partial\Omega$ if $\psi \in C(\overline{\Omega})$. Then these conditions are fulfilled if

$$(3.2.9)\qquad \begin{aligned} (\partial_i f, u) &= -G(\partial_i u, \Delta u), \quad 1 \le i \le n, \\ (\partial_i \partial_j f, u) &= \tfrac{1}{2} G(T_{ij}(u) - \Delta h_{ij}), \quad 1 \le i \le j \le n. \end{aligned}$$

The vectors $\partial_i f(x)$, $\partial_i \partial_j f(x)$ with $1 \le i \le j \le n$ form a basis for $\mathbf{R}^N$ for every $x \in \overline{\Omega}$ since f is free, so we have for every $U \in \mathbf{R}^N$

$$U = \sum_{1 \le i \le n} \varphi_i (\partial_i f, U) + \sum_{1 \le i \le j \le n} \varphi_{ij} (\partial_i \partial_j f, U),$$

where φ_i, φ_{ij} are analytic in $\overline{\Omega}$. Thus the equations (3.2.9) are equivalent to $u = H + T(u)$ where

$$H = -\tfrac{1}{2} \sum_{1 \le i \le j \le n} \varphi_{ij} h_{ij}, \quad T(u) = -\sum_{1 \le i \le n} \varphi_i G(\partial_i u, \Delta u) + \tfrac{1}{2} \sum_{1 \le i \le j \le n} \varphi_{ij} G T_{ij}(u).$$

Standard Hölder estimates in the theory of elliptic equations show that $T(u)$ is continuous in $C^{2+\varrho}$ and a contraction operator in the convex subset where $\sum_{|\alpha| \le 2} \sup |\partial^\alpha u|$ is small enough. In view of (3.2.7) it follows that the equation $u = H + T(u)$ has a unique solution there if ε is small enough. Arguing by standard elliptic theory we conclude that u is as smooth as h if h has additional regularity. Admitting these basic facts without proof here, we obtain:

Theorem 3.2.4. *If $\sum g_{jk}(x) dx^j dx^k$ is a Riemannian metric in a neighborhood of* 0 *in $\mathbf{R}^n$ with coefficients in $C^{2+\varrho}$ for some non-integer $\varrho > 0$, then there is a local $C^{2+\varrho}$ free isometric embedding in $\mathbf{R}^N$ with $N = n(n+3)/2$.*

Note that the solution we have found to the first order equations (3.2.8) is not smoother than the right hand side. This is unavoidable because of the Gauss equations and gives rise to the analytical difficulties of the problem which were first overcome by Nash [1]. A great deal is known also about global isometric embeddings. We must confine ourselves here to referring to Gromov and Rohlin [1], Berger, Bryant and Griffiths [1], and the references in these papers.

We shall finally justify the statement made above that a smooth isometric embedding in $\mathbf{R}^N$ does not exist in general unless $N \ge n(n+1)/2$. It is notationally more convenient to prove a general form of this statement. Let u denote a C^∞ function defined in a neighborhood of 0 in $\mathbf{R}^n$ with values in $\mathbf{R}^N$. Denote by $J_m u$ the m-jet of u,

$$J_m u = \{\partial^\alpha u\}_{|\alpha| \le m},$$

and let $\Phi(x, J_m u)$ be a C^∞ function of x and $J_m u$ with values in $\mathbf{R}^\nu$. Thus the equation $\Phi(x, J_m u) = v$ is a general system of ν differential equations of order m for N unknowns.

Theorem 3.2.5. *If for all v in an open subset of $C^\infty(\mathbf{R}^n, \mathbf{R}^\nu)$ one can find a function $u \in C^\infty(\mathbf{R}^n, \mathbf{R}^N)$ such that $\Phi(x, J_m u) = v$ in a neighborhood of 0, then $N \ge \nu$.*

Proof. Let μ be some large integer to be chosen later. We shall not really use the full differential equation but only the much weaker condition

$$J_\mu \Phi(x, J_m u)(0) = J_\mu v(0). \tag{3.2.10}$$

The space E_μ of μ-jets at 0 of functions in $C^\infty(\mathbf{R}^n, \mathbf{R}^\nu)$, to which the right-hand side belongs, has dimension $\nu\binom{\mu+n}{n}$, for $\binom{\mu+n}{n}$ is the number of multi-indices $\alpha = (\alpha_1, \dots, \alpha_n)$ with $\alpha_1 + \dots + \alpha_n \le \mu$. The left-hand side is a C^∞ function of $J_{\mu+m}u(0)$, which belongs to a space of dimension $N\binom{\mu+m+n}{n}$. By the Morse-Sard theorem the range is of measure 0 in E_μ if

$$N\binom{\mu+m+n}{n} < \nu\binom{\mu+n}{n}, \quad \text{that is, } N < \nu \prod_{j=1}^{n} \frac{\mu+j}{\mu+m+j}.$$

If $N < \nu$ then this condition is fulfilled for large μ, which proves the theorem and even more: If $N < \nu$ then the equation $\Phi(x, J_m u) = v$ cannot even be satisfied to order μ at a fixed point when μ is large, unless $J_\mu v$ is exceptional there in the sense that it belongs to a set of measure 0.

For the isometric embedding problem one can improve the estimate of μ obtained from the preceding proof by using the special properties of the equations. (See Exercise 2.3.1 and Gromov-Rokhlin [1].) Whether local isometric embedding of low regularity is possible in dimensions below the Janet-Cartan dimension $n(n+1)/2$ does not seem to be known except for the result of Nash [2], Kuipers [1]; they showed that even a global C^1 isometric embedding is possible in essentially the same dimension where a C^1 embedding exists, hence always in $2n+1$ dimensions. However, for C^k embeddings with $k \ge 2$ the situation is quite different as shown by (2.3.5), because the Gauss equations are then available, and the problem does not seem to have been studied then.

3.3. Spaces of constant curvature. Let M be a connected Riemannian manifold of dimension n, and recall from Section 2.3 that the sectional curvature of M at a point $x \in M$ for the two plane spanned by $t^1, t^2 \in T_x M$ is defined by

$$R(t^1, t^2, t^1, t^2)/(g(t^1, t^1)g(t^2, t^2) - g(t^1, t^2)^2). \tag{3.3.1}$$

If this is a function $f(x)$ independent of the direction of the two plane then

$$R(t^1, t^2, t^3, t^4) = f(x)(g(t^1, t^3)g(t^2, t^4) - g(t^1, t^4)g(t^2, t^3))$$

by the uniqueness statement in Theorem 2.3.1, for the right-hand side has all the symmetry properties defining $\mathcal{T}$. Thus the Weyl tensor and the traceless Ricci tensor are equal to 0. If this is true for every point in M, then M is an Einstein manifold, and if $n > 2$ it follows from Theorem 3.1.8 that f is a constant, so the sectional curvature is independent of x also. (This is a classical theorem of F. Schur, far older than Theorem 3.1.8.)

Definition 3.3.1. A connected Riemannian manifold M is said to have constant curvature K if the curvature tensor is given by

$$R_{ijkl} = K(g_{ik}g_{jl} - g_{il}g_{jk}), \tag{3.3.2}$$

with a constant K.

From (3.3.2) it follows that the Ricci tensor is $(n-1)Kg_{ij}$ and that the scalar curvature is $n(n-1)K$. Thus a manifold of constant curvature is an Einstein manifold but the converse is not true when $n > 3$ since the Weyl tensor may not be equal to 0 then. Since the covariant derivative of the metric tensor is equal to 0, this is also true for the curvature tensor (3.3.2).

An obvious example of a manifold of constant zero curvature is $\mathbf{R}^n$ with the standard Euclidean metric. We shall now prove that locally there are no others.

Theorem 3.3.2. *A Riemannian manifold M with curvature tensor identically equal to* 0 *is flat in the sense that at every point one can choose local coordinates such that the metric is $\sum(dx^j)^2$.*

Proof. Choose first some arbitrary local coordinates $x = (x^1, \dots, x^n)$ varying over a ball Ω with center at the origin. Set $\nabla_i = \nabla_{\partial_i}$ where $\partial_i = \partial/\partial x^i$. Since $[\partial_i, \partial_j] = 0$ it follows from (3.1.6) and the hypothesis that $[\nabla_i, \nabla_j] = 0$. For every $v_0 \in \mathbf{R}^n$ we can therefore find a unique vector field v in Ω with $v(0) = v_0$ and $\nabla_i v = 0$, $i = 1, \dots, n$. In fact, we shall prove inductively for $\nu = 1, \dots, n$ that there is such a vector field defined in $\Omega_\nu = \{x \in \Omega; x^j = 0, j > \nu\}$ with $\nabla_i v = 0$, $i \le \nu$. This is obvious when $\nu = 1$, for the equation $\nabla_1 v = 0$ is just a linear system of differential equations with leading term $\partial v/\partial x^1$ which we can solve with initial value v_0. If $\nu > 1$ and $\tilde v$ has the required property with ν replaced by $\nu - 1$, then we can find v defined in Ω_ν with $\nabla_\nu v = 0$ and $v = \tilde v$ when $x_\nu = 0$. If $\mu < \nu$ it follows that $\nabla_\nu \nabla_\mu v = \nabla_\mu \nabla_\nu v = 0$ in Ω_ν, and since $\nabla_\mu v = \nabla_\mu \tilde v = 0$ when $x_\nu = 0$, it follows that $\nabla_\mu v = 0$ in Ω_ν. When $\nu = n$, the claim is proved.

Now choose vector fields $v_1, \dots, v_n$ in Ω such that $\nabla v_j = 0$ for $j = 1, \dots, n$ and $v_1, \dots, v_n$ form an orthonormal basis at 0. Since $d(v_j, v_k) = (\nabla v_j, v_k) + (v_j, \nabla v_k) = 0$, this follows at every point in Ω. We have

$$[v_j, v_k] = \nabla_{v_j} v_k - \nabla_{v_k} v_j = 0,$$

so there are new coordinates y in a neighborhood of 0 such that $v_j = \partial/\partial y^j$, $j = 1, \dots, n$. But this means that the metric is the Euclidean metric in these coordinates.

An equally obvious example of a manifold of constant positive curvature K is a sphere with radius $R = K^{-\frac{1}{2}}$ in $\mathbf{R}^{n+1}$. In fact, $\mathbf{O}(n+1)$ is a transitive group of isometries, and the isotropy group $\mathbf{O}(n)$ leaving a given point $x \in S^n$ fixed is transitive on the two planes in the tangent plane at x. Analytically the sphere is defined by

$|x|^2 + (x^{n+1})^2 = R^2$ where $x = (x^1, \dots, x^n)$. In the hemisphere where $x^{n+1} > 0$ we can use x as parameter, and noting that $\langle x, dx\rangle + x^{n+1}dx^{n+1} = 0$ we obtain

$$ds^2 = |dx|^2 + \langle x, dx\rangle^2/(x^{n+1})^2 = |dx|^2 + K\langle x, dx\rangle^2 + O(|x|^4)$$

where $K = 1/R^2$. Thus $g_{ij} = \delta_{ij} + Kx^i x^j + O(|x|^4)$, and we obtain when $x = 0$ by (3.1.14)

$$R_{ijkl} = \tfrac{K}{2}(\partial_j\partial_k x^i x^l + \partial_i\partial_l x^j x^k - \partial_i\partial_k x^j x^l - \partial_j\partial_l x^i x^k) = K(\delta_{jl}\delta_{ik} - \delta_{il}\delta_{kj}).$$

This we knew already of course, for when $n = 2$ the total curvature is $1/R^2 = K$.

For the restriction to the hyperboloid $H = \{x \in \mathbf{R}^{n+1}; (x^{n+1})^2 = |x|^2 + R^2, x^{n+1} > 0\}$ of the hyperbolic (Lorentz) metric $|dx|^2 - |dx^{n+1}|^2$ we have

$$ds^2 = |dx|^2 - \langle x, dx\rangle^2/(x^{n+1})^2 = |dx|^2 + K\langle x, dx\rangle^2 + O(|x|^4)$$

where $K = -1/R^2 < 0$ now. The metric is positive definite since $|x|^2 < (x^{n+1})^2$ on H. The preceding calculation for the sphere gives that

$$R_{ijkl} = K(\delta_{jl}\delta_{ik} - \delta_{il}\delta_{kj})$$

at the origin. To prove that the curvature is constant it suffices to note that the group of Lorentz transformations, that is, linear transformations preserving the Lorentz form $\sum_1^n (x^j)^2 - (x^{n+1})^2$ acts isometrically and transitively on H.

The parametrization of S^n above covers only a half sphere. We can cover the whole sphere minus one point by means of the stereographic projection. If $X \in S^n$ then the stereographic projection $x \in \mathbf{R}^n$ from the point $(0, \dots, 0, -R)$ to the tangent plane at the antipodal point is obtained from the equations

$$X = (0, \dots, 0, -R) + t(x, 2R), \quad |X| = R, \quad \text{that is, } t = 4R^2/(4R^2 + |x|^2).$$

Thus

$$|dX|^2 = |tdx + xdt|^2 + 4R^2dt^2 = t^2|dx|^2 + (4R^2 + |x|^2)dt^2 + 2tdt\langle x, dx\rangle = t^2|dx|^2,$$

for $dt(4R^2 + |x|^2) + 2t\langle x, dx\rangle = 0$. Hence the metric is

$$ds^2 = (1 + |x|^2/4R^2)^{-2}|dx|^2 = |dx|^2(1 - \tfrac{K}{2}|x|^2) + O(|x|^4)$$

where $K = 1/R^2$ again. We leave as an exercise to calculate the curvature at 0 using this expression for the metric.

For the hyperboloid H we can similarly use stereographic projection from $(0, \dots, 0, -R)$ on the tangent plane at $(0, \dots, 0, R)$. This gives the equations

$$X = (0, \dots, 0, -R) + t(x, 2R), \quad X \in H, \quad \text{that is, } t = 4R^2/(4R^2 - |x|^2).$$

The stereographic projection just fills the open ball with radius $2R$ now. We obtain

$$\sum_1^n |dX^j|^2 - |dX^{n+1}|^2 = |tdx + xdt|^2 - 4R^2dt^2$$
$$= t^2|dx|^2 + (|x|^2 - 4R^2)dt^2 + 2tdt\langle x, dx\rangle = t^2|dx|^2.$$

With $K = -1/R^2$ we therefore obtain the metric

$$ds^2 = (1 + K|x|^2/4)^{-2}|dx|^2. \tag{3.3.3}$$

For $K = -1$ this is the Poincaré model of non-Euclidean geometry. By an inversion we pass to the half space model. First we move the boundary point $(0, \ldots, 0, -2R)$ to 0 by introducing

$$y = x + (0, \ldots, 0, 2R); \quad 4R^2 - |x|^2 = -|y|^2 + 4Ry^n.$$

With $z = y/|y|^2$ we make an inversion at the origin and obtain

$$|dx|^2 = |dy|^2 = |dz/|z|^2 - 2z\langle z, dz\rangle/|z|^4|^2 = |dz|^2/|z|^4,$$
$$4Ry^n - |y|^2 = (4Rz^n - 1)/|z|^2$$

so the metric is $16R^4(4Rz^n - 1)^{-2}|dz|^2$ or after a translation, $x = z - (0, \ldots, 0, 1/4R)$

$$ds^2 = R^2|dx|^2/(x^n)^2. \tag{3.3.4}$$

Exercise 3.3.1. Verify directly that the metric (3.3.4) in the half space where $x_n > 0$ has constant curvature $-1/R^2$.

Exercise 3.3.2. Find geodesic coordinates at 0 for the metric (3.3.3).

We obtained the Poincaré model starting from an immersion in Minkowski space. Schur [1] gave an isometric immersion of a part of the n dimensional hyperbolic space with curvature -1 in $\mathbf{R}^{2n-1}$ as follows. Set

$$X^1 + iX^2 = e^{ix^1}/x^n, \ldots, X^{2n-3} + iX^{2n-2} = e^{ix^{n-1}}/x^n.$$

Then

$$\sum_1^{2n-1}(dX^j)^2 = (x^n)^{-2}\sum_1^{n-1}(dx^j)^2 + (n-1)(x^n)^{-4}(dx^n)^2 + (dX^{2n-1})^2$$

which is equal to the Poincaré metric if

$$(dX^{2n-1})^2 = (dx^n)^2((x^n)^2 - (n-1))(x^n)^{-4}.$$

Choosing an integral of this differential equation we obtain an isometric immersion of the half space where $x^n > \sqrt{n-1}$. According to a reference in Gromov and Rokhlin [1] it was proved in Liber [1] that a local isometric embedding is not possible in $\mathbf{R}^{2n-2}$. A classical result of Hilbert states that for $n = 2$ it is not possible to embed the whole hyperbolic plane in $\mathbf{R}^3$; this was extended by Efimov [1] to arbitrary complete two dimensional surfaces with a negative upper bound for the curvature.

We shall now prove an extension of Theorem 3.3.2 to any constant curvature:

Theorem 3.3.3. *A Riemannian manifold with constant curvature K is isometric near any point to a sphere, if $K > 0$, to $\mathbf{R}^n$ if $K = 0$ and to hyperbolic space if $K < 0$.*

Proof. Assume that we have geodesic coordinates in a ball Ω with center at 0, $g_{ij}(0) = \delta_{ij}$. Then the theorem states that the metric tensor g_{ij} is uniquely determined by K in a neighborhood of the origin. This will be proved by deriving differential equations along the rays, that is, the integral curves of the radial vector field

$$\varrho = \sum_1^n x^j \partial_j.$$

We shall prove in Lemma 3.3.4 below that for $j = 1, \dots, n$ there is a unique vector field e_j such that

$$\nabla_\varrho e_j = 0, \quad e_j(0) = \partial_j. \tag{3.3.5}$$

(This is not quite obvious since the radial vector field vanishes at 0.) Since

$$\nabla_\varrho (e_j, e_k) = (\nabla_\varrho e_j, e_k) + (e_j, \nabla_\varrho e_k) = 0,$$

where $(\cdot, \cdot)$ denotes scalar product in the Riemannian metric, it follows that $e_1, \dots, e_n$ are an orthonormal system at every point in Ω. Let $\theta^1, \dots, \theta^n$ be the one forms biorthogonal to $e_1, \dots, e_n$, that is,

$$\langle \theta^j, e_k \rangle = \delta^j_k, \quad j, k = 1, \dots, n.$$

Then $t = \sum_i \langle \theta^i, t \rangle e_i$ for every tangent vector, so

$$\sum g_{jk} t^j t^k = \sum_i \langle \theta^i, t \rangle^2.$$

Writing

$$\theta^i = \sum A^i_j dx^j, \quad \text{that is, } A^i_j = \langle \theta^i, \partial_j \rangle,$$

we obtain $g_{jk} = \sum_i A^i_j A^i_k$, so it is enough to find the coefficients A^i_j.

Since $\langle \nabla_\varrho \theta^j, e_k \rangle = 0$ for $j, k = 1, \dots, n$, we have $\nabla_\varrho \theta^j = 0$ also. Thus, by (3.1.3),

$$\varrho A^i_j = \langle \theta^i, \nabla_\varrho \partial_j \rangle = \langle \theta^i, \nabla_j \varrho - \partial_j \rangle,$$

for $[\varrho, \partial_j] = -\partial_j$. Now $\langle \theta^i, \varrho \rangle = x^i$, for $\nabla_\varrho \varrho = \varrho$ since the radial direction is parallel, so $\varrho \langle \theta^i, \varrho \rangle = \langle \theta^i, \varrho \rangle$, which means that $\langle \theta^i, \varrho \rangle$ is homogeneous of degree 1. At the origin we have $\theta^i = dx^i$, so the assertion follows. It means that $\varrho = \sum x^i e_i$. Thus

$$(\varrho + 1) A^i_j = \langle \theta^i, \nabla_j \varrho \rangle = \partial_j x^i - \langle \nabla_j \theta^i, \varrho \rangle = \delta^i_j - \sum x^k \langle \nabla_j \theta^i, e_k \rangle = \delta^i_j + \sum x^k B^i_{kj}.$$

Here we have introduced

$$B^i_{jk} = \langle \theta^i, \nabla_k e_j \rangle$$

and must now obtain a differential equation for this new quantity. By (3.1.6)

$$\varrho B^i_{jk} = \langle \theta^i, \nabla_\varrho \nabla_k e_j \rangle = -\langle \theta^i, \nabla_k e_j \rangle + R(e_i, e_j, \varrho, \partial_k),$$

for $[\varrho, \partial_k] = -\partial_k$ and $\nabla_\varrho e_j = 0$. Thus

$$(\varrho + 1) B^i_{jk} = \sum_l R(e_i, e_j, \varrho, e_l) \langle \theta^l, \partial_k \rangle.$$

Since the covariant derivative of R is equal to 0, we have

$$\varrho R(e_i, e_j, \varrho, e_l) = R(e_i, e_j, \nabla_\varrho \varrho, e_l) = R(e_i, e_j, \varrho, e_l),$$

so this is a homogeneous function of degree 1. At the origin we have

$$R(e_i, e_j, \partial_m, e_l) = R_{ijml} = K(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}),$$

so it follows that

$$R(e_i, e_j, \varrho, e_l) = \sum_m x^m R_{ijml}.$$

Summing up and introducing polar coordinates $x = r\omega$, we have the differential equations

$$\begin{aligned} \frac{d}{dr}(rA^i_j) &= \delta^i_j + \sum r B^i_{kj} \omega^k, \\ \frac{d}{dr}(rB^i_{jk}) &= \sum_{l,m} R_{ijml} \omega^m r A^l_k \end{aligned}$$

for the functions rA^i_j and rB^i_{jk} which vanish at the origin. This determines them uniquely and proves the theorem when we have established the following

Lemma 3.3.4. *Let ϱ be the radial vector field in the ball Ω in a geodesic system of coordinates. For every $v_0 \in \mathbf{R}^n$ one can then find a unique C^∞ vector field v with $v(0) = v_0$ and $\nabla_\varrho v = 0$. It follows that $\nabla v = 0$ at the origin.*

Proof. We have to solve a singular Cauchy problem of the form

$$\sum x^j \partial_j v + Av = 0, \quad v(0) = v_0, \tag{3.3.6}$$

where $(Av)^j = \sum \Gamma_{ik}{}^j x^k v^i$ has coefficients vanishing of second order at the origin. If we introduce polar coordinates, $x = r\omega$, the problem takes the form

$$\partial v^j / \partial r + \sum \Gamma_{ik}{}^j (r\omega) \omega^k v^i = 0, \quad v = v_0 \text{ when } r = 0.$$

It is immediately clear that we have a unique solution in $[0, R) \times S^{n-1}$, if Ω is the ball with radius R, and it is a C^∞ function there. We have to verify that it is a C^∞

function in the original variables $x^1, \dots, x^n$. To do so we observe that (3.3.6) can be solved in terms of formal power series

$$v = v_0 + v_1 + v_2 + \dots$$

where v_j is homogeneous of degree j. In fact, (3.3.6) can be written

$$v_1 + 2v_2 + \dots + A(v_0 + v_1 + \dots) = 0$$

and gives $v_1 = 0$, $2v_2^j + \sum \gamma_{ik}{}^j x^k v_0^i$ and so on, where $\gamma_{ik}{}^j$ is the first order Taylor expansion of $\Gamma_{ik}{}^j$. Now write $v = v_0 + \dots + v_{N-1} + w$ for some large N. Then $w = 0$ when $r = 0$, and

$$\partial w^j/\partial r + \sum \Gamma_{ik}{}^j(r\omega)\omega^k w^i = f^j(r, \omega)$$

where $f = O(r^{N-1})$. This implies that $w = O(r^N)$, and since $\partial/\partial x^\nu = \omega_\nu \partial/\partial r + \langle \partial\omega/\partial x^\nu, \partial/\partial\omega \rangle$, where $\partial\omega/\partial x^\nu$ is homogeneous of degree -1, it follows that $w \in C^{N-1}$ as a function of the original variables, the derivatives vanishing at 0. This completes the proof of the lemma and of the theorem.

The tools introduced to prove Theorem 3.3.3 also give another useful result on the connection between the curvature tensor and the metric tensor.

Theorem 3.3.5. *Let $\sum g_{jk}(x)dx^j dx^k$ be the metric form for a geodesic coordinate system centered at 0, thus $\sum g_{jk}(x)x^k = x^j$. Any derivative $\partial^\alpha g_{jk}(0)$ can then be expressed as a polynomial in the components of the curvature tensor and its covariant derivatives of order $\le |\alpha| - 2$.*

Proof. It suffices to prove the theorem with covariant derivatives replaced by derivatives ∂^α with respect to the coordinates. In fact, the difference

$$R_{ijkl,\alpha_1\dots\alpha_m} - \partial_{\alpha_m} \dots \partial_{\alpha_1} R_{ijkl} \tag{3.3.7}$$

is at the origin a polynomial in derivatives of R of order $< m$ and derivatives of the metric tensor g of order $\le m$ since the Christoffel symbol just contains the first order derivatives of g. If we already know that these derivatives of g can be expressed in terms of derivatives of R of order $\le m - 2$, it follows that (3.3.7) can be expressed in terms of derivatives of R of order $< m$. By induction with respect to m we therefore conclude that every polynomial in derivatives of R of order $\le m$ can be written as a polynomial in covariant derivatives of R of order $\le m$.

Thus it suffices to examine how the Taylor expansion of R determines that of g. Since $(g_{ij}(0))$ is the identity it suffices in fact to consider the inverse matrix g^{ij}, which is somewhat more convenient since for every $t \in \mathbf{R}^n$, regarded as a covector,

$$\sum g^{ij} t_i t_j = \sum \langle e_i, t \rangle^2, \quad \text{hence } g^{ij} = \sum_m e_m{}^i e_m{}^j.$$

Here e_m are the vector fields in the proof of Theorem 3.3.3, said to form a *synchronous frame* for TM. If we prove that derivatives of e_m of order $\le \nu$ at 0 are polynomials in those of R of order $\le \nu - 2$, the theorem will be proved.

In the proof of Theorem 3.3.3 we saw that

$$\langle \theta^i, (\nabla_\varrho + 1)\nabla_l e_m\rangle = R(e_i, e_m, \varrho, \partial_l), \quad \text{that is, } ((\nabla_\varrho + 1)\nabla_l e_m)^i = \sum_{j,k} R^i{}_{jkl} e_m^j x^k.$$

The operator $(\varrho + 1)$ multiplies a homogeneous function by the degree of homogeneity plus 1. If we take the homogeneous terms of order μ in the Taylor expansion we find for $\mu \geq 0$ that the derivatives of order μ of $\nabla_l e_m$ at 0 are polynomials in those of R and of e_m, of total order at most $\mu - 1$. Since the first order derivatives of e_m vanish at 0 it follows inductively that all derivatives of e_m at 0 of order $\mu + 1$ are polynomials in those of R of order $\leq \mu - 1$, which completes the proof. Taking the first order term in $\nabla_l e_m$ above we also obtain

$$(\nabla_l e_m)^i = \tfrac{1}{2}\sum_k R^i{}_{mkl}(0)x^k + O(|x|^2),$$

which gives the following result which will be needed in Section 6.10:

Theorem 3.3.6. *If $e_1, \dots, e_n$ is the synchronous frame, equal to the basis vectors at the origin of a geodesic coordinate system, then*

$$\nabla_l e_j = \tfrac{1}{2}\sum_{i,k} R^i{}_{jkl}(0)x^k e_i + O(|x|^2),$$

and the same result holds for the dual frame in T^*M.

Exercise 3.3.3. Show that with the notation in Theorem 3.3.6 $\partial_k(\nabla_l e_j - \Gamma_{jl}{}^i e_i)(0)$ is symmetric in k and l. Express $\sum g_{jk}(x)dx^j dx^k$ in terms of $R(0)$ with an error $O(|x|^3)$.

Remark. For a manifold of constant curvature, the curvature tensor is a polynomial in the metric tensor. Hence it follows inductively from Theorem 3.3.5 that we can calculate all derivatives of the metric tensor at the origin of a geodesic system of coordinates. If we already knew that the metric is analytic, this would give another proof of Theorem 3.3.3, so the connection between Theorems 3.3.3 and 3.3.5 is quite close.

Exercise 3.3.4. Let M and N be Riemannian manifolds of the same constant curvature K such that N is complete. Let $\gamma : [0,1] \ni t \mapsto x(t) \in M$ be a smooth arc, and let f_0 be an isometry of a neighborhood of $x(0)$ on a neighborhood of some point $y(0) \in N$.

(1) Prove that, restricting f_0 if necessary, one can find a smooth arc $\tilde\gamma : [0,1] \ni t \mapsto y(t) \in N$ and for every $t \in [0,1]$ an isometry f_t of a neighborhood $U(t)$ of $x(t)$ on a neighborhood of $y(t)$ such that $f_t = f_s$ in $U(t) \cap U(s)$ for all $s, t \in [0,1]$.
(2) Show that $\tilde\gamma$ and the germ of f_1 at $x(1)$ are uniquely determined by γ and the germ of f_0 at $x(0)$.
(3) Prove that if M is simply connected, then $y(1)$ and the germ of f_1 are uniquely determined by $x(0)$, $x(1)$ and f_0.
(4) Deduce that if M is complete and simply connected, then f_0 can be extended to a locally isometric map $f : M \to N$, and that M is then for $K > 0$, $K = 0$ or $K < 0$ isomorphic to a sphere, to $\mathbf{R}^n$, or to hyperbolic space.

3.4. Conformal geometry. A Riemannian metric allows one to measure both angles and distances. In some cases only the angles are interesting, so metrics differing by a scalar factor give the same results. Also in analytical contexts where the metric originates from the principal part of a second order differential operator, one is often able to simplify by multiplication of the equation and therefore the metric by a non-vanishing factor.

Definition 3.4.1. If M and $\widetilde{M}$ are Riemannian manifolds with metrics g and $\tilde{g}$, then a diffeomorphism $f : M \to \widetilde{M}$ is said to be conformal if $f^*\tilde{g} = e^{2\varphi}g$ for some smooth φ. When $M = \widetilde{M}$ we shall assume that f is the identity unless otherwise stated.

The two dimensional case is very special so we shall discuss it first.

Theorem 3.4.2. *Every Riemannian manifold M of dimension 2 is locally conformal to the flat space $\mathbf{R}^2$.*

Proof. A metric $\sum_{j,k=1}^2 g_{jk}dx^j dx^k$ in a neighborhood of 0 in $\mathbf{R}^2$ is conformal to the Euclidean metric if and only if $g_{12} = 0$ and $g_{11} = g_{22}$. This means precisely that the one forms dx^1 and dx^2 are orthogonal and of equal pointwise norm.

Assuming that $g_{jk} = \delta_{jk} + O(|x|^2)$, as we may, we set $\omega_1 = dx^1$ and

$$\omega_2 = (g^{11}dx^2 - g^{12}dx^1)/\sqrt{g^{11}g^{22} - (g^{12})^2}.$$

Then ω_1 and ω_2 are orthogonal and have the same norm; $\omega_2 - dx^2$ is $O(|x|^2)$. We can choose a complex valued function u with $u(0) \neq 0$ such that $d(u(\omega_1 + i\omega_2)) = 0$, for this means that

$$du \wedge (\omega_1 + i\omega_2) + ud(\omega_1 + i\omega_2) = 0,$$

which reduces to $\partial u/\partial x^1 + i\partial u/\partial x^2 = 0$ at the origin, so it is an elliptic differential equation. In a connected neighborhood of the origin we can now write

$$u(\omega_1 + i\omega_2) = dy^1 + idy^2$$

where y^1 and y^2 are real valued and dy^1, dy^2 are orthogonal and of equal norm, for dy^1 and dy^2 are obtained from ω_1 and ω_2 by an orthogonal transformation and a rotation. If we take y^1 and y^2 as new coordinates, the metric is a multiple of the Euclidean metric $(dy^1)^2 + (dy^2)^2$, which proves the theorem.

Remark. What we have used here is that an oriented two dimensional Riemannian manifold has an analytic structure; we have chosen $y^1 + iy^2$ analytic with respect to it.

From now on we assume that the dimension n is at least 3.

Theorem 3.4.3. *If $\tilde{g} = e^{2\varphi}g$, then we have for the corresponding curvature tensors*

$$e^{-2\varphi}\widetilde{R}_{ijkl} = R_{ijkl} + g_{il}\varphi_{jk} + g_{jk}\varphi_{il} - g_{ik}\varphi_{jl} - g_{jl}\varphi_{ik} + (g_{il}g_{jk} - g_{ik}g_{jl})|\nabla\varphi|^2 \tag{3.4.1}$$

where $\varphi_{ij} = \varphi_{,ij} - \partial_i\varphi\partial_j\varphi$, calculated in terms of the metric g.

A direct verification is possible but laborious and uninteresting. We shall therefore prepare the proof with an elementary lemma which allows us to avoid messy calculations.

Lemma 3.4.4. *If $a \in \mathbf{R}^n$ then the metric $(1+2\langle x,a\rangle + |x|^2|a|^2)^{-2}|dx|^2$ in $\mathbf{R}^n$ is flat except at the singularity where $|a|^2x = -a \neq 0$.*

Proof. If we take $y = |a|^2x + a$, the metric becomes $|y|^{-4}|dy|^2$, and the inversion $z = y/|y|^2$ reduces it to $|dz|^2$, so it is flat. (See also the argument leading to (3.3.4).)

Proof of Theorem 3.4.3. If φ is a constant, the statement is obvious, so we may assume that $\varphi(0) = 0$. Next assume that g is the Euclidean metric in $\mathbf{R}^n$ and that

$$\varphi(x) = \varphi_a(x) = -\log(1+2\langle x,a\rangle + |x|^2|a|^2) = -2\langle x,a\rangle - |x|^2|a|^2 + 2\langle x,a\rangle^2 + O(|x|^3)$$

as for the conformal factor in Lemma 3.4.4. Then we have

$$\varphi_{ij} = 2(2a_ia_j - \delta_{ij}|a|^2 - 2a_ia_j) = -2\delta_{ij}|a|^2$$

and it follows at once that the right-hand side of (3.4.1) vanishes, so Theorem 3.4.3 is valid in this case. In the general case we write $\varphi(x) = \varphi_a(x) + q(x) + O(|x|^3)$ where q is quadratic, and we assume that $g_{jk}(x) = \delta_{jk} + G_{jk}(x) + O(|x|^3)$ where G_{jk} is a quadratic form. Then

$$\tilde{g} = e^{2\varphi_a}g_0 + G + 2qg_0 + O(|x|^3); \qquad g_0 = \sum (dx^j)^2.$$

Only the first term contributes to the Christoffel symbols $\widetilde{\Gamma}_{ijk}$ for the metric $\tilde{g}$ at 0. From the linearity of the formula (2.1.13)$'$ in the second order derivatives we conclude that $\widetilde{R}$ is the sum of the curvature tensor for the metric $e^{2\varphi_a}g_0$, which is 0, the curvature tensor R and the curvature tensor for the metric $g_0 + 2qg_0$. The latter has the $ijkl$ component

$$g_{il}\partial_j\partial_k q + g_{jk}\partial_i\partial_l q - g_{jl}\partial_i\partial_k q - g_{ik}\partial_j\partial_l q$$

which agrees with the terms in (3.4.1) of second order in φ. The proof is complete.

If we contract (3.4.1) in the indices jl we obtain the transformation law for the Ricci curvature:

$$\widetilde{R}_{ik} = R_{ik} + (2-n)\varphi_{ik} - \Big(\sum g^{jl}\varphi_{jl} + (n-1)|\nabla\varphi|^2\Big)g_{ik}. \tag{3.4.2}$$

Another contraction gives the transformation rule for the scalar curvature:

$$e^{2\varphi}\widetilde{S} = S + 2(1-n)\Delta\varphi - (n-1)(n-2)|\nabla\varphi|^2, \tag{3.4.3}$$

where we have introduced the Laplace-Beltrami operator

$$\Delta\varphi = \sum g^{ij}\varphi_{,ij}. \tag{3.4.4}$$

With this notation we can rewrite (3.4.2) in the form

$$\widetilde{R}_{ij} = R_{ij} + (2-n)\varphi_{ij} - (\Delta\varphi + (n-2)|\nabla\varphi|^2)g_{ij}. \tag{3.4.2$'$}$$

If we multiply (3.4.3) by $g_{ij}/(2-2n)$ and add to (3.4.2)$'$, we eliminate $\Delta\varphi$ and obtain

$$\widetilde{R}_{ij} + \widetilde{S}\tilde{g}_{ij}/(2-2n) = R_{ij} + Sg_{ij}/(2-2n) + (2-n)\Phi_{ij}, \quad \Phi_{ij} = \varphi_{ij} + \tfrac{1}{2}|\nabla\varphi|^2 g_{ij}.$$

If we note that (3.4.1) can be written

$$e^{-2\varphi}\widetilde{R}_{ijkl} = R_{ijkl} + g_{il}\Phi_{jk} + g_{jk}\Phi_{il} - g_{ik}\Phi_{jl} - g_{jl}\Phi_{ik} \tag{3.4.1$'$}$$

it follows that

$$\begin{aligned} e^{-2\varphi}\Big(\widetilde{R}_{ijkl} + \frac{\tilde{g}_{il}\widetilde{R}_{jk} + \tilde{g}_{jk}\widetilde{R}_{il} - \tilde{g}_{ik}\widetilde{R}_{jl} - \tilde{g}_{il}\widetilde{R}_{ik}}{n-2} - \frac{\widetilde{S}(\tilde{g}_{il}\tilde{g}_{jk} - \tilde{g}_{ik}\tilde{g}_{jl})}{(n-1)(n-2)}\Big) \\ = R_{ijkl} + \frac{g_{il}R_{jk} + g_{jk}R_{il} - g_{ik}R_{jl} - g_{il}R_{ik}}{n-2} - \frac{S(g_{il}g_{jk} - g_{ik}g_{jl})}{(n-1)(n-2)}. \end{aligned}$$

Comparison with (2.3.7)$'$ shows that this means precisely that

$$e^{-2\varphi}\widetilde{W}_{ijkl} = W_{ijkl} \tag{3.4.5}$$

for the corresponding Weyl tensors. Hence we have the following corollary to Theorem 3.4.3:

Corollary 3.4.4. *Under the hypotheses of Theorem 3.4.3 we have the transformation law* (3.4.5) *for the corresponding Weyl tensors.*

Exercise 3.4.1. Show that if $\tilde{g} = e^{2\varphi}g$ then the corresponding covariant derivatives are related by

$$\widetilde{\nabla}_X Y = \nabla_X Y + (X\varphi)Y + (Y\varphi)X - (X,Y)(d\varphi)^\sharp,$$

where the scalar product and $\sharp$ are defined by the metric g.

A metric is called *conformally flat* if it is conformal to a flat metric. From Corollary 3.4.4 it follows at once that a necessary condition for this is that the Weyl tensor vanishes. (When H. Weyl introduced this tensor he called it the *conformal curvature tensor.* This term is still used occasionally, and the term Weyl tensor may then have a different meaning. This is the case in Gerretsen [1], where somewhat different proofs of the following results can be found.) We shall now examine if it is a sufficient condition too.

If the Weyl tensor of the Riemannian manifold M with metric tensor g_{ij} is equal to 0, then the Weyl tensor of the manifold $\widetilde{M} = M$ with metric tensor $\tilde{g}_{ij} = e^{2\varphi}g_{ij}$ is 0 for arbitrary φ, so the curvature tensor of $\widetilde{M}$ vanishes if and only if the Ricci tensor is equal to 0, that is by (3.4.2)$'$

$$R_{ij} + (2-n)\varphi_{ij} - (\Delta\varphi + (n-2)|\nabla\varphi|^2)g_{ij} = 0, \quad i,j = 1,\dots,n.$$

To clarify the role of the Weyl tensor we shall discuss these equations for the vanishing of the Ricci tensor $\widetilde{R}_{ij}$ without assuming at first that $W = 0$. Let Φ be the one form $\nabla\varphi = d\varphi = \sum \partial_i \varphi dx^i$. Then the preceding equations can be written

$$R_{ij} + (2-n)(\Phi_{i,j} - \Phi_i\Phi_j) - (\sum_{k,l} g^{kl}\Phi_{k,l} + (n-2)|\Phi|^2)g_{ij} = 0. \tag{3.4.6}$$

Conversely, if we can find a one form satisfying these equations in an open set $\Omega \subset M$, we obtain using (3.1.2)$'$

$$0 = \Phi_{i,j} - \Phi_{j,i} = \partial_j\Phi_i - \partial_i\Phi_j$$

since $\Gamma_{ij}{}^l$ is symmetric in i, j. Thus $d\Phi = 0$ in Ω, so $\Phi = d\varphi$ for some φ if Ω is simply connected, and φ will then have the desired properties.

Contraction of (3.4.6) gives (see (3.4.3))

$$S + 2(1-n)\sum_{k,l} g^{kl}\Phi_{k,l} - (n-2)(n-1)|\Phi|^2 = 0,$$

or equivalently

$$2(n-1)(\sum_{k,l} g^{kl}\Phi_{k,l} + (n-2)|\Phi|^2) = S + (n-2)(n-1)|\Phi|^2.$$

We can therefore rewrite the equations (3.4.6) in the form

$$\Phi_{i,j} = \Phi_i\Phi_j + \omega_{ij} - \tfrac{1}{2}|\Phi|^2 g_{ij}, \tag{3.4.7}$$

$$\omega_{ij} = (R_{ij} - Sg_{ij}/(2n-2))/(n-2). \tag{3.4.8}$$

Note that (2.3.7)$'$ can be written

$$R_{ijkl} = W_{ijkl} + \omega_{ik}g_{jl} - \omega_{il}g_{jk} + \omega_{jl}g_{ik} - \omega_{jk}g_{il}. \tag{3.4.9}$$

If we introduce the definition $\Phi_{i,j} = \partial_j\Phi_i - \sum_l \Gamma_{ij}{}^l\Phi_l$ in (3.4.7), we obtain a system of the form (C.6) discussed in Theorem C.3 in the appendix, with Φ in the role of the y variables there. To find the Frobenius integrability condition (C.7) we should apply ∂_k and subtract the equation with j and k interchanged, and then use the equations (3.4.6) to express the derivatives of Φ. Clearly this gives the same result as if we take the covariant derivative of (3.4.6) and use that

$$\Phi_{i,jk} - \Phi_{i,kj} = \sum_\nu R^\nu{}_{ijk}\Phi_\nu. \tag{3.4.10}$$

Here the right-hand side can be expressed using (3.4.9), and we have

$$\Phi_{i,jk} = \Phi_{i,k}\Phi_j + \Phi_i\Phi_{j,k} + \omega_{ij,k} - \tfrac{1}{2}g_{ij}\partial_k|\Phi|^2.$$

(3.4.6) implies $\Phi_{j,k} = \Phi_{k,j}$, as already observed, and

$$\Phi_{i,k}\Phi_j - \Phi_{i,j}\Phi_k = \omega_{ik}\Phi_j - \omega_{ij}\Phi_k - \tfrac{1}{2}|\Phi|^2(g_{ik}\Phi_j - g_{ij}\Phi_k). \tag{3.4.11}$$

Furthermore,

$$\begin{aligned}\tfrac{1}{2}\nabla_k|\Phi|^2 = (\nabla_k\Phi, \Phi) &= \sum_{l,m} g^{lm}\Phi_{l,k}\Phi_m \\ &= \sum_{l,m} g^{lm}(\Phi_l\Phi_k + \omega_{lk} - \tfrac{1}{2}g_{lk}|\Phi|^2)\Phi_m = \sum_{l,m} g^{lm}\omega_{lk}\Phi_m + \tfrac{1}{2}\Phi_k|\Phi|^2.\end{aligned} \tag{3.4.12}$$

With Φ^ν denoting the coordinates of $\Phi^\sharp$, it follows from (3.4.9) that the right-hand side of (3.4.10) can be written

$$\sum_\nu W_{\nu ijk}\Phi^\nu + g_{ik}\sum_\nu \omega_{\nu j}\Phi^\nu - g_{ij}\sum_\nu \omega_{\nu k}\Phi^\nu + \omega_{ik}\Phi_j - \omega_{ij}\Phi_k.$$

The left-hand side is

$$\begin{aligned}\omega_{ik}\Phi_j - \omega_{ij}\Phi_k - \tfrac{1}{2}|\Phi|^2(g_{ik}\Phi_j - g_{ij}\Phi_k) + \omega_{ij,k} - \omega_{ik,j} \\ - g_{ij}(\sum_\nu \omega_{\nu k}\Phi^\nu + \tfrac{1}{2}|\Phi|^2\Phi_k) + g_{ik}(\sum_\nu \omega_{\nu j}\Phi^\nu + \tfrac{1}{2}|\Phi|^2\Phi_j).\end{aligned}$$

After cancellation the compatibility conditions (3.4.10) therefore simplify to

$$\sum_\nu W_{\nu ijk}\Phi^\nu = \omega_{ij,k} - \omega_{ik,j}, \quad i,j,k = 1,\dots,n. \tag{3.4.10$'$}$$

If $W = 0$, then these conditions no longer involve Φ at all, so they are necessary for the existence of any solution at all to (3.4.6), and they imply the local existence of a solution with given value at a point. On the other hand, if there is a solution with $\nabla\varphi$ given at a point, then it follows that both sides of (3.4.10)$'$ must vanish, since the right-hand side is independent of Φ, hence $W = 0$ and

$$\omega_{ij,k} - \omega_{ik,j} = 0, \quad i,j,k = 1,\dots,n. \tag{3.4.10$''$}$$

We shall now show that when $n > 3$ there is a second miracle: the conditions (3.4.10)$''$ are always fulfilled if the Weyl tensor is equal to 0. (This is not true when $n = 3$ which is not surprising since the Weyl tensor is always 0 then.) In fact,

$$(n-3)(\omega_{jk,i} - \omega_{ik,j}) = \sum_l W^l{}_{kij,l}. \tag{3.4.13}$$

(I owe the following calculations to Anders Melin.) First note that the definition of the Ricci tensor and the second Bianchi identity give

$$R_{jk,i} - R_{ik,j} = \sum_l R^l{}_{klj,i} - \sum_l R^l{}_{kli,j} = \sum_l (R^l{}_{klj,i} + R^l{}_{kil,j}) = \sum_l R^l{}_{kij,l}.$$

By (3.4.9) it follows that
(3.4.14)

$$\begin{aligned}R_{jk,i} - R_{ik,j} &= \sum_l W^l{}_{kij,l} + \sum_{lm} g^{lm}(\omega_{mi,l}g_{kj} - \omega_{mj,l}g_{ki} + \omega_{kj,l}g_{mi} - \omega_{ki,l}g_{mj}) \\ &= \sum_l W^l{}_{kij,l} + \omega_{kj,i} - \omega_{ki,j} + \sum_l \omega^l{}_{i,l}g_{kj} - \sum_l \omega^l{}_{j,l}g_{ki},\end{aligned}$$

for $\nabla g = 0$. Since

$$R_{ij} = (n-2)\omega_{ij} + Sg_{ij}/(2n-2), \tag{3.4.8$'$}$$

we have

$$R_{ij,l} = (n-2)\omega_{ij,l} + S_{,l}\, g_{ij}/(2n-2).$$

If we multiply by g^{il} and contract in i, l, recalling (3.1.25), we obtain

$$\tfrac{1}{2}S_{,j} = (n-2)\sum \omega^l{}_{j,l} + S_{,j}/(2n-2), \quad \text{hence } \sum \omega^l{}_{j,l} = S_{,j}/(2n-2).$$

(3.4.13) follows if we use this result and (3.4.8)$'$ in (3.4.14), for S drops out. Hence we have proved (3.4.13) and the following

Theorem 3.4.5. *A Riemannian manifold of dimension* > 3 *is conformally flat if and only if the Weyl tensor vanishes. A Riemannian manifold of dimension* 3 *is conformally flat if and only if the integrability conditions* (3.4.10)$'$ *are valid.*

We shall finally discuss how the Laplace-Beltrami operator defined by (3.4.4) is changed when one passes to a conformal metric. First we give an equivalent and often more useful definition:

Proposition 3.4.6. *For every* $u \in C^2$ *we have*

$$\sum_{i,j} g^{ij}u_{,ij} = \sum_{i,j} g^{-\frac{1}{2}}\partial_j(g^{\frac{1}{2}}g^{ij}\partial_i u), \tag{3.4.15}$$

where $g = \det(g_{ij})$.

Proof. If u, v are in C^2 with support in the same coordinate patch, then

$$\int \sum_{i,j} g^{ij}\partial_i u \partial_j v g^{\frac{1}{2}}\, dx$$

is invariantly defined, dx denoting the Lebesgue measure in the local coordinates, for $g^{\frac{1}{2}}\, dx$ is the invariant volume element $dvol(x)$. Since integration by parts shows that the integral is equal to

$$-\int v \sum g^{-\frac{1}{2}}\partial_j(g^{\frac{1}{2}}g^{ij}\partial_i u)\, dvol(x)$$

we conclude that the right-hand side of (3.4.15) is invariantly defined, and so is the left-hand side. Both are equal at the center of a geodesic coordinate system, which proves the proposition.

Exercise 3.4.2. Prove (3.4.15) by explicit computation using Exercise 2.1.1.

Denote by $\widetilde{\Delta}$ the Laplace-Beltrami operator defined using the conformal metric $\tilde{g} = e^{2\varphi}g$. When $n = 2$ it follows from the second expression for Δ in (3.4.15) that $\widetilde{\Delta} = e^{-2\varphi}\Delta$; hence the harmonic functions satisfying the homogeneous Laplace-Beltrami equation are the same for two conformal metrics. This is not true when $n > 2$, but we shall now prove a substitute result in that case. With a constant a to be chosen later we have

$$\widetilde{\Delta}(e^{a\varphi}u) = e^{-n\varphi}g^{-\frac{1}{2}}\sum \partial_j(e^{(n-2)\varphi}g^{\frac{1}{2}}g^{jk}\partial_k(e^{a\varphi}u))$$
$$= e^{(a-2)\varphi}(\Delta u + (n-2+2a)\sum g^{jk}\partial_j\varphi\partial_k u) + Fu,$$

where F does not depend on u. Now we choose $a = 1 - \frac{n}{2}$ so that the first order terms disappear. If the formula is applied with $u = e^{-a\varphi}$, it follows that

$$-F = e^{2(a-1)\varphi}\Delta(e^{-a\varphi}) = e^{(a-2)\varphi}(a^2|\nabla\varphi|^2 - a\Delta\varphi);$$

the last equality is justified using geodesic coordinates. By (3.4.3) we have

$$a|\nabla\varphi|^2 - \Delta\varphi = \tfrac{1}{2}((2-n)|\nabla\varphi|^2 - 2\Delta\varphi) = (e^{2\varphi}\widetilde{S} - S)/(2n-2),$$

so we have since $-a/(2n-2) = (n-2)/(4n-4)$

$$\widetilde{\Delta}(e^{a\varphi}u) = e^{(a-2)\varphi}(\Delta u + (n-2)/(4n-4)(e^{2\varphi}\widetilde{S} - S)u), \text{ or}$$

$$(\widetilde{\Delta} - \frac{\widetilde{S}(n-2)}{4n-4})(e^{\varphi(2-n)/2}u) = e^{-\varphi(n+2)/2}(\Delta - \frac{S(n-2)}{4n-4})u. \tag{3.4.16}$$

The operator $\Delta - S(n-2)/(4n-4)$ is called the *conformal Laplacian.* The transformation law (3.4.16) makes it easy to pass from solutions for one metric to solutions for another conformal one. Note that in a flat space, where $S = 0$, the conformal Laplacian is the standard Laplacian, so we shall be able to study its solutions by studying the conformal Laplacian in a conformally equivalent situation. In the Euclidean case one may for example use the stereographic projection to obtain a compact situation. The applications are perhaps even more striking in pseudo-Riemannian geometry where a related conformal map can be used to map Minkowski space into a bounded part of the Einstein universe, but that will not be discussed in the present notes.

If $\widetilde{S}$ is prescribed, then (3.4.3) is a non-linear elliptic equation for φ which can be solved locally without difficulty. Thus there always exists locally a conformal metric with zero scalar curvature. The corresponding global questions are very hard, however. The *Yamabe problem* to find for a given compact Riemannian manifold a conformal metric with constant scalar curvature was not completely solved until 1984, when R. Schoen cleared up the hardest exceptional cases. We hope to return to this in another chapter.

3.5. The curvature of a submanifold. Chapter II was restricted to the study of submanifolds M of $\mathbf{R}^N$, but in this chapter we have extended all results given there on the interior geometry of M to abstract Riemannian manifolds. We shall now extend the results related to the embedding in $\mathbf{R}^N$ also by discussing submanifolds of a general Riemannian manifold.

Let $\widetilde{M}$ be a Riemannian manifold and let M be a smooth submanifold. We shall denote the covariant differentiations in M and in $\widetilde{M}$ by ∇ and $\widetilde{\nabla}$ respectively.

Theorem 3.5.1. *If X and Y are vector fields in M then*

$$\widetilde{\nabla}_X Y - \nabla_X Y = h(X,Y), \tag{3.5.1}$$

where h is a symmetric bilinear form in T_xM with values in the normal plane $N_xM = T_x\widetilde{M} \ominus T_xM$ for every $x \in M$.

Proof. The theorem states in particular that $\tau\widetilde{\nabla}_X Y = \nabla_X Y$ if τ is the map $T\widetilde{M}|_M \to TM$ defined by orthogonal projection in each fiber. We begin by proving this weaker result. To do so we note that the identities (3.1.3), (3.1.4) for $\widetilde{\nabla}$ show that if X, Y, Z are vector fields in M, then we have there

$$\tau\widetilde{\nabla}_X Y - \tau\widetilde{\nabla}_Y X = \tau[X,Y] = [X,Y],$$
$$(\tau\widetilde{\nabla}_Z X, Y) + (X, \tau\widetilde{\nabla}_Z Y) = (\widetilde{\nabla}_Z X, Y) + (X, \widetilde{\nabla}_Z Y) = Z(X,Y),$$

hence $\tau\widetilde{\nabla}_X Y = \nabla_X Y$ by the uniqueness statement in Theorem 3.1.2.

From (3.1.3) applied to ∇ and $\widetilde{\nabla}$ we obtain

$$\widetilde{\nabla}_X Y - \nabla_X Y = \widetilde{\nabla}_Y X + [X,Y] - \nabla_Y X - [X,Y] = \widetilde{\nabla}_Y X - \nabla_Y X,$$

which proves that the left-hand side of (3.5.1) is symmetric in X and Y. If $\varphi \in C_0^\infty$ then

$$\widetilde{\nabla}_X(\varphi Y) - \nabla_X(\varphi Y) = \widetilde{\nabla}_{\varphi Y} X - \nabla_{\varphi Y} X = \varphi(\widetilde{\nabla}_Y X - \nabla_Y X) = \varphi(\widetilde{\nabla}_X Y - \nabla_X Y).$$

Hence the left-hand side of (3.5.1) at $x \in M$ is a symmetric bilinear form in $X(x)$ and $Y(x)$, with values in $N_x(M)$, which completes the proof.

Definition 3.5.2. The symmetric bilinear map h in T_xM with values in N_xM defined by (3.5.1) is called the second fundamental form of M with respect to $\widetilde{M}$.

In the particular case where $\widetilde{M} = \mathbf{R}^N$ Theorem 3.1.5 gives back the expression (2.1.11) for the covariant derivative. We shall now also give an extension of the Gauss equations (2.1.13), (2.1.15). Let X, Y, Z, W be vector fields in M. By (3.1.6) we have

$$\big((\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})Z, W\big) = R(W,Z,X,Y),$$

where R is the Riemann curvature tensor of M, and we have a similar formula with ∇ and R replaced by $\widetilde{\nabla}$ and $\widetilde{R}$. (We can extend X, Y, Z, W to vector fields on $\widetilde{M}$, but on M the result is independent of the extension.) By Theorem 3.5.1 we have

$$(\nabla_X \nabla_Y Z, W) = (\widetilde{\nabla}_X \nabla_Y Z, W) = (\widetilde{\nabla}_X \widetilde{\nabla}_Y Z, W) - (\widetilde{\nabla}_X h(Y, Z), W).$$

Since $(h(Y, Z), W) = 0$ on M, it follows from (3.1.4) and Theorem 3.5.1 that

$$-(\widetilde{\nabla}_X h(Y, Z), W) = (h(Y, Z), \widetilde{\nabla}_X W) = \big(h(Y, Z), h(X, W)\big).$$

Hence

$$(\nabla_X \nabla_Y Z, W) = (\widetilde{\nabla}_X \widetilde{\nabla}_Y Z, W) + \big(h(Y, Z), h(X, W)\big),$$

and since $(\nabla_{[X,Y]} - \widetilde{\nabla}_{[X,Y]})Z$ is a normal vector, we obtain

$$R(W, Z, X, Y) = \widetilde{R}(W, Z, X, Y) + \big(h(Y, Z), h(X, W)\big) - \big(h(X, Z), h(Y, W)\big),$$

which proves the following extension of the Gauss equations (2.1.13), (2.1.15):

Theorem 3.5.3. *If M is a smooth submanifold of a Riemannian manifold $\widetilde{M}$, with Riemann curvature tensors R and $\widetilde{R}$, respectively, and $t_1, t_2, t_3, t_4 \in T_x M$, then*

$$R(t_1, t_2, t_3, t_4) = \widetilde{R}(t_1, t_2, t_3, t_4) + \big(h(t_1, t_3), h(t_2, t_4)\big) - \big(h(t_1, t_4), h(t_2, t_3)\big), \tag{3.5.2}$$

where h is the second fundamental form of M with respect to $\widetilde{M}$.

CHAPTER IV

EXTERIOR DIFFERENTIAL CALCULUS IN RIEMANNIAN GEOMETRY

Summary. Chapters II and III have been based on the tensor calculus of Ricci. We shall now discuss an alternative approach due to E. Cartan using the calculus of exterior differential forms and moving orthogonal frames systematically. In Section 4.1 we reconsider the study in Chapter II of submanifolds of Euclidean space, and in Section 4.2 we discuss abstract Riemannian manifolds, as in Chapter III. Using the tools developed we then return to the Gauss-Bonnet theorem in Section 4.3 and prove its analogue in higher dimensions following S. S. Chern. Pontrjagin classes are then defined in Section 4.4.

4.1. Submanifolds of a Euclidean space. In the proof of the Gauss-Bonnet theorem (Theorem 3.1.6) we have seen a convincing example of the usefulness of working with arbitrary orthonormal frames in the tangent bundle of a Riemannian manifold. We shall now use such methods also to study submanifolds of a Euclidean vector space V, of dimension N. Let $F_0(V)$ be the set of all orthonormal frames $e_1,\dots,e_N\in V$. If we choose a fixed frame $(e_1^0,\dots,e_N^0)\in F_0(V)$, then any other orthonormal frame can be written uniquely in the form $Oe_1^0,\dots,Oe_N^0$ where $O\in\mathbf{O}(N)$, the orthogonal group. (See Section 1.4.) Another choice of $e_1^0,\dots,e_N^0$ gives an identification with $\mathbf{O}(N)$ which only differs by a right translation in $\mathbf{O}(N)$; in particular we see that $F_0(V)$ is a C^∞ (in fact real analytic) manifold.

In $F_0(V)$ we have N functions

$$e_j : F_0(V)\ni(e_1,\dots,e_N)\mapsto e_j\in V.$$

The differential de_j is a linear form on the tangent space of $F_0(V)$ with values in V, so it can be written

$$de_j=\sum_{k=1}^N\omega_{kj}e_k, \tag{4.1.1}$$

where $\omega_{jk}=(e_k,de_j)$ are scalar one forms on $F_0(V)$. Since

$$0=d(e_j,e_k)=(de_j,e_k)+(e_j,de_k),$$

we have

$$\omega_{jk}+\omega_{kj}=0,\quad j,k=1,\dots,N. \tag{4.1.2}$$

Since

$$0=d^2e_j=\sum_{k=1}^N(d\omega_{kj})e_k-\sum_{k=1}^N\omega_{kj}\wedge de_k=\sum_{k=1}^N(d\omega_{kj})e_k-\sum_{k,l=1}^N\omega_{lj}\wedge\omega_{kl}e_k,$$

we obtain

$$d\omega_{kj} + \sum_{l=1}^{N} \omega_{kl} \wedge \omega_{lj} = 0, \quad j, k = 1, \dots, N. \tag{4.1.3}$$

The equations (4.1.2), (4.1.3) are of course just another way of writing the results of Section 1.4 such as (1.4.1).

If $W \subset V$ is a subspace of dimension n, the set of orthonormal frames in $F_0(V)$ such that $e_1, \dots, e_n \in W$, hence $e_{n+1}, \dots, e_N \in W^\perp$, is a submanifold $F_0(V, W)$ which we can identify with the group $\mathbf{O}(n) \times \mathbf{O}(N-n)$. It is clear that the restriction of the form ω_{jk} above to $F_0(V, W)$ vanishes if $j \le n$ and $k > n$. When $j \le n$ and $k \le n$, it is equal to the analogous form on $F_0(W)$ pulled back to $F_0(V, W)$ by the obvious surjective map $F_0(V, W) \to F_0(W)$.

Now let M be a C^∞ submanifold of V, of dimension n, and set

$$F_x = F_0(V, T_x M), \quad x \in M, \tag{4.1.4}$$

which is the set of orthonormal frames at x such that $e_1, \dots, e_n$ span the tangent space $T_x M$ while $e_{n+1}, \dots, e_N$ span the normal space. We shall use the notation $\alpha, \beta, \dots$ for indices running from 1 to n, the notation $r, s, \dots$ for indices running from $n+1, \dots, N$, and $A, B, \dots$ for indices running from 1 to N. It is clear that $F(M) = \cup_{x \in M} \{x\} \times F_x$, as a subset of $M \times F_0(V)$ is a C^∞ fiber space over M. Composing the projection $p : F(M) \to M$ with the embedding $M \to V$ we get a map $F(M) \to V$, which we also denote by p. In addition we have N maps $e_A : F(M) \to V$. Since the range of the differential dp at a point in F_x is equal to $T_x M$, we can write

$$dp = \sum_{\alpha=1}^{n} \omega_\alpha e_\alpha, \tag{4.1.5}$$

and we have as in (4.1.1), (4.1.2)

$$de_A = \sum_{B=1}^{N} \omega_{BA} e_B, \quad A = 1, \dots, N, \tag{4.1.6}$$

$$\omega_{AB} + \omega_{BA} = 0, \quad A, B = 1, \dots, N. \tag{4.1.7}$$

Here ω_α and ω_{AB} are scalar differential forms on $F(M)$. From (4.1.5) we obtain

$$0 = d^2 p = \sum (d\omega_\alpha) e_\alpha - \sum \omega_\alpha \wedge de_\alpha = \sum (d\omega_\alpha) e_\alpha - \sum \omega_\alpha \wedge \omega_{B\alpha} e_B$$

or if we separate components with indices $\le n$ and $> n$:

$$d\omega_\alpha + \sum_{\beta=1}^{n} \omega_\beta \wedge \omega_{\beta\alpha} = 0, \quad \alpha = 1, \dots, n, \tag{4.1.8}$$

$$\sum_{\alpha=1}^{n} \omega_\alpha \wedge \omega_{\alpha r} = 0, \quad r = n+1, \dots, N. \tag{4.1.9}$$

Differentiation of (4.1.6) gives

$$0 = \sum (d\omega_{BA})e_B - \sum \omega_{CA} \wedge de_C = \sum (d\omega_{BA})e_B + \sum \omega_{BC} \wedge \omega_{CA} e_B,$$

hence

$$d\omega_{BA} + \sum_{C=1}^{N} \omega_{BC} \wedge \omega_{CA} = 0, \quad A, B = 1, \dots, N. \tag{4.1.10}$$

In particular,

(4.1.11)

$$d\omega_{\alpha\beta} + \sum_{\gamma=1}^{n} \omega_{\alpha\gamma} \wedge \omega_{\gamma\beta} = \Omega_{\alpha\beta}, \quad \alpha, \beta = 1, \dots, n, \quad \text{where}$$

$$\Omega_{\alpha\beta} = - \sum_{r=n+1}^{N} \omega_{\alpha r} \wedge \omega_{r\beta}, \quad \alpha, \beta = 1, \dots, n. \tag{4.1.12}$$

The restriction of ω_α to a fiber F_x is equal to 0 for $\alpha = 1, \dots, n$, since p is constant in F_x, and since p is surjective these forms are at every point in F_x a basis for covectors orthogonal to F_x. Now the restriction of $\omega_{\alpha r}$ to a fiber F_x must vanish, for when x is fixed we have the situation discussed for $F_0(V, W)$ above. Hence we can write

$$\omega_{\alpha r} = \sum_{\beta=1}^{n} \lambda_{r\alpha\beta} \omega_\beta$$

with uniquely determined coefficients $\lambda_{r\alpha\beta} \in C^\infty(F(M))$. Here $\lambda_{r\alpha\beta} = \lambda_{r\beta\alpha}$ according to (4.1.9). Hence

$$\Omega_{\alpha\beta} = \sum_{r=n+1}^{N} \omega_{\alpha r} \wedge \omega_{\beta r} = \sum_{r=n+1}^{N} \sum_{\varrho,\sigma=1}^{n} \lambda_{r\alpha\varrho} \lambda_{r\beta\sigma} \omega_\varrho \wedge \omega_\sigma$$

is also a form with vanishing restriction to the fibers. The forms ω_α, $\omega_{\alpha\beta}$ and $\Omega_{\alpha\beta}$, $\alpha, \beta = 1, \dots, n$, can be obtained by pulling back forms to $F(M)$ from the fiber space $P(M)$ over M with fiber at x consisting of the orthonormal frames in $T_x M$, for the definitions above are already applicable in $P(M)$ without reference to the normal vectors.

To clarify the meaning of the preceding equations we shall express them in terms of our earlier notation. Thus consider a subset of M with a local parametrization $x \mapsto f(x)$, where x varies in an open subset of $\mathbf{R}^n$. We can consider the coordinates $x^1, \dots, x^n$ as functions on M and lift them to $F(M)$. With the notation $f_\alpha = \partial f / \partial x^\alpha$, which is a vector field on M, we then obtain

$$dp = \sum_{\beta=1}^{n} f_\beta(x) dx^\beta = \sum_{\alpha=1}^{n} \omega_\alpha e_\alpha, \quad \text{thus } \omega_\alpha = \sum_{\beta=1}^{n} (f_\beta, e_\alpha) dx^\beta.$$

Here dx^β is a differential form on M pulled back to $F(M)$, and the coefficients are functions on $F(M)$. Let s be a local section of $F(M)$ and set $s^*e_\alpha = E_\alpha$. Then we have

$$s^*\omega_{\beta\alpha} = (dE_\alpha, E_\beta) = \sum_{i=1}^n (\partial E_\alpha/\partial x^i, E_\beta)dx^i = \sum_{i=1}^n (\nabla_{f_i} E_\alpha, E_\beta)dx^i.$$

By the product rule (3.1.4) for ∇ and (3.1.6) we obtain

$$\begin{aligned} s^*d\omega_{\beta\alpha} &= ds^*\omega_{\beta\alpha} = \sum \big((\nabla_{f_j}\nabla_{f_i}E_\alpha, E_\beta) + (\nabla_{f_i}E_\alpha, \nabla_{f_j}E_\beta)\big)dx^j \wedge dx^i \\ &= \tfrac12 \sum((\nabla_{f_j}\nabla_{f_i} - \nabla_{f_i}\nabla_{f_j})E_\alpha, E_\beta)dx^j \wedge dx^i + \sum(\nabla_{f_i}E_\alpha, E_\gamma)(\nabla_{f_j}E_\beta, E_\gamma)dx^j \wedge dx^i \\ &= \tfrac12 \sum R(E_\alpha, E_\beta, f_i, f_j)dx^j \wedge dx^i + s^* \sum \omega_{\gamma\alpha} \wedge \omega_{\beta\gamma}. \end{aligned}$$

If we compare this result with (4.1.11), it follows that

$$\Omega_{\alpha\beta} = \tfrac12 \sum_{i,j=1}^n R(e_\alpha, e_\beta, f_i, f_j)dx^i \wedge dx^j. \tag{4.1.13}$$

If we use that $dp = \sum f_i dx^i = \sum \omega_\gamma e_\gamma$, and evaluate (4.1.13) on a pair of tangent vectors, we conclude that

$$\Omega_{\alpha\beta} = \tfrac12 \sum_{\gamma,\delta=1}^n R(e_\alpha, e_\beta, e_\gamma, e_\delta)\omega_\gamma \wedge \omega_\delta. \tag{4.1.13$'$}$$

Thus the forms $\Omega_{\alpha\beta}$ contain exactly the same information as the Riemann curvature tensor. From (4.1.13)$'$ we also see again that $\Omega_{\alpha\beta}$ is a linear combination with coefficients in $C^\infty(P(M))$ of forms pulled back from M to $P(M)$. By the expression (2.1.13)$'$ for the curvature tensor it also follows that $\Omega_{\alpha\beta}$ is independent of the imbedding of M. In Section 4.2 we shall give an independent proof of that by reconsidering the preceding formulas for an abstract Riemannian manifold.

We shall now also calculate the forms $\omega_{\alpha r}$. With the same notation as above and (2.1.5), we have

$$\begin{aligned} s^*\omega_{r\alpha} = (dE_\alpha, E_r) = \sum(\partial E_\alpha/\partial x^j, E_r)dx^j &= \sum(h(f_j, E_\alpha), E_r)dx^j \\ &= \sum_\beta (h(E_\beta, E_\alpha), E_r)s^*\omega_\beta, \end{aligned}$$

since $dp = \sum f_j dx^j = \sum e_\beta \omega_\beta$. Hence we obtain

$$\omega_{r\alpha} = \sum_{\beta=1}^n (h(e_\beta, e_\alpha), e_r)\omega_\beta, \quad \alpha = 1, \ldots, n, \ r = n+1, \ldots, N. \tag{4.1.14}$$

Using (4.1.12) we therefore obtain

$$\Omega_{\alpha\beta} = \sum_{r=1}^{n} \omega_{\alpha r} \wedge \omega_{\beta r} = \sum_{\gamma,\delta} (h(e_\gamma, e_\alpha), h(e_\delta, e_\beta))\omega_\gamma \wedge \omega_\delta, \tag{4.1.15}$$

which in view of (4.1.13) shows that (4.1.12) is equivalent to the Gauss equations.

We shall finally express the degree of the Gauss mapping for a hypersurface in terms of differential forms (see (3.1.23)). Thus consider a compact oriented hypersurface $M \subset \mathbf{R}^{n+1}$. We can then restrict ourselves to positively oriented frames in $F(M)$ such that e_{n+1} has the positive normal direction. The product

$$\omega_{1,n+1} \wedge \cdots \wedge \omega_{n,n+1} \tag{4.1.16}$$

is the pullback to $F(M)$ of a differential form of degree n on M. It is clear that it is locally equal to such a form multiplied by a function on $F(M)$, so we just have to show that it does not change if we replace e_α by $\sum_{\beta=1}^{n} O_{\beta\alpha} e_\beta$, $\alpha = 1, \ldots, n$, where $O \in \mathbf{SO}(n)$. An easy calculation which we leave as an exercise shows that (4.1.16) is just multiplied by the determinant of O, which is equal to 1. To compute (4.1.16) we may therefore choose $e_1, \ldots, e_n$ as directions of principal curvature, which means that

$$\omega_{\alpha,n+1} = (de_{n+1}, e_\alpha) = -K_\alpha \omega_\alpha$$

where K_α are the principal curvatures. Hence (4.1.16) is equal to $(-1)^n K\omega$ where $K = \prod_1^n K_\alpha$ is the total curvature and $\omega = \omega_1 \wedge \cdots \wedge \omega_n$ is the volume form on M (pulled back to $F(M)$). If D is the degree of the reflected Gauss map $\check{\gamma}$, we obtain in view of (3.1.23)

$$\begin{aligned} D vol(S^n) = (-1)^n \int_M \omega_{1,n+1} \wedge \cdots \wedge \omega_{n,n+1} \\ = (-1)^n \int_M \sum \varepsilon_{i_1 \ldots i_n} \omega_{i_1,n+1} \wedge \cdots \wedge \omega_{i_n,n+1}/n! \end{aligned}$$

where $\varepsilon_{i_1 \ldots i_n}$ is $+1$ for even and -1 for odd permutations of $1, \ldots, n$ and 0 otherwise. (The integrands here and below should be read as the forms on M which pull back to them.) If n is *even* we can pair the factors using the Gauss equations (4.1.12), which gives

$$D vol(S^n) = (-1)^{\frac{n}{2}} \int_M \sum \varepsilon_{i_1 \ldots i_n} \Omega_{i_1 i_2} \wedge \Omega_{i_3 i_4} \wedge \cdots \wedge \Omega_{i_{n-1} i_n}/n!. \tag{4.1.17}$$

The forms in the right-hand side are well defined for any abstract Riemannian manifold, and we shall take (4.1.17) as our starting point when generalizing the Gauss-Bonnet theorem to higher dimensions.

4.2. Abstract Riemannian manifolds. In this section M will denote a general abstract C^∞ Riemannian manifold of dimension n. The fiber bundle $P(M)$ over M with fiber P_x at $x \in M$ consisting of orthonormal frames for T_xM is still well defined. We denote the projection $P(M) \to M$ by p. In a neighborhood U of any point in M we can choose a C^∞ orthonormal basis for TU and use it to identify $P(U)$ with $U \times \mathbf{O}(n)$. For $i = 1, \dots, n$ we have C^∞ maps $e_i : P(M) \to T(M)$ mapping an element in $P(M)$ to the ith element in the frame, in the tangent space at the base point. We denote by ω_j the one forms on $P(M)$ which vanish on the fibers such that

$$dp = \sum_1^n \omega_j e_j. \tag{4.2.1}$$

(Since all indices run from 1 to n now, we shall often omit the range.) We shall now extend (4.1.8).

Theorem 4.2.1. *There are uniquely determined one forms ω_{ij} on $P(M)$, $i, j = 1, \dots, n$, such that*

$$\omega_{ij} + \omega_{ji} = 0, \quad d\omega_j + \sum_{k=1}^n \omega_k \wedge \omega_{kj} = 0, \quad i, j = 1, \dots, n. \tag{4.2.2}$$

Proof. Uniqueness. Let $\tilde\omega_{ij}$ be forms with

$$\tilde\omega_{ij} + \tilde\omega_{ji} = 0, \quad \sum \omega_k \wedge \tilde\omega_{kj} = 0.$$

Since $\omega_1, \dots, \omega_n$ are linearly independent at every point, it follows if we extend to a basis for $T^*P(M)$ at a point that $\tilde\omega_{kj}$ must be a linear combination of $\omega_1, \dots, \omega_n$, that is,

$$\tilde\omega_{kj} = \sum \lambda_{kji}\omega_i, \quad \text{where } \lambda_{kji} - \lambda_{ijk} = 0.$$

Since also $\lambda_{kji} = -\lambda_{jki}$ by assumption, we obtain

$$\lambda_{ijk} = -\lambda_{jik} = -\lambda_{kij},$$

so a circular permutation of the indices changes the sign. After three circular permutations we conclude that $\lambda_{ijk} = 0$.

Existence. In view of the uniqueness it suffices to prove existence locally, so we may assume that M is an open subset of $\mathbf{R}^n$, with coordinates denoted by $(x^1, \dots, x^n)$, and we identify the tangent space of $\mathbf{R}^n$ with $\mathbf{R}^n$. Since

$$dp = (dx^1, \dots, dx^n) = \sum \omega_j e_j,$$

where p is a function with values in $\mathbf{R}^n$, we obtain

$$\sum (d\omega_j)e_j - \sum \omega_j \wedge de_j = 0, \quad \text{that is, } d\omega_j = \sum \omega_k \wedge (de_k, e_j).$$

(The scalar product is of course taken in the Riemannian metric.) The restriction of (de_k, e_j) to a fiber of $P(M)$ is antisymmetric in j and k, so we can find one forms B_{jk} which are also antisymmetric in j, k such that $(de_k, e_j) + B_{jk}$ vanishes on the fibers. Then we obtain

$$d\omega_i = \sum B_{ik} \wedge \omega_k + \sum A_{ijk}\omega_j \wedge \omega_k$$

where A_{ijk} is antisymmetric in j and k. Now set

$$\omega_{ki} = B_{ik} + \sum_j (A_{kij} + A_{ijk} + A_{jik})\omega_j.$$

Since

$$A_{kij} + A_{ijk} + A_{ikj} + A_{kji} = 0$$

it follows that ω_{ik} is antisymmetric in its indices, and

$$\sum \omega_i \wedge \omega_{ki} = \sum \omega_i \wedge (B_{ik} + \sum_j A_{kij}\omega_j) = d\omega_k,$$

because $A_{ijk} + A_{jik}$ is symmetric in i and j. This proves the existence.

Exercise 4.2.1. Show that for every section s of $P(M)$ we have

$$s^*\omega_{kj} = -(\nabla s^* e_k, s^* e_j) = (\nabla s^* e_j, s^* e_k),$$

and that this determines ω_{kj} uniquely.

Next we shall extend (4.1.11) and (4.1.13)$'$:

Theorem 4.2.2. *If ω_j and ω_{jk} are the forms on $P(M)$ defined in Theorem 4.2.1, then the two forms*

$$\Omega_{ik} = d\omega_{ik} + \sum_{j=1}^{n} \omega_{ij} \wedge \omega_{jk}, \quad i, k = 1, \dots, n, \tag{4.2.3}$$

are antisymmetric in the indices i and k, we have

$$\sum_{k=1}^{n} \Omega_{ik} \wedge \omega_k = 0, \quad i = 1, \dots, n, \tag{4.2.4}$$

and there are C^∞ functions R_{ijkl} on $P(M)$ such that

$$\Omega_{ik} = \tfrac{1}{2} \sum_{l,j=1}^{n} R_{ikjl}\omega_j \wedge \omega_l, \quad i, k = 1, \dots, n. \tag{4.2.5}$$

R_{ijkl} is antisymmetric in i, j and in k, l.

Proof. Exterior differentiation of the second part of (4.2.2) gives

$$0 = \sum d\omega_{jk} \wedge \omega_k - \sum \omega_{jk} \wedge d\omega_k = \sum (d\omega_{jk} + \sum_i \omega_{ji} \wedge \omega_{ik}) \wedge \omega_k = \sum \Omega_{jk} \wedge \omega_k.$$

This proves (4.2.4). From the discussion of vector spaces at the beginning of Section 4.1 we know that the restriction of Ω_{ik} to any fiber must vanish. (See (4.1.3).) We can therefore write

$$\Omega_{ik} = \sum \theta_{ikj} \wedge \omega_j + \tfrac{1}{2} \sum R_{ikjl} \omega_j \wedge \omega_l.$$

Here R_{ikjl} are C^∞ functions on $P(M)$ and θ_{ikj} are linear combinations with such coefficients of forms extending $\omega_1, \ldots, \omega_n$ to a basis for one forms at a point in $P(M)$. The forms θ_{ikj} are skew symmetric in i and k since Ω_{ik} are and θ_{ikj} is unique. By (4.2.4) we must have

$$\sum_{k,j} \theta_{ikj} \wedge \omega_k \wedge \omega_j = 0$$

and we can conclude that θ_{ikj} are skew symmetric in k, j. Recalling the first part of the proof of Theorem 4.2.1 we can therefore conclude that $\theta_{ikj} = 0$. Hence we obtain (4.2.5), where the coefficients are uniquely determined if we require skew symmetry in j, l. The proof is complete.

If we enter the expression (4.2.5) in (4.2.4) we find that

$$R_{i[kjl]} = 0,$$

where the notation means summation over the circular permutations of k, j, l. In view of (4.1.13)$'$ we can identify R_{ijkl} with the Riemann curvature tensor, expressed in the basis $e_1, \ldots, e_n$ corresponding to a point in $P(M)$, so this is again the *first Bianchi identity* (2.3.4). To derive the *second Bianchi identity* we take the exterior differential of (4.2.3) and eliminate $d\omega_{ij}$ and $d\omega_{jk}$ using the same equation. Two sums $\sum \omega_{il} \wedge \omega_{lj} \wedge \omega_{lk}$ cancel and we obtain

$$d\Omega_{ik} - \sum_{j=1}^n \Omega_{ij} \wedge \omega_{jk} + \sum_{j=1}^n \omega_{ij} \wedge \Omega_{jk} = 0. \tag{4.2.6}$$

To show that this is equivalent to the second Bianchi identity (3.1.24) we choose a system of geodesic coordinates $x^1, \ldots, x^n$ with center at a given point. By Lemma 3.3.4 there are unique orthonormal vector fields $E_1, \ldots, E_n$ near 0 such that $E_j = \partial/\partial x^j$ at the origin and the covariant derivative along the radial vector field vanishes. Thus we obtain a section s of $P(M)$ which we can use to pull back (4.2.6) from $P(M)$ to M. By Lemma 3.3.4 the first order derivatives of E_j vanish at 0, and since the forms $s^*\omega_i$ are biorthogonal, their first derivatives also vanish at 0. Hence it follows from (4.2.2) that $\sum s^*\omega_{ij} \wedge s^*\omega_j = 0$ at 0. Since $s^*\omega_{ij}$ is skew symmetric in i, j, it follows (see the proof of uniqueness in Theorem 4.2.1) that $s^*\omega_{ij} = 0$ at 0, so (4.2.6) gives $ds^*\Omega_{ik} = 0$. By (4.2.5) we have

$$s^*\Omega_{ik} = \tfrac{1}{2} \sum R(E_i, E_k, E_j, E_l) s^*\omega_j \wedge s^*\omega_l,$$

where R is the Riemann curvature tensor, so we obtain

$$\sum \partial R_{ikjl}/\partial x^r dx^r \wedge dx^j \wedge dx^l = 0$$

at the origin, if we recall that the derivatives of $s^*\omega_j$ and E_j vanish there. At the center of a geodesic coordinate system the Christoffel symbols vanish so covariant derivatives are equal to the corresponding partial derivatives, and we obtain

$$R_{ik[jl,r]} = 0,$$

which is the second Bianchi identity (3.1.24).

Summing up, the equations (4.2.1), (4.2.2), (4.2.3) define the one forms ω_i, ω_{ij} and the two forms Ω_{ij} on $P(M)$ uniquely. The equations (4.2.4) and (4.2.6) expressing the first and the second Bianchi identity are valid, and the curvature forms Ω_{ik} have the form (4.2.5) where R denotes the components of the Riemann curvature tensor in the moving frame in $P(M)$.

4.3. The Gauss-Bonnet theorem in higher dimensions. In Section 3.1 we proved the Gauss-Bonnet theorem by showing that the curvature form in the right-hand side of (3.1.21) lifted to the sphere (circle) bundle is exact. In this section we shall prove a similar result for the form suggested by (4.1.17). Thus let M be a compact oriented Riemannian manifold of *even* dimension n, let $P_+(M)$ be the positively oriented part of the frame bundle, and let $S(M)$ be the unit sphere bundle $\{(x,e); x \in M, e \in T_xM, (e,e) = 1\}$. We can regard $P_+(M)$ as a fiber space over $S(M)$ with the projection $(e_1, \dots, e_n) \mapsto e_n$. Indices running from 1 to $n-1$ will be denoted by α, β. The restrictions of the differential forms $\omega_{\alpha n}$ to the fibers of $P_+(M) \to S(M)$ are equal to 0 by the discussion of $F_0(V,W)$ in Section 4.1, with $V = T_xM$ and $W = \mathbf{R}e_n$. We can therefore regard them as linear combinations with coefficients depending on $e_1, \dots, e_{n-1}$ of forms lifted from $S(M)$. The same is true for Ω_{jk} in view of (4.2.5). If a linear combination of products of these forms is invariant under $\mathbf{SO}(n-1)$ operating on $e_1, \dots, e_{n-1}$, it follows that it is the pullback of a form on $S(M)$.

Lemma 4.3.1. *If the pullback F of a form f on $S(M)$ to $P_+(M)$ can be written $F = \sum_{\alpha,\beta=1}^{n-1} \omega_{\alpha\beta} \wedge F_{\alpha\beta}$, then it is equal to* 0.

Proof. We may assume that $\alpha < \beta$ in the sum. The restriction of these forms $\omega_{\alpha\beta}$ to a fiber of $P(M) \to S(M)$ are linearly independent (cf. Section 1.4). If π is the projection $P(M) \to S(M)$ we can therefore for every tangent vector t of $S(M)$ find a unique tangent vector τ of $P(M)$ such that $\pi'\tau = t$ and $\omega_{\alpha\beta}(\tau) = 0$, $\alpha < \beta$. Since the multilinear form F vanishes on the linear space of such tangent vectors at any $p \in P(M)$, it follows that f vanishes on the tangent space of $S(M)$ at $\pi(q)$, which proves the lemma.

The form we wish to integrate is the one in (4.1.17):

$$\Delta_0 = \sum \varepsilon_{i_1 \dots i_n} \Omega_{i_1 i_2} \wedge \Omega_{i_3 i_4} \wedge \dots \wedge \Omega_{i_{n-1} i_n}. \tag{4.3.1}$$

We shall first verify that it is the pullback of a form on M. Since (4.2.5) shows that Δ_0 is a linear combination with coefficients in $C^\infty(P(M))$ of pullbacks of forms on M, it suffices to show that Δ_0 is invariant under the map of $P(M)$ into itself replacing a frame $e_1, \dots, e_n$ by $e'_k = \sum_i O_{ik} e_i$ where $O \in \mathbf{SO}(n)$ is constant. Since $\sum \omega_k e_k = \sum \omega'_k e'_k$ we have $\omega_i = \sum O_{ik}\omega'_k$ then, hence

$$d\omega_i = \sum_k O_{ik} d\omega'_k = -\sum_{k,l} O_{ik}\omega'_{kl} \wedge \omega'_l = -\sum_{j,k,l} O_{ik}\omega'_{kl} \wedge \omega_j O_{jl},$$

$$\text{so} \quad \omega_{ij} = \sum_{k,l} O_{ik} O_{jl} \omega'_{kl}, \qquad \Omega_{ij} = \sum_{k,l} O_{ik} O_{jl} \Omega'_{kl},$$

and we obtain

$$\Delta_0 = \sum \varepsilon_{i_1 \dots i_n} O_{i_1 j_1} \dots O_{i_n j_n} \Omega'_{j_1 j_2} \wedge \dots \wedge \Omega'_{j_{n-1} j_n}.$$

The sum over $i_1, \dots, i_n$ of the coefficients is the determinant of n columns in the orthogonal matrix O; and since $O \in \mathbf{SO}(n)$ it follows that it is equal to $\varepsilon_{j_1 \dots j_n}$. This proves the invariance of Δ_0, so it is the pullback to $P(M)$ of a form of degree n on M, which we also denote by Δ_0. Hence it is a closed form.

In the same way we verify for $k \le \frac{n}{2} - 1$ that the $n-1$ form

$$\Phi_k = \sum \varepsilon_{\alpha_1 \dots \alpha_{n-1}} \Omega_{\alpha_1 \alpha_2} \wedge \dots \wedge \Omega_{\alpha_{2k-1}\alpha_{2k}} \wedge \omega_{\alpha_{2k+1} n} \wedge \dots \wedge \omega_{\alpha_{n-1} n} \tag{4.3.2}$$

is the pullback to $P(M)$ of a form on $S(M)$. It suffices to note that a rotation of the first $n-1$ elements in a frame to $e'_\alpha = \sum O_{\beta\alpha} e_\beta$ keeping $e_n = e'_n$ transforms $\Omega_{\alpha\beta}$ as in the discussion of Δ_0 above while $\omega_{\alpha n} = \sum O_{\alpha\beta}\omega'_{\beta n}$. We leave the details of the repetition of the argument as an exercise and form the differential of Φ_k. We claim that

$$\begin{aligned} d\Phi_k &= k \sum \varepsilon_{\alpha_1 \dots \alpha_{n-1}} \Omega_{\alpha_1 \alpha_2} \wedge \dots \wedge d\Omega_{\alpha_{2k-1}\alpha_{2k}} \wedge \omega_{\alpha_{2k+1} n} \wedge \dots \wedge \omega_{\alpha_{n-1} n} \\ &+ (n-2k-1) \sum \varepsilon_{\alpha_1 \dots \alpha_{n-1}} \Omega_{\alpha_1 \alpha_2} \wedge \dots \wedge \Omega_{\alpha_{2k-1}\alpha_{2k}} \wedge d\omega_{\alpha_{2k+1} n} \wedge \omega_{\alpha_{2k+2} n} \wedge \dots \wedge \omega_{\alpha_{n-1} n}. \end{aligned} \tag{4.3.3}$$

Here we have used that there will be no change of sign in terms where d acts on a factor Ω. Differential forms of even degree commute, so the differential of such a factor can be moved to the right of the others without any sign change, and this corresponds to an even number of inversions of the indices α_j. Thus we obtain k identical contributions when differentiating on the Ω factors. When the factor $\omega_{\alpha_{2k+j} n}$ is differentiated we obtain a sign factor $(-1)^{j-1}$ which is compensated by $j-1$ inversions if it is moved to the left of the other factors ω, so we get $n-2k-1$ identical contributions of this kind, which proves (4.3.3).

We shall now calculate the right-hand side of (4.3.3) using that

$$d\Omega_{\alpha\beta} = \sum_j \Omega_{\alpha j} \wedge \omega_{j\beta} - \sum_j \omega_{\alpha j} \wedge \Omega_{j\beta},$$

$$d\omega_{\alpha n} = \Omega_{\alpha n} - \sum_j \omega_{\alpha j} \wedge \omega_{jn}.$$

In view of Lemma 4.3.1 we may omit terms containing a factor $\omega_{\alpha\beta}$ when simplifying (4.3.3), so we may replace $d\Omega_{\alpha\beta}$ by $\Omega_{\alpha n}\wedge\omega_{n\beta}-\omega_{\alpha n}\wedge\Omega_{n\beta}=-\Omega_{\alpha n}\wedge\omega_{\beta n}+\Omega_{\beta n}\wedge\omega_{\alpha n}$ and $d\omega_{\alpha n}$ by $\Omega_{\alpha n}$. To handle the first sum in (4.3.3) we now introduce for $0\le k<\frac{n}{2}$ the n forms

(4.3.4)
$$\Psi_k=2(k+1)\sum\varepsilon_{\alpha_1\dots\alpha_{n-1}}\Omega_{\alpha_1\alpha_2}\wedge\dots\wedge\Omega_{\alpha_{2k-1}\alpha_{2k}}\wedge\Omega_{\alpha_{2k+1}n}\wedge\omega_{\alpha_{2k+2}n}\wedge\dots\wedge\omega_{\alpha_{n-1}n}.$$

As for the forms Φ_k one verifies at once that they are forms on $S(M)$ pulled back to $P(M)$. In view of Lemma 4.3.1 we obtain

$$d\Phi_k=-\Psi_{k-1}-((n-2k-1)/(2k+2))\Psi_k, \tag{4.3.5}$$

for the difference between the two sides is the pullback to $P(M)$ of a form on $S(M)$, and every term contains a factor $\omega_{\alpha\beta}$. For $k=0$ we must interpret Ψ_{-1} as 0, for then there are no factors Ω to differentiate.

When $k=p-1$ where $p=\frac{n}{2}$ (recall that n is even) we obtain

$$\Psi_{p-1}=n\sum\varepsilon_{\alpha_1\dots\alpha_{n-1}}\Omega_{\alpha_1\alpha_2}\wedge\dots\wedge\Omega_{\alpha_{n-1}n}=\Delta_0,$$

so we can write Δ_0 as a differential by successive elimination of Ψ_k, $k=p-1,p-2,\dots,0$ using (4.3.5). This gives

$$\Delta_0=d\Phi,\quad \Phi=\sum_{k=0}^{p-1}(-1)^{p-k}\Phi_k\prod_{j=k}^{p-1}\big((2j+2)/(n-2j-1)\big). \tag{4.3.6}$$

Using (4.3.6) instead of (3.1.21) we can now proceed as in the proof of Theorem 3.1.7. Choose $f\in C^\infty(M)$ with only non-degenerate critical points, and let $F=(df)^\sharp=\operatorname{grad}f$ be the corresponding vector field. We recall that

$$\Gamma=\{(x,F(x)/\|F(x)\|);F(x)\ne0\}\subset S(M),$$

is a manifold with boundary $\partial\Gamma$ consisting of the fibers S_x over the finitely many critical points of f, with positive or negative orientation when the sign of the Hessian of f at x is negative or positive, respectively. The integral of Δ_0 over M is equal to the integral over Γ of the pullback to $S(M)$, which by Stokes' formula and (4.3.6) is equal to the sum of the integrals of Φ over $S_x(M)$ when $f'(x)=0$. Since the restriction of $\Omega_{\alpha\beta}$ to $S_x(M)$ is equal to 0, we only get contributions from Φ_0,

$$(-1)^p\int_{S_x(M)}\Phi=2^pp!/(2p-1)!!\int_{S_x(M)}\Phi_0=-(2^pp!/(2p-1)!!)(n-1)!vol(S^{n-1})$$

by the calculation which led to (4.1.17). As in (4.1.17) we shall divide by $n!vol(S^n)$. Note that if B^n is the unit ball in $\mathbf{R}^n$ then

$$vol(S^n)/vol(S^{n-1})=((n+1)/n)vol(B^{n+1})/vol(B^n)=(n+1)/n\int_{-1}^1(1-t^2)^{\frac{n}{2}}\,dt$$
$$=(n+1)/n\int_0^1(1-s)^{\frac{n}{2}}s^{-1/2}\,ds=(2p+1)\Gamma(p+1)\Gamma(\tfrac12)/(2p\Gamma(p+\tfrac32))$$
$$=(2p+1)p!/(2p\tfrac12\dots(p+\tfrac12))=2^p(p-1)!/(2p-1)!!.$$

This gives

$$(-1)^p/(n!vol(S^n)) \int_{S_x(M)} \Phi = -\tfrac{1}{2}.$$

Here $S_x(M)$ has been oriented so that $e_1, \ldots, e_{n-1}$ is a positively oriented frame in the tangent space at $e_n \in S_x(M)$ if $e_1, \ldots, e_{n-1}, e_n$, is positively, that is, $e_n, e_1, \ldots, e_{n-1}$ is negatively oriented, so the sphere has been oriented as the boundary of its exterior. Hence we obtain

$$(-1)^p/(n!vol(S^n)) \int_M \Delta_0 = \tfrac{1}{2} \sum_{f'(x)=0} \varepsilon_x, \tag{4.3.7}$$

where ε_x is the sign of the Hessian of f at x. (We can check the sign by noting that for $M = S^n$, embedded in $\mathbf{R}^{n+1}$ as the unit sphere, the right-hand side is equal to 1 if f is one of the coordinates in $\mathbf{R}^{n+1}$, so the sign agrees with (4.1.17).) Using a method of Chern we have now proved the following extension due to Allendoerfer, Fenchel and Weil of the Gauss-Bonnet theorem:

Theorem 4.3.2. *If M is a compact oriented Riemannian manifold of even dimension $n = 2p$, then* (4.3.7) *is valid with Δ_0 denoting the form on M with pullback to $P(M)$ defined by* (4.3.1) *and any $f \in C^\infty(M)$ with only non-degenerate critical points, ε_x denoting the sign of the Hessian at a critical point x.*

As observed after Theorem 3.1.7 we could also replace $(df)^\sharp$ by any C^∞ vector field F with only non-degenerate zeros, with the sum taken over the zeros of F and ε_x defined as the index of the zero. The right-hand side of (4.3.7) is the Euler characteristic of M, which we cannot define until Chapter VI though. Without having this concept available we can still remark that the right-hand side is completely independent of the Riemannian metric, so the integral in (4.3.7) can only depend on the differentiable structure of M.

4.4. Pontrjagin forms and classes. The proof that the form Δ_0 defined by (4.3.1) is the pullback of a form on M suggests the definition of other forms on $P(M)$ with the invariance properties which guarantees that they are such pullbacks. The most important ones come from the determinant

$$\det(\delta_{ij} + \lambda\Omega_{ij})_{i,j=1}^n, \tag{4.4.1}$$

where λ is an indeterminate. The determinant is well defined since differential forms of even order commute. If as in the discussion of Δ_0 we make a rotation of $P(M)$ with $O \in \mathbf{O}(n)$, then Ω_{ij} is replaced by $\sum_{k,l} O_{ki}O_{lj}\Omega_{kl}$, which means that (4.4.1) is multiplied by the square of $\det O$, which is equal to 1. This proves that the coefficient of any power of λ in (4.4.1) is the pullback to $P(M)$ of a form in M. (Note that we have not assumed here that M is oriented.) We can write the coefficient of λ^k in the form

$$\sum_{I,J} \operatorname{sgn} \binom{I}{J} \Omega_{i_1j_1} \wedge \cdots \wedge \Omega_{i_kj_k}/k! \tag{4.4.2}$$

where $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$ are sequences of k different indices between 1 and n, and $\operatorname{sgn}\binom{I}{J}$ is defined as 0 if they are not permutations of each other and as the sign of the permutation otherwise. The sum (4.4.2) is equal to 0 if k is odd, for if we exchange I and J then $\operatorname{sgn}\binom{I}{J}$ does not change but the following differential forms are multiplied by -1. It follows from the Bianchi identities (4.2.6) that the forms (4.4.2) are closed. In fact, since the restrictions of the forms ω_{jk}, $j < k$, to the fibers of $P(M)$ are linearly independent, an obvious analogue of Lemma 4.3.1 shows that we may calculate the differential as if $d\Omega_{jk}$ were equal to 0. With a normalization which will be explained later on, we introduce:

Definition 4.4.1. The Pontrjagin form P_k of a Riemannian manifold M is the form on M of degree $4k$ which pulled back to the frame bundle $P(M)$ is the term of degree $4k$ in $\det(\delta_{ij} + \Omega_{ij}/2\pi)$ where Ω_{ij} are the curvature forms (4.2.3). Thus P_k is given by (4.4.2) with k replaced by $2k$, divided by $(2\pi)^{2k}$. The de Rham cohomology class of the closed form P_k is called the Pontrjagin class of M.

The definition of the Pontrjagin form depends heavily on the Riemannian metric, but we shall now prove that the Pontrjagin class is independent of it. On any C^∞ manifold one can define a Riemannian metric by $\sum \varphi_j g_j$ where $\{\varphi_j\}$ is a partition of unity subordinate to a covering with coordinate patches and g_j is the Euclidean metric in the local coordinates. Thus the Pontrjagin classes are well defined on arbitrary compact C^∞ manifolds.

Theorem 4.4.2. *If M_1 and M_2 are Riemannian manifolds with metrics g_1 and g_2, respectively, and if $M_1 \times M_2$ is given the metric $g_1 + g_2$, then the sum $\sum P_k^M$ of the Pontrjagin forms of M is equal to $p_1^* \sum P_k^{M_1} \wedge p_2^* \sum P_k^{M_2}$, where p_j is the projection $M_1 \times M_2 \to M_j$. The Pontrjagin class of a differentiable manifold is independent of the metric.*

Proof. On $P(M_1) \times P(M_2) \subset P(M_1 \times M_2)$ it is clear that the matrices ω_{jk} and Ω_{jk} are block matrices $\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$ where the diagonal entries are the corresponding matrices for M_1 and M_2, lifted to $P(M_1 \times M_2)$. This proves the first statement. Now denote by M a C^∞ manifold with two different metrics g_0 and g_1. Choose an increasing function $\chi \in C^\infty(\mathbf{R})$ with $\chi = 0$ in $(-\infty, 1/3)$ and $\chi = 1$ in $(2/3, \infty)$. On $M \times \mathbf{R}$ we introduce the metric

$$g = \chi(t) g_1 + (1 - \chi(t)) g_0 + (dt)^2,$$

where t is the coordinate in $\mathbf{R}$. From the first part of the proof it follows that on $M \times (-\infty, 1/3)$ (resp. $M \times (2/3, \infty)$) the Pontrjagin forms P_k on $M \times \mathbf{R}$ are equal to the Pontrjagin forms P_k^0 (resp. P_k^1) of (M, g_0) (resp. (M, g_1)) pulled back to $M \times \mathbf{R}$. If

$$i_t(x) = (x, t) \in M \times \mathbf{R}, \qquad x \in M, \ t \in \mathbf{R},$$

then $P_k^t = i_t^* P_k$, $t = 0, 1$. Since i_0 and i_1 are homotopic and P_k is closed, it follows that P_k^0 and P_k^1 are in the same cohomology class. The proof is complete.

CHAPTER V

CONNECTIONS, CURVATURES AND CHERN CLASSES

Summary. This chapter is not restricted to Riemannian geometry but rather a discussion of analogues for a general vector bundle of the study in Chapters III and IV of the tangent bundle of a Riemannian manifold. In Section 5.1 we introduce connections as differential operators on sections of a vector bundle and define the curvature as in Section 3.1. We also give an analogue of the arguments in Chapter IV using exterior differential forms, for this is essential in Section 5.2 when we define Chern classes, generalizing the Pontrjagin classes in Section 4.4. A brief introduction to Lie groups is given in Section 5.3 in preparation for a study of principal bundles and associated vector bundles in Section 5.4.

5.1. Connections in vector bundles. The essential point in the expression for the Riemann curvature tensor in (3.1.6) was the covariant differentiation ∇. Recall that for a vector field u on M, that is, a section of the tangent bundle TM, ∇u is a section of $T^*M \otimes TM$ such that for $\varphi \in C^\infty(M)$

$$\nabla(\varphi u) = \varphi \nabla u + (d\varphi) \otimes u, \tag{5.1.1}$$

and $u \mapsto \nabla u$ is a linear operator. Here TM can be replaced by any other vector bundle:

Definition 5.1.1. Let M be a C^∞ manifold and E a C^∞ vector bundle over M. Then a linear map $\nabla : C^\infty(M,E) \to C^\infty(M, T^*M \otimes E)$ is called a *connection* if (5.1.1) is valid for $u \in C^\infty(M,E)$ and $\varphi \in C^\infty(M)$.

In this definition E may be a real or a complex vector bundle provided that φ is in the same category. From (5.1.1) it follows at once that $\nabla u = 0$ in a neighborhood of x if $u = 0$ in a neighborhood of x, for we can then choose $\varphi \in C^\infty(M)$ equal to 0 in a neighborhood of x so that $u = \varphi u$. Thus ∇ is a local operator. If $e_1, \dots, e_N \in C^\infty(M,E)$ form a basis in E_x for all x in an open set $\omega \subset M$, then

$$\nabla(\sum \varphi_j e_j) = \sum \varphi_j \nabla e_j + \sum d\varphi_j \otimes e_j.$$

which proves that ∇ is a first order differential operator.

To put Definition 5.1.1 in the right context we digress to discuss linear differential operators between sections of vector bundles. If M is a C^∞ manifold and E, F two vector bundles over M, then a linear map $P : C^\infty(M,E) \to C^\infty(M,F)$ from sections of E to sections of F is a differential operator if $\operatorname{supp} Pu \subset \operatorname{supp} u$ for all u. By a theorem of Peetre this means that if $x^1, \dots, x^n$ are local coordinates in an open subset ω of M where E is identified with $\omega \times \mathbf{R}^N$ and F is identified with $\omega \times \mathbf{R}^{N'}$, then P is given by a matrix of partial differential operators in the usual sense:

$$(Pu)_j(x) = \sum_1^N P_{jk} u_k(x), \quad j = 1, \dots, N'. \tag{5.1.2}$$

(We could also have taken (5.1.2) as our definition for it is obvious that this property is invariant under change of local coordinates and bases for E and F.) If all P_{jk} are of order m, then P is said to be of order m, and the principal part is defined as the matrix of principal parts of P_{jk}, obtained when $(\partial_1, \dots, \partial_n)$ is replaced by $(\xi_1, \dots, \xi_n)$ in the terms of order m. An invariant formulation is that if $0 \neq \xi \in T_x^*$ and we choose $\varphi \in C^\infty$ with $d\varphi = \xi$ at x, then

$$p(\xi)u(x) = \lim_{t\to\infty} t^{-m} e^{-t\varphi} P(e^{t\varphi}u)(x) \in F_x. \tag{5.1.3}$$

In fact, using the local representation (5.1.2) we see immediately that this limit exists and is linear in $u(x)$ and a homogeneous polynomial of degree m in ξ. The invariance of (5.1.3) shows that $p(\xi) \in L(E_x, F_x)$ is invariantly defined. In the particular case of a first order operator, p is also linear in ξ, so p can be considered as a vector bundle map from $T_x^*M \otimes E$ to F, and then we have

$$P(\varphi u) = \varphi P u + p(d\varphi \otimes u). \tag{5.1.4}$$

A connection in E is therefore a first order differential operator from sections of E to sections of $T^*M \otimes E$ whose *principal symbol is the identity map in* $T^*M \otimes E$.

Proposition 5.1.2. *There exist connections in any vector bundle E. If ∇ is a connection in E, then every connection in E can be written in the form $\nabla + A$ where A is a section of* $\operatorname{Hom}(E, T^*M \otimes E) \cong T^*M \otimes E \otimes E^*$.

Proof. If E is a trivial vector bundle $M \times \mathbf{C}^N$, say, then we can define

$$\nabla(u_1, \dots, u_N) = (du_1, \dots, du_N).$$

If we take a covering of M with open subsets U_ν where E is a trivial bundle, the trivialization gives a connection ∇_ν for E restricted to U_ν. If χ_ν is a subordinate partition of unity, it is clear that $\nabla = \sum \chi_\nu \nabla_\nu$ is a connection. If A is the difference between two connections, it follows from (5.1.1) that

$$A(\varphi u) = \varphi A u,$$

so A is a bundle map, as claimed.

If X is a vector field and $u \in C^\infty(M, E)$ we shall just as in the Riemannian case write $\nabla_X u = \langle X, \nabla u\rangle \in C^\infty(M, E)$. Motivated by (3.1.6) we take another vector field Y and form

$$R^\nabla(X, Y)u = (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})u. \tag{5.1.5}$$

$R^\nabla(X, Y)u(x)$ depends only on the values of X, Y and u at x, for the proof of this fact for covariant differentiation only used that it is a connection. Thus (5.1.5) is at x a skew symmetric bilinear form in the tangent vectors $X(x)$, $Y(x)$ and a linear form in $u(x) \in E_x$, with values in E_x, so R^∇ is a two form with values in $E_x \otimes E_x^*$.

Definition 5.1.3. The curvature R^∇ of the connection ∇ in E is the two form with values in $E \otimes E^*$ defined by (5.1.5).

In the spirit of Chapter IV we can also argue more elegantly by extending the connection to forms, that is, sections of $\wedge^p T^*M \otimes E$. Note that if u is such a section and ψ is a scalar q form, then $\psi \wedge u \in \wedge^{q+p} T^*M \otimes E$ can be uniquely defined so that the multiplication is bilinear and $\psi \wedge (\varphi \otimes v) = (\psi \wedge \varphi) \otimes v$, if v is a section of E and φ is a scalar p form.

Proposition 5.1.4. *Given a connection ∇ in the vector bundle E, there is a unique differential operator $d^\nabla : C^\infty(M, \wedge^p T^*M \otimes E) \to C^\infty(M, \Lambda^{p+1} T^*M \otimes E)$, for any p, such that*

$$d^\nabla(\varphi \wedge u) = (-1)^p \varphi \wedge \nabla u + (d\varphi) \wedge u, \tag{5.1.6}$$

if φ is a scalar p form and u is a section of E. We have $d^\nabla = \nabla$ if $p = 0$, and (5.1.6) *remains valid if u is a section of $\wedge^q T^*M \otimes E$ and ∇u is replaced by $d^\nabla u$.*

Proof. Uniqueness is obvious so it suffices to prove the other statements locally, assuming that E is a trivial bundle. If also ∇ is trivial, the statement is just standard calculus of exterior differential forms. If we add to the trivial connection an operator A of order 0, as in Proposition 5.1.2, then $d^\nabla u$ just gets an additional term $(-1)^p Au$ if u is a section of $\wedge^p T^*M \otimes E$, and these contributions on the two sides of (5.1.6) cancel.

If u is a section of E, it follows from (5.1.6) for scalar φ that

$$d^\nabla d^\nabla(\varphi u) = d^\nabla(\varphi \nabla u + d\varphi \wedge u) = \varphi d^\nabla d^\nabla u + d\varphi \wedge \nabla u - d\varphi \wedge \nabla u = \varphi d^\nabla d^\nabla u,$$

which shows that $d^\nabla d^\nabla$ is a two form with values in $\operatorname{Hom}(E, E) \cong E \otimes E^*$. We claim that

$$d^\nabla d^\nabla = R^\nabla. \tag{5.1.7}$$

For the proof we note that in terms of local coordinates $x^1, \dots, x^n$ in M, with $\varepsilon_j = \partial/\partial x^j$, we have

$$d^\nabla u = \sum dx^i \nabla_{\varepsilon_i} u, \quad d^\nabla d^\nabla u = -\sum dx^i \wedge dx^j \nabla_{\varepsilon_j} \nabla_{\varepsilon_i};$$

if $X = \sum X^i \varepsilon_i$ and $Y = \sum Y^j \varepsilon_j$ this means that

$$\begin{aligned}(d^\nabla d^\nabla u)(X, Y) &= -\sum (X^i Y^j - X^j Y^i) \nabla_{\varepsilon_j} \nabla_{\varepsilon_i} u \\ &= \sum X^j Y^i (\nabla_{\varepsilon_j} \nabla_{\varepsilon_i} - \nabla_{\varepsilon_i} \nabla_{\varepsilon_j}) u = R^\nabla(X, Y) u,\end{aligned}$$

which proves (5.1.7).

One can also argue as in Sections 4.1 and 4.2. Since we are only interested in forms on the base manifold we shall study pullbacks of forms analogous to those studied there. Note that if s is a local section of $P(M)$ and we write $s^*e_\alpha = E_\alpha$, which is a local basis for TM, we saw in the proof of (4.1.13) that $s^*\omega_{\alpha\beta} = (\nabla E_\alpha, E_\beta)$, that is, $\nabla E_\alpha = \sum (s^*\omega_{\alpha\beta})E_\beta$. We take this as our starting point now.

Let $e_1, \dots, e_N$ be a local basis for the vector bundle E, that is, sections of E over an open set U which form a basis in the fiber E_x for every $x \in U$. Omitting the tensor product sign $\otimes$, we can then write

$$\nabla e_i = \sum_{j=1}^N \omega_{ji} e_j, \quad i = 1, \dots, N, \tag{5.1.8}$$

where ω_{ij} are one forms. Thus

$$d^\nabla d^\nabla e_i = \sum_{j=1}^N (d\omega_{ji})e_j - \sum_{k=1}^N \omega_{ki} \nabla e_k = \sum_{j=1}^N \Omega_{ji} e_j, \quad i = 1, \dots, N, \tag{5.1.9}$$

$$\Omega_{ji} = d\omega_{ji} + \sum_{k=1}^N \omega_{jk} \wedge \omega_{ki}, \quad i, j = 1, \dots, N. \tag{5.1.10}$$

If $u = \sum u_i e_i$ is a general local section of E, it follows that

$$d^\nabla d^\nabla u = \sum_{i,j=1}^N u_i \Omega_{ji} e_j = \sum_{i=1}^N \big(\sum_{j=1}^N \Omega_{ij} u_j\big) e_i,$$

so (Ω_{ij}) is the matrix of R^∇.

Exercise 5.1.1. Verify by direct computation that if $e'_i = \sum b_{ji} e_j$ is another local basis, then $\Omega' = B^{-1}\Omega B$ for the corresponding matrices of curvature forms and $B = (b_{ij})$.

From (5.1.10) we obtain the differential Bianchi identity

$$d\Omega_{ji} = \sum_{k=1}^N \Omega_{jk} \wedge \omega_{ki} - \sum_{k=1}^N \omega_{jk} \wedge \Omega_{ki} \tag{5.1.11}$$

by repeating the proof of (4.2.6). With matrix notation this can be written more succinctly

$$d\Omega = [\Omega, \omega], \tag{5.1.11$'$}$$

where the matrix multiplication is made using the wedge product of course. (See also the discussion following (5.3.8).)

The characterization of the covariant differentiation in Theorem 3.1.2 involved the conditions (3.1.3) and (3.1.4). The first is only meaningful for the tangent bundle, but the second has a general analogue:

Definition 5.1.5. A real (complex) vector bundle is called Euclidean (Hermitian) if in each fiber there is given a Euclidean (Hermitian) scalar product $(\cdot, \cdot)$ such that (e, e) is a C^∞ function on E.

Proposition 5.1.6. *In every real (complex) vector bundle it is possible to introduce a Euclidean (Hermitian) structure.*

Proof. In a coordinate patch U where the bundle is $\cong U \times \mathbf{R}^N$ ($\cong U \times \mathbf{C}^N$), we can just take the standard Euclidean (Hermitian) form in $\mathbf{R}^N$ (in $\mathbf{C}^N$). By piecing together with a partition of unity as in the proof of Proposition 5.1.2, we obtain a global form.

Proposition 5.1.7. *In a real (complex) vector bundle with Euclidean (Hermitian) structure, it is always possible to find a connection ∇ such that*

$$d\,(u, v) = (\nabla u, v) + (u, \nabla v), \tag{5.1.12}$$

when u and v are sections of E. For the corresponding curvature matrices in terms of orthonormal bases this implies that $\Omega_{ij} = -\overline{\Omega_{ji}}$, thus $\Omega_{ij} = -\Omega_{ji}$ in the real case.

Proof. By the Gram-Schmidt orthogonalization procedure one can in some neighborhood U of any point find an orthonormal basis for E, which means that the identification of E over U with $U \times \mathbf{R}^N$ (resp. $U \times \mathbf{C}^N$) carries the Euclidean (Hermitian) form in the bundle to the standard form in $\mathbf{R}^N$ (in $\mathbf{C}^N$). The trivial connection in $U \times \mathbf{R}^N$ (in $U \times \mathbf{C}^N$) obviously has the property (5.1.12), so we have such a connection in the restriction of the bundle E to U. Piecing together such local connections by a partition of unity as in the proof of Proposition 5.1.2, we obtain a global connection such that (5.1.12) holds.

If $e_1, \dots, e_N$ is a local orthonormal basis, it follows from (5.1.8) that

$$\omega_{ji} = (\nabla e_i, e_j).$$

Since $(e_i, e_j) = \delta_{ij}$, we obtain using (5.1.12) that

$$\omega_{ij} + \overline{\omega_{ji}} = 0, \quad \text{hence} \quad \overline{\Omega_{ji}} = -d\omega_{ij} - \sum_{k=1}^{N} \omega_{kj} \wedge \omega_{ik} = -\Omega_{ij},$$

which completes the proof.

So far we have only discussed connections in a fixed bundle. However, applying a connection to a section of a bundle one gets a section of another bundle, the tensor product with the cotangent bundle, so we must also examine how to obtain a connection there. Thus consider now two C^∞ vector bundles E and F over M, with connections denoted ∇^E and ∇^F. In $E \oplus F$ we get an obvious connection: A section consists of one section of E and one section of F, and we can apply ∇^E and ∇^F to them. The tensor product is not quite so obvious:

Proposition 5.1.8. *Given connections ∇^E and ∇^F in E and in F, there is one and only one connection ∇ in $E \otimes F$ such that*

$$\nabla(u \otimes v) = (\nabla^E u) \otimes v + u \otimes (\nabla^F v), \tag{5.1.13}$$

if u and v are sections of E and of F.

Proof. The uniqueness is obvious so it suffices to prove existence locally. Thus we may assume that there is a global basis $v_1, \dots, v_N$ for sections of F. Every section s of $E \otimes F$ can then uniquely be written in the form $s = \sum_1^N u_j \otimes v_j$, where u_j is a section of E. If (5.1.13) holds we must have

$$\nabla s = \sum_1^N (\nabla^E u_j) \otimes v_j + \sum_1^N u_j \otimes (\nabla^F v_j),$$

and we define ∇ now by this equation. If $v = \sum_1^N \varphi_j v_j$ with scalar $\varphi_j \in C^\infty$, and if u is a section of E, we have by definition

$$\begin{aligned} \nabla(u \otimes v) &= \sum \nabla^E(\varphi_j u) \otimes v_j + \sum \varphi_j u \otimes (\nabla^F v_j) \\ &= \nabla^E u \otimes \sum \varphi_j v_j + \sum d\varphi_j (u \otimes v_j) + \sum u \otimes \varphi_j (\nabla^F v_j) = (\nabla^E u) \otimes v + u \otimes (\nabla^F v), \end{aligned}$$

where we have used first that ∇^E is a connection, then that ∇^F is a connection. Thus (5.1.13) is valid which shows that the construction is independent of the choices made. That ∇ is a connection follows at once from (5.1.13) if we replace u by φu and use that ∇^E is a connection.

If E is a vector bundle with connection ∇ and E^* is the dual bundle, with fiber E_x^* equal to the dual space of E_x, then there is a unique connection ∇^* in E^* such that

$$d\langle u, v\rangle = \langle \nabla u, v\rangle + \langle u, \nabla^* v\rangle \tag{5.1.14}$$

for arbitrary sections u and v of E and E^*. We leave for the reader the simple verification that (5.1.14) defines a connection ∇^*.

5.2. Chern classes. If T is a linear transformation in a finite dimensional vector space V, then the determinant of T can be defined as the determinant of the matrix A representing T with respect to a coordinate system in V. This is independent of the choice of coordinates, for if $x' = Bx$, where B is an invertible matrix, is another system of coordinates, then the matrix BAB^{-1} for T in the new coordinates has the same determinant.

Now let φ be any polynomial in the vector space $\mathbf{M}(N)$ of $N \times N$ (complex) matrices which is invariant in the sense that

$$\varphi(A) = \varphi(B^{-1}AB) \tag{5.2.1}$$

for every $B \in \mathbf{GL}(N)$, the group of invertible matrices in $\mathbf{M}(N)$. This is of course equivalent to

$$\varphi(AB) = \varphi(BA) \tag{5.2.1$'$}$$

for all $A, B \in \mathbf{M}(N)$. An example is $\varphi(A) = \det(I + \lambda A)$ for an arbitrary λ, hence also the coefficient of λ^k for any k. (Every invariant polynomial is in fact a polynomial in these special ones, but we shall not give the proof here.)

Lemma 5.2.1. *If φ is an invariant polynomial, then the derivative of $\varphi(A)$ in the direction $[A, C]$ vanishes for any $C \in \mathbf{M}(N)$.*

Proof. If $B = I + \varepsilon C$, then $B^{-1}AB = A + \varepsilon[A, C] + O(\varepsilon^2)$. Since $(\varphi(BAB^{-1}) - \varphi(A))/\varepsilon = 0$, the statement follows when $\varepsilon \to 0$.

When we use the lemma below it is convenient to assume that φ is homogeneous, of degree m, and use the symmetric multilinear form $\Phi(A_1, \dots, A_m)$ such that $\varphi(A) = \Phi(A, \dots, A)$. The conclusion in the lemma is then that

$$m\Phi([A, C], A, \dots, A) = 0, \quad A, C \in \mathbf{M}(N). \tag{5.2.2}$$

Now let Ω be the matrix of curvature forms for a vector bundle E with fiber dimension N, defined by means of a local basis. Since all two forms commute, we can form $\varphi(\Omega)$ unambiguously, and since another choice of local basis just replaces Ω by $B^{-1}\Omega B$ for some invertible matrix B (cf. Exercise 5.1.1), it follows that $\varphi(\Omega)$ is independent of the choice of local basis, so we obtain a globally defined differential form for which we shall also use the notation $\varphi(\Omega)$.

Theorem 5.2.2. *If φ is an invariant polynomial in $\mathbf{M}(N)$ then $\varphi(\Omega)$ is a closed differential form, and its cohomology class does not depend on the choice of connection.*

Proof. Locally we write

$$\varphi(\Omega) = \Phi(\Omega, \dots, \Omega)$$

where the curvature matrix Ω satisfies the Bianchi identity (5.1.11)$'$. Since the forms Ω_{ij} have even degree, we can apply d separately to each argument in the multilinear form Φ. Two forms commute with arbitrary differential forms so we obtain using (5.2.2)

$$d\varphi(\Omega) = m\Phi(d\Omega, \Omega \dots, \Omega) = m\Phi([\Omega, \omega], \Omega, \dots, \Omega) = 0.$$

The proof that the cohomology class is independent of the choice of connection is essentially the same as the proof of Theorem 4.4.2. We keep the notation used there. If ∇^0 and ∇^1 are two connections for E over M, and $\widetilde{E}$ is the vector bundle E pulled back to $M \times \mathbf{R}$ by the projection $M \times \mathbf{R} \to M$, then

$$\widetilde{\nabla} = (1 - \chi)\nabla^0 + \chi\nabla^1 + \partial/\partial t\, dt$$

is a connection for $\widetilde{E}$ on $M \times \mathbf{R}$. Note that if $x \in M$ then the fiber of $\widetilde{E}$ over (x, t) is always equal to E_x, so the derivative of a section with respect to t is well defined. It is now obvious that the matrices of curvature forms for the connections ∇^0 and ∇^1 on M are pullbacks from M to $M \times \mathbf{R}$ by the homotopic maps i_0 and i_1 of the matrices of curvature forms for $\widetilde{\nabla}$, which proves that the cohomology class of $\varphi(\Omega)$ is independent of the choice of connection.

Definition 5.2.3. If E is a complex vector bundle over M and Ω is the (local) matrix of curvature forms with respect to a connection, then the cohomology class of

$$\det(I + i\Omega/2\pi), \tag{5.2.3}$$

where I is the unit matrix, is called the *total Chern class* of E. It can be written $1 + c_1 + c_2 + \dots$, where c_j is of degree $2j$ and is called the jth Chern class.

If E is the complexification of the tangent bundle of a Riemannian manifold with the covariant differentiation as connection, it is clear that $(-1)^j c_{2j}$ is the Pontrjagin class of degree $4j$ and that $c_{2j+1} = 0$. More generally, if E is the complexification of a real vector bundle, it follows from Propositions 5.1.6 and 5.1.7 that we can choose a connection such that $\Omega_{ij} = -\Omega_{ji}$ is a real form. That was all we used to prove that (4.4.2) vanishes for odd k. For a general complex vector bundle it follows from Propositions 5.1.6 and 5.1.7 that we can choose a connection such that $\overline{\Omega_{ij}} = -\Omega_{ji}$, which implies that complex conjugation preserves (5.2.3). Hence we obtain:

Proposition 5.2.4. *The Chern classes of any complex vector bundle are real. For the complexification of a real vector bundle the Chern classes c_j vanish when j is odd.*

This proposition motivates the factor i in (5.2.3). To explain the factor 2π we shall now discuss the case of a complex line bundle L in a manifold M. We introduce a Hermitian structure in L and choose an open covering $\{U_j\}$ such that L is trivial, that is, has a section s_j with norm 1 at every point in U_j. An arbitrary section can then in U_j be written in the form $u_j s_j$. In $U_j \cap U_k$ we have $u_j s_j = u_k s_k$ then, hence writing $g_{jk} s_j = s_k$ we obtain $|g_{jk}(x)| = 1$, $x \in U_j \cap U_k$, and

$$u_j = g_{jk} u_k.$$

The functions g_{jk} are the *transition functions* of L. With a connection ∇ in L satisfying (5.1.12) we now set

$$\nabla s_j = \Gamma_j s_j$$

where $i\Gamma_j$ is a real one form in U_j by (5.1.12). The curvature form is then

$$\Omega = d\Gamma_j + \Gamma_j \wedge \Gamma_j = d\Gamma_j \quad \text{in } \ U_j,$$

so c_1 is the cohomology class of the two form equal to $id\Gamma_j/(2\pi)$ in U_j for every j. Note that

$$\Gamma_k s_k = \nabla s_k = \nabla(g_{jk} s_j) = g_{jk}\Gamma_j s_j + s_j dg_{jk} \quad \text{in } U_j \cap U_k,$$

which means that

$$\Gamma_k = \Gamma_j + dg_{jk}/g_{jk}, \quad d\Gamma_j = d\Gamma_k \qquad \text{in} \quad U_j \cap U_k.$$

If the cohomology class is equal to 0, then there is a global one form Γ in M such that $d(\Gamma - \Gamma_j) = 0$ in U_j for every j. If U_j has been chosen simply connected, it follows that we can find a function φ_j in U_j such that $\Gamma - \Gamma_j = d\varphi_j$ in U_j; we can choose Γ and φ_j purely imaginary. Since $\Gamma_k - \Gamma_j = d\varphi_j - d\varphi_k$ in $U_j \cap U_k$, this means that

$$dg_{jk}/g_{jk} = d\varphi_j - d\varphi_k,$$

so $c_{jk} = g_{jk} e^{\varphi_k - \varphi_j}$ is locally constant, that is, constant in each component of $U_j \cap U_k$. If we replace the sections s_j by $s_j e^{\varphi_j}$, we get new transition functions which are locally constant. But in a line bundle with locally constant transition functions we can define a connection ∇ by $\nabla s_j = 0$ in U_j, and then all Γ_j become 0. Thus we have proved:

Proposition 5.2.5. *The Chern class of a line bundle vanishes if and only if it can be defined by locally constant transition functions.*

We shall now consider the example where M is the sphere S^2 of dimension 2. If L is a complex line bundle on S^2, we can find a non-zero section s_1 in a neighborhood U_1 of the upper hemisphere and another s_2 in a neighborhood U_2 of the lower one. We write as before $\nabla s_j = \Gamma_j s_j$ and recall that the Chern form is $id\Gamma_j/2\pi$ in U_j,

$$\Gamma_1 = \Gamma_2 + dg_{21}/g_{21} \quad \text{in} \;\; U_1 \cap U_2.$$

Choose a one form Γ in S^2 equal to Γ_2 in a neighborhood of the lower hemisphere. Then the Chern class contains the form $c = i(d\Gamma_1 - d\Gamma)/2\pi$ in U_1 which vanishes in the lower hemisphere, and by Stokes' theorem

$$\int_{S^2} c = i/2\pi \int (\Gamma_1 - \Gamma) = i/2\pi \int (\Gamma_1 - \Gamma_2) = -i/2\pi \int dg_{12}/g_{12},$$

the last three integrals taken over the equator. This is the winding number of g_{12} on the equator, thus an integer. In this example the effect of the factor 2π is therefore that $\int c$ can precisely take integer values, so that the Chern class is an integer cohomology class. It would take us too far into a discussion of cohomology theory to give a general form of this result here; the example should suffice as a motivation.

5.3. Lie groups. In Section 4.1 we worked systematically with forms on the orthonormal frame bundle of a Riemannian manifold. Each fiber becomes isomorphic to the orthogonal group $\mathbf{O}(n)$ when a point in it is distinguished. In Section 5.4 we shall clarify the results by putting them in a more general context. In order to cover the results in Sections 5.1 and 5.2 we shall also need other groups than $\mathbf{O}(n)$. Subgroups of the full linear groups suffice for our purposes, but the arguments become conceptually more transparent in an abstract setting where $\mathbf{O}(n)$ is replaced by an arbitrary Lie group, so we shall give a brief introduction to this concept here.

Definition 5.3.1. A group $\mathbf{G}$ which has also the structure of a C^∞ manifold is called a Lie group if

$$\mathbf{G} \times \mathbf{G} \ni (a, b) \mapsto ab^{-1} \in \mathbf{G}$$

is a C^∞ map.

In particular, the hypothesis implies that $b \mapsto b^{-1}$ and $(a, b) \mapsto ab$ are C^∞ maps. It would suffice to make the latter assumption:

Exercise 5.3.1. Show that if $\mathbf{G}$ is a group which is also a C^∞ manifold, then $\mathbf{G}$ is a Lie group if $\mathbf{G} \times \mathbf{G} \ni (a, b) \mapsto ab \in \mathbf{G}$ is a C^∞ map.

Let L_a be the *left translation* $L_a x = ax$, $a, x \in \mathbf{G}$, which maps the identity $e \in \mathbf{G}$ to a, thus $L'_a : T_e \to T_a$. If $X_0 \in T_e$, we define a C^∞ vector field X on $\mathbf{G}$ by

$$X(a) = L'_a X_0 \in T_a.$$

For arbitrary $a, b \in \mathbf{G}$ we have $L'_a X(b) = L'_a L'_b X_0 = L'_{ab} X_0 = X(ab)$ because $L_a L_b = L_{ab}$ by the associative law. Thus X is a *left invariant vector field.* If $\omega_0 \in T_e^*$ we define similarly a C^∞ one form on $\mathbf{G}$ so that ω is equal to ω_0 at the identity and $\omega(X)$ is a constant if X is a left invariant vector field. Then we have $L_a^* \omega = \omega$ for every $a \in \mathbf{G}$, and this left invariance together with the value ω_0 at e determine ω.

Let $X_1, \dots, X_n$ be left invariant vector fields such that $X_1(e), \dots, X_n(e)$ form a basis for T_e, and let $\omega^1, \dots, \omega^n$ be the left invariant one forms which are biorthogonal at e and hence at any $a \in \mathbf{G}$. Then we have

$$d\omega^i = \tfrac{1}{2} \sum_{j,k=1}^{n} c^i_{jk} \omega^j \wedge \omega^k, \tag{5.3.1}$$

for some constants c^i_{jk} with $c^i_{jk} = -c^i_{kj}$. In fact, this is obviously true at the identity, and both sides are left invariant, hence equal everywhere. The constants c^i_{jk} are called the *structural constants* of the Lie group. (They depend in an obvious way on the choice of basis in T_e^*.) We can also express (5.3.1) in terms of the left invariant vector fields if we use that by (C.4)

$$\langle X_j \wedge X_k, d\omega^i \rangle = X_j \langle X_k, \omega^i \rangle - X_k \langle X_j, \omega^i \rangle + \langle [X_k, X_j], \omega^i \rangle = \langle [X_k, X_j], \omega^i \rangle.$$

The left-hand side is c^i_{jk}, so we obtain

$$[X_k, X_j] = \sum c^i_{jk} X_i. \tag{5.3.1$'$}$$

Proposition 5.3.2. *In the tangent space $\mathfrak{g}$ of a Lie group $\mathbf{G}$ at the identity, there is a natural bilinear antisymmetric product $\mathfrak{g} \times \mathfrak{g} \ni X, Y \mapsto [X, Y]$ defined by taking the commutator of the left invariant vector fields extending X and Y. We have the Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0, \quad X, Y, Z \in \mathfrak{g}. \tag{5.3.2}$$

This means that the structural constants satisfy the conditions

$$\sum_{\nu=1}^{n} (c^\mu_{\nu i} c^\nu_{jk} + c^\mu_{\nu j} c^\nu_{ki} + c^\mu_{\nu k} c^\nu_{ij}) = 0, \quad \mu = 1, \dots, n. \tag{5.3.2$'$}$$

One calls $\mathfrak{g}$ a Lie algebra.

Proof. Only the Jacobi identity remains to prove, and it follows since it is valid for commutators of arbitrary vector fields just by the associativity of the products.

In Section 1.4 we discussed one parameter subgroups of the orthogonal group and of the general linear group. We shall now do the same for an arbitrary Lie group, so let us assume that $\varphi : \mathbf{R} \to \mathbf{G}$ is a C^1 function such that

$$\varphi(s + t) = \varphi(s)\varphi(t), \quad s, t \in \mathbf{R}. \tag{5.3.3}$$

This implies that $\varphi(0) = e$, and differentiation with respect to t gives when $t = 0$, if we write $\varphi(s)\varphi(t) = L_{\varphi(s)}\varphi(t)$,

$$\varphi'(s) = L'_{\varphi(s)} X_0 = X(\varphi(s)),$$

where $X_0 = \varphi'(0)$ and X is the left invariant vector field with $X(e) = X_0$. This system of differential equations in the local coordinates of $\mathbf{G}$ at e has a unique solution with $\varphi(0) = e$, when $|s| < 2\delta$, say. We claim that (5.3.3) is then valid if $|s| < \delta$ and $|t| < \delta$. Fix s and denote the left (right) side of (5.3.3) by $\varphi_1(t)$ (resp. $\varphi_2(t)$). Then $\varphi_1(0) = \varphi_2(0) = \varphi(s)$, and

$$\varphi_1'(t) = X(\varphi_1(t)), \quad \varphi_2'(t) = L'_{\varphi(s)} X(\varphi(t)) = X(\varphi(s)\varphi(t)) = X(\varphi_2(t)),$$

so the uniqueness theorems for ordinary differential equations show that $\varphi_1(t) = \varphi_2(t)$ when $|t| < \delta$. Now we can define $\varphi(s)$ uniquely for arbitrary s by setting

$$\varphi(s) = \varphi(s/N)^N$$

where N is a positive integer so large that $|s/N| < \delta$. This does not depend on N, for $\varphi(s/N) = \varphi(s/NM)^M$ for every positive integer M, hence $\varphi(s/N)^N = \varphi(s/M)^M$ if $|s/M| < \delta$ too. It is clear that (5.3.3) is valid for all s, t.

Proposition 5.3.3. *There is a unique C^∞ map* exp *from the Lie algebra $\mathfrak{g}$ of $\mathbf{G}$ to $\mathbf{G}$ such that the differential at the origin is the identity $\mathfrak{g} \to T_e(\mathbf{G}) = \mathfrak{g}$ and $\mathbf{R} \ni t \mapsto \exp(tX)$ is a one parameter subgroup for every $X \in \mathfrak{g}$. For X, Y in a neighborhood of 0 in $\mathfrak{g}$ we have $\exp X \exp Y = \exp \varphi(X, Y)$ where $\varphi(X, Y)$ is a C^∞ function with values in $\mathfrak{g}$ defined in a neighborhood of 0 in $\mathfrak{g} \times \mathfrak{g}$ such that*

$$\varphi(X, Y) = X + Y + \tfrac{1}{2}[X, Y] + O(|X||Y|(|X| + |Y|)). \tag{5.3.4}$$

Proof. Let $X_1, \dots, X_n$ be a basis for the left invariant vector fields. The one parameter subgroup with derivative $\sum a^j X_j(e)$ at 0 is the solution of the Cauchy problem

$$d\varphi(t)/dt = \sum_1^n a^j X_j(\varphi(t)), \quad \varphi(0) = e.$$

There is a unique solution for $|t| < \delta$ if $|a| \le 1$. It is a C^∞ function of a and t depending only on ta which we define to be $\exp(ta)$. The extension of the map is done as before by writing $\exp X = (\exp(X/N))^N$ for large N.

By the implicit function theorem the exponential map is a diffeomorphism of a neighborhood of the origin in $\mathfrak{g}$ on a neighborhood of the identity in $\mathbf{G}$. The function φ is therefore well defined as the composition of $\exp X \exp Y$ with the inverse of the exponential map, $\varphi \in C^\infty$, and $\varphi(sX, tX) = (s + t)X$ for small sX, tX. Hence it follows from Taylor's formula applied first in X, then in Y, that

$$\begin{aligned}\varphi(X, Y) - X - Y = \iint_{0 \le s,t \le 1} \varphi''_{XY}(sX, tY)(X, Y)\, ds\, dt& \\ = B(X, Y) + O(|X||Y|(|X| + |Y|))&\end{aligned}$$

where $B(X,Y)$ is a bilinear $\mathfrak{g}$ valued form with $B(X,X)=0$, hence skew symmetric. If $\widetilde{X}$ is the left invariant vector field equal to X at the identity, pushed back to $\mathfrak{g}$ from $\mathbf{G}$ by the inverse of the exponential map, then

$$\widetilde{X}(Z) = L'_Z(0)X, \quad L_Z = \varphi(Z,\cdot), \quad \text{hence}$$
$$\widetilde{X}(Z) = X + B(Z,X) + O(|Z|^2), \quad \text{as } Z \to 0.$$

If $\widetilde{Y}$ is a left invariant vector field equal to Y at 0, then we have a similar formula for $\widetilde{Y}$, hence

$$[\widetilde{X},\widetilde{Y}](0) = B(X,Y) - B(Y,X) = 2B(X,Y).$$

The left-hand side is $[X,Y]$, which completes the proof.

Note that for the full linear group the definition of the exponential map here agrees with the exponential of matrices used in Section 1.4.

Proposition 5.3.4. *If $X \in \mathfrak{g}$ and $a \in \mathbf{G}$, then*

$$a(\exp X)a^{-1} = \exp(\operatorname{Ad}(a)X), \tag{5.3.5}$$

where $\operatorname{Ad}(a) \in \mathbf{GL}(\mathfrak{g})$ *is the differential at e of the map $\mathbf{G} \ni x \mapsto axa^{-1} \in \mathbf{G}$. In a neighborhood of the identity* $\operatorname{Ad}(a)$ *is given by*

$$\operatorname{Ad}(\exp Y) = \exp(\operatorname{ad} Y), \quad \operatorname{ad} Y X = [Y,X].$$

Proof. The map $\mathbf{R} \ni t \mapsto a(\exp(tX))a^{-1} \in \mathbf{G}$ is a one parameter group, hence equal to $t \mapsto \exp(tZ)$ where Z is the derivative at 0 of $t \mapsto a(\exp(tX))a^{-1} \in \mathbf{G}$, hence equal to $\operatorname{Ad}(a)X$, which is a linear function of X. It is also clear that $\mathbf{R} \ni t \mapsto \operatorname{Ad}(\exp(tY)) \in \mathbf{GL}(\mathfrak{g})$ is a one parameter subgroup, hence equal to $\exp(t\operatorname{ad}(Y))$ for some linear transformation $\operatorname{ad}(Y)$ in $\mathfrak{g}$.To determine $\operatorname{ad}(Y)$ we use that by Proposition 5.3.3

$$\exp(tY)\exp(sX)\exp(-tY) = \exp(sX + st[Y,X] + O(st(|s|+|t|))),$$

and this implies that $\operatorname{Ad}(\exp(tY))X = X + t[Y,X] + O(t^2)$, hence that $\operatorname{ad}(Y)X = [Y,X]$. The proof is complete.

Formula (5.3.4) can be refined to the Campbell-Hausdorff formula

$$\varphi(X,Y) = \sum_{\nu=1}^{\infty}(-1)^{\nu+1}\nu^{-1} \sum_{\alpha_i+\beta_i\neq 0} (\operatorname{ad} X)^{\alpha_1}(\operatorname{ad} Y)^{\beta_1}\dots(\operatorname{ad} X)^{\alpha_\nu}(\operatorname{ad} Y)^{\beta_\nu - 1}Y/c_{\alpha\beta},$$
$$c_{\alpha\beta} = \sum_{i=1}^{\nu}(\alpha_i+\beta_i)\prod_{i=1}^{\nu}\alpha_i!\beta_i!,$$

for X and Y in a neighborhood of 0. (When $\beta_\nu = 0$ the last factor should be replaced by $(\operatorname{ad} X)^{\alpha_\nu - 1}X$.) In particular this shows that the Lie algebra determines the group, and using (5.3.4) one can also show that there is a local Lie group with any given Lie algebra. We shall not prove the Campbell-Hausdorff formula here but shall give another proof that the group operation is uniquely determined near e by the Lie algebra.

Introducing in a neighborhood of the identity in $\mathbf{G}$ the *canonical coordinates* in $\mathfrak{g}$ provided by the exponential map has the advantage that the one parameter groups become rays through the origin. Identifying $\mathfrak{g}$ with $\mathbf{R}^n$, we denote by $X_1, \dots, X_n$ the left invariant vector fields which are equal to the unit vectors in $\mathbf{R}^n$ at the origin. Then we know that for $a \in \mathbf{R}^n$, $s \in \mathbf{R}$ and $|a||s|$ small enough the map $s \mapsto sa$ is a one parameter subgroup. The derivative is a, and the left invariant vector field with this value at the origin is $\sum a^j X_j$, so it follows that $a = \sum a^j X_j(sa)$, that is,

$$x = \sum x^j X_j(x), \quad |x| < \delta. \tag{5.3.6}$$

In the left-hand side x denotes of course the radial vector field $\varrho = \sum x^j \partial/\partial x^j$. Taking the scalar product with the left invariant forms ω^k gives $\langle \varrho, \omega^k \rangle = x^k$. We shall derive differential equations for the determination of ω^i by writing down (5.3.1) explicitly. With $\omega^i = \sum \omega^i_\nu dx^\nu$ the equation means that

$$\partial\omega^i_\mu/\partial x^\nu - \partial\omega^i_\nu/\partial x^\mu = \sum_{j,k=1}^n c^i_{jk}\omega^j_\nu\omega^k_\mu, \quad i,\nu,\mu = 1,\dots,n, \tag{5.3.1$''$}$$

or if we multiply by x^ν and add, recalling that $\sum x^\nu \omega^i_\nu = x^i$,

$$(\varrho + 1)\omega^i_\mu = \delta^i_\mu + \sum c^i_{jk} x^j \omega^k_\mu.$$

This means that

$$\frac{\partial}{\partial t}\big(t\omega^i_\mu(t\theta)\big) = \delta^i_\mu + \sum c^i_{jk}\theta^j t\omega^k_\mu(t\theta)$$

which is a system of ordinary differential equations with constant coefficients for $t\omega^i_\mu(t\theta)$ from which ω^i_μ can be uniquely obtained. It follows that ω^i_μ is uniquely determined and analytic in a neighborhood of 0 in the canonical coordinates. We shall now prove that also the multiplication law is uniquely determined and analytic in a neighborhood of the identity. (Additional arguments using the conditions (5.3.2)$'$ show that all equations (5.3.1)$'$ are fulfilled and that there is a local Lie group with arbitrarily prescribed Lie algebra. However, we shall not give the details here but content ourselves with the local uniqueness of the group.) To do so we note that if $\varphi(t) = a\exp(tX)$, where $a \in \mathbf{G}$ and $X \in \mathfrak{g}$, then

$$\varphi'(t) = L'_{\varphi(t)}X = \widetilde{X}(\varphi(t)), \quad \varphi(0) = a,$$

if $\widetilde{X}$ is the left invariant vector field equal to X at the identity. This has a unique solution which is analytic in a and tX, when a is in a neighborhood of e in $\mathbf{G}$ and tX is in a neighborhood of 0 in $\mathfrak{g}$. The proof that the structure constants determine the multiplication uniquely in a neighborhood of the identity is now complete. Note that it would have been enough to assume that $\mathbf{G}$ is a C^3 manifold with C^3 group operations to conclude that it is actually an analytic manifold. Hilbert's fifth problem was to show that it suffices to assume that $\mathbf{G}$ is a topological manifold with continuous group operations. The proof was completed in the early 1950's by Gleason, Montgomery and Zippin. The proof also showed that any topological group having a neighborhood of the identity containing no subgroup $\neq \{e\}$ is a Lie group. In particular, every closed subgroup of a Lie group is a Lie group, which is also easily proved directly (cf. Exercise 1.4.5).

The special case of formula (5.3.1) for the orthogonal group is (4.1.3). Before returning to a discussion of connections we shall put (5.3.1) in another form which avoids the explicit use of a basis in $\mathfrak{g}$. Recall that having chosen a basis $X_1, \dots, X_n$ for $\mathfrak{g} = T_e\mathbf{G}$, we defined the left invariant form ω^j at $a \in \mathbf{G}$ by

$$\omega^j(L'_a X) = x^j, \quad \text{if } X = \sum_1^n x^j X_j \in T_e\mathbf{G}.$$

Thus $\sum X_j \omega^j(L'_a X) = X$, $X \in T_e\mathbf{G}$, so the differential form $\omega = \sum X_j \omega^j$ with values in $\mathfrak{g}$ is simply defined by

$$\omega(L'_a X) = X, \quad X \in \mathfrak{g} = T_e\mathbf{G}. \tag{5.3.7}$$

In $T_a\mathbf{G}$ it is therefore the pullback by $L_{a^{-1}}$ of the identity in $\mathfrak{g} = T_e\mathbf{G}$.

We can now write (5.3.1) in the compact form

$$d\omega = -\tfrac{1}{2}[\omega \wedge \omega], \tag{5.3.1)'''$$

where by (5.3.1) and (5.3.1)$'$

$$[\omega \wedge \omega] = -\sum c^i_{jk} X_i \omega^j \wedge \omega^k = \sum [X_j, X_k]\omega^j \wedge \omega^k. \tag{5.3.8}$$

To justify the notation we note that in general, if ω_1, ω_2 are differential forms with values in finite dimensional vector spaces W_1 and W_2, then we can define $\omega_1 \wedge \omega_2$ uniquely as a form with values in the tensor product $W_1 \otimes W_2$ so that it is equal to $(w_1 \otimes w_2)\sigma_1 \wedge \sigma_2$ if $\omega_j = w_j \sigma_j$, $j = 1, 2$, with $w_j \in W_j$ and scalar differential forms σ_j. In fact, this expression is bilinear in w_1, σ_1 and in w_2, σ_2, and the space of forms with values in W_j is the tensor product of W_j and the space of scalar forms. If we have a bilinear map $\beta : W_1 \times W_2 \to W_3$, a third vector space, then composition with the corresponding map $\tilde\beta : W_1 \otimes W_2 \to W_3$ gives a form $\tilde\beta \omega_1 \wedge \omega_2$ with values in W_3. In particular, if $W_1 = W_2 = \mathfrak{g}$ and β is the Lie bracket we get precisely the definition of $[\omega \wedge \omega]$ in (5.3.8). Of course the equation (5.3.1)$'''$ has the same contents as (5.3.1) or (5.3.1)$''$ but it is much more compact thanks to the coordinate free notation.

In $\mathfrak{g}$ there is a natural symmetric bilinear form

$$B(X, Y) = \operatorname{Tr}(\operatorname{ad} X \ \operatorname{ad} Y), \quad X, Y \in \mathfrak{g}, \tag{5.3.9}$$

called the *Killing form*. It is invariant under the adjoint action,

$$B(\operatorname{Ad}(a)X, \operatorname{Ad}(a)Y) = B(X, Y), \quad a \in \mathbf{G},\ X, Y \in \mathfrak{g}. \tag{5.3.10}$$

In fact, it follows from (5.3.4) that

$$[\mathrm{Ad}(a)X, \mathrm{Ad}(a)Z] = \mathrm{Ad}(a)[X, Z], \tag{5.3.11}$$

if we expand $a \exp X \exp(-Z) \exp X \exp Z a^{-1}$ at $X = Z = 0$ inserting factors $a^{-1}a$ between the exponentials. With $A = \mathrm{Ad}(a)$ (5.3.11) means that $(\mathrm{ad}(AX))A = A \,\mathrm{ad}\, X$, hence

$$B(AX, AY) = \mathrm{Tr}\,(\mathrm{ad}(AX)\ \mathrm{ad}(AY)) = \mathrm{Tr}\,(A(\mathrm{ad}\, X\ \mathrm{ad}\, Y)A^{-1}) = B(X, Y),$$

which proves (5.3.10). In many cases of interest to us the Killing form gives a natural Euclidean metric in $\mathfrak{g}$, invariant under $\mathrm{Ad}\, a$, $a \in \mathbf{G}$:

Proposition 5.3.5. *If* $\mathbf{G}$ *is a compact Lie group, then the Killing form is negative semi-definite; we have* $B(X, X) = 0$ *if and only if* $\mathrm{ad}\, X = 0$.

Proof. The compactness of $\mathbf{G}$ guarantees that there is a Euclidean scalar product (X, Y), $X, Y \in \mathfrak{g}$ which is invariant under $\mathrm{Ad}\, a$ for every $a \in \mathbf{G}$. (If (X, Y) is an arbitrary Euclidean scalar product we just have to replace it by the integral of $(\mathrm{Ad}\, aX, \mathrm{Ad}\, aY)$ over $\mathbf{G}$ with respect to the invariant measure on the group.) With corresponding orthonormal coordinates in $\mathfrak{g}$, the matrix of $\mathrm{Ad}\, a$ is then orthogonal for every $a \in \mathbf{G}$, hence the matrix (X_{jk}) of $\mathrm{ad}\, X$ is skew symmetric for every $X \in \mathfrak{g}$, as proved in Section 1.4. But then $\mathrm{Tr}(\mathrm{ad}\, X\ \mathrm{ad}\, X) = \sum X_{jk}X_{kj} = -\sum X_{jk}^2 \le 0$, with equality only if $\mathrm{ad}\, X = 0$.

Definition 5.3.6. A Lie group is called *semi-simple* if the Killing form is non-degenerate.

For a semi-simple compact Lie group changing the sign of the Killing form therefore gives a natural Euclidean metric in the Lie algebra. Let us look at the examples which will be needed later on; to deal with them we do not really need Proposition 5.3.5 since everything is done quite explicitly. The first example is the orthogonal group $\mathbf{O}(N)$. As we saw in Section 1.4, the Lie algebra consists of the skew symmetric matrices so the Killing form is negative definite (see the proof of Proposition 5.3.5). Let us now consider instead the unitary group $\mathbf{U}(N)$ of $N \times N$ unitary matrices. The Lie algebra consists of the matrices iH where H is hermitian symmetric. If $\mathrm{ad}\, Hh = [H, h] = 0$ for every Hermitian symmetric matrix h, then H is a multiple of the identity. This is clear if H has diagonal form, and we can always reduce to that case by a unitary transformation. Hence the Killing form vanishes in a one dimensional space, so $\mathbf{U}(N)$ is not semi-simple. However, the subgroup $\mathbf{SU}(N)$ is semi-simple, for its Lie algebra consists of the matrices iH where H is Hermitian with zero trace. If $[H, h] = 0$ for all h satisfying the same condition, we conclude again that H is a multiple of the identity, and since the trace vanishes it follows that $H = 0$. At least the groups $\mathbf{SU}(N)$ with $N = 2, 3, 4, 5, 8, 16$ have been proposed by physicists in connection with unified field theories.

5.4. Principal and associated bundles. Let us begin by recalling the definition of a C^∞ complex vector bundle E with fiber dimension N over a C^∞ manifold M. First of all, E is supposed to be a C^∞ manifold with

(i) a C^∞ map $\pi : E \to M$, called the projection;
(ii) a vector space structure in each fiber $E_x = \pi^{-1}(x)$.

These data are required to be *locally trivial* which means that

(iii) for every point in M there is a neighborhood U and a C^∞ diffeomorphism $\varphi : \pi^{-1}(U) \to U \times \mathbf{C}^N$ such that φ restricts to a linear isomorphism $E_x \to \{x\} \times \mathbf{C}^N \cong \mathbf{C}^N$ for every $x \in M$.

Let $\{U_i\}_{i\in I}$ be an open covering of M such that for every $i \in I$ there is a diffeomorphism $\varphi_i : \pi^{-1}(U_i) \to U_i \times \mathbf{C}^N$ with the properties listed above. Then we can regard

$$g_{ij} = \varphi_i \varphi_j^{-1}$$

as a C^∞ map $U_i \cap U_j \to \mathbf{GL}(N, \mathbf{C})$, called a *transition function*, and we have the *cocycle conditions*

$$\begin{aligned} g_{ij} g_{ji} &= \text{identity} \quad \text{in } U_i \cap U_j \\ g_{ij} g_{jk} g_{ki} &= \text{identity} \quad \text{in } U_i \cap U_j \cap U_k. \end{aligned} \tag{5.4.1}$$

The bundle E can be reconstructed using the system of transition functions g_{ij}. In fact, let $\widetilde{E}$ be the set of all $(i, x, w) \in I \times M \times \mathbf{C}^N$ with $x \in U_i$ and the equivalence relation

$$(i, x, w) \sim (j, x', w') \quad \text{if } x = x' \text{ and } w' = g_{ji} w. \tag{5.4.2}$$

The reflexivity and symmetry of this relation follow from the first part of (5.4.1), and the second part of (5.4.1) gives the transitivity. For any given family of transition functions with values in $\mathbf{GL}(N, \mathbf{C})$ satisfying (5.4.1), it is clear that the space of equivalence classes of $\widetilde{E}$, with the projection induced by the map $\widetilde{E} \ni (i, x, w) \to x \in M$ is a C^∞ complex vector bundle of fiber dimension N. (The linear structure is inherited from $\mathbf{C}^N$ of course.) If the transition matrices are obtained from a given vector bundle E as above, it is also clear that we get back an isomorphic bundle.

For a real vector bundle there is no real change in the argument except that the transition functions g_{ij} will take their values in the smaller group $\mathbf{GL}(N, \mathbf{R})$. The group can be reduced further if we recall that by Proposition 5.1.6 it is always possible to introduce a Hermitian (Euclidean) structure in a complex (real) vector bundle. The local trivializations can then be chosen so that they respect the structure, for the local frames in an arbitrary local trivialization can be orthonormalized using the Gram-Schmidt procedure. If only such trivializations are used, then the transition functions take their values in $\mathbf{U}(N)$ (resp. $\mathbf{O}(N)$). If we have a real vector bundle which is oriented, that is, each fiber is oriented and there exist local trivializations respecting the orientation, then g_{ij} takes its values in $\mathbf{SO}(N)$.

To avoid looking at a large number of cases we now consider an arbitrary Lie group $\mathbf{G}$. Assume that we have an open covering $\{U_i\}_{i\in I}$ of the C^∞ manifold M and C^∞ transition functions $g_{ij} : U_i \cap U_j \to \mathbf{G}$ satisfying (5.4.1). Let ϱ be any representation of $\mathbf{G}$ in a finite dimensional vector space F, that is, $\varrho(g)$ is for every $g \in \mathbf{G}$ a linear transformation in F such that

$$\varrho(g_1)(\varrho(g_2) w) = \varrho(g_1 g_2) w, \quad w \in F,\ g_1, g_2 \in \mathbf{G}. \tag{5.4.3}$$

Following the reconstruction of E above we now form the set of all $(i, x, w) \in I \times M \times F$ with $x \in U_i$. Then

$$(5.4.2)' \qquad (i, x, w) \sim (j, x', w') \quad \text{if } x = x' \text{ and } w' = \varrho(g_{ji})w,$$

is an equivalence relation, for if $(i, x, w) \sim (j, x, w') \sim (k, x, w'')$, then

$$w'' = \varrho(g_{kj})w' = \varrho(g_{kj})(\varrho(g_{ji})w) = \varrho(g_{kj}g_{ji})w = \varrho(g_{ki})w,$$

so $(i, x, w) \sim (k, x, w'')$. The set of equivalence classes is a vector bundle F_ϱ with fiber dimension $\dim F$ and the projection induced by the map $(i, x, w) \mapsto x \in M$. The restriction $\pi^{-1}(U_i)$ is identified with $U_i \times F$ by the map

$$U_i \times F \ni (x, w) \mapsto (i, x, w)^\sim \in F_\varrho,$$

and the corresponding transition functions are $\varrho(g_{ij})$. If the representation ϱ of $\mathbf{G}$ on F is *faithful*, so that $\varrho(g)$ =identity implies $g = e$, this means that the transition functions g_{ij} can be recovered from F_ϱ. It is clear that changing coverings and trivializations gives an isomorphic bundle F_ϱ, for if we have two such sets then each defines a bundle which is isomorphic to that defined by the union.

In the preceding discussion we assumed that F is a vector space, but the discussion of the equivalence relation used only (5.4.3), so we could use any space where $\mathbf{G}$ acts. In particular, taking $F = \mathbf{G}$ and ϱ equal to left multiplication, we obtain a *principal bundle* P. Note that if $(i, x, g) \sim (j, x, g')$, then $(i, x, ga) \sim (j, x, g'a)$, so right translation by $a \in \mathbf{G}$ will be defined on the set P of equivalence classes. Thus the conditions in the following definition are all fulfilled:

Definition 5.4.1. Let P and M be C^∞ manifolds, let $\pi : P \to M$ be a C^∞ map, and let $\mathbf{G}$ be a Lie group. Then P is called a principal $\mathbf{G}$ bundle over M with projection π if

(i) $\mathbf{G}$ acts *freely* to the right on P, that is, we have a map $P \times \mathbf{G} \ni (p, g) \mapsto pg \in P$ such that $(pg_1)g_2 = p(g_1g_2)$ if $g_1, g_2 \in \mathbf{G}$, $pe = p$, and $pg \neq p$ for every $p \in P$ if $\mathbf{G} \ni g \neq e$;

(ii) For every $p \in P$ the set $p\mathbf{G} = \{pa; a \in \mathbf{G}\}$, which is in one to one correspondence with $\mathbf{G}$, is equal to the fiber $\pi^{-1}(\pi(p))$ containing p;

(iii) P is locally trivial, that is, for every point in M there is a neighborhood U and a C^∞ diffeomorphism $\varphi : \pi^{-1}(U) \to U \times \mathbf{G}$ such that

$$\varphi(pg) = \varphi(p)g, \quad p \in \pi^{-1}(U),\ g \in \mathbf{G},$$

if we define $(x, a)g = (x, ag)$ when $x \in M$ and $a, g \in \mathbf{G}$.

For the principal bundle constructed above using the transition functions g_{ij} we have obvious trivializations of $\pi^{-1}(U_i)$ which give back the transition functions. Now we can pass directly from the principal bundle to any one of the bundles F_ϱ; they are said to be *associated* to the principal bundle. To do so we note that $\mathbf{G}$ acts to the right (see Definition 5.4.1, (i)) on $P \times F$ by

$$(P \times F) \times \mathbf{G} \ni ((p, w), a) \mapsto (pa, \varrho(a^{-1})w) = \psi_a(p, w).$$

In fact,

$$\psi_{ab}(p, w) = (p(ab), \varrho((ab)^{-1}w) = ((pa)b, \varrho(b^{-1})(\varrho(a^{-1}w)) = \psi_b(\psi_a(p, w)),$$

which is just the definition of action to the right. We define two elements in $P \times F$ to be equivalent if they are carried into each other by some element in $\mathbf{G}$, thus $(pa, w) \sim (p, \varrho(a)w)$, and form the quotient $(P \times F)/\mathbf{G}$ by this equivalence relation. The map

$$U_i \times \mathbf{G} \times F \ni (x, g, w) \mapsto ((x, i, g)^{\sim}, w) \in (P \times F)_{U_i}$$

is a bijection, and each equivalence class under the right action of $\mathbf{G}$ contains precisely one element with $g = e$; (x, i, g, w) is equivalent to $(x, i, e, \varrho(g)w)$ under the $\mathbf{G}$ action. If $x \in U_j$ also, then (x, i, e, w) and (x, j, g_{ji}, w) define the same element in $P \times F$, and under the $\mathbf{G}$ action the latter is equivalent to $(x, j, e, \varrho(g_{ji})w)$. Thus our trivializations of $(P \times F)/\mathbf{G}$ over U_j and U_i are related by the transition function $\varrho(g_{ji})$, which means that $(P \times F)/\mathbf{G}$ is isomorphic to F_ϱ. We have now proved:

Proposition 5.4.2. *If P is a principal $\mathbf{G}$ bundle over M, and ϱ is a representation of $\mathbf{G}$ on a vector space F, then the quotient of $P \times F$ by the right $\mathbf{G}$ action defined by*

$$(P \times F) \times \mathbf{G} \ni ((p, w), g) \mapsto (pg, \varrho(g^{-1})w)$$

is a vector bundle with fiber dimension $\dim F$, said to be associated with P and ϱ, denoted by $P \times_\varrho F$.

Let us now as an example consider the case of a real vector bundle E of fiber dimension N over M. Let P be the frame bundle over M with fiber P_x consisting of all bases $(t_1, \dots, t_N)$ for E_x. The full linear group $\mathbf{GL}(N, \mathbf{R})$ acts to the right on P, mapping $(t_1, \dots, t_N) \in P_x$ and $A = (a_{jk}) \in \mathbf{GL}(N, \mathbf{R})$ to $(t'_1, \dots, t'_N)$, where $t'_j = \sum a_{kj} t_k$, and it is clear that P is a principal $\mathbf{GL}(N, \mathbf{R})$ bundle. The vector bundle associated to P and the natural action of $\mathbf{GL}(N, \mathbf{R})$ in $\mathbf{R}^N$ is isomorphic to E. In fact, the map

$$P_x \times \mathbf{R}^N \ni (t_1, \dots, t_N, x^1, \dots, x^N) \mapsto \sum_1^N x^j t_j \in E_x$$

is constant on the $\mathbf{G}$ orbits, for

$$\sum_{j=1}^N x^j t'_j = \sum_{j,k=1}^N x^j a_{kj} t_k = \sum_{k=1}^N (Ax)^k t_k.$$

Hence it defines an isomorphism $(P \times \mathbf{R}^N)/\mathbf{G} \to E$. If we instead take $\varrho(A) = {}^tA^{-1}$, then the associated bundle is the dual bundle of E.

After introducing a Euclidean structure in E, we can instead consider the orthonormal frame bundle P which is a principal $\mathbf{O}(N)$ bundle. The bundle associated to the natural action of $\mathbf{O}(N)$ on $\mathbf{R}^N$ is again E.

We shall now carry the notion of connection over from a real vector bundle E to the corresponding frame bundle P. This will lead to a natural definition of a connection in a principal bundle which induces a connection in the sense already defined on every associated vector bundle. Let ∇ be a connection in E, choose a basis $e_1, \dots, e_N$ for the sections of E on a coordinate patch $U \subset M$, and write

$$\nabla e_i = \sum_{j=1}^N \omega_{ji} e_j, \quad i = 1, \dots, N,$$

as in (5.1.8). For a general section $u = \sum_1^N u_i e_i$ we have

$$\nabla u = \sum_{i=1}^N du_i e_i + \sum_{i,j=1}^N u_i \omega_{ji} e_j = \sum_{i=1}^N (du_i + \sum_{j=1}^N u_j \omega_{ij}) e_i.$$

The equation $\nabla u = 0$ can therefore be written

$$du_i + \sum_{j=1}^N \omega_{ij} u_j = 0. \tag{5.4.4}$$

We say that u is parallel along a curve $t \mapsto \gamma(t)$ if $(\nabla_{\gamma(t)} u)(\gamma(t)) \equiv 0$. This is a linear system of N differential equations for N unknowns, so the unique solvability of the Cauchy problem means that there is a unique parallel section of E along γ with prescribed value at a point on γ.

We can apply this to make a parallel transport of a frame $E_j = \sum_{i=1}^N a_{ij} e_i$, $j = 1, \dots, N$, along γ. By (5.4.4) the equations $\nabla E_i = 0$ can be written

$$da_{ij} + \sum_{k=1}^N \omega_{ik} a_{kj} = 0, \quad i, j = 1, \dots, N. \tag{5.4.5}$$

If $\sum \lambda_j E_j = 0$ at some point on γ then this is true everywhere, so if we prescribe linearly independent initial data we get a frame along γ.

Now we can regard the local coordinates $x^1, \dots, x^n$ in M together with the coordinates a_{ij} with $\det(a_{ij}) \neq 0$ as coordinates in the frame bundle P over U. The equations (5.4.5) have an invariant meaning independent of the choice of these coordinates. If they are satisfied for the restriction to a regular curve $\Gamma \subset P$, then Γ cannot be a tangent to a fiber, for if $dx = 0$ it follows from (5.4.5) that $da_{ij} = 0$, contradicting that the curve is regular. Thus the equations (5.4.5) define at every point $p \in P$ a plane in T_pP of dimension n which is mapped bijectively on $T_{\pi(p)}M$ by the differential of the base projection $\pi : P \to M$. The right action of an element $g = (g_{jk}) \in \mathbf{GL}(N, \mathbf{R})$ maps the frame $E_1, \dots, E_N$ to $\sum g_{kj} E_k = \sum_i (\sum_k a_{ik} g_{kj}) e_i$, $j = 1, \dots, N$, which just means that the coordinates $A = (a_{ik})$ are replaced by Ag. The equations (5.4.5) are obviously invariant under this right multiplication, so the "horizontal" planes defined by (5.4.5) have the properties in the following definition:

Definition 5.4.3. If P is a principal $\mathbf{G}$ bundle over the n dimensional manifold M, with projection $\pi : P \to M$, then a connection on P is a differentiable assignment $P \ni p \mapsto H_p \subset T_p P$, where H_p is an n dimensional subspace such that

(i) $\pi' : H_p \to T_{\pi(p)} M$ is bijective;
(ii) $H_{pa} = R'_a H_p$, $a \in \mathbf{G}$, where R_a denotes the right action on P of $a \in G$.

One calls H_p the *horizontal* space at p.

In the example of the frame bundle of a real vector bundle E with fiber dimension N, we saw above that the horizontal space is defined by the vanishing of N^2 differential forms (5.4.5), and N^2 is the dimension of the group $\mathbf{GL}(N, \mathbf{R})$. We shall now show that in the general case the horizontal space is defined by means of a natural differential form on P with values in the Lie algebra $\mathfrak{g}$ of $\mathbf{G}$. This is a consequence of the following two facts.

(a) If $p \in P$ then the tangent space $T_p P$ is by (i) in Definition 5.4.3 the direct sum of H_p and the kernel of $\pi' : H_p \to T_{\pi(p)} M$, that is, the tangent space $T_p^0 P$ at p of the fiber $\pi^{-1}(\pi(p))$ through p. Thus we have a well defined projection

$$v_p : T_p \to T_p^0 P, \quad v_p H_p = 0, \ (\mathrm{Id} - v_p) T_p^0 = 0. \tag{5.4.6}$$

(b) The right action $\mathbf{G} \ni a \mapsto pa$ is by Definition 5.4.1 for every $p \in P$ a diffeomorphism of $\mathbf{G}$ on the fiber of P through p. Hence the differential at the identity is a linear bijection

$$\gamma(p) : \mathfrak{g} \to T_p^0 P. \tag{5.4.7}$$

The composition $\omega_p = \gamma(p)^{-1} v_p$ of the projection (5.4.6) with the inverse $\gamma(p)^{-1}$ is thus a linear map from $T_p P$ to $\mathfrak{g}$, that is, a one form on P with values in $\mathfrak{g}$, which vanishes precisely in the horizontal spaces H_p. To define it we have used part (i) of Definition 5.4.3, but we must also express condition (ii) there in terms of ω. To do so we note that the condition means that

$$v_{pa} = R'_a v_p (R'_a)^{-1}, \quad \text{hence } \omega_{pa} = \gamma(pa)^{-1} R'_a v_p (R'_a)^{-1} = \gamma(pa)^{-1} R'_a \gamma(p) \omega_p (R'_a)^{-1}.$$

If $X, Y \in \mathfrak{g}$ and

$$R'_a \gamma(p) X = \gamma(pa) Y$$

then $t \mapsto p \exp(tX) a$ and $t \mapsto pa \exp(tY)$ have the same derivative for $t = 0$, hence $t \mapsto p \exp(tX)$ and $t \mapsto pa \exp(tY) a^{-1} = p \exp(\mathrm{Ad}\, a\, (tY))$ have the same derivatives for $t = 0$. By the injectivity of $\gamma(p)$ we conclude that $X = \mathrm{Ad}\, a\, Y$, that is, $Y = \mathrm{Ad}(a^{-1}) X$, so

$$\omega_{pa} R'_a = \mathrm{Ad}(a^{-1}) \omega_p.$$

The left-hand side is the pullback $R_a^* \omega$ in T_p. We have now proved the first half of the following:

Theorem 5.4.4. *For a connection in a principal* $\mathbf{G}$ *bundle* P*, the projection on the tangent space of the fiber along the horizontal space followed by the inverse of* (5.4.7) *is a one form on* P *with values in the Lie algebra* $\mathfrak{g}$ *of* $\mathbf{G}$ *such that*

$$R_a^*\omega = \mathrm{Ad}(a^{-1})\omega, \quad a \in \mathbf{G}, \tag{5.4.8}$$

$$\omega_p(\gamma(p)X) = X, \quad X \in \mathfrak{g}, \tag{5.4.9}$$

where R_a *is the right action of* a *on* P*. Conversely, if* ω *is a one form on* P *with values in* $\mathfrak{g}$ *satisfying* (5.4.8) *and* (5.4.9)*, then* $H_p = \{t \in T_pP; \langle t, \omega\rangle = 0\}$ *defines a connection in* P *in the sense of Definition 5.4.3. One calls* ω *the connection form.*

Proof. If ω satisfies (5.4.9) then H_p is transversal to the fiber and of codimension $\dim \mathfrak{g}$ in T_pP, which proves (i) in Definition 5.4.3. Condition (ii) follows from (5.4.8) by reversing the proof of (5.4.8).

The next goal is to show that a connection in a principal G bundle P gives rise to a unique connection in every associated vector bundle. This must be done so that in the motivating example of a real vector bundle above we get back the connection we started from. Let s be a section over an open set $U \subset M$ of the vector bundle F_ϱ associated to P, F and the representation ϱ of $\mathbf{G}$ on F, and let $t \in T_xM$, $x \in U$. We may assume that there is a section p of P in U, and this defines a map $w : U \to F$ such that s is the image $(p, w)^\sim$ of the section (p, w) of $P \times F$. (Recall that P_ϱ is the set of equivalence classes of $P \times F$ under the right $\mathbf{G}$ action $(p, w) \mapsto (pa, \varrho(a^{-1})w)$, $a \in \mathbf{G}$.) If $p'(x)t \in T_{p(x)}$ were horizontal, we would in the special case of the frame bundle of a vector bundle E have $p = (e_1, \dots, e_N)$ with $\nabla_t e_j(x) = 0$, and $w = (w_1, \dots, w_N)$, so we would have $\nabla_t s = \sum\langle t, dw_j\rangle e_j$. This leads us to require that

$$\nabla_t s(x) = (p, \langle t, dw\rangle)^\sim, \quad \text{if } p'(x)t \in H_{p(x)}. \tag{5.4.10}$$

($\langle t, dw\rangle$ is well defined since w is a function with values in a vector space.) If $p'(x)t$ is not horizontal, we replace p and w by $\tilde{p} = p\exp(-\varphi)$ and $\tilde{w} = \varrho(\exp\varphi)w$ where $\varphi : U \to \mathfrak{g}$, $\varphi(x) = 0$. If $\langle t, \varphi'(x)\rangle = X \in \mathfrak{g}$, then

$$\tilde{p}'(x)t = p'(x)t - \gamma(p)X, \quad \langle t, d\tilde{w}\rangle = \langle t, dw\rangle + \varrho'(X)w,$$

where $\varrho' = \varrho'_e$ is a linear map from $\mathfrak{g}$ to linear maps in F. Thus $\tilde{p}'(x)t \in H_{p(x)}$ if and only if $v_p p'(x)t = \gamma(p)X$, that is, $X = \langle p'(x)t, \omega_{p(x)}\rangle = \langle t, p^*\omega\rangle$, so (5.4.10) leads to

$$\nabla_t s(x) = \big(p(x), \langle t, dw(x)\rangle + \varrho'(\langle t, p^*\omega\rangle)w(x)\big)^\sim. \tag{5.4.11}$$

The preceding arguments show that (5.4.11) is the only definition satisfying (5.4.10), but it remains to show that it does not depend on $p(x)$. To do so we must prove that (5.4.11) does not change if p is replaced by $pa = R_a p$ and w is replaced by $\varrho(a^{-1})w$ with a constant $a \in \mathbf{G}$. If we do that then $\langle t, p^*\omega\rangle$ is replaced by

$$\langle t, p^* R_a^*\omega\rangle = \langle t, p^*\,\mathrm{Ad}(a^{-1})\omega\rangle,$$

by (5.4.8), and since $\varrho(a^{-1}ba) = \varrho(a^{-1})\varrho(b)\varrho(a)$ we have

$$\varrho'(\mathrm{Ad}(a^{-1})X) = \varrho(a^{-1})\varrho'(X)\varrho(a), \quad X \in \mathfrak{g}.$$

Thus the second component in (5.4.11) is replaced by

$$\langle t, d\varrho(a^{-1})w\rangle + \varrho(a^{-1})\varrho'(\langle t, p^*\omega\rangle)\varrho(a)\varrho(a^{-1})w$$

which is precisely $\varrho(a^{-1})$ times the value in (5.4.11). The first component is replaced by $p(x)a$, so the equivalence class under the $\mathbf{G}$ action is unchanged. Hence (5.4.11) gives a unique definition of $\nabla_t s(x)$, $t \in T_x M$, when s is a section of F_ϱ. It is obvious that $\nabla_t s(x)$ is linear in t and linear in s, and we have for $\varphi \in C^\infty$

$$\nabla_t(\psi s) = \psi \nabla_t s + \langle t, d\psi\rangle s,$$

so ∇ is a connection in F_ϱ.

To motivate the definition of the curvature form of a principal bundle with connection we shall write the preceding formulas explicitly in the case of the frame bundle P of a vector bundle, which is a principal $\mathbf{G} = \mathbf{GL}(N, \mathbf{R})$ bundle. As above we choose a basis $e_1, \dots, e_N$ for the sections of E in a coordinate patch $U \subset M$ and define forms ω_{ij} by (5.1.8). As coordinates in $P|_U$ we use the local coordinates $x^1, \dots, x^n$ in U and the components $A = (a_{ij}) \in \mathbf{G}$ of a general frame

$$E_j = \sum a_{ij} e_i, \quad j = 1, \dots, N.$$

The Lie algebra $\mathfrak{g}$ is identified with the space $\mathbf{M}(N)$ of $N \times N$ matrices, which is mapped bijectively to the tangent to the fiber of P at $E_1, \dots, E_N$ by

$$\mathbf{M}(N) \ni X \mapsto \frac{d}{dt}(Ae^{tX})|_{t=0} = AX.$$

The horizontal space is defined by the equations (5.4.5), hence

$$(A\omega)_{ij} = da_{ij} + \sum_{k=1}^{N} \omega_{ik} a_{kj}$$

if ω denotes the connection form in Definition 5.1.3, with values in $\mathbf{M}(N)$, for both sides vanish in the horizontal space and are equal in the fiber direction. If ω^e denotes the matrix (ω_{ij}), this means that the connection form ω satisfies

$$A\omega = dA + \omega^e A,$$

which should be compared to the solution of Exercise 5.1.1. Hence

$$dA \wedge \omega + Ad\omega = (d\omega^e)A - \omega^e \wedge dA,$$

or after insertion of $dA = A\omega - \omega^e A$

$$A(d\omega + \omega \wedge \omega) = (d\omega^e + \omega^e \wedge \omega^e)A. \tag{5.4.12}$$

The matrix of two forms $d\omega^e + \omega^e \wedge \omega^e$ in the right-hand side of (5.4.12) is the matrix Ω in (5.1.10).

It is now clear that we should define the curvature form Ω for a principal $\mathbf{G}$ bundle with $\mathfrak{g}$ valued connection form ω by

$$\Omega = d\omega + \tfrac{1}{2}[\omega \wedge \omega]. \tag{5.4.13}$$

Here the right-hand side is defined as explained in Section 5.3, using the bracket in $\mathfrak{g}$. When $\mathbf{G} = \mathbf{GL}(N, \mathbf{R})$, $\mathfrak{g} = \mathbf{M}(N)$, this is equal to $\omega \wedge \omega$. We shall now prove a general analogue of (5.4.12) where transformation by the matrix A is replaced by the adjoint action.

Theorem 5.4.5. *Let P be a principal $\mathbf{G}$ bundle with connection form ω. The curvature form Ω defined by* (5.4.13) *is then for every $p \in P$ an antisymmetric bilinear form in T_pP/T_p^0P with values in $\mathfrak{g}$, and*

$$R_a^*\Omega = \mathrm{Ad}(a^{-1})\Omega, \quad a \in \mathbf{G}, \tag{5.4.14}$$

$$\Omega(t_1, t_2) = d\omega(t_1, t_2), \quad \textit{if } t_1, t_2 \in H_p. \tag{5.4.15}$$

Proof. (5.4.15) is obvious since

$$[\omega \wedge \omega](t_1, t_2) = [\omega(t_1), \omega(t_2)] = 0 \quad \text{when } t_1 \in H_p.$$

For $a \in \mathbf{G}$ we have in view of (5.4.8) and (5.3.11)

$$R_a^*\Omega = dR_a^*\omega + \tfrac{1}{2}[R_a^*\omega, R_a^*\omega] = \mathrm{Ad}(a^{-1})\Omega$$

by (5.4.8). What remains is to prove that $\Omega(t_1, t_2) = 0$ if $t_1 \in T_pP$ and $t_2 \in T_p^0P$. If also $t_1 \in T_p^0P$ then this follows from (5.3.1)$'''$, for if we identify a fiber of P with $\mathbf{G}$, then it follows from (5.4.9) that the restriction of the connection form ω to the fiber is the form ω in Section 5.3. Hence it suffices to prove that

$$\Omega(t_1, t_2) = 0, \quad \text{if } t \in H_pP,\ t_2 \in T_p^0P.$$

Since $\omega(t_1) = 0$ we have $[\omega \wedge \omega](t_1, t_2) = 0$, so the claim is that $d\omega(t_1, t_2) = 0$. To prove this we choose a horizontal vector field h with $h(p) = t_1$, write $t_2 = \gamma(p)X$, and denote by $\widetilde{X}$ the vector field equal to $\gamma(q)X$ at $q \in P$, corresponding to the infinitesimal right action of $\mathbf{G}$ in the direction X. By (C.5) we must show that

$$-\langle [h, \widetilde{X}], \omega\rangle + h\langle \widetilde{X}, \omega\rangle - \widetilde{X}\langle h, \omega\rangle = 0.$$

Now $\langle h, \omega \rangle = 0$ since h is horizontal, and $\langle \widetilde{X}, \omega \rangle = X$ is constant by (5.4.9), so the last two terms vanish. The proof will be complete if we prove that $[h, \widetilde{X}]$ is horizontal. Now

$$[h, \widetilde{X}] = \lim_{\varepsilon \to 0} \big((R_{\exp(\varepsilon X)})_* h - h\big)/\varepsilon$$

as is immediately seen in a coordinate system where the vector field $\tilde{X}$ is constant. Here

$$(R_a)_* h(p) = R'_a h(pa^{-1})$$

is horizontal for every $a \in \mathbf{G}$, by condition (ii) in Definition 5.4.3, which proves that $[h, \widetilde{X}]$ is also horizontal.

Theorem 5.4.5 means that at every $p \in P$ we can regard the curvature form Ω as the pullback of a two form with values in $\mathfrak{g}$ at the base point $\pi(p)$, and this form transforms by $\mathrm{Ad}(a^{-1})$ under right translation by a. Thus we have obtained an extension of (5.4.12) with considerable precision added when we can reduce the group $\mathbf{G}$ to a subgroup.

Exercise 5.4.1. Show using (5.4.14) that if $\varrho : \mathbf{G} \to F$ is a representation and $\varrho' = \varrho'_e$, then $\varrho'(p^*\Omega)$ is a two form which at $x \in M$ only depends on $p(x)$, and that

$$\varrho'((pa)^*\Omega) = \varrho(a^{-1})\varrho'(p^*\Omega)\varrho(a),$$

which means that $\varrho'(p^*\Omega)$ defines a two form on M with values in $\mathrm{End}(P \times_\varrho F)$.

For a Riemannian manifold of dimension n the tangent bundle is associated to the principal $\mathbf{O}(n)$ bundle of orthonormal frames. So are all tensor bundles and more generally the subbundles associated to $\mathbf{O}(n)$ invariant subspaces of tensor products of a number of factors $\mathbf{R}^n$. All the covariant differentiations discussed in Section 3.1.5, including that in Exercise 3.1.5, are therefore induced by the same Riemannian connection in the orthonormal frame bundle, and its curvature form is essentially equivalent to the Riemann curvature tensor by (4.2.5).

CHAPTER VI

LINEAR DIFFERENTIAL OPERATORS IN RIEMANNIAN GEOMETRY

Summary. In Section 6.1 we discuss second order elliptic operators on sections of a vector bundle such that the principal part is a multiple of the identity. Such operators occur frequently in geometry, and their analysis is particularly elementary thanks to a classical construction of a parametrix due to Hadamard. In the later sections we shall apply the conclusions to the de Rham complex and to Dirac operators. Section 6.2 is a digression where we discuss from a purely algebraic point of view topics such as the $*$ operator on differential forms. Section 6.3 is then devoted to Hodge theory, culminating in a discussion of the Hirzebruch signature operator, Weitzenböck decomposition and Bochner's vanishing theorem. The corresponding heat equations are studied in Section 6.4. Gilkey's theorem on invariant forms with non-negative weights is proved in Section 6.5; the required background in invariant theory is developed in Appendix D. This theorem implies that there is a local index formula for the Hirzebruch signature operator. Rather than determining coefficients we then pass to a discussion of Dirac operators, started in Section 6.6 and leading in Section 6.10 to Getzler's completely constructive proof of the local index theorem for such operators. This requires a great deal of background material concerning Clifford algebras and spinors given in Sections 6.7 and 6.8. We also need the classical Mehler formula for Hermite polynomials and some extensions presented in Section 6.9.

6.1. Metric elliptic operators. Already in (3.4.4) and Proposition 3.4.6 we defined the Laplace operator acting on a scalar function on a Riemannian manifold. It is a second order operator with principal symbol $\xi \mapsto \sum g^{jk}\xi_j\xi_k$. (See (5.1.3) for the definition of the principal symbol.) However, in geometrical questions the most common objects are sections of vector bundles, such as tensor bundles, rather than scalar functions. We shall therefore consider a more general class of operators here.

Definition 6.1.1. If E is a C^∞ vector bundle on a C^∞ manifold M, then a second order differential operator $P : C^\infty(M,E) \to C^\infty(M,E)$ is said to be *metric* if the principal symbol $p(x,\xi) : E_x \to E_x$ is a positive multiple $p_0(x,\xi)I_{E_x}$ of the identity in E_x for every $\xi \in T_xM$.

The positive definite quadratic form p_0 defines a symmetric tensor of type $2,0$, and its dual tensor, of type $0,2$, is a Riemannian metric. With local coordinates $x^1,\dots,x^n$ in M and corresponding coordinates $x^1,\dots,x^n,\xi_1,\dots,\xi_n$ in T^*M, we have $p_0(x,\xi) = \sum g^{jk}(x)\xi_j\xi_k$, and the Riemannian metric is $ds^2 = \sum g_{jk}(x)dx^jdx^k$, where $(g_{jk}) = (g^{jk})^{-1}$. Thus a metric operator is associated with a unique Riemannian metric, hence the term "metric operator" (which is used by some authors but not universally).

An example of a metric operator is obtained if we choose a connection ∇^E in E and define

$$Pu = \sum_{i,j=1}^{n} g^{ij} (\widetilde{\nabla}^E \nabla^E u)_{ij}, \tag{6.1.1}$$

where $\widetilde{\nabla}^E$ is the connection in $E \otimes T^*M$ defined by ∇^E and the Levi-Civita connection ∇ (covariant differentiation) in T^*M according to Proposition 5.1.8. If $\varphi \in C^\infty(M)$ is a scalar function, then with $(\cdot)_s$ denoting symmetrization in $T^*M \times T^*M$

$$\widetilde{\nabla}^E \nabla^E(\varphi u) = \widetilde{\nabla}^E(\varphi \nabla^E u + (d\varphi)u) = \varphi \widetilde{\nabla}^E \nabla^E u + 2((d\varphi)\nabla^E u)_s + (\nabla d\varphi)u,$$

hence

$$P(\varphi u) = \varphi P u + 2\langle (d\varphi)^\sharp, \nabla^E u\rangle + (\Delta\varphi)u. \tag{6.1.2}$$

Now suppose that P is any metric operator in E, introduce the associated Riemannian metric, and define for $t \in T_x M$

$$\langle t, \nabla^E u\rangle = \tfrac{1}{2}(P(\varphi u) - \varphi P u - (\Delta\varphi)u), \tag{6.1.3}$$

if $\varphi \in C^\infty(M)$ and $\varphi'(x) = t^\flat$. We claim that the right-hand side only depends on t and that (6.1.3) defines a connection. With local coordinates $x^1, \dots, x^n$ in M and a local basis for E, we can write

$$Pu = \sum_{i,j=1}^{n} g^{ij}\partial_i\partial_j u + \sum_{j=1}^{n} c^j \partial_j u + cu,$$

where $u = (u_1, \dots, u_N)$ and $c^1, \dots, c^n, c$ are $N \times N$ matrices. Then the right-hand side of (6.1.3) can be written

$$\sum_{i,j=1}^{n} g^{ij}\partial_i\varphi\partial_j u + \tfrac{1}{2}\sum_{i=1}^{n}(c^i - g^{-\frac{1}{2}}\sum_{j=1}^{n}\partial_j(g^{\frac{1}{2}}g^{ij}))\partial_i\varphi\, u$$

which is a linear function of φ', hence of $t = \varphi'^\sharp$, with principal part equal to the identity. If we write

$$P = \sum g^{ij}(\widetilde{\nabla}^E \nabla^E u)_{ij} + Q,$$

it follows from (6.1.2) and the way we have defined ∇^E that $Q(\varphi u) = \varphi Q u$, $\varphi \in C^\infty$, so Q is a differential operator of order 0. We have proved:

Proposition 6.1.2. *If P is a metric differential operator in the vector bundle E on M, then there is a Riemannian metric in M, a connection ∇^E in E, and a section c of* $\operatorname{Hom} E = E \otimes E^*$*, such that*

$$Pu = \sum g^{ij}(\widetilde{\nabla}^E \nabla^E u)_{ij} + cu, \quad u \in C^\infty(M, E), \tag{6.1.4}$$

where $\widetilde{\nabla}^E$ is the connection in $E \otimes T^*M$ defined by ∇^E and the Levi-Civita connection. The Riemannian metric, ∇^E and c are uniquely determined by P.

Note that no first order terms are visible in (6.1.4); they have been absorbed in the connection ∇^E. We shall see many explicit formulas of the type (6.1.4) where c is related to the Riemannian curvature tensor.

Using a classical approach due to Hadamard we shall now construct a parametrix for any metric operator, that is, a distribution section F_a of E such that $PF - a\delta_x$ for given $x \in M$ and $a \in E_x$ is in C^∞ or at least highly differentiable. Here the Dirac function is defined in terms of coordinates such that $g = \det(g_{jk}) = 1$ at x. For general local coordinate $x^1, \dots, x^n$ it must be defined as $\delta(x^1, \dots, x^n)/\sqrt{g}$; division by the Riemannian density changes the distribution density $\delta(x^1, \dots, x^n)$ to a distribution independent of the coordinates. In our construction we shall use geodesic coordinates centered at x, so $g_{jk}(0) = \delta_{jk}$ which makes $g(0) = 1$. To simplify the construction one should also use a frame in E which is adapated to the geodesic coordinates.

With local coordinates $x^1, \dots, x^n$ in M and a local frame for E we write $u = (u_1, \dots, u_N)$ and

$$Pu = \sum_{j,k=1}^{n} g^{-\frac12}\partial_j(g^{\frac12} g^{jk}\partial_k u) + \sum_{j=1}^{n} c^j \partial_j u + cu, \tag{6.1.5}$$

where u is differentiated componentwise as a function in $\mathbf{R}^n$ with values in $\mathbf{R}^N$. Then (6.1.3) means that

$$\langle t, \nabla^E u\rangle = \tfrac12(P(\varphi u) - \varphi Pu - (\Delta\varphi)u) = (d\varphi, du) + \tfrac12 \sum_{j=1}^{n} c^j(\partial_j\varphi)u,$$

hence

$$\nabla^E u = du + \tfrac12 \sum_{j,k=1}^{n} c^j g_{jk} dx^k u.$$

If $\varrho = \sum_1^n x^j\partial_j$ is the radial vector field, we obtain

$$\nabla^E_\varrho u = \varrho u + \tfrac12 \sum_{k=1}^{n} c_k x^k u, \quad c_k = \sum_{j=1}^{n} c^j g_{jk}. \tag{6.1.6}$$

Lemma 6.1.3. *In a neighborhood of the origin in the local coordinates one can choose the local frame* $e_1, \dots, e_N$ *for* E *so that* $\nabla^E_\varrho e_\nu = 0$, $\nu = 1, \dots, N$, *where* ϱ *is the radial vector field. Such a frame is called synchronous; it is uniquely determined by* $e_1(0), \dots, e_N(0)$.

Proof. By (6.1.6) we want all e_ν to be solutions of the system of differential equations

$$\sum_{j=1}^n x^j \partial_j u + \tfrac12 \sum_{k=1}^n c_k x^k u = 0.$$

We shall prove that there is a unique solution with given value at the origin. At the origin the equations imply that $\partial_j u(0) = \frac12 c_j(0) u(0)$, so we set

$$v = \exp(\tfrac12 \sum_{j=1}^n c_j(0) x^j) u.$$

Then

$$\sum x^j \partial_j v = \exp(\tfrac12 \sum_{j=1}^n c_j(0) x^j)\big(\sum_{j=1}^n x^j \partial_j u + \tfrac12 \sum_{j=1}^n c_j(0) x^j u + O(|x|^2) u\big),$$

so the equation becomes

$$\sum_{j=1}^n x^j \partial_j v + Rv = 0,$$

where R vanishes of second order at the origin. But this equation is of the form discussed in the proof of Lemma 3.3.4, which completes the proof.

Remark. We have assumed aboved that the vector bundle is real, but there is no difference in statements or proofs if it is complex.

In the following construction we shall use geodesic coordinates and a frame for E satisfying the conclusions in Lemma 6.1.3, that is,

$$\sum_{j,k=1}^n g_{jk}(x) c^j(x) x^k = \sum_{j,k=1}^n g_{jk}(0) c^j(x) x^k = \sum_{k=1}^n c_k(x) x^k = 0. \tag{6.1.7}$$

The operator is then so well approximated by the constant coefficient Laplacian Δ that a fundamental solution of Δ is a good first approximation for a parametrix of P. To get higher order approximations we shall also have to use fundamental solutions of powers of Δ. To avoid some minor technical difficulties and at the same time prepare for a study of heat equations later on, we shall consider $\Delta + z$ and $P + z$ instead of Δ and P for suitable $z \in \mathbf{C}$.

Lemma 6.1.4. *If $z \in \mathbf{C}\backslash\overline{\mathbf{R}}_+$ then there is for every integer $\nu \ge 0$ a unique distribution $F_\nu \in \mathcal{S}'(\mathbf{R}^n)$ such that*

$$(\Delta + z)^{\nu+1} F_\nu = \delta_0. \tag{6.1.8}$$

$x^\beta \partial^\alpha F_\nu$ is a bounded continuous function if $|\alpha| < |\beta| + 2(\nu+1) - n$, and

$$(\Delta + z) F_\nu = F_{\nu-1}, \quad \nu > 0, \tag{6.1.9}$$

$$2\nu \partial F_\nu / \partial x = x F_{\nu-1}, \quad \nu > 0. \tag{6.1.10}$$

Proof. The equation (6.1.8) means that $(z - |\xi|^2)^{\nu+1} \widehat{F}_\nu = 1$ if $\widehat{F}_\nu$ is the Fourier transform of F_ν. Hence

$$F_\nu(x) = (2\pi)^{-n} \int e^{i\langle x,\xi\rangle} (z - |\xi|^2)^{-\nu-1}\, d\xi$$

with the integral taken in the sense of distribution theory. Thus $x^\beta \partial^\alpha F_\nu$ is the inverse Fourier transform of $(i\partial_\xi)^\beta (i\xi)^\alpha (z - |\xi|^2)^{-\nu-1}$ which is integrable if $|\alpha| - |\beta| - 2(\nu + 1) < -n$, so $x^\beta \partial^\alpha F_\nu$ is then a bounded continuous function. (6.1.9) is an obvious consequence of the fact that F_ν is the unique solution of (6.1.8). The Fourier transform of $\partial F_\nu / \partial x_j$ is

$$i\xi_j (z - |\xi|^2)^{-\nu-1} = i\partial_{\xi_j} (z - |\xi|^2)^{-\nu} / 2\nu,$$

which proves (6.1.10) and ends the proof of the lemma.

The uniqueness implies that F_ν is a function of $|x| = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}}$, and it would be easy to express F_ν in terms of Bessel functions. We shall not do so but note that $F_\nu(x) = O(|x|^{2(\nu+1)-n} \log |x|)$ as $x \to 0$, where the logarithm can be dropped if n is odd or $2(\nu + 1) < n$. This follows from the corresponding well known facts for the Laplacian. In particular, $F_\nu \in L^1$.

If $f \in \mathcal{D}'(\mathbf{R})$, we have for $x \ne 0$

$$\begin{aligned}\sum_{k=1}^n g^{jk}(x) \partial_k f(|x|^2) = 2f'(|x|^2) \sum_{k=1}^n g^{jk}(x) x^k \\ = 2f'(|x|^2) x^j = \partial_j f(|x|^2), \quad j = 1, \ldots, n,\end{aligned}$$

for $\sum_{k=1}^n g_{jk}(x) x^k = x^j$ since the coordinates are geodesic. When $f(|x|^2)$ is replaced by F_ν the first and last expressions are equal as distributions in $\mathbf{R}^n$ since both are locally integrable functions. Hence it follows from (6.1.9), (6.1.10) that

$$\begin{aligned}&\sum_{k=1}^n g^{jk}(x) \partial_k F_\nu = (2\nu)^{-1} x^j F_{\nu-1}, \quad \text{if } \nu > 0;\\ &\sum_{j,k=1}^n \partial_j (g^{jk}(x) \partial_k F_\nu) + z F_\nu = F_{\nu-1}, \quad \text{if } \nu \ge 0,\end{aligned}$$

where $F_{-1} = \delta_0$. If $u_\nu \in C^\infty$ with values in $\mathbf{C}^N$, it follows now from (6.1.5) that

$$\begin{aligned}(P+z)(u_\nu F_\nu) &= (Pu_\nu)F_\nu + (u_\nu + \nu^{-1}(hu_\nu + \sum x^j\partial_j u_\nu))F_{\nu-1}, \quad \nu > 0;\\ (P+z)(u_0F_0) &= (Pu_0)F_0 + u_0(0)\delta_0 + 2(hu_0 + \sum x^j\partial_j u_0)f'(|x|^2),\end{aligned} \tag{6.1.11}$$

where $F_0 = f(|x|^2)$. Here we have used (6.1.7) and introduced the notation

$$h = \tfrac{1}{2}\sum x^j g^{-\frac{1}{2}}\partial_j g^{\frac{1}{2}}. \tag{6.1.12}$$

If we sum the equations (6.1.11) for $\nu = 0, \dots, \mu$, we obtain

$$(P+z)\sum_{\nu=0}^{\mu} u_\nu F_\nu = u_0(0)\delta_0 + (Pu_\mu)F_\mu, \tag{6.1.13}$$

if the coefficients u_ν satisfy the differential equations

$$\nu u_\nu + hu_\nu + \sum_{j=1}^{n} x^j\partial_j u_\nu + \nu P u_{\nu-1} = 0, \quad \nu = 0, \dots, \mu; \tag{6.1.14}$$

here the undefined term Pu_{-1} should be omitted when $\nu = 0$. These equations are chosen so that the terms involving $F_0, \dots, F_{\mu-1}$ drop out. Note that the error term $(Pu_\mu)F_\mu$ in the right-hand side of (6.1.13) is as smooth as desired if μ is chosen large. The first equation (6.1.14) can be integrated explicitly, for

$$h = \tfrac{1}{2}\sum_{j=1}^{n} x^j g^{-\frac{1}{2}}\partial_j g^{\frac{1}{2}} = g^{-\frac{1}{4}}\sum_{j=1}^{n} x^j\partial_j(g^{\frac{1}{4}}). \tag{6.1.12$'$}$$

This means that

$$u_0(x) = u_0(0)g(x)^{-\frac{1}{4}}, \tag{6.1.15}$$

and that the other equations (6.1.13) can be written in the form

$$(\nu + \sum_{j=1}^{n} x^j\partial_j)(g^{\frac{1}{4}}u_\nu) + \nu g^{\frac{1}{4}}Pu_{\nu-1} = 0, \quad \nu = 1, \dots, \mu. \tag{6.1.14$'$}$$

A differential equation of the form

$$(\nu + \sum_{1}^{n} x^j\partial_j)v = f \tag{6.1.16}$$

in a starshaped neighborhood of the origin in $\mathbf{R}^n$ has a unique smooth solution if ν is a positive integer and f is smooth. In fact, with polar coordinates it can be written

$$(\nu + r\partial/\partial r)v = f, \quad \text{thus } \partial(r^\nu v)/\partial r = r^{\nu-1}f,$$

which gives the unique smooth solution

$$v(x) = \int_0^1 t^{\nu-1} f(tx)\, dt. \tag{6.1.17}$$

Thus the equations (6.1.14) have a unique solution when $u_0(0)$ is prescribed.

The functions u_ν obtained depend linearly on $u_0(0) = e(0)$, so we can write them in the form $\tilde{u}_\nu(x)e(0)$, where $\tilde{u}(x)$ is a $N \times N$ matrix. The corresponding linear transformation from E_0 to E_x, both of which have been identified with $\mathbf{C}^N$, is independent of the choice of basis made. In fact, by the uniqueness in Lemma 6.1.3 the basis is uniquely determined up to a linear transformation independent of x. From now on we drop tilde from the notation, so $u_\nu(x) \in \mathcal{L}(E_0, E_x) \cong E_x \otimes E_0^*$ and $u_0(0)$ is the identity in E_0. Note that u_ν is independent of z but the distributions F_ν depend on z.

For every $y \in M$ the exponential map $T_yM \ni X \mapsto \exp_y(X)$ at y gives geodesic coordinates centered at y when we introduce an orthonormal basis in T_yM. Using Lemma 6.1.3 we then construct a synchronous local frame for E, so that P can be written in the form (6.1.5) with (6.1.7) fulfilled, in a neighborhood of the origin. The functions u_ν obtained depend of course in a C^∞ fashion on the parameters y. Now the distributions $F_\nu(X)$ pull back to distributions $F_\nu(x, y)$ defined in a neighborhood of the diagonal in $M \times M$, which only depend on the geodesic distance between x and y, and are locally integrable in y for fixed x, hence also in x for fixed y. We multiply the functions u_ν by a C^∞ cutoff function which is 1 near the center of the geodesic coordinates and has support in a small neighborhood and pull them back to $M \times M$. Since $F_\nu(X) \in C^\infty$ when $X \in \mathbf{R}^n \setminus \{0\}$, this does not introduce any new singularities in (6.1.13). We have then obtained C^∞ functions $u_\nu \in C^\infty(M \times M, E \boxtimes E^*)$ where $E \boxtimes E^*$ is the vector bundle with fiber $E_x \otimes E_y^*$ at $(x, y) \in M \times M$, such that $F_\nu(x, y)$ is defined in a neighborhood of $\operatorname{supp} u_\nu$ and for every μ

$$(P_x + z) \sum_{\nu=0}^{\mu} u_\nu F_\nu - \delta_{\text{diag}} \in C^{2\mu+1-n}(M \times M, E \boxtimes E^*).$$

Here δ_{diag} is the kernel of the identity map in $C^\infty(M, E)$, that is, the distribution in $\mathcal{D}'(M \times M, E \boxtimes E^*)$ such that

$$\int \delta_{\text{diag}}(x, y)\varphi(y)\, dvol(y) = \varphi(x), \quad \varphi \in C_0^\infty(M, E),$$

which means that in terms of local coordinates $\delta_{\text{diag}}(x, y)$ is equal to $\delta(x - y)/\sqrt{g(y)}$ times the identity matrix. We can choose $F \in \mathcal{D}'(M \times M, E \boxtimes E^*)$ so that

$$F - \sum_{\nu=0}^{\mu-1} u_\nu F_\nu \in C^{2\mu+1-n}(M \times M, E \boxtimes E^*), \quad \forall \mu.$$

In fact, since $u_\nu F_\nu \in C^{2\nu+1-n}$ we can choose $\psi_\nu \in C^\infty(M \times M, E \boxtimes E^*)$ so that all seminorms in $C^{2\nu+1-n}$ of $u_\nu F_\nu - \psi_\nu$ over a compact set K_ν in $M \times M$ are $< 2^{-\nu}$. If

K_ν contains an arbitrary compact set in $M \times M$ when ν is large enough, this implies that the series $F = \sum_0^\infty (u_\nu F_\nu - \psi_\nu)$ converges in $\mathcal{D}'(M \times M, E \boxtimes E^*)$, and F has the required properties. Thus

$$(6.1.18) \qquad (P_x + z)F(x,y) = \delta + R(x,y), \quad R \in C^\infty(M \times M, E \boxtimes E^*).$$

If M is not compact it is useful to take the cutoffs in the definition so close to the diagonal that the maps

$$\operatorname{supp} F \ni (x,y) \mapsto x \in M \quad \text{and} \quad \operatorname{supp} F \ni (x,y) \mapsto y \in M$$

are proper. One calls F a *proper right parametrix* then.

If $h \in C^\infty(M,E)$ then

$$x \mapsto \int F(x,y)h(y)\, dvol(y) = (Fh)(x) \quad \text{is in } C^\infty(M,E).$$

Since F is in C^∞ outside the diagonal, it is sufficient to verify this in local coordinates. Introducing geodesic coordinates Y centered at x as integration variables we see that the terms $u_\nu(x,y)F_\nu(x,y)$ give a C^∞ contribution, for with these coordinates F_ν will only depend on Y and not on x, and u_ν and h are in C^∞ as functions of Y and x. This proves the assertion. In the same way we see that the operator F^* defined by the kernel $F(y,x)$ maps $C^\infty(M,E^*)$ to $C^\infty(M,E^*)$. We have $\langle F\varphi, \psi\rangle = \langle \varphi, F^*\psi\rangle$, $\varphi \in C_0^\infty(M,E)$, $\psi \in C_0^\infty(M,E^*)$, where $\langle \cdot, \cdot \rangle$ is the integral with respect to $dvol(x)$ of the scalar product between the fibers E_x and E_x^*.Hence both F and F^* have continuous extensions to $\mathcal{D}'$.

The adjoint operator $P^* : C^\infty(M,E^*) \to C^\infty(M,E^*)$ is defined similarly so that $\langle P\varphi, \psi\rangle = \langle \varphi, P^*\psi\rangle$. From the equation (6.1.18), which for the corresponding operators means that

$$(6.1.18)' \qquad (P+z)F\varphi = \varphi + R\varphi, \quad \varphi \in C_0^\infty(M,E),$$

it follows by taking adjoints that F^* is a left parametrix of $P^* + z$,

$$(6.1.18)'' \qquad F^*(P^*+z)\psi = \psi + R^*\psi, \quad \psi \in C_0^\infty(M,E^*).$$

The principal symbol of P^* is $p_0(x,\xi)$ times the identity in E^*, so if we apply $(6.1.18)''$ with P replaced by P^* we obtain a distribution $G \in C^\infty(M \times M, E \boxtimes E^*)$ with the same regularity properties as F which is a proper left parametrix of P. Then

$$F\varphi - G\varphi = (G(P+z) - S)F\varphi - G((P+z)F - R)\varphi = (GR - SF)\varphi, \quad \varphi \in C_0^\infty(M,E),$$

where R and S have C^∞ kernels. Hence $F - G$ is the kernel of $GR - SF$, which is in C^∞ since one of the factors in each term is. (The compositions are well defined since we take properly supported parametrices.) We may therefore conclude that F is also a left parametrix, so (6.1.18) can be strengthened to

$$(6.1.18)''' \quad (P+z)F\varphi = \varphi + R_1\varphi, \quad F(P+z)\varphi = \varphi + R_2\varphi, \qquad \varphi \in C_0^\infty(M,E),$$

where both R_1 and R_2 have kernels in $C^\infty(M \times M, E \boxtimes E^*)$. We have now forged all the tools required to discuss existence and uniqueness theorems for the operator P when M is a compact manifold. However, we first observe that the condition $z \notin \overline{\mathbf{R}}_+$ is easily removed here. In fact, for any z we can modify the definition of F_ν in Lemma 6.1.4 to the inverse Fourier transform of $(1-\chi(\xi))(z-|\xi|^2)^{-\nu-1}$ where $\chi \in C_0^\infty$ is equal to 1 in a neighborhood of the real zeros of $z-|\xi|^2$. Then δ_0 is replaced by $\delta_0 - R$ in (6.1.8), where $\widehat{R} = \chi$, hence $R \in \mathcal{S}$; (6.1.9) remains valid, and in (6.1.10) we just get an additional term in $\mathcal{S}$. The new smooth terms do not affect the arguments above, so (6.1.18)$'''$ holds for arbitrary z. In what follows we take $z = 0$, but we shall use general z later on to handle heat equations.

Theorem 6.1.5. *If M is a compact C^∞ manifold, E a C^∞ vector bundle over M, and P a smooth metric differential operator in $C^\infty(M,E)$, then*

(i) $\operatorname{Ker} P = \{u \in C^\infty(M,E); Pu = 0\}$ *is finite dimensional;*
(ii) $u \in \mathcal{D}'(M,E), Pu \in C^\infty(M,E)$ *implies* $u \in C^\infty(M,E)$*;*
(iii) *the equation $Pu = f \in \mathcal{D}'(M,E)$ has a solution $u \in \mathcal{D}'(M,E)$ if and only if $\langle f, v\rangle = 0$ for all $v \in \operatorname{Ker} P^*$, which is a finite dimensional subspace of $C^\infty(M,E^*)$.*

Proof. To prove (ii) we note that $Ff = FPu = u + R_2 u$ by (6.1.18)$'''$, and $Ff \in C^\infty$, $R_2 u \in C^\infty$ since $f \in C^\infty$ and $R_2 \in C^\infty$. Hence $u \in C^\infty$. It follows also that if $u \in \operatorname{Ker} P$, then $u + R_2 u = 0$, so Fredholm theory for the operator R_2 with C^∞ kernel proves that $\operatorname{Ker} P$ is finite dimensional. If $Pu = f$ then

$$\langle f, v\rangle = \langle Pu, v\rangle = \langle u, P^* v\rangle = 0, \quad \text{if } v \in \operatorname{Ker} P^*,$$

and $\operatorname{Ker} P^*$ is finite dimensional by (i) applied to P^*. To solve the equation $Pu = f$ with $f \in \mathcal{D}'(M,E)$ we set $u = v + Ff$ and get the equivalent equation $Pv = f - PFf = -R_1 f$. Since $R_1 f \in C^\infty$ it follows that it suffices to discuss solvability of $Pu = f$ when $f \in C^\infty$. With $u = Fw$ the equation becomes $w + R_1 w = f$, which by Fredholm theory can be solved for all f such that $\langle f, h_j\rangle = 0$, $j = 1, \dots, N$, where $h_1, \dots, h_N$ is a basis in $C^\infty(M, E^*)$ for solutions of $h + R_1^* h = 0$. Thus the equation $Pu = f$ can be solved if $\langle f, h_j\rangle = 0$, $j = 1, \dots, N$. Hence it can be solved if $(\langle f, h_1\rangle, \dots, \langle f, h_N\rangle) \in V$,

$$V = \{\langle Pu, h_j\rangle; u \in C^\infty(M,E)\} \subset \mathbf{C}^N.$$

If $(\lambda_1, \dots, \lambda_N)$ is a normal of the linear space V, then $\sum_1^N \lambda_j h_j \in \operatorname{Ker} P^*$, so $\sum \lambda_j \langle f, h_j\rangle = 0$ if $\langle f, \operatorname{Ker} P^*\rangle = 0$. Hence $(\langle f, h_1\rangle, \dots, \langle f, h_N\rangle) \in V$, which completes the proof.

It is easy to give other spaces than C^∞ and $\mathcal{D}'$, such as Sobolev spaces and Hölder spaces, for which (ii) and (iii) in Theorem 6.1.5 are true, but we shall postpone introducing these spaces until we need them.

Theorem 6.1.5 has an important corollary for first order operators:

Corollary 6.1.6. *Let M be a compact C^∞ manifold, E and F two C^∞ vector bundles over M, and*

$$P : C^\infty(M,E) \to C^\infty(M,F), \qquad Q : C^\infty(M,F) \to C^\infty(M,E)$$

first order differential operators such that QP is a metric operator. Then

(i) $\operatorname{Ker} P = \{u \in C^\infty(M,E); Pu = 0\}$ *is finite dimensional;*

(ii) $u \in \mathcal{D}'(M,E)$, $Pu \in C^\infty(M,F)$ *implies* $u \in C^\infty(M,E)$;

(iii) *the equation $Qv = g \in \mathcal{D}'(M,E)$ has a solution $v \in \mathcal{D}'(M,F)$ if and only if $\langle g, u\rangle = 0$ for all $u \in \operatorname{Ker} Q^*$, which is a finite dimensional subspace of $C^\infty(M,E^*)$.*

Proof. $Pu = 0$ implies $QPu = 0$, and $Pu \in C^\infty(M,F)$ implies $QPu \in C^\infty(M,E)$, so (i), (ii) follow from (i), (ii) in Theorem 6.1.5. Since $P^*Q^* = (QP)^*$ is a metric elliptic operator, the finite dimensionality of $\operatorname{Ker} Q^*$ follows from (i). The equation $Qv = g$ has a solution if the equation $QPu = g$ can be solved, which by Theorem 6.1.5 is possible if g is orthogonal to the finite dimensional space $\operatorname{Ker}(QP)$. As at the end of the proof of Theorem 6.1.5 it follows that the equation $Qv = g$ can be solved precisely when $\langle v, \operatorname{Ker} Q^*\rangle = 0$.

If p and q are the principal symbols of P and Q, then the hypothesis implies that

$$q(\xi)p(\xi) = r(\xi)I_{E_x}, \quad \text{where } r(\xi) > 0,\ 0 \neq \xi \in T_x^*M. \tag{6.1.19}$$

Conversely, if P is given and such a linear symbol q exists, then we can choose Q so that the hypotheses in the theorem are fulfilled. When E and F have the same fiber dimension, then (6.1.19) implies that $p(\xi)q(\xi) = r(\xi)I_{F_x}$, for $p(\xi)q(\xi)p(\xi) = r(\xi)p(\xi)$ and $p(\xi)$ is invertible. Hence the roles of P and Q may be interchanged, so the equation $Pu = f \in \mathcal{D}'(M,F)$ has a solution $u \in \mathcal{D}'(M,E)$ if and only if u is orthogonal to $\operatorname{Ker} P^*$, which is a finite dimensional subspace of $C^\infty(M,F^*)$; we have statements analogous to (i) and (ii) for Q.

6.2. The exterior algebra of a Euclidean vector space. As a preparation for the Hodge theory in Section 6.3 we shall discuss here some elementary *algebraic* aspects of exterior differential forms. Let V be a vector space of dimension $n < \infty$ over $\mathbf{R}$, and denote its dual space by V^*. Recall that the space $\wedge^p V^*$ of alternating p linear forms on V is spanned by the forms $\theta_1 \wedge \cdots \wedge \theta_p$, $\theta_j \in V^*$, defined by

$$\underbrace{V \times \cdots \times V}_{p \text{ times}} \ni (v_1, \dots, v_p) \mapsto \det\langle v_k, \theta_j\rangle_{j,k=1}^p. \tag{6.2.1}$$

In fact, if L is such a form and $e_1, \dots, e_n$ is a basis for V, $\theta_1, \dots, \theta_n$ the dual basis for V^*, then

$$\begin{aligned} L(v_1, \dots, v_p) &= L\big(\sum \langle v_1, \theta_{j_1}\rangle e_{j_1}, \dots, \sum \langle v_p, \theta_{j_p}\rangle e_{j_p}\big) \\ &= \sum \langle v_1, \theta_{j_1}\rangle \dots \langle v_p, \theta_{j_p}\rangle L(e_{j_1}, \dots, e_{j_p}) = \sum \det\langle v_k, \theta_{j_l}\rangle_{k,l=1}^p L(e_{j_1}, \dots, e_{j_p})/p!. \end{aligned}$$

Thus the forms (6.2.1) give a basis for $\wedge^p V^*$ even if $\theta_1, \dots, \theta_p$ are restricted to elements in a fixed basis for V^*.

The forms

$$\underbrace{V^* \times \cdots \times V^*}_{p \text{ times}} \ni (\theta_1, \dots, \theta_p) \mapsto \det\langle v_k, \theta_j\rangle_{j,k=1}^p, \tag{6.2.1$'$}$$

span $\wedge^p V$ for reasons of symmetry. We have used the duality between $\wedge^p V$ and $\wedge^p V^*$ such that

$$\langle v_1 \wedge \cdots \wedge v_p, \theta_1 \wedge \cdots \wedge \theta_p\rangle \mapsto \det\langle v_k, \theta_j\rangle_{j,k=1}^p. \tag{6.2.2}$$

(Many authors use another definition where the right-hand side is divided by $p!$, for that is the duality inherited from the natural duality of the tensor products. See Sternberg [1, p. 19] for a discussion of this point. Kobayashi and Nomizu [1] use the division by $p!$, which should be kept in mind when comparing identities.) The existence and uniqueness of a bilinear form on $\wedge^p V \times \wedge^p V^*$ such that (6.2.2) holds is obvious, for there is a unique bilinear form such that (6.2.2) holds when v_k are chosen in a fixed basis for V and θ_j in a fixed basis for V^*, and in view of the multilinearity and skew symmetry of the two sides they must then be equal for arbitrary $v_1, \dots, v_p, \theta_1, \dots, \theta_p$.

Assume now that V is a Euclidean vector space with scalar product denoted by $(\cdot, \cdot)$. The linear form $v^\flat : V \ni v' \mapsto (v, v')$ is an element $\neq 0$ in V^*, so we have a bijection $V \ni v \mapsto v^\flat \in V^*$ with inverse denoted by $\sharp$; these are the musical isomorphisms we used already in Chapter III. In particular, this gives a norm in V^*, which is of course the dual norm

$$\|\theta\| = \sup_{0 \neq v \in V} |\langle v, \theta\rangle| / \|v\|,$$

and the scalar product of $\theta_1, \theta_2 \in V^*$ is

$$(\theta_1, \theta_2) = \langle \theta_1^\sharp, \theta_2\rangle = (\theta_1^\sharp, \theta_2^\sharp)$$

where we have used the convention that $(\cdot, \cdot)$ denotes scalar product in a Euclidean space while $\langle \cdot, \cdot\rangle$ denotes the bilinear form in a space and its dual. The map $\theta \mapsto \theta^\sharp$ extends to a map $\wedge^p V^* \ni \varphi \mapsto \varphi^\sharp \in \wedge^p V$, and we define

$$(\varphi, \psi) = \langle \varphi^\sharp, \psi\rangle, \quad \varphi, \psi \in \wedge^p V^*. \tag{6.2.3}$$

If $\varepsilon_1, \dots, \varepsilon_n$ is an orthonormal basis in V^* then $e_j = \varepsilon_j^\sharp$ is an orthonormal basis for V which is biorthogonal to the basis in V^*, so we obtain

$$(\varphi, \psi) = \sum_{|I|=p} \varphi_I \psi_I / p!, \quad \text{if } \varphi = \sum_{|I|=p} \varphi_I \varepsilon_I / p!, \ \psi = \sum_{|I|=p} \psi_I \varepsilon_I / p!, \tag{6.2.4}$$

where φ_I and ψ_I are antisymmetric in the p indices of $I = (i_1, \dots, i_p)$, all ranging from 1 to n, and $\varepsilon_I = \varepsilon_{i_1} \wedge \cdots \wedge \varepsilon_{i_p}$. In fact, $\langle \varepsilon_I^\sharp, \varepsilon_J\rangle$ equals 0 unless the indices in

I are different and J is a permutation of them; in that case, we have the sign of the permutation. This means that for every I we get $p!$ equal non-zero terms $\varphi_I\psi_J\langle\varepsilon_I^\sharp, \varepsilon_J\rangle$ which proves (6.2.4). By (6.2.4) it is clear that (6.2.3) gives a Euclidean structure in $\wedge^p V^*$. We can suppress the factorials in (6.2.4) if we sum only for strictly increasing multiindices I, that is, $i_1 < \cdots < i_p$; when we do that we shall use the notation $\sum'$. Note that our definition (6.2.3) did not use a basis for V or V^*. If O is an orthogonal transformation in V and we define

$$(O^*\varphi)(v_1, \ldots, v_p) = \varphi(Ov_1, \ldots, Ov_p), \quad v_1, \ldots, v_p \in V,$$

it follows that we have orthogonal invariance:

$$(O^*\varphi, O^*\psi) = (\varphi, \psi), \quad \varphi, \psi \in \wedge^p V^*. \tag{6.2.5}$$

$\wedge^n V^*$ is a one dimensional Euclidean vector space so it contains precisely two elements ε with $\|\varepsilon\| = 1$. A choice of one of them means choosing an *orientation* of V^*; a basis $\varepsilon_1, \ldots, \varepsilon_n$ for V^* is defined to be positively oriented if $\varepsilon_1 \wedge \cdots \wedge \varepsilon_n$ is a positive multiple of ε. If we fix an orientation ε then

$$\varphi \wedge \psi = B(\varphi, \psi)\varepsilon, \quad \varphi \in \wedge^p V^*, \ \psi \in \wedge^{n-p} V^*$$

defines a non-degenerate bilinear form on $\wedge^p V^* \times \wedge^{n-p} V^*$, hence a duality between these spaces. Thus we obtain a linear isomorphism of $\wedge^p V^*$ on the dual of $\wedge^{n-p} V^*$, which we have identified with $\wedge^{n-p} V^*$ using (6.2.3), so we have an isomorphism $*$: $\wedge^p V^* \to \wedge^{n-p} V^*$ such that

$$\varphi \wedge \psi = \varepsilon(*\varphi, \psi), \quad \varphi \in \wedge^p V^*, \ \psi \in \wedge^{n-p} V^*. \tag{6.2.6}$$

If $\varepsilon_1, \ldots, \varepsilon_n$ is an orthonormal basis in V^* with $\varepsilon_1 \wedge \cdots \wedge \varepsilon_n = \varepsilon$, then

$$*(\varepsilon_{i_1} \wedge \cdots \wedge \varepsilon_{i_p}) = \sigma \varepsilon_{i_{p+1}} \wedge \cdots \wedge \varepsilon_{i_n}, \tag{6.2.7}$$

if $(i_1, \ldots, i_n)$ is a permutation with sign σ of $1, \ldots, n$.

Exercise 6.2.1. Show that $*(v \wedge w)$ is the usual vector product if v, w are vectors in $\mathbf{R}^3$.

It is obvious that the square of the $*$ operator must be an automorphism of $\wedge^p V^*$ for every p. Under the hypotheses of (6.2.7) we obtain

$$*(\varepsilon_{p+1} \wedge \cdots \wedge \varepsilon_n) = (-1)^{p(n-p)} \varepsilon_1 \wedge \cdots \wedge \varepsilon_p,$$

for the permutation $(p+1, \ldots, n, 1, \ldots, p)$ of $(1, \ldots, n)$ has $p(n-p)$ inversions. This means that

$$** \varphi = (-1)^{p(n-p)}, \quad \varphi \in \wedge^p. \tag{6.2.8}$$

This is independent of the choice of orientation, for changing the orientation means replacing $*$ by $-*$, which preserves any formula where there is an even number of factors $*$.

It is often convenient to write (6.2.6) in a different way. If we replace φ by $*\varphi$, we obtain

$$\varepsilon(\varphi, \psi) = (-1)^{p(n-p)}(*\varphi) \wedge \psi = \psi \wedge *\varphi,$$

hence

$$\psi \wedge *\varphi = \varepsilon(\varphi, \psi), \quad \varphi, \psi \in \wedge^{n-p} V^*. \tag{6.2.6$'$}$$

Using this formula it is easy to write down the $*$ operator explicitly for an arbitrary basis $\varepsilon_1, \dots, \varepsilon_n$ in V^*. Set $g^{jk} = (\varepsilon_j, \varepsilon_k)$, and let I, J be multiindices of length p, $n-p$, such that I, J is a positive permutation of $1, \dots, n$. If the basis $\varepsilon_1, \dots, \varepsilon_n$ is positively oriented, that is, $\varepsilon = \varepsilon_1 \wedge \cdots \wedge \varepsilon_n / \|\varepsilon_1 \wedge \cdots \wedge \varepsilon_n\|$, then taking $\psi = \varepsilon_I$ in (6.2.6)$'$ gives (6.2.9)

$$(*\varphi)_J = (\varphi, \varepsilon_I)/\|\varepsilon_1 \wedge \cdots \wedge \varepsilon_n\| = {\sum}'|K| = p\varphi_K \det\big(g^{k_\nu I_\mu}\big)^p_{\nu,\mu=1} / \big(\det(g^{\nu\mu})^n_{\nu,\mu=1}\big)^{\frac{1}{2}}.$$

If n is even, $n = 2l$, then $*$ maps $\wedge^l V^*$ into itself, and the square is $(-1)^{l^2}$. For a positively oriented orthonormal basis $\theta_1, \dots, \theta_{2l}$ in V^* we have

$$*(\theta_1 \wedge \cdots \wedge \theta_l + c\theta_{l+1} \wedge \cdots \wedge \theta_{2l}) = \theta_{l+1} \wedge \cdots \wedge \theta_{2l} + (-1)^{l^2} c\theta_1 \wedge \cdots \wedge \theta_l.$$

If l is even, then taking $c = \pm 1$ we see using such elements in $\wedge^l V^*$ that $\wedge^l V^* = \wedge_+ \oplus \wedge_-$ where $\wedge_+$ and $\wedge_-$ are subspaces of the same dimension $\frac{1}{2}\binom{2l}{l}$ and $*\lambda = \pm\lambda$ when $\lambda \in \wedge_\pm$. This is the general form of the decomposition we noticed in Lemma 2.3.3 when $n = 4$. If l is odd, we must take $c = \pm i$, so it is the complexification $\wedge^l V^*_{\mathbf{C}}$ which is the direct sum of the eigenspaces of $*$ with eigenvectors $\pm i$.

To generalize this to $\wedge^* V^* = \oplus_{p=0}^n \wedge^p V^*$, still with $n = 2l$, we define

$$\tau(\varphi) = i^{p(p-1)+l} * \varphi, \quad \varphi \in \wedge^p V^*_{\mathbf{C}}. \tag{6.2.10}$$

Since $p(p-1)$ is even this is real if l is even and purely imaginary if l is odd, so it is necessary to go to the complexification then. When $\varphi \in \wedge^p_{\mathbf{C}}$ we have

$$\tau^2\varphi = i^{p(p-1)+l+(2l-p)(2l-p-1)+l}(-1)^{p(2l-p)}\varphi.$$

The exponent of -1 here is congruent to $p^2 \mod 2$ and the exponent of i is $4l^2 - 4lp + 2p^2 \equiv 2p^2 \mod 4$, so $\tau^2\varphi = \varphi$. If $p \neq l$ it follows that $\wedge^p V^*_{\mathbf{C}} \oplus \wedge^{2l-p} V^*_{\mathbf{C}}$ is the direct sum of the eigenspaces of τ corresponding to the eigenvalues ± 1, and these must have the same dimension $\binom{2l}{p}$ since none intersects $\wedge^p V^*_{\mathbf{C}} \oplus \{0\}$. Summing up, we have proved:

Proposition 6.2.1. *If the oriented Euclidean vector space V is of even dimension $2l$, then the complexification $\wedge^* V^*_{\mathbf{C}}$ of $\wedge^* V^*$ is the direct sum of the eigenspaces $\wedge_\pm$ of the operator τ defined by* (6.2.10) *corresponding to the eigenvalues ± 1. $\wedge_+$ and $\wedge_-$ have the same dimension, and $\wedge_\pm$ is the direct sum of its intersections with $\wedge^p V^*_{\mathbf{C}} + \wedge^{2l-p} V^*_{\mathbf{C}}$ when $0 \le p \le l$.*

6.3. The Hodge decomposition. Let M be a C^∞ Riemannian manifold. Then the exterior powers $\wedge^p T_x^* M$ are the fibers of a vector bundle $\wedge^p T^* M$ with the obvious trivializations provided by the basis $dx^1, \dots, dx^n$ in T^*M over a local coordinate patch. We shall denote by λ^p the space of sections $C^\infty(M, \wedge^p T^* M)$, that is, the C^∞ p forms on M. If $\varphi, \psi \in \lambda^p$, then $(\varphi, \psi)(x)$ is a C^∞ function on M as defined in Section 6.2 using the Euclidean metric in $T_x M$. We set

$$(\varphi, \psi) = \int_M (\varphi, \psi)(x)\, dvol(x), \quad \varphi, \psi \in \lambda^p, \operatorname{supp}\varphi \cap \operatorname{supp}\psi \Subset M. \tag{6.3.1}$$

If M is oriented we can use $(6.2.6)'$ to write the integral in a more convenient way in terms of the $*$ operator. With local coordinates $x^1, \dots, x^n$ which are positively oriented we have

$$\varepsilon = dx^1 \wedge \dots \wedge dx^n / \|dx^1 \wedge \dots \wedge dx^n\| = (\det g_{jk})^{\frac{1}{2}} dx^1 \wedge \dots \wedge dx^n,$$

which means that

$$(\varphi, \psi) = \int_M \varphi \wedge *\psi, \quad \varphi, \psi \in \lambda^p, \ \operatorname{supp}\varphi \cap \operatorname{supp}\psi \Subset M. \tag{6.3.1$'$}$$

This formula makes it easy to compute the formal adjoint of the exterior differential operator $d : \lambda^p \to \lambda^{p+1}$. It is defined by

$$(d\varphi, \psi) = (\varphi, d^*\psi), \quad \varphi \in \lambda^p, \ \psi \in \lambda^{p+1}, \quad \operatorname{supp}\varphi \cap \operatorname{supp}\psi \Subset M.$$

By Stokes' formula we obtain

$$\begin{aligned}(d\varphi, \psi) = \int (d\varphi) \wedge *\psi = \int d(\varphi \wedge *\psi) - (-1)^p \int \varphi \wedge d * \psi \\ = (-1)^{p+1+p(n-p)} \int \varphi \wedge * * d * \psi.\end{aligned}$$

Since $p^2 - p = p(p-1)$ is an even number, it follows that

$$d^* = (-1)^{np+1} * d * \quad \text{on } \lambda^{p+1}. \tag{6.3.2}$$

Since the right-hand side contains two factors $*$ it does not change if the orientation is changed, so the formula can also be used for non-oriented manifolds if one keeps in mind that intermediate products such as $d*$ make no sense.

Assume now that M is a *compact* Riemannian manifold. Recall that the de Rham cohomology groups are defined by

$$H^p(M) = \{\psi \in \lambda^p; d\psi = 0\}/d\lambda^{p-1}.$$

Since λ^p is a pre-Hilbert space with the scalar product (6.3.1) it is natural to try to find in each class an element with minimal norm, for if it exists it must be unique. If $d\psi = 0$ and ψ minimizes the norm in its cohomology class, then

$$\|\psi\|^2 \le \|\psi + d\varphi\|^2, \quad \varphi \in \lambda^{p-1},$$

or equivalently,

$$(\psi, d\varphi) = 0, \quad \varphi \in \lambda^{p-1}.$$

Thus the minimum property is equivalent to $d^*\psi = 0$.

Theorem 6.3.1. *On a compact Riemannian manifold the equations* $d\psi = 0$*,* $d^*\psi = 0$ *for* $\psi \in \lambda^p(M)$ *are equivalent to* $\Delta\psi = 0$*, where* Δ *is the Hodge Laplacian*

$$\Delta = dd^* + d^*d : \lambda^p \to \lambda^p.$$

Proof. It is obvious that $\Delta\psi = 0$ if $d\psi = 0$ and $d^*\psi = 0$. On the other hand, if $\Delta\psi = 0$, then

$$0 = (\Delta\psi, \psi) = (dd^*\psi, \psi) + (d^*d\psi, \psi) = (d^*\psi, d^*\psi) + (d\psi, d\psi),$$

so $d^*\psi = d\psi = 0$.

Definition 6.3.2. The differential forms satisfying the equation $\Delta\psi = 0$ are called harmonic forms.

To justify the notation Δ and the term “harmonic” we shall calculate the Hodge Laplacian, starting with the case where $\varphi \in \lambda^p(\mathbf{R}^n)$ and the metric in $\mathbf{R}^n$ is the standard Euclidean. Write

$$\varphi = \sum\nolimits' |I| = p\varphi_I dx^I$$

where $\sum'$ denotes summation over increasing sequences $I = (i_1, \dots, i_p)$ and $dx^I = dx^{i_1} \wedge \dots \wedge dx^{i_p}$. Then

$$d\varphi = \sum\nolimits' I, i\partial_i\varphi_I dx^i \wedge dx^I, \quad *\varphi = \sum\nolimits' I, J\varphi_I dx^J \operatorname{sgn}(I, J)$$

where J is an increasing sequence of $n-p$ indices and I, J is a permutation of $1, \dots, n$. For the Dirichlet integral we now obtain if $\varphi \in \lambda^p$ has compact support

$$(\Delta\varphi, \varphi) = \|d\varphi\|^2 + \|d^*\varphi\|^2 = \int \Big(\sum\nolimits' \partial_i\varphi_I\partial_{i'}\varphi_{I'} \operatorname{sgn}\begin{pmatrix} i & I \\ i' & I' \end{pmatrix}$$
$$+ \partial_j\varphi_I\partial_{j'}\varphi_{I'} \operatorname{sgn}\begin{pmatrix} j & J \\ j' & J' \end{pmatrix} \operatorname{sgn}\begin{pmatrix} I & J \\ I' & J' \end{pmatrix}\Big)\, dx.$$

If $I \neq I'$ then the contribution is 0 unless the difference is in just one index. Assume for example that $I = (2, \dots, p+1)$, $I' = (1, \dots, p)$. Then we must have $i = j' = 1$, $i' = j = p+1$, and we obtain

$$\operatorname{sgn}\begin{pmatrix} i & I \\ i' & I' \end{pmatrix} = \operatorname{sgn}\begin{pmatrix} 1 & 2 & \dots & p+1 \\ p+1 & 1 & \dots & p \end{pmatrix} = (-1)^p,$$
$$\operatorname{sgn}\begin{pmatrix} j & J \\ j' & J' \end{pmatrix} = \operatorname{sgn}\begin{pmatrix} p+1 & 1 & p+2 & \dots \\ 1 & p+1 & p+2 & \dots \end{pmatrix} = -1$$
$$\operatorname{sgn}\begin{pmatrix} I & J \\ I' & J' \end{pmatrix} = \operatorname{sgn}\begin{pmatrix} 2 & \dots & p+1 & 1 & p+2 & \dots \\ 1 & \dots & p & p+1 & p+2 & \dots \end{pmatrix} = (-1)^p.$$

If we integrate by parts moving derivatives to the first factor the two terms will therefore give contributions which cancel each other, and the only remaining terms are those with $I = I'$ and $i = i' \in J$ or $j = j' \in I$, hence

$$(\Delta\varphi, \varphi) = \|d\varphi\|^2 + \|d^*\varphi\|^2 = \int {\sum}'_I \Big(\sum_{i=1}^n -\partial_i^2 \varphi_I\Big)\varphi_I \, dx.$$

Since Δ is formally self-adjoint, polarization gives that

$$\Delta\varphi = {\sum}'_I \sum_{i=1}^n -\partial_i^2 \varphi_I \, dx^I. \tag{6.3.3}$$

Thus the Hodge Laplacian operates componentwise as minus the classical Laplacian; the minus sign is natural since the Hodge Laplacian is a positive operator by its definition. We shall write Δ_H whenever confusion seems possible.

To compute Δ generally we can use the formula

$$\Delta = (-1)^{n(p-1)+1} d * d * + (-1)^{np+1} * d * d \quad \text{on } \lambda^p,$$

which follows from (6.3.2) and can be used locally even if M is not orientable. Choose geodesic coordinates $x^1, \dots, x^n$ centered at the point where we want to calculate $\Delta\varphi$. Then $g_{jk}(x) = \delta_{jk} + O(|x|^2)$, by (3.1.15) the second derivatives of g_{jk} are linear functions of the Riemann curvature tensor, and $g_{jk} + g^{jk} - 2\delta_{jk} = O(|x|^3)$. We can use (6.2.9) to express the $*$ operator. When we compute $\Delta_H\varphi(x)$ the result will be the same as for the Euclidean metric unless two derivatives fall on a coefficient involving g^{jk}. Hence (6.3.3) only needs modification by a term of order 0.

Since φ can also be regarded as a skew symmetric tensor field, we could also form another Laplacian by the contraction $\sum g^{jk}\varphi_{,jk}$. At the center of a geodesic coordinate system this will also apart from lower order terms be the classical Laplacian acting on each component of φ, so

$$\Delta_H\varphi + \sum g^{jk}\varphi_{,jk} \quad \text{is of order } 0.$$

When $p = 0$ there is no term $d * d*$ so no second order derivatives of g_{jk} can occur, hence $-\Delta_H\varphi = \sum g^{jk}\varphi_{,jk}$ is the standard Laplace-Beltrami operator. We shall return to the exact calculation for $p \neq 0$ later on in this section. For the moment it suffices for us just to observe that the calculation of the leading term just made shows that the principal symbol of the Hodge Laplacian is $-|\xi|^2$. Hence Δ_H (or rather $-\Delta_H$) satisfies the hypotheses of Theorem 6.1.5. Thus the space $\mathcal{H}^p$ of harmonic p forms is finite dimensional, and every $\psi \in \lambda^p$ can be written in the form

$$\psi = h + \Delta\varphi$$

where h is the orthogonal projection of ψ on $\mathcal{H}^p$ and $\varphi \in \lambda^p$ is uniquely determined mod $\mathcal{H}^p$. Here we have of course used that Δ_H is formally self-adjoint. With $\varphi_+ = d\varphi$ and $\varphi_- = d^*\varphi$, it follows that

$$\psi = h + d\varphi_- + d^*\varphi_+, \quad h \in \mathcal{H}^p,\ \varphi_- \in \lambda^{p-1},\ \varphi_+ \in \lambda^{p+1}. \tag{6.3.4}$$

The three terms in this decomposition are mutually orthogonal, for $dh = d^*h = 0$ by Theorem 6.3.1 and

$$(h, d\varphi_-) = (d^*h, \varphi_-) = 0, \quad (h, d^*\varphi_+) = (dh, \varphi_+) = 0,$$
$$(d\varphi_-, d^*\varphi_+) = (d^2\varphi_-, \varphi_+) = 0.$$

Hence the decomposition (6.3.4) is unique (but φ_- and φ_+ are not). If $d\psi = 0$ then $dd^*\varphi_+ = 0$, which implies

$$(dd^*\varphi_+, \varphi_+) = (d^*\varphi_+, d^*\varphi_+) = 0,$$

so $d\psi = 0$ is equivalent to $d^*\varphi_+ = 0$. Hence we have proved:

Theorem 6.3.3 (Hodge). *Every C^∞ p form on a compact C^∞ Riemannian manifold has a unique decomposition* (6.3.4) *as the sum of one harmonic p form, one form in the range of d, and one form in the range of d^*. Every de Rham cohomology class contains exactly one harmonic form, so $H^p(M)$ can be identified with the space $\mathcal{H}^p$ of harmonic p forms.*

An immediate consequence is of course that $H^p(M)$ is finite dimensional. If M is oriented we also conclude that $H^p(M)$ is isomorphic to $H^{n-p}(M)$ (Poincaré duality). In fact, we have $\Delta * = *\Delta$ since

$$(\Delta\varphi, \psi) = (d\varphi, d\psi) + (*d*\varphi, *d*\psi) = (*d**\varphi, *d**\psi) + (d*\varphi, d*\psi) = (\Delta*\varphi, *\psi),$$

which implies

$$(\Delta*\varphi, \psi) = (\Delta**\varphi, *\psi) = (**\Delta\varphi, *\psi) = (*\Delta\varphi, \psi).$$

Hence $*$ gives an isomorphism $\mathcal{H}^p \to \mathcal{H}^{n-p}$.

Assume now that M is a *compact oriented* C^∞ Riemannian manifold. Then the operator

$$D = d + d^* : \lambda^* \to \lambda^*$$

is defined, and it satisfies the hypotheses of Corollary 6.1.6 since

$$D^2 = dd + dd^* + d^*d + d^*d^* = dd^* + d^*d = \Delta. \tag{6.3.5}$$

From (6.3.5) it follows that the kernel consists of the harmonic forms, and D is formally selfadjoint by its definition. Hence the range is the orthogonal space of the harmonic forms, and the index is equal to 0. However, one can obtain operators with non-trivial index by restricting D to suitable subbundles of $\wedge^*T^*M$.

First note that D maps forms of even (odd) degree to forms of odd (even) degree. If we write

$$\lambda^{\text{even}} = \oplus\lambda^{2p}, \lambda^{\text{odd}} = \oplus\lambda^{2p+1},$$

it follows that the restriction D^{eo} to λ^{even} is a differential operator from λ^{even} to λ^{odd} with adjoint equal to the restriction of D to λ^{odd}, with values in λ^{even}. Hence the kernel of D^{eo} is $\oplus\mathcal{H}^{2p}$, and the range is the orthogonal space of $\oplus\mathcal{H}^{2p+1}$ in λ^{odd}. This means that

$$\text{(6.3.6)} \qquad \operatorname{ind} D^{\text{eo}} = \sum_{p=0}^{n}(-1)^p \dim \mathcal{H}^p = \sum_{p=0}^{n}(-1)^p \dim H^p(M),$$

which is the Euler characteristic of M. If we accept the fact that the Euler characteristic is equal to the number of critical points of functions on M, counted with signs, then Theorem 4.3.2 means that the index of D^{eo} can be expressed in terms of the curvature forms. This is the model for the index theorems which will be a major subject of this chapter.

Assume now that M is a *compact oriented* C^∞ Riemannian manifold of *even dimension* $n = 2l$. By Proposition 6.2.1 the map τ defined in $\wedge^* T^*_x$ for every $x \in M$ gives a decomposition

$$(\wedge^* T^* M)_{\mathbf{C}} = \wedge_+ \oplus \wedge_-,$$

where $\wedge_\pm$ are complex vector bundles of fiber dimension 2^{n-1}. We claim that

$$\text{(6.3.7)} \qquad D\tau = -\tau D,$$

which will prove that D maps sections of $\wedge_+$ $(\wedge_-)$ to sections of $\wedge_-$ $(\wedge_+)$. To prove (6.3.7) we let φ be a p form. Since the dimension n is even, we have $d^* = -*d*$, hence

$$\begin{aligned}
D\tau\varphi &= i^{p(p-1)+l}(d - *d*) * \varphi = i^{p(p-1)+l}(d * \varphi - (-1)^{p^2} * d\varphi)\\
\tau D\varphi &= \tau(d\varphi - *d * \varphi) = i^{(p+1)p+l} * d\varphi - i^{(p-1)(p-2)+l} * *d * \varphi\\
&= i^{p(p-1)+l}\big((-1)^p * d\varphi - (-1)^{p-1}(-1)^{2l-p+1} d * \varphi\big)\\
&= i^{p(p-1)+l}\big((-1)^p * d\varphi - d * \varphi\big).
\end{aligned}$$

This proves (6.3.7).

We shall now determine the index of the restriction $D^+ : C^\infty(M, \wedge_+) \to C^\infty(M, \wedge_-)$ of D. The adjoint of D^+ is the restriction $D^- : C^\infty(M, \wedge_-) \to C^\infty(M, \wedge_+)$, so the kernels $\mathcal{H}^\pm$ of $D^\pm$ are the harmonic forms h with $\tau h = \pm h$, and by definition

$$\operatorname{ind} D^+ = \dim \mathcal{H}^+ - \dim \mathcal{H}^-.$$

$\mathcal{H}^k_{\mathbf{C}} \oplus \mathcal{H}^{2l-k}_{\mathbf{C}}$ is mapped to itself by τ, and if $k < l$ then the eigenspaces with eigenvalues ± 1 are

$$\{h \pm \tau h; h \in \mathcal{H}^k_{\mathbf{C}}\}$$

so they have the same dimension. What remains is to examine how τ splits $\mathcal{H}^l_{\mathbf{C}}$. There we have $\tau = i^{l^2} *$ which is equal to $*$ if l is even and equal to $i*$ if l is odd. For odd l this means that $\mathcal{H}^l_{\mathbf{C}}$ is split according to the eigenspaces of the operator $*$ corresponding to eigenvalues $\pm i$, and since $*$ is a real operator in $\mathcal{H}^l$, they have the same dimension which gives no contribution to the index. On the other hand, if $l = 2k$ is even, that is,

$n \equiv 0 \mod 4$, then τ is a real operator and it suffices to consider the eigenspaces of $*$ acting in the real vector space $\mathcal{H}^l$ corresponding to eigenvalues ± 1. If $\varphi \in \mathcal{H}^l$ then

$$(\varphi, \tau\varphi) = \int_M \varphi \wedge * * \varphi = \int \varphi \wedge \varphi \tag{6.3.8}$$

is positive (negative) when φ is in the eigenspace of $*$ with eigenvalue $+1$ (-1), so the index of D^+ is equal to the index of the quadratic form

$$\mathcal{H}^l \ni \varphi \mapsto \int_M \varphi \wedge \varphi. \tag{6.3.9}$$

(If l is odd then $\varphi \wedge \varphi = 0$ since $\varphi \wedge \varphi = (-1)^{l^2} \varphi \wedge \varphi$, so (6.3.9) vanishes then.)

We can give (6.3.9) an interpretation which does not involve harmonic forms. To do so we note that the bilinear form

$$(\varphi, \psi) \mapsto \int_M \varphi \wedge \psi \tag{6.3.10}$$

on $\{\theta \in \wedge^l; d\theta = 0\}$ induces a bilinear form on $H^l(M)$. In fact, if $\varphi = d\theta$, then $\varphi \wedge \psi = d(\theta \wedge \psi)$ so (6.3.10) vanishes then, and the same is true if $\psi = d\theta$. Interchanging φ and ψ multiplies the integrand in (6.3.10) by $(-1)^{l^2}$, so the form is symmetric if l is even and it is skew symmetric if l is odd.

Definition 6.3.4. The signature of a compact oriented Riemannian manifold M of even dimension $2l$ is the signature of the quadratic form induced in $H^l(M)$ by (6.3.10), if l is even, and it is 0 if l is odd.

With this definition we have proved

$$\operatorname{ind} D^+ = \text{the signature of } M. \tag{6.3.11}$$

The expression of $\operatorname{ind} D^+$ in terms of Pontrjagin classes will be discussed later on.

We shall close this section by making a few additional calculations involving $*$, d^* and Δ. Let φ be a one form and write $\varphi = \sum \varphi_j dx^j$ in local coordinates. By (6.2.9)

$$\begin{aligned} * \varphi = \sum_{k,i=1}^{n} \varphi_k g^{ki} (-1)^{i-1} dx^1 \wedge \ldots \underbrace{\wedge dx^i}_{\text{omit}} \wedge \cdots \wedge dx^n \sqrt{g} \\ = \sum_{i=1}^{n} \Phi^i (-1)^{i-1} dx^1 \wedge \ldots \underbrace{\wedge dx^i}_{\text{omit}} \wedge \cdots \wedge dx^n \sqrt{g}, \end{aligned}$$

where $\Phi = \varphi^\sharp$. Hence $d * \varphi = \sum_i (\partial_i (\Phi^i \sqrt{g}) dx^1 \wedge \cdots \wedge dx^n$, so

$$d^* \varphi = - * d * \varphi = -g^{-\frac{1}{2}} \sum_{i=1}^{n} \partial_i (\Phi^i g^{\frac{1}{2}}) = -\operatorname{div} \Phi, \tag{6.3.12}$$

where the last equality is a definition. The invariance of $\operatorname{div}\Phi$ when Φ is a vector field is also clear since for $\chi \in C_0^\infty(M)$ we have

$$-\int_M \chi \operatorname{div}\Phi \, dvol = \int_M \langle \Phi, d\chi\rangle \, dvol. \tag{6.3.13}$$

This is also true if $\chi = 1$ and Φ has compact support, that is, the integral of the divergence of a vector field with compact support is equal to 0. We can also write

$$\operatorname{div}\Phi = \sum \Phi^i{}_{,i}, \tag{6.3.14}$$

for (6.3.14) is true in a geodesic coordinate system. (See also the proof of (3.4.15) and Exercise 3.4.2.) Thus

$$\int_M \sum_{i=1}^n \Phi^i{}_{,i} \, dvol = 0, \tag{6.3.14$'$}$$

which is a convenient formula to use for partial integration of expressions involving covariant differentiation.

For the one form φ we have

$$d\varphi = \sum \partial_i \varphi_j dx^i \wedge dx^j = \sum \varphi_{j,i} dx^i \wedge dx^j,$$

since $\varphi_{j,i} = \partial_i \varphi_j + \sum_k \Gamma_{ij}{}^k \varphi_k$ and $\Gamma_{ij}{}^k = \Gamma_{ji}{}^k$. However, one must remember when using this expression that $\varphi_{j,i}$ is not antisymmetric in general. We shall now compute $d^* d\varphi$ by taking another one form ψ with compact support and calculating $(d\varphi, d\psi)$. At a point where $g_{ij} = \delta_{ij}$ the pointwise scalar product is

$$\tfrac{1}{2} \sum_{i,j=1}^n (\varphi_{j,i} - \varphi_{i,j})(\psi_{j,i} - \psi_{i,j}) = \sum_{i,j=1}^n (\varphi_{j,i}\psi_{j,i} - \varphi_{i,j}\psi_{j,i}) = (\nabla\varphi, \nabla\psi) - \sum_{i,j=1}^n \Phi^i{}_{,j} \Psi^j{}_{,i},$$

where $\Phi = \varphi^\sharp$ and $\Psi = \psi^\sharp$. We have written

$$(\nabla\varphi, \nabla\psi)(x) = \sum_{i,j,i',j'=1}^n g^{ii'} g^{jj'} \varphi_{j,i} \varphi_{j',i'}$$

for the natural scalar product in $T_x^*M \otimes T_x^*M$. The last expression is invariant so it is always equal to $(d\varphi, d\psi)(x)$, which proves that

$$(d\varphi, d\psi) = \int_M (\nabla\varphi, \nabla\psi) \, dvol - \int_M \sum_{i,j=1}^n \Phi^i{}_{,j} \Psi^j{}_{,i} \, dvol.$$

In the last integral we can integrate by parts using (6.3.14)$'$ applied to the vector field $\sum_j \Phi^i{}_{,j} \Psi^j$, which gives

$$d^* d\varphi = \nabla^* \nabla \varphi + \sum_{i,j}^n \Phi^i{}_{,ji} dx^j.$$

Here ∇^* is the adjoint of the operator ∇ from one forms, that is, sections of T^*M, to sections of $T^*M \otimes T^*M$, with the scalar product above. Now we add

$$dd^*\varphi = -d\sum_{i=1}^n \Phi^i{}_{,i} = -\sum_{i,j=1}^n \Phi^i{}_{,ij}dx^j$$

It follows from (3.1.6) (see Exercise 3.1.4) that

$$\sum_{i=1}^n (\Phi^i{}_{,ji} - \Phi^i{}_{,ij}) = \sum_{i,l=1}^n R^i{}_{lij}\Phi^l = \sum_{l=1}^n R_{lj}\Phi^l,$$

where the last R is the Ricci tensor. Thus we have proved

$$\Delta_H\varphi = \nabla^*\nabla\varphi + \sum_{i,j=1}^n R_{ij}\Phi^i dx^j. \tag{6.3.15}$$

This makes the remarks following (6.3.3) explicit. The identity (6.3.15) is of the form (6.1.4), corresponding to the Levi-Civita connection in T^*M, and c has been identified as the Ricci curvature. There are similar formulas for forms of degree $p > 1$, due to Weitzenböck [1]; see also de Rham [1, p. 131]. Formulas of the same structure as (6.3.15) are called Weitzenböck decompositions in general.

To show the importance of (6.3.15) we now assume that M is compact and connected, and take the scalar product of (6.3.15) with φ. This gives

$$(\Delta_H\varphi, \varphi) = \|\nabla\varphi\|^2 + \int_M \sum R_{ij}\Phi^i\Phi^j\, dvol, \quad \varphi \in \lambda^1, \Phi = \varphi^\sharp. \tag{6.3.16}$$

If φ is a harmonic one form, that is, $\Delta\varphi = 0$, and the Ricci tensor is non-negative, it follows from (6.3.16) that both terms in the right-hand side must vanish. Hence the covariant differential of φ is equal to 0, and $\varphi^\sharp$ is at every point contained in the kernel of the Ricci tensor. If the Ricci tensor is strictly positive definite at some point, then φ must vanish in a neighborhood. But $\|\varphi(x)\|$ is a constant since $\nabla\varphi = 0$, so we obtain the following theorem of Bochner:

Theorem 6.3.5. *If h is a harmonic one form on a compact, connected Riemannian manifold M with non-negative Ricci tensor, then $\nabla h = 0$ and $h^\sharp$ is in the kernel of R at every point. Hence* $\dim H^1(M) \le n$. *If in addition R is strictly positive definite at some point, then $h = 0$, hence $H^1(M) = 0$ then.*

Theorem 6.3.5 has been the starting point of much work on the connection between the topology of a manifold and its curvature.

6.4. Heat equations. The first step to a calculation of the index of the operators D^{eo} and D^+ introduced in Section 6.3 is to study the heat equations associated with the corresponding metric elliptic operators. Let P be a metric differential operator acting on sections of the C^∞ vector bundle E over the compact C^∞ Riemannian manifold

M. By E we shall also denote the lifting of E to $\mathbf{R}\times M$. We want to solve the Cauchy problem to find $u \in C^\infty(\overline{\mathbf{R}}_+ \times M, E)$ such that

$$\partial_t u = Pu \quad \text{in } [0,\infty)\times M, \qquad u(0) = v, \tag{6.4.1}$$

where v is given in $C^\infty(M,E)$. Formally this means that $u = e^{tP}v$, $t \ge 0$, where

$$e^{tP} = (2\pi i)^{-1}\int_\Gamma \frac{e^{-tz}}{z+P}\,dz$$

with an integration contour Γ such as a sector $< \pi$ of the unit circle with the two tangents in the direction of the bisector of the first and fourth quadrants. (We have defined the principal symbol without the customary factors i, so our conventions make P bounded above.) The distributions $F_\nu(x)$ in (6.1.8), which depend on $z \in \mathbf{C}\setminus\overline{\mathbf{R}}_+$, are inverse Fourier transforms of $\xi \mapsto (z-|\xi|^2)^{-\nu-1}$, and

$$(2\pi i)^{-1}\int_\Gamma (z-|\xi|^2)^{-\nu-1}e^{-tz}\,dz = e^{-t|\xi|^2}(-t)^\nu/\nu!.$$

If we multiply F_ν by $e^{-tz}/(2\pi i)$ and integrate over Γ, we thus obtain at least formally the inverse Fourier transform of $(-t)^\nu e^{-t|\xi|^2}/\nu!$, that is,

$$H_\nu(t,x) = (-t)^\nu H_0(t,x)/\nu!, \quad \text{where } H_0(t,x) = (4\pi t)^{-n/2}e^{-|x|^2/4t},\ t>0, \tag{6.4.2}$$

is the fundamental solution of the scalar heat equation. The identities (6.1.9), (6.1.10) are transformed to the obvious identities

$$\begin{gathered}(\Delta-\partial_t)((-t)^\nu H_0(t,x)/\nu!) = (-t)^{\nu-1}H_0(t,x)/(\nu-1)!,\\ -2t\partial H_0(t,x)/\partial x = xH_0(t,x).\end{gathered}$$

Recall that in (6.1.13) the coefficients u_ν were independent of z. Repeating the proof of (6.1.13) we now obtain with the same coefficients u_ν

$$(P-\partial_t)\sum_0^\mu u_\nu H_\nu = (Pu_\mu)H_\mu. \tag{6.4.3}$$

Since $ze^{-tz} = -\partial_t e^{-tz}$, this can easily be justified also by an integration of (6.1.13) multiplied by e^{-tz}. Since $H_0(t,\cdot)\to\delta_0$ as $t\to+0$, it is clear that $\sum_0^\mu u_\nu H_\nu \to \delta_0$ as $t\to+0$. If we define $H_\nu(t,x)=0$ when $t\le 0$, then $H_\nu\in C^k(\mathbf{R}\times U)$, if $\nu > k+n/2$ and U is a geodesic coordinate patch in M. Every H_ν is C^∞ outside $(0,0)$, so $\sum_0^\mu u_\nu H_\nu(t,x)$ is a good parametrix for $(\partial_t - P)$ when μ is large.

Using a parametrix it is easy to solve the Cauchy problem (6.4.1). At first we put the inhomogeneity in the equation instead of the data and consider the Cauchy problem

$$(\partial_t - P)u = f, \quad u(t,\cdot)=0 \text{ when } t<0, \tag{6.4.1$'$}$$

where $f(t,\cdot)=0$ when $t<0$ and f is smooth. To do so we define $\widetilde{H}$ locally by cutoffs of the sum $\sum_0^\infty u_\nu H_\nu$ with the terms approximated by C^∞ functions vanishing in the lower half space, just as in the parametrix F in Section 6.1. Thus $\widetilde{H}$ is a section of $E \boxtimes E^*$ on $\mathbf{R}\times M\times M$ with $t\ge 0$ in the support and

$$(\partial_t - P_x)\widetilde{H}(t,x,y) = R(t,x,y) \quad \text{when } t>0; \quad \widetilde{H}(t,\cdot,\cdot)\to\delta_{\text{diag}} \quad \text{as } t\to +0,$$

where $R\in C^\infty(\mathbf{R}\times M\times M, E\boxtimes E^*)$ and $R(t,x,y)=0$ when $t<0$. Now set

$$u(t,x)=\iint_{s<t}\widetilde{H}(t-s,x,y)w(s,y)\,ds\,dvol(y).$$

The integral with respect to y converges to $w(t,x)$ as $s\to t$, so we obtain

$$(\partial_t - P)u(t,x) = w(t,x) + \iint_{s<t} R(t-s,x,y)w(s,y)\,ds\,dvol(y).$$

The equation $(6.4.1)'$ now becomes a Volterra integral equation

$$w(t,x)+\iint_{s<t}R(t-s,x,y)w(s,y)\,ds\,dvol(y)=f(t,x)$$

with C^∞ kernel, translation invariant in t, so by simple iteration one obtains a unique solution

$$w(t,x)=f(t,x)+\iint_{s<t}K(t-s,x,y)f(s,y)\,ds\,dvol(y).$$

Here $K\in C^\infty(\mathbf{R}\times M\times M, E\boxtimes E^*)$ also vanishes when $t<0$. Hence

$$H(t,x,y)=\widetilde{H}(t,x,y)+\iint_{0<s<t}\widetilde{H}(t-s,x,z)K(s,z,y)\,ds\,dvol(z)$$

is an exact fundamental solution for the Cauchy problem. (The product here is the duality between E_z^* and E_z.) The correction term is infinitely differentiable and vanishes for $t<0$, so the asymptotic behavior at the diagonal is still given by the sum in (6.4.3).

Now the Cauchy problem (6.4.1) has the exact solution

$$u(t,x)=\int H(t,x,y)v(y)\,dvol(y).$$

There is no other solution. In fact, for the scalar product with a solution U of the adjoint equation $(\partial_t+P^*)U=0$ we have

$$\partial_t\langle u,U\rangle = \langle \partial_t u, U\rangle + \langle u, \partial_t U\rangle = \langle Pu, U\rangle + \langle u, -P^*U\rangle = 0.$$

We can choose U as a solution for $t < T$ which is equal to a given section of E^* when $t = T$ by the preceding existence proof, for the change of sign in front of ∂_t changes the direction of "time". Since

$$\langle u(T,\cdot), U(T,\cdot\rangle = \langle u(0,\cdot), U(0,\cdot)\rangle$$

it follows that $u(T,\cdot)$ is uniquely determined by the initial data v. Hence H is also uniquely determined. (It would therefore have been enough to start from an approximation $\sum_0^\mu u_\nu H_\nu$ with a fixed large μ; the resulting H is then independent of μ.) In particular it follows that

$$H(t,x,x) \sim \sum_0^\infty t^{j-n/2} h_j(x), \quad t \to +0, \tag{6.4.4}$$

where $h_j \in C^\infty(M, E \otimes E^*)$. If one introduces geodesic coordinates at $y \in M$ and a corresponding synchronous frame for E, with the connection defined in Proposition 6.1.2, then the coefficients in the expansion at y are given by (6.1.15) and the recursion formulas (6.1.14). For operators such as D^+ which are entirely determined by a Riemannian geometry in M, it follows from Theorem 3.3.5 that $h_j(y)$ can be expressed as a polynomial in the Riemann curvature tensor and its covariant derivatives at y. (The trivialization of the bundles $\wedge_\pm$ is then given by the Levi Civita radial parallel translation.) This will be the starting point for the discussion of the corresponding index in the following section.

6.5. Hirzebruch's index formula and Gilkey's theorem. Let M be an oriented Riemannian manifold of even dimension, and consider the operator D^+ from sections of $\wedge_+$ to $\wedge_-$ defined in Section 6.3, using the decomposition of the exterior algebra by the eigenvalues of the map τ defined in Section 6.2. Let $H^\pm(t,x,y)$ be the fundamental solutions of $\partial_t + D^{\pm *} D^\pm$ constructed in Section 6.4; recall that $D^\mp$ is the adjoint of $D^\pm$. In (6.3.11) we expressed the index of D^+ in terms of the intersection form of the cohomology in the middle dimension. We shall now give an analytical expression in terms of the heat kernels:

Proposition 6.5.1. *For every $t > 0$ we have*

$$\operatorname{ind} D^+ = \int (\operatorname{Tr} H^+(t,x,x) - \operatorname{Tr} H^-(t,x,x))\, dvol(x). \tag{6.5.1}$$

Here $H^\pm(t,x,x)$ is a linear transformation in $\wedge_{\pm x}$.

Proof. Let us consider the decomposition of $C^\infty(M,\wedge_+)$ given by the eigenspaces of the self-adjoint operator $D^{+*}D^+$. The spectrum is discrete, for if

$$\|D^{+*}D^+ u\|_{L^2} + \|u\|_{L^2} \le 1,$$

then u belongs to a compact subset of L^2. In fact, there is a parametrix F such that

$$u = FD^{+*}D^+ u + Ru$$

where R is an integral operator with C^∞ kernel, and F is continuous from L^2 to the space of sections with first derivatives in L^2, because first derivatives of F are integrable. The set Γ_λ^+ of eigenfunctions with eigenvalue λ is a subset of $C^\infty(M, \wedge_+)$, and

$$L^2(M, \wedge_+) = \oplus_\lambda \Gamma_\lambda^+(\wedge_+).$$

If $\varphi \in \Gamma_\lambda^+$ then

$$\varphi = F\lambda\varphi + R\varphi = \cdots = (F\lambda)^N\varphi + (I + F\lambda + \cdots + (F\lambda)^{N-1})R\varphi.$$

When $N > n/4$ the kernel of F^N is square integrable in each variable, so we obtain

$$\sup|\varphi(x)| \le C(1+\lambda)^{1+\frac{N}{4}}\|\varphi\|_{L^2}, \quad \varphi \in \Gamma_\lambda. \tag{6.5.2}$$

The eigenspaces Γ_λ^- of $D^{-*}D^- = D^+D^{+*}$ have similar properties and their orthogonal direct sum is $L^2(M, \wedge_-)$.

Since $D^{+*}D^+u = \lambda u$ implies $D^+D^{+*}(D^+u) = \lambda D^+u$, and $D^+u = 0$ implies $u = 0$ if $\lambda \ne 0$, the restriction of D^+ to Γ_λ^+ is then an injective map into Γ_λ^-. In the same way we see that D^{+*} defines an injective map $\Gamma_\lambda^- \to \Gamma_\lambda^+$, so these spaces have the same dimension if $\lambda \ne 0$. Since $\Gamma_0^\pm = \operatorname{Ker} D^\pm$, it follows that

$$\operatorname{ind} D^+ = \dim\Gamma_0^+ - \dim\Gamma_0^- = \sum \chi(\lambda)(\dim\Gamma_\lambda^+ - \dim\Gamma_\lambda^-), \tag{6.5.3}$$

if $\chi(0) = 1$. We shall prove that $\sum\chi(\lambda)\dim\Gamma_\lambda^\pm$ converges when $\chi(\lambda) = e^{-t\lambda}$, $t > 0$, and that the sum is equal to $\int \operatorname{Tr} H^\pm(t,x,x)\, dvol(x)$. This implies (6.5.1).

If $\varphi \in \Gamma_\lambda^+$ then $(\partial_t + D^{+*}D^+)(e^{-t\lambda}\varphi) = 0$, hence

$$e^{-t\lambda}\varphi(x) = \int H^+(t,x,y)\varphi(y)\, dvol(y).$$

If $\alpha \in \wedge_{+x}$ it follows that

$$e^{-t\lambda}(\varphi(x), \alpha) = \int (\varphi(y), H^+(t,x,y)^*\alpha)\, dvol(y).$$

If we let φ run through a complete orthonormal system of eigenfunctions φ_j, with corresponding eigenvalues λ_j, then Parseval's formula gives

$$\sum e^{-2t\lambda_j}|(\varphi_j, \alpha)|^2 \le Ct^{-n}\|\alpha\|^2, \quad 0 < t < 1,$$

and if we sum over all α in an orthonormal basis in $\wedge_{+x}$ we get for $0 < t < 1$

$$\sum e^{-2t\lambda_j}|\varphi_j(x)|^2 \le C't^{-n}, \quad \text{hence} \quad \sum e^{-2t\lambda}\dim\Gamma_\lambda^+ \le C''t^{-n}. \tag{6.5.4}$$

The last bound follows by integrating the first over M, and it proves the convergence of the sum in (6.5.3) when $\chi(\lambda) = e^{-t\lambda}$, $t > 0$.

If f is a finite linear combination of the eigenfunctions φ_j, then

$$\int H^+(t,x,y)f(y)\,dvol(y) = \sum e^{-t\lambda_j}\varphi_j(x)(f,\varphi_j).$$

Since such sections of $\wedge_+$ are dense in the continuous sections it follows in view of (6.5.2) and (6.5.4) that for every $f \in \wedge_{+x}$

$$H^+(t,x,x)f = \sum e^{-t\lambda_j}\varphi_j(x)(f,\varphi_j(x)).$$

Let f_ν, $\nu = 1,\dots,2^{n-1}$, be an orthonormal basis in $\wedge_{+x}$. If we choose $f = f_\nu$, take the scalar product with f_ν in $\wedge_{+x}$ and sum, it follows that

$$\operatorname{Tr} H^+(t,x,x) = \sum_{j,\nu} e^{-t\lambda_j}|(\varphi_j(x),f_\nu)|^2 = \sum_j e^{-t\lambda_j}|\varphi_j(x)|^2.$$

Integration over M gives

$$\int \operatorname{Tr} H^+(t,x,x)\,dvol(x) = \sum e^{-t\lambda}\dim\Gamma^+_\lambda.$$

We have an analogous formula for H^-, so (6.5.1) follows from (6.5.3).

By (6.4.4) there is an asymptotic expansion

$$\operatorname{Tr} H^+(t,x,x) - \operatorname{Tr} H^-(t,x,x) \sim \sum_0^\infty t^{j-\frac{n}{2}}h_j(x),$$

where h_j is a polynomial in the components of the Riemann curvature tensor and its covariant derivatives. The product $h_j(x)\,dvol(x)$ is a density on the oriented manifold M. If we reverse the orientation, then τ is replaced by $-\tau$, which means that the spaces $\wedge_+$ and $\wedge_-$ and therefore the kernels H^+ and H^- are interchanged. Thus h_j is replaced by $-h_j$, which means that we can regard $h_j(x)\,dvol(x)$ as a n form ω_j. From (6.5.1) it follows then that

(6.5.5) $$\int_M \omega_j = 0, \quad j < n/2,$$

(6.5.6) $$\int_M \omega_{n/2} = \operatorname{ind} D^+.$$

Our aim now is to prove that $\omega_j = 0$ when $j < n/2$, which is a very much stronger statement than the integral condition (6.5.5), and that $\omega_{n/2}$ is a polynomial in the Pontrjagin forms. The following simple observation is crucial:

Lemma 6.5.2. *If the metric ds^2 is replaced by $\lambda^2 ds^2$ where λ is a positive constant, then Δ is multiplied by λ^{-2} and ω_j is multiplied by λ^{n-2j}.*

Proof. In a local coordinate system g_{jk} is replaced by $\lambda^2 g_{jk}$ and g^{jk} by $\lambda^{-2}g^{jk}$, which means that Δ is replaced by $\lambda^{-2}\Delta$. Now

$$\lambda^{-2}D^{+*}D^{+} + \partial_t = \lambda^{-2}(D^{+*}D^{+} + \partial_\tau),$$

if $\tau = \lambda^{-2}t$. Hence the initial value problem for $\lambda^{-2}D^{+*}D^{+} + \partial_t$ has the solution

$$\int H^{+}(t/\lambda^2, x, y) f(y)\, dvol(y)$$

where H^{+} and $dvol(y)$ belong to the metric ds^2. For the metric $\lambda^2 ds^2$ the volume element is $\lambda^n dvol(y)$, so the new fundamental solution is $H^{+}(t/\lambda^2, x, y)\lambda^{-n}$. The product of $H^{\pm}(t/\lambda^2, x, x)\lambda^{-n}$ by the new volume element $\lambda^n dvol(x)$ is equal to $H^{+}(t/\lambda^2, x, x)\, dvol(x)$, so

$$\sum t^{j-\frac{n}{2}}\omega_j \quad \text{is replaced by} \quad \sum \lambda^{n-2j} t^{j-\frac{n}{2}}\omega_j,$$

which proves the lemma.

When $j < n/2$ the weight of ω_j is positive in the sense that the power of λ in Lemma 6.5.2 is positive, and the weight is 0 when $j = n/2$. We shall prove a theorem of Gilkey which states that this suffices to make the conclusions about ω_j mentioned above. Of course we also have to use that if $x^1, \ldots, x^n$ is a geodesic coordinate system centered at p, then

$$\omega_j(p) = \Phi_j(R, R', \ldots, R^{(k)})\, dx^1 \wedge \cdots \wedge dx^n,$$

where Φ_j is a polynomial in the curvature tensor R and its covariant derivatives of order $\le k$ when $x = 0$. Note that the local definition of ω_j makes these forms defined for every Riemannian manifold of even dimension.

Definition 6.5.3. A q form invariant of Riemannian manifolds M of dimension n is a function which to every such manifold assigns a q form ω on M such that for every $p \in M$ and geodesic coordinate system $x^1, \ldots, x^n$ centered at p we have for some κ

$$\omega(p) = \sum{}' |I| = q\Phi_I(R, \ldots, R^{(\kappa)})\, dx^I, \tag{6.5.7}$$

where Φ_I is a polynomial in the curvature tensor and its covariant derivatives of order $\le k$ at p, which is independent of M and p. The invariant is said to have weight k if ω is multiplied by λ^k when the first fundamental form is multiplied by λ^2.

Let V be the Euclidean vector space $\mathbf{R}^n$. We can regard the curvature tensor R as an element of $\bigotimes^4 V$ and $R^{(j)}$ as an element of $\bigotimes^{4+j} V$, so (6.5.7) is a polynomial map

$$\Phi : W \to \wedge^q V, \qquad W = \bigoplus_{j=0}^{k} W_j, \quad W_j = \bigotimes^{4+j} V.$$

The orthogonal group operates both in W and in $\wedge^q V$, and we denote the operation of O by O^* in both cases. Since an orthogonal transformation of $x^1, \dots, x^n$ gives a new geodesic system of coordinates centered at p, we must have

$$\Phi(O^* w) = O^* \Phi(w)$$

if $w = (R, \dots, R^{(k)})$ for some choice of a Riemannian metric. (Note that already the Bianchi identities show that R is not an arbitrary element in $\bigoplus^4 V$.) Hence we have

$$\Phi(w) = (O^{-1})^* \Phi(O^* w)$$

for all such w. Taking the average of the right-hand side over the orthogonal group, we obtain a polynomial $\widetilde{\Phi}$ which can still be used in (6.5.7) and has the advantage of being *equivariant*, that is,

$$\widetilde{\Phi}(O^* w) = O^* \widetilde{\Phi}(w), \quad \forall w \in W.$$

To simplify notation we assume in what follows that Φ already has this property. It is clear that if we split Φ into a sum of polynomials which are homogeneous in each of the variables $w_j \in W_j$, then all the terms will be equivariant.

A polynomial map $\Psi(w)$ of degree ν in $w \in W_j$ can be written in one and only one way in the form $\Psi_p(w, \dots, w)$ where $\Psi_p(w_1, \dots, w_\nu)$ is a symmetric multilinear map in $w_1, \dots, w_\nu \in W_j$; the *polarization* Ψ_p is given by

$$\Psi_p(w_1, \dots, w_\nu) = \prod_{j=1}^{\nu} \langle w_j, \partial/\partial w \rangle \Psi(w)/\nu!.$$

Ψ_p defines a linear map on $\bigotimes^p W_j$ which is equivariant if Ψ is. If we use polarization for each W_j we may conclude that Φ is a sum of polynomials each of which is induced by a linear equivariant map

$$\varphi : \bigotimes^N V \to \wedge^q V.$$

The fundamental theorem on $\mathbf{O}(n)$ invariants, Theorem D.1, gives a complete description of such maps when $q = 0$. One calls a linear form φ on $\bigotimes^N V$ elementary if $N = 2k$ and

$$\varphi(v_1 \otimes \cdots \otimes v_{2k}) = (v_1, v_2) \dots (v_{2k-1}, v_{2k}),$$

that is,

$$\varphi(\xi) = \sum \xi_{\alpha_1 \alpha_1 \alpha_2 \alpha_2 \dots \alpha_k \alpha_k}, \quad \xi \in \bigotimes^N V,$$

or if φ differs from this linear form just by a permutation of the indices.

Theorem 6.5.4. *Every invariant linear form $\bigotimes^N V \to \mathbf{R}$ is a linear combination of elementary forms; in particular there are no such forms $\neq 0$ unless N is even.*

This is just a reformulation of Theorem D.1, so we pass to:

Corollary 6.5.5. *If $\varphi : \bigotimes^N V \to \wedge^q V$ is equivariant and not identically* 0*, then $N-q$ is an even integer $2r \geq 0$, and φ is a linear combination of elementary maps, that is, maps which apart from a permutation of the indices are of the form*

$$\varphi(\xi) = \sum_{\alpha,\beta} \xi_{\alpha_1\alpha_1\alpha_2\alpha_2\ldots\alpha_r\alpha_r[\beta_1\ldots\beta_q]} dx^{\beta_1} \wedge \cdots \wedge dx^{\beta_q},$$

where $[\ldots]$ denotes alternation over the enclosed indices, that is, summation over all permutations after multiplication by the sign of the permutation.

Proof. We define an invariant form $\tilde{\varphi} : \bigotimes^{N+q} V \to \mathbf{R}$ by

$$\tilde{\varphi}(v_1 \otimes \cdots \otimes v_{N+q}) = (\varphi(v_1, \ldots, v_N), v_{N+1} \wedge \cdots \wedge v_{N+q}).$$

It follows from Theorem 6.5.4 that $\tilde{\varphi} = 0$ unless $N+q$ is even. If $N+q$ is even then $\tilde{\varphi}(v_1, \ldots, v_{N+q})$ is a linear combination of elementary forms and is preserved by alternation of the last q vectors followed by division by $q!$. This eliminates elementary forms containing a scalar product of two vectors $v_{N+1}, \ldots, v_{N+q}$, for if they are interchanged the sign of the permutation is changed but the term is not affected otherwise. Hence $\tilde{\varphi} = 0$ if $q > N$. If $N = q + 2r$, where $r \geq 0$ is even, then $\tilde{\varphi}$ is a linear combination of forms of the type

$$(v_1, \ldots, v_{N+q}) \mapsto (v_1, v_2) \ldots (v_{2r-1}, v_{2r})(v_{N+1-q}, v_{N+1}) \ldots (v_N, v_{N+q}).$$

Alternation gives the forms in the corollary.

We shall now apply the corollary to the q form invariant ω in Definition 6.5.3. When $\alpha = (\alpha_1, \ldots, \alpha_\nu)$ is a sequence of $|\alpha| = \nu$ indices between 1 and n, we shall denote by R_α the corresponding components of the $\nu - 4$th covariant derivative of the Riemann curvature tensor R_{ijkl}. By an *elementary monomial* of degree r in R we shall mean an expression of the form

$$m(R) = \sum\nolimits_q^{*} R_{\alpha^1} R_{\alpha^2} \ldots R_{\alpha^r}, \tag{6.5.8}$$

where the summation indicates alternation of precisely q indices and pairwise contraction of the others in the total index sequence $\alpha^1\alpha^2\ldots\alpha^r$, where the number of indices shall exceed q by an even number (or 0). What we have proved so far is that with coefficients $a_m \in \mathbf{R}$ we have

$$\omega(p) = \sum a_m m(R). \tag{6.5.7$'$}$$

The number of possible terms is reduced if one examines the weights:

Lemma 6.5.6. *The weight of* $m(R)$ *is* $2r + q - |\alpha|$ *where* $|\alpha| = |\alpha^1| + \dots |\alpha^r|$.

Proof. With a fixed system of local coordinates the Christoffel symbols Γ_{ikj} are multiplied by λ^2 if the metric is multiplied by λ^2, but the Christoffel symbols $\Gamma_{ik}{}^j$ remain unchanged. By (2.1.13) it follows that R_{ijkl} is multiplied by λ^2, and so are the covariant derivatives. If $t_1, \dots, t_q$ are tangent vectors, then $m(R)$ evaluated with respect to these in a fixed coordinate system is

$$\sum_q{}^* R_{\alpha^1} \dots R_{\alpha^r} g^{i_1 i_2} \dots t_1^{j_1} \dots t_q^{j_q}, \tag{6.5.8$'$}$$

where for each pair $i_1 i_2$ of indices to be contracted there is a factor $g^{i_1 i_2}$ while $j_1, \dots, j_q$ correspond to indices to be alternated. In fact, (6.5.8)$'$ is invariant under coordinate changes and agrees with (6.5.8) at the center of a geodesic coordinate system. The number of indices in the contraction is $|\alpha| - q$, so replacing g by $\lambda^2 g$ multiplies (6.5.8)$'$ by $\lambda^{2r}\lambda^{-(|\alpha|-q)}$, which proves the lemma.

Write $|\alpha^i| = 4 + \varepsilon_i$ where ε_i is the number of covariant differentiations which occur in R_{α^i}, and let $\varepsilon = \sum \varepsilon_i$ be the total number of them. Then it follows from Lemma 6.5.6 that for the terms in (6.5.7)$'$ the weight of $m(R)$ is equal to $q - 2r - \varepsilon$. Thus one must alternate with respect to a larger number of indices to get a higher weight, which explains why invariants of higher weight have a simpler structure.

Theorem 6.5.7 (Gilkey). *Every* q *form invariant of positive weight for Riemannian manifolds is equal to* 0*, and every* q *form invariant of weight* 0 *is contained in the ring generated by the Pontrjagin forms, which all have weight* 0*.*

Proof. The curvature tensor and its covariant derivative have the symmetry properties

$$R_{ijkl} = -R_{jikl}, \quad R_{ijkl} = -R_{ijlk}, \quad R_{ijkl} = R_{klij}, \tag{6.5.9}$$

$$R_{i[jkl]} = 0, \quad R_{ij[kl,m]} = 0 \quad \text{(the Bianchi identities)}. \tag{6.5.10}$$

Such identities are preserved by covariant differentiation since it commutes with the permutation group acting on the tensor product. Thus differentiation of the first Bianchi identity gives

$$R_{ijkl,m} + R_{iljk,m} + R_{iklj,m} = 0$$

if we write out all terms explicitly. The alternation of R_{ijkl} over three arbitrary indices is zero by the first Bianchi identity and the symmetries (6.5.9), and this is also true for $R_{ijkl,m}$. When one alternates over mij or mkl this is the second Bianchi identity. If one alternates over jkm and notes that

$$R_{ijkl,m} - R_{ikjl,m} + R_{iljk,m} = 0,$$

it follows that two times the alternation is 0, for exchanging k and j changes the sign of the permutation. This proves the claim. By covariant differentiation we conclude that the alternation of R_α over three of the first four or five indices is always equal to 0.

If the term $m(R)$ is not zero it follows that the alternation in (6.5.8) involves at most two of the first four indices in each factor. Altogether we can then alternate in at most $2r+\varepsilon$ indices, hence $q \le 2r+\varepsilon$, which means that the weight of $m(R)$ is ≤ 0. This proves the first part of Gilkey's theorem. If the weight is 0 we have also found that $q = 2r+\varepsilon$ and that we must alternate over all indices corresponding to covariant differentiation and in addition two of the first four indices in each factor. If a covariant differentiation occurs in some factor, it follows that we alternate over three of the first five indices in it, which implies that $m(R) = 0$. Hence $\varepsilon = 0$, that is, *no covariant derivatives occur.* Furthermore we have alternation with respect to precisely two indices in each factor R_{ijkl}. If they are the first two, we can use the symmetry $R_{ijkl} = R_{klij}$ to replace them by the last two. If they are the middle two, we can use that

$$R_{ijkl} - R_{ikjl} = R_{ilkj} = (R_{ilkj} - R_{iljk})/2$$

to replace them by the last two indices, at the expense of a factor $\frac{1}{2}$. Hence it follows that in a geodesic coordinate system ω is a sum of products of expressions of the form

$$\sum_{i,j} R_{i_1 i_2 j_1 j_2} R_{i_2 i_3 j_3 j_4} \dots R_{i_k i_1 j_{2k-1} j_{2k}} dx^{j_1} \wedge \dots \wedge dx^{j_{2k}}.$$

With the notation in (4.1.13) we recognize this as 2^k times the form

$$\Omega_{i_1 i_2} \wedge \Omega_{i_2 i_3} \wedge \dots \wedge \Omega_{i_k i_1}. \tag{6.5.11}$$

Hence the Gilkey theorem is now a consequence of the following:

Lemma 6.5.8. *The Pontrjagin forms are universal polynomials in the forms* (6.5.11)*.*

Proof. We proved in Section 4.4 that the forms (6.5.11) are closed forms on M, lifted to $P(M)$. The Pontrjagin forms are the forms of degree $4, 8, \dots$ in

$$\det(\delta_{ij} + (2\pi)^{-1}\Omega_{ij}).$$

Let us now note that for $n \times n$ matrices A_{ij} with complex coefficients we have

$$\det(\lambda\delta_{ij} - A_{ij}) = \prod_1^n (\lambda - \lambda_j) = \sum_0^n \lambda^{n-j}(-1)^j c_j,$$

where λ_j are the eigenvalues and c_j the elementary symmetric functions of them. We have for every positive integer k

$$s_k = \sum_{j=1}^n \lambda_j^k = \sum A_{i_1 i_2} A_{i_2 i_3} \dots A_{i_k i_1},$$

for the sum is invariant under conjugation of the matrix A, and the formula is valid for diagonal matrices. Since $c_1 = s_1$, and by Newton's formulas

$$s_k = c_1 s_{k-1} - c_2 s_{k-2} + \dots + (-1)^k k c_k, \quad 2 \le k \le n,$$

we can for every $k \le n$ express s_k as a polynomial in $c_1, \dots, c_k$ and express c_k as a polynomial in $s_1, \dots, s_k$. But these polynomial identities in the coefficients of A_{ij} remain valid if we replace A_{ij} by the commuting forms Ω_{ij}, which completes the proof of the lemma and of the Gilkey theorem.

For an oriented compact manifold M of dimension $4k$ we can apply the Gilkey theorem to the forms ω_j in (6.5.5), (6.5.6) which occur in the expansion

$$\operatorname{Tr}\big(H^+(t,x,x) - H^-(t,x,x)\big)\, dvol(x) \sim \sum_{j=0}^{\infty} t^{j-2k}\omega_j.$$

The conclusion is that

$$\omega_j = 0, \quad j < 2k; \tag{6.5.5$'$}$$

$$\omega_{2k} = L_k(p_1, \dots, p_k). \tag{6.5.6$'$}$$

Here L_k is a polynomial of weight k if p_j are indeterminates of weight j, so replacing p_j by the Pontrjagin form of degree $4j$ gives a a form of highest degree $4k$ which is equal to ω_{2k}. Integration of (6.5.5)$'$ gives back (6.5.5), and integration of (6.5.6)$'$ gives by (6.5.6) and (6.3.11)

$$\operatorname{ind} D^+ = \operatorname{sign} M = \int_M L_k(p_1, \dots, p_k) \tag{6.5.6$''$}$$

for every oriented manifold M of dimension $4k$. One can now determine the coefficients of L by specializing M to manifolds for which the signature and the Pontrjagin classes are easy to describe, such as products of complex projective spaces. We refer to Atiyah, Bott and Patodi [1] for the calculation of L_k and only give the result. Set

$$\sum L_k = \prod x_j / \tanh x_j, \quad \text{where} \prod (1 + x_j^2) = \sum p_k. \tag{6.5.12}$$

This should be understood as follows. To calculate L_k we take $m \ge k$ variables x_j and note that $x_j / \tanh x_j$ is a power series in x_j^2. Collecting the terms of degree $2k$ in the product for $j = 1, \dots, m$ gives a polynomial L_k in the elementary symmetric functions $p_1, p_2, \dots$ of the $x_1^2, \dots, x_m^2$ which has weight k when p_j is given the weight j. The polynomial is independent of the choice of $m \ge k$, and one verifies that it is the only polynomial which makes (6.5.6)$''$ valid for products of complex projective spaces. Apart from the verification of this we have now proved

Theorem 6.5.9 (Hirzebruch). *The signature of an oriented Riemannian manifold M of dimension $4k$ is given by* (6.5.6)$''$ *where the polynomial L_k is defined by* (6.5.12) *and $p_1, \dots$ are the Pontrjagin forms of M.*

We also refer to Atiyah, Bott and Patodi [1] for an extension of Theorem 6.5.9 to the signature operator with coefficients in a vector bundle, and for the arguments required to go from there to the general index theorem for elliptic pseudo-differential operators.

We have followed that paper here apart from substituting the Hadamard construction of a parametrix for pseudo-differential operator theory. The Hadamard construction is not only more elementary, it fits precisely with the differential geometric context. For the operator D^{eo} having the Euler characteristic as index, it was first proved by Patodi [1] that a corresponding phenomenon occurs, leading to the Gauss-Bonnet-Chern formula. A variant of his proof using ideas of supersymmetry from physics is given in Cycon, Froese, Kirsch and Simon [1, Chapter 12]. (See page 258 for a criticism of the methods used here which do not go all the way to a direct computation of the coefficients in the index formula.) More recently, a direct computational proof has been obtained for twisted Dirac operators by Bismut [1] and Getzler [1,2]. Our next aim is to give an exposition of the methods of Getzler.

6.6. Operators of Dirac type. We started this chapter with a study of second order metric elliptic operators, but all the geometric applications in Sections 6.3 and 6.5 concerned first order operators D such that $-D^*D$ is metric. (The minus sign comes from the convention for defining the principal symbol introduced in Section 5.1, which does not contain the factor i which is customary in pseudo-differential operator theory.) The rest of this chapter will be devoted to a more systematic search for such operators.

Definition 6.6.1. If M is a C^∞ manifold and E_0, E_1 two C^∞ Hermitian vector bundles on M with the same fiber dimension, then a first order differential operator

$$D : C^\infty(M, E_0) \to C^\infty(M, E_1)$$

is said to be of Dirac type if $-D^*D : C^\infty(M, E_0) \to C^\infty(M, E_0)$ is a metric differential operator.

Here the formal adjoint D^* is defined using the hermitian metrics in E_0, E_1, and some positive density in M; the definition is independent of the choice of density since it only influences terms of lower order. If $\sigma(x,\xi) : E_{0x} \to E_{1x}$ is the symbol of D at $(x,\xi) \in T^*M$, then $\sigma(x,\xi)$ is a linear transformation depending linearly on ξ, and the definition means that

$$\sigma(x,\xi)^*\sigma(x,\xi) = p(x,\xi)I_{E_{0x}}, \tag{6.6.1}$$

where $\sigma(x,\xi)^* : E_{1x} \to E_{0x}$ is the adjoint with respect to the Hermitian metrics and $p(x,\xi)$ is a positive definite quadratic form in T_x^*M. Since E_{0x} and E_{1x} have the same dimension, it follows from (6.6.1) that

$$\sigma(x,\xi)\sigma(x,\xi)^* = p(x,\xi)I_{E_{1x}}, \tag{6.6.1$'$}$$

so D^* is also of Dirac type. When D is an operator of Dirac type on M, *we give M the Riemannian structure defined by the dual of the quadratic form $p(x,\xi)$.*

A Dirac type operator $D : C^\infty(M, E_0) \to C^\infty(M, E_1)$ is called *symmetric* if $E_0 = E_1$ and $D^* = D$. Thus $-D^2$ is then a metric operator. This is the context in which Dirac originally introduced what is now known as a Dirac operator. He wanted to write the Klein-Gordon equation

$$\partial^2/\partial t^2 - \partial^2/\partial x^2 - \partial^2/\partial y^2 - \partial^2/\partial z^2 + m^2$$

as the square of a first order operator in order to obtain a relativistically invariant operator similar to the Schrödinger equation, and found that this could be done using a 4×4 system of first order operators. Our Dirac operators are analogous but correspond to a positive definite metric rather than one of Lorentz signature. In part of Section 6.7 we shall avoid making assumptions on the signature in order to cover the original case also.

For a general operator $D : C^\infty(M, E_0) \to C^\infty(M, E_1)$ of Dirac type, a symmetric operator of Dirac type $C^\infty(E_0 \oplus E_1) \to C^\infty(E_0 \oplus E_1)$ is defined by

$$\begin{pmatrix} 0 & D^* \\ D & 0 \end{pmatrix}; \quad \text{the symbol is } \begin{pmatrix} 0 & \sigma(x,\xi)^* \\ \sigma(x,\xi) & 0 \end{pmatrix}. \tag{6.6.2}$$

It maps sections of $E_0 \subset E_0 \oplus E_1$ to sections of $E_1 \subset E_0 \oplus E_1$ and vice versa. Conversely, a symmetric Dirac type operator with this property in $C^\infty(M, E)$, $E = E_0 \oplus E_1$, is always obtained from a Dirac type operator $C^\infty(M, E_0) \to C^\infty(M, E_1)$ as in (6.6.2).

If $\sigma(x, \xi)$ is the principal symbol of a symmetric Dirac type operator in $C^\infty(M, E)$, then

$$\sigma(x,\xi) = \sigma(x,\xi)^*, \qquad \sigma(x,\xi)^2 = p(x,\xi) I_{E_x}. \tag{6.6.3}$$

The first step in the study of (symmetric) Dirac operators is to determine matrices $\sigma(x, \xi)$ depending linearly in ξ, which satisfy (6.6.3) for a *fixed* x. This leads to the definition of Clifford algebras and spinors, which will be studied in Section 6.7. We can then define (twisted) Dirac operators in Section 6.8, where we also prove an analogue of the Weitzenböck formula due to Lichnerowicz. Section 6.9 is an analytical interlude devoted to the classical Mehler formula giving the heat kernel for the harmonic oscillator explicitly, and to some extensions needed here. We are then prepared to prove the local index theorem for Dirac operators in Section 6.10.

6.7. Clifford and spinor algebra. Let V be a real vector space, q a quadratic form in V, E a complex vector space, and $\sigma : V \to \mathrm{End}(E) = \mathcal{L}(E, E)$ a linear map such that

$$\sigma(v)^2 = q(v) I_E, \tag{6.7.1}$$

as in (6.6.3). Every linear map $\sigma : V \to \mathrm{End}(E)$ can be uniquely extended to a linear map $\tilde\sigma$ from the tensor algebra $\bigoplus_0^\infty \otimes^k V$ to $\mathrm{End}(E)$ with

$$\tilde\sigma(v_1 \otimes \cdots \otimes v_k) = \sigma(v_1)\ldots\sigma(v_k), \quad v_1, \ldots, v_k \in V,$$

for the right-hand side is a multilinear function of $v_1, \ldots, v_k$. The condition (6.7.1) means that for $v \in V$ and arbitrary tensors t_1, t_2

$$\tilde\sigma(t_1 \otimes v \otimes v \otimes t_2) - \tilde\sigma(t_1 \otimes q(v) \otimes t_2) = 0,$$

that is, $\tilde\sigma$ vanishes on the two sided ideal $\mathcal{I}$ generated in the tensor algebra by all elements of the form $v\otimes v-q(v)\cdot 1$ with $v\in V$, so $\tilde\sigma$ induces an algebra homomorphism from the quotient algebra to $\operatorname{End}(E)$. Note that with the notation q also for the polarized form

$$(v_1\pm v_2)\otimes(v_1\pm v_2)-q(v_1\pm v_2)\cdot 1\in\mathcal{I}\implies v_1\otimes v_2+v_2\otimes v_1-2q(v_1,v_2)\cdot 1\in\mathcal{I}.$$

If we choose a basis $\varepsilon_1,\dots,\varepsilon_n$ in V which diagonalizes q,

$$q(\sum_1^n \xi_j\varepsilon_j)=\sum_1^n q_j\xi_j^2,\quad \xi\in\mathbf{R}^n, \tag{6.7.2}$$

it follows that $\mathcal{I}$ is generated by all elements of the form

$$\varepsilon_j\otimes\varepsilon_k+\varepsilon_k\otimes\varepsilon_j,\ j\neq k;\quad \varepsilon_j\otimes\varepsilon_j-q_j\cdot 1;\quad j,k=1,\dots,n. \tag{6.7.3}$$

Every element in the tensor algebra is a linear combination of tensor products of $\varepsilon_1,\dots,\varepsilon_n$, so using (6.7.3) we see that modulo $\mathcal{I}$ it is congruent to an element of the form

$$\sum_{j\le n} a_{i_1,\dots,i_j}\varepsilon_{i_1}\otimes\dots\otimes\varepsilon_{i_j},\quad i_1<i_2<\dots<i_j. \tag{6.7.4}$$

Here $a_{i_1,\dots,i_j}$ is a linear form on the tensor algebra vanishing in $\mathcal{I}$, defined by

$$a_{i_1,\dots,i_j}(\varepsilon_{k_1}\otimes\dots\otimes\varepsilon_{k_\nu})=(-1)^\iota\prod_{i=1}^n q_i^{\mu_i},$$

if the indices $i_1,\dots,i_j$ occur $1+2\mu_{i_1},\dots,1+2\mu_{i_j}$ times among $k_1,\dots,k_\nu$, the other indices $i\le n$ occur $2\mu_i$ times, and ι is the number of index pairs in $k_1,\dots,k_\nu$ which occur in the wrong order. Otherwise $a_{i_1,\dots,i_j}(\varepsilon_{k_1}\otimes\dots\otimes\varepsilon_{k_\nu})=0$. This follows since the definition makes the form $a_{i_1,\dots,i_j}$ vanish on $\mathcal{I}$ while it gives the desired coefficient for tensors of the form (6.7.4).

Definition 6.7.1. If V is a real vector space of dimension n and q a quadratic form in V, then the *Clifford algebra* $\mathrm{Cl}(V,q)$ is defined to be the quotient of the full tensor algebra $\bigoplus_0^\infty\otimes^k V$ by the twosided ideal $\mathcal{I}$ generated by the elements $v\otimes v-q(v)\cdot 1$, $v\in V$. Elements in V are identified with their images in $\mathrm{Cl}(V,q)$, and the product of two elements x and y in $\mathrm{Cl}(V,q)$ is denoted $x\cdot y$.

We have already proved most of the following result:

Theorem 6.7.2. *If $\varepsilon_1,\dots,\varepsilon_n$ is a basis for V giving q the form* (6.7.2)*, then every element in the associative Clifford algebra* $\mathrm{Cl}(V,q)$ *can be written in one and only one way in the form* (6.7.4)*, with coefficients in* $\mathbf{R}$*, so the dimension of* $\mathrm{Cl}(V,q)$ *over* $\mathbf{R}$ *is* 2^n*. The images* $\mathrm{Cl}^0(V,q)$ *and* $\mathrm{Cl}^1(V,q)$ *in* $\mathrm{Cl}(V,q)$ *of the tensors of even (odd) rank* *are represented by sums* (6.7.4) *with j even (odd), so the dimensions are 2^{n-1}. We have*

$$\mathrm{Cl}^i(V,q)\cdot \mathrm{Cl}^j(V,q)\subset \mathrm{Cl}^{i+j \mod 2}(V,q),$$

which gives a $\mathbf{Z}_2$ grading of the algebra $\mathrm{Cl}(V,q)$*. The set* $\mathrm{Cl}^{[j]}(V,q)$ *of elements of the form* (6.7.4) *with j fixed* $\in\{0,\dots,n\}$ *is a linear subspace of* $\mathrm{Cl}(V,q)$*, equal to the linear span of all images of elements of the form*

$$\sum_\pi \operatorname{sgn}\pi v_{\pi(1)}\otimes\cdots\otimes v_{\pi(j)},\quad v_1,\dots,v_j\in V, \tag{6.7.5}$$

where π is any permutation of $1,\dots,j$*; thus it is independent of the choice of diagonalizing basis* $\varepsilon_1,\dots,\varepsilon_n$*. We have*

$$\mathrm{Cl}^0(V,q)=\bigoplus_{j\text{ even}}\mathrm{Cl}^{[j]}(V,q),\quad \mathrm{Cl}^1(V,q)=\bigoplus_{j\text{ odd}}\mathrm{Cl}^{[j]}(V,q).$$

There is a unique linear map $\mathrm{Cl}(V,q)\ni x\mapsto {}^tx\in \mathrm{Cl}(V,q)$ *(transposition) such that* ${}^t(x\cdot y)={}^ty\cdot{}^tx$ *and* ${}^tx=x$ *if* $x\in V$*. The constant term $Q(x)$ in ${}^tx\cdot x$ is equal to the constant term in $x\cdot{}^tx$ and defines invariantly a quadratic form Q on* $\mathrm{Cl}(V,q)$ *such that* $Q(1)=1$ *and*

$$Q(v)=q(v),\quad Q(x)=Q({}^tx),\quad Q(v\cdot x)=Q(x\cdot v)=q(v)Q(x),$$

if $v\in V$, $x\in\mathrm{Cl}(V,q)$*. If x is the class of* (6.7.4) *and* (6.7.2) *holds, then*

$$Q(x)=\sum_{j\le n} q_{i_1}\dots q_{i_j}a^2_{i_1\dots i_j}.$$

Proof. The transposition is inherited from the maps $v_1\otimes\cdots\otimes v_k\to v_k\otimes\cdots\otimes v_1$ in the tensor algebra, for they vanish on $\mathcal{I}$. Only the statement about $\mathrm{Cl}^{[j]}$ remains to be verified. If j is fixed in (6.7.4) we can extend the sum to all indices $i_1,\dots,i_j\in\{1,\dots,n\}$ provided that we divide by $j!$ and extend the definition of the coefficients in a skew symmetric way. Hence every such element is a linear combination of elements of the form (6.7.5). On the other hand, suppose that $v_i=\sum_{k=1}^n v_{i,k}\varepsilon_k$, $i=1,\dots,j$. Then the tensor (6.7.5) can be written

$$\sum_{k_1,\dots,k_j}\Big(\sum_\pi \operatorname{sgn}\pi v_{\pi(1),k_1}\dots v_{\pi(j),k_j}\Big)\varepsilon_{k_1}\otimes\cdots\otimes\varepsilon_{k_j},$$

and since the sum over π is skew symmetric in $k_1,\dots,k_j$, we have an element of the form (6.7.4) with j fixed.

If x is the class of the form (6.7.4) with a basis $\varepsilon_1,\dots,\varepsilon_n$ satisfying (6.7.2), then $x\cdot{}^tx$ and ${}^tx\cdot x$ both have the constant term $\sum a^2_{i_1\dots i_j}q_{i_1}\dots q_{i_j}$, for $\varepsilon_{i_1}\cdots\varepsilon_{i_j}\cdot\varepsilon_{k_1}\cdots\varepsilon_{k_l}$ can only be a constant if there is some $\nu\in\{i_1,\dots,i_j\}\cap\{k_1,\dots,k_l\}$ with $q_\nu=0$ or $\{i_1,\dots,i_j\}=\{k_1,\dots,k_l\}$. The definition of Q is obviously invariant since it makes no reference to the choice of a basis, and the other statements are obvious since for example ${}^t(v\cdot x)\cdot(v\cdot x) = {}^tx\cdot v\cdot v\cdot x$.

Note that $\mathrm{Cl}^{[j]}\cdot\mathrm{Cl}^{[k]}$ is not contained in $\mathrm{Cl}^{[j+k]}$ but may also have components in $\mathrm{Cl}^{[i]}$ with $i<j+k$.

Exercise 6.7.1. Show that when $v_1,\dots,v_4\in V$ we have with all products taken in $\mathrm{Cl}(V,q)$

$$[v_1\cdot v_2 - v_2\cdot v_1, v_3\cdot v_4 - v_4\cdot v_3] = 4\big(q(v_1,v_4)(v_2\cdot v_3 - v_3\cdot v_2) - q(v_1,v_3)(v_2\cdot v_4 - v_4\cdot v_2) + q(v_2,v_3)(v_1\cdot v_4 - v_4\cdot v_1) - q(v_2,v_4)(v_1\cdot v_3 - v_3\cdot v_1)\big).$$

Thus $\mathrm{Cl}^{[2]}(V,q)$ is a Lie algebra. Deduce that if q is positive definite and $\varepsilon_1,\dots,\varepsilon_n$ is an orthonormal basis, then

$$[\sum_{i,j} a_{ij}\varepsilon_i\cdot\varepsilon_j, \sum_{k,l} b_{k,l}\varepsilon_k\cdot\varepsilon_l] = 4\sum_{i,j}[a,b]_{ij}\varepsilon_i\cdot\varepsilon_j$$

if a and b are skew symmetric matrices. (Compare with the Lie algebra of $\mathbf{SO}(n)$.)

When $q=0$, the Clifford algebra becomes the exterior algebra $\wedge^*V$ on V. However, the case of interest to us here is primarily the case where q is positive definite. The negative definite case is often preferred but the difference is not important in the arguments below where we introduce the complexification of Cl, for all non-degenerate quadratic forms in a complex vector space are equivalent. However, before doing so we introduce an element in the Clifford algebra which will be very important later on.

Theorem 6.7.3. *If V is oriented and q is non-degenerate, then*

$$\nu = \varepsilon_1\cdots\varepsilon_n/\prod_1^n\sqrt{|q_j|} \tag{6.7.6}$$

is independent of the choice of positively oriented diagonalizing basis $\varepsilon_1,\dots,\varepsilon_n$ for V. We have

$$\nu^2 = (-1)^{\frac12 n(n-1)}\prod_1^n \operatorname{sgn} q_j,\quad \nu u = (-1)^{j(n-1)}u\nu,\quad \textit{if } u\in\mathrm{Cl}^{[j]}(V,q). \tag{6.7.7}$$

Proof. As in the proof of Theorem 6.7.2 we have

$$\nu = \frac{1}{n!}\sum_\pi \operatorname{sgn}\pi\varepsilon_{\pi(1)}\cdots\varepsilon_{\pi(n)}/\prod_1^n\sqrt{|q_j|}.$$

If $\tilde\varepsilon_i = \sum c_{i,k}\varepsilon_k$ is another diagonalizing positively oriented basis, then $\det(c_{ik})>0$ and

$$\begin{aligned} n!\prod_1^n\sqrt{|\tilde q_j|}\tilde\nu &= \sum_\pi \operatorname{sgn}\pi\tilde\varepsilon_{\pi(1)}\cdots\tilde\varepsilon_{\pi(n)} \\ &= \sum_{k_1,\dots,k_n}\sum_\pi \operatorname{sgn}\pi c_{\pi(1),k_1}\dots c_{\pi(n),k_n}\varepsilon_{k_1}\cdots\varepsilon_{k_n} \\ &= \det(c_{i,k})\sum \operatorname{sgn}(k_1,\dots,k_n)\varepsilon_{k_1}\cdots\varepsilon_{k_n} \end{aligned}$$

where all $k_1, \ldots, k_n$ are different in the last sum. Now we have

$$\sum q_j \xi_j^2 = q(\sum \xi_j \varepsilon_j) = q(\sum \tilde{\xi}_j \tilde{\varepsilon}_j) = \sum \tilde{q}_j \tilde{\xi}_j^2, \quad \xi_k = \sum \tilde{\xi}_j c_{j,k}.$$

Hence $\prod_1^n \tilde{q}_j = \prod_1^n q_j |\det(c_{i,k})|^2$ which proves that

$$\det(c_{i,k}) \prod_1^n \sqrt{|q_j|} = \prod_1^n \sqrt{|\tilde{q}_j|},$$

since the determinant is positive. Thus $\tilde{\nu} = \nu$. The proof of (6.7.7) is straightforward:

$$\nu^2 = \varepsilon_1 \cdots \varepsilon_n \varepsilon_1 \cdots \varepsilon_n \prod_1^n |q_j| = (-1)^{n-1+\cdots+1} \varepsilon_1^2 \cdots \varepsilon_n^2 / \prod_1^n |q_j|$$

$$= (-1)^{\frac{1}{2}n(n-1)} \prod_1^n \frac{q_j}{|q_j|},$$

$$\prod_1^n \sqrt{|q_j|} \nu \varepsilon_j = (-1)^{n-j} \varepsilon_1 \cdots \varepsilon_{j-1} q_j \varepsilon_{j+1} \cdots \varepsilon_n,$$

$$\prod_1^n \sqrt{|q_j|} \varepsilon_j \nu = (-1)^{j-1} \varepsilon_1 \cdots \varepsilon_{j-1} q_j \varepsilon_{j+1} \cdots \varepsilon_n.$$

Hence $\nu \varepsilon_j = (-1)^{n-1} \varepsilon_j \nu$, which immediately gives the second part of (6.7.7).

A complete description of the real Clifford algebras is fairly complicated (see Atiyah, Bott and Shapiro [1]) so we content ourselves with discussing the complexification $\mathrm{Cl}_{\mathbf{C}}(V, q)$ obtained by allowing the coefficients in (6.7.4) to be complex, or equivalently, replacing V by the complexification $V_{\mathbf{C}}$ in the definition by the tensor algebra. All non-degenerate quadratic forms in a complex vector space are equivalent under complex linear coordinate transformations, and we shall only discuss $\mathrm{Cl}_{\mathbf{C}}(n) = \mathrm{Cl}_{\mathbf{C}}(\mathbf{R}^n, e)$ where e is the Euclidean quadratic form in $\mathbf{R}^n$.

Theorem 6.7.4. *There are complex algebra isomorphisms*

$$\mathrm{Cl}_{\mathbf{C}}(2k) \cong \mathrm{End}(\mathbf{C}^{2^k}), \quad \mathrm{Cl}_{\mathbf{C}}(2k+1) \cong \mathrm{End}(\mathbf{C}^{2^k}) \oplus \mathrm{End}(\mathbf{C}^{2^k}). \tag{6.7.8}$$

Proof. We start with dimensions 1 and 2. The algebra $\mathrm{Cl}_{\mathbf{C}}(1)$ consists of all $(\alpha, \beta) \in \mathbf{C}^2$ with componentwise addition and

$$(\alpha, \beta)(\alpha', \beta') = (\alpha\alpha' + \beta\beta', \alpha\beta' + \beta\alpha').$$

Since

$$\alpha\alpha' + \beta\beta' \pm (\alpha\beta' + \beta\alpha') = (\alpha \pm \beta)(\alpha' \pm \beta'),$$

the map $(\alpha,\beta)\to(\alpha+\beta,\alpha-\beta)$ changes the operations in $\mathrm{Cl}_{\mathbf{C}}(1)$ to the coordinatewise operations in $\mathbf{C}\oplus\mathbf{C}$.

To determine $\mathrm{Cl}_{\mathbf{C}}(2)$ we define a linear map $\sigma:\mathbf{R}^2\to\mathrm{End}(\mathbf{C}^2)$ by

$$\sigma(\varepsilon_1)=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \sigma(\varepsilon_2)=\begin{pmatrix}0&i\\-i&0\end{pmatrix},$$

where $\varepsilon_1,\varepsilon_2$ are the basis vectors in $\mathbf{R}^2$. Since

$$\sigma(\varepsilon_1)^2=\sigma(\varepsilon_2)^2=I,\quad \sigma(\varepsilon_1)\sigma(\varepsilon_2)=-\sigma(\varepsilon_2)\sigma(\varepsilon_1)=\begin{pmatrix}0&i\\i&0\end{pmatrix},$$

it follows that σ extends to a bijective homomorphism $\mathrm{Cl}_{\mathbf{C}}(2)\to\mathrm{End}(\mathbf{C}^2)$.

For larger n we shall obtain (6.7.8) inductively if we prove that

$$\mathrm{Cl}_{\mathbf{C}}(n+2)\cong\mathrm{Cl}_{\mathbf{C}}(n)\otimes\mathrm{Cl}_{\mathbf{C}}(2),\quad n\ge 1. \tag{6.7.9}$$

To prove (6.7.9) we define a linear map $\sigma:\mathbf{R}^{n+2}\to\mathrm{Cl}_{\mathbf{C}}(n)\otimes\mathrm{Cl}_{\mathbf{C}}(2)$ by

$$\sigma(\varepsilon_j)=\begin{cases} i\varepsilon_j\otimes\varepsilon_{n+1}\cdot\varepsilon_{n+2}, & \text{if } j\le n,\\ 1\otimes\varepsilon_j, & \text{if } j=n+1,n+2.\end{cases}$$

Since $\sigma(\varepsilon_j)^2=1\otimes 1$ for $j\le n+2$ and $\sigma(\varepsilon_j)\sigma(\varepsilon_k)+\sigma(\varepsilon_k)\sigma(\varepsilon_j)=0$ if $j\ne k$, the map σ induces a homomorphism $\tilde\sigma:\mathrm{Cl}_{\mathbf{C}}(n+2)\to\mathrm{Cl}_{\mathbf{C}}(n)\otimes\mathrm{Cl}_{\mathbf{C}}(2)$. The range of $\tilde\sigma$ is a subalgebra containing $1\otimes\mathrm{Cl}_{\mathbf{C}}(2)$ and $\mathrm{Cl}_{\mathbf{C}}(n)\otimes 1$, so $\tilde\sigma$ is surjective, henced bijective because the dimensions agree. Now the isomorphisms (6.7.8) follow inductively since for arbitrary (complex) vector spaces E and F we have

$$(\mathrm{End}\,E)\otimes(\mathrm{End}\,F)\cong\mathrm{End}(E\otimes F).$$

In fact, $\mathrm{End}\,E\cong E\otimes E^*$, $\mathrm{End}\,F\cong F\otimes F^*$, $\mathrm{End}(E\otimes F)\cong E\otimes F\otimes E^*\otimes F^*$, and the resulting isomorphism ι respects the product structure for if $S_1,S_2\in\mathrm{End}(E)$, $T_1,T_2\in\mathrm{End}\,F$, then $\iota(S_1\otimes T_1)$ times $\iota(S_2\otimes T_2)$ and $\iota(S_1S_2\otimes T_1T_2)$ both map $e\otimes f\in E\otimes F$ to $(S_1S_2e)\otimes(T_1T_2f)$. The proof is complete.

If we identify $\mathbf{R}^{2k}$ with $\mathbf{C}^k$, we get a natural isomorphism

$$\mu:\mathrm{Cl}_{\mathbf{C}}(2k)\cong\mathrm{End}(\wedge^*\mathbf{C}^k), \tag{6.7.10}$$

where $\wedge^*\mathbf{C}^k$ denotes the full complex exterior algebra over $\mathbf{C}^k$. (Since $\wedge^*\mathbf{C}^k$ has complex dimension 2^k, the two sides of (6.7.10) are isomorphic by Theorem 6.7.4.) To define the isomorphism we must recall a construction made in Section 6.2. For $w\in W=\mathbf{C}^k$ we denote by T_w the exterior multiplication $T_wu=w\wedge u$, $u\in\wedge^*W$, and by T_w^* we denote the adjoint with respect to the natural hermitian scalar product in $\wedge^*W$. If w is the basis vector $e_1\in W$, then

$$T_wu=e_1\wedge u_0,\quad \text{if } u=u_0+e_1\wedge u_1,$$

where u_0, u_1 are in the subalgebra generated by $e_2, \dots, e_k$. If u is of degree ν and $v = v_0 + e_1 \wedge v_1$ is of degree $\nu + 1$, then

$$(T_w u, v) = (e_1 \wedge u_0, v_0 + e_1 \wedge v_1) = (e_1 \wedge u_0, e_1 \wedge v_1) = (u_0, v_1),$$

which means that $T_w^* v = v_1$. Hence

$$\|T_w v\|^2 + \|T_w^* v\|^2 = \|e_1 \wedge v_0\|^2 + \|v_1\|^2 = \|v_0\|^2 + \|v_1\|^2 = \|v\|^2.$$

By unitary invariance and homogeneity we conclude that

$$T_w^* T_w + T_w T_w^* = |w|^2 I, \quad w \in W,$$

where I is the identity in $\wedge^* W$, and $|w|^2$ is the square of the norm of w as an element in $\mathbf{C}^k$, or equivalently as an element in $\mathbf{R}^{2k}$. Since $T_w^2 = 0$ and $T_w^{*\,2} = 0$, it follows that

$$\mu(w)^2 = |w|^2 I, \quad \text{if } \mu(w) = T_w + T_w^*. \tag{6.7.11}$$

It is clear that μ is real linear, so we can extend μ first to an algebra homomorphism $\mathrm{Cl}(2k) \to \mathrm{End}(\wedge^* \mathbf{C}^k)$, and then to a complex algebra homomorphism (6.7.10). It is actually an isomorphism, for both sides are isomorphic to $\mathrm{End}(\mathbf{C}^{2^k})$, by Theorem 6.7.4, and the kernel of μ must vanish by the following well known lemma:

Lemma 6.7.5. *If E is a vector space and J is a two sided ideal in* $\mathrm{End}(E)$*, then* $J = \{0\}$ *or* $J = \mathrm{End}(E)$.

Proof. Let $E = \mathbf{C}^N$ and $T \in J$, $\operatorname{rank} T = r > 0$. The product $ATB \in J$ consists of an invertible matrix T_1 in the upper left $r \times r$ corner with zeros elsewhere, if A, B are non-singular, B maps the last $N - r$ basis vectors to the kernel of T and A maps the range to the plane spanned by the first r basis vectors. Multiplying once more to the left we make T_1 equal to the identity. Adding such matrices for different choices of basis vectors we conclude that the identity is in J, so $J = \mathrm{End}(E)$.

Summing up, we have proved:

Theorem 6.7.6. *The complex extension of μ, defined by* (6.7.11) *gives a natural isomorphism* (6.7.10).

Recall that our purpose is to find linear maps $V \to \mathrm{End}(E)$ satisfying (6.7.1). We have found that the existence of such a map means precisely that E is a $\mathrm{Cl}(V, q)$ module. Theorem 6.7.6 gives an example, for the map μ makes $\wedge^*(\mathbf{C}^k)$ a $\mathrm{Cl}(\mathbf{R}^{2k}, e)$ module. To define an associated vector bundle on a Riemannian manifold we would need to have a representation of $\mathbf{O}(2k)$ on $\wedge^*(\mathbf{C}^k)$. To define the appropriate representations we digress to discuss the relation between the orthogonal group and the Clifford algebra.

Proposition 6.7.7. *If q is a positive definite quadratic form in the real vector space* $\mathbf{R}$*, then*

$$\mathrm{Pin}(V, q) = \{v_1 \cdots v_j; v_i \in V, q(v_i) = 1, \ \textit{for } i = 1, \dots, j\} \subset \mathrm{Cl}(V, q),$$

$$\mathrm{Spin}(V, q) = \{v_1 \cdots v_{2j}; v_i \in V, q(v_i) = 1, \ \textit{for } i = 1, \dots, 2j\} = \mathrm{Pin}(V, q) \cap \mathrm{Cl}^0(V, q),$$

are multiplicative groups.

Proof. The transpose $v_j \cdots v_1$ of $v_1 \cdots v_j$ is an inverse.

If $v \in V$, $q(v) = 1$, then

$$v \cdot x \cdot v = v \cdot (-v \cdot x + 2q(v,x)) = -x + 2q(v,x)v, \quad x \in V.$$

Now $x - 2q(v,x)v$ is the orthogonal reflection of x in the plane orthogonal to v, so $x \mapsto v \cdot x \cdot v$ is this reflection followed by reflection in the origin. If we define

$$u^* = (-1)^j {}^t u \quad \text{if } u \in \mathrm{Cl}^j(V,q),\ j = 0,1,$$

thus $(v_1 \cdots v_l)^* = (-1)^l v_l \cdots v_1,\ v_1, \dots, v_l \in V$, it follows if $u \in \mathrm{Pin}(V,q)$ that

$$V \ni x \mapsto u \cdot x \cdot u^*, \quad x \in V, \tag{6.7.12}$$

is a product of orthogonal reflections in V, and it preserves orientation if and only if $u \in \mathrm{Spin}(V,q)$. Hence (6.7.12) defines a homomorphism

$$\tau : \mathrm{Pin}(V,q) \to \mathbf{O}(V,q),$$

such that $\mathrm{Spin}(V,q)$ is the inverse image of $\mathbf{SO}(V,q)$. The maps $\mathrm{Pin}(V,q) \to \mathbf{O}(V,q)$ and $\mathrm{Spin}(V,q) \to \mathbf{SO}(V,q)$ are surjective by the following lemma, which is very close to Exercise 1.4.3:

Lemma 6.7.8. *Every element in* $\mathbf{O}(n)$ *is a product of at most* n *orthogonal reflections in* $\mathbf{R}^n$.

Proof. We can extend an orthogonal transformation O in $\mathbf{R}^n$ to a unitary transformation in $\mathbf{C}^n$. The projection in $\mathbf{R}^n$ of an eigenvector in $\mathbf{C}^n$ is an invariant subspace V_1 for O of dimension 1 or 2, and the orthogonal space $V_1^\perp$ is also invariant since

$$(OV_1^\perp, V_1) = (V_1^\perp, O^{-1}V_1) = 0.$$

Hence it suffices to prove the lemma in dimension 1 and dimension 2. In the first case an orthogonal transformation is a reflection or the identity, in the second case it is a reflection or a rotation by an angle θ and therefore the product of reflections in two lines with an angle $\theta/2$ between them.

Theorem 6.7.9. *If* q *is a positive definite quadratic form in* V*, then the homomorphisms* $\tau : \mathrm{Pin}(V,q) \to \mathbf{O}(V,q)$ *and* $\tau : \mathrm{Spin}(V,q) \to \mathbf{SO}(V,q)$ *are surjective with kernel* $\{\pm 1\}$*. Both* $\mathrm{Pin}(V,q)$ *and* $\mathrm{Spin}(V,q)$ *are compact subsets of* $\mathrm{Cl}(V,q)$*,* $\mathrm{Spin}(V,q)$ *is arcwise connected and so is* $\mathrm{Pin}(V,q) \setminus \mathrm{Spin}(V,q)$*. If* $\dim V \geq 3$ *then* $\mathrm{Spin}(V,q)$ *is simply connected.*

Proof. We have already proved surjectivity, and it is clear that ± 1 is in the kernel. Suppose that $u \in \mathrm{Pin}(V,q)$ and that $u \cdot x \cdot u^* = x$, $x \in V$. Then $u \in \mathrm{Spin}(V,q)$ since $\tau(u)$ preserves the orientation, so $u^* \cdot u = 1$, hence $u \cdot x = x \cdot u$ and

$$x \cdot u \cdot x = q(x)u.$$

With an orthonormal basis $\varepsilon_1, \dots, \varepsilon_n$ in V, we can write (see (6.7.4)

$$u = \sum a_{i_1,\dots,i_j} \varepsilon_{i_1} \cdots \varepsilon_{i_j}, \quad i_1 < \cdots < i_j, \quad j \text{ even}.$$

If we take $x = \varepsilon_l$ and note that $\varepsilon_l \cdot \varepsilon_{i_1} \cdots \varepsilon_{i_j} \cdot \varepsilon_l = \pm \varepsilon_{i_1} \cdots \varepsilon_{i_j}$ with the minus sign when $l \in \{i_1, \dots, i_j\}$, we conclude that only a term with $j = 0$ can occur. This means that u is a real number, and since $uxu = x$, $x \in V$, we have $u = \pm 1$.

The compactness follows from the bound for the number of reflections in Lemma 6.7.8. To prove that $\mathrm{Spin}(V, g)$ is connected it suffices to note that $v_1 \cdots v_{2k}$, where $v_j \in V$ and $q(v_j) = 1$, is connected to $v_1 \cdot v_1 \cdots v_1 = 1$ if each v_j is connected to v_1 in the connected unit sphere in V.

What remains is to prove that $\mathrm{Spin}(V, q)$ is simply connected if $\dim V \geq 3$. In the proof we shall need that τ is a local homeomorphism. It suffices to prove that at ± 1. Take disjoint compact neighborhoods $U_\pm$ of ± 1 in $\mathrm{Spin}(V, q)$ with $U_- = -1 \cdot U_+$ and, using the compactness, a compact neighborhood U of the identity in $\mathbf{SO}(V, q)$ such that $\tau^{-1} U \subset U_+ \cup U_-$. By the results already proved it follows that τ maps the neighborhoods $U'_\pm = U_\pm \cap \tau^{-1} U$ of ± 1 bijectively, hence homeomorphically, on U. For the proof we shall also need the following elementary lemma:

Lemma 6.7.10. *Every closed loop in $\mathbf{SO}(n)$, $n \geq 2$, is homotopic to a loop in $\mathbf{SO}(2)$, embedded in $\mathbf{SO}(n)$ as $\mathbf{SO}(2) \times I_{n-2}$.*

Proof. Let $\mathbf{R}/\mathbf{Z} \ni t \mapsto O(t) \in \mathbf{SO}(n)$ be the given loop, $n \geq 3$. If $\varepsilon_n = (0, \dots, 0, 1)$ is left fixed by $O(t)$ for every t, then $O(t) \in \mathbf{SO}(n-1) \times 1$, and the lemma follows by induction from lower dimensions. In the general case we first regularize $O(t)$ to a C^1 map. Then the curve $\{O(t)\varepsilon_n; t \in \mathbf{R}/\mathbf{Z}\}$ is of measure 0 in S^{n-1}, so we can choose $\xi \in S^{n-1}$ such that $\pm \xi \notin \{O(t); t \in \mathbf{R}/\mathbf{Z}\}$. Let $O_1 \in \mathbf{SO}(n)$ map ξ to ε_n. Our loop is homotopic to the loop $t \mapsto O_1 O(t) = \widetilde{O}(t)$, and $\widetilde{O}(t)\varepsilon_n \neq \pm \varepsilon_n$ for every t. Hence

$$\widetilde{O}(t)\varepsilon_n = \varepsilon_n \cos\theta(t) + \xi(t) \sin\theta(t)$$

with uniquely determined continuous $\theta(t) \in (0, \pi)$ and $\xi(t) \in S^{n-1}$ orthogonal to ε_n.Denote by $O_2(t, s)$ the rotation by the angle $-s\theta(t)$ in the $\varepsilon_n \xi(t)$ plane. Then

$$(\mathbf{R}/\mathbf{Z}) \times [0, 1] \ni (t, s) \mapsto O_2(t, s)\widetilde{O}(t)$$

is a homotopy connecting the loop $t \mapsto \widetilde{O}(t)$ to the loop $t \mapsto O_2(t, 1)\widetilde{O}(t)$ which leaves ε_n fixed. As observed at the beginning of the proof, the lemma is then proved by induction.

End of proof of Theorem 6.7.9. Let $\mathbf{R}/\mathbf{Z} \ni t \mapsto x(t) \in \mathrm{Spin}(\mathbf{R}^n, e)$ be a closed loop. Then $t \mapsto \tau(x(t))$ is a closed loop in $\mathbf{SO}(n)$, so by Lemma 6.7.10 there is a homotopy to a loop in $t \mapsto O(t) \times I_{n-2}$ where $O(t) \in \mathbf{SO}(2)$ consists of rotation in $\mathbf{R}^2$ by the angle $2\pi kt$. We can lift the homotopy to one in $\mathrm{Spin}(\mathbf{R}^n, e)$ connecting the given loop to the loop

$$t \mapsto (1, 0, \dots, 0) \cdot (\cos(\pi kt), -\sin(\pi kt), 0, \dots, 0).$$

Since it is closed it follows that k is even. When $n \geq 3$ a homotopy to a constant loop is given by

$$(t,s) \mapsto (1,0,\dots,0)\cdot(\cos(\tfrac{1}{2}\pi s)\cos(\pi kt), -\cos(\tfrac{1}{2}\pi s)\sin(\pi kt), \sin(\tfrac{1}{2}\pi s), 0,\dots,0)$$

where $t \in \mathbf{R}/\mathbf{Z}$, $s \in [0,1]$. Hence $\operatorname{Spin}(\mathbf{R}^n, e)$ is simply connected.

Remarks. 1. Comparison of the result with Exercise 1.4.6 shows that $\operatorname{Spin}(\mathbf{R}^3, e) \cong \mathbf{SU}(2)$. We could also have proved that $\operatorname{Spin}(\mathbf{R}^n, e)$ is simply connected starting from this fact.

2. Any closed loop $\gamma : \mathbf{R}/\mathbf{Z} \ni t \mapsto O(t) \in \mathbf{SO}(n)$, $n \geq 3$, can be lifted to an arc $[0,1] \ni t \mapsto x(t) \in \operatorname{Spin}(\mathbf{R}^n, e)$, and $x(1) = \pm x(0)$ by Theorem 6.7.9. If $x(1) = x(0)$ we have a closed loop in $\operatorname{Spin}(\mathbf{R}^n, e)$, so it is homotopic to a point, hence γ is trivial. However, if $x(1) = -x(0)$ we do not have a closed loop in $\operatorname{Spin}(\mathbf{R}^n, e)$, and this remains true for the lifting of any loop homotopic to γ. This proves that the fundamental group of $\mathbf{SO}(n)$ is $\mathbf{Z}_2$, with the element $\neq 0$ represented by the projection of any arc in $\operatorname{Spin}(\mathbf{R}^n, e)$ connecting 1 to -1. The proof of Theorem 6.7.9 would have been shorter if we had assumed the fundamental group of $\mathbf{SO}(n)$ known, but we have chosen to determine it at the same time.

3. Since $\mathbf{SO}(2) \cong S(1) \cong \operatorname{Spin}(\mathbf{R}^2, e)$, the fundamental group is equal to $\mathbf{Z}$ and the correspondence is given by the winding number of a loop, when $n = 2$.

Using Exercise 1.4.5 it is easy to see that $\operatorname{Spin}(V,q)$ and $\operatorname{Pin}(V,q)$ are analytic manifolds in $\operatorname{Cl}(V,q)$. However, we prefer to give a direct proof which also gives an explicit correspondence between the Lie algebras of $\operatorname{Spin}(V,q)$ and $\mathbf{SO}(V,q)$, which of course are isomorphic since the groups are locally isomorphic.

Theorem 6.7.11. *If $A = (a_{jk})$ is a real skew symmetric $n \times n$ matrix, and $\varepsilon_1, \dots, \varepsilon_n$ is an orthonormal basis in V, then*

$$\exp\Big(\sum_{j,k=1}^{n} a_{jk}\varepsilon_j \cdot \varepsilon_k\Big), \tag{6.7.13}$$

defined by the Taylor expansion which converges in $\operatorname{Cl}(V,q)$, is an element in $\operatorname{Spin}(V,q)$. The image in $\mathbf{SO}(V,q)$ has the matrix $\exp(4A)$ in the basis $\varepsilon_1, \dots, \varepsilon_n$. Thus (6.7.13) gives a bijection of a neighborhood of 0 in the space of skew symmetric matrices on a neighborhood of the identity in $\operatorname{Spin}(V,q)$, which is therefore an analytic manifold with tangent space $\operatorname{Cl}^{[2]}(V,q)$ at the identity, and

$$\mathfrak{so}(n) \ni 4(a_{jk})_{j,k=1}^{n} \mapsto \sum_{j,k=1}^{n} a_{jk} e_j e_k \in \operatorname{Cl}^{[2]}(V,q)$$

is the isomorphism of Lie algebras corresponding to the local isomorphism of $\mathbf{SO}(V,q)$, identified with $\mathbf{SO}(n)$, and $\operatorname{Spin}(V,q)$.

Note that this explains Exercise 6.7.1, including the factor 4.

Proof. Choose a norm $\|\cdot\|$ in $\mathrm{Cl}(V,q)$ such that $\|x\cdot y\| \le \|x\|\|y\|$. The formal expansion $\exp x = \sum_0^\infty x^\nu/\nu!$ converges for every $x \in \mathrm{Cl}(V,q)$, for the norm of the term is bounded by $\|x\|^\nu/\nu!$. The sum is an analytic function of x, and we obtain $\exp x \cdot \exp(-x) = \exp 0 = 1$ by rearranging the terms in the product. For small $t \in \mathbf{R}$ let

$$P(t) = \prod_{i,j=1}^{n} \cdot P_{ij}(t), \quad P_{ij}(t) = \varepsilon_i \cdot (\sqrt{1 - a_{ij}^2 t^2}\varepsilon_i + a_{ij} t \varepsilon_j),$$

with the product taken say in the lexicographical order. We have $P_{jj} = 1$, $P_{ij}(t) \in \mathrm{Spin}(V,q)$ for any i, j, and

$$P_{ij}(t) = 1 + a_{ij} t \varepsilon_i \cdot \varepsilon_j + O(t^2), \quad \text{hence } P(t) = 1 + tx + O(t^2),\ x = \sum_{i,j=1}^{n} a_{ij}\varepsilon_i \cdot \varepsilon_j,$$

which implies that $P(1/\nu)^\nu \to \exp x$ as $\nu \to \infty$, for

$$\|P(1/\nu)^\nu - (1 + x/\nu)^\nu\| \le (1 + \|x\|/\nu)^\nu((1 + O(1/\nu^2))^\nu - 1) \to 0,$$
$$(1 + x/\nu)^\nu = \sum_0^\nu x^\mu (1 - 1/\nu) \dots (1 - (\mu - 1)/\nu)/\mu! \to \exp x.$$

Hence $\exp x \in \mathrm{Spin}(V,q)$. To find the corresponding orthogonal transformation we note that by (6.7.12)

$$\tau(P_{ij}(t))\varepsilon_k = P_{ij}(t)\varepsilon_k P_{ij}(t)^* = \varepsilon_k + \begin{cases} 0, & \text{if } k \ne i, k \ne j \\ -2a_{kj} t \varepsilon_j + O(t^2), & \text{if } k = i \ne j \\ 2a_{ik} t \varepsilon_i + O(t^2), & \text{if } k = j \ne i \end{cases}, \quad \text{hence}$$
$$\tau(P(t))\varepsilon_k = \varepsilon_k - 2\sum_j a_{kj} t \varepsilon_j + 2\sum_i a_{ik} t \varepsilon_i + O(t^2) = \varepsilon_k + 4t\sum_i a_{ik}\varepsilon_i + O(t^2).$$

This implies that $\tau(P(1/\nu)^\nu) = (\tau(P(1/\nu)))^\nu \to e^{4A}$ where A is the skew symmetric operator in V with $A\varepsilon_k = \sum_i a_{ik}\varepsilon_i$, that is, with matrix (a_{jk}) in the basis $\varepsilon_1, \dots, \varepsilon_n$. The proof is complete.

From now on we shall use the abbreviated notation $\mathrm{Pin}(n)$ and $\mathrm{Spin}(n)$ for $\mathrm{Pin}(\mathbf{R}^n, e)$ and $\mathrm{Spin}(\mathbf{R}^n, e)$. The composition

(6.7.14) $$\varrho : \mathrm{Pin}(2k) \to \mathrm{Cl}(2k) \to \mathrm{Cl}_{\mathbf{C}}(2k) \xrightarrow{\mu} \mathrm{End}(\wedge^* \mathbf{C}^k))$$

with μ as in (6.7.10) is a representation since it is a homomorphism mapping 1 to the identity. It is not quite the desired representation of $\mathbf{O}(2k)$, but just a projective representation, determined up to a factor ± 1; we shall come back to this point in Section 6.8.

Proposition 6.7.12. *The representation* ϱ *of* $\mathrm{Pin}(2k)$ *defined by* (6.7.14) *is unitary and irreducible.*

Proof. That ϱ is unitary follows from (6.7.11), for $\mu(v)$ is self-adjoint with $\mu(v)^2 = 1$ if $v \in V$ and $|v| = 1$. Irreducibility is a consequence of the fact that $\mathrm{Pin}(2k)$ generates the algebra $\mathrm{Cl}(2k)$, for μ is an isomorphism.

The map $\mu(v)$ in (6.7.11) takes forms of even (odd) degree to forms of odd (even) degree. Hence the degree of $\mu(z)w$ has the same parity as $w \in \wedge^*(\mathbf{C}^k)$ if $z \in \mathrm{Cl}^0(2k)$ while it is opposite if $z \in \mathrm{Cl}^1_{\mathbf{C}}(2k)$. This means that the restriction of ϱ to $\mathrm{Spin}(2k)$ splits into the direct sum of two representations

(6.7.15) $$D^{\pm}_{\frac{1}{2}} : \mathrm{Spin}(2k) \to \mathrm{Aut}(S_{\pm}(2k)), \quad S_+(2k) = \wedge^{\mathrm{even}}(\mathbf{C}^k),\ S_-(2k) = \wedge^{\mathrm{odd}}(\mathbf{C}^k),$$

called the *half-spin representations.* One calls (6.7.14) the spin representation.

Proposition 6.7.13. *The half-spin representations are unitary and irreducible.*

Proof. If $W \subset S_+(2k)$ is an invariant subspace for $D^+_{\frac{1}{2}}$, then $W \oplus (\varrho(\varepsilon_1)W) \subset \wedge^*(\mathbf{C}^k)$ is an invariant subspace for ϱ, so $W = \{0\}$ or $W = S_+(2k)$ by Proposition 6.7.11, which also shows that the representations are unitary.

In the odd dimensional case we can use the isomorphism $\mathrm{Cl}_{\mathbf{C}}(2k-1) \to \mathrm{Cl}^0_{\mathbf{C}}(2k)$ mapping $v \in \mathbf{R}^{2k-1}$ to $iv \cdot \varepsilon_{2k}$, where ε_{2k} is the last basis vector in $\mathbf{R}^{2k}$. This means that the elements in $C^{[j]}_{\mathbf{C}}(2k-1)$ are multiplied by i^{j^2}, followed by right multiplication by ε_{2k} if j is odd. Hence we get an inclusion $\mathrm{Spin}(2k-1) \to \mathrm{Spin}(2k)$, so (6.7.15) gives a representation

$$\mathrm{Spin}(2k-1) \to \mathrm{Spin}(2k) \xrightarrow{D^+_{\frac{1}{2}}} \mathrm{Aut}(S_+(2k)),$$

which is also denoted by $D^+_{\frac{1}{2}}$. Similarly we get a representation $D^-_{\frac{1}{2}}$ in $S_-(2k)$, but since

$$D^-_{\frac{1}{2}}\mu(\varepsilon_{2k}) = \mu(\varepsilon_{2k})D^+_{\frac{1}{2}} \quad \text{on } S_+(2k),$$

it is an equivalent representation.

For reasons analogous to Proposition 6.5.1 we shall have to calculate the difference between the traces of an endomorphism K_+ of $S_+(2k)$ and one K_- of $S_-(2k)$. We combine them to

$$K = \begin{pmatrix} K_+ & 0 \\ 0 & K_- \end{pmatrix} \in \mathrm{End}(S(2k)) = \mathrm{End}(\wedge^*(\mathbf{C}^k)) \cong \mathrm{Cl}_{\mathbf{C}}(2k),$$

and we write

(6.7.16) $$\operatorname{Str} K = \operatorname{Tr} K_+ - \operatorname{Tr} K_-.$$

One calls $\operatorname{Str} K$ the *supertrace* of K; it is defined for any endomorphism in a space S with a given decomposition $S = S_+ \oplus S_-$.

Theorem 6.7.14 (The Berezin-Patodi formula). *If $z \in \mathrm{Cl}_{\mathbf{C}}(2k)$ and ν is the "volume element" $\varepsilon_1 \cdots \varepsilon_{2k}$ (as in (6.7.6)), then the component of z in $\mathrm{Cl}^{[2k]}_{\mathbf{C}}$ is $((2i)^{-k} \operatorname{Str} \mu(z))\nu$.*

Before the proof we shall discuss some properties of ν in addition to those given in Theorem 6.7.3. In doing so we must make our identification of $\mathbf{R}^{2k}$ with $\mathbf{C}^k$ explicit. We shall use the identification

$$\mathbf{R}^{2k} \ni (\xi_1, \dots, \xi_{2k}) \mapsto (\xi_1 + i\xi_2, \dots, \xi_{2k-1} + i\xi_{2k}) \in \mathbf{C}^k$$

which is customary in complex analysis.

Lemma 6.7.15. *The center of* $\mathrm{Spin}(2k)$ *is equal to* $\{\pm 1, \pm\nu\}$*. The restriction of* $\varrho(\nu)$ *to* $S_\pm(2k)$ *is equal to* $\pm i^k$*.*

Proof. The image $\tau(x)$ in $\mathbf{SO}(2k)$ of an element x in the center of $\mathrm{Spin}(2k)$ commutes with every $O \in \mathbf{SO}(2k)$. For every $\omega \in S^{2k-1}$ it follows that $O(\tau(x)\omega) = \tau(x)O\omega = \tau(x)\omega$, if $O \in \mathbf{SO}(2k)$ leaves ω fixed. Hence $\tau(x)\omega = \pm\omega$, and it follows that $\tau(x)$ is $\pm$ the identity. By Theorem 6.7.9 we conclude that there are at most four elements in the center. From the second part of (6.7.7) it follows that $\nu u = u\nu$ if $u \in \mathrm{Spin}(2k)$, so $\pm 1, \pm\nu$ are in the center and it can contain no other elements.

Since $\nu \in \mathrm{Spin}(2k)$ we know that $S_\pm(2k)$ is invariant under $\varrho(\nu)$, and since the restriction of ϱ to $\mathrm{Spin}(2k)$ is an irreducible representation on $S_\pm(2k)$ commuting with $\varrho(\nu)$, it follows that $\varrho(\nu)$ is a constant in each of these spaces. From (6.7.7) we know that $\varrho(\nu)^2 = \varrho(\nu^2) = (-1)^k$, so the constant values must be $\pm i^k$. By (6.7.7) we have $\varepsilon_1 \cdot \nu = -\nu \cdot \varepsilon_1$, hence $\varrho(\varepsilon_1)\varrho(\nu) = -\varrho(\nu)\varrho(\varepsilon_1)$ which proves that we have opposite signs in $S_\pm(2k)$. To find the sign in $S_+(2k)$ it suffices to calculate $\varrho(\nu)1$. Recall that $\varepsilon_1, \dots, \varepsilon_{2k}$ are the basis vectors in $\mathbf{R}^{2k}$, and that we have identified ε_{2j-1} with e_j and ε_{2j} with ie_j, if $e_1, \dots, e_k$ are the basis vectors in $\mathbf{C}^k$. By induction for decreasing j it follows from (6.7.11) that

$$\varrho(\varepsilon_{2j} \cdots \varepsilon_{2k})1 = i^{k-j+1} e_j \wedge \cdots \wedge e_k,$$

for $T^*_{e_{j-1}}$ annihilates the right-hand side. Starting from the case $j = 1$ we obtain by induction for increasing j

$$\varrho(\varepsilon_{2j-1} \cdots \varepsilon_1 \cdot \varepsilon_2 \cdots \varepsilon_{2k})1 = i^k e_{j+1} \wedge \cdots \wedge e_k,$$

for $T_{e_{j+1}}$ annihilates the right hand side while $T^*_{e_{j+1}}$ removes the factor e_{j+1}. When $j = k$ we obtain $\varrho(\nu)1 = i^k$, for

$$\varepsilon_{2k-1} \cdots \varepsilon_1 \cdot \varepsilon_2 \cdots \varepsilon_{2k} = \varepsilon_1 \cdot \varepsilon_2 \cdot \varepsilon_3 \cdot \varepsilon_4 \cdots \varepsilon_{2k-1} \cdot \varepsilon_{2k}$$

since $\varepsilon_1 \cdot \varepsilon_2, \varepsilon_3 \cdot \varepsilon_4, \dots$ commute with the other factors which allows us to to move them out to the left starting from the middle. The lemma is proved.

Proof of Theorem 6.7.14. From Lemma 6.7.15 it follows that

$$i^k \operatorname{Str} \mu(z) = \operatorname{Tr}(\mu(\nu)\mu(z)) = \operatorname{Tr} \mu(\nu z). \tag{6.7.17}$$

If we set $w = \nu z$, then $z = (-1)^k \nu w$ by (6.7.7), so the component of z in $\mathrm{Cl}_{\mathbf{C}}^{[2k]}$ is $(-1)^k w_0 \nu$ where w_0 is the constant term in w. We shall prove that

$$\operatorname{Tr}\mu(w) = 2^k w_0. \tag{6.7.18}$$

In view of (6.7.17) this will show that $(-1)^k w_0 = (-2)^{-k} \operatorname{Tr}\mu(w) = (2i)^{-k} \operatorname{Str}\mu(z)$, which is the Berezin-Patodi formula. To prove (6.7.18) we note first that $\mu(1)$ is the identity in a space of dimension 2^k, so the trace is 2^k. We also have to show that for $j = 1, \dots, 2k$ we have $\operatorname{Tr}\mu(w) = 0$ if

$$w = \varepsilon_{i_1} \cdots \varepsilon_{i_j}, \quad 1 \le i_1 < \cdots < i_j \le 2k.$$

If j is even we write $w = \varepsilon_{i_1} \cdot w_1 = -w_1 \cdot \varepsilon_{i_1}$ and obtain

$$\operatorname{Tr}\mu(w) = \operatorname{Tr}(\mu(\varepsilon_{i_1})\mu(w_1)) = \operatorname{Tr}(\mu(w_1)\mu(\varepsilon_{i_1})) = -\operatorname{Tr}\mu, \quad \text{hence } \operatorname{Tr}\mu(w) = 0,$$

for $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. If j is odd we choose $i \in \{1, \dots, 2k\} \setminus \{i_1, \dots, i_j\}$. Then we have $w = \varepsilon_i \cdot \varepsilon_i \cdot w = -\varepsilon_i \cdot w \cdot \varepsilon_i$, and it follows that

$$\operatorname{Tr}\mu(w) = \operatorname{Tr}(\mu(\varepsilon_i)^2 \mu(w)) = \operatorname{Tr}(\mu(\varepsilon_i)\mu(w)\mu(\varepsilon_i)) = -\operatorname{Tr}\mu(w),$$

hence $\operatorname{Tr}\mu(w) = 0$. The proof is complete.

The important feature of the Berezin-Patodi formula is that the information about the supertrace is contained in the highest part of the Clifford algebra. As we shall see in Section 6.10, the parametrix construction will show that the jth term has no component above the level $\mathrm{Cl}_{\mathbf{C}}^{[2j]}$, and that the term there can be calculated for $j \le k$; we shall precisely need the kth term to compute an index.

6.8. Clifford and spinor analysis. Let M be a C^∞ Riemannian manifold of dimension n. For every $x \in M$ we have defined in Section 6.7 a Clifford algebra $\mathrm{Cl}_x(M) = \mathrm{Cl}(T_x^*, g_x^*)$, where g_x^* is the quadratic form in T_x^*M dual to the metric form g_x in T_xM. It is clear that $\mathrm{Cl}_x(M)$ is the fiber at x of a vector bundle $\mathrm{Cl}(M)$, which is a quotient of $\oplus_{j \le n} \otimes^j T^*M$. We can also view $\mathrm{Cl}(M)$ as a bundle associated to the orthonormal frame bundle and the representation of the orthogonal group on $\mathrm{Cl}(\mathbf{R}^n, e)$ induced by the natural action of $\mathbf{O}(n)$ on $\mathbf{R}^n$. In either way we see that the Levi-Civita connection is defined in $\mathrm{Cl}(M)$.

As motivated in Section 6.6, Dirac operators involve bundles with a special structure:

Definition 6.8.1. A (complex) vector bundle E on M is called a Clifford bundle if for every $x \in M$ we have a map

$$\mathrm{Cl}_x(M) \times E_x \ni (v, \varphi) \mapsto v \cdot \varphi \in E_x$$

which makes E_x a $\mathrm{Cl}_x(M)$ module, and the map $\mathrm{Cl}(M) \oplus E \to E$ is in C^∞. We shall say that E is a *Hermitian Clifford bundle* if E is provided with a Hermitian metric such that the map $\sigma(v) : \varphi \mapsto v \cdot \varphi$ in $\operatorname{End} E_x$ is self-adjoint for every $v \in T_x^*M$.

Note that the last assumption implies that $(\sigma(v)\varphi, \sigma(v)\varphi) = (\sigma(v)^2\varphi, \varphi) = |v|_x^2(\varphi, \varphi)$, where $|v|_x$ is the norm in T_x^*M, so $\sigma(v)$ is an isometry if $|v|_x = 1$. This implies that the whole group $\mathrm{Pin}(T_x^*, g_x^*)$ acts as a group of isometries on E_x.

Proposition 6.8.2. *In a Hermitian Clifford bundle one can always define a connection ∇^E compatible with the metric in the sense of* (5.1.12) *such that for every vector field X, one form v and section φ of E we have*

$$\nabla^E_X(v \cdot \varphi) = (\nabla_X v) \cdot \varphi + v \cdot \nabla^E_X \varphi. \tag{6.8.1}$$

Here ∇ is the Levi-Civita connection.

A connection compatible with the metric and satisfying (6.8.1) is called a *Clifford connection.*

Proof. If (6.8.1) is valid and $v \in T^*_x$, $|v|_x = 1$, then it follows that

$$\nabla^E_X \varphi(x) = v \cdot \nabla^E_X (V \cdot \varphi)(x),$$

if the one form V is chosen with $\nabla V = 0$ and $V = v$ at x. This determines V up to second order terms which do not affect the right-hand side. Repeating this argument we obtain

$$\nabla^E_X \varphi(x) = u \cdot \nabla^E_X (u^* \cdot \varphi)(x), \quad u \in \mathrm{Pin}(T_x, g^*_x), \tag{6.8.2}$$

where u^* should be extended so that $\nabla u^* = 0$ at x. By Proposition 5.1.7 we can always choose a connection ∇^E which is compatible with the Hermitian metric. There is no reason why it should satisfy (6.8.2), but we can force it to do so by passing to the connection

$$\widetilde{\nabla}^E_X \varphi(x) = \int u \cdot \nabla^E_X (u^* \cdot \varphi)(x)\, du$$

where du is the invariant measure on the compact group Pin. Since the metric is invariant under Pin by assumption, this is still a connection compatible with the metric, and (6.8.2) is now valid for $\widetilde{\nabla}^E$, which means that (6.8.1) holds at x when $\nabla v = 0$ at x. In general we can for any given point x write $v = \sum_1^n \psi_j v_j$ in a neighborhood of x, where $\nabla v_j = 0$ at x and $\psi_j \in C^\infty$. Then we obtain

$$\widetilde{\nabla}^E_X (v \cdot \varphi) = \sum (X\psi_j) v_j \cdot \varphi + \sum \psi_j v_j \cdot \widetilde{\nabla}^E_X \varphi = (\nabla_X v) \cdot \varphi + v \cdot \widetilde{\nabla}^E_X \varphi,$$

where we have used first that $\widetilde{\nabla}^E_X$ is a connection, then that ∇ is a connection. This completes the proof.

For any connection ∇^E in a Clifford bundle E we can define a first order differential operator D in $C^\infty(M, E)$ by composing

$$\nabla^E : C^\infty(M, E) \to C^\infty(M, T^* \otimes E)$$

with the Clifford multiplication map

$$m : C^\infty(M, T^* \otimes E) \to C^\infty(M, E)$$

extending the bilinear map $T_x^* \times E_x \ni (v, \varphi) \mapsto v \cdot \varphi \in E_x$. For $\psi \in C^\infty(M)$ and $\varphi \in C^\infty(M, E)$ we have

$$e^{-t\psi} D(e^{t\psi}\varphi) = m((d\psi) \otimes \varphi) + D\varphi$$

so the principal symbol as defined in Section 5.1 is at $(x, \xi) \in T_x^*$ the map

$$E_x \ni w \mapsto \xi \cdot w.$$

In other words,

$$D(\psi\varphi) = \psi D\varphi + (d\psi) \cdot \varphi, \quad \psi \in C^\infty(M), \ \varphi \in C^\infty(M, E). \tag{6.8.3}$$

The square of the symbol is equal to $|\xi|^2$ times the identity in E_x, and if E is a Hermitian Clifford bundle it follows that D is of Dirac type. We can say more if the connection is well chosen:

Proposition 6.8.3. *If ∇^E is a Clifford connection in the Hermitian Clifford bundle E, then D is skew symmetric.*

Proof. We must show that if φ and ψ are in $C_0^\infty(M, E)$, then

$$\int_M \big((D\varphi, \psi) + (\varphi, D\psi)\big)\, dvol = 0. \tag{6.8.4}$$

At any point we can find an orthonormal basis $e_1, \dots, e_n$ for the vector fields, hence a dual orthonormal frame $\varepsilon_1, \dots, \varepsilon_n$ for the one forms, with $\nabla \varepsilon_j = 0$ for all j at the chosen point. Since

$$\nabla^E \varphi = \sum_1^n \varepsilon_j \otimes \nabla^E_{e_j} \varphi,$$

we have

$$D\varphi = \sum_1^n \varepsilon_j \cdot \nabla^E_{e_j} \varphi.$$

Define a (complex) vector field X by

$$X(v) = (v \cdot \varphi, \psi) = (\varphi, v \cdot \psi),$$

where v is a one form. Taking $v = \varepsilon_j$ we obtain at a point where $\nabla \varepsilon_j = 0$ for every j, using (6.8.1) and (5.1.12),

$$\sum_1^n (\nabla_{e_j} X)(\varepsilon_j) = \sum_1^n (\varepsilon_j \cdot \nabla^E_{e_j} \varphi, \psi) + (\varepsilon_j \cdot \varphi, \nabla^E_{e_j} \psi) = (D\varphi, \psi) + (\varphi, D\psi).$$

Since X has compact support and the divergence in the left-hand side is independent of the choice of frames (cf. (6.3.14)), we conclude using (6.3.14)$'$ that (6.8.4) holds.

Remark. Usually one includes a factor i in the definition to make D symmetric. However, we have stayed here with the definition of principal symbol given in Section 5.1 and hope to be consistent.

We shall say that a Clifford module E is graded if a direct sum decomposition $E = E^0 \oplus E^1$ is given which is compatible with the grading in $\mathrm{Cl}(M)$, that is, $\mathrm{Cl}_x^i \cdot E_x^j \subset E_x^{i+j \mod 2}$. The Clifford connection can clearly be chosen so that it respects this decomposition, and then the Dirac operator takes the form (6.6.2).

Exercise 6.8.1. Prove that if ∇^E is a Clifford connection and v is a one form, then

$$D(v \cdot \varphi) = -v \cdot D\varphi + 2\nabla^E_{v^\sharp}\varphi + (\mathcal{D}v) \cdot \varphi \tag{6.8.5}$$

where $\mathcal{D}$ is the Dirac type operator corresponding to the Clifford module $\mathrm{Cl}(M)$, that is, with the notation in the proof of Proposition 6.8.3,

$$\mathcal{D}v = \sum \varepsilon_j \cdot \nabla_{e_j} v.$$

Prove that $\mathcal{D}(d\psi) = \Delta\psi$ if $\psi \in C^\infty(M)$, and conclude that

$$D^2(\psi\varphi) = \psi D^2\varphi + 2\nabla_{(d\psi)^\sharp}\varphi + (\Delta\psi)\varphi, \quad \psi \in C^\infty(M),\ \varphi \in C^\infty(M, E). \tag{6.8.6}$$

We shall now prove Weitzenböck type formulas, similar to (6.3.15), for a Dirac type operator D corresponding to a Hermitian Clifford bundle E with a Clifford connection. Then D is skew adjoint and the principal symbol of D^2 is $|\xi|^2$. For the connection $\nabla : C^\infty(M, E) \to C^\infty(M, T^* \otimes E)$ we can form the adjoint with respect to the metric in E and the metric in $T^* \otimes E$ such that

$$\|\sum_1^n \varepsilon_j \otimes w_j\|^2 = \sum_1^n \|w_j\|^2, \quad w_j \in E_x,$$

if $\varepsilon_1, \dots, \varepsilon_n$ is an orthonormal basis in T_x^*. (This is the Hilbert-Schmidt metric; prove as an exercise that it is independent of the choice of basis $\varepsilon_1, \dots, \varepsilon_n$.) Then the principal symbol of $\nabla^*\nabla$ is $-|\xi|^2$, which proves that $D^2 + \nabla^*\nabla$ is of order ≤ 1. We shall now show that it is of order 0, which is no surprise in view of Proposition 6.1.2. For the proof we take $\chi \in C_0^\infty(M, \mathbf{R})$, $\varphi \in C^\infty(M, E)$ and recall from Exercise 6.8.1 that

$$D^2(\chi\varphi) = \chi D^2\varphi + 2\nabla_X\varphi + (\Delta\chi)\varphi, \quad X = (d\chi)^\sharp. \tag{6.8.7}$$

Since ∇ is a connection we have

$$\nabla(\chi\varphi) = \chi\nabla\varphi + (d\chi) \otimes \varphi.$$

If $\psi \in C^\infty(M, \mathbf{R})$ it follows that

$$\begin{aligned}(\nabla^*\nabla(\chi\varphi), \psi) &= (\nabla(\chi\varphi), \nabla\psi) = (\chi\nabla\varphi, \nabla\psi) + ((d\chi) \otimes \varphi, \nabla\psi) = (\nabla\varphi, \nabla(\chi\psi)) \\ &- (\nabla\varphi, (d\chi) \otimes \psi) + ((d\chi) \otimes \varphi, \nabla\psi) = (\chi\nabla^*\nabla\varphi, \psi) - (\nabla_X\varphi, \psi) + (\varphi, \nabla_X\psi).\end{aligned}$$

Since the connection is compatible with the Hermitian metric we have with scalar products in the fiber at x

$$(\varphi, \nabla_X \psi)(x) + (\nabla_X \varphi, \psi)(x) = \nabla_X (\varphi, \psi)(x),$$

so using (6.3.13) to integrate by parts we obtain

$$(\nabla^* \nabla(\chi\varphi), \psi) = (\chi \nabla^* \nabla \varphi, \psi) - 2(\nabla_X \varphi, \psi) - ((\operatorname{div} X)\varphi, \psi),$$

that is,

$$\nabla^* \nabla(\chi\varphi) = \chi \nabla^* \nabla \varphi - 2\nabla_X \varphi - (\operatorname{div} X)\varphi.$$

Combining this result with (6.8.7) we conclude that

$$(D^2 + \nabla^* \nabla)(\chi\varphi) = \chi(D^2 + \nabla^* \nabla)\varphi,$$

which means that $D^2 + \nabla^* \nabla$ is of order 0. More precisely, we have

Proposition 6.8.4. *If D is the Dirac type operator in a Hermitian Clifford bundle E over M corresponding to a Clifford connection ∇, then*

$$D^2 \varphi + \nabla^* \nabla \varphi = \tfrac{1}{2} \sum_{j,k=1}^{n} \varepsilon_k \cdot \varepsilon_j \cdot R^{\nabla}(e_k, e_j)\varphi, \quad \varphi \in C^\infty(M, E), \tag{6.8.8}$$

where $\varepsilon_1, \dots, \varepsilon_n$ and $e_1, \dots, e_n$ are local dual orthonormal frames for the one forms and the vector fields, and R^{∇} is the curvature of E with the connection ∇, defined by (5.1.5).

Proof. Writing ∇^E for the connection in E to distinguish it from the Levi-Civita connection in T^*M, we have

$$D\varphi = \sum_1^n \varepsilon_j \cdot \nabla^E_{e_j} \varphi, \quad \text{hence}$$

$$D^2 \varphi = \sum_{j,k=1}^{n} \varepsilon_k \cdot \nabla^E_{e_k} (\varepsilon_j \cdot \nabla^E_{e_j} \varphi) = \sum_{j,k=1}^{n} \varepsilon_k \cdot (\nabla_{e_k} \varepsilon_j) \cdot \nabla^E_{e_j} \varphi + \sum_{j,k=1}^{n} \varepsilon_k \cdot \varepsilon_j \cdot \nabla^E_{e_k} \nabla^E_{e_j} \varphi,$$

since ∇^E is a Clifford connection. The second sum is equal to

$$\sum_{j=1}^{n} \nabla^E_{e_j} \nabla^E_{e_j} \varphi + \sum_{j>k} \varepsilon_k \cdot \varepsilon_j \cdot (\nabla^E_{e_k} \nabla^E_{e_j} - \nabla^E_{e_j} \nabla^E_{e_k})\varphi, \quad \text{and}$$

$$(\nabla^E_{e_k} \nabla^E_{e_j} - \nabla^E_{e_j} \nabla^E_{e_k})\varphi = R^{\nabla^E}(e_k, e_j)\varphi + \nabla^E_{[e_k, e_j]}\varphi$$

by (5.1.5). Since $\nabla \varphi = \sum \varepsilon_j \otimes \nabla^E_{e_j} \varphi$ we obtain if $\psi \in C_0^\infty(M, E)$ has support in the coordinate patch U where the local frames are defined

$$(\nabla^* \nabla \varphi, \psi) = (\nabla \varphi, \nabla \psi) = \sum (\nabla^E_{e_j} \varphi, \nabla^E_{e_j} \psi).$$

When $x \in U$ we have with scalar product taken only in the fiber at x

$$(\nabla^E_{e_j}\varphi, \nabla^E_{e_j}\psi)(x) = \nabla^E_{e_j}(\nabla^E_{e_j}\varphi, \psi)(x) - (\nabla^E_{e_j}\nabla^E_{e_j}\varphi, \psi)(x).$$

If $\chi \in C_0^\infty(U)$ then

$$\int_U \nabla_{e_j}\chi \, dvol = \int_U \langle e_j, d\chi\rangle \, dvol = -\int \chi \operatorname{div} e_j \, dvol,$$

which proves that

$$\nabla^*\nabla\varphi = -\sum_{j=1}^n \nabla^E_{e_j}\nabla^E_{e_j}\varphi - \sum_{j=1}^n (\operatorname{div} e_j)\nabla^E_{e_j}\varphi.$$

Summing up, we have found that

$$\begin{aligned}(D^2 + \nabla^*\nabla)\varphi &- \tfrac{1}{2}\sum_{j,k=1}^n \varepsilon_k \cdot \varepsilon_j \cdot R^{\nabla^E}(e_k, e_j)\varphi \\ &= \sum_{j,k=1}^n \varepsilon_k \cdot (\nabla_{e_k}\varepsilon_j) \cdot \nabla^E_{e_j}\varphi - \sum_{j=1}^n \operatorname{div} e_j \nabla^E_{e_j}\varphi + \tfrac{1}{2}\sum_{j,k=1}^n \varepsilon_k \cdot \varepsilon_j \cdot \nabla^E_{[e_k, e_j]}\varphi.\end{aligned}$$

For any point $x \in U$ we can choose φ with given value and $\nabla\varphi = 0$ at x, which makes the right-hand side equal to 0 at x. But we have already proved that the left-hand side only depends on $\varphi(x)$, which completes the proof of (6.8.8).

As an example we shall derive (6.3.16) again by applying Proposition 6.8.4 to the Dirac type operator $d + d^*$ in $C^\infty(M, \wedge^*_{\mathbf{C}} M)$ (see (6.3.5)). Recall that by Exercise 3.1.2 and (5.1.5) the matrix of $R^\nabla(X, Y)$ for the cotangent bundle with the Levi-Civita connection is $-\sum_{k,l} R^j{}_{ikl} X^k Y^l$, $i, j = 1, \ldots, n$, in the coordinate frame. With $R^i{}_{jkl}$ denoting the components of the Riemann curvature tensor in the frame $\varepsilon_1, \ldots, \varepsilon_n$ instead, we obtain if $\varphi = \sum \varphi_l \varepsilon_l$ is a one form

$$R^\nabla(e_k, e_j)\varphi = -\sum_{l,i} R^i{}_{lkj}\varphi_i \varepsilon_l.$$

Hence the right-hand side of (6.8.8) becomes

$$-\tfrac{1}{2}\sum_{i,j,k,l=1}^n \varepsilon_k \cdot \varepsilon_j \cdot \varepsilon_l R^i{}_{lkj}\varphi_i. \tag{6.8.9}$$

If j, k, l are different then $\varepsilon_k \cdot \varepsilon_j \cdot \varepsilon_l$ is invariant under circular permutations, so the first Bianchi identity shows that the sum of such terms is equal to 0. If $k = j$ then

$R^i{}_{lkj} = 0$ so the only contributions come when $k = l$ or $j = l$; the terms common to these cases vanish. Hence (6.8.9) is equal to

$$(6.8.10) \qquad -\tfrac{1}{2} \sum_{i,j,k=1}^{n} (-\varepsilon_j R^i{}_{kkj} + \varepsilon_k R^i{}_{jkj})\varphi_i = - \sum_{i,j,k=1}^{n} \varepsilon_j R^i{}_{kjk}\varphi_i = - \sum \varepsilon_j R^i{}_j \varphi_i,$$

where R denotes the Ricci tensor in the last formula. This gives (6.3.16). The Weitzenböck formulas for forms of higher degree could be obtained in the same way, but then we would have to work out the curvature form in the higher exterior powers using Exercise 5.4.1.

We shall now pass to the main examples of Dirac operators, based on the half spin representations $D^{\pm}_{\frac{1}{2}}$ in (6.7.15). These are not representations of $\mathbf{SO}(2k)$ but of the covering group $\mathrm{Spin}(2k)$, so we cannot define associated vector bundles for an arbitrary Riemannian manifold.

Definition 6.8.5. An oriented Riemannian manifold M of dimension n has a *spin structure* if there exists a principal $\mathrm{Spin}(n)$ bundle $\widetilde{P}$ on M, which is a double cover of the oriented orthonormal frame bundle P, such that with the map $\mathrm{Spin}(n) \to \mathbf{SO}(n)$ defined by (6.7.12) we have a commutative diagram, with horizontal arrows defined by the right group action,

$$\begin{array}{ccc} \widetilde{P} \times \mathrm{Spin}(n) & \longrightarrow & \widetilde{P} \\ \Big\downarrow & & \Big\downarrow \\ P \times \mathbf{SO}(n) & \longrightarrow & P. \end{array}$$

In a neighborhood U of any point in M we can choose a spin structure just by identifying P with $U \times \mathbf{SO}(n)$ and $\widetilde{P}$ with $U \times \mathrm{Spin}(n)$. However, the existence of a spin structure in the large requires the vanishing of an element in $H^2(M, \mathbf{Z}_2)$ (the Stiefel-Whitney class), and there may exist inequivalent spin structures when this condition is fulfilled.

Assume now that M is an oriented manifold of dimension $n = 2k$ with a spin structure. Then the half spin bundles

$$S_{\pm}(\widetilde{P}) = \widetilde{P} \times_{D^{\pm}_{\frac{1}{2}}} S_{\pm}(2k)$$

and the full spin bundle $S(\widetilde{P}) = S_+(\widetilde{P}) \oplus S_-(\widetilde{P})$ are defined. The latter is associated with the direct sum ϱ of the representations $D^{\pm}_{\frac{1}{2}}$, which is the restriction of (6.7.14) to $\mathrm{Spin}(2k)$. The Clifford bundle can be regarded as the bundle associated with the representation of $\mathrm{Spin}(2k)$ on $\mathrm{Cl}(2k)$ by the representation

$$\mathrm{Spin}(2k) \times \mathrm{Cl}(2k) \ni (a, w) \mapsto awa^* \in \mathrm{Cl}(2k)$$

for this is a representation since a^* is the inverse of $a \in \mathrm{Spin}(2k)$, and it gives the orthogonal transformation $\tau(a)$ when applied to an element in $\mathbf{R}^{2k}$.

Proposition 6.8.6. *The spin bundle $S(\widetilde{P})$ is a Hermitian Clifford bundle, and the Levi-Civita connection is a Clifford connection in $S(\tilde{P})$. For the corresponding Dirac operator we have the Lichnerowicz formula*

$$D^2 + \nabla^*\nabla = -S/4, \tag{6.8.11}$$

where S is the scalar curvature.

Proof. Since the spin representation is unitary, the Hermitian metric in $S(2k)$ induces one in $S(\widetilde{P})$. The map

$$\mathrm{Cl}(2k) \times S(2k) \to S(2k)$$

defined by the identification of $\mathrm{Cl}_{\mathbf{C}}(2k)$ with $\mathrm{End}(S(2k))$, induces for every $x \in M$ a Clifford module structure $\mathrm{Cl}_x(M) \times S_x(\widetilde{P}) \to S_x(\widetilde{P})$. In fact, when

$$(p, x, w) \in \widetilde{P}_x \times \mathrm{Cl}(2k) \times S(2k)$$

and (pa^{-1}, x', w') with $a \in \mathrm{Spin}(2k)$ define the same element in $\mathrm{Cl}_x(M) \times S_x(\widetilde{P})$, then $x' = a \cdot x \cdot a^*$ and $w' = a \cdot w$, which implies that $x' \cdot w' = a \cdot (x \cdot w)$, hence that $(p, x \cdot w)$ and $(pa^{-1}, x' \cdot w') = (pa^{-1}, a \cdot (x \cdot w))$ define the same element in $S(\tilde{P})$. It is clear that the multiplication so defined depends smoothly on x. Multiplication by a vector $v \in T_x^* M$ is self-adjoint, since by our original definition it is the sum of an operator and its adjoint, so $S(\widetilde{P})$ is a Clifford module. At the center of a geodesic coordinate system the Levi-Civita connection cannot be distinguished from the flat connection in any of the bundles involved, so it is obvious that we have a Clifford connection.

With $R^i{}_{jkl}$ denoting the components of the Riemannian curvature tensor in the oriented orthonormal frame $\varepsilon_1, \dots, \varepsilon_n$ of T_x^*, with dual frame $e_1, \dots, e_n$ in T_x, we have seen above that the curvature $R^\nabla(e_k, e_j)$ of the cotangent bundle has the skew symmetric matrix $-(R^l{}_{ikj})_{i=1,\dots,n}^{l=1,\dots,n} = -(R_{likj})_{i=1,\dots,n}^{l=1,\dots,n}$. By Theorem 6.7.11 this skew symmetric matrix as an element in $\mathfrak{so}(n)$ corresponds in the Lie algebra of $\mathrm{Spin}(n)$ to $-\frac{1}{4}\sum_{i,l} R_{likj}\varepsilon_i \cdot \varepsilon_l$. Hence the right-hand side of (6.8.8) becomes

$$-\tfrac{1}{8}\sum_{i,j,k,l} R_{likj}\varepsilon_k \cdot \varepsilon_j \cdot \varepsilon_i \cdot \varepsilon_l = -\tfrac{1}{4}\sum \varepsilon_j R_{lj} \cdot \varepsilon_l = -\tfrac{1}{4}S,$$

where we have used (6.8.10). This proves (6.8.11).

The Lichnerowicz formula has a corollary completely parallel to Bochner's Theorem 6.3.5:

Corollary 6.8.7. *If M is a compact Riemannian manifold of dimension $2k$ with spin structure and non-negative not identically vanishing scalar curvature, then there are no harmonic spinors in M, that is, if $\varphi \in \mathcal{D}'(M, S)$ and $D\varphi = 0$, then $\varphi = 0$. If the scalar curvature is identically 0, then the space of harmonic spinors has dimension $\leq 2^k$.*

Proof. If $D\varphi = 0$ it follows that $\varphi \in C^\infty$ since D is elliptic, and taking the scalar product of (6.8.11) with φ we obtain

$$\|\nabla\varphi\|^2 + \tfrac{1}{4}(S\varphi, \varphi) = 0.$$

Hence φ is parallel, and if S is not identically 0 we obtain $\varphi = 0$ since φ must vanish in some open set. If S is identically 0 we can just conclude that φ is uniquely determined by its values at a point.

If F is a Hermitian vector bundle over the Riemannian manifold M of dimension $2k$ with spin structure defined by the principal $\mathrm{Spin}(2k)$ bundle $\widetilde{P}$, then the tensor product $E = S(\tilde{P}) \otimes F$ is a Clifford bundle with $\mathrm{Cl}(M)$ acting on the first factor. If F has a connection ∇^F compatible with the metric, then the connection ∇^E in E, given according to Proposition 5.1.8 by the Levi-Civita connection in $S(\widetilde{P})$ and ∇^F in F, is a Clifford connection. The easy verification is left as an exercise. Hence we get a *twisted Dirac operator* D_F in $C^\infty(M,E) = C^\infty(M,E_+) \oplus C^\infty(M,E_-)$, interchanging sections of $E_\pm = S_\pm(\widetilde{P}) \otimes F$. Since the curvature of E is the sum of the curvature of $S(\widetilde{P})$ and that of F, tensored respectively with the identity in F and that in $S(\widetilde{P})$, the Weitzenböck formula (6.8.8) now takes the form

$$D_F^2\varphi + \nabla^*\nabla\varphi = -\tfrac{1}{4}S\varphi - \tfrac{1}{2}\sum \varepsilon_i \cdot \varepsilon_j I \otimes R^{\nabla^F}(e_i, e_j), \tag{6.8.12}$$

where I is the identity in the spin bundle. We shall need (6.8.12) in Section 6.10.

6.9. Hermite polynomials and Mehler's formula. The proof of the local index formula for twisted Dirac operators will culminate in an identification of the difference of the traces of the heat kernels involved with the heat kernel belonging to an operator closely related to the harmonic oscillator. As a preparation we shall now give the classical background.

The Hermite polynomial $H_n(x)$, $x \in \mathbf{R}$, of order n is defined by

$$H_n(x) = e^{x^2}(-d/dx)^n(e^{-x^2}) = (2x)^n + \dots, \tag{6.9.1}$$

which by Taylor's formula for $e^{-(x-z)^2}$ implies that we have the generating function

$$\sum_{n=0}^{\infty} \frac{H_n(x)}{n!} z^n = e^{2xz - z^2}. \tag{6.9.2}$$

They are orthogonal with respect to the weight function e^{-x^2}, for if $n < m$ we have

$$\int_{\mathbf{R}} H_n(x)H_m(x)e^{-x^2}\,dx = \int_{\mathbf{R}} H_n(x)(-d/dx)^m e^{-x^2}\,dx = 0$$

since the integrand vanishes after $n+1 \le m$ integrations by parts. When $n = m$ we obtain instead

$$\int_{\mathbf{R}} H_n(x)H_n(x)e^{-x^2}\,dx = \int_{\mathbf{R}} H_n^{(n)} e^{-x^2}\,dx = 2^n n!\sqrt{\pi}.$$

Summing up,

$$\int_{\mathbf{R}} H_n(x)H_m(x)e^{-x^2}\,dx = \delta_{nm}2^n n!\sqrt{\pi} \tag{6.9.3}$$

Hence the functions $e^{-\frac{1}{2}x^2}H_n(x)/\sqrt{2^n n!\sqrt{\pi}}$ form an orthonormal system in $L^2(\mathbf{R})$. It is complete, for if $u \in L^2(\mathbf{R})$ is orthogonal to all of them, then $U(x) = u(x)e^{-\frac{1}{2}x^2}$ is orthogonal to all polynomials; the Fourier transform is then an entire analytic function with all derivatives equal to 0 at 0, hence identically 0, which proves that $u = 0$.

If in the identity

$$\sum_0^\infty H_n(x)e^{-\frac{1}{2}x^2}z^n/n! = e^{-\frac{1}{2}x^2+2xz-z^2}$$

we apply the differential operator

$$L = -d^2/dx^2 + x^2 - 1, \tag{6.9.4}$$

noting that $-(2z-x)^2 + x^2 = 2z(2x-2z)$, it follows that

$$\sum_0^\infty L(H_n(x)e^{-\frac{1}{2}x^2})z^n/n! = 2z\frac{d}{dz}e^{\frac{1}{2}x^2+2xz-z^2} = \sum_0^\infty H_n(x)e^{-\frac{1}{2}x^2}2nz^n/n!,$$

which means that

$$(L-2n)(H_n(x)e^{-\frac{1}{2}x^2}) = 0, \tag{6.9.5}$$

so we have a complete set of eigenfunctions of L, with eigenvalues $2n$.

To calculate the polynomials we introduce a convenient formalism. We can write

$$e^{-z^2} = \sum_0^\infty h_n z^n/n!, \quad h_{2n+1} = 0,\ h_{2n} = (-1)^n(2n)!/n!. \tag{6.9.6}$$

The right-hand side is obtained from the power series $\sum h^n z^n/n! = e^{hz}$ when h^n is replaced by h_n. Such a substitution is legitimate in any equality between two power series in h provided that no convergence difficulties occur, for the coefficients of corresponding powers of h are identical. If we apply this to the identity

$$e^{2xz}e^{hz} = e^{2xz+hz} = \sum_0^\infty (2x+h)^n z^n/n!$$

we obtain with $\mathcal{E}$ denoting the substitution of h_n for h^n

(6.9.7)

$$H_n(x) = \mathcal{E}(2x+h)^n = \sum_{2k\le n}(2x)^{n-2k}h_{2k}\binom{n}{2k} = n!\sum_{2k\le n}(2x)^{n-2k}(-1)^k/((n-2k)!k!).$$

Now we wish to determine the kernel of e^{-tL} when $t > 0$, which is equal to

$$\sum e^{-2tn} H_n(x)H_n(y)e^{-\frac{1}{2}(x^2+y^2)}/(2^n n!\sqrt{\pi}).$$

This means that we must calculate

$$\sum_0^\infty H_n(x)H_n(y)(s/2)^n/n!, \quad \text{when } 0 < s < 1. \tag{6.9.8}$$

By (6.9.7) and (6.9.2) the left-hand side is $\mathcal{E}$ applied to

$$\sum_0^\infty (h+2x)^n H_n(y)(s/2)^n/n! = e^{y(h+2x)s-(h+2x)^2s^2/4}, \tag{6.9.9}$$

where the exponent simplifies to

$$-h^2s^2/4 + h(ys - xs^2) + 2xys - x^2s^2.$$

The expansion of the right-hand side is not obvious unless the coefficient of h vanishes, that is, $y = xs$, but then we can use that

$$\mathcal{E}e^{-h^2s^2/4} = \sum_0^\infty (-1)^n h_{2n}(s^2/4)^n/n! = \sum_0^\infty \frac{(2n)!}{n!^2}(s^2/4)^n.$$

Since $(2n)!/(4^n n!^2) = \frac{1}{2}\frac{3}{2}\dots(\frac{1}{2}+n-1)/n! = (-1)^n\binom{-\frac{1}{2}}{n}$, we conclude that

$$\mathcal{E}e^{-h^2s^2/4} = 1/\sqrt{1-s^2}. \tag{6.9.10}$$

We have now proved that

$$\sum H_n(x)H_n(y)(s/2)^n/n! = e^{2xys-x^2s^2}/\sqrt{1-s^2}, \quad \text{if } y = xs.$$

In view of the symmetry we may exchange x and y in our earlier calculations which gives

$$\mathcal{E}e^{-h^2s^2/4+h(xs-ys^2)+2xys-y^2s^2} = e^{2xys-x^2s^2}/\sqrt{1-s^2}, \quad \text{if } y = xs.$$

With the notation $v = xs - ys^2 = xs(1-s^2)$ we have $(y^2-x^2)s^2 = x^2(s^2-1)s^2 = -v^2/(1-s^2)$ and conclude that for any v

$$\mathcal{E}e^{-h^2s^2/4+hv} = e^{-v^2/(1-s^2)}/\sqrt{1-s^2}. \tag{6.9.11}$$

This is the expansion we needed, so we now obtain for arbitrary x, y

$$\sum H_n(x)H_n(y)(s/2)^n/n! = e^{-(ys-xs^2)^2/(1-s^2)+2xys-x^2s^2}/\sqrt{1-s^2}.$$

This simplifies to

$$\sum_0^\infty H_n(x)H_n(y)(s/2)^n/n! = e^{-(x^2s^2+y^2s^2-2xys)/(1-s^2)}/\sqrt{1-s^2}, \tag{6.9.12}$$

which is known as *Mehler's formula.* I owe the trick in the proof to Marcel Riesz who showed it to me in a private conversation with him in 1952. My notes state that the proof goes back to the 1920's. I have inserted the operator $\mathcal{E}$ to justify the formal argument.

From (6.9.12) it follows that

$$\pi^{-\frac12}\sum_0^\infty H_n(x)H_n(y)(s/2)^n e^{-\frac12(x^2+y^2)}/n! \\ = e^{-\frac12((x^2+y^2)(1+s^2)-4xys)/(1-s^2)}/\sqrt{\pi(1-s^2)},$$

and with $s=e^{-2t}$ we obtain the kernel of e^{-tL} as

$$\exp\big(-(\tfrac12(x^2+y^2)\cosh 2t - xy)/\sinh 2t\big)e^t/\sqrt{2\pi\sinh 2t}.$$

The factor e^t is caused by the term -1 in $L=-d^2/dx^2+x^2-1$. If we define $L_1=-d^2/dx^2+x^2$, it follows that the kernel of e^{-tL_1} is

$$K(t,x,y)=\exp\big(-\tfrac12((x^2+y^2)\cosh 2t-2xy)/\sinh 2t)\big)/\sqrt{2\pi\sinh 2t}.$$

To find the kernel of the operator $L_a=-d^2/dx^2+a^2x^2$ we take $\tilde x = x\sqrt a$ as new variable so that $L_a=a\widetilde{L}_1$. Then

$$(e^{-tL_a}f)(x)=(e^{-ta\widetilde{L}_1}\tilde f)(\tilde x)=\int K(ta,\tilde x,\tilde y)f(\tilde y/\sqrt a)d\tilde y \\ = \int K(ta,x\sqrt a,y\sqrt a)f(y)\sqrt a\,dy,$$

so the kernel of e^{-tL_a} is $\sqrt a K(ta,x\sqrt a,y\sqrt a)$, that is,

$$\frac{1}{\sqrt{4\pi t}}\sqrt{\frac{2at}{\sinh 2at}}\exp\Big(-\frac{1}{4t}\frac{2at}{\sinh 2at}\big((x^2+y^2)\cosh 2at-2xy\big)\Big). \tag{6.9.13}$$

This formula gives immediately the kernel of a harmonic oscillator

$$-\Delta+\langle Ax,x\rangle$$

when A is a positive diagonal matrix, the only difference is that we get a product of such factors. By the orthogonal invariance the kernel is therefore in general, when the dimension is n,

$$\frac{1}{(4\pi t)^{\frac n2}}\sqrt{\det\frac{2\sqrt A t}{\sinh 2\sqrt A t}} \\ \times\exp\Big(-\frac{1}{4t}\Big(\big\langle\frac{2\sqrt A t}{\tanh 2\sqrt A t}x,x\big\rangle+\big\langle\frac{2\sqrt A t}{\tanh 2\sqrt A t}y,y\big\rangle-2\big\langle\frac{2\sqrt A t}{\sinh 2\sqrt A t}x,y\big\rangle\Big)\Big). \tag{6.9.14}$$

This is how the result is stated by Getzler.

We shall encounter a somewhat different operator, namely

$$L = -\sum_{j=1}^{n}(-\partial_j - i\sum_{k=1}^{n}\Omega_{jk}x^k)^2, \tag{6.9.15}$$

where Ω is a real skew symmetric matrix and i is the imaginary unit. At least formally we can write $L = L_0 + L_1$, where

$$L_0 = -\sum_1^n \partial_j^2 + \langle \Omega x, \Omega x\rangle, \quad L_1 = -2i\sum \Omega_{jk}x^k\partial_j.$$

Formally the operators commute, for

$$[\Delta, L_1] = -4i\sum \Omega_{jk}\partial_j\partial_k = 0,$$
$$[L_1, \langle \Omega^2 x, x\rangle] = -4i\sum \Omega_{jk}x^k(\Omega^2 x)_j = -4i\langle \Omega x, \Omega(\Omega x)\rangle = 0.$$

Formally this means that $e^{-tL} = e^{-tL_0}e^{-tL_1}$, but this is rather delicate since L_1 is a formally self-adjoint operator not bounded from below. To proceed, at first formally, we shall now assume that $n = 2$, which is no essential restriction since every skew symmetric Ω is the direct sum of such operators and trivial ones.

With $n = 2$ we now assume that

$$\Omega = \begin{pmatrix} 0 & a \\ -a & 0\end{pmatrix}, \quad \text{hence } -\Omega^2 = \begin{pmatrix} a^2 & 0 \\ 0 & a^2\end{pmatrix},$$
$$L_0 = -\Delta + a^2|x|^2, \quad L_1 = -2i\langle \Omega x, \partial\rangle.$$

The kernel of e^{-tL_0} is

$$\frac{1}{4\pi t}\frac{2at}{\sinh(2at)}\exp\Big(-\frac{1}{4t}\frac{2at}{\sinh(2at)}\big((|x|^2+|y|^2)\cosh(2at) - 2\langle x, y\rangle\big)\Big).$$

The vector field $\langle \Omega x, \partial\rangle = ax_2\partial/\partial x_1 - ax_1\partial/\partial x_2$ is infinitesimal generator of rotation with the speed a, so for real s we conclude that $e^{s\langle \Omega x, \partial\rangle}$ is rotation by the angle sa, that is, composition with the map $x \mapsto (x_1\cos(sa) - x_2\sin(sa), x_1\sin(sa) + x_2\cos(sa))$, so we can calculate $e^{s\langle Ox, \partial\rangle}e^{-tL_0}$ by just composing in the x variables with this map. By analytic continuation to $s = 2it$ this gives for the inner parenthesis in the exponent

$$\cosh(2at)(|x|^2+|y|^2-2x_1y_1-2x_2y_2) + 2i\sinh(2at)(x_2y_1 - x_1y_2)$$
$$= \cosh(2at)|x-y|^2 + 2i\sinh(2at)\langle \Omega x, y\rangle/a,$$

and leads to the kernel

$$\frac{1}{4\pi t}\frac{2at}{\sinh(2at)}\exp\Big(-\frac{1}{4t}\Big(\frac{2at}{\tanh 2at}|x-y|^2 + 4it\langle \Omega x, y\rangle\Big)\Big).$$

By analytic continuation it follows at once that $\partial/\partial t + L_x$ annihilates the kernel, and taking $x = y + z\sqrt{t}$ we find at once that it converges to δ as $t \to 0$, which easily justifies that we have indeed found the fundamental solution. With the suggestive notation $|\Omega| = \sqrt{-\Omega^2}$ it follows that the heat kernel for the operator (6.9.15) with a general real skew symmetric Ω is

$$(6.9.16)\quad \frac{1}{(4\pi t)^{\frac{n}{2}}}\sqrt{\det \frac{2|\Omega|t}{\sinh 2|\Omega|t}}\exp\Big(-\frac{1}{4t}\Big(\langle\frac{2|\Omega|t}{\tanh 2|\Omega|t}(x-y), x-y\rangle + 4it\langle\Omega x, y\rangle\Big)\Big).$$

Note that $z/\sinh z$ and $z/\tanh z$ are analytic functions of z^2 in a neighborhood of the real axis. For small t the right-hand side of (6.9.16) is therefore well defined even if Ω_{jk} are not real valued but take their values in some commutative finite dimensional algebra, such as $\wedge^*\mathbf{C}^k$, and it will still be a fundamental solution. This will be important in Section 6.10.

6.10. The local index formula for twisted Dirac operators. Let M be a Riemannian manifold of even dimension n with a spin structure, and let F be a Hermitian vector bundle on M with a connection compatible with the metric. At the end of Section 6.8 we defined the skew adjoint twisted Dirac operator D_F in $C^\infty(M, E)$ where $E = E_+ \oplus E_-$ and $E_\pm = S_\pm(\widetilde{P}) \otimes F$. It maps sections of $E_\pm$ to sections of $E_\mp$. By Proposition 6.5.1 we know that the index of $D_F^+ : C^\infty(M, E_+) \to C^\infty(M, E_-)$ for $t > 0$ is equal to the difference between the trace of $e^{tD_F^2}$ on $C^\infty(M, E_+)$ and the trace on $C^\infty(M, E_-)$. Thus it is the supertrace of $e^{tD_F^2}$ on $C^\infty(M, E)$ with the grading by E_+ and E_-. Let $K(t, x, y) \in \operatorname{Hom}(E_y, E_x)$ be the kernel of $e^{tD_F^2}$. We want to show in analogy to (6.5.5)$'$ and (6.5.6)$'$ that the supertrace of $K(t, x, x)$ for fixed x, without integration over M, has an asymptotic expansion in non-negative integer powers of t and calculate the constant term. The crucial point is the Berezin-Patodi formula (Theorem 6.7.14), which has an obvious extension to $\mathrm{Cl}_x(M)_{\mathbf{C}} \otimes F_x$. Let

$$\mathcal{L}_x^{(j)} = \oplus_{\nu \le \min(j, n/2)} \mathrm{Cl}_x^{[2\nu]}(M)_{\mathbf{C}} \otimes \operatorname{End} F_x,$$

be the natural filtration of the even part of $\operatorname{End} E$; recall that $\operatorname{End} S(\tilde{P}) \cong \mathrm{Cl}(M)_{\mathbf{C}}$. When $j \ge n/2$ then $\mathcal{L}_x^{(j)}$ is the whole even part.

We have seen in Section 6.4 that using geodesic coordinates centered at y and synchronous frames for E and T^*M, hence for $S(\widetilde{P})$, we have for fixed y an asymptotic expansion

$$K(t, x, y) \sim H_0(t, x)\sum_0^\infty (-t)^\nu u_\nu(x)/\nu!$$

where H_0 is the fundamental solution of the scalar heat equation, $u_0 = g(x)^{-\frac{1}{4}}$ and $u_1, u_2, \dots$ are determined successively by integrating the equations (6.1.14) (or equivalently (6.1.14)$'$). What makes the supertrace on the diagonal accessible to explicit calculation is the following fact:

Lemma 6.10.1. *The coefficients $u_\nu(0)$ are in $\mathcal{L}_y^{(\nu)}$ for $\nu = 0, 1, \dots$*

This is clear when $\nu = 0$ and will be proved by induction for increasing ν. The statement is void when $\nu \geq n/2$, but in the course of the proof we shall isolate the terms which are not better than the lemma states, and when $\nu = n/2$ this will give u_ν mod L_y^{k-1}, which is all that we need. Before the proof we need some preliminaries. The first point is to express D_F^2 in terms of geodesic coordinates and synchronous frames $\varepsilon_1, \dots, \varepsilon_n$ for T^*M, $f_1, \dots, f_N$ for F. With the dual frame $e_1, \dots, e_n$ in TM we write for $l = 1, \dots, n$

$$\nabla_{e_l}\varepsilon_j = \sum_{k=1}^n \widetilde{\Gamma}_{lj}{}^k \varepsilon_k,\ j = 1, \dots, n, \quad \nabla_{e_l} f_j = \sum_{k=1}^N \Gamma^{F\,k}_{lj} f_k,\ j = 1, \dots, N. \tag{6.10.1}$$

Here $\widetilde{\Gamma}_{lj}{}^k = -\widetilde{\Gamma}_{lk}{}^j$ since the frame is orthonormal, and by Theorem 3.3.6 we have

$$\widetilde{\Gamma}^k_{lj}(x) = \tfrac{1}{2}\sum_{i=1}^n R_{kjil}(0)x^i + O(|x|^2). \tag{6.10.2}$$

Note that since $\nabla_k \varepsilon_j = \partial_k \varepsilon_j$ at 0, it follows that $\varepsilon_j(x) - \varepsilon_j(0) = O(|x|^2)$. ($\widetilde{\Gamma}_{lj}{}^k$ are not quite the Christoffel symbols since they refer to the synchronous frame for T^*M and not the frame $dx^1, \dots, dx^n$.) Since $\nabla_\varrho f_j = \sum_k x^k \nabla_{\partial_k} f_j(x) = 0$, taking the first order terms in the Taylor expansion shows that $\nabla_{\partial_k} f_j(0) = 0$ for all j and k, hence

$$\Gamma^{F\,k}_{lj}(0) = 0, \quad \text{for } l = 1, \dots, n; \quad j, k = 1, \dots, N. \tag{6.10.3}$$

In the cotangent bundle we have

$$\nabla_{e_l}\Big(\sum \varphi_j \varepsilon_j\Big) = \langle e_l, \partial\rangle \varphi_j + \sum_{j,k=1}^n \widetilde{\Gamma}_{lj}{}^k \varphi_j \varepsilon_k.$$

The matrix $(\widetilde{\Gamma}_{lj}{}^k)_{k=1,\dots,n}^{j=1,\dots,n}$ is in $\mathfrak{so}(n)$ for fixed l, and as in the proof of Proposition 6.8.6 it corresponds to multiplication by $-\frac{1}{4}\sum_{j,k} \widetilde{\Gamma}_{lj}{}^k \varepsilon_j \cdot \varepsilon_k$ in the Clifford algebra, so

$$\nabla\varphi = \sum_{l=1}^n \varepsilon_l \otimes \Big(\langle e_l, \partial\rangle - \tfrac{1}{4}\sum_{j,k=1}^n \widetilde{\Gamma}_{lj}{}^k \varepsilon_j \cdot \varepsilon_k + \sum_{j,k=1}^N \Gamma^{F\,k}_{lj}\Big) \cdot \varphi. \tag{6.10.4}$$

To find the adjoint acting on $\psi = \sum \varepsilon_l \otimes \psi_l$ we form

$$(\psi, \nabla\varphi) = \sum_{l=1}^n \Big(\psi_l, \Big(\langle e_l, \partial\rangle - \tfrac{1}{4}\sum_{j,k=1}^n \widetilde{\Gamma}_{lj}{}^k \varepsilon_j \cdot \varepsilon_k + \sum_{j,k=1}^N \Gamma^{F\,k}_{lj}\Big) \cdot \varphi\Big).$$

Here the scalar product is of course taken using the Riemannian volume form, which is $1+O(|x|^2)$ times the Lebesgue measure. Hence we obtain by partial integration

$$\nabla^*\psi = -\sum \big(\langle e_l, \partial\rangle - \tfrac14 \sum_{j,k=1}^{n} \tilde{\tilde{\Gamma}}_{lj}{}^k \varepsilon_j\cdot\varepsilon_k + \sum_{j,k=1}^{N} \tilde{\Gamma}_{lj}^{F\,k}\big)\psi_l,$$

where $\tilde{\tilde{\Gamma}}_{lj}{}^k$ and $\tilde{\Gamma}_{lj}^{F\,k}$ also satisfy (6.10.2), (6.10.3) but include the terms arising when derivatives fall on e_l or on the Riemannian volume density. Summing up, we obtain in view of (6.8.12)

$$D_F^2 = -\tfrac14 S - \tfrac12 \sum_{i,j=1}^{n} \varepsilon_i\cdot\varepsilon_j R^{\nabla^F}(e_i,e_j) \tag{6.10.5}$$

$$+\sum_{l=1}^{n} \big(\langle e_l, \partial\rangle - \tfrac14 \sum_{j,k=1}^{n} \tilde{\tilde{\Gamma}}_{lj}{}^k \varepsilon_j\cdot\varepsilon_k + \sum_{j,k=1}^{N} \tilde{\Gamma}_{lj}^{F\,k}\big)\big(\langle e_l, \partial\rangle - \tfrac14 \sum_{j,k=1}^{n} \tilde{\Gamma}_{lj}{}^k \varepsilon_j\cdot\varepsilon_k + \sum_{j,k=1}^{N} \Gamma_{lj}^{F\,k}\big).$$

The coefficients can be viewed as elements in $\mathrm{Cl}_{\mathbf{C}}^{[2]} \otimes \operatorname{End} F_y$ or just $\operatorname{End} F_y$.

Proof of Lemma 6.10.1. We are now ready to prove that for the coefficients u_ν constructed for D_F^2 as in Section 6.1 we have

$$u_\nu = \sum_{\mu\le 2\nu} u_{\nu\mu}, \quad u_{\nu\mu}\in \mathcal{L}_y^{(\mu)}, \quad u_{\nu\mu} = O(|x|^{2\mu-2\nu}), \text{ if } \mu>\nu. \tag{6.10.6}$$

This is obvious when $\nu = 0$. Assuming that (6.10.6) holds for a certain ν we shall prove that

$$D_F^2 u_\nu = \sum_{\mu\le 2(\nu+1)} v_{\nu\mu}, \quad v_{\nu\mu}\in \mathcal{L}_y^{(\mu)}, \quad v_{\nu\mu} = O(|x|^{2\mu-2\nu-2}), \text{ if } \mu>\nu+1. \tag{6.10.6$'$}$$

This is obvious for the terms coming from the terms in (6.10.5) after the first sum. To handle the others we observe that multiplication by $\Gamma_{lj}^{F\,k}$ does not raise the Clifford degree but increases the order of the zeros. Multiplication by $\tilde{\Gamma}_{lj}{}^k\varepsilon_j\cdot\varepsilon_k$ may raise the Clifford degree by at most 2, but at the same time one gets a factor vanishing at 0. Differentiation reduces the order of the zero but does not affect the Clifford degree. Altogether, in the terms coming from the first sum in (6.10.5) where the Clifford degree is raised 4 units we also get a compensating factor $O(|x|^2)$, and $O(|x|^{2\mu-2\nu+2}) = O(|x|^{2(\mu+2)-2(\nu+1)})$. In those where the Clifford degree is raised 2 units we do not lose any zero at 0, and $O(|x|^{2\mu-2\nu}) = O(|x|^{2(\mu+1)-2(\nu+1)})$. In the terms where the Clifford degree is not raised we cannot lose more than two zeros, and $O(|x|^{2\mu-2\nu-2}) = O(|x|^{2\mu-2(\nu+1)})$, which completes the proof of (6.10.6)$'$.

The integration of the transport equations (6.1.14) for $u_{\nu+1}$, using (6.1.14)$'$ and (6.1.17), does not affect the order of the zeros at 0, so (6.10.6) follows by induction, which proves the lemma.

In the proof the estimate for the order of the zeros was always too low except for terms coming from the simplified operator

$$\sum_{l=1}^{n}(\partial_l - \tfrac{1}{8}\sum_{i,j,k=1}^{n} R_{kjil}(y)x^i\varepsilon_j\cdot\varepsilon_k)^2 - \tfrac{1}{2}\sum_{i,j=1}^{n}\varepsilon_i\cdot\varepsilon_j R^{\nabla^F}(e_i,e_j)(y), \tag{6.10.7}$$

where we have put in the argument y instead of 0 since 0 was the geodesic coordinate of y. Hence $\lim_{t\to 0}\operatorname{Str} K(t,y,y)$ exists and depends only on the Riemann curvature tensor and the curvature forms of F at y.

An explicit computation can be obtained from the results related to Mehler's formula given in Section 6.9. As a first step we observe that in the computation of the component of $u_{n/2}$ in $\mathrm{Cl}_y^{[n]}{}_{\mathbf{C}}\otimes\operatorname{End}F_y$ there will be no contributions where two equal factors ε_i in the Clifford algebra have been multiplied, and different ε_i anticommute. The component is therefore equal to the term of degree n obtained if one replaces Clifford multiplication by exterior multiplication throughout. With the notation

$$\Omega_{li} = -\tfrac{1}{8}\sum_{j,k=1}^{n} R_{kjil}(y)dx^j\wedge dx^k,\quad \Omega^F = -\tfrac{1}{2}R^{\nabla^F}(e_i,e_j)(y)dx^i\wedge dx^j, \tag{6.10.8}$$

we must therefore determine the form of degree n in the fundamental solution of $\partial/\partial t - L$ where

$$L = \sum_{l=1}^{n}(\partial_j + \sum_{k=1}^{n}\Omega_{jk}x^k)^2 + \Omega^F \tag{6.10.9}$$

acts on functions in $\mathbf{R}^n$ with values in $(\wedge^*\mathbf{C}^n)\otimes F_y$. The second term acts only in the factor F_y and commutes with the first, so it only contributes to e^{tL} a factor $e^{t\Omega^F}$, and for the other part we get the fundamental solution from (6.9.16) with Ω replaced by Ω/i. This means that $|\Omega|^2$ is replaced by Ω^2, so the kernel with the pole at 0 is

$$\frac{1}{(4\pi t)^{\frac{n}{2}}}\sqrt{\det\frac{2\Omega t}{\sinh 2\Omega t}}\exp\Big(-\frac{1}{4t}\Big\langle\frac{2\Omega t}{\tanh 2\Omega t}x,x\Big\rangle\Big)e^{t\Omega^F}, \tag{6.10.10}$$

which simplifies to

$$(4\pi t)^{-\frac{n}{2}}\sqrt{\det\frac{2\Omega t}{\sinh 2\Omega t}}e^{t\Omega^F} \tag*{(6.10.10)$'$}$$

when $x=0$. By the Berezin-Patodi formula the limit as $t\to 0$ of $\operatorname{Str}K(t,y,y)$ is $(2i)^{n/2}$ times the coefficient of $dx^1\wedge\cdots\wedge dx^n$ here, which does not depend on t. Recall now that the total Chernclass of F is $\det(I+i\Omega^F/2\pi)$; one calls

$$\operatorname{Ch}(F) = \operatorname{Tr}\exp(i\Omega^F/2\pi)$$

the *Chern character* of F. The term independent of t in (6.10.10)′ can be calculated using any convenient choice of t, and we choose $t = i/2\pi$. After multiplication by $(2i)^{n/2}$ this gives the form

$$\sqrt{\det \frac{(\Omega i/\pi)}{\sinh(\Omega i/\pi)}}\,\mathrm{Ch}(F).$$

Introducing the Riemann curvature matrix

$$\Omega^R_{li} = \tfrac{1}{2}\sum R_{kjil}\,dx^j \wedge dx^k, \quad (\Omega^R)^l_i = \sum g^{lk}\Omega^R_{ki} \tag{6.10.11}$$

instead of Ω we obtain finally the form

$$\sqrt{\det \frac{(\Omega^R/4\pi i)}{\sinh(\Omega^R/4\pi i)}}\,\mathrm{Ch}(F). \tag{6.10.12}$$

We have proved the local index theorem:

Theorem 6.10.2. *The index of the twisted Dirac operator* D^+_F *in the even dimensional spin manifold* M *is the integral of the form* (6.10.12) *over* M. *More precisely, the term in* (6.10.12) *of maximal degree divided by the normalized positively oriented* n *form is equal to the limit as* $t \to 0$ *of the supertrace of the kernel of* e^{tD_F} *on the diagonal.*

We refer to Atiyah-Bott-Patodi [1] for the topological arguments required to derive the general index theorem from Theorem 6.10.2; see also Roe [1] for a discussion of the passage to the Hirzebruch signature theorem.

APPENDIX A. PREREQUISITES FROM MULTILINEAR ALGEBRA

Let V be a vector space over $K = \mathbf{R}$ or $K = \mathbf{C}$ of finite dimension n. Then we can choose a *basis* $e_1, \dots, e_n \in V$ so that every $x \in V$ is a linear combination

$$x = \sum_1^n x_j e_j$$

with uniquely determined *coordinates* $x_j \in K$. If e'_j is another basis then

$$e'_k = \sum_1^n c_{jk} e_j$$

for some $c_{jk} \in K$ with non-zero determinant; if we write

$$x = \sum x'_k e'_k$$

it follows that

$$x_j = \sum_1^n c_{jk} x'_k$$

so bases and coordinates are transformed by inverse transposed matrices.

The *dual vector space* V' (sometimes denoted V^*) is the space of linear forms on V, with values in K. With a basis in V as above a linear form L can be expanded as

$$L(x) = L(\sum_1^n x_k e_k) = \sum_1^n L(e_k) x_k = \sum L(e_k) \varepsilon_k(x)$$

where

$$\varepsilon_k(\sum x_j e_j) = x_k, \quad \text{that is, } \varepsilon_k(e_j) = \delta_{jk}.$$

Thus $\varepsilon_1, \dots, \varepsilon_n$ form a *dual basis* in V' which is also a vector space of dimension n with the natural definition of addition of linear forms and multiplication of them by scalars. When $\xi \in V'$ and $x \in V$ we usually write $\langle x, \xi \rangle$ for the linear form $\xi(x)$ to emphasize that it is bilinear, that is, linear in x for fixed ξ (because ξ is a linear form) and linear in ξ for fixed x (by definition of the vector operations in V'). The bilinear form is *non-degenerate*, that is

$$\langle x, \xi \rangle = 0 \ \forall x \in V \implies \xi = 0; \quad \langle x, \xi \rangle = 0 \ \forall \xi \in V' \implies x = 0.$$

By the symmetry here we can identify $(V')'$ with V.

If $T : V_1 \to V_2$ is a linear map, then a unique linear map $T' : V_2' \to V_1'$ (also denoted T^* or ${}^tT$) called the *adjoint* or *transposed* map is defined by

$$\langle Tx, \eta\rangle = \langle x, T'\eta\rangle, \quad x \in V_1, \eta \in V_2',$$

for the left-hand side is for fixed η a linear form on V_1, hence an element $T'\eta \in V_1'$, and it depends linearly on η. T_1 and T_1' have the same *rank*.

If V and W are two vector spaces and $B : V \times W \ni x, y \mapsto B(x, y) \in K$ is a bilinear form, then we get in the same way a map $W \to V'$ and an adjoint map $V \to W'$. The form is called *non-degenerate* if one (and therefore both) are bijective. Thus the bilinear forms on $V \times W$ can be identified either with the linear maps $L(W, V')$ or the linear maps $L(V, W')$.

Every bilinear form B on $V \times V$ can in one and only one way be written $B = B_0 + B_1$ where

$$B_0(x, y) = B_0(y, x),\ x, y \in V; \quad B_1(x, y) = -B_1(y, x),\ x, y \in V;$$

one calls B_0 *symmetric* and B_1 *skew symmetric*, and we have

$$B_0(x, y) = (B(x, y) + B(y, x))/2, \quad B_1(x, y) = (B(x, y) - B(y, x))/2;\ x, y \in V.$$

With a basis $e_1, \dots, e_n$ for V and corresponding coordinates

$$B(\sum_1^n x_j e_j, \sum_1^n y_k e_k) = \sum_{j,k=1}^n b_{jk} x_j y_k, \quad b_{jk} = B(e_j, e_k).$$

From the matrix (b_{jk}) for B one obtains the matrices for B_0 and B_1 as $(b_{jk} \pm b_{kj})/2$. Only B_0 is determined by the *quadratic form* $B(x, x) = \sum b_{jk} x_j x_k$; we have

$$B_0(x, y) = (B(x + y, x + y) - B(x - y, x - y))/4 \quad \text{(polarization)}.$$

Thus we have a one to one correspondence between quadratic forms and symmetric bilinear forms. A quadratic form Q is called non-degenerate if the corresponding symmetric bilinear form B is non-degenerate. If the bilinear form is degenerate we can find $z \neq 0$ so that $B(V, z) = 0$, $B(z, V) = 0$, which implies that $Q(x + tz) = B(x+tz, x+tz) = B(x, x) = Q(x)$ for any $t \in K$, so Q is defined in the quotient space V/Kz. In what follows we shall use the same notation for a quadratic form and the corresponding symmetric bilinear form.

If G is a non-degenerate quadratic form in V, then the polarized form defines a bijection $g : V \to V'$ and we can define a *dual quadratic form* G' in V' by

$$G'(gx) = G(x), x \in V.$$

If we introduce dual bases in V and V' and (g_{jk}) is the symmetric matrix for G, then the matrix for G' is (g^{jk}) if (g^{jk}) is the inverse of the matrix g_{jk}.

If H is another quadratic form in V, then a linear transformation $T : V \to V$ is defined by the identity

$$H(x, y) = G(Tx, y), \quad x, y \in V,$$

for the corresponding symmetric bilinear forms. With the notation g, h for the maps $V \to V'$ defined by G and H, we have $T = g^{-1}h$. The map T is self-adjoint with respect to the bilinear form G since H is symmetric.

There is a unique linear form on $L(V, V)$, called the trace and denoted Tr, such that for all $x, \xi \in V \times V'$

$$\operatorname{Tr} T_{x,\xi} = \langle x, \xi\rangle, \quad \text{if } T_{x,\xi} \text{ is defined by } T_{x,\xi}y = x\langle y, \xi\rangle, \quad y \in V.$$

In fact, if e_j and ε_j are dual bases and $T \in L(V, V)$, then

$$Ty = T(\sum_1^n e_j\langle y, \varepsilon_j\rangle) = \sum_1^n Te_j\langle y, \varepsilon_j\rangle = \sum T_{Te_j,\varepsilon_j} y,$$

so the condition requires that

$$\operatorname{Tr} T = \sum_1^n \langle Te_j, \varepsilon_j\rangle.$$

If we define $\operatorname{Tr} T$ by this sum for a fixed basis, we obtain for arbitrary $x = \sum_1^n x_j e_j \in V$ and $\xi = \sum_1^n \xi_j \varepsilon_j \in V'$

$$\operatorname{Tr} T_{x,\xi} = \sum_1^n \langle x\langle e_j, \xi\rangle, \varepsilon_j\rangle = \sum_1^n \langle x, \varepsilon_j\rangle\langle e_j, \xi\rangle = \langle \sum_1^n e_j\langle x, \varepsilon_j\rangle, \xi\rangle = \langle x, \xi\rangle$$

which proves the existence of the form Tr with the required properites, independent of the choice of basis. Note that if T_{jk} is the matrix of T in terms of a basis then $\operatorname{Tr} T = \sum_1^n T_{jj}$.

Let V and W be two finite dimensional vector spaces over K. The bilinear forms on $V' \times W'$ form another vector space called the *tensor product* of V and W and denoted $V \otimes W$, and we have a bilinear map $V \times W \to V \otimes W$ mapping $x, y \in V \times W$ to the bilinear form

$$V' \times W' \ni \xi, \eta \mapsto \langle x, \xi\rangle\langle y, \eta\rangle$$

which is denoted $x \otimes y$. (Recall that V is identified with the dual of V' and W with the dual of W', which is behind the definition.) If $e_1, \dots, e_n$ is a basis of V with dual basis $\varepsilon_1, \dots, \varepsilon_n$, and $f_1, \dots, f_N$, $\varphi_1, \dots, \varphi_N$ are dual bases for W and W', then a bilinear form on $V' \times W'$ can be written

$$B(\xi, \eta) = B(\sum_1^n \langle e_j, \xi\rangle\varepsilon_j, \sum_1^N \langle f_k, \eta\rangle\varphi_k)$$
$$= \sum_{j=1}^n \sum_{k=1}^N B(\varepsilon_j, \varphi_k)\langle e_j, \xi\rangle\langle f_k, \eta\rangle, \quad \xi, \eta \in V' \times W',$$

which means that

$$B = \sum_{j=1}^{n}\sum_{k=1}^{N} B(\varepsilon_j, \varphi_k) e_j \otimes f_k.$$

This proves that $V \otimes W$ is a vector space of dimension $\dim V \dim W$ with basis $e_j \otimes f_k$, $j = 1, \dots, n$, $k = 1, \dots, N$. If $x_1, \dots, x_n$ are the coordinates of an element in V in the basis $e_1, \dots, e_n$ and $y_1, \dots, y_N$ are the coordinates of an element in W in the basis $f_1, \dots, f_N$, then $x_j y_k$ are the coordinates of $x \otimes y$.

If T is a linear map from $V \otimes W$ to a third vector space Z, then

$$S : V \times W \ni x, y \mapsto T(x \otimes y) \in Z$$

is a bilinear map. Every bilinear map $S : V \times W \to Z$ has a unique representation of this form: If $e_1, \dots, e_n$ and $f_1, \dots, f_N$ are bases in V and W we define T as the linear map with $T(e_j \otimes f_k) = S(e_j, f_k)$ on the basis elements in $V \otimes W$, and conclude that if $x = \sum x_j e_j$, $y = \sum y_k f_k$ then

$$\begin{aligned} S(x, y) = \sum x_j y_k T(e_j \otimes f_k) = T(\sum x_j y_k e_j \otimes f_k) \\ = T((\sum x_j e_j) \otimes (\sum y_k f_k)) = T(x \otimes y). \end{aligned}$$

Thus the bilinear map $V \times W \ni x, y \mapsto x \otimes y \in V \otimes W$ is universal in the sense that all other bilinear maps from $V \times W$ can be factored through it. The dual space of $V \otimes W$ is identified with $V' \otimes W'$ in a natural way so that $\langle x \otimes y, \xi \otimes \eta \rangle = \langle x, \xi \rangle \langle y, \eta \rangle$ if $x \in V$, $\xi \in V'$, $y \in W$ and $\eta \in W'$.

The linear transformations $L(V, W)$ from V to W were identified above with bilinear forms on $W' \times V$, that is, elements of $W \otimes V'$,

$$L(V, W) \cong W \otimes V'; \quad \text{in particular, } L(V, V) \cong V \otimes V'.$$

We can now look at the trace in a new way. The bilinear form

$$V \times V' \ni x, \xi \mapsto \langle x, \xi \rangle$$

defines a linear map $L(V, V) \cong V \otimes V' \to K$, which is precisely the trace.

We can continue to define tensor products of more than two vector spaces. For example $V_1 \otimes V_2 \otimes V_3$ is the space of trilinear forms on $V_1' \times V_2' \times V_3'$; it is isomorphic to $(V_1 \otimes V_2) \otimes V_3$ and to $V_1 \otimes (V_2 \otimes V_3)$ and is universal for trilinear maps in $V_1 \times V_2 \times V_3$. The verification is left as an exercise. Also the trace map can be generalized: If W and V are finite dimensional vector spaces then the trilinear map

$$W \times V \times V' \ni w, x, \xi \mapsto w\langle x, \xi \rangle \in W$$

defines a linear map $W \otimes V \otimes V' \to W$ called *contraction*. If T_{ijk}, $i = 1, \dots, N$, $j, k = 1, \dots, n$, are the coordinates of an element in $W \otimes V \otimes V'$ with respect to a basis $f_1, \dots, f_N$ in W and dual bases $e_1, \dots, e_n$ and $\varepsilon_1, \dots, \varepsilon_n$ in V and V', then the coordinates of the contraction are $\sum_{j=1}^n T_{ijj}$. This contraction operation can of course be applied to any tensor product

$$W_1 \otimes \cdots \otimes V \otimes \cdots \otimes V' \otimes W_k$$

containing two dual vector spaces; contraction gives an element in the tensor product where they are both removed.

In Riemannian geometry where V is the tangent space and V' the cotangent space, it is customary to put coordinate indices corresponding to factors V in a tensor product as superscripts and to put indices corresponding to factors V' as subscripts. Contraction then means putting a superscript equal to a subscript and summing over it.

If B is a bilinear form in $V \otimes V$ and a non-degenerate quadratic form G is given in V, then we can define the trace of B with respect to G as follows: We have $B \in V' \otimes V'$, and G gives an identification $V \cong V'$, so B can be identified with an element in $V \otimes V'$ for which the trace is defined. In terms of the matrices (b_{jk}) and (g_{jk}) for B and G with respect to a basis, the trace is $\sum b_{jk} g^{kj}$ where (g^{jk}) is the inverse of the matrix (g_{jk}).

We have defined the tensor product $V \otimes V$ as the space of bilinear forms on V'. Now every bilinear form on V' can be written as the sum of one symmetric and one skew symmetric bilinear form in one and only one way. We shall denote by $S^2(V) \subset V \otimes V$ the space of symmetric bilinear forms on V' and by $\wedge^2 V \subset V \otimes V$ the space of skew symmetric bilinear forms on V'. Thus the tensor product $V \otimes V$ is the direct sum of the *symmetric tensor product* $S^2(V)$ and the *exterior product* $\wedge^2 V$,

$$V \otimes V = S^2(V) \oplus \wedge^2 V.$$

For an arbitrary positive integer k we define the *symmetric* tensor product $S^k(V)$ as the space of *symmetric* k linear forms L on $V' \times \cdots \times V'$, that is, forms such that for every permutation π of $1, \dots, k$

$$L(\xi_1, \dots, \xi_k) = L(\xi_{\pi(1)}, \dots, \xi_{\pi(k)}), \quad \xi_1, \dots, \xi_k \in V'.$$

If $L \in S^k(V)$, then

$$\widetilde{L} : V' \ni \xi \mapsto L(\xi, \dots, \xi)$$

is a homogeneous polynomial of degree k in the sense that

$$(t_1, \dots, t_j) \mapsto \widetilde{L}(t_1\xi_1 + \cdots + t_j\xi_j)$$

for arbitrary $\xi_1, \dots, \xi_j \in V'$ is a polynomial in $(t_1, \dots, t_j) \in \mathbf{R}^j$; it suffices of course to have this for a basis in V'. We can recover $L(\xi_1, \dots, \xi_k)$ as the coefficient of $k! t_1 \dots t_k$ in $\widetilde{L}(t_1\xi_1 + \cdots + t_k\xi_k)$. Conversely, given a homogeneous polynomial $\widetilde{L}$ in V' of degree k, we can define $L(\xi_1, \dots, \xi_k)$ as the coefficient of $k! t_1 \dots t_k$ in $\widetilde{L}(t_1\xi_1 + \cdots + t_k\xi_k)$. The definition is symmetric in $\xi_1, \dots, \xi_k$. Since

$$\widetilde{L}(t_1\xi + \cdots + t_k\xi) = (t_1 + \cdots + t_k)^k \widetilde{L}(\xi),$$

this implies that $L(\xi,\dots,\xi)=\widetilde{L}(\xi)$, and since $\widetilde{L}(t_1\xi_1+\cdots+t_{k+1}\xi_{k+1})$ has no terms except

$$k!t_1\dots t_kL(\xi_1,\dots,\xi_k)+\cdots+k!t_2\dots t_{k+1}L(\xi_2,\dots,\xi_{k+1})$$

not containing the square of any t_j, it follows by taking $t_{k+1}=t_1$ that $L(\xi_1,\dots,\xi_k)$ is linear in ξ_1. Thus $L\in S^k(V)$, so we have identified $S^k(V)$ with the space of homogeneous polynomials of degree k in V'; this is the general meaning of the polarization discussed above for quadratic forms.

Exercise A.1. Determine the dimension of $S^k(V)$ in terms of $\dim V$ and k.

Similarly we define $\wedge^kV$ as the space of *alternating* forms L on V', that is, forms with

$$L(\xi_1,\dots,\xi_k)=\operatorname{sgn}\pi L(\xi_{\pi(1)},\dots,\xi_{\pi(k)}),\quad \xi_1,\dots,\xi_k\in V'.$$

If as before $e_1,\dots,e_n$ is a basis of V and $\varepsilon_1,\dots,\varepsilon_n$ the dual basis of V', then we can write

$$\begin{aligned}L(\xi_1,\dots,\xi_k)&=L(\sum\langle\xi_1,e_{j_1}\rangle\varepsilon_{j_1},\dots,\sum\langle\xi_k,e_{j_k}\rangle\varepsilon_{j_k})\\&=\sum\langle\xi_1,e_{j_1}\rangle\dots\langle\xi_k,e_{j_k}\rangle L(\varepsilon_{j_1},\dots,\varepsilon_{j_k})=\sum\det\langle\xi_i,e_{j_l}\rangle_{i,l=1}^kL(\varepsilon_{j_1},\dots,\varepsilon_{j_k})/k!.\end{aligned}$$

Thus the forms

$$V^k\ni(\xi_1,\dots,\xi_k)\mapsto\det\langle x_i,\xi_j\rangle_{i,j=1}^k. \tag{A.1}$$

give a basis for $\wedge^kV$ if we restrict $x_1,\dots,x_k$ to $e_{j_1},\dots,e_{j_k}$ where $e_1,\dots,e_n$ is a fixed basis for V and $1\le j_1<\cdots<j_k\le n$. (The linear independence of the forms follows since (A.1) is the only one which is not 0 on $\varepsilon_{j_1},\dots,\varepsilon_{j_k}$.) Thus the dimension of $\wedge^kV$ is $\binom{n}{k}$, defined as 0 if $k>n$.

The form (A.1) is denoted by $x_1\wedge\cdots\wedge x_k$. The map

$$V^k\ni(x_1,\dots,x_k)\mapsto x_1\wedge\cdots\wedge x_k$$

is obviously multilinear and alternating. If W is another vector space and $T:\wedge^kV\to W$ is a linear map, then

$$S:V^k\ni(x_1,\dots,x_k)\mapsto T(x_1\wedge\cdots\wedge x_k)\in W \tag{A.2}$$

is also multilinear and alternating. Conversely, every alternating multilinear map $S:V^k\to W$ has as unique representation of the form (A.2). For if $e_1,\dots,e_n$ is a basis for V, with dual basis $\varepsilon_1,\dots,\varepsilon_n$ for V', then

$$\begin{aligned}S(x_1,\dots,x_k)&=S(\sum\langle x_1,\varepsilon_{j_1}\rangle e_{j_1},\dots,\sum\langle x_k,\varepsilon_{j_k}\rangle e_{j_k})\\&=\sum_{1\le j_1<\cdots<j_k\le n}S(e_{j_1},\dots,e_{j_k})\det\langle x_i,\varepsilon_{j_l}\rangle_{i,l=1}^k,\\x_1\wedge\cdots\wedge x_k&=\sum_{1\le j_1<\cdots<j_k\le n}e_{j_1}\wedge\cdots\wedge e_{j_k}\det\langle x_i,\varepsilon_{j_l}\rangle_{i,l=1}^k\end{aligned}$$

so S is of the form (A.2) if T is the linear map $\wedge^k V \to W$ defined by

$$T(e_{j_1} \wedge \cdots \wedge e_{j_k}) = S(e_{j_1}, \ldots, e_{j_k}), \quad 1 \le j_1 < \cdots < j_k \le n,$$

for the basis elements in $\wedge^k V$. (However, S is of course independent of the choice of basis.)

The alternating multilinear map

$$V^{k+\kappa} \ni (x_1, \ldots, x_{k+\kappa}) \mapsto x_1 \wedge \cdots \wedge x_k \wedge x_{k+1} \wedge \cdots \wedge x_{k+\kappa} \in \wedge^{k+\kappa} V$$

is for fixed $x_{k+1}, \ldots, x_{k+\kappa}$ an alternating multilinear map from V^k, so it defines a linear map

$$T(x_{k+1}, \ldots, x_{k+\kappa}) : \wedge^k V \to \wedge^{k+\kappa} V.$$

For every $w \in \wedge^k V$ the map

$$V^\kappa \ni (x_{k+1}, \ldots, x_{k+\kappa}) \mapsto T(x_{k+1}, \ldots, x_{k+\kappa})w \in \wedge^{k+\kappa} V$$

is alternating and multilinear so it defines a linear map $T'(w) : \wedge^\kappa V \to \wedge^{k+\kappa} V$. It depends linearly on w, so

$$\wedge^k V \times \wedge^\kappa V \ni (w, w') \mapsto T'(w)w' \in \wedge^{k+\kappa} V$$

is a bilinear map. It is characterized by the fact that it maps $x_1 \wedge \cdots \wedge x_k \in \wedge^k V$ and $x_{k+1} \wedge \cdots \wedge x_{k+\kappa} \in \wedge^\kappa V$ to $x_1 \wedge \cdots \wedge x_{k+\kappa} \in \wedge^{k+\kappa} V$. With the notation $\wedge$ also for this map, we have thus found that there is a uniquely defined bilinear map

$$\wedge^k V \times \wedge^\kappa V \ni (w, w') \mapsto w \wedge w' \in \wedge^{k+\kappa} V \tag{A.3}$$

such that

$$(x_1 \wedge \cdots \wedge x_k) \wedge (x_{k+1} \wedge \cdots \wedge x_{k+\kappa}) = x_1 \wedge \cdots \wedge x_{k+\kappa}, \quad x_1, \ldots, x_{k+\kappa} \in V. \tag{A.4}$$

It is called the exterior product. It makes

$$\wedge^* V = \bigoplus_{k=0}^n \wedge^k V,$$

where $\wedge^0 V = K$, an associative algebra. Since

$$x_1 \wedge \cdots \wedge x_k \wedge x_{k+1} \wedge \cdots \wedge x_{k+\kappa} = (-1)^{k\kappa} x_{k+1} \wedge \cdots \wedge x_{k+\kappa} \wedge x_1 \wedge \cdots \wedge x_k,$$

the algebra is not commutative but we have

$$w \wedge w' = (-1)^{k\kappa} w' \wedge w, \quad \text{if} \quad w \in \wedge^k V, \ w' \in \wedge^\kappa V. \tag{A.5}$$

Note that w and w' commute unless the degrees k and κ of w and w' are *both odd*; then they anticommute.

If V_1 and V_2 are finite dimensional vector spaces, then every linear map $A : V_1 \to V_2$ induces a linear map $\wedge^k A : \wedge^k V_1 \to \wedge^k V_2$. In fact, the map

$$V_1^k \ni (x_1, \dots, x_k) \mapsto (Ax_1) \wedge \dots \wedge (Ax_k) \in \wedge^k V_2$$

is alternating and multilinear so there is a unique linear map $\wedge^k A : \wedge^k V_1 \to \wedge^k V_2$ such that

$$(\wedge^k A)(x_1 \wedge \dots \wedge x_k) = (Ax_1) \wedge \dots \wedge (Ax_k), \quad x_1, \dots, x_k \in V. \tag{A.6}$$

We can define the exterior powers $\wedge^k V'$ of the dual V' of V in the same way. The multilinear map

$$V^k \times V'^k \ni (x_1, \dots, x_k, \xi_1, \dots, \xi_k) \mapsto \det\langle x_i, \xi_j\rangle_{i,j=1}^k$$

is alternating in the first k and in the last k variables. Hence it induces a bilinear form on $\wedge^k V \times \wedge^k V'$. It defines a duality for if $e_1, \dots, e_n$ and $\varepsilon_1, \dots, \varepsilon_n$ are dual bases for V and for V', then the bases $e_{j_1} \wedge \dots \wedge e_{j_k}$ and $\varepsilon_{i_1} \wedge \dots \wedge \varepsilon_{i_k}$ with $1 \le j_1 < \dots < j_k \le n$ and $1 \le i_1 < \dots < i_k \le n$ are dual with respect to this form. Thus we have a natural duality between $\wedge^k V$ and $\wedge^k V'$ such that

$$\langle x_1 \wedge \dots \wedge x_k, \xi_1 \wedge \dots \wedge \xi_k\rangle = \det\langle x_i, \xi_j\rangle_{i,j=1}^k, \quad x_1, \dots, x_k \in V,\ \xi_1, \dots, \xi_k \in V'. \tag{A.7}$$

(Many authors use another definition where the right-hand side is divided by $k!$, for that is the duality inherited from the natural duality of the tensor products. See Sternberg [1, p. 19] for a discussion of this point. Kobayashi and Nomizu [1] use the division by $k!$, which should be kept in mind when comparing identities.)

APPENDIX B. THE CALCULUS OF DIFFERENTIAL FORMS

Let M be a C^∞ manifold of dimension n. Then the exterior power $\wedge^k T_x^* M$ of the cotangent space at $x \in M$ is a real vector space of dimension $\binom{n}{k}$. If $M \subset \mathbf{R}^n$, then $T_x^* M \cong \mathbf{R}^n$, and $\wedge^k T_x^* M$ has the basis $dx^{j_1} \wedge \cdots \wedge dx^{j_k}$, where $1 \le j_1 < \cdots < j_k \le n$; this is the alternating multilinear form

$$(\mathbf{R}^n)^k \ni (t_1, \ldots, t_k) \mapsto \det({t_i}^{j_l})_{i,l=1}^k.$$

For any C^1 map $f : N \to M$, where N is another smooth manifold, the adjoint of the differential

$$f'(y) : T_y N \to T_{f(y)} M$$

is a map $T_{f(y)}^* M \to T_y^* N$ inducing a map

$$\wedge^k T_{f(y)}^* M \to \wedge^k T_y^* N, \tag{B.1}$$

which is bijective if $f'(y)$ is bijective. In particular, if f is a coordinate system, $\mathbf{R}^n \supset \omega \to f(\omega) \subset M$, we obtain an identification of $\bigcup_{x \in f(\omega)} \wedge^k T_x^* M$ and $\omega \times \wedge^k \mathbf{R}^n$. Thus the vector spaces $\wedge^k T_x^* M$ define a vector bundle $\wedge^k T^* M$ over M. In a coordinate patch with local coordinates $x^1, \ldots, x^n$, a local frame for $\wedge^k T^* M$ is given by the exterior products

$$dx^{j_1} \wedge \cdots \wedge dx^{j_k}, \quad 1 \le j_1 < \cdots < j_k \le n.$$

We shall denote the space of sections of $\wedge^k T^* M$ by $\lambda^k(M)$; in terms of local coordinates a section u can thus be written

$$u = \sum_{1 \le j_1 < \cdots < j_k \le n} a_{j_1 \ldots j_k}(x) dx^{j_1} \wedge \cdots \wedge dx^{j_k}. \tag{B.2}$$

Exterior multiplication of forms is done just by formal multiplication observing the anticommutation rule $dx^i \wedge dx^j = -dx^j \wedge dx^i$.

If $f : N \to M$ is a smooth map, then the map (B.1) defines for every k a linear map

$$f^* : \lambda^k(M) \to \lambda^k(N).$$

One calls $f^* u \in \lambda^k(N)$ the pullback of $u \in \lambda^k(M)$, from M to N. If x and y are local coordinates in M and N and u has the form (B.2), $x = f(y)$, then

$$f^* u = \sum_{1 \le j_1 < \cdots < j_k \le n} a_{j_1 \ldots j_k}(f(y)) df^{j_1}(y) \wedge \cdots \wedge df^{j_k}(y), \tag{B.3}$$

where $df^j(y) = \sum \partial f^j(y)/\partial y^i dy^i$. In fact, if φ is a function on M (such as a local coordinate), then

$$\langle f^* d\varphi, t\rangle = \langle d\varphi, f' t\rangle = \langle d(\varphi \circ f), t\rangle = \langle d(f^*\varphi), t\rangle, \quad t \in TN,$$

by the chain rule, which means that $f^* d\varphi = d(f^*\varphi)$. We have natural rules of computation such as $(fg)^* = g^* f^*$; the notation with the upper star is meant to be a reminder of this reversal of factors.

We shall now define the *exterior differential* of a k form. We do this first in a fixed local coordinate system and verify the independence of the chosen coordinates afterwards. For the form u in (B.2) we thus define

$$du = \sum_{1 \le j_1 < \cdots < j_k \le n} da_{j_1 \ldots j_k}(x) \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}. \tag{B.4}$$

It follows at once from this definition that

$$d(u \wedge v) = (du) \wedge v + (-1)^k u \wedge dv, \quad u \in \lambda^k,\ v \in \lambda^\kappa; \tag{B.5}$$

in view of the linearity it suffices to verify (B.5) when $u = adx^{j_1} \wedge \cdots \wedge dx^{j_k}$ and $v = bdx^{j_{k+1}} \wedge \cdots \wedge dx^{j_{k+\kappa}}$, and then it follows from the fact that $d(ab) = adb + bda$; commutation of db through u gives the sign $(-1)^k$. If u is a smooth function we have

$$d^2 u = 0. \tag{B.6}$$

In fact, $du = \sum \partial_j u dx^j$, so

$$d^2 u = \sum d(\partial_j u) \wedge dx^j = \sum \partial_k \partial_j u dx^k \wedge dx^j = 0$$

since $dx^k \wedge dx^j$ is antisymmetric in k and j while $\partial_k \partial_j u$ is symmetric. In view of (B.2) and (B.5) it follows at once that (B.6) is valid for $u \in \lambda^k$ for any k.

Repeated use of (B.5) shows that for arbitrary smooth functions $f_0, \ldots, f_k$ we have

$$d(f_0 df_1 \wedge \cdots \wedge df_k) = df_0 \wedge df_1 \wedge \cdots \wedge df_k.$$

If $u \in \lambda^k(\omega)$, where ω is an open subset of $\mathbf{R}^n$, and $f : \omega' \to \omega$ is a smooth map from an open subset ω' of $\mathbf{R}^{n'}$, it follows that

$$f^*(du) = df^* u, \quad u \in \lambda^k(\omega). \tag{B.7}$$

In fact, if u is given by (B.2), then du is given by (B.4) and $f^* u$ is given by (B.3), hence

$$df^* u(x) = \sum_{1 \le j_1 < \cdots < j_k \le n} da_{j_1 \ldots j_k}(f(y)) \wedge df^{j_1}(y) \wedge \cdots \wedge df^{j_k}(y),$$

which proves (B.7). In particular, it follows from (B.7) that our definition (B.4) of the exterior differential is independent of the choice of local coordinates.

Locally there is a converse of (B.6):

THEOREM B.1 (POINCARÉ'S LEMMA). *Let $v \in \lambda^{k+1}(X)$ where X is a convex open set in $\mathbf{R}^n$, assume that the coefficients of v are in C^μ and that $dv = 0$. Then there is a form $u \in \lambda^k(X)$ with C^μ coefficients such that $du = v$.*

PROOF. We may assume that $0 \in X$. Then $\widehat{X} = \{(x,t) \in \mathbf{R}^n \times \mathbf{R}; tx \in X\}$ is an open set containing $X \times [0,1]$, and $f(x,t) = tx$ is a C^∞ map from $\widehat{X}$ to X, so we can form

$$f^*v = f_t^*v + dt \wedge w_t$$

where f_t^* is the pullback of v when t is regarded as a parameter, so it is a differential form which only involves the differentials of the coordinates in X. Since $df^*v = f^*dv = 0$ it follows that

$$0 = dt \wedge (\partial(f_t^*v)/\partial t - d_x w_t) + \dots$$

where the dots indicate a form which has no factor dt and $d_x w_t$ is the differential of w_t for fixed t. Hence

$$\partial(f_t^*v)/\partial t = d_x w_t,$$

and integration from 0 to 1 gives

$$f_1^*v - f_0^*v = du, \quad u = \int_0^1 w_t dt.$$

But $f_1^*v = v$ and $f_0^*v = 0$ since f_1 is the identity and f_0 maps X to $\{0\}$, which proves the theorem.

Poincaré's lemma reflects the fact that the topology of a convex set in $\mathbf{R}^n$ is very simple; in general there is a topological obstruction:

DEFINITION B.2. If X is a smooth manifold then the quotient $H^k(X)$ of the *closed* forms $\{u \in \lambda^k(X); du = 0\}$ by the linear subspace $\{dv; v \in \lambda^{k-1}(X)\}$ of *exact* forms, $k > 0$, is called the de Rham cohomology of degree k. The residue class in $H^k(X)$ of a form $u \in \lambda^k(X)$ with $du = 0$ is called the cohomology class of u.

EXERCISE B.1. Prove that if M and N are smooth manifolds, $f : M \times [0,1] \to N$ is a smooth map, and $u \in \lambda^k(N)$ is a closed form, then the forms $f(\cdot,t)^*u \in \lambda^k(M)$ are in the same cohomology class for all $t \in [0,1]$.

Poincaré's lemma is closely related to Stokes' formula, which we shall now discuss. First we define the integral of a form $u \in \lambda^n(M)$ with compact support over M when M is an *oriented* manifold of dimension n. To do so we first assume that the support is contained in a local coordinate patch with positively oriented local coordinates $x \in \omega \subset \mathbf{R}^n$. In terms of these coordinates we can then write

$$u = a(x)dx^1 \wedge \dots \wedge dx^n$$

and we define

$$\int u = \int a(x)dx,$$

where dx is the Lebesgue measure. If the support of u is also contained in a coordinate patch with positively oriented local coordinates y, then

$$u = a(x(y))dx^1(y) \wedge \cdots \wedge dx^n(y) = a(x(y)) \det(\partial x^i(y)/\partial y^j)_{i,j=1}^n$$

so in terms of these coordinates $\int u$ should be

$$\int a(x(y)) \det(\partial x^i(y)/\partial y^j)_{i,j=1}^n dy.$$

The positive orientation means precisely that the Jacobian here is positive, so the definitions in terms of the two coordinate systems agree. Now we can define $\int u$ in general by using a partition of unity $1 = \sum \varphi_j$ to split u into a finite sum $u = \sum \varphi_j u$ where each term has support in a local coordinate patch and using these local coordinates set

$$u = \sum \int \varphi_j u.$$

The definition is independent of the choice of partition of unity and corresponding local coordinates, for if we have another partition of unity $1 = \sum \psi_k$ with each term supported by a coordinate patch, then $\int \psi_k \varphi_j u$ has the same value if we use the coordinates associated with φ_j or those associated with ψ_k, and summing over k or over j we conclude that the definition using the partition of unity $\{\varphi_j\}$ is equivalent to that using $\{\psi_k\}$.

Remark. We have not given a precise definition of orientation above. One way to do so is to say that an orientation on a manifold of dimension n is a n form o, the orientation form, which is everywhere different from 0; orientation forms differing by multiplication with a positive function are considered equivalent. A system of local coordinates $x^1, \dots, x^n$ is then positively oriented if with these coordinates $o = g(x)dx^1 \wedge \cdots \wedge dx^n$ where $g(x) > 0$. The manifold is said to be oriented by $o > 0$.

The integral over M of a k form is defined as 0 if $k < n$, so we have a definition of the integral of any form with compact support. The definition of the integral over a submanifold N of M is an immediate consequence: For a submanifold we have an embedding $i : N \to M$, and for a form u on M we define $\int_N u = \int_N i^* u$ if $i^* u$ has compact support.

Now assume that X is an open subset of the oriented manifold M, with a C^∞ boundary ∂X for the sake of simplicity. At any boundary point we can then choose local coordinates in M varying over the unit ball B in $\mathbf{R}^n$, say, such that the points in X correspond to $B_- = \{x \in B; x_1 < 0\}$. If $u \in \lambda^{n-1}(B)$ has compact support in B, it follows that

$$\int_{B_0} u = \int_{B_-} du, \tag{B.8}$$

where $B_0 = \{x \in B; x_1 = 0\}$. In fact,

$$u = \sum_1^n (-1)^{j-1} u_j(x) dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n,$$

$$du = (\sum_1^n \partial_j u_j) dx^1 \wedge \cdots \wedge dx^n,$$

so the statement is that with $x' = (x^2, \dots, x^n)$

$$\int u_1(0, x')\, dx' = \int_{x_1<0} \sum_1^n \partial_j u_j \, dx,$$

which is obvious.

Now we can orient ∂X uniquely by using at points in ∂X positively oriented local coordinates in M such that X is located as just described where $x_1 < 0$; the coordinates $x_2, \dots, x_n$ in ∂X are then by definition positively oriented. From the local formula (B.8) it follows by using a partition of unity that

$$\int_{\partial X} u = \int_X du, \quad \text{Stokes' formula,} \tag{B.9}$$

if u is a form with compact support in M. The same formula holds if X is an open subset with C^1 boundary in an oriented submanifold of M; this is an immediate consequence in view of the definition of integration over a submanifold.

Given a $k+1$ form v with $dv = 0$ it follows from (B.9) that there does not exist a k form u with $du = v$ unless $\int_X v = 0$ for any compact oriented submanifold X without boundary. This is the starting point of homology theory, but it would take us too far to pursue the matter here.

APPENDIX C. THE FROBENIUS THEOREM

To study flat structures we shall need the Frobenius existence theorem for first order systems of differential equations. One of the standard forms of this result is as follows:

THEOREM C.1. *Let $v_1, \dots, v_r$ be C^∞ vector fields in a neighborhood of* 0 *in* $\mathbf{R}^n$ *such that*

$$v_1(0), \dots, v_r(0) \quad \text{are linearly independent,} \tag{C.1}$$

$$[v_i, v_j] = \sum_{k=1}^{r} c_{ijk} v_k, \quad i, j = 1, \dots, r, \tag{C.2}$$

where $[v, w]$ denotes the commutator of v and w, and $c_{ijk} \in C^\infty$. Then there exist new local coordinates $y_1, \dots, y_n$ in a neighborhood of 0 *such that*

$$\partial/\partial y_i = \sum_{j=1}^{r} b_{ij} v_j, \quad i = 1, \dots, r.$$

Thus the solutions of the equations $v_j u = 0$, $j = 1, \dots, r$ are in a neighborhood of the origin precisely the functions of $y_{r+1}, \dots, y_n$.

PROOF. The proof, which can be found in Hörmander [1, Appendix C1], is by induction with respect to n. Invariance under change of variables and non-singular recombinations of the v_j shows that one may assume that $v_1 = \partial/\partial x_1$ while

$$v_j = \sum_{l=2}^{n} v_{jl} \partial/\partial x_l, \quad j = 2, \dots, r.$$

By the inductive hypothesis we may also assume that $v_{jl} = 0$ for $j = 2, \dots, r$ if $l > r$ and $x_1 = 0$. By (C.1)

$$\partial v_{jl}/\partial x_1 = [v_1, v_j] x_l = \sum_{k=2}^{r} c_{1jk} v_k x_l = \sum_{k=2}^{r} c_{1jk} v_{kl}, \qquad j = 2, \dots, r.$$

Hence $v_{jl} = 0$ in a neighborhood of 0 if $l > r$ since this is true when $x_1 = 0$, which proves the theorem.

The geometric interpretation is as follows: The vector fields $v_1, \dots, v_r$ span at every point x a linear space F_x of dimension r in the tangent space. The condition (C.2) is the necessary and sufficient condition in order that through every point there is a manifold M of dimension r such that the tangent plane of M is F_x at every $x \in M$. Now we can also define the linear space F_x by means of its orthogonal spaces $F_x^\perp$ in the cotangent space, that it, by $n - r$ conditions

$$\omega_j = 0, \quad j = r+1, \dots, n, \tag{C.3}$$

where $\omega_{r+1}, \dots, \omega_n$ are linearly independent differential one forms at the origin. We want to rewrite the Frobenius condition (C.2) in terms of these forms. To do so we need a lemma:

LEMMA C.2. *If X and Y are C^1 vector fields and ω is a C^1 one form, then*

$$\langle X \wedge Y, d\omega \rangle = \langle [Y, X], \omega \rangle + X\langle Y, \omega \rangle - Y\langle X, \omega \rangle. \tag{C.4}$$

PROOF. If $\omega = du$ for some u then (C.4) is the definition of the commutator vector field. If ω is replaced by $\varphi\omega$, then both sides are multiplied by φ and in addition we get on the left-hand side a term $\langle X \wedge Y, d\varphi \wedge \omega \rangle$ which is equal to the additional term

$$\langle X, d\varphi \rangle \langle Y, \omega \rangle - \langle Y, d\varphi \rangle \langle X, \omega \rangle$$

in the right-hand side. Hence (C.4) follows in general.

Let us now return to the Frobenius theorem. In a neighborhood of the origin we can extend the system of vector fields $v_1, \dots, v_r$ to a basis $v_1, \dots, v_n$. Let ω_j be the biorthogonal basis of one forms, that is,

$$\langle v_j, \omega_k \rangle = \delta_{jk}, \quad j, k = 1, \dots, n.$$

Then

$$[v_i, v_j] = \sum_{k=1}^{n} c_{ijk} v_k, \quad i, j = 1, \dots, n \implies d\omega_k = -\tfrac{1}{2} \sum_{i,j=1}^{n} c_{ijk} \omega_i \wedge \omega_j, \tag{C.5}$$

for it follows from (C.4) that

$$\begin{aligned} \langle v_j \wedge v_i, d\omega_k \rangle = \sum c_{ijl} \langle v_l, \omega_k \rangle + v_j \langle v_i, \omega_k \rangle - v_i \langle v_j, \omega_k \rangle \\ = c_{ijk} = \tfrac{1}{2} \langle v_i \wedge v_j, \sum c_{\mu\nu k} \omega_\mu \wedge \omega_\nu \rangle. \end{aligned}$$

We can reverse the implication in (C.5) if we demand that c_{ijk} is antisymmetric in i, j (which is automatic in the left-hand side). The Frobenius condition (C.2) stating that $c_{ijk} = 0$ if $k > r$ and $i, j \le r$ is therefore equivalent to

(C.2)′ $\quad d\omega_j \quad$ is in the ideal of forms generated by $\omega_{r+1}, \dots, \omega_n$, $\quad$ if $j > r$.

Another equivalent formulation of the Frobenius theorem concerns complete integrability of a system of differential equations

$$\partial y_\mu(x)/\partial x_\nu = F_{\mu\nu}(x,y), \quad \nu = 1,\dots,n,\ \mu = 1,\dots,m, \tag{C.6}$$

where $F_{\mu\nu}$ are smooth functions defined in a neighborhood of $(x_0, y_0) \in \mathbf{R}^{n+m}$. The system is called *completely integrable* if there is a smooth solution in a neighborhood of x_0 with $y(x_0)$ given in a neighborhood of y_0. Since the equations give

$$\partial^2 y_\mu/\partial x_\nu \partial x_\kappa = \partial F_{\mu\nu}/\partial x_\kappa + \sum_{\sigma=1}^{m} \partial F_{\mu\nu}/\partial y_\sigma F_{\sigma\kappa}$$

a necessary condition is that in a neighborhood of (x_0, y_0) we have for $\mu = 1,\dots,m$ and $\nu, \kappa = 1,\dots,n$

$$\partial F_{\mu\nu}/\partial x_\kappa + \sum_{\sigma=1}^{m} \partial F_{\mu\nu}/\partial y_\sigma F_{\sigma\kappa} = \partial F_{\mu\kappa}/\partial x_\nu + \sum_{\sigma=1}^{m} \partial F_{\mu\kappa}/\partial y_\sigma F_{\sigma\nu}. \tag{C.7}$$

This condition is also sufficient. In fact, the equations (C.6) mean precisely that the graph of $y(x)$ shall be defined by the equations

$$\omega_\mu = dy_\mu - \sum_{\nu=1}^{n} F_{\mu\nu} dx_\nu = 0.$$

Since

$$\begin{aligned} d\omega_\mu = -\sum_{\kappa,\nu} \partial F_{\mu\nu}/\partial x_\kappa dx_\kappa \wedge dx_\nu - \sum \partial F_{\mu\nu}/\partial y_\sigma (\omega_\sigma + \sum F_{\sigma\kappa} dx_\kappa) \wedge dx_\nu \\ = -\sum \partial F_{\mu\nu}/\partial y_\sigma \omega_\sigma \end{aligned}$$

when (C.7) holds, the condition (C.2)$'$ is fulfilled. Hence we have:

THEOREM C.3. *For the system* (C.6) *to have a solution in a neighborhood of* x_0 *with* $y(x_0)$ *arbitrarily prescribed near* y_0 *it is necessary and sufficient that the integrability condition* (C.7) *is valid near* (x_0, y_0)*. The solution is then unique.*

APPENDIX D. INVARIANTS OF $\mathbf{O}(n)$

A polynomial $f(x_1,\dots,x_N)$ in $x_j \in V = \mathbf{R}^n$, $j=1,\dots,N$, is called a $\mathbf{O}(n)$ invariant if

$$f(Ox_1,\dots,Ox_N) = f(x_1,\dots,x_N), \quad O \in \mathbf{O}(n). \tag{D.1}$$

Sometimes f is then called an *even* invariant, and one calls f an *odd* invariant if

$$f(Ox_1,\dots,Ox_N) = \det O f(x_1,\dots,x_N), \quad O \in \mathbf{O}(n). \tag{D.2}$$

In both cases we have

$$f(Ox_1,\dots,Ox_N) = f(x_1,\dots,x_N), \quad O \in \mathbf{SO}(n). \tag{D.3}$$

where $\mathbf{SO}(n) = \{O \in \mathbf{O}(n); \det O = 1\}$. On the other hand, it follows from (D.3) that

$$f(Ox_1,\dots,Ox_N) = g(x_1,\dots,x_N), \quad O \in \mathbf{O}(n),\ \det O = -1,$$

where g is independent of the choice of O and satisfies (D.3). Hence $f = \frac{1}{2}(f+g) + \frac{1}{2}(f-g)$ is the sum of one even and one odd invariant.

THEOREM D.1. *Every even invariant $f(x_1,\dots,x_N)$ is a polynomial in the scalar products (x_j,x_k), $j,k=1,\dots,n$. Every odd invariant is a sum of such polynomials multiplied by the determinant $\det(x_{i_1},\dots,x_{i_n})$ of n vectors x_i. In particular, there are no odd invariants $\neq 0$ if $N<n$.*

PROOF. The proof is by induction. The statement is obvious when $n=1$ so we assume that $n>1$ and that the theorem has already been proved when n is replaced by $n-1$. Let $V_1 = \mathbf{R}^{n-1}$, identified with the plane $\{x \in \mathbf{R}^n; x_n = 0\}$. If we choose for O the reflection in V_1 which just changes the sign of x_n and leaves V_1 fixed, it follows that $f(x_1,\dots,x_N)=0$ if all $x_j \in V_1$ and f is an odd invariant; if f is an even invariant then $f(x_1,\dots,x_N) = p((x_j,x_k)_{j,k=1,\dots,N})$ by the inductive hypothesis. Since every hyperplane can be carried to the special position V_1 by a transformation in $\mathbf{SO}(n)$, these statements remain true for arbitrary $x_1,\dots,x_N$ which do not span V. In particular, the theorem is true if $N<n$. When $N \geq n$ the proof requires an identity due to Capelli which will now be discussed.

The Capelli identity concerns the differential operators

$$D_{j,k} = \langle x_j, \partial/\partial x_k \rangle$$

which operate on polynomials $f(x_1,\dots,x_N)$ in N vectors $x_1,\dots,x_N \in V = \mathbf{R}^n$. If f is homogeneous in each vector x_j, then $D_{j,k}$ increases the degree in x_j by 1 and decreases that in x_k by 1, if $j \neq k$, but changes no degree if $j = k$. We shall write $D_{j,k}D_{j',k'}$ for the standard composition of the operators, but write $D_{j,k}\#D_{j',k'}$ for the formal (commutative) multiplication, defined as if the coefficients were constant. Capelli's identity states that

$$\text{(D.4)}\quad \begin{vmatrix} D_{N,N}+N-1 & D_{N,N-1} & \dots & D_{N,1} \\ D_{N-1,N} & D_{N-1,N-1}+N-2 & \dots & D_{N,1} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1,N} & D_{1,N-1} & \dots & D_{1,1} \end{vmatrix} f \\ = \begin{cases} 0, & \text{if } N > n \\ \det(x_1,\dots,x_n)\det(\partial/\partial x_1,\dots,\partial/\partial x_n)f, & \text{if } N = n. \end{cases}$$

The diagonal elements are here $D_{k,k}+(k-1)$, the off diagonal elements are $D_{j,k}$, and the expansion of the determinant shall be made so that the elements are multiplied in the order of their columns. To prove (D.4) we first consider the "minors" formed from the last two columns. We have

$$D_{j,2}D_{k,1} - D_{k,2}D_{j,1} = D_{j,2}\#D_{k,1} - D_{k,2}\#D_{j,1} + \delta_{k,2}D_{j,1} - \delta_{j,2}D_{k,1},$$

that is,

$$(D_{j,2}+\delta_{j,2})D_{k,1} - (D_{k,2}+\delta_{k,2})D_{j,1} = D_{j,2}\#D_{k,1} - D_{k,2}\#D_{j,1}.$$

Thus the minors of (D.4) in the left-hand side are the same as in a determinant with diagonal elements $D_{k,k}$ expanded using the formal product. We shall prove by induction for increasing k that this is true for the minors taken from the last $k \ge 2$ columns. When $k = N$ this will prove (D.4), for

$$\det\langle x_j,\xi_k\rangle_{j,k=1}^N = \begin{cases} 0, & \text{if } N > n \\ \det(x_1,\dots,x_n)\det(\xi_1,\dots,\xi_n), & \text{if } N = n. \end{cases}$$

By the inductive hypothesis we have to calculate

$$\begin{aligned}(D_{i_k,k} &+ (k-1)\delta_{i_k,k})D_{i_{k-1},k-1}\#\dots\#D_{i_1,1} \\ &= D_{i_k,k}\#D_{i_{k-1},k-1}\#\dots\#D_{i_1,1} + (k-1)\delta_{i_k,k}D_{i_{k-1},k-1}\#\dots\#D_{i_1,1} \\ &\quad + \sum_{\nu=1}^{k-1}\delta_{i_\nu,k}D_{i_{k-1},k-1}\#\dots\#D_{i_{\nu+1},\nu+1}\#D_{i_k,\nu}\#D_{i_{\nu-1},\nu-1}\#\dots\#D_{i_1,1}.\end{aligned}$$

We shall let $i_1,\dots,i_k$ run through all permutations of k fixed indices, multiply by the sign of the permutation and sum. Now the permutations

$$i_\nu, i_{k-1},\dots,i_{\nu+1}, i_k, i_{\nu-1},\dots,i_1 \quad \text{and} \quad i_k, i_{k-1},\dots,i_1$$

have opposite signs since they differ by one inversion. Hence the terms in the last sum will cancel the preceding one, which completes the proof of the inductive statement and hence of the Capelli identity.

End of Proof of Theorem D.1. With the Capelli identity available we can now finish the proof. By the first part of the proof we may assume that $N \geq n$ and that the theorem has already been proved when N is replaced by $N-1$. Furthermore we may assume that f is homogeneous in each of the vectors $x_1, \dots, x_N$, for if we split f into a sum of polynomials of such separate homogeneities, then each term must be an invariant. Let r_j be the degree of homogeneity with respect to x_j, and let $|r| = \sum_1^N r_j$ be the total degree of homogeneity. We order all multiindices $r = (r_1, \dots, r_N)$ first for increasing $|r|$, and then lexocographically for fixed $|r|$, that is, first according to increasing r_1, then for fixed r_1 according to increasing r_2, and so on. If $r_1 = 0$ then f depends on $N-1$ vectors so the theorem is true then by inductive hypothesis. By still another inductive argument we may therefore assume that $r_1 > 0$ and that the theorem has already been proved for lower values of r.

Now consider (D.4). If $N = n$, then $g = \det(\partial/\partial x_1, \dots, \partial/\partial x_n)f$ is an odd (even) invariant if f is an even (odd) invariant, and the degree of g is lower than the degree of f, so we know that g has the form stated in the theorem. Since

$$\det(x_1, \dots, x_n)^2 = \det(x_j, x_k)_{j,k=1}^n,$$

it follows that the right-hand side has the stated form for any $N \geq n$. In the left-hand side of (D.4) the diagonal term is cf where $c = r_1(r_2+1)\dots(r_N+N-1) \neq 0$ since $r_1 \neq 0$. All other terms contain some factor $D_{j,k}$ with $j \neq k$. For a given term we choose the factor $D_{j,k}$ with $j \neq k$ which is furthest to the right. To the right of it we then have only the factors $D_{k-1,k-1}\dots D_{1,1}$ which multiply f by $r_1 \dots (r_{k-1}+k-2)$. We must have $j > k$ since two factors must not be chosen in the same row. The full expression of the term is of the form $a_{jk}D_{j,k}f$ where a_{jk} is a product of operators $D_{p,q}$. But $D_{j,k}f$ is also an invariant and its total degree is equal to that of f while its lexicographic order is lower since $D_{j,k}$ lowers the degree in x_k at the expense of a raise of degree in x_j for some $j > k$. Thus $D_{j,k}$ has the desired form by inductive hypothesis. So has $a_{jk}D_{j,k}f$, for

$$D_{p,q}(x_r, x_s) = (x_p, x_s)\delta_{q,r} + (x_p, x_r)\delta_{q,s},$$
$$D_{p,q}\det(x_{i_1}, \dots, x_{i_n}) = \sum \delta_{q,i_j}\det(x_{i_1}, \dots, x_{i_{j-1}}, x_p, x_{i_{j+1}}, \dots, x_{i_n}).$$

Hence this is also true of f, which completes the proof.

NOTES AND REFERENCES

Chapter 1. Starting with the notion of curvature for a curve, which is as old as calculus, we study the tangent and normals associated to a curve which leads up to a first encounter with the idea of moving orthogonal frames. It goes back to Euler and is basic for the methods of E. Cartan presented in Chapter IV. We introduce it in the elementary context of curves in a Euclidean space and take the opportunity to give at the same time a first introduction to Lie groups, restricted to subgroups of the linear group. A more extensive discussion is given in Section 5.3, and the reader who wants more information is referred to Chevalley [1], Helgason [1], Sternberg [1], Warner [1] and Weyl [1].

Chapters 2, 3. Principal curvatures of surfaces in $\mathbf{R}^3$ were also known to Euler, but it was Gauss [1] who discovered that the total curvature is an inner property of the surface. Reading Gauss [1] is still very pleasant, and one will recognize many of the ideas presented in Chapters 2 and 3 there. The extension to higher dimensions outlined in Riemann [1], with emphasis on ideas rather than formulas, is equally enjoyable to read. The later formal development of his ideas by Bianchi, Christoffel, Ricci and others is the main theme of Chapters 2 and 3. Our presentation is in the same spirit as Klingenberg [1,2] and to some extent the classic Eisenhart [1,2]. One can find a good presentation in Berger, Gauduchon and Mazet [1], and excellent summaries are given in Aubin [1], Besse [1]. Chapter 2 ends with some detailed analysis of the curvature tensor at a point taken from Atiyah, Hitchin and Singer [1], where much more information is available.

To give a bridge between the study of submanifolds of $\mathbf{R}^n$ in Chapter 2 and abstract manifolds in Chapter 3 we have devoted much space in Chapter 3 to the problem of embedding an abstract manifold in a Euclidean space. After the classical results of Cartan and Janet for the analytic case, the big advance on the problem was made by Nash [1,2]; recently a technically simpler variation has been found by Günther [1,2]. For additional information on global embedding theorems we refer to Gromov and Rohlin [1], Griffiths [1], Griffiths and Jensen [1], Jacobowitz [1,2] and references in these papers.

The discussion of spaces of constant curvature in Section 3.3 has been taken to a large extent from Kobayashi and Nomizu [1]. This is a very careful and useful presentation of basic differential geometry, though provided with very little motivation for the reader.

Chapter 4. In this chapter we leave the Ricci tensor calculus which dominated Chapter 3 and introduce differential forms in the spirit of E. Cartan. The reader can consult Chern [1,2], Kobayashi and Nomizu [1] and Sternberg [1] for a more thorough discussion of these topics. Section 4.3 is almost entirely taken from Chern [1].

Chapter 5. We have started the chapter with a discussion of connections in a vector bundle, as they are encountered naturally by an analyst. However, at the end we also give the more geometrical approach using connection forms on a principal bundle. The purpose of this material is to give the language required to state index theorems for elliptic differential operators and to give the basic concepts required in gauge theory. Also in this chapter one can consult Kobayashi and Nomizu [1] or Sternberg [1] for most of the topics covered.

Chapter 6. The term "metric operator" is taken from Günther [3], and Proposition 6.1.2 can also be found there. The Hadamard construction of parametrices (Hadamard [1], see also e.g. Hörmander [1, Chapter 17]) is made more transparent by using the connection provided by the operator itself, and this is useful in the proof of the local index theorem in Section 6.10. Section 6.2 is devoted to the algebra of differential forms, and Hodge theory is presented in Section 6.3. For a more detailed discussion of differential forms on a Riemannian manifold including Hodge theory, the reader can also consult de Rham [1] or Warner [1]. In Section 6.4 the Hadamard construction of parametrices is adapted to heat equations associated to metric operators. This gives the analytical tools required for the discussion of Hirzebruch's signature theorem in Section 6.5. (For additional background to this result one should consult Hirzebruch [1].) The presentation here follows Atiyah, Bott and Patodi [1] in principle, but we have substituted the parametrix construction of Hadamard for the application of pseudodifferential operators. Sections 6.6–6.10 are devoted to a direct proof of the local index formula for Dirac operators, following Bismut [1] and particularly Getzler [1, 2]. The presentation owes much to Roe [1], and we have also benefited from some lecture notes of M. Taylor.

Unwritten chapters. One of the possible directions for a continuation is to cover the solution of the Yamabe problem, following the excellent exposition in Lee and Parker [1]. After this introduction to non-linear problems in geometry one might study some gauge theory. Another natural direction is to study pseudo-Riemannian manifolds of Lorentz signature. There are a number of papers exploiting conformal invariance to prove global existence theorems for non-linear hyperbolic systems, most recently Christodoulou and Klainerman [1] where the stability of Minkowski space under small perturbations is proved. This requires of course some preparations concerning general relativity theory. The classical survey by Pauli [1] is still very readable. More recent information can be found in Bergmann [1], but this reference is directed towards mathematical physicists rather than mathematicians interested in physics. With some background from these sources, including the Schwarzschild solution, the way is also open to discuss the interesting proof of the positive mass conjecture in Schoen and Yau [1,2], Witten [1] and Parker and Taubes [1].

Bibliography

M. F. Atiyah [1], *Classical groups and classical differential operators on manifolds*, Differential operators on manifolds, Edition Cremonese, Roma, 1975, CIME conference in Varenna September 1975..

M. F. Atiyah, R. Bott and V. K. Patodi [1], *On the heat equation and the index theorem*, Inv. math. **19** (1973), 279–330, Correction **28** (1975), 277–280..

M. F. Atiyah, R. Bott and A. Shapiro, *Clifford modules*, Topology **3** (1964), 3–38, Suppl. 1..

M. F. Atiyah, N. J. Hitchin and I. M. Singer [1], *Self-duality in four-dimensional Riemannian geometry*, Proc. R. Soc. Lond. A. **362** (1978), 425–461.

T. Aubin [1], *Nonlinear analysis on manifolds. Monge-Ampère equations*, Springer-Verlag, New York, Heidelberg, Berlin, 1982.

E. Berger, R. Bryant and Ph. Griffiths [1], *The Gauss equations and rigidity of isometric embeddings*, Duke Math. J. **50** (1983), 803–892.

M. Berger, P. Gauduchon and E. Mazet [1], *Le spectre d'une variété Riemannienne*, Springer-Verlag, 1971, SLN 194..

P. G. Bergmann [1], *The general theory of relativity*, Handbuch der Physik IV, Springer-Verlag, 1962.

A. L. Besse [1], *Einstein Manifolds*, Springer-Verlag, 1987, Erg. d. Math. Band 10..

J.M. Bismut [1], *The Atiyah-Singer theorem for classical elliptic operators: a probabilistic approach. I. The index theorem.*, J. Func. Anal. **57** (1984), 56–99.

Yu. D. Burago and V. A. Zalgaller [1], *Geometric inequalities*, Springer-Verlag, 1988, Grundl.d.Math. Wiss. 285..

S. S. Chern [1], *On the curvature integral in a Riemannian manifold*, Ann. of Math. **46** (1945), 674–684.

S.S. Chern [2], *Topics in differential geometry*, Inst. for Adv. Study, Princeton, N.J., 1951.

C. Chevalley [1], *Theory of Lie groups I*, Princeton University Press, Princeton, N.J., 1946.

D. Christodoulou and S. Klainerman [1], *The global nonlinear stability of the Minkowski space*, Ann. of Math., Manuscript of 565 pages, to appear..

H. L. Cycon, R. G. Froese, W. Kirsch and B. Simon [1], *Schrödinger operators with applications to quantum mechanics and global geometry*, Springer-Verlag, 1987, Texts and Monographs in Physics..

N. V. Efimov [1], *Generation of singularities on surfaces of negative curvature*, Mat. Sb. **64** (1964), 286–320, (Russian; translation in AMS translations.).

L. Eisenhart [1], *Introduction to differential geometry*, Princeton University Press, Princeton, N. J., 1940.

L. Eisenhart [2], *Riemannian geometry*, Princeton University Press, Princeton, N. J., 1926.

C. F. Gauss [1], *Disquisitiones generales circa superficies curvas*, Comm. soc. reg. scient. Gött. **6** (1823–27), German translation in Ostwald's Klassiker der exakten Wissenschaften 5(1889), English translation in "General investigations of curved surfaces", Raven Press, Hewlett, N.Y. (1965), which also contains an earlier version..

J. C. H. Gerretsen [1], *Lectures on tensor calculus and differential geometry*, P. Noordhoff N. V., Groningen, 1962.

E. Getzler [1], *Pseudodifferential operators on supermanifolds and the Atiyah-Singer index theorem*, Comm. Math. Phys. **92** (1983), 163–178.

E. Getzler [2], *A short proof of the local Atiyah-Singer index theorem*, Topology **25** (1986), 111–117.

P. Gilkey [1], *Curvature and the eigenvalues of the Laplacian for elliptic complexes*, Adv. in Math. **10** (1973), 344–382.

Ph. A. Griffiths [1], *Exterior differential systems and the calculus of variations*, Birkhäuser, Boston, Basel, Stuttgart, 1983, Progress in Math. 25..

Ph. A. Griffiths and G. R. Jensen [1], *Differential systems and isometric embeddings*, Princeton University Press, Princeton, N. J., 1987, Annals of Math. Studies 114..

M. L. Gromov and V. A. Rokhlin [1], *Embeddings and immersions in Riemannian geometry.*, Russian Math. Surveys **25:5** (1970), 1–57, Russian original in Usp. Mat. Nauk 25:5(155), 3–62 (1970)..

M. Günther [1], *On the perturbation problem associated to isometric embeddings of Riemannian manifolds*, Ann. Global Anal. Geom. **7** (1989), 69–77.

M. Günther [2], *Zum einbettungssatz von J. Nash*, Math. Nachr. **144** (1989), 165–187.

P. Günther [3], *Huyghens' principle and hyperbolic equations*, Academic Press, Boston and London, 1988.

J. Hadamard, *Le problème de Cauchy et les équations aux dérivées partielles hyperboliques*, Hermann, Paris, 1932.

S. Helgason [1], *Differential geometry and symmetric spaces*, Academic Press, New York, London, 1962.

F. Hirzebruch [1], *Topological methods in algebraic geometry*, Springer-Verlag, Berlin, Göttingen, Heidelberg, 1966, Grundl. d. Math. Wiss. 131..

L. Hörmander [1], *The analysis of linear partial differential operators III*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985.

N. Kuipers [1], *On C^1 isometric imbeddings*, Indag. Math. **18** (1955), 545–556, 683–689.

H. Jacobowitz [1], *Local isometric embeddings*, Seminar on differential geometry, Princeton University Press, 1982, Editor S.-T. Yau, Ann. of Math. Studies 102..

H. Jacobowitz [2], *Local isometric embeddings of surfaces into Euclidean four space*, Ind. Univ. Math. J. **21** (1971), 249–254.

W. Klingenberg [1], *Eine Vorlesung über Differentialgeometrie*, Springer-Verlag, Berlin, Göttingen, Heidelberg, 1973, Heidelberger Taschenbücher..

W. Klingenberg [2], *Riemannian geometry*, Walter de Gruyter, Berlin, 1982.

S. Kobayashi and K. Nomizu [1], *Foundations of differential geometry I-II*, Interscience Publ, New York, N. Y., London, 1963, 1969.

J.M. Lee and T. H. Parker [1], *The Yamabe problem*, Bull. Amer. Math. Soc. **17** (1987), 37–91.

A. E. Liber [1], *On a class of Riemannian spaces having constant negative curvature*, Uch. Zap. Saratov. Gos. Univ. Ser. Fiz.-Mat. **1:2** (1938), 105–122.

J. Milnor [1], *On total curvature of closed space curves*, Math. Scand. **1** (1953), 289–296.

J. Nash [1], *The imbedding problem for Riemannian manifolds*, Ann. of Math. **63** (1956), 20–63.

J. Nash [2], *C^1 isometric imbeddings*, Ann. of Math. **60** (1954), 383–396.

T. Parker and C. H. Taubes [1], *On Witten's proof of the positive energy theorem*, Comm. Math. Phys. **84** (1982), 223–238.

V. K. Patodi [1], *Curvature and the eigenforms of the Laplace operator*, J. Diff. Geometry **5** (1971), 233–249.

W. Pauli [1], *Relativitätstheorie*, Teubner Verlag, Leipzig, Berlin, 1921, Reprint from Enc. d. math. Wiss. V 19..

G. de Rham, *Variétés différentiables*, Hermann, Paris, 1960, Actualités scient. et ind. 1222.

B. Riemann [1], *Über die Hypothesen welche der Geometrie zu Grunde liegen*, Werke, ed. H. Weber, Teubner Verlag, Leipzig, 1892, pp. 372–387.

J. Roe, *Elliptic operators, topology and asymptotic methods*, Longman Scientific & Technical, Harlow, Essex and New York, 1988, Pitman Research Notes in Mathematics 179..

R. Schoen and S.-T. Yau [1], *On the proof of the positive dmass conjecture in general relativity theory*, Comm. math. phys. **65** (1979), 45–76.

R. Schoen and S.-T. Yau [2], *Proof of the positive mass theorem*, Comm. math. phys. **79** (1981), 231–260.

F. Schur [1], *Ueber die Deformation der Räume constanten Riemann'schen Krümmingsmasses*, Math. Ann. **27** (1886), 163–176.

S. Sternberg [1], *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, N.J., 1964.

F. W. Warner [1], *Foundations of differentiable manifolds and Lie groups*, Scott, Foresman and Co., Glenview, Ill., London, 1971.

R. Weitzenböck, *Invariantentheorie*, Nordhoff, Groningen, 1923.

H. Weyl [1], *Classical groups*, Princeton University Press, Princeton, N. J., 1946.

E. Witten [1], *A new proof of the positive energy theorem*, Comm. math. phys. **80** (1981), 381–402.

Index