2

Vector and tensor fields

• From elementary physics we know vectors as being arrows, exhibiting direction and length. This means that they have both a head *as well as a tail*, the latter being drawn as a point *of the same space* in which the physics is enacted. A vector, then, is equivalent to an ordered *pair of points* in the space. Such a conception works perfectly on the common plane as well as in three-dimensional (Euclidean) space.

However, in general this idea presents difficulties. One can already perceive them clearly on "curved" two-dimensional surfaces (consider, as an example, such a "vector" on a sphere S^2 in the case when its length equals the length of the equator). Recall, however, the various contexts in which vectors enter the physics. One comes to the conclusion that the "tail" point of the vector has no "invariant" meaning; only the *head* point of the vector makes sense as a point of the space. Take as a model case the concept of the (instantaneous) *velocity* vector $\mathbf{v}$ of a point mass at some definite instant of time t. Its meaning is as follows: if the point is at position $\mathbf{r}$ at time t, then it will be at position $\mathbf{r} + \epsilon\mathbf{v}$ at time $t + \epsilon$. However long the vector $\mathbf{v}$ is, the point mass will be *only infinitesimally* remote from its original position. The (instantaneous) velocity vector $\mathbf{v}$ thus evidently carries only "local" information and it is related in no reasonable way to any "tail" point at *finite* distance from its head.

And the transition from (say) a plane to a sphere (or any other curved surface) changes practically nothing in this reasoning: although we may visualize the velocity as an arrow *touching* the surface at a given place, it makes no sense to take seriously its tail as a second point on the surface (within a finite distance from the first one), since all the velocity vector informs us about is the behavior of the trajectory within the nearest (infinitesimal) time interval and over such a short time interval all that we manage to do is to move to a point infinitesimally near to the first one. Consequently, the *second point* (the tail of the vector) *plays no invariant role* in this business. The velocity vector is thus to be treated as a concept which is strictly *confined to a point*. A similar analysis of other vectors in physics (acceleration, force, etc.) leads to the same result. Vectors are objects which are to be treated as being "point-like" entities, i.e. as existing at *a single point*.

That means, however, that our approach to vectors on a manifold has to take into account this essential piece of information. Fortunately, such an approach does exist; in fact, there are even several equivalent ways of reaching this goal, as described in Section 2.2.

Before doing this, we undertake a short digression on the concepts of a curve and a function on a manifold, since they play (in addition to being important enough in themselves) essential roles in the construction of a vector. The simple machinery of multilinear algebra (see Section 2.4) then makes it possible to take a (long) step forward, introducing objects of great importance in physics as well as in mathematics – tensor fields on a manifold.

2.1 Curves and functions on M

• A *curve* on a manifold M is a (smooth) map

$$\gamma : \mathbb{R}[t] \to M \qquad t \mapsto \gamma(t) \in M$$

or, more generally,

$$\gamma : I \to M$$

$I \equiv (a, b)$ being an open interval on $\mathbb{R}[t]$. Note that a definite *parametrization* of points from $\mathrm{Im}\,\gamma \subset M$ is inherent in the definition of a curve, and two curves which differ by the parametrization alone are to be treated as being different (in spite of the fact that their image sets $\mathrm{Im}\,\gamma$ on the manifold M coincide). If

$$\varphi : \mathcal{O} \to \mathbb{R}^n[x^1, \dots, x^n]$$

is a chart (i.e. x^i are local coordinates on $\mathcal{O} \subset M$), one obtains a *coordinate presentation* of a curve γ,

$$\hat{\gamma} \equiv \varphi \circ \gamma : \mathbb{R}[t] \to \mathbb{R}^n[x^1, \dots, x^n]$$

i.e. a curve on $\mathbb{R}^n$

$$t \mapsto (x^1(t), \dots, x^n(t)) \equiv (x^1(\gamma(t)), \dots, x^n(\gamma(t)))$$

In general, a curve may convey several coordinate patches, so that several coordinate presentations are sometimes needed for a single curve.

A *function* on a manifold M is a (smooth) map

$$f : M \to \mathbb{R} \qquad x \mapsto f(x) \in \mathbb{R}$$

If

$$\varphi : \mathcal{O} \to \mathbb{R}^n[x^1, \dots, x^n]$$

is a chart, one obtains a *coordinate presentation* of a function f

$$\hat{f} \equiv f \circ \varphi^{-1} \ : \ \mathbb{R}^n \to \mathbb{R}$$

i.e. a function on (a part of) $\mathbb{R}^n$

$$(x^1, \dots, x^n) \mapsto \hat{f}(x^1, \dots, x^n) \in \mathbb{R}$$

so that $\hat{f}$ is a common "function of n variables." We will frequently identify the function with its coordinate presentation in what follows. What will be "really" meant should be clear from the context (the same holds for curves).

[2.1.1] Show that the prescription

$$A \mapsto \det A \equiv f(A)$$

defines a smooth function on the manifold of all real $n \times n$ matrices ($\sim \mathbb{R}^{n^2}$).

Hint: The determinant is a polynomial in the matrix elements. □

2.2 Tangent space, vectors and vector fields

• The concept of a vector in a point $x \in M$ is undoubtedly one of the most fundamental notions in differential geometry, serving as the basis from which the whole machinery of tensor fields (in particular, differential forms) on a manifold is developed with the aid of the standard methods of multilinear algebra (to be explained in Section 2.4).

A word of caution is in order. Although the *actual computations* with vectors (as well as vector and tensor fields) are very simple and indeed "user friendly," the *definition* of a vector is, in contrast, a fairly subtle and tricky matter for the beginner and it might need some time to grasp the ideas involved in full detail. Our recommendation is not to be in a hurry and reserve due time to digest all the details of the exposition. A clear understanding of what a vector *is* in differential geometry saves time later, when vectors are used in more advanced applications.

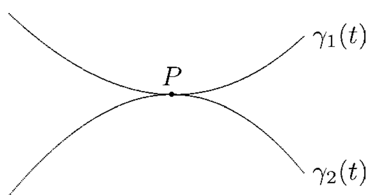
There are several (equivalent) ways in which the concept of a vector at a point $x \in M$ may be introduced. In what follows we mention four of them. In different contexts different definitions turn out to be the most natural. That is why it is worth being familiar with *all* of them.

Each approach reveals the key fact that one can naturally associate an n-dimensional vector space with each point P on an n-dimensional manifold M. The elements of this vector space (the tangent space at P) are then treated as vectors at the point $P \in M$.

The first approach generalizes naturally the concept of the instantaneous velocity $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$ of a point mass moving along a trajectory $\mathbf{r}(t)$, mentioned at the beginning of the chapter. The essential idea is that of tangent curves.

Definition Given two curves γ_1, γ_2 on M, we say that γ_1 is *tangent* to γ_2 at the point $P \in M$ if

1. $\gamma_1(0) = \gamma_2(0) = P$
2. $\frac{d}{dt}\big|_0 x^i(\gamma_1(t)) = \frac{d}{dt}\big|_0 x^i(\gamma_2(t))$



(x^i being arbitrary local coordinates in the neighborhood of P). When expressed in the terminology of analytical mechanics, the definition says that at the moment $t = 0$ the positions of two fictitious points in a configuration space M, moving along trajectories $\gamma_1(t)$ and $\gamma_2(t)$ respectively, happen to coincide (they are both in P) and, in addition, the values of their generalized velocities are the same. The curves (trajectories), which are tangent at $t = 0$, thus have (at $t = 0$) the same values of *both* generalized *coordinates and velocities*. It is clear, then, that the motions along these trajectories are *up to the first order* in time (within the interval from 0 to ϵ) *equal*. (Note that the particular choice $t = 0$ actually plays no distinguished role in this concept; the curves may be tangent at any other "time" as well.)

2.2.1 Show that

(i) the definition does not depend on the choice of local coordinates in a neighborhood of P
(ii) the relation "to be tangent in P" is an equivalence on the set of curves on M obeying $\gamma(0) = P$
(iii) the Taylor expansion (the class of smoothness C^ω is assumed here) of equivalent curves in a neighborhood of $t = 0$ is as follows:

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

where $x^i(P), a^i \in \mathbb{R}$ are common for the whole equivalence class.

Hint: (i) $\frac{d}{dt}\Big|_0 x'^i(\gamma(t)) = \frac{\partial x'^i}{\partial x^j}(P)\, \frac{dx^j(\gamma(t))}{dt}\Big|_0$, i.e. $a'^i = J^i_j(P)a^j$. □

• It turns out that the *equivalence classes* $\dot\gamma := [\gamma]$ of curves γ are endowed with a natural linear structure, which may be introduced by means of representatives.

2.2.2 Given T_PM the set of equivalence classes in the sense of (2.2.1), let $v, w \in T_PM$ and γ, σ be two representatives of these classes ($v = \dot\gamma \equiv [\gamma]$, $w = \dot\sigma \equiv [\sigma]$), such that

$$x^i(\gamma(t)) = x^i(P) + ta^i + o(t)$$
$$x^i(\sigma(t)) = x^i(P) + tb^i + o(t)$$

Show that the prescription

$$v + \lambda w \equiv [\gamma] + \lambda[\sigma] := [\gamma + \lambda\sigma]$$

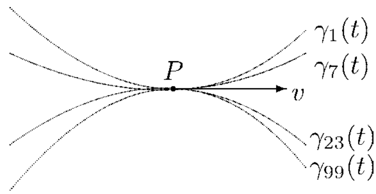
where

$$x^i((\gamma + \lambda\sigma)(t)) := x^i(P) + t(a^i + \lambda b^i) + o(t)$$

introduces by means of representatives into T_PM the well-defined structure of an n-dimensional linear space, i.e. that the definition does not depend on

(i) the choice of local coordinates
(ii) the choice of representatives γ, σ of the classes v, w. □

• Because of this result we may for good reasons (and justly indeed) call the elements $v \in T_PM$ (tangent) *vectors* at the point $P \in M$; the space T_PM itself is called the *tangent space* at the point P. From the definition of linear combination in (2.2.2) one can see that *all* vectors at the point P share the same values of $x^i(P)$ and the property by which they can be



distinguished from one another is by the values of the coefficients $a^i \equiv \dot{x}^i(0)$. Note that a vector "uses" only the first two terms of the Taylor expansion of its coordinate presentation (the zeroth and the first derivatives), the higher terms being completely arbitrary. This means that a single vector corresponds to an *infinite number* of curves which represent this particular vector (which should be clear in advance from the intuitive vision of *all* the curves being tangent to one another), so that there are an infinite number of representatives of each equivalence class. If we would like to visualize the concept of a vector in the sense of an equivalence class of tangent curves, we should assign something like a "bunch" or a "sheaf" of curves, all of them firmly bound together at the point P. And a good old arrow, which cannot be thought of apart from the vector, could be put at P in the direction of this bunch, too (so that it does not feel sick at heart that it had been forgotten because of some dubious novelties).

2.2.3 Verify that

(i) if $\dim M = n$, then $T_P M$ is an n-dimensional space
(ii) equivalence classes of *coordinate curves* $\gamma_j(t)$, i.e. the curves obeying $x^i(\gamma_j(t)) = x^i(P) + \delta^i_j t$ (the value of the jth coordinate is the only one that varies (namely linearly) with t) constitute a basis of $T_P M$.

Hint: (i) $v \leftrightarrow a^i$ is an isomorphism $T_P M \leftrightarrow \mathbb{R}^n$; (ii) check that $v \equiv [\gamma] = a^i[\gamma_i]$. □

• The definition of a vector in terms of curves is intuitively easy to grasp. From the point of view of practical manipulations with vectors (and tensors) another one proves to be convenient, too. It is based on the idea of the *directional derivative* of a function and leans heavily on *algebraic* properties of functions and their directional derivatives.

2.2.4 Let $\mathcal{F}(M) := \{f : M \to \mathbb{R}\}$ denote the set of (smooth) functions on M, $f \in \mathcal{F}(M)$, $v \in T_P M$. Define the map (derivative of f in the direction of v)

$$\hat{v} : \mathcal{F}(M) \to \mathbb{R} \qquad f \mapsto \hat{v}(f) := \left.\frac{d}{dt}\right|_0 f(\gamma(t)) \qquad v = [\gamma]$$

Prove that $\hat{v}$ does not depend on the representative γ in the class $[\gamma] = v$ (i.e. correctness of the definition). □

• It turns out that this map has interesting algebraic properties, enabling one to give an alternative definition of the concept of a vector at the point $P \in M$.

2.2.5 Check that

(i) the prescriptions

$$(f + \lambda g)(x) := f(x) + \lambda g(x)$$
$$(fg)(x) := f(x)g(x)$$

($f, g \in \mathcal{F}(M), \lambda \in \mathbb{R}$) endow $\mathcal{F}(M)$ naturally with the structure of an (∞-dimensional) *associative algebra* (Appendix A.2) and that this algebra turns out to be *commutative* ($fg = gf$) for each manifold; it is called the *algebra of functions on a manifold M*

(ii) the map

$$\hat{v} : \mathcal{F}(M) \to \mathbb{R}$$

from exercise (2.2.4) is a *linear functional* on $\mathcal{F}(M)$, i.e. it behaves on linear combination according to the rule

$$\hat{v}(f + \lambda g) = \hat{v}(f) + \lambda \hat{v}(g)$$

(iii) in addition this functional has the property (behavior on a product)

$$\hat{v}(fg) = \hat{v}(f)g(P) + f(P)\hat{v}(g) \qquad \text{(Leibniz's rule)}$$

(iv) such linear functionals (obeying Leibniz's rule associated with the point P) constitute a linear space (we denote it as $\hat{T}_P M$, here), if one defines

$$(\hat{v} + \lambda \hat{w})(f) := \hat{v}(f) + \lambda \hat{w}(f)$$

(v) the map

$$\psi : T_P M \to \hat{T}_P M \qquad v \mapsto \hat{v}$$

is linear and bijective (i.e. it is an *isomorphism*).

Hint: (v) surjectivity: if $\hat{v}x^i =: a^i$, the inverse image is $v = a^i[\gamma_i]$. □

• Because of the existence of the (canonical) isomorphism $T_P M \leftrightarrow \hat{T}_P M$, these spaces are completely equivalent, so that one may alternatively *define* a vector at the point $P \in M$ as a linear functional on $\mathcal{F}(M)$, behaving according to Leibniz's rule on the product, too.

2.2.6 Define the elements $e_i \in \hat{T}_P M, i = 1, \ldots, n$ as follows:

$$e_i(f) := \left.\frac{\partial f}{\partial x^i}\right|_P \equiv \partial_i|_P f$$

or symbolically

$$e_i := \partial_i|_P$$

Check that

(i) the e_i belong to $\hat{T}_P M$, indeed
(ii) the e_i happen to be just the images of vectors $[\gamma_i] \in T_P M$ (which constitute a basis of $T_P M$) with respect to the map ψ from exercise (2.2.5)
(iii) any vector $\hat{v} \in \hat{T}_P M$ may be uniquely written in the form

$$\hat{v} = a^i e_i \qquad \text{where} \quad a^i = \hat{v}x^i$$

(iv) under the change of coordinates $x^i \mapsto x'^i(x)$, the quantities a^i and e_i transform as follows:

$$a^i \mapsto a'^i = J^i_j a^j \qquad e_i \mapsto e'_i = (J^{-1})^j_i e_j$$

where

$$J^i_j = \frac{\partial x'^i}{\partial x^j}(P) \equiv J^i_j(P) = \text{Jacobian matrix of the change of coordinates}$$

(v) the "whole" $\hat{v} \equiv a^i e_i$ is not altered under the change of coordinates (it is invariant)

$$a^i e_i = a'^i e'_i = \hat{v}$$

(vi) the transformation rules for a^i as well as e_i meet the consistency condition on the intersection of *three* charts (coordinate patches): the composition $x \mapsto x' \mapsto x''$ is to give the same result as the direct way $x \mapsto x''$. □

• These results enable one to introduce immediately *another two* definitions of a vector at the point P (and the tangent space as well). The first possibility is to declare as a vector a first-order differential operator with constant coefficients, i.e. an expression $a^i \; \partial_i|_P$, with linear combinations being given by

$$a^i \; \partial_i|_P + \lambda b^i \; \partial_i|_P := (a^i + \lambda b^i) \; \partial_i|_P$$

The second possibility is the definition adopted by classical differential geometry: a vector at a point $P \in M$ is an n-tuple of real numbers $a^i, i = 1, \ldots, n$, associated with the coordinates x^i in a neighborhood of P; under change of coordinates the n-tuple should transform (by definition) according to the rule

$$x^i \mapsto x'^i(x) \quad \Rightarrow \quad a^i \mapsto J^i_j(P) a^j$$

Altogether we gave *four equivalent* definitions (one can even add more) of a vector: a vector as being

1. an equivalence class of curves (with respect to the equivalence relation "being tangent at the point P")
2. a linear functional on $\mathcal{F}(M)$, which behaves on a product according to Leibniz's rule
3. a first-order differential operator (together with the evaluation of the result at the point P)
4. an n-tuple of real numbers a^i, which transform in a specific way under the change of coordinates.

[2.2.7] Check in detail their equivalence: given a vector in any of these four ways, associate with it corresponding vectors in the other three senses. In particular, make explicit the correspondence between the *basis* vectors in all four languages. □

• Taking into account the equivalence of the four definitions mentioned above, we may regard a vector as being given in *any* of the possible realizations, from now on. The corresponding tangent space will be denoted by a *common* symbol $T_P M$, as well. The basis $e_i \equiv \partial_i|_P \leftrightarrow [\gamma_i] \leftrightarrow \ldots$ is said to be the *coordinate basis* in $T_P M$ and the numbers a^i constitute the *components* of a vector v with respect to the basis e_i.

(Note that the linear combination has only been defined for vectors sitting *at the same point* of a manifold (i.e. in a single tangent space $T_P M$). The spaces $T_P M$ and $T_{P'} M$ for $P \neq P'$ are to be regarded as *different* vector spaces. It is true that they are isomorphic (both being n-dimensional), but there is no *canonical* isomorphism (there exist *infinitely many* isomorphisms, but none is distinguished, in general) so that there is no *natural* (preferred) correspondence between vectors sitting at different points. The fact that vectors *are routinely* linearly combined, in spite of sitting at different points, in physics (the momenta of a collection of particles are *added* in order to obtain the total momentum vector of the system, to give an example) is justified by *particular additional* structure inherent in the *Euclidean* space – so-called complete parallelism (to be discussed in Chapter 15).)

We say that a *vector field* on M has been defined if a rule is given which enables one to choose exactly one vector residing at each point of a manifold M. Only the fields which "do not differ too much" at two "neighboring" points will be of interest for us in what follows (what we need is smoothness of the field). It turns out that this property is most easily formulated after one learns how vector fields act on (the algebra of) functions, i.e. by looking at the matter from an algebraic perspective.

One can apply a vector *field* V to a function f so that at each point $P \in M$ the *vector* $V_P \in T_P M$ (regarded as a linear functional on $\mathcal{F}(M)$, here) is applied to f. In this way we get a *number* $V_P(f)$ residing at each point P of a manifold M, i.e. a *new function* altogether. A vector field thus may be regarded *as a map* (operator)

$$V : \mathcal{F}(M) \to \mathcal{F}(M) \qquad f \mapsto Vf \qquad (Vf)(P) := V_P(f)$$

V is said to be a *smooth vector field* ($\equiv C^\infty$-field) if the image of the map V above is indeed in $\mathcal{F}(M)$, that is to say, if a smooth function results whenever acted on a smooth function by V. The set of (smooth) vector fields on M will be denoted by $\mathfrak{X}(M) \equiv \mathcal{T}_0^1(M)$ (the reason for the second notation will be elucidated in Section 2.5).

2.2.8 Show that the map

$$V : \mathcal{F}(M) \to \mathcal{F}(M) \qquad f \mapsto Vf$$

obeys

$$V(f + \lambda g) = Vf + \lambda Vg$$
$$V(fg) = (Vf)g + f(Vg)$$

($f, g \in \mathcal{F}(M), \lambda \in \mathbb{R}$). The first property alone says that V is a *linear operator* on $\mathcal{F}(M)$; when taken both together they say that V is a *derivation of the algebra of functions* $\mathcal{F}(M)$ (in the sense of Appendix A.2).[13] □

• As is the case for vectors, components may be assigned to vector fields, too. In a given coordinate patch $\mathcal{O}$ with coordinates x^i, a vector field V may be written, according to (2.2.6), in the form

$$V = V^i(x)\partial_i \equiv V^i(x)\frac{\partial}{\partial x^i}$$

[13] The converse is true, too: given *any* derivation D of the algebra of functions $\mathcal{F}(M)$, there exists a vector field V such that $D = V$. This makes it possible to *identify* vector fields on M with derivations of the algebra of functions $\mathcal{F}(M)$.

since the coefficients of a decomposition of a vector with respect to the coordinate basis may be, in general, different at different points (a^i, denoted here as V^i, depend on x). The functions $V^i(x)$ are called the *components* of the field V. The vector fields (!) ∂_i are called the *coordinate basis* of vector fields.

We came to the conclusion, then, that a first-order differential operator with *non-constant* coefficients corresponds to a vector *field* and the action of V on f in coordinates may be expressed simply as

$$f(x) \mapsto (Vf)(x) = V^i(x)(\partial_i f)(x) \equiv V^i(x)\frac{\partial f(x)}{\partial x^i}$$

2.2.9 Prove that V is smooth if and only if its components $V^i(x)$ are smooth functions and that this criterion does not depend on the choice of local coordinates.

Hint: smooth functions are closed with respect to linear combinations and product (= operations in $\mathcal{F}(\mathcal{O})$) elements of $J^i_j(x)$ are smooth. □

2.2.10 Show that under the change of coordinates $x \mapsto x'(x)$ the components of a vector field transform as follows:

$$V'^i(x') = J^i_j(x)V^j(x)$$

Hint: see (2.2.6); $V'^i(x')\partial'_i = V^i(x)\partial_i$. □

2.2.11 Write down the vector field $V = \partial_\varphi$ (in polar coordinates in the plane $\mathbb{R}^2$) in Cartesian coordinates and try to visualize at various points the direction of the vectors given by this field.

Hint: see (2.2.10); $(V = \partial_\varphi = x\partial_y - y\partial_x)$. □

• One should understand clearly the difference between the algebraic properties of a vector and a vector field: a vector is a linear *functional* on $\mathcal{F}(M)$ (a map into $\mathbb{R}$), a vector field is a linear *operator* on $\mathcal{F}(M)$ (a map into $\mathcal{F}(M)$). We have learned in exercise (2.2.5) that the linear functionals on $\mathcal{F}(M)$ comprise a vector space over $\mathbb{R}$, i.e. linear combinations with coefficients from $\mathbb{R}$ are permitted. This kind of combination is permitted for vector fields as well (so that they comprise a real (albeit ∞-dimensional) vector space, too). It turns out, however, that the life of vector fields is *considerably richer*; in particular, one can form linear combinations with coefficients from the *algebra* $\mathcal{F}(M)$. This means (Appendix A.4) that vector fields actually comprise a *module* over the algebra of functions $\mathcal{F}(M)$.

2.2.12 Given $V, W \in \mathfrak{X}(M)$ and $f \in \mathcal{F}(M)$, check that a linear combination $V + fW$ is a vector field, too, if one defines it in terms of a *pointwise combination* of the constituent vectors

$$(V + fW)_P := V_P + f(P)W_P$$

or equivalently (in terms of the action on functions) as

$$(V + fW)g := Vg + f(Wg)$$

□

• If we say, then, that the fields ∂_i constitute a *basis* for vector fields, what we have in mind is that this is a basis *in the sense of a module* (as opposed to a linear space over $\mathbb{R}$). This means that any vector field in a coordinate patch $\mathcal{O} \leftrightarrow x^i$ (this may not hold for the manifold as a whole) may be uniquely decomposed with respect to ∂_i as $V = V^i \partial_i$, the coefficients of the decomposition (components) V^i being, however, from the *algebra* $\mathcal{F}(\mathcal{O})$ ($\mathbb{R}$ is not enough, in general). Thus $\mathfrak{X}(\mathcal{O})$ is an ∞-dimensional linear space over $\mathbb{R}$, but it is, at the same time, *finitely generated* as a module over $\mathcal{F}(\mathcal{O})$. Namely, it has n *generators* (∂_i, for example), from which it may be generated completely by means of the algebra $\mathcal{F}(\mathcal{O})$ in full analogy with an n-dimensional linear space, which may be generated from an arbitrary basis $e_1, \dots, e_n$ with the help of the *field*[14] of real numbers $\mathbb{R}$.

2.2.13* Let L be an n-dimensional linear space over $\mathbb{R}$. Show that

(i) there exists the canonical (independent of the choice of basis in L) isomorphism of L itself and a tangent space $T_x L$ (x being an arbitrary point in L), so that a linear space L may be canonically identified with the tangent space at an arbitrary point

(ii) if a fixed vector $v \in L$ is successively mapped into *all* tangent spaces in this way, the vector *field* V is obtained on L; explicitly (in coordinates introduced in (1.4.11), $v = v^a e_a$) it reads $V = v^a \partial_a$.

Hint: (i) $L \ni v \mapsto (d/dt)_0 (x + tv)$; a picture might be helpful in order to visualize what is going on. □

2.2.14 Let $M \times N$ be a manifold, which is the Cartesian product of two other manifolds M and N. Show that

(i) there is a canonical decomposition of tangent spaces at any point (m, n) into the sum of two subspaces, each of them being isomorphic to the tangent spaces at points m and n respectively of the initial manifolds

$$T_{(m,n)}(M \times N) = T_m M \oplus T_n N$$

(ii) any vector field V on $M \times N$ may be uniquely decomposed into the sum of two vector fields $V = V_M + V_N$, where V_M "is tangent to" M and V_N "is tangent to" N.

Hint: (i) consider the curves $t \mapsto (m(t), n)$ and $t \mapsto (m, n(t))$; in coordinates from (1.3.3) the subspaces span ∂_i and ∂_a; (ii) pointwise realization of (i); $V = A^i(x, y)\partial_i + B^a(x, y)\partial_a \equiv V_M + V_N$. □

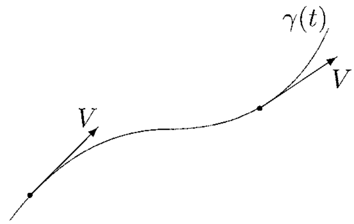
2.3 Integral curves of a vector field

• *Lines of force field* provide an aid for visualizing the field; they are essentially a map of the field. A momentary glance at the pattern of lines provides rich information concerning the field itself, since if $\mathbf{F}(\mathbf{r})$ is the field in question, we know that (by definition) the vector $\mathbf{F}$

[14] The field $\mathbb{R}$ is hidden in the algebra $\mathcal{F}(\mathcal{O})$ in terms of *constant functions*, so that the algebra $\mathcal{F}(\mathcal{O})$ is a much richer object than $\mathbb{R}$ is – this is the reason why far fewer generators are needed to reach the same goal.

at a point **r** is *tangent* to the line of force at **r**. The concept of an integral curve adds a definite *parametrization* to this idea (it is a *curve* rather than a *line*), the latter being irrelevant in the case of force lines: its orientation is all they need.

A vector field V on M determines a vector $V_P \in T_P M$ at each point $P \in M$. On the other hand, we know from Section 2.1 that a vector V_P may be regarded as an equivalence class of curves, each representative of the class "being the same" in the immediate vicinity of the point P (up to order ϵ). An *integral curve* of a vector field V is then the curve γ on M, such that at each point of its image, the equivalence class $[\gamma]$ given by the curve, coincides with the class V_P, given by the value of the field V in P. Put another way, from each point it reaches, it moves away exactly in the direction (as well as with the speed) dictated by[15] the vector V_P. All this may be written as a succinct geometrical equation



$$\dot{\gamma} = V \qquad \text{i.e.} \quad \dot{\gamma}(P) = V_P$$

(this is the equation for finding an integral curve γ of a vector field V in a "coordinate-free" form), where the symbol $\dot{\gamma}(P)$ denotes the *tangent vector* to the curve γ at the point P (i.e. the equivalence class $[\gamma]$, given by the curve γ at the point P). If the vectors on both sides of this equation are decomposed with respect to a coordinate basis, a system of *differential equations* for the functions $x^i(t) \equiv x^i(\gamma(t))$ (for the coordinate presentation of the curve to be found) is obtained.

2.3.1 Show that the differential equations for finding an integral curve γ of a vector field V have the form

$$\dot{x}^i = V^i(x) \qquad i = 1, \dots, n$$

i.e. in more detail

$$\dot{x}^1(t) = V^1(x^1, \dots, x^n)$$
$$\dots$$
$$\dot{x}^n(t) = V^n(x^1, \dots, x^n)$$

Hint: $\dot{\gamma}(\gamma(t)) = \dot{x}^i(t)\, \partial_i|_{\gamma(t)}$, $V_{\gamma(t)} = V^i(x(t))\, \partial_i|_{\gamma(t)}$. □

2.3.2 Write down and solve the equations for integral curves of the field V from exercise (2.2.11), both in polar and in Cartesian coordinates. Draw the solutions ($\dot{r} = 0, \dot{\varphi} = 1;\ \dot{x} = -y, \dot{y} = x$). □

[15] Like a well-disciplined hiker, always walking in the direction of arrows on destination signs and obediently following the instructions concerning time indications given there (how many minutes he or she would need to reach the next arrow).

2.3.3 Find integral curves of the field $V = \partial_x + 2\partial_\varphi$ on $\mathbb{R}[x] \times S^1[\varphi]$ (the surface of a cylinder). Draw the results. □

• We may see that, in general, one has to do with a system of n first-order ordinary differential equations for n unknown functions $x^i(t)$. Moreover, the system is *quasi-linear* (linear in the highest (= the first, here) derivatives), *autonomous* (functions on the right-hand side do not depend explicitly on the variables with respect to which the unknown functions are differentiated (t here)) and, in general, coupled. Since the functions on the right-hand side are smooth (2.2.9), the theory of equations of this type guarantees that there exists a unique solution in some neighborhood of the point, which corresponds to the initial conditions. There then exists a unique integral curve of a field V, which starts at (any given) $P \in M$ in $t = 0$. However, it is not, in general, possible to extend this curve for all values of the parameter $t \in (-\infty, \infty)$.

2.3.4 A vector field V on M is said to be *complete* if for any point $P \in M$ the integral curve $\gamma(t)$, which starts from P, may be extended to all values of the parameter t. Show that the vector fields $V = \partial_x$ on $M = (-1, 1)$ and $W = x^2\partial_x$ on $N = \mathbb{R}$ are *not complete* (and learn a lesson from these two examples, what some problems with such an extension might look like). □

2.3.5 Given $\gamma(t)$, an integral curve of a vector field V on M, let $\hat{\gamma}(t) := \gamma(\sigma(t))$ be a *reparametrized* curve. Find the most general dependence $\sigma(t)$, so that $\hat{\gamma}$ will be an integral curve of the vector field V, too.

Hint: $(d/dt)\, f(\gamma(\sigma(t))) = \sigma'(t)(d/dt)\, f(\gamma(t))$, so that $\dot{\hat{\gamma}} = \sigma'\dot{\gamma}$; $[\sigma(t) = t + \text{constant}]$. □

• This result is easy to understand. Consider $\gamma(t)$ as being a trajectory. Then $\hat{\gamma}$ is another trajectory, such that we traverse the same set of points on M at different moments of time. Put another way, the *path* remains unchanged, but the (instantaneous) *speed* of traversing the path may be different.[16] Just how much different depends on the point and the result of the exercise shows that the new speed is $\sigma'(t)$ times the old one at any point $\gamma(t)$. (As an example, for $\sigma(t) = 2t$, the new speed is *twice* the old one at each point.) Since the velocity vector of an integral curve *may not* be changed (it is given by V uniquely), $\sigma'(t) = 1$ results. This means that the only possibility to change the trajectory is to traverse the same path either *sooner* or *later*. This freedom ($t \mapsto t +$ constant) enables one to set an arbitrary value of the parameter t (time) at the starting point P.

2.3.6 Let γ be an integral curve of a vector field V on M, which starts from $P \equiv \gamma(0) \in M$. Show that the integral curve (of the same field V) $\hat{\gamma}$, which starts from $Q \equiv \gamma(a)$, is $\hat{\gamma}(t) := \gamma(t + a)$.

Hint: see (2.3.5). □

[16] In fact, we have not enough structure, yet, to speak of the "speed" (a metric tensor, to be introduced later, is needed for it). In spite of this, we *can* speak of the *ratio* of two speeds, since our velocity vectors are *proportional*.

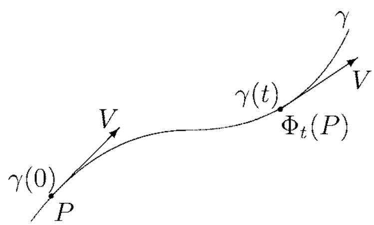
• The result of (2.3.1) admits a different interpretation, too. It shows that *each* system of equations of the type (2.3.1) may be regarded as a system for finding integral curves of the particular vector field (we read out its components from the right-hand side of the equations). This is the important observation, since it provides a key to the investigation of properties of solutions of such equations by powerful geometrical and topological methods – corresponding vector fields (or other objects associated with them) are studied instead of the equations themselves. We will see this, for example, in Chapter 14, where Hamiltonian systems will be discussed.

2.3.7 Find a vector field V on $\mathbb{R}^{2n}[q^1, \ldots, q^n, p_1, \ldots, p_n]$, which corresponds to the *Hamilton equations*

$$\dot{q}^a = \frac{\partial H}{\partial p_a} \qquad \dot{p}_a = -\frac{\partial H}{\partial q^a} \quad a = 1, \ldots, n$$

$(V = (\partial H/\partial p_a)\partial/\partial q^a - (\partial H/\partial q^a)\partial/\partial p_a)$. □

• A vector field V on a manifold M gives rise to a new and interesting structure, a *congruence* of integral curves on M: the manifold M is "densely" filled by a system of (infinitely many) curves, which never intersect and the "speed" of motion along them is completely determined by the field V. This situation may be conveniently visualized as the *flow of a river*. This flow is *stationary* (the velocity vector in a given point being always the same; in particular, the river does not flow at the points where the field vanishes) and for particular types of fields (e.g. for Hamiltonian fields) the fluid is in addition *incompressible* (14.3.6). Integral curves correspond to the streamlines of the flow. If one fine (and hot) afternoon we do not resist the temptation and let ourselves waft downstream, we get from $P \equiv \gamma(0) \in M$ to the point $Q \equiv \gamma(t) \in M$; naturally a one-parameter class of mappings



$$\Phi_t : M \to M \qquad P \equiv \gamma(0) \mapsto \gamma(t)$$

arises, called a (local) *flow* generated by the vector field V. We will return to this important concept in more detail later, in Chapter 4 and beyond.

2.3.8 Justify the statement mentioned above, that integral curves never intersect (nor are tangent to one another).

Hint: from a point P one has to make a move in the direction of V_P (uniquely). □

2.3.9 Express the results of exercises (2.3.2) and (2.3.3) in the form of a flow $\Phi_t : x^i \mapsto x^i(t) \equiv \Phi_t(x^i)$ $((r, \varphi) \mapsto (r, \varphi + t)$ or $(x, y) \mapsto (x \cos t - y \sin t, x \sin t + y \cos t)$; $(x, \varphi) \mapsto (x + t, \varphi + 2t))$. □

2.4 Linear algebra of tensors (multilinear algebra)

• It turns out that each linear space L automatically gives rise to a fairly rich algebraic structure "above" L – a whole infinite cascade of further and further linear spaces, the spaces of *tensors in* L and an ∞-dimensional associative graded algebra, the *tensor algebra* $T(L)$, associated with them. In this section we will become familiar with tensors at the level of linear algebra, and in the next section we shift to manifolds and introduce the concept of a tensor field.

Within this section we consider arbitrary n-dimensional linear space L over the field of real numbers $\mathbb{R}$.

First, we observe that *linear forms* on L, i.e. linear maps such that

$$\alpha : L \to \mathbb{R} \qquad \alpha(v + \lambda w) = \alpha(v) + \lambda\alpha(w) \quad v, w \in L, \lambda \in \mathbb{R}$$

form a linear space in its own right, the *dual space* L^*. Its elements are called *covectors* in L.

2.4.1 Check that the prescription

$$(\alpha + \lambda\beta)(v) := \alpha(v) + \lambda\beta(v)$$

introduces a linear structure in L^* (i.e. check that the linear combination is indeed a linear map $L \to \mathbb{R}$). □

• The resulting value of $\alpha(v) \in \mathbb{R}$ will be denoted, as a rule, in the form

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

Given a basis e_a in L, there already exists the distinguished basis in L^* (tailored to the basis e_a in L).

2.4.2 Let e_a be a basis in L and let $v = \sum_{b=1}^{n} v^b e_b \equiv v^b e_b$. Verify that

(i) the maps

$$e^a : L \to \mathbb{R} \qquad e^a(v^b e_b) := v^a \quad a = 1, \dots, n$$

are covectors and, in addition, they constitute a basis in L^* (called the *dual basis* with respect to e_a):

$$\alpha = \alpha_a e^a \qquad \alpha_a := \langle \alpha, e_a \rangle$$

(ii) an equivalent definition of the dual basis is

$$\langle e^a, e_b \rangle = \delta^a_b$$

(iii) a change of the basis in L given by a matrix A results in the change of the dual basis given by the *inverse* matrix A^{-1}

$$e_a \mapsto e'_a = A^b_a e_b \quad \Rightarrow \quad e^a \mapsto e'^a = (A^{-1})^a_b e^b$$

(iv) the dimension of a dual space equals the dimension of the original space: $\dim L^* = \dim L\ (= n)$.

Hint: (i) check that α and $\langle \alpha, e_a \rangle e^a$ are equal linear maps; (iv) consider the number of elements of the dual basis. □

• Since L^* is an n-dimensional vector space in its own right, the whole story may be repeated again and one can construct the dual space $(L^*)^*$. It turns out, however, that this space is (for finite-dimensional L) in a sense redundant. The reason is that it is *canonically* isomorphic to the original space L. What do we mean by this and how can one profit from it?

In general, any two n-dimensional linear spaces are isomorphic, but there are an infinite number of equally good isomorphisms available ($e_a \mapsto E_a$, for *arbitrary* choice of basis E_a), so that there is no reasonable (independent of arbitrary choices) way to choose a preferred one. This is true, in particular, for the relation $L \leftrightarrow L^*$. (Try, for example, to describe your favorite isomorphism to a remote extraterrestrial, who is well educated in linear algebra and understands all the steps you dictate.) Exercise (2.4.3) shows, however, that for $L \to (L^*)^*$ the situation is essentially different. In this case, there is a *distinguished* isomorphism f, which *can* be described to our remote extraterrestrial friend and he or she or it *will know* what maps into what. This isomorphism suggests using a standard mathematical trick – *identification* of the spaces L and $(L^*)^*$, and, by analogy then, the nth with the $(n-2)$th dual spaces. Only the first two members, L and L^*, thus survive from the threatening looking, potentially infinite chain of still higher and higher dual spaces. (This, in a moment, will result in the fact that we will make do with only *two kinds* of indices, "lower" and "upper," on general tensors.)[17] If a non-degenerate bilinear form were *added* to L, the situation would change significantly, since it *would* be possible already to identify L with L^* in a *canonical* way (via the "raising and lowering of indices" procedure, see (2.4.13).)

2.4.3 Prove that the space $(L^*)^*$ is canonically isomorphic to the space L.

Hint: the canonical isomorphism $f : L \to (L^*)^*$ is $\langle f(v), \alpha \rangle := \langle \alpha, v \rangle$. □

2.4.4 Imagine we have defined a "canonical" isomorphism $L \leftrightarrow L^*$ with the help of *dual* bases by

$$f(e_a) := e^a$$

(i.e. $v \leftrightarrow \alpha$, if they have equal coefficients of decomposition with respect to e_a and e^a respectively). Check that if we change the basis as $e_a \mapsto A^b_a e_b$, the isomorphism above will be *changed* (and since in general L all bases are equally good, no distinguished f is given in this way). □

[17] This step saves the huge number of higher dual spaces as well as various kinds of indices for future generations, so it can be regarded as highly satisfactory far-sighted behavior from an ecological point of view; one should not lavishly waste any non-renewable resources, including mathematical structures.

• Let us have a look at one aspect, common for linear spaces L, L^* and $\mathbb{R}$. One may, in all three cases, regard their elements as *linear maps into* $\mathbb{R}$, namely

1. $\alpha \in L^*$ maps $v \mapsto \langle \alpha, v \rangle \in \mathbb{R}$ $(v \in L)$
2. $v \in L$ maps $\alpha \mapsto \langle \alpha, v \rangle \in \mathbb{R}$ $(\alpha \in L^*)$
3. $a \in \mathbb{R}$ maps $\emptyset \mapsto a \in \mathbb{R}$ (*no* input and a real number as output).

Although item 3 might look fairly far-fetched, it proves convenient to incorporate it as a gear-wheel into a device, which in general operates as follows: *several vectors* as well as *covectors* are inserted and (after a crank is turned, of course) a real number drops out. Moreover, if this number depends *linearly* on *each* argument (which holds for all three cases, albeit trivially for the third case), we get a *tensor*.

Definition Let L be an n-dimensional linear space and L^* its dual space. A *tensor of type* $\binom{p}{q}$ in L is a *multilinear* ($\equiv$ *polylinear* := linear in each argument) map

$$t : \underbrace{L \times \cdots \times L}_{q} \times \underbrace{L^* \times \cdots \times L^*}_{p} \to \mathbb{R}$$

$$(\underbrace{v, \ldots, w}_{q}; \underbrace{\alpha, \ldots, \beta}_{p}) \mapsto t(v, \ldots, w; \alpha, \ldots, \beta) \in \mathbb{R}$$

$$t(\ldots, v + \lambda w, \ldots) = t(\ldots, v, \ldots) + \lambda t(\ldots, w, \ldots)$$

(and similarly for an arbitrary covector argument). A collection of tensors of type $\binom{p}{q}$ in L will be denoted by $T^p_q(L)$, and for $p = q = 0$ we set $T^0_0(L) := \mathbb{R}$.

2.4.5 Check that

(i) for $t, \tau \in T^p_q(L)$, $\lambda \in \mathbb{R}$, the rule

$$(t + \lambda\tau)(v, \ldots; \alpha, \ldots) := t(v, \ldots; \alpha, \ldots) + \lambda\tau(v, \ldots; \alpha, \ldots)$$

introduces a linear structure into $T^p_q(L)$ (i.e. the linear combination displayed above indeed happens to be a multilinear map)

(ii) some special instances are given by

$$T^0_0(L) = \mathbb{R} \qquad T^0_1(L) = L^* \qquad T^1_0(L) \approx L$$

$$T^1_1(L) \approx \mathrm{Hom}\,(L, L) \approx \mathrm{Hom}\,(L^*, L^*) \qquad T^0_2(L) = \mathcal{B}_2(L)$$

where $\mathrm{Hom}\,(L_1, L_2)$ denotes all linear maps from L_1 into L_2, $\mathcal{B}_2(L)$ are bilinear forms on L and $\approx$ denotes canonical isomorphism.

Hint: $\binom{0}{0}$, $\binom{0}{1}$ and $\binom{0}{2}$ definitions, $\binom{1}{0}$ (2.4.3); $\binom{1}{1}$: the isomorphisms $\mathrm{Hom}(L, L) \to T^1_1(L)$ and $\mathrm{Hom}\,(L^*, L^*) \to T^1_1(L)$ read

$$t(v; \alpha) := \langle \alpha, A(v) \rangle \qquad \text{and} \qquad t(v; \alpha) := \langle B(\alpha), v \rangle$$

or, equivalently (in the opposite direction),

$$A(v) := t(v; \cdot\,) \qquad B(\alpha) := t(\,\cdot\,; \alpha)$$

□

• Taking into account (multi)linearity, a tensor $t \in T^p_q(L)$ is known completely if we know its values on all possible combinations of basis vectors e_a and covectors e^a. This collection of numbers

$$t^{c\dots d}_{a\dots b} := t(e_a, \dots, e_b; e^c, \dots, e^d)$$

is said to form the *components* of the tensor t with respect to e_a. The mnemonic rule of the notation $\binom{p}{q}$ should finally be clear: a tensor t is in the space $T^p_q(L)$ if its components have p upper indices and q lower indices.

2.4.6 Check that

(i) in components, the rule for performing linear combinations from (2.4.5) reduces to

$$(t + \lambda\tau)^{c\dots d}_{a\dots b} = t^{c\dots d}_{a\dots b} + \lambda\tau^{c\dots d}_{a\dots b}$$

(ii) $\dim T^p_q(L) = n^{p+q} \equiv (\dim L)^{p+q}$ (the number $(p+q)$ is known as the *rank* of a tensor)

(iii) under the change of basis in L, components of a tensor transform as follows:

$$e_a \mapsto A^b_a e_b \equiv e'_a \quad \Rightarrow \quad t'^{c\dots d}_{a\dots b} = (A^{-1})^c_k \dots (A^{-1})^d_l A^r_a \dots A^s_b t^{k\dots l}_{r\dots s}$$

(iv) if $v = v^a e_a, \alpha = \alpha_a e^a \dots$ represent the decompositions of arguments, then

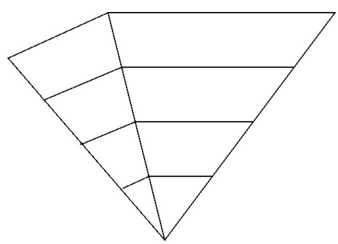
$$t(v, \dots, w; \alpha, \dots, \beta) = t^{c\dots d}_{a\dots b} v^a \dots w^b \alpha_c \dots \beta_d$$

(v) three different applications of a $\binom{1}{1}$-type tensor t from (2.4.5) in components look like

$$(v^b, \alpha_a) \mapsto t^a_b v^b \alpha_a \qquad v^a \mapsto t^a_b v^b \qquad \alpha_a \mapsto t^b_a \alpha_b$$

Hint: (ii) $t \mapsto t^{c\dots d}_{a\dots b}$ is the isomorphism $T^p_q(L) \to \mathbb{R}^{n^{p+q}}$ (each of $(p+q)$ indices takes n values); (iii) $t'^{c\dots d}_{a\dots b} := t(e'_a, \dots, e'^d)$ + linearity in each argument. □

• Thus we have learned that L induces an infinite number of further linear spaces – for each pair (p, q) of non-negative integers there is the n^{p+q}-dimensional space $T^p_q(L)$. (This means that if we envisage tensor spaces as a "tower," the tower dilates in the upward direction, like a pyramid does on a photograph snapped in Giza by a distrait yogi, forgetting he has just performed a headstand.)



If we combine components with a suitable basis, we get "complete" tensors. It turns out that a suitable basis may be constructed out of the basis for vectors and covectors, if an additional operation on tensors is introduced, the *tensor product*. It may be regarded as a map

$$\otimes : T^p_q(L) \times T^{p'}_{q'}(L) \to T^{p+p'}_{q+q'}(L)$$

i.e. two tensors of *arbitrary* types $\binom{p}{q}$ and $\binom{p'}{q'}$ are multiplied – contrary to linear combination, where both types have to be equal – and the resulting tensor is of type $\binom{p+p'}{q+q'}$. The definition is as follows:

$$(t \otimes \sigma)(v_1, \dots, v_q, w_1, \dots, w_{q'}; \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_{p'})$$
$$:= t(v_1, \dots v_q; \alpha_1, \dots, \alpha_p)\sigma(w_1, \dots, w_{q'}; \beta_1, \dots, \beta_{p'})$$

(here the indices label *complete* vectors and covectors, rather than their components!). Stated in words, we first insert the arguments of both types into the first (left) tensor, until it is filled completely; the rest we put into the second (right) one. The resulting two *numbers* are then simply multiplied.

2.4.7 Verify that

(i) the result of the multiplication $t \otimes \sigma$ is a tensor, indeed (i.e. check multilinearity)
(ii) at the level of components the multiplication $\otimes$ gives

$$(t \otimes \sigma)^{c\dots dr\dots s}_{a\dots bk\dots l} = t^{c\dots d}_{a\dots b}\sigma^{r\dots s}_{k\dots l}$$

(iii) the multiplication $\otimes$ is associative (we need not bother about brackets in *multiple* products), bilinear and *non*-commutative
(iv) tensors of type (p, q)

$$e^a \otimes \dots \otimes e^b \otimes e_c \otimes \dots \otimes e_d \in T^p_q(L)$$

constitute the basis of $T^p_q(L)$ with respect to which components have been defined above, i.e. an arbitrary tensor $t \in T^p_q(L)$ may be decomposed as

$$t = t^{c\dots d}_{a\dots b} e^a \otimes \dots \otimes e^b \otimes e_c \otimes \dots \otimes e_d \qquad t^{c\dots d}_{a\dots b} := t(e_a, \dots, e_b; e^c, \dots, e^d)$$

Hint: (iv) one has to check that the "original" tensor and its decomposition represent *the same map*; since they are (multi)linear, it is enough to check it for the basis; as an example

$$\left(t^c_d e^d \otimes e_c\right)(e_a; e^b) = t^c_d \langle e^d, e_a\rangle\langle e^b, e_c\rangle = t^b_a = t(e_a; e^b)$$

thus the equality $t^c_d e^d \otimes e_c = t$ *of maps* (= tensors) has been proved. □

• The result (2.4.7) shows that all tensors constitute an (∞-dimensional non-commutative) *associative algebra* (Appendix A.2), called the *tensor algebra* $T(L)$. As a linear space, it is a *direct sum* of all spaces $T^p_q(L)$

$$T(L) := \bigoplus_{r,s=0}^{\infty} T^r_s(L)$$
$$\equiv T^0_0(L) \oplus T^1_0(L) \oplus T^0_1(L) \oplus T^2_0(L) \oplus T^1_1(L) \oplus T^0_2(L) \oplus \cdots$$

(up to infinity), i.e. an element from $T(L)$ may be regarded as a linear combination of tensors of all types $\binom{p}{q}$. Multiplication $\otimes$ is defined as a linear extension of the definition of $\otimes$ on *homogeneous terms* (terms with fixed $\binom{p}{q}$), i.e. according to the rule "everybody with everybody”:[18]

$$(k + v + \alpha + \cdots) \otimes (q + w + \beta + \cdots) := k \otimes q + k \otimes w + k \otimes \beta + \cdots \\ + v \otimes q + v \otimes w + \cdots$$

Furthermore, this algebra is $(\mathbb{Z} \times \mathbb{Z})$-*graded* (Appendix A.5): its “homogeneous” subspaces $T^p_q(L)$ are labelled by a *pair* of integers (p, q), i.e. (we define $T^p_q(L) := 0$ for negative p, q) by an element of group $\mathbb{Z} \times \mathbb{Z}$, and multiplication in algebra $T(L)$ is compatible with the grading: the product of any two elements from the subspaces $\leftrightarrow (p, q)$ and $(p', q') \in \mathbb{Z} \times \mathbb{Z}$ is homogeneous, too, belonging to the subspace which corresponds to a *product in the sense* of $\mathbb{Z} \times \mathbb{Z}$, i.e. $(p + p', q + q')$.

Operations producing tensors from tensors, are said to be *tensor operations*. So far we have met linear combination and tensor product. One further important tensor operation is provided by *contraction*. It is defined (for $p, q \geq 1$) as follows:

$$C : T^p_q(L) \to T^{p-1}_{q-1}(L) \qquad t \mapsto Ct := t(\dots, e_a, \dots; \dots, e^a, \dots)$$

where the exact position of arguments e_a and e^a is to be specified – it forms a part of the definition (there are several (pq) various possible contractions, in general, and one has to state *which one* is to be performed).

2.4.8 Check that

(i) the result is indeed a tensor (multilinearity)

(ii) C does not depend on the choice of the basis e_a (when e_a has been fixed, however, e^a is to be the dual)

(iii) in components the rule for C looks like[19]

$$t^{\cdots\cdots}_{\cdots\cdots} \mapsto t^{\cdots a \cdot}_{\cdot a \cdots} \qquad \text{i.e.} \quad \text{as a } \textit{summation} \text{ with respect to a pair of upper and lower indices}$$

(iv) independence of a choice of basis results from the component formula, too.

Hint: (ii) see (2.4.2); (iv) see (2.4.6). □

2.4.9 Show that

(i) the prescription

$$\hat{1}(V; \alpha) := \langle \alpha, V \rangle$$

defines a $\binom{1}{1}$-type tensor, the *unit* tensor

(ii) its components with respect to *any* basis e_a (e^a being dual, as usual) are given by

$$\hat{1}^a_b = \delta^a_b \qquad \text{so that} \quad \hat{1} = e^a \otimes e_a$$

[18] The maximum promiscuity rule.

[19] Each contraction thus unloads a tensor by two indices. It breathes with fewer difficulties immediately (fewer indices = fewer worries), it feels like after a rejuvenation cure. This human aspect of the matter is reflected sensitively in German terminology, where the word *Verjüngung* (rejuvenescence) is used.

(iii) it realizes the *unit* operator ($v \mapsto v, \alpha \mapsto \alpha$) if it is interpreted as a map

$$\hat{1} : L \to L \qquad \text{and} \qquad \hat{1} : L^* \to L^*$$

respectively

(iv) its contraction (2.4.8) gives

$$C\hat{1} = n \equiv \dim L$$

Hint: (iii) see (2.4.5). □

2.4.10 Show that the evaluation of a tensor on arguments may be regarded as a composition of tensor product and contractions; as an example, for a $\binom{1}{1}$-type tensor it is

$$t(v, \alpha) = CC(t \otimes v \otimes \alpha) = (t \otimes v \otimes \alpha)^{ab}_{ba} \equiv (t \otimes v \otimes \alpha)(e_b, e_a; e^a, e^b)$$

In particular, (see exercise 2.4.8),

$$\hat{1}(v, \alpha) \equiv \langle \alpha, v \rangle = C(\alpha \otimes v)$$ □

- A *metric tensor* in L is a symmetric non-degenerate tensor of type $\binom{0}{2}$, i.e. $g \in T^0_2(L)$ such that

$$g(v, w) = g(w, v) \qquad \text{symmetric}$$
$$g(v, w) = 0 \text{ for all } w \Rightarrow v = 0 \qquad \text{non-degenerate}$$

2.4.11 Check that

(i)

$$g_{ab} = g_{ba} \qquad \det g_{ab} \neq 0$$

(ii) conditions in (i) do not depend on the choice of basis e_a. □

- Sometimes one demands that g meets stronger requirements, namely to be *positive definite*,[20] so that

$$g(v, v) \geq 0 \qquad \text{(and equality holds } only \text{ for } v = 0)$$

and (metric) tensors, which *are not* positive definite, are said to be *pseudo-metric* tensors. We will use, in what follows, the nomenclature *metric* tensor also for g, which is *not* positive definite,[21] and if some statement relies heavily on the positive definiteness of the latter (i.e. "true" metric tensor), it will be specially emphasized.

As is well known from linear algebra, one can bring a matrix of a general symmetric bilinear form by a suitable (non-unique) choice of basis e_a to the canonical form

$$b_{ab} = \text{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}, \underbrace{0, \ldots, 0}_{l})$$

[20] Then $(v, w) := g(v, w)$ has the properties of a *scalar product* in L, see (2.4.13).

[21] This is the case both in special and in general relativity, where one speaks of a "metric" in situations where in finer terminology *pseudo*-metric tensor (or even tensor *field*) should be used.

where the numbers (r, s, l) are inherent properties of the form (*Sylvester's theorem*). Non-degeneracy adds $l = 0$ (why?), so that the canonical form of a *metric tensor* reads as

$$g_{ab} = \eta_{ab} \equiv \text{diag}\,(\underbrace{1, \dots, 1}_{r}, \underbrace{-1, \dots, -1}_{s})$$

or, in other words,

$$\begin{aligned} g &= g_{ab} e^a \otimes e^b \\ &= e^1 \otimes e^1 + \cdots + e^r \otimes e^r - e^{r+1} \otimes e^{r+1} - \cdots - e^{r+s} \otimes e^{r+s} \end{aligned}$$

In this case we will speak about a metric tensor with *signature* (r, s).[22] Thus, the positive definite case corresponds to $s = 0$ (terms with a minus sign are not present in the canonical form). Any basis $e_a \leftrightarrow e^a$ in which this canonical form of g is obtained is called an *orthonormal basis*.

2.4.12 Given e_a an arbitrary basis and $g_{ab} = g(e_a, e_b)$, define g^{ab} as elements of the *inverse matrix* to g_{ab}, i.e.

$$g^{ac} g_{cb} := \delta^a_b$$

Prove that

(i) g^{ab} constitute the components of a (symmetric) $\binom{2}{0}$-type tensor (so that they indeed deserve two upper indices)

$$g \equiv g_{ab} e^a \otimes e^b \in T^0_2(L) \quad \Rightarrow \quad g^{-1} := g^{ab} e_a \otimes e_b \in T^2_0(L)$$

(ii) matrix g^{ab} is non-singular.

Hint: (i) check the transformation law of g^{ab} under a change of basis. □

2.4.13 Consider the maps $\flat_g$ and $\sharp_g$ given by

$$\begin{aligned} \flat_g &: L \to L^* & v &\mapsto \flat_g v := g(v, \cdot) \\ \sharp_g &: L^* \to L & \alpha &\mapsto \sharp_g \alpha := g^{-1}(\alpha, \cdot) \end{aligned}$$

Check that

(i) they are linear (and canonical) isomorphisms
(ii) when expressed in bases and in components, they look like

$$\begin{aligned} \flat_g &: \quad e_a \mapsto g_{ab} e^b \qquad v^a \mapsto v_a := g_{ab} v^b \qquad v^a e_a \mapsto v_a e^a \\ \sharp_g &: \quad e^a \mapsto g^{ab} e_b \qquad \alpha_a \mapsto \alpha^a := g^{ab} \alpha_b \qquad \alpha_a e^a \mapsto \alpha^a e_a \end{aligned}$$

(iii) they are inverse to each other:

$$\flat_g \circ \sharp_g = \text{id}_{L^*} \qquad \sharp_g \circ \flat_g = \text{id}_L$$

[22] Sometimes, the number $r - s$ is called the signature, too.

(iv) if *scalar products* in L and L^* are introduced[23] by

$$(v, w) := g(v, w) \equiv g_{ab} v^a w^b \qquad (\alpha, \beta) := g^{-1}(\alpha, \beta) \equiv g^{ab} \alpha_a \beta_b$$

then both $\flat_g$ and $\sharp_g$ are *isometries*, i.e. $(\flat_g v, \flat_g w) = (v, w)$, $(\sharp_g \alpha, \sharp_g \beta) = (\alpha, \beta)$. □

• The maps $\flat_g$ and $\sharp_g$ are known as *lowering* and *raising of indices* (with the help of g), respectively. The quantities v_a, v^a are often called covariant and contravariant *components* of (the same) vector v. We will not adopt this nomenclature, however. We will always strictly discriminate between a *vector* $v = v^a e_a$ and a *covector* $v_a e^a$ (as being elements of L and L^*) and interpret the operations of raising and lowering of indices as maps between two *different* spaces $L \leftrightarrow L^*$. Note that the graphical expressions used for these maps originate from well-known musical symbols.[24]

The metric tensor makes it possible to change the position of indices on higher rank tensors, too, for example

$$t^a_{\ bc} \mapsto t_{abc} := g_{ad}\, t^d_{\ bc} \qquad R^{ab}_{\ \ cd} \mapsto R_{abcd} := g_{ae}\, g_{bf}\, R^{ef}_{\ \ cd}$$

This belongs to basic exercises of *index gymnastics*.[25]

2.4.14 Prove the validity of the exercise

$$t^{....a.}_{..a...} = t^{..a...}_{....a.}$$

Hint: do you intend to base your proof upon the fact that the total potential energy remains unchanged? (Red herring.) □

• There are several possibilities of how to raise or lower indices on second or higher rank tensors, differing in the order of the indices on the resulting tensor. As an example, there are four places below where one can lower the index on the fourth rank tensor $R^a_{\ bcd}$

$$R_{abcd} := g_{aj} R^j_{\ bcd} \qquad R_{abcd} := g_{bj} R^j_{\ acd} \qquad \dots$$

The indices are sometimes written so as to have *only one* index on each vertical line, being *either* upper *or* lower, e.g. $R_a{}^b{}_{cd}$. Within this particular convention, it is always clear where exactly any upper index should be lowered.

It is useful to realize that symmetry of the metric tensor g is of no importance for raising and lowering of indices, the *only* property that matters being its *non-degeneracy*. These operations might as well be defined by virtue of an *anti*symmetric tensor $\omega_{ab} = -\omega_{ba}$, provided that it happens to be non-degenerate ($\det \omega_{ab} \neq 0$). We will see in what follows that this possibility is indeed exploited, the most prominent applications being in symplectic geometry (to be discussed in Chapter 14 and beyond) and in the theory of two-component spinors (12.5.3).

[23] They are positive definite for *Euclidean* g only!
[24] Namely "flat" and "sharp." Thoughtful graduates of schools of music might recall that *no* g *was present* on sharps and flats they had read in sheets of music – this is simply because the validity of *Euclidean* geometry is normally assumed in concert halls, so that musical flats and sharps are conventionally associated with *this Euclidean* g (and are not indicated explicitly).
[25] It should be performed, as is the case for arbitrary gymnastics, at an open window, never directly after a substantial meal.

Finally, let us contemplate whether the lowering and raising of indices does change the numerical values of components. The formula $v^a \mapsto v_a \equiv g_{ab}v^b$ shows that the numbers v^a and v_a are the same only in the case where g is given, in a given basis, by the *identity* matrix, $g_{ab} = \delta_{ab}$. This (only) happens to be true in the *positive definite* case in the *orthonormal* basis; in the indefinite case, this happens *in no basis*. Therefore, when working with vectors in Euclidean spaces E^2 or E^3, one may safely ignore the detailed position (upper/lower) of indices with respect to an *orthonormal basis*.[26] On the other hand, one should pay due attention to this issue in all cases when non-orthonormal bases or indefinite metrics are used. In Minkowski space, for example, the lowering and raising of indices *always* changes numerical values of (some) components; in an orthonormal basis this change reduces to the change of a *sign* (of some of them), but it may be more complicated in general.

2.4.15 Check that raising and lowering of indices

(i) are tensor operations
(ii) may be regarded as compositions of a tensor product (with the tensor g) and contractions.

Hint: e.g. $\flat_g v \equiv g(v, \cdot) = C(g \otimes v)$. □

• The last tensor operations to be mentioned are symmetrizations and antisymmetrizations in various subgroups of indices. Let us illustrate this on just two indices.

2.4.16 Given $t \in T^0_2(L)$, define

$$t^{\mathrm{S}} := \frac{1}{2}(t_{ab} + t_{ba})e^a \otimes e^b \equiv t_{(ab)}e^a \otimes e^b$$
$$t^{\mathrm{A}} := \frac{1}{2}(t_{ab} - t_{ba})e^a \otimes e^b \equiv t_{[ab]}e^a \otimes e^b$$

(symmetric and antisymmetric part of the tensor t respectively). Check that

(i)

$$t \mapsto t^{\mathrm{S}} \equiv \pi^{\mathrm{S}} t \qquad t \mapsto t^{\mathrm{A}} \equiv \pi^{\mathrm{A}} t$$

are tensor operations, independent of the choice of e_a
(ii) tensors, for which $t = t^{\mathrm{S}}$ or $t = t^{\mathrm{A}}$ is true, constitute subspaces in $T^0_2(L)$
(iii) π^{S} and π^{A} satisfy

$$\pi^{\mathrm{S}} \circ \pi^{\mathrm{S}} = \pi^{\mathrm{S}} \qquad \pi^{\mathrm{S}} \circ \pi^{\mathrm{A}} = \pi^{\mathrm{A}} \circ \pi^{\mathrm{S}} = 0$$
$$\pi^{\mathrm{A}} \circ \pi^{\mathrm{A}} = \pi^{\mathrm{A}} \qquad \pi^{\mathrm{S}} + \pi^{\mathrm{A}} = \hat{1}$$

so that they serve as projection operators on the subspaces of the symmetric and antisymmetric tensors mentioned above, the whole space $T^0_2(L)$ being the direct sum of these two subspaces (only). □

• Finally, two more useful concepts will be introduced at the end of this section on multilinear algebra, namely those of a dual map and an induced metric tensor.

[26] That is, at the level of components one is allowed to make no difference between a vector and the associated covector, like the gradient as a covector and a gradient as a vector, see the end of Section 2.6.

2.4.17 Let $A : L_1 \to L_2$ be a linear map, e_i a basis of L_1 and e_a a basis of L_2. The *rank* of the map A is defined as a dimension of the image of the space L_1 in L_2, i.e. rank $A := \dim \operatorname{Im} A$. Show that

(i) by the prescription

$$\langle A^*(\alpha_2), v_1\rangle := \langle \alpha_2, A(v_1)\rangle \qquad \alpha_2 \in L_2^*, v_1 \in L_1$$

a linear map

$$A^* : L_2^* \to L_1^*$$

is defined (*dual map*)

(ii) on the basis it gives

$$e_i \overset{A}{\mapsto} A_i^a e_a \quad \Rightarrow \quad e^a \overset{A^*}{\mapsto} A_i^a e^i$$

i.e. matrices of the maps A, A^* are *transposes* of each other

(iii)

$$\text{rank } A = \text{rank of the } \textit{matrix} \text{ of a map } A$$
$$\text{rank } A^* = \text{rank of the } \textit{matrix} \text{ of a map } A^*$$

(iv) rank $A =$ rank A^* ($\Rightarrow$ that the *row* and *column* ranks of a matrix happen to coincide).

Hint: (iv) use *adapted* bases: a part of e_i is a basis of the *kernel* Ker A of the map (those v for which $v \mapsto 0 \in L_2$), the rest are chosen arbitrarily to complete a basis; in L_2 take images of the remaining part (they span Im A) + complete a basis. □

2.4.18 Given $A : L_1 \to (L_2, h)$, $\dim L_1 \leq \dim L_2$ a maximum rank linear map (2.4.17) (h being a metric tensor in L_2), show that

(i) by the rule

$$g := A^* h \qquad (A^* h)(v, w) := h(Av, Aw)$$

a metric tensor g in L_1 is defined (*induced metric tensor*)

(ii) if $e_i \in L_1$ and $e_a \in L_2$ are bases, then

$$g_{ij} = A_i^a h_{ab} A_j^b \qquad Ae_i =: A_i^a e_a \qquad \text{(in matrix notation } g = A^{\mathrm{T}} h A)$$

Hint: (i) (among others) one has to check the maximum rank (2.4.13) of the map

$$\flat_g : L_1 \to L_1^* \qquad v \mapsto g(v, \cdot\,) \equiv \tilde{v} \equiv \flat_g v$$

($\equiv$ non-degeneracy of g). This map is a composition of

$$\flat_g = A^* \circ \flat_h \circ A \qquad L_1 \overset{A}{\to} L_2 \overset{\flat_h}{\to} L_2^* \overset{A^*}{\mapsto} L_1^* \qquad (A^* \text{ in the sense of (2.4.17)})$$

(since $e_i \mapsto g_{ij} e^j = A_i^a h_{ab} A_j^b e^j$), all factors in the composition do have maximum rank and $\dim L_1 \leq \dim L_2 \Rightarrow \flat_g$ is a maximum rank ($= \dim L_1$) map, too. □

2.4.19 Let V be a linear space with a *distinguished subspace* $W \subset V$. Show that in the dual space V^* the associated distinguished subspace $\hat{W} \subset V^*$ of dimension dim V minus dim W is given canonically; it is said to be an *annihilator* of the subspace W.

Hint: consider covectors $\sigma \in V^*$ annihilated by vectors from W, i.e. such that $\langle \sigma, w \rangle = 0$ for all $w \in W$ (see also (10.1.13)). □

2.5 Tensor fields on M

• In Section 2.2 we showed that there is a vector space associated with each point P of a manifold M, the tangent space $T_P M$. In Section 2.4 we learned how to construct tensors of type $\binom{p}{q}$, starting from an *arbitrary* finite-dimensional vector space L. If we now take L to be the space $T_P M$, we immediately get (with practically no labor – it simply suffices to harvest the crop sown earlier in Section 2.4) tensors at the point $P \in M$. In particular, the dual space to $T_P M$, the space of *covectors* in $P \in M$, is called the *cotangent space* in P and it is denoted by $T_P^* M$.

Equally naturally the concept of a *tensor field* of type $\binom{p}{q}$ on M appears. In full analogy with the special case of a vector field, one has to choose exactly one tensor of type $\binom{p}{q}$ residing at each point of a manifold M. Once again, we restrict to fields which vary smoothly from point to point. In order to formulate this succinctly, an algebraic perspective is useful. In particular, one should realize what kind of *maps* tensor fields actually are.

An individual tensor of type $\binom{p}{q}$ in $P \in M$ takes as its arguments vectors and covectors in P, and the result is a number which depends linearly on each of the arguments. At the level of fields, this happens in each point $P \in M$. It is convenient to regard it as if we inserted vector and covector *fields* as arguments of a tensor *field*, obtaining a number at each point, i.e. a *function*. Since at each point linearity *over* $\mathbb{R}$ is required, one has to demand linearity *over* $\mathcal{F}(M)$ for fields. Let us clarify this subtle point in more detail. Consider a covector field α. At each point P we have α_P, and the value V_P of a vector field V is inserted in it as an argument. In this way we obtain a function

$$\langle \alpha, V \rangle \in \mathcal{F}(M) \qquad \langle \alpha, V \rangle(P) := \langle \alpha_P, V_P \rangle \in \mathbb{R}$$

Since α_P is a covector, for any $\lambda \in \mathbb{R}$ it holds that

$$\langle \alpha_P, V_P + \lambda W_P \rangle = \langle \alpha_P, V_P \rangle + \lambda \langle \alpha_P, W_P \rangle$$

At a different point $Q \neq P$ we have

$$\langle \alpha_Q, V_Q + \lambda W_Q \rangle = \langle \alpha_Q, V_Q \rangle + \lambda \langle \alpha_Q, W_Q \rangle$$

Both results should be valid, however, for *arbitrary* λ, so that λ present in the formula corresponding to the point P may be completely *different* from λ in the formula corresponding to the point Q – a "constant" λ may depend on a point, and therefore for any *function*

$f \in \mathcal{F}(M)$ we must have

$$\langle \alpha, V + fW \rangle = \langle \alpha, V \rangle + f \langle \alpha, W \rangle$$

This is said to be the $\mathcal{F}(M)$-*linearity* of the map α, which should be contrasted with the weaker requirement of $\mathbb{R}$-*linearity*. At the same time, we see the important fact that the property of being $\mathcal{F}(M)$-linear ultimately springs from the *pointwise* character of the construction (the expression $\langle \alpha, V \rangle$ is in fact $\langle \alpha_P, V_P \rangle$ performed in each point P). The $\mathcal{F}(M)$-linearity means that the arguments (vector fields in the case of a covector field) constitute a module over the algebra $\mathcal{F}(M)$ and the map

$$\alpha : \mathcal{T}_0^1(M) \to \mathcal{F}(M)$$

is linear *in the sense of modules*.

In terms of these maps the smoothness of a covector field is easily stated: α is said to be *smooth* (of class C^∞) if the function $\langle \alpha, V \rangle$ is smooth for any smooth vector field V. Smooth covector fields on M will be denoted by $\mathcal{T}_1^0(M)$.

$\boxed{2.5.1}$ Given $\alpha, \beta \in \mathcal{T}_1^0(M)$, $f \in \mathcal{F}(M)$, check that also $\alpha + f\beta \in \mathcal{T}_1^0(M)$, if the linear combination is defined as

$$\langle \alpha + f\beta, V \rangle := \langle \alpha, V \rangle + f \langle \beta, V \rangle.$$

□

• This means that not only vector fields, but also covector fields constitute an $\mathcal{F}(M)$-module. Now, it is clear from this perspective that a tensor field of type $\binom{p}{q}$ may be regarded *as a map*

$$t : \underbrace{\mathcal{T}_0^1(M) \times \cdots \times \mathcal{T}_0^1(M)}_{q} \times \underbrace{\mathcal{T}_1^0(M) \times \cdots \times \mathcal{T}_1^0(M)}_{p} \to \mathcal{F}(M)$$

which is $\mathcal{F}(M)$-linear in each argument. If the resulting function happens to be smooth for arbitrary smooth arguments, the field t is said to be smooth. Smooth tensor fields of type $\binom{p}{q}$ on M will be denoted by $\mathcal{T}_q^p(M)$, the case of $\mathcal{T}_0^0(M)$ being identified with $\mathcal{F}(M)$. (This makes the notation $\mathcal{T}_0^1(M)$ comprehensible for vector fields, too.)

$\boxed{2.5.2}$ Check that each $\mathcal{T}_q^p(M)$ is naturally endowed with the structure of an $\mathcal{F}(M)$-module. □

• If we make a comparison between tensors in L and tensor fields on M, we can say that virtually everything goes the same way, if we substitute $T_q^p(L)$ by $\mathcal{T}_q^p(M)$, linear spaces by $\mathcal{F}(M)$-modules and $\mathbb{R}$-linearity by $\mathcal{F}(M)$-linearity.

In particular, let us look more closely at the properties of tensor algebra. This concept may be readily transferred to a manifold, after performing the substitutions mentioned above: one takes the direct sum of all *modules* $\mathcal{T}_q^p(M)$

$$\mathcal{T}(M) := \bigoplus_{p,q=0}^{\infty} \mathcal{T}_q^p(M)$$

(it is an $\mathcal{F}(M)$-module, too) and defines there a pointwise product $\otimes$, just like in Section 2.4. This algebra, the *algebra of tensor fields on* M, is ∞-dimensional (which looks much the same as for $T(L)$), but here already each homogeneous part $\mathcal{T}^p_q(M)$ is ∞-dimensional (over $\mathbb{R}$; the most salient difference occurs for the lowest degree $\binom{0}{0}$: $\mathbb{R} \leftrightarrow \mathcal{F}(M)$). On higher degrees, the situation is repeated in the form we met already in Section 2.2: although the spaces $\mathcal{T}^p_q(\mathcal{O})$ are ∞-dimensional even on "sufficiently small" domains $\mathcal{O} \subset M$ (e.g. in coordinate patches $\mathcal{O} \leftrightarrow x^i$), when regarded as linear spaces, they are finitely generated, when regarded as modules. And what do the basis tensor fields actually look like, with respect to which decomposition is to be performed?

We have seen in Section 2.4 that the most natural basis in L^*, with respect to a given basis e_a in L, is the dual basis e^a. At the same time, for vector fields we know a *coordinate* basis ∂_i. What does a basis for covector fields look like which is dual (in each point) to this particular basis?

2.5.3 Let $f \in \mathcal{F}(M)$, and let x^i be local coordinates in $\mathcal{O} \subset M$. Check that

(i) by the prescription

$$\langle df, V\rangle := Vf$$

a *covector field* df on M is defined. This field is called the *gradient* of the function f

(ii) gradients of coordinates (= functions!) $dx^i \in \mathcal{T}^0_1(\mathcal{O})$ constitute a basis for covector fields on $\mathcal{O}$, i.e. any $\alpha \in \mathcal{T}^0_1(\mathcal{O})$ may be decomposed in the form

$$\alpha = \alpha_i(x)\,dx^i \qquad \alpha_i(x) := \langle \alpha, \partial_i\rangle \ (\equiv \textit{components} \text{ with respect to the basis } dx^i)$$

and, in particular, for a gradient we have

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

(iii) covectors $dx^i|_P$ constitute a basis for covectors in P, which is dual to the coordinate basis $\partial_i|_P$ for vectors in P (the basis dx^i is said to be a *coordinate basis*, too)

(iv)

$$\langle \alpha, V\rangle = \alpha_i(x)V^i(x)$$

(v) under the change of coordinates one has (J being, as usual, the Jacobian matrix)

$$x^i \mapsto x'^i(x) \ \Rightarrow \ dx^i \mapsto dx'^i = J^i_j(x)\,dx^j \qquad \text{and} \qquad \alpha_i(x) \mapsto \alpha'_i(x') = (J^{-1})^j_i(x)\alpha_j(x)$$

Hint: (i) see (2.2.12); (v) set $f = x'^i$ in (ii). □

• Since we already have the dual basis dx^i to ∂_i, we may write down component decompositions of arbitrary tensor fields.

2.5.4 Check that if $t \in \mathcal{T}^p_q(M)$, then

(i) locally (in $\mathcal{O} \leftrightarrow x^i$) it holds that

$$t = t^{i\dots j}_{k\dots l}(x)\,dx^k \otimes \cdots \otimes dx^l \otimes \partial_i \otimes \cdots \partial_j$$

(ii) under the change of coordinates $x \mapsto x'$ components transform according to the formula

$$x^i \mapsto x'^i(x) \quad \Rightarrow \quad t^{i\dots j}_{k\dots l}(x) \mapsto t'^{i\dots j}_{k\dots l}(x') \\ \equiv J^i_r(x)\dots J^j_s(x)(J^{-1})^u_k(x)\dots(J^{-1})^v_l(x)t^{r\dots s}_{u\dots v}(x)$$

□

2.5.5 Prove that the module $\mathcal{T}^p_q(\mathcal{O})$ has n^{p+q} generators.

Hint: see (2.5.4) and (2.4.6); $t^{i\dots j}_{k\dots l} \in \mathcal{F}(\mathcal{O})$. □

• The result given in (2.5.4) might serve as a basis for an independent *definition* of a tensor field on M (definition of classical differential geometry; refer to definition no. 4 of a vector in Section 2.2): the tensor field of type $\binom{p}{q}$ on M is a collection of *functions* $t^{i\dots j}_{k\dots l}(x)$ associated with coordinates x^i defined in patches $\mathcal{O} \leftrightarrow x^i$, transforming under the changes of coordinates according to the rule given in (2.5.4). Note that a *global* object on M is defined here in terms of its pieces (components $t^{i\dots j}_{k\dots l}(x)$ on $\mathcal{O} \subset M$) as well as a rule of how to *globalize* them, i.e. how to glue these pieces together consistently so as to obtain a desired whole. In order to make this method work, one has to ensure that the rule for transition from one piece to another satisfies a consistency condition on *triple overlap* of charts (see (2.2.6)): two steps $x \mapsto x' \mapsto x''$ are to lead to the same result as a single one $x \mapsto x''$. This may be regarded actually as a requirement, namely that the rule should have particular *group properties* – coordinate changes on triple overlaps are naturally endowed with the structure of a group (multiplication being realized as a composition of the two transformations involved) and the transformation rules are to have the properties of "action" of the group (in particular, its *representation* in linear spaces, as is the case here; see Section 12.1). Some of these rules may be fairly complicated (e.g. the rule for Christoffel symbols of a linear connection, see (15.2.3)), but the property of group action is necessary for a globally defined object (and sufficient as well).

2.5.6 Check that the rule given in (2.5.4) for transformation of components of a tensor field *meets* the requirement of consistency on triple overlaps of charts.

Hint: consider the behavior of Jacobian matrices for the transitions $x \mapsto x' \mapsto x''$. □

2.5.7 Prove that a tensor field is smooth if and only if its components happen to be smooth (and this does not depend on the choice of coordinates). □

2.6 Metric tensor on a manifold

• On a manifold M, tensor fields of arbitrary type $\binom{p}{q}$ may be introduced. The only *canonical* (existing automatically) tensor field on a general manifold is the *unit* tensor field $\hat{1}$ of type $\binom{1}{1}$ (its other names being the *contraction tensor* or *canonical pairing*; note that the tensor product of several copies of this tensor as well as all possible symmetrizations and antisymmetrizations of such products are canonical, too)

$$\hat{1}(V,\alpha) := \langle \alpha, V\rangle \qquad \text{i.e.} \quad \hat{1}(V,\,\cdot\,) = V, \quad \hat{1}(\,\cdot\,,\alpha) = \alpha$$

2.6.1 Check that

(i) in coordinates

$$\hat{1} = dx^i \otimes \partial_i \qquad \text{i.e.} \quad \hat{1}^i_j = \delta^i_j$$

(ii) the expression in (i) does not depend on the choice of coordinates (see (2.4.9)). □

• All other tensor fields on a manifold have to be specially defined and they provide *additional* structure on M. What particular manifold we choose and what tensor fields it is endowed with depend ultimately on the physical context in which the tools of differential geometry are intended to be used (they represent input data, which characterize the problem in geometric language). In the majority of physically interesting applications of geometry (although not in all of them) a metric tensor on a manifold enters the scene, i.e. a field $g \in \mathcal{T}^0_2(M)$ such that for each point P it is a metric tensor in $T_P M$ in the sense of (2.4.11). It is a fairly "strong" structure, indeed, which enables one to perform various operations directly (such as lowering and raising of indices, association of lengths and angles with vectors, etc.), but it also *induces* various additional structures (linear connection, volume form, etc.) as well. A manifold endowed with a metric tensor, i.e. a pair (M, g), is said to be the *Riemannian manifold* and the branch of geometry which treats such manifolds is *Riemannian geometry*. If g is not positive definite (see the text just after (2.4.11)), one sometimes speaks about the *pseudo-Riemannian* manifold and geometry and, in particular, about the *Lorentzian* manifold and geometry for signature $(+, -, \cdots -)$ or $(-, +, \cdots +)$.

2.6.2 Check that in the coordinate basis it holds that

$$\flat_g(V^i \partial_i) = V_i\, dx^i \qquad \sharp_g(\alpha_i\, dx^i) = \alpha^i \partial_i$$

where

$$V_i := g_{ij} V^j \qquad \alpha^i := g^{ij}\alpha_j$$

Hint: see (2.4.13). □

The simplest n-dimensional manifold is given by Cartesian space $\mathbb{R}^n$. Here the *standard (flat) metric tensor* of signature (r, s) $(r + s = n)$ is introduced; by definition, in *Cartesian* coordinates we put

$$g_{ij} = \eta_{ij} \equiv \text{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s})$$

i.e.

$$g = \eta_{ij}\, dx^i \otimes dx^j = dx^1 \otimes dx^1 + \cdots + dx^r \otimes dx^r - dx^{r+1} \otimes dx^{r+1} - \cdots - dx^n \otimes dx^n$$

This manifold will be denoted by $(\mathbb{R}^n, \eta_{ij}) \equiv E^{r,s}$ from now on (and called the *pseudo-Euclidean space*), and, in particular, in the positive definite case $(\mathbb{R}^n, \delta_{ij}) \equiv E^n$ (the *Euclidean space*).

Let us have a closer look at the motivation for this definition in the most mundane spaces E^2 and E^3. In a common plane E^2 it says, for example, that the length of the two vectors ∂_x and ∂_y is (at each point) 1 and that these vectors are orthogonal to each other. For $|\partial_x|^2 = g(\partial_x, \partial_x) = (dx \otimes dx + dy \otimes dy)(\partial_x, \partial_x) = 1$, the rest similarly. This shows that the definition nicely matches our intuitive conception of metric conditions in the usual plane.

2.6.3 Write down the metric tensor in the common plane E^2 in *polar* coordinates. ($g = \underbrace{dx \otimes dx + dy \otimes dy}_{\text{Cartesian}} = \underbrace{dr \otimes dr + r^2 d\varphi \otimes d\varphi}_{\text{polar}}$.) □

2.6.4 Write down the metric tensor g in the common three-dimensional space E^3 in Cartesian, cylindrical and spherical polar coordinates.
Result:

$$\begin{aligned} g &= dx \otimes dx + dy \otimes dy + dz \otimes dz && \text{Cartesian coordinates} \\ &= dr \otimes dr + r^2 d\varphi \otimes d\varphi + dz \otimes dz && \text{cylindrical coordinates} \\ &= dr \otimes dr + r^2\, d\vartheta \otimes d\vartheta + r^2 \sin^2 \vartheta\, d\varphi \otimes d\varphi && \text{spherical polar coordinates} \end{aligned}$$

□

• This kind of computation can be done either making use of transformational properties of tensor components (i.e. reading components from its expression in Cartesian coordinates, using (2.5.4) or (2.4.18) and "gluing together" a new coordinate basis with new components), or computing new "differentials" (= gradients of coordinates), first, according to (2.5.3), e.g. in (2.6.3) $dx = x_{,r}\, dr + x_{,\varphi}\, d\varphi = \cos\varphi\, dr - r \sin\varphi\, d\varphi$, and then exploiting bilinearity of the tensor product. As a rule, this alternative method is quicker for simple metric tensors. In elementary situations (like that mentioned above) one can see, after a bit of practice, the result directly from the visual conception of what the geometry is about on a particular manifold, see (3.2.11) and (3.2.12).

2.6.5 Check that the non-Cartesian coordinate bases in (2.6.3) and (2.6.4) are ortho*gonal*, but they are not ortho*normal*.

Hint: see the text prior to (2.4.12). □

• If some local coordinates on (M, g) induce at each point the orthogonal coordinate basis of the tangent space, they are said to be *orthogonal coordinates*. We have learned above that, besides Cartesian coordinates, also polar coordinates in E^2 and spherical polar as well as cylindrical coordinates in E^3 (and various others, too; e.g. see (3.2.2)–(3.2.7)) deserve to be titled by this prestigious nomenclature.

A manifold $(\mathbb{R}^4, \eta_{ij}) \equiv E^{1,3}$ with signature (1, 3) is called *Minkowski space* and it plays a featured role in the special theory of relativity (being the *space-time* there; see more in

Chapter 16). Cartesian coordinates are usually labelled in this particular case as (x^0, x^i), $i = 1, 2, 3$, $x^0 = t$ being time and x^i corresponding to Cartesian coordinates in our good old $\mathbb{R}^3$ (the choice of units with $c = 1$ is adopted).

2.6.6 Write down the Minkowski metric η in spherical polar and cylindrical coordinates (i.e. $(t, r, \vartheta, \varphi)$ and (t, r, φ, z) respectively instead of (t, x, y, z)). ($\eta = dt \otimes dt - h$, h from (2.6.4).) □

• An important metric tensor is unobtrusively hidden in the expression for the kinetic energy of a system of particles.

2.6.7 Given $(\mathbf{r}_1(t), \dots, \mathbf{r}_N(t))$ a trajectory of a system of N point masses in mechanics, we may regard it as a *curve* $\Gamma(t)$ on a manifold $M \equiv \mathbb{R}^3 \times \dots \times \mathbb{R}^3 = \mathbb{R}^{3N}$. Check that the *kinetic energy* of this system induces the particular metric tensor $h \in \mathcal{T}^0_2(M)$ on $\mathbb{R}^{3N}$ (being different from the standard one, in general) by

$$\text{kinetic energy } \equiv T = \frac{1}{2} h(\dot\Gamma, \dot\Gamma)$$

Hint: if (x_k, y_k, z_k) are Cartesian coordinates of the kth point, then $h = m_1 h_1 + \dots + m_N h_N$, where $h_k := dx_k \otimes dx_k + dy_k \otimes dy_k + dz_k \otimes dz_k$. □

2.6.8 Write down the kinetic energy of *a single* point mass in Cartesian, cylindrical and spherical polar coordinates.

Hint: see (2.6.7) and (2.6.4); for a single point mass, h is *only a multiple* of the standard metric tensor; one obtains

$$\begin{aligned} T &= \frac{1}{2} m(\dot x^2 + \dot y^2 + \dot z^2) && \text{Cartesian coordinates} \\ &= \frac{1}{2} m(\dot r^2 + r^2 \dot\varphi^2 + \dot z^2) && \text{cylindrical coordinates} \\ &= \frac{1}{2} m(\dot r^2 + r^2 \dot\vartheta^2 + r^2 \sin^2\vartheta\, \dot\varphi^2) && \text{spherical polar coordinates} \end{aligned}$$

□

• The metric tensor turns out to be the essential element for introducing the concept of the *length of a curve* on (M, g), too. Let us begin in E^3. If a point moves along a trajectory $\mathbf{r}(t)$ in our usual space E^3, it traverses (to first order in ϵ) the distance $ds = |\mathbf{v}|\epsilon = \epsilon\sqrt{\dot x^2 + \dot y^2 + \dot z^2}$ within the time interval between t and $t + \epsilon$ (according to the theorem of Pythagoras; this is the place, of course, where the metric tensor in E^3 is hidden). Note, however, that one can write this as $\epsilon\sqrt{g(\dot\gamma, \dot\gamma)}$ for $\gamma \leftrightarrow (x(t), y(t), z(t))$. The length of a finite segment between $P = \gamma(t_1)$ and $Q = \gamma(t_2)$ is given by $\int dt\, \sqrt{g(\dot\gamma, \dot\gamma)}$. The most interesting feature of this expression consists in the fact that one cannot see from it that $(M, g) = E^3$ and Cartesian coordinates are used. It is then natural to use this very expression for the definition of the length of a curve *in general*. One should understand that even in this general case its meaning remains just the same – *for small pieces*, the relation "distance = speed × time interval" is used, and the result is summed over all small pieces (i.e. integrated).

It is a suitable time now to contemplate the visual meaning of the concept of the length of a vector V itself. The following is meant by this notion: if we proceed a parametric distance ϵ along the vector V, we travel (in the positive definite case) a distance (in the sense of the length of the curve)[27] $\epsilon|V| \equiv \epsilon\sqrt{g(V,V)}$. Keeping this in mind one is often able to derive explicit forms of metric tensors on two-dimensional surfaces in E^3 simply by a "rule of thumb" (see (3.2.11); the same is true for curves in E^3 as well, being fairly useful, for example, in computing line integrals of the first kind, see (7.7.4)).

There is an alternative way of displaying the metric tensor, which is frequently used in general relativity, and may be ultimately traced back to the connection between the length of a curve and a metric tensor. In this convention one writes directly the "square of the distance" dl^2 between two points which are infinitesimally close to one another (i.e. points with values of coordinates being x^i and $x^i + dx^i$ respectively), where dx^i denote infinitesimal *increments* of the values of coordinates (so that *they are not* our base covector fields (!)). For metric tensors from exercise (2.6.4), as an example, we have

$$\begin{aligned} dl^2 &= dx^2 + dy^2 + dz^2 \\ &= dr^2 + r^2\, d\varphi^2 + dz^2 \\ &= dr^2 + r^2\, d\vartheta^2 + r^2 \sin^2\vartheta\, d\varphi^2 \end{aligned}$$

Although we will not, as a rule, use this convention in the course of the book, it is fairly common in texts on relativity and one should understand clearly its precise meaning.

2.6.9 Let $t \mapsto t(\sigma)$ be a *reparametrization* of a curve γ, i.e. $\hat\gamma(\sigma) := \gamma(t(\sigma))$. Check that the functional of the *length of a curve* (refer to (4.6.1), (7.7.5) and (15.4.8))

$$\text{length of a curve } \gamma \ \equiv l[\gamma] := \int_{t_1}^{t_2} dt\, \sqrt{g(\dot\gamma, \dot\gamma)}$$

is *reparametrization invariant*, $l[\gamma] = l[\hat\gamma]$, i.e. this expression depends on the image set of a curve (i.e. on the *path*; recall that the curve is *a map*) rather than on a particular parametrization of this set (on a *curve*).

Hint: according to (2.3.5) $\hat\gamma' = (dt/d\sigma)\dot\gamma$, therefore $d\sigma\sqrt{g(\hat\gamma', \hat\gamma')} = dt\sqrt{g(\dot\gamma, \dot\gamma)}$. □

• Finally, we mention the possibility of introducing the *gradient* as a *vector* field. The gradient df as a *covector* field has been defined in (2.5.3). If a metric tensor is available, we can find a *vector* field, simply by raising the index on the covector df. The resulting vector field is called the gradient (of a function f), too, and will be denoted by grad f or ∇f

$$\text{grad} f \equiv \nabla f := \sharp_g df \equiv g^{-1}(df, \cdot\,) \qquad \text{i.e.} \quad (\nabla f)^i := g^{ij}(df)_j \equiv g^{ij} f_{,j}$$

A well-known example is provided by the *potential force field* in mechanics. It is the gradient of the (by definition negative) *potential energy* of a system. Here, indices are raised by means

[27] Remember that the vector V officially resides as a whole at a single point x and its length is $g_x(V,V)$. This length (in the sense of a scalar product in $T_x M$) now becomes related with a formally *different* length, namely the length of a small piece of a *curve* $\gamma(t)$ defined by the vector, the representative of a class specified by the vector V. Both computations need g and the definitions are intentionally designed *so as to* make the results coincide.

of the *standard* metric tensor on $M \equiv \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 = \mathbb{R}^{3N}$, that is by $h_1 + \cdots + h_N$ (as opposed to (2.6.7), where masses are present, too).

2.6.10 Find the lines of electric field of a point charge and of an elementary dipole.

Hint: first, write down equations for *integral curves* of the electric field $\mathbf{E} = -\nabla\Phi$, i.e.

$$\dot{x}^i = -g^{ij}\Phi_{,j}$$

for

$$\Phi(r, \vartheta, \varphi) = \frac{\alpha}{r} \qquad \text{resp.} \quad \Phi(r, \vartheta, \varphi) = \alpha\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} \equiv (\alpha p)\frac{\cos\vartheta}{r^2}$$

($\alpha \in \mathbb{R}$) and then disregard parametrization (eliminate dt in the separation of variables procedure; see also (8.5.13)). □

Summary of Chapter 2

For each point x of an n-dimensional manifold M there is the canonically defined n-dimensional linear space $T_x M$, the tangent space at the point x. Its elements are called vectors at x. There are several mutually equivalent definitions of this concept, useful in different contexts. A vector *field* on a manifold M is a smooth assignment of a vector to each point $x \in M$. The integral curve of a vector field is the curve whose motion at each point is just that dictated by the vector of the field at this point. Standard constructions of multilinear algebra (construction of tensors of type $\binom{p}{q}$ for a given vector space L) lead to the notion of a *tensor field* of type $\binom{p}{q}$ on a manifold. In particular, one has functions (type $\binom{0}{0}$), vector and covector fields (type $\binom{1}{0}$ and $\binom{0}{1}$), fields of bilinear form (type $\binom{0}{2}$, in the symmetric non-degenerate case the metric tensor) and linear operators (type $\binom{1}{1}$).

$\gamma : \mathbb{R} \to M$	A curve γ on a manifold M	Sec. 2.1
$f : M \to \mathbb{R}$	A function f on a manifold M	Sec. 2.1
$e_i := \partial_i\|_P$	Coordinate basis of $T_P M$	(2.2.6)
$a^i \mapsto a'^i = J^i_j(P)a^j$	Transformation of components of a vector in P	(2.2.6)
$V(fg) = (Vf)g + f(Vg)$	Leibniz rule for action of vector fields	(2.2.8)
$\dot{x}^i = V^i(x) \quad (\dot{\gamma} = V)$	Equations for finding integral curves of V	(2.3.1)
$v = \sum_{b=1}^{n} v^b e_b \equiv v^b e_b$	Summation convention	(2.4.2)
$\langle e^a, e_b\rangle = \delta^a_b$	The base e^a is dual with respect to e_a	(2.4.2)
$t^{c\dots d}_{a\dots b} := t(e_a, \dots, e_b; e^c, \dots, e^d)$	Components of tensor $t \in T^p_q(L)$	(2.4.6)
$v_a := g_{ab}v^b\ ,\ \alpha^a := g^{ab}\alpha_b$	Lowering and raising of indices by means of g	(2.4.13)
$\langle df, V\rangle := Vf$	Gradient of a function f as a covector field	(2.5.3)
$T = \frac{1}{2}h(\dot{\Gamma}, \dot{\Gamma})$	Kinetic energy of a system of N point masses	(2.6.7)
$l[\gamma] := \int_{t_1}^{t_2} dt\sqrt{g(\dot{\gamma}, \dot{\gamma})}$	Functional of the length of a curve γ	(2.6.9)
$(\nabla f)^i := g^{ij}f_{,j} \quad (\nabla f := \sharp_g\, df)$	Gradient of a function f as a vector field	Sec. 2.6