Chapter 4

Covariant derivatives.

4.1 Differentiation of tensors.

Let us calculate the derivative of a contravariant vector field v^α after it had been transformed from the $\{x\}$-coordinates to the $\{x'\}$-coordinates. We have

$$v^{\alpha'}{}_{,\beta'} = \left(x^{\alpha'}{}_{,\alpha}\, v^\alpha\right)_{,\beta'} = x^{\alpha'}{}_{,\alpha\beta'}\, v^\alpha + x^{\alpha'}{}_{,\alpha}\, v^\alpha{}_{,\beta'} = x^{\alpha'}{}_{,\alpha\beta}\, x^\beta{}_{,\beta'}\, v^\alpha + x^{\alpha'}{}_{,\alpha}\, x^\beta{}_{,\beta'}\, v^\alpha{}_{,\beta}\,. \tag{4.1}$$

This is not a tensor, in consequence of the term $x^{\alpha'}{}_{,\alpha\beta}\, x^\beta{}_{,\beta'}\, v^\alpha$. An analogous result would be obtained for most other tensors. The derivative of an *arbitrary* tensor field transforms like a tensor only under linear transformations, for which $x^{\alpha'}{}_{,\alpha\beta} = 0$. There are only a few special cases in which the derivatives of tensor fields are themselves tensors with respect to arbitrary coordinate transformations. One example we already know – it is the derivative of a scalar field, which is a covariant vector. The three other examples are:

1. The derivatives of the Levi-Civita symbols and of all the Kronecker deltas are identically equal to zero.

2. If $T_{\alpha_1\ldots\alpha_k}$ is a tensor (of weight 0), then $T_{[\alpha_1\ldots\alpha_k,\alpha_{k+1}]}$ is a tensor, too. This quantity is a generalisation of the rotation of a vector field.

3. If $T^{\alpha_1\cdots\alpha_k}$ is a tensor density of weight -1, and is completely antisymmetric in all the indices, then $T^{\alpha_1\cdots\alpha_k}{}_{,\alpha_k}$ is also a tensor density of weight -1. This is a generalisation of the divergence of a vector field.

The first example is trivial, while the second one is easy to verify (hint: we consider only functions of class C^2, for which second derivatives commute). We will verify the third example because it provides an application of multidimensional deltas in a calculation.

By assumption, when the coordinates are transformed from $\{x\}$ to $\{x'\}$, $T^{\alpha_1\ldots\alpha_k}$ transforms as follows

$$T^{\alpha'_1\ldots\alpha'_k} = \frac{\partial(x)}{\partial(x')}\, x^{\alpha'_1}{}_{,\alpha_1} \ldots x^{\alpha'_k}{}_{,\alpha_k}\, T^{\alpha_1\ldots\alpha_k}. \tag{4.2}$$

Let us differentiate this by $x^{\alpha'_k}$ and contract the result by α'_k:

$$
\begin{aligned}
T^{\alpha'_1\ldots\alpha'_k}{}_{,\alpha'_k} &= \left(\frac{\partial(x)}{\partial(x')}\right)_{,\alpha'_k} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_k}{}_{,\alpha_k}\, T^{\alpha_1\ldots\alpha_k} \\
&\quad+\frac{\partial(x)}{\partial(x')}\sum_{i=1}^{k-1} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_i}{}_{,\alpha_i\alpha'_k}\ldots x^{\alpha'_k}{}_{,\alpha_k}\, T^{\alpha_1\ldots\alpha_k} \\
&\quad+\frac{\partial(x)}{\partial(x')} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_{k-1}}{}_{,\alpha_{k-1}}\, x^{\alpha'_k}{}_{,\alpha_k\alpha'_k}\, T^{\alpha_1\ldots\alpha_k} \\
&\quad+\frac{\partial(x)}{\partial(x')} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_k}{}_{,\alpha_k}\, T^{\alpha_1\ldots\alpha_k}{}_{,\rho}\, x^{\rho}{}_{,\alpha'_k}\,.
\end{aligned} \tag{4.3}
$$

We leave the first term unchanged. In the second term, each component of the sum contains $x^{\alpha'_i}{}_{,\alpha_i\alpha'_k}\, x^{\alpha'_k}{}_{,\alpha_k} \equiv x^{\alpha'_i}{}_{,\alpha_i\alpha_k}$. This expression is symmetric in (α_i, α_k) and is contracted with respect to both (α_i, α_k) with $T^{\alpha_1\cdots\alpha_k}$, which is antisymmetric in these same two indices. Such a contraction is always identically zero, hence the second term is zero.

In the third term, we note that $x^{\alpha'_k}{}_{,\alpha_k}$ is an element of the inverse matrix to $[x^{\alpha}{}_{,\alpha'}]$. Consequently, $x^{\alpha'_k}{}_{,\alpha_k}$ is equal to the cofactor of the element transposed to $\left({}^{\alpha'_k}{}_{\alpha_k}\right)$ in the matrix $[x^{\alpha}{}_{,\alpha'}]$, divided by the determinant of $[x^{\alpha}{}_{,\alpha'}]$. The element transposed to $\left({}^{\alpha'_k}{}_{\alpha_k}\right)$ in $[x^{\alpha}{}_{,\alpha'}]$ is $x^{\alpha_k}{}_{,\alpha'_k}$, so its cofactor is, by (3.36)

$$
\frac{1}{(n-1)!}\delta^{\alpha'_k\mu'_1\ldots\mu'_{n-1}}_{\alpha_k\nu_1\ldots\nu_{n-1}}\, x^{\nu_1}{}_{,\mu'_1}\ldots x^{\nu_{n-1}}{}_{,\mu'_{n-1}}\,, \tag{4.4}
$$

while $\det\,[x^{\alpha}{}_{,\alpha'}] = \frac{\partial(x)}{\partial(x')}$. Hence we have

$$
\left(x^{\alpha'_k}{}_{,\alpha_k}\right)_{,\alpha'_k} = \left[\left(\frac{\partial(x)}{\partial(x')}\right)^{-1}\frac{1}{(n-1)!}\delta^{\alpha'_k\mu'_1\ldots\mu'_{n-1}}_{\alpha_k\nu_1\ldots\nu_{n-1}}\, x^{\nu_1}{}_{,\mu'_1}\ldots x^{\nu_{n-1}}{}_{,\mu'_{n-1}}\right]_{,\alpha'_k}. \tag{4.5}
$$

The differentiation of $x^{\nu_i}{}_{,\mu'_i}$ by $x^{\alpha'_k}$ will give zero contributions because $x^{\nu_i}{}_{,\mu'_i\alpha'_k}$, are symmetric in (μ'_i, α'_k) and will be contracted with the delta which is antisymmetric in the same indices. The only nonzero contribution will be from the derivative of the determinant, so

$$
\begin{aligned}
\left(x^{\alpha'_k}{}_{,\alpha_k}\right)_{,\alpha'_k} &= -\left(\frac{\partial(x)}{\partial(x')}\right)^{-2}\left(\frac{\partial(x)}{\partial(x')}\right)_{,\alpha'_k} \\
&\quad\times\ \frac{1}{(n-1)!}\delta^{\alpha'_k\mu'_1\ldots\mu'_{n-1}}_{\alpha_k\nu_1\ldots\nu_{n-1}}\, x^{\nu_1}{}_{,\mu'_1}\ldots x^{\nu_{n-1}}{}_{,\mu'_{n-1}}\,.
\end{aligned} \tag{4.6}
$$

The last line above is just $\left(\frac{\partial(x)}{\partial(x')}x^{\alpha'_k}{}_{,\alpha_k}\right)$. Using this in the third term of (4.3) we obtain

$$
\text{third term} = -\left(\frac{\partial(x)}{\partial(x')}\right)_{,\alpha'_k} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_k}{}_{,\alpha_k}\, T^{\alpha_1\ldots\alpha_k}.
$$

We see that terms I and III cancel out. The final result in (4.3) is thus the last term. Using $x^{\alpha'_k}{}_{,\alpha_k}\, x^{\rho}{}_{,\alpha'_k} = x^{\rho}{}_{,\alpha_k} = \delta^{\rho}{}_{\alpha_k}$, it becomes

$$
T^{\alpha'_1\ldots\alpha'_k}{}_{,\alpha'_k} = \frac{\partial(x)}{\partial(x')} x^{\alpha'_1}{}_{,\alpha_1}\ldots x^{\alpha'_{k-1}}{}_{,\alpha_{k-1}}\, T^{\alpha_1\ldots\alpha_k}{}_{,\alpha_k}\,. \tag{4.7}
$$

Hence, $T^{\alpha_1 \dots \alpha_k}{}_{,\alpha_k}$ is a tensor density of weight -1.

Note that the statement proved above is also correct for the case when T^α has just one index and is a contravariant vector density. In that case, terms I and III still cancel each other while term II in (4.3) simply does not exist.

The fact that derivatives of tensor fields are not tensor fields is unfortunate because laws of physics are usually formulated as differential equations. Hence, those equations are not tensorial; they will change when coordinates are transformed. But we would like the laws of physics to have the form (a tensor) = 0, since such an equation would hold in all the coordinate systems. This suggests the following idea: let us define a "generalised differentiation", which will yield tensor fields when acting on tensor fields, and will coincide with ordinary differentiation when acting on scalars and the Kronecker deltas, for which the partial derivative does not destroy the tensor property. Then, we will replace the partial derivatives with the generalised derivatives in the laws of physics. We guess that this generalised differentiation, called **covariant differentiation**, will reduce to ordinary differentiation in certain privileged coordinate systems.

4.2 Axioms of the covariant derivative.

We want the covariant differentiation to have all the algebraic properties of an ordinary differentiation, but in addition we want it to yield tensor densities when acting on tensor densities. We will denote the covariant derivative by ∇_α or ${}_{|\alpha}$ or $D/\partial x^\alpha$. The symbols $T_i[w, k, l]$ will denote tensor densities whose indices we do not need to write out explicitly.

Specifically, we want the ∇_α to have the following properties:

1. To be distributive with respect to addition:

$$\nabla_\alpha \left(T_1[w, k, l] + T_2[w, k, l] \right) = \nabla_\alpha \left(T_1[w, k, l] \right) + \nabla_\alpha \left(T_1[w, k, l] \right). \tag{4.8}$$

2. To obey the Leibniz rule when acting on a tensor product:

$$\begin{aligned} &\nabla_\alpha \left(T_1[w_1, k_1, l_1] \otimes T_2[w_2, k_2, l_2] \right) \\ &= \left(\nabla_\alpha T_1[w_1, k_1, l_1] \right) \otimes T_2[w_2, k_2, l_2] + \left(T_1[w_1, k_1, l_1] \right) \otimes \left(\nabla_\alpha T_2[w_2, k_2, l_2] \right). \end{aligned} \tag{4.9}$$

3. To reduce to the partial derivative when acting on a scalar:

$$\nabla_\alpha \Phi = \Phi_{,\alpha}\,. \tag{4.10}$$

4. To yield zero when acting on the Levi-Civita symbols and Kronecker deltas:

$$\begin{aligned} \nabla_\alpha \epsilon^{\alpha_1 \dots \alpha_n} &= 0, \\ \nabla_\alpha \epsilon_{\alpha_1 \dots \alpha_n} &= 0, \\ \nabla_\alpha \delta^\alpha{}_\beta &= 0. \end{aligned} \tag{4.11}$$

The last equation implies at once that

$$\nabla_\alpha \delta^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_k} = 0 \tag{4.12}$$

for any k. It also implies that ∇_α commutes with contraction.

5. When acting on a tensor density of type $[w, k, l]$, it produces a tensor density of type $[w, k, l+1]$, thus

$$\nabla_\alpha \left(T_1[w, k, l]\right) = T_2[w, k, l+1].$$

Only the last property is different for the covariant and for the partial derivative.

From these requirements we will now derive an operational formula for the covariant derivative.

4.3 A field of vector bases on a manifold and scalar components of tensors.

In every tangent space to an n-dimensional manifold M_n we can choose a set of n linearly independent contravariant vectors, $\{{e_1}^\alpha, \ldots, {e_n}^\alpha\}$. The indices a, b, c, ... will label vectors (as opposed to Greek indices that label coordinate components of tensors). After such a basis of the tangent space is chosen at every $x \in M_n$, let us consider the n vector *fields*:

$$x \to {e_a}^\alpha(x), \qquad a = 1, \ldots, n.$$

The collection of quantities $\{{e_a}^\alpha(x)\}$, $\alpha = 1, \ldots, n$, $a = 1, \ldots, n$ forms a matrix whose elements are functions on the manifold. Since all the vectors are linearly independent at every x, the matrix is nonsingular, so there exists an inverse matrix ${e^a}_\alpha$ that obeys

$${e^a}_\alpha {e_a}^\beta = {\delta^\beta}_\alpha. \tag{4.13}$$

Subsets of the matrix $||{e^a}_\alpha||$ defined by a fixed a are then covariant vectors that form a **dual basis** to $\{{e_a}^\alpha(x)\}$, $a = 1, \ldots, n$. One can verify that, in virtue of the $\{{e_a}^\alpha(x)\}$ being linearly independent, eq. (4.13) implies the following:

$${e^a}_\alpha {e_b}^\alpha = {\delta^a}_b. \tag{4.14}$$

It follows that for any tensor field (i.e. of weight 0) $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}$, the collection of quantities

$$T^{a_1\ldots a_k}_{b_1\ldots b_l} \overset{\mathrm{def}}{=} {e^{a_1}}_{\alpha_1} \ldots {e^{a_k}}_{\alpha_k} {e_{b_1}}^{\beta_1} \ldots {e_{b_l}}^{\beta_l} T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}, \tag{4.15}$$

labelled by the indices $a_1, \ldots, a_k, b_1, \ldots, b_l = 1, \ldots, n$ is a set of n^{k+l} scalar fields that uniquely represents the set of n^{k+l} coordinate components of the tensor field $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}$. This is because, in consequence of (4.13) – (4.14), an inverse formula to (4.15) exists that allows one to calculate $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}$ when $T^{a_1\ldots a_k}_{b_1\ldots b_l}$ are given:

$$T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} = {e_{a_1}}^{\alpha_1} \ldots {e_{a_k}}^{\alpha_k} {e^{b_1}}_{\beta_1} \ldots {e^{b_l}}_{\beta_l} T^{a_1\ldots a_k}_{b_1\ldots b_l}. \tag{4.16}$$

Let us denote

$$e := \det ||{e_a}^\alpha|| = \frac{1}{n!} \epsilon^{a_1\ldots a_n} \epsilon_{\alpha_1\ldots\alpha_n} {e_{a_1}}^{\alpha_1} \ldots {e_{a_n}}^{\alpha_n}. \tag{4.17}$$

Now, $\epsilon_{\alpha_1 \ldots \alpha_n}$ is a tensor density of weight $+1$, while $\epsilon^{a_1 \ldots a_n}$ is a set of scalars because it depends on the choice of basis in the vector space, and not on the coordinate system. Hence, e is a scalar density of weight $+1$.

The quantity e, together with the bases $\{e_a{}^\alpha\}$ and $\{e^a{}_\alpha\}$ can be used to represent arbitrary tensor densities by sets of scalars. Let $T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l}$ now be a tensor density of type $[w, k, l]$; then each element of the set

$$T^{a_1 \ldots a_k}_{b_1 \ldots b_l} := e^{-w} e^{a_1}{}_{\alpha_1} \ldots e^{a_k}{}_{\alpha_k} e_{b_1}{}^{\beta_1} \ldots e_{b_l}{}^{\beta_l} T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l} \tag{4.18}$$

is a scalar. The set $T^{a_1 \ldots a_k}_{b_1 \ldots b_l}$ uniquely defines $T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l}$ via (4.16) with the factor e^{+w} added. The weight w has to be given as an extra bit of information, since the set of scalars alone does not define the weight.

4.4 The affine connection.

We now define the set of quantities:

$$\Gamma^\alpha{}_{\beta\gamma} = -e_s{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) e^s{}_\beta. \tag{4.19}$$

The elements of this set are the coefficients of **affine connection**.When specified explicitly, they tell us how the covariant derivative acts on the basis vector fields. Later we will consider manifolds in which these coefficients can be calculated from more basic objects (see Chapter 7). For now, we consider manifolds in which the $\Gamma^\alpha{}_{\beta\gamma}$ are just given.

Equation (4.19) can be rewritten in an equivalent form

$$\nabla_\gamma e^a{}_\beta = \partial_\gamma e^a{}_\beta - \Gamma^\alpha{}_{\beta\gamma} e^a{}_\alpha.$$

We will verify that the $\Gamma^\alpha{}_{\beta\gamma}$ do not depend on the choice of basis. Let us assume that $\{e_a{}^\alpha\}$ and $\{e_{a'}{}^\alpha\}$ are two different bases. The vectors of the second basis can then be decomposed in the first basis

$$e_{a'}{}^\alpha = A^b{}_{a'} e_b{}^\alpha, \tag{4.20}$$

and the elements of the transformation matrix

$$A^b{}_{a'} = e^b{}_\alpha e_{a'}{}^\alpha \tag{4.21}$$

are scalar fields. Hence, $A^b{}_{a'|\alpha} = A^b{}_{a',\alpha}$ and $\left(A^{-1}\right)^{c'}{}_{d|\alpha} = \left(A^{-1}\right)^{c'}{}_{d,\alpha}$. Then, calculating the $\Gamma^\alpha{}_{\beta\gamma}$ in the basis $\{e_{a'}{}^\alpha\}$ we have

$$\begin{aligned}
\left(\Gamma^\alpha{}_{\beta\gamma}\right)_{e'} &= -e_{s'}{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) e^{s'}{}_\beta = -A^r{}_{s'} e_r{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) \left[\left(A^{-1}\right)^{s'}{}_s e^s{}_\beta\right] \\
&= -A^r{}_{s'} \left(A^{-1}\right)^{s'}{}_s e_r{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) e^s{}_\beta = -\delta^r{}_s e_r{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) e^s{}_\beta \\
&= -e_s{}^\alpha \left(\nabla_\gamma - \partial_\gamma\right) e^s{}_\beta = \left(\Gamma^\alpha{}_{\beta\gamma}\right)_e .
\end{aligned} \tag{4.22}$$

Now let us note that the $\Gamma^\alpha{}_{\beta\gamma}$ *are not tensor fields.* When coordinates are transformed, these coefficients change as follows

$$\Gamma^{\alpha'}{}_{\beta'\gamma'} = -e_s{}^{\alpha'} \left(\nabla_{\gamma'} - \partial_{\gamma'}\right) e^s{}_{\beta'} = x^{\alpha'}{}_{,\alpha} x^\beta{}_{,\beta'} x^\gamma{}_{,\gamma'} \Gamma^\alpha{}_{\beta\gamma} + x^{\alpha'}{}_{,\rho} x^\rho{}_{,\beta'\gamma'} . \tag{4.23}$$

However, the antisymmetric part of $\Gamma^{\alpha}{}_{\beta\gamma}$

$$\Omega^{\alpha}{}_{\beta\gamma} \stackrel{\text{def}}{=} \Gamma^{\alpha}{}_{[\beta\gamma]}, \tag{4.24}$$

is a tensor, called the **torsion tensor**, since $x^{\rho}{}_{,[\beta'\gamma']} = 0$.

4.5 The explicit formula for the covariant derivative of tensor densities.

In order to obtain the explicit formula for the covariant derivative, we need to know two other properties of the connection coefficients:

$$(I) \quad \Gamma^{\alpha}{}_{\beta\gamma} = e^{s}{}_{\beta} \left(\nabla_{\gamma} - \partial_{\gamma}\right) e_{s}{}^{\alpha}. \tag{4.25}$$

The verification of this is an easy exercise.

$$(II) \quad \nabla_{\alpha} \left(e^{w}\right) = w e^{w-1} \nabla_{\alpha} e. \tag{4.26}$$

This can be verified in the following way. Let us consider the quantity

$$F_{\alpha}(w) := e^{-w} \nabla_{\alpha} \left(e^{w}\right). \tag{4.27}$$

Using the postulated properties of the covariant derivative we obtain:

$$\begin{aligned} F_{\alpha}\left(w_1 + w_2\right) &= e^{-w_1} e^{-w_2} \left[\left(\nabla_{\alpha} e^{w_1}\right) e^{w_2} + e^{w_1} \left(\nabla_{\alpha} e^{w_2}\right)\right] \\ &= e^{-w_1} \left(\nabla_{\alpha} e^{w_1}\right) + e^{-w_2} \left(\nabla_{\alpha} e^{w_2}\right) = F_{\alpha}(w_1) + F_{\alpha}(w_2). \end{aligned} \tag{4.28}$$

Every continuous function that has the property $f\left(w_1 + w_2\right) = f\left(w_1\right) + f\left(w_2\right)$ for all real w_1 and w_2 also has the property $f(w) = f(1) w$. Hence

$$e^{-w} \nabla_{\alpha} \left(e^{w}\right) = w e^{-1} \nabla_{\alpha} e,$$

which is equivalent to (4.26).

Equation (4.26) holds also for partial derivatives, so

$$e^{-w} \left(\nabla_{\gamma} - \partial_{\gamma}\right) \left(e^{w}\right) = w e^{-1} \left(\nabla_{\gamma} - \partial_{\gamma}\right) e. \tag{4.29}$$

Now, using (3.32) and (3.36), we obtain

$$\begin{aligned} \left(\nabla_{\gamma} - \partial_{\gamma}\right) e &= \frac{1}{(n-1)!} \delta^{a_1 \ldots a_n}_{\rho_1 \ldots \rho_n} e_{a_1}{}^{\rho_1} \ldots e_{a_{n-1}}{}^{\rho_{n-1}} \left(\nabla_{\gamma} - \partial_{\gamma}\right) e_{a_n}{}^{\rho_n} \\ &= e e^{a_n}{}_{\rho_n} \left(\nabla_{\gamma} - \partial_{\gamma}\right) e_{a_n}{}^{\rho_n} = e \Gamma^{\rho}{}_{\rho\gamma}. \end{aligned} \tag{4.30}$$

At this point, we are prepared to deduce the general formula for the covariant derivative of an arbitrary tensor density. As an introductory exercise we do it first for contravariant and covariant vector densities.

Let us convert a contravariant vector density A^α of weight w to a set of scalars by (4.18). Since $A^a = e^{-w} e^a{}_\alpha A^\alpha$ are scalars, axiom 3 implies

$$\left(\nabla_\gamma - \partial_\gamma\right) A^a = 0. \tag{4.31}$$

On the other hand, using now axiom 2 and (4.30), we apply $(\nabla_\gamma - \partial_\gamma)$ to the right-hand side of (4.18) and obtain

$$\begin{aligned}&\left(\nabla_\gamma - \partial_\gamma\right) A^a = \\ &-we^{-w}\Gamma^\rho{}_{\rho\gamma} e^a{}_\alpha A^\alpha + e^{-w}\left[\left(\nabla_\gamma - \partial_\gamma\right) e^a{}_\alpha\right] A^\alpha + e^{-w} e^a{}_\alpha \left(\nabla_\gamma - \partial_\gamma\right) A^\alpha.\end{aligned} \tag{4.32}$$

Now let us convert this equation back to coordinate components, by contracting it with $e^w e_a{}^\alpha$ (first change the summation index α!). Using eqs. (4.19) and (4.25), we get

$$\left(\nabla_\gamma - \partial_\gamma\right) A^\alpha = w\Gamma^\rho{}_{\rho\gamma} A^\alpha + \Gamma^\alpha{}_{\rho\gamma} A^\rho. \tag{4.33}$$

From here, finally

$$\nabla_\gamma A^\alpha = \partial_\gamma A^\alpha + w\Gamma^\rho{}_{\rho\gamma} A^\alpha + \Gamma^\alpha{}_{\rho\gamma} A^\rho. \tag{4.34}$$

By similar calculations we obtain for a *covariant* vector density B_α of weight w:

$$\nabla_\gamma B_\alpha = \partial_\gamma B_\alpha + w\Gamma^\rho{}_{\rho\gamma} B_\alpha - \Gamma^\rho{}_{\alpha\gamma} B_\rho, \tag{4.35}$$

and for tensors of rank 2:

$$T^{\alpha\beta}{}_{|\gamma} = T^{\alpha\beta}{}_{,\gamma} + \Gamma^\alpha{}_{\rho\gamma} T^{\rho\beta} + \Gamma^\beta{}_{\rho\gamma} T^{\alpha\rho}; \tag{4.36}$$

$$T_{\alpha\beta|\gamma} = T_{\alpha\beta,\gamma} - \Gamma^\rho{}_{\alpha\gamma} T_{\rho\beta} - \Gamma^\rho{}_{\beta\gamma} T_{\alpha\rho}; \tag{4.37}$$

$$T^\alpha{}_{\beta|\gamma} = T^\alpha{}_{\beta,\gamma} + \Gamma^\alpha{}_{\rho\gamma} T^\rho{}_\beta - \Gamma^\rho{}_{\beta\gamma} T^\alpha{}_\rho. \tag{4.38}$$

For a general tensor density of weight w, $T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l}$, still by the same reasoning, we obtain

$$\begin{aligned}\nabla_\gamma T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l} &= \partial_\gamma T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l} + w\Gamma^\rho{}_{\rho\gamma} T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l} + \sum_{i=1}^{k} \Gamma^{\alpha_i}{}_{\rho_i\gamma} T^{\alpha_1 \ldots \rho_i \ldots \alpha_k}_{\beta_1 \ldots \beta_l} \\ &- \sum_{j=1}^{l} \Gamma^{\rho_j}{}_{\beta_j\gamma} T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \rho_j \ldots \beta_l},\end{aligned} \tag{4.39}$$

where the sums run through all the positions of the respective indices.

Note that, unlike a partial derivative, the covariant derivative does not act on single components of tensor densities. It is an operator that acts on the whole tensor density and produces another tensor density.

4.6 Exercises.

1. What is the condition for the "covariant rotation" $T_{[\alpha|\beta]}$ of a covariant vector field T_α to coincide with the ordinary rotation $T_{[\alpha,\beta]}$?

2. Let $g_{\alpha\beta} = g_{(\alpha\beta)}$ be a doubly covariant tensor that is nonsingular, i.e. $\det ||g_{\alpha\beta}|| \neq 0$. Let $g^{\alpha\beta}$ be its inverse matrix, i.e.

$$g^{\alpha\rho} g_{\rho\beta} = \delta^{\alpha}{}_{\beta}.$$

Show that the object defined as follows

$$\left\{ \begin{matrix} \alpha \\ \beta\gamma \end{matrix} \right\} = \frac{1}{2} g^{\alpha\rho} \left(g_{\beta\rho,\gamma} + g_{\gamma\rho,\beta} - g_{\beta\gamma,\rho} \right)$$

transforms under coordinate transformations by the same law as the coefficients of affine connection. What is the torsion in this case?

Chapter 5

Parallel transport and geodesic lines.

5.1 Parallel transport.

Let a curve C be given in a manifold with affine connection, and let v be a field of vectors defined along the curve, $C \ni x \to v^\alpha(x)$. In a Euclidean space, in Cartesian coordinates, the vectors of the field are parallel when

$$\frac{\mathrm{d}v^\alpha}{\mathrm{d}\tau} = \frac{\partial v^\alpha}{\partial x^\beta}\,\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0, \tag{5.1}$$

where τ is a parameter along C, while $x^\alpha(\tau)$ are coordinates of a point on C. Then, for any two points on C corresponding to the parameter values $\tau = \tau_1$ and $\tau = \tau_2$ we have

$$v^\alpha(\tau_1) \underset{*}{=} v^\alpha(\tau_2). \tag{5.2}$$

(An asterisk below the equality sign means that the equation holds only in some specific coordinate systems. For example, (5.2) does not hold for parallel vectors in polar coordinates on a Euclidean plane.) We now generalise the definition of parallelism along a curve C in such a way that it is independent of the coordinates used:

$$\frac{\mathrm{D}v^\alpha}{\mathrm{d}\tau} \overset{\mathrm{def}}{=} \left(\nabla_\rho v^\alpha\right)\,\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0, \tag{5.3}$$

where $\mathrm{d}x^\rho/\mathrm{d}\tau$ are components of the tangent vector to C. Equation (5.3) is at the same time the definition of a covariant derivative along a curve. Using (4.34) with $w = 0$, (5.3) is equivalent to

$$v^\alpha{}_{,\rho}\,\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} + \Gamma^\alpha{}_{\sigma\rho} v^\sigma\,\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0. \tag{5.4}$$

This can be written as

$$\frac{\mathrm{d}v^\alpha}{\mathrm{d}\tau} = -\Gamma^\alpha{}_{\sigma\rho} v^\sigma \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}. \tag{5.5}$$

Thus, the vector $v^\alpha(\tau_1)$, while being parallely transported from the point $\tau = \tau_1$ to the point $\tau = \tau_2$ along C, changes in the following way

$$v_\parallel^\alpha(\tau_2) = v^\alpha(\tau_1) - \underset{C}{\int_{\tau_1}^{\tau_2}} \Gamma^\alpha{}_{\sigma\rho}(\tau) v^\sigma(\tau) \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\mathrm{d}\tau = v^\alpha(\tau_1) - \underset{C}{\int_{x(\tau_1)}^{x(\tau_2)}} \Gamma^\alpha{}_{\sigma\rho}(x) v^\sigma(x)\mathrm{d}x^\rho. \tag{5.6}$$

The integrand in (5.6) in general is not a perfect differential, so the result of integration in general depends on the curve C. Conclusion: the parallel transport so defined is not universal; its result depends on the path of transport. In particular, for parallel transport along a closed curve, in general we will have

$$v_{\|}^{\alpha}\left(\tau_1\right) \stackrel{\text{def}}{=} v^{\alpha}\left(\tau_1\right) - \oint_C \Gamma^{\alpha}{}_{\sigma\rho}(\tau) v^{\sigma}(\tau) \frac{\mathrm{d}x^{\rho}}{\mathrm{d}\tau} \mathrm{d}\tau \neq v^{\alpha}\left(\tau_1\right). \tag{5.7}$$

The conditions under which the result of parallel transport is independent of the path, i.e. under which the effect (5.7) does not occur, will be given in Chap. 6.

Parallel transport of an arbitrary tensor density $T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_l}$ is defined analogously to (5.3):

$$\frac{\mathrm{D}}{\mathrm{d}\tau} T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_l} = T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_l|\rho} \frac{\mathrm{d}x^{\rho}}{\mathrm{d}\tau} = 0. \tag{5.8}$$

5.2 Geodesic lines.

We call a **geodesic line** (briefly just **geodesic**) such a curve G whose tangent vector, after being parallely transported along it from a point $x\left(\tau_0\right) \in G$ to an arbitrary point $x^{\alpha}(\tau) \in G$, is collinear with the tangent vector that is defined at $x^{\alpha}(\tau)$. Hence

$$v_{\|}^{\alpha}(\tau)\Big|_{\tau_0 \to \tau} = \lambda(\tau) v^{\alpha}(\tau), \tag{5.9}$$

or, according to (5.6)

$$v_{\|}^{\alpha}(\tau)\Big|_{\tau_0 \to \tau} = v^{\alpha}\left(\tau_0\right) - \underset{G}{\int_{\tau_0}^{\tau}} \Gamma^{\alpha}{}_{\sigma\rho}(t) \lambda(t) v^{\sigma}(t) v^{\rho}(t) \mathrm{d}t = \lambda(\tau) v^{\alpha}(\tau). \tag{5.10}$$

An example of a geodesic is a straight line in a Euclidean space. In that case, *in Cartesian coordinates*, the integral in (5.10) is zero and, *if* τ *is the length of arc*, $\lambda(\tau) = 1$. A geodesic line is a generalisation of the notion of a straight line to any manifold with affine connection.

Let us differentiate (5.10) by τ. The result is

$$-\Gamma^{\alpha}{}_{\sigma\rho}(\tau) \lambda(\tau) v^{\sigma}(\tau) v^{\rho}(\tau) = \frac{\mathrm{d}\lambda}{\mathrm{d}\tau} v^{\alpha}(\tau) + \lambda(\tau) \frac{\mathrm{d}v^{\alpha}}{\mathrm{d}\tau}.$$

This can be rewritten as follows

$$\lambda \left(\frac{\mathrm{d}^2 x^{\alpha}}{\mathrm{d}\tau^2} + \Gamma^{\alpha}{}_{\sigma\rho}(\tau) \frac{\mathrm{d}x^{\sigma}}{\mathrm{d}\tau} \frac{\mathrm{d}x^{\rho}}{\mathrm{d}\tau} \right) = -\frac{\mathrm{d}\lambda}{\mathrm{d}\tau} \frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}. \tag{5.11}$$

Now let us change the parameter as follows

$$\tau \to s(\tau) \stackrel{\text{def}}{=} \int_{\tau_0}^{\tau} \frac{c}{\lambda(t)} \mathrm{d}t, \tag{5.12}$$

where c and τ_0 are arbitrary constants. Then, eq. (5.11) becomes:

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}s^2} + \Gamma^\alpha{}_{\sigma\rho}(s)\frac{\mathrm{d}x^\sigma}{\mathrm{d}s}\frac{\mathrm{d}x^\rho}{\mathrm{d}s} = 0, \tag{5.13}$$

or, equivalently

$$\frac{\mathrm{D}}{\mathrm{d}s}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\right) = 0. \tag{5.14}$$

The form (5.13) of the **geodesic equation** is privileged in that, in the parametrisation (5.12), the tangent vector transported parallely along the geodesic is not only collinear with the locally defined tangent vector, but coincides with it. The parameter s that has this property exists for any $\lambda(\tau)$ such that $\lambda(\tau) \neq 0$ at every τ, and it is called the **affine parameter**. It is defined up to the linear transformations

$$s' = as + b, \qquad a, b = \text{constant}. \tag{5.15}$$

Equation (5.13) allows us to prove the following theorem:

Theorem 5.1 *On a manifold M_n with an affine connection, at its every point x and for every vector v tangent to M_n at x, there exists a geodesic line passing through x that is tangent to v.*

This is so because a solution of a second-order differential equation is uniquely determined by its value at one point and its first derivative at that point.

However, it is not the case that on a manifold with affine connection any two points can be connected by one and only one geodesic. A geodesic joining two points may not exist, like on a two-sheeted hyperboloid. On the other hand, any two points on the surface of a cylinder (where the geodesic lines are straight lines, circles and screw-lines) can be connected by an infinite number of geodesics. (Imagine a screw-line that connects two points p and q by the shortest arc, then another one that runs one extra time around the cylinder between p and q, then another one that runs two times around the cylinder, and so on.)

Note that only the symmetric part of the connection gives a nonzero contribution to the geodesic equation.

5.3 Exercises.

1. Consider a vector on a Euclidean plane being transported parallely along a straight line. Find how its components change when they are given in polar coordinates.

2. Do the same for a vector in a 3-dimensional Euclidean space when its components are given in spherical coordinates. From the result, read out the connection coefficients of the Euclidean space in spherical coordinates.