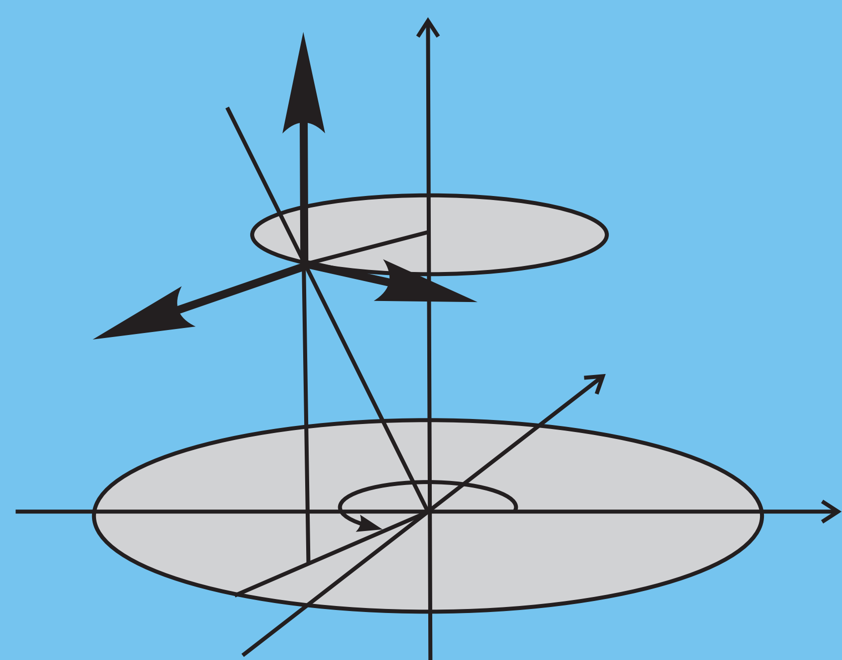
R. A. Sharipov

Quick Introduction to Tensor Analysis



MSC 97U20
PACS 01.30.Pp

R. A. Sharipov. **Quick Introduction to Tensor Analysis**: lecture notes. Freely distributed on-line. Is free for individual use and educational purposes. Any commercial use without written consent from the author is prohibited.

This book was written as lecture notes for classes that I taught to undergraduate students majoring in physics in February 2004 during my time as a guest instructor at The University of Akron, which was supported by Dr. Sergei F. Lyuksyutov's grant from the National Research Council under the COBASE program. These 4 classes have been taught in the frame of a regular Electromagnetism course as an introduction to tensorial methods.

I wrote this book in a "do-it-yourself" style so that I give only a draft of tensor theory, which includes formulating definitions and theorems and giving basic ideas and formulas. All other work such as proving consistence of definitions, deriving formulas, proving theorems or completing details to proofs is left to the reader in the form of numerous exercises. I hope that this style makes learning the subject really quick and more effective for understanding and memorizing.

I am grateful to Department Chair Prof. Robert R. Mallik for the opportunity to teach classes and thus to be involved fully in the atmosphere of an American university. I am also grateful to

Mr. M. Boiwka (mboiwka@hotmail.com)
Mr. A. Calabrese (ajc10@uakron.edu)
Mr. J. Comer (funnybef@lycos.com)
Mr. A. Mozinski (arm5@uakron.edu)
Mr. M. J. Shepard (sheppp2000@yahoo.com)

for attending my classes and reading the manuscript of this book. I would like to especially acknowledge and thank Mr. Jeff Comer for correcting the grammar and wording in it.

Contacts to author.

Office: Mathematics Department, Bashkir State University,
32 Frunze street, 450074 Ufa, Russia
Phone: 7-(3472)-23-67-18
Fax: 7-(3472)-23-67-74

Home: 5 Rabochaya street, 450003 Ufa, Russia
Phone: 7-(917)-75-55-786
E-mails: R_Sharipov@ic.bashedu.ru,
r-sharipov@mail.ru,
ra_sharipov@hotmail.com,
URL: http://www.geocities.com/r-sharipov

CONTENTS.

CHAPTER I

PRELIMINARY INFORMATION.

§ 1. Geometrical and physical vectors.

Vector is usually understood as a segment of straight line equipped with an arrow. Simplest example is displacement vector $\mathbf{a}$. Say its length is 4 cm, i. e.

$$|\mathbf{a}| = 4\,\text{cm}.$$

You can draw it on the paper as shown on Fig. 1a. Then it means that point B is 4 cm apart from the point A in the direction pointed to by vector $\mathbf{a}$. However, if

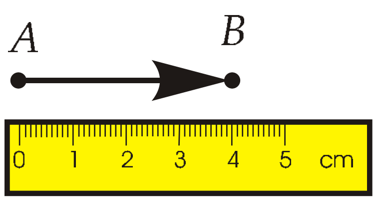


Fig. 1a.

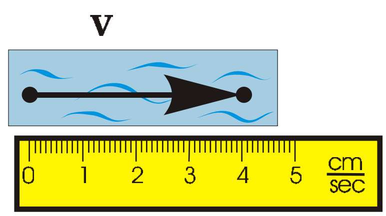


Fig. 1b.

you take velocity vector $\mathbf{v}$ for a stream in a brook, you cannot draw it on the paper immediately. You should first adopt a scaling convention, for example, saying that 1 cm on paper represents 1 cm/sec (see Fig. 1b).

Conclusion 1.1. *Vectors with physical meaning other than displacement vectors have no unconditional geometric presentation. Their geometric presentation is conventional; it depends on the scaling convention we choose.*

Conclusion 1.2. *There are plenty of physical vectors, which are not geometrically visible, but can be measured and then drawn as geometric vectors.*

One can consider unit vector $\mathbf{m}$. Its length is equal to unity not 1 cm, not 1 km, not 1 inch, and not 1 mile, but simply number 1:

$$|\mathbf{m}| = 1.$$

Like physical vectors, unit vector $\mathbf{m}$ cannot be drawn without adopting some scaling convention. The concept of a unit vector is a very convenient one. By multiplying $\mathbf{m}$ to various scalar quantities, we can produce vectorial quantities of various physical nature: velocity, acceleration, force, torque, etc.

CONCLUSION 1.3. *Along with geometrical and physical vectors one can imagine vectors whose length is a number with no unit of measure.*

§ 2. Bound vectors and free vectors.

All displacement vectors are bound ones. They are bound to those points whose displacement they represent. Free vectors are usually those representing global physical parameters, e. g. vector of angular velocity $\boldsymbol{\omega}$ for Earth rotation about its axis. This vector produces the Coriolis force affecting water streams in small

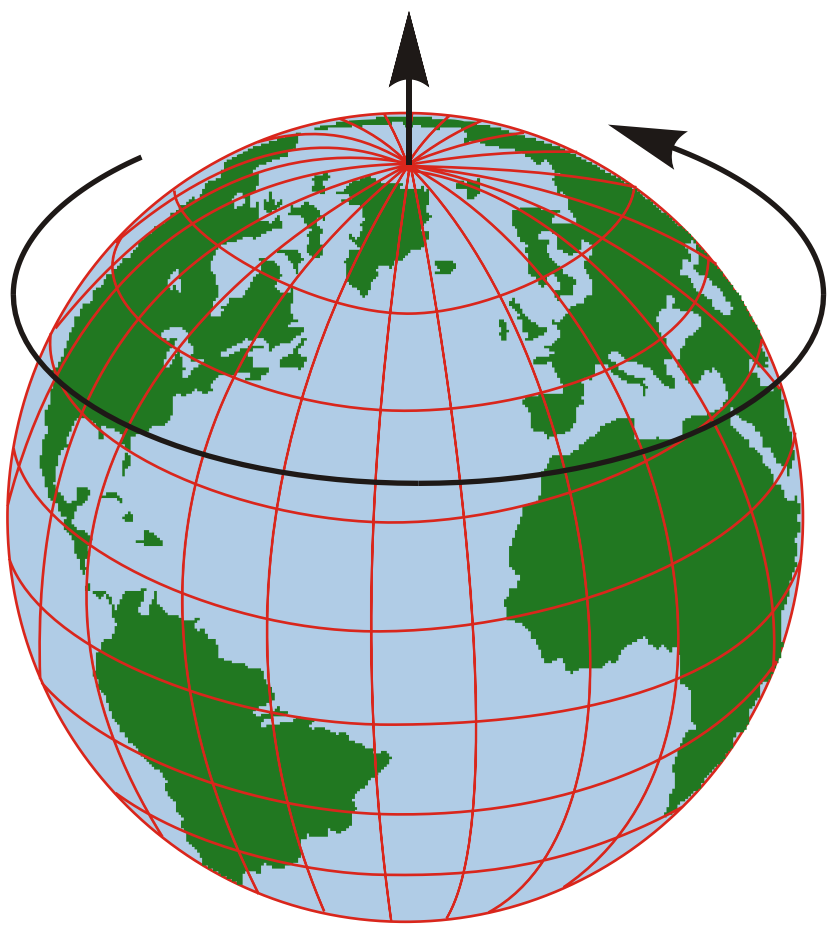


Fig 2.

rivers and in oceans around the world. Though it is usually drawn attached to the North pole, we can translate this vector to any point along any path provided we keep its length and direction unchanged.

The next example illustrates the concept of a *vector field*. Consider the water flow in a river at some fixed instant of time t. For each point P in the water the velocity of the water jet passing through this point is defined. Thus we have a function

$$\mathbf{v} = \mathbf{v}(t, P). \tag{2.1}$$

Its first argument is time variable t. The second argument of function (2.1) is not numeric. It is geometric object — a point. Values of a function (2.1) are also not numeric: they are vectors.

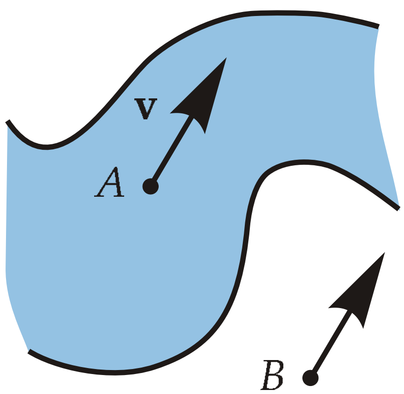


Fig. 3.

Definition 2.1. A vector-valued function with point argument is called vector field. If it has an additional argument t, it is called a time-dependent vector field.

Let $\mathbf{v}$ be the value of function (2.1) at the point A in a river. Then vector $\mathbf{v}$ is a bound vector. It represents the velocity of the water jet at the point A. Hence, it is bound to point A. Certainly, one can translate it to the point B on the bank of the river (see Fig. 3). But there it loses its original purpose, which is to mark the water velocity at the point A.

Conclusion 2.1. *There exist functions with non-numeric arguments and non-numeric values.*

Exercise 2.1. *What is a scalar field? Suggest an appropriate definition by analogy with definition 2.1.*

Exercise 2.2 (for deep thinking). *Let $y = f(x)$ be a function with a non-numeric argument. Can it be continuous ? Can it be differentiable ? In general, answer is negative. However, in some cases one can extend the definition of continuity and the definition of derivatives in a way applicable to some functions with non-numeric arguments. Suggest your version of such a generalization. If no versions, remember this problem and return to it later when you gain more experience.*

Let A be some fixed point (on the ground, under the ground, in the sky, or in outer space, wherever). Consider all vectors of some physical nature bound to this point (say all force vectors). They constitute an infinite set. Let's denote it V_A. We can perform certain algebraic operations over the vectors from V_A:

(1) we can add any two of them;
(2) we can multiply any one of them by any real number $\alpha \in \mathbb{R}$;

These operations are called linear operations and V_A is called a linear vector space.

Exercise 2.3. *Remember the parallelogram method for adding two vectors (draw picture). Remember how vectors are multiplied by a real number α. Consider three cases: $\alpha > 0$, $\alpha < 0$, and $\alpha = 0$. Remember what the zero vector is. How it is represented geometrically ?*

Exercise 2.4. *Do you remember the exact mathematical definition of a linear vector space? If yes, write it. If no, visit Web page of Jim Hefferon*

http://joshua.smcvt.edu/linearalgebra/

and download his book [1]. *Keep this book for further references. If you find it useful, you can acknowledge the author by sending him e-mail:* jim@joshua.smcvt.edu .

Conclusion 2.2. *Thus, each point A of our geometric space is not so simple, even if it is a point in a vacuum. It can be equipped with linear vector spaces of various natures (such as a space of force vectors in the above example). This idea, where each point of vacuum space is treated as a container for various physical fields, is popular in modern physics. Mathematically it is realized in the concept of bundles: vector bundles, tensor bundles, etc.*

Free vectors, taken as they are, do not form a linear vector space. Let's denote by V the set of all free vectors. Then V is union of vector spaces V_A associated with all points A in space:

$$V = \bigcup_{A\in E} V_A. \tag{2.2}$$

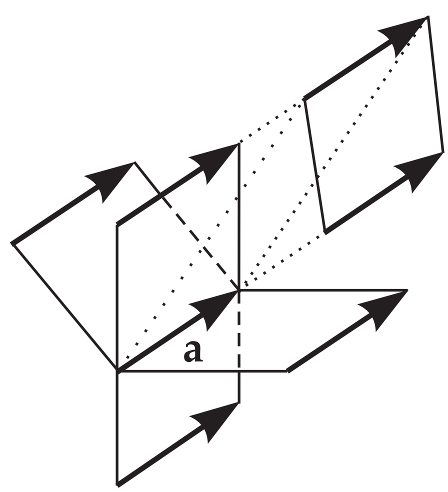


Fig. 4.

The free vectors forming this set (2.2) are too numerous: we should work to make them confine the definition of a linear vector space. Indeed, if we have a vector $\mathbf{a}$ and if it is a free vector, we can replicate it by parallel translations that produce infinitely many copies of it (see Fig. 4). All these clones of vector $\mathbf{a}$ form a class, the class of vector $\mathbf{a}$. Let's denote it as $\mathrm{Cl}(\mathbf{a})$. Vector $\mathbf{a}$ is a representative of its class. However, we can choose any other vector of this class as a representative, say it can be vector $\tilde{\mathbf{a}}$. Then we have

$$\mathrm{Cl}(\mathbf{a}) = \mathrm{Cl}(\tilde{\mathbf{a}}).$$

Let's treat $\mathrm{Cl}(\mathbf{a})$ as a whole unit, as one indivisible object. Then consider the set of all such objects. This set is called a factor-set, or quotient set. It is denoted as $V/\sim$. This quotient set $V/\sim$ satisfies the definition of linear vector space. For the sake of simplicity further we shall denote it by the same letter V as original set (2.2), from which it is produced by the operation of factorization.

Exercise 2.5. *Have you heard about binary relations, quotient sets, quotient groups, quotient rings and so on? If yes, try to remember strict mathematical definitions for them. If not, then have a look to the references* [2], [3], [4]. *Certainly, you shouldn't read all of these references, but remember that they are freely available on demand.*

3. Euclidean space.

What is our geometric space? Is it a linear vector space? By no means. It is formed by points, not by vectors. Properties of our space were first systematically described by Euclid, the Greek mathematician of antiquity. Therefore, it is called Euclidean space and denoted by E. Euclid suggested 5 axioms (5 postulates) to describe E. However, his statements were not satisfactorily strict from a modern point of view. Currently E is described by 20 axioms. In memory of Euclid they are subdivided into 5 groups:

(1) axioms of incidence;
(2) axioms of order;
(3) axioms of congruence;
(4) axioms of continuity;
(5) axiom of parallels.

20-th axiom, which is also known as 5-th postulate, is most famous.

Exercise 3.1. *Visit the following Non-Euclidean Geometry web-site and read a few words about non-Euclidean geometry and the role of Euclid's 5-th postulate in its discovery.*

Usually nobody remembers all 20 of these axioms by heart, even me, though I wrote a textbook on the foundations of Euclidean geometry in 1998. Furthermore, dealing with the Euclidean space E, we shall rely only on common sense and on our geometric intuition.

§ 4. Bases and Cartesian coordinates.

Thus, E is composed by points. Let's choose one of them, denote it by O and consider the vector space V_O composed by displacement vectors. Then each point $B \in E$ can be uniquely identified with the displacement vector $\mathbf{r}_B = \overrightarrow{OB}$. It is called the radius-vector of the point B, while O is called origin. Passing from points to their radius-vectors we identify E with the linear vector space V_O. Then, passing from the vectors to their classes, we can identify V with the space of free vectors. This identification is a convenient tool in studying E without referring to Euclidean axioms. However, we should remember that such identification is not unique: it depends on our choice of the point O for the origin.

Definition 4.1. We say that three vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ form a non-coplanar triple of vectors if they cannot be laid onto the plane by parallel translations.

These three vectors can be bound to some point O common to all of them, or they can be bound to different points in the space; it makes no difference. They also can be treated as free vectors without any definite binding point.

Definition 4.2. Any non-coplanar ordered triple of vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is called a basis in our geometric space E.

Exercise 4.1. *Formulate the definitions of bases on a plane and on a straight line by analogy with definition 4.2.*

Below we distinguish three types of bases: orthonormal basis (ONB), orthogonal basis (OB), and skew-angular basis (SAB). Orthonormal basis is formed by three mutually perpendicular unit vectors:

$$\begin{aligned}&\mathbf{e}_1 \perp \mathbf{e}_2,\\ &\mathbf{e}_2 \perp \mathbf{e}_3,\\ &\mathbf{e}_3 \perp \mathbf{e}_1,\end{aligned} \tag{4.1}$$

$$\begin{aligned}&|\mathbf{e}_1| = 1,\\ &|\mathbf{e}_2| = 1,\\ &|\mathbf{e}_3| = 1.\end{aligned} \tag{4.2}$$

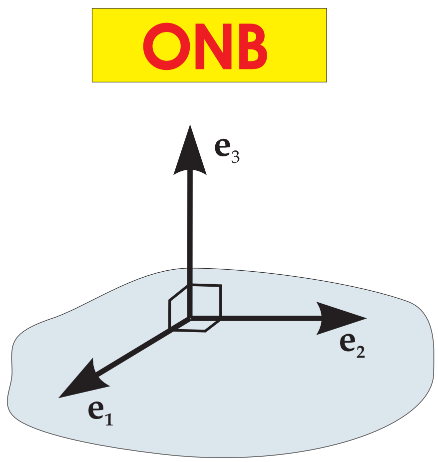


Fig. 5.

For orthogonal basis, the three conditions (4.1) are fulfilled, but lengths of basis vectors are not specified.

And skew-angular basis is the most general case. For this basis neither angles nor lengths are specified. As we shall see below, due to its asymmetry SAB can reveal a lot of features that are hidden in symmetric ONB.

Let's choose some basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in E. In the general case this is a skew-angular basis. Assume that vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are bound to a common point O as shown on Fig. 6 below. Otherwise they can be brought to this position by means of parallel translations. Let $\mathbf{a}$ be some arbitrary vector. This vector also can be translated to the point O. As a result we have four vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, and $\mathbf{a}$ beginning at the same point O. Drawing additional lines and vectors as shown on Fig. 6, we get

$$\mathbf{a} = \overrightarrow{OD} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}. \tag{4.3}$$

Then from the following obvious relationships

$$\begin{aligned}&\mathbf{e}_1 = \overrightarrow{OE_1}, && \mathbf{e}_2 = \overrightarrow{OE_2}, && \mathbf{e}_3 = \overrightarrow{OE_3},\\ &\overrightarrow{OE_1} \parallel \overrightarrow{OA}, && \overrightarrow{OE_2} \parallel \overrightarrow{OB}, && \overrightarrow{OE_3} \parallel \overrightarrow{OC}\end{aligned}$$

we derive

$$\overrightarrow{OA} = \alpha\,\mathbf{e}_1, \qquad \overrightarrow{OB} = \beta\,\mathbf{e}_2, \qquad \overrightarrow{OC} = \gamma\,\mathbf{e}_3, \tag{4.4}$$

where α, β, γ are scalars. Now from (4.3) and (4.4) we obtain

$$\mathbf{a} = \alpha\,\mathbf{e}_1 + \beta\,\mathbf{e}_2 + \gamma\,\mathbf{e}_3. \tag{4.5}$$

Exercise 4.2. *Explain how, for what reasons, and in what order additional lines on Fig. 6 are drawn.*

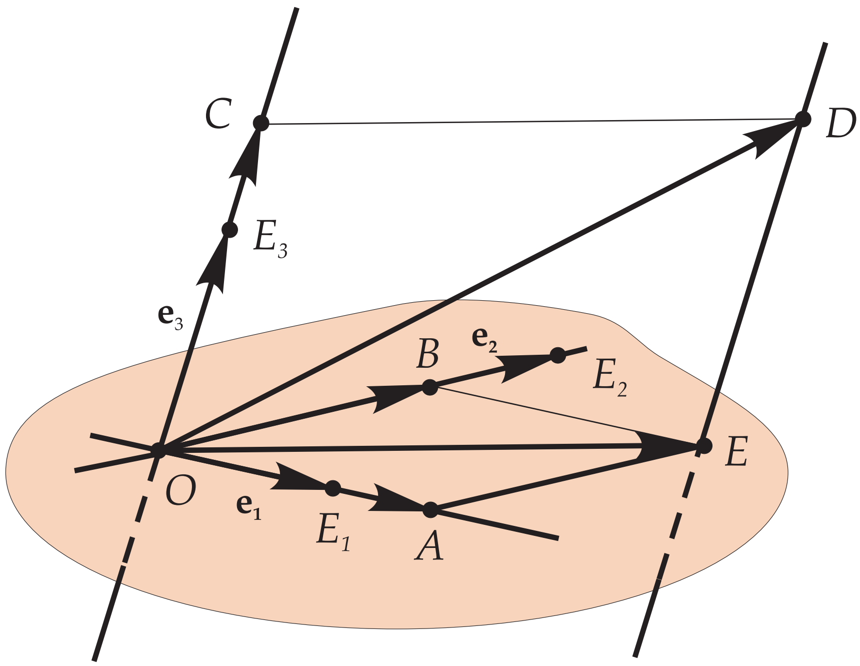


Fig. 6.

Formula (4.5) is known as the expansion of vector $\mathbf{a}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, while α, β, γ are coordinates of vector $\mathbf{a}$ in this basis.

Exercise 4.3. *Explain why α, β, and γ are uniquely determined by vector* $\mathbf{a}$.

Hint: remember what linear dependence and linear independence are. Give exact mathematical statements for these concepts. Apply them to exercise 4.3.

Further we shall write formula (4.5) as follows

$$\mathbf{a} = a^1\,\mathbf{e}_1 + a^2\,\mathbf{e}_2 + a^3\,\mathbf{e}_3 = \sum_{i=1}^{3} a^i\,\mathbf{e}_i, \tag{4.6}$$

denoting $\alpha = a^1$, $\beta = a^2$, and $\gamma = a^3$. Don't confuse upper indices in (4.6) with power exponentials, a^1 here is not a, a^2 is not a squared, and a^3 is not a cubed. Usage of upper indices and the implicit summation rule were suggested by Einstein. They are known as Einstein's tensorial notations.

Once we have chosen the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ (no matter ONB, OB, or SAB), we can associate vectors with columns of numbers:

$$\mathbf{a} \longleftrightarrow \begin{Vmatrix} a^1 \\ a^2 \\ a^3 \end{Vmatrix}, \qquad\qquad \mathbf{b} \longleftrightarrow \begin{Vmatrix} b^1 \\ b^2 \\ b^3 \end{Vmatrix}. \tag{4.7}$$

We can then produce algebraic operations with vectors, reducing them to arithmetic operations with numbers:

$$\begin{aligned}\mathbf{a}+\mathbf{b} &\longleftrightarrow \begin{Vmatrix} a^1 \\ a^2 \\ a^3 \end{Vmatrix} + \begin{Vmatrix} b^1 \\ b^2 \\ b^3 \end{Vmatrix} = \begin{Vmatrix} a^1+b^1 \\ a^2+b^2 \\ a^3+b^3 \end{Vmatrix},\\ \alpha\,\mathbf{a} &\longleftrightarrow \alpha \begin{Vmatrix} a^1 \\ a^2 \\ a^3 \end{Vmatrix} = \begin{Vmatrix} \alpha\,a^1 \\ \alpha\,a^2 \\ \alpha\,a^3 \end{Vmatrix}.\end{aligned}$$

Columns of numbers framed by matrix delimiters like those in (4.7) are called vector-columns. They form linear vector spaces.

EXERCISE 4.4. *Remember the exact mathematical definition for the real arithmetic vector space* $\mathbb{R}^n$*, where* n *is a positive integer.*

DEFINITION 4.1. The Cartesian coordinate system is a basis complemented with some fixed point that is called the origin.

Indeed, if we have an origin O, then we can associate each point A of our space with its radius-vector $\mathbf{r}_A = \overrightarrow{OA}$. Then, having expanded this vector in a basis, we get three numbers that are called the Cartesian coordinates of A. Coordinates of a point are also specified by upper indices since they are coordinates of the radius-vector for that point. However, unlike coordinates of vectors, they are usually not written in a column. The reason will be clear when we consider curvilinear coordinates. So, writing $A(a^1, a^2, a^3)$ is quite an acceptable notation for the point A with coordinates a^1, a^2, and a^3.

The idea of cpecification of geometric objects by means of coordinates was first raised by French mathematician and philosopher René Descartes (1596-1650). Cartesian coordinates are named in memory of him.

§ 5. What if we need to change a basis ?

Why could we need to change a basis? There may be various reasons: we may dislike initial basis because it is too symmetric like ONB, or too asymmetric like SAB. Maybe we are completely satisfied; but the wisdom is that looking on how something changes we can learn more about this thing than if we observe it in a static position. Suppose we have a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, let's call it **the old basis**, and suppose we want to change it to a **new** one $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$. Let's take the first vector of the new basis $\mathbf{e}_1$. Being isolated from the other two vectors $\tilde{\mathbf{e}}_2$ and $\tilde{\mathbf{e}}_3$, it is nothing, but quite an ordinary vector of space. In this capacity, vector $\tilde{\mathbf{e}}_1$ can be expanded in the old basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\tilde{\mathbf{e}}_1 = S^1\,\mathbf{e}_1 + S^2\,\mathbf{e}_2 + S^3\,\mathbf{e}_3 = \sum_{j=1}^{3} S^j\,\mathbf{e}_j. \tag{5.1}$$

Compare (5.1) and (4.6). Then we can take another vector $\tilde{\mathbf{e}}_2$ and also expand it in the old basis. But what letter should we choose for denoting the coefficients of this expansion? We can choose another letter; say the letter "R":

$$\tilde{\mathbf{e}}_2 = R^1\,\mathbf{e}_1 + R^2\,\mathbf{e}_2 + R^3\,\mathbf{e}_3 = \sum_{j=1}^{3} R^j\,\mathbf{e}_j. \tag{5.2}$$

However, this is not the best decision. Indeed, vectors $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$ differ only in number, while for their coordinates we use different letters. A better way is to add an extra index to S in (5.1). This is the lower index coinciding with the number of the vector:

$$\tilde{\mathbf{e}}_1 = S^1_1\,\mathbf{e}_1 + S^2_1\,\mathbf{e}_2 + S^3_1\,\mathbf{e}_3 = \sum_{j=1}^{3} S^j_1\,\mathbf{e}_j \tag{5.3}$$

Color is of no importance; it is used for highlighting only. Instead of (5.2), for the second vector $\mathbf{e}_2$ we write a formula similar to (5.3):

$$\tilde{\mathbf{e}}_2 = S^1_2\,\mathbf{e}_1 + S^2_2\,\mathbf{e}_2 + S^3_2\,\mathbf{e}_3 = \sum_{j=1}^{3} S^j_2\,\mathbf{e}_j. \tag{5.4}$$

And for third vector as well:

$$\tilde{\mathbf{e}}_3 = S^1_3\,\mathbf{e}_1 + S^2_3\,\mathbf{e}_2 + S^3_3\,\mathbf{e}_3 = \sum_{j=1}^{3} S^j_3\,\mathbf{e}_j. \tag{5.5}$$

When considered jointly, formulas (5.3), (5.4), and (5.5) are called **transition formulas**. We use a left curly bracket to denote their union:

$$\begin{cases} \tilde{\mathbf{e}}_1 = S^1_1\,\mathbf{e}_1 + S^2_1\,\mathbf{e}_2 + S^3_1\,\mathbf{e}_3, \\ \tilde{\mathbf{e}}_2 = S^1_2\,\mathbf{e}_1 + S^2_2\,\mathbf{e}_2 + S^3_2\,\mathbf{e}_3, \\ \tilde{\mathbf{e}}_3 = S^1_3\,\mathbf{e}_1 + S^2_3\,\mathbf{e}_2 + S^3_3\,\mathbf{e}_3. \end{cases} \tag{5.6}$$

We also can write transition formulas (5.6) in a more symbolic form

$$\tilde{\mathbf{e}}_i = \sum_{j=1}^{3} S_i^j \, \mathbf{e}_j. \tag{5.7}$$

Here index i runs over the range of integers from 1 to 3.

Look at index i in formula (5.7). It is a free index, it can freely take any numeric value from its range: 1, 2, or 3. Note that i is the lower index in both sides of formula (5.7). This is a general rule.

Rule 5.1. *In correctly written tensorial formulas free indices are written on the same level (upper or lower) in both sides of the equality. Each free index has only one entry in each side of the equality.*

Now look at index j. It is summation index. It is present only in right hand side of formula (5.7), and it has exactly two entries (apart from that $j = 1$ under the sum symbol): one in the upper level and one in the lower level. This is also general rule for tensorial formulas.

Rule 5.2. *In correctly written tensorial formulas each summation index should have exactly two entries: one upper entry and one lower entry.*

Proposing this rule 5.2, Einstein also suggested not to write the summation symbols at all. Formula (5.7) then would look like $\tilde{\mathbf{e}}_i = S_i^j \, \mathbf{e}_j$ with implicit summation with respect to the double index j. Many physicists (especially those in astrophysics) prefer writing tensorial formulas in this way. However, I don't like omitting sums. It breaks the integrity of notations in science. Newcomers from other branches of science would have difficulties in understanding formulas with implicit summation.

Exercise 5.1. *What happens if $\tilde{\mathbf{e}}_1 = \mathbf{e}_1$? What are the numeric values of coefficients S_1^1, S_1^2, and S_1^3 in formula* (5.3) *for this case?*

Returning to transition formulas (5.6) and (5.7) note that coefficients in them are parameterized by two indices running independently over the range of integer numbers from 1 to 3. In other words, they form a two-dimensional array that usually is represented as a table or as a matrix:

$$S = \begin{Vmatrix} S_1^1 & S_2^1 & S_3^1 \\ S_1^2 & S_2^2 & S_3^2 \\ S_1^3 & S_2^3 & S_3^3 \end{Vmatrix} \tag{5.8}$$

Matrix S is called **a transition matrix** or **direct transition matrix** since we use it in passing from old basis to new one. In writing such matrices like S the following rule applies.

Rule 5.3. *For any double indexed array with indices on the same level (both upper or both lower) the first index is a row number, while the second index is a column number. If indices are on different levels (one upper and one lower), then the upper index is a row number, while lower one is a column number.*

Note that according to this rule 5.3, coefficients of formula (5.3), which are written in line, constitute first column in matrix (5.8). So lines of formula (5.6) turn into columns in matrix (5.8). It would be worthwhile to remember this fact.

If we represent each vector of the new basis $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ as a column of its coordinates in the old basis just like it was done for $\mathbf{a}$ and $\mathbf{b}$ in formula (4.7) above

$$\mathbf{e}_1 \longleftrightarrow \begin{Vmatrix} S_1^1 \\ S_1^2 \\ S_1^3 \end{Vmatrix}, \qquad \mathbf{e}_2 \longleftrightarrow \begin{Vmatrix} S_2^1 \\ S_2^2 \\ S_2^3 \end{Vmatrix}, \qquad \mathbf{e}_3 \longleftrightarrow \begin{Vmatrix} S_3^1 \\ S_3^2 \\ S_3^3 \end{Vmatrix}, \tag{5.9}$$

then these columns (5.9) are exactly the first, the second, and the third columns in matrix (5.8). This is the easiest way to remember the structure of matrix S.

Exercise 5.2. *What happens if* $\tilde{\mathbf{e}}_1 = \mathbf{e}_1$, $\tilde{\mathbf{e}}_2 = \mathbf{e}_2$, *and* $\tilde{\mathbf{e}}_3 = \mathbf{e}_3$ *? Find the transition matrix for this case. Consider also the following two cases and write the transition matrices for each of them:*

(1) $\tilde{\mathbf{e}}_1 = \mathbf{e}_1$, $\tilde{\mathbf{e}}_2 = \mathbf{e}_3$, $\tilde{\mathbf{e}}_3 = \mathbf{e}_2$;
(2) $\tilde{\mathbf{e}}_1 = \mathbf{e}_3$, $\tilde{\mathbf{e}}_2 = \mathbf{e}_1$, $\tilde{\mathbf{e}}_3 = \mathbf{e}_2$.

Explain why the next case is impossible:

(3) $\tilde{\mathbf{e}}_1 = \mathbf{e}_1 - \mathbf{e}_2$, $\tilde{\mathbf{e}}_2 = \mathbf{e}_2 - \mathbf{e}_3$, $\tilde{\mathbf{e}}_3 = \mathbf{e}_3 - \mathbf{e}_1$.

Now let's swap bases. This means that we are going to consider $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ as the old basis, $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ as the new basis, and study the inverse transition. All of the above stuff applies to this situation. However, in writing the transition formulas (5.6), let's use another letter for the coefficients. By tradition here the letter "T" is used:

$$\begin{cases} \mathbf{e}_1 = T_1^1\,\tilde{\mathbf{e}}_1 + T_1^2\,\tilde{\mathbf{e}}_2 + T_1^3\,\tilde{\mathbf{e}}_3, \\ \mathbf{e}_2 = T_2^1\,\tilde{\mathbf{e}}_1 + T_2^2\,\tilde{\mathbf{e}}_2 + T_2^3\,\tilde{\mathbf{e}}_3, \\ \mathbf{e}_3 = T_3^1\,\tilde{\mathbf{e}}_1 + T_3^2\,\tilde{\mathbf{e}}_2 + T_3^3\,\tilde{\mathbf{e}}_3. \end{cases} \tag{5.10}$$

Here is the short symbolic version of transition formulas (5.10):

$$\mathbf{e}_i = \sum_{j=1}^{3} T_i^j\,\tilde{\mathbf{e}}_j. \tag{5.11}$$

Denote by T the transition matrix constructed on the base of (5.10) and (5.11). It is called **the inverse transition matrix** when compared to the direct transition matrix S:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \underset{T}{\overset{S}{\rightleftarrows}} (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3). \tag{5.12}$$

Theorem 5.1. *The inverse transition matrix* T *in* (5.12) *is the inverse matrix for the direct transition matrix* S, *i. e.* $T = S^{-1}$.

Exercise 5.3. *What is the inverse matrix? Remember the definition. How is the inverse matrix* A^{-1} *calculated if* A *is known? (Don't say that you use a computer package like Maple, MathCad, or any other; remember the algorithm for calculating* A^{-1}*).*

Exercise 5.4. *Remember what is the determinant of a matrix. How is it usually calculated? Can you calculate* $\det(A^{-1})$ *if* $\det A$ *is already known?*

Exercise 5.5. *What is matrix multiplication? Remember how it is defined. Suppose you have a rectangular* 5×3 *matrix* A *and another rectangular matrix* B *which is* 4×5. *Which of these two products* $A\,B$ *or* $B\,A$ *you can calculate?*

Exercise 5.6. *Suppose that* A *and* B *are two rectangular matrices, and suppose that* $C = A\,B$. *Remember the formula for the components in matrix* C *if the components of* A *and* B *are known (they are denoted by* A_{ij} *and* B_{pq}*). Rewrite this formula for the case when the components of* B *are denoted by* B^{pq}. *Which indices (upper, or lower, or mixed) you would use for components of* C *in the last case (see rules 5.1 and 5.2 of Einstein's tensorial notation).*

Exercise 5.7. *Give some examples of matrix multiplication that are consistent with Einstein's tensorial notation and those that are not (please, do not use examples that are already considered in exercise 5.6).*

Let's consider three bases: basis one $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, basis two $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$, and basis three $\tilde{\tilde{\mathbf{e}}}_1, \tilde{\tilde{\mathbf{e}}}_2, \tilde{\tilde{\mathbf{e}}}_3$. And let's consider the transition matrices relating them:

$$(\mathbf{e}_1,\,\mathbf{e}_2,\,\mathbf{e}_3) \;\underset{T}{\overset{S}{\rightleftarrows}}\; (\tilde{\mathbf{e}}_1,\,\tilde{\mathbf{e}}_2,\,\tilde{\mathbf{e}}_3) \;\underset{\tilde{T}}{\overset{\tilde{S}}{\rightleftarrows}}\; (\tilde{\tilde{\mathbf{e}}}_1,\,\tilde{\tilde{\mathbf{e}}}_2,\,\tilde{\tilde{\mathbf{e}}}_3). \tag{5.13}$$

Denote by $\tilde{\tilde{S}}$ and $\tilde{\tilde{T}}$ transition matrices relating basis one with basis three in (5.13):

$$(\mathbf{e}_1,\,\mathbf{e}_2,\,\mathbf{e}_3) \;\underset{\tilde{\tilde{T}}}{\overset{\tilde{\tilde{S}}}{\rightleftarrows}}\; (\tilde{\tilde{\mathbf{e}}}_1,\,\tilde{\tilde{\mathbf{e}}}_2,\,\tilde{\tilde{\mathbf{e}}}_3). \tag{5.14}$$

Exercise 5.8. *For matrices* $\tilde{\tilde{S}}$ *and* $\tilde{\tilde{T}}$ *in* (5.14) *prove that* $\tilde{\tilde{S}} = S\,\tilde{S}$ *and* $\tilde{\tilde{T}} = \tilde{T}\,T$. *Apply this result for proving theorem 5.1.*

§ 6. What happens to vectors when we change the basis?

The answer to this question is very simple. Really nothing! Vectors do not need a basis for their being. But their coordinates, they depend on our choice of basis. And they change if we change the basis. Let's study how they change. Suppose we have some vector $\mathbf{x}$ expanded in the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$:

$$\mathbf{x} = x^1\,\mathbf{e}_1 + x^2\,\mathbf{e}_2 + x^3\,\mathbf{e}_3 = \sum_{i=1}^{3} x^i\,\mathbf{e}_i. \tag{6.1}$$

Then we keep vector $\mathbf{x}$ and change the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ to another basis $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$. As we already learned, this process is described by transition formula (5.11):

$$\mathbf{e}_i = \sum_{j=1}^{3} T^j_i\,\tilde{\mathbf{e}}_j.$$

Let's substitute this formula into (6.1) for $\mathbf{e}_i$:

$$\begin{aligned}\mathbf{x}&=\sum_{i=1}^{3}x^i\left(\sum_{j=1}^{3}T^j_i\,\mathbf{e}_j\right)=\sum_{i=1}^{3}\sum_{j=1}^{3}x^i\,T^j_i\,\tilde{\mathbf{e}}_j=\sum_{j=1}^{3}\sum_{i=1}^{3}x^i\,T^j_i\,\tilde{\mathbf{e}}_j=\\&=\sum_{j=1}^{3}\left(\sum_{i=1}^{3}T^j_i\,x^i\right)\tilde{\mathbf{e}}_j=\sum_{j=1}^{3}\tilde{x}^j\,\tilde{\mathbf{e}}_j,\ \text{ where }\ \tilde{x}^j=\sum_{i=1}^{3}T^j_i\,x^i.\end{aligned}$$

Thus we have calculated the expansion of vector $\mathbf{x}$ in the new basis and have derived the formula relating its new coordinates to its initial ones:

$$\tilde{x}^j=\sum_{i=1}^{3}T^j_i\,x^i. \tag{6.2}$$

This formula is called **a transformation formula**, or **direct transformation formula**. Like (5.7), it can be written in expanded form:

$$\begin{cases}\tilde{x}^1=T^1_1\,x^1+T^1_2\,x^2+T^1_3\,x^3,\\ \tilde{x}^2=T^2_1\,x^1+T^2_2\,x^2+T^2_3\,x^3,\\ \tilde{x}^3=T^3_1\,x^1+T^3_2\,x^2+T^3_3\,x^3.\end{cases} \tag{6.3}$$

And the transformation formula (6.2) can be written in matrix form as well:

$$\begin{Vmatrix}\tilde{x}^1\\ \tilde{x}^2\\ \tilde{x}^3\end{Vmatrix}=\begin{Vmatrix}T^1_1 & T^1_2 & T^1_3\\ T^2_1 & T^2_2 & T^2_3\\ T^3_1 & T^3_2 & T^3_3\end{Vmatrix}\begin{Vmatrix}x^1\\ x^2\\ x^3\end{Vmatrix}. \tag{6.4}$$

Like (5.7), formula (6.2) can be inverted. Here is **the inverse transformation formula** expressing the initial coordinates of vector $\mathbf{x}$ through its new coordinates:

$$x^j=\sum_{i=1}^{3}S^j_i\,\tilde{x}^i. \tag{6.5}$$

Exercise 6.1. *By analogy with the above calculations derive the inverse transformation formula* (6.5) *using formula* (5.7).

Exercise 6.2. *By analogy with* (6.3) *and* (6.4) *write* (6.5) *in expanded form and in matrix form.*

Exercise 6.3. *Derive formula* (6.5) *directly from* (6.2) *using the concept of the inverse matrix* $S=T^{-1}$.

Note that the direct transformation formula (6.2) uses the inverse transition matrix T, and the inverse transformation formula (6.5) uses direct transition matrix S. It's funny, but it's really so.

§ 7. What is the novelty about the vectors that we learned knowing transformation formula for their coordinates ?

Vectors are too common, too well-known things for one to expect that there are some novelties about them. However, the novelty is that the method of their treatment can be generalized and then applied to less customary objects. Suppose, we cannot visually observe vectors (this is really so for some kinds of them, see section 1), but suppose we can measure their coordinates in any basis we choose for this purpose. What then do we know about vectors ? And how can we tell them from other (non-vectorial) objects ? The answer is in formulas (6.2) and (6.5). Coordinates of vectors (and only coordinates of vectors) will obey transformation rules (6.2) and (6.5) under a change of basis. Other objects usually have a different number of numeric parameters related to the basis, and even if they have exactly three coordinates (like vectors have), their coordinates behave differently under a change of basis. So transformation formulas (6.2) and (6.5) work like detectors, like a sieve for separating vectors from non-vectors. What are here non-vectors, and what kind of geometrical and/or physical objects of a non-vectorial nature could exist — these are questions for a separate discussion. Furthermore, we shall consider only a part of the set of such objects, which are called tensors.

CHAPTER II

TENSORS IN CARTESIAN COORDINATES.

§ 8. Covectors.

In previous 7 sections we learned the following important fact about vectors: a vector is a physical object in each basis of our three-dimensional Euclidean space E represented by three numbers such that these numbers obey certain transformation rules when we change the basis. These certain transformation rules are represented by formulas (6.2) and (6.5).

Now suppose that we have some other physical object that is represented by three numbers in each basis, and suppose that these numbers obey some certain transformation rules when we change the basis, but these rules are different from (6.2) and (6.5). Is it possible? One can try to find such an object in nature. However, in mathematics we have another option. We can construct such an object mentally, then study its properties, and finally look if it is represented somehow in nature.

Let's denote our hypothetical object by $\mathbf{a}$, and denote by a_1, a_2, a_3 that three numbers which represent this object in the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. By analogy with vectors we shall call them **coordinates**. But in contrast to vectors, we intentionally used lower indices when denoting them by a_1, a_2, a_3. Let's prescribe the following transformation rules to a_1, a_2, a_3 when we change $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ to $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$:

$$\tilde{a}_j = \sum_{i=1}^{3} S^i_j \, a_i, \tag{8.1}$$

$$a_j = \sum_{i=1}^{3} T^i_j \, \tilde{a}_i. \tag{8.2}$$

Here S and T are the same transition matrices as in case of the vectors in (6.2) and (6.5). Note that (8.1) is sufficient, formula (8.2) is derived from (8.1).

EXERCISE 8.1. *Using the concept of the inverse matrix $T = S^{-1}$ derive formula* (8.2) *from formula* (8.1). *Compare exercise 8.1 and exercise 6.3.*

DEFINITION 8.1. A geometric object $\mathbf{a}$ in each basis represented by a triple of coordinates a_1, a_2, a_3 and such that its coordinates obey transformation rules (8.1) and (8.2) under a change of basis is called **a covector**.

Looking at the above considerations one can think that we arbitrarily chose the transformation formula (8.1). However, this is not so. The choice of the transformation formula should be self-consistent in the following sense. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$ be two bases and let $\tilde{\tilde{\mathbf{e}}}_1, \tilde{\tilde{\mathbf{e}}}_2, \tilde{\tilde{\mathbf{e}}}_3$ be the third basis in the space. Let's call them basis one, basis two and basis three for short. We can pass from basis one to basis three directly, see the right arrow in (5.14). Or we can use basis two as an intermediate basis, see the right arrows in (5.13). In both cases the ultimate result for the coordinates of a covector in basis three should be the same: this is the self-consistence requirement. It means that coordinates of a geometric object should depend on the basis, but not on the way that they were calculated.

EXERCISE 8.2. *Using* (5.13) *and* (5.14), *and relying on the results of exercise 5.8 prove that formulas* (8.1) *and* (8.2) *yield a self-consistent way of defining the covector.*

EXERCISE 8.3. *Replace S by T in* (8.1) *and T by S in* (8.2). *Show that the resulting formulas are not self-consistent.*

What about the physical reality of covectors? Later on we shall see that covectors do exist in nature. They are the nearest relatives of vectors. And moreover, we shall see that some well-known physical objects we thought to be vectors are of covectorial nature rather than vectorial.

§ 9. Scalar product of vector and covector.

Suppose we have a vector $\mathbf{x}$ and a covector $\mathbf{a}$. Upon choosing some basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, both of them have three coordinates: x^1, x^2, x^3 for vector $\mathbf{x}$, and a_1, a_2, a_3 for covector $\mathbf{a}$. Let's denote by $\langle\mathbf{a}, \mathbf{x}\rangle$ the following sum:

$$\langle\mathbf{a}, \mathbf{x}\rangle = \sum_{i=1}^{3} a_i\, x^i. \tag{9.1}$$

The sum (9.1) is written in agreement with Einstein's tensorial notation, see rule 5.2 in section 5 above. It is a number depending on the vector $\mathbf{x}$ and on the covector $\mathbf{a}$. This number is called the scalar product of the vector $\mathbf{x}$ and the covector $\mathbf{a}$. We use angular brackets for this scalar product in order to distinguish it from the scalar product of two vectors in E, which is also known as the dot product.

Defining the scalar product $\langle\mathbf{a}, \mathbf{x}\rangle$ by means of sum (9.1) we used the coordinates of vector $\mathbf{x}$ and of covector $\mathbf{a}$, which are basis-dependent. However, the value of sum (9.1) does not depend on any basis. Such numeric quantities that do not depend on the choice of basis are called **scalars** or **true scalars**.

EXERCISE 9.1. *Consider two bases $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3$, and consider the coordinates of vector $\mathbf{x}$ and covector $\mathbf{a}$ in both of them. Relying on transformation rules* (6.2), (6.5), (8.1), *and* (8.2) *prove the equality*

$$\sum_{i=1}^{3} a_i\, x^i = \sum_{i=1}^{3} \tilde{a}_i\, \tilde{x}^i. \tag{9.2}$$

Thus, you are proving the self-consistence of formula (9.1) *and showing that the scalar product $\langle\mathbf{a}, \mathbf{x}\rangle$ given by this formula is a true scalar quantity.*

Exercise 9.2. *Let α be a real number, let $\mathbf{a}$ and $\mathbf{b}$ be two covectors, and let $\mathbf{x}$ and $\mathbf{y}$ be two vectors. Prove the following properties of the scalar product* (9.1)*:*

(1) $\langle \mathbf{a}+\mathbf{b},\, \mathbf{x}\rangle = \langle \mathbf{a},\, \mathbf{x}\rangle + \langle \mathbf{b},\, \mathbf{x}\rangle$; (3) $\langle \mathbf{a},\, \mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{a},\, \mathbf{x}\rangle + \langle \mathbf{a},\, \mathbf{y}\rangle$;
(2) $\langle \alpha\,\mathbf{a},\, \mathbf{x}\rangle = \alpha\,\langle \mathbf{a},\, \mathbf{x}\rangle$; (4) $\langle \mathbf{a},\, \alpha\,\mathbf{x}\rangle = \alpha\,\langle \mathbf{a},\, \mathbf{x}\rangle$.

Exercise 9.3. *Explain why the scalar product $\langle \mathbf{a},\, \mathbf{x}\rangle$ is sometimes called the bilinear function of vectorial argument $\mathbf{x}$ and covectorial argument $\mathbf{a}$. In this capacity, it can be denoted as $f(\mathbf{a}, \mathbf{x})$. Remember our discussion about functions with non-numeric arguments in section 2.*

Important note. The scalar product $\langle \mathbf{a},\, \mathbf{x}\rangle$ is not symmetric. Moreover, the formula

$$\langle \mathbf{a},\, \mathbf{x}\rangle = \cancel{\langle \mathbf{x},\, \mathbf{a}\rangle}$$

is incorrect in its right hand side since the first argument of scalar product (9.1) by definition should be a covector. In a similar way, the second argument should be a vector. Therefore, we never can swap them.

§ 10. Linear operators.

In this section we consider more complicated geometric objects. For the sake of certainty, let's denote one of such objects by $\mathbf{F}$. In each basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, it is represented by a square 3×3 matrix F^i_j of real numbers. Components of this matrix play the same role as coordinates in the case of vectors or covectors. Let's prescribe the following transformation rules to F^i_j:

$$\tilde F^i_j = \sum_{p=1}^{3}\sum_{q=1}^{3} T^i_p\, S^q_j\, F^p_q, \tag{10.1}$$

$$F^i_j = \sum_{p=1}^{3}\sum_{q=1}^{3} S^i_p\, T^q_j\, \tilde F^p_q. \tag{10.2}$$

Exercise 10.1. *Using the concept of the inverse matrix $T = S^{-1}$ prove that formula* (10.2) *is derived from formula* (10.1).

If we write matrices F^i_j and $\tilde F^p_q$ according to the rule 5.3 (see in section 5), then (10.1) and (10.2) can be written as two matrix equalities:

$$\tilde F = T\,F\,S, \qquad\qquad F = S\,\tilde F\,T. \tag{10.3}$$

Exercise 10.2. *Remember matrix multiplication (we already considered it in exercises 5.5 and 5.6) and derive* (10.3) *from* (10.1) *and* (10.2).

Definition 10.1. A geometric object $\mathbf{F}$ in each basis represented by some square matrix F^i_j and such that components of its matrix F^i_j obey transformation rules (10.1) and (10.2) under a change of basis is called **a linear operator**.

EXERCISE 10.3. *By analogy with exercise 8.2 prove the self-consistence of the above definition of a linear operator.*

Let's take a linear operator $\mathbf{F}$ represented by matrix F^i_j in some basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and take some vector $\mathbf{x}$ with coordinates x^1, x^2, x^3 in the same basis. Using F^i_j and x^j we can construct the following sum:

$$y^i = \sum_{j=1}^{3} F^i_j \, x^j. \tag{10.4}$$

Index i in the sum (10.4) is a free index; it can deliberately take any one of three values: $i = 1$, $i = 2$, or $i = 3$. For each specific value of i we get some specific value of the sum (10.4). They are denoted by y^1, y^2, y^3 according to (10.4). Now suppose that we pass to another basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ and do the same things. As a result we get other three values $\tilde{y}^1$, $\tilde{y}^2$, $\tilde{y}^3$ given by formula

$$\tilde{y}^p = \sum_{q=1}^{3} \tilde{F}^p_q \, \tilde{x}^q. \tag{10.5}$$

EXERCISE 10.4. *Relying upon* (10.1) *and* (10.2) *prove that the three numbers* y^1, y^2, y^3 *and the other three numbers* $\tilde{y}^1$, $\tilde{y}^2$, $\tilde{y}^3$ *are related as follows:*

$$\tilde{y}^j = \sum_{i=1}^{3} T^j_i \, y^i, \qquad\qquad y^j = \sum_{i=1}^{3} S^j_i \, \tilde{y}^i. \tag{10.6}$$

EXERCISE 10.5. *Looking at* (10.6) *and comparing it with* (6.2) *and* (6.5) *find that the* y^1, y^2, y^3 *and* $\tilde{y}^1$, $\tilde{y}^2$, $\tilde{y}^3$ *calculated by formulas* (10.4) *and* (10.5) *represent the same vector, but in different bases.*

Thus formula (10.4) defines the vectorial object $\mathbf{y}$, while exercise 10.5 assures the correctness of this definition. As a result we have vector $\mathbf{y}$ determined by a linear operator $\mathbf{F}$ and by vector $\mathbf{x}$. Therefore, we write

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) \tag{10.7}$$

and say that $\mathbf{y}$ is obtained by applying linear operator $\mathbf{F}$ to vector $\mathbf{x}$. Some people like to write (10.7) without parentheses:

$$\mathbf{y} = \mathbf{F}\,\mathbf{x}. \tag{10.8}$$

Formula (10.8) is a more algebraistic form of formula (10.7). Here the action of operator $\mathbf{F}$ upon vector $\mathbf{x}$ is designated like a kind of multiplication. There is also a matrix representation of formula (10.8), in which $\mathbf{x}$ and $\mathbf{y}$ are represented as columns:

$$\begin{Vmatrix} y^1 \\ y^2 \\ y^3 \end{Vmatrix} = \begin{Vmatrix} F^1_1 & F^1_2 & F^1_3 \\ F^2_1 & F^2_2 & F^2_3 \\ F^3_1 & F^3_2 & F^3_3 \end{Vmatrix} \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix}. \tag{10.9}$$

Exercise 10.6. *Derive* (10.9) *from* (10.4).

Exercise 10.7. *Let α be some real number and let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two vectors. Prove the following properties of a linear operator* (10.7)*:*

(1) $\mathbf{F}(\mathbf{x}+\mathbf{y})=\mathbf{F}(\mathbf{x})+\mathbf{F}(\mathbf{y})$,
(2) $\mathbf{F}(\alpha\,\mathbf{x})=\alpha\,\mathbf{F}(\mathbf{x})$.

Write these equalities in the more algebraistic style introduced by (10.8). *Are they really similar to the properties of multiplication?*

Exercise 10.8. *Remember that for the product of two matrices*

$$\det(A\,B)=\det A\,\det B. \tag{10.10}$$

Also remember the formula for $\det(A^{-1})$. *Apply these two formulas to* (10.3) *and derive*

$$\det F=\det\tilde{F}. \tag{10.11}$$

Formula (10.10) means that despite the fact that in various bases linear operator $\mathbf{F}$ is represented by various matrices, the determinants of all these matrices are equal to each other. Then we can define the determinant of linear operator $\mathbf{F}$ as the number equal to the determinant of its matrix in any one arbitrarily chosen basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\det\mathbf{F}=\det F. \tag{10.12}$$

Exercise 10.9 (for deep thinking). *Square matrices have various attributes: eigenvalues, eigenvectors, a characteristic polynomial, a rank (maybe you remember some others). If we study these attributes for the matrix of a linear operator, which of them can be raised one level up and considered as basis-independent attributes of the linear operator itself? Determinant* (10.12) *is an example of such attribute.*

Exercise 10.10. *Substitute the unit matrix for* F^i_j *into* (10.1) *and verify that* $\tilde{F}^i_j$ *is also a unit matrix in this case. Interpret this fact.*

Exercise 10.11. *Let* $\mathbf{x}=\mathbf{e}_i$ *for some basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space. Substitute this vector* $\mathbf{x}$ *into* (10.7) *and by means of* (10.4) *derive the following formula:*

$$\mathbf{F}(\mathbf{e}_i)=\sum_{j=1}^{3}F^j_i\,\mathbf{e}_j. \tag{10.13}$$

Compare (10.13) *and* (5.7). *Discuss the similarities and differences of these two formulas. The fact is that in some books the linear operator is determined first, then its matrix is introduced by formula* (10.13). *Explain why if we know three vectors* $\mathbf{F}(\mathbf{e}_1)$, $\mathbf{F}(\mathbf{e}_2)$, *and* $\mathbf{F}(\mathbf{e}_3)$, *then we can reconstruct the whole matrix of operator* $\mathbf{F}$ *by means of formula* (10.13).

Suppose we have two linear operators $\mathbf{F}$ and $\mathbf{H}$. We can apply $\mathbf{H}$ to vector $\mathbf{x}$ and then we can apply $\mathbf{F}$ to vector $\mathbf{H}(\mathbf{x})$. As a result we get

$$\mathbf{F}\circ\mathbf{H}(\mathbf{x})=\mathbf{F}(\mathbf{H}(\mathbf{x})). \tag{10.14}$$

Here $\mathbf{F}\circ\mathbf{H}$ is new linear operator introduced by formula (10.14). It is called **a composite operator**, and the small circle sign denotes **composition**.

Exercise 10.12. *Find the matrix of composite operator* $\mathbf{F} \circ \mathbf{H}$ *if the matrices for* $\mathbf{F}$ *and* $\mathbf{H}$ *in the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *are known.*

Exercise 10.13. *Remember the definition of the identity map in mathematics (see* on-line Math. Encyclopedia) and define the identity operator **id**. Find the matrix of this operator.

Exercise 10.14. *Remember the definition of the inverse map in mathematics and define inverse operator* $\mathbf{F}^{-1}$ *for linear operator* $\mathbf{F}$*. Find the matrix of this operator if the matrix of* $\mathbf{F}$ *is known.*

§ 11. Bilinear and quadratic forms.

Vectors, covectors, and linear operators are all examples of tensors (though we have no definition of tensors yet). Now we consider another one class of tensorial objects. For the sake of clarity, let's denote by a one of such objects. In each basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ this object is represented by some square 3×3 matrix a_{ij} of real numbers. Under a change of basis these numbers are transformed as follows:

$$\tilde{a}_{ij} = \sum_{p=1}^{3} \sum_{q=1}^{3} S_i^p \, S_j^q \, a_{pq}, \tag{11.1}$$

$$a_{ij} = \sum_{p=1}^{3} \sum_{q=1}^{3} T_i^p \, T_j^q \, \tilde{a}_{pq}. \tag{11.2}$$

Transformation rules (11.1) and (11.2) can be written in matrix form:

$$\tilde{a} = S^{\top} a \, S, \qquad\qquad a = T^{\top} \tilde{a} \, T. \tag{11.3}$$

Here by $S^{\top}$ and $T^{\top}$ we denote the transposed matrices for S and T respectively.

Exercise 11.1. *Derive* (11.2) *from* (11.1)*, then* (11.3) *from* (11.1) *and* (11.2).

Definition 11.1. A geometric object a in each basis represented by some square matrix a_{ij} and such that components of its matrix a_{ij} obey transformation rules (11.1) and (11.2) under a change of basis is called **a bilinear form**.

Let's consider two arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$. We use their coordinates and the components of bilinear form a in order to write the following sum:

$$a(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \, x^i \, y^j. \tag{11.4}$$

Exercise 11.2. *Prove that the sum in the right hand side of formula* (11.4) *does not depend on the basis, i. e. prove the equality*

$$\sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \, x^i \, y^j = \sum_{p=1}^{3} \sum_{q=1}^{3} \tilde{a}_{pq} \, \tilde{x}^p \, \tilde{y}^q.$$

This equality means that $a(\mathbf{x}, \mathbf{y})$ is a number determined by vectors $\mathbf{x}$ and $\mathbf{y}$ irrespective of the choice of basis. Hence we can treat (11.4) as a scalar function of two vectorial arguments.

Exercise 11.3. *Let α be some real number, and let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be three vectors. Prove the following properties of function* (11.4)*:*

(1) $a(\mathbf{x}+\mathbf{y},\mathbf{z}) = a(\mathbf{x},\mathbf{z})+a(\mathbf{y},\mathbf{z})$; (3) $a(\mathbf{x},\mathbf{y}+\mathbf{z}) = a(\mathbf{x},\mathbf{y})+a(\mathbf{x},\mathbf{z})$;
(2) $a(\alpha\,\mathbf{x},\mathbf{y}) = \alpha\,a(\mathbf{x},\mathbf{y})$; (4) $a(\mathbf{x},\alpha\,\mathbf{y}) = \alpha\,a(\mathbf{x},\mathbf{y})$.

Due to these properties function (10.4) *is called a bilinear function or a bilinear form. It is linear with respect to each of its two arguments.*

Note that scalar product (9.1) is also a bilinear function of its arguments. However, there is a crucial difference between (9.1) and (11.4). The arguments of scalar product (9.1) are of a different nature: the first argument is a covector, the second is a vector. Therefore, we cannot swap them. In bilinear form (11.4) we can swap arguments. As a result we get another bilinear function

$$b(\mathbf{x},\mathbf{y}) = a(\mathbf{y},\mathbf{x}). \tag{11.5}$$

The matrices of a and b are related to each other as follows:

$$b_{ij} = a_{ji}, \qquad\qquad b = a^{\top}. \tag{11.6}$$

Definition 11.2. A bilinear form is called symmetric if $a(\mathbf{x},\mathbf{y}) = a(\mathbf{y},\mathbf{x})$.

Exercise 11.4. *Prove the following identity for a symmetric bilinear form:*

$$a(\mathbf{x},\mathbf{y}) = \frac{a(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) - a(\mathbf{x},\mathbf{x}) - a(\mathbf{y},\mathbf{y})}{2}. \tag{11.7}$$

Definition 11.3. A quadratic form is a scalar function of one vectorial argument $f(\mathbf{x})$ produced from some bilinear function $a(\mathbf{x},\mathbf{y})$ by substituting $\mathbf{y} = \mathbf{x}$:

$$f(\mathbf{x}) = a(\mathbf{x},\mathbf{x}). \tag{11.8}$$

Without a loss of generality a bilinear function a in (11.8) can be assumed symmetric. Indeed, if a is not symmetric, we can produce symmetric bilinear function

$$c(\mathbf{x},\mathbf{y}) = \frac{a(\mathbf{x},\mathbf{y}) + a(\mathbf{y},\mathbf{x})}{2}, \tag{11.9}$$

and then from (11.8) due to (11.9) we derive

$$f(\mathbf{x}) = a(\mathbf{x},\mathbf{x}) = \frac{a(\mathbf{x},\mathbf{x}) + a(\mathbf{x},\mathbf{x})}{2} = c(\mathbf{x},\mathbf{x}).$$

This equality is the same as (11.8) with a replaced by c. Thus, each quadratic function f is produced by some symmetric bilinear function a. And conversely, comparing (11.8) and (11.7) we get that a is produced by f:

$$a(\mathbf{x},\mathbf{y}) = \frac{f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x}) - f(\mathbf{y})}{2}. \tag{11.10}$$

Formula (11.10) is called **the recovery formula**. It recovers bilinear function a from quadratic function f produced in (11.8). Due to this formula, in referring to a quadratic form we always imply some symmetric bilinear form like the geometric tensorial object introduced by definition 11.1.

§ 12. General definition of tensors.

Vectors, covectors, linear operators, and bilinear forms are examples of tensors. They are geometric objects that are represented numerically when some basis in the space is chosen. This numeric representation is specific to each of them: vectors and covectors are represented by one-dimensional arrays, linear operators and quadratic forms are represented by two-dimensional arrays. Apart from the number of indices, their position does matter. The coordinates of a vector are numerated by one upper index, which is called the contravariant index. The coordinates of a covector are numerated by one lower index, which is called the covariant index. In a matrix of bilinear form we use two lower indices; therefore bilinear forms are called **twice-covariant tensors**. Linear operators are tensors of **mixed type**; their components are numerated by one upper and one lower index. The number of indices and their positions determine the transformation rules, i. e. the way the components of each particular tensor behave under a change of basis. In the general case, any tensor is represented by a multidimensional array with a definite number of upper indices and a definite number of lower indices. Let's denote these numbers by r and s. Then we have **a tensor of the type** (r, s), or sometimes the term **valency** is used. A tensor of type (r, s), or of valency (r, s) is called **an r-times contravariant** and **an s-times covariant** tensor. This is terminology; now let's proceed to the exact definition. It is based on the following general transformation formulas:

$$X^{i_1\ldots\, i_r}_{j_1\ldots\, j_s} = \sum^{3}_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\sum^{3} S^{i_1}_{h_1}\ldots S^{i_r}_{h_r} T^{k_1}_{j_1}\ldots T^{k_s}_{j_s}\,\tilde{X}^{h_1\ldots\, h_r}_{k_1\ldots\, k_s}, \tag{12.1}$$

$$\tilde{X}^{i_1\ldots\, i_r}_{j_1\ldots\, j_s} = \sum^{3}_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\sum^{3} T^{i_1}_{h_1}\ldots T^{i_r}_{h_r} S^{k_1}_{j_1}\ldots S^{k_s}_{j_s}\,X^{h_1\ldots\, h_r}_{k_1\ldots\, k_s}. \tag{12.2}$$

Definition 12.1. A geometric object $\mathbf{X}$ in each basis represented by $(r+s)$-dimensional array $X^{i_1\ldots\, i_r}_{j_1\ldots\, j_s}$ of real numbers and such that the components of this array obey the transformation rules (12.1) and (12.2) under a change of basis is called **tensor** of type (r, s), or of valency (r, s).

Formula (12.2) is derived from (12.1), so it is sufficient to remember only one of them. Let it be the formula (12.1). Though huge, formula (12.1) is easy to remember. One should strictly follow the rules 5.1 and 5.2 from section 5.

Indices $i_1, \ldots, i_r$ and $j_1, \ldots, j_s$ are free indices. In right hand side of the equality (12.1) they are distributed in S-s and T-s, each having only one entry and each keeping its position, i. e. upper indices $i_1, \ldots, i_r$ remain upper and lower indices $j_1, \ldots, j_s$ remain lower in right hand side of the equality (12.1).

Other indices $h_1, \ldots, h_r$ and $k_1, \ldots, k_s$ are summation indices; they enter the right hand side of (12.1) pairwise: once as an upper index and once as a lower index, once in S-s or T-s and once in components of array $\tilde{X}^{h_1\ldots\, h_r}_{k_1\ldots\, k_s}$.

When expressing $X^{i_1\ldots\, i_r}_{j_1\ldots\, j_s}$ through $\tilde{X}^{h_1\ldots\, h_r}_{k_1\ldots\, k_s}$ each upper index is served by direct transition matrix S and produces one summation in (12.1):

$$X^{\ldots\, i_\alpha\, \ldots}_{\ldots\,\ldots\,\ldots} = \sum\ldots\sum^{3}_{h_\alpha=1}\ldots\sum\ldots S^{i_\alpha}_{h_\alpha}\ldots\tilde{X}^{\ldots\, h_\alpha\,\ldots}_{\ldots\,\ldots\,\ldots}. \tag{12.3}$$

In a similar way, each lower index is served by inverse transition matrix T and also produces one summation in formula (12.1):

$$X^{\cdots\,\cdots\,\cdots}_{\cdots\, j_\alpha \,\cdots} = \sum \ldots \sum_{k_\alpha=1}^{3} \ldots \sum \ldots T^{k_\alpha}_{j_\alpha} \ \ldots\ \tilde{X}^{\cdots\,\cdots\,\cdots}_{\cdots\, k_\alpha \,\cdots}. \tag{12.4}$$

Formulas (12.3) and (12.4) are the same as (12.1) and used to highlight how (12.1) is written. So tensors are defined. Further we shall consider more examples showing that many well-known objects undergo the definition 12.1.

Exercise 12.1. *Verify that formulas* (6.5), (8.2), (10.2), *and* (11.2) *are special cases of formula* (12.1). *What are the valencies of vectors, covectors, linear operators, and bilinear forms when they are considered as tensors.*

Exercise 12.2. *Let a_{ij} be the matrix of some bilinear form a. Let's denote by b^{ij} components of inverse matrix for a_{ij}. Prove that matrix b^{ij} under a change of basis transforms like matrix of twice-contravariant tensor. Hence it determines tensor b of valency* $(2, 0)$. *Tensor b is called* **a dual bilinear form** *for a.*

§ 13. Dot product and metric tensor.

The covectors, linear operators, and bilinear forms that we considered above were artificially constructed tensors. However there are some tensors of natural origin. Let's remember that we live in a space with measure. We can measure distance between points (hence we can measure length of vectors) and we can measure angles between two directions in our space. Therefore for any two vectors $\mathbf{x}$ and $\mathbf{y}$ we can define their natural scalar product (or dot product):

$$(\mathbf{x},\,\mathbf{y}) = |\mathbf{x}|\,|\mathbf{y}|\,\cos(\varphi), \tag{13.1}$$

where φ is the angle between vectors $\mathbf{x}$ and $\mathbf{y}$.

Exercise 13.1. *Remember the following properties of the scalar product* (13.1)*:*

(1) $(\mathbf{x}+\mathbf{y},\,\mathbf{z}) = (\mathbf{x},\,\mathbf{z}) + (\mathbf{y},\,\mathbf{z})$; (3) $(\mathbf{x},\,\mathbf{y}+\mathbf{z}) = (\mathbf{x},\,\mathbf{y}) + (\mathbf{x},\,\mathbf{z})$;
(2) $(\alpha\,\mathbf{x},\,\mathbf{y}) = \alpha\,(\mathbf{x},\,\mathbf{y})$; (4) $(\mathbf{x},\,\alpha\,\mathbf{y}) = \alpha\,(\mathbf{x},\,\mathbf{y})$;
(5) $(\mathbf{x},\,\mathbf{y}) = (\mathbf{y},\,\mathbf{x})$;
(6) $(\mathbf{x},\,\mathbf{x}) \geqslant 0$ *and* $(\mathbf{x},\,\mathbf{x}) = 0$ *implies* $\mathbf{x} = 0$.

These properties are usually considered in courses on analytic geometry or vector algebra, see Vector Lessons on the Web.

Note that the first four properties of the scalar product (13.1) are quite similar to those for quadratic forms, see exercise 11.3. This is not an occasional coincidence.

Exercise 13.2. *Let's consider two arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$ expanded in some basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. This means that we have the following expressions for them:*

$$\mathbf{x} = \sum_{i=1}^{3} x^i\,\mathbf{e}_i, \qquad\qquad \mathbf{y} = \sum_{j=1}^{3} x^j\,\mathbf{e}_j. \tag{13.2}$$

Substitute (13.2) *into* (13.1) *and using properties* (1)–(4) *listed in exercise 13.1 derive the following formula for the scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$*:*

$$(\mathbf{x},\,\mathbf{y})=\sum_{i=1}^{3}\sum_{j=1}^{3}(\mathbf{e}_i,\,\mathbf{e}_j)\,x^i\,y^j. \tag{13.3}$$

Exercise 13.3. *Denote* $g_{ij}=(\mathbf{e}_i,\,\mathbf{e}_j)$ *and rewrite formula* (13.3) *as*

$$(\mathbf{x},\,\mathbf{y})=\sum_{i=1}^{3}\sum_{j=1}^{3}g_{ij}\,x^i\,y^j. \tag{13.4}$$

Compare (13.4) *with formula* (11.4)*. Consider some other basis* $\tilde{\mathbf{e}}_1,\,\tilde{\mathbf{e}}_2,\,\tilde{\mathbf{e}}_3$*, denote* $\tilde{g}_{pq}=(\tilde{\mathbf{e}}_p,\,\tilde{\mathbf{e}}_q)$ *and by means of transition formulas* (5.7) *and* (5.11) *prove that matrices* g_{ij} *and* $\tilde{g}_{pq}$ *are components of a geometric object obeying transformation rules* (11.1) *and* (11.2) *under a change of base. Thus you prove that the Gram matrix*

$$g_{ij}=(\mathbf{e}_i,\,\mathbf{e}_j) \tag{13.5}$$

defines tensor of type $(0,2)$*. This is very important tensor; it is called* **the metric tensor***. It describes not only the scalar product in form of* (13.4)*, but the whole geometry of our space. Evidences for this fact are below.*

Matrix (13.5) is symmetric due to property (5) in exercise 13.1. Now, comparing (13.4) and (11.4) and keeping in mind the tensorial nature of matrix (13.5), we conclude that the scalar product is a symmetric bilinear form:

$$(\mathbf{x},\,\mathbf{y})=g(\mathbf{x},\mathbf{y}). \tag{13.6}$$

The quadratic form corresponding to (13.6) is very simple: $f(\mathbf{x})=g(\mathbf{x},\mathbf{x})=|\mathbf{x}|^2$. The inverse matrix for (13.5) is denoted by the same symbol g but with upper indices: g^{ij}. It determines a tensor of type $(2,0)$, this tensor is called **dual metric tensor** (see exercise 12.2 for more details).

§ 14. Multiplication by numbers and addition.

Tensor operations are used to produce new tensors from those we already have. The most simple of them are **multiplication by number** and **addition**. If we have some tensor $\mathbf{X}$ of type (r,s) and a real number α, then in some base $\mathbf{e}_1,\,\mathbf{e}_2,\,\mathbf{e}_3$ we have the array of components of tensor X; let's denote it $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$. Then by multiplying all the components of this array by α we get another array

$$Y^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}=\alpha\,X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}. \tag{14.1}$$

Choosing another base $\tilde{\mathbf{e}}_1,\,\tilde{\mathbf{e}}_2,\,\tilde{\mathbf{e}}_3$ and repeating this operation we get

$$\tilde{Y}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}=\alpha\,\tilde{X}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}. \tag{14.2}$$

Exercise 14.1. *Prove that arrays* $\tilde{Y}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ *and* $Y^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ *are related to each other in the same way as arrays* $\tilde{X}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ *and* $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$*, i.e. according to transformation formulas* (12.1) *and* (12.2)*. In doing this you prove that formula* (14.1) *applied in all bases produces new tensor* $\mathbf{Y} = \alpha\,\mathbf{X}$ *from initial tensor* $\mathbf{X}$.

Formula (14.1) defines **the multiplication of tensors by numbers**. In exercise 14.1 you prove its consistence. The next formula defines **the addition of tensors**:

$$X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s} + Y^{i_1\ldots\,i_r}_{j_1\ldots\,j_s} = Z^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}. \tag{14.3}$$

Having two tensors $\mathbf{X}$ and $\mathbf{Y}$ both of type (r,s) we produce a third tensor $\mathbf{Z}$ of the same type (r,s) by means of formula (14.3). It's natural to denote $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$.

Exercise 14.2. *By analogy with exercise 14.1 prove the consistence of formula* (14.3).

Exercise 14.3. *What happens if we multiply tensor* $\mathbf{X}$ *by the number zero and by the number minus one? What would you call the resulting tensors?*

§ 15. Tensor product.

The tensor product is defined by a more tricky formula. Suppose we have tensor $\mathbf{X}$ of type (r,s) and tensor $\mathbf{Y}$ of type (p,q), then we can write:

$$Z^{i_1\ldots\,i_{r+p}}_{j_1\ldots\,j_{s+q}} = X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}\,Y^{i_{r+1}\ldots\,i_{r+p}}_{j_{s+1}\ldots\,j_{s+q}}. \tag{15.1}$$

Formula (15.1) produces new tensor $\mathbf{Z}$ of the type $(r+p, s+q)$. It is called **the tensor product** of $\mathbf{X}$ and $\mathbf{Y}$ and denoted $\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y}$. Don't mix the tensor product and the cross product. They are different.

Exercise 15.1. *By analogy with exercise 14.1 prove the consistence of formula* (15.1).

Exercise 15.2. *Give an example of two tensors such that* $\mathbf{X} \otimes \mathbf{Y} \neq \mathbf{Y} \otimes \mathbf{X}$.

§ 16. Contraction.

As we have seen above, the tensor product increases the number of indices. Usually the tensor $\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y}$ has more indices than $\mathbf{X}$ and $\mathbf{Y}$. Contraction is an operation that decreases the number of indices. Suppose we have tensor $\mathbf{X}$ of the type $(r+1, s+1)$. Then we can produce tensor $\mathbf{Z}$ of type (r,s) by means of the following formula:

$$Z^{i_1\ldots\,i_r}_{j_1\ldots\,j_s} = \sum_{\rho=1}^{n} X^{i_1\ldots\,i_{m-1}\,\rho\,i_m\ldots\,i_r}_{j_1\ldots\,j_{k-1}\,\rho\,j_k\ldots\,j_s}. \tag{16.1}$$

What we do? Tensor $\mathbf{X}$ has at least one upper index and at least one lower index. We choose the m-th upper index and replace it by the summation index ρ. In the same way we replace the k-th lower index by ρ. Other r upper indices and s lower indices are free. They are numerated in some convenient way, say as in formula (16.1). Then we perform summation with respect to index ρ. The contraction is over. This operation is called **a contraction with respect to** m-th **upper and** k-th **lower indices**. Thus, if we have many upper an many lower indices in tensor $\mathbf{X}$, we can perform various types of contractions to this tensor.

Exercise 16.1. *Prove the consistence of formula* (16.1).

Exercise 16.2. *Look at formula* (9.1) *and interpret this formula as the contraction of the tensor product* $\mathbf{a}\otimes\mathbf{x}$. *Find similar interpretations for* (10.4), (11.4), *and* (13.4).

§ 17. Raising and lowering indices.

Suppose that $\mathbf{X}$ is some tensor of type (r, s). Let's choose its α-th lower index: $X^{\dots\,\dots\,\dots}_{\dots\,k\,\dots}$. The symbols used for the other indices are of no importance. Therefore, we denoted them by dots. Then let's consider the tensor product $\mathbf{Y} = g\otimes\mathbf{X}$:

$$Y^{\dots\,p\,q\,\dots}_{\dots\,k\,\dots} = g^{pq}\,X^{\dots\,\dots\,\dots}_{\dots\,k\,\dots}. \tag{17.1}$$

Here g is the dual metric tensor with the components g^{pq} (see section 13 above). In the next step let's contract (17.1) with respect to the pair of indices k and q. For this purpose we replace them both by s and perform the summation:

$$X^{\dots\,p\,\dots}_{\dots\,\dots\,\dots} = \sum_{s=1}^{3} g^{ps}\,X^{\dots\,\dots\,\dots}_{\dots\,s\,\dots}. \tag{17.2}$$

This operation (17.2) is called **the index raising procedure**. It is invertible. The inverse operation is called **the index lowering procedure**:

$$X^{\dots\,\dots\,\dots}_{\dots\,p\,\dots} = \sum_{s=1}^{3} g_{ps}\,X^{\dots\,s\,\dots}_{\dots\,\dots\,\dots}. \tag{17.3}$$

Like (17.2), the index lowering procedure (17.3) comprises two tensorial operations: the tensor product and contraction.

§ 18. Some special tensors and some useful formulas.

Kronecker symbol is a well known object. This is a two-dimensional array representing the unit matrix. It is determined as follows:

$$\delta^i_j = \begin{cases} 1 & \text{for } \ i = j, \\ 0 & \text{for } \ i \neq j. \end{cases} \tag{18.1}$$

We can determine two other versions of Kronecker symbol:

$$\delta^{ij} = \delta_{ij} = \begin{cases} 1 & \text{for } \ i = j, \\ 0 & \text{for } \ i \neq j. \end{cases} \tag{18.2}$$

Exercise 18.1. *Prove that definition* (18.1) *is invariant under a change of basis, if we interpret the Kronecker symbol as a tensor. Show that both definitions in* (18.2) *are not invariant under a change of basis.*

Exercise 18.2. *Lower index i of tensor* (18.1) *by means of* (17.3). *What tensorial object do you get as a result of this operation?*

Exercise 18.3. *Likewise, raise index J in* (18.1).

Another well known object is the Levi-Civita symbol. This is a three-dimensional array determined by the following formula:

$$\epsilon_{jkq} = \epsilon^{jkq} = \begin{cases} 0, & \text{if among } j,\ k,\ q, \text{ there are at least two equal numbers;} \\ 1, & \text{if } (j\,k\,q) \text{ is even permutation of numbers } (1\,2\,3); \\ -1, & \text{if } (j\,k\,q) \text{ is odd permutation of numbers } (1\,2\,3). \end{cases} \tag{18.3}$$

The Levi-Civita symbol (18.3) is not a tensor. However, we can produce two tensors by means of Levi-Civita symbol. The first of them

$$\omega_{ijk} = \sqrt{\det(g_{ij})}\,\epsilon_{ijk} \tag{18.4}$$

is known as **the volume tensor**. Another one is **the dual volume tensor**:

$$\omega^{ijk} = \sqrt{\det(g^{ij})}\,\epsilon^{ijk}. \tag{18.5}$$

Let's take two vectors $\mathbf{x}$ and $\mathbf{y}$. Then using (18.4) we can produce covector $\mathbf{a}$:

$$a_i = \sum_{j=1}^{3}\sum_{k=1}^{3}\omega_{ijk}\,x^j\,y^k. \tag{18.6}$$

Then we can apply index raising procedure (17.2) and produce vector $\mathbf{a}$:

$$a^r = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} g^{ri}\,\omega_{ijk}\,x^j\,y^k. \tag{18.7}$$

Formula (18.7) is known as formula for the vectorial product (cross product) in skew-angular basis.

Exercise 18.4. *Prove that the vector* $\mathbf{a}$ *with components* (18.7) *coincides with cross product of vectors* $\mathbf{x}$ *and* $\mathbf{y}$, *i.e.* $\mathbf{a} = [\mathbf{x},\,\mathbf{y}]$.

CHAPTER III

TENSOR FIELDS. DIFFERENTIATION OF TENSORS.

§ 19. Tensor fields in Cartesian coordinates.

The tensors that we defined in section 12 are free tensors. Indeed, their components are arrays related to bases, while any basis is a triple of free vectors (not bound to any point). Hence, the tensors previously considered are also not bound to any point.

Now suppose we want to bind our tensor to some point in space, then another tensor to another point and so on. Doing so we can fill our space with tensors, one per each point. In this case we say that we have a tensor field. In order to mark a point P to which our particular tensor is bound we shall write P as an argument:

$$\mathbf{X} = \mathbf{X}(P). \tag{19.1}$$

Usually the valencies of all tensors composing the tensor field are the same. Let them all be of type (r, s). Then if we choose some basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, we can represent any tensor of our tensor field as an array $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ with $r+s$ indices:

$$X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s} = X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(P). \tag{19.2}$$

Thus, the tensor field (19.1) is a tensor-valued function with argument P being a point of three-dimensional Euclidean space E, and (19.2) is the basis representation for (19.1). For each fixed set of numeric values of indices $i_1, \ldots, i_r, j_1, \ldots, j_s$ in (19.2), we have a numeric function with a point-valued argument. Dealing with point-valued arguments is not so convenient, for example, if we want to calculate derivatives. Therefore, we need to replace P by something numeric. Remember that we have already chosen a basis. If, in addition, we fix some point O as an origin, then we get Cartesian coordinate system in space and hence can represent P by its radius-vector $\mathbf{r}_P = \overrightarrow{OP}$ and by its coordinates x^1, x^2, x^3:

$$X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s} = X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(x^1, x^2, x^3). \tag{19.3}$$

Conclusion 19.1. *In contrast to free tensors, tensor fields are related not to bases, but to whole coordinate systems (including the origin). In each coordinate system they are represented by functional arrays, i. e. by arrays of functions (see (19.3)).*

A functional array (19.3) is a coordinate representation of a tensor field (19.1). What happens when we change the coordinate system? Dealing with (19.2), we need only to recalculate the components of the array $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ in the basis by means of (12.2):

$$\tilde{X}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(P)=\sum^{3}_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\!\!\sum^{3} T^{i_1}_{h_1}\ldots\,T^{i_r}_{h_r}S^{k_1}_{j_1}\ldots\,S^{k_s}_{j_s}\,X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(P). \tag{19.4}$$

In the case of (19.3), we need to recalculate the components of the array $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ in the new basis

$$\tilde{X}^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3)=\sum^{3}_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\!\!\sum^{3} T^{i_1}_{h_1}\ldots\,T^{i_r}_{h_r}S^{k_1}_{j_1}\ldots\,S^{k_s}_{j_s}\,X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(x^1,x^2,x^3), \tag{19.5}$$

using (12.2), and we also need to express the old coordinates x^1, x^2, x^3 of the point P in right hand side of (19.5) through new coordinates of the same point:

$$\begin{cases} x^1=x^1(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3),\\ x^2=x^2(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3),\\ x^3=x^3(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3). \end{cases} \tag{19.6}$$

Like (12.2), formula (19.5) can be inverted by means of (12.1):

$$X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(x_1,x_2,x_3)=\sum^{3}_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\!\!\sum^{3} S^{i_1}_{h_1}\ldots\,S^{i_r}_{h_r}T^{k_1}_{j_1}\ldots\,T^{k_s}_{j_s}\,\tilde{X}^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3). \tag{19.7}$$

But now, apart from (19.7), we should have inverse formulas for (19.6) as well:

$$\begin{cases} \tilde{x}^1=x^1(x^1,x^2,x^3),\\ \tilde{x}^2=x^2(x^1,x^2,x^3),\\ \tilde{x}^3=x^3(x^1,x^2,x^3). \end{cases} \tag{19.8}$$

THe couple of formulas (19.5) and (19.6), and another couple of formulas (19.7) and (19.8), in the case of tensor fields play the same role as transformation formulas (12.1) and (12.2) in the case of free tensors.

§ 20. Change of Cartesian coordinate system.

Note that formulas (19.6) and (19.8) are written in abstract form. They only indicate the functional dependence of new coordinates of the point P from old ones and vice versa. Now we shall specify them for the case when one Cartesian coordinate system is changed to another Cartesian coordinate system. Remember that each Cartesian coordinate system is determined by some basis and some fixed point (the origin). We consider two Cartesian coordinate systems. Let the origins of the first and second systems be at the points O and $\tilde{O}$, respectively. Denote by $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ the basis of the first coordinate system, and by $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ the basis of the second coordinate system (see Fig. 7 below).

Let P be some point in the space for whose coordinates we are going to derive the specializations of formulas (19.6) and (19.8). Denote by $\mathbf{r}_P$ and $\tilde{\mathbf{r}}_P$ the

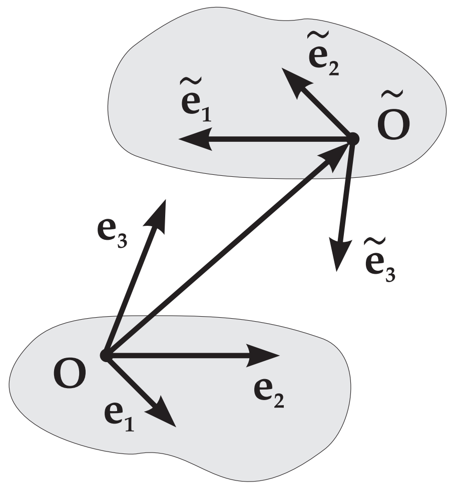


Fig. 7.

radius-vectors of this point in our two coordinate systems. Then $\mathbf{r}_P = \overrightarrow{OP}$ and $\tilde{\mathbf{r}}_P = \overrightarrow{\tilde{O}P}$. Hence

$$\mathbf{r}_P = \overrightarrow{O\tilde{O}} + \tilde{\mathbf{r}}_P. \tag{20.1}$$

Vector $\overrightarrow{O\tilde{O}}$ determines the origin shift from the old to the new coordinate system. We expand this vector in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\mathbf{a} = \overrightarrow{O\tilde{O}} = \sum_{i=1}^{3} a^i\,\mathbf{e}_i. \tag{20.2}$$

Radius-vectors $\mathbf{r}_P$ and $\tilde{\mathbf{r}}_P$ are expanded in the bases of their own coordinate systems:

$$\begin{aligned} \mathbf{r}_P &= \sum_{i=1}^{3} x^i\,\mathbf{e}_i,\\ \tilde{\mathbf{r}}_P &= \sum_{i=1}^{3} \tilde{x}^i\,\tilde{\mathbf{e}}_i, \end{aligned} \tag{20.3}$$

Exercise 20.1. *Using* (20.1), (20.2), (20.3), *and* (5.7) *derive the following formula relating the coordinates of the point P in the two coordinate systems in Fig. 7:*

$$x^i = a^i + \sum_{j=1}^{3} S^i_j\,\tilde{x}^j. \tag{20.4}$$

Compare (20.4) *with* (6.5). *Explain the differences in these two formulas.*

Exercise 20.2. *Derive the following inverse formula for* (20.4)*:*

$$\tilde{x}^i = \tilde{a}^i + \sum_{j=1}^{3} T^i_j\,x^j. \tag{20.5}$$

Prove that a^i in (20.4) *and $\tilde{a}^i$ in* (20.5) *are related to each other as follows:*

$$\tilde{a}^i = -\sum_{j=1}^{3} T^i_j\,a^j, \qquad\qquad a^i = -\sum_{j=1}^{3} S^i_j\,\tilde{a}^j. \tag{20.6}$$

Compare (20.6) *with* (6.2) *and* (6.5). *Explain the minus signs in these formulas.*

Formula (20.4) can be written in the following expanded form:

$$\begin{cases} x^1 = S^1_1\,\tilde{x}^1 + S^1_2\,\tilde{x}^2 + S^1_3\,\tilde{x}^3 + a^1,\\ x^2 = S^2_1\,\tilde{x}^1 + S^2_2\,\tilde{x}^2 + S^2_3\,\tilde{x}^3 + a^2,\\ x^3 = S^3_1\,\tilde{x}^1 + S^3_2\,\tilde{x}^2 + S^3_3\,\tilde{x}^3 + a^3. \end{cases} \tag{20.7}$$

This is the required specialization for (19.6). In a similar way we can expand (20.5):

$$\begin{cases} \tilde{x}^1 = T^1_1\,x^1 + T^1_2\,x^2 + T^1_3\,x^3 + \tilde{a}^1, \\ \tilde{x}^2 = T^2_1\,x^1 + T^2_2\,x^2 + T^2_3\,x^3 + \tilde{a}^2, \\ \tilde{x}^3 = T^3_1\,x^1 + T^3_2\,x^2 + T^3_3\,x^3 + \tilde{a}^3. \end{cases} \tag{20.8}$$

This is the required specialization for (19.8). Formulas (20.7) and (20.8) are used to accompany the main transformation formulas (19.5) and (19.7).

§ 21. Differentiation of tensor fields.

In this section we consider two different types of derivatives that are usually applied to tensor fields: differentiation with respect to spacial variables x^1, x^2, x^3 and differentiation with respect to external parameters other than x^1, x^2, x^3, if they are present. The second type of derivatives are simpler to understand. Let's consider them to start. Suppose we have tensor field $\mathbf{X}$ of type (r, s) and depending on the additional parameter t (for instance, this could be a time variable). Then, upon choosing some Cartesian coordinate system, we can write

$$\frac{\partial X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}}{\partial t} = \lim_{h\to 0} \frac{X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(t+h, x^1, x^2, x^3) - X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(t, x^1, x^2, x^3)}{h}. \tag{21.1}$$

The left hand side of (21.1) is a tensor since the fraction in right hand side is constructed by means of tensorial operations (14.1) and (14.3). Passing to the limit $h \to 0$ does not destroy the tensorial nature of this fraction since the transition matrices S and T in (19.5), (19.7), (20.7), (20.8) are all time-independent.

Conclusion 21.1. *Differentiation with respect to external parameters (like t in* (21.1)*) is a tensorial operation producing new tensors from existing ones.*

Exercise 21.1. *Give a more detailed explanation of why the time derivative* (21.1) *represents a tensor of type (r, s).*

Now let's consider the spacial derivative of tensor field $\mathbf{X}$, i. e. its derivative with respect to a spacial variable, e. g. with respect to x^1. Here we also can write

$$\frac{\partial X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}}{\partial x^1} = \lim_{h\to 0} \frac{X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(x^1+h, x^2, x^3) - X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(x^1, x^2, x^3)}{h}, \tag{21.2}$$

but in numerator of the fraction in the right hand side of (21.2) we get the difference of two tensors bound to different points of space: to the point P with coordinates x^1, x^2, x^3 and to the point P' with coordinates $x^1 + h$, x^2, x^3. To which point should be attributed the difference of two such tensors? This is not clear. Therefore, we should treat partial derivatives like (21.2) in a different way.

Let's choose some additional symbol, say it can be q, and consider the partial derivative of $X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}$ with respect to the spacial variable x^q:

$$Y^{i_1\ldots\,i_r}_{q\,j_1\ldots\,j_s} = \frac{\partial X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}}{\partial x^q}. \tag{21.3}$$

Partial derivatives (21.2), taken as a whole, form an $(r + s + 1)$-dimensional array with one extra dimension due to index q. We write it as a lower index in $Y^{i_1\ldots\,i_r}_{q\,j_1\ldots\,j_s}$ due to the following theorem concerning (21.3).

THEOREM 21.1. *For any tensor field* $\mathbf{X}$ *of type* (r,s) *partial derivatives* (21.3) *with respect to spacial variables* x^1, x^2, x^3 *in any Cartesian coordinate system represent another tensor field* $\mathbf{Y}$ *of the type* $(r,s+1)$.

Thus differentiation with respect to x^1, x^2, x^3 produces new tensors from already existing ones. For the sake of beauty and convenience this operation is denoted by the nabla sign: $\mathbf{Y} = \nabla\mathbf{X}$. In index form this looks like

$$Y^{i_1\ldots\,i_r}_{q\,j_1\ldots\,j_s} = \nabla_q X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}. \tag{21.4}$$

Simplifying the notations we also write:

$$\nabla_q = \frac{\partial}{\partial x^q}. \tag{21.5}$$

WARNING 21.1. *Theorem 21.1 and the equality* (21.5) *are valid only for Cartesian coordinate systems. In curvilinear coordinates things are different.*

EXERCISE 21.2. *Prove theorem 21.1. For this purpose consider another Cartesian coordinate system* $\tilde x^1$, $\tilde x^2$, $\tilde x^3$ *related to* x^1, x^2, x^3 *via* (20.7) *and* (20.8). *Then in the new coordinate system consider the partial derivatives*

$$\tilde Y^{i_1\ldots\,i_r}_{q\,j_1\ldots\,j_s} = \frac{\partial \tilde X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}}{\partial \tilde x^q} \tag{21.6}$$

and derive relationships binding (21.6) *and* (21.3).

§ 22. Gradient, divergency, and rotor. Laplace and d'Alambert operators.

The tensorial nature of partial derivatives established by theorem 21.1 is a very useful feature. We can apply it to extend the scope of classical operations of vector analysis. Let's consider **the gradient**, $\operatorname{grad} = \nabla$. Usually the gradient operator is applied to scalar fields, i. e. to functions $\varphi = \varphi(P)$ or $\varphi = \varphi(x^1, x^2, x^3)$ in coordinate form:

$$a_q = \nabla_q\varphi = \frac{\partial\varphi}{\partial x^q}. \tag{22.1}$$

Note that in (22.1) we used a lower index q for a_q. This means that $\mathbf{a} = \operatorname{grad}\varphi$ is a covector. Indeed, according to theorem 21.1, the nabla operator applied to a scalar field, which is tensor field of type $(0,0)$, produces a tensor field of type $(0,1)$. In order to get the vector form of the gradient one should raise index q:

$$a^q = \sum_{i=1}^{3} g^{qi}\,a_i = \sum_{i=1}^{3} g^{qi}\,\nabla_i\varphi. \tag{22.2}$$

Let's write (22.2) in the form of a differential operator (without applying to φ):

$$\nabla^q = \sum_{i=1}^{3} g^{qi}\,\nabla_i. \tag{22.3}$$

In this form the gradient operator (22.3) can be applied not only to scalar fields, but also to vector fields, covector fields and to any other tensor fields.

Usually in physics we do not distinguish between the vectorial gradient ∇^q and the covectorial gradient ∇_q because we use orthonormal coordinates with ONB as a basis. In this case dual metric tensor is given by unit matrix ($g^{ij} = \delta^{ij}$) and components of ∇^q and ∇_q coincide by value.

Divergency is the second differential operation of vector analysis. Usually it is applied to a vector field and is given by formula:

$$\operatorname{div}\mathbf{X} = \sum_{i=1}^{3} \nabla_i X^i. \tag{22.4}$$

As we see, (22.4) is produced by contraction (see section 16) from tensor $\nabla_q X^i$. Therefore we can generalize formula (22.4) and apply divergency operator to arbitrary tensor field with at least one upper index:

$$(\operatorname{div}\mathbf{X})^{\dots\,\dots\,\dots}_{\dots\,\dots\,\dots} = \sum_{s=1}^{3} \nabla_s X^{\dots\,.s.\,\dots}_{\dots\,\dots\,\dots}. \tag{22.5}$$

The Laplace operator is defined as the divergency applied to a vectorial gradient of something, it is denoted by the triangle sign: $\triangle = \operatorname{div}\operatorname{grad}$. From (22.3) and (22.5) for Laplace operator $\triangle$ we derive the following formula:

$$\triangle = \sum_{i=1}^{3}\sum_{j=1}^{3} g^{ij}\,\nabla_i\,\nabla_j. \tag{22.6}$$

Denote by $\square$ the following differential operator:

$$\square = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle. \tag{22.7}$$

Operator (22.7) is called **the d'Alambert operator** or **wave operator**. In general relativity upon introducing the additional coordinate $x^0 = c\,t$ one usually rewrites the d'Alambert operator in a form quite similar to (22.6) (see my book [5], it is free for download from http://samizdat.mines.edu/).

And finally, let's consider **the rotor operator** or **curl operator** (the term "rotor" is derived from "rotation" so that "rotor" and "curl" have approximately the same meaning). The rotor operator is usually applied to a vector field and produces another vector field: $\mathbf{Y} = \operatorname{rot}\mathbf{X}$. Here is the formula for the r-th coordinate of $\operatorname{rot}\mathbf{X}$:

$$(\operatorname{rot}\mathbf{X})^r = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} g^{ri}\,\omega_{ijk}\,\nabla^j X^k. \tag{22.8}$$

The volume tensor $\boldsymbol{\omega}$ in (22.8) is given by formula (18.4), while the vectorial gradient operator ∇^j is defined in (22.3).

EXERCISE 22.1. *Formula* (22.8) *can be generalized for the case when* $\mathbf{X}$ *is an arbitrary tensor field with at least one upper index. By analogy with* (22.5) *suggest your version of such a generalization.*

Note that formulas (22.6) and (22.8) for the Laplace operator and for the rotor are different from those that are commonly used. Here are standard formulas:

$$\triangle = \left(\frac{\partial}{\partial x^1}\right)^2 + \left(\frac{\partial}{\partial x^2}\right)^2 + \left(\frac{\partial}{\partial x^3}\right)^2, \tag{22.9}$$

$$\operatorname{rot}\mathbf{X} = \det\begin{Vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \dfrac{\partial}{\partial x^1} & \dfrac{\partial}{\partial x^2} & \dfrac{\partial}{\partial x^3} \\ X^1 & X^2 & X^3 \end{Vmatrix}. \tag{22.10}$$

The truth is that formulas (22.6) and (22.8) are written for a general skew-angular coordinate system with a SAB as a basis. The standard formulas (22.10) are valid only for orthonormal coordinates with ONB as a basis.

EXERCISE 22.2. *Show that in case of orthonormal coordinates, when* $g^{ij} = \delta^{ij}$, *formula* (22.6) *for the Laplace operator* $\triangle$ *reduces to the standard formula* (22.9).

The coordinates of the vector $\operatorname{rot}\mathbf{X}$ in a skew-angular coordinate system are given by formula (22.8). Then for vector $\operatorname{rot}\mathbf{X}$ itself we have the expansion:

$$\operatorname{rot}\mathbf{X} = \sum_{r=1}^{3} (\operatorname{rot}\mathbf{X})^r\,\mathbf{e}_r. \tag{22.11}$$

EXERCISE 22.3. *Substitute* (22.8) *into* (22.11) *and show that in the case of a orthonormal coordinate system the resulting formula* (22.11) *reduces to* (22.10).

CHAPTER IV

TENSOR FIELDS IN CURVILINEAR COORDINATES.

§ 23. General idea of curvilinear coordinates.

What are coordinates, if we forget for a moment about radius-vectors, bases and axes? What is the pure idea of coordinates? The pure idea is in representing points of space by triples of numbers. This means that we should have one to one map $P \leftrightarrows (y^1, y^2, y^3)$ in the whole space or at least in some domain, where we are going to use our coordinates y^1, y^2, y^3. In Cartesian coordinates this map $P \leftrightarrows (y^1, y^2, y^3)$ is constructed by means of vectors and bases. Arranging other coordinate systems one can use other methods. For example, in **spherical coordinates** $y^1 = r$ is a distance from the point P to the center of sphere, $y^2 = \theta$ and $y^3 = \varphi$ are two angles. By the way, spherical coordinates are one of the simplest examples of curvilinear coordinates. Furthermore, let's keep in mind spherical coordinates when thinking about more general and hence more abstract curvilinear coordinate systems.

§ 24. Auxiliary Cartesian coordinate system.

Now we know almost everything about Cartesian coordinates and almost nothing about the abstract curvilinear coordinate system y^1, y^2, y^3 that we are going to study. Therefore, the best idea is to represent each point P by its radius-vector $\mathbf{r}_P$ in some auxiliary Cartesian coordinate system and then consider a map $\mathbf{r}_P \leftrightarrows (y^1, y^2, y^3)$. The radius-vector itself is represented by three coordinates in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ of the auxiliary Cartesian coordinate system:

$$\mathbf{r}_P = \sum_{i=1}^{3} x^i\, \mathbf{e}_i. \tag{24.1}$$

Therefore, we have a one-to-one map $(x^1, x^2, x^3) \leftrightarrows (y^1, y^2, y^3)$. Hurrah! This is a numeric map. We can treat it numerically. In the left direction it is represented by three functions of three variables:

$$\begin{cases} x^1 = x^1(y^1, y^2, y^3),\\ x^2 = x^2(y^1, y^2, y^3),\\ x^3 = x^3(y^1, y^2, y^3). \end{cases} \tag{24.2}$$

In the right direction we again have three functions of three variables:

$$\begin{cases} y^1 = y^1(x^1, x^2, x^3),\\ y^2 = y^2(x^1, x^2, x^3),\\ y^3 = y^3(x^1, x^2, x^3). \end{cases} \tag{24.3}$$

Further we shall assume all functions in (24.2) and (24.3) to be differentiable and consider their partial derivatives. Let's denote

$$S^i_j = \frac{\partial x^i}{\partial y^j}, \qquad\qquad T^i_j = \frac{\partial y^i}{\partial x^j}. \tag{24.4}$$

Partial derivatives (24.4) can be arranged into two square matrices S and T respectively. In mathematics such matrices are called Jacobi matrices. The components of matrix S in that form, as they are defined in (24.4), are functions of y^1, y^2, y^3. The components of T are functions of x^1, x^2, x^3:

$$S^i_j = S^i_j(y^1, y^2, y^3), \qquad\qquad T^i_j = T^i_j(x^1, x^2, x^3). \tag{24.5}$$

However, by substituting (24.3) into the arguments of S^i_j, or by substituting (24.2) into the arguments of T^i_j, we can make them have a common set of arguments:

$$S^i_j = S^i_j(x^1, x^2, x^3), \qquad\qquad T^i_j = T^i_j(x^1, x^2, x^3), \tag{24.6}$$

$$S^i_j = S^i_j(y^1, y^2, y^3), \qquad\qquad T^i_j = T^i_j(y^1, y^2, y^3), \tag{24.7}$$

When brought to the form (24.6), or when brought to the form (24.7) (but not in form of (24.5)), matrices S and T are inverse of each other:

$$T = S^{-1}. \tag{24.8}$$

This relationship (24.8) is due to the fact that numeric maps (24.2) and (24.3) are inverse of each other.

Exercise 24.1. *You certainly know the following formula:*

$$\frac{df(x^1(y), x^2(y), x^3(y))}{dy} = \sum_{i=1}^{3} f'_i(x^1(y), x^2(y), x^3(y))\,\frac{dx^i(y)}{dy}, \quad \text{where} \quad f'_i = \frac{\partial f}{\partial x^i}.$$

It's for the differentiation of composite function. Apply this formula to functions (24.2) *and derive the relationship* (24.8).

§ 25. Coordinate lines and the coordinate grid.

Let's substitute (24.2) into (24.1) and take into account that (24.2) now assumed to contain differentiable functions. Then the vector-function

$$\mathbf{R}(y^1, y^2, y^3) = \mathbf{r}_P = \sum_{i=1}^{3} x^i(y^1, y^2, y^3)\,\mathbf{e}_i \tag{25.1}$$

is a differentiable function of three variables y^1, y^2, y^3. The vector-function $\mathbf{R}(y^1, y^2, y^3)$ determined by (25.1) is called **a basic vector-function** of a curvilinear coordinate system. Let P_0 be some fixed point of space given by its curvilinear coordinates y^1_0, y^2_0, y^3_0. Here zero is not the tensorial index, we use it in order to emphasize that P_0 is fixed point, and that its coordinates y^1_0, y^2_0, y^3_0

are three fixed numbers. In the next step let's undo one of them, say first one, by setting

$$y^1 = y_0^1 + t, \qquad y^2 = y_0^2, \qquad y^3 = y_0^3. \tag{25.2}$$

Substituting (25.2) into (25.1) we get a vector-function of one variable t:

$$\mathbf{R}_1(t) = \mathbf{R}(y_0^1 + t, y_0^2, y_0^3), \tag{25.3}$$

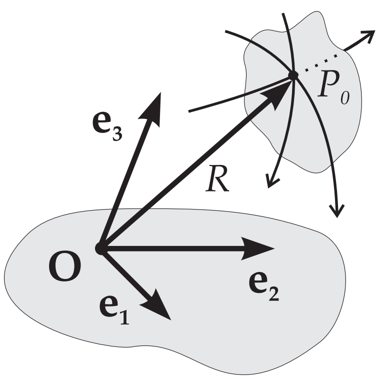


Fig. 8.

If we treat t as time variable (though it may have a unit other than time), then (25.3) describes a curve (the trajectory of a particle). At time instant $t = 0$ this curve passes through the fixed point P_0. Same is true for curves given by two other vector-functions similar to (25.4):

$$\mathbf{R}_2(t) = \mathbf{R}(y_0^1,\, y_0^2 + t,\, y_0^3), \tag{25.4}$$

$$\mathbf{R}_3(t) = \mathbf{R}(y_0^1,\, y_0^2,\, y_0^3 + t). \tag{25.5}$$

This means that all three curves given by vector-functions (25.3), (25.4), and (25.5) are intersected at the point P_0 as shown on Fig. 8. Arrowheads on these lines indicate the directions in which parameter t increases. Curves (25.3), (25.4), and (25.5) are called **coordinate lines**. They are subdivided into three families. Curves within one family do not intersect each other. Curves from different families intersect so that any regular point of space is an intersection of exactly three coordinate curves (one per family).

Coordinate lines taken in whole form **a coordinate grid**. This is an infinitely dense grid. But usually, when drawing, it is represented as a grid with finite density. On Fig. 9 the coordinate grid of curvilinear coordinates is compared to that of the Cartesian coordinate system.

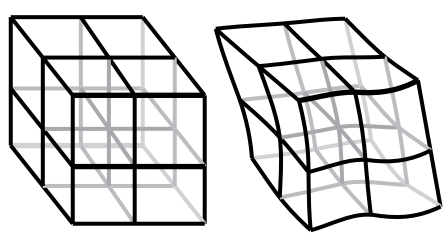


Fig. 9.

Another example of coordinate grid is on Fig. 2. Indeed, meridians and parallels are coordinate lines of a spherical coordinate system. The parallels do not intersect, but the meridians forming one family of coordinate lines do intersect at the North and at South Poles. This means that North and South Poles are singular points for spherical coordinates.

Exercise 25.1. *Remember the exact definition of spherical coordinates and find all singular points for them.*

§ 26. Moving frame of curvilinear coordinates.

Let's consider the three coordinate lines shown on Fig. 8 again. And let's find tangent vectors to them at the point P_0. For this purpose we should differentiate vector-functions (25.3), (25.4), and (25.5) with respect to the time variable t and then substitute $t = 0$ into the derivatives:

$$\mathbf{E}_i = \left.\frac{d\mathbf{R}_i}{dt}\right|_{t=0} = \left.\frac{\partial\mathbf{R}}{\partial y^i}\right|_{\text{at the point } P_0}. \tag{26.1}$$

Now let's substitute the expansion (25.1) into (26.1) and remember (24.4):

$$\mathbf{E}_i = \frac{\partial\mathbf{R}}{\partial y^i} = \sum_{j=1}^{3} \frac{\partial x^j}{\partial y^i}\,\mathbf{e}_j = \sum_{j=1}^{3} S^j_i\,\mathbf{e}_j. \tag{26.2}$$

All calculations in (26.2) are still in reference to the point P_0. Though P_0 is a fixed point, it is an arbitrary fixed point. Therefore, the equality (26.2) is valid at any point. Now let's omit the intermediate calculations and write (26.2) as

$$\mathbf{E}_i = \sum_{i=1}^{3} S^j_i\,\mathbf{e}_j. \tag{26.3}$$

And then compare (26.3) with (5.7). They are strikingly similar, and $\det S \neq 0$ due to (24.8). Formula (26.3) means that tangent vectors to coordinate lines $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ form a basis (see Fig. 10), matrices (24.4) are transition matrices to this basis and back to the Cartesian basis.

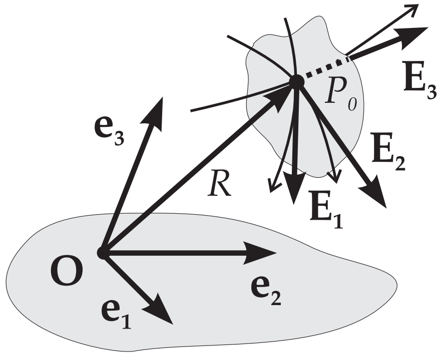


Fig. 10.

Despite obvious similarity of the formulas (26.3) and (5.7), there is some crucial difference of basis $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ as compared to $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Vectors $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ are not free. They are bound to that point where derivatives (24.4) are calculated. And they move when we move this point. For this reason basis $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ is called **moving frame** of the curvilinear coordinate system. During their motion the vectors of the moving frame $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ are not simply translated from point to point, they can change their lengths and the angles they form with each other. Therefore, in general the moving frame $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ is a skew-angular basis. In some cases vectors $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ can be orthogonal to each other at all points of space. In that case we say that we have an orthogonal curvilinear coordinate system. Most of the well known curvilinear coordinate systems are orthogonal, e. g. spherical, cylindrical, elliptic, parabolic, toroidal, and others. However, there is no curvilinear coordinate system with the moving frame being ONB! We shall not prove this fact since it leads deep into differential geometry.

§ 27. Dynamics of moving frame.

Thus, we know that **the moving frame** moves. Let's describe this motion quantitatively. According to (24.5) the components of matrix S in (26.3) are functions of the curvilinear coordinates y^1, y^2, y^3. Therefore, differentiating $\mathbf{E}_i$ with respect to y^j we should expect to get some nonzero vector $\partial\mathbf{E}_i/\partial y^j$. This vector can be expanded back in moving frame $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$. This expansion is written as

$$\frac{\partial\mathbf{E}_i}{\partial y^j}=\sum_{k=1}^{3}\Gamma^k_{ij}\,\mathbf{E}_k. \tag{27.1}$$

Formula (27.1) is known as **the derivational formula**. Coefficients Γ^k_{ij} in (27.1) are called **Christoffel symbols** or **connection components**.

Exercise 27.1. *Relying upon formula* (25.1) *and* (26.1) *draw the vectors of the moving frame for cylindrical coordinates.*

Exercise 27.2. *Do the same for spherical coordinates.*

Exercise 27.3. *Relying upon formula* (27.1) *and results of exercise 27.1 calculate the Christoffel symbols for cylindrical coordinates.*

Exercise 27.4. *Do the same for spherical coordinates.*

Exercise 27.5. *Remember formula* (26.2) *from which you derive*

$$\mathbf{E}_i=\frac{\partial\mathbf{R}}{\partial y^i}. \tag{27.2}$$

Substitute (27.2) *into left hand side of the derivational formula* (27.1) *and relying on the properties of mixed derivatives prove that the Christoffel symbols are symmetric with respect to their lower indices:* $\Gamma^k_{ij}=\Gamma^k_{ji}$.

Note that Christoffel symbols Γ^k_{ij} form a three-dimensional array with one upper index and two lower indices. However, they do not represent a tensor. We shall not prove this fact since it again leads deep into differential geometry.

§ 28. Formula for Christoffel symbols.

Let's take formula (26.3) and substitute it into both sides of (27.1). As a result we get the following equality for Christoffel symbols Γ^k_{ij}:

$$\sum_{q=1}^{3}\frac{\partial S^q_i}{\partial y^j}\,\mathbf{e}_q=\sum_{k=1}^{3}\sum_{q=1}^{3}\Gamma^k_{ij}\,S^q_k\,\mathbf{e}_q. \tag{28.1}$$

Cartesian basis vectors $\mathbf{e}_q$ do not depend on y^j; therefore, they are not differentiated when we substitute (26.3) into (27.1). Both sides of (28.1) are expansions in the base $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ of the auxiliary Cartesian coordinate system. Due to the uniqueness of such expansions we have the following equality derived from (28.1):

$$\frac{\partial S^q_i}{\partial y^j}=\sum_{k=1}^{3}\Gamma^k_{ij}\,S^q_k. \tag{28.2}$$

Exercise 28.1. *Using concept of the inverse matrix ($T = S^{-1}$) derive the following formula for the Christoffel symbols Γ^k_{ij} from* (28.2)*:*

$$\Gamma^k_{ij} = \sum_{q=1}^{3} T^k_q \, \frac{\partial S^q_i}{\partial y^j}. \tag{28.3}$$

Due to (24.4) this formula (28.3) can be transformed in the following way:

$$\Gamma^k_{ij} = \sum_{q=1}^{3} T^k_q \, \frac{\partial S^q_i}{\partial y^j} = \sum_{q=1}^{3} T^k_q \, \frac{\partial^2 x^q}{\partial y^i \, \partial y^j} = \sum_{q=1}^{3} T^k_q \, \frac{\partial S^q_j}{\partial y^i}. \tag{28.4}$$

Formulas (28.4) are of no practical use because they express Γ^k_{ij} through an external thing like transition matrices to and from the auxiliary Cartesian coordinate system. However, they will help us below in understanding the differentiation of tensors.

§ 29. Tensor fields in curvilinear coordinates.

As we remember, tensors are geometric objects related to bases and represented by arrays if some basis is specified. Each curvilinear coordinate system provides us a numeric representation for points, and in addition to this it provides the basis. This is the moving frame. Therefore, we can refer tensorial objects to curvilinear coordinate systems, where they are represented as arrays of functions:

$$X^{i_1 \ldots i_r}_{j_1 \ldots j_s} = X^{i_1 \ldots i_r}_{j_1 \ldots j_s}(y^1, y^2, y^3). \tag{29.1}$$

We also can have two curvilinear coordinate systems and can pass from one to another by means of transition functions:

$$\begin{cases} \tilde{y}^1 = \tilde{y}^1(y^1, y^2, y^3), \\ \tilde{y}^2 = \tilde{y}^2(y^1, y^2, y^3), \\ \tilde{y}^3 = \tilde{y}^3(y^1, y^2, y^3), \end{cases} \qquad \begin{cases} y^1 = y^1(\tilde{y}^1, \tilde{y}^2, \tilde{y}^3), \\ y^2 = y^2(\tilde{y}^1, \tilde{y}^2, \tilde{y}^3), \\ y^3 = y^3(\tilde{y}^1, \tilde{y}^2, \tilde{y}^3). \end{cases} \tag{29.2}$$

If we call $\tilde{y}^1$, $\tilde{y}^2$, $\tilde{y}^3$ the new coordinates, and y^1, y^2, y^3 the old coordinates, then transition matrices S and T are given by the following formulas:

$$S^i_j = \frac{\partial y^i}{\partial \tilde{y}^j}, \qquad T^i_j = \frac{\partial \tilde{y}^i}{\partial y^j}. \tag{29.3}$$

They relate moving frames of two curvilinear coordinate systems:

$$\tilde{\mathbf{E}}_i = \sum_{j=1}^{3} S^j_i \, \mathbf{E}_j, \qquad \mathbf{E}_j = \sum_{i=1}^{3} T^i_j \, \tilde{\mathbf{E}}_i. \tag{29.4}$$

EXERCISE 29.1. *Derive* (29.3) *from* (29.4) *and* (29.2) *using some auxiliary Cartesian coordinates with basis* $\mathbf{e}_1,\ \mathbf{e}_2, \mathbf{e}_3$ *as intermediate coordinate system:*

$$(\mathbf{E}_1,\,\mathbf{E}_2,\,\mathbf{E}_3)\ \underset{T}{\overset{S}{\rightleftarrows}}\ (\mathbf{e}_1,\,\mathbf{e}_2,\,\mathbf{e}_3)\ \underset{\tilde T}{\overset{\tilde S}{\rightleftarrows}}\ (\tilde{\mathbf{E}}_1,\,\tilde{\mathbf{E}}_2,\,\tilde{\mathbf{E}}_3) \tag{29.5}$$

Compare (29.5) *with* (5.13) *and explain differences you have detected.*

Transformation formulas for tensor fields for two curvilinear coordinate systems are the same as in (19.4) and (19.5):

$$\tilde X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(\tilde y^1,\tilde y^2,\tilde y^3)=\sum^3_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\!\sum^3 T^{i_1}_{h_1}\ldots\,T^{i_r}_{h_r}S^{k_1}_{j_1}\,\ldots\,S^{k_s}_{j_s}\,X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(y^1,y^2,y^3), \tag{29.6}$$

$$X^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(y_1,y_2,y_3)=\sum^3_{\substack{h_1,\,\ldots\,,\,h_r\\ k_1,\,\ldots\,,\,k_s}}\!\!\ldots\!\sum^3 S^{i_1}_{h_1}\ldots\,S^{i_r}_{h_r}T^{k_1}_{j_1}\,\ldots\,T^{k_s}_{j_s}\,\tilde X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(\tilde y^1,\tilde y^2,\tilde y^3). \tag{29.7}$$

But formulas (19.6) and (19.8) should be replaced by (29.2).

§ 30. Differentiation of tensor fields in curvilinear coordinates.

We already know how to differentiate tensor fields in Cartesian coordinates (see section 21). We know that operator ∇ produces tensor field of type $(r,s+1)$ when applied to a tensor field of type (r,s). The only thing we need now is to transform ∇ to a curvilinear coordinate system. In order to calculate tensor $\nabla\mathbf{X}$ in curvilinear coordinates, let's first transform $\mathbf{X}$ into auxiliary Cartesian coordinates, then apply ∇, and then transform $\nabla\mathbf{X}$ back into curvilinear coordinates:

$$\begin{array}{ccc}
X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(y^1,y^2,y^3) & \xrightarrow{S,T} & X^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(x^1,x^2,x^3)\\
\Big\downarrow{\scriptstyle\nabla_p} & & \Big\downarrow{\scriptstyle\nabla_q=\partial/\partial x^q}\\
\nabla_pX^{i_1\ldots\,i_r}_{j_1\ldots\,j_s}(y^1,y^2,y^3) & \xleftarrow{T,S} & \nabla_qX^{h_1\ldots\,h_r}_{k_1\ldots\,k_s}(x^1,x^2,x^3)
\end{array} \tag{30.1}$$

Matrices (24.4) are used in (30.1). From (12.3) and (12.4) we know that the transformation of each index is a separate multiplicative procedure. When applied to the α-th upper index, the whole chain of transformations (30.1) looks like

$$\nabla_pX^{\ldots\,i_\alpha\,\ldots}_{\ldots\,\ldots\,\ldots}=\sum^3_{q=1}S^q_p\ldots\sum^3_{h_\alpha=1}T^{i_\alpha}_{h_\alpha}\,\ldots\,\nabla_q\,\ldots\sum^3_{m_\alpha=1}S^{h_\alpha}_{m_\alpha}\ldots\,X^{\ldots\,m_\alpha\,\ldots}_{\ldots\,\ldots\,\ldots}. \tag{30.2}$$

Note that $\nabla_q=\partial/\partial x^q$ is a differential operator and due to (24.4) we have

$$\sum^3_{q=1}S^q_p\,\frac{\partial}{\partial x^q}=\frac{\partial}{\partial y^p}. \tag{30.3}$$

Any differential operator when applied to a product produces a sum with as many summands as there were multiplicand in the product. Here is the summand produced by term $S^{h_\alpha}_{m_\alpha}$ in formula (30.2):

$$\nabla_p X^{\dots\, i_\alpha\, \dots}_{\dots\,\dots\,\dots} = \ldots + \sum_{m_\alpha=1}^{3}\sum_{h_\alpha=1}^{3} T^{i_\alpha}_{h_\alpha}\,\frac{S^{h_\alpha}_{m_\alpha}}{\partial y^p}\,X^{\dots\, m_\alpha\, \dots}_{\dots\,\dots\,\dots} + \ldots . \tag{30.4}$$

Comparing (30.4) with (28.3) or (28.4) we can transform it into the following equality:

$$\nabla_p X^{\dots\, i_\alpha\, \dots}_{\dots\,\dots\,\dots} = \ldots + \sum_{m_\alpha=1}^{3} \Gamma^{i_\alpha}_{p m_\alpha}\,X^{\dots\, m_\alpha\, \dots}_{\dots\,\dots\,\dots} + \ldots . \tag{30.5}$$

Now let's consider the transformation of the α-th lower index in (30.1):

$$\nabla_p X^{\dots\,\dots\,\dots}_{\dots\, j_\alpha\, \dots} = \sum_{q=1}^{3} S^q_p \ldots \sum_{k_\alpha=1}^{3} S^{k_\alpha}_{j_\alpha}\ \ldots\ \nabla_q\ \ldots \sum_{n_\alpha=1}^{3} T^{n_\alpha}_{k_\alpha} \ldots\ X^{\dots\,\dots\,\dots}_{\dots\, n_\alpha\, \dots}. \tag{30.6}$$

Applying (30.3) to (30.6) with the same logic as in deriving (30.4) we get

$$\nabla_p X^{\dots\,\dots\,\dots}_{\dots\, j_\alpha\, \dots} = \ldots + \sum_{n_\alpha=1}^{3}\sum_{k_\alpha=1}^{3} S^{k_\alpha}_{j_\alpha}\,\frac{T^{n_\alpha}_{k_\alpha}}{\partial y^p}\,X^{\dots\,\dots\,\dots}_{\dots\, n_\alpha\, \dots} + \ldots . \tag{30.7}$$

In order to simplify (30.7) we need the following formula derived from (28.3):

$$\Gamma^k_{ij} = -\sum_{q=1}^{3} S^q_i\,\frac{\partial T^k_q}{\partial y^j}. \tag{30.8}$$

Applying (30.8) to (30.7) we obtain

$$\nabla_p X^{\dots\,\dots\,\dots}_{\dots\, j_\alpha\, \dots} = \ldots - \sum_{n_\alpha=1}^{3} \Gamma^{n_\alpha}_{p j_\alpha}\,X^{\dots\,\dots\,\dots}_{\dots\, n_\alpha\, \dots} + \ldots . \tag{30.9}$$

Now we should gather (30.5), (30.9), and add the term produced when ∇_q in (30.2) (or equivalently in (30.4)) acts upon components of tensor $\mathbf{X}$. As a result we get the following general formula for $\nabla_p X^{i_1\dots\, i_r}_{j_1\dots\, j_s}$:

$$\begin{aligned}\nabla_p X^{i_1\dots\, i_r}_{j_1\dots\, j_s} = \frac{\partial X^{i_1\dots\, i_r}_{j_1\dots\, j_s}}{\partial y^p} &+ \sum_{\alpha=1}^{r}\sum_{m_\alpha=1}^{3} \Gamma^{i_\alpha}_{p m_\alpha}\,X^{i_1\dots\, m_\alpha\, \dots\, i_r}_{j_1\dots\,\dots\,\dots\, j_s} - \\ &- \sum_{\alpha=1}^{s}\sum_{n_\alpha=1}^{3} \Gamma^{n_\alpha}_{p j_\alpha}\,X^{i_1\dots\,\dots\,\dots\, i_r}_{j_1\dots\, n_\alpha\, \dots\, j_s}.\end{aligned} \tag{30.10}$$

The operator ∇_p determined by this formula is called **the covariant derivative**.

EXERCISE 30.1. *Apply the general formula* (30.10) *to a vector field and calculate the covariant derivative* $\nabla_p X^q$.

EXERCISE 30.2. *Apply the general formula* (30.10) *to a covector field and calculate the covariant derivative* $\nabla_p X_q$.

EXERCISE 30.3. *Apply the general formula* (30.10) *to an operator field and find* $\nabla_p F^q_m$. *Consider special case when* ∇_p *is applied to the Kronecker symbol* δ^q_m.

EXERCISE 30.4. *Apply the general formula* (30.10) *to a bilinear form and find* $\nabla_p a_{qm}$.

EXERCISE 30.5. *Apply the general formula* (30.10) *to a tensor product* $\mathbf{a} \otimes \mathbf{x}$ *for the case when* $\mathbf{x}$ *is a vector and* $\mathbf{a}$ *is a covector. Verify formula* $\nabla(\mathbf{a} \otimes \mathbf{x}) = \nabla\mathbf{a} \otimes \mathbf{x} + \mathbf{a} \otimes \nabla\mathbf{x}$.

EXERCISE 30.6. *Apply the general formula* (30.10) *to the contraction* $C(\mathbf{F})$ *for the case when* $\mathbf{F}$ *is an operator field. Verify the formula* $\nabla C(\mathbf{F}) = C(\nabla\mathbf{F})$.

EXERCISE 30.7. *Derive* (30.8) *from* (28.3).

§ 31. Concordance of metric and connection.

Let's remember that we consider curvilinear coordinates in Euclidean space E. In this space we have the scalar product (13.1) and the metric tensor (13.5).

EXERCISE 31.1. *Transform the metric tensor* (13.5) *to curvilinear coordinates using transition matrices* (24.4) *and show that here it is given by formula*

$$g_{ij} = (\mathbf{E}_i,\, \mathbf{E}_j). \tag{31.1}$$

In Cartesian coordinates all components of the metric tensor are constant since the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are constant. The covariant derivative (30.10) in Cartesian coordinates reduces to differentiation $\nabla_p = \partial/\partial x^p$. Therefore,

$$\nabla_p g_{ij} = 0. \tag{31.2}$$

But ∇g is a tensor. If all of its components in some coordinate system are zero, then they are identically zero in any other coordinate system (explain why). Therefore the identity (31.2) is valid in curvilinear coordinates as well.

EXERCISE 31.2. *Prove* (31.2) *by direct calculations using formula* (27.1).

The identity (31.2) is known as **the concordance condition** for the metric g_{ij} and connection Γ^k_{ij}. It is very important for general relativity.

Remember that the metric tensor enters into many useful formulas for the gradient, divergency, rotor, and Laplace operator in section 22. What is important is that all of these formulas remain valid in curvilinear coordinates, with the only difference being that you should understand that ∇_p is not the partial derivative $\partial/\partial x^p$, but the covariant derivative in the sense of formula (30.10).

EXERCISE 31.3. *Calculate* $\operatorname{rot}\mathbf{A}$, $\operatorname{div}\mathbf{H}$, $\operatorname{grad}\varphi$ *(vectorial gradient) in cylindrical and spherical coordinates.*

EXERCISE 31.4. *Calculate the Laplace operator* $\triangle\varphi$ *applied to the scalar field* φ *in cylindrical and in spherical coordinates.*

REFERENCES.

1. Hefferon J., *Linear algebra*, Electronic textbook, free for downloading from Web site of Saint Michael's College, Colchester, VM 05439, USA; Download [PDF] or [PS] file.
2. Lehnen A. P., *An elementary introduction to logic and set theory*, On-line resource, Madison Area Technical College, Madison, WI 53704, USA.
3. Konstantopoulos T., *Basic background for a course of information and cryptography*, On-line materials, Feb. 2000, Electrical & Computer Engineering Dept., University of Texas at Austin, Austin, Texas 78712, USA.
4. Vaughan-Lee M., *B2 ring theory preliminaries*, On-line lecture materials, Sept. 2000, University of Oxford, Math. Institute, Oxford, OX1 3LB, UK.
5. Sharipov R. A., *Classical electrodynamics and theory of relativity*, Bashkir State University, Ufa, Russia, 1997; English tr., 2003, physics/0311011 in Electronic Archive http://arXiv.org.