Marián Fecko
Differential Geometry and Lie Groups for Physicists
Some additional material to the book
Version from April 15, 2021. In progress.

Here one finds some comments, improvements, more detailed hints and additional (solved) problems. All this should hopefully make the book more useful.

[1.3] p.8: In the definition of a chart a homeomorphism $\varphi : \mathcal{O} \to \mathbb{R}^n$ is mentioned. In fact, a homeomorphism on an *open subset* of $\mathbb{R}^n$ might be more precise statement.

[1.3.1] p.9: In (i) the general (n-dimensional) case might be treated as follows.

Let the sphere S^n be realized as

$$\boldsymbol{\rho}^2 + z^2 = R^2 \qquad \boldsymbol{\rho} \equiv (x^1, \dots, x^n)$$

The South pole and the North pole, respectively, read

$$S = (\mathbf{0}, -R) \qquad N = (\mathbf{0}, R)$$

Let $P \equiv (\boldsymbol{\rho}, z)$ be a point on S^n (other then S or N). Then the line from S through P may be written as

$$\begin{pmatrix} \boldsymbol{\rho}(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ -R \end{pmatrix} + \left[\begin{pmatrix} \boldsymbol{\rho} \\ z \end{pmatrix} - \begin{pmatrix} \mathbf{0} \\ -R \end{pmatrix} \right] t$$

i.e.

$$\boldsymbol{\rho}(t) = \boldsymbol{\rho} t$$
$$z(t) = (R + z)t - R$$

This line crosses, in a unique point $(\mathbf{r}, R)$, where $\mathbf{r}$ is to be determined, the (hyper)plane touching the North pole, i.e. the plane $z = R$. A simple computation leads to

$$\mathbf{r} = \frac{2R}{R+z} \boldsymbol{\rho}$$

▼ The crossing happens for $t = t_1$ such that $z(t_1) \equiv (R + z)t_1 - R = R$, so that $t_1 = 2R/(R + z)$. Then

$$\mathbf{r} \equiv \boldsymbol{\rho}(t_1) = \boldsymbol{\rho} t_1 = \frac{2R}{R+z} \boldsymbol{\rho}$$

▲

From a similar computation we learn, that the line from N through P crosses, in a unique point $(\mathbf{r}', R)$, where $\mathbf{r}'$ is to be determined, the (hyper)plane touching the South pole, i.e. the plane $z = -R$. Here, we get

$$\mathbf{r}' = \frac{2R}{R-z} \boldsymbol{\rho}$$

So, we see that the two vectors are parallel

$$\mathbf{r}' = \lambda \mathbf{r}$$

and one only has to determine λ. In order to do that, notice that

$$rr' = \frac{2R}{R-z} \frac{2R}{R+z} \rho^2 = (2R)^2 \frac{\rho^2}{R^2 - z^2} = (2R)^2$$

from where we easily get $\lambda = (2R)^2/r^2$, or, finally

$$\mathbf{r}' = \frac{(2R)^2}{r} \frac{\mathbf{r}}{r}$$

1.3.2 p.10: In (iv) the fact $S^2 = \mathbb{C}P^1$ is to be proved. The following simple observation makes it easy: consider in general two n-dimensional manifolds M, N. Suppose each of them may be covered by exactly two charts, (x_1, x_2) and (y_1, y_2) respectively, such that "change of coordinates" expressions on the corresponding intersections, i.e. $x_2(x_1)$ and $y_2(y_1)$, happen to be given *by the same formulas.*

[Here x_1 denotes the whole n-tuple $(x_1^1, x_1^2, \dots, x_1^n)$ and similarly for x_2, y_1, y_2. Then $x_2(x_1)$ means in detail

$$x_2^a(x_1^1, x_1^2, \dots, x_1^n) \qquad\qquad a = 1, \dots, n$$

and similarly for $y_2(y_1)$.]

Then, M and N *are diffeomorphic*, the coordinate expression for the diffeomorphism being simply (check)

$$y_1^a = x_1^a \qquad\qquad y_2^i = x_2^i$$

In particular, concerning $\mathbb{C}P^1$ and S^2, the coordinates introduced in 1.3.2(ii) (for $p = 1$) and 1.3.5 *do have* the needed properties. (The common "change of coordinates" formula is, in complex language, of the form $z(w) = 1/w$.)

The same trick clearly works for more than just two charts per manifold as well.

1.5 p.16: There is a flaw in the Theorem (assumptions are not stated correctly).

The Theorem deals with constraints in Cartesian space $\mathbb{R}^n$ as a tool for defining a manifold.

In particular it says that, given *independent* constrains $\phi^1(x) = \dots = \phi^m(x) = 0$, the resulting manifold M is $(n - m)$-dimensional, i.e.

each constraint reduces the dimension by one unit.

More generally, according to the theorem, if the rank of the Jacobian matrix

$$J_i^a(x) := \frac{\partial \phi^a(x)}{\partial x^i}$$

is *constant on the set* M (i.e. at points satisfying the constrains), the dimension of the resulting manifold is $n - k$, where $k :=$ Rank J. (The case of *maximum* rank, i.e. $k = m$, corresponds to independent constraints mentioned above.)

So we can state more precisely that, if the assumptions of the theorem are fulfilled,

each constraint reduces the dimension *at most* by one unit.

Well, consider the following simple example: Take $n = 5$ and

$$\begin{aligned}
\phi^1(x^1, x^2, x^3, x^4, x^5) &= x^1 \\
\phi^2(x^1, x^2, x^3, x^4, x^5) &= x^2 \\
\phi^3(x^1, x^2, x^3, x^4, x^5) &= x^3 \\
\phi^4(x^1, x^2, x^3, x^4, x^5) &= x^4
\end{aligned}$$

What we get is clearly $M = \{\text{the } x^5\text{-axis in } \mathbb{R}^5\}$ as an *one*-dimensional manifold. (The constraints *are* independent.)

Now, on the same starting $\mathbb{R}^5$, consider *just a single* constraint

$$\phi = (\phi^1)^2 + (\phi^2)^2 + (\phi^3)^2 + (\phi^4)^2$$

so that

$$\phi(x^1, x^2, x^3, x^4, x^5) = (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2$$

It is clear that this constraint *alone* produces the same *one*-dimensional resulting $M =$ the x^5-axis in $\mathbb{R}^5$. So this particular constraint reduces the dimension by as many as *four* units!

Then, perhaps, the assumptions of the theorem are not satisfied. Let us check:

The Jacobian matrix reduces to a single-raw matrix

$$J = 2(x^1, x^2, x^3, x^4, 0)$$

and, on $M =$ the x^5-axis in $\mathbb{R}^5$, it takes the value

$$J = (0, 0, 0, 0, 0) \qquad \text{on } M$$

so that $k = \text{rank } J = 0$ and $\dim M = n - k = 5 - 0 = 5$. In spite of the fact that the rank of the Jacobian matrix *is* constant on M (it is namely vanishing there, $k = 0$), the statement $\dim M = n - k$ does *not* hold, here. So,

the theorem (as it stands in the book) *is not true in general.*

It turns out that what one should really assume is that

the rank is constant on *a neighbourhood* of M

rather than just on M alone!

What this improved version of the theorem says for our example?

We can see that although the rank of the Jacobian matrix is $k = 0$ right on M, we have $k = 1$ at *any* point *outside* M. So the rank k is *not* constant on a neighborhood of M, the assumptions of the (correct version of the) theorem are *not* fulfilled and therefore the theorem *says nothing* about the number of dimensions reduced by our particular constraint. Reduction by four units is perfectly compatible with the theorem (although unusual and slightly counter-intuitive).

Actually, the situation, when a single function (used as a constraint) reduces dimension by more than one unit, is well known in the context of the standard *radial coordinate* r (in polar coordinates (r, φ) in plane as well as in cylindrical or spherical polar coordinates (r, φ, z) or (r, ϑ, φ) in space). Namely we all know that the coordinate r is *defective* when $r = 0$. In particular, consider the family of constraints

$$\phi(r, \varphi) = r - c$$

in the plane, where c is a nonnegative constant. For c positive, the constraint reduces dimension by one unit (resulting in *circle* with radius c). For $c = 0$, however, the dimension is reduced by as many as *two* units (resulting in a *point*). The same constraint

$$\phi(r, \vartheta, \varphi) = r - c$$

regarded as a function in the space reduces (for $c = 0$) the dimension even by *three* units (resulting in a *point* again). And this is exactly the reason why one should *exclude* the point where $r = 0$ from the part of (say) the plane, where r may be used as a coordinate. *Coordinate* functions should have the property of reducing the dimension by *exactly one* unit. This leads, then, to the fact that *points* (zero-dimensional subsets) are labeled by exactly n values of coordinate functions (we need n steps to reduce the dimension from n to zero).

Moral: One should be careful about the character of some innocently looking (smooth) constraints. They may behave rather unexpectedly.

1.5.9 p.19: One can check explicitly, that the formula leads (via the idea from the Exercise (1.5.6)) to the following equivalences in the (u, v)-plane:

$$(u, v) \sim (u, v + 2\pi)$$
$$(u, v) \sim (u + 2\pi, 2\pi - v)$$

This means that

- we can *forget* about the points of the (u, v)-plane *outside* the square $\langle 0, 2\pi\rangle \times \langle 0, 2\pi\rangle$
- at the *boundary* of the square, there is equivalence of the type K^2 in (1.5.11)

The only thing to check is that all points *inside* the square are *inequivalent.* Try to prove it yourself, i.e. prove that, *inside* the square, if $(u, v) \neq (u', v')$, then resulting points in $\mathbb{R}^4$ are different (injectivity of the map $(0, 2\pi) \times (0, 2\pi) \to \mathbb{R}^4$).

2.2.1 - 2.2.3 p.24-25: Here we recapitulate general strategy of how to introduce, in terms of the concept of mutually tangent curves, the tangent space T_PM as a *vector* space *canonically* associated with a point $P \in M$.

We think of curves $\gamma(t)$ such that $\gamma(0) = P$ (no other curves are of interest for us). In a neighborhood of P, choose coordinates x^i. Then, near P, we have first order approximation

$$x^i(\gamma(t)) = x^i(P) + ta^i + \dots \qquad \text{where} \qquad a^i = \frac{d}{dt}\bigg|_0 x^i(\gamma(t))$$

So, with respect to coordinates x^i, the only data needed to first order approximation of a curve (within the family of curves $\gamma(t)$ such that $\gamma(0) = P$) consist in n-tuple of real numbers $(a^1, \dots, a^n)$.

Two curves with identical n-tuples a^i (i.e. such that their first order approximations coincide) are called *tangent* to one another.

▼ This fact does not depend on the choice of coordinates (i.e. the concept is well-defined). Indeed, if $x^i \leftrightarrow a^i$ and $x^{i'} \leftrightarrow a^{i'}$, then $a^{i'} = J^{i'}_j a^j$, where $J^{i'}_j = (\partial x^{i'}/\partial x^j)(P)$ is the *Jacobi matrix* of the change of coordinates $x'(x)$, evaluated at P. Therefore, if for two curves γ_1 and γ_2 holds $a^i_1 = a^i_2$ (i.e. they are tangent to one another with respect to x), then it also holds $a^{i'}_1 = a^{i'}_2$ (i.e. they are also tangent to one another with respect to x'; one just multiplies the first equation by the Jacobi matrix). ▲

One easily checks that the concept of tangency of curves introduces *equivalence relation* into the family of curves of our interest. An equivalence class is denoted $[\gamma]$ and the set of all equivalence classes is denoted T_PM (the *tangent space* at $P \in M$).

▼ All this is also well-defined: equivalent curves have equal n-tuples $(a^1, \dots, a^n)$ and equality of n-tuples does not depend on the choice of coordinates. ▲

So, for *each* choice of coordinates, there is a *bijection* between the equivalence classes of curves and n-tuples a^i (elements of $\mathbb{R}^n$):

$$\begin{aligned} T_PM \ni [\gamma] \;&\leftrightarrow\; a \in \mathbb{R}^n \qquad && \text{w.r.t. } x \\ &\leftrightarrow\; a' \in \mathbb{R}^n && \text{w.r.t. } x' \end{aligned}$$

Now recall that $\mathbb{R}^n$ has natural *linear* structure

$$(a^1, \dots, a^n) + \lambda(b^1, \dots, b^n) := (a^1 + \lambda b^1, \dots, a^n + \lambda b^n)$$

Therefore, we can steal, via a bijection, the linear structure from $\mathbb{R}^n$ and donate it to T_PM (well, in official parlance we *induce* the linear structure).

A small complication in this robbery presents the fact that there is no *canonical* bijection, but rather *infinitely* many bijections, all of them of exactly the same value (each local coordinate system provides a bijection). Fortunately, it turns out that each choice of coordinates (and, consequently, of the corresponding bijection) leads to the same linear structure in T_PM. So, there is *the* linear structure in T_PM.

▼ Indeed, let $[\gamma]$ and $[\sigma]$ be two elements from T_PM such that

$$\begin{aligned} [\gamma] \leftrightarrow a \quad &\text{and} \quad [\sigma] \leftrightarrow b \qquad && \text{w.r.t. } x \\ \leftrightarrow a' \quad &\text{and} \quad \leftrightarrow b' && \text{w.r.t. } x' \end{aligned}$$

Then, *by definition*,

$$\begin{aligned} [\gamma] + \lambda[\sigma] &\leftrightarrow a + \lambda b \qquad && \text{w.r.t. } x \\ &\leftrightarrow a' + \lambda b' && \text{w.r.t. } x' \end{aligned}$$

Do both prescriptions (via x and via x') lead to the same result? They do, if (and only if)

$$J(a + \lambda b) = a' + \lambda b'$$

(i.e. if first performing the linear combination via x and then translating the result into the language of the x'-tribe gives the same object as performing the linear combination directly via x'). But this *is* clearly so, since

$$J(a + \lambda b) = Ja + \lambda Jb = a' + \lambda b'$$

This shows that we can induce linear structure into $T_P M$ using *any* bijection $T_P M \leftrightarrow \mathbb{R}^n$ (i.e. using *any* local coordinate system x). The result is always the same: there is *canonical* linear structure in $T_P M$. ▲

Finally, let's have a look at an appropriate *basis* in $T_P M$. Well, since there is a standard ("natural") basis E_j in $\mathbb{R}^n$ (number 1 at the j-th place of the n-tuple and 0 at all remaining places), the bijection via coordinates x gives us unique equivalence class $[\gamma_j]$ which corresponds to E_j. And then, for general $[\gamma] \leftrightarrow a = a^j E_j$, we get

$$[\gamma] = a^j [\gamma_j]$$

And what about the *representative* γ_j (i.e. the j-th *curve* itself) generating the *class* $[\gamma_j]$? Well, we need to return back to the general formula

$$x^i(\gamma(t)) = x^i(P) + ta^i + \dots$$

and set a^i corresponding to the basis vector E_j, i.e. $a^i = \delta^i_j$. We get

$$x^i(\gamma_j(t)) = x^i(P) + t\delta^i_j + \dots$$

Any curve $\gamma_j(t)$ with this first order approximation induces the class $[\gamma_j]$. The simplest one is clearly the curve

$$x^i(\gamma_j(t)) = x^i(P) + t\delta^i_j$$

(i.e. such that there are "no dots at all" at the end of the expansion). Such curves $\gamma_j(t)$ (for all $j = 1, \dots, n$) are known as *coordinate curves.* (For j-th coordinate curve, $\gamma_j(t)$, the only coordinate whose value changes as t increases is x^j, namely the coordinate x^j increases *linearly.*) So, in this approach to vectors on a manifold, we can take as a basis of the tangent space $T_P M$ the vectors $[\gamma_j(t)]$, equivalence classes generated by coordinate curves $\gamma_j(t)$.

Recall that there is an alternative notation $\dot\gamma$ for the equivalence class $[\gamma]$ (see p.24 in the book; actually $\dot\gamma$ is used much more frequently than $[\gamma]$). Using both notations, we can write for a general vector v

$$v = \dot\gamma \equiv [\gamma] = a^i[\gamma_i] = a^i \dot\gamma_i$$

$\boxed{2.4.3}$ and $\boxed{2.4.4}$ p.35: A slightly more detailed discussion of higher dual spaces might be useful. In particular, we show here, first, how *natural bases* are introduced in all those spaces. And second, how one easily gains, from elementary properties of those bases, an alternative insight into the canonical isomorphism $L \to L^{**}$ from (2.4.3).

Recall that the n-th dual space of L is simply the dual space of the $(n-1)$-st one. That's why the construction of the dual basis from (2.4.2) may be *repeated* for each pair of adjacent duals. So, as soon as we pick up a basis e_a in L, the whole (infinite) chain of bases in *all* higher duals is canonically defined. Now, a change of basis $e_a \mapsto e'_a = A^b_a e_b$ in L by means of a non-singular matrix A results, according to (2.4.2), in the change of basis $e^a \mapsto e'^a = (A^{-1})^a_b e^b$ by means of A^{-1} in L^*. Then, however, we also get $E_a \mapsto E'_a = ((A^{-1})^{-1})^b_a E_b \equiv A^b_a E_b$, $E^a \mapsto E'^a = (((A^{-1})^{-1})^{-1})^a_b E^b \equiv (A^{-1})^a_b E^b$ etc.

So, the overall situation reads as follows:

space	L	L^*	L^{**}	L^{***}	L^{****}	L^{*****}	etc.
basis	e_a	e^a	E_a	E^a	$\mathcal{E}_a$	$\mathcal{E}^a$	etc.
scrambled by	A	A^{-1}	A	A^{-1}	A	A^{-1}	etc.

where

$$\delta^a_b = \langle e^a, e_b\rangle = \langle E_b, e^a\rangle = \langle E^a, E_b\rangle = \langle \mathcal{E}_b, E^a\rangle = \langle \mathcal{E}^a, \mathcal{E}_b\rangle = \text{ etc.}$$

Now it is clear that the correspondence

$$e_a \;\leftrightarrow\; E_a \;\leftrightarrow\; \mathcal{E}_a \;\leftrightarrow\; \text{etc.}$$

provides the *canonical* isomorphisms between members of the (infinite) chain of the spaces

$$L \;\leftrightarrow\; L^{**} \;\leftrightarrow\; L^{****} \;\leftrightarrow\; \text{etc.}$$

First, it is clearly a bijection (i.e. a *linear isomorphism*, when extended by linearity from the basis to the entire space L). Moreover, and this is crucial observation, it *does not* depend on the choice of a basis in L, since the bases within all members of the chain are scrambled in *the same way* - in terms of the matrix A.

[Compare this with the similarly looking correspondence $e_a \leftrightarrow e^a$. It gives a perfect isomorphism of L and L^* as long as we keep the basis e_a fixed. However, as soon as we change the basis e_a in L, the isomorphism is replaced by a *different* one, given by the *new* correspondence $A^b_a e_b \leftrightarrow (A^{-1})^a_b e^b$.]

And, perhaps, one should not be overly surprised that the isomorphism $L \leftrightarrow L^{**}$ introduced here is nothing but the one proposed in the hint to (2.4.3), i.e. the isomorphism f given by the formula

$$f : L \to L^{**} \qquad \langle f(v), \alpha\rangle := \langle \alpha, v\rangle$$

Indeed, in terms of the bases e_a and E_a in L and L^{**} respectively we have

$$f(e_a) = B^b_a E_b$$

with a matrix B. Then

$$\begin{aligned} \langle f(e_a), e^b\rangle &\overset{1.}{=} \langle B^c_a E_c, e^b\rangle = B^c_a \langle E_c, e^b\rangle = B^b_a \\ &\overset{2.}{=} \langle e^b, e_a\rangle = \delta^b_a \end{aligned}$$

so that

$$f(e_a) = E_a$$

2.4.18 p.44: For a general linear map $A : L_1 \to L_2$, the statement is only true for *positive* (or negative) *definite* metric tensor h in L_2. If h fails to be positive (or negative) definite, induced tensor A^*h still may be "good", i.e. non-degenerate, but one must be careful in choosing "good" map A.)

Consider, to illustrate the matter, the two-dimensional "Minkowski" (linear) space, i.e. the (indefinite) space (L_2, h) with

$$h = e^1 \otimes e^1 - e^2 \otimes e^2$$

Let L_1 be one-dimensional, spanned by $\hat{e}_1$. A general linear map $A : L_1 \to L_2$ is given as

$$A : \hat{e}_1 \mapsto u e_1 + v e_2 \qquad\qquad u, v \in \mathbb{R}$$

Then (check)

$$A^*h = \cdots = (u^2 - v^2)\ \hat{e}^1 \otimes \hat{e}^1$$

So the resulting induced tensor A^*h is

- positive definite (= non-degenerate) for $u^2 - v^2 > 0$
- negative definite (= non-degenerate) for $u^2 - v^2 < 0$
- vanishing (most degenerate possible)) for $u^2 - v^2 = 0$ (i.e. for $u = \pm v$)

The problem lies in the *third* possibility. Although, say,

$$A : \hat{e}_1 \mapsto e_1 + e_2$$

is a perfect *maximum-rank* linear map $L_1 \to L_2$ (the rank being 1 :-), the resulting induced tensor A^*h vanishes, i.e., it is definitely *not* a *metric* tensor in L_1. (Actually, it is induced from an *isotropic subspace* of L_2.)

2.6 p.52: At the end of the page, grad$f$, the gradient of f, is introduced as a *vector* field (rather than the *covector* field df). It's important property should be mentioned: at any point x, the vector grad$f$ is *orthogonal* to the level surface of f passing through x (i.e. the hypersurface $f =$ const.). [Meaning it is orthogonal to any vector v *tangent* to the surface.] Indeed, for any vector v we have

$$g(\text{grad}f, v) = \langle df, v\rangle = vf$$

due to

$$g(\text{grad}f, v) = g_{ij}(\text{grad}f)^i v^j = g_{ij} g^{ik} (df)_k v^j = (df)_k v^k = \langle df, v\rangle$$

(the first equal sign) and 2.5.3(i) (the second equal sign). Then

$$g(\mathrm{grad} f, v) = 0 \quad \Leftrightarrow \quad vf = 0 \quad \Leftrightarrow \quad v \text{ is tangent to the hypersurface } \{f = \text{ const.}\}$$

There is a "skew" analog of all this (see Chapter 14): the "skew gradient" of f (= the "hamiltonian field" ζ_f generated by f) is "skew-orthogonal" to the hypersurface $f =$ const., meaning

$$\omega(\zeta_f, v) = 0 \quad \Leftrightarrow \quad vf = 0 \quad \Leftrightarrow \quad v \text{ is tangent to the hypersurface } \{f = \text{ const.}\})$$

("Skew-orthogonality" is defined with respect to ω.)

2.6 p.52: The components of the gradient (as a *vector* field) are given by $(\mathrm{grad} f)^i = g^{ik}\partial_k f$. You may be disappointed with this result, however, trying to compute it, say, in spherical polar coordinates and then checking it versus the result in your favorite book on, say, electrodynamics. The two results *differ*. Why? Because in your electrodynamics book *orthonormal* components are displayed whereas the result mentioned here refers to *coordinate* components, respectively. See more details in Chapter 8 (e.g. 8.5.10).

3.1.7 p.59: In order to prove

$$f^* \circ C = C \circ f^*$$

for a diffeomorphism $f : M \to N$, one should realize that contractions are performed on *different manifolds* (M and N, respectively) where, consequently, *different frame fields* are used.

Assume that e_i and e^i is a (mutually dual) frame and co-frame field on (a part of) M and the same is true for e_a and e^a on (a part of) N. Then

$$\begin{aligned}(f^* \circ Ct)(V, \dots; \alpha, \dots) &= (Ct)(f_* V, \dots; f_* \alpha, \dots) \\ &= t(f_* V, \dots, e_a, \dots; f_* \alpha, \dots, e^a, \dots)\end{aligned}$$

whereas

$$\begin{aligned}(C \circ f^* t)(V, \dots; \alpha, \dots) &= (f^* t)(V, \dots, e_i, \dots; \alpha, \dots, e^i, \dots) \\ &= t(f_* V, \dots, f_* e_i, \dots; f_* \alpha, \dots, f_* e^i, \dots)\end{aligned}$$

So, the second result may be obtained from the first one by the replacement $e_a \mapsto f_* e_i$ and $e^a \mapsto f_* e^i$. Notice however that, f being a *diffeomorphism*, $f_* e_i$ and $f_* e^i$ is an equally good mutually dual frame/co-frame field on N as is e_a and e^a. And since in (2.4.8) we learned that contraction does not depend on the choice of the frame/co-frame field (provided that they satisfy the duality condition), we are done.

4.3.6 p.74: A useful fact about commutator of vector fields is that it is *preserved* by maps. Let us explain in more detail what does it mean and prove it is really the case.

Consider two vector fields U, V on M. Let $f : M \to N$ be a map and suppose the fields U, V are *projectable* with respect to f (see the text before 3.1.3). Denote the images $\tilde{U}, \tilde{V}$ (they are vector fields on N)

$$f_* U = \tilde{U} \qquad f_* V = \tilde{V}$$

Notice that the fact that U is projectable (and $f_* U = \tilde{U}$) may also be expressed in terms of action of fields on functions as follows (recall that $(U\chi)(m) = U_m \chi$)

$$\begin{aligned} &(\tilde{U}\psi) \circ f = U(\psi \circ f) && (\equiv U(f^* \psi)) \\ \text{or} \quad &\tilde{U}_{f(m)} \psi = U_m(\psi \circ f) && (\equiv U_m(f^* \psi)) \end{aligned}$$

where ψ is a function on N. The same is clearly true for V. Then, the straightforward calculation shows that

$$\begin{aligned}(f_*[U, V]_m)\psi &= [U, V]_m(\psi \circ f) \\ &= U_m(V(\psi \circ f)) - V_m(U(\psi \circ f)) \\ &= U_m((\tilde{V}\psi) \circ f) - V_m((\tilde{U}\psi) \circ f) \\ &= \tilde{U}_{f(m)}(\tilde{V}\psi) - \tilde{V}_{f(m)}(\tilde{U}\psi) \\ &= [\tilde{U}, \tilde{V}]_{f(m)}\psi\end{aligned}$$

i.e.

$$f_*[U,V]_m = [\tilde{U},\tilde{V}]_{f(m)}$$

or, finally, we get the preservation of commutator of (f-projectable, otherwise the r.h.s. makes no sense) vector fields in the form

$$\boxed{f_*[U,V] = [f_*U, f_*V]}$$

Note: this fact is (implicitly) also present for the particular case of a *diffeomorphism* $f : M \to M$ (where $f^* = f_*^{-1}$) in the result $f^*\mathcal{L}_V = \mathcal{L}_{f^*V} f^*$ of 8.3.7. Indeed, when applied on U we get

$$f^*(\mathcal{L}_V U) \equiv f^*[V,U] = \mathcal{L}_{f^*V} f^*U \equiv [f^*V, f^*U]$$

The computation displayed above shows its validity for a *general* (smooth) map $f : M \to N$.

$\boxed{6.2.8, 6.2.9}$ p.131: the hint for the proof might be even simpler. We are to prove (in 6.2.8, say) the Cartan's "magic formula"

$$\mathcal{L}_V = i_V d + d i_V$$

It is written in the hint, that the formula to be proved is an equality of two derivations (of degree 0) of the algebra $\Omega(M)$ and therefore it is enough to verify it on degrees 0 and 1, "where it is easy (e.g. in components"). Now, for the degree 1 case, no components are needed. Actually 1-forms of the structure $d\psi$ (i.e. *exact* 1-forms) *suffice*. (Observe that all p-forms are *sums* of expressions of the structure $f d\psi \wedge \cdots \wedge d\chi$, i.e. *products* of 0-forms and *exact* 1-forms. This is, btw. used in the hint for 6.2.11!) So, one is only to prove

$$\mathcal{L}_V f = (i_V d + d i_V) f$$
$$\mathcal{L}_V df = (i_V d + d i_V) df$$

which is evident (just definitions are needed).

$\boxed{7.2.4}$ p.148: Some more details to the hint: Denote $P_{j\mu}$ and $Q_{j\mu}$ the j-th coordinate of the points P_μ and Q_μ respectively ($j = 1,\ldots,n$, $\mu = 0,1,\ldots,n$). Then the fact, that $x \mapsto Ax + a$ maps P_μ into Q_μ (for each $\mu = 0,1,\ldots,n$) leads to the system of equations

$$A_{ij} P_{j\mu} + a_i = Q_{i\mu} \qquad \mu = 0,1,\ldots,n$$

We are to show that, given the points P_μ and Q_μ, a unique pair (A,a) emerges. If we choose $P_0 = (0,\ldots,0), P_1 = (1,\ldots,0),\ldots, P_n = (0,\ldots,1)$, so that $P_{j0} = 0, P_{jk} = \delta_{jk}$, the equations read

$$a_i = Q_{i0} \qquad A_{ij}\delta_{jk} + a_i = Q_{ik}$$

so that

$$a_i = Q_{i0} \qquad A_{ij} = Q_{ij} - Q_{i0}$$

(one can also obtain the same result without any computation after some contemplation). So, any simplex may be obtained from the *standard* one by means of a *unique affine* transformation. Consequently, any two simplices may be mapped into one another by means of a *unique affine* transformation. [Map the first simplex to the standard one (by the inverse of (A,a) given above) and then, in the second step, map the standard one to the second simplex.]

$\boxed{7.2.6}$ p.149: In part (ii) the third, more explicit (and useful), characterization of the standard simplex $\bar{s}_p$ may be mentioned:

$$\bar{s}_p : \qquad (x^1,\ldots,x^p) \qquad \text{such that} \qquad \begin{array}{l} 0 \le x^1 \le 1 \\ 0 \le x^2 \le 1 - x^1 \\ 0 \le x^3 \le 1 - x^1 - x^2 \\ \qquad \cdots \\ 0 \le x^p \le 1 - x^1 - \cdots - x^{p-1} \end{array}$$

Therefore, integration over $\bar{s}_p$ amounts to

$$\int_{\bar{s}_p} (\ldots) = \int_0^1 dx^1 \int_0^{1-x^1} dx^2 \cdots \int_0^{1-(x^1+\cdots+x^{p-1})} dx^p \; (\ldots)$$

Notice that

$$\int_{\bar{s}_{p+1}} (\dots) = \int_{s_p} \int_0^{1-(x^1+\dots+x^p)} dx^{p+1}\ (\dots)$$

where s_p at the r.h.s. is the *base* s_p of the $\bar{s}_{p+1}$ (such face of the boundary, where $x^{p+1} = 0$). So, as an example,

$$\int_{\bar{s}_3} (\dots) = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-(x+y)} dz\ (\dots) = \int_{s_2} \int_0^{1-(x+y)} dz\ (\dots)$$

where s_2 at the r.h.s. is the *base* s_2 of the $\bar{s}_3$ (such face of the boundary, where $z = 0$; the "floor" of the 3-simplex).

7.5.3 p.153: more details for *arbitrary* p:
1.: The boundary of the standard simplex $\bar{s}_{p+1} \equiv (P_0, P_1, \dots, P_{p+1})$ contains $p + 1$ terms. Among them, there are $p - 1$ "internal" terms, where the points $P_1, \dots, P_p$ are omitted (in each term a single point). Realize that when P_k (the k-th one) is omitted, the corresponding p-simplex does not "stick into the k-th dimension" (for all the points $x^k = 0$ holds). This means, however, that the pull-back of the differential dx^k with respect to the map Φ_k (into the standard simplex $\bar{s}_p$) vanishes ($\Phi_k^* dx^k = d0 = 0$). This leads to vanishing of the *complete* form $\Phi_k^* \eta = \Phi_k^*(f dx^1 \wedge \dots \wedge dx^p)$ to be integrated over *this particular* $\bar{s}_p$. Since this is true for all "internal" points $P_1, \dots, P_p$ mentioned above, there are *just two* terms of the boundary which *do* contribute to the integral - those in which *the first* and *the last* points are omitted respectively. They read explicitly

$$\begin{aligned}\partial \bar{s}_{p+1} &= \partial(P_0, P_1, \dots, P_{p+1}) \\ &= (P_1, \dots, P_{p+1}) + (-1)^{p+1}(P_0, P_1, \dots, P_p) \text{ (plus irrelevant terms)} \\ &= (-1)^{p+1}[(P_0, P_1, \dots, P_p) - (P_{p+1}, P_1, \dots, P_p)]\end{aligned}$$

2.: So, the "boundary" side of the Stokes theorem is

$$\int_{\partial \bar{s}_{p+1}} \eta = (-1)^{p+1}[\int_{(P_0,P_1,\dots,P_p)} \eta - \int_{(P_{p+1},P_1,\dots,P_p)} \eta]$$

where

$$\eta := f(x^1, \dots, x^p, x^{p+1}) dx^1 \wedge \dots \wedge dx^p$$

3.: The "bulk" side of the Stokes theorem may be expressed, as mentioned in the book (see also comment to (7.2.6) in this text), in the form

$$\int_{\bar{s}_{p+1}} d\eta = (-1)^{p+1}[\int_{s_p} dx^1 \dots dx^p \{f(x^1, \dots, x^p, 0) - f(x^1, \dots, x^p, 1 - (x^1 + \dots + x^p))\}]$$

where

$$\int_{s_p} dx^1 \dots dx^p\ (\dots) \equiv \int_0^1 dx^1 \int_0^{1-x^1} dx^2 \cdots \int_0^{1-(x^1+\dots+x^{p-1})} dx^p\ (\dots)$$

4.: Then, the statement $\int_{\bar{s}_{p+1}} d\eta = \int_{\partial \bar{s}_{p+1}} \eta$ is true if

$$\int_{(P_0,P_1,\dots,P_p)} f(x^1, \dots, x^p, x^{p+1}) dx^1 \wedge \dots \wedge dx^p = \int_{s_p} dx^1 \dots dx^p f(x^1, \dots, x^p, 0)$$

$$\int_{(P_{p+1},P_1,\dots,P_p)} f(x^1, \dots, x^p, x^{p+1}) dx^1 \wedge \dots \wedge dx^p = \int_{s_p} dx^1 \dots dx^p f(x^1, \dots, x^p, 1 - (x^1 + \dots + x^{p-1}))$$

5.: The p-simplex $(P_0, P_1, \dots, P_p)$ forms the *base* (floor) of the standard $(p + 1)$-simplex $\bar{s}_{p+1}$ (so $x^{p+1} = 0$ on it). It is clearly the Φ-image of the standard p-simplex $\bar{s}_p$:

$$(P_0, P_1, \dots, P_p) = \Phi(\bar{s}_p) \qquad \Phi : (u^1, \dots, u^p) \mapsto (u^1, \dots, u^p, 0)$$

Using the definition $\int_{\Phi(\bar{s}_p)} \eta = \int_{\bar{s}_p} \Phi^* \eta$ we immediately get the *first* equality in 4.
6.: The p-simplex $(P_{p+1}, P_1, \dots, P_p)$ forms the *slant* face of the standard $(p+1)$-simplex $\bar{s}_{p+1}$. (One obtains $(P_{p+1}, P_1, \dots, P_p)$ from $(P_0, P_1, \dots, P_p)$ by lifting the vertex P_0 to the height 1 along x^{p+1} and a general point $(x^1, \dots, x^p)$ inside to the height $1 - (x^1 + \dots + x^p)$; draw the pictures for $p = 1, 2$.) It is the Φ-image of the standard p-simplex $\bar{s}_p$:

$$(P_{p+1}, P_1, \dots, P_p) = \Phi(\bar{s}_p) \qquad \Phi : (u^1, \dots, u^p) \mapsto (u^1, \dots, u^p, 1 - (u^1 + \dots + u^p))$$

Using again the definition $\int_{\Phi(\bar{s}_p)} \eta = \int_{\bar{s}_p} \Phi^* \eta$ for *this* Φ, we get the *second* equality in 4.

8.1.3 p.165: Here we learned the (generalized) "integration by parts" formula

$$\int_D d\alpha \wedge \beta = -\int_D \hat{\eta}\alpha \wedge d\beta + \int_{\partial D} \alpha \wedge \beta$$

which holds for any two forms α and β such that their degrees fulfil $\deg \alpha + \deg \beta + 1 = n$ and D being an n-dimensional domain, and, as a simplest special case, the "ordinary" 1-dimensional formula

$$\int_a^b f'(x) g(x) dx = -\int_a^b f(x) g'(x) dx + [fg]_a^b$$

Let us mention (here) what the general formula offers in *vector analysis* context (see Section 8.5. for treatment of vector analysis in terms of forms).

Here (in E^3), we have $\deg \alpha + \deg \beta = 2$, leaving just *two* cases of interest,

$$\begin{array}{lll} 1. & \deg \alpha = 0 & \deg \beta = 2 \\ 2. & \deg \alpha = 1 & \deg \beta = 1 \end{array}$$

From (8.5.2) then

$$\begin{array}{lll} 1. & \alpha = f & \beta = \mathbf{A} \cdot d\mathbf{S} \\ 2. & \alpha = \mathbf{A} \cdot d\mathbf{r} & \beta = \mathbf{B} \cdot d\mathbf{r} \end{array}$$

so that (from (8.5.4))

$$\begin{array}{lll} 1. & d\alpha = (\nabla f).d\mathbf{r} & d\beta = (\operatorname{div} \mathbf{A})\ dV \\ 2. & d\alpha = \operatorname{curl} \mathbf{A} \cdot d\mathbf{S} & d\beta = \operatorname{curl} \mathbf{B} \cdot d\mathbf{S} \end{array}$$

Therefore, taking int account

$$(\mathbf{A}.d\mathbf{r}) \wedge (\mathbf{B}.d\mathbf{r}) = (\mathbf{A} \times \mathbf{B}).d\mathbf{S}$$
$$(\mathbf{A}.d\mathbf{r}) \wedge (\mathbf{B}.d\mathbf{S}) = (\mathbf{A}.\mathbf{B})dV$$

from (8.5.8), the "general by parts" formula gives us the following two integral identities:

$$\begin{array}{ll} 1. & \int_D (\nabla f \cdot \mathbf{A})\ dV = -\int_D f(\operatorname{div} \mathbf{A})\ dV + \int_{\partial D} (f\mathbf{A}) \cdot d\mathbf{S} \\ 2. & \int_D (\mathbf{B} \cdot \operatorname{curl} \mathbf{A})\ dV = \int_D (\mathbf{A} \cdot \operatorname{curl} \mathbf{B})\ dV + \int_{\partial D} (\mathbf{A} \times \mathbf{B}) \cdot d\mathbf{S} \end{array}$$

8.1.4 p.166: see additional material to (8.5.6) here

8.3.6 p.174: It is useful to realize that j_V depends actually on as many as *two* tensor fields, V *and* g. So, the more precise way is to write it as $j_{V,g}$:

$$j_{V,g}\alpha := g(V, \ .\ ) \wedge \alpha$$

Then, the fact that it is natural with respect to diffeomorphisms reads

$$\boxed{f^* j_{V,g}\alpha = j_{f^*V, f^*g} f^* \alpha}$$

This is indeed readily verified:

$$f^*(j_{V,g}\alpha) = f^*(g(V,\ .\ ) \wedge \alpha) = f^*(g(V,\ .\ )) \wedge f^*\alpha = (f^*g)(f^*V,\ .\ ) \wedge f^*\alpha \equiv j_{f^*V,f^*g} f^*\alpha$$

Or, alternatively, one can use the flat symbol $\flat_g$ to express

$$j_{V,g}\alpha := (\flat_g V) \wedge \alpha$$

and write

$$f^*(j_{V,g}\alpha) = f^*((\flat_g V) \wedge \alpha) = (f^*(\flat_g V)) \wedge (f^*\alpha) = (\flat_{f^*g} f^*V) \wedge (f^*\alpha) \equiv j_{f^*V,f^*g} f^*\alpha$$

where we used naturalness of $\flat_g$ (see 8.3.8)

8.5.4 p.181: *Two*-dimensional vector analysis (in E^2 rather than E^3) is mentioned (in a footnote, p.179) as a simpler realization of the usual one. Perhaps some additional remarks might be useful concerning this topics. (See also additional material to (8.5.6) here.)

First, rewrite the diagram (8.5.4) (valid for three dimensions)

$$\begin{array}{ccccccc}
\Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) & \xrightarrow{d} & \Omega^3(M) \\
\downarrow \scriptstyle{\text{id}} & & \downarrow \scriptstyle{\sharp} & & \downarrow \scriptstyle{\sharp *} & & \downarrow \scriptstyle{*} \\
\mathcal{F}(M) & \xrightarrow[\text{grad}]{} & \mathfrak{X}(M) & \xrightarrow[\text{curl}]{} & \mathfrak{X}(M) & \xrightarrow[\text{div}]{} & \mathcal{F}(M)
\end{array}$$

in the form

$$\begin{array}{ccccccc}
\Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) & \xrightarrow{d} & \Omega^3(M) \\
\downarrow \scriptstyle{a_0} & & \downarrow \scriptstyle{a_1} & & \downarrow \scriptstyle{a_2} & & \downarrow \scriptstyle{a_3} \\
\mathcal{F}(M) & \xrightarrow[A_1]{} & \mathfrak{X}(M) & \xrightarrow[A_2]{} & \mathfrak{X}(M) & \xrightarrow[A_3]{} & \mathcal{F}(M)
\end{array}$$

where a_0, a_1, a_2, a_3 denote isomorphisms encoding scalar and vector fields into forms of various degrees and A_1, A_2, A_3 are the effective differential operators from vector analysis. Then, clearly

$$\begin{aligned}
A_1 &= a_1\ d\ a_0^{-1} \\
A_2 &= a_2\ d\ a_1^{-1} \\
A_3 &= a_3\ d\ a_2^{-1}
\end{aligned}$$

so that

$$A_2\ A_1 = 0 \qquad A_3\ A_2 = 0$$

These are, in reality, the statements

$$\text{curl grad} = 0 \qquad \text{div rot} = 0$$

Now, the corresponding general counterpart in *two* dimensions reads

$$\begin{array}{ccccc}
\Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\
\downarrow \scriptstyle{a_0} & & \downarrow \scriptstyle{a_1} & & \downarrow \scriptstyle{a_2} \\
\mathcal{F}(M) & \xrightarrow[A_1]{} & \mathfrak{X}(M) & \xrightarrow[A_2]{} & \mathcal{F}(M)
\end{array}$$

$$A_1 = a_1\ d\ a_0^{-1} \qquad A_2 = a_2\ d\ a_1^{-1} \qquad A_2\ A_1 = 0$$

Notice, however, that on a two-dimensional M the star operator $*$, when applied to "middle degree forms", is an operator *on* Ω^1 (so it does not connect forms of different degrees, $\Omega^1(M) \leftrightarrow \Omega^2(M)$, as is the case in E^3). Therefore, there are as many as *two distinct* natural identifications of vector fields with one-forms, $\Omega^1(M) \leftrightarrow \mathfrak{X}(M)$: via $\sharp : \Omega^1(M) \to \mathfrak{X}(M)$ and via $\sharp * : \Omega^1(M) \to \mathfrak{X}(M)$. Or, when looking from the other side, there are two ways how vector fields may be encoded into *one*-forms, via $\flat : \mathfrak{X}(M) \to \Omega^1(M)$ and $*\flat : \mathfrak{X}(M) \to \Omega^1(M)$.

Therefore we get, instead of the single diagram in three dimensions, as many as *two* corresponding general diagrams in *two*-dimensional vector analysis:

$$\begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ a_0 \big\downarrow & & \big\downarrow a_1 & & \big\downarrow a_2 \\ \mathcal{F}(M) & \xrightarrow[A_1]{} & \mathfrak{X}(M) & \xrightarrow[A_2]{} & \mathcal{F}(M) \end{array} \qquad \text{and} \qquad \begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ b_0 \big\downarrow & & \big\downarrow b_1 & & \big\downarrow b_2 \\ \mathcal{F}(M) & \xrightarrow[B_1]{} & \mathfrak{X}(M) & \xrightarrow[B_2]{} & \mathcal{F}(M) \end{array}$$

$$A_1 = a_1 \ d \ a_0^{-1} \qquad A_2 = a_2 \ d \ a_1^{-1} \qquad A_2 \ A_1 = 0$$

$$B_1 = b_1 \ d \ b_0^{-1} \qquad B_2 = b_2 \ d \ b_1^{-1} \qquad B_2 \ B_1 = 0$$

When concrete isomorphisms a_i and b_i are inserted, we get

$$\begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \mathrm{id} \big\downarrow & & \big\downarrow \sharp & & \big\downarrow * \\ \mathcal{F}(M) & \xrightarrow[\mathrm{grad}]{} & \mathfrak{X}(M) & \xrightarrow[\mathcal{D}_2]{} & \mathcal{F}(M) \end{array} \qquad \text{and} \qquad \begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \mathrm{id} \big\downarrow & & \big\downarrow \sharp * & & \big\downarrow * \\ \mathcal{F}(M) & \xrightarrow[\mathcal{D}_1]{} & \mathfrak{X}(M) & \xrightarrow[-\mathrm{div}]{} & \mathcal{F}(M) \end{array}$$

and

$$\mathrm{grad} \ = \sharp \ d \qquad \mathcal{D}_2 = * \ d \ \flat \qquad \mathcal{D}_2 \ \mathrm{grad} \ = 0$$

$$\mathcal{D}_1 = \sharp * \ d \qquad \mathrm{div} \ = *^{-1} \ d \ * \flat \qquad \mathrm{div} \ \mathcal{D}_1 = 0$$

(We have, on vector fields in E^2,

$$*d(\sharp *)^{-1} = *d *^{-1} \flat = - * d * \flat = - * \hat{\eta} d * \flat = - *^{-1} \ d * \flat = *^{-1} d * \hat{\eta} \flat = \delta \flat = -\mathrm{div}$$

due to $*^{-1} = *\hat{\eta}$, see 5.8.2, and $\delta \tilde{V} = -\mathrm{div}\, V$, see 8.3.4).

So, in comparison with the three-dimensional case, there are still, in two dimensions, the two good old operations grad and div, but, first, curl (as an operation vector field $\mapsto$ vector field) is missing and, second, two "new" differential operations (denoted here as $\mathcal{D}_1$ and $\mathcal{D}_2$) emerged.[1] The first, $\mathcal{D}_1$, is of type scalar $\mapsto$ vector (like the gradient) whereas the second, $\mathcal{D}_2$, is of type vector $\mapsto$ scalar (like the divergence). In particular, in Cartesian coordinates in E^2 we have

$$\begin{array}{rcl} \mathrm{grad} : \qquad f & \mapsto & (\partial_1 f, \partial_2 f) \\ \mathcal{D}_1 : \qquad f & \mapsto & (-\partial_2 f, \partial_1 f) \\ \mathcal{D}_2 : (A_1, A_2) & \mapsto & (\partial_1 A_2 - \partial_2 A_1) \\ \mathrm{div} : (A_1, A_2) & \mapsto & (\partial_1 A_1 + \partial_2 A_2) \end{array}$$

Observe that the composition of two operators belonging to the same diagram indeed vanishes (since it is just masked $dd = 0$)

$$\begin{array}{rcl} \mathcal{D}_2 \ \mathrm{grad} : \qquad f & \mapsto & (\partial_1 f, \partial_2 f) \mapsto (\partial_1 \partial_2 f - \partial_2 \partial_1 f) = 0 \\ \mathrm{div} \ \mathcal{D}_1 : \qquad f & \mapsto & (-\partial_2 f, \partial_1 f) \mapsto (\partial_1(-\partial_2 f) + \partial_2(\partial_1 f)) = 0 \end{array}$$

Composition of two operators belonging to distinct diagrams leads (for both possibilities) to the *Laplace* operator. Abstractly

$$\mathrm{div} \ \mathrm{grad} \ = \mathcal{D}_2 \ \mathcal{D}_1 = *d * d = \triangle$$

[1]Notice on component expressions below, however, that $\mathcal{D}_2 \mathbf{A}$ is just the 3-rd component of the three-dimensional curl f. Also $\mathcal{D}_1 f$ may be regarded as *hamiltonian* vector field generated by f, since the (metric) volume form ω can be viewed as a *symplectic* form (see below; also Chapter 14).

and in particular, in Cartesian coordinates

$$\begin{aligned}
\text{div grad}: &\qquad f \mapsto (\partial_1 f, \partial_2 f) \mapsto (\partial_1\partial_1 f + \partial_1\partial_1 f) = \triangle f \\
\mathcal{D}_2\ \mathcal{D}_1: &\qquad f \mapsto (-\partial_2 f, \partial_1 f) \mapsto (\partial_1(\partial_1 f) - \partial_2(-\partial_2 f)) = \triangle f
\end{aligned}$$

There is still another (and perhaps the most natural) point of view: the manifold E^2 is, at the same time, Riemannian and *symplectic* manifold (see Chapter 14).[2] The symplectic form is just $dx \wedge dy$, the standard (metric) *volume* form in E^2. Therefore, there are as many as *two* natural ways of identifying vectors with 1-forms (i.e. two natural ways of raising and lowering of indices) - in terms of the metric tensor g and in terms of the symplectic form ω. Therefore,

- two vector fields may be associated with gradient df as a covector field, the *gradient* $\nabla f \equiv \operatorname{grad} f$ as well as the *Hamiltonian field* ζ_f; the latter is nothing but $\mathcal{D}_1 f$:-)

- two ways of lowering indices on a vector field W result in two distinct 1-forms associated with the vector field; then the exterior derivative d produces two distinct 2-forms, both of them proportional, however, to the (metric) volume form; so, the "proportionality coefficients" may be regarded as (two distinct) functions associated to the vector field; they turn out to be (check) $\mathcal{D}_2 W$ (for the metric version) and $\operatorname{div} W$ (for the symplectic version) respectively.

As a simple (and perhaps useful) application, consider 2-*dimensional incompressible fluid flow.* Incompressibility gives (from continuity equation)

$$\operatorname{div} v = 0$$

So, at least locally

$$v = \mathcal{D}_1 \psi \equiv \zeta_\psi$$

for some function ψ (see the right diagram above). This function ("potential" of the velocity field) is known as *stream function* in 2D-hydrodynamics. Why? Since $\mathcal{D}_1\psi$ is nothing but the Hamiltonian field generated by ψ, we get

$$v\psi = \zeta_\psi \psi = \{\psi, \psi\} = 0$$

(Poisson bracket of two equal entries). So ψ is constant along stream-lines. Put it differently, the fluid flow is everywhere *parallel to contours* (level lines) of ψ. So, by plotting level sets (lines, here) of ψ, we can "see", modulo direction, the fluid flow.

Notice also that when Laplace operator is applied on ψ, we get *vorticity function* $\hat{\omega}$

$$\Delta\psi = \hat{\omega}$$

Indeed, we get

$$\Delta\psi = \mathcal{D}_2\ \mathcal{D}_1\psi = \mathcal{D}_2 v = *\ d\ \flat v = *\ (d\tilde{v})$$

Now $d\tilde{v}$ is just vorticity *form*, which, being a *two*-form, is necessarily proportional to the volume form, the proportionality factor being (by definition) the vorticity *function* $\hat{\omega}$. And since Hodge star gives just 1 when applied on the volume form, we get the desired result.

8.5.6 p.182: When the general *Stokes* theorem $\int_D d\alpha = \oint_{\partial D} \alpha$ is applied on statements obtained from three individual blocks (three appearances of d) of the combined diagram

$$\begin{array}{ccccccc}
\Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) & \xrightarrow{d} & \Omega^3(M) \\
\downarrow \scriptstyle{\text{id}} & & \downarrow \scriptstyle{\sharp} & & \downarrow \scriptstyle{\sharp *} & & \downarrow \scriptstyle{*} \\
\mathcal{F}(M) & \xrightarrow[\text{grad}]{} & \mathfrak{X}(M) & \xrightarrow[\text{curl}]{} & \mathfrak{X}(M) & \xrightarrow[\text{div}]{} & \mathcal{F}(M)
\end{array}$$

we get the following *triple* of "classical" integral theorems

$$\begin{array}{ccc}
\Omega^0(M) & \xrightarrow{d} & \Omega^1(M) \\
\downarrow \scriptstyle{\text{id}} & & \downarrow \scriptstyle{\sharp} \\
\mathcal{F}(M) & \xrightarrow[\text{grad}]{} & \mathfrak{X}(M)
\end{array} \qquad \leftrightarrow \qquad (\nabla f).d\mathbf{r} = df \qquad \leftrightarrow \qquad \int_C (\nabla f).d\mathbf{r} = f(B) - f(A)$$

[2] Manifolds like E^2, where both Riemannian and symplectic structure is available (and, in addition, these structures are related in a specific way) are known as *Kähler manifolds*.

$$\begin{array}{ccc} \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \downarrow \sharp & & \downarrow \sharp * \\ \mathfrak{X}(M) & \xrightarrow[\text{curl}]{} & \mathfrak{X}(M) \end{array} \qquad \leftrightarrow \qquad (\text{curl}\,\mathbf{A}).d\mathbf{S} = d(\mathbf{A}.d\mathbf{r}) \qquad \leftrightarrow \qquad \int_S (\text{curl}\,\mathbf{A}).d\mathbf{S} = \oint_{\partial S} \mathbf{A}.d\mathbf{r}$$

$$\begin{array}{ccc} \Omega^2(M) & \xrightarrow{d} & \Omega^3(M) \\ \downarrow \sharp * & & \downarrow * \\ \mathfrak{X}(M) & \xrightarrow[\text{div}]{} & \mathcal{F}(M) \end{array} \qquad \leftrightarrow \qquad (\text{div}\,\mathbf{A})dV = d(\mathbf{A}.d\mathbf{S}) \qquad \leftrightarrow \qquad \int_D (\text{div}\,\mathbf{A})dV = \oint_{\partial D} \mathbf{A}.d\mathbf{S}$$

(the *gradient* one, the *Stokes* one and the *Gauss* one). Now, we can mimic the procedure for *both two*-dimensional diagrams (see comment to 8.5.4 here in this text)

$$\begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \text{id} \downarrow & & \downarrow \sharp & & \downarrow * \\ \mathcal{F}(M) & \xrightarrow[\text{grad}]{} & \mathfrak{X}(M) & \xrightarrow[\mathcal{D}_2]{} & \mathcal{F}(M) \end{array} \qquad \text{and} \qquad \begin{array}{ccccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \text{id} \downarrow & & \downarrow \sharp * & & \downarrow * \\ \mathcal{F}(M) & \xrightarrow[\mathcal{D}_1]{} & \mathfrak{X}(M) & \xrightarrow[-\text{div}]{} & \mathcal{F}(M) \end{array}$$

What we get is the following *quadruple* of "classical" (two-dimensional) integral theorems:

$$\begin{array}{ccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) \\ \text{id} \downarrow & & \downarrow \sharp \\ \mathcal{F}(M) & \xrightarrow[\text{grad}]{} & \mathfrak{X}(M) \end{array} \qquad \leftrightarrow \qquad (\nabla f).d\mathbf{r} = df \qquad \leftrightarrow \qquad \int_C (\nabla f).d\mathbf{r} = f(B) - f(A)$$

$$\begin{array}{ccc} \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) \\ \text{id} \downarrow & & \downarrow \sharp * \\ \mathcal{F}(M) & \xrightarrow[\mathcal{D}_1]{} & \mathfrak{X}(M) \end{array} \qquad \leftrightarrow \qquad *(\mathcal{D}_1 f).d\mathbf{r} = -df \qquad \leftrightarrow \qquad \int_C *(\mathcal{D}_1 f).d\mathbf{r} = f(A) - f(B)$$

$$\begin{array}{ccc} \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \downarrow \sharp & & \downarrow * \\ \mathfrak{X}(M) & \xrightarrow[\mathcal{D}_2]{} & \mathcal{F}(M) \end{array} \qquad \leftrightarrow \qquad (\mathcal{D}_2 \mathbf{A})dS = d(\mathbf{A}.d\mathbf{r}) \qquad \leftrightarrow \qquad \int_S (\mathcal{D}_2 \mathbf{A})dS = \oint_{\partial S} \mathbf{A}.d\mathbf{r}$$

$$\begin{array}{ccc} \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) \\ \downarrow \sharp * & & \downarrow * \\ \mathfrak{X}(M) & \xrightarrow[-\text{div}]{} & \mathcal{F}(M) \end{array} \qquad \leftrightarrow \qquad (\text{div}\,\mathbf{A})dS = d(*(\mathbf{A}.d\mathbf{r})) \qquad \leftrightarrow \qquad \int_S (\text{div}\,\mathbf{A})dS = \oint_{\partial S} *(\mathbf{A}.d\mathbf{r})$$

The first and the second integral statements just say (in cartesian coordinates) the same (and well-known) thing, namely that (check)

$$\int_C (\partial_x f)dx + (\partial_y f)dy = f(B) - f(A)$$

The third and the fourth integral statements say the following (check):

$$\int_S (\partial_x A_y - \partial_y A_x)dxdy = \oint_{\partial S} A_x dx + A_y dy$$

$$\int_S (\partial_x A_x + \partial_y A_y)dxdy = \oint_{\partial S} A_x dy - A_y dx$$

Although these two statements perhaps look different at first sight, a bit closer look shows (just *rename* $(A_x, A_y) \mapsto (-A_y, A_x)$ in the first one and You get the second one) that it is actually a single statement again, namely the classical *Green theorem*

$$\int_S (\partial_x g - \partial_y f) dx dy = \oint_{\partial S} f dx + g dy$$

(see 8.1.4). So, there are altogether no more than *two independent* statements emerging from $\int_D d\alpha = \oint_{\partial D} \alpha$ in two-dimensional vector analysis.

8.5.6 p.182: see additional material to (8.1.3) here

8.5.8 p.183: The following immediate consequences from the results listed in (ii) might be useful:

$$(\boldsymbol{A}.d\boldsymbol{r})(\boldsymbol{B}) = \boldsymbol{A} \cdot \boldsymbol{B}$$
$$(\boldsymbol{A}.d\boldsymbol{S})(\boldsymbol{B}, \boldsymbol{C}) = \boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})$$
$$(h dV)(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) = h(\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C}))$$

10.2.6 p.212: The group of the upper-triangular matrices

$$A(a,b,c) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

mentioned here, with the composition law (check)

$$A(a,b,c)A(\tilde{a},\tilde{b},\tilde{c}) = A(a+\tilde{a}, b+\tilde{b}, c+\tilde{c}+a\tilde{b})$$

is also known as the *Heisenberg group.* Why? Let us compute its Lie algebra. Using the method of computation introduced in (11.7.5) we can immediately see that it is a vector space consisting of the matrices of the form

$$X(x,y,z) = \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = xE_1 + yE_2 + zE_3$$

where

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The commutation relations read

$$[E_1, E_2] = E_3 \qquad [E_1, E_3] = 0 \qquad [E_2, E_3] = 0$$

Now compare these relations with the celebrated Heisenberg ("canonical") commutation relations (CCR)

$$[\hat{p}, \hat{x}] = -i\hbar\hat{1} \qquad [\hat{p}, i\hbar\hat{1}] = 0 \qquad [\hat{x}, i\hbar\hat{1}] = 0$$

We can see that this is *the same* (abstract) Lie algebra.

In order to match the Heisenberg commutation relations for *more* canonical variables

$$[\hat{p}_i, \hat{x}_j] = -i\hbar\delta_{ij}\hat{1} \qquad [\hat{p}_j, i\hbar\hat{1}] = 0 \qquad [\hat{x}_j, i\hbar\hat{1}] = 0$$

we promote x, y to become n-dimensional *columns* (z remains to be a number). Namely, a Lie algebra element is, in general,

$$X(x,y,z) = \begin{pmatrix} 0 & x^T & z \\ 0 & 0_n & y \\ 0 & 0 & 0 \end{pmatrix} = x_i E_i + y_i F_i + zG$$

Then one checks that

$$[E_i, F_j] = \delta_{ij} G \qquad [E_i, G] = 0 \qquad [F_i, G] = 0$$

and so we are done.

In the coordinates introduced so far (for $n = 1$, to be specific), we have the following obvious *possibility* to relate the coordinates on the group and on the Lie algebra respectively:

$$A(a, b, c) = \mathbb{I} + X(x, y, z)$$

i.e.

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \qquad (*)$$

i.e.

$$\begin{aligned} a(x, y, z) &= x \\ b(x, y, z) &= y \\ c(x, y, z) &= z \end{aligned}$$

So, the coordinates on the Lie algebra can *directly* serve as coordinates on the group.

There is also *another standard* parametrization of the group. Note that for a general Lie algebra element we have

$$X = \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \qquad X^2 = \begin{pmatrix} 0 & 0 & xy \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad X^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so that the *exponential* of X is simply

$$e^X = \mathbb{I} + X + \frac{1}{2} X^2 = \begin{pmatrix} 1 & x & z + \frac{1}{2} xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \qquad (**)$$

which provides an *alternative parametrization* of upper-triangular matrices: We can, again, use (x, y, x) as coordinates on the *group* (so we can use coordinates on the Lie algebra as coordinates on the group). However, their relation to "original" (a, b, c)-type coordinates is more complicated, namely

$$\begin{aligned} a(x, y, z) &= x \\ b(x, y, z) &= y \\ c(x, y, z) &= z + \frac{1}{2} xz \end{aligned}$$

One can check that the composition law in the group within *these* particular ("exponential") coordinates (x, y, z) reads

$$(x, y, z) \circ (\tilde{x}, \tilde{y}, \tilde{z}) = (z + \tilde{x}, y + \tilde{y}, z + \tilde{z} + \frac{1}{2}(x\tilde{y} - y\tilde{x}))$$

11.2.2 p.223: We introduced the commutator $[X, Y]$ of two vectors from $\mathcal{G} \equiv T_e G$ by the formula

$$[X, Y] := [L_X, L_Y](e) \qquad (*)$$

Bilinearity and skew-symmetry is clear. Here, let us discuss in more detail how *Jacobi identity* emerges (i.e. why the commutator is well-defined by this formula :-)

So, let us have $X, Y, Z \in \mathcal{G}$. Construct corresponding left-invariant fields L_X, L_Y, L_Z. Since Jacobi indentity *does hold* for commutator of vector *fields* (in general, see 4.3.6), we have

$$[[L_X, L_Y], L_Z] + [[L_Z, L_X], L_Y] + [[L_Y, L_Z], L_X] = 0$$

Evaluating both sides in $e \in G$ we get

$$[[L_X, L_Y], L_Z](e) + [[L_Z, L_X], L_Y](e) + [[L_Y, L_Z], L_X](e) = 0 \qquad (**)$$

Now, $[L_X, L_Y]$ is a left-invariant field due to 8.3.7 (or 4.3.6 - see the additional material to the latter). Then, due to 11.1.4i), it is L_W for some $W \in \mathcal{G}$. From (*) we see, that $W = [X, Y]$, so $[L_X, L_Y] = L_{[X,Y]}$. Therefore we can rewrite (**) as

$$[L_{[X,Y]}, L_Z](e) + [L_{[Z,X]}, L_Y](e) + [L_{[Y,Z]}, L_X](e) = 0 \qquad (***)$$

Using once more (*) we are done:

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

$\boxed{11.4.4}$ p.229: It might be instructive to treat in parallel *both additive and multiplicative* groups of real numbers, i.e. the groups

$$\begin{aligned} & (\mathbb{R}, +) && : m(x, y) = x + y \\ \text{and} \quad & GL(1, \mathbb{R}) && : m(x, y) = xy \end{aligned}$$

The corresponding basis left-invariant vector fields are ∂_x and $x\partial_x$ respectively, so we get

group	left-invariant fields	one-parameter subgroups	exponential map
$(\mathbb{R}, +)$	$k\partial_x$	$x(t) = kt$	$k \mapsto x(1) = k$
$GL(1, \mathbb{R})$	$kx\partial_x$	$x(t) = e^{kt}$	$k \mapsto x(1) = e^k$

Now think of the *additive* group and its Lie algebra. Note that the Lie algebra structure just *adds* the multiplication by scalars and the (trivial) commutator to the structure of the group. Therefore, there is a *natural identification* (!) of the group manifold $\mathbb{R}$ with the Lie algebra manifold $\mathbb{R}$, namely $x \leftrightarrow x$. After the identification, the exponential map comes out to be just the "*identity*" map $k \mapsto k$ on $\mathbb{R}$.

For the multiplicative group, the Lie algebra manifold $\mathbb{R}$ is mapped (via the "ordinary" exponential map $k \mapsto e^k$, as mentioned in the assignment of the exercise) onto the *sub*group $GL_+(1, \mathbb{R})$ of the group $GL(1, \mathbb{R})$ (the connected component of the identity).

$\boxed{11.5.2}$ p.230: in (ii) we are to inspect that $L_{f'(X)} = f_* L_X$. It is computed straightforwardly. First note, that homomorphism property of f may be written as $f \circ L_g = L_{f(g)} \circ f$. Then

$$L_{f'(X)}(f(g)) = L_{f(g)*} f'(X) = L_{f(g)*} f_* X = (L_{f(g)} \circ f)_* X = (f \circ L_g)_* X = f_* L_X(g)$$

which is just detailed version of the statement $L_{f'(X)} = f_* L_X$.

$\boxed{11.7.20, 21}$ p.241: Let us discuss the method of computation of the canonical 1-form on matrix groups even more explicitly (the method described here is very convenient so one should not feel sorry for additional few minutes spent on full understanding of the matter).

So, first, consider Lie groups G and H and a smooth injective homomorphism

$$f : G \to H$$

Let

$$f' : \mathcal{G} \to \mathcal{H}$$

be the corresponding derived homomorphism of Lie algebras. If E_i is a basis in $\mathcal{G}$ and E_a is a basis in $\mathcal{H}$, then the matrix f_i^a of f' is given by

$$f'(E_i) = f_i^a E_a \qquad (1)$$

The fact that f is homomorphism

$$f(g_1 g_2) = f(g_1) f(g_2)$$

may be rewritten as a condition

$$f \circ L_g = \hat{L}_{f(g)} \circ f \qquad (2)$$

relating left translations L_g and $\hat{L}_h$ on G and H respectively.

Now recall (see 11.2.6) that the definition of the canonical 1-forms θ^G and θ^H (on G and H respectively) reads

$$\theta^G_g := L_{g*} : T_g G \to T_{e_G} G \equiv \mathcal{G}$$
$$\theta^H_h := L_{h*} : T_h H \to T_{e_H} H \equiv \mathcal{H}$$

(Here θ^G_g denotes the 1-form θ^G at point $g \in G$ and e_G is the unit element on G. The Lie algebra $\mathcal{G}$ of G is identified with the tangent space $T_{e_G} G$ at e_G.) Therefore, taking $(\dots)_*$ of both sides of (2) we get

$$f_* \circ \theta^G_g = \theta^H_{f(g)} \circ f_* \tag{3}$$

The canonical 1-forms may be decomposed (see 11.2.6) as follows:

$$\theta^G = e^i E_i \tag{4}$$
$$\theta^H = e^a E_a \tag{5}$$

Then (3) gives

$$f_* \circ (e^i_g E_i) = (e^a_{f(g)} E_a) \circ f_* \tag{6}$$

Notice that both side are mappings $T_g G \to \mathcal{H}$. On argument $v \in T_g G$ we obtain

$$(f_* \circ (e^i_g E_i))(v) := f_*(\langle e^i_g, v \rangle E_i) = \langle e^i_g, v \rangle f'(E_i) = (e^i_g f'(E_i))(v)$$
$$((e^a_{f(g)} E_a) \circ f_*)(v) := \langle e^a_{f(g)}, f_* v \rangle E_a = \langle (f^* e^a)_g, v \rangle E_a = ((f^* e^a)_g E_a)(v)$$

or

$$e^i f'(E_i) = f^*(e^a E_a) \tag{7}$$

This can be briefly written as

$$f^* \theta^H = f'(\theta^G) \qquad \text{where} \qquad f'(\theta^G) \equiv f'(e^i E_i) := e^i f'(E_i) \tag{8}$$

So, performing pull-back f^* of the canonical 1-form θ^H we get "almost" canonical 1-form θ^G, the difference lies in that the resulting 1-form takes values in $f'(\mathcal{G}) \subset \mathcal{H}$ rather than in $\mathcal{G}$ itself.

Now, let us specify H to be $GL(n, \mathbb{R})$. Then $f \equiv \rho$ becomes *representation* of G in $\mathbb{R}^n$

$$\rho : G \to GL(n, \mathbb{R})$$

Natural ("Weyl") basis in the Lie algebra $gl(n, \mathbb{R})$ is labeled by as many as *two* indices, so (1) becomes

$$\rho'(E_i) = \rho^b_{ai} E^a_b \tag{9}$$

and (5) takes the form

$$\theta^{GL(n,\mathbb{R})} = e^a_b E^b_a \equiv (x^{-1} dx)^a_b E^b_a$$

where we already used the explicit coordinate expression of the canonical 1-form on $GL(n, \mathbb{R})$ known from 11.7.19. In this particular case we get from (7)

$$(\rho^*(x^{-1} dx)^b_a) E^a_b = e^i \rho'(E_i) = (\rho^b_{ai} e^i) E^a_b \tag{10}$$

Taking into account that the representation $\rho : G \to GL(n, \mathbb{R})$ has coordinate expression

$$z \mapsto x(z) \qquad \text{i.e. in detail} \qquad z^\mu \mapsto x^b_a(z) \tag{11}$$

we get from (10)

$$(x^{-1})^b_c(z) dx^c_a(z) = \rho^b_{ai} e^i \tag{12}$$

Finally, recalling that

$$\rho^b_{ai} = (\mathcal{E}_i)^b_a \tag{13}$$

are nothing but $(\dots)^b_a$-elements of *generators*

$$\mathcal{E}_i := \rho'(E_i) \tag{14}$$

of the representation ρ (see 12.1.4), we can rewrite (12) as

$$\boxed{x^{-1}(z)dx(z) = e^i\mathcal{E}_i} \tag{15}$$

This is the final formula enabling one to compute basis left-invariant 1-forms e^i (and, consequently, the canonical 1-form $\theta^G = e^i E_i$ on matrix Lie group G). One has to perform the following steps

- write down the formula (11) (i.e. parametrize the matrix group G)
- compute the expression $x^{-1}(z)dx(z)$
- identify coefficients (1-forms) e^i standing by the generators $\mathcal{E}_i$

Normally, one uses the "identity" representation $\rho = \mathrm{id}$, $\rho(A) = A$, of the matrix group G. See explicit examples in 11.7.22 and 11.7.23.

12.1.13 p.251: More general and elegant result may be obtained virtually with the same effort: If we start with a bilinear form (not necessarily symmetric!) given by a *general* real matrix, the procedure $(1/2\pi)\int_0^{2\pi} d\alpha$ (averaging over $SO(2)$) leads to (check)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{a+d}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{b-c}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

This nicely shows that the only invariant bilinear forms are (any multiple of) the ordinary scalar product and (any multiple of) the ordinary *volume* form. You can also check that repeated averaging already does nothing (the map behaves as a projector)

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

See also 12.2.8.

12.2.5 p.254: Here, it might be not clear what precisely N_j mean (although a reader may deduce it from description of $so(r,s)$ in 11.7.6) so I add a more detailed account of $so(1,3)$, which is important in its own right.

Consider matrices C from the Lie algebra $so(1,3)$ (the Lie algebra of the *Lorentz* group), i.e. (see 11.7.6) real 4×4 matrices, satisfying

$$(\eta C)^T + \eta C = 0 \qquad \eta \equiv \begin{pmatrix} 1 & 0^T \\ 0 & -\mathbb{I} \end{pmatrix}$$

(0 is the 3-dimensional null column and $\mathbb{I}$ is the 3×3 unit matrix). Check that
i) in the 1+3 block form (similar to the form of the matrix η) we can parameterize the matrices C as follows

$$C(u,X) = \begin{pmatrix} 0 & u^T \\ u & X \end{pmatrix} \qquad X^T = -X$$

(u is a 3-dimensional real column and X is a 3×3 real antisymmetric matrix, i.e. $X \in so(3)$)
ii) the natural decomposition of the resulting most general matrix

$$\begin{pmatrix} 0 & u^T \\ u & X \end{pmatrix} = \begin{pmatrix} 0 & 0^T \\ 0 & X \end{pmatrix} + \begin{pmatrix} 0 & u^T \\ u & 0 \end{pmatrix}$$

i.e.

$$C(u,X) = C(0,X) + C(u,0)$$

may be regarded as a decomposition of the total 6-dimensional linear space $so(1,3)$ into the direct sum of two 3-dimensional subspaces

$$so(1,3) = so(1,3)_{\text{rot}} + so(1,3)_{\text{boo}}$$

and the elements of the corresponding subspaces generate *rotations* and *boosts* (= „genuine" Lorentz transformations). So, if $x = x^\mu = t, x^1, x^2, x^3$ is a point in the Minkowski space, then the rules

$$x \mapsto (\hat{1} + \epsilon C(0,X))x \qquad \text{and} \qquad x \mapsto (\hat{1} + \epsilon C(u,0))x$$

describe infinitesimal rotations and infinitesimal boosts respectively
iii) the commutation relations in the Lie algebra read

$$[C(u_1, X_1), C(u_2, X_2)] = C(X_1 u_2 - X_2 u_1, [X_1, X_2] + u_1 u_2^T - u_2 u_1^T)$$

and, in particular

$$\begin{aligned} [C(0,X), C(0,Y)] &= C(0, [X,Y]) \\ [C(u,0), C(v,0)] &= C(0, uv^T - vu^T) \\ [C(0,X), C(u,0)] &= C(Xu, 0) \end{aligned}$$

From this we see that $so(1,3)_{\rm rot}$ is a (Lie) *subalgebra* (isomorphic to $so(3)$), but $so(1,3)_{\rm boo}$ *fails* to be a subalgebra (being just a subspace)

$$\begin{aligned} [so(1,3)_{\rm rot}, so(1,3)_{\rm rot}] &\subset so(1,3)_{\rm rot} \\ [so(1,3)_{\rm boo}, so(1,3)_{\rm boo}] &\subset so(1,3)_{\rm rot} \\ [so(1,3)_{\rm rot}, so(1,3)_{\rm boo}] &\subset so(1,3)_{\rm boo} \end{aligned}$$

iv) if we choose natural bases in the spaces of X's and u's (for X's the matrices l_j from exercise 11.7.13 and for u's the columns e_j with a single 1 and all others 0, formally $(e_j)_k := \delta_{jk}$), we get as a basis of the Lie algebra $so(1,3)$ the matrices

$$s_j = C(0, l_j) = \begin{pmatrix} 0 & 0^T \\ 0 & l_j \end{pmatrix} \qquad N_j = C(e_j, 0) = \begin{pmatrix} 0 & e_j^T \\ e_j & 0 \end{pmatrix} \qquad j = 1,2,3$$

Their commutation relations read

$$\begin{aligned} [s_i, s_j] &= \epsilon_{ijk} s_k \\ [N_i, N_j] &= -\epsilon_{ijk} s_k \\ [s_i, N_j] &= \epsilon_{ijk} N_k \end{aligned}$$

v) to find a representation of the Lie algebra $so(1,3)$, then, means to find any operators (matrices) $\hat{s}_i, \hat{N}_i$ whose commutation relations exactly copy those for s_i, N_i from *iv*)

Now, I hope, the problem 12.2.5 should be crystal clear (and easy).

12.3.16, 13.1.10 pp.267,292: More on *conjugation* (and conjugation *classes*, in particular for the group $SO(3)$).

Let G act on a set M. Let an element g perform $x \mapsto y \equiv gx$. What kind of transformation does then the element hgh^{-1}? Since

$$(hgh^{-1})(hx) = h(gx) = hy$$

we can deduce the following (simple albeit important) piece of wisdom:

if action of	g	performs	$x \mapsto y$
then action of	hgh^{-1}	performs	$hx \mapsto hy$

and read it in a more wordy way as follows: *In general*, the h-conjugated element hgh^{-1} does "the same job" on h-transformed set M as the original element g does on the original set M.

Now, let us apply it to the group $SO(3)$ acting in the standard way (via rotations) in the 3-dimensional Euclidean space:

if action of	A	performs rotation	$\mathbf{x} \mapsto \mathbf{y}$
then action of	BAB^{-1}	performs rotation	$B\mathbf{x} \mapsto B\mathbf{y}$

Let A perform the rotation through angle α about $\mathbf{n}$. So $A = e^{\alpha \mathbf{n}.\mathbf{l}}$. If $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{n})$ is an orthonormal right-handed bases, then

$$A \text{ performs} \qquad \begin{aligned} \mathbf{e}_1 &\mapsto \mathbf{e}_1 \cos\alpha + \mathbf{e}_2 \sin\alpha \\ \mathbf{e}_2 &\mapsto -\mathbf{e}_1 \sin\alpha + \mathbf{e}_2 \cos\alpha \\ \mathbf{n} &\mapsto \mathbf{n} \end{aligned}$$

Then, according to the general wisdom,

$$BAB^{-1} \text{ performs} \qquad \begin{aligned} B\mathbf{e}_1 &\mapsto B\mathbf{e}_1 \cos\alpha + B\mathbf{e}_2 \sin\alpha \\ B\mathbf{e}_2 &\mapsto -B\mathbf{e}_1 \sin\alpha + B\mathbf{e}_2 \cos\alpha \\ B\mathbf{n} &\mapsto B\mathbf{n} \end{aligned}$$

Since $(B\mathbf{e}_1, B\mathbf{e}_2, B\mathbf{n})$ is an orthonormal right-handed bases, too, this means that BAB^{-1} performs the rotation through *the same* angle α about *the transformed* vector $B\mathbf{n}$. In formulas:

$$\boxed{Be^{\alpha \mathbf{n}.\mathbf{l}}B^{-1} = e^{\alpha (B\mathbf{n}).\mathbf{l}}} \qquad \text{or} \qquad \boxed{B(\mathbf{n}.\mathbf{l})B^{-1} = (B\mathbf{n}).\mathbf{l}}$$

We see, that the *conjugacy class* of an element $A \equiv e^{\alpha \mathbf{n}.\mathbf{l}}$ consists of *all* rotations through *the same angle* α about *arbitrary* vectors $\mathbf{n}$.

12.4.6 p.272, an alternative (and instructive) hint: Modify the solution of 12.2.5 *presented here* (see above). Actually, one can treat both cases, $so(1,3)$ as well as $so(4)$, *in parallel* starting from the following more general condition

$$(\eta C)^T + \eta C = 0 \qquad \eta \equiv \begin{pmatrix} 1 & 0^T \\ 0 & -\lambda \mathbb{I} \end{pmatrix}$$

where

$$\begin{aligned} \lambda &= 1 && \text{for } so(1,3) \\ \lambda &= -1 && \text{for } so(4) \end{aligned}$$

Then, modifying the calculations to the λ-case we obtain

$$\begin{aligned} [s_i, s_j] &= \epsilon_{ijk} s_k \\ [N_i, N_j] &= -\lambda \epsilon_{ijk} s_k \\ [s_i, N_j] &= \epsilon_{ijk} N_k \end{aligned}$$

In order to split the Lie algebra to the direct sum of two Lie subalgebras (spanned by A_i and B_i respectively) we introduce)

$$\begin{aligned} A_i &:= s_i + aN_i \\ B_i &:= s_i + bN_i \end{aligned}$$

Then

$$[A_i, B_j] = \epsilon_{ijk} \ ((1 - \lambda ab) \ s_k + (a + b) \ N_k))$$

If this is to vanish (as it should for the direct sum), one has to choose

$$a + b = 0 \qquad 1 - \lambda ab = 0$$

or, equivalently

$$\begin{aligned} A_i &:= s_i + aN_i \\ B_i &:= s_i - aN_i \qquad \lambda = -a^2 \end{aligned}$$

Then the complete set of commutation relations becomes

$$\begin{aligned} [A_i, A_j] &= 2 \ \epsilon_{ijk} \ A_k \\ [B_i, B_j] &= 2 \ \epsilon_{ijk} \ B_k \\ [A_i, B_j] &= 0 \end{aligned}$$

so that $(s,N)\mapsto(A,B)$ provides a desired direct sum decomposition. This only occurs, however, when

$$\lambda = -a^2$$

is satisfied.

In the case of $so(4)$, i.e. when $\lambda=-1$, this is easily fulfilled by $a=1$

$$A_j := s_j + N_j$$
$$B_j := s_j - N_j$$

and we get the decomposition needed in 12.4.6 .

In the case of $so(1,3)$, however, when $\lambda=+1$, we badly need $a=i$ (or $a=-i$)

$$A_j := s_j + iN_j$$
$$B_j := s_j - iN_j \qquad\qquad i\equiv\sqrt{-1}$$

and this is *not acceptable* as long as we understand $so(1,3)$ to denote the (usual) *real* Lie algebra (where, of course, complex linear combinations are not allowed). So, $so(1,3)$ (regarded as a real Lie algebra) *cannot* be (contrary to $so(4)$) decomposed in this way.

[The needed complex combinations *can* be used if we think of $so(1,3)$ as a *complex* Lie algebra (the complexification of the real $so(1,3)$). (This is also standardly done in physics, when 2-component relativistic ("dotted" and "undotted") spinors are introduced. The *matrices*, representing s_i and N_i act in *complex* spaces and, consequently, their *complex* linear combinations are as good as real ones.)]

12.5.4 p.281: Maybe a bit more systematic formulation: η^{ab} and η_{ab} are invariant tensors for $O(r,s)$, $\epsilon^{a\dots b}$ and $\epsilon_{a\dots b}$ are invariant tensors for $SL(n)$. Then, consequently, *all of them* are invariant tensors for $SO(r,s)$.

12.6.2 p.283: The complex constructed here is often called the *Chevalley-Eilenberg complex* (and corresponding Lie algebra cohomologies then become *Chevalley-Eilenberg cohomologies*)

13.1.10 p.292: see 12.3.16, p.267 (in this text)

13.2.7 p.296: In the hint to (i) the map f given as

$$\pi(g)\mapsto\pi'(g) \qquad \text{or, in another notation} \qquad [g]\mapsto[g]'$$

is, unfortunately, not well defined. Indeed, take two elements of gH, say g and gh. They both, by definition of π, project to $\pi(g)$. So they both can play the role of "g" in the formula (the role of a representative). However, $\pi'(g)\neq\pi'(gh)$, since $h\notin H'\equiv kHk^{-1}$ in general. So we have to invent another map to prove the statement :-(

A way to find it might be based on using more general result mentioned in ii): *each* homogeneous space $(M,\hat L_g)$ is isomorphic to some canonical space $(G/H,L_g)$ (with canonical action $L_g[\tilde g]\equiv g\cdot\tilde g := [g\tilde g]$).

▼ Indeed, choose $m\in M$ an denote H_m its stabilizer ($\hat L_h m = m, h\in H_m$). Arbitrary point in M is $g\cdot m$. (In addition to g, the same job does any gh, $h\in H_m$, see additional material to 13.2.7). Define

$$\varphi_m : M\to G/H_m \qquad\qquad g\cdot m\mapsto[g]\leftrightarrow gH_m$$

(Because of $(gh)\cdot m\mapsto[gh]=[g]$, φ_m *is* well-defined.) Then it is clearly bijective and it also fulfills

$$L_g\circ\varphi_m = \varphi_m\circ\hat L_g$$

since

$$(L_g\circ\varphi_m)(\tilde g\cdot m) = L_g[\tilde g] = [g\tilde g] = \varphi_m((g\tilde g)\cdot m)) = \varphi_m(g\cdot(\tilde g\cdot m)) = (\varphi_m\circ\hat L_g)(\tilde g\cdot m)$$

So, φ_m provides *isomorphism* of an *arbitrary* homogeneous space $(M,\hat L_g)$ to the *canonical* homogeneous space $(G/H,L_g)$ with $H:=H_m$. ▲

Notice that the isomorphism φ_m depends on the choice of $m\in M$. If

$$m'\equiv k\cdot m$$

is another point in M, we get, in the same way, *another* isomorphism

$$\varphi_{m'} : M \to G/H_{m'} \qquad\qquad g \cdot m' \mapsto [g]' \leftrightarrow gH_{m'}$$

Now it is clear, that we can obtain an isomorphism of two *canonical* homogeneous spaces simply by *composition* of φ_m^{-1} and $\varphi_{m'}$

$$\psi : G/H_m \to G/H_{m'} \qquad\qquad \psi := \varphi_{m'} \circ \varphi_m^{-1}$$

or

$$\begin{array}{ccc} M & \xrightarrow{\text{identity}} & M \\ \varphi_m \big\downarrow & & \big\downarrow \varphi_{m'} \\ G/H_m & \xrightarrow[\psi]{} & G/H_{m'} \end{array}$$

Let us compute explicit formula. We have

$$\begin{aligned} \varphi_m &: g \cdot m \mapsto gH_m \\ \varphi_{m'} &: g \cdot m' \mapsto gH_{m'} \qquad\qquad m' \equiv k \cdot m \\ \varphi_m^{-1} &: gH_m \mapsto g \cdot m \end{aligned}$$

(we use notation gH_m instead of $[g]$ for the *point* of G/H_m in order to make computation more transparent). Then

$$\psi \equiv \varphi_{m'} \circ \varphi_m^{-1} : gH_m \mapsto g \cdot m \equiv g \cdot (k^{-1} \cdot m') \equiv (gk^{-1}) \cdot m' \mapsto (gk^{-1})H_{m'}$$

so that

$$\psi(gH_m) = (gk^{-1})H_{m'}$$

or, in square bracket notation,

$$\psi[g] = [gk^{-1}]' \qquad\qquad \text{(rather than } \psi[g] = [g]' \text{ from the hint in the book!)}$$

It is clearly a bijection from G/H_m to $G/H_{m'}$. Let us check it is also equivariant:

$$\tilde{g} \cdot (\psi(gH_m)) = \tilde{g} \cdot ((gk^{-1})H_{m'})) = (\tilde{g}gk^{-1})H_{m'} = ((\tilde{g}g)k^{-1})H_{m'} = \psi((\tilde{g}g)H_m) = \psi(\tilde{g} \cdot (gH_m))$$

so that

$$L'_{\tilde{g}} \circ \psi = \psi \circ L_{\tilde{g}}$$

▼ Let us also check *directly* that the "k^{-1}-correction" to the original (wrong) formula makes the mapping already *well-defined* (see the beginning of this additional text): Indeed, if we use representative gh instead of g, we get

$$[gh] \mapsto [(gh)k^{-1}]' = [ghk^{-1}]' = [(gk^{-1})(khk^{-1})]' = [(gk^{-1})h']' = [gk^{-1}]'$$

regardless of $h \in H$ (since $khk^{-1} = h' \in H' \equiv kHk^{-1}$). ▲

The map ψ may be expressed even more simply. Notice that ψ is fully specified by ascribing its value to a *single* point (the value on any other point is already fixed by equivariancy). Then, taking this point to be $H_m \leftrightarrow [e]$, we get

$$\psi(H_m) = k^{-1}H_{m'} \qquad \text{or, equivalently} \qquad \psi[e] = [k^{-1}]'$$

We can also proceed in the following alternative way: Let H, H' be *any* two subgroups of G and let

$$\psi : G/H \to G/H'$$

be an *isomorphism.* Because of *equivariancy* of ψ, it is enough to tell the result of $\psi[e]$. Since the result is in G/H', it has to have the form

$$\psi[e] = [\hat{g}]'$$

for some (yet unknown) $\hat{g} \in G$. Now, the map ψ should not depend on representatives (should be *well-defined*). This gives

$$\psi[e] = \psi[h] = \psi(h \cdot [e]) = h \cdot (\psi[e]) = h \cdot [\hat{g}]' = [h\hat{g}]' \stackrel{!}{=} [\hat{g}]'$$

for any $h \in H$. This gives the following restriction on $\hat{g}$: for each $h \in H$,

$$h\hat{g} = \hat{g}h' \qquad h' \in H'$$

or, equivalently,

$$\hat{g}^{-1}h\hat{g} = h' \qquad h' \in H'$$

This means, that *conjugation* of H by $\hat{g}^{-1}$ (which gives clearly a subgroup of G *isomorphic* to H itself) is to be a *subgroup* of H'. Since we know, however, that $\#H = \#H'$ (otherwise $\psi : G/H \to G/H'$ cannot be *bijection*),[3] we get *equality*

$$\hat{g}^{-1}H\hat{g} = H'$$

This may be reformulated as follows: $\psi : G/H \to G/H'$ is an isomorphism (if and) only if H' is a *conjugated* subgroup w.r.t. H, i.e. $H' = kHk^{-1}$ for some $k \in G$, and then

$$\psi[e] = [k^{-1}]'$$

13.2.7 p.296: Here we learned that any homogeneous space $(M, \hat{L}_g)$ is isomorphic to the canonical homogeneous space $(G/G_x, L_g)$. The following (simple) useful fact remained, however, somehow hidden (it is not mentioned explicitly):

Let $x \in \mathcal{O}_x$ and let $y \equiv gx \in \mathcal{O}_x$ be another point of the orbit $\mathcal{O}_x$. As we see, starting from x, we can reach y by the action of g. The question then arises: Are there also other group elements which do the same job (perform $x \mapsto y$)? The answer says that the set of all those elements is just the coset gG_x. (This gives a useful interpretation of the coset gG_x.)

[Indeed, it is clear that elements from gG_x do shift x to gx, since G_x "does nothing" and g shifts x to gx. Now,let k be another group element with the same property, so that $kx = gx$. Then $g^{-1}k \in G_x$ and, consequently, $k \in gG_x$.]

It might seem that we can enlarge the set gG_x to $G_y gG_x$ (since G_y at the end "does nothing" when acting on y in the same way as G_x "does nothing" when acting on x). However, we learned in 13.1.10 that $G_y = gG_xg^{-1}$ and therefore

$$G_y gG_x = (gG_xg^{-1})gG_x = gG_xG_x = gG_x$$

13.2.9 p.297: we learned in this problem that the formula

$$z \mapsto \frac{az+b}{cz+d} \equiv L_A z \qquad\qquad A \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2,\mathbb{C})$$

(the *Möbius* or *linear-fractional* transformation) defines a left action of $GL(2,\mathbb{C})$ in the (extended) complex plane and that A and λA result in the same action (so that "only $SL(2,\mathbb{C})$ matters"). Here we present a useful interpretation of the formula in more geometric terms. The new point of view shows clearly *why* the strange looking formula works (why it provides an action). It also leads straightforwardly to a more dimensional generalization.

Consider the complex vector space $\mathbb{C}^n$ (the needed modification for the real case will be evident). There is a natural action of $GL(n,\mathbb{C})$ on $\mathbb{C}^n$ given by

$$\hat{L}_A u := Au$$

We know from 1.3.2(ii) that the *rays* in $\mathbb{C}^n$ (complex lines in $\mathbb{C}^n$ passing through the origin) form a new set (manifold) called the complex *projective space* $\mathbb{C}P^{n-1}$. Now it is intuitively clear that the action on vectors passes to the action on rays (by means of the representatives). Namely, if $[u]$ denotes the ray given by the vector u, we have

$$\boxed{L_A[u] := [\hat{L}_A u] \equiv [Au]}$$

Indeed, the formal check reads

$$L_{AB}[u] = [(AB)u] = [A(Bu)] = L_A[Bu] = (L_A L_B)[u]$$

[3] Or, using *symmetry* argument, by the same computation with H and H' *interchanged* we get that by conjugation of H' we get a *subgroup* of H :-)

Note that

$$L_{\lambda A}[u] = [(\lambda A)u] = [\lambda(Au)] = [Au] = L_A[u]$$

so that

$$L_{\lambda A} = L_A$$

So, the lesson is that

i) $GL(n, \mathbb{C})$ (as well as any subgroup) not only naturally acts on the vector space $\mathbb{C}^n$ (the action being *linear*, i.e. a representation), but also on a more complicated manifold, the *projective* space $\mathbb{C}P^{n-1}$ (here, of course, the action is *not* linear)

ii) in this action, A produces the same result as λA does

Let us compute explicitly the coordinate expression of the action in the simplest case, on $\mathbb{C}P^1$. The original action of $GL(2, \mathbb{C})$ on $\mathbb{C}^2$ reads

$$\begin{pmatrix} z^1 \\ z^2 \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z^1 \\ z^2 \end{pmatrix}$$

Introduce the standard two charts on $\mathbb{C}P^1$ given by (complex) coordinates ξ^1 or η^1 (see the hint to 1.3.2)

$$\begin{pmatrix} z^1 \\ z^2 \end{pmatrix} \sim \begin{pmatrix} 1 \\ \xi^1 \end{pmatrix} \sim \begin{pmatrix} \eta^1 \\ 1 \end{pmatrix} \qquad \text{where} \quad \eta^1 = 1/\xi^1$$

We get

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ \xi^1 \end{pmatrix} = \begin{pmatrix} a + b\xi^1 \\ c + d\xi^1 \end{pmatrix} = (a + b\xi^1) \begin{pmatrix} 1 \\ \hat{\xi}^1 \end{pmatrix} \sim \begin{pmatrix} 1 \\ \hat{\xi}^1 \end{pmatrix} \qquad \hat{\xi}^1 = \frac{c + d\xi^1}{a + b\xi^1}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \eta^1 \\ 1 \end{pmatrix} = \begin{pmatrix} a\eta^1 + b \\ c\eta^1 + d \end{pmatrix} = (a\eta^1 + b) \begin{pmatrix} \hat{\eta}^1 \\ 1 \end{pmatrix} \sim \begin{pmatrix} \hat{\eta}^1 \\ 1 \end{pmatrix} \qquad \hat{\eta}^1 = \frac{a\eta^1 + b}{c\eta^1 + d}$$

So, effectively,

$$\xi^1 \mapsto \hat{\xi}^1 = \frac{c + d\xi^1}{a + b\xi^1}$$

$$\eta^1 \mapsto \hat{\eta}^1 = \frac{a\eta^1 + b}{c\eta^1 + d}$$

We see that the expression of the transformation in terms of the variable η^1 is given by the Möbius (linear-fractional) transformation. Concerning ξ^1, we obtained a modification of the Möbius (linear-fractional) transformation and one can easily check that it (i.e. the composition rule) really works.

It is also instructive to check, that both formulas describe *the same* abstract transformation

$$L_A : \mathbb{C}P^1 \to \mathbb{C}P^1 \qquad \text{(or, equivalently, } S^2 \to S^2\text{)}$$

We have to check that

$$\eta^1 = 1/\xi^1 \qquad \text{implies} \qquad \hat{\eta}^1 = 1/\hat{\xi}^1$$

which is straightforward:

$$\hat{\eta}^1 = \frac{a\eta^1 + b}{c\eta^1 + d} = \frac{a + b/\eta^1}{c + d/\eta^1} = \frac{a + b\xi^1}{c + d\xi^1} = \frac{1}{\hat{\xi}^1}$$

The reader is invited to show that the explicit formula for the transformation of $\mathbb{C}P^{n-1}$ in terms of the coordinates $(\eta^1, \dots, \eta^{n-1})$ reads

$$\hat{\eta}^a = \frac{A^a_b \eta^b + A^a_n}{A^n_b \eta^b + A^n_n}$$

[Note: we obtained an action of $SL(2, \mathbb{C})$ on S^2. One can check that the action of the *unitary* subgroup, $SU(2)$, gives nothing but *rotations* of the (standard round) sphere. This can be ultimately traced back to the fact that unitary group preserves the natural "hermitian" scalar product in $\mathbb{C}^n$ and, consequently, the induced so called *Fubini-Study* scalar product on $\mathbb{C}P^n$ (not mentioned in the book)]

13.3.1 p.300: in (ii) a method is described in which a covering group steals a representation from the covered group (by composition with the covering homomorphism f, i.e. $\rho = \tilde{\rho} \circ f$). It should be clear that *any* homomorphism can be used for this purpose (if f is *a* homomorphism $G \to \tilde{G}$ and $\tilde{\rho}$ is a representation of $\tilde{G}$, then $\tilde{\rho} \circ f$ is a representation of G; if $\tilde{\rho}$ is irreducible, the same is true for ρ).

$\boxed{13.4.7}$ p.315 In the hint to 13.4.7iv) a reference to 4.6.10 might be useful

$\boxed{13.4.7 - 10}$ p.315-16 Prove that the trace and the determinant are the only (independent) invariants of a symmetric 2×2 matrix A with respect to $A \mapsto B^{-1}AB$, B special orthogonal.

Solution: The problem is from linear algebra. Nevertheless, a useful idea might be to treat the problem infinitesimally, introduce the *fundamental field* of the action and find (in this way!) the most general invariant function.

For infinitesimal $B = \mathbb{I} + \epsilon C$, where $C^T = -C$, we have

$$A \mapsto B^{-1}AB = (\mathbb{I} - \epsilon C)A(\mathbb{I} + \epsilon C) = A + \epsilon(AC - CA) + \dots$$

Then,

$$\begin{aligned} A \equiv \begin{pmatrix} u & v \\ v & w \end{pmatrix} &\mapsto \begin{pmatrix} u & v \\ v & w \end{pmatrix} + \epsilon \left(\begin{pmatrix} u & v \\ v & w \end{pmatrix} \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix} - \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix} \begin{pmatrix} u & v \\ v & w \end{pmatrix} \right) \\ &= \begin{pmatrix} u & v \\ v & w \end{pmatrix} + \epsilon k \begin{pmatrix} -2v & u-w \\ u-w & 2v \end{pmatrix} \end{aligned}$$

so that

$$(u, v, w) \mapsto (u(\epsilon), v(\epsilon), w(\epsilon)) = (u + k\epsilon(-2v), v + k\epsilon(u-w), w + k\epsilon(2v))$$

The *generator* (fundamental field) ξ of the action reads

$$\begin{aligned} \xi &= -2v\partial_u + (u-w)\partial_v + 2v\partial_w \\ &= 2v(\partial_w - \partial_u) + (u-w)\partial_v \end{aligned}$$

We are to find the most general function $\Phi(u, v, w)$, which is *invariant* with respect to the flow of ξ, i.e. such that

$$\xi\Phi \overset{!}{=} 0$$

This greatly simplifies in appropriate coordinates. We see from the structure of ξ that variables $u - w, v$ play an important role. So try a change of coordinates $(u, v, w) \to (x, y, z)$ given by

$$\begin{aligned} \lambda(u-w) &= x \\ v &= y \\ \lambda(u+w) &= z \end{aligned}$$

Then we get

$$\xi = (1/\lambda)x\partial_y - (4\lambda)y\partial_x$$

or, for $\lambda = 1/2$,

$$\xi = 2(x\partial_y - y\partial_x)$$

This field clearly generates rotations in xyz-space about z-axis. So, introducing further standard *cylindrical* coordinates r, φ, z according to

$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \\ z &= z \end{aligned}$$

we come to the most simple coordinate form of the field ξ:

$$\xi = 2\partial_\varphi$$

Then, the invariance condition reads

$$\xi\Phi \overset{!}{=} 0 \qquad \Leftrightarrow \qquad \frac{\partial\Phi(r, \varphi, z)}{\partial\varphi} \overset{!}{=} 0$$

and the obvious general solution is

$$\Phi(r,\varphi,z) = f(r,z)$$

But note, that

$$\operatorname{Tr} A \equiv T(u,v,w) = u + w = 2z \equiv \hat{T}(r,z)$$
$$\det A \equiv D(u,v,w) = uw - v^2 = (z+x)(z-x) - y^2 = z^2 - (x^2+y^2) = z^2 - r^2 \equiv \hat{D}(r,z)$$

so that $\hat{T}, \hat{D}$ are just two independent combinations of r, z.
[Since $d\hat{T} \wedge d\hat{D} = \cdots = -4r dz \wedge dr$, the combinations $\hat{T}, \hat{D}$ *are independent*. One can easily express *explicitly* r, z in terms of $\hat{T}, \hat{D}$.]

This means, that *any invariant* function of matrix elements of A is of the form $f(r,z)$ or, equivalently, of just *two* particular combinations of matrix elements of A, namely $\operatorname{Tr} A$ and $\det A$. Q.E.D.

14.2.4 p.339 In this exercise, we study canonical transformations as such coordinate tranformations,

$$(q^a, p_a) \mapsto (Q^a(q,p), P_a(q,p))$$

which preserve canonical (Darboux) expression of the *symplectic form*, i.e. such that

$$\boxed{\omega = dp_a \wedge dq^a \overset{!}{=} dP_a \wedge dQ^a \equiv dP_a(q,p) \wedge dQ^a(q,p)}$$

holds. This is, however, equivalent to several alternative ways of expressing the same fact. First, the (canonical) form of the corresponding *Poisson tensor* is preserved

$$\boxed{\mathcal{P} = \frac{\partial}{\partial p_a} \wedge \frac{\partial}{\partial q^a} = \frac{\partial}{\partial P_a} \wedge \frac{\partial}{\partial Q^a}}$$

(since $\mathcal{P} \circ \omega = -\hat{1}$), resulting then in equally simple (canonical) explicit formulas of Poisson bracket in old and new coordinates

$$\boxed{\{f,g\} = \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q^a} - \frac{\partial f}{\partial q^a}\frac{\partial g}{\partial p_a} = \frac{\partial f}{\partial P_a}\frac{\partial g}{\partial Q^a} - \frac{\partial f}{\partial Q^a}\frac{\partial g}{\partial P_a}}$$

This is, in turn, equivalent to the statement, that the *component matrix* $\mathcal{P}^{ij}$ has, in both coordinates, (q^a, p_a) and (Q^a, P_a) the canonical form

$$\mathcal{P}^{ij}(q,p) = \begin{pmatrix} 0_n & -\mathbb{I}_n \\ \mathbb{I}_n & 0_n \end{pmatrix} = \mathcal{P}^{ij}(Q,P)$$

Now, because

- components are just the values of a tensor on basis arguments, i.e. $\mathcal{P}^{ij}(z) = \mathcal{P}(dz^i, dz^j)$
- the Poisson bracket is defined as $\{f,g\} = \mathcal{P}(df,dg)$, so that $\mathcal{P}^{ij}(z) = \{z^i, z^j\}$

we see that we can rewrite the equation above in the form

$$\mathcal{P}^{ij}(q,p) = \begin{pmatrix} \{q^a,q^b\} & \{q^a,p_b\} \\ \{p_a,q^b\} & \{p_a,p_b\} \end{pmatrix} = \begin{pmatrix} 0 & -\delta^a_b \\ \delta^b_a & 0 \end{pmatrix} = \begin{pmatrix} \{Q^a,Q^b\} & \{Q^a,P_b\} \\ \{P_a,Q^b\} & \{P_a,P_b\} \end{pmatrix} = \mathcal{P}^{ij}(Q,P)$$

And this, at last, can be rephrased as the statement, that canonical transformations are such coordinate tranformations,

$$(q^a, p_a) \mapsto (Q^a(q,p), P_a(q,p))$$

which preserve *canonical Poisson brackets* between the coordinate functions, i.e. for which

$$\boxed{\{q^a,q^b\} = \{Q^a,Q^b\} = 0}$$

$$\boxed{\{p_a,q^b\} = \{P_a,Q^b\} = \delta^b_a}$$

$$\boxed{\{p_a,p_b\} = \{P_a,P_b\} = 0}$$

In quantum mechanics, after replacing of the Poisson bracket (of functions on phase space) by a constant multiple of the *commutator* (of operators in the corresponding Hilbert space), canonical transformations of operators are such transformations,

$$(\hat{q}^a, \hat{p}_a) \mapsto (\hat{Q}^a(\hat{q}, \hat{p}), \hat{P}_a(\hat{q}, \hat{p}))$$

which preserve the *canonical commutation relations* (CCR)

$$[\hat{q}^a, \hat{q}^b] = [\hat{Q}^a, \hat{Q}^b] = 0$$

$$[\hat{p}_a, \hat{q}^b] = [\hat{P}_a, \hat{Q}^b] = -i\hbar\delta^b_a$$

$$[\hat{p}_a, \hat{p}_b] = [\hat{P}_a, \hat{P}_b] = 0$$

Or, finally, when expressed in terms of annihilation and creation operators

$$a_a := \frac{1}{\sqrt{2\hbar}} (\hat{q}^a + i\hat{p}_a) \qquad a^+_a := \frac{1}{\sqrt{2\hbar}} (\hat{q}^a - i\hat{p}_a)$$

$$A_a := \frac{1}{\sqrt{2\hbar}} \left(\hat{Q}^a + i\hat{P}_a\right) \qquad A^+_a := \frac{1}{\sqrt{2\hbar}} \left(\hat{Q}^a - i\hat{P}_a\right)$$

as such transformation of operators

$$(a_a, a^+_a) \mapsto (A_a(a, a^+), A^+_a(a, a^+))$$

which preserve the CCR in the form

$$\boxed{[a_a, a_b] = [A_a, A_b] = 0}$$

$$\boxed{[a_a, a^+_b] = [A_a, A^+_b] = \delta_{ab}}$$

$$\boxed{[a^+_a, a^+_b] = [A^+_a, A^+_b] = 0}$$

14.3.4 p.342 The concept of *relative invariance* of a form was introduced in 14.3.2 and 14.3.3, but I somehow forgot to mention the corresponding version of integral invariants :-(. Here the gap is filled.

Consider an invariant p-form α and a relative invariant p-form β. This means

$$\mathcal{L}_V \alpha = 0$$
$$\mathcal{L}_V \beta = d\rho$$

for some $(p-1)$-form ρ. Let $\Phi_t \leftrightarrow V$ be the corresponding flow. Then, because of

$$\Phi^*_t = \hat{1} + t\mathcal{L}_V + (t^2/2)\mathcal{L}_V\mathcal{L}_V + \dots$$

we have

$$\Phi^*_t \alpha = \alpha$$
$$\Phi^*_t \beta = \beta + d\phi_t \qquad \phi_t = t\rho + (t^2/2)\mathcal{L}_V\rho + \dots$$

Upon integration over a p-chain c we get

$$\int_{\Phi_t(c)} \alpha = \int_c \alpha$$
$$\int_{\Phi_t(c)} \beta = \int_c \beta + \int_{\partial c} \phi_t$$

So, we see that α leads to an integral invariant, but the situation with β is more delicate. There is a term which spoils the "integral invariant property". But still there is a possibility to construct an integral invariant. We should just restrict ourself to *closed chains* c. Indeed, clearly

$$\partial c = 0 \qquad \Rightarrow \qquad \int_{\Phi_t(c)} \beta = \int_c \beta$$

So, whenever we find a form β obeying $\mathcal{L}_V\beta = d(\dots)$ for some V, there is a *relative integral invariant* given by the integral

$$\int_c \beta \qquad c \text{ such that} \qquad \partial c = 0$$

Now back to Hamiltonian mechanics (with time-independent Hamiltonian $H(q,p)$). We know (from 14.1.6iii and 14.3.3) that

$$\begin{aligned} \mathcal{L}_{\zeta_f}\omega &= 0 \\ \mathcal{L}_{\zeta_f}(\omega\wedge\omega) &= 0 \\ \mathcal{L}_{\zeta_f}(\omega\wedge\omega\wedge\omega) &= 0 \\ &\text{etc.} \end{aligned}$$

So, we have the (absolute) integral invariants described in the book. Now, for θ such that $\omega = d\theta$, we get

$$\mathcal{L}_{\zeta_f}\theta = i_{\zeta_f}d\theta + di_{\zeta_f}\theta = d(i_{\zeta_f}\theta - f) \equiv d(\dots)$$

But then, we have

$$\begin{aligned} \mathcal{L}_{\zeta_f}\theta &= d(\dots) \\ \mathcal{L}_{\zeta_f}(\theta\wedge\omega) &= d(\dots) \\ \mathcal{L}_{\zeta_f}(\theta\wedge\omega\wedge\omega) &= d(\dots) \\ \mathcal{L}_{\zeta_f}(\theta\wedge\omega\wedge\omega\wedge\omega) &= d(\dots) \\ &\text{etc.} \end{aligned}$$

since, for example

$$\mathcal{L}_{\zeta_f}(\theta\wedge\omega) = d(i_{\zeta_f}\theta - f)\wedge\omega = d[(i_{\zeta_f}\theta - f)\wedge\omega] \equiv d(\dots)$$

Therefore we have, in time-independent Hamiltonian mechanics, absolute integral invariants

$$\begin{aligned} I_2 &\equiv \int_{D_2}\omega && \text{for } \textit{any } D_2 \\ I_4 &\equiv \int_{D_4}\omega\wedge\omega && \text{for } \textit{any } D_4 \\ I_6 &\equiv \int_{D_6}\omega\wedge\omega\wedge\omega && \text{for } \textit{any } D_6 \\ &\text{etc.} \end{aligned}$$

and relative integral invariants

$$\begin{aligned} I_1 &\equiv \int_{c_1}\theta && \text{for any } \textit{closed } c_1 \ (\partial c_1 = 0) \\ I_3 &\equiv \int_{c_3}\theta\wedge\omega && \text{for any } \textit{closed } c_3 \ (\partial c_3 = 0) \\ I_5 &\equiv \int_{c_5}\theta\wedge\omega\wedge\omega && \text{for any } \textit{closed } c_5 \ (\partial c_5 = 0) \\ &\text{etc.} \end{aligned}$$

Interested reader might find useful to consult my (mostly tutorial) article ”Modern geometry in not-so-high echelons of physics: Case studies”, Acta Physica Slovaca 63, No.5, 261 359 (2013) (99 pages) (available online at http://www.physics.sk/aps or http://arxiv.org/abs/1406.0078)

14.5.3 p.351 Here we show a concrete example of a moment map. Namely, the moment map of a standard action of the rotation group $SO(3)$ on the common two-dimensional ”round” sphere S^2 (treated as a symplectic manifold, see 14.2.3).

The first way of computation:

The generators of the action are displayed in 13.4.6 (or in any textbook of Quantum Mechanics :-):

$$\begin{aligned}\xi_{l_1} &= -\sin\varphi\partial_\vartheta - \operatorname{ctg}\vartheta\cos\varphi\partial_\varphi\\ \xi_{l_2} &= \cos\varphi\partial_\vartheta - \operatorname{ctg}\vartheta\sin\varphi\partial_\varphi\\ \xi_{l_3} &= \partial_\varphi;\end{aligned}$$

(Here $E_j = l_j$ is the standard basis in $so(3)$, see 11.7.13). The volume = symplectic form (on *unit* sphere) reads

$$\omega = \sin\vartheta d\vartheta \wedge d\varphi$$

Therefore

$$\begin{aligned}i_{\xi_{l_1}}\omega &= \cdots = -d(-\sin\vartheta\cos\varphi)\\ i_{\xi_{l_2}}\omega &= \cdots = -d(-\sin\vartheta\sin\varphi)\\ i_{\xi_{l_3}}\omega &= \cdots = -d(-\cos\vartheta)\end{aligned}$$

The functions in the brackets are, however, nothing but (restriction to the sphere of) minus standard coordinate functions $(x, y, z) \equiv (x_1, x_2, x_3)$ in the ambient space E^3. So, we get

$$i_{\xi_{E_j}}\omega = -d(-x_j(\vartheta, \varphi))$$

Since $P_j(\vartheta, \varphi)$ should be after $-d$ symbol, we get

$$P_j(\vartheta, \varphi) = -x_j(\vartheta, \varphi)$$

or

$$P \equiv P_j E^j = -x_j(\vartheta, \varphi)E^j$$

The second way of computation:

Recall that orbits of the co-adjoint action for $G = SO(3)$ are just spheres (see 14.6.7). On the unit sphere, the canonical symplectic form (see 14.6.3) is a 2-form

$$\omega = f(\vartheta, \varphi)d\vartheta \wedge d\varphi$$

obeying

$$\omega_{X^*}(\xi_{l_1}, \xi_{l_2}) = \langle X^*, [l_1, l_2]\rangle = \langle X^*, l_3\rangle = x_3 = \cos\vartheta$$

Plugging explicit expressions for ξ_{l_1} and ξ_{l_2} gives $f(\vartheta, \varphi) = \sin\vartheta$. So, this is exactly the standard volume form.

But there is an explicit result for the moment map for any co-adjoint orbit in 14.6.5iv):

$$P : \mathcal{O} \to \mathcal{G}^* \qquad X^* \mapsto X^*$$

So, a point on the sphere with spherical coordinates (ϑ, φ), i.e. sitting in

$$X^* \leftrightarrow (x_1(\vartheta, \varphi), x_2(\vartheta, \varphi), x_3(\vartheta, \varphi) \equiv (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$$

of $\mathcal{G}^* \equiv so(3)^* \equiv \mathbb{R}^3$ maps to the point

$$X^* \leftrightarrow x_j E^j$$

in $\mathcal{G}^* \equiv so(3)^* \equiv \mathbb{R}^3$. Put it another way, the three components of the moment map are just the three *Cartesian* coordinates of the point. This is, modulo sign, in agreement with the first way of computation. (The sign probably comes from the freedom in the sign of the symplectic form on S^2, i.e. ω is equally good as $-\omega$. But I am not sure, the reader is invited to fill this gap herself :-)

Two simple checks of the result:

1. Consider a Hamiltonian on S^2 invariant with respect to the action. So, it is (completely!) rotational-invariant. So *constant*. There are three conserved quantities, x_1, x_2, x_3. This means, however, there is *no motion* on the sphere. But this is OK for constant Hamiltonian :-)

2. Consider a Hamiltonian on S^2 invariant just with respect to rotation about the x_3-axis. So, it only depends on ϑ, $H = H(\vartheta)$. The relevant part of the action is the action of $SO(2)$ generated by ξ_{l_3}. Then $P_3 = x_3$ is conserved (and nothing more). But this is OK for Hamiltonian dependent on just ϑ, since $\dot{\vartheta} = \cdots = 0$, so that $\vartheta(t) =$ const., so that x_3 is const.

14.5.7 p.354 We learned that, for each dimension (each basis element E_i) of the Lie algebra $\mathcal{G}$, we get conserved quantity $P_i(x)$. A natural question arises whether these conserved quantities are *functionally independent.* Put it another way, whether each dimension of the symmetry Lie algebra $\mathcal{G}$ adds a fully-fledged conserved quantity rather than quantities, which are just functions of those already obtained before.

Imagine $f(x)$ and $g(x)$ are two conserved quantities ($\dot{f} = 0 = \dot{g}$). Then $h(f(x), g(x))$ is clearly conserved quantity as well. Indeed,

$$\dot{h} = \frac{\partial h}{\partial f}\dot{f} + \frac{\partial h}{\partial g}\dot{g} = 0$$

Note that, in this case,

$$dh = \frac{\partial h}{\partial f}df + \frac{\partial h}{\partial g}dg$$

so the *differentials* of the three functions f, g, h are linearly dependent. On the contrary, linear independence of $df_1, \ldots, df_n$ implies functional independence of $f_1(x), \ldots, f_n(x)$, which means there is no F such that

$$F(f_1(x), \ldots, f_n(x)) = 0$$

or equivalently, one cannot express any particular f_k in terms of the rest members f_j.

Therefore, in the case of our interest, we are interested in (point-wise) linear independence of the differentials

$$dP_1(x), \ldots, dP_n(x) \qquad\qquad n = \dim \mathcal{G}$$

Now, hamiltonian character of symmetry generators

$$i_{\xi_{E_j}}\omega = -dP_j$$

(see the text between 14.5.1 and 14.5.2, p.350) and non-degeneracy of ω shows that linear independence of the differentials is the same thing as linear independence of the generators ξ_{E_j}. So, when the generators are (point-wise) linearly independent? The answer: the action is to be *free.* Indeed, otherwise elements X of the Lie sub-algebra, which corresponds to (non-vanishing) stabilizer of a point x result in *vanishing* generators ξ_X. So, in an adapted basis E_j, some of ξ_{E_j} vanish and the (whole set of) generators ξ_{E_j} cease to be linearly independent.

14.7.1 p.361 Some more details concerning the proof of the statement are presented here.

First, let us see in more detail, why "the complement of the complement turns out to be the initial subspace once again". (This is a well-known fact in *Euclidean* case, but perhaps it is less clear for the case of *symplectic* orthogonality.)

So, let $L_1 \subset L$ be a k-dimensional subspace of an n-dimensional linear space L. Then, there is a distinguished $(n-k)$-dimensional subspace $\hat{L}_1 \subset L^*$ in the *dual* space, the *annihilator* of L_1 (see 2.4.9, 10.1.13)

$$\langle \hat{L}_1, L_1 \rangle = 0 \qquad\qquad L_1 \subset L\ , \quad \hat{L}_1 \subset L^*$$

(If (e_a, e_i) is an adapted basis in L (i.e. $e_a \in L_1$), then i-part of the dual basis (e^a, e^i) of L^* serves as a basis of $\hat{L}_1$.)

Now, we have two (mutually inverse) linear isomorphisms

$$\omega : L \to L^* \qquad\qquad v \mapsto \omega(v,\ .\ )$$
$$\omega^{-1} : L^* \to L \qquad\qquad \alpha \mapsto \omega^{-1}(\alpha,\ .\ )$$

(They are *iso*morphisms because of non-degeneracy of ω.) Let $L_2 \subset L$ be the image of $\hat{L}_1$ w.r.t. ω^{-1}

$$L_2 := \omega^{-1}(\hat{L}_1) \subset L$$

Since ω^{-1} is isomorphism, the dimension of L_2 is the same as the dimension of $\hat{L}_1$, i.e. $(n-k)$ (complementary dimension w.r.t. L_1)

$$\dim L_1 = k \qquad \Rightarrow \qquad \dim L_2 = n - k \qquad\qquad (n \equiv \dim L)$$

Now, let $v \in L_1$ and $w \equiv \omega^{-1}(\alpha, \ .\ ) \in L_2$. Then

$$\omega(v,w) = \omega(v, \omega^{-1}(\alpha, \ .\ )) = \omega_{ij} v^i \omega^{kj} \alpha_k = -\alpha_k v^k \equiv -\langle \alpha, v \rangle$$

This shows that L_2 is nothing but the *symplectic-orthogonal complement* $L_1^\perp$ of L_1. (Vanishing of the LHS holds for exactly those w which are images of $\alpha \in \hat{L}_1$.) So, the dimension of the complement $L_1^\perp$ is always *complementary* to the dimension of L_1

$$\dim L = \dim L_1 + \dim L_1^\perp \qquad\qquad n = k + (n-k)$$

Consequently, the dimension of the complement *of the complement* is complementary to the dimension of the complement and this is nothing than the dimension of the original subspace L_1:

$$\dim (L_1^\perp)^\perp = \dim L - \dim L_1^\perp = \dim L_1$$

(since $n - (n-k) = k$:-). Now, by definition (of $L_1^\perp$)

$$\omega(L_1, L_1^\perp) = 0$$

The same formula, however, shows (!) that $L_1 \subset (L_1^\perp)^\perp$ (only the fact that it is a *subspace* is clear at once). But since we already know that L_1 and $(L_1^\perp)^\perp$ have the same dimension, the spaces are actually *equal*:

$$(L_1^\perp)^\perp = L_1$$

So, in general, making orthogonal *complement* (of a subspace) *twice* is *doing nothing*, we return to the subspace again

$$\boxed{U^{\perp\perp} = U}$$

[Looking at the proof we see, that this fact is true for both symmetric and skew-symmetric "scalar products" in L. The symmetric case needs not be positive definite, it works in pseudo-Euclidean cases as well. What really matters is *non-degeneracy* of the product. Notice, however, that only these two cases - symmetry and skew-symmetry - make the very notion of orthogonality (and, consequently, the orthogonal complement) meaningful! For a "general" (even though non-degenerate) bilinear form, the value of the "scalar product" $B(v,w)$ depends on the order of the pair, so it can happen $B(v,w) = 0$ and $B(w,v) \neq 0$. So, orthogonality depends on the order.]

B.t.w., the fact that the *dimensions* of L_1 and $L_1^\perp$ add to the full dimension of L (i.e. $\dim L = \dim L_1 + \dim L_1^\perp$) does *not* mean, that the subspaces themselves do add (i.e. that $L = L_1 \oplus L_1^\perp$)! Consider, as a simple example, two dimensional symplectic space with $\omega = e^1 \wedge e^2$. Let $L_1 = \text{Span } e_1$. Then $L_1^\perp = L_1$ - it is the subspace itself (!). Such subspaces are called *Lagrangian* and they play important role in the theory.

Finally, let us mention, just for fun, a *direct* proof of the opposite implication (not based on the fact that $U^{\perp\perp} = U$; notice that *this very* implication is crucial in the proof of 14.7.6). What we want to prove is the following statement: any vector which is symplectic-orthogonal to the subspace tangent to M_p is necessarily of the form ξ_X for some $X \in \mathcal{G}$. Well, being tangent to M_p is the same thing as annihilate all 1-forms dP_i. So, the *annihilator* of the subspace of vectors tangent to M_p is *spanned* by the covectors dP_i. The orthogonal complement to the subspace of vectors v tangent to M_p consists of all vectors w such that $\omega(w,v) = 0$ (by definition). Equivalently, of such w that the 1-form $\omega(w, \ .\ )$ belongs to the annihilator. Since the latter is spanned by dP_i, there are numbers X^i such that

$$\omega(w, \ .\ ) = X^i dP_i$$

Applying "raising of indices" operation on both sides of the equation we get

$$w = X^i \zeta_{P_i} = X^i \xi_{E_i} = \xi_X$$

And that's all :-)

15.3.4 p.385 The coordinate-free expression of the RLC-connection from (iii) is called the *Koszul formula*

15.3.4 p.385, 15.3.14 p.389, 15.6.9 p.411 and 22.5.4 p.665:

In 15.3.4, the classical (coordinate) expression

$$\Gamma^i_{jk} = \frac{1}{2} g^{il}(g_{lj,k} + g_{lk,j} - g_{jk,l}) \qquad \text{i.e.} \qquad 2\Gamma_{ijk} = g_{ij,k} + g_{ik,j} - g_{jk,i} \tag{a1}$$

for Christoffel symbols of the *metric and symmetric* (= RLC) connection is derived.

In 15.3.14, it is generalized to just *metric* connection (torsion allowed, still coordinate frame).

In 15.6.9, *metric* connection in *orthonormal* frame is discussed (torsion allowed).

In 22.5.4, *metric* connection coefficients in *orthonormal* frame enter *Dirac operator* in "curved space".

It might be of interest to learn what the situation looks like for the *most general* case, i.e. when

i) *torsion* as well as *non-metricity* may not vanish (connection is neither metric nor symmetric)

ii) the frame field e_a is *general* (neither orthonormal nor coordinate; we have $[e_a, e_b] = c^c_{ab} e_c$).

Recall that we found the above mentioned expressions for Γ_{ijk} (for the RLC case!) by solving system of equations

$$\begin{aligned} \Gamma_{ijk} + \Gamma_{jik} &= g_{ij,k} && \text{since it is metric} \\ \Gamma_{ijk} - \Gamma_{ikj} &= 0 && \text{since it is symmetric} \end{aligned}$$

Now, the system is to be generalized to

$$\Gamma_{abc} + \Gamma_{bac} = A_{abc} \qquad A_{abc} = A_{bac} \tag{a2}$$

$$\Gamma_{abc} - \Gamma_{acb} = B_{abc} \qquad B_{abc} = -B_{acb} \tag{a3}$$

where the following abbreviations are used:

$$\begin{aligned} A_{abc} &\equiv g_{ab,c} - g_{ab;c} \\ B_{abc} &\equiv -(c_{abc} + T_{abc}) \end{aligned}$$

and

$$\begin{aligned} g_{ab,c} &:= e_c g_{ab} \\ g_{ab;c} &:= (\nabla_c g)_{ab} && \textit{non-metricity} \text{ tensor} \end{aligned}$$

▼ Indeed:

$$\begin{aligned} g_{ab;c} &= (\nabla_c g)(e_a, e_b) \\ &= \nabla_c (g(e_a, e_b)) - g(\nabla_c e_a, e_b) - g(e_a, \nabla_c e_b) \\ &= e_c g_{ab} - \Gamma^d_{ac} g_{db} - \Gamma^d_{bc} g_{ad} \\ &= g_{ab,c} - \Gamma_{bac} - \Gamma_{abc} \end{aligned}$$

$$\begin{aligned} T^c_{ab} e_c &= T(e_a, e_b) \\ &= \nabla_a e_b - \nabla_b e_a - [e_a, e_b] \\ &= \Gamma^c_{ba} e_c - \Gamma^c_{ab} e_c - c^c_{ab} e_c \end{aligned}$$

so that

$$T_{cab} \equiv g_{cd} T^d_{ab} = g_{cd}(\Gamma^d_{ba} - \Gamma^d_{ab} - c^d_{ab}) = \Gamma_{cba} - \Gamma_{cab} - c_{cab}$$

▲

Notice that the non-metricity tensor is a natural measure of *deviation* from metricity of a connection. We will *not* assume this tensor to vanish, now. Rather we will only assume this tensor to be *given*.

Now, what was *the* trick which enabled us to solve the equations (a2) and (a3)? We see from the result (a1) that what very much helps is to make the following combination

$$g_{ij,k} + g_{ki,j} - g_{jk,i}$$

from (a2) and then use (a3) whenever needed. (Here, the expression is written in a form when each term begins with a different index.)

In order to formalize (and then to repeat) the trick, one can introduce, for any three-index quantity Q_{abc}, the *contortion operation* (also known[4] as the *Schouten braces*)

$$\boxed{Q_{\{abc\}} := Q_{abc} + Q_{cab} - Q_{bca}} \tag{a4}$$

Then, clearly, (a1) reads

$$2\Gamma_{ijk} = g_{\{ij,k\}} \qquad \text{for the RLC-connection} \tag{a5}$$

But the real goal to introduce Schouten braces, as indicated above, is to use them to solve the general system (a2) and (a3). So, let us repeat the trick there:

Apply the Schouten braces to (a2) and, when needed, use (a3).

In this way we get unique solution:

$$2\Gamma_{abc} = A_{\{abc\}} - B_{\{acb\}}$$

▼ Indeed:

$$\begin{aligned} A_{\{abc\}} &\equiv A_{abc} + A_{cab} - A_{bca} \\ &= (\Gamma_{abc} + \Gamma_{bac}) + (\Gamma_{cab} + \Gamma_{acb}) - (\Gamma_{bca} + \Gamma_{cba}) \\ &= (\Gamma_{abc} + \Gamma_{acb}) + (\Gamma_{bac} - \Gamma_{bca}) + (\Gamma_{cab} - \Gamma_{cba}) \\ &= (\Gamma_{abc} + (\Gamma_{abc} - B_{abc})) + (B_{bac}) + (B_{cab}) \\ &= 2\Gamma_{abc} - B_{abc} + B_{bac} + B_{cab} \end{aligned}$$

so that

$$2\Gamma_{abc} = A_{\{abc\}} + B_{abc} - B_{bac} - B_{cab}$$

The last three terms clearly resemble Schouten braces of (perhaps) B_{abc}. Unfortunately,

$$B_{abc} - B_{bac} - B_{cab} \neq B_{\{abc\}} \qquad : -($$

(check). But, fortunately, there still *is* a possibility to write it as

$$B_{abc} - B_{bac} - B_{cab} = -B_{\{acb\}} \qquad : -)$$

(check, use skew symmetry $B_{abc} = -B_{acb}$). So, in order to express the result in terms of Schouten braces, one also has to perform some *index reshuffling*. This is a darker side of life, since the resulting formula is then harder to remember. ▲

Inserting concrete expressions for A_{abc} and B_{abc} we get the following expression for Christoffel symbols of general linear connection on Riemannian manifold (M, g):

$$2\Gamma_{abc} = g_{\{ab,c\}} - g_{\{ab;c\}} + T_{\{acb\}} + c_{\{acb\}} \tag{a6}$$

The third term is often expressed in terms of specifically introduced *contortion tensor*

$$K_{abc} := \frac{1}{2} T_{\{acb\}} \equiv \frac{1}{2}(T_{acb} + T_{bac} - T_{cba}) = -\frac{1}{2}(T_{abc} + T_{bca} - T_{cab}) \tag{a7}$$

[4]See page 132, formula 3.7 in J.A.Schouten: Ricci Calculus, Springer-Verlag 1954; Schouten *braces* are not to be confused with a completely different (and much more profound!) object, the Schouten *bracket* (known also as *Schouten-Nijenhuis bracket*).)

(for the last equality sign we used skew-symmetry $T_{abc} = -T_{acb}$). Then, the decomposition formula (a6) reads

$$\boxed{\Gamma_{abc} = \frac{1}{2}(g_{ab,c} + g_{ac,b} - g_{bc,a}) - \frac{1}{2}(g_{ab;c} + g_{ac;b} - g_{bc;a}) + K_{abc} - \frac{1}{2}(c_{abc} + c_{bca} - c_{cab})} \tag{a8}$$

So, in general, coefficients of arbitrary linear connection may be uniquely written as a *sum of four terms*:

The first plus the fourth term correspond to good old RLC-connection.

The second term is given by (minus) Schouten braces of non-metricity tensor. (It vanishes for metric connection).

The third term is given by the contortion tensor, i.e. by a combination of Schouten braces and index reshuffling of the torsion tensor (it vanishes for symmetric connection).

Symbolically (in words)

$$\boxed{\text{general} = \text{RLC} + \text{non-metricity part} + \text{contortion part}} \tag{a9a}$$

so that

$$\boxed{\text{metric} = \text{RLC} + \text{contortion part}} \tag{a9b}$$

and

$$\boxed{\text{metric and symmetric} = \text{RLC}} \tag{a9c}$$

Particular cases of (a8):

1. In *coordinate* (= holonomic) frame, the last term (containing anholonomy coefficients c_{abc}) vanishes and what remains is

$$\Gamma_{ijk} = \frac{1}{2}(g_{ij,k} + g_{ik,j} - g_{jk,i}) - \frac{1}{2}(g_{ij;k} + g_{ik;j} - g_{jk;i}) + K_{ijk} \tag{a10}$$

2. In *orthonormal* frame $e_c g_{ab} = 0$ (since $g_{ab} = \eta_{ab} =$ const.). If the connection is, in addition, *metric*, i.e. $g_{ab;c} = 0$, we are left with

$$\Gamma_{abc} = -\frac{1}{2}(c_{abc} + c_{bca} - c_{cab}) + K_{abc} \tag{a11}$$

(This is the *correct* result of 15.6.20iii) - the contortion term is forgotten, there. In 15.6.9 it is ok, since only RLC case is discussed.)

Btw, from (a10) we see, that Christoffel symbols of any linear connection are given as a sum of a unique connection (namely the RLC one) plus a tensor. Therefore, "the difference of any two connections is a tensor".

Let us remark, finally, that Schouten braces are "invertible" - one can express original quantity Q_{ijk} in terms of appropriate combinations of its Schouten braces. Namely, the explicit formula reads

$$Q_{ijk} = \frac{1}{2}(Q_{\{ijk\}} + Q_{\{jki\}}) \tag{a12}$$

In particular, it means that one can reconstruct the "original" torsion tensor from "its" contortion tensor:

$$T_{ijk} = -(K_{ijk} + K_{kij}) \tag{a13}$$

▼ Indeed, we can write a matrix relation

$$\begin{pmatrix} Q_{\{ijk\}} \\ Q_{\{kij\}} \\ Q_{\{jki\}} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} Q_{ijk} \\ Q_{kij} \\ Q_{jki} \end{pmatrix}$$

(The first line is definition (a4), remaining lines arise by just renaming indices.) The inverse relation is

$$\begin{pmatrix} Q_{ijk} \\ Q_{kij} \\ Q_{jki} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} Q_{\{ijk\}} \\ Q_{\{kij\}} \\ Q_{\{jki\}} \end{pmatrix}$$

The first line gives (a12).

Similarly, from (a7) we have

$$\begin{pmatrix} K_{ijk} \\ K_{kij} \\ K_{jki} \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} T_{ijk} \\ T_{kij} \\ T_{jki} \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} T_{ijk} \\ T_{kij} \\ T_{jki} \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ -1 & 0 & -1 \end{pmatrix}\begin{pmatrix} K_{ijk} \\ K_{kij} \\ K_{jki} \end{pmatrix}$$

The first line gives (a13). ▲

15.3.9 p.386 Remarkable insight is gained through the approach discussed in the additional material to the problem (15.6.10). It is shown there that the resulting angle of rotation of the vector may be expressed in terms of the solid angle subtended by the path. (Here the angle is clearly 1/8 of the total solid angle 4π.)

15.3.14 p.389: see the comment to 15.3.4, p.385 (here, in the Additional material).

15.4.3 p.391 There is an error in part (i), both in the assignment and in the hint. A more detailed general treatment of how (velocities and) accelerations are related for two curves, differing just in parametrization, might be useful to understand the situation.

Ok, let $\hat{\gamma}(t) = \gamma(\sigma(t))$. So, we are given two curves, $\hat{\gamma}$ and γ, passing a common *set* of points on M, but differing in "times" visiting particular point. (As an example, if $\sigma(t) = t^3$, then $\hat{\gamma}(1) = \gamma(1)$, $\hat{\gamma}(2) = \gamma(8)$), $\hat{\gamma}(3) = \gamma(27)$, etc.). In a point $P \equiv \hat{\gamma}(t) = \gamma(\sigma(t))$, two relevant vectors may be associated with each curve, their velocity and acceleration. So, altogether, as many as *four* vectors are to be considered in (each) P:

$$\hat{v} := \dot{\hat{\gamma}}(t) \qquad \hat{a} := (\nabla_{\dot{\hat{\gamma}}}\dot{\hat{\gamma}})(t) \qquad v := \dot{\gamma}(\sigma(t)) \qquad a := (\nabla_{\dot{\gamma}}\dot{\gamma})(\sigma(t))$$

The task is to express hatted quantities in terms of original (unhatted) ones. From (2.3.5) we know how hatted *velocity* is expressed in terms of the original one:

$$\hat{v} = \sigma' v \qquad\qquad \sigma' \equiv \sigma'(t)$$

so that $\hat{v}$ is just an appropriate *multiple* of v.

▼

$$\hat{v}f = \frac{d}{dt}f(x^i(\sigma(t))) = \frac{\partial f}{\partial x^i}\frac{dx^i}{d\sigma}\frac{d\sigma}{dt} = \left(\sigma'(t)\frac{dx^i}{d\sigma}\partial_i\right)f = (\sigma'(t)\dot{\gamma}(\sigma(t)))\,f = (\sigma'(t)v)f$$

▲

Let us proceed to the acceleration, now. We get

$$\hat{a} = \nabla_{\hat{v}}\hat{v} = \nabla_{\hat{v}}(\sigma' v) = (\hat{v}\sigma')v + \sigma'\nabla_{\hat{v}}v = (\hat{v}\sigma')v + \sigma'\nabla_{\sigma' v}v = (\hat{v}\sigma')v + (\sigma')^2\nabla_v v = \sigma'' v + (\sigma')^2 a$$

or, in the notation used in the book,

$$\nabla_{\dot{\hat{\gamma}}}\dot{\hat{\gamma}} = \sigma''\dot{\gamma} + (\sigma')^2\nabla_{\dot{\gamma}}\dot{\gamma}$$

▼ We can regard $\sigma(t)$ (and, consequently, also $\sigma'(t)$ and $\sigma''(t)$) as a function on the curve $\hat{\gamma}(t)$. Then, applying $\dot{\hat{\gamma}}$ on σ', giving by definition derivative of σ' along the curve $\hat{\gamma}(t)$, results in nothing but the derivative of σ' with respect to t

$$(\hat{v}\sigma')(t) = (\dot{\hat{\gamma}}\sigma')(t) = (d/dt)\sigma'(t) = \sigma''(t)$$

▲

So, the answer to the question, how the hatted quantities are related to the original ones, reads:

$$\hat{v} = \sigma' v$$
$$\hat{a} = \sigma'' v + (\sigma')^2 a$$

Then, the answer to (15.4.3) part (i) follows readily: if γ is an affinely parameterized geodesics (i.e. $a = 0$), then $\hat{a} = \sigma'' v$. Therefore, requiring $\hat{\gamma}$ be an affinely parameterized geodesics as well, needs $\sigma'' = 0$ (i.e. $\sigma(t) = k_1 t + k_2$).

Example 1: Dimensional check. Let dimension of γ, $\hat{\gamma}$ (or corresponding coordinates x^i) be A, dimension of t be B and dimension of σ be C. Then the dimensions of relevant quantities read:

$$\hat{v} \leftrightarrow A/B \qquad \hat{a} \leftrightarrow A/B^2 \qquad v \leftrightarrow A/C \qquad a \leftrightarrow A/C^2 \qquad \sigma' \leftrightarrow C/B \qquad \sigma'' \leftrightarrow C/B^2$$

and we see that both equations are dimensionally correct.

Example 2: A simple random (and dull) check of the validity of the two formulas above: Consider a one-dimensional motion of a point on a straight line given by $x(t) = t^4$ and a reparametrization of the form $\sigma(t) = t^3$. Then

$$\begin{array}{lll} \sigma(t) = t^3 & \sigma'(t) = 3t^2 & \sigma''(t) = 6t \\ x(t) = t^4 & v \equiv \dot{x}(\sigma(t)) = 4t^9 & a = \ddot{x}(\sigma(t)) = 12t^6 \\ \hat{x}(t) = t^{12} & \hat{v} \equiv \dot{\hat{x}}(t) = 12t^{11} & \hat{a} = \ddot{\hat{x}}(t) = 122t^{10} \end{array}$$

The question is whether

$$12t^{11} \stackrel{?}{=} 3t^2 4t^9$$
$$122t^{10} \stackrel{?}{=} 6t4t^9 + (3t^2)^2 12t^6$$

is true. Yes, it is :-)

Example 3: Consider (as $\gamma(t)$) a *uniform* motion of a point along a circle. Then, as we know, the acceleration a is oriented to the center (and the velocity is tangent to the circle, the vectors are perpendicular). The magnitude of both velocity and acceleration is constant. Now let, say, $\sigma(t) = t^2$. Then $\hat{\gamma}(t)$ describes a motion which is (still along the same circle but) non-uniform - we get $\hat{v} = 2tv$, so the magnitude of the new velocity grows. This is compatible with $\hat{a} = 2v + 4t^2 a$ - we see that the new acceleration exhibits tangent (constant) component $2v$, so the magnitude of the velocity grows linearly.

15.4.4, 18.4.6 pp.392,511: In 15.4.4 we learned how to use *Lagrange* equations for finding geodesics (of the RLC-connection) - simply take

$$L(x, v) = (1/2)g_{ab}(x)v^a v^b$$

(the *kinetic* energy alone) and write down Lagrange equations. From 18.4.6 we know how to compute the Hamiltonian "corresponding" to a given Lagrangian. For our L we get

$$H(x, p) = (1/2)g^{ab}(x)p_a p_b$$

(the same *kinetic* energy alone). So, we can also write *geodesic* equations in the form of appropriate *Hamiltonian* equations - simply take $H(x, p) = (1/2)g^{ab}(x)p_a p_b$ and write down Hamiltonian equations. We get explicitly (check)

$$\begin{aligned} \dot{x}^a &= g^{ab}(x)p_b \\ \dot{p}_a &= -(1/2)g^{bc}{}_{,a}(x)p_b p_c \end{aligned}$$

Once we solve these $2n$ first order equations, we get curves on T^*M, in coordinates $x^a(t), p_a(t)$. Their projections to M, i.e. $x^a(t)$, are then geodesics on (M, g).

One can also easily check explicitly, that these (first order) equations are indeed equivalent to the (second order) *geodesic* equations. Indeed, from the first equation we have $p_a = g_{ab}(x)\dot{x}^b$. Plugging this into the second one leads to

$$g_{ab}\ddot{x}^b + g_{ab,c}\dot{x}^b\dot{x}^c = -\frac{1}{2}g_{br}g^{bc}{}_{,a}g_{cs}\dot{x}^r\dot{x}^s$$

Differentiating the identity $g^{bc}g_{cd} = \delta^b_d$ with respect to x^a we get

$$-g_{br}g^{bc}{}_{,a}g_{cs} = g_{rs,a}$$

and, therefore

$$g_{ab}\ddot{x}^b + \Gamma_{abc}\dot{x}^b\dot{x}^c = 0 \qquad\qquad \Gamma_{abc} := \frac{1}{2}(g_{ab,c} + g_{ac,b} - g_{bc,a})$$

15.5.4*i* p.403: There is a tricky point (potential *sign* error) in the action of the $\binom{1}{1}$-tensor A of a general derivation $D \equiv \mathcal{L}_V + A$.

Recall that a general $\binom{1}{1}$-tensor A was, up to now (see 2.4.5ii), considered in *three* roles:

- as a mapping $v \mapsto A(v;\ .\ ) \equiv A(v)$ (vector to vector)
- as a mapping $\alpha \mapsto A(\ .\ ;\alpha) \equiv A(\alpha)$ (covector to covector)
- as a mapping $(v,\alpha) \mapsto A(v;\alpha)$ (vector and covector to a number)

Here it is considered in a *fourth* role:

- as a mapping $t \mapsto A\cdot t$ ($\binom{p}{q}$-tensor to $\binom{p}{q}$-tensor)

In particular, the tensor t may be a vector or a covector. In these two particular cases we get

$$A\cdot v = A(v) \tag{1}$$
$$A\cdot \alpha = -A(\alpha) \tag{2}$$

This means, on a basis vectors an (dual) covectors

$$A\cdot e_a = A^b_a e_b \tag{1a}$$
$$A\cdot e^a = -A^a_b e^b \tag{2a}$$

So, there is an important *difference* between $A(\alpha)$ and $A\cdot\alpha$, a potential danger to *forget* about the *minus sign* in $A\cdot(\ )$-operation on *co*vectors.

▼ By definition, the mapping

$$t \mapsto A\cdot t$$

is to be a *derivation* of tensor algebra which, in addition, vanishes on $\binom{0}{0}$-tensors (scalars) and commutes with contractions. Then we get

$$\begin{aligned} A\cdot\langle\alpha,v\rangle &\overset{1.}{=} 0 \\ &\overset{2.}{=} \langle A\cdot\alpha,v\rangle + \langle\alpha,A\cdot v\rangle \end{aligned}$$

The first line comes from vanishing on scalars, the second line holds because of

$$A\cdot\langle\alpha,v\rangle \equiv A\cdot(C(\alpha\otimes v)) = C(A\cdot(\alpha\otimes v)) = C((A\cdot\alpha)\otimes v + \alpha\otimes(A\cdot v)) = \langle A\cdot\alpha,v\rangle + \langle\alpha,A\cdot v\rangle$$

So, we get

$$\langle A\cdot\alpha,v\rangle = -\langle\alpha,A\cdot v\rangle$$

or, choosing $v = e_a$ and $\alpha = e^b$,

$$(A\cdot e^b)_a = -(A\cdot e_a)^b$$

This means that if we *choose*

$$A\cdot v := A(v)$$

we inevitably have to accept

$$A\cdot\alpha = -A(\alpha)$$

since definitions of $A(v)$ and $A(\alpha)$, see (1) and (2), result in

$$\langle A(\alpha),v\rangle = \langle\alpha,A(v)\rangle$$

(Btw. we could equally well choose plus sign on covectors and then accept minus sign on vectors. So an invariant statement is that there is a sign difference on either vectors or covectors in action of A *as a derivation* of tensor algebra in comparison with its natural action on the latter stemming from mere definition of A *as a* $\binom{1}{1}$*-tensor*.) ▲

15.5.4*ii* p.403: The formula $\nabla_V = \mathcal{L}_V + (\nabla V)$ to be proved in part (ii) is not completely general. Rather, if *torsion* does not vanish, it should read

$$\boxed{\nabla_V = \mathcal{L}_V + (\nabla V) + T(V,\ .\ )}$$

Indeed, let $A := \nabla_V - \mathcal{L}_V$. Then, from the definition of the torsion tensor

$$T(V,W) := \nabla_V W - \nabla_W V - [V,W]$$

we infer

$$A(W) = (\nabla_V - \mathcal{L}_V)(W) = \nabla_V W - [V,W] = \nabla_W V + T(V,W) \equiv (\nabla V + T(V,\ .\ ))(W)$$

So, in components,

$$A^\mu_\nu = V^\mu{}_{;\nu} + T^\mu_{\rho\nu} V^\rho$$

15.5.6i p.404: From the result derived in i)

$$R(\partial_i,\partial_j)\partial_k = R^l_{kij}\partial_l$$

combined with

$$\tau_{A,A} = (\hat{1} - \epsilon^2 R(U,V))$$

from 15.5.1 one can envisage a situation in which the R^l_{kij}-component of the Riemann tensor is "in action". Namely, perform parallel transport of ∂_k around the $\epsilon \times \epsilon$ coordinate loop in the ij-plane (spanned on ∂_i and ∂_j). Then the resulting vector differs from the original ∂_k by a vector, whose l-th component is just $-\epsilon^2 R^l_{kij}$:

$$\partial_k \mapsto \partial_k - \epsilon^2 R^l_{kij}\partial_l$$

15.5.6ii p.404: Here, explicit formula for $A^{k\dots l}_{r\dots s;i;j} - A^{k\dots l}_{r\dots s;j;i}$ is presented. As an example it says that for *vector* field W we get

$$W^k{}_{;i;j} - W^k{}_{;j;i} = -R^k_{mij}W^m$$

Well, there is a problem with this formula. Namely, one can interpret the expression present at the l.h.s. in (at least) two ways:

$$W^k{}_{;i;j} \overset{1.}{=} (\nabla_j(\nabla_i W))^k \qquad \text{my book} \qquad (1)$$

$$W^k{}_{;i;j} \overset{2.}{=} (W^k{}_{;i})_{\ ;j} \qquad \text{virtually } \textit{all other} \text{ sources :-(} \qquad (2)$$

And, unfortunately, it makes a difference provided that there is a *non-vanishing torsion*. Let's compute:

In my interpretation, $\nabla_i W = W^k{}_{;i}\partial_k$ is still a *vector* and the same is true for

$$\nabla_j(\nabla_i W) = \nabla_j(W^k{}_{;i}\otimes\partial_k) = (\partial_j W^k{}_{;i})\partial_k + W^k{}_{;i}\nabla_j\partial_k = ((W^k{}_{;i})_{\ ,j} + \Gamma^k_{mj}W^m{}_{;i})\partial_k \qquad (A)$$

So *a single* Γ-term occurs (coming from $\nabla_j\partial_k = \Gamma^m_{kj}\partial_m$).

In the "rest of the world" interpretation, $W^k{}_{;i}$ already corresponds to a $\binom{1}{1}$-*tensor* $B = \nabla W$

$$B = B^i_j dx^j \otimes \partial_i \qquad B^i_j := W^i{}_{;j} \qquad (B)$$

so that

$$(W^k{}_{;i})_{\ ;j} \equiv B^k_i{}_{;j} = B^k_i{}_{,j} + \Gamma^k_{mj}B^m_i - \Gamma^m_{ij}B^k_m \qquad (C)$$

and as many as *two* Γ-terms do occur (also the new lower index i adds its Γ-term).

Now, from (A), (B) and (C) we see that

$$(\nabla_j(\nabla_i W))^k = B^k_i{}_{,j} + \Gamma^k_{mj}B^m_i$$
$$(W^k{}_{;i})_{\ ;j} = B^k_i{}_{,j} + \Gamma^k_{mj}B^m_i - \Gamma^m_{ij}B^k_m$$

i.e. the two interpretations ((1) versus (2)) are related as follows:

$$(W^k{}_{;i})_{\ ;j} = (\nabla_j(\nabla_i W))^k - \Gamma^m_{ij}W^k{}_{;m}$$

Since we know from 15.5.6i) that antisymmetrization leads in *my* interpretation to

$$(\nabla_j(\nabla_i W))^k - (\nabla_i(\nabla_j W))^k = -R^k_{mij}W^m$$

we can write

$$(W^k_{\ ;i})_{\ ;j} - (W^k_{\ ;j})_{\ ;i} = -R^k_{mij} W^m - (\Gamma^m_{ij} W^k_{\ ;m} - \Gamma^m_{ji} W^k_{\ ;m})$$

And because of

$$T^k_{ij} = -(\Gamma^k_{ij} - \Gamma^k_{ji})$$

(see 15.3.3), we get finally

$$\boxed{(W^k_{\ ;i})_{\ ;j} - (W^k_{\ ;j})_{\ ;i} = -R^k_{mij} W^m + T^m_{ij} W^k_{\ ;m}}$$

We see that, sadly enough, that there is a *torsion term* present in the expression, in general. (In "virtually all other sources" interpretation of the expression $W^k_{\ ;i;j}$)

15.6.9 p.411: see the comment to 15.3.4, p.385 (here, in the Additional material).

15.6.10 p.411: We learned how, within the Cartan formalism, the *Gaussian curvarure* K arises through the only independent curvature 2-form

$$\beta = d\alpha = K e^1 \wedge e^2 \equiv K dS$$

dS being the (canonical metric) area element on the surface under consideration. Here we add another useful interpretation of this 2-form in terms of *holonomy.*

Consider a loop γ which is the boundary of a domain S, $\gamma = \partial S$. Fix an orthonormal frame field e_a on S. Then, with respect to the frame field, $\omega_{ab} = \epsilon_{ab}\alpha$. Take a *unit* vector v and perform the parallel transport of the vector around the loop. The parallel transport equations for v read

$$\dot{v}^a + \langle \omega^a_b, \dot{\gamma} \rangle v^b = \dot{v}^a + \epsilon_{ab} \langle \alpha, \dot{\gamma} \rangle v^b = 0$$

or in detail

$$\dot{v}^1 + \langle \alpha, \dot{\gamma} \rangle v^2 = 0$$
$$\dot{v}^2 - \langle \alpha, \dot{\gamma} \rangle v^1 = 0$$

Let $\varphi(t)$ denote the angle between $e_1(t)$ and $v(t)$. Since v is unit vector, we may parameterize the components of v as follows

$$v^1 = \cos\varphi \qquad v^2 = \sin\varphi$$

Then we get a *single* equation for $\varphi(t)$

$$\dot{\varphi} = \langle \alpha, \dot{\gamma} \rangle$$

So, in time dt the increment of $d\varphi$ is

$$d\varphi = \langle \alpha, \dot{\gamma} \rangle dt$$

an the total net angle summed in the course of the loop γ is

$$[\varphi]_{\circlearrowleft} = \int_0^1 \langle \alpha, \dot{\gamma} \rangle dt \equiv \oint_\gamma \alpha = \int_S d\alpha = \int_S K dS$$

The net angle is, however, nothing but the holonomy for the loop. (The group element of the rotation group $SO(2)$ is given in terms of the angle $[\varphi]_{\circlearrowleft}$ modulo 2π.) So, the holonomy is given simply as the *total Gaussian curvature*

$$\boxed{\text{holonomy for } \partial S = \int_S K dS \equiv \text{ total Gaussian curvature (over } S)}$$

A remarkable observation indeed! Take, in particular, *constant* Gaussian curvature K_0. Then

$$\text{holonomy for } \partial S = K_0 \int_S dS = K_0 S$$

So the holonomy is now proportional to the total *area* of the domain S. This is the case if we take the sphere S^2, where (see 15.6.11) $K(x) = K_0 = 1/\rho^2$. We get

$$\text{holonomy for } \partial S = K_0 S = (1/\rho^2) S = (1/\rho^2)\rho^2 \hat{S}_1 \equiv \hat{S}_1$$

where $\hat{S}^1$ denotes the area of the corresponding *unit* sphere, i.e. the *solid angle* subtended by the loop.

$$\boxed{\text{holonomy for loop } \gamma = \textit{solid angle} \text{ subtended by } \gamma}$$

The reader is invited to check this elegant result for the particular case of the Foucault angle [(15.3.10, 15.6.22) - here the solid angle subtended by the *parallel* $\vartheta_0 = \text{const.}$ is $2\pi(1-\cos\vartheta_0)$] as well as for the geodesic triangle in (15.3.9) [here we get 1/8 of the unit sphere's area, i.e. $4\pi/8 = \pi/2$].

15.6.16 p.415: Here, various versions of *Bianchi identity* and *Ricci identity* are discussed.

In part *i*) we are asked to prove their "form version", i.e.

$$d\Omega + \omega\wedge\Omega - \Omega\wedge\omega = 0 \qquad \textit{Bianchi identity}$$
$$dT + \omega\wedge T = \Omega\wedge e \qquad \textit{Ricci identity}$$

▼ This is remarkably easy: Let's apply d on *Cartan structure equations* (15.6.7):

$$de + \omega\wedge e = T \qquad \Rightarrow \qquad d\omega\wedge e - \omega\wedge de = dT$$
$$d\omega + \omega\wedge\omega = \Omega \qquad \Rightarrow \qquad d\omega\wedge\omega - \omega\wedge d\omega = d\Omega$$

Using, on the right, Cartan structure equations themselves, we get

$$(\Omega - \omega\wedge\omega)\wedge e - \omega\wedge(T - \omega\wedge e) = dT$$
$$(\Omega - \omega\wedge\omega)\wedge\omega - \omega\wedge(\Omega - \omega\wedge\omega) = d\Omega$$

or, after some canceling,

$$\Omega\wedge e - \omega\wedge T = dT$$
$$\Omega\wedge\omega - \omega\wedge\Omega = d\Omega$$

which is exactly what we need. ▲

In part *iv*) interesting point is *coordinate* (index) version of both identities for *RLC*-connection (used standardly in general relativity), i.e.

$$d\Omega + \omega\wedge\Omega - \Omega\wedge\omega = 0 \qquad \Leftrightarrow \qquad R^i{}_{j[kl;m]} = 0 \qquad \text{Bianchi identity}$$
$$\Omega\wedge e = 0 \qquad \Leftrightarrow \qquad R^i{}_{[jkl]} = 0 \qquad \text{Ricci identity}$$

▼ We want to derive coordinate (index) versions of the identities from the corresponding form versions. Let's express the forms w.r.t. coordinate coframes (15.6.1), (15.6.3):

$$e^i = dx^i \qquad \omega^i_j = \Gamma^i_{jm}dx^m \qquad \Omega^i_j = \frac{1}{2}R^i{}_{jkl}dx^k\wedge dx^l$$

Then

$$d\Omega^i_j = \frac{1}{2}R^i{}_{jkl,m}dx^k\wedge dx^l\wedge dx^m$$
$$\omega^i_a\wedge\Omega^a_j = \frac{1}{2}\Gamma^i_{am}R^a{}_{jkl}dx^k\wedge dx^l\wedge dx^m$$
$$-\Omega^i_a\wedge\omega^a_j = -\frac{1}{2}\Gamma^a_{jm}R^i{}_{akl}dx^k\wedge dx^l\wedge dx^m$$

and therefore, summing all three pieces, we get

$$(d\Omega + \omega\wedge\Omega - \Omega\wedge\omega)^i_j = \frac{1}{2}(R^i{}_{jkl,m} + \Gamma^i_{am}R^a{}_{jkl} - \Gamma^a_{jm}R^i{}_{akl})dx^k\wedge dx^l\wedge dx^m$$

Now comparing the terms in the bracket with

$$R^i{}_{jkl;m} \equiv R^i{}_{jkl,m} + \Gamma^i_{am}R^a{}_{jkl} - \Gamma^a_{jm}R^i{}_{akl} - \Gamma^a_{km}R^i{}_{jal} - \Gamma^a_{lm}R^i{}_{jka}$$

we see that two terms are missing in the bracket to become the complete covariant derivative $R^i{}_{jkl;m}$. However, combining of the two missing terms with $dx^k \wedge dx^l \wedge dx^m$ actually gives *zero*

$$(-\Gamma^a_{km} R^i{}_{jal} - \Gamma^a_{lm} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m = 0$$

because of the *symmetry* of Christoffel symbols w.r.t. lower indices! So we can freely add them to the bracket with no consequence and get

$$(d\Omega + \omega \wedge \Omega - \Omega \wedge \omega)^i_j = \frac{1}{2} R^i{}_{jkl;m} dx^k \wedge dx^l \wedge dx^m$$

Then, vanishing of the l.h.s. (i.e. the form version of Bianchi idendity) may be also expressed as

$$R^i{}_{jkl;m} dx^k \wedge dx^l \wedge dx^m = 0$$

or, finally, as

$$R^i{}_{j[kl;m]} = 0$$

which is exactly what we need.

Ricci identity is even simpler: we have

$$(\Omega \wedge e)^i = \Omega^i_m \wedge dx^m = \frac{1}{2} R^i{}_{jkl} dx^k \wedge dx^l \wedge dx^m$$

Then vanishing of the expression (i.e. Ricci idendity) may be also expressed as

$$R^i{}_{jkl} dx^k \wedge dx^l \wedge dx^m = 0$$

or, finally, as

$$R^i{}_{[jkl]} = 0$$

which is, again, exactly what we need. ▲

In part *iii*) *coordinate* (index) version of both identities for a *general* linear connection (so non-vanishing *torsion* may be present) is found to be

$d\Omega + \omega \wedge \Omega - \Omega \wedge \omega = 0$	$\Leftrightarrow$	$R^i{}_{j[kl;m]} + R^i{}_{jr[k} T^r_{lm]} = 0$	Bianchi identity
$\Omega \wedge e - (dT + \omega \wedge T) = 0$	$\Leftrightarrow$	$R^i{}_{[jkl]} = T^i_{[jk;l]} + T^i_{r[j} T^r_{kl]}$	Ricci identity

▼ Let's show the difference (w.r.t. part *iv*) discussed above) on Bianchi identity (Ricci identity is then proved in a similar way). So, we proceed as it is discussed in part *iv*) above. There we come to the point where two terms are missing in the bracket to become the complete covariant derivative $R^i{}_{jkl;m}$. Combining them with $dx^k \wedge dx^l \wedge dx^m$ actually gave zero

$$(-\Gamma^a_{km} R^i{}_{jal} - \Gamma^a_{lm} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m = 0$$

because of the *symmetry* of Christoffel symbols w.r.t. lower indices. Now the difference starts, since the symmetry is no longer true. Rather, we have

$$\Gamma^i_{jk} = \Gamma^i_{(jk)} + \Gamma^i_{[jk]} = \Gamma^i_{(jk)} - \frac{1}{2} T^i_{jk} \qquad \text{(see (15.3.3))}$$

Therefore, what we get now is

$$\begin{aligned}
(\Gamma^a_{km} R^i{}_{jal} + \Gamma^a_{lm} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m &= (\Gamma^a_{[km]} R^i{}_{jal} + \Gamma^a_{[lm]} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m \\
&= -\frac{1}{2} (T^a_{km} R^i{}_{jal} + T^a_{lm} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m \\
&= -\frac{1}{2} (R^i{}_{ja[l} T^a_{km]} - R^i{}_{ja[k} T^a_{lm]}) dx^k \wedge dx^l \wedge dx^m \\
&= \frac{1}{2} (R^i{}_{ja[k} T^a_{lm]} + R^i{}_{ja[k} T^a_{lm]}) dx^k \wedge dx^l \wedge dx^m \\
&= R^i{}_{ja[k} T^a_{lm]}) dx^k \wedge dx^l \wedge dx^m
\end{aligned}$$

Therefore we can write, now,

$$(d\Omega + \omega \wedge \Omega - \Omega \wedge \omega)^i_j = \frac{1}{2}(R^i{}_{jkl;m} + \Gamma^a_{km} R^i{}_{jal} + \Gamma^a_{lm} R^i{}_{jka}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(R^i{}_{jkl;m} + R^i{}_{ja[k} T^a_{lm]}) dx^k \wedge dx^l \wedge dx^m$$

and, consequently,

$$R^i{}_{j[kl;m]} + R^i{}_{ja[k} T^a_{lm]} = 0$$

For Ricci identity, the new part (namely $dT + \omega \wedge T$) reads

$$(dT + \omega \wedge T)^i = dT^i + \omega^i_j \wedge T^j \qquad\qquad T^i = \frac{1}{2} T^i_{kl} dx^k \wedge dx^l$$
$$= \frac{1}{2}(T^i_{kl,m} + \Gamma^i_{jm} T^j_{kl}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(T^i_{kl;m} + \Gamma^a_{km} T^i_{al} + \Gamma^a_{lm} T^i_{ka}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(T^i_{kl;m} + \Gamma^a_{[km]} T^i_{al} + \Gamma^a_{[lm]} T^i_{ka}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(T^i_{kl;m} - T^a_{km} T^i_{al}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(T^i_{kl;m} - T^i_{al} T^a_{km}) dx^k \wedge dx^l \wedge dx^m$$
$$= \frac{1}{2}(T^i_{[kl;m]} + T^i_{a[k} T^a_{lm]}) dx^k \wedge dx^l \wedge dx^m$$

Then, because of still valid fact

$$(\Omega \wedge e)^i = \Omega^i_m \wedge dx^m = \frac{1}{2} R^i{}_{[klm]} dx^k \wedge dx^l \wedge dx^m$$

we come to

$$R^i{}_{[klm]} = T^i_{[kl;m]} + T^i_{a[k} T^a_{lm]}$$

which is, as usually, exactly what we need. ▲

15.6.20 p.417: see the comment to 15.3.4, p.385 (here, in the Additional material).

16.4.6 p.453: Here we learned that a term of the form

$$\langle \alpha, \alpha \rangle$$

in an action leads to energy-momentum tensor

$$T_{ab} = 2(i_a \alpha, i_b \alpha) - g_{ab}(\alpha, \alpha)$$

Now, let us concentrate on a rather specific property of the latter, namely the *sign* of *energy density* corresponding to a general $\langle \alpha, \alpha \rangle$ term, i.e. the sign of T_{00} component of the energy-momentum tensor. From the general formula we get, in particular

$$T_{00} = 2(i_0 \alpha, i_0 \alpha) - (\alpha, \alpha)$$

In order to decide whether this is positive or negative definite (or perhaps indefinite), the results of (16.3.6) come in handy. The problem teaches us that for

$$\alpha = dt \wedge \hat{s} + \hat{r}$$

we get

$$(\alpha, \alpha)_{E^{1,3}} = (\hat{\eta} \hat{s}, \hat{s})_{E^3} + (\hat{\eta} \hat{r}, \hat{r})_{E^3}$$

Recall that both expressions (scalar products) on the right-hand side of the expression for T_{00} are of $E^{1,3}$-type. We need, using the formula above, to express them in terms of E^3-type scalar products. (The reason is, that

$E^{1,3}$-type scalar product is indefinite, whereas E^3-type scalar product is *positive definite*: $(\hat{s},\hat{s})_{E^3} \geq 0$ for any $\hat{s}$.) So, we need to express $(i_0\alpha, i_0\alpha)$ in terms of E^3-type scalar products. Now,

$$\alpha = dt \wedge \hat{s} + \hat{r} \qquad \Rightarrow \qquad i_0\alpha = \hat{s}$$

and so

$$(i_0\alpha, i_0\alpha)_{E^{1,3}} = (\hat{\eta}\hat{s}, \hat{s})_{E^3}$$

Therefore,

$$\begin{aligned} T_{00} &= 2(i_0\alpha, i_0\alpha)_{E^{1,3}} - (\alpha,\alpha)_{E^{1,3}} \\ &= 2(\hat{\eta}\hat{s},\hat{s})_{E^3} - ((\hat{\eta}\hat{s},\hat{s})_{E^3} + (\hat{\eta}\hat{r},\hat{r})_{E^3}) \\ &= (\hat{\eta}\hat{s},\hat{s})_{E^3} - (\hat{\eta}\hat{r},\hat{r})_{E^3} \end{aligned}$$

The last formula clearly displays *all we need.*[5] Let the degree of α be even. Then $\hat{\eta}\hat{r} = \hat{r}$, $\hat{\eta}\hat{s} = -\hat{s}$ and so

$$T_{00} = -(\hat{s},\hat{s})_{E^3} - (\hat{r},\hat{r})_{E^3} \leq 0$$

On the contrary, for the degree of α being odd, $\hat{\eta}\hat{r} = -\hat{r}$, $\hat{\eta}\hat{s} = \hat{s}$ and

$$T_{00} = (\hat{s},\hat{s})_{E^3} + (\hat{r},\hat{r})_{E^3} \geq 0$$

So, altogether, we see the following simple picture:

degree of α is even	$T_{00} \leq 0$
degree of α is odd	$T_{00} \geq 0$

Since physics only likes the case $T_{00} \geq 0$, the result above is to be understood as follows: the term $\langle\alpha,\alpha\rangle$ should enter the action with *plus* sign iff α has odd degree and with *minus* sign iff α has even degree:

degree of α is even	$-\langle\alpha,\alpha\rangle$ in action integral
degree of α is odd	$+\langle\alpha,\alpha\rangle$ in action integral

This concise rule is indeed satisfied by all field theory action integrals mentioned in the book. Check, in particular, (each individual term in) the following cases:

free (massive) scalar field	$S = \frac{1}{2}\langle d\phi, d\phi\rangle - \frac{1}{2}m^2\langle\phi,\phi\rangle$	(16.3.7)
free (massive) vector field	$S = -\frac{1}{2}\langle dW, dW\rangle + \frac{1}{2}m^2\langle W,W\rangle$	(16.3.8)
free electromagnetic field	$S = -\frac{1}{2}\langle dA, dA\rangle$	(16.3.2)

16.5.3 p.460: Here, the *Hilbert action* is discussed, and in particular its form in *two-dimensional* case (which is exceptional) is shown.

And what about the *three-dimensional* case? Well, in orthonormal frame we have

$$\Omega_{ab} \wedge *(e^a \wedge e^b) = \Omega_{ab} \wedge \omega^{ab}{}_c e^c = \Omega^{ab} \wedge \omega_{abc} e^c = \epsilon_{cab}\Omega^{ab} \wedge e^c = \Omega_a \wedge e^a$$

so that

$$R_g \omega_g = \Omega_a \wedge e^a$$

where we used, in 3D, a possibility to express *curvature forms* in terms of three two-forms

$$\Omega_a := \epsilon_{abc}\Omega^{bc}$$

[5]Except for *love*, of course. Sorry, John/Paul.

(*not* to be confused with *Ricci 1-forms* $R_a \equiv i_b \Omega^b_a = R_{ab} e^b$, see 16.5.1).

18.4.6 p.511: see 15.4.4, p.392 (in this text)

18.4.10 p.514 The result

$$\mathcal{L}_V A = d\chi$$

might look unnatural (one should expect $\mathcal{L}_V A = 0$ as the "invariance condition" for A). Note, however, that according to 18.4.9(iv) there is a ("gauge") *freedom*

$$A \mapsto A' = A + df$$

in choosing A. So, if some particular A obeys $\mathcal{L}_V A = d\chi$, then the "new" one, A' obeys

$$\mathcal{L}_V A' = d(\chi - Vf)$$

So when we choose f such that

$$Vf = \chi$$

(in coordinates, where V straightens out, we are to solve the equation $\partial f / \partial x^1 = \chi$) we get

$$\mathcal{L}_V A' = 0$$

The moral is that we cannot ask "nice" symmetry condition $\mathcal{L}_V A = 0$ being valid *for all* A's from the equivalence class $A \sim A + df$, since the transformation $A \mapsto A + df$ *spoils* it. In general, just the weaker condition $\mathcal{L}_V A = d\chi$ holds and only for a *carefully chosen* representative A this simplifies to $\mathcal{L}_V A = 0$. (Since there is no comparable freedom in g and ϕ, we have always "nice" invariance conditions $\mathcal{L}_V g = 0$ and $\mathcal{L}_V \phi = 0$ respectively for those fields.)

19.3.10 p.534. Just prior to this problem the following formulation of *Frobenius integrability criterion* is presented without proof: In terms of constraint 1-forms θ^i, $\mathcal{D}$ is integrable if and only if there are $(n-k)^2$ 1-forms σ^i_j such that $d\theta^i = \sigma^i_j \wedge \theta^j$, i.e.

$$\mathcal{D} \text{ is integrable} \qquad \Leftrightarrow \qquad \{\exists \sigma^i_j : d\theta^i = \sigma^i_j \wedge \theta^j\}$$

Let us have a look how this particular formulation follows from the one mentioned after 19.3.7:

$$\mathcal{D} \text{ is integrable} \quad \Leftrightarrow \quad \{\theta^i\big|_{\mathcal{D}} = 0 \ \Rightarrow \ d\theta^i\big|_{\mathcal{D}} = 0\} \ \text{ i.e. } \ \{U, V \in \mathcal{D} \ \Rightarrow \ d\theta^i(U, V) = 0\}$$

Well, consider a co-frame (e^i, e^a) *adapted to* $\mathcal{D}$, i.e. $e^i := \theta^i$ and e^a realize a completion to a co-frame. Also consider the *dual frame* (e_i, e_a). So, $\mathcal{D} = \text{Span}\ \{e_a\}$.

Then, since $d\theta^i$ is a 2-form, we can decompose it w.r.t. (e^i, e^a) as follows:

$$d\theta^i = A^i_{kj} e^k \wedge e^j + B^i_{aj} e^a \wedge e^j + C^i_{ab} e^a \wedge e^b \qquad\qquad A^i_{kj}, \ B^i_{aj}, \ C^i_{ab} \ \text{ functions}$$

Now, the previous formulation says (for $U = e_a, V = e_b$)

$$\mathcal{D} \text{ is integrable} \quad \Leftrightarrow \quad d\theta^i(e_a, e_b) = 0$$

But

$$d\theta^i(e_a, e_b) = 2C^i_{ab}$$

So integrabilty *kills* C^i_{ab} (and vice versa) and we have

$$\mathcal{D} \text{ is integrable} \quad \Leftrightarrow \quad d\theta^i = A^i_{kj} e^k \wedge e^j + B^i_{aj} e^a \wedge e^j \equiv (A^i_{kj} e^k + B^i_{aj} e^a) \wedge e^j$$

So, defining

$$A^i_{kj} e^k + B^i_{aj} e^a =: \sigma^i_j$$

(and recalling that $e^j := \theta^j$) we finally get

$$\mathcal{D} \text{ is integrable} \quad \Leftrightarrow \quad d\theta^i = \sigma^i_j \wedge \theta^j$$

19.3.12 p.535. In part (iii) we study, as an example to integrability conditions

$$f^i_{[\alpha,\beta]} = f^j_{[\alpha} f^i_{\beta],j} \tag{a}$$

for system of partial differential equations

$$y^i{}_{,\alpha} = f^i_\alpha(x,y) \qquad\qquad x^\alpha \text{ independent}, \ y^i \text{ dependent}$$

the following simple system

$$\frac{\partial f}{\partial x} = f \sin y \tag{b1}$$
$$\frac{\partial f}{\partial y} = \lambda f x \cos y \tag{b2}$$

(So that $y^1 \equiv f$, $(x^1, x^2) \equiv (x, y)$, $f^1_1 \equiv f \sin y$, $f^1_2 \equiv \lambda f x \cos y$.)

Let's have a look on the system in elementary way, without using (a).

First, notice that we have a solution $f = 0$ for *any* λ. (So, the statement in the book that the system has a solution only for $\lambda = 1$ is indeed wrong, see Errata.)

Now, imagine $f \neq 0$. Then, for $F \equiv \ln f$, we get slightly simpler equations

$$\frac{\partial F}{\partial x} = \sin y \tag{c1}$$
$$\frac{\partial F}{\partial y} = \lambda x \cos y \tag{c2}$$

So, if $F(x,y)$ exists, its dF reads

$$dF = \sin y \ dx + \lambda x \cos y \ dy$$

Then, necessarily,

$$0 = ddF = d(\sin y \ dx + \lambda x \cos y \ dy) = (\lambda - 1) \ \cos y \ dx \wedge dy$$

Therefore we inevitably need $\lambda = 1$. If $\lambda = 1$ holds, we easily get from (c1) and (c2)

$$F(x,y) = x \ \sin y + A(y) = x \ \sin y + B(x)$$

so that $A = B = k =$ const. and finally (only for $\lambda = 1$!)

$$F(x,y) = x \ \sin y + k \qquad \text{i.e.} \qquad f(x,y) = e^F = K \ e^{x \sin y}$$

If we want to use criterion (a), it reads, here (check!)

$$(1 - \lambda) f \cos y = 0 \tag{d}$$

for each (x, y, f). This is equivalent to $\lambda = 1$.

19.4.4 p.540 Correct version of the expression of H_i reads

$$H_i \equiv \partial^h_i := \partial_i - \langle \hat\omega^a_b, \partial_i \rangle y^b_c \partial^c_a \equiv \partial_i - \langle \omega^a_b, \partial_i \rangle \xi_{E^b_a}$$

Notice the *hat symbol* on the first ω^a_b - it is very important. Indeed, according to 19.1.4

$$\xi_{E^b_a} = y^c_a \partial^b_c$$

so that the above equality *cannot* be true with two unhatted ω^a_b (the summation within the $y\partial$ part is different). However, the relation between ω^a_b and $\hat\omega^a_b$ is (see 19.2.1)

$$\omega^a_b := (y^{-1})^a_c (\pi^* \hat\omega^c_d) y^d_b + (y^{-1})^a_c dy^c_b$$

so that

$$\langle \omega^a_b, \partial_i \rangle = (y^{-1})^a_c \langle \hat{\omega}^c_d, \partial_i \rangle y^d_b$$

Then,

$$\begin{aligned} H_i &= \partial_i - \langle \omega^a_b, \partial_i \rangle \xi_{E^b_a} \\ &= \partial_i - \langle \omega^a_b, \partial_i \rangle y^c_a \partial^b_c \\ &= \partial_i - (y^{-1})^a_r \langle \hat{\omega}^r_d, \partial_i \rangle y^d_b y^c_a \partial^b_c \\ &= \partial_i - \langle \hat{\omega}^a_b, \partial_i \rangle y^b_c \partial^c_a \end{aligned}$$

or, consequently

$$\langle \omega^a_b, \partial_i \rangle y^c_a \partial^b_c = \langle \hat{\omega}^a_b, \partial_i \rangle y^b_c \partial^c_a$$

We see that the replacement $\omega^a_b \mapsto \hat{\omega}^a_b$ indeed produces the change in summation in the $y\partial$ expression.

$\boxed{19.6.1-3}$ p.545-547: Here we present *coordinate* versions (in coordinates (x, y) introduced in 19.1.1) of the "abstract" functions $\Phi^B : LM \to (V, \rho)$.

Let's start with a vector field. We have, then, equivariant $\Phi^v : LM \to (\mathbb{R}^n, \rho^1_0)$, i.e. in coordinates some

$$\hat{v}^a(x, y) \qquad \text{obeying} \qquad \hat{v}^a(x, yA) = (A^{-1})^a_b \hat{v}^b(x, y)$$

In particular, for $y = \mathbb{I}$ (the *unit* matrix), we get $\hat{v}^a(x, A) = (A^{-1})^a_b \hat{v}^b(x, \mathbb{I})$, i.e.

$$\hat{v}^a(x, y) = (y^{-1})^a_b V^b(x) \qquad V^b(x) := \hat{v}^b(x, \mathbb{I})$$

So we see that the *family* of functions $\hat{v}^a(x, y)$ obeying $\hat{v}^a(x, yA) = (A^{-1})^a_b \hat{v}^b(x, y)$ is *parametrized* by functions $V^a(x)$ of *variables x alone.* (There is already *no* freedom in the dependance on variables y.)

What is the meaning of those functions $V^a(x)$? Well, recall that the space $(\mathbb{R}^n, \rho^1_0)$ is just the space of *components* of vectors, so (think it over) $V^a(x)$ are just components

- of the vector field $V \equiv V^a(x) e_a \leftrightarrow \Phi^v \leftrightarrow \hat{v}^a(x, y)$
- w.r.t. the frame $e_a \leftrightarrow y = \mathbb{I}$, i.e.
- w.r.t. the frame $e_a \leftrightarrow$ the coordinates y^a_b (see 19.1.1)

In a similar way we get, as an example, the following functions on LM:

$$\hat{\alpha}_a(x, y) = y^b_a \alpha_b(x) \qquad \text{satisfying} \qquad \hat{\alpha}_a(x, yA) = A^b_a \hat{\alpha}_b(x, y)$$

corresponding to a *covector* field $\alpha = \alpha_a(x) e^a$,

$$\hat{g}_{ab}(x, y) = y^c_a y^d_b g_{cd}(x) \qquad \text{satisfying} \qquad \hat{g}_{ab}(x, yA) = A^c_a A^d_b \hat{g}_{cd}(x, y)$$

corresponding to *metric tensor* field $g = g_{ab}(x) e^a \otimes e^b$,

$$\hat{t}^c_{ab}(x, y) = (y^{-1})^c_i y^j_a y^k_b T^i_{jk}(x) \qquad \text{satisfying} \qquad \hat{t}^c_{ab}(x, yA) = (A^{-1})^c_i A^j_a A^k_b \hat{t}^i_{jk}(x, y)$$

corresponding to *torsion tensor* field $T = T^c_{ab}(x) e^a \otimes e^b \otimes e_c$ or

$$\hat{r}^d_{abc}(x, y) = (y^{-1})^d_i y^j_a y^k_b y^l_c R^i_{jkl}(x) \qquad \text{satisfying} \qquad \hat{r}^d_{abc}(x, yA) = (A^{-1})^d_i A^j_a A^k_b A^l_c \hat{r}^i_{jkl}(x, y)$$

corresponding to *Riemann curvature tensor* field $R = R^d_{abc}(x) e^a \otimes e^b \otimes e^c \otimes e_d$.

$\boxed{19.6.6}$ p.549: Here we speak of an equivariant function Φ which is defined on the fibers over the curve $\gamma(t)$ and which corresponds to field of quantities of type ρ (e.g. a tensor field B) on γ. We are to check that its derivative along the horizontal lift γ^h of the curve γ corresponds (just in the sense in that Φ corresponds to B) to the *covariant* derivative $\nabla_{\dot{\gamma}} B$ of the field B

$$\Phi \leftrightarrow B \qquad \Rightarrow \qquad (\dot{\gamma})^h \Phi \leftrightarrow \nabla_{\dot{\gamma}} B$$

Let us check, first, that if Φ is of type ρ, the same holds for the derivative $(\dot{\gamma})^h\Phi$.

Recall (see 19.5.2 *i*)) that the horizontal lift γ^h_e (i.e. the lift from x to e over x) is uniquely given by

$$\pi \circ \gamma^h_e = \gamma \qquad \gamma^h_e(0) = e \qquad \langle \omega, (\dot{\gamma^h_e}) \rangle = 0$$

Now it is useful to notice that it is $GL(n,\mathbb{R})$-*invariant* in the sense

$$R_A \circ \gamma^h_e = \gamma^h_{eA}$$

(so that R_A-image of the lift starting at e is a lift again, starting at eA).

▼ Indeed, let us denote $\Gamma := R_A \circ \gamma^h_e$. It is clearly a curve on LM enjoying the following properties:

$$\pi \circ \Gamma = \gamma \qquad \Gamma(0) = eA \qquad \langle \omega, \dot{\Gamma} \rangle = 0$$

(We used $\pi \circ R_A = \pi$ for the first property and $\dot{\Gamma} = R_{A*}(\dot{\gamma^h_e})$ as well as $R^*_A\omega = A^{-1}\omega A$ for the last one.) The three properties define, however, exactly the curve γ^h_{eA}. ▲

Now, define (again on the fibers over the curve $\gamma(t)$) the function

$$\chi := (\dot{\gamma^h})\Phi \qquad \text{i.e.} \qquad \chi(e) := (\dot{\gamma^h_e})\Phi$$

(So, at *each* point e of the fiber over x, *the corresponding* lift of $\dot{\gamma}$ is applied on Φ.)

Then

$$\chi(eA) = (\dot{\gamma^h_{eA}})\Phi = R_{A*}(\dot{\gamma^h_e})\Phi = \langle d\Phi, R_{A*}(\dot{\gamma^h_e}) \rangle = \langle d(R^*_A\Phi), (\dot{\gamma^h_e}) \rangle = \rho(A^{-1})\langle d\Phi, (\dot{\gamma^h_e}) \rangle = \rho(A^{-1})\chi(e)$$

So $\chi \equiv (\dot{\gamma})^h\Phi$ *is* an equivariant function (of type ρ) as well, i.e. it indeed corresponds to a quantity "of the same type" as B does.

Which one? The derivative of Φ along γ^h corresponds to

- derivative of *components* (since the space (V,ρ) of values of Φ is the space of components of B)
- of the quantity B (since $\Phi \leftrightarrow B$)
- w.r.t. *autoparallel* frame (since γ^h corresponds to autoparallel frames).

And this corresponds (see 19.6.4) to *covariant* derivative $\nabla_{\dot{\gamma}}B$ of B.

20.2.5 p.562: Here we learn that two crucial properties of connection form read

$$R^*_g\omega = \mathrm{Ad}_{g^{-1}}\omega \qquad \langle \omega, \xi_X \rangle = X$$

In 20.2.6 we add that *any* ω possessing the two properties is a connection form. And then, in 20.2.7, we also came to infinitesimal version of the former property, namely

$$\mathcal{L}_{\xi_X}\omega = -[X,\omega]$$

Here, in the Additional material, we just mention that the two properties may also be equivalently rewritten into the following convenient form:

$$\boxed{\mathcal{L}_{\xi_i}\omega^j = -c^j_{ik}\omega^k} \qquad \boxed{\langle \omega^i, \xi_j \rangle = \delta^i_j} \tag{a}$$

where $\xi_j \equiv \xi_{E_j}$ are generators of the action R_g and c^j_{ik} are the structure constants of the Lie algebra w.r.t. E_i (i.e. $[E_i, E_k] = c^j_{ik}E_j$).

▼ Indeed, set $X = E_i$; then

$$\mathcal{L}_{\xi_{E_i}}\omega = (\mathcal{L}_{\xi_i}\omega^j)E_j \stackrel{!}{=} -[E_i,\omega] = -[E_i,\omega^k E_k] = -[E_i,E_k]\omega^k = -c^j_{ik}\omega^k E_j$$

$$\langle \omega, \xi_i \rangle = \langle \omega^j, \xi_i \rangle E_j \stackrel{!}{=} E_i$$

▲

Example: consider a $GA(1,\mathbb{R})$-principal bundle $\pi : P \to M$ and let ω be a connection form. Since a possible basis E_1, E_2 in the Lie algebra $ga(1,\mathbb{R})$ is

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

(see 11.7.10 - notice however an error mentioned in Errata), any connection form may be written as

$$\omega = \omega^j E_j = \omega^1 E_1 + \omega^2 E_2 = \begin{pmatrix} \omega^1 & \omega^2 \\ 0 & 0 \end{pmatrix}$$

The right translation on $GA(1,\mathbb{R})$ is

$$R_{(a,b)}(u,v) = (u,v) \circ (a,b) = (ua, ab + a)$$

(see 10.2.6). One easily finds the corresponding generators:

$$\xi_1 = u\partial_u \qquad \xi_2 = u\partial_v \tag{b}$$

So, if we use a local trivialization of the bundle, we have $P \sim \mathcal{O} \times GA(1,\mathbb{R})$ with coordinates (x^μ, u, v), where x^μ are local coordinates on $\mathcal{O} \subset M$, and the generators of the action of $GA(1,\mathbb{R})$ on (the corresponding part of) P look exactly like those in (b).

Therefore if we want to find the most general connection form on this particular principal bundle, we can start with the most general expressions

$$\begin{aligned} \omega^1 &= A_\mu(x,u,v)dx^\mu + B(x,u,v)du + C(x,u,v)dv \\ \omega^2 &= \tilde{A}_\mu(x,u,v)dx^\mu + \tilde{B}(x,u,v)du + \tilde{C}(x,u,v)dv \end{aligned}$$

and then restrict to those of them, which satisfy (a), i.e., in our particular case,

$$\begin{aligned} \mathcal{L}_{\xi_1}\omega^1 &= 0 & \langle \omega^1, \xi_1 \rangle &= 1 \\ \mathcal{L}_{\xi_1}\omega^2 &= -\omega^2 & \langle \omega^1, \xi_2 \rangle &= 0 \\ \mathcal{L}_{\xi_2}\omega^1 &= 0 & \langle \omega^2, \xi_1 \rangle &= 0 \\ \mathcal{L}_{\xi_2}\omega^2 &= \omega^1 & \langle \omega^2, \xi_2 \rangle &= 1 \end{aligned}$$

The right ("algebraic") column alone quickly restricts ω^1 and ω^2 to

$$\begin{aligned} \omega^1 &= \frac{du}{u} + A_\mu(x,u,v)dx^\mu \\ \omega^2 &= \frac{dv}{u} + \tilde{A}_\mu(x,u,v)dx^\mu \end{aligned}$$

The left ("differential") column then restricts $A_\mu(x,u,v)$ and $\tilde{A}_\mu(x,u,v)$ further to

$$A_\mu(x,u,v) = f_\mu(x) \qquad \tilde{A}_\mu(x,u,v) = \frac{1}{u}(vf_\mu(x) + g_\mu(x))$$

(check :-). So, the most general ω^1 and ω^2 are

$$\begin{aligned} \omega^1 &= \frac{du}{u} + f_\mu(x)dx^\mu \\ \omega^2 &= \frac{dv}{u} + \frac{1}{u}(vf_\mu(x) + g_\mu(x))dx^\mu \end{aligned}$$

i.e. the most general connection form looks (locally) as follows:

$$\omega = \begin{pmatrix} \omega^1 & \omega^2 \\ 0 & 0 \end{pmatrix} = \frac{1}{u}\begin{pmatrix} du + uf_\mu(x)dx^\mu & dv + (vf_\mu(x) + g_\mu(x))dx^\mu \\ 0 & 0 \end{pmatrix} \tag{c}$$

It is parameterized by *arbitrary* functions $f_\mu(x)$ and $g_\mu(x)$. For example, for all of them vanishing we get

$$\omega = \begin{pmatrix} \omega^1 & \omega^2 \\ 0 & 0 \end{pmatrix} = \frac{1}{u}\begin{pmatrix} du & dv \\ 0 & 0 \end{pmatrix}$$

Locally we can use "canonical" section

$$\sigma : x^\mu \mapsto (x^\mu, u(x), v(x)) = (x^\mu, 1, 0)$$

(abstractly $m \mapsto (m, g(m)) = (m, e)$, where e is the unit element of the group; see the proof of 20.1.3). In this particular gauge, the *gauge potential* reads

$$A \equiv \sigma^*\omega = \begin{pmatrix} A^1 & A^2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} f_\mu(x)dx^\mu & g_\mu(x)dx^\mu \\ 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} A^1_\mu(x)dx^\mu & A^2_\mu(x)dx^\mu \\ 0 & 0 \end{pmatrix}$$

So, the functions $f_\mu(x)$ and $g_\mu(x)$ parameterizing the general expression (c) may be identified with gauge potentials in the canonical gauge. In particular, the example with vanishing $f_\mu(x)$ and $g_\mu(x)$ corresponds to $A = 0$ (in the particular gauge) and, therefore, to $F \equiv dA + A \wedge A = 0$. This means (due to the fact, that $F = 0$ is already gauge-independent statement), that the example corresponds to *flat* connection ($\Omega = 0$).

21.3.3 p.602: A useful information about *holonomy* may be gained here.

We learned here, that the quantity of type ρ_n (with respect to the group $U(1)$) gets multiplied, when parallel transported along γ, by the factor $e^{-in\int_\gamma A}$

$$z(0) \mapsto z(1) = e^{-in\int_\gamma A} z(0)$$

Let γ be a *loop*, now. Then $\gamma(0) = \gamma(1)$ and the holonomy associated with the loop is the group element $g \in U(1)$ such that

$$\gamma^h(1) = R_g \gamma^h(0) \equiv \gamma^h(0) g$$

We assumed z to be a quantity of type ρ_n. Therefore it behaves in general as follows

$$z(pg) = \rho_n(g^{-1}) z(p)$$

($R_g^*\Phi = \rho(g^{-1})\Phi$). In our particular case

$$p = \gamma^h(0) \qquad g = e^{i\alpha} \text{ (= unknown, wanted, the holonomy)} \qquad \rho_n(e^{i\alpha}) = e^{in\alpha}$$

and also

$$z(0) = z(\gamma^h(0))$$

Therefore

$$\begin{aligned} z(1) \equiv z(\gamma^h(1)) &\overset{1.}{=} z(\gamma^h(0)e^{i\alpha}) = e^{-in\alpha} z(\gamma^h(0)) = e^{-in\alpha} z(0) \\ &\overset{2.}{=} e^{-in\oint_\gamma A} z(0) \end{aligned}$$

whence

$$\text{the } \textit{holonomy} \text{ for } \gamma \;\equiv\; g \;\equiv\; e^{i\alpha} \;=\; e^{i\oint_\gamma A}$$

If γ is the *boundary* of a 2-dimensional surface S, then $\oint_\gamma A = \int_S F$ and so, alternatively, we can write down the holonomy in terms of the *curvature* 2-form

$$\text{the } \textit{holonomy} \text{ for } \gamma \;\equiv\; g \;\equiv\; e^{i\alpha} \;=\; e^{i\int_S F}$$

Particularly interesting is the case of a bundle over *two*-dimensional Riemannian base manifold (M, g) (a surface) such that the curvature form coincides (possibly up to a constant multiple) with the metric area (= volume) form dS

$$F = dS$$

Then, clearly

$$\int_S F = \text{ the } \textit{area} \text{ of } S$$

so that

$$\text{the } \textit{holonomy} \text{ for } \partial S \;=\; e^{i\ (\text{area of } S)}$$

[This is to be compared with the additional text to 15.6.10.]

To give a concrete (and important)[6] example, consider the *Hopf bundle* $\pi : S^3 \to S^2$ studied in 20.1.7-10. We can introduce a $U(1)$-connection by defining the subspace spanned by *left-invariant* fields (e_1, e_2) on $SU(2)$ (see 11.7.23) to be the horizontal subspace $\text{Hor}_{\,p}P$ (whereas the span of e_3 is the vertical one $\text{Ver}_{\,p}P$, see the figure in 20.1.10; this means that the horizontal subspace is defined as the *orthogonal* complement, w.r.t. the Killing metric, to the vertical one). Alternatively (check), we can fix the connection via the connection form

$$\omega := ie^3$$

(it is to be $u(1) \equiv i\mathbb{R}$-valued, that's why the i). Then

$$\Omega = d\omega = ide^3 = i(-\sin\vartheta d\vartheta \wedge d\varphi)$$

and so

$$iF \equiv \sigma^*\Omega = i(-\sin\vartheta d\vartheta \wedge d\varphi)$$

(for *any* σ, check) so that

$$F = -\sin\vartheta d\vartheta \wedge d\varphi \equiv -dS$$

where dS is the standard round area form on S^2. Therefore

$$\boxed{\text{the } \textit{holonomy} \text{ for } \partial S \;=\; e^{-i\ (\text{area of } S)} \;=\; e^{-i\ (\textit{solid angle} \text{ subtended by } \gamma)}}$$

22.5.4 p.665: In the 4-th expression of the Dirac operator, Γ_{abc} in the 2-nd expression was replaced by the expression

$$\Gamma_{abc} = -\frac{1}{2}(c_{abc} + c_{bca} - c_{cab})$$

from 15.6.20iii. From *Additional material* to 15.6.20iii we know, however, that the full expression reads

$$\Gamma_{abc} = -\frac{1}{2}(c_{abc} + c_{bca} - c_{cab}) + K_{abc}$$

where

$$K_{abc} := -\frac{1}{2}(T_{abc} + T_{bca} - T_{cab})$$

is the *contortion tensor*. So, the contortion term is to be included also in the Dirac operator, in general.

Recall that *spin* connection corresponds to a *metric*, but *not necessarily symmetric* connection on (M, g), see the book, p.658 in Sec.22.4. Therefore, whenever we need covariantly differentiate spinor fields (and we *do* need it within the construction called the Dirac operator), *torsion* of our connection is to be *specified.* If the connection under consideration has non-vanishing torsion, the contortion term has to be present in the corresponding Dirac operator:

$$\begin{aligned} \not{D}\psi &\overset{2.}{=} \gamma^a(e_a\psi + \frac{1}{4}\Gamma_{bca}\gamma^b\gamma^c\psi) \\ &\overset{4.}{=} \gamma^a\left(e_a\psi + \frac{1}{8}\left((c_{abc} + c_{cba} - c_{bca}) + 2K_{bca}\right)\gamma^b\gamma^c\psi\right) \end{aligned}$$

Appendix A p.673: The concept of the *direct sum* $V \oplus W$ of vector spaces V and W was introduced in the beginning of Appendix A. An interesting particular case is the space $V \oplus V^*$ (i.e. when W is the *dual* of V).

[6] This stuff is used in the context of the *Berry phase* for the spin 1/2. The phase is intimately connected with the holonomy of this particular bundle.

Why? Because there are two important canonical structures available on this space: a *metric tensor* (with n pluses and n minuses in canonical form) g and a *symplectic form* ω. Both are constructed in terms of *canonical pairing* $\langle\ .\ ,\ .\ \rangle$ between V and V^* alone. They are given as follows:

$$\begin{aligned} g((v,\alpha),(w,\beta)) &:= \langle \alpha, w\rangle + \langle \beta, v\rangle \\ \omega((v,\alpha),(w,\beta)) &:= \langle \alpha, w\rangle - \langle \beta, v\rangle \end{aligned}$$

In particular, if we choose the basis E_a, E^a in $V \oplus V^*$, where

$$\begin{aligned} V \ni e_a &\leftrightarrow (e_a, 0) \equiv E_a \in V \oplus V^* \\ V^* \ni e^a &\leftrightarrow (0, e^a) \equiv E^a \in V \oplus V^* \end{aligned}$$

where, of course,

$$\langle e^a, e_b\rangle = \delta^a_b$$

we get the component expressions

$$\begin{aligned} g &\leftrightarrow \begin{pmatrix} g(E_a,E_b) & g(E_a,E^b) \\ g(E^a,E_b) & g(E^a,E^b) \end{pmatrix} = \begin{pmatrix} 0 & \delta^b_a \\ \delta^a_b & 0 \end{pmatrix} \\ \omega &\leftrightarrow \begin{pmatrix} \omega(E_a,E_b) & \omega(E_a,E^b) \\ \omega(E^a,E_b) & \omega(E^a,E^b) \end{pmatrix} = \begin{pmatrix} 0 & -\delta^b_a \\ \delta^a_b & 0 \end{pmatrix} \end{aligned}$$

So, the symplectic form is already in canonical form in the basis. Metric tensor turns out to be indefinite, with n pluses and n minuses.

▼ Determinant of the matrix of g is

$$\det \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} = \det \begin{pmatrix} \mathbb{I} & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} = \det \begin{pmatrix} \mathbb{I} & 0 \\ \mathbb{I} & -\mathbb{I} \end{pmatrix} = \det \mathbb{I}\ \det(-\mathbb{I}) = (-1)^n \neq 0$$

so that there are no zeroes on diagonal in canonical form. What is on the diagonal is a (non-zero :-) solution of the characteristic equation

$$0 = \det \begin{pmatrix} -\lambda\mathbb{I} & \mathbb{I} \\ \mathbb{I} & -\lambda\mathbb{I} \end{pmatrix} = \det \begin{pmatrix} -\lambda\mathbb{I} & 0 \\ \mathbb{I} & (1/\lambda - \lambda)\mathbb{I} \end{pmatrix} = \det(-\lambda\mathbb{I})\ \det(1/\lambda - \lambda)\mathbb{I} = (\lambda-1)^n(\lambda+1)^n$$

Therefore, the canonical form is diag $(+1, \dots, +1, -1, \dots, -1)$. ▲

If V carries a representation ρ of a group G, then (see 12.1.8) V^* carries the corresponding *dual* representation $\check{\rho}$ and, consequently, $V \oplus V^*$ carries their direct sum $\rho \oplus \check{\rho}$ (see 12.4.10). Now, remarkably (albeit technically trivially), both g and ω are *invariant* with respect to $\rho \oplus \check{\rho}$.

▼ Indeed, for the metric tensor g, say, we have (for any group element $k \in G$)

$$\begin{aligned} g((\rho\oplus\check{\rho})(k)(v,\alpha),(\rho\oplus\check{\rho})(k)(w,\beta)) &:= g(\rho(k)v, \check{\rho}(k)\alpha),(\rho(k)w,\check{\rho}(k)\beta)) && \text{(see 12.4.10)} \\ &= \langle \check{\rho}(k)\alpha, \rho(k)w\rangle + \langle \check{\rho}(k)\beta, \rho(k)v\rangle \\ &= \langle \alpha, w\rangle + \langle \beta, v\rangle && \text{(see 12.1.8)} \\ &= g((v,\alpha),(w,\beta)) \end{aligned}$$

and similarly for the symplectic form ω. ▲

So, starting with just *a* representation (V, ρ) of G, we get, in an appropriately *enlarged* space, much stronger structure, *the G-invariant* metric g as well as symplectic form ω.