Solution Manual
for
Manifolds, Tensors, and Forms

Paul Renteln

Department of Physics
California State University
San Bernardino, CA 92407
and
Department of Mathematics
California Institute of Technology
Pasadena, CA 91125
prenteln@csusb. edu

CAMBRIDGE
UNIVERSITY PRESS

Contents

1
Linear algebra

1.1 We have

$$0 = c_1(1,1) + c_2(2,1) = (c_1 + 2c_2, c_1 + c_2)$$
$$\Rightarrow \quad c_2 = -c_1 \quad \Rightarrow \quad c_1 - 2c_1 = 0 \quad \Rightarrow \quad c_1 = 0 \quad \Rightarrow \quad c_2 = 0,$$

so $(1, 1)$ and $(2, 1)$ are linearly independent. On the other hand,

$$0 = c_1(1,1) + c_2(2,2) = (c_1 + 2c_2, c_1 + 2c_2)$$

can be solved by choosing $c_1 = 2$ and $c_1 = -1$, so $(1, 1)$ and $(2, 2)$ are linearly dependent (because c_1 and c_2 are not necessarily zero).

1.2 Subtracting gives

$$0 = \sum_i v_i e_i - \sum_i v_i' e_i = \sum_i (v_i - v_i') e_i.$$

But the e_i's are a basis for V, so they are linearly independent, which implies $v_i - v_i' = 0$.

1.3 Let $V = U \oplus W$, and let $E := \{e_i\}_{i=1}^n$ be a basis for U and $F := \{f_j\}_{j=1}^m$ a basis for W. Define a collection of vectors $G := \{g_k\}_{k=1}^{n+m}$ where $g_i = e_i$ for $1 \leq i \leq n$ and $g_{n+i} = f_i$ for $1 \leq i \leq m$. Then the claim follows if we can show G is a basis for V. To that end, assume

$$0 = \sum_{i=1}^{n+m} c_i g_i = \sum_{i=1}^{n} c_i e_i + \sum_{i=1}^{m} c_i f_i.$$

The first sum in the rightmost expression lives in U and the second sum lives in W, so by the uniqueness property of direct sums, each sum must vanish by itself. But then by the linear independence of E and F, all the constants c_i must vanish. Therefore G is linearly independent. Moreover, every vector $v \in V$ is of the form $v = u + w$ for some $u \in U$ and $w \in W$, each of which can be written as a linear combination of the g_i's. Hence the g_i's form a basis for V.

1.4 Let S be any linearly independent set of vectors with $|S| < n$. The claim is that we can always find a vector $v \in V$ so that $S \cup \{v\}$ is linearly independent. If not, consider the sum

$$cv + \sum_{i=1}^{|S|} c_i s_i = 0,$$

where $s_i \in S$. Then some of the c_i's are nonzero. We cannot have $c = 0$, because S is linearly independent. Therefore v lies in the span of S, which says that $\dim V = |S| < n$, a contradiction.

1.5 Let $S, T : V \to W$ be two linear maps, and let $\{e_i\}$ be a basis for V. Assume $Se_i = Te_i$ for all i, and that $v = \sum_i a_i e_i$. Then $Sv = \sum_i a_i Se_i = \sum_i a_i Te_i = Tv$.

1.6 Let $v_1, v_2 \in \ker T$. Then $T(av_1 + bv_2) = aTv_1 + bTv_2 = 0$, so $\ker T$ is closed under linear combinations. Moreover $\ker T$ contains the zero vector of V. All the other vector space properties are easily seen to follow, so $\ker T$ is a subspace of V. Similarly, let $w_1, w_2 \in \operatorname{im} T$ and consider $aw_1 + bw_2$. There exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$, so $T(av_1 + bv_2) = aTv_1 + bTv_2 = aw_1 + bw_2$, which shows that $\operatorname{im} T$ is closed under linear combinations. Moreover, $\operatorname{im} T$ contains the zero vector, so $\operatorname{im} T$ is a subspace of W.

1.7 For any two vectors v_1 and v_2 we have

$$Tv_1 = Tv_2 \quad \Rightarrow \quad T(v_1 - v_2) = 0 \quad \Rightarrow \quad v_1 - v_2 = 0 \quad \Rightarrow \quad v_1 = v_2.$$

Assume the kernel of T consists only of the zero vector. Then for any two vectors v_1 and v_2, $T(v_1 - v_2) = 0$ implies $v_1 - v_2 = 0$, which is equivalent to saying that $Tv_1 = Tv_2$ implies $v_1 = v_2$, namely that T is injective. The converse follows similarly.

1.8 Let V and W be two vector spaces of the same dimension, and choose a basis $\{e_i\}$ for V and a basis $\{f_i\}$ for W. Let $T : V \to W$ be the map that sends e_i to f_i, extended by linearity. Then the claim is that T is an isomorphism. Let $v = \sum_i a_i e_i$ be a vector in V. If $v \in \ker T$, then $0 = Tv = \sum_i a_i Te_i = \sum_i a_i f_i$. By linear independence, all the a_i's vanish, which means that the kernel of T consists only of the zero vector, and hence by Exercise 1.7, T is injective. Also, if $w = \sum_i a_i f_i$, then $w = \sum_i a_i Te_i = T \sum_i a_i e_i$, which shows that T is also surjective.

1.9 a. Let $v \in V$ and define $w := \pi(v)$ and $u := (1 - \pi)(v)$. Then $\pi(u) = (\pi - \pi^2)(v) = 0$, so $v = w + u$ with $w \in \operatorname{im}\pi$ and $u \in \ker\pi$. Now suppose $x \in \ker \pi \cap \operatorname{im} \pi$. Then there is a $y \in V$ such that $x = \pi(y)$. But then $0 = \pi(x) = \pi^2(y) = \pi(y) = x$.

b. Let $\{f_i\}$ be a basis for W, and complete it to a basis of V by adding a linearly independent set of vectors $\{g_j\}$. Let U be the subspace of V spanned by the g_i's. With these choices, any vector $v \in V$ can be written uniquely as $v = w + u$, where $w \in W$ and $u \in U$. Define a linear map $\pi : V \to V$ by $\pi(v) = w$. Obviously $\pi(w) = w$, so $\pi^2 = \pi$.

1.10 Clearly, $T0 = 0$, so $T^{-1}0 = 0$. Let $Tv_1 = v_1'$ and $Tv_2 = v_2'$. Then

$$aT^{-1}v_1' + bT^{-1}v_2' = av_1 + bv_2 = (T^{-1}T)(av_1 + bv_2) = T^{-1}(av_1' + bv_1'),$$

which shows that T^{-1} is linear.

1.11 The identity map $I : V \to V$ is clearly an automorphism. If $S \in \operatorname{Aut} V$ then $S^{-1}S = SS^{-1} = I$. Finally, if $S, T \in \operatorname{Aut} V$, then ST is invertible, with inverse $(ST)^{-1} = T^{-1}S^{-1}$. (Check.) This implies that $ST \in \operatorname{Aut} V$. (Associativity is automatic.)

1.12 By exactness, the kernel of φ_1 is the image of φ_0. But the image of φ_0 consists only of the zero vector (as its domain consists only of the zero vector). Hence the kernel of φ_1 is trivial, so by Exercise 1.7, φ_1 must be injective. Again by exactness, the kernel of φ_3 is the image of φ_2. But φ_3 maps everything to zero, so $V_3 = \ker \varphi_1$, and hence $V_3 = \operatorname{im} V_2$, which says that φ_2 is surjective. The converse follows by reversing the preceding steps. As for the last assertion, φ is both injective and surjective, so it is an isomorphism.

1.13 If T is injective then $\ker T = 0$, so by the rank/nullity theorem $\operatorname{rk} T = \dim V = \dim W$, which shows that T is surjective as well.

1.14 The rank of a linear map is the dimension of its image. There is no way that the image of ST can be larger than that of either S or T individually, because the dimension of the image of a map cannot exceed the dimension of its domain.

1.15 If $v' \in [v]$ then $v' = v + u$ for some $u \in U$. By linearity $\varphi(v') = \varphi(v) + w$ for some $w \in W$, so $[\varphi(v')] = [\varphi(v) + w] = [\varphi(v)]$.

1.16 Pick a basis $\{e_i\}$ for V. Then,

$$\sum_i (ST)_{ij} e_i = (ST)e_j = S(\sum_k T_{kj} e_k) = \sum_k T_{kj} S e_k = \sum_{ik} T_{kj} S_{ik} e_i.$$

Hence

$$(ST)_{ij} = \sum_k S_{ik} T_{kj} = (\boldsymbol{ST})_{ij},$$

which shows that $ST \to \boldsymbol{ST}$.

1.17 The easiest way to see this is just to observe that the identity automorphism I is represented by the identity matrix $\boldsymbol{I}$ (in any basis). Suppose T^{-1} is represented by $\boldsymbol{U}$ in some basis. Then by the results of Exercise 1.16,

$$TT^{-1} \to \boldsymbol{TU}.$$

But $TT^{-1} = I$, so $\boldsymbol{TU} = \boldsymbol{I}$, which shows that $\boldsymbol{U} = \boldsymbol{T}^{-1}$.

1.18 Choose a basis $\{e_i\}$ for V. Then by definition,

$$Te_j = \sum_i T_{ij} e_i.$$

It follows that Te_j is represented by the j^{th} column of $\boldsymbol{T}$, so the maximum number of linearly dependent vectors in the image of T is precisely the maximum number of linearly independent columns of $\boldsymbol{T}$.

1.19 Suppose $\sum_i c_i \theta_i = 0$. By linearity of the dual pairing,

$$0 = \left\langle e_j, \sum_i c_i \theta_i \right\rangle = \sum_i c_i \left\langle e_j, \theta_i \right\rangle = \sum_i c_i \delta_{ij} = c_j,$$

so the θ_j's are linearly independent.

Now let $f \in V^*$. Define $f(e_j) =: a_j$ and introduce a linear functional $g := \sum_i a_i \theta_i$. Then

$$g(e_j) = \langle g, e_j \rangle = \sum_i a_i \delta_{ij} = a_j,$$

so $f = g$ (two linear functionals that agree on a basis agree everywhere). Hence the θ_j's span.

1.20 Suppose $f(v) = 0$ for all v. Let $f = \sum_i f_i \theta_i$ and $v = e_j$. Then $f(v) = f(e_j) = f_j = 0$. This is true for all j, so $f = 0$. The other proof is similar.

1.21 Let $w \in W$ and $\theta_1, \theta_2 \in \operatorname{Ann} W$. Then

$$(a\theta_1 + b\theta_2)(w) = a\theta_1(w) + b\theta_2(w) = 0,$$

so $\operatorname{Ann} W$ is closed under linear combinations. Moreover, the zero functional (which sends every vector to zero) is clearly in $\operatorname{Ann} W$, so $\operatorname{Ann} W$ is a subspace of V^*.

Conversely, let $U^* \subseteq V^*$ be a subspace of V^*, and define

$$W := \{v \in V : f(v) = 0, \text{ for all } f \in U^*\}.$$

If $f \in U^*$ then $f(v) = 0$ for all $v \in W$, so $f \in \operatorname{Ann} W$. It therefore suffices to prove that $\dim U^* = \dim \operatorname{Ann} W$. Let $\{f_i\}$ be a basis for U^*, and let $\{e_i\}$ be its dual basis, satisfying $f_i(e_j) = \delta_{ij}$. Obviously, $e_i \notin W$. Thus $\dim W = \dim V - \dim U^*$. On the other hand, let $\{w_i\}$ be a basis for W and complete it to a basis for V: $\{w_1, \ldots, w_{\dim W}, e_{\dim W+1}, \ldots, e_{\dim V}\}$. Let $\{u_i\}$ be a basis for $\operatorname{Ann} W$. Then $u_i(e_j) \neq 0$, else $e_j \in W$. So $\dim \operatorname{Ann} W = \dim V - \dim W$.

1.22 a. The map is well defined, because if $[v'] = [v]$ then $v' = v + w$ for some $w \in W$, so $\varphi(f)([v']) = f(v') = f(v + w) = f(v) + f(w) = f(v) = \varphi(f)([v])$. Moreover, if $\varphi(f) = \varphi(g)$ then for any $v \in V$, $0 = \varphi(f - g)([v]) = (f - g)(v)$, so $f = g$. But the proof of Exercise 1.21 shows that $\dim \operatorname{Ann} W = \dim(V/W) = \dim(V/W)^*$, so φ is an isomorphism.

b. Suppose $[g] = [f]$ in $V^*/\operatorname{Ann} W$. Then $g = f + h$ for some $h \in \operatorname{Ann} W$. So $\pi^*([g])(v) = g(\pi(v)) = f(\pi(v)) + h(\pi(v)) = f(\pi(v)) = \pi^*([f])(v)$. Moreover, if $\pi^*([f]) = \pi^*([g])$ then $f(\pi(v)) = g(\pi(v))$ or $(f - g)(\pi(v)) = 0$, so $f = g$ when restricted to W. Dimension counting shows that π^* is an isomorphism.

1.23 Let g be the standard inner product on $\mathbb{C}^n$ and let $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$. Then

$$\begin{aligned} g(u, av + bw) &= \sum_i \overline{u_i}(av_i + bw_i) \\ &= a \sum_i \overline{u_i} v_i + b \sum_i \overline{u_i} w_i \\ &= a g(u, v) + b g(u, w). \end{aligned}$$

Also,

$$g(v, u) = \sum_i \overline{v_i} u_i = \overline{\sum_i \overline{u_i} v_i} = \overline{g(u, v)}.$$

Assume $g(u, v) = 0$ for all v. Let v run through all the vectors $v^{(i)} = (0, \ldots, 1, \ldots, 0)$, where the '1' is in the i^{th} place. Plugging into the definition of g gives $u_i = 0$ for all i, so $u = 0$. Thus g is indeed an inner product. The same proof works equally well for the Euclidean and Lorentzian inner products.

Again consider the standard inner product on $\mathbb{C}^n$. Then

$$g(u, u) = \sum_i \overline{u_i} u_i = \sum_i |u_i|^2 \geq 0,$$

because the modulus squared of a complex number is always nonnegative, so g is nonnegative definite. Moreover, the only way we could have $g(u, u) = 0$ is if each u_i were zero, in which case we would have $u = 0$. Thus g is positive definite. The same proof applies in the Euclidean case, but fails in the Lorentzian case because then

$$g(u, u) = -u_0^2 + \sum_{i=1}^{n-1} u_i^2,$$

and it could happen that $g(u, u) = 0$ but $u \neq 0$. (For example, let $u = (1, 1, 0, \ldots, 0)$.)

1.24 We have

$$\begin{aligned}(A^*(af + bg))(v) &= (af + bg)(Av) = af(Av) + bg(Av) \\ &= a(A^* f)(v) + b(A^* g)(v) = (aA^* f + bA^* g)(v),\end{aligned}$$

so A^* is linear. (The other axioms are just as straightforward.)

1.25 We have

$$\langle A^* e_j^*, e_i \rangle = \sum_k \langle (A^*)_{kj} e_k^*, e_i \rangle = \sum_k (A^*)_{kj} \delta_{ki} = (A^*)_{ij},$$

while

$$\langle e_j^*, Ae_i \rangle = \sum_k \langle e_j^*, A_{ki} e_k \rangle = \sum_k A_{ki} \delta_{jk} = A_{ji},$$

so the matrix representing A^* is just the transpose of the matrix representing A.

1.26 We have

$$\langle A^\dagger e_j, e_i \rangle = \sum_k \langle (A^\dagger)_{kj} e_k, e_i \rangle = \sum_k \overline{(A^\dagger)_{kj}} \delta_{ki} = \overline{(A^\dagger)_{ij}},$$

while

$$\langle e_j^*, Ae_i \rangle = \sum_k \langle e_j^*, A_{ki} e_k \rangle = \sum_k A_{ki} \delta_{jk} = A_{ji},$$

which gives

$$(A^\dagger)_{ij} = \overline{A}_{ji}.$$

1.27 Let $w = \sum_i a_i v_i$ (where not all the a_i's vanish) and suppose $\sum_i c_i v_i + cw = 0$. The latter equation may be solved by choosing $c = 1$ and $c_i = -a_i$, so the set $\{v_1, \ldots, v_n, w\}$ is linearly dependent. Conversely, suppose $\{v_1, \ldots, v_n, w\}$ is linearly dependent. Then the equations $\sum_i c_i v_i + cw = 0$ have a nontrivial solution $(c, c_1, \ldots, c_n)$. We must have $c \neq 0$ else the set $\{v_i\}$ is not linearly independent. But then $w = -\sum_i (c_i/c) v_i$.

1.28 Obviously, the monomials span V, so we need only check linear independence. Assume

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 = 0.$$

The zero on the right side represents the zero vector, namely the polynomial that is zero for all values of x. In other words, this equation must hold for all values of x. In particular, it must hold for $x = 0$. Plugging in gives $c_0 = 0$. Next let $x = 1$ and $x = -1$, giving $c_1 + c_2 + c_3 = 0$ and $-c_1 + c_2 - c_3 = 0$. Adding and subtracting the latter two equations gives $c_2 = 0$ and $c_1 + c_3 = 0$. Finally, choose $x = 2$ to get $2c_1 + 8c_3 = 0$. Combining this with $c_1 + c_3 = 0$ gives $c_1 = c_3 = 0$.

1.29 We must show exactness at each space. Clearly the sequence is exact at $\ker T$, because the inclusion map $\iota : \ker T \to V$ is injective, so only zero gets sent to zero. By definition, the kernel of T is $\ker T$, namely the image of ι, so the sequence is exact at V. Let $\pi : W \to \operatorname{coker} T$ be the projection map onto the quotient $W/\operatorname{im} T$. Then by definition π kills everything in $\operatorname{im} T$, so the sequence is exact at W. Finally, π is surjective onto the quotient, so the sequence is exact at $\operatorname{coker} T$.

1.30 Write the exact sequence together with its maps

$$0 \longrightarrow V_0 \xrightarrow{\varphi_0} V_1 \xrightarrow{\varphi_1} \cdots \xrightarrow{\varphi_{n-1}} V_n \longrightarrow 0$$

and set $\varphi_{-1} = \varphi_n = 0$. By exactness, $\operatorname{im} \varphi_{i-1} = \ker \varphi_i$. But the rank/nullity theorem gives

$$\dim V_i = \dim \ker \varphi_i + \dim \operatorname{im} \varphi_i .$$

Hence,

$$\begin{aligned}
\sum_i (-1)^i \dim V_i &= \sum_i (-1)^i (\dim \ker \varphi_i + \dim \operatorname{im} \varphi_i) \\
&= \sum_i (-1)^i (\dim \operatorname{im} \varphi_{i-1} + \dim \operatorname{im} \varphi_i) \\
&= 0,
\end{aligned}$$

because the sum is telescoping.

1.31 An arbitrary term of the expansion of $\det A$ is of the form

$$(-1)^\sigma A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{n\sigma(n)}. \tag{1}$$

As each number from 1 to n appears precisely once among the set $\sigma(1), \sigma(2), \dots, \sigma(n)$, the product may be rewritten (after some rearrangement) as

$$(-1)^\sigma A_{\sigma^{-1}(1)1} A_{\sigma^{-1}(2)2} \dots A_{\sigma^{-1}(n)n}, \tag{2}$$

where σ^{-1} is the inverse permutation to σ. For example, suppose $\sigma(5) = 1$. Then there would be a term in (1) of the form $A_{5\sigma(5)} = A_{51}$. This term appears first in (2), as $\sigma^{-1}(1) = 5$. Since a permutation and its inverse both have the same sign (because $\sigma\sigma^{-1} = e$ implies $(-1)^{\sigma}(-1)^{\sigma^{-1}} = 1$), Equation (2) may be written

$$(-1)^{\sigma^{-1}} A_{\sigma^{-1}(1)1} A_{\sigma^{-1}(2)2} \dots A_{\sigma^{-1}(n)n}. \tag{3}$$

Hence

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma^{-1}} A_{\sigma^{-1}(1)1} A_{\sigma^{-1}(2)2} \dots A_{\sigma^{-1}(n)n}. \tag{4}$$

As σ runs over all the elements of S_n, so does σ^{-1}, so (4) may be written

$$\det A = \sum_{\sigma^{-1} \in \mathfrak{S}_n} (-1)^{\sigma^{-1}} A_{\sigma^{-1}(1)1} A_{\sigma^{-1}(2)2} \dots A_{\sigma^{-1}(n)n}. \tag{5}$$

But this is just $\det A^T$.

1.32 By (1.46) the coefficient of A_{11} in $\det A$ is

$$\sum_{\sigma' \in \mathfrak{S}_n} (-1)^{\sigma'} A_{2\sigma'(2)} \dots A_{n\sigma'(n)}, \tag{1}$$

where σ' means a general permutation in S_n that fixes $\sigma(1) = 1$. But this means the sum in (1) extends over all permutations of the numbers $\{2, 3, \dots, n\}$, of which there are $(n-1)!$. A moment's reflection reveals that (1) is nothing more than the determinant of the matrix obtained from A by removing the first row and first column, namely $A(1|1)$.

Now consider a general element A_{ij}. What is its coefficient in $\det A$? Well, consider the matrix A' obtained from A by moving the i^{th} row up to the first row. To get A' we must execute $i - 1$ adjacent row flips, so $\det A' = (-1)^{i-1} \det A$. Now consider the matrix A'' obtained from A' by moving the j^{th} column left to the first column. Again we have $\det A'' = (-1)^{j-1} \det A'$. So $\det A'' = (-1)^{i+j} \det A$. The element A_{ij} appears in the (11) position in A'', so by the reasoning used above, its coefficient in $\det A''$ is just $\det A''(1|1) = \det A(i|j)$. Hence, the coefficient of A_{ij} in $\det A$ is $(-1)^{i+j} \det A(i|j) = \widetilde{A}_{ij}$.

Next consider the expression

$$A_{11}\widetilde{A}_{11} + A_{12}\widetilde{A}_{12} + \dots + A_{1n}\widetilde{A}_{1n}, \tag{2}$$

which is (1.57) with $i = 1$. Thinking of the A_{ij} as independent variables, each term in (2) is distinct (because, for example, only the first term contains A_{11}, *etc.*). Moreover, each term appears in (2) precisely as it appears in $\det A$ (with the correct sign and correct products of elements of A). Finally, (2) contains $n(n-1)! = n!$ terms, which is the number that appear in $\det A$. So (2) must *be*

$\det A$. As there was nothing special about the choice $i = 1$, (1.57) is proved. Equation (1.58) is proved similarly.

1.33 Suppose we begin with a matrix A and substitute for its i^{th} row a new row of elements labeled B_{ij}, where j runs from 1 to n. Now, the cofactors of the B_{ij} in the new matrix are obviously the same as those of the A_{ij} in the old matrix, so we may write the determinant of the new matrix as, for instance,

$$B_{i1}\widetilde{A}_{i1} + B_{i2}\widetilde{A}_{i2} + \cdots + B_{in}\widetilde{A}_{in}. \tag{1}$$

Of course, we could have substituted a new j^{th} column instead, with similar results.

If we were to let the B_{ij} be the elements of any row of A other than the i^{th}, then the expression in Equation (1) would vanish, as the determinant of any matrix with two identical rows is zero. This gives us the following result:

$$A_{k1}\widetilde{A}_{i1} + A_{k2}\widetilde{A}_{i2} + \cdots + A_{kn}\widetilde{A}_{in} = 0, \qquad k \neq i. \tag{2}$$

Again, a similar result holds for columns. (The cofactors appearing in (1) are called *alien cofactors*, because they are the cofactors properly corresponding to the elements A_{ij}, $j = 1, \dots, n$, of the i^{th} row of A rather than the k^{th} row.) We may summarize (2) by saying that expansions in terms of alien cofactors vanish identically.

Consider the ik^{th} element of $A(\operatorname{adj} A)$:

$$[A(\operatorname{adj} A)]_{ik} = \sum_{j=1}^{n} A_{ij}(\operatorname{adj} A)_{jk} = \sum_{j=1}^{n} A_{ij}\widetilde{A}_{kj}.$$

If $i \neq k$ this is an expansion in terms of alien cofactors and vanishes. If $i = k$ then this is just the determinant of A. Hence $[A(\operatorname{adj} A)]_{ik} = (\det A)\delta_{ik}$. This proves the first half. To prove the second half, note that $(\operatorname{adj} A)^T = (\operatorname{adj} A^T)$. That is, the transpose of the adjugate is the adjugate of the transpose. (Just trace back the definitions.) Hence, using the result (whose easy proof is left to the reader) that $(AB)^T = B^T A^T$ for any matrices A and B,

$$[(\operatorname{adj} A)A)]^T = A^T(\operatorname{adj} A)^T = A^T \operatorname{adj} A^T = (\det A^T)I = (\det A)I. \tag{3}$$

1.34 By (1.59),

$$\boldsymbol{A}(\operatorname{adj} \boldsymbol{A}) = (\operatorname{adj} \boldsymbol{A})\boldsymbol{A} = (\det \boldsymbol{A})\boldsymbol{I},$$

so if $\boldsymbol{A}$ is nonsingular, then the inverse of $\boldsymbol{A}$ is $\operatorname{adj} \boldsymbol{A}/\det \boldsymbol{A}$, and if $\boldsymbol{A}$ is invertible, then multiplying both sides of this equation by $\boldsymbol{A}^{-1}$ gives $\operatorname{adj} \boldsymbol{A} = (\det \boldsymbol{A})\boldsymbol{A}^{-1}$, which implies $\boldsymbol{A}$ is nonsingular (because the adjugate cannot vanish identically). Next, suppose $\boldsymbol{A}\boldsymbol{v} = 0$. If $\boldsymbol{A}$ were invertible, then multiplying both sides of this equation by $\boldsymbol{A}^{-1}$ would give $\boldsymbol{v} = 0$. So $\boldsymbol{v}$ is nontrivial if and only if $\boldsymbol{A}$ is not invertible, which holds if and only if $\det \boldsymbol{A} = 0$.

1.35 $\boldsymbol{A}$ is nonsingular, so

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \frac{1}{\det \boldsymbol{A}}(\operatorname{adj} \boldsymbol{A})\boldsymbol{b}.$$

But expanding by the i^{th} column gives

$$\det \boldsymbol{A}^{(i)} = \sum_j b_j \widetilde{A}_{ji} = \sum_j (\operatorname{adj} \boldsymbol{A})_{ij} b_j,$$

and therefore

$$x_i = \frac{\det \boldsymbol{A}^{(i)}}{\det \boldsymbol{A}}.$$

1.36 From (1.57),

$$\frac{\partial}{\partial A_{12}}(\det \boldsymbol{A}) = \frac{\partial}{\partial A_{12}}(A_{11}\widetilde{A}_{11} + A_{12}\widetilde{A}_{12} + \cdots) = \widetilde{A}_{12},$$

because A_{12} only appears in the second term. A similar argument shows that, in general,

$$\frac{\partial}{\partial A_{ij}}(\det \boldsymbol{A}) = \widetilde{A}_{ij}.$$

But from (1.59), $\operatorname{adj} \boldsymbol{A} = (\det \boldsymbol{A})\boldsymbol{A}^{-1}$, so

$$\widetilde{A}_{ij} = (\operatorname{adj} \boldsymbol{A})_{ji} = (\det \boldsymbol{A})(A^{-1})_{ji}.$$

1.37 a. If T is an automorphism then it is surjective. Hence its rank equals $\dim V$.

b. If T is an automorphism then it is invertible. Suppose T^{-1} is represented by the matrix $\boldsymbol{S}$. Then $I = TT^{-1}$ is represented by the matrix $\boldsymbol{T}\boldsymbol{S}$. But any basis, the identity automorphism I is represented by the identity matrix $\boldsymbol{I}$, so $\boldsymbol{T}\boldsymbol{S} = \boldsymbol{I}$, which shows that $\boldsymbol{T}$ is invertible, and hence nonsingular.

1.38 a. Suppose $\{v_i\}$ is an orthonormal basis. Then

$$g(Rv_i, Rv_j) = g(v_i, v_j) = \delta_{ij},$$

whence we see that $\{Rv_i\}$ is again orthonormal. Conversely, if $\{Tv_i\}$ is orthonormal, then

$$g(Tv_i, Tv_j) = \delta_{ij} = g(v_i, v_j).$$

If $v = \sum_i a_i v_i$ and $w = \sum_j b_j v_j$ then

$$g(Tv, Tw) = \sum_{ij} a_i b_j g(Tv_i, Tv_j) = \sum_{ij} a_i b_j g(v_i, v_j) = g(v, w),$$

so T is orthogonal.

b. By orthogonality of R, for any $u, v \in V$,

$$g(v, w) = g(Rv, Rw) = g(R^\dagger Rv, w).$$

It follows that $R^\dagger R = I$, where I is the identity map. (Just let v and w run through all the basis elements.) By the discussion following Exercise 1.26, $R^\dagger$ is represented by $\boldsymbol{R}^T$, so $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}$. As a left inverse must also be a right inverse, $\boldsymbol{R}\boldsymbol{R}^T = \boldsymbol{I}$. Tracing the steps backwards yields the converse.

c. We have $\boldsymbol{I} = \boldsymbol{R}^T\boldsymbol{R}$, so by Exercise 1.31 and Equation (2.54), $1 = \det \boldsymbol{R}^T \det \boldsymbol{R} = (\det \boldsymbol{R})^2$.

d. Let $\boldsymbol{R}$ be orthogonal so that $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}$. In components, $R_{ik}R_{jk} = \delta_{ij}$. *A priori* this looks like n^2 conditions (the number of entries in the identity matrix), but δ_{ij} is symmetric, so the independent conditions arise from those pairs (i, j) for which $i \leq j$. To count these we observe that there are n pairs (i, j) with $i = j$, and $\binom{n}{2} = n(n-1)/2$ pairs with $i < j$. Adding these together gives $n(n+1)/2$ constraints. Therefore the number of independent parameters is $n - n(n+1)/2 = n(n-1)/2$.

1.39 From (2.54) we get

$$1 = \det \boldsymbol{I} = \det \boldsymbol{A}\boldsymbol{A}^{-1} = (\det \boldsymbol{A})(\det \boldsymbol{A}^{-1}),$$

so

$$\det \boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1}.$$

1.40 In our shorthand notation we can write

$$Ae_j = \sum_i e_i A_{ij} \quad \Rightarrow \quad A\boldsymbol{e} = \boldsymbol{e}\boldsymbol{A}, \tag{1}$$

and similarly,

$$Ae'_j = \sum_i e'_i A'_{ij} \quad \Rightarrow \quad A\boldsymbol{e}' = \boldsymbol{e}'\boldsymbol{A}'. \tag{2}$$

Substituting into (2) we get

$$A\boldsymbol{e}\boldsymbol{S} = \boldsymbol{e}\boldsymbol{S}\boldsymbol{A}' \quad \Rightarrow \quad A\boldsymbol{e} = \boldsymbol{e}\boldsymbol{S}\boldsymbol{A}'\boldsymbol{S}^{-1},$$

so comparing with (1) (and using the fact that $\boldsymbol{e}$ is a basis) gives

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{A}'\boldsymbol{S}^{-1} \quad \text{or} \quad \boldsymbol{A}' = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}.$$

1.41 Assume $\boldsymbol{A}$ has n linearly independent eigenvectors $\{v_1, v_2, \dots, v_n\}$ with corresponding eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, and let $\boldsymbol{S}$ be a matrix whose columns are the vectors v_i, $i = 1, \dots, n$. Then $\boldsymbol{S}$ is clearly nonsingular (because its rank is maximal), and multiplication reveals that $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is the diagonal matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ with the eigenvalues of $\boldsymbol{A}$ along the diagonal. It follows that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Sigma}$. Conversely, if there exists a nonsingular matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Sigma}$, then as $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Sigma}$, the columns of $\boldsymbol{S}$ are the eigenvectors of $\boldsymbol{A}$ (which are linearly independent because $\boldsymbol{S}$ is nonsingular), and the diagonal elements of $\boldsymbol{\Sigma}$ are the eigenvalues of $\boldsymbol{A}$.

1.42 The equation $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ holds if and only if $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{v} = 0$, which has a nontrivial solution for $\boldsymbol{v}$ if and only if $\boldsymbol{A} - \lambda\boldsymbol{I}$ is singular, and this holds if and only if $\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$. So the roots of the characteristic polynomial are the eigenvalues of $\boldsymbol{A}$.

1.43 Let $p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I})$ be the characteristic polynomial of $\boldsymbol{A}$. Then

$$\begin{aligned} p_{\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}} &= \det(\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} - \lambda\boldsymbol{I}) \\ &= \det(\boldsymbol{S}^{-1}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{S}) \\ &= (\det\boldsymbol{S})^{-1} p_{\boldsymbol{A}}(\lambda)\det\boldsymbol{S} \\ &= p_{\boldsymbol{A}}. \end{aligned}$$

It follows that the eigenvalues of $\boldsymbol{A}$ are similarity invariants.

1.44 Let $p_{\boldsymbol{A}}(\lambda)$ be the characteristic polynomial of $\boldsymbol{A}$. Then we can write

$$p_{\boldsymbol{A}}(\lambda) = (-1)^n(\lambda - \mu_1)(\lambda - \mu_2)\cdots(\lambda - \mu_n),$$

where the roots (eigenvalues) μ_i are not necessarily distinct. By expanding out the product we see that the constant term in this polynomial is the product of the eigenvalues, but the constant term is also $p_{\boldsymbol{A}}(0) = \det\boldsymbol{A}$.

Again by expanding, we see that the coefficient of the term of order λ^{n-1} is the sum of the eigenvalues times $(-1)^{n-1}$. Now consider $\det(\boldsymbol{A} - \lambda\boldsymbol{I})$. Of all the terms in the Laplace expansion, only one contains $n-1$ powers of λ, namely the product of all the diagonal elements. (In order to contain $n-1$ powers of λ the term must contain at least $n-1$ diagonal elements, which forces it to contain the last diagonal element as well.) But the product of all the diagonal elements is

$$(A_{11} - \lambda)(A_{22} - \lambda)\cdots(A_{nn} - \lambda) = (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}\sum_i A_{ii} + \cdots,$$

where the missing terms are of lower order in λ.

1.45 For any two matrices $\boldsymbol{A}$ and$\boldsymbol{B}$,

$$\operatorname{tr}\boldsymbol{A}\boldsymbol{B} = \sum_{ij} A_{ij}B_{ji} = \sum_{ij} B_{ji}A_{ij} = \operatorname{tr}\boldsymbol{B}\boldsymbol{A}.$$

The general case follows by setting $\boldsymbol{A} = \boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{n-1}$ and $\boldsymbol{B} = \boldsymbol{A}_n$.

1.46 a.

$$\begin{aligned}\prod_{j=1}^{\infty}(1+tx_j) &= (1+tx_1)(1+tx_2)\cdots \\ &= 1+t(x_1+x_2+\cdots)+t^2(x_1x_2+x_1x_3+\cdots+x_2x_3+\cdots) \\ &\qquad +\cdots+t^n(x_1x_2\cdots) \\ &= \sum_{j=1}^{\infty} t^j e_j.\end{aligned}$$

b.

$$\begin{aligned}\sum_j p_j t^{j-1} &= \sum_j (\sum_i x_i^j) t^{j-1} \\ &= \sum_i x_i \sum_j (tx_i)^{j-1} \\ &= \sum_i \frac{x_i}{1-tx_i}.\end{aligned}$$

c. We have

$$\frac{dE}{dt} = \sum_j x_j \prod_{k\neq j}(1+x_k t),$$

so

$$\frac{1}{E}\frac{dE}{dt} = \sum_j \frac{x_j}{1+x_j t}.$$

From $dE/dt = E(t)P(-t)$ we get

$$\begin{aligned}\sum_k k e_k t^{k-1} &= (\sum_i e_i t^i)(\sum_j p_j(-t)^{j-1}) \\ &= \sum_j (-1)^{j-1}\sum_i e_i p_j t^{i+j-1}.\end{aligned}$$

Equating powers of t on both sides gives

$$k e_k = \sum_{j=1}^{k}(-1)^{j-1} e_{k-j} p_j.$$

d. Write down the first n Newton identities in matrix form to get

$$\begin{pmatrix} 1 & 0 & \cdots & & \cdots \\ s_1 & 2 & 0 & \cdots & \cdots \\ s_2 & s_1 & 3 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \vdots \\ s_{n-1} & s_{n-2} & \cdots & s_1 & n \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{pmatrix} = - \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_n \end{pmatrix},$$

where $s_i := (-1)^i p_i$. Then Cramer's rule gives $e_n = \det \boldsymbol{A}/n!$, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & \cdots & & \cdots & -s_1 \\ s_1 & 2 & 0 & \cdots & \cdots & -s_2 \\ s_2 & s_1 & 3 & 0 & \cdots & -s_3 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ s_{n-2} & s_{n-3} & \cdots & s_1 & n-1 & \vdots \\ s_{n-1} & s_{n-2} & \cdots & & s_1 & -s_n \end{pmatrix}$$

Taking the determinant of this, commuting the last column to the front, and pulling out a sign gives

$$\det \boldsymbol{A} = (-1)^n \begin{vmatrix} s_1 & 1 & 0 & \cdots & \cdots \\ s_2 & s_1 & 2 & 0 & \cdots \\ s_3 & s_2 & s_1 & \cdots & \cdots \\ \vdots & \vdots & & \ddots & \vdots \\ s_n & s_{n-1} & s_{n-2} & \cdots & s_1 \end{vmatrix}.$$

Now multiply the odd columns by -1 and the even rows by -1 to get

$$\det \boldsymbol{A} = \begin{vmatrix} p_1 & 1 & 0 & \cdots & \cdots \\ p_2 & p_1 & 2 & 0 & \cdots \\ p_3 & p_2 & p_1 & \cdots & \cdots \\ \vdots & \vdots & & \ddots & \vdots \\ p_n & p_{n-1} & p_{n-2} & \cdots & p_1 \end{vmatrix}.$$

e. If $\boldsymbol{A}v = \lambda v$ then $\boldsymbol{A}^k v = \lambda^k v$. As the trace is the sum of the eigenvalues, $\operatorname{tr} \boldsymbol{A}^k = p_k(\lambda_1, \ldots, \lambda_n)$, where $\lambda_1, \ldots, \lambda_n$ are the n eigenvalues of $\boldsymbol{A}$. The determinant is the product of the eigenvalues, which is $e_n(\lambda_1, \ldots, \lambda_n)$. Thus,

$$4!e_4 = \begin{vmatrix} p_1 & 1 & 0 & 0 \\ p_2 & p_1 & 2 & 0 \\ p_3 & p_2 & p_1 & 3 \\ p_4 & p_3 & p_2 & p_1 \end{vmatrix} = p_1^4 - 6p_1^2 p_2 + 3p_2^2 + 8p_1 p_3 - 6p_4,$$

and so

$$\det \boldsymbol{A} = \frac{1}{4!}\left[(\operatorname{tr}\boldsymbol{A})^4 - 6(\operatorname{tr}\boldsymbol{A})^2(\operatorname{tr}\boldsymbol{A}^2) + 3(\operatorname{tr}\boldsymbol{A}^2)^2 \right.$$
$$\left. + \, 8(\operatorname{tr}\boldsymbol{A})(\operatorname{tr}\boldsymbol{A}^3) - 6\operatorname{tr}\boldsymbol{A}^4\right].$$

1.47 This follows immediately from the results of Exercises 1.40, 1.43 and 1.44.

1.48 First assume $\boldsymbol{A}$ to be diagonalizable with eigenvectors $e_1, e_2, \dots, e_n$ and corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then

$$\begin{aligned} e^{\boldsymbol{A}} e_i &= \left(1 + \boldsymbol{A} + \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{3!}\boldsymbol{A}^3 + \cdots\right) e_i, \\ &= \left(1 + \lambda_i + \frac{1}{2}\lambda_i^2 + \frac{1}{3!}\lambda_i^3 + \cdots\right) e_i, \\ &= e^{\lambda_i} e_i. \end{aligned}$$

It follows that $e^{\boldsymbol{A}}$ is diagonalizable with eigenvectors $e_1, e_2, \dots, e_n$ and corresponding eigenvalues $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}$. [If we had not assumed diagonalizability, we could not say that we had gotten *all* the eigenvectors of $\det e^{\boldsymbol{A}}$ this way.] The result now follows from Exercise 1.44, because

$$e^{\sum_i \lambda_i} = \prod_i e^{\lambda_i}.$$

Next, suppose $\boldsymbol{A}$ is not necessarily diagonal. Because the trace and the determinant are both similarity invariants we may assume, using Schur's theorem, that $\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{N}$, where $\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_n)$. Observe that

$$\boldsymbol{A}^2 = (\boldsymbol{D} + \boldsymbol{N})^2 = \boldsymbol{D}^2 + \boldsymbol{D}\boldsymbol{N} + \boldsymbol{N}\boldsymbol{D} + \boldsymbol{N}^2.$$

But $\boldsymbol{D}^2$ is diagonal and $\boldsymbol{D}\boldsymbol{N}$, $\boldsymbol{N}\boldsymbol{D}$, and $\boldsymbol{N}^2$ are all strictly upper triangular, so we can write

$$\boldsymbol{A}^2 = \boldsymbol{D}^2 + \boldsymbol{N}',$$

for some strictly upper triangular matrix $\boldsymbol{N}'$. By induction, it follows that

$$e^{\boldsymbol{A}} = e^{\boldsymbol{D}} + \boldsymbol{N}'',$$

where $\boldsymbol{N}''$ is some other strictly upper triangular matrix. The matrix on the right is upper triangular, so by the Laplace expansion its determinant is just the product of its diagonal elements. Thus,

$$\det e^{\boldsymbol{A}} = \prod_i e^{d_i}.$$

But $\operatorname{tr}\boldsymbol{A} = \operatorname{tr}\boldsymbol{D}$, so

$$e^{\operatorname{tr}\boldsymbol{A}} = e^{\operatorname{tr}\boldsymbol{D}} = e^{\sum_i d_i},$$

whereupon the claim follows.

1.49 By positive definiteness,

$$0 \le g(u+\alpha v, u+\alpha v) = g(u,u) + 2\alpha g(u,v) + \alpha^2 g(v,v).$$

Minimizing the right side with respect to α gives

$$\alpha = -\frac{g(u,v)}{g(v,v)},$$

and plugging back in gives

$$0 \le g(u,u) - 2\frac{g(u,v)^2}{g(v,v)} + \frac{g(u,v)^2}{g(v,v)}$$

or

$$0 \le g(u,u)g(v,v) - g(u,v)^2.$$

Moreover, equality holds if and only if $u + \alpha v = 0$, or $u = -\alpha v$.

1.50 Symmetry is obvious, and bilinearity follows from the linearity of integration. For example,

$$\begin{aligned}(f, ag+bh) &= \int_{-\infty}^{\infty} f(x)(ag(x)+bh(x))\,dx \\ &= \int_{-\infty}^{\infty} (af(x)g(x)+bf(x)h(x))\,dx \\ &= a(f,g)+b(f,h).\end{aligned}$$

Finally,

$$(f,f) = \int_{-\infty}^{\infty} f^2(x)\,dx \ge 0$$

because $f^2 \ge 0$. The integral vanishes if and only if $f = 0$, so the map $(\cdot,\cdot)$ is positive definite.

1.51 Suppose $v_1, \dots, v_n$ are linearly dependent. Then there exist constants c_j, not all zero, such that

$$c_1 v_1 + c_2 v_2 + \dots c_n v_n = 0. \tag{1}$$

Take the inner product of (1) with each of the vectors v_i to get

$$\begin{array}{ccccccc} g(v_1,v_1)c_1 & + & g(v_1,v_2)c_2 & \cdots & + & g(v_1,v_n)c_n & = 0, \\ g(v_2,v_1)c_1 & + & g(v_2,v_2)c_2 & \cdots & + & g(v_2,v_n)c_n & = 0, \\ \vdots & & \vdots & & & \vdots & \\ g(v_n,v_1)c_1 & + & g(v_n,v_2)c_2 & \cdots & + & g(v_n,v_n)c_n & = 0. \end{array} \tag{2}$$

Regarding this as a set of linear equations for the constants c_j, we see that the Grammian must vanish.

Conversely, suppose the Grammian of $\{v_1, \dots, v_n\}$ is zero. Then the system (2) has a nonzero solution. Multiplying the equations in (2) by each c_j in succession and then adding them all together gives

$$\|c_1 v_1 + c_2 v_2 + \cdots + c_m v_m\| = 0$$

where $\|v\|^2 := g(v, v)$. Equation (1) now follows by virtue of the nondegeneracy of the inner product, so the vectors are linearly dependent.

1.52 Define $v_i = x^i$ for $i = 0, 1, 2, 3$. Then

$$g(v_0, v_0) = \int_{-1}^{1} dx = 2,$$

so

$$e_0 := \frac{1}{\sqrt{2}}.$$

Next,

$$e_1' = x - g\left(x, \frac{1}{\sqrt{2}}\right) \cdot \frac{1}{\sqrt{2}} = x - \frac{1}{2}\int_{-1}^{1} x\,dx = x - \left.\frac{x^2}{4}\right|_{-1}^{1} = x.$$

Thus,

$$g(e_1', e_1') = g(x, x) = \int_{-1}^{1} x^2\,dx = \left.\frac{x^3}{3}\right|_{-1}^{1} = \frac{2}{3}.$$

Hence

$$e_1 = \frac{e_1'}{g(e_1', e_1')^{1/2}} = \sqrt{\frac{3}{2}}x.$$

Next we have

$$e_2' = x^2 - e_0 g(e_0, x^2) - e_1 g(e_1, x^2).$$

The last inner product vanishes by a simple parity argument, so we only need to compute the second term, which is

$$e_0 g(e_0, x^2) = \frac{1}{2}\int_{-1}^{1} x^2\,dx = \frac{1}{3}.$$

Thus,

$$e_2' = x^2 - \frac{1}{3}.$$

Now we normalize.

$$g(e_2', e_2') = g(x^2-1/3, x^2-1/3) = g(x^2, x^2)-(2/3)g(x^2, 1)+(1/9)g(1, 1).$$

The only inner product we haven't done yet is the first, which is

$$g(x^2, x^2) = \int_{-1}^{1} x^4\, dx = \frac{2}{5}.$$

Hence

$$g(e_2', e_2') = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45},$$

whereupon we obtain

$$e_2 = \frac{e_2'}{g(e_2', e_2')^{1/2}} = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right).$$

Lastly, we have

$$e_3' = x^3 - e_0 g(e_0, x^3) - e_1 g(e_1, x^3) - e_2 g(e_2, x^3).$$

Again by parity we need only compute the third term on the right, which is

$$e_1 g(e_1, x^3) = \frac{3}{2}x \int_{-1}^{1} x^4\, dx = \frac{3}{5}x.$$

Thus,

$$e_3' = x^3 - \frac{3}{5}x.$$

The next step is to normalize. We have

$$g(e_3', e_3') = g(x^3-(3/5)x, x^3-(3/5)x) = g(x^3, x^3)-\frac{6}{5}g(x^3, x)+\frac{9}{25}g(x, x).$$

Having done this many times by now, we can pretty much read off the answer:

$$g(e_3', e_3') = \frac{2}{7} - \frac{12}{25} + \frac{6}{25} = \frac{8}{175}.$$

Hence

$$e_3 = \sqrt{\frac{175}{8}}\left(x^3 - \frac{3}{5}x\right).$$

1.53 For the first, we have, by the definition above,

$$f \in \ker T^* \Leftrightarrow T^* f = 0 \Leftrightarrow fT = 0 \Leftrightarrow f \in \operatorname{Ann} \operatorname{im} T.$$

For the second, if $f \in \operatorname{im} T^*$ then $f = T^* g$ for some $g \in W^*$. So, if $v \in \ker T$,

$$f(v) = T^* g(v) = gT(v) = g(0) = 0,$$

so $f \in \operatorname{Ann} \ker T$. Conversely, let $f \in \operatorname{Ann} \ker T$. By Theorem 1.3 we can write

$$V = \ker T \oplus S(\operatorname{im} T),$$

where S is a section of T. Define an element $g \in W^*$ by

$$g(w) := \begin{cases} f(S(w)), & \text{if } w \in \operatorname{im} T, \\ 0, & \text{otherwise.} \end{cases}$$

For any $v \in V$ we have

$$g(Tv) = f(v), \tag{1}$$

because if $v \in \ker T$ then both sides of (1) are zero, and if $v \notin \ker T$ then both sides of (1) are equal by virtue of the definition of g (and the fact that $T \circ S = 1$). We conclude that $f = T^* g \in \operatorname{im} T^*$.

1.54 Suppose that $v = \sum_k v_k e_k$. Then, as the basis is orthonormal,

$$g(v, e_i) = \sum_k v_k g(e_k, e_i) = g(e_i, e_i) v_i.$$

Hence,

$$v_{f_v} = \sum_i g(e_i, e_i) f_{v,i} e_i = \sum_i g(e_i, e_i) g(v, e_i) e_i = \sum_i g(e_i, e_i)^2 v_i e_i = v.$$

Also,

$$f_{v_f}(e_j) = g(v_f, e_j) = \sum_i g(e_i, e_i) f_i g(e_i, e_j) = g(e_j, e_j)^2 f_j = f_j,$$

so $f_{v_f} = f$ (two linear maps that agree on a basis agree everywhere). Therefore, the two maps $v \to f_v$ and $f \to v_f$ are indeed inverses of one another.

2

Multilinear algebra

2.1 Let $\boldsymbol{e}$ and $\boldsymbol{e}'$ be two bases related by the change of basis matrix $\boldsymbol{A}$, so that

$$e_{i'} = \sum_i e_i A^i{}_{i'}.$$

Then the components of T in the two bases are related by

$$T_{i'j'} = \sum_{ij} A^i{}_{i'} A^j{}_{j'} T_{ij}.$$

Suppose $T_{ij} = T_{ji}$. Then

$$T_{j'i'} = \sum_{ij} A^i{}_{j'} A^j{}_{i'} T_{ij} = \sum_{ij} A^i{}_{j'} A^j{}_{i'} T_{ji} = \sum_{ij} A^j{}_{j'} A^i{}_{i'} T_{ij} = T_{i'j'},$$

where in the penultimate step we changed the names of the dummy indices from i and j to j and i, respectively. The antisymmetric case is similar and is left to the reader.

2.2 We have

$$\sum_{ij} A_{ij} B^{ij} = \sum_{ij} A_{ji} B^{ij} = -\sum_{ij} A_{ji} B^{ji} = -\sum_{ij} A_{ij} B^{ij} = 0.$$

In the first equality we are allowed to switch i and j in A_{ij} because A is symmetric. In the second equality we swap i and j in B^{ij} at the cost of a minus sign, because B is antisymmetric. In the third equality we change the names of the dummy indices from i and j to j and i, respectively. The last equality follows because the only number equal to its negative is zero.

2.3 Just repeat what was done in the text, without the signs.

2.4 We have

$$\begin{aligned}
&\sum_{i_1,\dots,i_p} a^{[i_1\dots i_p]} e_{i_1}\wedge\cdots\wedge e_{i_p}\\
&\quad= \frac{1}{p!}\sum_{i_1,\dots,i_p}\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma a^{i_{\sigma(1)}\dots i_{\sigma(p)}} e_{i_1}\wedge\cdots\wedge e_{i_p} &(1)\\
&\quad= \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}\sum_{i_1,\dots,i_p} a^{i_{\sigma(1)}\dots i_{\sigma(p)}} e_{i_{\sigma(1)}}\wedge\cdots\wedge e_{i_{\sigma(p)}} &(2)\\
&\quad= \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}\sum_{i_{\sigma(1)},\dots,i_{\sigma(p)}} a^{i_{\sigma(1)}\dots i_{\sigma(p)}} e_{i_{\sigma(1)}}\wedge\cdots\wedge e_{i_{\sigma(p)}} &(3)\\
&\quad= \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}\sum_{i_1,\dots,i_p} a^{i_1\dots i_p} e_{i_1}\wedge\cdots\wedge e_{i_p} &(4)\\
&\quad= \sum_{i_1,\dots,i_p} a^{i_1\dots i_p} e_{i_1}\wedge\cdots\wedge e_{i_p}. &(5)
\end{aligned}$$

Equality (1) is just the definition (2.21), while in (2) we have used (2.39) and flipped the order of summation. In (3) we have changed the name of the dummy indices from $i_1,\dots,i_p$ to $i_{\sigma(1)},\dots,i_{\sigma(p)}$, and then in (4) we have changed the dummy indices back to $i_1,\dots,i_p$. Finally, (5) holds because $\mathfrak{S}_p$ has $p!$ elements.

2.5 By linearity is suffices to prove Property (3) for two monomials. So let $\lambda = v_1\wedge\cdots\wedge v_p$ and $\mu = w_1\wedge\cdots\wedge w_q$. Then by moving the vectors v_i successively through the w_i's (of which there are q), we get

$$\begin{aligned}
\mu\wedge\lambda &= w_1\wedge\cdots\wedge w_q\wedge v_1\wedge\cdots\wedge v_p\\
&= (-1)^q v_1\wedge w_1\wedge\cdots\wedge w_q\wedge v_2\wedge\cdots\wedge v_p\\
&= (-1)^{2q} v_1\wedge v_2\wedge w_1\wedge\cdots\wedge w_q\wedge v_3\wedge\cdots\wedge v_p\\
&= \qquad\qquad \vdots\\
&= (-1)^{pq} v_1\wedge\cdots\wedge v_p\wedge w_1\wedge\cdots\wedge w_q\\
&= (-1)^{pq}\lambda\wedge\mu.
\end{aligned}$$

2.6 By multilinearity is suffices to prove everything for monomials. So, let $T = e_1\otimes\cdots\otimes e_p$, say, and define an action of an element $\sigma\in\mathfrak{S}_p$ on tensors by $T^\sigma := e_{\sigma(1)}\otimes\cdots\otimes e_{\sigma(p)}$, extended by linearity. Note that, by definition, $(T^\sigma)^\tau = T^{\tau\sigma}$. With this notation we have

$$\operatorname{alt}(T) = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma T^\sigma.$$

Hence (*cf.,* Exercise 2.3),

$$\begin{aligned}\operatorname{alt}(T^\tau) &= \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma (T^\tau)^\sigma = \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma T^{\sigma\tau} = \frac{1}{p!}\sum_{\pi\tau^{-1}\in\mathfrak{S}_p}(-1)^{\pi\tau^{-1}}T^\pi \\ &= (-1)^\tau\frac{1}{p!}\sum_{\pi\tau^{-1}\in\mathfrak{S}_p}(-1)^\pi T^\pi = (-1)^\tau\frac{1}{p!}\sum_{\pi\in\mathfrak{S}_p}(-1)^\pi T^\pi = (-1)^\tau \operatorname{alt}(T).\end{aligned}$$

Therefore

$$\begin{aligned}\operatorname{alt}(\operatorname{alt}(T)) &= \frac{1}{p!}\sum_{\sigma\in\mathfrak{S}_p}(-1)^\sigma(\operatorname{alt}(T))^\sigma = \left(\frac{1}{p!}\right)^2\sum_{\sigma\in\mathfrak{S}_p}\sum_{\tau\in\mathfrak{S}_p}(-1)^\sigma(-1)^\tau(T^\tau)^\sigma \\ &= \left(\frac{1}{p!}\right)^2\sum_{\sigma\in\mathfrak{S}_p}\sum_{\tau\in\mathfrak{S}_p}(-1)^{\sigma\tau}T^{\sigma\tau} = \left(\frac{1}{p!}\right)^2\sum_{\sigma\in\mathfrak{S}_p}\sum_{\sigma^{-1}\pi\in\mathfrak{S}_p}(-1)^\pi T^\pi \\ &= \left(\frac{1}{p!}\right)^2\sum_{\sigma\in\mathfrak{S}_p}\sum_{\pi\in\mathfrak{S}_p}(-1)^\pi T^\pi = \left(\frac{1}{p!}\right)\sum_{\pi\in\mathfrak{S}_p}(-1)^\pi T^\pi = \operatorname{alt}(T).\end{aligned}$$

This proves (i).

For simplicity, we prove (ii) for two rank one tensors, as the general case follows by similar considerations. Suppose $S = v_1\otimes\cdots\otimes v_p$ and $T = v_{p+1}\otimes\cdots\otimes v_{p+q}$. Then

$$\operatorname{alt}(S\otimes T) = \frac{1}{(p+q)!}\sum_{\sigma\in\mathfrak{S}_{p+q}}(-1)^\sigma v_{\sigma(1)}\otimes v_{\sigma(2)}\otimes\cdots\otimes v_{\sigma(p+q)}. \tag{1}$$

The key point is that $\mathfrak{S}_{p+q}$ naturally decomposes into pieces, and each term in the sum vanishes on each of the pieces. Specifically, contained in $\mathfrak{S}_{p+q}$ is a subgroup isomorphic to $\mathfrak{S}_p$ that permutes the first p numbers and leaves the remaining q numbers fixed. For all such permutations, the right side of (1) vanishes by the hypothesis $\operatorname{alt}(S) = 0$ (because the sign of such a permutation just equals the sign of the permutation of the first p numbers, and the remaining q vectors pull out of the sum).

To show the rest of the terms vanish, we need a little bit of group theory. If G is a group and H a subgroup, a right coset of H in G is any subset of the form Hg for $g\in G$. The right cosets partition G. (Proof. If $x\in Hg_1\cap Hg_2$ then $x = h_1g_1 = h_2g_2$ for some $h_1, h_2\in H$. So $g_1 = hg_2$ for $h = h_1^{-1}h_2\in H$, which shows that $Hg_1 = Hg_2$. In other words, if two cosets are not disjoint they coincide.)

Returning to our problem, let $\{H\tau_1, H\tau_2, \dots, H\tau_k\}$ be a partition of $\mathfrak{S}_{p+q}$ into right cosets of $H = \mathfrak{S}_p$, the subgroup permuting the first p numbers.

Then

$$\begin{aligned}\operatorname{alt}(S \otimes T) &= \frac{1}{(p+q)!}\sum_{i=1}^{k}\sum_{\sigma \in H\tau_i}(-1)^{\sigma} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(p+q)} \\ &= \frac{1}{(p+q)!}\sum_{i=1}^{k}(-1)^{\tau_i}\sum_{\pi \in H}(-1)^{\pi} v_{\pi\tau_i(1)} \otimes v_{\pi\tau_i(2)} \otimes \cdots \otimes v_{\pi\tau_i(p+q)}.\end{aligned}$$

A moment's thought shows that, for each i, the inner sum vanishes for precisely the same reason as before, because the only effect of τ_i is to renumber the indices. A similar argument shows that $T \wedge S = 0$.

By multilinearity,

$$\begin{aligned}(R \wedge S) \wedge T - \operatorname{alt}(R \otimes S \otimes T) &= \operatorname{alt}((R \wedge S) \otimes T) - \operatorname{alt}(R \otimes S \otimes T) \\ &= \operatorname{alt}((R \wedge S - R \otimes S) \otimes T) \\ &= (R \wedge S - R \otimes S) \wedge T.\end{aligned}$$

But this vanishes by (ii), because by (i), $\operatorname{alt}(R \wedge S) = \operatorname{alt}(\operatorname{alt}(R \otimes S)) = \operatorname{alt}(R \otimes S)$, so

$$\operatorname{alt}(R \wedge S - R \otimes S) = 0.$$

Similar reasoning shows that

$$R \wedge (S \wedge T) = \operatorname{alt}(R \otimes S \otimes T),$$

whereupon we conclude that the wedge product defined by alt is indeed associative. Wow. All that just to prove a fact that is obvious when viewed from the axiomatic perspective. Well, *chacun à son gout*.

2.7 We have

$$\begin{aligned}\textstyle\bigwedge^3 T(e_1 \wedge e_2 \wedge e_3) &= Te_1 \wedge Te_2 \wedge Te_3 \\ &= (e_1 + 2e_2) \wedge (3e_2 + 2e_3) \wedge (e_1 + e_3) \\ &= (e_1 + 2e_2) \wedge (3e_2 \wedge e_1 + 3e_2 \wedge e_3 + 2e_3 \wedge e_1) \\ &= 7e_1 \wedge e_2 \wedge e_3.\end{aligned}$$

On the other hand, the matrix representing T is

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

and its determinant is—you guessed it—7.

2.8 Pick a basis $\{e_1, \ldots, e_n\}$, and suppose that T is represented by $\boldsymbol{T}$ in that basis. Then (ignoring index placement, as it is irrelevant here),

$$\begin{aligned}(\textstyle\bigwedge^n T)e_1 \wedge \cdots \wedge e_n &= Te_1 \wedge \cdots \wedge Te_n \\ &= \sum_{i_1,\ldots,i_n} T_{i_1 1} \cdots T_{i_n n}\, e_{i_1} \wedge \cdots e_{i_n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma T_{\sigma(1)1} \cdots T_{\sigma(n)n}\, e_1 \wedge \cdots \wedge e_n,\end{aligned}$$

because the only terms contributing to the sum are permutations of $\{1, \ldots, n\}$, and by definition of the sign,

$$e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = (-1)^\sigma e_1 \wedge \cdots \wedge e_n.$$

2.9 We have

$$\begin{aligned}(\textstyle\bigwedge^2 T)(e_i \wedge e_j) &= Te_i \wedge Te_j \\ &= \sum_k T_{ki} e_k \wedge \sum_\ell T_{\ell j} e_\ell \\ &= \sum_{k\ell} T_{ki} T_{\ell j} (e_k \wedge e_\ell) \\ &= \sum_{k<\ell} \big(T_{ki} T_{\ell j} - T_{\ell i} T_{kj}\big)\, (e_k \wedge e_\ell).\end{aligned}$$

For example,

$$\begin{aligned}(\textstyle\bigwedge^2 T)(e_1 \wedge e_2) = (T_{11}T_{22} - T_{21}T_{12})(e_1 \wedge e_2) \\ + (T_{11}T_{32} - T_{31}T_{12})(e_1 \wedge e_3) \\ + (T_{21}T_{32} - T_{31}T_{22})(e_2 \wedge e_3).\end{aligned}$$

Let $T^{(2)}$ denote the matrix representation of the operator $\bigwedge^2 T$, and arrange the basis elements of $\bigwedge^2 V$ in lexicographic order: $e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3$. Then similar calculations reveal that

$$\begin{pmatrix} T_{11}T_{22} - T_{21}T_{12} & T_{11}T_{23} - T_{21}T_{13} & T_{12}T_{23} - T_{22}T_{13} \\ T_{11}T_{32} - T_{31}T_{12} & T_{11}T_{33} - T_{31}T_{13} & T_{12}T_{33} - T_{32}T_{13} \\ T_{21}T_{32} - T_{31}T_{22} & T_{21}T_{33} - T_{31}T_{23} & T_{22}T_{33} - T_{32}T_{23} \end{pmatrix}.$$

2.10 With the setup of the hint, we have

$$\det \boldsymbol{T} = \det T = (\textstyle\bigwedge^n T)(e_1 \wedge \cdots \wedge e_n) = Te_1 \wedge \cdots \wedge Te_n = v_1 \wedge \cdots \wedge v_n.$$

It is now painfully obvious from the properties of wedge products that, (1) swapping two columns of $\boldsymbol{T}$ flips the sign of the determinant, (2) setting two columns equal kills the determinant, (3) adding a multiple of a column to another column leaves the determinant unchanged because, by multilinearity,

$$\begin{aligned} &v_1 \wedge \cdots \wedge (v_i + \lambda v_j) \wedge \cdots \wedge v_j \wedge \cdots \wedge v_n \\ &\quad = v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_j \wedge \cdots \wedge v_n + \lambda v_1 \wedge \cdots \wedge v_j \wedge \cdots \wedge v_j \wedge \cdots \wedge v_n \\ &\quad = v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_j \wedge \cdots \wedge v_n, \end{aligned}$$

and (4) multiplying a column vector by a scalar multiplies the entire determinant by that scalar. The corresponding statements with the word 'column' replaced by the word 'row' follow by appealing to (1.56).

2.11 Assume the same setup as in the proof of (2.60). By linearity we may assume $\eta = e_I$. Now $\lambda = \sum_K a^K e_K$, but $g(e_I, e_K) = 0$ unless $K = I$, so we may as well assume $\lambda = e_I$. Then using (2.63) and (2.65) we have

$$\begin{aligned} \eta \wedge \star\lambda &= e_I \wedge g(e_J, e_J) e_J = g(e_J, e_J)\sigma = (-1)^d g(e_I, e_I)\sigma \\ &= (-1)^d g(\eta, \lambda)\sigma. \end{aligned}$$

The other equality follows from the symmetry of the inner product.

2.12 Complete $\{v_1, \dots, v_k\}$ to a basis $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$. Then

$$w_i = \sum_{j=1}^n a_{ij} v_j,$$

for some coefficients a_{ij}. Hence

$$0 = \sum_{i=1}^k v_i \otimes \sum_{j=1}^n a_{ij} v_j = \sum_{i=1}^k \sum_{j=1}^n a_{ij} v_i \otimes v_j.$$

But $v_i \otimes v_j$, $1 \le i, j, \le n$ is a basis for the tensor product space $V \otimes V$. Therefore all the a_{ij} must vanish.

2.13 By definition,

$$(A \otimes B)(e_i \otimes f_j) = \sum_{k\ell} (A \otimes B)_{k\ell,ij} (e_k \otimes f_\ell),$$

where $(A \otimes B)_{k\ell,ij}$ is the $(k\ell, ij)^{\text{th}}$ component of the matrix representing $A \otimes B$ in the basis $\{e_i \otimes f_j\}$. But also,

$$(A \otimes B)(e_i \otimes f_j) = Ae_i \otimes Bf_j = \sum_{k\ell} A_{ki} B_{\ell j} (e_k \otimes f_\ell).$$

It follows that

$$(A \otimes B)_{k\ell,ij} = A_{ki} B_{\ell j}.$$

But this is just $(\boldsymbol{A} \otimes \boldsymbol{B})_{k\ell,ij}$, as one can see by unpacking the definition of the Kronecker product.

2.14 Assume $v_1, v_2, \dots, v_p$ are linearly dependent. Then there exist constants, not all zero, such that

$$c_1 v_1 + \cdots c_p v_p = 0.$$

By renumbering the vectors if necessary, we may take $c_p \neq 0$. Then

$$v_p = -\frac{1}{c_p}(c_1 v_1 + \cdots + c_{p-1} v_{p-1}).$$

By the multilinearity and antisymmetry properties of the wedge product, the expression $v_1 \wedge \cdots \wedge v_p$ is a sum of terms, each of which involves the wedge product of two copies of the same vector, so it must vanish.

Conversely, suppose $v_1, v_2, \dots, v_p$ are linearly independent. Then they form a basis for the p dimensional subspace $W \subseteq V$ that they span. The p-vector $v_1 \wedge \cdots \wedge v_p$ is a basis for the one dimensional space $\bigwedge^p W$, and therefore cannot vanish.

2.15 Following the hint, let $\{v_1, v_2, \dots, v_p, v_{p+1}, \dots, v_n\}$ be a basis of V. Since any vector can be expanded in terms of the basis, we can write

$$w_i = \sum_{j=1}^{p} A_{ij} v_j + \sum_{j=p+1}^{n} B_{ij} v_j$$

for some matrices A and B. Thus,

$$\begin{aligned} 0 = \sum_{i=1}^{p} v_i \wedge w_i &= \sum_{i=1}^{p} v_i \wedge \left(\sum_{j=1}^{p} A_{ij} v_j + \sum_{j=p+1}^{n} B_{ij} v_j \right) \\ &= \sum_{i,j=1}^{p} A_{ij} (v_i \wedge v_j) + \sum_{i=1}^{p} \sum_{j=p+1}^{n} B_{ij} (v_i \wedge v_j). \end{aligned}$$

Each term on the right side must vanish separately, because they involve linearly independent bivectors. The first term can be written

$$\sum_{1<i<j<p} (A_{ij} - A_{ji})(v_i \wedge v_j),$$

from which it follows that $A_{ij} = A_{ji}$. The only way the second term can vanish is if $B_{ij} = 0$ for all i and j for which it is defined.

2.16 To compute the determinant of $\bigwedge^p A$, we compute its eigenvalues and take their product. The map $\bigwedge^p A$ acts on the vector space $\bigwedge^p V$ consisting of all p forms of V. As A is assumed diagonal, we may choose a basis $\{e_1, \dots, e_n\}$ for V consisting of eigenvectors of A. Set $A e_i = \lambda_i e_i$. Then

$$\begin{aligned}(\textstyle\bigwedge^p A)(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}) &= Ae_{i_1} \wedge Ae_{i_2} \wedge \cdots \wedge Ae_{i_p} \\ &= \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_p}(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}). \qquad (1)\end{aligned}$$

In other words, the eigenvectors of $\bigwedge^p A$ acting on $\bigwedge^p V$ are products of eigenvalues of A whose indices range over all the p subsets of $\{1, 2, \ldots, n\}$.

Multiplying all these eigenvalues of $\bigwedge^p A$ together, the only question is how often each individual λ appears. Let's do a simple example. Suppose $p = 3$ and $n = 5$. Then the following subsets appear:

$$123, 124, 125, 134, 135, 145, 234, 235, 245, 345.$$

How many times does the '1' appear? Well, to construct the subsets containing 1, we first chose 1, and then the rest of the numbers in $\binom{5-1}{3-1} = \binom{4}{2} = 6$ ways. This holds in general, so each index appears in the product $\binom{n-1}{p-1}$ times, which means that

$$\det(\textstyle\bigwedge^p A) = (\det A)^{\binom{n-1}{p-1}}.$$

2.17 As in Example 2.4 we get (using the fact that $d = 1$ for a Euclidean signature)

$$\star e_i = (-1)^{i-1} e_1 \wedge \cdots \wedge \widehat{e_i} \wedge \cdots \wedge e_n,$$

where the hat means that e_i is omitted. Thus

$$\star\mu = \sum_i (-1)^{i-1} a_i\, e_1 \wedge \cdots \wedge \widehat{e_i} \wedge \cdots \wedge e_n.$$

2.18 From (2.59) and (2.61) we get

$$(-1)^d g(\lambda, \mu)\sigma = \lambda \wedge \star\mu = g(\star\lambda, \star\mu)\sigma,$$

from which the result follows.

2.19 Let $\{e_i\}$, $1 \le i \le m$ be a basis for V and $\{f_j\}$, $1 \le j \le n$ a basis for W. Then $\{(e_i, 0), (0, f_j)\}$, $1 \le i \le m$, $1 \le j \le n$ is a basis for $V \oplus W$. The vector space $\bigwedge^n(V \oplus W)$ is thus generated (as a direct sum) by all basis elements of the form

$$(e_{i_1}, 0) \wedge \cdots \wedge (e_{i_k}, 0) \wedge (0, f_{j_1}) \wedge \cdots \wedge (0, f_{j_{n-k}}),$$

where k runs from 0 to n (and all the indices are in increasing—but not necessarily sequential—order). Under the linear map defined in the problem,

$$(e_{i_1} \wedge \cdots e_{i_k}) \otimes (f_{j_1} \wedge \cdots \wedge f_{j_{n-k}}) \mapsto (e_{i_1}, 0) \wedge \cdots \wedge (e_{i_k}, 0) \wedge (0, f_{j_1}) \wedge \cdots \wedge (0, f_{j_{n-k}}),$$

so

$$\bigoplus_{k=0}^{n} \left(\textstyle\bigwedge^k V \otimes \bigwedge^{n-k} W\right) \cong \textstyle\bigwedge^n (V \oplus W)$$

(because the induced map on the direct sum is clearly bijective).

2.20 a. If A is diagonalizable with eigenvalues $\{\lambda_i\}$ then $I + zA$ is diagonalizable with eigenvalues $\{1 + \lambda_i z\}$. So (denoting the kth elementary symmetric function by E_k)

$$\det(I + zA) = \prod_k (1 + \lambda_k z) = \sum_k E_k(\lambda_1, \dots, \lambda_n) z^k.$$

On the other hand, if $\{e_i\}$ is a basis for V,

$$\begin{aligned}(\textstyle\bigwedge^k A)(e_{i_1} \wedge \cdots \wedge e_{i_k}) &= Ae_{i_1} \wedge \cdots \wedge Ae_{i_k} \\ &= (\lambda_{i_1} \cdots \lambda_{i_k})(e_{i_1} \wedge \cdots \wedge e_{i_k}),\end{aligned}$$

so

$$\operatorname{tr} \textstyle\bigwedge^k A = \displaystyle\sum_{i_1 < i_2 < \cdots < i_k} \lambda_{i_1} \cdots \lambda_{i_k} = E_k(\lambda_1, \dots, \lambda_n).$$

b. Define

$$B := I + zA.$$

From the definition of the determinant,

$$\textstyle\bigwedge^n B(e_1 \wedge e_2 \wedge \cdots \wedge e_n) = (\det B) e_1 \wedge e_2 \wedge \cdots \wedge e_n.$$

Thus,

$$\operatorname{tr} \textstyle\bigwedge^n B = \det B,$$

because the trace just sums all the components of the linear operator that map the basis elements to themselves. (Basically, the trace counts the 'multiplicity' of the fixed points of the action of the linear operator on all the basis elements.) But, from the definition,

$$\begin{aligned}\textstyle\bigwedge^n B(e_1 \wedge e_2 \wedge \cdots \wedge e_n) &= Be_1 \wedge Be_2 \wedge \cdots \wedge Be_n \\ &= (I + zA)e_1 \wedge (I + zA)e_2 \wedge \cdots \wedge (I + zA)e_n.\end{aligned}$$

Now we are finished, because the coefficient of z^r consists of a sum of all possible (ordered) wedge products of r Ae_i's and $n - r$ e_i's. More precisely, each term in the sum involves the action of $\bigwedge^r A$ on all possible basis elements of $\bigwedge^r V$. Adding up all these contributions yields exactly $\operatorname{tr} \bigwedge^r A$.

We can see this explicitly when $n = 3$:

$$\begin{aligned}&\textstyle\bigwedge^n B(e_1 \wedge e_2 \wedge e_3) \\ &\quad = (I + zA)e_1 \wedge (I + zA)e_2 \wedge (I + zA)e_3 \\ &\quad = e_1 \wedge e_2 \wedge e_3 \\ &\qquad + z[Ae_1 \wedge e_2 \wedge e_3 + e_1 \wedge Ae_2 \wedge e_3 + e_1 \wedge e_2 \wedge Ae_3] \\ &\qquad + z^2[Ae_1 \wedge Ae_2 \wedge e_3 + Ae_1 \wedge e_2 \wedge Ae_3 + e_1 \wedge Ae_2 \wedge Ae_3]\end{aligned}$$

$$\begin{aligned}
&\quad + z^3[Ae_1 \wedge Ae_2 \wedge Ae_3] \\
&= e_1 \wedge e_2 \wedge e_3 \\
&\quad + z[\textstyle\bigwedge^1 Ae_1 \wedge e_2 \wedge e_3 + e_1 \wedge \bigwedge^1 Ae_2 \wedge e_3 + e_1 \wedge e_2 \wedge \bigwedge^1 Ae_3] \\
&\quad + z^2[\textstyle\bigwedge^2 A(e_1 \wedge e_2) \wedge e_3 + \bigwedge^2 A(e_1 \wedge |e_2| \wedge e_3) \\
&\quad + e_1 \wedge \textstyle\bigwedge^2 A(e_2 \wedge e_3)] \\
&\quad + z^3[\textstyle\bigwedge^3 A(e_1 \wedge e_2 \wedge e_3)] \\
&= (1 + z \operatorname{tr} \textstyle\bigwedge^1 A + z^2 \operatorname{tr} \bigwedge^2 A + z^3 \operatorname{tr} \bigwedge^3 A)(e_1 \wedge e_2 \wedge e_3),
\end{aligned}$$

where the straight bracket symbol means that the term is omitted from the action of the exterior algebra operator. Here we are using the fact that, for example,

$$\begin{aligned}
\textstyle\bigwedge^2 A(e_1 \wedge e_2) &= \alpha_{12}(e_1 \wedge e_2) + \dots \\
\textstyle\bigwedge^2 A(e_1 \wedge e_3) &= \alpha_{13}(e_1 \wedge e_3) + \dots \\
\textstyle\bigwedge^2 A(e_2 \wedge e_3) &= \alpha_{23}(e_2 \wedge e_3) + \dots,
\end{aligned}$$

and so

$$\operatorname{tr} \textstyle\bigwedge^2 A = \alpha_{12} + \alpha_{13} + \alpha_{23}.$$

2.21 a. The question asks for the number of multisets of size p chosen from a set of size n, where a multiset is just a set where repetitions are allowed. For example, $\{1, 1, 1, 2, 2, 4\}$ is a multiset of size 6 chosen from a set of 4 objects, corresponding to the basis element $e_1 \odot e_1 \odot e_1 \odot e_2 \odot e_2 \odot e_4$. The slickest way to count these objects is to use a "stars and bars" argument. We observe that we can represent the multiset above by the following picture:

$$* * *| * *||*,$$

where the number of stars in each "compartment" determined by the bars equals the number of times a particular element appears in the multiset. For a multiset of size p we must have p stars, while n is the number of compartments, so there must be $n-1$ bars. Altogether there are $p+n-1$ symbols, and a multiset is a choice of which $n-1$ of those symbols will be bars. By definition, the number of such choices is $\binom{n+p-1}{n-1}$.

b. The natural map $\psi : \operatorname{Sym} V \to \mathbb{F}[e_1, e_2, \dots, e_n]$ given by $e_1 \odot \dots \odot e_p \mapsto e_1 e_2 \cdots e_p$ and extended by linearity, provides the natural isomorphism we seek. The tedious but entirely straightforward details are left to the reader.

c. Just follow the proof of Exercise 2.19.

d. If A is diagonalizable with eigenvalues $\{\lambda_i\}$ then $I - zA$ is diagonalizable with eigenvalues $\{1 - \lambda_i z\}$. So

$$\frac{1}{\det(I - zA)} = \prod_k (1 - \lambda_k z)^{-1} = \sum_k H_k(\lambda_1, \ldots, \lambda_n) z^k,$$

where H_k is the k^{th} homogeneous symmetric function. On the other hand, if $\{e_i\}$ is a basis for V,

$$\begin{aligned}(\mathrm{Sym}^k A)(e_{i_1} \odot \cdots \odot e_{i_k}) &= Ae_{i_1} \odot \cdots \odot Ae_{i_k} \\ &= (\lambda_{i_1} \cdots \lambda_{i_k})(e_{i_1} \odot \cdots \odot e_{i_k}),\end{aligned}$$

so

$$\operatorname{tr} \mathrm{Sym}^k A = \sum_{i_1 \le i_2 < \cdots \le i_k} \lambda_{i_1} \cdots \lambda_{i_k} = H_k(\lambda_1, \ldots, \lambda_n).$$

2.22 Let $\{v_i\}$ be a basis for $\mathcal{H}$, so that $\{v_{i_1} \odot \cdots \odot v_{i_p}\}$ with $1 \le i_1 \cdots \le i_p \le \dim V$ is a basis for $\mathrm{Sym}^p \mathcal{H}$ and $\{v_{i_1} \wedge \cdots \wedge v_{i_p}\}$ with $1 < i_1 \cdots < i_p < \dim V$ is a basis for $\bigwedge^p \mathcal{H}$.

(i).

$$\begin{aligned}&\widehat{a}(u)\widehat{a}(v) v_1 \odot \cdots \odot v_p \\ &\quad = \widehat{a}(u) \sum_{i=1}^{p} (v, v_i) v_1 \odot \cdots \odot \widehat{v_i} \odot \cdots \odot v_p \\ &\quad = \sum_{i=1}^{p} (v, v_i) \sum_{\substack{j=1 \\ j \ne i}}^{p-1} (u, v_j) v_1 \odot \cdots \odot \widehat{v_i} \odot \cdots \odot \widehat{v_j} \cdots \odot v_p.\end{aligned}$$

This expression is symmetric in u and v, so $[\widehat{a}(u), \widehat{a}(v)]_- = 0$.

(ii).

$$\begin{aligned}\widehat{a}^{\dagger}(u)\widehat{a}^{\dagger}(v) v_1 \odot \cdots \odot v_p &= \widehat{a}^{\dagger}(u) v \odot v_1 \odot \cdots \odot v_p \\ &= u \odot v \odot v_1 \odot \cdots \odot v_p.\end{aligned}$$

This is manifestly symmetric in u and v, so $[\widehat{a}^{\dagger}(u), \widehat{a}^{\dagger}(v)]_- = 0$.

(iii).

$$\begin{aligned}\widehat{a}(u)\widehat{a}^{\dagger}(v) v_1 \odot \cdots \odot v_p &= \widehat{a}(u) v \odot v_1 \odot \cdots \odot v_p \\ &= {\sum_i}' (u, v_i) v \odot v_1 \odot \cdots \odot \widehat{v_i} \odot \cdots v_p,\end{aligned}$$

where the prime indicates that the sum includes the case in which $v_i = v$. On the other hand,

$$\begin{aligned}\widehat{a}^{\dagger}(v)\widehat{a}(u)v_1 \odot \cdots \odot v_p &= \widehat{a}^{\dagger}(v)\sum_i (u, v_i)v_1 \odot \cdots \odot \widehat{v_i} \odot \cdots v_p \\ &= \sum_i (u, v_i)v \odot v_1 \odot \cdots \odot \widehat{v_i} \odot \cdots v_p.\end{aligned}$$

Therefore,

$$[\widehat{a}(u), \widehat{a}^{\dagger}(v)]_- v_1 \odot \cdots \odot v_p = (u, v)v_1 \odot \cdots \odot v_p,$$

from which we conclude that

$$[\widehat{a}(u), \widehat{a}^{\dagger}(v)]_- = (u, v)\widehat{1}.$$

(iv).

$$\begin{aligned}&\widehat{b}(u)\widehat{b}(v)v_1 \wedge \cdots \wedge v_p \\ &\quad = \widehat{b}(u)\sum_i (-1)^i (v, v_i)v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_p \\ &\quad = \sum_{j<i} (v, v_i)(u, v_j)(-1)^{i+j} v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge \widehat{v_i} \cdots \wedge v_p \\ &\qquad + \sum_{j>i} (v, v_i)(u, v_j)(-1)^{i+j-1} v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge \widehat{v_j} \cdots \wedge v_p.\end{aligned} \tag{1}$$

Now change the names of the dummy indices in both terms from i and j to j and i, respectively, to get

$$\begin{aligned}&\widehat{b}(u)\widehat{b}(v)v_1 \wedge \cdots \wedge v_p \\ &\quad = \sum_{i<j} (v, v_j)(u, v_i)(-1)^{i+j} v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge \widehat{v_j} \cdots \wedge v_p \\ &\qquad + \sum_{i>j} (v, v_j)(u, v_i)(-1)^{i+j-1} v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge \widehat{v_i} \cdots \wedge v_p.\end{aligned} \tag{2}$$

If we now flip u and v in (2) and add it to (1) we get zero because the terms in the sums cancel pairwise. Hence, $[\widehat{b}(u), \widehat{b}(v)]_+ = 0$.

(v).

$$\begin{aligned}\widehat{b}^{\dagger}(u)\widehat{b}^{\dagger}(v)v_1 \wedge \cdots \wedge v_p &= \widehat{b}^{\dagger}(u)v \wedge v_1 \wedge \cdots \wedge v_p \\ &= u \wedge v \wedge v_1 \wedge \cdots \wedge v_p.\end{aligned}$$

Obviously, $[\widehat{b}^{\dagger}(u), \widehat{b}^{\dagger}(v)]_+ = 0$.

(vi).

$$\widehat{b}(u)\widehat{b}^{\dagger}(v)v_1 \wedge \cdots \wedge v_p = \widehat{b}(u)v \wedge v_1 \wedge \cdots \wedge v_p$$
$$= \sum_i{}' (-1)^i (u, v_i) v \wedge v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_p, \tag{3}$$

where the prime means that the sum includes the case $i = 0$ corresponding to $v_i = v$. On the other hand,

$$\widehat{b}^{\dagger}(v)\widehat{b}(u)v_1 \wedge \cdots \wedge v_p = \widehat{b}^{\dagger}(v)\sum_i (-1)^i (u, v_i) v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_p$$
$$= \sum_i (-1)^i (u, v_i) v \wedge v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_p. \tag{4}$$

If we add (3) and (4) all the terms cancel except the $i = 0$ term in (3), so

$$[\widehat{b}(u), \widehat{b}^{\dagger}(v)]_+ v_1 \wedge \cdots \wedge v_p = (u, v) v_1 \wedge \cdots \wedge v_p,$$

whereupon we conclude that $[\widehat{b}(u), \widehat{b}^{\dagger}(v)]_+ = (u, v)\widehat{1}$.

3

Differentiation on manifolds

3.1 For every point $y \in Y$ let $U(y) \subset Y$ be the open set whose existence is guaranteed by the hypothesis. Then the claim is that $Y = \cup_y U(y)$, which implies that Y is open in X. Clearly, $\cup_y U(y) \subseteq Y$. But if $y \in Y$ then $y \in U(y)$, so $Y \subseteq \cup_y U(y)$.

3.2 The first two properties hold by virtue of de Morgan's laws.

1. Let $\mathcal{C}$ be a collection of closed sets. Then

$$\overline{\bigcap_{C \in \mathcal{C}} C} = \bigcup_{C \in \mathcal{C}} \overline{C},$$

   which is open in X.

2. Let $\mathcal{C}$ be a finite collection of closed sets. Then

$$\overline{\bigcup_{C \in \mathcal{C}} C} = \bigcap_{C \in \mathcal{C}} \overline{C},$$

   which is open in X.

3. The empty set and X are both open, so they are also both closed.

3.3 Let U be a neighborhood of y. If $U \cap Y = \emptyset$ then there is a closed set, namely $\overline{U}$, that contains Y but does not contain y, so $y \notin \operatorname{cl} Y$. Conversely, suppose $y \notin \operatorname{cl} Y$, and let C be a closed set containing Y that does not contain y. Then $\overline{C}$ is a neighborhood of y that misses Y.

3.4 Let Y be closed and suppose x is an accumulation point of Y. If $x \notin Y$ then $x \in \operatorname{cl}(Y - \{x\}) = \operatorname{cl} Y = Y$, a contradiction. Conversely, suppose Y contains all its accumulation points, and let $x \in \overline{Y}$. Then x is not an accumulation point, so $x \notin \operatorname{cl}(Y - \{x\}) = \operatorname{cl} Y$. By the previous exercise, this means there is an open neighborhood of x that does not meet Y, so $\overline{Y}$ must be open (being the union of open sets). Therefore, Y is closed.

3.5 Let X be Hausdorff, and suppose $x \in X$. If $y \in \overline{X}$, then there is an open set U containing y but not x. Therefore $\overline{X}$ is open (being the union of open sets), so x is closed.

3.6 Let X be compact, and let $Y \subseteq X$ be closed. Let $\mathcal{U}$ be an open cover of Y. Then $\mathcal{U} \cup \overline{Y}$ is an open cover of X and so admits a finite subcover $\mathcal{U}'$ of X. But Y is not covered by $\overline{Y}$, so it must be covered by all the sets in $\mathcal{U}'$ distinct from $\overline{Y}$, namely by a finite subcollection of the sets in $\mathcal{U}$.

3.7 Let K be closed in Y. Then $Y \backslash K$ is open. Now, $f^{-1}(Y \backslash K) = X \backslash f^{-1}(K)$, because the inverse image of Y is X itself, so everything not in $f^{-1}(K)$ maps to $Y \backslash K$. By hypothesis, $f^{-1}(K)$ is closed, so $X \backslash f^{-1}(K)$ is open. As every open set is the complement of some closed set, the inverse image of every open set is open, so f is continuous.

3.8 Label the two types of continuity C1 (for the inverse image definition) and C2 (for the epsilon-delta definition). Assume f is C1, and let $p \in \mathbb{R}$. The interval $I := (f(p) - \epsilon, f(p) + \epsilon)$ is an open subset of $\mathbb{R}$, so $f^{-1}(I)$ is open as well. But $f^{-1}(I)$ contains p, so it must contain an open interval (a, b) containing p (because every open set in $\mathbb{R}$ is a union of open intervals). Now let $\delta = \min\{p - a, b - p\}$. Then the condition $|x - p| < \delta$ guarantees that $x \in (a, b)$, and hence that $f(x) \in I$, which means that $|f(x) - f(p)| < \epsilon$. Hence f is C2.

To prove the converse it suffices to show that the inverse image of a single open interval is open, because the inverse image of a union of sets is the union of the inverse images. So let $I = (c, d)$, and let $p \in f^{-1}(I)$ so that $f(p) \in I$. Choose $\epsilon = \min\{f(p) - c, d - f(p)\}$. As f is C2, there exists a δ such that the open interval $(p - \delta, p + \delta)$ is in $f^{-1}(I)$. Therefore $f^{-1}(I)$ is the union of open sets, which shows that f is C1.

3.9 Let $\mathcal{U}$ be an open cover of Y. By continuity and surjectivity, $\{f^{-1}(U) : U \in \mathcal{U}\}$ is an open cover of X, which therefore has a finite subcover $\mathcal{U}'$. By construction the images of the open sets in the finite set $\mathcal{U}'$ are elements of $\mathcal{U}$, and by surjectivity they cover Y.

3.10 Let U be open in X. Then $\iota^{-1}(U) = Y \cap U$, which is clearly open in τ_Y, so ι is continuous in the subspace topology. On the other hand, if ι is continuous in some topology σ on Y, then for any open set U in X, $\iota^{-1}(U) = Y \cap U$ is open in σ, so if $Y \cap U$ were not in σ then ι would not be continuous.

3.11 Let X be Hausdorff and let Y be a subspace of X. If p and q are distinct points of Y then they are distinct in X, so there exist subsets U and V with $U \cap V = \emptyset$ and $p \in U$ and $q \in V$. But then $(Y \cap U) \cap (Y \cap V) = Y \cap (U \cap V) = \emptyset$ as well, so Y is Hausdorff.

3.12 By Exercise 3.5, y is closed in Y, so by Exercise 3.7, $f^{-1}(y)$ is closed. But by Exercise 3.6, the closed subset of a compact set is compact.

3.13 We construct f as follows. Let $[x] \in Y$ and choose some $x \in [x]$. Define $f([x]) := g(x)$. This is well defined, because if x' is any other point in $[x]$ then by hypothesis, $g(x') = g(x)$. Moreover, it is determined uniquely by g. Lastly, if U is open in Z then by the continuity of g, $g^{-1}(U) = \pi^{-1} \circ f^{-1}(U)$ is open in X. So by the definition of the quotient topology, $f^{-1}(U)$ is open in Y, so f is continuous.

3.14 It suffices to show that the Jacobian matrix equals the matrix $(f^i{}_j)$ representing f itself relative to the standard bases. By linearity,

$$f(x) = f(\sum_j x^j e_j) = \sum_{ij} e_i f^i{}_j x^j,$$

so the i^{th} coordinate function is $f^i(x) = \sum_j f^i{}_j x^j$. Taking partial derivatives gives

$$\frac{\partial f^i}{\partial x^j} = f^i{}_j.$$

3.15 Following the hint, we observe that $f \circ f^{-1} = id$ implies that $(Df)(x) \cdot (Df^{-1})(y) = I$, which shows that $(Df^{-1})(y) = (Df(x))^{-1}$. Well, ok, that skips a bunch a steps. To fill in a few of those steps, we have to show that the chain rule applied to multivariate functions just gives the product of the Jacobian matrices, and that the derivative of the identity map (as a multivariate function) is the identity map (as a linear operator). For the first, we have

$$\begin{aligned}[(D(f \circ g))(x)]_{ij} &= \frac{\partial (f \circ g)^i}{\partial x_j} = \frac{\partial f^i}{\partial x^k}(g(x)) \cdot \frac{\partial g^k}{\partial x^j}(x) \\ &= [(Df)(g(x))]_{ik}[(Dg)(x)]_{kj} = [(Df)(g(x)) \cdot (Dg)(x)]_{ij}.\end{aligned}$$

For the second, note that $id(x) = x$, so in components $(id)^i(x) = x^i$. Applying the definition of the derivative gives

$$(D\,id)(x) = \frac{\partial (id)^i}{\partial x^j} = \frac{\partial x^i}{\partial x^j} = \delta^i_j.$$

[Incidentally, this is why it is a bit dangerous to write ‘1’ for the identity, because one might think that the derivative of ‘1’ should be 0!]

3.16 We consider the nine coordinate charts suggested in the problem hint. First we observe that these patches do indeed cover the torus. Next, we must show that the transition maps between patches are diffeomorphisms. We do this for two overlaps, leaving the rest to your imagination. On $U_1 \cap U_2$ we have $(\varphi_1 \circ \varphi_2^{-1})(x, y) = (x, y)$ for the same reason as in the Möbius case. On $U_2 \cap U_8$ (where $U_8 = \{[x, y] : 2/3 < x < 4/3, 1/3 < y < 1\}$) we have, for example, $\varphi_2^{-1}(1/6, 1/2) = [1/6, 1/2] = [7/6, 1/2]$, so

$(\varphi_8 \circ \varphi_2^{-1})(1/6, 1/2) = (7/6, 1/2)$. In general, $(\varphi_8 \circ \varphi_2^{-1})(x, y) = (x+1, y)$. In fact, we see that all the transition functions are either the identity map or else a translation through 1 in one of the coordinate directions. Therefore, the Jacobians all equal $+1$, and the torus is seen to be a smooth orientable manifold.

3.17 First we observe that every point of $\mathbb{RP}^n$ is in some U_i, because the zero vector is not in $\mathbb{RP}^n$, so at least one entry in $[x^0, x^1, \ldots, x^n]$ must be nonzero. We must show that the charts are compatible. To do this, we show that $\varphi_i \circ \varphi_j^{-1}$ is a diffeomorphism of $\varphi_i(U_i \cap U_j)$ and $\varphi_j(U_i \cap U_j)$. Let $p \in U_i \cap U_j$. By construction, $p = [x^0, x^1, \ldots, x^n]$ where $x^i \neq 0$ and $x^j \neq 0$. We have

$$\varphi_i(p) = \left(\frac{x^0}{x^i}, \frac{x^1}{x^i}, \ldots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i} \ldots, \frac{x^n}{x^i}\right)$$

and

$$\varphi_j(p) = \left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \ldots, \frac{x^{j-1}}{x^j}, \frac{x^{j+1}}{x^j} \ldots, \frac{x^n}{x^j}\right).$$

We construct an inverse map from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}$ for $0 \leq j \leq n$ by

$$\varphi_j^{-1}(y^1, \ldots, y^n) = (y^1, \ldots, \underbrace{1}_{j^{\text{th}}\text{ position}}, \ldots, y^n)$$

Let's verify that this is indeed an inverse. One direction is trivial. We have

$$\begin{aligned}(\varphi_j \circ \varphi_j^{-1})(y^1, \ldots, y^n) &= \varphi_j(y^1, \ldots, \underbrace{1}_{j^{\text{th}}\text{ position}}, \ldots, y^n)\\ &= (y^1, \ldots, y^n).\end{aligned}$$

For the other direction, we have

$$\begin{aligned}(\varphi_j^{-1} \circ \varphi_j)(p) &= \varphi_j^{-1}\left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \ldots, \frac{x^{j-1}}{x^j}, \frac{x^{j+1}}{x^j} \ldots, \frac{x^n}{x^j}\right)\\ &= \left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \ldots, \frac{x^{j-1}}{x^j}, 1, \frac{x^{j+1}}{x^j} \ldots, \frac{x^n}{x^j}\right)\\ &= p,\end{aligned}$$

because p represents the equivalence class of $[x^0, x^1, \ldots, x^n]$.

To show that $\varphi_i \circ \varphi_j^{-1}$ is a diffeomorphism, we must show it is a smooth bijection. If p is as above, then

$$
\begin{aligned}
(\varphi_i \circ \varphi_j^{-1}) & \left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \dots, \frac{x^{j-1}}{x^j}, \frac{x^{j+1}}{x^j} \dots, \frac{x^n}{x^j}\right) \\
&= \varphi_i \left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \dots, \frac{x^{j-1}}{x^j}, 1, \frac{x^{j+1}}{x^j} \dots, \frac{x^n}{x^j}\right) \\
&= \frac{1}{x^i/x^j} \left(\frac{x^0}{x^j}, \frac{x^1}{x^j}, \dots, \widehat{\frac{x^i}{x^j}}, \dots, \underbrace{1}_{j^{\text{th}}\text{ place}}, \dots, \frac{x^n}{x^j}\right) \\
&= \left(\frac{x^0}{x^i}, \frac{x^1}{x^i}, \dots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i} \dots, \frac{x^n}{x^i}\right).
\end{aligned}
$$

(The caret means that the term is excised.) This is clearly a bijection from $\phi_j(U_i \cap U_j)$ to $\phi_i(U_i \cap U_j)$.

To show $\varphi_i \circ \varphi_j^{-1}$ is smooth, we proceed as follows. If $i = j$ then $\varphi_i \circ \varphi_j^{-1}$ is just the identity map, so there is nothing to prove. (The identity map is always a diffeomorphism.) So we examine the cases in which $i \neq j$. For concreteness, let us choose $i = 0$ and $j = 1$, as the other cases are similar. Then, by the above discussion, we have (with $x^0 \neq 0$ and $x^1 \neq 0$)

$$
(\varphi_0 \circ \varphi_1^{-1}) \left(\frac{x^0}{x^1}, \frac{x^2}{x^1}, \dots, \frac{x^n}{x^1}\right) = \left(\frac{x^1}{x^0}, \frac{x^2}{x^0}, \dots, \frac{x^n}{x^0}\right).
$$

That is,

$$
(\varphi_0 \circ \varphi_1^{-1})(y^1, y^2, \dots, y^n) = \left(\frac{1}{y^1}, \frac{y^2}{y^1}, \dots, \frac{y^n}{y^1}\right),
$$

where $y_1 \neq 0$. This map is infinitely differentiable (smooth), so we are done.

The Jacobian of this map is

$$
\begin{pmatrix}
-1/(y^1)^2 & 0 & 0 & \cdots & 0 \\
-y^2/(y^1)^2 & 1/y^1 & 0 & \cdots & 0 \\
-y^3/(y^1)^2 & 0 & 1/y^1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-y^n/(y^1)^2 & 0 & 0 & \cdots & 1/y^1
\end{pmatrix}.
$$

It has determinant $-1/(y^1)^{n+1}$. If n is even, this varies with the sign of y^1, so no consistent orientation is possible. If n is odd, this is negative for every pair of overlaps, so a consistent choice of orientation is possible. We conclude that if n is odd, $\mathbb{RP}^n$ is orientable, but if n is even we cannot say anything (because we do not know if some other choice of coordinates would make the manifold orientable). It turns out, in fact, that for n even, the manifold is not orientable.

3.18 Define $f : \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$f(x) = \sum_i (x^i)^2.$$

As $\mathbb{R}$ and $\mathbb{R}^{n+1}$ are obviously manifolds and as S^n is $f^{-1}(1)$, we need only show that 1 is a regular value of f. For this, we compute the Jacobian matrix, which is a 1 by $n+1$ matrix whose j^{th} entry is $\partial f/\partial x^j = 2x^j$. This matrix is rank deficient (i.e., less than one) only at $x = 0$, but fortunately for us, $x = 0$ is not on the sphere, so all is well.

3.19 Applying the chain rule gives

$$\begin{aligned}
\frac{\partial}{\partial r} &= \left(\frac{\partial x}{\partial r}\right)\frac{\partial}{\partial x} + \left(\frac{\partial y}{\partial r}\right)\frac{\partial}{\partial y} + \left(\frac{\partial z}{\partial r}\right)\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \theta} &= \left(\frac{\partial x}{\partial \theta}\right)\frac{\partial}{\partial x} + \left(\frac{\partial y}{\partial \theta}\right)\frac{\partial}{\partial y} + \left(\frac{\partial z}{\partial \theta}\right)\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \phi} &= \left(\frac{\partial x}{\partial \phi}\right)\frac{\partial}{\partial x} + \left(\frac{\partial y}{\partial \phi}\right)\frac{\partial}{\partial y} + \left(\frac{\partial z}{\partial \phi}\right)\frac{\partial}{\partial z}.
\end{aligned}$$

Substituting into these equations using

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \sin\theta\cos\phi & \frac{\partial y}{\partial r} &= \sin\theta\sin\phi & \frac{\partial z}{\partial r} &= \cos\theta \\
\frac{\partial x}{\partial \theta} &= r\cos\theta\cos\phi & \frac{\partial y}{\partial \theta} &= r\cos\theta\sin\phi & \frac{\partial z}{\partial \theta} &= -r\sin\theta \\
\frac{\partial x}{\partial \phi} &= -r\sin\theta\sin\phi & \frac{\partial y}{\partial \phi} &= r\sin\theta\cos\phi & \frac{\partial z}{\partial \phi} &= 0,
\end{aligned}$$

gives

$$\begin{aligned}
\frac{\partial}{\partial r} &= \sin\theta\cos\phi\,\frac{\partial}{\partial x} + \sin\theta\sin\phi\,\frac{\partial}{\partial y} + \cos\theta\,\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \theta} &= r\cos\theta\cos\phi\,\frac{\partial}{\partial x} + r\cos\theta\sin\phi\,\frac{\partial}{\partial y} - r\sin\theta\,\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \phi} &= -r\sin\theta\sin\phi\,\frac{\partial}{\partial x} + r\sin\theta\cos\phi\,\frac{\partial}{\partial y}.
\end{aligned}$$

Now we have to normalize the vectors by dividing by their lengths. $\partial/\partial r$ is already a unit vector, so we just identify $e_r = \partial/\partial r$. $\partial/\partial\theta$ has length r, so $e_\theta = (1/r)\partial/\partial\theta$. Finally, $\partial/\partial/\phi$ has length $r\sin\theta$, so $e_\phi = (1/r\sin\theta)\partial/\partial\phi$.

3.20 a. The product of linear maps is linear, as is the sum, so $[X, Y]$ is linear. Also

$$\begin{aligned}
[X,Y](fg) &= (XY)(fg) - (YX)(fg) \\
&= X(gYf + fYg) - Y(gXf + fXg) \\
&= (Yf)(Xg) + g(XYf) + (Yg)(Xf) + f(XYg) \\
&\quad - (Xf)(Yg) - g(YXf) - (Xg)(Yf) - f(YXg) \\
&= g[X,Y]f + f[X,Y]g,
\end{aligned}$$

so $[X, Y]$ is also a derivation.

b. In the given local coordinates,

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

Subtracting a term just like this with X and Y interchanged gives

$$[X,Y]f = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right)\frac{\partial f}{\partial x^j},$$

because the terms with the mixed partial derivatives of f cancel. Hence

$$[X,Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right)\frac{\partial}{\partial x^j},$$

as desired.

c.

$$\begin{aligned}
&[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \\
&\quad = [X, YZ - ZY] + [Y, ZX - XZ] + [Z, XY - YX], \\
&\quad = [X, YZ] - [X, ZY] + [Y, ZX] - [Y, XZ] + [Z, XY] - [Z, YX], \\
&\quad = XYZ - YZX - XZY + ZYX + YZX - ZXY \\
&\qquad\qquad - YXZ + XZY + ZXY - XYZ - ZYX + YXZ, \\
&\quad = 0.
\end{aligned}$$

3.21 Plug (3.37) into (3.38) to get

$$a_i(x) = a_k(x)\frac{\partial x^k}{\partial y^j}\frac{\partial y^j}{\partial x^i} = a_k(x)\delta_i^k = a_i(x).$$

Similarly we obtain the identity $b_j(y) = b_j(y)$ if we plug (3.38) into (3.37), so the two formulae are inverses of one another.

Next, using $X = X^i(x)\partial/\partial x^i = X^j(y)\partial/\partial y^j$, $\alpha = a_i(x)\,dx^i = b_j(y)\,dy^j$ and (3.24) gives

$$X^i(x)a_i(x) = \langle X, \alpha\rangle = Y^j(y)b_j(y) = b_j(y)X^i(x)\frac{\partial y^j}{\partial x^i}.$$

Stripping off X^i from both sides we get (3.38).

3.22 Under a change of basis from x to y, $X = X^{i'}(\partial/\partial y^{i'})$, where

$$X^{i'} = X^i\frac{\partial y^{i'}}{\partial x^i}.$$

It follows that

$$\frac{\partial X^{i'}}{\partial y^{j'}} = \frac{\partial X^i}{\partial x^j}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial y^{i'}}{\partial x^i} + X^i\frac{\partial^2 y^{i'}}{\partial x^k \partial x^i}\frac{\partial x^k}{\partial y^{j'}},$$

which differs from (3.53) by the second term.

3.23 Set

$$dx_{U_2} = dx_{U_1} = dx^1$$

on the overlap $U_1 \cap U_2$. But on the overlap,

$$dx^1 = \frac{\partial x^1}{\partial x^2}dx^2 = dx^2,$$

because the Jacobian 'matrix' is just unity. So, if we choose $dx_{U_2} = dx^2$ everywhere on U_2, then it agrees with dx_{U_1} on the overlap. Similar reasoning shows that, if we choose $dx_{U_i} = dx^i$ on the i^{th} patch, the 1-forms all (smoothly) agree on the overlaps, so together they define a global 1-form dx. The same argument works for dy.

3.24

$$\begin{aligned}
df \wedge df &= \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right) \wedge \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)\\
&= \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}(dx \wedge dy + dy \wedge dx) + \frac{\partial f}{\partial y}\frac{\partial f}{\partial z}(dy \wedge dz + dz \wedge dy)\\
&\quad + \frac{\partial f}{\partial z}\frac{\partial f}{\partial x}(dz \wedge dx + dx \wedge dz)\\
&= 0.
\end{aligned}$$

3.25

$$\begin{aligned}
d\alpha &= 6xy^3z\,dx \wedge dx + 9x^2y^2z\,dy \wedge dx + 3x^2y^3\,dz \wedge dx\\
&\quad + y^3z^2\,dx \wedge dy + 3xy^2\,dy \wedge dy + 2xy^3z\,dz \wedge dy\\
&\quad + 2z^2\,dy \wedge dz + 4yz\,dz \wedge dz\\
&= (y^3z^2 - 9x^2y^2z)\,dx \wedge dy + (2z^2 - 2xy^3z)\,dy \wedge dz + 3x^2y^3\,dz \wedge dx,
\end{aligned}$$

and

$$\begin{aligned} d^2\alpha &= -18xy^2z\,dx \wedge dx \wedge dy + (3y^2z^2 - 18x^2yz)\,dy \wedge dx \wedge dy \\ &\quad + (2y^3z - 9x^2y^2)\,dz \wedge dx \wedge dy - 2y^3z\,dx \wedge dy \wedge dz \\ &\quad - 6xy^2z\,dy \wedge dy \wedge dz + (4z - 2xy^3)\,dz \wedge dy \wedge dz \\ &\quad + 6xy^3\,dx \wedge dz \wedge dx + 9x^2y^2\,dy \wedge dz \wedge dx \\ &= 0. \end{aligned}$$

3.26

$$d\omega = (2x^2y - 3yz^3)\,dx \wedge dy \wedge dz,$$

and $d^2\omega = 0$ because all the 4-forms vanish.

3.27

$$d^2(a\,dx^I) = d\sum_k \frac{\partial a}{\partial x^k}\,dx^k \wedge dx^I = \sum_{k\ell} \frac{\partial^2 a}{\partial x^\ell \partial x^k}\,dx^\ell \wedge dx^k \wedge dx^I = 0,$$

because mixed partial derivatives commute, but wedge products anticommute. (See Exercise 2.2.)

3.28 Starting again from

$$\begin{aligned} F &= -E_x\,dt \wedge dx - E_y\,dt \wedge dy - E_z\,dt \wedge dz \\ &\quad + B_x\,dy \wedge dz + B_y\,dz \wedge dx + B_z\,dx \wedge dy, \end{aligned}$$

Applying the exterior derivative operator and collecting terms we get

$$\begin{aligned} dF &= \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} + \frac{\partial B_z}{\partial t}\right)(dt \wedge dx \wedge dy) \\ &\quad + \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} + \frac{\partial B_x}{\partial t}\right)(dt \wedge dy \wedge dz) \\ &\quad + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} + \frac{\partial B_y}{\partial t}\right)(dt \wedge dy \wedge dz) \\ &\quad + \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right)(dx \wedge dy \wedge dz). \end{aligned}$$

If $dF = 0$ then each term vanishes individually, whereupon we obtain Maxwell's equations without sources.

3.29 Starting from

$$\begin{aligned} \star J &= -\rho\,dx \wedge dy \wedge dz + J_x\,dt \wedge dy \wedge dz \\ &\quad + J_y\,dt \wedge dz \wedge dx + J_z\,dt \wedge dx \wedge dy. \end{aligned}$$

we get

$$d \star J = \left(-\frac{\partial \rho}{\partial t} - \frac{\partial J_x}{\partial x} - \frac{\partial J_y}{\partial y} - \frac{\partial J_z}{\partial z} \right) dt \wedge dx \wedge dy \wedge dz.$$

Taking 'd' of the Maxwell's equation with sources gives

$$0 = d^2 \star F = 4\pi d \star J,$$

whence we obtain the law of charge conservation.

3.30 a. Let $X = X^i(\partial/\partial x^i)$ (implicit sum). Then, as

$$df = \frac{\partial f}{\partial x^i} dx^i,$$

property (2) of the interior product and the linearity of the dual pairing, yield

$$i_X(df) = df(X) = \langle df, X \rangle = X^i \frac{\partial f}{\partial x^i} = Xf.$$

b. The first thing to observe is that the dual pairing of a one form ω and a vector field X is function linear:

$$\langle \omega, fX \rangle = f \langle \omega, X \rangle ,$$

which follows immediately from its definition. Hence, on one forms,

$$i_{fX}\omega = \langle \omega, fX \rangle = f \langle \omega, X \rangle = f i_X \omega$$

Now let η be a $p-1$ form, so that $\lambda = \omega \wedge \eta$ is a p form. Then

$$\begin{aligned} i_{fX}\lambda &= i_{fX}\omega \wedge \eta - \omega \wedge i_{fX}\eta \\ &= (f i_X \omega) \wedge \eta - f\omega \wedge i_X \eta \\ &= f i_X \lambda, \end{aligned}$$

where the second step follows by the induction hypothesis. As any p form can be written as a sum of monomials, the claim is proved.

c. Let $\omega \wedge \eta$ be a p-form. Then by Property (iii) of the interior product,

$$\begin{aligned} i_X i_Y(\omega \wedge \eta) &= i_X(i_Y\omega \wedge \eta + (-1)^{\deg \omega}\omega \wedge i_Y\eta) \\ &= i_X i_Y \omega \wedge \eta + (-1)^{\deg \omega - 1} i_Y\omega \wedge i_X\eta \\ &\qquad + (-1)^{\deg \omega}[i_X\omega \wedge i_Y\eta + (-1)^{\deg \omega}\omega \wedge i_X i_Y \eta] \\ &= i_X i_Y\omega \wedge \eta + \omega \wedge i_X i_Y \eta \\ &\qquad + (-1)^{\deg \omega}[i_X\omega \wedge i_Y\eta - i_Y\omega \wedge i_X\eta]. \end{aligned}$$

By induction on p, the first two terms of this expression are antisymmetric under the interchange of X and Y. As the last term is manifestly antisymmetric under the interchange of X and Y, we have shown that, on p forms,

$$i_X i_Y = -i_Y i_X,$$

which proves the claim.

d. Using the properties of the interior product we get

$$\begin{aligned}
\omega(X_1, X_2, X_3) &= i_{X_3} i_{X_2} i_{X_1} \omega \\
&= i_{X_3} i_{X_2} i_{X_1} (\lambda \wedge \mu \wedge \nu) \\
&= i_{X_3} i_{X_2} (i_{X_1} \lambda \wedge \mu \wedge \nu - \lambda \wedge i_{X_1} (\mu \wedge \nu)) \\
&= i_{X_3} i_{X_2} (i_{X_1} \lambda \wedge \mu \wedge \nu - \lambda \wedge i_{X_1} \mu \wedge \nu + \lambda \wedge \mu \wedge i_{X_1} \nu) \\
&= i_{X_3} i_{X_2} (\lambda(X_1) \mu \wedge \nu - \mu(X_1) \lambda \wedge \nu + \nu(X_1) \lambda \wedge \mu) \\
&= i_{X_3} (\lambda(X_1)(i_{X_2} \mu \wedge \nu - \mu \wedge i_{X_2} \nu) \\
&\qquad - \mu(X_1)(i_{X_2} \lambda \wedge \nu - \lambda \wedge i_{X_2} \nu) \\
&\qquad + \nu(X_1)(i_{X_2} \lambda \wedge \mu - \lambda \wedge i_{X_2} \mu)) \\
&= i_{X_3} (\lambda(X_1)(\mu(X_2)\nu - \nu(X_2)\mu) \\
&\qquad - \mu(X_1)(\lambda(X_2)\nu - \nu(X_2)\lambda) \\
&\qquad + \nu(X_1)(\lambda(X_2)\mu - \mu(X_2)\lambda)) \\
&= \lambda(X_1)\mu(X_2)\nu(X_3) - \lambda(X_1)\mu(X_3)\nu(X_2) \\
&\qquad - \lambda(X_2)\mu(X_1)\nu(X_3) + \lambda(X_3)\mu(X_1)\nu(X_2) \\
&\qquad + \lambda(X_2)\mu(X_3)\nu(X_1) - \lambda(X_3)\mu(X_2)\nu(X_1).
\end{aligned}$$

e. We see immediately that $c_3 = 1$, and in general $c_p = 1$.

3.31 We choose the easy way. If $\{X_1, X_2, \dots, X_n\}$ are linearly dependent, then we may write

$$X_n = \sum_{i=1}^{n-1} a_i X_i$$

for some constants a_i, not all of which are zero. By multilinearity, we have

$$\omega(X_1, X_2, \dots, X_{n-1}, X_n) = \sum_{i=1}^{n-1} a_i \omega(X_1, X_2, \dots, X_{n-1}, X_i).$$

But, by antisymmetry, each summand vanishes, because in each term at least two of the vectors are the same.

3.32 We compute

$$\begin{aligned}
\phi^* \omega &= 3(t^2)(t)(2t\,dt) - 7(2t-1)^2(t)(2\,dt) + 2(t^2)^2(2t-1)(dt) \\
&= (6t^4 - 14t(2t-1)^2 + 2t^4(2t-1))\,dt \\
&= (4t^5 + 4t^4 - 56t^3 + 56t^2 - 14t)\,dt.
\end{aligned}$$

3.33 We compute

$$\begin{aligned}\phi^*\theta &= -(u^3\cos v)(3u^2\sin v\,du + u^3\cos v\,dv)\wedge(2u\,du)\\ &\qquad - (u^3\sin v)(2u\,du)\wedge(3u^2\cos v\,du - u^3\sin v\,dv)\\ &\qquad + 2(u^2)(3u^2\cos v\,du - u^3\sin v\,dv)\wedge(3u^2\sin v\,du + u^3\cos v\,dv)\\ &= (2u^7\cos^2 v + 2u^7\sin^2 v + 6u^7\cos^2 v + 6u^7\sin^2 v)du\wedge dv\\ &= 8u^7\,du\wedge dv.\end{aligned}$$

3.34 We have

$$\phi^*\omega = 4v\,du + 2uv\,dv + 2u\,du + 2v\,dv = 2(u+2v)\,du + 2v(u+1)\,dv,$$

so

$$d\phi^*\omega = 4\,dv\wedge du + 2v\,du\wedge dv = (2v-4)\,du\wedge dv.$$

On the other hand,

$$d\omega = 4\,dy\wedge dx + 2y\,dx\wedge dy = (2y-4)\,dx\wedge dy,$$

so

$$\phi^*d\omega = (2v-4)\,du\wedge dv,$$

and $d\phi^*\omega = \phi^*d\omega$.

3.35 Let $X = \partial/\partial x^i$. Then

$$X_p(f^*g) = \left.\frac{\partial(g\circ f)}{\partial x^i}\right|_p = \left.\frac{\partial g}{\partial y^j}\right|_{f(p)}\left.\frac{\partial f^j}{\partial x^i}\right|_p = \left.\frac{\partial y^j}{\partial x^i}\right|_p\left.\frac{\partial g}{\partial y^j}\right|_{f(p)}.$$

Also,

$$(f_*X_p)_{f(p)}(g) = \left(f_*\frac{\partial}{\partial x^i}\right)_{f(p)}(g),$$

so we conclude

$$f_*\left.\frac{\partial}{\partial x^i}\right|_{f(p)} = \left.\frac{\partial y^j}{\partial x^i}\right|_p\left.\frac{\partial}{\partial y^j}\right|_{f(p)}.$$

But recall that the components of the matrix representing the linear transformation $A : U \to V$ relative to the bases $\{e_i\}$ of U and $\{f_i\}$ of V are defined by

$$Ae_i = A_{ji}f_j.$$

It follows that the matrix representing f_* relative to the local bases $\partial/\partial x^i$ and $\partial/\partial y^j$ has components $(f_*)_{ji} = (\partial y^j/\partial x^i)_p$, whence the conclusion follows.

3.36 Suppose $h : M \to \mathbb{R}$ is a function. Then if f is the identity map, $(f^*h)(x) = h(f(x)) = h(x)$, which shows that f^* is just the identity. Similarly, $f(x) = x$ implies $y^i = f^i(x) = x^i$, and the Jacobian matrix $(\partial y^i/\partial x^j)$ representing f_* is just the identity matrix.

If $f : U \to V$ and $g : V \to W$ and $h : W \to \mathbb{R}$, then

$$(g \circ f)^*h = h \circ (g \circ f) = (h \circ g) \circ f = f^*(h \circ g) = f^*g^*h,$$

so $(g \circ f)^* = f^* \circ g^*$. Lastly, by Exercise 3.35, f_* is represented locally by the Jacobian of the transformation, so by the chain rule, $(g \circ f)_* = g_* f_*$.

3.37 4. We have

$$\phi_t^* \langle \omega, Y \rangle \big|_p = \langle \omega, Y \rangle|_q = \langle \omega_q, Y_q \rangle = \omega(Y)_q.$$

But also

$$\begin{aligned} \left\langle \phi_t^* \omega, \phi_t^* Y \right\rangle\big|_p &= \left\langle (\phi_t^*\omega)_p, (\phi_t^* Y)_p \right\rangle = \left\langle (\phi_t^*\omega)_p, (\phi_{-t*} Y)_p \right\rangle \\ &= \omega(\phi_{t*}\phi_{-t*}Y)|_q = \omega(Y)_q, \end{aligned}$$

so,

$$\phi_t^* \langle \omega, Y \rangle = \left\langle \phi_t^* \omega, \phi_t^* Y \right\rangle,$$

i.e., pullback commutes with contraction. Applying this result gives

$$\begin{aligned} \mathcal{L}_X \langle \omega, Y \rangle &= \lim_{t\to 0} \frac{\phi_t^* \langle \omega, Y \rangle - \langle \omega, Y \rangle}{t} \\ &= \lim_{t\to 0} \frac{\left\langle \phi_t^*\omega, \phi_t^* Y \right\rangle - \langle \omega, Y \rangle}{t} \\ &= \lim_{t\to 0} \frac{\left\langle \phi_t^*\omega, \phi_t^* Y \right\rangle - \left\langle \phi_t^*\omega, Y \right\rangle + \left\langle \phi_t^*\omega, Y \right\rangle - \langle \omega, Y \rangle}{t} \\ &= \lim_{t\to 0} \left\langle \phi_t^*\omega, \frac{\phi_t^* Y - Y}{t} \right\rangle + \lim_{t\to 0} \left\langle \frac{\phi_t^*\omega - \omega}{t}, Y \right\rangle \\ &= \langle \omega, \mathcal{L}_X Y \rangle + \langle \mathcal{L}_X \omega, Y \rangle . \end{aligned}$$

5. Taking the hint, we observe that $\mathcal{L}_X d = d\mathcal{L}_X$ because pullback commutes with d. So, invoking 4 we get

$$\begin{aligned} X(Y(f)) = \mathcal{L}_X(Y(f)) &= \mathcal{L}_X \langle df, Y \rangle \\ &= \langle \mathcal{L}_X df, Y \rangle + \langle df, \mathcal{L}_X Y \rangle \\ &= \langle d\mathcal{L}_X f, Y \rangle + (\mathcal{L}_X Y)(f) \\ &= \langle dX(f), Y \rangle + (\mathcal{L}_X Y)(f) \\ &= Y(X(f)) + (\mathcal{L}_X Y)(f), \end{aligned}$$

whereupon we conclude that

$$\mathcal{L}_X Y = XY - YX = [X, Y].$$

6. By 5 and the Jacobi identity,

$$\begin{aligned}
[\mathcal{L}_X, \mathcal{L}_Y]Z &= \mathcal{L}_X \mathcal{L}_Y Z - \mathcal{L}_Y \mathcal{L}_X Z \\
&= \mathcal{L}_X [Y, Z] - \mathcal{L}_Y [X, Z] \\
&= [X, [Y, Z]] - [Y, [X, Z] \\
&= [X, [Y, Z]] + [Y, [Z, X]] \\
&= [[X, Y], Z] \\
&= \mathcal{L}_{[X,Y]} Z.
\end{aligned}$$

3.38 Bisect the circle by the line $L = \{(x, 0) \in \mathbb{R}^2$ and let $U = S^1 - \{N\}$ and $V = S^1 - \{S\}$, where the 'north pole' is $N = (0, 1)$ and the 'south pole' is $S = (0, -1)$. Let φ be the projection onto L from N and ψ be the projection onto L from S. Let $\varphi(P) = Q$ and let the coordinate function on U be u. The parametric line from N to $Q \in L$ intersects the circle at

$$(x_P, y_P) = (0, 1) + t(u, -1) = (tu, 1 - t).$$

We have

$$t^2 u^2 + (1 - t)^2 = 1 \quad \Rightarrow \quad t = \frac{2}{1 + u^2},$$

so

$$x_P = \frac{2u}{1 + u^2} \qquad \text{and} \qquad y_P = -\frac{1 - u^2}{1 + u^2}.$$

Projecting from the south pole instead gives

$$x_P = \frac{2v}{1 + v^2} \qquad \text{and} \qquad y_P = \frac{1 - v^2}{1 + v^2}.$$

where v is the coordinate function on V. Equating x_P and y_P in both coordinate systems gives

$$v(u) = (\psi \circ \varphi^{-1})(u) = \frac{1}{u}.$$

This is clearly a diffeomorphism on the overlap (where $u \neq 0$), and the Jacobian of the transformation is $-1/u^2 < 0$, so the circle is orientable.

3.39 The patches $U_i \times V_j$ clearly cover $M \times N$, so we need only show that the coordinate maps are compatible. First we observe that

$$(\varphi_i \times \varphi_j)^{-1}(x, y) = (\varphi_i^{-1}(x), \varphi_j^{-1}(y)).$$

On the overlap $(U_i \times V_j) \cap (U_k \times V_\ell)$,

$$(\varphi_i \times \varphi_j) \circ (\varphi_k \times \varphi_\ell)^{-1}(x, y) = (\varphi_i \circ \varphi_k^{-1}(x), \varphi_j \circ \varphi_\ell^{-1}(y)).$$

This map is a diffeomorphism, as each component of the map is a diffeomorphism. To show orientability, it suffices to show that the sign of the Jacobian of the composite map $f \times g$ is the product of the signs of the Jacobians of the component maps f and g. For this, we have, symbolically,

$$\begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{vmatrix} = \left[\frac{\partial f}{\partial x}\right]\left[\frac{\partial g}{\partial y}\right],$$

whence the assertion follows.

3.40 a. In spherical polar coordinates

$$\begin{aligned} r &= (x^2 + y^2 + z^2), \\ \theta &= \cos^{-1}(z/r), \\ \phi &= \tan^{-1}(y/x). \end{aligned}$$

For future use we observe that

$$\frac{\partial r}{\partial x} = \frac{1}{2}(x^2 + y^2 + z^2)^{-1/2} \cdot 2x = \frac{x}{r},$$

and similarly for y and z. Also,

$$\begin{aligned} \frac{\partial}{\partial x}\left(\frac{z}{r}\right) &= -\frac{z}{r^2} \cdot \frac{x}{r} = -\frac{xz}{r^3}, \\ \frac{\partial}{\partial y}\left(\frac{z}{r}\right) &= -\frac{z}{r^2} \cdot \frac{y}{r} = -\frac{yz}{r^3}, \\ \frac{\partial}{\partial z}\left(\frac{z}{r}\right) &= \frac{r - z(z/r)}{r^2} = \frac{r^2 - z^2}{r^3} = \frac{x^2 + y^2}{r^3}. \end{aligned}$$

Therefore

$$\begin{aligned} d\theta &= -\frac{1}{\sqrt{1 - (z/r)^2}}\left(\frac{\partial(z/r)}{\partial x}\,dx + \frac{\partial(z/r)}{\partial y}\,dy + \frac{\partial(z/r)}{\partial z}\,dz\right) \\ &= -\frac{1}{\sqrt{1 - (z/r)^2}}\left(-\frac{xz}{r^3}\,dx - \frac{yz}{r^3}\,dy + \frac{x^2 + y^2}{r^3}\,dz\right). \end{aligned}$$

Also,

$$\begin{aligned} d\phi &= \frac{1}{1+(y/x)^2}\left(-\frac{y}{x^2}\,dx + \frac{1}{x}\,dy\right) \\ &= \frac{1}{x^2+y^2}(-y\,dx + x\,dy). \end{aligned}$$

Note that

$$\sin\theta = \sqrt{1-\cos^2\theta} = \sqrt{1-(z/r)^2}.$$

Putting this all together gives

$$\begin{aligned} f^*\sigma = -\left(\frac{1}{x^2+y^2}\right)&\frac{1}{r^3}[(-x^2z - y^2z)(dx \wedge dy) \\ &- (x^2+y^2)(x\,dy \wedge dz + y\,dz \wedge dx)] \\ = \frac{1}{r^3}&(x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy), \end{aligned}$$

as advertised.

b. We first compute σ_U. To avoid drowning in superscripts, let $a := u^1$ and $b := u^2$. Then

$$x = \frac{2a}{1+\eta}, \quad y = \frac{2b}{1+\eta}, \quad \text{and} \quad z = \frac{\eta-1}{\eta+1},$$

with $\eta = a^2 + b^2$, so

$$\begin{aligned} dx &= \frac{2(1+\eta)-(2a)(2a)}{(1+\eta)^2}\,da - \frac{(2a)(2b)}{(1+\eta)^2}\,db \\ &= \frac{2}{(1+\eta)^2}[(1+b^2-a^2)\,da - 2ab\,db], \\ dy &= -\frac{(2b)(2a)}{(1+\eta)^2}\,da + \frac{2(1+\eta)-(2b)(2b)}{(1+\eta)^2}\,db \\ &= \frac{2}{(1+\eta)^2}[-2ab\,da + (1+a^2-b^2)\,db], \end{aligned}$$

and

$$\begin{aligned} dz &= \frac{2a(\eta+1)-(\eta-1)(2a)}{(1+\eta)^2}\,da + \frac{2b(\eta+1)-(\eta-1)(2b)}{(1+\eta)^2}\,db \\ &= \frac{4}{(1+\eta)^2}(a\,da + b\,db). \end{aligned}$$

Thus,

$$x\,dy \wedge dz = \frac{16a}{(1+\eta)^5}[(-2ab)b - (1+a^2-b^2)a]\,da \wedge db$$
$$= \frac{-16a^2}{(1+\eta)^4}\,da \wedge db.$$

Similarly, we find

$$y\,dz \wedge dx = \frac{-16b^2}{(1+\eta)^4}\,da \wedge db$$

and

$$z\,dx \wedge dy = \frac{4(\eta-1)}{(1+\eta)^5}[(1+(b^2-a^2))(1-(b^2-a^2)) - 4a^2b^2]\,da \wedge db$$
$$= \frac{4(\eta-1)}{(1+\eta)^5}[(1-(b^2-a^2)^2) - 4a^2b^2]\,da \wedge db$$
$$= \frac{4(\eta-1)}{(1+\eta)^5}(1-b^4+2a^2b^2-a^4-4a^2b^2)\,da \wedge db$$
$$= \frac{4(\eta-1)}{(1+\eta)^5}(1-\eta^2)\,da \wedge db.$$

Adding everything together we get

$$\frac{-16\eta(1+\eta)}{(1+\eta)^5} - \frac{4(1-\eta-\eta^2+\eta^3)}{(1+\eta)^5} = -\frac{4}{(1+\eta)^2}$$

times $da \wedge db$, as advertised.

The calculation for σ_V is similar. Evidently $\sigma_U = 4\omega_U$ and $\sigma_V = 4\omega_V$, where ω_U and ω_V were defined in Exercise 3.10.

3.41 Note that $f : \mathbb{R}^2 \to \mathbb{R}$, and Σ is $f^{-1}(0)$. The Jacobian matrix of f is $(-3x^2 - a \quad 2y)$, so the map is rank deficient if $y = 0$ and $3x^2 + a = 0$. But if $y = 0$ then $x^3 + ax + b = 0$, which means that x is a double root of $x^3 + ax + b$. This can only happen if the discriminant vanishes. Thus, as long as the discriminant is nonzero, Σ is an embedded submanifold of the plane.

3.42 f is constant on X so the differential f_* must vanish everywhere on X. Hence, $T_pX \subseteq \ker f_*$. Now, f_* maps T_pM surjectively onto T_pN, so by the rank/nullity theorem $\dim \ker f_* = m - n$. But by the regular value theorem, $\dim T_pX = \dim X = m - n$ as well, so $T_pX = \ker f_*$.

3.43 a. The determinant of a matrix is a polynomial in its entries, so it is a smooth function from $M_n(\mathbb{R})$ to $\mathbb{R}$. The inverse image of 0 is closed in $M_n(\mathbb{R})$, so its complement is open. Any open subset of Euclidean space is a smooth manifold (just take a single coordinate chart consisting of the open subset itself). A single entry of a product of two matrices is a sum of products of the entries of each individual matrix, which is clearly a smooth map. Finally, the inverse of a matrix is the ratio of two polynomials in the matrix entries, namely the adjugate and the determinant, which is also smooth whenever the determinant does not vanish.

b. An element of $M_n^+(\mathbb{R})$ is determined by $n(n+1)/2$ numbers: the n diagonal elements and the $n(n-1)/2$ elements above the diagonal. Thus, it is naturally a submanifold of $M_n(\mathbb{R})$ diffeomorphic to $\mathbb{R}^{n(n+1)/2}$. The map $\varphi : M_n(\mathbb{R}) \to M_n^+(\mathbb{R})$ is smooth (because each entry of the image is a quadratic polynomial in the matrix entries), and $O(n) = \varphi^{-1}(I)$, where I is the identity matrix. To show I is a regular value we must examine the derivative map φ_*, which maps the tangent space of $M_n(\mathbb{R})$, which we identify with $M_n(\mathbb{R})$ (because it is a vector space), to the tangent space to $M_n^+(\mathbb{R})$, which is identified with $M_n^+(\mathbb{R})$. Here is one way to do it. (For another, see [33], pp. 22-23.) We view φ as a map on $\mathbb{R}^{n_2}$ by considering its effect on individual matrix entries. Thus,

$$\varphi_{ij} := \varphi(A_{ij}) = A_{ik}A_{jk}.$$

So

$$\frac{\partial \varphi_{ij}}{\partial A_{mn}} = \delta_{im}\delta_{kn}A_{jk} + A_{ik}\delta_{jm}\delta_{kn}.$$

This is the Jacobian matrix. Viewing B as a "vector" in the tangent space to $M_n(\mathbb{R})$ we get

$$\frac{\partial \varphi_{ij}}{\partial A_{mn}} B_{mn} = B_{ik}A_{jk} + A_{ik}B_{jk},$$

or

$$\varphi_{*A}B = BA^T + AB^T.$$

To check that I is a regular value of φ we must check that, for every $A \in \varphi^{-1}(I) = O(n)$ and for any symmetric matrix C there exists a matrix B such that $\varphi_{*A}B = C$. Every symmetric matrix is of the form

$$C = \frac{1}{2}(C + C^T),$$

so it suffices to solve the equation $BA^T = \frac{1}{2}C$. But A is orthogonal, so multiplying both sides on the right by A gives $B = \frac{1}{2}CA$. Therefore, by Theorem 3.2, $O(n)$ is a smooth manifold. As it inherits the smooth group operations from $GL(n, \mathbb{R})$, it is a Lie group. Again by Theorem 3.2, the dimension of $O(n)$ is $n^2 - n(n+1)/2 = n(n-1)/2$.

c. The determinant is a smooth map from $O(n)$ to $\mathbb{R}$, and $SO(n)$ is the inverse image of $+1$. From the result of Exercise 1.36, we have

$$\frac{\partial}{\partial A_{ij}}(\det A) = (\det A)(A^{-1})_{ji} \tag{1}$$

whenever $\det A \neq 0$, which is the case for $A \in O(n)$. As this never vanishes, the derivative map is surjective everywhere, so by Theorem 3.2 $SO(n)$ is a smooth submanifold of $O(n)$. The other component of $O(n)$ is not a Lie group for the simple reason that it is not a group, because $\det AB = \det A \det B$, so if A and B both have determinant -1 then AB has determinant $+1$.

d. Start with a real matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We have

$$AA^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix},$$

so the condition $AA^T = I$ gives

$$\begin{aligned} 1 &= a^2+b^2 = c^2+d^2, \\ 0 &= ac+bd. \end{aligned}$$

From $\det A = 1$ we get

$$ad - bc = 1.$$

The second, third and fourth equations yield

$$\begin{pmatrix} c & d \\ d & -c \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} a \\ b \end{pmatrix} = -\begin{pmatrix} -c & -d \\ -d & c \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} d \\ -c \end{pmatrix},$$

so a general element of $SO(2)$ can be written

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

where $a^2+b^2 = 1$.

The circle S^1 is the locus of points $(x, y) \in \mathbb{R}^2$ where $x^2 + y^2 = 1$. The map $\varphi : \mathbb{R}^2 \to \mathbb{R}^4$ given by

$$(x, y) \mapsto \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

is obviously smooth, and it restricts to a bijection from S^1 to $SO(2)$. Hence, by the fact supplied in the hint, φ restricts to a diffeomorphism from S^1 to $SO(2)$.

3.44 a. L_g is a diffeomorphism, so by (3.87),

$$L_{g*}[X, Y] = [L_{g*}X, L_{g*}Y] = [X, Y].$$

b. If X is left invariant, then for all $g \in G$, $X_g = L_{g*}X_e$. Conversely, given Y we define a vector field X by $X_g = L_{g*}Y$. Then $L_{g*}X_h = L_{g*}L_{h*}Y = L_{gh*}Y = X_{gh}$, so X is left invariant.

c. Given a one parameter subgroup γ, its derivative $\gamma'(0)$ is an element of $T_eG = \mathfrak{g}$. Conversely, given a vector $X_e \in T_eG$, consider the left invariant vector field X it determines. Define γ to be the unique integral curve of X through e satisfying $\gamma'(0) = X_e$, so that $\gamma'(t) = X_{\gamma(t)} = L_{\gamma(t)*}X_e$ for $|t| < \epsilon$. Fix s with $|s| < \epsilon$ and $|t + s| < \epsilon$ and define two new curves: $\eta_1(t) := \gamma(s + t)$ and $\eta_2(t) := \gamma(s)\gamma(t)$. We have

$$\eta_1'(t) = \gamma'(s + t) = X_{s+t} = X_{\eta_1(t)}.$$

Now view $\eta_2(t)$ as a composition: $(L_{\gamma(s)} \circ \gamma)(t)$. Then as the pushforward behaves nicely with respect to composition (chain rule), we have

$$\begin{aligned}\eta_2'(t) &= \frac{d}{dt}\eta_2(t) = \eta_{2*}(d/dt) = L_{\gamma(s)*}\gamma_*(d/dt) = L_{\gamma(s)*}\gamma'(t)\\ &= L_{\gamma(s)*}X_{\gamma(t)} = X_{\gamma(s)\gamma(t)} = X_{\eta_2(t)}.\end{aligned}$$

In particular, η_1 and η_2 are both integral curves of X and they satisfy the same initial condition

$$\eta_1(0) = \gamma(s) = \eta_2(0),$$

so they must be equal. Finally, we can extend $\gamma(t)$ to all values of t using the group law, by defining $\gamma(t) := \gamma(t/n)^n$ for some large n.

3.45 (ii) By definition, for any matrices A and B, $L_A B = AB$ (matrix multiplication). Thus,

$$(L_A^* x)(B) = x(L_A B) = x(AB) = AB,$$

so

$$L_A^* x = Ax.$$

On the other hand,

$$(L_A^* x^{-1})(B) = x^{-1}(L_A B) = (AB)^{-1} = B^{-1}A^{-1},$$

and therefore

$$L_A^* x^{-1} = x^{-1} A^{-1}.$$

Hence,

$$\begin{aligned} L_A^* \Theta = L_A^*(x^{-1}\, dx) &= (L_A^* x^{-1})\, d(L_A^* x) \\ &= x^{-1} A^{-1} A\, dx = x^{-1}\, dx = \Theta. \end{aligned}$$

(iii) First note that left multiplication by g is a transitive action on any group G, meaning that, for any two points x and y in G, there exists a g such that $gx = y$. (Just take $g = yx^{-1}$.) Thus, it suffices to show that, for any A, $L_A^* \langle \Theta, X \rangle = \langle \Theta, X \rangle$, as that will show B_X has the same value at any point. But as pullback commutes with contraction,

$$L_A^* \langle \Theta, X \rangle = \left\langle L_A^* \Theta, L_A^* X \right\rangle = \langle \Theta, X \rangle .$$

It follows that we may as well evaluate $\langle \Theta, X \rangle$ at the identity. But then we get a map from an element of $T_e GL(n, \mathbb{R}) = g\ell(n, \mathbb{R})$ to a matrix in $M_n(\mathbb{R})$. It is obviously linear by the linearity of the dual pairing.

(iv) Applied to the current situation, (3.123) gives

$$d\Theta(X, Y) = X\Theta(Y) - Y\Theta(X) - \Theta([X, Y]).$$

But we already showed that if X and Y are left invariant, $\Theta(X)$ and $\Theta(Y)$ are constant functions, so the first two terms on the right must vanish.

(v) Differentiating both sides of

$$xx^{-1} = 1$$

gives

$$dx \cdot x^{-1} + x \cdot d(x^{-1}) = 0,$$

which shows that

$$d(x^{-1}) = -x^{-1}\, dx\, x^{-1}.$$

Thus,

$$d\Theta = d(x^{-1}\, dx) = -x^{-1}\, dx\, x^{-1} \wedge dx = -\Theta \wedge \Theta.$$

(vi) We have

$$\Theta([X, Y]) = B_{[X,Y]}$$

and

$$(\Theta \wedge \Theta)(X, Y) = \Theta(X)\Theta(Y) - \Theta(Y)\Theta(X) = [B_X, B_Y],$$

so we conclude that $X \to B_X$ is indeed a Lie algebra homomorphism.

3.46 $SO(n)$ is the set of all n by n real matrices satisfying $AA^T = I$, and by the previous exercise we may identify $\mathfrak{so}(n)$ with a subset of $M_n(\mathbb{R})$. Consider a curve $A : (a, b) \to SO(n)$ given by $t \mapsto A(t)$, with $A(0) = e$. Differentiating $I = AA^T$ with respect to t at $t = 0$ gives

$$0 = \left(\frac{dA}{dt} A^T + A \frac{dA^T}{dt} \right)\bigg|_0 = \frac{dA}{dt}\bigg|_0 + \frac{dA^T}{dt}\bigg|_0 = \frac{dA}{dt}\bigg|_0 + \left(\frac{dA}{dt}\bigg|_0 \right)^T,$$

because $A(0) = e$. Elements of $\mathfrak{so}(n)$ are tangent vectors to the curve at the identity and therefore may be identified with the derivatives dA/dt at $t = 0$. (We are using the fact that $M_n(\mathbb{R})$ is a vector space, so it coincides with its tangent space.) Therefore, $\mathfrak{so}(n)$ consists of skew symmetric matrices. Moreover, if A and B are skew symmetric then

$$\begin{aligned} [A, B]^T &= (AB - BA)^T = (AB)^T - (BA)^T = B^T A^T - A^T B^T \\ &= (-B)(-A) - (-A)(-B) = [B, A] = -[A, B], \end{aligned}$$

so $\mathfrak{so}(n)$ is indeed a subalgebra of $M_n(\mathbb{R})$.

3.47 $M_n(\mathbb{R})$ is a vector space, so the differential of a map just equals the map itself. Hence $L_{A*} = L_A$. The differential equation defining the integral curve therefore becomes

$$\frac{d\gamma(t)}{dt} = \gamma(t)A,$$

where $A = \gamma'(0)$. To find the solution to the differential equation, rewrite it as an integral equation,

$$\gamma(t) = I + \int_0^t \gamma(t)A\, dt.$$

where $\gamma(0) = I$. Now iterate by repeatedly substituting the left side into the right side to get

$$\gamma(t) = I + tA + \frac{1}{2!}(tA)^2 + \cdots = e^{tA}.$$

But $\operatorname{Exp} tA$ is also an integral curve whose tangent vector at the identity is A, so the conclusion follows from the uniqueness of integral curves.

3.48 Proceeding as in the solution of Exercise 3.46 we take a curve $A(t)$ on $SU(n)$ through the identity and differentiate $I = AA^\dagger$ to get

$$0 = \dot{A}(0) + (\dot{A}(0))^\dagger,$$

which shows that elements of $\mathfrak{su}(n)$ must be anti-Hermitian. (The overdot means time derivative.) In the real case the condition that $\dot{A}(0)$ be traceless is automatic from the requirement that it be skew symmetric, but in the complex case we need to work a little harder. By virtue of Exercise 3.47 every element of $SU(n)$ can be written in the form $e^{\dot{A}(0)}$ for some element $\dot{A}(0)$ in $\mathfrak{su}(n)$. But by Exercise 1.48 $\det e^{\dot{A}(0)} = e^{\operatorname{tr} \dot{A}(0)}$, so the condition that $e^{\dot{A}(0)}$ have determinant one is exactly the condition that $\dot{A}(0)$ be traceless.

3.49 a. $\varphi(g)$ is a smooth bijection with smooth inverse $\varphi(g)^{-1} = \varphi(g^{-1})$. Therefore it is a diffeomorphism, and its derivative is a local isomorphism. Let X and Y be left invariant vector fields on G. By (3.87),

$$\varphi(g)_{*,e}[X_e, Y_e] = \varphi(g)_{*,e}[X, Y]_e = [\varphi(g)_{*,e}X, \varphi(g)_{*,e}Y]_e,$$

so $\operatorname{Ad} g$ is a Lie algebra automorphism.

b. In terms of matrices, the conjugation map is

$$\varphi(A)B = ABA^{-1}.$$

This is linear in B, so its derivative equals itself.

3.50 We write

$$\begin{aligned}\frac{d}{dt}\operatorname{Ad} e^{tX}(Y) &= \frac{d}{dt}\left(e^{tX}Ye^{-tX}\right)\\ &= \left(\frac{d}{dt}e^{tX}\right)Ye^{-tX} + e^{tX}Y\left(\frac{d}{dt}e^{-tX}\right)\\ &= e^{tX}XYe^{-tX} - e^{tX}YXe^{-tX}\\ &= e^{tX}[X, Y]e^{-tX}\\ &= \operatorname{Ad} e^{tX}[X, Y].\end{aligned}$$

But also (as $\operatorname{ad} X$ commutes with itself)

$$\frac{d}{dt}e^{t \operatorname{ad} X}(Y) = e^{t \operatorname{ad} X}(\operatorname{ad} X)(Y) = e^{t \operatorname{ad} X}[X, Y].$$

Hence both $\operatorname{Ad} e^{tX}$ and $e^{t \operatorname{ad} X}$ satisfy the same differential equation with the same initial condition (both reduce to the identity at $t = 0$), so by uniqueness they must coincide.

3.51 Assuming (3.114), we get, by setting $t = 1$,

$$e^{A+B} = e^A e^B e^{-(1/2)[A,B]},$$

from which (3.111) follows, because $[A, B]$ commutes with A and B.

Differentiating both sides of (3.111) with respect to t gives

$$(A + B)e^{t(A+B)} = Ae^{tA}f(t) + e^{tA}f'(t),$$

so using (3.113) we get

$$(A + B)e^{tA}f(t) = Ae^{tA}f(t) + e^{tA}f'(t),$$

or, simplifying,

$$Be^{tA}f(t) = e^{tA}f'(t).$$

Multiplying both sides on the left by e^{-tA} yields

$$e^{-tA}Be^{tA}f(t) = f'(t),$$

so applying (3.110) we obtain

$$f'(t) = (\operatorname{Ad} e^{-tA}B)f(t) = (e^{-t\,\operatorname{ad} A}B)f(t).$$

Using the fact that $[A, [A, B]] = 0$ we thus get

$$f'(t) = (B - t[A, B])f(t),$$

together with the initial condition $f(0) = 1$.

On the other hand, mindful of the fact that $[B, [A, B]] = 0$, we see that

$$\begin{aligned}\frac{d}{dt}e^{tB}e^{-(t^2/2)[A,B]} &= Be^{tB}e^{-(t^2/2)[A,B]} - e^{tB}t[A, B]e^{-(t^2/2)[A,B]}\\ &= (B - t[A, B])e^{tB}e^{-(t^2/2)[A,B]}.\end{aligned}$$

As $e^{tB}e^{-(t^2/2)[A,B]}|_{t=0} = 1$, the proof is complete.

3.52 a. $\operatorname{ad} X$ is a linear map, so this follows from the cyclicity of the trace.

b. By definition, an automorphism T of $\mathfrak{g}$ must carry brackets to brackets:

$$T[X, Y] = [TX, TY].$$

Writing $Z = TY$ this becomes

$$T[X, T^{-1}Z] = [TX, Z]$$

which just says that $\operatorname{ad} TX = T \circ \operatorname{ad} X \circ T^{-1}$. Thus, by cyclicity of the trace,

$$\begin{aligned}(TX, TY) = \operatorname{tr}(\operatorname{ad} TX \circ \operatorname{ad} TY) &= \operatorname{tr}(T \circ \operatorname{ad} X \circ \operatorname{ad} Y \circ T^{-1})\\ &= \operatorname{tr}(\operatorname{ad} X \circ \operatorname{ad} Y) = (X, Y).\end{aligned}$$

But according to Exercise 3.49a $\operatorname{Ad} g$ is an automorphism of $\mathfrak{g}$ for any $g \in G$.

c. By the Jacobi identity,

$$\begin{aligned}\operatorname{ad}[X, Y](Z) = [[X, Y], Z] &= [X, [Y, Z]] + [Y, [Z, X]]\\ &= \operatorname{ad} X \circ \operatorname{ad} Y(Z) - \operatorname{ad} Y \circ \operatorname{ad} X(Z)\\ &= [\operatorname{ad} X, \operatorname{ad} Y](Z),\end{aligned}$$

so by the cyclicity of the trace again,

$$\begin{aligned}(\operatorname{ad} Z(X), Y) &= ([Z, X], Y) = \operatorname{tr}(\operatorname{ad}[Z, X] \circ \operatorname{ad} Y)\\ &= \operatorname{tr}([\operatorname{ad} Z, \operatorname{ad} X] \circ \operatorname{ad} Y)\\ &= \operatorname{tr}(\operatorname{ad} Z \circ \operatorname{ad} X \circ \operatorname{ad} Y) - \operatorname{tr}(\operatorname{ad} X \circ \operatorname{ad} Z \circ \operatorname{ad} Y)\\ &= \operatorname{tr}(\operatorname{ad} Y \circ \operatorname{ad} Z \circ \operatorname{ad} X) - \operatorname{tr}(\operatorname{ad} Z \circ \operatorname{ad} Y \circ \operatorname{ad} X)\\ &= -(\operatorname{ad} Z(Y), X).\end{aligned}$$

3.53 a. We have

$$E_1E_2 = \begin{pmatrix} 0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 0 & -1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}$$

and

$$E_2E_1 = \begin{pmatrix} 0 & 0 & -1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 1 & 0\end{pmatrix},$$

so

$$E_1E_2 - E_2E_1 = \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 1\\ 0 & -1 & 0\end{pmatrix} = E_3.$$

Similarly we find

$$[E_2, E_3] = E_1 \qquad \text{and} \qquad [E_3, E_1] = E_2,$$

from which the result follows.

b. The jk^{th} element of the matrix representing $\operatorname{ad} E_i$ in the given basis is $-\epsilon_{ijk}$, because (using the summation convention),

$$(\operatorname{ad} E_i)(E_j) = (\operatorname{ad} E_i)_{kj} E_k = \epsilon_{ijk} E_k.$$

Therefore,

$$(E_i, E_j) = \operatorname{tr}(\operatorname{ad} E_i \circ \operatorname{ad} E_j) = (-\epsilon_{imn})(-\epsilon_{jnm}) = -2\delta_{ij}.$$

The metric is negative definite, so it is nondegenerate and its signature is $(-, -, -)$.

3.54 Elements of $\mathfrak{su}(2)$ consist of 2×2 complex traceless skew-Hermitian matrices with determinant one. But A is anti-Hermitian if and only if iA is Hermitian, so we may as well find all 2×2 traceless Hermitian matrices with determinant one. By Hermiticity

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix} = \begin{pmatrix} \bar a & \bar c\\ \bar b & \bar d\end{pmatrix},$$

which gives

$$a = \bar a, \quad d = \bar d, \quad \text{and} \quad b = \bar c.$$

In particular, a and d are real, and b and c are complex conjugates of one another. The condition that the matrix be traceless gives $a+d=0$. Therefore the most general element of $\mathfrak{su}(2)$ is of the form

$$\begin{pmatrix} x & y-iz \\ y+iz & -x \end{pmatrix} = x\sigma_1 + y\sigma_2 + z\sigma_3 = \boldsymbol{x}\cdot\boldsymbol{\sigma},$$

where x, y, and z are real. It follows that the matrices $\tau_k := (-i/2)\sigma_k$ form a basis for $\mathfrak{su}(2)$.

To show that $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ are isomorphic it suffices to show that they have the same structure constants. The commutation relations for the Pauli matrices are well known and easy to prove by multiplying out the matrices. Using the summation convention again we get

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k.$$

Hence,

$$[\tau_i, \tau_j] = \epsilon_{ijk}\tau_k,$$

and the claim is proved.

3.55 a. We want to show that

$$\mathcal{L}_X = i_X d + d i_X \tag{1}$$

when acting on p forms. The result holds for zero forms, because

$$(i_X d + d i_X) f = i_X df = Xf = \mathcal{L}_X f.$$

Now suppose it holds for $p-1$ forms, and let $\eta\wedge\lambda$ be a p-form, with $\deg \eta = r$. Then by the induction hypothesis and the properties of the Lie derivative,

$$\begin{aligned}\mathcal{L}_X(\eta\wedge\lambda) &= \mathcal{L}_X\eta\wedge\lambda + \eta\wedge\mathcal{L}_X\lambda \\ &= (i_X d + d i_X)\eta\wedge\lambda + \eta\wedge(i_X d + d i_X)\lambda.\end{aligned}$$

But we also have, by the properties of the differential and the interior product,

$$\begin{aligned}(i_X d + d i_X)(\eta\wedge\lambda) &= (i_X d)(\eta\wedge\lambda) + (d i_X)(\eta\wedge\lambda) \\ &= i_X(d\eta\wedge\lambda + (-1)^r\eta\wedge d\lambda) \\ &\qquad + d(i_X\eta\wedge\lambda + (-1)^r\eta\wedge i_X\lambda) \\ &= i_X d\eta\wedge\lambda + (-1)^{r+1} d\eta\wedge i_X\lambda \\ &\qquad + (-1)^r[i_X\eta\wedge d\lambda + (-1)^r\eta\wedge i_X d\lambda] \\ &\qquad + d i_X\eta\wedge\lambda + (-1)^{r-1} i_X\eta\wedge d\lambda \\ &\qquad + (-1)^r[d\eta\wedge i_X\lambda + (-1)^r\eta\wedge d i_X\lambda] \\ &= (i_X d + d i_X)\eta\wedge\lambda + \eta\wedge(i_X d + d i_X)\lambda.\end{aligned}$$

This proves (1).

b. We want to show that, for ω a p-form,

$$\mathcal{L}_X i_Y \omega - i_Y \mathcal{L}_X \omega = i_{\mathcal{L}_X Y} \omega. \tag{2}$$

Let $p = 1$. By property (iii) of the Lie derivative,

$$\mathcal{L}_X \langle Y, \omega \rangle - \langle Y, \mathcal{L}_X \omega \rangle = \langle \mathcal{L}_X Y, \omega \rangle .$$

But this is just

$$\mathcal{L}_X i_Y \omega - i_Y \mathcal{L}_X \omega = i_{\mathcal{L}_X Y} \omega,$$

so (2) holds for $p = 1$. Now assume (2) holds for $p-1$, and let $\omega = \eta \wedge \lambda$ with deg $\eta = r$. Then, using the induction hypothesis we get

$$\begin{aligned}
\mathcal{L}_X i_Y (\eta \wedge \lambda) - i_Y \mathcal{L}_X (\eta \wedge \lambda) &= \mathcal{L}_X [i_Y \eta \wedge \lambda + (-1)^r \eta \wedge i_Y \lambda] \\
&\quad - i_Y [\mathcal{L}_X \eta \wedge \lambda + \eta \wedge \mathcal{L}_X \lambda] \\
&= \mathcal{L}_X i_Y \eta \wedge \lambda + \cancel{i_Y \eta \wedge \mathcal{L}_X \lambda} \\
&\quad + \cancel{(-1)^r \mathcal{L}_X \eta \wedge i_Y \lambda} + (-1)^r \eta \wedge \mathcal{L}_X i_Y \lambda \\
&\quad - i_Y \mathcal{L}_X \eta \wedge \lambda - \cancel{(-1)^r \mathcal{L}_X \eta \wedge i_Y \lambda} \\
&\quad \cancel{- i_Y \eta \wedge \mathcal{L}_X \lambda} - (-1)^r \eta \wedge i_Y \mathcal{L}_X \lambda \\
&= i_{\mathcal{L}_X Y} \eta \wedge \lambda + (-1)^r \eta \wedge i_{\mathcal{L}_X Y} \lambda \\
&= i_{\mathcal{L}_X Y} (\eta \wedge \lambda).
\end{aligned}$$

c. We want to show

$$\begin{aligned}
(\mathcal{L}_{X_0} \omega)(X_1, \ldots, X_p) &= \mathcal{L}_{X_0} (\omega(X_1, \ldots, X_p)) \\
&\quad - \sum_{i=1}^{p} \omega(X_1, \ldots, \mathcal{L}_{X_0} X_i, \ldots, X_p).
\end{aligned} \tag{3}$$

For $p = 1$, this is

$$(\mathcal{L}_{X_0} \omega)(X_1) = \mathcal{L}_{X_0} (\omega(X_1)) - \omega(\mathcal{L}_{X_0} X_1),$$

or

$$i_{X_1} \mathcal{L}_{X_0} \omega = \mathcal{L}_{X_0} i_{X_1} \omega - i_{\mathcal{L}_{X_0} X_1} \omega,$$

which is just the case $p = 1$ of the formula proved in Part (b). Now assume (3) holds for $p - 1$. Then, by induction and the result of Part (b) again,

$$\begin{aligned}
\mathcal{L}_{X_0} (\omega(X_1, \ldots, X_p)) &= \mathcal{L}_{X_0} ((i_{X_1} \omega)(X_2, \ldots, X_p)) \\
&= (\mathcal{L}_{X_0} (i_{X_1} \omega))(X_2, \ldots, X_p) \\
&\quad + \sum_{i=2}^{p} (i_{X_1} \omega)(X_2, \ldots, \mathcal{L}_{X_0} X_i, \ldots, X_p).
\end{aligned}$$

$$
\begin{aligned}
&= (i_{X_1}\mathcal{L}_{X_0}\omega + i_{\mathcal{L}_{X_0}X_1}\omega)(X_2, \ldots, X_p) \\
&\qquad + \sum_{i=2}^{p} \omega(X_1, X_2, \ldots, \mathcal{L}_{X_0}X_i, \ldots, X_p). \\
&= (\mathcal{L}_{X_0}\omega)(X_1, \ldots, X_p) \\
&\qquad + \sum_{i=1}^{p} \omega(X_1, X_2, \ldots, \mathcal{L}_{X_0}X_i, \ldots, X_p).
\end{aligned}
$$

d. By the inductive hypothesis and the antisymmetry of the contraction mapping,

$$
\begin{aligned}
&(di_{X_0}\omega)(X_1, \ldots, X_p) \\
&= \sum_{i=1}^{p}(-1)^{i-1}\mathcal{L}_{X_i}((i_{X_0}\omega)(X_1, \ldots, \widehat{X}_i, \ldots, X_p)) \\
&\qquad + \sum_{1\le i<j\le p} (-1)^{i+j}(i_{X_0}\omega)(\mathcal{L}_{X_i}X_j, X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_p) \\
&= \sum_{i=1}^{p}(-1)^{i-1}\mathcal{L}_{X_i}(\omega(X_0, X_1, \ldots, \widehat{X}_i, \ldots, X_p)) \\
&\qquad + \sum_{1\le i<j\le p} (-1)^{i+j}\omega(X_0, \mathcal{L}_{X_i}X_j, X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_p) \\
&= \sum_{i=1}^{p}(-1)^{i-1}\mathcal{L}_{X_i}(\omega(X_0, \ldots, \widehat{X}_i, \ldots, X_p)) \\
&\qquad + \sum_{1\le i<j\le p} (-1)^{i+j+1}\omega(\mathcal{L}_{X_i}X_j, X_0, X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_p).
\end{aligned}
\tag{4}
$$

Equations (1) and (3) yield

$$
\begin{aligned}
&(i_{X_0}d\omega + di_{X_0}\omega)(X_1, \ldots, X_p) \\
&\qquad = \mathcal{L}_{X_0}(\omega(X_1, \ldots, X_p)) \\
&\qquad\qquad - \sum_{i=1}^{p} \omega(X_1, \ldots, \mathcal{L}_{X_0}X_i, \ldots, X_p) \\
&\qquad = \mathcal{L}_{X_0}(\omega(X_1, \ldots, X_p)) \\
&\qquad\qquad - \sum_{i=0, 1\le j\le p}^{p} (-1)^{i+j-1}\omega(\mathcal{L}_{X_0}X_j, X_1, \ldots, \widehat{X_j}, \ldots, X_p),
\end{aligned}
\tag{5}
$$

where the last equality follows by changing dummy indices and by the antisymmetry of the contraction map. Combining (4) and (5) we get

$$\begin{aligned}
&d\omega(X_0, X_1, \dots, X_p) \\
&\quad = \mathcal{L}_{X_0}(\omega(X_1, \dots, X_p)) \\
&\qquad - \sum_{i=0, 1\le j\le p} (-1)^{i+j-1}\omega(\mathcal{L}_{X_0}X_j, X_1, \dots, \widehat{X_j}, \dots, X_p) \\
&\qquad - \sum_{i=1}^{p}(-1)^{i-1}\mathcal{L}_{X_i}(\omega(X_0, \dots, \widehat{X}_i, \dots, X_p)) \\
&\qquad - \sum_{1\le i<j\le p} (-1)^{i+j+1}\omega(\mathcal{L}_{X_i}X_j, X_0, X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_p) \\
&\quad = \sum_{i=0}^{p}(-1)^{i}\mathcal{L}_{X_i}(\omega(X_0, \dots, \widehat{X}_i, \dots, X_p)) \\
&\qquad + \sum_{0\le i<j\le p} (-1)^{i+j}\omega(\mathcal{L}_{X_i}X_j, X_0, X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_p).
\end{aligned}$$

The second formula in the problem follows because for any function f and vector fields X and Y, $\mathcal{L}_X f = Xf$ and $\mathcal{L}_X Y = [X, Y]$.

3.56 By compatibility of the Lie derivative with the dual pairing,

$$X\,\langle\alpha, \partial_i\rangle = \langle\mathcal{L}_X\alpha, \partial_i\rangle + \langle\alpha, \mathcal{L}_X\partial_i\rangle\,. \tag{1}$$

From (3.125) we get

$$\mathcal{L}_X\partial_i = -X^k{}_{,i}\,\partial_k. \tag{2}$$

Combining this with (1) gives

$$X^j\alpha_{i,j} = (\mathcal{L}_X\alpha)_i - \alpha_j X^j{}_{,i},$$

which is (3.126). In particular, if $\alpha = dx^j$ we get

$$\mathcal{L}_X dx^j = X^j{}_{,k}\,dx^k. \tag{3}$$

Now we have

$$T = T^{i_1\dots i_r}{}_{j_1\dots j_s}\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s},$$

so applying the Leibniz rule and using (2) and (3) yields

$$\begin{aligned}
\mathcal{L}_X T = {}& X(T^{i_1\dots i_r}{}_{j_1\dots j_s})\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s} \\
&+ T^{i_1\dots i_r}{}_{j_1\dots j_s}(-X^k{}_{,i_1})\partial_k\otimes\partial_{i_2}\otimes\cdots\otimes\partial_{i_r}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s} \\
&+ \cdots \\
&+ T^{i_1\dots i_r}{}_{j_1\dots j_s}(-X^k{}_{,i_r})\partial_{i_1}\otimes\cdots\otimes\partial_{i_{r-1}}\otimes\partial_k\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s} \\
&+ T^{i_1\dots i_r}{}_{j_1\dots j_s}(+X^{j_1}{}_{,k})\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes dx^k\otimes dx^{j_2}\otimes\cdots\otimes dx^{j_s} \\
&+ \cdots \\
&+ T^{i_1\dots i_r}{}_{j_1\dots j_s}(+X^{j_s}{}_{,k})\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes dx^{i_1}\otimes\cdots\otimes dx^{j_{s-1}}\otimes dx^k.
\end{aligned}$$

Upon changing dummy indices and stripping off

$$\partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$$

from each term we arrive at (3.124).

3.57 a. We have

$$\omega = dq^1 \wedge dp_1 + \cdots + dq^n \wedge dp_n,$$

so

$$\begin{aligned}\omega^n &= (dq^1 \wedge dp_1 + \cdots + dq^n \wedge dp_n)^n \\ &= (dq^1 \wedge dp_1) \wedge \cdots \wedge (dq^n \wedge dp_n) \\ &= (-1)^{n(n-1)/2} dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n,\end{aligned}$$

where the sign is obtained by permuting all the dq^i's to the left in succession, starting with dq^2. This gives $\sum_{i=1}^{n-1} i = n(n-1)/2$ sign flips. Thus, ω^n is manifestly nonvanishing. As ω^n is expressible in the same way in every coordinate patch, it extends globally to all of M.

b. As the pullback map is a ring homomorphism, $f^*\nu^n = \mu^n$. In particular, the Jacobian determinant of f is unity, which means that the differential f_* is an isomorphism of tangent spaces. Hence, by the inverse function theorem, f is a local diffeomorphism.

c. Let $Y = a^i(\partial/\partial q^i) + b_i(\partial/\partial p_i)$ be an arbitrary vector field in local coordinates. Then by definition,

$$\omega(X_f, Y) = i_Y i_{X_f} \omega = i_Y df = df(Y) = Yf = a^i \frac{\partial f}{\partial q^i} + b_i \frac{\partial f}{\partial p_i}.$$

But also

$$\begin{aligned}\omega(X_f, Y) &= (dq^i \wedge dp_i)(X_f, Y) \\ &= dq^i(X_f)dp_i(Y) - dq^i(Y)dp_i(X_f) \\ &= Y(p_i)X_f(q^i) - Y(q^i)X_f(p_i) \\ &= b_i X_f(q^i) - a^i X_f(p_i).\end{aligned}$$

As Y was arbitrary,

$$X_f(p_i) = -\frac{\partial f}{\partial q^i} \quad \text{and} \quad X_f(q^i) = \frac{\partial f}{\partial p_i},$$

whence the conclusion follows.

d. From Part (c),

$$X_H(q^i) = \frac{\partial H}{\partial p_i} \qquad \text{and} \qquad X_H(p_i) = -\frac{\partial H}{\partial q^i}.$$

But the tangent vector to the integral curve γ can be written

$$\dot{\gamma} = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} = X_H = X_H(q^i)\frac{\partial}{\partial q^i} + X_H(p_i)\frac{\partial}{\partial p_i},$$

whence Hamilton's equations follow.

e. By definition,

$$\begin{aligned} \left.\frac{dH}{dt}\right|_{\gamma(t)} &= \frac{d}{dt}\gamma^* H = \gamma_*(d/dt)H = X_H H \\ &= \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial H}{\partial p_i} = 0. \end{aligned}$$

In old fashioned notation,

$$\frac{dH}{dt} = \frac{\partial H}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial H}{\partial p_i}\frac{dp_i}{dt},$$

which gives the same thing.

f. Following the hint, we have

$$\mathcal{L}_{X_H}\omega = i_{X_H} d\omega + d i_{X_H}\omega = 0,$$

because ω is closed, and $i_{X_H}\omega = dH$. From the definition of the Lie derivative, it follows that

$$\frac{d}{dt}\varphi_t^*\omega = 0,$$

which shows that $\varphi_t^*\omega$ is a constant, independent of t. But $\varphi_0 = id$, whence we obtain $\varphi_t^*\omega = \omega$.

g. Using Part (c) we get

$$\omega(X_f, X_g) = df(X_g) = X_g(f) = \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q^i} - \frac{\partial g}{\partial q^i}\frac{\partial f}{\partial p_i}.$$

h. From (3.122) we have

$$\begin{aligned} &d\omega(X_f, X_g, X_h) \\ &\quad = \mathcal{L}_{X_f}(\omega(X_g, X_h)) - \mathcal{L}_{X_g}(\omega(X_f, X_h)) + \mathcal{L}_{X_h}(\omega(X_f, X_g)) \\ &\qquad - \omega([X_f, X_g], X_h) - \omega([X_g, X_h], X_f) + \omega([X_f, X_h], X_g). \quad (1) \end{aligned}$$

The symplectic form ω is closed, so the left side vanishes. We also have

$$\{f, g\} = \omega(X_f, X_g) = -\omega(X_g, X_f) = -dg(X_f) = -X_f(g) = -\mathcal{L}_{X_f} g,$$

so the second line in (1) is

$$-\{f, \{g, h\}\} + \{g, \{f, h\}\} - \{h, \{f, g\}\}. \quad (2)$$

Cartan's second formula says that

$$[\mathcal{L}_X, i_Y]\omega = i_{[X,Y]}\omega.$$

Thus, using Cartan's first formula, the fact that ω is closed, and the definition of X_f we get

$$\begin{aligned}
\omega([X_f, X_g], X_h) &= i_{X_h} i_{[X_f, X_g]}\omega \\
&= (i_{X_h}\mathcal{L}_{X_f} i_{X_g} - i_{X_h} i_{X_g}\mathcal{L}_{X_f})\omega \\
&= (i_{X_h}(i_{X_f} d + d i_{X_f}) i_{X_g} - i_{X_h} i_{X_g}(i_{X_f} d + d i_{X_f})\omega \\
&= i_{X_h} i_{X_f} d(dg) + i_{X_h} d\{g, f\} - i_{X_h} i_{X_g} d(df) \\
&= d\{g, f\}(X_h) \\
&= -\{h, \{g, f\}\}.
\end{aligned}$$

It follows that the third line in (1) is

$$\{h, \{g, f\}\} + \{f, \{h, g\}\} - \{g, \{h, f\}\}. \tag{3}$$

Adding (2) and (3) gives

$$0 = -2(\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\}).$$

4

Homotopy and de Rham cohomology

4.1 This is just a matter of unpacking the definitions. If X is contractible then there exist maps $f : X \to \{p\}$ and $g : \{p\} \to X$ satisfying $f \circ g = id_p$ and $g \circ f \sim id_X$. This, in turn, means that there is a homotopy $F : I \times X \to X$ such that $F(0, x) = g(f(x)) = x_0$ and $F(1, x) = id_X(x) = x$ for some $x_0 \in X$ and all $x \in X$. But this just says that id_X is null homotopic, because the existence of F shows that id_X is homotopic to the one point map $g \circ f : X \to x_0$.

Conversely, if id_X is null homotopic, then there exists a continuous map $F : I \times X \to X$ such that $F(0, x) = x_0$ and $F(1, x) = id_X(x)$. Pick a point $p \in X$ and define maps $f : X \to \{p\}$ and $g : \{p\} \to X$ by $g(p) = x_0$. (Note that both maps are automatically continuous, because the inverse image of every closed set is closed.) Then $f \circ g = id_p$, $F(0, x) = x_0 = g(f(x))$, and $F(1, x) = id_X$, so X is contractible.

4.2 We show that id_{CX} is homotopic to the one point map $f : CX \to \{p\}$ sending every point to the apex of the cone. Thus, consider the continuous map $F : I \times CX \to CX$ given by $F(t, (s, x)) = ((1-s)t+s, x)$. Then $F(0, (s, x)) = (s, x) = id_{CS}(s, x)$ and $F(1, (s, x)) = (1, x) = p = f(s, x)$.

4.3 a. We have

$$\begin{aligned}
F^*\omega &= A(tx, ty, tz)\, d(ty) \wedge d(tz) + B(tx, ty, tz)\, d(tz) \wedge d(tx) \\
&\qquad + C(tx, ty, tz)\, d(tx) \wedge d(ty) \\
&= A(tx, ty, tz)(y\, dt + t\, dy) \wedge (z\, dt + t\, dz) \\
&\quad + B(tx, ty, tz)(z\, dt + t\, dz) \wedge (x\, dt + t\, dx) \\
&\qquad + C(tx, ty, tz)(x\, dt + t\, dx) \wedge (y\, dt + t\, dy) \\
&= A(tx, ty, tz)(ty\, dt \wedge dz - tz\, dt \wedge dy) \\
&\quad + B(tx, ty, tz)(tz\, dt \wedge dx - tx\, dt \wedge dz) \\
&\qquad + C(tx, ty, tz)(tx\, dt \wedge dy - ty\, dt \wedge dx) \\
&\qquad\quad + \text{terms without } dt.
\end{aligned}$$

Hence,

$$\begin{aligned}
hF^*\omega = &\left(\int_0^1 A(tx, ty, tz)\, t\, dt\right)(y\, dz - z\, dy) \\
&+ \left(\int_0^1 B(tx, ty, tz)\, t\, dt\right)(z\, dx - x\, dz) \\
&+ \left(\int_0^1 C(tx, ty, tz)\, t\, dt\right)(x\, dy - y\, dx).
\end{aligned}$$

b. For ease of writing, define

$$e(x, y, z) := \int_0^1 A(tx, ty, tz)\, t\, dt \tag{1}$$

$$f(x, y, z) := \int_0^1 B(tx, ty, tz)\, t\, dt \tag{2}$$

$$g(x, y, z) := \int_0^1 C(tx, ty, tz)\, t\, dt, \tag{3}$$

so that

$$\alpha := hF^*\omega = e(y\, dz - z\, dy) + f(z\, dx - x\, dz) + g(x\, dy - y\, dx).$$

Then

$$\begin{aligned}
d\alpha &= \left(-\frac{\partial(ez)}{\partial x} - \frac{\partial(fz)}{\partial y} + \frac{\partial(gx)}{\partial x} + \frac{\partial(gy)}{\partial y}\right) dx \wedge dy + \text{cyclic} \\
&= \left(-\frac{\partial e}{\partial x}z - \frac{\partial f}{\partial y}z + \frac{\partial g}{\partial x}x + g + \frac{\partial g}{\partial y}y + g\right) dx \wedge dy + \text{cyclic}. \qquad (4)
\end{aligned}$$

Now add and subtract $(\partial g/\partial z)z$ inside the parenthetical term in Equation (4). Observe that

$$\frac{\partial e}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial g}{\partial z} = 0. \tag{5}$$

This follows from

$$\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} = 0,$$

which holds because ω is closed. Also,

$$\begin{aligned}
\frac{\partial g}{\partial x}x + \frac{\partial g}{\partial y}y + \frac{\partial g}{\partial z}z &= (\boldsymbol{x} \cdot \nabla)g \\
&= \int_0^1 (\boldsymbol{x} \cdot \nabla)C(tx, ty, tz)\, t\, dt \qquad (6)
\end{aligned}$$

$$= \int_0^1 \frac{d}{dt} C(tx, ty, tz)\, t^2\, dt \tag{7}$$

$$= \Big[C(tx, ty, tz)\, t^2 \Big]_0^1 - 2\int_0^1 C(tx, ty, tz)\, t\, dt \tag{8}$$

$$= C(x, y, z) - 2g. \tag{9}$$

Equation (7) holds by virtue of the chain rule. The easiest way to see this is to write

$$\nabla = t\left(\frac{\partial}{\partial(tx)}, \frac{\partial}{\partial(ty)}, \frac{\partial}{\partial(tz)}\right)$$

in Equation (6). Equation (8) follows by integration by parts, and Equation (9) from definition (3). Combining (4), (5), and (9) gives

$$d\alpha = C\, dx \wedge dy + \text{cyclic} = \omega,$$

as was to be shown.

4.4 Following the hint, we consider the homology class $[a] \in H_i(A)$. As φ is a chain map and $a \in Z(A^i)$ is closed,

$$d_i \varphi_i a = \varphi_{i+1} d_i a = 0,$$

so $\varphi_i(a) \in Z(B^i)$. Hence it makes sense to consider its cohomology class. So define $h_i([a]) := [\varphi_i(a)]$. We need only show that the map is independent of class representative. So, suppose $a' \in [a]$. Then $a - a' = d_{i-1}\gamma$ for some $\gamma \in A^{i-1}$. But φ is a chain map, so

$$\varphi_i(a - a') = \varphi_i d_{i-1}\gamma = d_{i-1}\varphi_{i-1}\gamma.$$

This means that $[\varphi_i(a - a')]$ is zero in $H_i(B)$, which, by linearity, means that $[\varphi_i(a)] = [\varphi_i(a')]$.

4.5 Begin again with $[c] \in H^i(C)$. Choose an element $c' \in [c]$ different from c and follow the same steps as before. This gives a $b' \in B^i$ satisfying $\psi_i b' = c'$ and an $a' \in A^{i+1}$ such that $\varphi_{i+1} a' = d_i b'$. We must show that $[a] = [a']$. As $[c] = [c']$, $c - c' = d_{i-1} w$ for some $w \in C^{i-1}$. By exactness, there is a $v \in B^{i-1}$ such that $\psi_{i-1} v = w$. Thus $c - c' = d_{i-1}\psi_{i-1} v = \psi_i d_{i-1} v$. It follows that $\psi_i(b - b') = c - c' = \psi_i d_{i-1} v$ or $\psi_i(b - b' - d_{i-1} v) = 0$. By exactness, there exists a $u \in A^i$ such that $b - b' - d_{i-1} v = \varphi_i u$, so $\varphi_{i+1}(a - a') = d_i(b - b') = d_i \varphi_i u = \varphi_{i+1} d_i u$, or $\varphi_{i+1}(a - a' - d_i u) = 0$. By exactness, we conclude $a - a' = d_i u$, so $[a] = [a']$.

4.6 We show that the sequence is exact at $H_i(A)$, $H_i(B)$, and $H_i(C)$, in that order.

1. ($\operatorname{im} \delta_{i-1} \subseteq \ker \alpha_i$.) Using the same notation as in the text, we have

$$\alpha_i \delta_{i-1}[c] = \alpha_i[a] = [\varphi_i a] = [d_{i-1} b] = 0.$$

($\ker \alpha_i \subseteq \operatorname{im} \delta_{i-1}$.) If $\alpha_i[a] = [\varphi_i a] = 0$ then $\varphi_i a = d_{i-1} b$ for some $b \in B^{i-1}$. Define $c := \psi_{i-1} b$. Then $\delta_{i-1}[c] = [a]$.

2. ($\operatorname{im} \alpha_i \subseteq \ker \beta_i$.) By exactness,

$$\beta_i \alpha_i [a] = \beta_i [\varphi_i a] = [\psi_i \varphi_i a] = 0.$$

($\ker \beta_i \subseteq \operatorname{im} \alpha_i$.) If $\beta_i[b] = [\psi_i b] = 0$ then $\psi_i b = d_{i-1} c$ for some $c \in C^{i-1}$. By exactness, there is a $b' \in B^{i-1}$ such that $\psi_{i-1} b' = c$. But then $\psi_i b = d_{i-1} c = d_{i-1} \psi_{i-1} b' = \psi_i d_{i-1} b'$, so $\psi_i (b - d_{i-1} b') = 0$. By exactness, there exists an $a \in A^i$ such that $\varphi_i a = b - d_{i-1} b'$. But then $\alpha_i[a] = [\varphi_i a] = [b]$.

3. ($\operatorname{im} \beta_i \subseteq \ker \delta_i$.)

$$\delta_i \beta_i [b] = \delta_i [\psi_i b] = [a],$$

where $\varphi_{i+1} a = d_i b'$ for any b' with $\psi_i b' = \psi_i b$. So choose $b' = b$. But $d_i b = 0$, so $\varphi_{i+1} a = 0$ and therefore by exactness $a = 0$.

($\ker \delta_i \subseteq \operatorname{im} \beta_i$.) If $\delta_i[c] = [a] = 0$ then $a = d_i u$ for some $u \in A^i$. So $d_i b = \varphi_{i+1} a = \varphi_{i+1} d_i u = d_i \varphi_i u$, which gives $d_i(b - \varphi_i u) = 0$. As $b - \varphi_i u$ is closed it determines a cohomology class in B^i, and by exactness, $\beta_i [b - \varphi_i u] = [\psi_i b] = [c]$.

4.7 By stereographic projection we can cover S^n by two patches. U covers everything but the south pole while V covers everything but the north pole, and each is homeomorphic to $\mathbb{R}^n$. The question is what happens on the overlap. In the case of the circle, the two patches overlap in two disconnected pieces, each of which is homotopic to a point—in other words, the intersection is homotopic to a zero-dimensional sphere. In the case of the two sphere, the two patches overlap in a fat band around the equator, which is homotopic to a one dimensional sphere. We claim that the same thing happens in any dimension: $U \cap V$ is homotopic to an $n-1$ dimensional sphere. There are many ways to see this. Here is one.

Define $S^n = \{x \in \mathbb{R}^{n+1} : (x^1)^2 + \cdots + (x^n)^2 = 1\}$. Let the north pole be $(0, \dots, 1)$ and the south pole $(0, \dots, -1)$. Using homotopy, shrink U and V to sets that just overlap along the equator $x^n = 0$. Then U and V overlap in an open set consisting of all points on the n-sphere satisfying $x^n \in (-\epsilon, \epsilon)$. This set is homotopic to the equator itself, which is clearly an $n-1$-sphere.

We proceed by induction on n. The base case $n = 1$ is the previous example. Assume it is true for $n-1$ for $n > 1$. We know that $H^0_{\text{dR}}(U) = H^0_{\text{dR}}(V) = \mathbb{R}$ and $H^k_{\text{dR}}(U) = H^k_{\text{dR}}(V) = 0$ for $k \neq 1$, and by hypothesis $H^0_{\text{dR}}(S^k) = \mathbb{R}$

for $k < n$. Therefore, the bottom two rows of the Mayer-Vietoris sequence look like this:

$$\begin{tikzcd}
H^1(S^n) \arrow[r] & 0 \arrow[r] & \cdots \\
H^0(S^n) \arrow[r] & \mathbb{R}\oplus\mathbb{R} \arrow[r, "\psi^*"] & \mathbb{R} \arrow[ull, "\delta_0"]
\end{tikzcd}$$

As in the case of the circle, $\operatorname{rk}\psi = \dim\ker\psi = 1$, so by exactness we conclude that $H^0_{\text{dR}}(S^n) = \mathbb{R}$ (which we already knew) and $H^1_{\text{dR}}(S^n) = 0$. Now, by hypothesis $H^{n-2}_{\text{dR}}(S^{n-1}) = 0$ and $H^{n-1}_{\text{dR}}(S^{n-1}) = \mathbb{R}$, so the top three rows look like this:

$$\begin{tikzcd}
H^n(S^n) \arrow[r] & 0 \arrow[r] & \cdots \\
H^{n-1}(S^n) \arrow[r] & 0 \arrow[r] & \mathbb{R} \arrow[ull, "\delta_1"] \\
& \cdots \arrow[r] & 0 \arrow[ull, "\delta_0"]
\end{tikzcd}$$

By exactness we therefore conclude $H^{n-1}_{\text{dR}}(S^n) = 0$ and $H^n_{\text{dR}}(S^n) = \mathbb{R}$. Finally, using the induction hypothesis again on the middle range of the Mayer-Vietoris diagram we get $H^k_{\text{dR}}(S^n) = 0$ if $k \neq 0$ and $k \neq n$.

4.8 a. $f \sim f$ via the homotopy $F(t, x) = f(x)$ for all t.

b. $f \sim g$ implies there exists a continuous map $F : I \times X \to Y$ such that $F(0, x) = f(x)$ and $F(1, x) = g(x)$. Define $G(t, x) = F(1 - t, x)$. This is clearly continuous, and provides the required homotopy from g to f.

c. $f \sim g$ and $g \sim h$ implies there exist continuous maps $F : I \times X \to Y$ and $G : I \times X \to Y$ such that $F(0, x) = f(x)$, $F(1, x) = g(x)$, $G(0, x) = g(x)$, and $G(1, x) = h(x)$. Define

$$H(t, x) = \begin{cases} F(2t, x), & \text{if } 0 \le t \le 1/2, \text{ and} \\ G(2t - 1, x) & \text{if } 1/2 \le t \le 1. \end{cases}$$

Then $H(t, x)$ is a homotopy from f to h. It is continuous for all x and for all $t \in [0, 1]$ because $F(1, x) = G(0, x)$.

4.9 If $f \sim g$ there exists a continuous $F : I \times X \to Y$ such that $F(0, x) = f(x)$ and $F(1, x) = g(x)$. Define a continuous map $G : I \times X \to Z$ by $G(t, x) = h \circ F(t, x)$. Then $G(0, x) = h(f(x))$ and $G(1, x) = h(g(x))$, so G is a homotopy between $h \circ f$ and $h \circ g$.

If $g \sim h$ there exists a continuous $F : I \times Y \to Z$ such that $F(0, y) = g(y)$ and $F(1, y) = h(y)$. Define a continuous map $G : I \times X \to Z$ by $G(t, x) = F(t, f(x))$. Then $G(0, x) = g(f(x))$ and $G(1, x) = h(f(x))$, so G is a homotopy between $g \circ f$ and $h \circ f$.

4.10 Recall that $X \sim Y$ if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \sim id_X$ and $f \circ g \sim id_Y$.

a. Set $f = g = id_X$ to show that $X \sim X$.

b. If $X \sim Y$ then clearly $Y \sim X$.

c. If $X \sim Y$ and $Y \sim Z$ then, in addition to f and g above, there exist continuous maps $h : Y \to Z$ and $k : Z \to Y$ such that $k \circ h \sim id_Y$ and $h \circ k \sim id_Z$. Define $u : X \to Z$ and $v : Z \to X$ by $u = h \circ f$ and $v = g \circ k$. Then by the results of Exercise 4.9, $v \circ u = g \circ k \circ h \circ f \sim id_X$ and $u \circ v = h \circ f \circ g \circ k \sim id_Z$. Thus $X \sim Z$.

4.11 a. This follows from the results of Exercise 4.8, because path-homotopy is a more restrictive kind of homotopy, namely one that preserves the endpoints.

b. Suppose $\alpha\beta$ is defined, so that $\alpha(1) = \beta(0)$. Let $\alpha' \in [\alpha]$ and $\beta' \in [\beta]$. Then $\alpha'(1) = \beta'(0)$ as well, so $\alpha'\beta'$ makes sense. There exist homotopies $F : I \times I \to X$ and $G : I \times I \to X$ such that $F(0, t) = \alpha(t)$, $F(1, t) = \alpha'(t)$, $F(s, 0) = \alpha(0) = \alpha'(0)$, $F(s, 1) = \alpha(1) = \alpha'(1)$, $G(0, t) = \beta(t)$, $G(1, t) = \beta'(t)$, $G(s, 0) = \beta(0) = \beta'(0)$, and $G(s, 1) = \beta(1) = \beta'(1)$. Define a continuous map $H : I \times I \to X$ by

$$H(s, t) = \begin{cases} F(s, 2t), & \text{if } 0 \le t \le 1/2, \\ G(s, 2t - 1), & \text{if } 1/2 \le t \le 1. \end{cases}$$

Then $H(0, t) = \alpha\beta(t)$, $H(1, t) = \alpha'\beta'(t)$, $H(s, 0) = \alpha\beta(0) = \alpha'\beta'(0)$, and $H(s, 1) = \alpha\beta(1) = \alpha'\beta'(1)$, so $\alpha'\beta'$ is path-homotopic to $\alpha\beta$ via the homotopy H.

c. Let $\alpha(0) = p$ and $\alpha(1) = q$, and suppose $\alpha' \in [\alpha]$. Then there exists a path-homotopy $F : I \times I \to X$ such that $F(0, t) = \alpha(t)$, $F(1, t) = \alpha'(t)$, $F(s, 0) = p$, and $F(s, 1) = q$. Define $G(s, t) = F(s, 1 - t)$. Then $G(0, t) = \alpha(1 - t) = \alpha^{-1}(t)$, $G(1, t) = \alpha'(1 - t) = \alpha'^{-1}(t)$, $G(s, 0) = q$, and $G(s, 1) = p$. So α'^{-1} is path-homotopic to α^{-1}, which shows that $[\alpha]^{-1}$ is independent of class representative.

d. i). If α and β are two loops based at p then $\alpha\beta$ is a loop based at p. So $[\alpha][\beta] = [\alpha\beta]$ exists and is a path-homotopy equivalence class of loops based at p.

ii). Let α, β, and γ be loops based at p. To show multiplication is associative we must show that $[(\alpha\beta)\gamma] = [\alpha(\beta\gamma)]$. Now

$$((\alpha\beta)\gamma)(t) = \begin{cases} \alpha(4t), & \text{if } 0 \le t \le 1/4, \\ \beta(4t-1), & \text{if } 1/4 \le t \le 1/2, \\ \gamma(2t-1), & \text{if } 1/2 \le t \le 1, \end{cases}$$

and

$$(\alpha(\beta\gamma))(t) = \begin{cases} \alpha(2t), & \text{if } 0 \le t \le 1/2, \\ \beta(4t-2), & \text{if } 1/2 \le t \le 3/4, \\ \gamma(4t-3), & \text{if } 3/4 \le t \le 1 \end{cases}$$

Define

$$F(s,t) = \begin{cases} \alpha\left(\frac{4t}{s+1}\right), & \text{if } 0 \le t \le (s+1)/4, \\ \beta(4t-(s+1)), & \text{if } (s+1)/4 \le t \le (s+2)/4, \\ \gamma\left(\frac{4t-(s+2)}{2-s}\right), & \text{if } (s+2)/4 \le t \le 1. \end{cases}$$

Then $F(0,t) = ((\alpha\beta)\gamma)(t)$, $F(1,t) = (\alpha(\beta\gamma))(t)$, and $F(s,0) = F(s,1) = p$, so F is the homotopy we desire.

iii). $[id_p]$ is the identity element of $\pi_1(X, p)$, for $[\alpha][id_p] = [\alpha \cdot id_p] = [\alpha]$ and $[id_p][\alpha] = [id_p \cdot \alpha] = [\alpha]$.

iv). If α is a loop based at p then so is α^{-1}. We must show that $[\alpha][\alpha^{-1}] = [id_p]$. For that, define a continuous family $\gamma_s(t)$ of loops based at p indexed by $s \in [0,1]$ according to

$$\gamma_s(t) := \begin{cases} \alpha(2st), & \text{if } 0 \le t \le 1/2, \\ \alpha(2s(1-t)), & \text{if } 1/2 \le t \le 1. \end{cases}$$

Define a map $F : I \times I \to X$ by $F(s,t) = \gamma_s(t)$. Then $F(0,t) = \alpha(0) = p = id_p(t)$, $F(1,t) = (\alpha\alpha^{-1})(t)$, and $F(s,0) = F(s,1) = \alpha(0) = p$. Hence F is a homotopy from $\alpha\alpha^{-1}$ to id_p.

e. Let γ be a path from p to q. If β is a loop at q then $\alpha := \gamma^{-1}\beta\gamma$ is a loop at p. So we get a map $\widetilde{\gamma} : \pi_1(X,q) \to \pi_1(X,p)$ given by $[\beta] \mapsto [\alpha]$. The claim is that $\widetilde{\gamma}$ is a group isomorphism. (It is well defined, because if β' is path homotopic to β (fixing q), then $\gamma^{-1}\beta\gamma$ is path homotopic to $\gamma^{-1}\beta'\gamma$ (by a homotopy that fixes p and q).)

i). $\widetilde{\gamma}$ is a homomorphism.

$$\begin{aligned} \widetilde{\gamma}([\beta_1][\beta_2]) = \widetilde{\gamma}([\beta_1\beta_2]) &= [\gamma^{-1}\beta_1\beta_2\gamma] = [\gamma^{-1}\beta_1\gamma\gamma^{-1}\beta_2\gamma] \\ &= [\gamma^{-1}\beta_1\gamma][\gamma^{-1}\beta_2\gamma] = (\widetilde{\gamma}[\beta_1])(\widetilde{\gamma}[\beta_2]). \end{aligned}$$

ii). $\widetilde{\gamma}$ is bijective. If $\widetilde{\gamma}[\beta_1] = \widetilde{\gamma}[\beta_2]$ then there is a homotopy from $\gamma^{-1}\beta_1\gamma$ to $\gamma^{-1}\beta_2\gamma$ that fixes p. By conjugating with γ (and, if necessary, changing the parameterization of the curves), we get a homotopy from β_1

to β_2 that fixes q, so $[\beta_1] = [\beta_2]$. Also, by shrinking the tail $\gamma^{-1}\gamma$, $\widetilde{\gamma}[\gamma\alpha\gamma^{-1}] = [\alpha]$, so $\widetilde{\gamma}$ is surjective.

f. See, *e.g.,* [55], p. 81 or [77], p. 61.

4.12 Obviously $f = id_Y \circ f$. But Y is contractible, so $id_Y \sim id_q$ for some $q \in Y$. So, by the results of Exercise 4.9, $f \sim id_q \circ f = id_q$.

If $f : X \to S^n$ is not surjective then we may view it as a map to S^n minus a point. But using stereographic projection we see that S^n minus a point is homotopic to $\mathbb{R}^n$, which, in turn, is contractible.

4.13 a. S^0 is a two point space, and B^1 is an interval of the real line. If f is 0-connected and $f(-1) = p$ and $f(1) = q$, then f can be extended to a map from the interval $[-1, 1]$ into X whose endpoints are p and q. In other words, X is path connected. Conversely, if X is path connected, then there is a path $f : [-1, 1] \to X$ between any two points, say $p = f(-1)$ and $q = f(1)$, which means that any map $f : \{-1, 1\} \to X$ can be extended to a map from $B^1 \cong [-1, 1]$ into X, so X is 0-connected.

b. Let $f : S^i \to X$ be given, where $S^i = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ is the standard sphere. Suppose f extends to a map $\hat{f}$ on the $(i+1)$-ball $B^{i+1} = \{x \in \mathbb{R}^{n+1} : \|x\| \le 1\}$. Consider the map $F : I \times B^{i+1} \to X$ given by

$$F(t, x) = \hat{f}(tx).$$

Then for any $x \in S^i$, $F(0, x) = \hat{f}(0)$ and $F(1, x) = \hat{f}(x) = f(x)$, so f is null homotopic.

Conversely, if f is null homotopic there exists a continuous map $F : I \times S^i \to X$ with $F(0, x) = q$ for some $q \in X$ and $F(1, x) = f(x)$. Define $\hat{f} : B^{i+1} \to X$ by

$$\hat{f}(x) = \begin{cases} F(\|x\|, x/\|x\|), & \text{if } x \neq 0, \text{ and} \\ q, & \text{if } x = 0. \end{cases}$$

Clearly, $\hat{f}$ restricted to the sphere is $F(1, x) = f(x)$, so it remains to show that $\hat{f}$ is continuous at $x = 0$. This is actually a bit subtle. Although it is true that $x/\|x\|$ does not approach a well-defined limit as $x \to 0$, it does not matter. Set $x = tv$ for some unit vector v. Then in the limit that $t \to 0$, $x/\|x\| = v$, so in that case

$$\lim_{x \to 0} F(\|x\|, x/\|x\|) = \lim_{t \to 0} F(t, v) = q.$$

As this holds independent of v, $\hat{f}$ is continuous.

c. If $X \sim Y$ there exist $f : X \to Y$ and $g : Y \to X$, both continuous, such that $g \circ f \sim id_X$ and $f \circ g \sim id_Y$. By virtue of the result of Part (b), we need only show that if every continuous map from S^i to X is null homotopic

$(1 \le i \le k)$ then every continuous map from S^i to Y is null homotopic. So let $h : S^i \to Y$ be given, and define $u := g \circ h : S^i \to X$. By hypothesis, u is null homotopic. But by Exercise 4.9, $f \circ u = f \circ g \circ h \sim id_Y \circ h = h$. As u is null homotopic, so is $f \circ u$, therefore h is null homotopic.

d. If S^n were n-connected, then every continuous map $f : S^n \to S^n$ would be null homotopic. In particular, the identity map would be null homotopic, which would say that the sphere is contractible. But it's not.

4.14 Let U and V be two overlapping cylinders covering the torus. They intersect in two disjoint cylinders. A cylinder is homotopic to a circle, so the Mayer-Vietoris sequence looks like this:

$$\begin{tikzcd}
H^2(T^2) \arrow[r] & 0 \arrow[r] & 0 \\
H^1(T^2) \arrow[r, "\varphi_1^*"] & \mathbb{R}\oplus\mathbb{R} \arrow[r, "\psi_1^*"] & \mathbb{R}\oplus\mathbb{R} \arrow[ull, "\delta_1"] \\
H^0(T^2) \arrow[r, "\varphi_0^*"] & \mathbb{R}\oplus\mathbb{R} \arrow[r, "\psi_0^*"] & \mathbb{R}\oplus\mathbb{R} \arrow[ull, "\delta_0"]
\end{tikzcd}$$

The torus is connected, so $H^0_{\mathrm{dR}}(T^2) = \mathbb{R}$. As in the case of the circle, the difference map ψ_0^* has rank one, so $\dim\ker\delta_0 = \operatorname{rk}\psi_0^* = 1$. By counting dimensions we get $\operatorname{rk}\delta_0 = 1$, so $\dim\ker\varphi_1^* = 1$. Assume for the moment that $\operatorname{rk}\psi_1^* = 1$. Then $\dim\ker\psi_1^* = 1$, so $\operatorname{rk}\varphi_1^* = 1$. Hence, $\operatorname{rk}\varphi_1^* + \dim\ker\varphi_1^* = 2$, so $H^1_{\mathrm{dR}}(T^2) = \mathbb{R}\oplus\mathbb{R}$. Also, $\dim\ker\delta_1 = \operatorname{rk}\psi_1^* = 1$ so $\operatorname{rk}\delta_1 = 1$. By exactness, δ_1 is surjective, so $H^2_{\mathrm{dR}}(T^2) = \mathbb{R}$.

It remains to show that $\operatorname{rk}\psi_1^* = 1$. Write $U \cap V = X \cup Y$, where X and Y are disjoint cylinders. X is a deformation retract of U (and of V), and similarly for Y, so the inclusion maps $i_U^* : H^\bullet_{\mathrm{dR}}(U) \to H^\bullet_{\mathrm{dR}}(X \cup Y)$ and $i_V^* : H^\bullet_{\mathrm{dR}}(V) \to H^\bullet_{\mathrm{dR}}(X \cup Y)$ are isomorphisms in cohomology. Thus $\psi_1^*(\omega, \nu) = (i_V^* - i_U^*)(\omega, \nu) = (\nu - \omega, \nu - \omega)$ is a cohomology class on $X \cup Y$, which shows that $\operatorname{rk}\psi_1^* = 1$.

Natural generators of cohomology in the various degrees are as follows. (1) Degree 0: the unit function; (2) Degree 1: the one forms dx and dy from Exercise 3.23; (3) Degree 2: the two form $dx \wedge dy$. (1) is immediate. As for (2), dx and dy are obviously closed, but neither is exact, because neither x nor y is single valued on the torus. Lastly, (3) holds because if we had $dx \wedge dy = d\theta$ for some θ then we would have to have $\theta = x\,dy$ or $\theta = -y\,dx$ or some linear combination, none of which are single-valued. (A slightly more rigorous argument based on integration can be found in [84], Proposition 28.2.)

4.15 We have $T = U \cup V$ and $U \cap V \cong S^1$, so the Mayer-Vietoris sequence becomes

$$
\begin{tikzcd}
\mathbb{R} \arrow[r, "\varphi_2^*"] & H^2(T') \arrow[r, "\psi_2^*"] & 0 \\
\mathbb{R}\oplus\mathbb{R} \arrow[r, "\varphi_1^*"] & H^1(T') \arrow[r, "\psi_1^*"] & \mathbb{R} \arrow[ull, "\delta_1"] \\
\mathbb{R} \arrow[r, "\varphi_0^*"] & H^0(T')\oplus\mathbb{R} \arrow[r, "\psi_0^*"] & \mathbb{R} \arrow[ull, "\delta_0"]
\end{tikzcd}
.
$$

T' is connected, so $H^0(T') = \mathbb{R}$. $\dim\ker\delta_0 = \operatorname{rk}\psi_0^* = 1$, so φ_1^* is injective, which means that $H^1(T')$ is at least two dimensional. But ψ_1^* is the zero map, so $H^1(T') = \mathbb{R}\oplus\mathbb{R}$ and δ_1 is injective. By counting dimensions δ_1 is also surjective, so $\dim\ker\varphi_2^* = 1$, which means that $\operatorname{rk}\varphi_2^* = 0$. Hence $H^2(T') = 0$.

4.16 We denote the genus 2 surface by T_2. (The subscript denotes the genus, not the dimension; the latter is always two in this exercise.) Let $U = V = T'$ be the punctured torus with the hole stretched to an open disk, so that $U \cap V$ is the circle. Then using the previous exercise the Mayer-Vietoris sequence looks like this:

$$
\begin{tikzcd}
H^2(T_2) \arrow[r, "\varphi_2^*"] & 0 \arrow[r, "\psi_2^*"] & 0 \\
H^1(T_2) \arrow[r, "\varphi_1^*"] & \mathbb{R}^2\oplus\mathbb{R}^2 \arrow[r, "\psi_1^*"] & \mathbb{R} \arrow[ull, "\delta_1"] \\
H^0(T_2) \arrow[r, "\varphi_0^*"] & \mathbb{R}\oplus\mathbb{R} \arrow[r, "\psi_0^*"] & \mathbb{R} \arrow[ull, "\delta_0"]
\end{tikzcd}
.
$$

T_2 is connected, so $H^0(T_2) = \mathbb{R}$. As before, $\dim\ker\delta_0 = \operatorname{rk}\psi_0^* = 1$ and ψ_1^* is the zero map, so by exactness we conclude that $H^1_{\mathrm{dR}}(T_2) = \mathbb{R}^4$ and $H^2_{\mathrm{dR}}(T_2) = \mathbb{R}$.

The obvious guess is that

$$
H^k_{\mathrm{dR}}(T_g) = \begin{cases} \mathbb{R}, & k = 0, \\ \mathbb{R}^{2g}, & k = 1, \\ \mathbb{R}, & k = 2, \text{ and} \\ 0, & k \geq 2. \end{cases}
$$

This turns out to be true, and follows by a simple inductive argument along the lines given above.

4.17 A suspension is just a double cone over the base. If we remove one of the endpoints we are left with a cone whose base is attached to a cylinder over the base. But we can shrink the cylinder back down to the base, whereupon we are left with just a cone, and this is contractible by Exercise 4.2. Similarly, $U \cap V$ is just the cylinder over M (because both endpoints are removed), which is homotopic to M. Hence, the Mayer-Vietoris sequence looks like this:

$$\begin{tikzcd}
H^{k+1}(\Sigma M) \arrow[r] & 0 \arrow[r] & \cdots \\
H^2(\Sigma M) \arrow[r] & 0 \cdots \quad \cdots 0 \arrow[r] & H^k(M) \arrow[ull, "\delta_k"'] \\
H^1(\Sigma M) \arrow[r] & 0 \arrow[r] & H^1(M) \arrow[ull, "\delta_1"'] \\
H^0(\Sigma M) \arrow[r] & \mathbb{R} \oplus \mathbb{R} \arrow[r, "\psi_0^*"] & H^0(M) \arrow[ull, "\delta_0"']
\end{tikzcd}.$$

Looking at the diagram, we see that for every $k \geq 1$ we get an exact sequence of the form

$$0 \longrightarrow H^k(M) \longrightarrow H^{k+1}(\Sigma M) \longrightarrow 0,$$

so that

$$H^k(M) \cong H^{k+1}(\Sigma M). \tag{1}$$

As for the lowest level, once again we have $\operatorname{rk} \psi_0^* = 1$. By exactness and dimension counting we get

$$H^0(\Sigma M) = \mathbb{R}. \tag{2}$$

(That's because ΣM is connected even if M is not.) Also, $\dim \ker \delta_0 = \operatorname{rk} \psi_0^* = 1$, so

$$H^0(M) = H^1(\Sigma M) \oplus \mathbb{R}. \tag{3}$$

It follows that

$$H^k(S^r) = H^{k+1}(\Sigma S^r) = H^{k+1}(S^{r+1}) \qquad (k \geq 1). \tag{4}$$

Now we proceed in steps. 1) $H^0(S^0) = \mathbb{R} \oplus \mathbb{R}$ because S^0 is just two points. 2) $H^0(S^n) = \mathbb{R}$ for any $n \geq 1$ by (2). 3) $H^1(S^1) = \mathbb{R}$ by (1) and (3). 4) $H^n(S^n) = \mathbb{R}$ for all $n > 1$ by (3) and (4). 5) $H^1(S^r) = 0$ for $r \geq 2$ by (2) and (3). 6) $H^k(S^r) = 0$ if $r - k \geq 1$ and $k > 1$ by (5) and (4).

4.18 By construction U and V are both open in $M \vee N$ (because they are the complements of closed sets). Moreover, U is homotopic to M (just shrink the ball), V is homotopic to N (ditto), and $U \cap V$ is contractible to $p = q$. The Mayer-Vietoris sequence therefore looks like this:

$$\begin{tikzcd}
H^2(M \vee N) \arrow[r] & H^2(M) \oplus H^2(N) \arrow[r] & 0 \cdots \\
H^1(M \vee N) \arrow[r] & H^1(M) \oplus H^1(N) \arrow[r] & 0 \arrow[ull, "\delta_1"'] \\
H^0(M \vee N) \arrow[r, "\varphi_0^*"] & H^0(M) \oplus H^0(N) \arrow[r, "\psi_0^*"] & \mathbb{R} \arrow[ull, "\delta_0"']
\end{tikzcd} .$$

Suppose M has r connected pieces and N has s. Then obviously $M \vee N$ has $r + s - 1$ pieces. So $H^0(M \vee N) \oplus \mathbb{R} = H^0(M) \oplus H^0(N)$. It follows that $\dim \ker \psi_0^* = \operatorname{rk} \varphi_0^* = r + s - 1$, so $\operatorname{rk} \psi_0^* = 1$. By exactness, δ_0 is the zero map. It follows immediately that $H^k(M \vee N) \cong H^k(M) \oplus H^k(N)$ for $k \geq 1$.

5

Elementary homology theory

5.1 Let K be the set of convex combinations of the points of S. Let $p, q \in K$. Then $p = \sum_i s_i p_i$ and $q = \sum_i t_i p_i$ with $s_i, t_i > 0$ and $\sum_i s_i = \sum_i t_i = 1$. Any point r on the line segment pq is $r = tp + (1-t)q = \sum_i r_i p_i$ where $0 \le t \le 1$ and $r_i := ts_i + (1-t)t_i$. As $r_i \ge 0$ and $\sum_i r_i = 1$, $r \in K$, so K is convex.

If $k = 2$ and $p \in K$ then $p = t_1 p_1 + t_2 p_2$ with $t_i \ge 0$ and $\sum_i t_i = 1$, which shows that p is in every convex set that contains p_1 and p_2—i.e., $p \in P$. If $|S| = n$, then $p \in K$ means

$$p = \sum_i t_i p_i = (1 - t_k)(s_1 p_1 + \cdots s_{k-1} p_{k-1}) + t_k p_k,$$

where $t_i \ge 0$, $\sum_{i=1}^k t_i = 1$, and $s_i = t_i/(1 - t_k)$. But $\sum_{i=1}^{k-1} s_i = 1$, so by induction p is on a segment joining two points of p, which means $p \in P$. Hence $K \subseteq P$.

But K contains the points $p_1, \dots, p_k$ and is convex, so $K \supseteq P$, which shows that $P = K$.

5.2 By Exercise 5.1, every point in $[S]$ can be written as a convex combination of its vertices. So we need only prove uniqueness. For that, suppose $(s_0, \dots, s_d)$ and $(t_0, \dots, t_d)$ represent the same point of $[S]$. Define $c_i := t_i - s_i$. Then

$$0 = \sum_{i=0}^{d} c_i p_i = (c_0 + c_1 + \dots c_d) p_0 + \sum_{i=1}^{d} c_i (p_i - p_0).$$

But the first term vanishes because $\sum_i c_i = \sum_i t_i - \sum_i s_i = 1 - 1 = 0$. By linear independence $c_i = 0$ for all $i > 0$, and when combined with $\sum_i c_i = 0$ we get $c_0 = 0$ as well.

5.3 The real homology groups of a space are of the form $H_k = \mathbb{R}^{\beta_k}$ where β_k is the k^{th} Betti number. As $Z_k = \ker \partial_k$, $B_k = \operatorname{im} \partial_{k+1}$, and $H_k = Z_k/B_k$, we have

$$\beta_k = \dim H_k = \dim \ker \partial_k - \dim \operatorname{im} \partial_{k+1}.$$

Thus, to find the homology groups of the tetrahedron K, we just need to compute the ranks and nullities of the boundary maps ∂.

Let $K = (P_0, P_1, P_2, P_3)$. Every 0-chain on K is of the form

$$c_0 = a_0 P_0 + a_1 P_1 + a_2 P_2 + a_3 P_3.$$

The boundary of a point is zero, so $\partial_0 c_0 = 0$. That is, ∂_0 is the zero map. In particular, $\dim\ker\partial_0 = \dim C_0 = 4$.

The most general 1-chain in K is

$$\begin{aligned} c_1 = b_{01}(P_0, P_1) + b_{02}(P_0, P_2) + b_{03}(P_0, P_3) \\ + b_{12}(P_1, P_2) + b_{13}(P_1, P_3) + b_{23}(P_2, P_3). \end{aligned} \tag{1}$$

Its boundary is

$$\begin{aligned} \partial_1 c_1 = b_{01}(P_1 - P_0) + b_{02}(P_2 - P_0) + b_{03}(P_3 - P_0) \\ + b_{12}(P_2 - P_1) + b_{13}(P_3 - P_1) + b_{23}(P_3 - P_2). \end{aligned}$$

Thus, $\partial_1 c_1 = c_0$ provided

$$\begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} b_{01} \\ b_{02} \\ b_{03} \\ b_{12} \\ b_{13} \\ b_{23} \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}. \tag{2}$$

The matrix in (2) represents the boundary map ∂_1. The maximum rank of the matrix is 4, because it has 4 rows, but it turns out that its rank is only 3. (This follows because the rows are not all linearly independent: their sum is zero.) Therefore

$$\beta_0 = \dim\ker\partial_0 - \dim\operatorname{im}\partial_1 = 4 - 3 = 1.$$

For use below, observe that $\dim\ker\partial_1 = 3$, which follows from the rank/nullity theorem.

Next we find β_1. The most general 2-cycle is of the form

$$c_2 = d_0(P_1, P_2, P_3) + d_1(P_0, P_2, P_3) + d_2(P_0, P_1, P_3) + d_3(P_0, P_1, P_2).$$

Its boundary is

$$\begin{aligned} \partial_2 c_2 = d_0[(P_1, P_2) - (P_1, P_3) + (P_2, P_3)] \\ + d_1[(P_0, P_2) - (P_0, P_3) + (P_2, P_3)] \\ + d_2[(P_0, P_1) - (P_0, P_3) + (P_1, P_3)] \\ + d_3[(P_0, P_1) - (P_0, P_2) + (P_1, P_2)]. \end{aligned}$$

We have $\partial_2 c_2 = c_1$ for c_1 as in (1) if

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} b_{01} \\ b_{02} \\ b_{03} \\ b_{12} \\ b_{13} \\ b_{23} \end{pmatrix}. \tag{3}$$

The maximum rank of ∂_2 is 4, because it has 4 columns, but the actual rank is 3, because columns one and three add to the same vector as columns two and four. It follows that

$$\beta_1 = \dim \ker \partial_1 - \dim \operatorname{im} \partial_2 = 3 - 3 = 0.$$

Note that $\dim \ker \partial_2 = 1$.

Lastly, consider ∂_3, which maps from three chains to two chains. There are no three chains in K (which comprises only the surface of a tetrahedron), so the image of ∂_3 is only the zero vector, which has dimension 0. Hence

$$\beta_2 = \dim \ker \partial_2 - \dim \operatorname{im} \partial_3 = 1 - 0 = 1.$$

The analogous answer for the n-sphere is that it has trivial homology in all degrees except 0 and n:

$$H_p(S^n) = \begin{cases} \mathbb{R}, & p = 0 \text{ or n}, \\ 0, & \text{otherwise}. \end{cases}$$

5.4 As x and A both have positive real entries, so does Ax. Moreover, by construction $\sum_i f^i = 1$, so f maps the simplex σ^{n-1} to itself. As σ^{n-1} is homeomorphic to the $(n-1)$-ball, Brouwer's theorem guarantees the existence of a fixed point, namely a vector satisfying $f(x) = x$. But then $Ax = |Ax|x$ and the conclusion follows.

5.5 Let P denote the crosspolytope, and let $F = P \cap H$ be a proper face. If F were to contain both e_i and $-e_i$ it would contain the origin 0 (as the intersection of convex sets is convex), so we would have $0 \in H$ as well. But F is proper, so there are vertices of P not in H, say e_j, $j \neq i$. We cannot have $-e_j \in H$ either, otherwise the segment $(-e_j, 0)$ would lie in H, and by linear extension so would the segment $(0, e_j)$. Hence there are points of P on both sides of H, a contradiction.

Every face of a polytope is the convex hull of some subcollection of the vertices of P, so we examine all such subcollections. Let $S = \{v_1, \dots, v_n\}$ be any maximal subset of the vertex set of P that does not contain both e_i and $-e_i$ for any i. S is linearly independent, so $\{v_2 - v_1, \dots, v_n - v_1\}$ are linearly independent. In other words, S is affinely independent, so it determines a simplex F in the affine subspace H spanned by S. We claim that F is a facet of P. To show this, assume that $S = \{e_1, \dots, e_n\}$ (the general case is similar). Suppose H were not a supporting hyperplane of P. Then there would exist a point $p \in P$ on the opposite side of H from the origin. Let q be the point of intersection of the ray $(0p)$ with H, so that $p = tq$ for some $t > 1$. As $q = \sum_{i>0} s_i e_i$ with $\sum_i s_i = 1$ and $s_i \geq 0$, we would have $p = \sum_{i>0} t s_i e_i$ with $t s_i \geq 0$. But $p \in P$, so we must have $\sum_i t s_i = 1$, a contradiction. A similar argument shows that every maximal subcollection of the vertex set of P not containing both e_i and $-e_i$ is a facet of P. All the other faces are then determined by subsets of these subcollections.

5.6 a. By simple counting we obtain the following table.

Platonic Solid	V	E	F
tetrahedron	4	6	4
cube	8	12	6
octahedron	6	12	8
dodecahedron	20	30	12
icosahedron	12	30	20

The face numbers of the cube and octahedron are the reverse of one another, as are the face numbers of the dodecahedron and the icosahedededron. This is because those figures are related by *duality*: putting a vertex in the middle of each face and taking the convex hull gives you the dual figure, reversing the face numbers. The tetrahedron is in its own class because it is *self-dual*.

b. We use the classic method of double counting. Let $|(e, f)|$ be the number of pairs with the edge e contained in the 2-face f. There are F 2-faces, and each 2-face contains p edges, so $|(e, f)| = pF$. But there are E edges, and each edge is contained in two 2-faces, so $|(e, f)| = 2E$. Hence $2E = pF$. Similarly, by counting pairs (v, e) of vertices v contained in edges e in two ways, we get $qV = 2E$. It follows that $pF = qV = 2E$. Euler's equation thus gives

$$2 = V - E + F = \frac{2E}{q} - E + \frac{2E}{p} \quad \Rightarrow \quad E = \left(\frac{1}{p} + \frac{1}{q} - \frac{1}{2}\right)^{-1}.$$

As $E > 0$,

$$\frac{1}{p} + \frac{1}{q} > \frac{1}{2}.$$

Now start plugging in numbers. We must have $p \geq 3$, because the triangle is the smallest regular polygon. Similarly, $q \geq 3$, otherwise the figure is not three dimensional. But the inequality implies that p and q cannot both exceed 5. This leaves only $\{3, 3\}$, $\{4, 3\}$, $\{3, 4\}$, $\{5, 3\}$, and $\{3, 5\}$, namely the Platonic solids.

5.7 From the figure, we count 9 vertices, 27 edges, and 18 2-faces. (Identified vertices are labeled, and edges with the same endpoints are identified. All the 2-faces are distinct.) The Euler characteristic is therefore

$$\chi = V - E + F = 9 - 27 + 18 = \boxed{0}.$$

5.8 a. The wedge sum of two tetrahedra has 7 vertices, 12 edges, and 8 2-faces, so $\chi = 7 - 12 + 8 = 3$.

b. We computed the homology of the tetrahedron in Exercise 5.3. It is $\mathbb{R}$ in dimensions 0 and 2, and 0 in dimension 1. So the homology of the wedge sum is $\mathbb{R}$ in dimension 0, 0 in dimension 1, and $\mathbb{R} \oplus \mathbb{R}$ in dimension 2. The alternating sum of Betti numbers is therefore $1 - 0 + 2 = 3$, as before.

5.9 a. We have

$$\begin{aligned}\chi(U \cup V) &= \chi((U \cup V)\backslash V) + \chi(V)\\ &= \chi(U\backslash(U \cap V)) + \chi(V)\\ &= \chi(U) - \chi(U \cap V) + \chi(V).\end{aligned}$$

b. The only question is what happens to the Euler characteristic of a surface if you remove an open disk D^o. Using the additivity of χ, we have

$$\chi(\text{open disk}) = \chi(\text{closed disk}) - \chi(\text{circle}) = 1 - 0 = 1.$$

Hence, removing an open disk from a surface reduces its Euler characteristic by 1. We show $\chi(T^{\#k}) = 2 - 2k$ by induction on k. For $k = 0$ we get the 2-sphere, whose Euler characteristic we already know to be 2. By the definition of connected sum, the additivity of the Euler characteristic, and induction, we get

$$\begin{aligned}\chi(T^{\#k}) &= \chi(T^{\#(k-1)}\backslash D^o) + \chi(T^2\backslash D^o) - \chi(S^1)\\ &= (2 - 2(k-1) - 1) + (0 - 1) - 0\\ &= 2 - 2k.\end{aligned}$$

In the solutions to Exercise 4.16 we asserted that the cohomology of the g-holed torus was $\mathbb{R}$ in dimensions 0 and 2 and $\mathbb{R}^{2g}$ in dimension 1. By definition, the Euler characteristic is the alternating sum of Betti numbers, so we ought to have $\chi = 1 - 2g + 1 = 2 - 2g$, which agrees with our answer above, as it must.

5.10 Following the hint, we consider the homology class $[\alpha_i] \in H_i(A)$. As ψ is a chain map and α_i is a cycle in A_i,

$$\partial_i \psi_i(\alpha_i) = \psi_{i-1}\partial_i(\alpha_i) = 0,$$

so $\psi_i(\alpha_i)$ is a cycle in B_i. Hence it makes sense to consider its homology class. So define $h_i([\alpha_i]) := [\psi_i(\alpha_i)]$. We need only show that the map is independent of class representative. So, suppose $[\alpha_i] = [\beta_i]$. Then $\alpha_i - \beta_i = \partial_{i+1}(\gamma_{i+1})$ for some $\gamma_{i+1} \in A_{i+1}$. But ψ is a chain map, so

$$\psi_i(\alpha_i - \beta_i) = \psi_i \partial_{i+1}(\gamma_{i+1}) = \partial_{i+1}\psi_{i+1}(\gamma_{i+1}).$$

This means that $[\psi_i(\alpha_i - \beta_i)]$ is zero in $H_i(B)$, which, by linearity, means that $[\psi_i(\alpha_i)] = [\psi_i(\beta_i)]$.

5.11 From the result of Exercise 1.30 and the given Mayer-Vietoris sequence we get

$$\begin{aligned} 0 = &\quad \dim H_0(U \cup V) - \dim H_0(U) - \dim H_0(V) + \dim H_0(U \cap V) \\ &- \dim H_1(U \cup V) + \dim H_1(U) + \dim H_1(V) - \dim H_1(U \cap V) \\ &+ \cdots, \end{aligned}$$

from which we conclude

$$\chi(U \cup V) - \chi(U) - \chi(V) + \chi(U \cap V) = 0.$$

5.12 The sequence

$$0 \longrightarrow Z_i \xrightarrow{\ \iota\ } C_i \xrightarrow{\ \partial_i\ } B_{i-1} \longrightarrow 0$$

is trivially exact at Z_i because $Z_i \subseteq C_i$, and at B_{i-1} by definition. It is exact at C_i because, by definition, i-cycles are annihilated by the boundary operator ∂_i. It follows that

$$\operatorname{tr}(\psi_i, C_i) = \operatorname{tr}(\psi_i, Z_i) + \operatorname{tr}(\psi_i, s(B_{i-1})),$$

where $\partial_i \circ s = 1$. But ψ is a chain map, so if $b_{i-1} \in B_{i-1}$ then

$$\partial_i \psi_i(s(b_{i-1})) = \psi_{i-1}\partial_i(s(b_{i-1})) = \psi_{i-1}(b_{i-1}).$$

Hence, $\partial_i \psi_i s = \psi_{i-1}$ restricted to B_{i-1}. In particular,

$$\operatorname{tr}(\psi_i, s(B_{i-1})) = \operatorname{tr}(\psi_{i-1}, B_{i-1}).$$

The sequence

$$0 \longrightarrow B_i \xrightarrow{\ \iota\ } Z_i \xrightarrow{\ \pi_i\ } H_i \longrightarrow 0$$

is exact because $H_i = Z_i/B_i$. It follows that

$$\operatorname{tr}(\psi_i, Z_i) = \operatorname{tr}(\psi_i, B_i) + \operatorname{tr}(\psi_i, H_i).$$

As we showed in Exercise 5.10, the map $h_i : H_i \to H_i$ is defined as follows. Let α_i be a cycle in C_i. It defines a cohomology class $[\alpha_i]$ in H_i, and we set $h_i([\alpha_i]) = [\psi_i(\alpha_i)]$. In other words, the map ψ_i, restricted to H_i, is precisely the map h_i. Hence $\mathrm{tr}(\psi_i, H_i) = \mathrm{tr}(h_i, H_i)$. Thus

$$\begin{aligned}
\sum_{i=0}^{n}(-1)^i \,\mathrm{tr}(\psi_i, C_i) &= \sum_{i=0}^{n}(-1)^i[\mathrm{tr}(\psi_i, Z_i) + \mathrm{tr}(\psi_{i-1}, B_{i-1})] \\
&= \sum_{i=0}^{n}(-1)^i[\mathrm{tr}(\psi_i, B_i) + \mathrm{tr}(\psi_i, H_i) + \mathrm{tr}(\psi_{i-1}, B_{i-1})] \\
&= \sum_{i=0}^{n}(-1)^i[\mathrm{tr}(\psi_i, B_i) + \mathrm{tr}(h_i, H_i) + \mathrm{tr}(\psi_{i-1}, B_{i-1})].
\end{aligned}$$

But the sum over the boundary spaces telescopes to zero:

$$\begin{aligned}
&\sum_{i=0}^{n}(-1)^i[\mathrm{tr}(\psi_i, B_i) + \mathrm{tr}(\psi_{i-1}, B_{i-1})] \\
&\qquad = \sum_{i=0}^{n}(-1)^i \,\mathrm{tr}(\psi_i, B_i) + \sum_{i=-1}^{n-1}(-1)^{i+1} \,\mathrm{tr}(\psi_i, B_i) \\
&\qquad = \sum_{i=0}^{n}(-1)^i \,\mathrm{tr}(\psi_i, B_i) + \sum_{i=0}^{n}(-1)^{i+1} \,\mathrm{tr}(\psi_i, B_i) \\
&\qquad = 0,
\end{aligned}$$

because $B_{-1} = B_n = 0$. Hence the claim is proved.

6

Integration on manifolds

6.1 We have

$$\gamma^*\omega = (t^2)(t^3)\,dt + (t^3)(t)(2t)\,dt + (t)(t^2)(3t^2)\,dt = 6t^5\,dt,$$

so

$$\int_\gamma \omega = \int_I \gamma^*\omega = 6\int_0^1 t^5\,dt = t^6\Big|_0^1 = 1.$$

6.2 We have

$$\begin{aligned}
\varphi^*(dy \wedge dz) &= \varphi^*(dy) \wedge \varphi^*(dz) \\
&= d(\varphi^* y) \wedge d(\varphi^* z) \\
&= d(y \circ \varphi) \wedge d(z \circ \varphi) \\
&= d\varphi^2 \wedge d\varphi^3 \\
&= \left(\frac{\partial \varphi^2}{\partial u}\,du + \frac{\partial \varphi^2}{\partial v}\,dv\right) \wedge \left(\frac{\partial \varphi^3}{\partial u}\,du + \frac{\partial \varphi^3}{\partial v}\,dv\right) \\
&= \left(\frac{\partial \varphi^2}{\partial u}\frac{\partial \varphi^3}{\partial v} - \frac{\partial \varphi^2}{\partial v}\frac{\partial \varphi^3}{\partial u}\right) du \wedge dv \\
&= \left(\frac{\partial \boldsymbol{\varphi}}{\partial u} \times \frac{\partial \boldsymbol{\varphi}}{\partial v}\right)^1 du \wedge dv \\
&= n^1\,du \wedge dv
\end{aligned}$$

and

$$\varphi^*\beta = (B_x \circ \varphi)\varphi^*(dy \wedge dz) + \text{cyclic}.$$

Thus

$$\varphi^*\beta = \boldsymbol{B}(\varphi(u, v)) \cdot \boldsymbol{n}(\varphi(u, v))\,du \wedge dv,$$

and

$$\int_\Sigma \beta = \int_c \varphi^* \beta = \int_c \boldsymbol{B}(\varphi(u, v)) \cdot \boldsymbol{n}(\varphi(u, v))\, du \wedge dv = \int_\Sigma \boldsymbol{B} \cdot d\boldsymbol{S}.$$

6.3 By the change of variables theorem,

$$\int_\Sigma d\omega = \int_U \varphi^* d\omega.$$

We have

$$\omega = -yz\,dx + xz\,dy \quad \Rightarrow \quad d\omega = 2z\,dx \wedge dy - x\,dy \wedge dz - y\,dz \wedge dx.$$

Hence

$$\begin{aligned}\varphi^* d\omega = 2(u^2)&(3u^2 \cos v\,du - u^3 \sin v\,dv) \wedge (3u^2 \sin v\,du + u^3 \cos v\,dv)\\ &- (u^3 \cos v)(3u^2 \sin v\,du + u^3 \cos v\,dv) \wedge (2u\,du)\\ &- (u^3 \sin v)(2u\,du) \wedge (3u^2 \cos v\,du - u^3 \sin v\,dv)\\ = 8&u^7\,du \wedge dv.\end{aligned}$$

This gives

$$\int_\Sigma d\omega = \int_0^{2\pi} dv \int_1^2 8u^7\,du = 2\pi u^8 \Big|_1^2 = 510\pi.$$

On the other hand,

$$\int_{\partial\Sigma} \omega = \int_{\partial U} \varphi^* \omega.$$

We have

$$\begin{aligned}\varphi^* \omega &= -(u^5 \sin v)(3u^2 \cos v\,du - u^3 \sin v\,dv)\\ &\qquad + (u^5 \cos v)(3u^2 \sin v\,du + u^3 \cos v\,dv)\\ &= u^8\,dv.\end{aligned}$$

The boundary of U is a square, so we need to compute the line integral around the square. This gives

$$\int_{\partial\Sigma} \omega = \oint_{\partial U} u^8\,dv = 2^8(2\pi - 0) + 1^8(0 - 2\pi) = 510\pi,$$

as expected. Hence Stokes' theorem is confirmed in this case.

6.4 Both results are direct consequences of Stokes' theorem. If ω is closed and c is any chain, then

$$\int_{\partial c} \omega = \int_c d\omega = 0.$$

If $\omega = d\mu$ is exact and the chain c has no boundary, then

$$\int_c \omega = \int_c d\mu = \int_{\partial c} \mu = 0.$$

6.5 Let ω be a k-form and c a $(k+2)$-chain. Then Stokes theorem applied twice gives

$$\int_c d^2\omega = \int_{\partial c} d\omega = \int_{\partial^2 c} \omega.$$

It follows that $d^2 = 0$ implies $\partial^2 = 0$ and vice versa.

6.6 Let $\{V_j\}$ be another coordinate cover of M, and let $\{\lambda_i\}$ be a partition of unity subordinate to $\{V_i\}$. Note that

$$\int_{U_i} \rho_i \lambda_j \omega = \int_{V_j} \rho_i \lambda_j \omega$$

because $\rho_i \lambda_j \omega$ has support in $U_i \cap V_j$. Therefore

$$\sum_i \int_{U_i} \rho_i \omega = \sum_{i,j} \int_{U_i} \rho_i \lambda_j \omega = \sum_{ij} \int_{V_j} \rho_i \lambda_j \omega = \sum_j \int_{V_j} \lambda_j \omega.$$

because $\sum_i \rho_i = \sum_j \lambda_j = 1$.

6.7 Stokes' theorem gives

$$\int_{\partial M} f\eta = \int_M d(f\eta) = \int_M df \wedge \eta + \int_M f \wedge d\eta.$$

6.8 For any ℓ-chain c,

$$\left(\int \omega\right)(c) = \left(\int d\tau\right)(c) = \int_c d\tau = \int_{\partial c} \tau = \left(\int \tau\right)(\partial c) = \left(\partial^* \int \tau\right)(c).$$

6.9 Let $f \in Z^\ell(M)$ be a cocycle. Then $f(\partial c) = (\partial^* f)(c) = 0$, so f vanishes on all boundaries. By de Rham's second theorem, there exists a closed form ω such that, for every cycle z, $\int_z \omega = f(z)$. This shows that $\int \omega$ and f agree on cycles, and hence define the same cocycle. In other words, every cocycle is of the form $\int \omega$ for some closed ω, which means that the map $[\omega] \mapsto [\int \omega]$ is surjective.

Next, suppose $\int \omega$ is zero in $H^\ell(M)$. Then $\int \omega \in B^\ell(M)$, which means there exists a τ such that $\int \omega = \partial^* \int \tau$. Hence, if z is a cycle,

$$\int_z \omega = \left(\int \omega\right)(z) = \left(\partial^* \int \tau\right)(z) = \left(\int \tau\right)(\partial z) = 0.$$

By de Rham's first theorem, ω is exact, so $[\omega] = 0$, whence we see that $[\omega] \mapsto [\int \omega]$ is injective.

6.10 We choose the orientation to be the one in which the variables are ordered (u^2, u^1) and (v^1, v^2), so that the Jacobian is positive on the overlap. As we see from Exercise 3.40, this is equivalent to choosing the area element on the sphere to be $+\sin\theta\, d\theta \wedge d\phi$. (Equivalently, we choose the 'outward pointing normal'.) Integrating over all of U gives

$$\begin{aligned}\int_U \omega_U &= \int_{U_0} \frac{du_2 \wedge du_1}{(1+\eta)^2} \\ &= \int_0^{2\pi} d\theta \int_0^\infty dr\, \frac{r\,dr\,d\theta}{(1+r^2)^2} \\ &= -2\pi \cdot \left.\frac{1/2}{1+r^2}\right|_0^\infty = \pi.\end{aligned}$$

Alternatively,

$$\int_{S^2} \omega = \int_{U_0} \omega_U + \int_{V_0} \omega_V,$$

where U_0 is the region of (u_1, u_2)-space corresponding to the southern hemisphere and V_0 is the region of (v_1, v_2)-space corresponding to the northern hemisphere. In both cases, this region is the interior of the unit circle in the plane. The first integral is

$$\begin{aligned}\int_{U_0} \omega_U &= \int_{U_0} \frac{du_2 \wedge du_1}{(1+\eta)^2} \\ &= \int_0^{2\pi} d\theta \int_0^1 dr\, \frac{r\,dr\,d\theta}{(1+r^2)^2} \\ &= -2\pi \cdot \left.\frac{1/2}{1+r^2}\right|_0^1 = \frac{\pi}{2}.\end{aligned}$$

A similar computation gives

$$\int_{V_0} \omega_V = \frac{\pi}{2},$$

and adding them together gives π, as before.

6.11 As in the previous exercise, it is easiest just to integrate $dx \wedge dy$ over U_1 and ignore the missing parts, which are sets of measure zero. But then the integral is about as simple as it gets:

$$\int_{U_1} dx \wedge dy = \int_0^1 dx \int_0^1 dy = 1.$$

6.12 a. We see immediately that the integral over the base must vanish, because $dz = 0$ there. A natural parameterization for the rest of the cone surface is

$$\varphi(u, v) = (u \cos v, u \sin v, uh/a), \qquad 0 \le u \le a, 0 \le v \le 2\pi.$$

This gives

$$\begin{aligned}\varphi^*\sigma &= (u \cos v)(\sin v\, du + u \cos v\, dv) \wedge ((h/a)du) \\ &\qquad + (u \sin v)((h/a)du) \wedge (\cos v\, du - u \sin v\, dv) \\ &= -(h/a)u^2\, du \wedge dv.\end{aligned}$$

Thus,

$$\int_{\partial V} \sigma = -\frac{h}{a} \int_0^{2\pi} dv \int_0^a u^2\, du = -\frac{2\pi}{3} ha^2.$$

The negative sign just means that our parameterization naturally selects the inward pointing normal rather than the outward pointing normal. To see this, look back at Exercise 6.2. With our parameterization,

$$\begin{aligned}\boldsymbol{n} &= (\cos v, \sin v, h/a) \times (-u \sin v, u \cos v, 0) \\ &= (-(uh/a) \cos v, -(uh/a) \sin v, u),\end{aligned}$$

which definitely points inwards. This is easily remedied, though, simply by flipping the sign of the integral.

b. By Stokes' theorem,

$$\int_{\partial V} \sigma = \int_V d\sigma = 2 \int_V dx \wedge dy \wedge dz,$$

which is twice the volume. Hence the volume of the cone is $\pi ha^2/3$.

6.13 The disk Σ meets the sphere in a circle $\partial\Sigma$ of radius $a/2$. We parameterize the circle by

$$\gamma(t) = (a/2)(\cos t, \sin t, \sqrt{3}), \qquad 0 \le t \le 2\pi,$$

so

$$\gamma^*\omega = (a/2)^2[-\sin^2 t + \sqrt{3} \cos t]\, dt$$

and

$$\int_{\partial\Sigma} \omega = \int_I \gamma^*\omega = -(a/2)^2 \int_0^{2\pi} \sin^2 t\, dt = -a^2\pi/4.$$

On the other hand, the disk can be parameterized by

$$\sigma(u, v) = (u \cos v, u \sin v, \sqrt{3}a/2), \qquad 0 \le u \le a/2, 0 \le v \le 2\pi.$$

Also,

$$d\omega = dy \wedge dx + dz \wedge dy + dx \wedge dz,$$

so

$$\begin{aligned}\sigma^* d\omega &= (\sin v\, du + u \cos v\, dv) \wedge (\cos v\, du - u \sin v\, dv) \\ &= -u\, du \wedge dv.\end{aligned}$$

Note that the unit normal to the disk with this parameterization points upwards, which is consistent with the choice of direction for $\partial\Sigma$. Integrating over the disk gives

$$\int_\Sigma d\omega = \int_R \sigma^* d\omega = -\int_0^{2\pi} dv \int_0^{a/2} u\, du\, dv = -\pi a^2/4.$$

Again, Stokes' theorem is verified.

6.14 Let D be the closed disk in M containing p whose boundary is C, and let D^0 be its interior. Then $M - \{p\}$ is homotopic to $M - D^0$, and as homotopic maps induce the same map in cohomology, it suffices to show that $\iota^* : H^1_{\mathrm{dR}}(M - D^0) \to H^1(C)$ is the zero map, where now ι is the inclusion $\iota : C \to M - D^0$. But $M - D^0$ is a compact oriented surface with boundary C, so by Stokes' theorem, for any $\omega \in H^1_{\mathrm{dR}}(M - D^0)$,

$$\int_C \iota^*\omega = \int_{\partial(M-D^0)} \iota^*\omega = \int_{M-D^0} d\omega = 0,$$

because ω is closed. By de Rham's theorem, integration gives an isomorphism between $H^1_{\mathrm{dR}}(C)$ and $\mathbb{R}$, so $\iota^*\omega$ must be zero.

6.15 Let (U_i, φ_i) be an atlas for a manifold M with boundary. We must show that, given $p \in \partial M$ there is an open neighborhood $U \subset \partial M$ containing p that is homeomorphic to $\mathbb{R}^{n-1}$. Let $p \in U_i$. By definition, $X := \varphi_i(U_i \cap \partial M) \subseteq \partial\mathbb{H}^n \cong \mathbb{R}^{n-1}$. Pick $q \in X$. If $\varphi_i^{-1}(q) \notin \partial M$, then $\varphi_i^{-1}(q)$ has a neighborhood in M homeomorphic to $\mathbb{R}^n$. But φ_i is a homeomorphism, so q has a neighborhood in X homeomorphic to $\mathbb{R}^n$, a contradiction. We may therefore take $U = \varphi_i^{-1}(X) = U_i \cap \partial M$, which is clearly open in ∂M.

7

Vector bundles

7.1 By definition,

$$(\varphi_V \circ \varphi_U^{-1})(p, y_U) = (p, y_V) = (p, g_{VU}(p)y_U).$$

If $U = V$ then $\varphi_V = \varphi_U$, so we get

$$(p, y_U) = (p, g_{UU}(p)y_U).$$

for all y_U, which gives $g_{UU}(p) = id$.

Similarly,

$$(\varphi_U \circ \varphi_V^{-1})(p, y_V) = (p, g_{UV}(p)y_V) = (p, g_{UV}g_{VU}y_U).$$

But

$$(\varphi_U \circ \varphi_V^{-1})(p, y_V) = (\varphi_U \circ \varphi_V^{-1})(\varphi_V \circ \varphi_U^{-1})(p, y_U) = (p, y_U),$$

whereupon we conclude

$$g_{UV}(p)g_{VU}(p) = 1.$$

Lastly, if U, V, and W are three overlapping neighborhoods,

$$(\varphi_V \circ \varphi_U^{-1})(p, y_U) = (p, y_V),$$
$$(\varphi_W \circ \varphi_V^{-1})(p, y_V) = (p, y_W),$$
$$(\varphi_U \circ \varphi_W^{-1})(p, y_W) = (p, y_U),$$

where $y_V = g_{VU}y_U$, $y_W = g_{WV}y_V$ and $y_U = g_{UW}y_W$. Combining these equations together with the ones above yields

$$g_{UW}g_{WV}g_{VU} = 1.$$

7.2 Let $\varphi : E \to M \times Y$ be a global trivialization and choose $\varphi_U = \varphi_V = \varphi$. This immediately implies $g_{UV} = id$.

7.3 Let X be a vector field on G. The map $L_{g*} : T_eG \to T_gG$ is an isomorphism of vector spaces, so the bundle isomorphism we seek is afforded by the map

$$X_g \mapsto (g, L_{g^{-1}*}X_g).$$

7.4 By definition of the curvature operator,

$$R(fX, gY)hZ = \nabla_{fX}\nabla_{gY}hZ - \nabla_{gY}\nabla_{fX}hZ - \nabla_{[fX,gY]}hZ. \tag{1}$$

Expanding each term on the right side of (1) in succession gives the following set of equations.

$$\begin{aligned}\nabla_{fX}\nabla_{gY}hZ &= f\nabla_X(g\nabla_Y(hZ))\\ &= f\nabla_X(gY(h)Z + gh\nabla_Y Z)\\ &= fX(gY(h))Z + fgY(h)\nabla_X Z\\ &\qquad + fX(gh)\nabla_Y Z + fgh\nabla_X\nabla_Y Z.\end{aligned} \tag{2}$$

$$\begin{aligned}\nabla_{gY}\nabla_{fX}hZ &= g\nabla_Y(f\nabla_X(hZ))\\ &= g\nabla_Y(fX(h)Z + fh\nabla_X Z)\\ &= gY(fX(h))Z + fgX(h)\nabla_Y Z\\ &\qquad + gY(fh)\nabla_X Z + fgh\nabla_Y\nabla_X Z.\end{aligned} \tag{3}$$

$$\begin{aligned}\nabla_{[fX,gY]}hZ &= \nabla_{fX(g)Y+fgXY-gY(f)X-fgYX}hZ\\ &= fX(g)\nabla_Y hZ - gY(f)\nabla_X hZ + fg\nabla_{[X,Y]}hZ\\ &= fX(g)Y(h)Z + fX(g)h\nabla_Y Z\\ &\qquad - gY(f)X(h)Z - gY(f)h\nabla_X Z\\ &\qquad + fg([X,Y]h)Z + fgh\nabla_{[X,Y]}Z.\end{aligned} \tag{4}$$

Combining (1), (2), (3), and (4) gives

$$\begin{aligned}R(fX, gY)hZ = &\underbrace{fX(gY(h))Z}_{a} + \underbrace{fgY(h)\nabla_X Z}_{b}\\ &+ \underbrace{fX(gh)\nabla_Y Z}_{c} + \underbrace{fgh\nabla_X\nabla_Y Z}_{d}\\ &- \underbrace{gY(fX(h))Z}_{a} - \underbrace{fgX(h)\nabla_Y Z}_{c}\\ &- \underbrace{gY(fh)\nabla_X Z}_{b} - \underbrace{fgh\nabla_Y\nabla_X Z}_{d}\\ &- \underbrace{fX(g)Y(h)Z}_{a} - \underbrace{fX(g)h\nabla_Y Z}_{c}\end{aligned}$$

$$+ \underbrace{gY(f)X(h)Z}_{a} + \underbrace{gY(f)h\nabla_X Z}_{b}$$
$$- \underbrace{fg([X,Y]h)Z}_{a} - \underbrace{fgh\nabla_{[X,Y]}Z}_{d}. \tag{5}$$

Close examination of this expression reveals that all the terms labeled 'a', 'b', and 'c' cancel, while the terms labeled 'd' combine to yield the desired result:

$$R(fX, gY)hZ = fghR(X,Y)Z.$$

7.5 A point of the unit tangent bundle $T_1(S^2)$ is a pair (x, y), where $x \in S^2$ and $y \in T_x S^2$. Such a pair determines a unique orthonormal triple (x, y, z), where $z = x \times y$. Take these to be the column vectors of a matrix R. By orthonormality, $RR^T = 1$, so $R \in SO(3)$. Conversely, given an element $R \in SO(3)$ (viewed as a subgroup of $GL(3, \mathbb{R})$) the columns form an orthonormal triple, and therefore we may choose the first and second entries to be a point of $T_1(S^2)$. This shows that the map $T_1(S^2) \mapsto SO(3)$ given by $(x, y) \mapsto R$ is bijective, and continuity is clear.

7.6 The three vector fields are mutually orthogonal. (Indeed, on S^4 they are orthonormal.) In particular, they are clearly linearly independent. Moreover, they are nowhere vanishing on S^4, because $(0,0,0,0) \notin S^4$. Smoothness is obvious, so there is a global frame field and the tangent bundle is indeed trivial.

7.7 The projection map π just sends $(E \oplus F)_p$ to p. If φ is a local trivialization of $E \to M$ and ψ is a local trivialization of F, then (φ, ψ) is a local trivialization of $E \oplus F$ according to

$$(\varphi, \psi)_U(q) = (p, (v, w)),$$

where $\pi(q) = p$, $\varphi_U(q) = (p, v)$, and $\psi_U(q) = (p, w)$. We have

$$[(\varphi, \psi)_V \circ (\varphi, \psi)_U^{-1}](p, (x, y)) = (p, (x', y')),$$

where $x' = g_{E,UV}x$ and $y' = g_{F,UV}y$. In other words, the transition functions for the direct sum are $g_{E\oplus F,UV} = g_{E,UV} \oplus g_{F,UV}$, where $(A \oplus B)(x, y) := (Ax, By)$.

7.8 To show that D_t is a connection, we need only show that it obeys the two axioms (because it clearly maps sections of E to one form valued sections of E, since each piece of it does). The two axioms are (i) linearity (which is obvious, because a convex combination is a special kind of linear combination), and (ii) the Leibniz rule. To show (ii), let s be a section and f a function. Then, by linearity,

$$\begin{aligned}
D_t(sf) &= (1-t)D(sf) + t\widetilde{D}(sf) \\
&= (1-t)\{Ds \cdot f + s \otimes df\} + t\{\widetilde{D}s \cdot f + s \otimes df\} \\
&= D_t s \cdot f + s \otimes df.
\end{aligned}$$

(It is worth noting that this also shows that an arbitrary linear combination of connections is not necessarily a connnection.)

7.9 a. This follows immediately from the cyclicity of the trace and the properties of matrix multiplication. We have $\Omega' = A^{-1}\Omega A$, so $\Omega'^2 = (A^{-1}\Omega A)(A^{-1}\Omega A) = A^{-1}\Omega^2 A$. Inductively, we have $\Omega'^k = A^{-1}\Omega^k A$. By the cyclicity of the trace,

$$\operatorname{tr}\Omega'^k = \operatorname{tr}(A^{-1}\Omega^k A) = \operatorname{tr}(\Omega^k A A^{-1}) = \operatorname{tr}\Omega^k.$$

b. We have

$$\begin{aligned}
d\operatorname{tr}\Omega^k &= \operatorname{tr} d\Omega^k = \operatorname{tr} d(\Omega \wedge \cdots \wedge \Omega) && (1)\\
&= \operatorname{tr}(d\Omega \wedge \Omega \wedge \cdots \wedge \Omega + \Omega \wedge d\Omega \wedge \cdots \wedge \Omega \\
&\qquad + \cdots + \Omega \wedge \cdots \wedge d\Omega) && (2)\\
&= k\operatorname{tr}(d\Omega \wedge \Omega^{k-1}) && (3)\\
&= k\operatorname{tr}(\{\Omega \wedge \omega - \omega \wedge \Omega\} \wedge \Omega^{k-1}) && (4)\\
&= k\operatorname{tr}(\omega \wedge \Omega^k - \omega \wedge \Omega^k) && (5)\\
&= 0. && (6)
\end{aligned}$$

Equation (1) follows because d is linear, so it commutes with the trace. Equation (2) holds by virtue of the fact that Ω is a matrix of two forms: d is an antiderivation, so $d(\lambda \wedge \eta) = d\lambda \wedge \eta + (-1)^{(\deg \lambda)(\deg \eta)}\lambda \wedge d\eta$. In our case, the degree of each entry of Ω is two, so the sign is always positive. Equation (3) is a bit subtle. Consider, for example, the second term in (2). We have (with the usual implicit summation)

$$\begin{aligned}
\operatorname{tr}(\Omega \wedge d\Omega \wedge \cdots \wedge \Omega) &= \Omega^{i_1}{}_{i_2} \wedge d\Omega^{i_2}{}_{i_3} \wedge \Omega^{i_3}{}_{i_4} \wedge \cdots \wedge \Omega^{i_k}{}_{i_1} \\
&= d\Omega^{i_2}{}_{i_3} \wedge \Omega^{i_3}{}_{i_4} \wedge \cdots \wedge \Omega^{i_k}{}_{i_1} \wedge \Omega^{i_1}{}_{i_2} \\
&= \operatorname{tr}(d\Omega \wedge \Omega^{k-1}),
\end{aligned}$$

where the middle step follows from the antisymmetry properties of the wedge. (Each element $\Omega^{i_1}{}_{i_2}$ is a two form, so it commutes with all other forms.) Every other term in (2) can be brought to the same form, and there are k of them. Equation (4) follows from the Bianchi identity. Finally, Equation (5) uses the same idea that we used to obtain Equation (3).

c. By an argument similar to that given in Part (b), we have

$$\frac{d}{dt}\operatorname{tr}\Omega_t^k = k\operatorname{tr}\left(\frac{d\Omega_t}{dt}\wedge\Omega_t^{k-1}\right).$$

But

$$\begin{aligned}\Omega_t &= d\omega_t + \omega_t\wedge\omega_t\\ &= d(\omega+t\eta)+(\omega+t\eta)\wedge(\omega+t\eta)\\ &= \Omega + t(d\eta+\eta\wedge\omega+\omega\wedge\eta)+t^2\eta\wedge\eta,\end{aligned}$$

so

$$\frac{d\Omega_t}{dt} = d\eta+\eta\wedge\omega+\omega\wedge\eta+2t\eta\wedge\eta,$$

and

$$\frac{d}{dt}\operatorname{tr}\Omega_t^k = k\operatorname{tr}\left(\{d\eta+\eta\wedge\omega+\omega\wedge\eta+2t\eta\wedge\eta\}\wedge\Omega_t^{k-1}\right).$$

Under a change of frame field, the connection matrix transforms according to

$$\omega' = A^{-1}\omega A + A^{-1}dA$$

so

$$\eta' = \widetilde{\omega}' - \omega' = A^{-1}\widetilde{\omega}A - A^{-1}\omega A = A^{-1}\eta A,$$

showing that

$$\alpha = \operatorname{tr}(\eta\wedge\Omega_t^{k-1})$$

is a global form on M.
We have

$$\begin{aligned}d\alpha &= d\operatorname{tr}(\eta\wedge\Omega_t^{k-1})\\ &= \operatorname{tr}(d\eta\wedge\Omega_t^{k-1} - \eta\wedge d\Omega_t^{k-1}).\end{aligned}$$

The second term in this expression is

$$\begin{aligned}\operatorname{tr}(\eta\wedge d\Omega_t^{k-1}) = \operatorname{tr}(\eta\wedge\{&d\Omega_t\wedge\Omega_t\wedge\cdots\wedge\Omega_t\\ &+\Omega_t\wedge d\Omega_t\wedge\Omega_t\wedge\cdots\wedge\Omega_t\\ &+\Omega_t\wedge\cdots\wedge\Omega_t\wedge d\Omega_t\}).\end{aligned}$$

Using the Bianchi identity

$$d\Omega_t = \Omega_t\wedge\omega_t - \omega_t\wedge\Omega_t$$

gives

$$\begin{aligned}\operatorname{tr}(\eta \wedge d\Omega_t^{k-1}) &= \operatorname{tr}(\eta \wedge \{(\Omega_t \wedge \omega_t - \omega_t \wedge \Omega_t) \wedge \Omega_t \wedge \cdots \wedge \Omega_t \\ &\qquad + \Omega_t \wedge (\Omega_t \wedge \omega_t - \omega_t \wedge \Omega_t) \wedge \Omega_t \wedge \cdots \wedge \Omega_t \\ &\qquad + \Omega_t \wedge \cdots \wedge \Omega_t \wedge (\Omega_t \wedge \omega_t - \omega_t \wedge \Omega_t)\}) \\ &= \operatorname{tr}(-\eta \wedge \omega_t \wedge \Omega_t^{k-1} + \eta \wedge \Omega_t^{k-1} \wedge \omega_t) \\ &= -\operatorname{tr}(\{\eta \wedge \omega_t + \omega_t \wedge \eta\} \wedge \Omega_t^{k-1}).\end{aligned}$$

(The second equality follows because the sum is *telescoping*, meaning that one term in each line cancels with one term in the next line. The last equality follows by commuting forms and using the cyclicity of the trace.) Hence

$$\begin{aligned}d\alpha &= \operatorname{tr}(\{d\eta + \eta \wedge \omega_t + \omega_t \wedge \eta\} \wedge \Omega_t^{k-1}) \\ &= \operatorname{tr}(\{d\eta + \eta \wedge \omega + \omega \wedge \eta + 2t\eta \wedge \eta\} \wedge \Omega_t^{k-1}) \\ &= \frac{1}{k}\frac{d}{dt} \operatorname{tr} \Omega_t^k.\end{aligned}$$

Integrating both sides with respect to t gives

$$\operatorname{tr} \widetilde{\Omega}^k - \operatorname{tr} \Omega^2 = \int_0^1 \frac{d}{dt} \operatorname{tr} \Omega_t^k \, dt = k \int_0^1 d\alpha \, dt = d\left(k \int_0^1 \alpha \, dt\right),$$

thereby establishing the desired result.

7.10 a.

$$\begin{aligned}D_i(g\psi) &= \partial_i(g\psi) + A_i g\psi \\ &= (\partial_i g)\psi + g\partial_i\psi + gg^{-1}A_i g\psi \\ &= g\left(g^{-1}\partial_i g + g^{-1}A_i g\right)\psi + g\partial_i\psi \\ &= gD_i'\psi.\end{aligned}$$

b. Starting from $A = A_i\, dx^i$ we get

$$\begin{aligned}F &= dA + A \wedge A \\ &= \partial_j A_i \, dx^j \wedge dx^i + A_i A_j \, dx^i \wedge dx^j \\ &= \left(\partial_i A_j + A_i A_j\right) dx^i \wedge dx^j \\ &= \frac{1}{2}\left(\partial_i A_j - \partial_j A_i + [A_i, A_j]\right) dx^i \wedge dx^j.\end{aligned}$$

Also,

$$\begin{aligned}
[D_i, D_j]\psi &= [\partial_i + A_i, \partial_j + A_j]\psi \\
&= ([\partial_i, \partial_j] + [\partial_i, A_j] + [A_i, \partial_j] + [A_i, A_j])\psi \\
&= \partial_i(A_j\psi) - A_j\partial_i\psi + A_i\partial_j\psi - \partial_j(A_i\psi) + [A_i, A_j]\psi \\
&= (\partial_i A_j - \partial_j A_i + [A_i, A_j])\psi.
\end{aligned}$$

c. The Bianchi identity (7.20) reads

$$dF - F \wedge A + A \wedge F = 0.$$

In local coordinates (after multiplying by 2), this becomes

$$\left(\partial_i F_{jk} - F_{ij} A_k + A_i F_{jk}\right) dx^i \wedge dx^j \wedge dx^k = 0.$$

Note that, by virtue of the antisymmetry of the wedge products,

$$B_{ijk}\, dx^i \wedge dx^j \wedge dx^k = 0$$

holds if and only if $B_{[ijk]}$, the totally antisymmetric part of B, vanishes. Using the fact that F_{ij} is already antisymmetric, we get (dropping a factor of 3),

$$\begin{aligned}
0 &= \partial_{[i} F_{jk]} - F_{[ij} A_{k]} + A_{[i} F_{jk]} \\
&= \partial_i F_{jk} + \partial_j F_{ki} + \partial_k F_{ij} - F_{ij} A_k - F_{jk} A_i - F_{ki} A_j \\
&\qquad + A_i F_{jk} + A_j F_{ki} + A_k F_{ij} \\
&= \partial_i F_{jk} + [A_i, F_{jk}] + \text{cyclic}.
\end{aligned}$$

But

$$\begin{aligned}
[D_i, F_{jk}]\psi &= [\partial_i + A_i, F_{jk}]\psi \\
&= \partial_i(F_{jk}\psi) - F_{jk}\partial_i\psi + [A_i, F_{jk}]\psi \\
&= (\partial_i F_{jk} + [A_i, F_{jk}])\psi,
\end{aligned}$$

so the conclusion follows.

8

Geometric manifolds

8.1

$$\begin{aligned}\tau(fX, gY) &= \nabla_{fX}(gY) - \nabla_{gY}(fX) - [fX, gY] \\ &= f(Xg)Y + fg\nabla_X Y - g(Yf)X - gf\nabla_Y X \\ &\qquad - f(Xg)Y - fgXY + g(Yf)X + gfYX \\ &= fg\tau(X, Y).\end{aligned}$$

8.2 From the definition,

$$\begin{aligned}\tau^a(e_b, e_c) &= \theta^a(\tau(e_b, e_c)) \\ &= \theta^a(\nabla_{e_b} e_c - \nabla_{e_c} e_b - [e_b, e_c]) \\ &= \Gamma^a{}_{bc} - \Gamma^a{}_{cb} - \theta^a([e_b, e_c]) \\ &= \omega^a{}_c(e_b) - \omega^a{}_b(e_c) - \theta^a([e_b, e_c]) \\ &= (\omega^a{}_d \wedge \theta^d)(e_b, e_c) + d\theta^a(e_b, e_c).\end{aligned}$$

The last line follows from Exercise 3.15 and Equation 3.122. Specifically,

$$\begin{aligned}(\omega^a{}_d \wedge \theta^d)(e_b, e_c) &= \omega^a{}_d(e_b)\theta^d(e_c) - \omega^a{}_d(e_c)\theta^d(e_b) \\ &= \omega^a{}_d(e_b)\delta^d_c - \omega^a{}_d(e_c)\delta^d_b \\ &= \omega^a{}_c(e_b) - \omega^a{}_b(e_c),\end{aligned}$$

and

$$\begin{aligned}d\theta^a(e_b, e_c) &= e_b\theta^a(e_c) - e_c\theta^a(e_b) - \theta^a([e_b, e_c]) \\ &= e_b(\delta^a_c) - e_c(\delta^a_b) - \theta^a([e_b, e_c]) \\ &= -\theta^a([e_b, e_c]).\end{aligned}$$

8.3 By Properties C4 and C5 of Section 7.2,

$$0 = \nabla_{e_a}\left\langle e_b, \theta^c\right\rangle = \left\langle \nabla_{e_a} e_b, \theta^c\right\rangle + \left\langle e_b, \nabla_{e_a}\theta^c\right\rangle = \Gamma^c{}_{ab} + \left\langle e_b, \nabla_{e_a}\theta^c\right\rangle,$$

so that $\nabla_{e_a}\theta^c = -\Gamma^c{}_{ab}\theta^b$.

8.4 Under a coordinate transformation $x^i \to y^{i'}$ with $\partial_i := \partial/\partial x^i$ and $\partial_{i'} := \partial/\partial y^{i'}$ we have

$$\partial_{i'} = J^i{}_{i'}\partial_i \quad \text{and} \quad \partial_i = (J^{-1})^{i'}{}_i \partial_{i'},$$

where

$$J^i{}_{i'} = \frac{\partial x^i}{\partial y^{i'}} \quad \text{and} \quad (J^{-1})^{i'}{}_i = \frac{\partial y^{i'}}{\partial x^i}.$$

Hence

$$\begin{aligned}\Gamma^{k'}{}_{i'j'}\partial_{k'} &= \nabla_{\partial_{i'}}\partial_{j'} = J^i{}_{i'}\nabla_{\partial_i}(J^j{}_{j'}\partial_j) = J^i{}_{i'}J^j{}_{j'}\nabla_{\partial_i}\partial_j + J^i{}_{i'}(\partial_i J^j{}_{j'})\partial_j \\ &= J^i{}_{i'}J^j{}_{j'}\Gamma^k{}_{ij}\partial_k + J^i{}_{i'}((J^{-1})^{\ell'}{}_i\partial_{\ell'}J^j{}_{j'})(J^{-1})^{k'}{}_j\partial_{k'} \\ &= \left(J^i{}_{i'}J^j{}_{j'}\Gamma^k{}_{ij}(J^{-1})^{k'}{}_k + \delta^{\ell'}_{i'}(\partial_{\ell'}J^j{}_{j'})(J^{-1})^{k'}{}_j\right)\partial_{k'},\end{aligned}$$

whereupon we conclude

$$\Gamma^{k'}{}_{i'j'} = \frac{\partial y^{k'}}{\partial x^k}\frac{\partial x^i}{\partial y^{i'}}\frac{\partial x^j}{\partial y^{j'}}\Gamma^k{}_{ij} + \frac{\partial y^{k'}}{\partial x^j}\frac{\partial^2 x^j}{\partial y^{i'}\partial y^{j'}}.$$

8.5 The Christoffel symbols are given by

$$\Gamma^m{}_{ij} = g^{mk}\Gamma_{kij},$$

where

$$\Gamma_{kij} = \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right)$$

and g^{mk} are the components of the inverse metric.

Let the index '1' be 'θ' and the index '2' be 'ϕ'. Then, written as a matrix, the metric tensor components are

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix},$$

so the inverse metric components are

$$g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \csc^2\theta \end{pmatrix}.$$

Now we compute, using the fact that the off diagonal terms of the metric vanish,

$$\begin{aligned}\Gamma_{\theta\theta\theta} &= \frac{1}{2}(\partial_\theta g_{\theta\theta}) = 0, \\ \Gamma_{\theta\theta\phi} &= \frac{1}{2}(\partial_\phi g_{\theta\theta}) = 0,\end{aligned}$$

$$\begin{aligned}
\Gamma_{\theta\phi\theta} &= \Gamma_{\theta\theta\phi} = 0,\\
\Gamma_{\theta\phi\phi} &= -\frac{1}{2}(\partial_\theta g_{\phi\phi}) = -\sin\theta\cos\theta,\\
\Gamma_{\phi\theta\theta} &= -\frac{1}{2}(\partial_\phi g_{\theta\theta}) = 0,\\
\Gamma_{\phi\theta\phi} &= \frac{1}{2}(\partial_\theta g_{\phi\phi}) = \sin\theta\cos\theta,\\
\Gamma_{\phi\phi\theta} &= \Gamma_{\phi\theta\phi} = \sin\theta\cos\theta,\\
\Gamma_{\phi\phi\phi} &= \frac{1}{2}(\partial_\phi g_{\phi\phi}) = 0.
\end{aligned}$$

Raising indices with the inverse metric tensor we get

$$\begin{aligned}
\Gamma^\theta{}_{\theta\theta} &= g^{\theta\theta}\Gamma_{\theta\theta\theta} + g^{\theta\phi}\Gamma_{\phi\theta\theta} = 0,\\
\Gamma^\theta{}_{\theta\phi} &= \Gamma^\theta{}_{\phi\theta} = g^{\theta\theta}\Gamma_{\theta\theta\phi} + g^{\theta\phi}\Gamma_{\phi\theta\phi} = 0,\\
\Gamma^\theta{}_{\phi\phi} &= g^{\theta\theta}\Gamma_{\theta\phi\phi} + g^{\theta\phi}\Gamma_{\phi\phi\phi} = -\sin\theta\cos\theta,\\
\Gamma^\phi{}_{\theta\theta} &= g^{\phi\theta}\Gamma_{\theta\theta\theta} + g^{\phi\phi}\Gamma_{\phi\theta\theta} = 0,\\
\Gamma^\phi{}_{\theta\phi} &= \Gamma^\phi{}_{\phi\theta} = g^{\phi\theta}\Gamma_{\theta\theta\phi} + g^{\phi\phi}\Gamma_{\phi\theta\phi} = \cot\theta,\\
\Gamma^\phi{}_{\phi\phi} &= g^{\phi\theta}\Gamma_{\theta\phi\phi} + g^{\phi\phi}\Gamma_{\phi\phi\phi} = 0.
\end{aligned}$$

8.6 We have $g(e_{\hat b}, e_{\hat c}) = \pm\delta_{\hat b\hat c}$, so by metric compatibility,

$$\begin{aligned}
0 = e_{\hat a} g(e_{\hat b}, e_{\hat c}) &= \nabla_{e_{\hat a}} g(e_{\hat b}, e_{\hat c})\\
&= g(\nabla_{e_{\hat a}} e_{\hat b}, e_{\hat c}) + g(e_{\hat b}, \nabla_{e_{\hat a}} e_{\hat c})\\
&= g(\Gamma^{\hat d}{}_{\hat a\hat b} e_{\hat d}, e_{\hat c}) + g(e_{\hat b}, \Gamma^{\hat d}{}_{\hat a\hat c} e_{\hat d})\\
&= \Gamma^{\hat d}{}_{\hat a\hat b} g(e_{\hat d}, e_{\hat c}) + \Gamma^{\hat d}{}_{\hat a\hat c} g(e_{\hat b}, e_{\hat d})\\
&= \omega^{\hat d}{}_{\hat b}(e_{\hat a}) g_{\hat d\hat c} + \omega^{\hat d}{}_{\hat c}(e_{\hat a}) g_{\hat b\hat d}\\
&= (\omega_{\hat c\hat b} + \omega_{\hat b\hat c})(e_{\hat a}).
\end{aligned}$$

As $e_{\hat a}$ is a basis, the conclusion follows.

The curvature matrix components are

$$\Omega^{\hat\ell}{}_{\hat n} = d\omega^{\hat\ell}{}_{\hat n} + \omega^{\hat\ell}{}_{\hat a} \wedge \omega^{\hat a}{}_{\hat n}.$$

The downstairs components are $\Omega_{\hat m\hat n} = g_{\hat\ell\hat m}\Omega^{\hat\ell}{}_{\hat n}$, so

$$\begin{aligned}
\Omega_{\hat m\hat n} &= d\omega_{\hat m\hat n} + \omega_{\hat m\hat a} \wedge \omega^{\hat a}{}_{\hat n}\\
&= d\omega_{\hat m\hat n} + \omega_{\hat m\hat a} \wedge \omega_{\hat b\hat n} g^{\hat a\hat b}.
\end{aligned}$$

Hence

$$\begin{aligned}\Omega_{\hat{n}\hat{m}} &= d\omega_{\hat{n}\hat{m}} + \omega_{\hat{n}\hat{a}} \wedge \omega_{\hat{b}\hat{m}} g^{\hat{a}\hat{b}} \\ &= -d\omega_{\hat{m}\hat{n}} + \omega_{\hat{a}\hat{n}} \wedge \omega_{\hat{m}\hat{b}} g^{\hat{a}\hat{b}} \\ &= -d\omega_{\hat{m}\hat{n}} - \omega_{\hat{m}\hat{b}} \wedge \omega_{\hat{a}\hat{n}} g^{\hat{a}\hat{b}} \\ &= -d\omega_{\hat{m}\hat{n}} - \omega_{\hat{m}\hat{a}} \wedge \omega^{\hat{a}}{}_{\hat{n}} \\ &= -\Omega_{\hat{m}\hat{n}}.\end{aligned}$$

8.7 We have $dA = A\omega$ and $dA^T = (dA)^T = \omega^T A^T = -\omega A^T$, so

$$d(AA^T) = dA \cdot A^T + A \cdot dA^T = A\omega A^T - A\omega A^T = 0.$$

Hence $AA^T =$ constant. But $AA^T(p) = I$, so $AA^T = I$ everywhere, i.e., A is orthogonal.

8.8 Equation (7.34) says that if $s = e\sigma$ is a section, then

$$R(X, Y) \circ s = e\Omega(X, Y)\sigma.$$

Let $X = e_c$, $Y = e_d$, and $s = Z = e_b Z^b$. Then

$$Z^b R(e_c, e_d) e_b = e_a \Omega^a{}_b(e_c, e_d) Z^b \quad \Rightarrow \quad \Omega^a{}_b(e_c, e_d) = R^a{}_{bcd},$$

from which the result follows. (Okay, maybe this isn't immediately obvious. The point is that $\Omega^a{}_b$ is a matrix of two forms, so it can be written

$$\Omega^a{}_b = \Omega^a{}_{bef}\, \theta^e \wedge \theta^f$$

for some coefficients $\Omega^a{}_{bef}$ which are necessarily antisymmetric in the last two indices. Then taking inner products gives

$$\begin{aligned}\Omega^a{}_b(e_c, e_d) = (\Omega^a{}_{bef}\, \theta^e \wedge \theta^f)(e_c, e_d) &= \Omega^a{}_{bef}(\theta^e(e_c)\theta^f(e_d) - \theta^e(e_d)\theta^f(e_c)) \\ &= \Omega^a{}_{bef}(\delta^e_c \delta^f_d - \delta^e_d \delta^f_c) = 2\Omega^a{}_{bcd}.\end{aligned}$$

Now perhaps the conclusion is a bit more apparent.)

8.9 Equation (7.36) gives

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

Choose $X = \partial_k$, $Y = \partial_\ell$, and $Z = \partial_j$. Then $[X, Y] = 0$, so we get

$$\begin{aligned}R(\partial_k, \partial_\ell)\partial_j &= \nabla_{\partial_k} \nabla_{\partial_\ell} \partial_j - (k \leftrightarrow \ell) \\ &= \nabla_{\partial_k}(\Gamma^m{}_{\ell j} \partial_m) - (k \leftrightarrow \ell) \\ &= \Gamma^m{}_{\ell j,k} \partial_m + \Gamma^m{}_{\ell j} \Gamma^n{}_{km} \partial_n - (k \leftrightarrow \ell) \\ &= \left(\Gamma^i{}_{\ell j,k} + \Gamma^i{}_{km} \Gamma^m{}_{\ell j}\right) \partial_i - (k \leftrightarrow \ell),\end{aligned}$$

which yields the desired result.

8.10 Start with the following four equations, each of which follows from (8.49):

$$
\begin{aligned}
g(W, R(X, Y)Z) + g(W, R(Y, Z)X) + g(W, R(Z, X)Y) &= 0, \\
g(X, R(Y, Z)W) + g(X, R(Z, W)Y) + g(X, R(W, Y)Z) &= 0, \\
g(Y, R(Z, W)X) + g(Y, R(W, X)Z) + g(Y, R(X, Z)W) &= 0, \\
g(Z, R(W, X)Y) + g(Z, R(X, Y)W) + g(Z, R(Y, W)X) &= 0.
\end{aligned}
$$

Adding the first two and subtracting the second two and applying (8.47) and (8.48) liberally gives

$$
\begin{aligned}
0 = {} & g(W, R(X, Y)Z) + \cancelto{a}{g(W, R(Y, Z)X)} + \cancelto{b}{g(W, R(Z, X)Y)} \\
& + \cancelto{a}{g(X, R(Y, Z)W)} + g(X, R(Z, W)Y) + \cancelto{c}{g(X, R(W, Y)Z)} \\
& - g(Y, R(Z, W)X) - \cancelto{d}{g(Y, R(W, X)Z)} - \cancelto{b}{g(Y, R(X, Z)W)} \\
& - \cancelto{d}{g(Z, R(W, X)Y)} - g(Z, R(X, Y)W) - \cancelto{c}{g(Z, R(Y, W)X)},
\end{aligned}
$$

where terms with the same label cancel. This leaves

$$0 = 2g(W, R(X, Y)Z) - 2g(X, R(W, Z)Y),$$

as desired.

8.11 Begin with (8.49) and take inner products to get

$$g(W, R(X, Y)Z + R(Y, Z)X + R(Z, X)Y) = 0.$$

According to equation (8.44)

$$g(\partial_i, R(\partial_k, \partial_\ell)\partial_j) = R_{ijk\ell},$$

so (8.49) is equivalent to the condition

$$R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk} = 0. \tag{1}$$

On the other hand, if we expand out the condition

$$R_{i[jk\ell]} = 0 \tag{2}$$

we get

$$R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk} - (R_{ij\ell k} + R_{ikj\ell} + R_{i\ell kj}) = 0,$$

or

$$R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk} = R_{ij\ell k} + R_{ikj\ell} + R_{i\ell kj}. \tag{3}$$

But by antisymmetry in the last two indices, we must have

$$R_{ij\ell k} + R_{ikj\ell} + R_{i\ell kj} = -(R_{ijk\ell} + R_{ik\ell j} + R_{i\ell jk}) \tag{4}$$

as well. Equations (3) and (4) together imply (1). It is obvious that (1) implies (2), so (1) and (2) are equivalent constraints.

8.12 Property 3 by itself implies that the last three indices of the Riemann tensor must be distinct. Naively, then, Property 3 imposes $n\binom{n-1}{3} = 4\binom{n}{4}$ constraints. But in fact it imposes fewer than that, because by Properties 1 and 4, the first index cannot equal any of the other three. For example, if we were to choose '1' for the first index and '234' for the last three indices, Property 3 would give

$$R_{1123} + R_{1231} + R_{1312} = 0.$$

But by Property 1, the first term vanishes, and the last equals $-R_{3112}$, so the equation would become

$$R_{1231} - R_{3112} = 0,$$

which is already true by Property 4. It follows that all four indices must be distinct, which explains why Property 3 imposes only $\binom{n}{4}$ constraints. (All these constraints are independent of the ones imposed by the other Properties, because each one reduces to a condition on the cyclic permutation of the last three indices, none of which follow from conditions on pairs of indices (or pairs of pairs).)

A symmetric N by N matrix has $N(N+1)/2$ independent components (where $N = (n^2 - n)/2$), so altogether there are

$$\begin{aligned}\frac{1}{2}(N^2 + N) - \binom{n}{4} &= \frac{1}{2}\left(\frac{1}{4}(n^2 - n)^2 + \frac{1}{2}(n^2 - n)\right) \\ &\qquad - \frac{n(n-1)(n-2)(n-3)}{4!} \\ &= \frac{1}{12}(n^4 - n^2)\end{aligned}$$

independent components.

8.13 By antisymmetry, the last two components of the Riemann tensor cannot be equal, so the only independent components are of the form $R^i{}_{j\theta\phi}$. We have

$$R^\theta{}_{\theta\theta\phi} = g^{\theta\theta} R_{\theta\theta\theta\phi} + g^{\theta\phi} R_{\phi\theta\theta\phi} = 0,$$

where the first term vanishes by virtue of the antisymmetry of the first two components, and the second term vanishes because there are no off-diagonal terms in the inverse metric. By similar reasoning we get

$$R^{\theta}{}_{\phi\theta\phi} = g^{\theta\theta} R_{\theta\phi\theta\phi} + g^{\theta\phi} R_{\phi\phi\theta\phi} = R_{\theta\phi\theta\phi},$$
$$R^{\phi}{}_{\theta\phi\theta} = g^{\phi\theta} R_{\theta\theta\phi\theta} + g^{\phi\phi} R_{\phi\theta\phi\theta} = \csc^2\theta\, R_{\theta\phi\theta\phi},$$

and

$$R^{\phi}{}_{\phi\phi\theta} = g^{\phi\theta} R_{\theta\phi\phi\theta} + g^{\phi\phi} R_{\phi\phi\phi\theta} = 0.$$

Therefore the only nonvanishing components of the Riemann tensor are $R^{\theta}{}_{\phi\theta\phi}$ and $R^{\phi}{}_{\theta\phi\theta}$, as advertised. Using Equation (8.46) and the Christoffel symbols given in Exercise 8.5 we get

$$\begin{aligned} R^{\phi}{}_{\theta\phi\theta} &= \Gamma^{\phi}{}_{\theta\theta,\phi} - \Gamma^{\phi}{}_{\phi\theta,\theta} + \Gamma^{\phi}{}_{\phi m}\Gamma^{m}{}_{\theta\theta} - \Gamma^{\phi}{}_{\theta m}\Gamma^{m}{}_{\phi\theta} \\ &= \csc^2\theta - \cot^2\theta = 1. \end{aligned}$$

It follows that

$$R^{\theta}{}_{\phi\theta\phi} = R_{\theta\phi\theta\phi} = \sin^2\theta\, R^{\phi}{}_{\theta\phi\theta} = \sin^2\theta.$$

8.14 By symmetry Property 4,

$$R_{ji} = g^{k\ell} R_{\ell jki} = g^{k\ell} R_{ki\ell j} = R^{\ell}{}_{i\ell j} = R_{ij}.$$

8.15

$$\begin{aligned} R_{\theta\theta} &= R^{k}{}_{\theta k\theta} = R^{\theta}{}_{\theta\theta\theta} + R^{\phi}{}_{\theta\phi\theta} = 1, \\ R_{\theta\phi} = R_{\phi\theta} &= R^{k}{}_{\theta k\phi} = R^{\theta}{}_{\theta\theta\phi} + R^{\phi}{}_{\theta\phi\phi} = 0, \\ R_{\phi\phi} &= R^{k}{}_{\phi k\phi} = R^{\theta}{}_{\phi\theta\phi} + R^{\phi}{}_{\phi\phi\phi} = \sin^2\theta. \end{aligned}$$

8.16 Fix $X = \partial_i$ and $Y = \partial_j$. The linear map T, given by

$$Z \mapsto R(Z, \partial_j)\partial_i$$

is then represented by the matrix $\boldsymbol{T}$, where

$$T\partial_k = T^{m}{}_{k}\partial_m = R(\partial_k, \partial_j)\partial_i = R^{m}{}_{ikj}\partial_m,$$

whence we conclude that the trace of T is

$$T^{m}{}_{m} = R^{m}{}_{imj} = R_{ij}.$$

8.17

$$R = g^{ij} R_{ij} = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = 1 \cdot 1 + \csc^2\theta \cdot \sin^2\theta = 2.$$

8.18 If $s = s(t)$ is another parameterization of γ, we have

$$\frac{d\gamma^i}{ds}\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^{k}{}_{ij} Y^j\right) = \frac{dt}{ds}\frac{d\gamma^i}{dt}\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^{k}{}_{ij} Y^j\right) = 0.$$

8.19 If Y and Z are both parallel translated along X, then $\nabla_X Y = \nabla_X Z = 0$. So the metric compatibility equation yields

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = 0,$$

which says that $g(Y, Z)$ is constant along the curve. Substituting Y for Z or Z for Y shows that the lengths of vectors are also preserved, so the angle between Y and Z is preserved.

8.20 Let $s = s(t)$ be a reparameterization. Then

$$\frac{dx^k}{ds} = \frac{dx^k}{dt}\frac{dt}{ds}$$

and

$$\frac{d^2x^k}{ds^2} = \frac{d^2x^k}{dt^2}\left(\frac{dt}{ds}\right)^2 + \frac{dx^k}{dt}\frac{d^2t}{ds^2}.$$

Thus,

$$\frac{d^2x^k}{ds^2} + \Gamma^k{}_{ij}\frac{dx^i}{ds}\frac{dx^j}{ds} = \left(\frac{d^2x^k}{dt^2} + \Gamma^k{}_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)\left(\frac{dt}{ds}\right)^2 + \frac{dx^k}{dt}\frac{d^2t}{ds^2}.$$

In order for the curve to be a geodesic with respect to both the s and t parameters simultaneously we must have $d^2t/ds^2 = 0$, which means that t is a linear function of s (and *vice versa*).

8.21 We basically solved this one already. If $X = \gamma_*(d/dt)$ is the tangent vector along the geodesic, then from Exercise 8.19 and the fact that X is parallel translated along itself, $\nabla_X g(X, X) = 0$, which is the equation above.

8.22 This is best illustrated by example, say $n = 3$. The domain of integration in the path ordered integral is a cube of side t that can be broken up into six (3!) chambers satisfying $t_1 > t_2 > t_3$, $t_1 > t_3 > t_2$, $t_2 > t_1 > t_3$, $t_2 > t_3 > t_1$, $t_3 > t_1 > t_2$, and $t_3 > t_2 > t_1$. (We can disregard the cases of equality, which are sets of measure zero.) Thus we get

$$\begin{aligned}
&\int_0^t dt_1 \int_0^t dt_2 \int_0^t dt_3\, \mathcal{P}(A(t_1)A(t_2)A(t_3)) \\
&\qquad = \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3\, A(t_1)A(t_2)A(t_3) \\
&\qquad\quad + \int_0^t dt_1 \int_0^{t_1} dt_3 \int_0^{t_3} dt_2\, A(t_1)A(t_3)A(t_2) \\
&\qquad\quad + \int_0^t dt_2 \int_0^{t_2} dt_1 \int_0^{t_1} dt_3\, A(t_2)A(t_1)A(t_3) \\
&\qquad\quad + \int_0^t dt_2 \int_0^{t_2} dt_3 \int_0^{t_3} dt_1\, A(t_2)A(t_3)A(t_1)
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^t dt_3 \int_0^{t_3} dt_1 \int_0^{t_1} dt_2\, A(t_3)A(t_1)A(t_2) \\
&+\int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1\, A(t_3)A(t_2)A(t_1) \\
&= 3! \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3\, A(t_1)A(t_2)A(t_3).
\end{aligned}
$$

The general case is similar.

8.23 From (8.84) we have

$$\vartheta(\gamma, \nabla)\vartheta(\epsilon_p, \nabla) = \vartheta(\epsilon_p, \nabla)\vartheta(\gamma, \nabla) = \vartheta(\gamma, \nabla)$$

so $\vartheta(\epsilon_p, \nabla) = id$.

For the second claim, if we reverse the direction of the curve we reverse the tangent vector, so $A \to -A$. We need to show that $\mathcal{P}e^{\int_0^t A(t)\,dt} = (\mathcal{P}e^{-\int_0^t A(t)\,dt})^{-1}$. It suffices to do this for an infinitessimal parameter distance, because every curve can be built from such pieces. So, let $t = \delta$ be a small quantity. Starting from the equation

$$Y(t) = Y(0) - \int_0^t dt_1 A(t_1)Y(t_1)$$

and using Taylor's theorem and keeping only terms of first order in δ we get

$$Y(\delta) \approx (1 - \delta A(0))Y(0).$$

Again by Taylor, $(1 - \delta A(0))^{-1} \approx 1 + \delta A(0)$, so

$$Y(0) \approx (1 + \delta A(0))Y(\delta),$$

which shows that reversing the sign of A carries us backwards from $Y(\delta)$ to $Y(0)$.

8.24 Let Γ be a loop based at q. Then $\gamma\Gamma\gamma^{-1}$ is a loop based at p, so

$$\vartheta(\gamma^{-1}, \nabla)\vartheta(\Gamma, \nabla)\vartheta(\gamma, \nabla) \in H(\nabla; p),$$

and therefore,

$$\vartheta(\Gamma, \nabla) \in \vartheta(\gamma, \nabla)H(\nabla; p)\vartheta(\gamma^{-1}, \nabla).$$

A similar argument yields the reverse inclusion.

8.25 Let $y^1, \dots, y^n$ be local coordinates on a neighborhood V that meets U. On the overlap we have

$$
\begin{aligned}
\sigma(x) &= \sqrt{G(x)}\, dx^1 \wedge \cdots \wedge dx^n \\
&= \sqrt{J^2 G'(y)} J^{-1}\, dy^1 \wedge \cdots \wedge dy^n \\
&= \sqrt{G'(y)}\, dy^1 \wedge \cdots \wedge dy^n,
\end{aligned}
$$

which shows that the definition of σ can be consistently extended across overlaps to all of M. (We assumed $J > 0$, which we may do by virtue of the fact that M is orientable.)

8.26 On U write $dx^i = A^i{}_j\theta^j$ for some matrix A of smooth functions. Then at any point of U,

$$g^{ij} = g(dx^i, dx^j) = A^i{}_k A^j{}_\ell g(\theta^k, \theta^\ell) = A^i{}_k A^j{}_\ell \delta^{k\ell} = (AA^T)^{ij}.$$

Taking determinants of both sides gives $G^{-1} = (\det A)^2$. Therefore

$$\begin{aligned}\sigma &= \sqrt{G}\, dx^1 \wedge \cdots \wedge dx^n \\ &= (\det A)^{-1}(\det A)\theta^1 \wedge \cdots \wedge \theta^n \\ &= \theta^1 \wedge \cdots \wedge \theta^n.\end{aligned}$$

8.27 This follows immediately from the fact that $\star^2 = \pm 1$ and $d^2 = 0$.

8.28 If f is a harmonic function it must be closed, so $df = 0$, which means f is constant on connected components. Now let $\omega = f\sigma$ be a harmonic top dimensional form. Assuming $\star\Delta = \Delta\star$ we get $0 = \Delta \star \omega = \Delta f$ (because $\star\sigma = 1$), so f is harmonic and therefore constant. To prove the assumption, let η be a k-form. Then using the fact that $\star^2 = (-1)^{k(n-k)}$ on a k-form (in positive definite signature), we find

$$\begin{aligned}\star(d\delta + \delta d)\eta &= (-1)^{nk+n+1} \star d \star d \star \eta + (-1)^{nk+2n+1}(-1)^{k(n-k)} d \star d\eta \\ &= (-1)^{nk+n+1} \star d \star d \star \eta + (-1)^{k^2+1} d \star d\eta\end{aligned}$$

and

$$\begin{aligned}(d\delta + \delta d)\star\eta &= (-1)^{n(n-k)+n+1}(-1)^{k(n-k)} d\star d\eta + (-1)^{n(n-k+1)+n+1} \star d\star d\star\eta \\ &= (-1)^{k^2+1} d \star d\eta + (-1)^{nk+n+1} \star d \star d \star \eta.\end{aligned}$$

8.29 Let $\omega = \eta + d\mu$ and $\pi = \lambda + d\nu$. Then (recalling that η and λ are closed),

$$\begin{aligned}([\omega], [\pi]) &= \int_M (\eta + d\mu) \wedge (\lambda + d\nu) = \int_M \eta\wedge\lambda + \eta\wedge d\nu + d\mu \wedge \lambda + d\mu\wedge d\nu \\ &= \int_M \eta \wedge \lambda + \int_M d((-1)^k \eta \wedge \nu + \mu \wedge \lambda + \mu \wedge d\nu) \\ &= ([\eta], [\lambda]),\end{aligned}$$

where the second integral in the penultimate line vanishes by Stokes' theorem, because M has no boundary.

8.30 By Poincaré duality, $\beta_n = \beta_0 = 1$, so M cannot be contractible.

8.31 In a coordinate basis, we have

$$T = T^i{}_j \partial_i \otimes dx^j.$$

Thus, from (8.33), (8.36), and Property (C6) of the covariant derivative operator,

$$\begin{aligned}\nabla_k T &= \partial_k({T^i}_j)\partial_i \otimes dx^j + {T^i}_j \nabla_k(\partial_i) \otimes dx^j + {T^i}_j \partial_i \otimes \nabla_k(dx^j)\\ &= {T^i}_{j,k}\partial_i \otimes dx^j + {T^i}_j {\Gamma^\ell}_{ki}\partial_\ell \otimes dx^j - {T^i}_j \partial_i \otimes {\Gamma^j}_{km}\, dx^m\\ &= \left({T^i}_{j,k} + {\Gamma^i}_{k\ell}{T^\ell}_j - {\Gamma^\ell}_{kj}{T^i}_\ell\right)\partial_i \otimes dx^j.\end{aligned}$$

The general case is similar.

8.32

$$\begin{aligned}&{T^{i_1\dots i_r}}_{j_1\dots j_s;k}X^k\\ &\qquad - {T^{ki_2\dots i_r}}_{j_1\dots j_s}{X^{i_1}}_{;k} - {T^{i_1 k\dots i_r}}_{j_1\dots j_s}{X^{i_2}}_{;k} - \cdots - {T^{i_1\dots i_{r-1}k}}_{j_1\dots j_s}{X^{i_r}}_{;k}\\ &\qquad + {T^{i_1\dots i_r}}_{kj_2\dots j_s}{X^k}_{;j_1} + {T^{i_1\dots i_r}}_{j_1 k\dots j_s}{X^k}_{;j_2} + \cdots + {T^{i_1\dots i_r}}_{j_1\dots j_{s-1}k}{X^k}_{;j_s}\\ &= (\mathcal{L}_X T)^{i_1\dots i_r}{}_{j_1\dots j_s}\\ &\qquad + \big({\Gamma^{i_1}}_{k\ell}{T^{\ell i_2\dots i_r}}_{j_1\dots j_s} + {\Gamma^{i_2}}_{k\ell}{T^{i_1\ell\dots i_r}}_{j_1\dots j_s} + \cdots + {\Gamma^{i_r}}_{k\ell}{T^{i_1\dots i_{r-1}\ell}}_{j_1\dots j_s}\\ &\qquad - {\Gamma^\ell}_{kj_1}{T^{i_1\dots i_r}}_{\ell j_2\dots j_s} - {\Gamma^\ell}_{kj_2}{T^{i_1\dots i_r}}_{j_1\ell\dots j_s} + \cdots - {\Gamma^\ell}_{kj_s}{T^{i_1\dots i_r}}_{j_1\dots j_{s-1}\ell}\big)X^k\\ &\qquad - \big({T^{ki_2\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_1}}_{k\ell} - {T^{i_1k\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_2}}_{k\ell} - \cdots - {T^{i_1\dots i_{r-1}k}}_{j_1\dots j_s}{\Gamma^{i_r}}_{k\ell}\\ &\qquad + {T^{i_1\dots i_r}}_{kj_2\dots j_s}{\Gamma^k}_{\ell j_1} + {T^{i_1\dots i_r}}_{j_1k\dots j_s}{\Gamma^k}_{\ell j_2} + \cdots + {T^{i_1\dots i_r}}_{j_1\dots j_{s-1}k}{\Gamma^k}_{\ell j_s}\big)X^\ell\\ &= (\mathcal{L}_X T)^{i_1\dots i_r}{}_{j_1\dots j_s}.\end{aligned}$$

In this computation we used the fact that the Christoffel symbols are symmetric in the two lower indices so all the terms in parentheses cancel. For example, the term

$${\Gamma^{i_1}}_{k\ell}{T^{\ell i_2\dots i_r}}_{j_1\dots j_s}X^k$$

cancels with the term

$${T^{ki_2\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_1}}_{k\ell}X^\ell$$

because changing dummy indices then switching indices in the Christoffel symbol gives

$${T^{ki_2\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_1}}_{k\ell}X^\ell = {T^{\ell i_2\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_1}}_{\ell k}X^k = {T^{\ell i_2\dots i_r}}_{j_1\dots j_s}{\Gamma^{i_1}}_{k\ell}X^k.$$

8.33 From equation (8.103) we get

$$g_{ij;k} = g_{ij,k} - {\Gamma^\ell}_{ki}g_{\ell j} - {\Gamma^\ell}_{kj}g_{i\ell}.$$

But the right hand side vanishes by virtue of the metric compatibility equation

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

with $X = \partial_k$, $Y = \partial_i$ and $Z = \partial_j$:

$$g_{ij,k} = g({\Gamma^\ell}_{ki}\partial_\ell, \partial_j) + g(\partial_i, {\Gamma^\ell}_{kj}\partial_\ell) = {\Gamma^\ell}_{ki}g_{\ell j} + {\Gamma^\ell}_{kj}g_{i\ell}.$$

8.34 There are a few ways to do this. One way would be to use (8.103) to write out the covariant derivatives on both sides, then multiply on the left by the metric and verify explicitly that both sides are the same. An easier way is just to note that the covariant derivative obeys a tensorial Leibniz rule and respects the dual pairing, so

$$T_{mj;\ell} = (g_{im}T^i{}_j)_{;\ell} = g_{im;\ell}T^i{}_j + g_{im}T^i{}_{j;\ell}.$$

But $g_{im;\ell} = 0$ by covariant constancy of the metric, so the result follows.

8.35 In a coordinate basis, the Bianchi identity (7.20) can be written

$$d\Omega^i{}_j - \Omega^i{}_p \wedge \omega^p{}_j + \omega^i{}_p \wedge \Omega^p{}_j = 0.$$

Substituting into this equation using (8.35) and (8.45) (and dropping an irrelevant factor of 1/2) gives

$$\begin{aligned}0 &= d(R^i{}_{jk\ell}\, dx^k \wedge dx^\ell) \\ &\quad - (R^i{}_{pmn}\, dx^m \wedge dx^n) \wedge \Gamma^p{}_{qj}\, dx^q + \Gamma^i{}_{qp}\, dx^q \wedge (R^p{}_{jn\ell}\, dx^n \wedge dx^\ell), \\ &= \left(R^i{}_{jk\ell,m} - R^i{}_{pmk}\Gamma^p{}_{\ell j} + R^p{}_{jk\ell}\Gamma^i{}_{mp}\right) dx^m \wedge dx^k \wedge dx^\ell, \\ &= R^i{}_{jk\ell;m}\, dx^m \wedge dx^k \wedge dx^\ell,\end{aligned}$$

from which $R^i{}_{j[k\ell;m]} = 0$ follows. The last equation above holds by virtue of the symmetry of the lower two Christoffel indices. Specifically, (8.103) yields

$$R^i{}_{jk\ell;m} = R^i{}_{jk\ell,m} + \Gamma^i{}_{mp}R^p{}_{jk\ell} - \Gamma^p{}_{mj}R^i{}_{pk\ell} - \Gamma^p{}_{mk}R^i{}_{jp\ell} - \Gamma^p{}_{m\ell}R^i{}_{jkp}.$$

Antisymmetrizing this expression on k, ℓ, and m kills the last two terms and leaves

$$R^i{}_{j[k\ell;m]} = R^i{}_{j[k\ell,m]} + \Gamma^i{}_{[m|p}R^p{}_{j|k\ell]} - \Gamma^p{}_{[m|j}R^i{}_{p|k\ell]},$$

where the indices between the straight brackets are not antisymmetrized. But

$$\Gamma^p{}_{[\ell|j}R^i{}_{p|mk]} = \Gamma^p{}_{[m|j}R^i{}_{p|k\ell]}$$

by cyclic permutation of indices.

The other expression is equivalent, because

$$\nabla_m \{[\nabla_k, \nabla_\ell]\partial_i\} = \nabla_m \left\{R^i{}_{jk\ell}\partial_i\right\} = R^i{}_{jk\ell;m}\partial_i,$$

and

$$[\nabla_m, [\nabla_k, \nabla_\ell]] + \text{cyclic} = 0 \quad \Leftrightarrow \quad \nabla_{[m}[\nabla_k, \nabla_{\ell]}] = 0.$$

8.36 We have

$$\nabla_\ell \Psi^{ab} = \Psi^{ab}{}_{,\ell} + \Gamma^a{}_{\ell p}\Psi^{pb} + \Gamma^b{}_{\ell p}\Psi^{ap}.$$

The individual pieces are not tensors, but the whole thing *is* a tensor (that's the *raison d'etre* for the covariant derivative), so (8.103) gives

$$\begin{aligned}\nabla_k(\nabla_\ell \Psi^{ab}) &= [\nabla_\ell \Psi^{ab}]_{,k} - \Gamma^m{}_{k\ell}(\nabla_m \Psi^{ab}) + \Gamma^a{}_{km}(\nabla_\ell \Psi^{mb}) + \Gamma^b{}_{km}(\nabla_\ell \Psi^{am}) \\ &= \Psi^{ab}{}_{,\ell k} + \Gamma^a{}_{\ell p,k}\Psi^{pb} + \Gamma^a{}_{\ell p}\Psi^{pb}{}_{,k} + \Gamma^b{}_{\ell p,k}\Psi^{ap} + \Gamma^b{}_{\ell p}\Psi^{ap}{}_{,k} \\ &\quad - \Gamma^m{}_{k\ell}[\Psi^{ab}{}_{,m} + \Gamma^a{}_{mp}\Psi^{pb} + \Gamma^b{}_{mp}\Psi^{ap}] \\ &\quad + \Gamma^a{}_{km}[\Psi^{mb}{}_{,\ell} + \Gamma^m{}_{\ell p}\Psi^{pb} + \Gamma^b{}_{\ell p}\Psi^{mp}] \\ &\quad + \Gamma^b{}_{km}[\Psi^{am}{}_{,\ell} + \Gamma^a{}_{\ell p}\Psi^{pm} + \Gamma^m{}_{\ell p}\Psi^{ap}].\end{aligned}$$

Now subtract the same thing with k and ℓ interchanged. Any term above that is symmetric in k and ℓ will cancel. This means we can throw away the entire term multiplying $\Gamma^m{}_{k\ell}$, because of the symmetry of the Christoffel symbols. Moreover, the terms involving a derivative of Ψ all disappear as well, because

$$\Psi^{ab}{}_{,\ell k} + \Gamma^a{}_{\ell p}\Psi^{pb}{}_{,k} + \Gamma^b{}_{\ell p}\Psi^{ap}{}_{,k} + \Gamma^a{}_{km}\Psi^{mb}{}_{,\ell} + \Gamma^b{}_{km}\Psi^{am}{}_{,\ell}$$

is symmetric under $k \leftrightarrow \ell$. Finally, the terms

$$\Gamma^a{}_{km}\Gamma^b{}_{\ell p}\Psi^{mp} + \Gamma^b{}_{km}\Gamma^a{}_{\ell p}\Psi^{pm}$$

also disappear for the same reason. We are left with

$$(\Gamma^a{}_{\ell p,k} + \Gamma^a{}_{km}\Gamma^m{}_{\ell p})\Psi^{pb} + (\Gamma^b{}_{\ell p,k} + \Gamma^b{}_{km}\Gamma^m{}_{\ell p})\Psi^{ap}$$

minus the same term with k and ℓ interchanged, which gives

$$R^a{}_{pk\ell}\Psi^{pb} + R^b{}_{pk\ell}\Psi^{ap}.$$

8.37 a. The Levi-Civita connection is torsion-free and metric compatible, so

$$\nabla_X Y - \nabla_Y X = [X, Y] \quad \text{and} \quad Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Hence

$$\begin{aligned}0 = (\mathcal{L}_X g)(Y, Z) &= Xg(Y, Z) - g([X, Y], Z) - g(Y, [X, Z]) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ &\qquad - g(\nabla_X Y - \nabla_Y X, Z) - g(Y, \nabla_X Z - \nabla_Z X) \\ &= g(\nabla_Y X, Z) + g(Y, \nabla_Z X).\end{aligned}$$

The claim now follows.

b. Let $X = X^k\partial_k$, $Y = \partial_i$, and $Z = \partial_j$. Then the result follows immediately from the result of Part (a) and the covariant constancy of the metric.

$$\begin{aligned}0 &= g(\nabla_i(X^k\partial_k), \partial_j) + g(\partial_i, \nabla_j(X^k\partial_k)), \\ &= g(X^k{}_{;i}\partial_k, \partial_j) + g(\partial_i, X^k{}_{;j}\partial_k), \\ &= X^k{}_{;i}g_{kj} + X^k{}_{;j}g_{ik} = X_{j;i} + X_{i;j}.\end{aligned}$$

c. By metric compatibility,

$$Yg(X, Y) = g(\nabla_Y X, Y) + g(X, \nabla_Y Y).$$

But the curve is a geodesic, so $\nabla_Y Y = 0$. Also, X is Killing, so

$$g(\nabla_Y X, Y) = -g(Y, \nabla_Y X) = 0,$$

whereupon we conclude $Yg(X, Y) = 0$.

d. By our previous results,

$$\nabla_{\partial i} \partial_j = \Gamma^k{}_{ij} \partial_k = g^{k\ell} \Gamma_{\ell ij} \partial_k,$$

where

$$\Gamma_{\ell ij} = \frac{1}{2}(g_{\ell i,j} + g_{\ell j,i} - g_{ij,\ell}).$$

Choosing coordinates so that $X = \partial_1$, $Y = Y^j \partial_j$, and $Z = Z^k \partial_k$ gives

$$\begin{aligned} &g(\nabla_Y X, Z) + g(Y, \nabla_Z X) \\ &\quad = g(Y^j \Gamma^m{}_{j1} \partial_m, Z^k \partial_k) + g(Y^j \partial_j, Z^k \Gamma^m{}_{k1} \partial_m) \\ &\quad = Y^j Z^k (\Gamma_{kj1} + \Gamma_{jk1}) \\ &\quad = Y^j Z^k g_{jk,1}. \end{aligned}$$

As Y and Z are arbitrary, it follows that $g_{jk,1} = 0$, as claimed.

e. Let ϕ_t be the flow corresponding to X. Then

$$\begin{aligned} \phi_{-t*} g_{\phi_t p} &= g_p + \int_0^t \frac{d}{ds}(\phi_{-s*} g_{\phi_s p})\, ds, \\ &= g_p + \int_0^t \frac{d}{dx}(\phi_{-(s+x)*} g_{\phi_{(s+x)} p})\Big|_{x=0} ds, \\ &= g_p + \int_0^t \phi_{-s*} \frac{d}{dx}(\phi_{-x*} g_{\phi_{(s+x)} p})\Big|_{x=0} ds, \\ &= g_p + \int_0^t \phi_{-s*} (\mathcal{L}_X g)_{\phi_s p}\, ds, \\ &= g_p, \end{aligned}$$

which is precisely the statement that ϕ_t is an isometry.

f. It suffices to show that Killing fields are closed under the Lie bracket. But this follows immediately from (3.102) applied to the metric. If X and Y are Killing fields, then

$$\mathcal{L}_{[X,Y]} g = \mathcal{L}_X \mathcal{L}_Y g - \mathcal{L}_Y \mathcal{L}_X g = 0,$$

so $[X, Y]$ is Killing as well.

8.38 a. Watch the birdie:

$$\begin{aligned}
\xi_{i,jk} &= \xi_{i,kj} && \text{(mixed partials commute)} \\
&= -\xi_{k,ij} && \text{(Killing's equation)} \\
&= -\xi_{k,ji} && \text{(mixed partials again)} \\
&= \xi_{j,ki} && \text{(Killing's equation)} \\
&= \xi_{j,ik} && \text{(mixed partials again)} \\
&= -\xi_{i,jk} && \text{(Killing's equation)} \\
&= 0 && (x = -x \Rightarrow x = 0).
\end{aligned}$$

Integrating $\xi_{i,jk} = 0$ gives

$$\xi_{i,j} = \alpha_{ij}$$

for some constant tensor which, by virtue of Killing's equation must be antisymmetric. Integrating again gives

$$\xi_i = \alpha_{ij}x^j + a_i$$

for some constants a_i, as advertised.

b. With $a_i = 0$ the integral curve equation reads $\dot{x}^i = \xi^i = \delta^{ij}\alpha_{jk}x^k$. Setting $\alpha_{12} = -\alpha_{21} = 1$ and all other components to zero we get

$$\begin{aligned}
\dot{x}^1 &= x^2 \\
\dot{x}^2 &= -x^1 \\
\dot{x}^3 &= 0.
\end{aligned}$$

The latter equation shows that x^3 = constant, so the curve is restricted to the xy-plane. The other two equations can be solved by diagonalizing a matrix (or appealing to complex numbers), but it is simpler just to differentiate the first and substitute the second to get

$$\ddot{x}^1 = \dot{x}^2 = -x^1 \qquad x^1 = A\cos(t + \delta).$$

Plugging this into the second equation and integrating gives

$$x^2 = -A\sin(t + \delta) + B,$$

but this is only compatible with the first equation if $B = 0$. Therefore the integral curves are circles, as promised.

c. This is an elementary exercise in change of variables, but here is one solution anyway. Recall that in spherical polar coordinates

$$\begin{aligned} x &= r \sin\theta \cos\phi & r &= (x^2 + y^2 + z^2)^{1/2} \\ y &= r \sin\theta \sin\phi & \theta &= \cos^{-1}(z/r) \\ z &= r \cos\theta & \phi &= \tan^{-1}(y/x). \end{aligned}$$

Applying the chain rule gives

$$\begin{aligned} \frac{\partial}{\partial x} &= \left(\frac{\partial r}{\partial x}\right)\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial x}\right)\frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial x}\right)\frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial y} &= \left(\frac{\partial r}{\partial y}\right)\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial y}\right)\frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial y}\right)\frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial z} &= \left(\frac{\partial r}{\partial z}\right)\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial z}\right)\frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial z}\right)\frac{\partial}{\partial \phi} \end{aligned}$$

Next we have to compute all those partial derivatives, then convert back to spherical polar coordinates. (We could also invert the Jacobian matrix of the inverse transformation, but that's just as irritating.) For example, we have

$$\begin{aligned} \frac{\partial r}{\partial x} &= \frac{1}{2}(x^2 + y^2 + z^2)^{-1/2}(2x) = \frac{x}{r} = \sin\theta\cos\phi, \\ \frac{\partial \theta}{\partial x} &= -(1 - (z/r)^2)^{-1/2}(-z/r^2)(x/r) = \frac{xz}{r^2(x^2 + y^2)^{1/2}} = \frac{\cos\theta\cos\phi}{r}, \\ \frac{\partial \phi}{\partial x} &= \frac{-y/x^2}{1 + (y/x)^2} = -\frac{y}{x^2 + y^2} = -\frac{\sin\phi}{r\sin\theta}. \end{aligned}$$

Continuing in this way gives

$$\begin{aligned} \frac{\partial}{\partial x} &= \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial \phi}, \\ \frac{\partial}{\partial y} &= \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial \phi}, \\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}. \end{aligned}$$

From the previous parts we have

$$\begin{aligned} \xi^{(1)} &= y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}, \\ \xi^{(2)} &= z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}, \\ \xi^{(3)} &= x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}, \end{aligned}$$

whereupon the desired result follows by substituting using the above computations. For example,

$$\begin{aligned}
\xi^{(1)} &= (r\sin\theta\sin\phi)\left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right) \\
&\quad - (r\cos\theta)\left(\sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}\right) \\
&= -\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}.
\end{aligned}$$

The other two vector fields follow similarly.

8.39 a. From Equation (3.124) we get

$$\mathcal{L}_\xi F = \left(F_{ij,k}\xi^k + F_{kj}\xi^k{}_{,i} + F_{ik}\xi^k{}_{,j}\right) dx^i \wedge dx^j$$

For $\xi^{(3)}$ we have

$$0 = \mathcal{L}_{\xi^{(3)}} F = F_{ij,\phi}\, dx^i \wedge dx^j,$$

which just yields

$$F_{ij,\phi} = 0. \tag{1}$$

Next, for $\xi^{(1)}$ we have

$$\begin{aligned}
0 = \mathcal{L}_{\xi^{(1)}} F = [&F_{ij,\theta}(-\sin\phi) + F_{ij,\phi}(-\cot\theta\cos\phi) \\
&+ F_{\theta j}(-\sin\phi)_{,i} + F_{\phi j}(-\cot\theta\cos\phi)_{,i} \\
&+ F_{i\theta}(-\sin\phi)_{,j} + F_{i\phi}(-\cot\theta\cos\phi)_{,j}]\, dx^i \wedge dx^j,
\end{aligned}$$

which yields the following equations (using (1)):

$$F_{\tau\rho,\theta} = 0, \tag{2}$$

$$F_{\tau\theta,\theta} - F_{\tau\phi}\csc^2\theta\cot\phi = 0, \tag{3}$$

$$F_{\tau\phi,\theta} + F_{\tau\theta}\cot\phi - F_{\tau\phi}\cot\theta = 0, \tag{4}$$

$$F_{\rho\theta,\theta} - F_{\rho\phi}\csc^2\theta\cot\phi = 0, \tag{5}$$

$$F_{\rho\phi,\theta} + F_{\rho\theta}\cot\phi - F_{\rho\phi}\cot\theta = 0, \tag{6}$$

$$F_{\theta\phi,\theta} - F_{\theta\phi}\cot\theta = 0. \tag{7}$$

Likewise,

$$\begin{aligned}
0 = \mathcal{L}_{\xi^{(2)}} F = [&F_{ij,\theta}(\cos\phi) + F_{ij,\phi}(-\cot\theta\sin\phi) \\
&+ F_{\theta j}(\cos\phi)_{,i} + F_{\phi j}(-\cot\theta\sin\phi)_{,i} \\
&+ F_{i\theta}(\cos\phi)_{,j} + F_{i\phi}(-\cot\theta\sin\phi)_{,j}]\, dx^i \wedge dx^j,
\end{aligned}$$

which yields the following equations (using (1) again):

$$F_{\tau\rho,\theta} = 0, \tag{8}$$

$$F_{\tau\theta,\theta} + F_{\tau\phi} \csc^2\theta \tan\phi = 0, \tag{9}$$

$$F_{\tau\phi,\theta} - F_{\tau\theta} \tan\phi - F_{\tau\phi} \cot\theta = 0, \tag{10}$$

$$F_{\rho\theta,\theta} + F_{\rho\phi} \csc^2\theta \tan\phi = 0, \tag{11}$$

$$F_{\rho\phi,\theta} - F_{\rho\theta} \tan\phi - F_{\rho\phi} \cot\theta = 0, \tag{12}$$

$$F_{\theta\phi,\theta} - F_{\theta\phi} \cot\theta = 0. \tag{13}$$

Combining (2)-(7) with (8)-(12) we have the following constraints:

$$
\begin{aligned}
F_{\tau\rho,\theta} &= 0 && \text{((2) and (8))}, && (14)\\
F_{\tau\theta,\theta} &= F_{\tau\phi} = 0 && \text{((3) and (9))}, && (15)\\
F_{\tau\theta} &= 0 && \text{((4) and (10) and (15))}, && (16)\\
F_{\rho\theta,\theta} &= F_{\rho\phi} = 0 && \text{((5) and (11))}, && (17)\\
F_{\rho\theta} &= 0 && \text{((6) and (12) and (17))}, && (18)\\
F_{\theta\phi} &= B(\tau, \rho) \sin\theta && \text{((7) and (13))}, && (19)
\end{aligned}
$$

where $B(\tau, \rho)$ is some arbitrary function. By (1) and (14) $F_{\tau\rho} = A(\tau, \rho)$ for some arbitrary function A. Thus

$$F = A(\tau, \rho)\, d\tau \wedge d\rho + B(\tau, \rho) \sin\theta\, d\theta \wedge d\phi,$$

as promised.

b. From (8.91) we get

$$
\begin{aligned}
(d\tau \wedge d\rho) \wedge (d\theta \wedge d\phi) &= g(\star(d\tau \wedge d\rho), d\theta \wedge d\phi)\sqrt{|G|}\, d\tau \wedge d\rho \wedge d\theta \wedge d\phi\\
&= \alpha\sqrt{|G|} g(d\theta \wedge d\phi, d\theta \wedge d\phi)\, d\tau \wedge d\rho \wedge d\theta \wedge d\phi,
\end{aligned}
$$

so

$$\alpha\sqrt{|G|} g(d\theta \wedge d\phi, d\theta \wedge d\phi) = 1.$$

As the determinant of a diagonal matrix is just the product of the diagonal elements,

$$\sqrt{|G|} = abr^2 \sin\theta.$$

Also, inverting the metric gives

$$g = -a^{-2}\partial_\tau^2 + b^{-2}\partial_\rho^2 + r^{-2}\left(\partial_\theta^2 + \sin^{-2}\partial_\phi^2\right),$$

where $\partial_\tau^2 := \partial_\tau \otimes \partial_\tau$ etc., so

$$g(d\theta \wedge d\phi, d\theta \wedge d\phi) = \begin{vmatrix} g(d\theta, d\theta) & g(d\theta, d\phi) \\ g(d\phi, d\theta) & g(d\phi, d\phi) \end{vmatrix}$$
$$= \begin{vmatrix} r^{-2} & 0 \\ 0 & r^{-2}\sin^{-2}\theta \end{vmatrix}$$
$$= r^{-4}\sin^{-2}\theta.$$

Hence

$$\star(d\tau \wedge d\rho) = \frac{r^2 \sin\theta}{ab}\, d\theta \wedge d\phi.$$

Using $\star^2 = (-1)^{k(n-k)+d}$ on k forms in n dimensions with $d = 1$ for a Lorentzian metric gives

$$\star(d\theta \wedge d\phi) = -\frac{ab}{r^2 \sin\theta}\, d\tau \wedge d\rho.$$

c. The first Maxwell equation gives

$$0 = dF = \frac{\partial B}{\partial \tau} \sin\theta\, d\tau \wedge d\theta \wedge d\phi + \frac{\partial B}{\partial \rho} \sin\theta\, d\rho \wedge d\theta \wedge d\phi,$$

so B must be constant. The second Maxwell equation gives

$$0 = d\star F = \frac{\partial(Ar^2/ab)}{\partial \tau} \sin\theta\, d\tau \wedge d\theta \wedge d\phi + \frac{\partial(Ar^2/ab)}{\partial \rho} \sin\theta\, d\rho \wedge d\theta \wedge d\phi,$$

so Ar^2/ab is a constant.

The metric is already diagonal, so by inspection we have

$$\theta^{\hat{0}} = a\, d\tau,$$
$$\theta^{\hat{1}} = b\, d\rho,$$
$$\theta^{\hat{2}} = r\, d\theta,$$
$$\theta^{\hat{3}} = r \sin\theta\, d\phi.$$

(For instance, $g(\theta^{\hat{1}}, \theta^{\hat{1}}) = a^2 g(d\tau, d\tau) = a^2(-a^{-2}) = -1$, *etc.*.) Therefore

$$F = \frac{1}{2} F_{\hat{i}\hat{j}} \theta^{\hat{i}} \wedge \theta^{\hat{j}}$$
$$= abF_{\hat{0}\hat{1}}\, d\tau \wedge d\rho + r^2 \sin\theta F_{\hat{2}\hat{3}}\, d\theta \wedge d\phi + \cdots .$$

We conclude that $F_{\hat{0}\hat{1}} = A/ab$ and $F_{\hat{2}\hat{3}} = B/r^2$, with all other components vanishing. It follows that $F_{\hat{0}\hat{1}} = \text{constant}/r^2$ and $F_{\hat{2}\hat{3}} = \text{constant}/r^2$. The magnitudes and signs of the constants follow from the fact that the electric and magnetic field components are related to the field strength tensor in a local orthonormal frame as they are related in flat spacetime, namely by $E^{\hat{j}} = F^{\hat{0}\hat{j}}$ and $F^{\hat{j}\hat{k}} = \varepsilon^{\hat{j}\hat{k}\hat{\ell}} B_{\hat{\ell}}$.

8.40 a. Simply observe that the metric tensor can be written

$$ds^2 = g_{\hat{a}\hat{b}} \theta^{\hat{a}} \otimes \theta^{\hat{b}}$$

where $g_{\hat{a}\hat{b}}$ is the Minkowski metric.

b. By the antisymmetry of the connection matrices with downstairs indices, we have (with Greek indices running from 0 to 3 and Latin indices running from 1 to 3):

$$\begin{aligned}
\omega^{\hat{0}}{}_{\hat{0}} &= g^{\hat{0}\hat{\mu}} \omega_{\hat{\mu}\hat{0}} = g^{\hat{0}\hat{0}} \omega_{\hat{0}\hat{0}} = 0 \\
\omega^{\hat{0}}{}_{\hat{i}} &= g^{\hat{0}\hat{\mu}} \omega_{\hat{\mu}\hat{i}} = g^{\hat{0}\hat{0}} \omega_{\hat{0}\hat{i}} = -\omega_{\hat{0}\hat{i}} \\
\omega^{\hat{i}}{}_{\hat{0}} &= g^{\hat{i}\hat{\mu}} \omega_{\hat{\mu}\hat{0}} = g^{\hat{i}\hat{i}} \omega_{\hat{i}\hat{0}} = \omega_{\hat{i}\hat{0}} = -\omega_{\hat{0}\hat{i}} = \omega^{\hat{0}}{}_{\hat{i}} \\
\omega^{\hat{i}}{}_{\hat{j}} &= g^{\hat{i}\hat{\mu}} \omega_{\hat{\mu}\hat{j}} = g^{\hat{i}\hat{i}} \omega_{\hat{i}\hat{j}} = \omega_{\hat{i}\hat{j}} = -\omega_{\hat{j}\hat{i}} = -\omega^{\hat{j}}{}_{\hat{i}}.
\end{aligned}$$

c. We have

$$\begin{aligned}
d\theta^{\hat{0}} &= \frac{1}{2} \Phi^{-1/2} \left(\frac{2m}{r^2} \right) dr \wedge dt = -\Phi^{-1/2} \left(\frac{m}{r^2} \right) \theta^{\hat{0}} \wedge \theta^{\hat{1}} \\
&= -\omega^{\hat{0}}{}_{\hat{a}} \wedge \theta^{\hat{a}} \\
&= -\omega^{\hat{0}}{}_{\hat{0}} \wedge \theta^{\hat{0}} - \omega^{\hat{0}}{}_{\hat{1}} \wedge \theta^{\hat{1}} - \omega^{\hat{0}}{}_{\hat{2}} \wedge \theta^{\hat{2}} - \omega^{\hat{0}}{}_{\hat{3}} \wedge \theta^{\hat{3}} \\
&= -\omega^{\hat{0}}{}_{\hat{1}} \wedge \theta^{\hat{1}}.
\end{aligned}$$

Now we guess that

$$\omega^{\hat{0}}{}_{\hat{1}} = \omega^{\hat{1}}{}_{\hat{0}} = \Phi^{-1/2} \left(\frac{m}{r^2} \right) \theta^{\hat{0}} = \frac{m}{r^2}\, dt,$$

which certainly satisfies the structure equation. (The reason this is a guess is because $\omega^{\hat{0}}{}_{\hat{1}}$ could have a term proportional to dr. It turns out that the other structure equations rule out this possibility, but we would have to write them all out in order to verify this.)

d. We have

$$\begin{aligned}
\Omega^{\hat{0}}{}_{\hat{1}} &= d\omega^{\hat{0}}{}_{\hat{1}} + \omega^{\hat{0}}{}_{\hat{\mu}} \wedge \omega^{\hat{\mu}}{}_{\hat{1}} \\
&= d\omega^{\hat{0}}{}_{\hat{1}} \\
&= \frac{2m}{r^3}\, dt \wedge dr \\
&= \frac{2m}{r^3}\, \theta^{\hat{0}} \wedge \theta^{\hat{1}},
\end{aligned}$$

and

$$
\begin{aligned}
\Omega^{\hat{1}}{}_{\hat{2}} &= d\omega^{\hat{1}}{}_{\hat{2}} + \omega^{\hat{1}}{}_{\hat{\mu}} \wedge \omega^{\hat{\mu}}{}_{\hat{2}} \\
&= d\omega^{\hat{1}}{}_{\hat{2}} + \omega^{\hat{1}}{}_{\hat{3}} \wedge \omega^{\hat{3}}{}_{\hat{2}} \\
&= -\frac{1}{2}\Phi^{-1/2}\left(-\frac{2m}{r^2}\right) dr \wedge d\theta \\
&= \frac{m}{r^3}\,\theta^{\hat{1}} \wedge \theta^{\hat{2}}.
\end{aligned}
$$

8.41 a. Begin with the parallel transport equation (8.63)

$$\frac{d\gamma^i}{dt}\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^k{}_{ij}Y^j\right) = 0.$$

We must choose a parameterization for γ. As the result is parameterization independent we choose the simplest one, namely

$$\theta(t) = \theta_0 \qquad \text{and} \qquad \phi(t) = t.$$

Then $\dot{\theta} = 0$ and $\dot{\phi} = 1$, so (8.63) reduces to

$$Y^k{}_{,\phi} + \Gamma^k{}_{\phi\theta}Y^\theta + \Gamma^k{}_{\phi\phi}Y^\phi = 0.$$

Plugging in the known values of the Christoffel symbols gives

$$Y^\theta{}_{,\phi} - \sin\theta_0\cos\theta_0 Y^\phi = 0$$

and

$$Y^\phi{}_{,\phi} + \cot\theta_0 Y^\theta = 0.$$

Now differentiate the second equation and substitute the first to get

$$Y^\phi{}_{,\phi\phi} + \cos\theta_0 Y^\phi = 0,$$

whose solution is

$$Y^\phi = A\cos(\phi\cos\theta_0) + B\sin(\phi\cos\theta_0).$$

Plug this back into the differential equation for Y^ϕ and solve for Y^θ. This gives

$$
\begin{aligned}
Y^\theta &= -\frac{1}{\cot\theta_0}Y^\phi{}_{,\phi} \\
&= \sin\theta_0[A\sin(\phi\cos\theta_0) - B\cos(\phi\cos\theta_0)].
\end{aligned}
$$

From the initial conditions we get

$$a = Y^\theta(\theta_0, 0) = -B\sin\theta_0$$

and

$$b = Y^\phi(\theta_0, 0) = A,$$

so the general solutions are

$$\begin{pmatrix} Y^\theta \\ Y^\phi \end{pmatrix} = \begin{pmatrix} \cos(\phi\cos\theta_0) & \sin\theta_0 \sin(\phi\cos\theta_0) \\ -\sin(\phi\cos\theta_0)/\sin\theta_0 & \cos(\phi\cos\theta_0) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

b. The angle between the parallel transported vector Y and the original vector Y_0 is

$$\cos\psi = \frac{g(Y, Y_0)}{\sqrt{g(Y, Y)g(Y_0, Y_0)}}.$$

We have

$$\begin{aligned} g(Y, Y) &= g_{\theta\theta}(Y^\theta)^2 + g_{\phi\phi}(Y^\phi)^2 \\ &= [a\cos(\phi\cos\theta_0) + b\sin\theta_0\sin(\phi\cos\theta_0)]^2 \\ &\qquad + \sin^2\theta_0[-a\sin(\phi\cos\theta_0)/\sin\theta_0 + b\cos(\phi\cos\theta_0)]^2 \\ &= a^2 + b^2\sin^2\theta_0 \\ &= g(Y_0, Y_0), \end{aligned}$$

which just confirms that parallel transport preserves the length of vectors. Also

$$\begin{aligned} g(Y, Y_0) &= g_{\theta\theta}Y^\theta a + g_{\phi\phi}Y^\phi b \\ &= a[a\cos(\phi\cos\theta_0) + b\sin\theta_0\sin(\phi\cos\theta_0)] \\ &\qquad + b\sin^2\theta_0[-a\sin(\phi\cos\theta_0)/\sin\theta_0 + b\cos(\phi\cos\theta_0)] \\ &= (a^2 + b^2\sin^2\theta_0)\cos(\phi\cos\theta_0), \end{aligned}$$

from which it follows that

$$\cos\psi = \cos(\phi\cos\theta_0),$$

or

$$\psi = \phi\cos\theta_0.$$

As we walk with an attitude around a latitude, ϕ turns through a full 2π, so the vector turns through an angle

$$\psi = 2\pi\cos\theta_0.$$

8.42 Rewrite (8.114) as

$$\ddot{\phi}/\dot{\phi} = -2(\cot\theta)\dot{\theta}.$$

Integrating both sides with respect to t gives

$$\ln\dot{\phi} = -2\ln\sin\theta + c$$

and exponentiating yields

$$(\sin^2\theta)\dot{\phi} = J,$$

where $J = e^c$. Substituting this into (8.113) gives

$$\ddot{\theta} = J^2\frac{\cos\theta}{\sin^3\theta}.$$

Multiply both sides by $\dot{\theta}$ and integrate with respect to t to get

$$\frac{1}{2}\dot{\theta}^2 = -\frac{1}{2}J^2\sin^{-2}\theta + c,$$

which gives (8.116).

Next we observe that (8.116) is separable:

$$\begin{aligned}
& (\sin^2\theta)\dot{\theta}^2 = a^2\sin^2\theta - J^2 \\
\Rightarrow \quad & (\sin\theta)\dot{\theta} = \sqrt{a^2\sin^2\theta - J^2} \\
\Rightarrow \quad & \int\frac{\sin\theta\,d\theta}{\sqrt{a^2\sin^2\theta - J^2}} = \int dt \\
\Rightarrow \quad & -\int\frac{d(\cos\theta)}{\sqrt{1-\cos^2\theta-(J/a)^2}} = at + c \\
\Rightarrow \quad & -\arcsin\left(\frac{\cos\theta}{\sqrt{1-(J/a)^2}}\right) = at + c \\
\Rightarrow \quad & \cos\theta = -\sqrt{1-(J/a)^2}\sin(at+c),
\end{aligned}$$

which is (8.117).

For (8.115) we have

$$\begin{aligned}
\dot{\phi} &= \frac{J}{\sin^2\theta} \\
&= \frac{J}{1-\cos^2\theta} \\
&= \frac{J}{1-(1-(J/a)^2)\sin^2(at+c)} \\
&= \frac{J}{\cos^2(at+c)+(J/a)^2\sin^2(at+c)} \\
&= \frac{J\sec^2(at+c)}{1+(J/a)^2\tan^2(at+c)},
\end{aligned}$$

or

$$\phi = \int \frac{J \sec^2(at + c)}{1 + (J/a)^2 \tan^2(at + c)}\, dt.$$

Set $x = (J/a)\tan(at + c)$ so $dx = J \sec^2(at + c)\, dt$ to get

$$\phi = \int \frac{dx}{1 + x^2} = \arctan x + \phi_0 = \arctan((J/a)\tan(at + c)) + \phi_0,$$

or

$$\tan(\phi - \phi_0) = (J/a)\tan(at + c),$$

which is (8.118).

For brevity, set $\beta := J/a$, $\alpha^2 = 1 - \beta^2$, and $b := at + c$. Then

$$\sin\theta = \sqrt{1 - \cos^2\theta} = \sqrt{\cos^2 b + \beta^2 \sin^2 b},$$

so

$$\begin{aligned}\cot\theta &= \frac{-\alpha \sin b}{\sqrt{\cos^2 b + \beta^2 \sin^2 b}} \\ &= -\frac{\alpha}{\beta^2 + \cot^2 b} \\ &= -\frac{\alpha/\beta}{1 + \cot^2(\phi - \phi_0)} \\ &= C \sin(\phi - \phi_0),\end{aligned}$$

where

$$C := -\frac{\alpha}{\beta} = -\sqrt{(a/J)^2 - 1}.$$

This yields (8.119).

Now multiply both sides of (8.119) by $\sin\theta$ and expand the right hand side to get

$$\cos\theta = C \sin\theta[\sin\phi \cos\phi_0 - \cos\phi \sin\phi_0].$$

Transforming to Cartesian coordinates gives

$$z = Ay - Bx,$$

where $A := C\cos\phi_0$ and $B := C\sin\phi_0$. This is the equation of a plane passing through the origin. The geodesics are the intersections of this plane with the sphere, namely great circles.

8.43 a. Let $X' = aX + bY$ and $Y' = cX + dY$. Then using the fact that $R(X, Y)Z$ is antisymmetric in X and Y we have

$$\begin{aligned} R(X', Y')Y' &= R(aX + bY, cX + dY)(cX + dY) \\ &= acdR(X, Y)X + ad^2R(X, Y)Y \\ &\qquad + bc^2R(Y, X)X + bcdR(Y, X)Y \\ &= (ad - bc)(cR(X, Y)X + dR(X, Y)Y). \end{aligned}$$

Also, $g(W, R(X, Y)Z)$ is antisymmetric in W and Z, so

$$\begin{aligned} g(X', R(X', Y')Y') &= (ad-bc)(bcg(Y, R(X, Y)X)+adg(X, R(X, Y)Y)) \\ &= (ad - bc)^2 g(X, R(X, Y)Y). \end{aligned}$$

Lastly,

$$\begin{aligned} &g(aX + bY, aX + bY)g(cX + dY, cX + dY) - g(aX + bY, cX + dY)^2 \\ &= (a^2g(X, X) + 2abg(X, Y) + b^2g(Y, Y)) \\ &\qquad \times (c^2g(X, X) + 2cdg(X, Y) + d^2g(Y, Y)) \\ &\qquad\qquad - (acg(X, X) + (ad + bc)g(X, Y) + bdg(Y, Y))^2 \\ &= (ad - bc)^2(g(X, X)g(Y, Y) - g(X, Y)^2). \end{aligned}$$

Thus,

$$\begin{aligned} &\frac{g(X', R(X', Y')Y')}{g(X', X')g(Y', Y') - g(X', Y')^2} \\ &\qquad = \frac{(ad - bc)^2 g(X, R(X, Y)Y)}{(ad - bc)^2(g(X, X)g(Y, Y) - g(X, Y)^2)} \\ &\qquad = K(\Pi). \end{aligned}$$

b. By hypothesis

$$g(X, R(X, Y)Y) = g(X, R'(X, Y)Y), \tag{1}$$

so

$$g(X + W, R(X + W, Y)Y) = g(X + W, R'(X + W, Y)Y).$$

Expanding both sides and using (1) gives

$$\begin{aligned} &g(X, R(W, Y)Y) + g(W, R(X, Y)Y) \\ &\qquad = g(X, R'(W, Y)Y) + g(W, R'(X, Y)Y). \end{aligned}$$

Now apply (8.50) to the first term on each side to get

$$2g(W, R(X, Y)Y) = 2g(W, R'(X, Y)Y). \tag{2}$$

In (2) send $Y \to Y + Z$ to get, after the dust has cleared,

$$\begin{aligned} &g(W, R(X, Y)Z) + g(W, R(X, Z)Y) \\ &= g(W, R'(X, Y)Z) + g(W, R'(X, Z)Y). \end{aligned}$$

Equivalently,

$$\begin{aligned} &g(W, R(X, Y)Z) - g(W, R'(X, Y)Z) \\ &= g(W, R(Z, X)Y) - g(W, R'(Z, X)Y). \end{aligned}$$

This says that

$$g(W, R(X, Y)Z) - g(W, R'(X, Y)Z)$$

is invariant under cyclic permutations of X, Y, and Z. So using (8.50) we get

$$3(g(W, R(X, Y)Z) - g(W, R'(X, Y)Z)) = 0,$$

which yields the desired result.

c. Assume

$$R_{ijk\ell} = K_p(g_{ik}g_{j\ell} - g_{i\ell}g_{jk}).$$

Then plugging into (8.120) gives $K(\Pi) = K_p$ for any two plane Π. Conversely, assume $K(\Pi) = K_p$ for any two plane. By definition,

$$g(X, R'(X, Y)Y) = g(X, X)g(Y, Y) - g(X, Y)^2,$$

so (8.120) gives, for any X and Y,

$$g(X, R(X, Y)Y) = K(\Pi)g(X, R'(X, Y)Y) = K_p g(X, R'(X, Y)Y).$$

But R and R' both satisfy the symmetry properties (8.47)-(8.50), so by the proof of Part (b) we must have $g(W, R(X, Y)Z) = K_p g(W, R'(X, Y)Z)$. Taking $X = \partial_k$, $Y = \partial_\ell$, $Z = \partial_j$ and $W = \partial_i$ yields the result.

d. From Part (c) we get, by isotropy,

$$R_{\hat{a}\hat{b}\hat{c}\hat{d}} = K(\delta_{\hat{a}\hat{c}}\delta_{\hat{b}\hat{d}} - \delta_{\hat{a}\hat{d}}\delta_{\hat{b}\hat{c}})$$

in an orthonormal basis. Raising indices with the Kronecker delta gives

$$R^{\hat{a}\hat{b}}{}_{\hat{c}\hat{d}} = \delta^{\hat{a}\hat{m}}\delta^{\hat{b}\hat{n}}R_{\hat{m}\hat{n}\hat{c}\hat{d}} = K(\delta^{\hat{a}}_{\hat{c}}\delta^{\hat{b}}_{\hat{d}} - \delta^{\hat{a}}_{\hat{d}}\delta^{\hat{b}}_{\hat{c}}).$$

Therefore from (8.45) we get

$$\Omega^{\hat{a}\hat{b}} = \frac{1}{2}K(\delta^{\hat{a}}_{\hat{c}}\delta^{\hat{b}}_{\hat{d}} - \delta^{\hat{a}}_{\hat{d}}\delta^{\hat{b}}_{\hat{c}})\theta^{\hat{c}} \wedge \theta^{\hat{d}} = K\,\theta^{\hat{a}} \wedge \theta^{\hat{b}}.$$

The same proof works backwards, so (8.122) implies (8.121). (You have to change coordinates at the last moment from an orthonormal basis to a general coordinate basis.)

e. The torsion free condition is $d\theta + \omega \wedge \theta = 0$, so

$$\begin{aligned} d\Omega^{\hat{a}\hat{b}} &= dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} + K\, d\theta^{\hat{a}} \wedge \theta^{\hat{b}} - K\, \theta^{\hat{a}} \wedge d\theta^{\hat{b}} \\ &= dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} - K\, \omega^{\hat{a}}{}_{\hat{c}} \wedge \theta^{\hat{c}} \wedge \theta^{\hat{b}} + K\, \theta^{\hat{a}} \wedge \omega^{\hat{b}}{}_{\hat{c}} \wedge \theta^{\hat{c}} \\ &= dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} - K\, \omega^{\hat{a}}{}_{\hat{c}} \wedge \theta^{\hat{c}} \wedge \theta^{\hat{b}} - K\, \theta^{\hat{a}} \wedge \theta^{\hat{c}} \wedge \omega^{\hat{b}}{}_{\hat{c}} \\ &= dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} - \omega^{\hat{a}}{}_{\hat{c}} \wedge \Omega^{\hat{c}\hat{b}} - \Omega^{\hat{a}\hat{c}} \wedge \omega^{\hat{b}}{}_{\hat{c}} \\ &= dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} - \omega^{\hat{a}}{}_{\hat{c}} \wedge \Omega^{\hat{c}\hat{b}} + \Omega^{\hat{a}\hat{c}} \wedge \omega_{\hat{c}}{}^{\hat{b}}. \end{aligned}$$

Applying the Bianchi identity gives

$$dK \wedge \theta^{\hat{a}} \wedge \theta^{\hat{b}} = 0.$$

Now $dK = \alpha_{\hat{a}}\theta^{\hat{a}}$ for some $\alpha_{\hat{a}}$'s, so by linear independence of $\theta^{\hat{a}} \wedge \theta^{\hat{b}} \wedge \theta^{\hat{c}}$ (this is where we need the dimension to be at least three) we must have $\alpha_{\hat{a}} = 0$, whence we conclude that $dK = 0$ and $K =$ constant.

8.44 In two dimensions there is only one linearly independent component of the Riemann tensor, which we can take to be R_{1212} in some coordinate basis. The Ricci curvature scalar is therefore

$$\begin{aligned} R = g^{ij} R_{ij} &= g^{ij} R^k{}_{ikj} = g^{ij} g^{k\ell} R_{\ell ikj} \\ &= 2g^{11} g^{22} R_{1212} + 2(g^{12})^2 R_{1221} \\ &= \left(2g^{11} g^{22} - 2(g^{12})^2\right) R_{1221} \\ &= 2G^{-1} R_{1212}, \end{aligned}$$

where G is the determinant of the metric tensor. On the other hand, the Gaussian curvature in two dimensions is, with $X = \partial_1$ and $Y = \partial_2$,

$$K = \frac{R_{1212}}{g_{11} g_{22} - g_{12}^2},$$

so $R = 2K$ (a basis independent result).

8.45 a. The metric components in the coordinate basis (x, y) are $g_{ij} = \delta_{ij}/y^2$. Consider the parameterized curve $\gamma : I \to \mathbb{H}_2^+$ from $(0, 0)$ to $(0, 1)$, say, given by $t \mapsto (0, t)$. Then

$$\int_\gamma ds = \int_0^1 \sqrt{g_{ij}(t) \frac{dx^i}{dt} \frac{dx^j}{dt}}\, dt = \int_0^1 \frac{dt}{t} = \ln t \Big|_0^1 = \infty.$$

b. In a coordinate basis (8.39) reads

$$\Gamma_{kij} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) = \frac{1}{2}(g_{jk,i} + g_{ik,j} - g_{ij,k}).$$

Then, for example,

$$\Gamma_{xxy} = \frac{1}{2} g_{xx,y} = -1/y^3.$$

Similarly we obtain

$$\Gamma_{xxx} = \Gamma_{xyy} = \Gamma_{yxy} = \Gamma_{yyx} = 0$$

and

$$\Gamma_{xxy} = \Gamma_{xyx} = -\Gamma_{yxx} = \Gamma_{yyy} = -1/y^3.$$

Next we raise indices with the inverse metric ($g^{ij} = y^2\delta^{ij}$) to get, for example,

$$\Gamma^x{}_{xx} = g^{xx}\Gamma_{xxx} + g^{xy}\Gamma_{yxx} = 0.$$

Continuing in this way we get

$$\Gamma^x{}_{xx} = \Gamma^x{}_{yy} = \Gamma^y{}_{xy} = \Gamma^y{}_{yx} = 0$$

and

$$\Gamma^x{}_{xy} = \Gamma^x{}_{yx} = -\Gamma^y{}_{xx} = \Gamma^y{}_{yy} = -1/y.$$

c. The geodesic equations read

$$\frac{d^2x^i}{dt^2} + \Gamma^i{}_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0.$$

Substituting from Part (b) we get

$$\ddot{x} - \frac{2}{y}\dot{x}\dot{y} = 0,$$
$$\ddot{y} + \frac{1}{y}\dot{x}^2 - \frac{1}{y}\dot{y}^2 = 0,$$

and simplifying gives

$$y\ddot{x} - 2\dot{x}\dot{y} = 0,$$
$$y\ddot{y} + \dot{x}^2 - \dot{y}^2 = 0.$$

From the first equation we get

$$\frac{\ddot{x}}{\dot{x}} = 2\frac{\dot{y}}{y}.$$

Integrating gives

$$\ln \dot{x} = 2 \ln y + c \Rightarrow \dot{x} = c y^2.$$

Using the hint we can write the second geodesic equation as

$$\dot{f} = \frac{y\ddot{y} - \dot{y}^2}{y^2} = -\frac{\dot{x}^2}{y^2} = -c\dot{x}.$$

There are two cases. If $c = 0$ we have $\dot{x} = 0$ and $f = \dot{y}/y =$ constant. The solutions to these equations are

$$x = \text{constant} \qquad \text{and} \qquad y = y_0 e^{bt},$$

for some constant b. These are straight vertical lines (with a funny parameterization).
If $c \neq 0$ we integrate to get

$$f = -cx + e$$

for some other constant e. Multiply through by $y^2 = \dot{x}/c$ to get

$$y\dot{y} + x\dot{x} = a\dot{x},$$

where $a := e/c$. Integrating both sides gives

$$\frac{1}{2}(y^2 + x^2) = ax + q$$

for some other constant q. This can be rewritten as

$$y^2 + (x - a)^2 = r^2,$$

which is the equation of a circle of radius $r = \sqrt{2q + a^2}$ centered at a. Using $y^2 = \dot{x}/c$ again gives

$$\dot{x} + c(x - a)^2 = cr^2,$$

which separates to

$$\frac{dx}{r^2 - (x - a)^2} = c\, dt.$$

Changing variables to $z = (x - a)/r$ gives

$$\frac{dz}{1 - z^2} = \frac{c}{r}\, dt \quad \Rightarrow \quad \tanh^{-1} z = \frac{ct}{r} + h,$$

and thus

$$x = r \tanh(ct/r + h) + a.$$

Plugging back into $\dot{x} = cy^2$ gives

$$cy^2 = \frac{c}{1 - (ct/r + h)^2}$$

or

$$y = (1 - (ct/r + h)^2)^{-1/2}.$$

d. The Gaussian curvature is the sectional curvature of any two plane at a point. At any point we can choose the basis ∂_x and ∂_y. Thus

$$K = \frac{R_{xyxy}}{g_{xx}g_{yy} - g_{xy}^2}.$$

We have

$$\begin{aligned} R_{xyxy} &= g_{xx} R^x{}_{yxy} \\ &= \frac{1}{y^2}(\Gamma^x{}_{yy,x} - \Gamma^x{}_{xy,y} + \Gamma^x{}_{xx}\Gamma^x{}_{yy} + \Gamma^x{}_{xy}\Gamma^y{}_{yy} \\ &\qquad - \Gamma^x{}_{yx}\Gamma^x{}_{xy} - \Gamma^x{}_{yy}\Gamma^y{}_{xy}) \\ &= -\frac{1}{y^4}. \end{aligned}$$

As the denominator is y^{-4}, we conclude that $K = -1$.

8.46 a. Let

$$T_{A_i}(z) = \frac{a_i z + b_i}{c_i z + d_i}$$

for $i = 1, 2$. Then

$$\begin{aligned} T_{A_2}(T_{A_1}(z)) &= \frac{a_2(a_1 z + b_1)/(c_1 z + d_1) + b_2}{c_2(a_1 z + b_1)/(c_1 z + d_1) + d_2} \\ &= \frac{a_2(a_1 z + b_1) + b_2(c_1 z + d_1)}{c_2(a_1 z + b_1) + d_2(c_1 z + d_1)} \\ &= \frac{(a_2 a_1 + b_2 c_1)z + (a_2 b_1 + b_2 d_1)}{(c_2 a_1 + d_2 c_1)z + (c_2 b_1 + d_2 d_1)}. \end{aligned}$$

But, observe that

$$\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_2 a_1 + b_2 c_1 & a_2 b_1 + b_2 d_1 \\ c_2 a_1 + d_2 c_1 & c_2 b_1 + d_2 d_1 \end{pmatrix}.$$

It follows that

$$T_{A_2} \circ T_{A_1} = T_{A_2 A_1}, \tag{1}$$

which shows that Möbius transformations are closed under composition. Composition is associative. Also, if $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ then $T_I(z) = z$, the identity map on $\mathbb{C}$. This is the identity element of $\mathcal{M}$ because

$$T_A \circ T_I(z) = T_A(z) \quad \text{and} \quad T_I \circ T_A(z) = T_A(z).$$

Lastly, from (1) we get

$$T_A \circ T_{A^{-1}} = T_{A^{-1}} \circ T_A = T_I,$$

which implies that

$$T_A^{-1}(z) = T_{A^{-1}}(z).$$

(Note that the inverse exists because $A \in GL(2, \mathbb{C})$.) Therefore all the group axioms are satisfied.

b. Equation (1) together with $T_I(z) = z$ shows that the map $A \mapsto T_A$ is a group homomorphism. The kernel consists of those matrices A that also map to T_I. In other words,

$$\frac{az + b}{cz + d} = z.$$

Cross multiplying gives

$$cz^2 + (d - a)z - b = 0,$$

and this holds for all z only if $c = b = 0$ and $a = d = \lambda \neq 0$. This shows that the kernel of the homomorphism consists of all nonzero complex multiples of the identity matrix.

c. The existence of the inverse map $T_A^{-1} = T_{A^{-1}}$ shows that the map is bijective. Moreover, because the entries of A are real, multiplying the numerator and denominator by the complex conjugate of the denominator gives

$$\frac{az + b}{cz + d} = \left(\frac{az + b}{cz + d}\right)\left(\frac{c\bar{z} + d}{c\bar{z} + d}\right) = \frac{ac|z|^2 + adz + bc\bar{z} + bd}{|cz + d|^2},$$

where as usual $|z|^2 = z\bar{z}$. Thus, as a map from $\mathbb{R}^2$ to itself T_A is given by

$$x \mapsto \frac{ac(x^2 + y^2) + (ad + bc)x + bd}{(cx + d)^2 + (cy)^2}$$

and

$$y \mapsto \frac{y}{(cx + d)^2 + (cy)^2}.$$

(For the second map we used the fact that $ad - bc = 1$.) Observe that the denominator appearing in these two expressions is always positive because c and d cannot vanish simultaneously without violating the determinant condition. It follows that the map carries $\mathbb{H}^2_+$ to $\mathbb{H}^2_+$ and is everywhere differentiable. (The differentiability of the inverse map follows similarly by sending A to A^{-1}.)

d. This one is a chore in either real or complex coordinates. A better way to do it is to decompose an arbitrary Möbius transformation into a product of a translation, an inversion, and a dilation and prove that each transformation is an isometry, but as we did not discuss such decompositions we will do the problem the long way instead.

First note that

$$dz = dx + i\,dy$$

and

$$d\bar{z} = dx - i\,dy,$$

so

$$dz\,d\bar{z} = dx^2 + dy^2.$$

Also,

$$y = \frac{z - \bar{z}}{2i}$$

so

$$ds^2 = \frac{dx^2 + dy^2}{y^2} = -4\frac{dz\,d\bar{z}}{(z - \bar{z})^2}.$$

In order to avoid confusing the differential dz with the product of d and z we will denote the differential dz by η (and its complex conjugate by $\overline{\eta} := d\bar{z}$). Thus,

$$\begin{aligned} T_A^*(\eta) = dT_A^* z &= \frac{a\eta(cz + d) - (az + b)(c\eta)}{(cz + d)^2} \\ &= \frac{\eta(ad - bc)}{(cz + d)^2} = \frac{\eta}{(cz + d)^2}. \end{aligned}$$

Similarly, because A is real,

$$T_A^*(\overline{\eta}) = dT_A^*\bar{z} = \overline{dT_A^* z} = \frac{\overline{\eta}}{(c\bar{z} + d)^2}.$$

Thus

$$T_A^*(\eta\overline{\eta}) = \frac{\eta\overline{\eta}}{|cz + d|^4}.$$

Lastly,

$$
\begin{aligned}
T_A^*(z - \bar{z}) &= \frac{az + b}{cz + d} - \frac{a\bar{z} + b}{c\bar{z} + d} \\
&= \frac{(az + b)(c\bar{z} + d) - (a\bar{z} + b)(cz + d)}{|cz + d|^2} \\
&= \frac{(z - \bar{z})}{|cz + d|^2},
\end{aligned}
$$

so

$$
T_A^*(z - \bar{z})^2 = \frac{(z - \bar{z})^2}{|cz + d|^4}.
$$

Putting it all together gives

$$
\begin{aligned}
T_A^* ds^2 &= -4 \frac{T_A^*(\eta) T_A^*(\overline{\eta})}{T_A^*(z - \bar{z})^2} \\
&= -4 \left(\frac{\eta \overline{\eta}}{|cz + d|^4} \right) \left(\frac{|cz + d|^4}{(z - \bar{z})^2} \right) \\
&= -4 \frac{\eta \overline{\eta}}{(z - \bar{z})^2} = ds^2.
\end{aligned}
$$

8.47 a. Following the hint we get

$$
h = \sigma^* g = \sigma^* dx \otimes \sigma^* dx + \sigma^* dy \otimes \sigma^* dy + \sigma^* dz \otimes \sigma^* dz.
$$

Now

$$
\sigma^* dx = d\sigma^* x = d(x \circ \sigma) = \frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv,
$$

and similarly for $\sigma^* dy$ and $\sigma^* dz$, so

$$
\begin{aligned}
h &= \left(\frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv \right) \otimes \left(\frac{\partial x}{\partial u}\, du + \frac{\partial x}{\partial v}\, dv \right) + \cdots \\
&= g(\sigma_u, \sigma_u)\, du^2 + 2g(\sigma_u, \sigma_v)\, du dv + g(\sigma_v, \sigma_v)\, dv^2.
\end{aligned}
$$

b. We have

$$
\begin{aligned}
\sigma_\theta &= (\cos\theta \cos\phi, \cos\theta \sin\phi, -\sin\theta), \\
\sigma_\phi &= (-\sin\theta \sin\phi, \sin\theta \cos\phi, 0),
\end{aligned}
$$

so

$$
\begin{aligned}
E &= g(\sigma_\theta, \sigma_\theta) = 1, \\
F &= g(\sigma_\theta, \sigma_\phi) = 0, \\
G &= g(\sigma_\phi, \sigma_\phi) = \sin^2\theta,
\end{aligned}
$$

and thus

$$g = d\theta^2 + \sin^2 \, d\phi^2,$$

as expected.

8.48 We have

$$\sigma_u = (-a \cosh v \sin u, a \cosh v \cos u, 0),$$
$$\sigma_v = (a \sinh v \cos u, a \sinh v \sin u, a).$$

The metric tensor components are therefore

$$\begin{aligned} E &= g(\sigma_u, \sigma_u) = a^2 \cosh^2 v, \\ F &= g(\sigma_u, \sigma_v) = 0, \\ G &= g(\sigma_v, \sigma_v) = a^2 \cosh^2 v. \end{aligned}$$

By inspection,

$$\theta^{\hat{1}} = a \cosh v \, du,$$
$$\theta^{\hat{2}} = a \cosh v \, dv.$$

Thus,

$$\begin{aligned} d\theta^{\hat{1}} &= a \sinh v \, dv \wedge du = -\frac{\sinh v}{a \cosh^2 v} \theta^{\hat{1}} \wedge \theta^{\hat{2}}, \\ d\theta^{\hat{2}} &= 0. \end{aligned}$$

There is only one independent connection form, namely $\omega^{\hat{1}}{}_{\hat{2}}$. Using the torsion free condition $d\theta = -\omega \wedge \theta$ we guess

$$\omega^{\hat{1}}{}_{\hat{2}} = \frac{\sinh v}{\cosh^2 v} \theta^{\hat{1}} = \tanh v \, du.$$

(The guess is consistent, because if we had added a term of the form $f(u, v) \, dv$ to $\omega^{\hat{1}}{}_{\hat{2}}$ it would have contradicted the equation for $d\theta^{\hat{2}}$.) The corresponding independent curvature two form component is

$$\begin{aligned} \Omega^{\hat{1}\hat{2}} = \Omega^{\hat{1}}{}_{\hat{2}} &= d\omega^{\hat{1}}{}_{\hat{2}} + \omega^{\hat{1}}{}_{\hat{2}} \wedge \omega^{\hat{2}}{}_{\hat{1}} \\ &= \operatorname{sech}^2 v \, dv \wedge du \\ &= -a^{-2} \operatorname{sech}^4 v \, \theta^{\hat{1}} \wedge \theta^{\hat{2}}, \end{aligned}$$

whereupon we conclude that $K = -1/a^2 \cosh^4 v$.

8.49 We have

$$\sigma_u = (-(b + a \cos v) \sin u, (b + a \cos v) \cos u, 0),$$
$$\sigma_v = (-a \sin v \cos u, -a \sin v \sin u, a \cos v).$$

The metric tensor components are therefore

$$\begin{aligned} E &= g(\sigma_u, \sigma_u) = (b + a\cos v)^2, \\ F &= g(\sigma_u, \sigma_v) = 0, \\ G &= g(\sigma_v, \sigma_v) = a^2. \end{aligned}$$

By inspection,

$$\begin{aligned} \theta^{\hat{1}} &= (b + a\cos v)\,du, \\ \theta^{\hat{2}} &= a\,dv. \end{aligned}$$

Thus,

$$\begin{aligned} d\theta^{\hat{1}} &= -a\sin v\,dv \wedge du, \\ d\theta^{\hat{2}} &= 0. \end{aligned}$$

There is only one independent connection form, namely $\omega^{\hat{1}}{}_{\hat{2}}$. Using the torsion free condition $d\theta = -\omega \wedge \theta$ we guess

$$\omega^{\hat{1}}{}_{\hat{2}} = \sin v\,du.$$

(The guess is consistent, because if we had added a term of the form $f(u, v)\,dv$ to $\omega^{\hat{1}}{}_{\hat{2}}$ it would have contradicted the equation for $d\theta^{\hat{2}}$.) The corresponding independent curvature two form component is

$$\begin{aligned} \Omega^{\hat{1}\hat{2}} = \Omega^{\hat{1}}{}_{\hat{2}} &= d\omega^{\hat{1}}{}_{\hat{2}} + \omega^{\hat{1}}{}_{\hat{2}} \wedge \omega^{\hat{2}}{}_{\hat{1}} \\ &= \cos v\,dv \wedge du \\ &= -\frac{\cos v}{a(b + a\cos v)}\theta^{\hat{1}} \wedge \theta^{\hat{2}}, \end{aligned}$$

whereupon we conclude that $K = -\cos v/a(b + a\cos v)$.

8.50 a. By definition, $\operatorname{Ad} g = (R_{g^{-1}} \circ L_g)_{*,e} = R_{g^{-1}*,g} \circ L_{g*,e}$. Thus

$$\begin{aligned} &(R^*_{g^{-1}}m)(X_h, Y_h)_h \\ &= m(R_{g^{-1}*}X_h, R_{g^{-1}*}Y_h)_{hg^{-1}} && \text{(Equation (3.90))}, \\ &= B(L_{gh^{-1}*}R_{g^{-1}*}X_h, L_{gh^{-1}*}R_{g^{-1}*}Y_h) && \text{(Equation 8.124)}, \\ &= B(R_{g^{-1}*}L_{gh^{-1}*}X_h, R_{g^{-1}*}L_{gh^{-1}*}Y_h) && \text{(R_a and L_b commute)}, \\ &= B(R_{g^{-1}*}L_{g*}L_{h^{-1}*}X_h, R_{g^{-1}*}L_{g*}L_{h^{-1}*}Y_h) && \text{(chain rule)}, \\ &= B(L_{h^{-1}*}X_h, L_{h^{-1}*}Y_h) && \text{(Ad invariance of B)}, \\ &= m(X_h, Y_h) && \text{(Equation 8.124)}. \end{aligned}$$

b. Let X, Y, and Z be left invariant vector fields. The Koszul formula reads

$$\begin{aligned}2(\nabla_X Y, Z) = X(Y, Z) + Y(X, Z) - Z(X, Y)\\ + ([X, Y], Z) - ([X, Z], Y) - ([Y, Z], X).\end{aligned}$$

By Exercise 3.52c the last two terms cancel, because

$$(\operatorname{ad} Z(X), Y) + (\operatorname{ad} Z(Y), X) = 0.$$

Also, by left invariance of X and Y,

$$(X_g, Y_g) = (L_{g^{-1}*}X_g, L_{g^{-1}*}Y_g)_e = (X_e, Y_e)$$

so (X, Y) is constant and all terms of the form $X(Y, Z)$ vanish. As Z was arbitrary (G is parallelizable, so we can always find a basis of left invariant vector fields at any point), we conclude that

$$2\nabla_X Y = [X, Y].$$

c. From Part (b) and the Jacobi identity,

$$\begin{aligned}R(X, Y)Z &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z\\ &= \frac{1}{4}([X, [Y, Z]] - [Y, [X, Z]]) - \frac{1}{2}[[X, Y], Z]\\ &= -\frac{1}{4}[[X, Y], Z].\end{aligned}$$

d. From Part (c) and Exercise 3.52c,

$$(X, R(X, Y)Y) = -\frac{1}{4}(X, [[X, Y], Y]) = \frac{1}{4}([X, Y], [X, Y]).$$

Now apply (8.120).

e. According to (8.53), the Ricci tensor is given by

$$\operatorname{Ric}(X, Y) = \operatorname{tr}(Z \to R(Z, Y)X),$$

so using the result of Part (c) gives

$$\begin{aligned}\operatorname{Ric}(X, Y) &= -\frac{1}{4}\operatorname{tr}(Z \to [[Z, Y], X])\\ &= -\frac{1}{4}\operatorname{tr}(Z \to \operatorname{ad} X \operatorname{ad} Y \circ Z)\\ &= -\frac{1}{4}\operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y)\\ &= -\frac{1}{4}(X, Y).\end{aligned}$$

8.51 a. We have

$$
\begin{aligned}
&d\theta^{\hat{1}} = 0,\\
&d\theta^{\hat{2}} = f'\,dr \wedge d\theta = \frac{f'}{f}\theta^{\hat{1}} \wedge \theta^{\hat{2}},\\
&d\theta^{\hat{3}} = f' \sin\theta\,dr \wedge d\phi + f \cos\theta\,d\theta \wedge d\phi = \frac{f'}{f}\,\theta^{\hat{1}} \wedge \theta^{\hat{3}} + \frac{1}{f}\cot\theta\,\theta^{\hat{2}} \wedge \theta^{\hat{3}}.
\end{aligned}
$$

From $d\theta^{\hat{a}} = -\omega^{\hat{a}}{}_{\hat{b}} \wedge \theta^{b}$ and the properties of the connection matrix in an orthonormal frame we guess and verify that

$$
\begin{aligned}
\omega^{\hat{1}}{}_{\hat{2}} &= -\omega^{\hat{2}}{}_{\hat{1}} = -\frac{f'}{f}\theta^{\hat{2}} = -f'\,d\theta,\\
\omega^{\hat{1}}{}_{\hat{3}} &= -\omega^{\hat{3}}{}_{\hat{1}} = -\frac{f'}{f}\theta^{\hat{3}} = -f'\sin\theta\,d\phi,\\
\omega^{\hat{2}}{}_{\hat{3}} &= -\omega^{\hat{3}}{}_{\hat{2}} = -\frac{1}{f}\cot\theta\,\theta^{\hat{3}} = -\cos\theta\,d\phi.
\end{aligned}
$$

b.

$$
\begin{aligned}
\Omega^{\hat{1}}{}_{\hat{2}} &= d\omega^{\hat{1}}{}_{\hat{2}} + \omega^{\hat{1}}{}_{\hat{3}} \wedge \omega^{\hat{3}}{}_{\hat{2}}\\
&= -f''\,dr \wedge d\theta = -\frac{f''}{f}\,\theta^{\hat{1}} \wedge \theta^{\hat{2}},\\
\Omega^{\hat{1}}{}_{\hat{3}} &= d\omega^{\hat{1}}{}_{\hat{3}} + \omega^{\hat{1}}{}_{\hat{2}} \wedge \omega^{\hat{2}}{}_{\hat{3}}\\
&= -f''\sin\theta\,dr \wedge d\phi - f'\cos\theta\,d\theta \wedge d\phi + f'\cos\theta\,d\theta \wedge d\phi\\
&= -\frac{f''}{f}\,\theta^{\hat{1}} \wedge \theta^{\hat{3}},\\
\Omega^{\hat{2}}{}_{\hat{3}} &= d\omega^{\hat{2}}{}_{\hat{3}} + \omega^{\hat{2}}{}_{\hat{1}} \wedge \omega^{\hat{1}}{}_{\hat{3}}\\
&= \sin\theta\,d\theta \wedge d\phi - (f')^2 \sin\theta\,d\theta \wedge d\phi\\
&= \frac{1 - f'^2}{f^2}\,\theta^{\hat{2}} \wedge \theta^{\hat{3}}.
\end{aligned}
$$

c. By Schur's theorem we must have

$$
-\frac{f''}{f} = K = \frac{1 - f'^2}{f^2}.
$$

If $K = -1$ the first equation gives $f'' - f = 0$ which has solution $f = a\sinh r + b\cosh r$. Applying the boundary condition requires $b = 0$. But the second equation requires $f'^2 - f^2 = 1$, so $a = \pm 1$ and $f = \pm\sinh r$. If $K = 0$ the first equation gives $f = ar + b$, while the second requires

$f' = \pm 1$, so $a = \pm 1$ and $f = \pm r$. Lastly, if $K = 1$ a similar argument gives $f = \pm \sin r$.

8.52 a. Laplace expansion and the definition of the Levi-Civita alternating symbol gives

$$\varepsilon_{i_1 \dots i_n} A^{i_1}{}_1 \cdots A^{i_n}{}_n = \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma A^{\sigma(1)}{}_1 \cdots A^{\sigma(n)}{}_n = \det A.$$

Now the claim is that

$$\varepsilon_{i_1 \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_n}{}_{j_n}$$

is totally antisymmetric under the interchange of the j's. We see this as follows:

$$\begin{aligned}
&\varepsilon_{i_1 \dots i_k \dots i_\ell \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_k}{}_{j_k} \cdots A^{i_\ell}{}_{j_\ell} \cdots A^{i_n}{}_{j_n} \\
&= \varepsilon_{i_1 \dots i_\ell \dots i_k \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_\ell}{}_{j_k} \cdots A^{i_k}{}_{j_\ell} \cdots A^{i_n}{}_{j_n} && \text{(flip dummy indices)} \\
&= \varepsilon_{i_1 \dots i_\ell \dots i_k \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_k}{}_{j_\ell} \cdots A^{i_\ell}{}_{j_k} \cdots A^{i_n}{}_{j_n} && \text{(commute components)} \\
&= -\varepsilon_{i_1 \dots i_k \dots i_\ell \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_k}{}_{j_\ell} \cdots A^{i_\ell}{}_{j_k} \cdots A^{i_n}{}_{j_n} && \text{(antisymmetry of } \varepsilon\text{).}
\end{aligned}$$

Moreover, when $(j_1, \dots, j_n) = (1, \dots, n)$ it gives $\det A$, so

$$\varepsilon_{i_1 \dots i_n} A^{i_1}{}_{j_1} \cdots A^{i_n}{}_{j_n} = \varepsilon_{j_1 \dots j_n} \det \boldsymbol{A}.$$

b. By the result of Part (a),

$$\frac{\partial x^{i_1}}{\partial y^{i'_1}} \cdots \frac{\partial x^{i_n}}{\partial y^{i'_n}} \varepsilon_{i_1 \dots i_n} = J^{-1} \varepsilon_{i'_1 \dots i'_n},$$

so $\xi = J$.

c. Under a coordinate transformation,

$$g_{i'j'} = \frac{\partial x^i}{\partial y^{i'}} \frac{\partial x^j}{\partial y^{j'}} g_{ij}.$$

Taking the determinant of both sides gives $G' = J^{-2} G$.

d. Under a coordinate transformation we get

$$\begin{aligned}
\epsilon_{i'_i \dots i'_n} &= \sqrt{|G'|} \varepsilon_{i'_1 \dots i'_n} \\
&= (|J|^{-1} \sqrt{|G|}) J \frac{\partial x^{i_1}}{\partial y^{i'_1}} \cdots \frac{\partial x^{i_n}}{\partial y^{i'_n}} \varepsilon_{i_1 \dots i_n} \\
&= (\operatorname{sgn} J) \frac{\partial x^{i_1}}{\partial y^{i'_1}} \cdots \frac{\partial x^{i_n}}{\partial y^{i'_n}} \epsilon_{i_1 \dots i_n}.
\end{aligned}$$

e. We have

$$\varepsilon_{i_1\dots i_n}\, dx^{i_1} \wedge \dots \wedge dx^{i_n} = n!\, dx^1 \wedge \dots \wedge dx^n,$$

so

$$\frac{1}{n!}\epsilon_{i_1\dots i_n}\, dx^{i_1} \wedge \dots \wedge dx^{i_n} = \sqrt{|G|}\, dx^1 \wedge \dots \wedge dx^n = \sigma.$$

f. Using Part (a) again gives

$$\frac{\partial y^{i'_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{i'_n}}{\partial x^{i_n}} \varepsilon^{i_1\dots i_n} = J\varepsilon^{i'_1\dots i'_n},$$

so $\eta = J^{-1}$.

g. Again by Part (a),

$$\begin{aligned}
\epsilon^{i_1\dots i_n} &= g^{i_1 j_1} \cdots g^{i_n j_n} \epsilon_{j_1\dots j_n} \\
&= \sqrt{|G|} g^{i_1 j_1} \cdots g^{i_n j_n} \varepsilon_{j_1\dots j_n} \\
&= \sqrt{|G|} G^{-1} \varepsilon^{j_1\dots j_n} \\
&= \pm(\operatorname{sgn} G)\epsilon^{j_1\dots j_n}.
\end{aligned}$$

Therefore we must choose $+1$ in the Euclidean case and -1 in the Lorentzian case.

h. The covariant derivative of the Levi-Civita tensor vanishes because it is built from the metric, and the Levi-Civita connection is metric compatible. For instance, we have

$$\nabla_k \epsilon_{a_1\dots a_n} = \nabla_k(\sqrt{G}\varepsilon_{a_1\dots a_n}) = \nabla_k(\sqrt{G})\varepsilon_{a_1\dots a_n} = 0,$$

because $\varepsilon_{a_1\dots a_n} = -1, 0, 1$ and the determinant of the metric is built from the metric components.

8.53 a. $\varepsilon^{i_1\dots i_n}\varepsilon_{j_1\dots j_n}$ and $\delta^{i_1\dots i_n}_{j_1\dots j_n}$ have the same symmetry properties, namely antisymmetry under the interchange of any pair of i or j indices. This is true for the Levi-Civita symbols by definition, and true for the determinant, because swapping two i indices swaps two columns, while swapping two j indices swaps to rows. Moreover, if $i_s = j_s = s$ then both sides equal 1. Hence two two quantities must be equal.

To prove the identity in general, proceed by reverse induction, assuming it holds for $k+1$ and proving it for k. By hypothesis,

$$\varepsilon^{i_1\dots i_{k+1} i_{k+2}\dots i_n}\varepsilon_{j_1\dots j_{k+1} i_{k+2}\dots i_n} = (n-k-1)!\delta^{i_1\dots i_{k+1}}_{j_1\dots j_{k+1}}.$$

This gives, by Laplace expansion on the last column,

$$
\begin{aligned}
&\frac{1}{(n-k-1)!}\varepsilon^{i_1\dots i_k i_{k+1}\dots i_n}\varepsilon_{j_1\dots j_k i_{k+1}\dots i_n} \\
&\quad = \det\left(\begin{array}{ccc|c}
\delta^{i_1}_{j_1} & \cdots & \delta^{i_1}_{j_k} & \delta^{i_1}_{i_{k+1}} \\
\vdots & \ddots & \vdots & \vdots \\
\delta^{i_k}_{j_1} & \cdots & \delta^{i_k}_{j_k} & \delta^{i_k}_{i_{k+1}} \\
\hline
\delta^{i_{k+1}}_{j_1} & \cdots & \delta^{i_{k+1}}_{j_k} & \delta^{i_{k+1}}_{i_{k+1}}
\end{array}\right) \\
&\quad = (-1)^k\big[\delta^{i_1}_{i_{k+1}}\delta^{\hat{i}_1 i_2\dots i_{k+1}}_{j_1\dots j_k} - \delta^{i_2}_{i_{k+1}}\delta^{i_1\hat{i}_2\dots i_{k+1}}_{j_1\dots j_k} \\
&\qquad \cdots + (-1)^{k-1}\delta^{i_k}_{i_{k+1}}\delta^{i_1\dots\hat{i}_k i_{k+1}}_{j_1\dots j_k} + (-1)^k\delta^{i_{k+1}}_{i_{k+1}}\delta^{i_1\dots i_k}_{j_1\dots j_k}\big] \\
&\quad = (-1)^k\big[\delta^{\hat{i}_1 i_2\dots i_k i_1}_{j_1\dots j_k} - \delta^{i_1\hat{i}_2\dots i_k i_2}_{j_1\dots j_k} \\
&\qquad \cdots + (-1)^{k-1}\delta^{i_1\dots i_k}_{j_1\dots j_k} + (-1)^k n\delta^{i_1\dots i_k}_{j_1\dots j_k}\big] \\
&\quad = \big[\underbrace{-1-1\cdots-1}_{k\text{ terms}}+n\big]\delta^{i_1\dots i_k}_{j_1\dots j_k} \\
&\quad = (n-k)\delta^{i_1\dots i_k}_{j_1\dots j_k}.
\end{aligned}
$$

In the third line the caret indicates the absence of that index. The signs are derived by permuting the indices in the δ symbols back to their canonical forms.

Actually, there is a much easier way to prove the identity in general. Begin again from result for $k = n$ (base case of the induction). Next, observe that for each fixed sequence of indices $(i_{k+1}, \dots, i_n)$, the remaining indices $(i_1, \dots, i_k)$ and $(j_1, \dots, j_k)$ must be permutations of one another, and so the base case implies that, for any fixed collection $(i_{k+1}, \dots, i_n)$ we have

$$\varepsilon^{i_1\dots i_k i_{k+1}\dots i_n}\varepsilon_{j_1\dots j_k i_{k+1}\dots i_n} = \delta^{i_1\dots i_k}_{j_1\dots j_k}.$$

(In this formula repeated indices are not summed.) Now we just observe that there are $(n-k)!$ choices for the index set $(i_{k+1}, \dots, i_n)$.

b. We have

$$\epsilon^{i_1\dots i_k i_{k+1}\dots i_n}\epsilon_{j_1\dots j_k i_{k+1}\dots i_n} = \pm(n-k)!\delta^{i_1\dots i_k}_{j_1\dots j_k},$$

where we must choose the plus sign in Riemannian case, and the minus sign in the Lorentzian case.

8.54 Recall the definition (8.91) of the Hodge dual:

$$\alpha \wedge \beta = g(\star\alpha, \beta)\sigma,$$

where σ is the canonical volume element. Let $\beta := dx^{i_{k+1}} \wedge \cdots \wedge dx^{i_n}$. Then

$$\begin{aligned}\alpha \wedge \beta &= \frac{1}{k!} a_{i_1 \dots i_k}\, dx^{i_1} \wedge \cdots \wedge dx^{i_n} \\ &= \frac{1}{k!} a_{i_1 \dots i_k} \varepsilon^{i_1 \dots i_n}\, dx^1 \wedge \cdots \wedge dx^n \\ &= \frac{1}{k!} \frac{1}{\sqrt{|G|}} a_{i_1 \dots i_k} \varepsilon^{i_1 \dots i_n} \sigma \\ &= \frac{1}{k!} \epsilon^{i_1 \dots i_n} a_{i_1 \dots i_k} \sigma.\end{aligned}$$

On the other hand,

$$\begin{aligned}g(\star\alpha, \beta)\sigma &= \frac{1}{(n-k)!} a^*_{i_{k+1} \dots i_n} g(\beta, \beta)\sigma \\ &= \frac{1}{(n-k)!} a^*_{i_{k+1} \dots i_n} g(\beta, \beta)\sigma \\ &= \frac{1}{(n-k)!} (a^*)^{j_{k+1} \cdots j_n} g_{j_{k+1} i_{k+1}} \cdots g_{j_n i_n} g(\beta, \beta)\sigma.\end{aligned}$$

But

$$\begin{aligned}g(\beta, \beta) &= \det \begin{pmatrix} g(dx^{i_{k+1}}, dx^{i_{k+1}}) & \cdots & g(dx^{i_{k+1}}, dx^{i_n}) \\ \vdots & \ddots & \vdots \\ g(dx^{i_n}, dx^{i_{k+1}}) & \cdots & g(dx^{i_n}, dx^{i_n}) \end{pmatrix} \\ &= \det \begin{pmatrix} g^{i_{k+1} i_{k+1}} & \cdots & g^{i_{k+1} i_n} \\ \vdots & \ddots & \vdots \\ g^{i_n i_{k+1}} & \cdots & g^{i_n i_n} \end{pmatrix},\end{aligned}$$

so

$$g_{j_{k+1} i_{k+1}} \cdots g_{j_n i_n} g(\beta, \beta) = \delta^{i_{k+1} \cdots i_n}_{j_{k+1} \cdots j_n}.$$

Hence

$$g(\star\alpha, \beta)\sigma = \frac{1}{(n-k)!} (a^*)^{j_{k+1}, \dots, j_n} \delta^{i_{k+1} \cdots i_n}_{j_{k+1} \cdots j_n} \sigma = (a^*)^{i_{k+1} \dots i_n} \sigma,$$

and the claim is proved.

8.55 From Exercises 8.52, 8.53 and 8.54 we have

$$\begin{aligned}\alpha \wedge \star\beta &= \frac{1}{k!(n-k)!} a_{i_1 \dots i_k} b^*_{i_{k+1} \dots i_n}\, dx^{i_1} \wedge \cdots \wedge dx^{i_n} \\ &= \frac{1}{(k!)^2 (n-k)!} a_{i_1 \dots i_k} b^{j_1 \cdots j_k} \epsilon_{j_1 \dots j_k i_{k+1} \dots i_n}\, dx^{i_1} \wedge \cdots \wedge dx^{i_n} \\ &= \frac{1}{(k!)^2 (n-k)!} a_{i_1 \dots i_k} b^{j_1 \cdots j_k} \sqrt{|G|} \varepsilon_{j_1 \dots j_k i_{k+1} \dots i_n} \varepsilon^{i_1 \dots i_n}\, dx^1 \wedge \cdots \wedge dx^n\end{aligned}$$

$$= \frac{1}{(k!)^2} a_{i_1\dots i_k} b^{j_1\dots j_k} \delta^{i_1\dots i_k}_{j_1\dots j_k} \sigma$$
$$= \frac{1}{k!} a_{i_1\dots i_k} b^{j_1\dots j_k} \sigma.$$

8.56 a. According to Exercise 1.36,

$$\frac{\partial G}{\partial g_{ij}} = G g^{ij}.$$

By the chain rule

$$\frac{\partial G}{\partial x^k} = \frac{\partial G}{\partial g_{ij}} \frac{\partial g_{ij}}{\partial x^k} = G g^{ij} g_{ij,k}.$$

Recall that

$$\Gamma_{kij} = \frac{1}{2}(g_{jk,i} + g_{ik,j} - g_{ij,k}),$$

and observe that

$$g^{ik} g_{jk,i} = g^{ik} g_{ji,k}.$$

Thus,

$$\Gamma^i{}_{ij} = g^{ik}\Gamma_{kij} = \frac{1}{2} g^{ik} g_{ik,j} = \frac{1}{2} G^{-1} \frac{\partial G}{\partial x^j}.$$

It follows that

$$\frac{\partial G^{1/2}}{\partial x^j} = \frac{1}{2} G^{-1/2} \frac{\partial G}{\partial x^j} = G^{1/2} \Gamma^i{}_{ij},$$

and therefore

$$\operatorname{div} X = \nabla_i X^i = \partial_i X^i + \Gamma^i{}_{ji} X^j = G^{-1/2}(G^{1/2} X^j)_{,j}.$$

b. We have

$$\mathcal{L}_X \sigma = \mathcal{L}_X(G^{1/2})\, dx^1 \wedge \cdots \wedge dx^n + G^{1/2} \mathcal{L}_X(dx^1 \wedge \cdots \wedge dx^n).$$

For the first term, we have

$$\mathcal{L}_X(G^{1/2}) = X(G^{1/2}) = X^j (G^{1/2})_{,j}.$$

For the second, we note that

$$\mathcal{L}_X(dx^j) = d(X(x^j)) = dX^j = \sum_k X^j{}_{,k}\, dx^k,$$

so

$$\mathcal{L}_X(dx^1 \wedge \cdots \wedge dx^n) = \sum_{j=1}^{n} dx^1 \wedge \cdots \wedge \mathcal{L}_X(dx^j) \wedge \cdots \wedge dx^n$$
$$= X^j{}_{,j}\, dx^1 \wedge \cdots \wedge dx^n.$$

Adding together the two terms and using the result of Part (a) gives

$$\mathcal{L}_X\sigma = (X^j(G^{1/2})_{,j} + G^{1/2}X^j{}_{,j})\, dx^1 \wedge \cdots \wedge dx^n = (\operatorname{div} X)\sigma.$$

c. From the Ricci identity,

$$\operatorname{div} \circ \operatorname{curl} X = \nabla_k(\epsilon^{ijk}\nabla_i X_j) = \epsilon^{ijk}\nabla_{[k}\nabla_{i]}X_j = -\frac{1}{2}\epsilon^{ijk}R^q{}_{jki}X_q = 0,$$

by virtue of the symmetry properties of the Riemann tensor. Also,

$$(\operatorname{curl} \circ \operatorname{grad} f)^k = \epsilon^{ijk}\nabla_i\nabla_j f = \epsilon^{ijk}\nabla_i\partial_j f = \epsilon^{ijk}(\partial_i\partial_j f - \Gamma^\ell{}_{ij}\partial_\ell f) = 0,$$

because mixed partials commute and the Christoffel symbols are symmetric in the lower indices.

8.57 a. This is just a change of variables problem. By definition,

$$g_{i'j'} = \frac{\partial x^i}{\partial y^{i'}}\frac{\partial x^j}{\partial y^{j'}}g_{ij},$$

where the primed coordinates are the spherical polar ones and $g_{ij} = \delta_{ij}$. We have

$$\begin{array}{lll}
\dfrac{\partial x}{\partial r} = \sin\theta\cos\phi & \dfrac{\partial y}{\partial r} = \sin\theta\sin\phi & \dfrac{\partial z}{\partial r} = \cos\theta \\
\dfrac{\partial x}{\partial \theta} = r\cos\theta\cos\phi & \dfrac{\partial y}{\partial \theta} = r\cos\theta\sin\phi & \dfrac{\partial z}{\partial \theta} = -r\sin\theta \\
\dfrac{\partial x}{\partial \phi} = -r\sin\theta\sin\phi & \dfrac{\partial y}{\partial \phi} = r\sin\theta\cos\phi & \dfrac{\partial z}{\partial \phi} = 0,
\end{array}$$

so the diagonal elements are

$$g_{rr} = \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 + \left(\frac{\partial z}{\partial r}\right)^2 = 1,$$
$$g_{\theta\theta} = \left(\frac{\partial x}{\partial \theta}\right)^2 + \left(\frac{\partial y}{\partial \theta}\right)^2 + \left(\frac{\partial z}{\partial \theta}\right)^2 = r^2,$$
$$g_{\phi\phi} = \left(\frac{\partial x}{\partial \phi}\right)^2 + \left(\frac{\partial y}{\partial \phi}\right)^2 + \left(\frac{\partial z}{\partial \phi}\right)^2 = r^2\sin^2\theta,$$

and all the off diagonal terms vanish.

b. We have, for example,

$$g(e_{\hat\phi}, e_{\hat\phi}) = \frac{1}{r^2 \sin^2\theta} g(\partial_\phi, \partial_\phi) = \frac{1}{r^2 \sin^2\theta} g_{\phi\phi} = 1.$$

In a similar fashion we can check that

$$g(e_{\hat i}, e_{\hat j}) = \delta_{\hat i \hat j},$$

so the basis is indeed orthonormal.

c. For future use we observe that

$$G^{1/2} = r^2 \sin\theta.$$

Also,

$$X = X^r \frac{\partial}{\partial r} + X^\theta \frac{\partial}{\partial\theta} + X^\phi \frac{\partial}{\partial\phi} = X^{\hat r} e_{\hat r} + X^{\hat\theta} e_{\hat\theta} + X^{\hat\phi} e_{\hat\phi},$$

so

$$X^r = X^{\hat r}, \quad X^\theta = \frac{1}{r} X^{\hat\theta}, \quad \text{and} \quad X^\phi = \frac{1}{r\sin\theta} X^{\hat\phi}.$$

Also,

$$\begin{aligned} X_r &= g_{rr} X^r = X^{\hat r}, \\ X_\theta &= g_{\theta\theta} X^\theta = r X^{\hat\theta}, \\ X_\phi &= g_{\phi\phi} X^\phi = r \sin\theta X^{\hat\phi}. \end{aligned}$$

First, the divergence. We have

$$\operatorname{div} X = G^{-1/2} (G^{1/2} X^j)_{,j}.$$

Now

$$\begin{aligned} (G^{1/2} X^r)_{,r} &= (r^2 X^{\hat r})_{,r} \sin\theta, \\ (G^{1/2} X^\theta)_{,\theta} &= r (\sin\theta\, X^{\hat\theta})_{,\theta}, \\ (G^{1/2} X^\phi)_{,\phi} &= X^{\hat\phi}{}_{,\phi}, \end{aligned}$$

and therefore

$$\operatorname{div} X = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 X^{\hat r} \right) + \frac{1}{r\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta X^{\hat\theta} \right) + \frac{1}{r\sin\theta} \frac{\partial}{\partial\phi} X^{\hat\phi}.$$

Next, the curl. We have

$$\begin{aligned} \operatorname{curl} X &= \epsilon^{ijk} (\nabla_i X_j) \partial_k = \epsilon^{ijk} (\partial_i X_j - \Gamma^m{}_{ij} X_m) \partial_k \\ &= \epsilon^{ijk} (\partial_i X_j) \partial_k = G^{-1/2} \varepsilon^{ijk} (\partial_i X_j) \partial_k, \end{aligned}$$

by virtue of the symmetry of the Christoffel symbols in the last two indices. Now

$$
\begin{aligned}
(X_r)_{,\theta} &= (X^{\hat{r}})_{,\theta}, \\
(X_r)_{,\phi} &= (X^{\hat{r}})_{,\phi}, \\
(X_\theta)_{,r} &= \left(rX^{\hat{\theta}}\right)_{,r}, \\
(X_\theta)_{,\phi} &= r\left(X^{\hat{\theta}}\right)_{,\phi}, \\
(X_\phi)_{,r} &= \sin\theta\left(rX^{\hat{\phi}}\right)_{,r}, \\
(X_\phi)_{,\theta} &= r\left(\sin\theta X^{\hat{\phi}}\right)_{,\theta}.
\end{aligned}
$$

Thus we have the following.

$$
\begin{aligned}
(\text{curl } X)^r &= G^{-1/2}(\partial_\theta X_\phi - \partial_\phi X_\theta) \\
&= \frac{1}{r^2\sin\theta}\left[r\left(\sin\theta X^{\hat{\phi}}\right)_{,\theta} - r\left(X^{\hat{\theta}}\right)_{,\phi}\right] \\
&= \frac{1}{r\sin\theta}\left[\left(\sin\theta X^{\hat{\phi}}\right)_{,\theta} - \left(X^{\hat{\theta}}\right)_{,\phi}\right] \\
&= (\text{curl } X)^{\hat{r}}.
\end{aligned}
$$

Also,

$$
\begin{aligned}
(\text{curl } X)^\theta &= G^{-1/2}(\partial_\phi X_r - \partial_r X_\phi) \\
&= \frac{1}{r^2\sin\theta}\left[X^{\hat{r}}{}_{,\phi} - \sin\theta\left(rX^{\hat{\phi}}\right)_{,r}\right] \\
&= \frac{1}{r}(\text{curl } X)^{\hat{\theta}},
\end{aligned}
$$

which implies

$$
(\text{curl } X)^{\hat{\theta}} = \frac{1}{r\sin\theta}\left[X^{\hat{r}}{}_{,\phi} - \sin\theta\left(rX^{\hat{\phi}}\right)_{,r}\right].
$$

Lastly,

$$
\begin{aligned}
(\text{curl } X)^\phi &= G^{-1/2}(\partial_r X_\theta - \partial_\theta X_r) \\
&= \frac{1}{r^2\sin\theta}\left[\left(rX^{\hat{\theta}}\right)_{,r} - X^{\hat{r}}{}_{,\theta}\right] \\
&= \frac{1}{r\sin\theta}(\text{curl } X)^{\hat{\phi}},
\end{aligned}
$$

so

$$
(\text{curl } X)^{\hat{\phi}} = \frac{1}{r}\left[\left(rX^{\hat{\theta}}\right)_{,r} - X^{\hat{r}}{}_{,\theta}\right].
$$

Putting everything together, we obtain

$$\begin{aligned}\operatorname{curl} X = \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta X^{\hat\phi}) - \frac{\partial X^{\hat\theta}}{\partial\phi}\right] e_r \\ + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial X^{\hat r}}{\partial\phi} - \frac{\partial}{\partial r}(rX^{\hat\phi})\right] e_\theta \\ + \frac{1}{r}\left[\frac{\partial}{\partial r}(rX^{\hat\theta}) - \frac{\partial X^{\hat r}}{\partial\theta}\right] e_\phi.\end{aligned}$$

For the gradient, we have

$$\begin{aligned}\operatorname{grad} f &= g^{ij}(\nabla_i f)\partial_j = g^{ij}(\partial_i f)\partial_j \\ &= (\partial_r f)e_{\hat r} + \frac{1}{r^2}(\partial_\theta f) r e_{\hat\theta} + \frac{1}{r^2\sin^2\theta}(\partial_\phi f)(r\sin\theta)e_{\hat\phi} \\ &= \frac{\partial f}{\partial r}e_{\hat r} + \frac{1}{r}\frac{\partial f}{\partial\theta}e_{\hat\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}e_{\hat\phi}.\end{aligned}$$

Finally, the Laplacian may be read off immediately by taking the divergence of the gradient of f.

8.58 a. We start with

$$\alpha = \frac{1}{k!}a_{i_1\dots i_k}\, dx^{i_1}\wedge\cdots\wedge dx^{i_k}$$

and apply the exterior differential to get

$$\begin{aligned}d\alpha &= \frac{1}{k!}a_{i_1\dots i_k,j}\, dx^j\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k} \\ &= \frac{1}{k!}\left(a_{i_1\dots i_k;j} + \Gamma^\ell{}_{ji_1}a_{\ell i_2\dots i_k} + \cdots + \Gamma^\ell{}_{ji_k}a_{i_1\dots i_{k-1}\ell}\right) \\ &\qquad \times\, dx^j\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k} \\ &= \frac{1}{k!}a_{i_1\dots i_k;j}\, dx^j\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k} \\ &= \frac{1}{k!}\nabla_j a_{i_1\dots i_k}\, dx^j\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k},\end{aligned}$$

where we used the symmetry of the Christoffel symbols in the last two indices.

The second expression for $d\alpha$ given in the problem holds by virtue of the antisymmetry of the wedge product, as discussed in Exercise 2.4.

To get the last expression, we start from

$$\nabla_{[i_1}a_{i_2\dots i_{k+1}]} = \frac{1}{(k+1)!}\sum_{\sigma\in\mathfrak{S}_{k+1}}(-1)^\sigma\nabla_{i_{\sigma(1)}}a_{i_{\sigma(2)}\dots i_{\sigma(k+1)}}$$

and note that $a_{i_1 \dots i_k}$ is already totally antisymmetric. We can split up the sum over $\mathfrak{S}_{k+1}$ by fixing a value for $\sigma(1)$ and then summing over all $k!$ permutations of the remaining indices. The sign of such a permutation is just the product of the sign of the identity permutation with one element moved to the front times the sign of the rest of the permutation. For example, the sign of the permutation 2143 is -1, the sign of the permutation 143 is -1, and their product is $+1$. Reasoning in this way, we get

$$\nabla_{[i_1} a_{i_2 \dots i_{k+1}]} = \frac{1}{(k+1)!} \sum_{r=1}^{k+1} (-1)^{r-1} \sum_{\pi \in \mathfrak{S}(r)} (-1)^{\pi} \nabla_{i_r} a_{i_{\pi(1)} \dots i_{\pi(k)}},$$

where $\mathfrak{S}(r)$ stands for the permutation subgroup that fixes the first element. Summing over the elements in each subgroup gives

$$\nabla_{[i_1} a_{i_2 \dots i_{k+1}]} = \frac{1}{(k+1)} \sum_{r=1}^{k+1} (-1)^{r-1} \nabla_{i_r} a_{i_1 \dots \hat{i}_r \dots i_{k+1}}.$$

b. By definition, $\delta\alpha = (-1)^{nk+n+1} \star d \star \alpha$. We write

$$\begin{aligned}
\alpha &= \frac{1}{k!} a_{i_1 \dots i_k} \, dx^{i_1} \wedge \dots \wedge dx^{i_k}, \\
\star\alpha &= \frac{1}{k!(n-k)!} a^{i_1 \dots i_k} \epsilon_{i_1 \dots i_n} \, dx^{i_{k+1}} \wedge \dots \wedge dx^{i_n}, \\
d \star \alpha &= \frac{1}{k!(n-k)!} \left(a^{i_1 \dots i_k} \epsilon_{i_1 \dots i_n} \right)_{;j} \, dx^j \wedge dx^{i_{k+1}} \wedge \dots \wedge dx^{i_n} \\
&= \frac{1}{k!(n-k)!} \nabla_{j_1} \left(a^{i_1 \dots i_k} \epsilon_{i_1 \dots i_k j_2 \dots j_{n-k+1}} \right) dx^{j_1} \wedge \dots \wedge dx^{j_{n-k+1}} \\
&= b_{j_1 \dots j_{n-k+1}} \, dx^{j_1} \wedge \dots \wedge dx^{j_{n-k+1}},
\end{aligned}$$

where

$$b_{j_1 \dots j_{n-k+1}} := \frac{1}{k!(n-k)!} \nabla_{j_1} \left(a^{i_1 \dots i_k} \epsilon_{i_1 \dots i_k j_2 \dots j_{n-k+1}} \right).$$

Thus we have

$$\begin{aligned}
\star d \star \alpha &= \frac{1}{[n-(n-k+1)]!} b^{j_1 \dots j_{n-k+1}} \epsilon_{j_1 \dots j_n} \, dx^{j_{n-k+2}} \wedge \dots \wedge dx^{j_n} \\
&= \frac{1}{(k-1)!} b_{p_1 \dots p_{n-k+1}} \epsilon_{j_1 \dots j_n} \\
&\qquad \times g^{p_1 j_1} \dots g^{p_{n-k+1} j_{n-k+1}} \, dx^{j_{n-k+2}} \wedge \dots \wedge dx^{j_n} \\
&= \frac{1}{k!(n-k)!(k-1)!} \nabla_{p_1} \left(a^{i_1 \dots i_k} \epsilon_{i_1 \dots i_k p_2 \dots p_{n-k+1}} \right) \epsilon_{j_1 \dots j_n} \\
&\qquad \times g^{p_1 j_1} \dots g^{p_{n-k+1} j_{n-k+1}} \, dx^{j_{n-k+2}} \wedge \dots \wedge dx^{j_n}
\end{aligned}$$

$$= \frac{1}{k!(n-k)!(k-1)!} g^{p_1 j_1} \nabla_{p_1} \left(a_{i_1 \dots i_k} \epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \right) \epsilon_{j_1 \dots j_n} \times dx^{j_{n-k+2}} \wedge \cdots \wedge dx^{j_n},$$

where we used metric compatibility. Applying the Leibniz rule we have

$$\begin{aligned} \nabla_{p_1} &\left(a_{i_1 \dots i_k} \epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \right) \epsilon_{j_1 \dots j_n} \\ &= \nabla_{p_1} (a_{i_1 \dots i_k}) \epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \epsilon_{j_1 \dots j_n} \\ &\quad + a_{i_1 \dots i_k} \nabla_{p_1} \left(\epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \right) \epsilon_{j_1 \dots j_n}. \end{aligned}$$

The second term vanishes, while the first term becomes, after permuting some indices,

$$\begin{aligned} \nabla_{p_1} &(a_{i_1 \dots i_k}) \epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \epsilon_{j_1 \dots j_n} \\ &= (-1)^{(n-k)(k-1)} \nabla_{p_1} (a_{i_1 \dots i_k}) \epsilon^{i_1 \dots i_k j_2 \cdots j_{n-k+1}} \epsilon_{j_1 j_{n-k+2} \cdots j_n j_2 \cdots j_{n-k+1}} \\ &= (-1)^{(n-k)(k-1)} (n-k)! \nabla_{p_1} (a_{i_1 \dots i_k}) \delta^{i_1 \dots i_k}_{j_1 j_{n-k+2} \cdots j_n} \\ &= (-1)^{(n-k)(k-1)} k!(n-k)! \nabla_{p_1} (a_{j_1 j_{n-k+2} \cdots j_n}). \end{aligned}$$

Therefore

$$\begin{aligned} \delta\alpha &= (-1)^{nk+n+1} \star d \star \alpha \\ &= -\frac{1}{(k-1)!} g^{p_1 j_1} \nabla_{p_1} (a_{j_1 j_{n-k+2} \dots j_n}) \\ &\qquad \times dx^{j_{n-k+2}} \wedge \cdots \wedge dx^{j_n}. \end{aligned}$$

c. Define

$$\beta = \delta\alpha = \frac{1}{(k-1)!} \beta_{i_2 \dots i_k} \, dx^{i_2} \wedge \cdots \wedge dx^{i_k},$$

where

$$\beta_{i_2 \dots i_k} := -\nabla^j a_{j i_2 \dots i_k}.$$

Then

$$\begin{aligned} d\delta\alpha &= d\beta \\ &= \frac{1}{k!} \sum_{r=1}^{k} (-1)^{r-1} \nabla_{i_r} \beta_{i_1 \dots \hat{i}_r \dots i_k} \, dx^{i_1} \wedge \cdots \wedge dx^{i_k} \\ &= \frac{1}{k!} \sum_{r=1}^{k} (-1)^{r} \nabla_{i_r} \nabla^j a_{j i_1 \dots \hat{i}_r \dots i_k} \, dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \end{aligned}$$

where we coalesced the signs.

Next, if

$$\beta = \frac{1}{(k+1)!} b_{i_1 \dots i_{k+1}} \, dx^{i_1} \wedge \cdots \wedge dx^{i_{k+1}},$$

then

$$\delta\beta = -\frac{1}{k!} \nabla^j b_{j i_1 \dots i_k} \, dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Defining $\beta := d\alpha$ we have

$$\begin{aligned} b_{i_1 \dots i_{k+1}} &= \sum_{r=1}^{k+1} (-1)^{r-1} \nabla_{i_r} a_{i_1 \dots \hat{i}_r \dots i_{k+1}} \\ &= \nabla_{i_1} a_{i_2 \dots i_{k+1}} + \sum_{r=2}^{k+1} (-1)^{r-1} \nabla_{i_r} a_{i_1 \dots \hat{i}_r \dots i_{k+1}}, \end{aligned}$$

and thus

$$\delta d\alpha = -\frac{1}{k!} \left(\nabla^2 a_{i_1 \dots i_k} + \sum_{r=1}^{k} (-1)^r \nabla^j \nabla_{i_r} a_{j i_1 \dots \hat{i}_r \dots i_k} \right) dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Adding the two expressions together we get

$$\begin{aligned} (\Delta a)_{i_1 \dots i_k} &= -\nabla^2 a_{i_1 \dots i_k} - \sum_{r=1}^{k} (-1)^r \left(\nabla^j \nabla_{i_r} - \nabla_{i_r} \nabla^j \right) a_{j i_1 \dots \hat{i}_r \dots i_k} \\ &= -\nabla^2 a_{i_1 \dots i_k} + \sum_{r=1}^{k} \left(\nabla^j \nabla_{i_r} - \nabla_{i_r} \nabla^j \right) a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_k}, \end{aligned}$$

where we permuted the j using the antisymmetry of the a's. Now we have to employ Ricci's identity (8.109), then massage this expression until it takes the desired form. To this end, we write

$$\begin{aligned} -[\nabla^j, \nabla_{i_r}] & a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_k} \\ &= \sum_{s=1}^{r-1} a_{i_1 \dots i_{s-1} p i_{s+1} \dots i_{r-1} j i_{r+1} \dots i_k} R^p{}_{i_s}{}^j{}_{i_r} \\ &\quad + a_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k} R^p{}_j{}^j{}_{i_r} \\ &\quad + \sum_{s=r+1}^{k} a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_{s-1} p i_{s+1} \dots i_k} R^p{}_{i_s}{}^j{}_{i_r}. \end{aligned}$$

By swapping indices we have

$$R^p{}_j{}^j{}_{i_r} = -R_j{}^{pj}{}_{i_r} = -R^{jp}{}_{j i_r} = -R^p{}_{i_r}.$$

Also,

$$R^{p}{}_{i_s}{}^{j}{}_{i_r} + R^{pj}{}_{i_r i_s} + R^{p}{}_{i_r i_s}{}^{j} = 0,$$

so antisymmetrizing on p and j gives

$$R^{[p}{}_{i_s}{}^{j]}{}_{i_r} = -\frac{1}{2} R^{pj}{}_{i_r i_s} = \frac{1}{2} R^{pj}{}_{i_s i_r}.$$

Hence

$$\begin{aligned}
[\nabla^j, \nabla_{i_r}] & a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_k} \\
&= a_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k} R^{p}{}_{i_r} \\
&\quad - \frac{1}{2} \sum_{s \neq r}^{k} a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_{s-1} p i_{s+1} \dots i_k} R^{jp}{}_{i_r i_s}.
\end{aligned}$$

Putting everything all together yields

$$\begin{aligned}
(\Delta a)_{i_1 \dots i_k} &= -\nabla^2 a_{i_1 \dots i_k} \\
&\quad + \sum_{r=1}^{k} a_{i_1 \dots i_{r-1} p i_{r+1} \dots i_k} R^{p}{}_{i_r} \\
&\quad - \frac{1}{2} \sum_{\substack{r=1 \dots k \\ s=1 \dots k \\ r \neq s}} a_{i_1 \dots i_{r-1} j i_{r+1} \dots i_{s-1} p i_{s+1} \dots i_k} R^{jp}{}_{i_r i_s},
\end{aligned}$$

which is the Weitzenböck formula.

8.59 a. α is harmonic, so

$$\begin{aligned}
0 = k!(\alpha, \Delta\alpha) &= \int_M a^{i_1 \dots i_k} (\Delta a)_{i_1 \dots i_k} \sigma \\
&= \int_M a^{i_1 \dots i_k} \Big(- \nabla^2 a_{i_1 \dots i_k} \\
&\quad + \sum_{r=1}^{k} a_{i_1 \dots i_{r-1} i i_{r+1} \dots i_k} R^{i}{}_{i_r} \\
&\quad - \frac{1}{2} \sum_{\substack{r=1 \dots k \\ s=1 \dots k \\ r \neq s}} a_{i_1 \dots i_{r-1} i i_{r+1} \dots i_{s-1} j i_{s+1} \dots i_k} R^{ij}{}_{i_r i_s} \Big) \sigma.
\end{aligned}$$

We have, by permuting indices,

$$\begin{aligned}
\sum_{r=1}^{k} & a^{i_1 \dots i_k} a_{i_1 \dots i_{r-1} i i_{r+1} \dots i_k} R^{i}{}_{i_r} \\
&= \sum_{r=1}^{k} a^{i_1 \dots i_{r-1} j i_{r+1} \dots i_k} a_{i_1 \dots i_{r-1} i i_{r+1} \dots i_k} R^{i}{}_{j}
\end{aligned}$$

$$= \sum_{r=1}^{k} a^{ji_1\ldots\hat{i}_r\ldots i_k} a_{ii_1\ldots\hat{i}_r\ldots i_k} R^i{}_j$$
$$= ka^{ji_2\ldots i_k} a_{ii_2\ldots\ldots i_k} R^i{}_j.$$

Similarly,

$$\frac{1}{2}\sum_{r\neq s} a^{i_1\ldots i_k} a_{i_1\ldots i_{r-1}ii_{r+1}\ldots i_{s-1}ji_{s+1}\ldots i_k} R^{ij}{}_{i_r i_s}$$
$$= \frac{1}{2}\sum_{r\neq s} a^{i_1\ldots i_r\ldots i_s\ldots i_k} a_{i_1\ldots i_{r-1}ii_{r+1}\ldots i_{s-1}ji_{s+1}\ldots i_k} R^{ij}{}_{i_r i_s}$$
$$= \frac{1}{2}\sum_{r\neq s} a^{pqi_3\ldots i_k} a_{iji_3\ldots i_k} R^{ij}{}_{pq}$$
$$= \frac{1}{2}k(k-1) a^{pqi_3\ldots i_k} a_{iji_3\ldots i_k} R^{ij}{}_{pq}.$$

Hence the claim is proved.

b. By Weitzenböck, for any function f,

$$\Delta f = -\nabla^2 f.$$

But $\Delta f = \delta df$ because $\delta f = 0$. (f is a zero form.) Thus,

$$\int_M (\nabla^2 f)\sigma = -(\delta df, 1) = -(df, d1) = 0.$$

Now let $f = a^{i_1\ldots i_k} a_{i_1\ldots i_k}$ and expand to get

$$0 = \int_M \nabla^2 (a^{i_1\ldots i_k} a_{i_1\ldots i_k})\sigma$$
$$= \int_M \nabla^i \nabla_i (a^{i_1\ldots i_k} a_{i_1\ldots i_k})\sigma$$
$$= \int_M \nabla^i \{(\nabla_i a^{i_1\ldots i_k}) a_{i_1\ldots i_k} + a^{i_1\ldots i_k} \nabla_i a_{i_1\ldots i_k}\}\sigma$$
$$= \int_M \{(\nabla^2 a^{i_1\ldots i_k}) a_{i_1\ldots i_k} + (\nabla_i a^{i_1\ldots i_k})(\nabla^i a_{i_1\ldots i_k})$$
$$+ (\nabla^i a^{i_1\ldots i_k})(\nabla_i a_{i_1\ldots i_k}) + a^{i_1\ldots i_k} \nabla^2 a_{i_1\ldots i_k}\}\sigma$$
$$= 2\int_M \{(\nabla^2 a^{i_1\ldots i_k}) a_{i_1\ldots i_k} + (\nabla^i a^{i_1\ldots i_k})(\nabla_i a_{i_1\ldots i_k})\}\sigma.$$

c. Combining (a) and (b) gives

$$\int_M \left\{(\nabla_i a_{i_1\ldots i_k})(\nabla^i a^{i_1\ldots i_k}) + kF(\alpha)\right\}\sigma = 0. \tag{1}$$

From (8.109) we have

$$R_{ij} = g^{pq} R_{piqj} = K g^{pq}(g_{pq} g_{ij} - g_{pj} g_{iq}) = K(n-1) g_{ij},$$

so

$$a^{j i_2 \dots i_k} a_{i i_2 \dots\dots i_k} R^i{}_j = K(n-1)k! \|\alpha\|^2.$$

Also,

$$\begin{aligned} a^{pq i_3 \dots i_k} a_{ij i_3 \dots i_k} R^{ij}{}_{pq} &= K a^{pq i_3 \dots i_k} a_{ij i_3 \dots i_k} (\delta^i_p \delta^j_q - \delta^i_q \delta^j_p) \\ &= 2Kk! \|\alpha\|^2. \end{aligned}$$

Hence

$$F(\alpha) = Kk!(n-k)\|\alpha\|^2,$$

as anticipated. By the nondegeneracy of the inner product, $\|\alpha\|^2$ is nonnegative, as is $\|\nabla_j a_{i_1 \dots i_k}\|^2$, so each term in the brackets in (1) must vanish. In particular, we must have $F(\alpha) = 0$, and hence $\alpha = 0$. Evidently there are no nonzero harmonic k forms on M for $0 < k < n$. By Hodge's theorem, the corresponding Betti numbers must vanish.

8.60 a.

$$V_{n+1} = \int_0^1 r^n A_n \, dr = \frac{1}{n+1} A_n.$$

b. We compute

$$d\sigma = (n+1)\omega,$$

so

$$\begin{aligned} A_n &= \int_{S^n} \sigma = c \int_{\partial B^{n+1}} \tau = c \int_{B^{n+1}} d\sigma = (n+1)c \int_{B^{n+1}} \omega \\ &= (n+1) c V_{n+1} = c A_n, \end{aligned}$$

and therefore $c = 1$.

c. We follow the proof in ([5], p. 411). The domain of integration for the $(n+1)$-ball is

$$(x^1)^2 + \cdots + (x^{n+1})^2 \le 1,$$

which is equivalent to

$$(x^3)^2 + \cdots + (x^{n+1})^2 \le 1 - (x^1)^2 + (x^2)^2 \qquad \text{and} \qquad (x^1)^2 + (x^2)^2 \le 1.$$

Thus,

$$
\begin{aligned}
V_{n+1} &= \int_{\|x\|^2 \le 1} \omega \\
&= \int_{(x^1)^2+(x^2)^2 \le 1} dx^1\, dx^2 \int_{\substack{(x^3)^2+\cdots+(x^{n+1})^2 \\ \le 1-(x^1)^2+(x^2)^2}} dx^3 \dots dx^{n+1}.
\end{aligned}
$$

The second integral is the volume of an $(n-1)$-ball of radius $(1-(x^1)^2+(x^2)^2)^{1/2}$, so

$$
\begin{aligned}
V_{n+1} &= V_{n-1} \int_{(x^1)^2+(x^2)^2 \le 1} (1-(x^1)^2+(x^2)^2)^{(n-1)/2}\, dx^1\, dx^2 \\
&= 2\pi V_{n-1} \int_0^1 (1-r^2)^{(n-1)/2}\, r\, dr \\
&= \frac{2\pi}{n+1} V_{n-1}.
\end{aligned}
$$

Equivalently,

$$\frac{V_n}{V_{n-2}} = \frac{2\pi}{n}.$$

Temporarily set

$$f_n := \frac{\pi^{n/2}}{\Gamma((n/2)+1)}.$$

Then by the recursive property of the gamma function,

$$\frac{f_n}{f_{n-2}} = \frac{2\pi}{n}.$$

Moreover, $V_1 = f_1 = 2$ and $V_2 = f_2 = \pi$. As V_n and f_n satisfy the same recurrence relation with the same initial conditions, we must have $V_n = f_n$.

8.61 a. The Laplacian in spherical polar coordinates in $\mathbb{R}^3$ is

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial f}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}.$$

Plugging in $f = r^n h$ gives

$$
\begin{aligned}
0 &= n(n+1)h + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial h}{\partial \theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 h}{\partial \phi^2} \\
&= n(n+1)h + \nabla^2 h \\
&= n(n+1)h - \Delta h.
\end{aligned}
$$

b. Let h_n and h_m be spherical harmonics of degrees n and m, respectively. Then

$$n(n+1)(h_n, h_m) = (\Delta h_n, h_m) = (h_n, \Delta h_m) = m(m+1)(h_n, h_m).$$

If $n \neq m$ then $(h_n, h_m) = 0$.

8.62 Let η be harmonic. Clearly, if $\lambda = \Delta\omega$ then

$$(\lambda, \eta) = (\Delta\omega, \eta) = (\omega, \Delta\eta) = 0.$$

Conversely, suppose $(\lambda, \eta) = 0$ for all η. By the Hodge decomposition, there exist α, β, and γ such that

$$\lambda = d\alpha + \delta\beta + \gamma.$$

Choosing $\eta = \gamma$ we get

$$0 = (\lambda, \gamma) = (d\alpha + \delta\beta + \gamma, \gamma) = (\gamma, \gamma),$$

because γ is closed and co-closed. By nondegeneracy, $\gamma = 0$.

Again by the Hodge decomposition,

$$\alpha = d\alpha_1 + \delta\beta_1 + \gamma_1,$$

so

$$d\alpha = d\delta\beta_1.$$

Once again we have

$$\beta_1 = d\alpha_2 + \delta\beta_2 + \gamma_2,$$

whereupon we find

$$d\alpha = d\delta\beta_1 = d\delta d\alpha_2 = \Delta(d\alpha_2).$$

A similar argument shows that $\delta\beta$ is harmonic, and the sum of two harmonic functions is harmonic.

9

The degree of a smooth map

9.1 We just extend the commutative diagram used to define 'Deg' to get

$$\begin{tikzcd} H^n(P) \arrow[r, "g^*"] \arrow[d, "\int_P"'] & H^n(N) \arrow[r, "f^*"] \arrow[d, "\int_N"'] & H^n(M) \arrow[d, "\int_M"'] \\ \mathbb{R} \arrow[r, "\mathrm{Deg}(g)"] & \mathbb{R} \arrow[r, "\mathrm{Deg}(f)"] & \mathbb{R}. \end{tikzcd} \tag{1}$$

The claim then follows from the fact that $(g \circ f)^* = f^* \circ g^*$.

9.2 The form dx is globally defined on the torus. It is closed but not exact, because x is not a single valued function on the torus. Nevertheless, in the neighborhood of a point x is a well-defined function, so $f^*(dx) = df^*x$ makes sense locally. Specifically, $f^*(dx)$ is a closed one-form on the sphere (because 'd' is a local operation). Let $\lambda := f^*(dx)$ and $\eta := f^*(dy)$. Every closed one-form on the 2-sphere is exact (because $H^1(S^2) = 0$), so $\lambda = dg$ and $\eta = dh$ for some (globally defined) functions g and h on the 2-sphere. But then

$$\int_M dg \wedge dh = \int_M d(g \wedge dh) = \int_{\partial M} g \wedge dh = 0,$$

so the degree of f is zero.

9.3 First note that choosing $x = 0$ gives $g = f$. Then

$$\int_{S^1} f^*(d\varphi) = (\mathrm{Deg}\, f) \int_{S^1} d\varphi = 2\pi(\mathrm{Deg}\, f).$$

But $f^*(d\varphi) = df^*\varphi = d(\varphi \circ f) = k\, d\theta$, so $\mathrm{Deg}\, f = k$. (OK, that's a little cavalier, because θ and φ are not globally defined functions, but the argument works because k is the Jacobian of the transformation f at every point, so the local argument extends to a global one.)

9.4 Let σ be the volume form (generator of top dimensional cohomology) on S^n. Flipping the sign of all the x^i's sends $\sigma \to (-1)^{n+1}\sigma$. So

$$(-1)^{n+1}\int_{S^n}\sigma = \int_{S^n}(-id)^*\sigma = (\mathrm{Deg}(-id))\int_{S^n}\sigma.$$

9.5 The map $F(t, x)$ is clearly a smooth homotopy between f and the antipodal map, provided it exists. But the denominator vanishes only if

$$(1-t)f(x) = tx \Leftrightarrow \|(1-t)f(x)\| = \|tx\| \Leftrightarrow 1-t = t \Leftrightarrow t = 1/2,$$

(because $x, f(x) \in S^n$). This holds provided $f(x) = x$, but f has no fixed points by hypothesis, so F is indeed well-behaved. The result now follows from the homotopy invariance of the degree.

9.6 By Exercise 4.12, f is null homotopic. This means that it is homotopic to a constant map. But the degree of the constant map is zero. (One way to see this is that pullback is the inverse of pushforward, and the derivative of a constant is zero. Alternatively, the degree is the sum of the indices at a regular point, and each index is a sign of a Jacobian determinant, and therefore zero.)

9.7 Suppose we had $g(x) = x$ for some x. By construction, g sends all the points on the sphere to the southern hemisphere S, so the fixed point would have to be in S. But then we would have $f(\varphi_S(x)) = \varphi_S(x)$, a contradiction, as f has no fixed points. Hence, by Exercise 9.5 the degree of g is $(-1)^{n+1}$. But g is a nonsurjective map of the sphere to itself, so by Exercise 9.6 its degree must be zero. This contradiction shows that f must have a fixed point.

9.8 The first part is straightforward:

$$\begin{aligned}
\lambda^n &= 2^{-n} Q_{i_1 i_2}\cdots Q_{i_{2n-1}i_{2n}}\,\theta^{i_1}\wedge\cdots\wedge\theta^{i_{2n}} \\
&= 2^{-n}\sum_{\sigma\in\mathfrak{S}_{2n}}(-1)^\sigma Q_{\sigma(1)\sigma(2)}\cdots Q_{\sigma(2n-1)\sigma(2n)}\,\sigma \\
&= n!\,\mathrm{pf}(Q)\,\sigma.
\end{aligned}$$

For the second part, under a change of basis $\theta' = A^{-1}\theta$ so $\theta = A\theta'$ or $\theta^j = A^j{}_{k'}\theta^{k'}$. Hence,

$$\begin{aligned}
\lambda &= \frac{1}{2}Q_{ij}\,\theta^i\wedge\theta^j \\
&= \frac{1}{2}Q_{ij}A^i{}_{k'}A^j{}_{\ell'}\,\theta^{k'}\wedge\theta^{\ell'} \\
&= \frac{1}{2}(A^T QA)_{k'\ell'}\,\theta^{k'}\wedge\theta^{\ell'} \\
&= \frac{1}{2}Q'_{k'\ell'}\,\theta^{k'}\wedge\theta^{\ell'},
\end{aligned}$$

so

$$Q' = A^T QA.$$

But under this change of basis,

$$\sigma \mapsto \sigma' = (\det A^{-1})\sigma = (\det A)^{-1}\sigma,$$

so

$$n!\,\mathrm{pf}(Q)\sigma = \lambda^n = n!\,\mathrm{pf}(Q')\sigma' = n!\,\mathrm{pf}(A^T QA)(\det A)^{-1}\,\sigma,$$

and therefore

$$\mathrm{pf}(A^T QA) = (\det A)\,\mathrm{pf}(Q).$$

Appendix D

Riemann normal coordinates

D.1 In the specified coordinates the Christoffel symbols vanish. From the definition of the Riemann tensor,

$$R^i{}_{jk\ell} = \Gamma^i{}_{\ell j,k} - \Gamma^i{}_{kj,\ell}.$$

Lowering all the indices with the Euclidean metric gives

$$\begin{aligned} R_{ijk\ell} &= \Gamma_{i\ell j,k} - \Gamma_{ikj,\ell} \\ &= \frac{1}{2}(g_{ji,\ell} + g_{\ell i,j} - g_{\ell j,i})_{,k} - (k \leftrightarrow \ell) \\ &= \frac{1}{2}(g_{\ell i,jk} - g_{\ell j,ik} - g_{ki,j\ell} + g_{kj,i\ell}) \\ &= \frac{1}{2}(g_{i\ell,jk} - g_{j\ell,ik} - g_{ik,j\ell} + g_{jk,i\ell}). \end{aligned}$$

D.2 There are n choices for i. By the symmetry of the partial derivatives we must count the number of multisets of size 3 chosen from a set of size n. By the solution to Exercise 2.21a, this is $\binom{n+2}{3}$, so $a = n\binom{n+2}{3}$.

The metric tensor is symmetric, so it has $n(n+1)/2 = \binom{n+1}{2}$ components, and by the symmetry of partial derivatives there are this many independent second derivatives as well, so there are $b = \binom{n+1}{2}^2$ independent second derivatives of the metric.

We have

$$\begin{aligned} b - a &= \left(\frac{(n+1)n}{2}\right)^2 - n\frac{(n+2)(n+1)n}{6} \\ &= n^2(n+1)\left[\frac{n+1}{4} - \frac{n+2}{6}\right] \\ &= \frac{n^2(n^2-1)}{12}. \end{aligned}$$

According to (8.51), this is precisely the number of independent components of the Riemann tensor, which shows that the Riemann tensor is completely determined by the second derivatives of the metric, or equivalently, that all the curvature information is contained in the second derivatives of the metric. Yet another way to put it is that, whereas we can always find a coordinate system in which the first derivatives of the metric vanish, we cannot change coordinates so as to make any of the second derivatives go away. (Physicists would say that the second derivatives of the metric are "physical", as opposed to "gauge degrees of freedom".)

Appendix F

Frobenius' theorem

F.1 First we show that (2) and (3) are equivalent. If $d\theta^i = \sum_{j=1}^m A^i{}_j \theta^j$ then wedging with any of the θ^j's $(1 \leq j \leq m)$ will kill it. Conversely, we can complete the θ^i's to a coframe field, in which case we can write $d\theta^i = \sum_{j,k=1}^n a^i{}_{jk} \theta^j \wedge \theta^k$. Then

$$0 = d\theta^i \wedge (\theta^1 \wedge \cdots \wedge \theta^m) = \sum_{j,k=1}^n a^i{}_{jk} \theta^j \wedge \theta^k \wedge \theta^1 \wedge \cdots \wedge \theta^m$$

If $m = n$ there is nothing to prove, so suppose $m < n$. Then by linear independence of the θ^j's we must have $a^i{}_{jk} = 0$ for $k = m - n, \ldots, n$, so (3) holds.

Next we show that (1) and (3) are equivalent. Let $X, Y \in \Delta$. By (3.123),

$$d\theta^i(X, Y) = X\theta^i(Y) - Y\theta^i(X) - \theta^i([X, Y]) = -\theta^i([X, Y]).$$

Completing the θ^i's to a basis we can write

$$\begin{aligned} d\theta^i(X, Y) &= \sum_{j,k=1}^n a^i{}_{jk} \theta^j \wedge \theta^k(X, Y) \\ &= \sum_{j,k=m-n}^n a^i{}_{jk} \theta^j \wedge \theta^k(X, Y), \end{aligned}$$

because $\theta^k(X) = 0$ if $k = 1, \ldots, m$.

If Δ is involutive then $\theta^i([X, Y]) = 0$, so $a^i{}_{jk} = 0$ for $j, k = m - n, \ldots, n$, and (3) holds. Conversely, if (3) holds then $d\theta^i(X, Y) = 0$, which means $\theta^i([X, Y]) = 0$ for $i = 1, \ldots, m$, and this implies $[X, Y] \in \Delta$.

Appendix G
The topology of electrical circuits

G.1 Let G be a tree. Pick an edge e with endpoints p and q. If removing e failed to disconnect G, there would be a path from p to q in $G - e$. But then adding e to that path would produce a cycle, a contradiction. Conversely, let G be a connected graph with the property that removing every edge disconnects G. Then G cannot have any cycles and must therefore be a tree.

Appendix H

Intrinsic and extrinsic curvature

H.1 We have

$$\widetilde{\nabla}_{fX}(gY) = f\widetilde{\nabla}_X(gY) = fX(g)Y + fg\widetilde{\nabla}_X Y$$

and

$$\nabla_{fX}(gY) = f\nabla_X(gY) = fX(g)Y + fg\nabla_X Y,$$

so

$$\alpha(fX, gY) = \widetilde{\nabla}_{fX}(gY) - \nabla_{fX}(gY) = fg(\widetilde{\nabla}_X Y - \nabla_X Y) = fg\alpha(X, Y).$$

H.2 From the definition of $R(X, Y)Z$ and Gauss's formula we obtain

$$\begin{aligned}
\widetilde{R}(X, Y)Z &= \widetilde{\nabla}_X\widetilde{\nabla}_Y Z - \widetilde{\nabla}_Y\widetilde{\nabla}_X Z - \widetilde{\nabla}_{[X,Y]}Z \\
&= \widetilde{\nabla}_X(\nabla_Y Z + \alpha(Y, Z)) - (X \leftrightarrow Y) - \nabla_{[X,Y]}Z - \alpha([X, Y], Z) \\
&= \nabla_X\nabla_Y Z + \alpha(X, \nabla_Y Z) + \widetilde{\nabla}_X\alpha(Y, Z) \\
&\qquad - \nabla_Y\nabla_X Z - \alpha(Y, \nabla_X Z) - \widetilde{\nabla}_Y\alpha(X, Z) \\
&\qquad\quad - \nabla_{[X,Y]}Z - \alpha([X, Y], Z),
\end{aligned}$$

where '$(X \leftrightarrow Y)$' means 'repeat the previous terms with X and Y interchanged'. Taking projections onto the tangent and normal spaces yields the Gauss and Codazzi-Mainardi equations immediately:

$$P(\widetilde{R}(X, Y)Z) = R(X, Y)Z + P(\widetilde{\nabla}_X\alpha(Y, Z) - \widetilde{\nabla}_Y\alpha(X, Z))$$

and

$$\begin{aligned}
(1 - P)(\widetilde{R}(X, Y)Z) &= \alpha(X, \nabla_Y Z) - \alpha(Y, \nabla_X Z) - \alpha([X, Y], Z) \\
&\qquad + (1 - P)(\widetilde{\nabla}_X\alpha(Y, Z) - \widetilde{\nabla}_Y\alpha(X, Z)).
\end{aligned}$$

H.3 The tangent space to M is orthogonal to the normal space to N, so by metric compatibility and Gauss's formula,

$$\begin{aligned} 0 = X(g(Y,\xi)) &= \widetilde{\nabla}_X(g(Y,\xi)) = g(\widetilde{\nabla}_X Y, \xi) + g(Y, \widetilde{\nabla}_X \xi) \\ &= g(\nabla_X Y + \alpha(X,Y), \xi) - g(Y, S_\xi(X)) \\ &= g(\alpha(X,Y),\xi) - g(S_\xi(X), Y). \end{aligned}$$

H.4 From Gauss's equation (H.7), the properties of projections, and metric compatibility,

$$\begin{aligned} &g(\widetilde{R}(X,Y)Z, W) \\ &\quad = g(R(X,Y)Z,W) + g(\widetilde{\nabla}_X \alpha(Y,Z) - \widetilde{\nabla}_Y \alpha(X,Z), W) \\ &\quad = g(R(X,Y)Z,W) - g(\alpha(Y,Z), \widetilde{\nabla}_X W) + g(\alpha(X,Z), \widetilde{\nabla}_Y W) \\ &\quad = g(R(X,Y)Z,W) - g(\alpha(Y,Z), \alpha(X,W)) + g(\alpha(X,Z), \alpha(Y,W)) \\ &\quad = g(R(X,Y)Z,W) + \mathrm{II}_\xi(X,Z)\mathrm{II}_\xi(Y,W) - \mathrm{II}_\xi(Y,Z)\mathrm{II}_\xi(X,W). \end{aligned}$$

Again, by metric compatibility,

$$g(\widetilde{\nabla}_X \alpha(Y,Z), \xi) = -g(\alpha(Y,Z), \widetilde{\nabla}_X \xi) = -\mathrm{II}_\xi(Y,Z) g(\xi, \widetilde{\nabla}_X \xi) = 0.$$

So by the Codazzi-Mainardi equation (H.8) and metric compatibility,

$$\begin{aligned} g(\widetilde{R}(X,Y)Z,\xi) &= \mathrm{II}_\xi(X, \nabla_Y Z) - \mathrm{II}_\xi(Y, \nabla_X Z) - \mathrm{II}_\xi([X,Y],Z) \\ &= g(S_\xi(X), \nabla_Y Z) - g(S_\xi(Y), \nabla_X Z) - g(S_\xi([X,Y]), Z) \\ &= g(\nabla_X S_\xi(Y) - \nabla_Y S_\xi(X) - S_\xi([X,Y]), Z). \end{aligned}$$

H.5 $\mathbb{R}^3$ is flat, so its curvature vanishes. By the Gauss equation,

$$g(R(X,Y)Y,X) = -\mathrm{II}_\xi(X,Y)\mathrm{II}_\xi(Y,X) + \mathrm{II}_\xi(Y,Y)\mathrm{II}_\xi(X,X).$$

Choose $X = e_1$ and $Y = e_2$ to be orthornormal at p. Then the left side of the above equation is just the Gaussian curvature K. As for the right side, we have

$$\mathrm{II}_\xi(e_j, e_i) = g(S_\xi(e_j), e_i) = S_{ij},$$

where S_{ij} is the matrix representation of the shape operator S in the basis $\{e_i\}$. So we get

$$K = -S_{12}S_{21} + S_{22}S_{11} = \det S.$$

H.6 If the ambient space is Euclidean then its curvature vanishes, and Gauss's equation becomes

$$g(R(X,Y)Z,W) = -\mathrm{II}_\xi(X,Z)\mathrm{II}_\xi(Y,W) + \mathrm{II}_\xi(Y,Z)\mathrm{II}_\xi(X,W).$$

By Equation (8.44),

$$R_{abcd} = g(R(e_c, e_d)e_b, e_a),$$

so choosing $X = \partial_k$, $Y = \partial_\ell$, $Z = \partial_j$, and $W = \partial_i$ gives

$$R_{ijk\ell} = b_{ik}b_{j\ell} - b_{i\ell}b_{jk}.$$

where we used the fact that $b_{ij} = b_{ji}$.

For the Codazzi-Mainardi equation, we first observe that the components of the second fundamental form are related to those of the shape operator by the metric function. Specifically, if we choose a coordinate basis and set

$$S(\partial_i) = S^k{}_i \partial_k,$$

then

$$b_{ij} = \mathrm{II}(\partial_i, \partial_j) = g(S(\partial_i), \partial_j) = S^k{}_i g_{kj} =: S_{ji}.$$

Second, recall that $[\partial_i, \partial_j] = 0$. Hence, by Equation (H.14) with $X = \partial_i$, $Y = \partial_j$, and $Z = \partial_k$,

$$\begin{aligned}
0 &= g(\nabla_{\partial_i} S_\xi(\partial_j) - \nabla_{\partial_j} S_\xi(\partial_i), \partial_k) \\
&= g(\nabla_{\partial_i}(S^\ell{}_j \partial_\ell) - \nabla_{\partial j}(S^\ell{}_i \partial_\ell), \partial_k) \\
&= g(S^\ell{}_{j;i} \partial_\ell - S^\ell{}_{i;j} \partial_\ell, \partial_k) \\
&= g_{\ell k}(S^\ell{}_{j;i} - S^\ell{}_{i;j}) \\
&= S_{kj;i} - S_{ki;j} \\
&= (S_{kj,i} - S_{mj}\Gamma^m{}_{ki} - S_{km}\Gamma^m{}_{ji}) - (i \leftrightarrow j) \\
&= (b_{kj,i} - b_{mj}\Gamma^m{}_{ki}) - (i \leftarrow j).
\end{aligned}$$

H.7 a. The ambient space is Euclidean, so the ambient connection $\widetilde{\nabla}$ just reduces to the ordinary derivative. In particular, if X is a tangent vector field on Σ,

$$S_n(X) = -\widetilde{\nabla}_X n = -X(n^k)\partial_k.$$

Now let $X = \sigma_*(\partial/\partial u)$. Then

$$\begin{aligned}
J = \mathrm{II}_n(X, X) &= g(S_n(X), X) \\
&= -X(n^k)(\sigma_u)^j g(\partial_k, \partial_j) = -(n_u)^k(\sigma_u)_k = -g(n_u, \sigma_u).
\end{aligned}$$

But σ_u is tangent to Σ and n is normal, so

$$0 = g(n, \sigma_u)_u = g(n_u, \sigma_u) + g(n, \sigma_{uu}),$$

and therefore

$$J = g(n, \sigma_{uu})$$

as well. The other terms are obtained similarly.

b. We can do this the simple way or the fancy way. First the simple way. Let $X = \sigma_* \partial_u$ and $Y = \sigma_* \partial_v$, as before. As X and Y span the tangent space to Σ,

$$\begin{aligned} S(X) &= aX + bY \\ S(Y) &= cX + dY, \end{aligned}$$

for some a, b, c, and d. Taking inner products gives

$$\begin{aligned} J &= g(S(X), X) = ag(X, X) + bg(Y, X) = aE + bF \\ K &= g(S(X), Y) = ag(X, Y) + bg(Y, Y) = aF + bG \\ K &= g(S(Y), X) = cg(X, X) + dg(Y, X) = cE + dF \\ L &= g(S(Y), Y) = cg(X, Y) + dg(Y, Y) = cF + dG, \end{aligned}$$

which we write as

$$\begin{pmatrix} J & K \\ K & L \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Inverting and multiplying gives

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} J & K \\ K & L \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \\ &= \frac{1}{EG - F^2} \begin{pmatrix} J & K \\ K & L \end{pmatrix} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \\ &= \frac{1}{EG - F^2} \begin{pmatrix} GJ - FK & EK - FJ \\ GK - FL & EL - FK \end{pmatrix}. \end{aligned}$$

But, relative to the basis $\{X, Y\}$, the matrix representing S is just the transpose of this matrix, by virtue of our conventions (*cf.,* (1.11)).

Alternatively, the fancy way is to observe that (using the symmetry of b_{ij})

$$b_{ij} = \mathrm{II}(\partial_i, \partial_j) = g(S(\partial_i), \partial_j) = S^k{}_i g_{kj} \quad \Rightarrow \quad S^k{}_i = g^{kj} b_{ji},$$

which gives the same result as before.

c. This follows immediately from the previous solution by taking determinants:

$$\det S = (\det \mathrm{II})(\det I)^{-1}.$$

d. Differentiate both sides of

$$g(\sigma_u, \sigma_u) = g(\sigma_v, \sigma_v)$$

with respect to u to get

$$g(\sigma_{uu}, \sigma_u) = g(\sigma_{vu}, \sigma_v).$$

But

$$0 = g(\sigma_u, \sigma_v)_v = g(\sigma_{uv}, \sigma_v) + g(\sigma_u, \sigma_{vv}),$$

so combining the two equations we get

$$g(\sigma_u, \sigma_{uu} + \sigma_{vv}) = 0.$$

A similar computation shows that

$$g(\sigma_v, \sigma_{uu} + \sigma_{vv}) = 0$$

as well, demonstrating that $\sigma_{uu} + \sigma_{vv}$ is normal to Σ. Taking inner products with n and using the results of Part (a) we get

$$g(n, \sigma_{uu} + \sigma_{vv}) = J + L.$$

On the other hand, from Part (b),

$$g(n, 2\lambda^2 H) = \lambda^2 \operatorname{tr} S = \lambda^2 \frac{GJ + EL - 2FK}{EG - F^2} = J + L.$$

because $E = G = \lambda^2$ and $F = 0$.

e. We have

$$\begin{aligned}
\sigma_u &= (-a \cosh v \sin u, a \cosh v \cos u, 0),\\
\sigma_{uu} &= (-a \cosh v \cos u, -a \cosh v \sin u, 0),\\
\sigma_v &= (a \sinh v \cos u, a \sinh v \sin u, a),\\
\sigma_{vv} &= (a \cosh v \cos u, a \cosh v \sin u, 0).
\end{aligned}$$

It follows that

$$\begin{aligned}
&g(\sigma_u, \sigma_u) = a^2 \cosh^2 v,\\
&g(\sigma_v, \sigma_v) = a^2 \sinh^2 v + a^2 = a^2(1 + \sinh^2 v) = a^2 \cosh^2 v,\\
&g(\sigma_u, \sigma_v) = 0,
\end{aligned}$$

so the parameterization is indeed isothermal. Moreover, $\sigma_{uu} + \sigma_{vv} = 0$, so by Part (d), the catenoid is a minimal surface. (It is the only minimal surface of revolution.)

H.8 We have

$$\begin{aligned}
\sigma_u &= (-a \sin u, a \cos u, 0),\\
\sigma_{uu} &= (-a \cos u, -a \sin u, 0),\\
\sigma_{uv} &= (0, 0, 0),\\
\sigma_v &= (0, 0, 1),\\
\sigma_{vv} &= (0, 0, 0),
\end{aligned}$$

and

$$n = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|} = \frac{(a\cos u, a\sin u, 0)}{a} = (\cos u, \sin u, 0).$$

Thus,

$$\begin{aligned} E &= (\sigma_u, \sigma_u) = a^2, \\ F &= (\sigma_u, \sigma_v) = 0, \\ G &= (\sigma_v, \sigma_v) = 1, \\ J &= (n, \sigma_{uu}) = -a, \\ K &= (n, \sigma_{uv}) = 0, \\ L &= (n, \sigma_{vv}) = 0. \end{aligned}$$

The shape operator is

$$S = (EG - F^2)^{-1}\begin{pmatrix} GJ - FK & GK - FL \\ EK - FJ & EL - FK \end{pmatrix} = \begin{pmatrix} -1/a & 0 \\ 0 & 0 \end{pmatrix},$$

which is already diagonalized. Therefore the Gaussian curvature (product of eigenvalues) is 0, while the mean curvature (average of eigenvalues) is $-1/2a$.

H.9 We have

$$\begin{aligned} \sigma_u &= (1 - u^2 + v^2, 2uv, 2u), \\ \sigma_{uu} &= (-2u, 2v, 2), \\ \sigma_{uv} &= (2v, 2u, 0), \\ \sigma_v &= (2uv, 1 - v^2 + u^2, -2v), \\ \text{and} \qquad \sigma_{vv} &= (2u, -2v, -2). \end{aligned}$$

Also, $n = \eta/\|\eta\|$ where

$$\eta = \sigma_u \times \sigma_v = \begin{vmatrix} \hat{\boldsymbol{i}} & \hat{\boldsymbol{j}} & \hat{\boldsymbol{k}} \\ 1 - u^2 + v^2 & 2uv & 2u \\ 2uv & 1 - v^2 + u^2 & -2v \end{vmatrix}.$$

To save writing we define $x := u^2 + v^2$ and $\alpha := 1 + x$. Then the individual components of η are:

$$\begin{aligned} \hat{\boldsymbol{i}}: \quad & (2uv)(-2v) - (2u)(1 - v^2 + u^2) = -4uv^2 - 2u + 2uv^2 - 2u^3 \\ & = -2u(1 + u^2 + v^2) = -2u\alpha \\ \hat{\boldsymbol{j}}: \quad & (2u)(2uv) - (-2v)(1 - u^2 + v^2) = 4u^2v + 2v - 2vu^2 - 2v^3 \\ & = 2v(1 + u^2 + v^2) = 2v\alpha \\ \hat{\boldsymbol{k}}: \quad & (1 - (u^2 - v^2))(1 + (u^2 - v^2)) - (2uv)^2 = 1 - (u^2 - v^2)^2 - 4u^2v^2 \\ & = 1 - (u^4 + 2u^2v^2 + v^4) = 1 - x^2. \end{aligned}$$

Therefore,

$$\begin{aligned}\|\eta\|^2 &= \eta \cdot \eta = (-2u\alpha)^2 + (2v\alpha)^2 + (1-x^2)^2 \\ &= 4(u^2+v^2)\alpha^2 + (1-2x^2+x^4) = 4x(1+x)^2 + 1 - 2x^2 + x^4 \\ &= 4x(1+2x+x^2) + 1 - 2x^2 + x^4 = 1 + 4x + 6x^2 + 4x^3 + x^4 \\ &= (1+x)^4 = \alpha^2,\end{aligned}$$

and

$$n = \frac{\eta}{\|\eta\|} = \frac{(-2u\alpha, 2v\alpha, 1-x^2)}{\alpha^2} = \left(-\frac{2u}{\alpha}, \frac{2v}{\alpha}, \frac{1-(u^2+v^2)^2}{\alpha^2}\right).$$

Using these results we get the following.

a.

$$\begin{aligned}E = (\sigma_u, \sigma_u) &= (1-u^2+v^2)^2 + (2uv)^2 + (2u)^2 \\ &= 1 + u^4 + v^4 - 2u^2 + 2v^2 - 2u^2v^2 + 4u^2v^2 + 4u^2 \\ &= 1 + u^2 + v^4 + 2u^2 + 2v^2 + 2u^2v^2 = (1+u^2+v^2)^2.\end{aligned}$$

Similar reasoning gives $F = 0$ and $G = E$.

b.

$$\begin{aligned}J = (n, \sigma_{uu}) &= \frac{1}{\alpha^2}[(-2u\alpha)(-2u) + (2v\alpha)(2v) + 2(1-x^2)] \\ &= \frac{1}{\alpha^2}[4x(1+x) + 2 - 2x^2] = \frac{2}{\alpha^2}(1+2x+x^2) = 2.\end{aligned}$$

Similar reasoning gives $K = 0$ and $L = -2$.

c. The shape operator is

$$S = (EG - F^2)^{-1}\begin{pmatrix} GJ - FK & GK - FL \\ EK - FJ & EL - FK \end{pmatrix} = \frac{2}{\alpha^2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is already diagonalized. Therefore the principal curvatures are $\lambda_1 = 2/\alpha^2$ and $\lambda_2 = -2/\alpha^2$. In particular, the mean curvature (average of the principal curvatures) is zero.

H.10 a. Parameterize the surface in the obvious way by

$$\sigma(x, y) = (x, y, f(x, y)).$$

Then

$$\sigma_x = (1, 0, f_x) \quad \text{and} \quad \sigma_y = (0, 1, f_y),$$

so

$$\begin{aligned}E &= g(\sigma_x, \sigma_x) = 1 + f_x^2, \\ F &= g(\sigma_x, \sigma_y) = f_x f_y, \\ G &= g(\sigma_y, \sigma_y) = 1 + f_y^2.\end{aligned}$$

b. We have

$$\sigma_x \times \sigma_y = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = (-f_x, -f_y, 1),$$

and

$$n = \frac{(-f_x, -f_y, 1)}{(1 + f_x^2 + f_y^2)}.$$

Also,

$$\sigma_{xx} = (0, 0, f_{xx}), \quad \sigma_{xy} = (0, 0, f_{xy}), \quad \text{and} \quad \sigma_{yy} = (0, 0, f_{yy}),$$

so

$$\begin{aligned} J = g(n, \sigma_{xx}) &= \frac{f_{xx}}{(1 + f_x^2 + f_y^2)}, \\ K = g(n, \sigma_{xy}) &= \frac{f_{xy}}{(1 + f_x^2 + f_y^2)}, \\ L = g(n, \sigma_{yy}) &= \frac{f_{yy}}{(1 + f_x^2 + f_y^2)}. \end{aligned}$$

c. The Gaussian curvature is the ratio of the determinants of the second and first fundamental forms, namely

$$\begin{aligned} \lambda_1 \lambda_2 &= \frac{\det \mathrm{II}}{\det I} = \frac{JL - K^2}{EG - F^2} \\ &= \frac{1}{(1 + f_x^2 + f_y^2)^2} \left(\frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2)(1 + f_y^2) - (f_x f_y)^2} \right) \\ &= \frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^3}. \end{aligned}$$

d. The mean curvature is the trace of the shape operator, namely

$$\lambda_1 + \lambda_2 = \frac{GJ + EL - 2FK}{EG - F^2}.$$

This vanishes when the numerator vanishes, so the equation of a minimal surface is

$$(1 + f_y^2) f_{xx} + (1 + f_x^2) f_{yy} - 2 f_x f_y f_{xy} = 0.$$