Lecture Notes on Vector and Tensor Algebra and Analysis

Ilya L. Shapiro

Departamento de Física – Instituto Ciências Exatas
Universidade Federal de Juiz de Fora,
Juiz de Fora, CEP 36036-330, MG, Brazil

Preface

These lecture notes are the result of teaching a half-semester course of tensors for undergraduates in the Department of Physics at the Federal University of Juiz de Fora. The same lectures were also given at the summer school in the Institute of Mathematics in the University of Brasilia, where I was kindly invited by Dra. Maria Emília Guimarães and Dr. Guy Grebot. Furthermore, I have used the first version of these notes to teach students of "scientific initiation" in Juiz de Fora. Usually, in the case of independent study, good students learn the material of the lectures in one semester.

Since the lectures have some original didactic elements, we decided to publish them. These lectures are designed for the second-third year undergraduate student and are supposed to help in learning such disciplines of the course of Physics as Classical Mechanics, Electrodynamics, Special and General Relativity. One of my purposes was, e.g., to make derivation of *grad*, *div* and *rot* in the curvilinear coordinates understandable for the student, and this seems to be useful for some of the students of Physics, Mathematics or Engineering. Of course, those students which are going to make career in Mathematics or Theoretical Physics, may and should continue their education using serious books on Differential Geometry like [1]. These notes are nothing but a simple introduction for beginners. As examples of similar books we can indicate [2, 3] and [4], but our treatment of many issues is much more simple. A more sophisticated and modern, but still relatively simple introduction to tensors may be found in [5]. Some books on General Relativity have excellent introduction to tensors, let us just mention famous example [6] and [7]. Some problems included into these notes were taken from the textbooks and collection of problems [8, 9, 10] cited in the Bibliography. It might happen that some problems belong to the books which were not cited there, author wants apologize for this occurrence.

In the preparation of these notes I have used, as a starting point, the short course of tensors given in 1977 at Tomsk State University (Russia) by Dr. Veniamin Alexeevich Kuchin, who died soon after that. In part, these notes may be viewed as a natural extension of what he taught us at that time.

The preparation of the manuscript would be impossible without an important organizational work of Dr. Flavio Takakura and his generous help in preparing the Figures. I am especially grateful to the following students of our Department: to Raphael Furtado Coelho for typing the first draft and to Flavia Sobreira and Leandro de Castro Guarnieri, who saved these notes from many typing mistakes.

The present version of the notes is published due to the kind interest of Prof. José Abdalla Helayël-Neto. We hope that this publication will be useful for some students. On the other hand, I would be very grateful for any observations and recommendations. The correspondence may be send to the electronic address *shapiro@fisica.ufjf.br* or by mail to the following address:

Ilya L. Shapiro
Departamento de Física, ICE, Universidade Federal de Juiz de Fora
CEP: 36036-330, Juiz de Fora, MG, Brazil

Contents:

Preliminary observations and notations

i) It is supposed that the student is familiar with the backgrounds of Calculus, Analytic Geometry and Linear Algebra. Sometimes the corresponding information will be repeated in the text of the notes. We do not try to substitute the corresponding courses here, but only supplement them.

ii) In this notes we consider, by default, that the space has dimension 3. However, in some cases we shall refer to an arbitrary dimension of space D, and sometimes consider $D = 2$, because this is the simplest non-trivial case. The indication of dimension is performed in the form like $3D$, that means $D = 3$.

iii) Some objects with indices will be used below. Latin indices run the values

$$(a, b, c, ..., i, j, k, l, m, n, ...) = (1, 2, 3)$$

in $3D$ and

$$(a, b, c, ..., i, j, k, l, m, n, ...) = (1, 2, ..., D)$$

for an arbitrary D.

Usually, the indices $(a,\ b,\ c, ...)$ correspond to the orthonormal basis and to the Cartesian coordinates. The indices $(i,\ j,\ k, ...)$ correspond to the an arbitrary (generally non-degenerate) basis and to arbitrary, possibly curvilinear coordinates.

iv) Following standard practice, we denote the set of the elements f_i as $\{f_i\}$. The properties of the elements are indicated after the vertical line. For example,

$$E = \{\, e \,|\, e = 2n,\ n \in N \,\}$$

means the set of even natural numbers. The comment may follow after the comma. For example,

$$\{\, e \,|\, e = 2n,\ n \in N\,, n \leq 3 \,\} = \{\, 2, 4, 6 \,\}\,.$$

v) The repeated upper and lower indices imply summation (Einstein convention). For example,

$$a^i b_i \ = \sum_{i=1}^{D} a^i b_i \ = \ a^1 b_1 + a^2 b_2 + ... + a^D b_D$$

for the D-dimensional case. It is important that the summation (umbral) index i here can be renamed in an arbitrary way, e.g.

$$C^i_i = C^j_j = C^k_k = \ ...\,.$$

This is completely similar to the change of the notation for the variable of integration in a definite integral:

$$\int_a^b f(x)dx = \int_a^b f(y)dy\,.$$

where, also, the name of the variable does not matter.

vi) We use the notations like a'^i for the components of the vectors corresponding to the coordinates x'^i. In general, the same objects may be also denoted as $a^{i'}$ and $x^{i'}$. In the text we do not distinguish these two notations, and one has to assume, e.g., that $a'^i = a^{i'}$. The same holds for any tensor indices, of course.

vii) We use abbreviation:

$$\textbf{Def.} \equiv \textbf{Definition}$$

viii) The exercises are dispersed in the text, and the reader is advised to solve them in order of appearance. Some exercises are in fact simple theorems which will be consequently used, many of them are very important statements. Most of the exercises are very simple, and they suppose to consume small time. But, there are a few exercises marked by *. Those are presumably more difficult, therefore they are recommended only for the students which are going to specialize in Mathematics or Theoretical Physics.

Chapter 1

Vectors and Tensors

In this chapter we shall introduce and discuss the basic notions, part of them belong to the Analytic Geometry.

1.1 Vector basis and its transformation

Def. 1. Consider the set of vectors in $3D$:

$$\{\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_n}\} = \{\mathbf{a_i}, i = 1, n\}$$

Vector $\mathbf{k} = \sum_{i=1}^{n} k^i \mathbf{a_i}$ is called a *linear combination* of these vectors.

Def. 2. Vectors $\{\mathbf{a_i},\ i = 1, ..., n\}$ are called linearly independent if

$$\mathbf{k} = \sum_{i=1}^{n} k^i \mathbf{a_i} = k^i \mathbf{a_i} = 0 \quad \Longrightarrow \quad \sum_{i=1}^{n} (k^i)^2 = 0\,.$$

Comment: For example, in $2D$ two vectors are linearly independent if and only if they are parallel, in $3D$ two vectors are linearly independent if and only if they belong to the same plane etc.

Def. 3. If the vectors $\{\mathbf{a_i}\}$ are not linearly independent, they are called linearly dependent.

Exercise 1. Prove that vectors $\{\mathbf{0}, \mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}, ..., \mathbf{a_n}\}$ are always linearly dependent.

Exercise 2. Prove that $\mathbf{a_1} = \mathbf{i} + \mathbf{k}$, $\mathbf{a_2} = \mathbf{i} - \mathbf{k}$, $\mathbf{a_3} = \mathbf{i} - 3\mathbf{k}$ are linearly dependent for $\forall\, \{\mathbf{i}, \mathbf{k}\}$

Exercise 3. Prove that $\mathbf{a_1} = \mathbf{i} + \mathbf{k}$ and $\mathbf{a_2} = \mathbf{i} - \mathbf{k}$ are linearly independent if and only if $\mathbf{i}$, $\mathbf{k}$ are linearly independent.

Def. 4. Any three linearly independent vectors in 3D form a basis. Let us consider three such vectors be $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$. Then, for any vector $\mathbf{a}$ we can expand

$$\mathbf{a} = a^i \mathbf{e}_i\,. \tag{1.1}$$

The numbers a^i are called contravariant components of the vector $\mathbf{a}$. Sometimes we say, for brevity, that the numbers $\{a^i\}$ form contravariant vector.

Exercise 4. Prove that the expansion (1.1) is unique. That is, for any given basis and for any vector **a**, its contravariant components a^i are defined in a unique way.

Observation: It is important that the contravariant components of the vector a^i do change when we turn from one basis to another. Let us consider, along with the original basis $\mathbf{e}_i$, another basis $\mathbf{e}'_i$. Since each vector of the new basis belongs to the same space, it can be expanded using the original basis as

$$\mathbf{e}'_i = \wedge^j_{i'} \mathbf{e}_j \,. \tag{1.2}$$

Then, the uniqueness of the expansion of the vector **a** leads to the following relation:

$$\mathbf{a} = a^i \mathbf{e}_i = a^{j'} \mathbf{e}_{j'} = a^{j'} \wedge^i_{j'} \mathbf{e}_i \,,$$

and hence

$$a^i = a^{j'} \wedge^i_{j'} \,. \tag{1.3}$$

This is a very important relation, because it shows us how the contravariant components of the vector transform from one basis to another.

Similarly, we can make the transformation inverse to (1.2). It is easy to see that the relation between the contravariant components can be also written as

$$a^{k'} = (\wedge^{-1})^{k'}_l a^l \,,$$

where the matrix $(\wedge^{-1})$ is inverse to $\wedge$:

$$(\wedge^{-1})^{k'}_l \cdot \wedge^l_{i'} = \delta^{k'}_{i'} \qquad \text{and} \qquad \wedge^l_{i'} \cdot (\wedge^{-1})^{i'}_k = \delta^l_k$$

In the last formula we introduced what is called the **Kronecker symbol**.

$$\delta^i_l = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \,. \tag{1.4}$$

Exercise 5. Verify and discuss the relations between the formulas

$$\mathbf{e}_{i'} = \wedge^j_{i'} \mathbf{e}_j \,, \qquad \mathbf{e}_i = (\wedge^{-1})^{k'}_i \mathbf{e}_{k'}$$

and

$$a^{i'} = (\wedge^{-1})^{i'}_k a^k \,, \qquad a^j = \wedge^j_{i'} a^{i'} \,.$$

Observation: We can interpret the relations (1.3) in a different way. Suppose we have a set of three numbers which characterize some physical, geometrical or other quantity. We can identify whether these numbers are components of a vector, by looking at their transformation rule. They form vector if and only if they transform according to (1.3).

Now we are in a position to introduce the coordinates of a point in the $3D$ space. Let us remark that the generalization to the general case of an arbitrary D is straightforward.

Def. 5. Three vectors $\mathbf{e}_i$, together with some initial point 0 form a system of coordinates. For any point M we can introduce a vector.

$$\mathbf{r} = \mathbf{0M} = x^i \mathbf{e}_i \,, \tag{1.5}$$

which is called radius-vector of this point. The coefficients x^i of the expansion (1.5) are called the coordinates of the point M. Indeed, the coordinates are the contravariant components of the radius-vector $\mathbf{r}$.

Observation: Of course, the coordinates x^i of a point depend on the choice of the basis $\{\mathbf{e}_i\}$, and also on the choice of the initial point O.

Similarly to the Eq. (1.5), we can expand the same vector $\mathbf{r}$ using another basis

$$\mathbf{r} = x^i \mathbf{e}_i = x^{j'} \mathbf{e}_{j'} = x^{j'} . \wedge^i_{j'} \mathbf{e}_i$$

Then, according to (1.3)

$$x^i = x^{j'} \wedge^i_{j'} \ .$$

The last formula has an important consequence. Taking the partial derivatives, we arrive at the relations

$$\wedge^i_{j'} = \frac{\partial x^i}{\partial x^{j'}} \qquad \text{and} \qquad (\wedge^{-1})^{k'}_l = \frac{\partial x^{k'}}{\partial x^l} \,.$$

Now we can understand better the sense of the matrix $\wedge$. In particular, the relation

$$(\wedge^{-1})^{k'}_l \cdot \wedge^i_{k'} = \frac{\partial x^{k'}}{\partial x^l} \frac{\partial x^i}{\partial x^{k'}} = \frac{\partial x^i}{\partial x^l} = \delta^i_l$$

is nothing but the chain rule for the partial derivatives.

1.2 Scalar, vector and tensor fields

Def. 6. The function $\varphi(x) \equiv \varphi(x^i)$ is called scalar field or simply scalar if it does not transform under the change of coordinates.

$$\varphi(x) = \varphi'(x') \,. \tag{1.6}$$

Observation 1. Geometrically, this means that when we change the coordinates, the value of the function $\varphi(x^i)$ remains the same in a given geometric point.

Due to the great importance of this definition, let us give clarifying example in $1D$. Consider some function, e.g. $y = x^2$. The plot is parabola, as we all know. Now, let us change the variables $x' = x + 1$. The function $y = (x')^2$, obviously, represents another parabola. In order to preserve the plot intact, we need to modify the *form of the function*, that is to go from φ to φ'. Then, the new function $y = (x' - 1)^2$ will represent the original parabola. Two formulas $y = x^2$ and $y = (x' - 1)^2$

represent the same function, because the change of the variable is completely compensated by the change of the form of the function.

Observation 2. An important physical example of scalar is the temperature T in the given point of space at a given instant of time. Other examples are pressure p and the density of air ρ.

Exercise 6. Discuss whether the three numbers T, p, ρ form a contravariant vector or not.

Observation 3. ¿From **Def. 6** follows that for the generic change of coordinates and scalar field $\varphi'(x) \neq \varphi(x)$ and $\varphi(x') \neq \varphi(x)$. Let us calculate these quantities explicitly for the special case of infinitesimal transformation $x'^i = x^i + \xi^i$ where ξ are constant coefficients. We shall take info account only 1-st order in ξ^i. Then

$$\varphi(x') = \varphi(x) + \frac{\partial \varphi}{\partial x^i}\,\xi^i \tag{1.7}$$

and

$$\varphi'(x^i) = \varphi'(x'^i - \xi^i) = \varphi'(x') - \frac{\partial \varphi'}{\partial x'^i} \cdot \xi^i\,.$$

Let us take into account that the partial derivative in the last term can be evaluated for $\xi = 0$. Then

$$\frac{\partial \varphi'}{\partial x'^i} = \frac{\partial \varphi(x)}{\partial x^i} + \mathcal{O}(\xi)$$

and therefore

$$\varphi'(x^i) = \varphi(x) - \xi^i \partial_i \varphi\,, \tag{1.8}$$

where we have introduced a useful notation $\partial_i = \frac{\partial}{\partial x^i}$.

Exercise 7. Using (1.7) and (1.8), verify relation (1.6).

Exercise 8.* Continue the expansions (1.7) and (1.8) until the second order in ξ, and consequently consider $\xi^i = \xi^i(x)$. Use (1.6) to verify your calculus.

Def. 7. The set of three functions $\{a^i(x)\} = \{a^1(x), a^2(x), a^3(x)\}$ form contravariant vector field (or simply vector field) if they transform, under the change of coordinates $\{x^i\} \to \{x'^j\}$, as

$$a^{j'}(x') = \frac{\partial x^{j'}}{\partial x^i} \cdot a^i(x) \tag{1.9}$$

at any point of space.

Observation 1. One can easily see that this transformation rule is different from the one for scalar field (1.6). In particular, the components of the vector in a given geometrical point of space do modify under the coordinate transformation while the scalar field does not. The transformation rule (1.9) corresponds to the geometric object **a** in $3D$. Contrary to the scalar φ, the vector **a** has direction - that is exactly the origin of the apparent difference between the transformation laws.

Observation 2. Some physical examples of quantities (fields) which have vector transformation rule are perhaps familiar to the reader: electric field, magnetic field, instantaneous velocities of particles in a stationary flux of a fluid etc. Later on we shall consider particular examples of the transformation of vectors under rotations and inversions.

Observation 3. The scalar and vector fields can be considered as examples of the more general objects called tensors. Tensors are also defined through their transformation rules.

Def. 8. The set of 3^n functions $\{a^{i_1 \dots i_n}(x)\}$ is called contravariant tensor of rank n, if these functions transform, under $x^i \to x'^i$, as

$$a^{i'_1 \dots i'_n}(x') = \frac{\partial x^{i'_1}}{\partial x^{j_1}} \dots \frac{\partial x^{i'_n}}{\partial x^{j_n}}\, a^{j_1 \dots j_n}(x)\,. \tag{1.10}$$

Observation 4. According to the above definitions the scalar and contravariant vector fields are nothing but the particular cases of the contravariant tensor field. Scalar is a tensor of rank 0, and vector is a tensor of rank 1.

Exercise 9. Show that the product $a^i \cdot b^j$ is a contravariant second rank tensor, if both a^i and b^i are contravariant vectors.

1.3 Orthonormal basis and Cartesian coordinates

Def. 9. Scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined in a usual way

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \cos(\widehat{\mathbf{a}, \mathbf{b}})$$

where $|\mathbf{a}|$, $|\mathbf{b}|$ are absolute values of the vectors $\mathbf{a}$, $\mathbf{b}$ and $(\widehat{\mathbf{a}, \mathbf{b}})$ is the angle between these two vectors.

Of course, the scalar product has the properties familiar from the Analytic Geometry, such as linearity

$$(\mathbf{a}, \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2) = \alpha_1 (\mathbf{a}, \mathbf{b}_1) + \alpha_2 (\mathbf{a}, \mathbf{b}_2)$$

and symmetry

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a})\,.$$

Observation. Sometimes, we shall use a different notation for the scalar product

$$(\mathbf{a}, \mathbf{b}) \equiv \mathbf{a} \cdot \mathbf{b}\,.$$

Until now, we considered an arbitrary basis $\{\mathbf{e}_i\}$. Indeed, we requested the vectors $\{\mathbf{e}_i\}$ to be point-independent, but the length of each of these vectors and the angles between them were restricted only by the requirement of their linear independence. It proves useful to define a special basis, which corresponds to the conventional Cartesian coordinates.

Def. 10. Special orthonormal basis $\{\hat{\mathbf{n}}_a\}$ is the one with

$$(\hat{\mathbf{n}}_a, \hat{\mathbf{n}}_b) = \delta_{ab} = \begin{cases} 1 & \text{if} \quad a = b \\ 0 & \text{if} \quad a \neq b \end{cases} \tag{1.11}$$

Exercise 10. Making transformations of the basis vectors, verify that the change of coordinates

$$x' = \frac{x+y}{\sqrt{2}} + 3\,, \qquad y' = \frac{x-y}{\sqrt{2}} - 5$$

does not modify the type of coordinates x', y', which remains Cartesian.

Exercise 11. Using the geometric intuition, discuss the form of coordinate transformations which do not violate the orthonormality of the basis.

Notations. We shall denote the coordinates corresponding to the $\{\hat{\mathbf{n}}_a\}$ basis as X^a. With a few exceptions, we reserve the indices $a, b, c, d, ...$ for these (Cartesian) coordinates, and use indices $i, j, k, l, m, ...$ for an arbitrary coordinates. For example, the radius-vector $\mathbf{r}$ in $3D$ can be presented using different coordinates:

$$\mathbf{r} = x^i \mathbf{e}_i = X^a \hat{\mathbf{n}}_a = x\hat{\mathbf{n}}_x + y\hat{\mathbf{n}}_y + z\hat{\mathbf{n}}_z\,. \tag{1.12}$$

Indeed, the last equality is identity, because

$$X^1 = x\,, \qquad X^2 = y\,, \qquad X^3 = z\,.$$

Sometimes, we shall also use notations

$$\hat{\mathbf{n}}_x = \hat{\mathbf{n}}_1 = \hat{\mathbf{i}}\,, \qquad \hat{\mathbf{n}}_y = \hat{\mathbf{n}}_2 = \hat{\mathbf{j}}\,, \qquad \hat{\mathbf{n}}_z = \hat{\mathbf{n}}_3 = \hat{\mathbf{k}}\,.$$

Observation 1. One can express the vectors of an arbitrary basis $\{\mathbf{e}_i\}$ as linear combinations of the components of the orthonormal basis

$$\mathbf{e}_i = \frac{\partial X^a}{\partial x^i}\,\hat{\mathbf{n}}_a.$$

An inverse relation has the form

$$\hat{\mathbf{n}}_a = \frac{\partial x^j}{\partial X^a}\,\mathbf{e}_j\,.$$

Exercise 12. Prove the last formulas using the uniqueness of the components of the vector $\mathbf{r}$ for a given basis.

Observation 2. It is easy to see that

$$X^a \hat{\mathbf{n}}_a = x^i \mathbf{e}_i = x^i \cdot \frac{\partial X^a}{\partial x^i}\,\hat{\mathbf{n}}_a \quad \Longrightarrow \quad X^a = \frac{\partial X^a}{\partial x^i} \cdot x^i\,. \tag{1.13}$$

Similarly,

$$x^i = \frac{\partial x^i}{\partial X^a} \cdot X^a\,. \tag{1.14}$$

Indeed, these relations holds only because the basis vectors $\mathbf{e}_i$ do not depend on the space-time points and the matrices

$$\frac{\partial X^a}{\partial x^i} \qquad \text{and} \qquad \frac{\partial x^i}{\partial X^a}$$

have constant elements. Later on, we shall consider more general (called nonlinear or local) coordinate transformations where the relations (1.13) and (1.14) do not hold.

Now we can introduce a conjugated covariant basis.

Def. 11. Consider basis $\{\mathbf{e}_i\}$. The conjugated basis is defined as a set of vectors $\{\mathbf{e}^j\}$, which satisfy the relations

$$\mathbf{e}_i \cdot \mathbf{e}^j \;=\; \delta_i^j\,. \tag{1.15}$$

Remark: The special property of the orthonormal basis is that $\hat{\mathbf{n}}^a \;=\; \hat{\mathbf{n}}_a$, therefore in this special case the conjugated basis coincides with the original one.

Exercise 13. Prove that only an orthonormal basis may have this property.

Def. 12. Any vector $\mathbf{a}$ can be expanded using the conjugated basis $\mathbf{a} = a_i \mathbf{e}^i$. The coefficients a_i are called **covariant** components of the vector $\mathbf{a}$. Sometimes we use to say that these coefficients form the covariant vector.

Exercise 14. Consider, in $2D$, $\mathbf{e}_1 = \hat{\mathbf{i}} + \hat{\mathbf{j}}$ and $\mathbf{e}_2 = \hat{\mathbf{i}}$. Find basis conjugated to $\{\mathbf{e}_1\,,\;\mathbf{e}_2\}$.

The next question is how do the coefficients a_i transform under the change of coordinates?

Theorem. If we change the basis of the coordinate system from $\mathbf{e}_i$ to $\mathbf{e}'_i$, then the covariant components of the vector $\mathbf{a}$ transform as

$$a'_i \;=\; \frac{\partial x^j}{\partial x'^i}\, a_j\,. \tag{1.16}$$

Remark. It is easy to see that in the case of a covariant vector components the transformation performs by means of the matrix inverse to the one for the contravariant components (1.9).

Proof: First of all we need to learn how do the vectors of the conjugated basis transform. Let us use the definition of the conjugated basis in both coordinate systems.

$$(\mathbf{e}_i, \mathbf{e}^j) = (\mathbf{e}'_i\,,\,\mathbf{e}'^j) = \delta_i^j\,.$$

Hence, since

$$\mathbf{e}'_i = \frac{\partial x^k}{\partial x'^i}\,\mathbf{e}_k\,,$$

and due to the linearity of a scalar product of two vectors, we obtain

$$\frac{\partial x^k}{\partial x'^i}\,(\mathbf{e}_k\,,\,\mathbf{e}'^j) = \delta_i^j\,.$$

Therefore, the matrix $(\mathbf{e}_k, \mathbf{e}^{j'})$ is inverse to the matrix $\left(\frac{\partial x^k}{\partial x'^i}\right)$. Using the uniqueness of the inverse matrix we arrive at the relation

$$(\mathbf{e}_k, \mathbf{e}^{j'}) = \frac{\partial x^{j'}}{\partial x^k} = \frac{\partial x^{j'}}{\partial x^l}\,\delta_k^l = \frac{\partial x^{j'}}{\partial x^l}(\mathbf{e}_k, \mathbf{e}^l) = (\mathbf{e}_k, \frac{\partial x^{j'}}{\partial x^l}\,\mathbf{e}^l), \qquad \forall \mathbf{e}_k.$$

Thus, $\mathbf{e}^{j'} = \frac{\partial x^{j'}}{\partial x^l}\,\mathbf{e}^l$. Using the standard way of dealing with vector transformations, we obtain

$$a_{j'}\mathbf{e}^{j'} = a_{j'}\,\frac{\partial x^{j'}}{\partial x^l}\,\mathbf{e}^l \quad \Longrightarrow \quad a_l = \frac{\partial x'^j}{\partial x^l}\,a'_j\,.$$

1.4 Covariant and mixed vectors and tensors

Now we are in a position to define covariant vectors, tensors and mixed tensors.

Def. 13. The set of 3 functions $\{A_i(x)\}$ form covariant vector field, if they transform from one coordinate system to another as

$$A'_i(x') = \frac{\partial x^j}{\partial x'^i}\,A_j(x)\,. \tag{1.17}$$

The set of 3^n functions $\{A_{i_1 i_2 \dots i_n}(x)\}$ form covariant tensor of rank n if they transform from one coordinate system to another as

$$A'_{i_1 i_2 \dots i_n}(x') = \frac{\partial x^{j_1}}{\partial x'^{i_1}}\,\frac{\partial x^{j_2}}{\partial x'^{i_2}}\,\dots\,\frac{\partial x^{j_n}}{\partial x'^{i_n}}\,A_{j_1 j_2 \dots j_n}(x)\,.$$

Exercise 15. Show that the product of two covariant vectors $A_i(x)B_j(x)$ transforms as a second rank covariant tensor. Discuss whether *any* second rank covariant tensor can be presented as such a product.

Hint. Try to evaluate the number of independent functions in both cases.

Def. 14. The set of 3^{n+m} functions $\{B_{i_1 \dots i_n}{}^{j_1 \dots j_m}(x)\}$ form tensor of the type (m, n) (another possible names are mixed tensor of covariant rank n and contravariant rank m or simply (m, n)-tensor), if these functions transform, under the change of coordinate basis, as

$$B_{i'_1 \dots i'_n}{}^{j'_1 \dots j'_m}(x') \;=\; \frac{\partial x^{j'1}}{\partial x^{l_1}}\,\dots\,\frac{\partial x^{j'm}}{\partial x^{l_m}}\,\frac{\partial x^{k_1}}{\partial x'^{i_1}}\,\dots\,\frac{\partial x^{k_n}}{\partial x'^{i_n}}\quad B_{k_1 \dots k_n}{}^{l_1 \dots l_m}(x)\,.$$

Exercise 16. Verify that the scalar, co- and contravariant vectors are, correspondingly, $(0, 0)$-tensor, $(0, 1)$-tensor and $(1, 0)$-tensor.

Exercise 17. Prove that if the Kronecker symbol transforms as a mixed $(1, 1)$ tensor, then in any coordinates x^i it has the same form

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

Observation. This property is very important, for it enables us to use the Kronecker symbol in any coordinates.

Exercise 18. Show that the product of co- and contra-variant vectors $A^i(x)B_j(x)$ transforms as a $(1, 1)$-type mixed tensor.

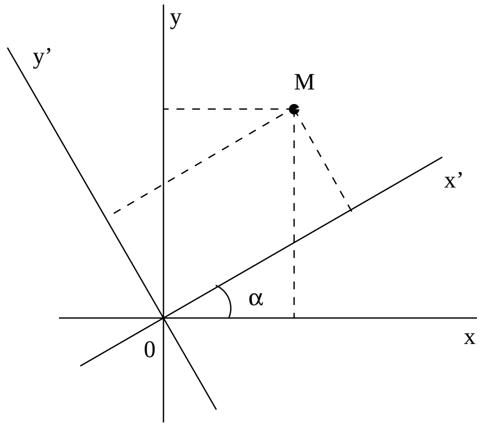

Figure 1.1: Illustration of the $2D$ rotation to the angle α.

Observation. The importance of tensors is due to the fact that they offer the opportunity of the *coordinate-independent* description of geometrical and physical laws. Any tensor can be viewed as a geometric object, independent on coordinates. Let us consider, as an example, the mixed $(1,1)$-type tensor with components A_i^j. The tensor (as a geometric object) can be presented as a contraction of A_i^j with the basis vectors

$$\mathbf{A} = A_i^j \, \mathbf{e}^i \otimes \mathbf{e}_j \, . \tag{1.18}$$

The operation $\otimes$ is called "direct product", it indicates that the basis for the tensor $\mathbf{A}$ is composed by the products of the type $\mathbf{e}^i \otimes \mathbf{e}_j$. In other words, the tensor $\mathbf{A}$ is a linear combination of such "direct products". The most important observation is that the Eq. (1.18) transforms as a scalar. Hence, despite the *components* of a tensor are dependent on the choice of the basis, the tensor itself is coordinate-independent.

Exercise 19. Discuss the last observation for the case of a vector.

1.5 Orthogonal transformations

Let us consider important particular cases of the global coordinate transformations, which are called orthogonal. The orthogonal transformations may be classified to rotations, inversions (parity transformations) and their combinations.

First we consider the rotations in the $2D$ space with the initial Cartesian coordinates (x, y). Suppose another Cartesian coordinates (x', y') have the same origin as (x, y) and the difference is the rotation angle α. Then, the same point M (see the Figure 1.1) has coordinates (x,y) and (x', y'), and the relation between the coordinates is

$$\begin{array}{l} x = x' \cos\alpha - y' \sin\alpha \\ y = x' \sin\alpha + y' \cos\alpha \end{array} \, . \tag{1.19}$$

Exercise 20. Check that the inverse transformation has the form or rotation to the same angle but in the opposite direction.

$$x' = x\cos\alpha + y\sin\alpha \\ y' = -x\sin\alpha + y\cos\alpha \,. \tag{1.20}$$

The above transformations can be seen from the $3D$ viewpoint and also may be presented in a matrix form, e.g.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \hat{\wedge}_z \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \qquad \text{where} \qquad \hat{\wedge}_z = \hat{\wedge}_z(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.21}$$

Here the index z indicates that the matrix $\hat{\wedge}_z$ executes the rotation around the $\hat{\mathbf{z}}$ axis.

Exercise 21. Write the rotation matrices around $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ axices with the angles γ and β correspondingly.

Exercise 22. Check the following properties of the matrix $\hat{\wedge}_z$:

$$\hat{\wedge}_z^T = \hat{\wedge}_z^{-1}\,. \tag{1.22}$$

Def. 15. The matrix $\hat{\wedge}_z$ which satisfies $\hat{\wedge}_z^{-1} = \hat{\wedge}_z^T$ and the corresponding coordinate transformation are called orthogonal [1].

In $3D$ any rotation of the rigid body may be represented as a combination of the rotations around the axis $\hat{\mathbf{z}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{x}}$ to the angles [2] α, β and γ. Hence, the general $3D$ rotation matrix $\hat{\wedge}$ may be represented as

$$\hat{\wedge} = \hat{\wedge}_z(\alpha)\,\hat{\wedge}_y(\beta)\,\hat{\wedge}_x(\gamma)\,. \tag{1.23}$$

If we use the known properties of the square invertible matrices

$$(A\cdot B)^T = B^T\cdot A^T \qquad \text{and} \qquad (A\cdot B)^{-1} = B^{-1}\cdot A^{-1}\,, \tag{1.24}$$

we can easily obtain that the general $3D$ rotation matrix satisfies the relation $\hat{\wedge}^T = \hat{\wedge}^{-1}$. Therefore we proved that an arbitrary rotation matrix is orthogonal.

Exercise 23. Investigate the following relations and properties of the matrices $\hat{\wedge}$:

(i) Check explicitly, that $\hat{\wedge}_z^T(\alpha) = \hat{\wedge}_z^{-1}(\alpha)$

(ii) Check that $\hat{\wedge}_z(\alpha)\cdot\hat{\wedge}_z(\beta) = \hat{\wedge}_z(\alpha+\beta)$ (group property).

[1] In case the matrix elements are allowed to be complex, the matric which satisfies the property $U^\dagger = U^{-1}$ is called unitary. The operation $U^\dagger$ is called Hermitian conjugation and consists in complex conjugation plus transposition $U^\dagger = (U^*)^{-1}$.

[2] One can prove that it is always possible to make an arbitrary rotation of the rigid body by performing the sequence of particular rotations $\hat{\wedge}_z(\alpha)\,\hat{\wedge}_y(\beta)\,\hat{\wedge}_z(-\alpha)$. The proof can be done using the Euler angles (see, e.g. [13], [12]).

(iii) * Write all three matrices $\hat{\wedge}_z(\alpha)$, $\hat{\wedge}_x(\gamma)$, and $\hat{\wedge}_y(\beta)$ in $3D$ for the case of infinitesimal angles α, β, γ. Find all three non-trivial commutators

$$\begin{aligned}
\left[\hat{\wedge}_z(\alpha)\,,\,\hat{\wedge}_y(\beta)\right] &= \hat{\wedge}_z(\alpha)\,\hat{\wedge}_y(\beta) - \hat{\wedge}_y(\beta)\,\hat{\wedge}_z(\alpha)\,, \\
\left[\hat{\wedge}_y(\beta)\,,\,\hat{\wedge}_x(\gamma)\right] &= \hat{\wedge}_y(\beta)\,\hat{\wedge}_x(\gamma) - \hat{\wedge}_x(\gamma)\,\hat{\wedge}_y(\beta)\,, \\
\left[\hat{\wedge}_z(\alpha)\,,\,\hat{\wedge}_x(\gamma)\right] &= \hat{\wedge}_z(\alpha)\,\hat{\wedge}_x(\gamma) - \hat{\wedge}_x(\gamma)\,\hat{\wedge}_z(\alpha)\,.
\end{aligned} \tag{1.25}$$

of these matrices for $\alpha = \beta = \gamma$. Try to establish the relations between these commutators and the matrices $\hat{\wedge}$ themselves.

Hint. When making commutations keep only the first order in the angle α. In this way, one may replace

$$\sin\alpha \simeq \alpha \qquad \text{and} \qquad \cos\alpha \simeq 1\,.$$

(iv) Verify that the determinant $\det \hat{\wedge}^z(\alpha) = 1$. Prove that the determinant is equal to one for any rotation matrix.

Observation. Since for the orthogonal matrix $\hat{\wedge}$ we have the equality

$$\hat{\wedge}^{-1} = \hat{\wedge}^T\,,$$

we can take determinant and arrive at $\det\hat{\wedge} = \det\hat{\wedge}^{-1}$ and therefore $\det\hat{\wedge} = \pm 1$. As far as any rotation matrix has determinant equal to one, there must be some other orthogonal matrices with the determinant equal to -1. An examples of such matrices are, e.g., the following:

$$\hat{\pi}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad \hat{\pi} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{1.26}$$

The last matrix corresponds to the transformation which is called inversion. This transformation means that we simultaneously change the direction of all the axis to the opposite ones. One can prove that an arbitrary orthogonal matrix can be presented as a product of inversion and rotation matrices.

Exercise 24. (i) Following the pattern of Eq. (1.26), construct $\hat{\pi}_y$ and $\hat{\pi}_x$.
(ii) Find rotation matrices which transform $\hat{\pi}_z$ into $\hat{\pi}_x$ and $\hat{\pi}$ into $\hat{\pi}_x$.
(iii) Suggest geometric interpretation of the matrices $\hat{\pi}_x$, $\hat{\pi}_y$ and $\hat{\pi}_z$.

Chapter 2

Operations over tensors, metric tensor

In this Chapter we shall define operations over tensors. These operations may involve one or two tensors. The most important point is that the result is always a tensor.

Operation 1. Summation is defined only for tensors of the same type. Sum of two tensors of type (m, n) is also a tensor type (m, n), which has components equal to the sums of the corresponding components of the tensors-summands.

$$(A+B)_{i_1 \dots i_n}{}^{j_1 \dots j_m} = A_{i_1 \dots i_n}{}^{j_1 \dots j_m} + B_{i_1 \dots i_n}{}^{j_1 \dots j_m} \,. \tag{2.1}$$

Exercise 1. Prove that the components of $(A + B)$ in (2.1) form a tensor, if $A_{i_1 \dots i_n}{}^{j_1 \dots j_m}$ and $B_{i_1 \dots i_n}{}^{j_1 \dots j_m}$ form tensors.

Hint. First try this proof for vectors and scalars.

Exercise 2. Prove the commutativity and associativity of the tensor summation.

Hint. Use the definition of tensor **Def. 1.14**.

Operation 2. Multiplication of a tensor by number produces the tensor of the same type. This operation is equivalent to the multiplication of all tensor components to the same number α.

$$(\alpha A)_{i_1 \dots i_n}{}^{j_1 \dots j_m} = \alpha \cdot A_{i_1 \dots i_n}{}^{j_1 \dots j_m} \,. \tag{2.2}$$

Exercise 3. Prove that the components of (αA) in (2.2) form a tensor, if $A_{i_1 \dots i_n}{}^{j_1 \dots j_m}$ are components of a tensor.

Operation 3. Multiplication of 2 tensors is defined for a couple of tensors of any type. The product of a (m, n)-tensor and a (t, s)-tensor results in the $(m + t, n + s)$-tensor, e.g.:

$$A_{i_1 \dots i_n}{}^{j_1 \dots j_m} \cdot C_{l_1 \dots l_s}{}^{k_1 \dots k_t} = D_{i_1 \dots i_n}{}^{j_1 \dots j_m}{}_{l_1 \dots l_s}{}^{k_1 \dots k_t} \,. \tag{2.3}$$

We remark that the order of indices is important here, because, say, a_{ij} may be different from a_{ji}.

Exercise 4. Prove, by checking the transformation law, that the product of the covariant vector a^i and the mixed tensor b_j^k is a mixed $(2, 1)$ - type tensor.

Exercise 5. Prove the following theorem: the product of an arbitrary (m,n)-type tensor and a scalar is a (m,n)-type tensor. Formulate and prove similar theorems concerning multiplication by co- and contravariant vectors.

Operation 4. Contraction reduces the (n,m)-tensor to the $(n-1,m-1)$-tensor through the summation over two (always upper and lower, of course) indices.

Example.

$$A_{ijk}{}^{ln} \to A_{ijk}{}^{lk} = \sum_{k=1}^{3} A_{ijk}{}^{lk} \tag{2.4}$$

Operation 5. Internal product of the two tensors consists in their multiplication with the consequent contraction over some couple of indices. Internal product of (m,n) and (r,s)-type tensors results in the $(m+r-1, n+s-1)$-type tensor.

Example:

$$A_{ijk} \cdot B^{lj} = \sum_{k=1}^{3} A_{ijk}\, B^{lj} \tag{2.5}$$

Observation. Contracting another couple of indices may give another internal product.

Exercise 6. Prove that the internal product $a_i \cdot b^i$ is a scalar if $a_i(x)$ and $b^i(x)$ are co- and contravariant vectors.

Now we can introduce one of the central notions of this course.

Def. 1. Consider some basis $\{\mathbf{e}_i\}$. The scalar product of two basis vectors

$$g_{ij} = (\mathbf{e}_i, \mathbf{e}_j) \tag{2.6}$$

is called metric.

Properties:

1. Symmetry $g_{ij} = g_{ji}$ follows from the symmetry of a scalar product $(\mathbf{e}_i, \mathbf{e}_j) = (\mathbf{e}_j, \mathbf{e}_i)$.

2. For the orthonormal basis $\hat{\mathbf{n}}_a$ the metric is nothing but the Kronecker symbol

$$g_{ab} = (\hat{\mathbf{n}}_a, \hat{\mathbf{n}}_b) = \delta_{ab}\,.$$

3. Metric is a $(2,0)$ - tensor. The proof of this statement is very simple:

$$g_{i'j'} = (\mathbf{e}'_i, \mathbf{e}'_j) = \Big(\frac{\partial x^l}{\partial x'^i}\mathbf{e}_l\,,\ \frac{\partial x^k}{\partial x'^j}\mathbf{e}_k\Big) = \frac{\partial x^l}{\partial x'^i}\,\frac{\partial x^k}{\partial x'^j}\,\Big(\mathbf{e}_l\,,\,\mathbf{e}_k\Big) = \frac{\partial x^l}{\partial x'^i}\,\frac{\partial x^k}{\partial x'^j}\cdot g_{kl}\,.$$

Sometimes we shall call metric as metric tensor.

4. The distance between 2 points: $M_1(x^i)$ and $M_2(y^i)$ is defined by

$$S_{12}^2 = g_{ij}(x^i - y^i)(x^j - y^j)\,. \tag{2.7}$$

Proof. Since g_{ij} is $(2,0)$ - tensor and $(x^i - y^i)$ is $(0,1)$ - tensor (contravariant vector), S_{12}^2 is a scalar. Therefore S_{12}^2 is the same in any coordinate system. In particular, we can use the $\hat{\mathbf{n}}_a$ basis, where the (Cartesian) coordinates of the two points are $M_1(X^a)$ and $M_2(Y^a)$. Making the transformation, we obtain

$$S_{12}^2 = g_{ij}(x^i - y^i)(x^j - y^j) = g_{ab}(X^a - Y^a)(X^b - Y^b) =$$

$$= \delta_{ab}(X^a - Y^a)(X^b - Y^b) = (X^1 - Y^1)^2 + (X^2 - Y^2)^2 + (X^3 - Y^3)^2\,,$$

that is the standard expression for the square of the distance between the two points in Cartesian coordinates.

Exercise 7. Check, by direct inspection of the transformationlaw, that the double internal product $g_{ij}z^iz^j$ is a scalar. In the proof of the Property 4, we used this for $z^i = x^i - y^i$.

Exercise 8. Use the definition of the metric and the relations

$$x^i = \frac{\partial x^i}{\partial X^a}\,X^a \qquad \text{and} \qquad \mathbf{e}_i = \frac{\partial X^a}{\partial x^i}\,\hat{\mathbf{n}}_a$$

to prove the Property 4, starting from $S_{12}^2 = g_{ab}(X^a - Y^a)(X^b - Y^b)$.

Def. 2. The conjugated metric is defined as

$$g^{ij} = (\mathbf{e}^i, \mathbf{e}^j) \qquad \text{where} \qquad (\mathbf{e}^i, \mathbf{e}_k) = \delta^i_k\,.$$

Indeed, $\mathbf{e}^i$ are the vectors of the basis conjugated to $\mathbf{e}_k$ (see **Def. 1.11**).

Exercise 9. Prove that g^{ij} is a contravariant second rank tensor.

Theorem 1.

$$g^{ik}g_{kj} = \delta^i_j \tag{2.8}$$

Due to this theorem, the conjugated metric g^{ij} is conventionally called inverse metric.

Proof. For the special orthonormal basis $\hat{\mathbf{n}}^a = \hat{\mathbf{n}}_a$ we have $g^{ab} = \delta^{ab}$ and $g_{bc} = \delta_{bc}$. Of course, in this case the two metrics are inverse matrices

$$g^{ab} \cdot g_{bc} = \delta^a_c\,.$$

Now we can use the result of the Excersize 9. The last equality can be transformed to an arbitrary basis $\mathbf{e}_i$, by multiplying both sides by $\frac{\partial x^i}{\partial X^a}$ and $\frac{\partial X^c}{\partial x^j}$ and inserting the identity matrix (in the parenthesis) as follows

$$\frac{\partial x^i}{\partial X^a}\,g^{ab} \cdot g_{bc}\,\frac{\partial X^c}{\partial x^j} = \frac{\partial x^i}{\partial X^a}\,g^{ab}\left(\frac{\partial x^k}{\partial X^b}\,\frac{\partial X^d}{\partial x^k}\right)g_{dc}\,\frac{\partial X^c}{\partial x^j} \tag{2.9}$$

The last expression can be rewritten as

$$\frac{\partial x^i}{\partial X^a}\, g^{ab}\, \frac{\partial x^k}{\partial X^b} \;\cdot\; \frac{\partial X^d}{\partial x^k}\, g_{dc}\, \frac{\partial X^c}{\partial x^j} \;=\; g^{ik}\, g_{kj}\,.$$

But, at the same time the same expression (2.9) can be presented in other form

$$\frac{\partial x^i}{\partial X^a}\, g^{ab} \cdot g_{bc}\, \frac{\partial X^c}{\partial x^j} = \frac{\partial x^i}{\partial X^a}\, \delta^a_c\, \frac{\partial X^c}{\partial x^j} = \frac{\partial x^i}{\partial X^a}\, \frac{\partial X^a}{\partial x^j} = \delta^i_j$$

that completes the proof.

Operation 5. Raising and lowering indices of a tensor. This operation consists in taking an appropriate internal product of a given tensor and the corresponding metric tensor.

Examples.
Lowering the index:

$$A_i(x) = g_{ij} A^j(x)\,, \qquad\qquad B_{ik}(x) = g_{ij} B^j{}_k(x)\,.$$

Raising the index:

$$C^l(x) = g^{lj}\, C_j(x)\,, \qquad\qquad D^{ik}(x) = g^{ij}\, D_j{}^k(x)\,.$$

Exercise 10. Prove the relations:

$$\mathbf{e}_i\, g^{ij} = \mathbf{e}^j\,, \qquad\qquad \mathbf{e}^k\, g_{kl} = \mathbf{e}_l\,. \tag{2.10}$$

Exercise 11 *. Try to solve the previous Exercise in several distinct ways.

The following Theorem clarifies the geometric sense of the metric and its relation with the change of basis / coordinate system. Furthermore, it has serious practical importance and will be extensively used below.

Theorem 2. Consider the metric g_{ij} corresponding to the basis $\{\mathbf{e}_k\}$ and to the coordinates x^l. The following relations between the determinants of the metric and of the matrix of the transformation to Cartesian coordinates holds:

$$g = \det\,(g_{ij}) = \det\left(\frac{\partial X^a}{\partial x^k}\right)^2, \qquad\qquad g^{-1} = \det\,(g^{kl}) = \det\left(\frac{\partial x^l}{\partial X^b}\right)^2. \tag{2.11}$$

Proof. The first observation is that, in the Cartesian coordinates the metric is identity matrix $g_{ab} = \delta_{ab}$ and therefore $\det\,(g_{ab}) = 1$. Next, since the metric is a tensor,

$$g_{ij} = \frac{\partial X^a}{\partial x^i}\, \frac{\partial X^b}{\partial x^j}\, g_{ab}\,.$$

Using the known property of determinants (you are going to prove this relation in the next Chapter as an Exercise)

$$\det\,(A \cdot B) = \det\,(A) \cdot \det\,(B)$$

we arrive at the first relation in (2.11). The proof of the second relation is similar and we leave it to the reader as an Exercise.

Observation. The possibility of lowering and raising the indices may help us in contracting two conta or two covariant indices of a tensor. For this end, we need just to raise (lower) one of two indices and then perform contraction.

Examples.

$$A^{ij} \quad \longrightarrow \quad A^{k}_{\cdot j} = A^{ij}\, g_{ik} \quad \longrightarrow \quad A^{k}_{\cdot k}\,.$$

Here the point indicates the position of the raised index. Sometimes, this is an important information, which helps to avoid an ambiguity. Let us see this using another simple example.

Suppose we need to contract the two first indices of the $(3,0)$-tensor B^{ijk}. After we lower the second index, we arrive at the tensor

$$B^{i}{}_{\cdot l \cdot}{}^{k} = B^{ijk}\, g_{jl} \tag{2.12}$$

and now can contract the indices i and l. But, if we forget to indicate the order of indices, we obtain, instead of (2.12), the formula B^{ik}_{l}, and it is not immediately clear which index was lowered and in which couple of indices one has to perform the contraction.

Exercise 12. Transform the vector $\mathbf{a} = a_1 \hat{\mathbf{i}} + a_2 \hat{\mathbf{j}}$ in $2D$ into another basis

$$\mathbf{e}_1 = 3\hat{\mathbf{i}} + \hat{\mathbf{j}}\,, \qquad \mathbf{e}_2 = \hat{\mathbf{i}} - \hat{\mathbf{j}}\,, \qquad \text{using}$$

(i) direct substitution of the basis vectors;
(ii) matrix elements of the corresponding transformation $\frac{\partial x^i}{\partial X^a}$.

Chapter 3

Symmetric, skew(anti) symmetric tensors and determinants

3.1 Symmetric and antisymmetric tensors

Def. 1. Tensor $A^{ijk...}$ is called symmetric in the indices i and j, if $A^{ijk...} = A^{jik...}$.

Example: As we already know, the metric tensor is symmetric, because

$$g_{ij} = (\mathbf{e}_i, \mathbf{e}_j) = (\mathbf{e}_j, \mathbf{e}_i) = g_{ji}\,. \tag{3.1}$$

Exercise 1. Prove that the inverse metric is also symmetric tensor $g^{ij} = g^{ji}$.

Exercise 2. Let a^i are the components of a contravariant vector. Prove that the product $a^i a^j$ is a symmetric contravariant $(2,0)$-tensor.

Exercise 3. Let a^i and b^j are components of two (generally different) contravariant vectors. As we already know, the products $a^i b^j$ and $a^j b^i$ are contravariant tensor. Construct a linear combination $a^i b^j$ and $a^j b^i$ such that it would be a symmetric tensor.

Exercise 4. Let c^{ij} be components of a contravariant tensor. Construct a linear combination of c^{ij} and c^{ji} which would be a symmetric tensor.

Observation 1. In general, the symmetry reduces the number of independent components of a tensor. For example, an arbitrary $(2,0)$-tensor has 9 independent components, while a symmetric tensor of 2-nd rank has only 6 independent components.

Exercise 5. Evaluate the number of independent components of a symmetric second rank tensor B^{ij} in D-dimensional space.

Def. 2. Tensor $A^{i_1 i_2 ... i_n}$ is called completely (or absolutely) symmetric in the indices $(i_1, i_2, \dots, i_n)$, if it is symmetric in any couple of these indices.

$$A^{i_1 ... i_{k-1} i_k i_{k+1} ... i_{l-1}\, i_l\, i_{l+1} ... i_n} = A^{i_1 ... i_{k-1}\, i_l\, i_{k+1} ... i_{l-1} i_k i_{l+1} ... i_n}\,. \tag{3.2}$$

Observation 2. Tensor can be symmetric in any couple of contravariant indices or in any couple of covariant indices, and not symmetric in other couples of indices.

Exercise 6. * Evaluate the number of independent components of an absolutely symmetric third rank tensor B^{ijk} in a D-dimensional space.

Def. 3. Tensor A^{ij} is calles skew-symmetric or antisymmetric, if

$$A^{ij} = -A^{ij} \,. \qquad (3.3)$$

Observation 3. We shall formulate the most of consequent definitions and theorems for the contravariant tensors. The statements for the covariant tensors are very similar and can be prepared by the reader as simple Exercise.

Observation 4. The above definitions and consequent properties and relations use only algebraic operations over tensors, and do not address the transformation from one basis to another. Hence, one can consider not symmetric or antisymmetric tensors, but just symmetric or antisymmetric objects provided by indices. The advantage of tensors is that their (anti)symmetry holds under the transformation from one basis to another. But, the same property is shared by a more general objects calles tensor densities (see **Def. 8** below).

Exercise 7. Let c_{ij} be components of a contravariant tensor. Construct a linear combination of c_{ij} and c_{ji} which would be an antisymmetric tensor.

Exercise 8. Let c_{ij} be the components of a contravariant tensor. Show that c_{ij} can be always presented as a sum of a symmetric tensor $c_{(ij)}$

$$c_{(ij)} = c_{(ji)} \qquad (3.4)$$

and an antisymmetric tensor $c_{[ij]}$

$$c_{[ij]} = -c_{[ji]} \,. \qquad (3.5)$$

Exercise 9. Discuss the peculiarities of this representation for the cases when the initial tensor c_{ij} is, by itself, symmetric or antisymmetric. Discuss whether there are any restrictions on the dimension D for this representation.

Hint 1. One can always represent the tensor c_{ij} as

$$c_{ij} \,=\, \frac{1}{2}\,(\,c_{ij} + c_{ji}\,) \,+\, \frac{1}{2}\,(\,c_{ij} - c_{ji}\,)\,. \qquad (3.6)$$

Hint 2. Consider particular cases $D = 2$, $D = 3$ and $D = 1$.

Def. 4. $(n, 0)$-tensor $A^{i_1...i_n}$ is called completely (or absolutely) antisymmetric, if it is antisymmetric in any couple of its indices

$$\forall (k, l)\,, \qquad A^{i_1...i_l...i_k...i_n} = -A^{i_1...i_k...i_l...i_n}\,.$$

In the case of an absolutely antisymmetric tensor, the sign changes when we perform permutation of any two indices.

Exercise 10. *(i)* Evaluate the number of independent components of an antisymmetric tensor A^{ij} in 2 and 3-dimensional space. *(ii)* Evaluate the number of independent components of absolutely antisymmetric tensors A^{ij} and A^{ijk} in a D-dimensional space, where $D > 3$.

Theorem 1. In a D-dimensional space all completely antisymmetric tensors of rank $n > D$ equal zero.

Proof. First we notice that if two of its indices are equal, an absolutely skew-symmetric tensor is zero:

$$A^{...i...i...} = -A^{...i...i...} \qquad \Longrightarrow \qquad A^{...i...i...} = 0\,.$$

Furthermore, in a D - dimensional space any index can take only D distinct values, and therefore among $n > D$ indices there are at least two equal ones. Therefore, all components of $A^{i_1 i_2 ... i_n}$ equal zero.

Theorem 2. An absolutely antisymmetric tensor $A^{i_1...i_D}$ in a D - dimensional space has only one independent component.

Proof. All components with two equal indices are zeros. Consider an arbitrary component with all indices disctinct from each other. Using permutations, it can be transformed into one single component $A^{12...D}$. Since each permulation changes the sign of the tensor, the non-zero components have the absolute value equal to $\left|A^{12...D}\right|$, and the sign dependent on whether the number of necessary permutations is even or odd.

Observation 5. We shall call an absolutely antisymmetric tensor with D indices a maximal absolutely antisymmetric tensor in the D-dimensional space. There is a special case of the maximal absolutely antisymmetric tensor $\varepsilon^{i_1 i_2 ... i_D}$ which is called **Levi-Civita tensor**. In the Cartesian coordinates the non-zero component of this tensor equals

$$\varepsilon^{123...D} \;=\; E^{123...D} = 1\,. \tag{3.7}$$

It is customary to consider the non-tensor absolutely antisymmetric object $E^{i_1 i_2 i_3 ... i_D}$, which has the same components (3.7) in *any* coordinates. The non-zero components of the tensor $\varepsilon^{i_1 i_2 ... i_D}$ in arbitrary coordinates will be derived below.

Observation 6. In general, the tensor $A^{ijk...}$ may be symmetric in some couples of indexes, antisymmetric in some other couples of indices and have no symmetry in some other couple of indices.

Exercise 11. Prove that for symmetric a^{ij} and antisymmetric b_{ij} the scalar internal product is zero $a^{ij} \cdot b_{ij} = 0$.

Observation 7. This is an extremely important statement!

Exercise 12. Suppose that for some tensor a_{ij} and some numbers x and y there is a relation:

$$x\, a_{ij} + y\, a_{ji} = 0$$

Prove that either

$$x = y \quad \text{and} \quad a_{ij} = -a_{ji}$$

or

$$x = -y \quad \text{and} \quad a_{ij} = a_{ji}\,.$$

Exercise 13. *(i)* In $2D$, how many distinct components has tensor c^{ijk}, if

$$c^{ijk} = c^{jik} = -c^{ikj}\ ?$$

(ii) Investigate the same problem in $3D$. *(iii)* Prove that $c^{ij}_{\cdot\cdot j} \equiv 0$ in $\forall D$.

Exercise 14. *(i)* Prove that if $b^{ij}(x)$ is a tensor and $a_l(x)$ and $c_k(x)$ are covariant vectors fields, then $f(x) = b^{ij} \cdot a_i(x) \cdot c_j(x)$ is a scalar field. *(ii)* Try to generalize this Exercise for arbitrary tensors. *(iii)* In case $a_i(x) \equiv c_i(x)$, formulate the condition for $b^{ij}(x)$ providing that the product $f(x) \equiv 0$.

Exercise 15.* *(i)* Prove that, in a D-dimensional space, with $D > 3$, the numbers of independent components of absolutely antisymmetric tensors $A^{k_1 k_2 \dots k_n}$ and $\tilde{A}_{i_1 i_2 i_3 \dots i_{D-n}}$ are equal, for any $n \leq D$. *(ii)* Using the absolutely antisymmetric tensor $\varepsilon_{j_1 j_2 \dots j_D}$, try to find an explicit relation (which is called dual correspondence) between the tensor $A^{k_1 k_2 \dots k_n}$ and some $\tilde{A}_{i_1 i_2 i_3 \dots i_{D-n}}$ and then invert this relation.

3.2 Determinants

There is a very deep relation between maximal absolutely antisymmetric tensors and determinants. Let us formulate the main aspects of this relation as a set of theorems. Some of the theorems will be first formulated in $2D$ and then generalized for the case of an $\forall D$.

Consider first $2D$ and special coordinates X^1 and X^2 corresponding to the orthonormal basis $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$.

Def. 5. We define a special maximal absolutely antisymmetric object in $2D$ using the relations (see **Observation 5**)

$$E_{ab} = -E_{ba} \qquad \text{and} \qquad E_{12} = 1\,.$$

Theorem 3. For any matrix $\|C^a_b\|$ we can write its determinant as

$$\det\|C^a_b\| \;=\; E_{ab} C^a_1 C^b_2\,. \tag{3.8}$$

Proof.

$$E_{ab} C^a_1 C^b_2 = E_{12} C^1_1 C^2_2 + E_{21} C^2_1 C^1_2 + E_{11} C^1_1 C^1_2 + E_{22} C^2_1 C^2_1 =$$

$$= C^1_1 C^2_1 - C^2_1 C^1_2 + 0 + 0 = \begin{vmatrix} C^1_1 & C^1_2 \\ C^2_1 & C^2_2 \end{vmatrix}\,.$$

Remark. This fact is not accidental. The sourse of the coincidence is that the $\det(C^a_b)$ is the sum of the products of the elements of the matrix C^a_b, taken with the sign corresponding to the parity of the permutation. The internal product with E_{ab} provides correct sign of all these products. In particular, the terms with equal indices a and b vanish, because $E_{aa} \equiv 0$. This exactly coresponds to the fact that the determinant with two equal lines vanish. But, the determinant with two equal raws also vanish, hence we can formulate a following more general theorem:

Theorem 4.

$$\det(C^a_b) \;=\; \frac{1}{2} E_{ab}\, E^{de}\, C^a_d\, C^b_e\,. \tag{3.9}$$

Proof. The difference between (3.8) and (3.9) is that the former admits two expressions $C^a_1 C^b_2$ and $C^a_2 C^b_1$, while the first only the first combination. But, since in both cases the indices run the same values $(a, b) = (1, 2)$, we have, in (3.9), two equal expressions compensated by the $1/2$ factor.

Exercise 16. Check the equality (3.9) by direct calculus.

Theorem 5.

$$E_{ab} C^a_e C^b_d = -E_{ab} C^b_e C^a_d \tag{3.10}$$

Proof.

$$E_{ab} C^a_e C^b_d = E_{ba} C^b_e C^a_d = -E_{ab} C^a_d C^b_e\,.$$

Here in the first equality we have exchanged the names of the umbral indices $a \leftrightarrow b$, and in the second one used antisymmetry $E_{ab} = -E_{ba}$.

Now we consider an arbitrary dimension of space D.

Theorem 6. Consider $a, b = 1, 2, 3, \dots, D$. $\forall$ matrix $\|C^a_b\|$, the determinant is

$$det\|C^a_b\| = E_{a_1 a_2 \dots a_D} \cdot C^{a_1}_1 C^{a_2}_2 \dots C^{a_D}_D\,. \tag{3.11}$$

where (remember we use special Cartesian coordinates here)

$$E_{a_1 a_2 \dots a_D} \qquad \text{is absolutely antisymmetric and} \qquad E_{123\dots D} = 1\,.$$

Proof. The determinant in the left-hand side (*l.h.s.*) of Eq. (3.11) is a sum of $D!$ terms, each of which is a product of elements from different lines and columns, taken with the positive sign for even and with the negative sign for odd parity of the permitations.

The expresion in the right-hand side (*r.h.s.*) of (3.11) also has $D!$ terms. In order to see this, let us notice that the terms with equal indices give zero. Hence, for the first index we have D choices, for the second $D-1$ choices, for the third $D-2$ choices etc. Each non-zero term has one element from each line and one element from each column, because two equal numbers of column give zero when we make a contraction with $E_{a_1 a_2 \ldots a_D}$. Finally, the sign of each term depends on the permutations parity and is identical to the sign of the corresponding term in the *l.h.s.* Finally, the terms on both sides are identical.

Theorem 7. The expression

$$\mathcal{A}_{a_1 \ldots a_D} = E^{b_1 \ldots b_D} A^{b_1}_{a_1} A^{b_2}_{a_2} \ldots A^{b_D}_{a_D} \tag{3.12}$$

is absolutely antisymmetric in the indices $\{b_1, \ldots, b_D\}$ and therefore is proportional to $E_{a_1 \ldots a_D}$ (see Theorem 2).

Proof. For any couple b_l, b_k $(l \neq k)$ the proof performs exactly as in the Theorem 5 and we obtain that the *r.h.s.* is antisymmetric in $\forall$ couple of indices. Due to the Def. 4 this proves the Theorem.

Theorem 8. In an arbitrary dimension D

$$det\|C^a_b\| = \frac{1}{D!} E_{a_1 a_2 \ldots a_D} E^{b_1 b_2 \ldots b_D} \cdot C^{a_1}_{b_1} C^{a_2}_{b_2} \ldots C^{a_D}_{b_D} . \tag{3.13}$$

Proof. The proof can be easily performed by analogy with the Theorems 7 and 4.

Theorem 9. The following relation is valid:

$$E^{a_1 a_2 \ldots a_D} \cdot E_{b_1 b_2 \ldots b_D} = \begin{vmatrix} \delta^{a_1}_{b_1} & \delta^{a_1}_{b_2} & \ldots & \delta^{a_1}_{b} \\ \delta^{a_2}_{b_1} & \delta^{a_2}_{b_2} & \ldots & \delta^{a_2}_{b_D} \\ \ldots & \ldots & \ldots & \ldots \\ \delta^{a_D}_{b_1} & \ldots & \ldots & \delta^{a_D}_{b_D} \end{vmatrix} \tag{3.14}$$

Proof. First of all, both *l.h.s.* and *r.h.s.* of Eq. (3.14) are completely antisymmetric in both $\{a_1, \ldots, a_D\}$ and $\{b_1, \ldots, b_D\}$. To see that this is true for the *r.h.s.*, we just change $a_l \leftrightarrow a_k$ for any $l \neq k$. This is equivalent to the permutation of the two lines in the determinant, $\Rightarrow$ the sign changes. Making such permutations, we can reduce the Eq. (3.14) to the obviosly correct equality

$$1 = E^{12 \ldots D} \cdot E_{12 \ldots D} = det\|\delta^i_j\| = 1 .$$

Consequence: For the special dimension $3D$ we have

$$E^{abc} E_{def} = \begin{vmatrix} \delta^a_d & \delta^b_d & \delta^c_d \\ \delta^a_e & \delta^b_e & \delta^c_e \\ \delta^a_f & \delta^b_f & \delta^c_f \end{vmatrix} . \tag{3.15}$$

Let us make a contraction of the indices c and f in the last equation. Then (remember that $\delta^c_c = 3$ in $3D$)

$$E^{abc} E_{dec} = \begin{vmatrix} \delta^a_d & \delta^b_d & \delta^c_d \\ \delta^a_e & \delta^b_e & \delta^c_e \\ \delta^a_c & \delta^b_c & 3 \end{vmatrix} = \delta^a_d \delta^b_e - \delta^a_e \delta^b_d . \tag{3.16}$$

The last formula is an extremely useful tool for the calculations in analytic geometry and vector calculus. We shall extensively use this formula in what follows.

Let us proceed and contract indices b and e. We obtain

$$E^{abc}E_{dbc} = 3\delta^a_d - \delta^a_d = 2\delta^a_d\,. \tag{3.17}$$

Finally, contracting the last remaining couple of indices, we arrive at

$$E^{abc}E_{abc} = 6\,. \tag{3.18}$$

Exercise 17. (a) Check the last equality by direct calculus.

(b) Without making detailed intermediate calculations, even without using the formula (3.14), prove that for an arbitrary dimension D the similar formula looks like $E^{a_1 a_2 \dots a_D} E_{a_1 a_2 \dots a_D} = D!$.

Exercise 18.* Using the properties of the maximal antisymmetric symbol $E^{a_1 a_2 \dots a_D}$, prove the rule for the product of matrix determinants

$$\det\left(\hat{A}\cdot\hat{B}\right) \,=\, \det\hat{A}\,\cdot\,\det\hat{B}\,.$$

Hint. It is recommended to consider the simplest case $D = 2$ first and only then proceed with an arbitrary dimension D. The proof may be performed in a most compact way by using the representation of the determinants from the Theorem 8 and, of course, Theorem 9.

Finally, let us consider two useful, but a bit more complicated statements concerning inverse matrix and derivative of the determinant.

Lemma. For any non-degenerate $D\times D$ matrix M^a_b with the determinant M, the elements of the inverse matrix $(M^{-1})^b_c$ are given by the expressions

$$(M^{-1})^b_a \,=\, \frac{1}{M\,(D-1)!}\, E_{a\,a_2 \dots a_D}\, E^{b\,b_2 \dots b_D}\, M^{a_2}_{b_2}\, M^{a_3}_{b_3} \dots M^{a_D}_{b_D}\,. \tag{3.19}$$

Proof. In order to prove this statement, let us multiply the expression (3.19) by the element M^c_b. We want to prove that $(M^{-1})^b_a\, M^c_b \,=\, \delta^c_a$, that is

$$\frac{1}{(D-1)!}\, E_{a\,a_2 \dots a_D}\, E^{b\,b_2 \dots b_D}\, M^c_b\, M^{a_2}_{b_2}\, M^{a_3}_{b_3} \dots M^{a_D}_{b_D} \,=\, M\,\delta^c_a\,. \tag{3.20}$$

The first observation is that, due to the Theorem 7, the expression

$$E^{b\,b_2 \dots b_D}\, M^c_b\, M^{a_2}_{b_2}\, M^{a_3}_{b_3} \dots M^{a_D}_{b_D} \tag{3.21}$$

is maximal absolutely antisymmetric object. Therefore it may be different from zero only if the index c is different from all the indices $a_2,\, a_3,\, \dots a_D$. The same is of course valid for the index a in (3.20). Therefore, the expression (3.20) is zero if $a \neq c$. Now we have to prove that the *l.h.s.* of this expression equals M for $a = c$. Consider, e.g., $a = c = 1$. It is easy to see that the coefficient of the term $\|M^1_b\|$ in the *l.h.s* of (3.20) is the $(D-1)\times(D-1)$ determinant of the matrix $\|M\|$

without the first line and without the row b. The sign is positive for even b and negative for odd b. Summing up over the index b, we meet exactly $M = \det \|M\|$. It is easy to see that the same situation holds for all $a = c = 1, 2, ..., D$. The proof is over.

Theorem 10. Consider the non-degenerate $D \times D$ matrix $\|M^a_b\|$ with the elements $M^a_b = M^a_b(\kappa)$ being functions of some parameter κ. In general, the determinant of this matrix $M = \det \|M^a_b\|$ also depends on κ. Suppose all functions $M^a_b(\kappa)$ are differentiable. Then the derivative of the determinant equals

$$\dot{M} = \frac{dM}{d\kappa} = M\,(M^{-1})^b_a\,\frac{d\,M^a_b}{d\kappa}\,, \tag{3.22}$$

where $(M^{-1})^b_a$ are the elements of the matrix inverse to (M^a_b) and the dot some function indicates its derivative with respect to κ.

Proof. Let us use the Theorem 8 as a starting point. Taking derivative of

$$M = \frac{1}{D!}\,E_{a_1 a_2 \dots a_D}\,E^{b_1 b_2 \dots b_D} \cdot M^{a_1}_{b_1}\,M^{a_2}_{b_2}\,\dots\,M^{a_D}_{b_D}\,, \tag{3.23}$$

we arrive at the expression

$$\dot{M} = \frac{1}{D!}E_{a_1 a_2 \dots a_D}E^{b_1 b_2 \dots b_D}\left(\dot{M}^{a_1}_{b_1}M^{a_2}_{b_2}\dots M^{a_D}_{b_D} + M^{a_1}_{b_1}\dot{M}^{a_2}_{b_2}\dots M^{a_D}_{b_D} + ... + M^{a_1}_{b_1}M^{a_2}_{b_2}\dots\dot{M}^{a_2}_{b_2}\right)\,,$$

It is easy to see that the last expression is nothing but the sum of D equal terms, therefore it can be rewrittem as

$$\dot{M} = \dot{M}^a_b\left[\frac{1}{(D-1)!}E_{a\,a_2 \dots a_D} \cdot E^{b\,b_2 \dots b_D} \cdot M^{a_2}_{b_2}\dots M^{a_D}_{b_D}\right]. \tag{3.24}$$

According to the last Lemma, the coefficient of the derivative $\dot{M}^a_b$ in (3.24) is $(M^{-1})^b_a/M$, that completes the proof.

3.3 Applications to Vector Algebra

Consider applications of the formula (3.16) to vector algebra. One has to remember that we are working in a special orthonormal basis $\hat{\mathbf{n}}_a$ (where $a = 1, 2, 3$), corresponding to the Cartesian coordinates $X^a = x, y, z$.

Def. 6. The vector product of the two vectors **a** and **b** in $3D$ is the third vector[1]

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = [\mathbf{a}\,,\,\mathbf{b}]\,, \tag{3.25}$$

which satisfies the following conditions:

i) $\mathbf{c} \perp \mathbf{a}$ and $\mathbf{c} \perp \mathbf{b}$;

ii) $c = |\mathbf{c}| = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin\varphi$, where $\varphi = (\widehat{\mathbf{a}, \mathbf{b}})$ is the angle between the two vectors;

iii) The direction of the vector product **c** is defined by the right-hand rule. That is, looking along **c** (from the beginning to the end) one observes the smallest rotation angle from the direction of **a** to the direction of **b** performed clockwise (see Figure 3.1).

[1]Exactly as in the case of the scalar product of two vectors, it proves useful to introduce two different notations for the vector product of two vectors.

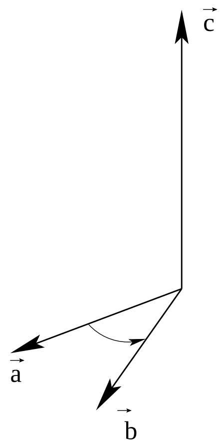

Figure 3.1: Positively oriented basis in $3D$.

The properties of the vector product are discussed in the courses of analytic geometry and we will not repeat them here. The property which is the most important for us is linearity

$$\mathbf{a} \times (\mathbf{b}_1 + \mathbf{b}_2) = \mathbf{a} \times \mathbf{b}_1 + \mathbf{a} \times \mathbf{b}_2\,.$$

Consider all possible products of the elements of the orthonormal basis $\hat{\mathbf{i}} = \hat{\mathbf{n}}_x = \hat{\mathbf{n}}_1$, $\hat{\mathbf{j}} = \hat{\mathbf{n}}_y = \hat{\mathbf{n}}_2$, $\hat{\mathbf{k}} = \hat{\mathbf{n}}_z = \hat{\mathbf{n}}_3$. It is easy to see that these product satisfy the cyclic rule:

$$\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}\,, \qquad \hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}\,, \qquad \hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$$

or, in other notations

$$[\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2] = \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2 = \hat{\mathbf{n}}_3\,, \qquad [\hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3] = \hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_3 = \hat{\mathbf{n}}_1\,,$$

$$[\hat{\mathbf{n}}_3, \hat{\mathbf{n}}_1] = \hat{\mathbf{n}}_3 \times \hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2\,, \tag{3.26}$$

while, of course, $[\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_i] = 0\ \forall i$. It is easy to see that all the three products (3.26 can be written in a compact form using index notations and the maximal anti-symmetric symbol E^{abc}

$$[\hat{\mathbf{n}}_a,\, \hat{\mathbf{n}}_b] \;=\; E_{abc} \cdot \hat{\mathbf{n}}^c\,. \tag{3.27}$$

Let us remind that in the special orthonormal basis $\hat{\mathbf{n}}_a = \hat{\mathbf{n}}^a$ and therefore we do not need to distinguish covariant and contravariant indices. In particular, we can use the equality $E^{abc} = E_{abc}$.

Theorem 11: Consider two vectors $\mathbf{V} = V^a \hat{\mathbf{n}}_a$ and $\mathbf{W} = W^a \hat{\mathbf{n}}_a$. Then

$$[\mathbf{V}, \mathbf{W}] = E_{abc} \cdot V^a \cdot W^b \cdot \hat{\mathbf{n}}^c\,.$$

Proof: Due to the linearity of the vector product $[\mathbf{V}, \mathbf{W}]$ we can write

$$[\mathbf{V}, \mathbf{W}] = \Big[V^a \hat{\mathbf{n}}_a\,,\, W^b \hat{\mathbf{n}}_b\Big] = V^a W^b\, [\hat{\mathbf{n}}_a\,,\, \hat{\mathbf{n}}_b] = V^a W^b E_{abc} \cdot \hat{\mathbf{n}}^c\,.$$

Observation 8. Together with the contraction formula for E_{abc}, the Theorem 10 provides a simple way to solve many problems of vector algebra.

Example 1. Consider the mixed product of the three vectors

$$(\mathbf{U},\mathbf{V},\mathbf{W}) = (\mathbf{U},[\mathbf{V},\mathbf{W}]) = U^a \cdot [\mathbf{V},\mathbf{W}]_a =$$

$$= U^a \,\cdot\, E_{abc}V^b\,W^c = E_{abc}U^a\,V^b\,W^c = \begin{vmatrix} U^1 & U^2 & U^3 \\ V^1 & V^2 & V^3 \\ W^1 & W^2 & W^3 \end{vmatrix} . \tag{3.28}$$

Exercise 19. Using the formulas above, prove the following properties of the mixed product:

(i) Cyclic identity:

$$(\mathbf{U},\mathbf{V},\mathbf{W}) = (\mathbf{W},\mathbf{U},\mathbf{V}) = (\mathbf{V},\mathbf{W},\mathbf{U}) \ .$$

(ii) Antisymmetry:

$$(\mathbf{U},\mathbf{V},\mathbf{W}) = -\,(\mathbf{V},\mathbf{U},\mathbf{W}) = -\,(\mathbf{U},\mathbf{W},\mathbf{V}) \ .$$

Example 2. Consider the vector product of the two vector products $[[\mathbf{U},\mathbf{V}]\times[\mathbf{W},\mathbf{Y}]]$. It proves useful to work with the components, e.g. with the component

$$[[\mathbf{U},\mathbf{V}]\times[\mathbf{W},\mathbf{Y}]]_a \ =$$

$$= E_{abc}\,[\mathbf{U},\mathbf{V}]^b\cdot[\mathbf{W},\mathbf{Y}]^c = E_{abc}\,\cdot\,E^{bde}\,U_d\,V_e\,\cdot\,E^{cfg}\,W_f\,Y_g =$$

$$= -E_{bac}\,E^{bde}\,\cdot\,E^{cfg}\,U_d\,V_e\,W_f\,Y_g = -\left(\delta^d_a\delta^e_c - \delta^e_a\delta^d_c\right)E^{cfg}\cdot\,U_d\,V_e\,W_f\,Y_g =$$

$$= E^{cfg}\,(\,U_c\,V_a\,W_f\,Y_g\,-\,U_a\,V_c\,W_f\,Y_g) = V_a\cdot(\mathbf{U},\mathbf{W},\mathbf{Y}) - U_a\cdot(\mathbf{V},\mathbf{W},\mathbf{Y}) \ .$$

In a vector form we proved that

$$[[\mathbf{U},\mathbf{V}]\times[\mathbf{W},\mathbf{Y}]] = \mathbf{V}\cdot(\mathbf{U},\mathbf{W},\mathbf{Y}) - \mathbf{U}\cdot(\mathbf{V},\mathbf{W},\mathbf{Y}) \ .$$

Warning: Do not try to reproduce this formula in a "usual" way without using the index notations, for it will take too much time.

Exercise 20.

(i) Prove the following dual formula:

$$[[\mathbf{U},\mathbf{V}]\ ,\ [\mathbf{W},\mathbf{Y}]] = \mathbf{W}\cdot(\mathbf{Y},\mathbf{U},\mathbf{V}) - \mathbf{Y}\cdot(\mathbf{W},\mathbf{U},\mathbf{V}) \ .$$

(ii) Using (i), prove the following relation:

$$\mathbf{V}\,(\mathbf{W},\mathbf{Y},\mathbf{U}) - \mathbf{U}\,(\mathbf{W},\mathbf{Y},\mathbf{V}) = \mathbf{W}\,(\mathbf{U},\mathbf{V},\mathbf{Y}) - \mathbf{Y}\,(\mathbf{U},\mathbf{V},\mathbf{W}) \ .$$

(iii) Using index notations and properties of the symbol E_{abc}, prove the identity

$$[\mathbf{U} \times \mathbf{V}] \cdot [\mathbf{W} \times \mathbf{Y}] = (\mathbf{U} \cdot \mathbf{W})(\mathbf{V} \cdot \mathbf{Y}) - (\mathbf{U} \cdot \mathbf{Y})(\mathbf{V} \cdot \mathbf{W}) .$$

(iv) Prove the identity

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] = \mathbf{B}\,(\mathbf{A}, \mathbf{C}) - \mathbf{C}\,(\mathbf{A}, \mathbf{B}) .$$

Exercise 21. Prove the Jacoby identify for the vector product

$$[\mathbf{U}, [\mathbf{V}, \mathbf{W}]] + [\mathbf{W}, [\mathbf{U}, \mathbf{V}]] + [\mathbf{V}, [\mathbf{W}, \mathbf{U}]] = 0 .$$

Def. 7 Till this moment we have considered the maximal antisymmetric symbol E_{abc} only in the special Cartesian coordinates $\{X^a\}$ associated with the orthonormal basis $\{\hat{\mathbf{n}}_a\}$. Let us now construct an absolutely antisymmetric tensor ε_{ijk} such that it coincides with E_{abc} in the special coordinates $\{X^a\}$. The receipt for constructing this tensor is obvios – we have to use the transformation rule for the covariant tensor of 3-d rank and start transformation from the special coordinates X^a. Then, for an arbitrary coordinates x^i we obtain the following components of ε:

$$\varepsilon_{ijk} = \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial X^c}{\partial x^k}\,E_{abc} . \tag{3.29}$$

Exercise 22. Using the definition (3.29) prove that

(i) Prove that if $\varepsilon_{l'm'n'}$, corresponding to the coordinates x'^l is defined through the formula like (3.29), it satisfies the tensor transformation rule

$$\varepsilon'_{lmn} = \frac{\partial x^i}{\partial x'^l}\frac{\partial x^j}{\partial x'^m}\frac{\partial x^k}{\partial x'^n}\,\varepsilon_{ijk} ;$$

(ii) ε_{ijk} is absolutely antisymmetric

$$\varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj} .$$

Hint. Compare this Exercise with the Theorem 7.

According to the last Exercise ε_{ijk} is a maximal absolutely antisymmetric tensor, and then Theorem 2 tells us it has only one relevant component. For example, if we derive the component ε_{123}, we will be able to obtain *all* other components which are either zeros (if any two indices coinside) or can be obtained from ε_{123} via the permutations of the indices

$$\varepsilon_{123} = \varepsilon_{312} = \varepsilon_{231} = -\varepsilon_{213} = -\varepsilon_{321} = -\varepsilon_{123} .$$

Remark. Similar property holds in any dimension D.

Theorem 12. The component ε_{123} is a square root of the metric determinant

$$\varepsilon_{123} = g^{1/2} , \quad \text{where} \quad g = \det \|g_{ij}\| . \tag{3.30}$$

Proof: Let us use the relation (3.29) for ε_{123}

$$\varepsilon_{123} = \frac{\partial X^a}{\partial x^1} \frac{\partial X^b}{\partial x^2} \frac{\partial X^c}{\partial x^3} E_{abc} \,. \tag{3.31}$$

According to the Theorem 6, this is nothing but

$$\varepsilon_{123} = \det \left\| \frac{\partial X^a}{\partial x^i} \right\| . \tag{3.32}$$

On the other hand, we can use Theorem 2.2 from the Chapter 2 and immediately arrive at (3.30).

Exercise 23. Define ε^{ijk} as a tensor which coincides with $E^{abc} = E_{abc}$ in Cartesian coordinates, and prove that

i) ε^{ijk} is absolutely antisymmetric.

ii) $\varepsilon'^{\,lmn} = \frac{\partial x'^l}{\partial x^i} \frac{\partial x'^m}{\partial x^j} \frac{\partial x'^n}{\partial x^k} E^{ijk}\,;$

iii) The unique non-trivial component is

$$\varepsilon^{123} = \frac{1}{\sqrt{g}}\,, \qquad \text{where} \quad g = \det \|g_{ij}\| \,. \tag{3.33}$$

Observation. E_{ijk} and E^{ijk} are examples of quantifies which transform in such a way, that their value, in a given coordinate system, is related with some power of $g = \det \|g_{ij}\|$, where g_{ij} is a metric corresponding to these coordinates. Let us remind that

$$E_{ijk} = \frac{1}{\sqrt{g}}\, \varepsilon_{ijk}\,,$$

where ε_{ijk} is a tensor. Of course, E_{ijk} is not a tensor. The values of a unique non-trivial component for both symbols are in general different

$$E_{123} = 1 \qquad \text{and} \qquad \varepsilon_{123} = \sqrt{g}\,.$$

Despite E_{ijk} is not a tensor, remarkably, we can easily control its transformation from one coordinate system to another. E_{ijk} is a particular example of objects which are called tensor densities.

Def. 8 The quantity $A_{i_1 \dots i_n}{}^{j_1 \dots j_m}$ is called tensor density of the (m,n)-type with the weight r, if the quantity

$$g^{-r/2} \cdot A_{i_1 \dots i_n}{}^{j_1 \dots j_m} \qquad \text{is a tensor}\,. \tag{3.34}$$

It is easy to see that the tensor density transforms as

$$(g')^{-r/2} A_{l'_1 \dots l'_n}{}^{k'_1 \dots k'_m}(x') = g^{-r/2} \frac{\partial x^{i_1}}{\partial x'^{l_1}} \dots \frac{\partial x^{i_n}}{\partial x'^{l_n}} \frac{\partial x^{k'_1}}{\partial x^{j_1}} \dots \frac{\partial x^{k'_m}}{\partial x^{j_m}} A_{i_1 \dots i_n}{}^{j_1 \dots j_m}(x)\,,$$

or, more explicit,

$$A_{l'_1 \dots l'_n}{}^{k'_1 \dots k'_m}(x') = \left(\det \left\| \frac{\partial x^i}{\partial x'^l} \right\| \right)^r \cdot \frac{\partial x^{i_1}}{\partial x'^{l_1}} \dots \frac{\partial x^{i_n}}{\partial x'^{l_n}} \frac{\partial x^{k'_1}}{\partial x^{j_1}} \dots \frac{\partial x^{k'_m}}{\partial x^{j_m}} A_{i_1 \dots i_n}{}^{j_1 \dots j_m}(x) \,. \tag{3.35}$$

Remark. All operations over densities can be defined in a obvious way. We shall leave the construction of these definitions as Exercises for the reader.

Exercise 24. Prove that a product of the tensor density of the weight r and another tensor density of the weight s is a tensor density of the weight $r + s$.

Exercise 25. *(i)* Construct an example of the covariant tensor of the 5-th rank, which symmetric in the two first indices and absolutely antisymmetric in other three.

(ii) Complete the same task using, as a building blocks, only the maximally antisymmetric tensor ε_{ijk} and the metric tensor g_{ij}.

Chapter 4

Curvilinear coordinates (local coordinate tranformations)

4.1 Curvilinear coordinates and change of basis

The purpose of this chapter is to consider an arbitrary change of coordinates

$$x'^{\alpha} = x'^{\alpha}\left(x^{\mu}\right) , \tag{4.1}$$

where $x'^{\alpha}(x)$ are not necessary linear functions as before. In is supposed that $x'^{\alpha}(x)$ are smooth functions (this means that all partial derivatives $\partial x'^{\alpha}/\partial x^{\mu}$ are continuous functions) maybe except some isolated points.

It is easy to see that all the definitions we have introduced before can be easily generalized for this, more general, case. Say, the scalar field may be defined as

$$\varphi'\left(x'\right) = \varphi\left(x\right) . \tag{4.2}$$

Also, vectors (contra- and covariant) are defined as a quantities which transform according to the rules

$$a'^{\,i}\left(x'\right) = \frac{\partial x'^{i}}{\partial x^{j}}\, a^{j}\left(x\right) \tag{4.3}$$

for the contravariant and

$$b'_{l}\left(x'\right) = \frac{\partial x^{k}}{\partial x'^{l}}\, b_{k}\left(x\right) \tag{4.4}$$

for the covariant cases. The same scheme works for tensors, for example we define the $(1,1)$-type tensor as an object, whose components are transforming according to

$$A^{i'}_{j'}\left(x'\right) = \frac{\partial x^{i'}}{\partial x^{k}}\,\frac{\partial x^{l}}{\partial x'^{j}}\, A^{k}_{l}\left(x\right) \qquad\qquad \text{etc.} \tag{4.5}$$

The metric tensor g_{ij} is defined as

$$g_{ij}\left(x\right) \;=\; \frac{\partial x^{a}}{\partial x^{i}}\,\frac{\partial x^{b}}{\partial x^{j}}\, g_{ab} \,, \tag{4.6}$$

where $g_{ab} = \delta_{ab}$ is a metric in Cartesian coordinates. Also, the inverse metric is

$$g^{ij} = \frac{\partial x^{i}}{\partial x^{a}}\,\frac{\partial x^{j}}{\partial x^{l}}\, \delta^{ab} \,,$$

where X^a (as before) are Cartesian coordinates corresponding to the orthonormal basis $\hat{\mathbf{n}}_a$. Of course, the basic vectors $\hat{\mathbf{n}}_a$ are the same in all points of the space.

It is so easy to generalize notion of tensors and algebraic operatons with them, because all these operations are defined in the same point of the space. Thus the only difference between general coordinate transformation $x'^{\alpha} = x'^{\alpha}(x)$ and the special one $x'^{\alpha} = \wedge^{\alpha'}_{\beta} x^{\beta} + B^{\alpha'}$ with $\wedge^{\alpha'}_{\beta} = \text{const}$ and $B^{\alpha'} = \text{const}$, is that, in the general case, the transition coefficients $\partial x^i / \partial x'^j$ are not necessary constants.

One of the important consequences is that the metric tensor g_{ij} also depends on the point. Another important fact is that antisymmetric tensor ε^{ijk} also depends on the coordinates

$$\varepsilon^{ijk} = \frac{\partial x^i}{\partial X^a} \frac{\partial x^j}{\partial X^b} \frac{\partial x^k}{\partial X^c} E^{abc} \,.$$

Using $E_{123} = 1$ and according to the Theorem 3.10 we have

$$\varepsilon_{123} = \sqrt{g} \qquad \text{and} \qquad \varepsilon^{123} = \frac{1}{\sqrt{g}}\,, \qquad \text{where} \qquad g = g(x) = \det \|g_{ij}(x)\| \,. \tag{4.7}$$

Before turning to the applications and particular cases, let us consider the definition of the basis vectors corresponding to the curvilinear coordinates. It is intuitively clear that these vector must be different in different points. Let us suppose that the vectors of the point-dependent basis $\mathbf{e}_i(x)$ are such that the vector $\mathbf{a} = a^i \mathbf{e}_i(x)$ is coordinate-independent geometric object. As far as the rule for the vector transformation is (4.3), the transformation for the basic vectors must have the form

$$\mathbf{e}'_i = \frac{\partial x^l}{\partial x'^i}\, \mathbf{e}_l \,. \tag{4.8}$$

In particular, we obtain the following relation between the point-dependent and permanent orthonormal basis

$$\mathbf{e}_i = \frac{\partial X^a}{\partial x^i}\, \hat{\mathbf{n}}_a \,. \tag{4.9}$$

Exercise 1. Derive the relation (4.8) starting from (4.3).

4.2 Polar coordinates on the plane

As an example of local coordinate transformations, let us consider the polar coordinates on the $2D$ plane. The polar coordinates r, φ are defined as follows

$$x = r\,\cos\varphi\,, \qquad\qquad y = r\,\sin\varphi\,, \tag{4.10}$$

where x, y are Cartesian coordinates corresponding to the basic vectors (orts) $\hat{\mathbf{n}}_x = \hat{\mathbf{i}}$ and $\hat{\mathbf{n}}_y = \hat{\mathbf{j}}$. Our purpose is to learn how to transform an arbitrary tensor to polar coordinates.

Let us denote, using our standard notations, $X^a = (x, y)$ and $x^i = (r, \varphi)$. The first step is to derive the two mutually inverse matrices

$$\left(\frac{\partial X^a}{\partial x^i}\right) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\,r\,\sin\varphi & r\,\cos\varphi \end{pmatrix} \qquad \text{and} \qquad \left(\frac{\partial x^j}{\partial X^b}\right) = \begin{pmatrix} \cos\varphi & -1/r\,\sin\varphi \\ \sin\varphi & 1/r\,\cos\varphi \end{pmatrix}. \tag{4.11}$$

Exercise 1. Derive the first matrix by taking derivatives of (4.10) and the second one by inverting the first. Interpret both transformations on the circumference $r = 1$ as rotations.

Now we are in a position to calculate the components of an arbitrary tensor in polar coordinates. Let us start from the basic vectors. Using the general formula (4.9) in the specific case of polar coordinates, we obtain

$$\mathbf{e}_r = \frac{\partial x}{\partial r}\,\hat{\mathbf{n}}_x + \frac{\partial y}{\partial r}\,\hat{\mathbf{n}}_y = \hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi\,,$$

$$\mathbf{e}_\varphi = \frac{\partial x}{\partial \varphi}\,\hat{\mathbf{n}}_x + \frac{\partial y}{\partial \varphi}\,\hat{\mathbf{n}}_y = -\,\hat{\mathbf{i}}\,r\sin\varphi + \hat{\mathbf{j}}\,r\cos\varphi\,. \tag{4.12}$$

Let us analyze these formulas in some details. The first observation is that the basic vectors $\mathbf{e}_r$ and $\mathbf{e}_\varphi$ are ortogonal $\mathbf{e}_r\cdot\mathbf{e}_\varphi = 0$. As a result, the metric in polar coordinates is diagonal

$$g_{ij} = \begin{pmatrix} g_{rr} & g_{r\varphi} \\ g_{\varphi r} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_r\cdot\mathbf{e}_r & \mathbf{e}_r\cdot\mathbf{e}_\varphi \\ \mathbf{e}_\varphi\cdot\mathbf{e}_r & \mathbf{e}_\varphi\cdot\mathbf{e}_\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{4.13}$$

Furthermore, we can see that the vector $\mathbf{e}_r$ is nothing but the normalized radius-vector of the point

$$\mathbf{e}_r = \frac{\mathbf{r}}{r}\,. \tag{4.14}$$

Correspondingly, the basic vector $\mathbf{e}_\varphi$ is orthogonal to the radius-vector $\mathbf{r}$.

The second observation is that only the basic vector $\mathbf{e}_r$ is normalized $|\mathbf{e}_r| = 1$, while the basic vector $\hat{\mathbf{e}}_\varphi$ is not $|\hat{\mathbf{e}}_\varphi| = r$ (if we are not on the circumference $x^2 + y^2 = r^2 = 1$). One can, of course, introduce another, normalized basis

$$\hat{\mathbf{n}}_r = \mathbf{e}_r\,, \qquad \hat{\mathbf{n}}_\varphi = \frac{1}{r}\,\mathbf{e}_\varphi\,. \tag{4.15}$$

Let us notice that the normalized basis $\hat{\mathbf{n}}_r$, $\hat{\mathbf{n}}_\varphi$ is not what is called coordinate basis. This means that there are no coordinates which correspond to this basis in the sense of the formula (4.8). The basis corresponding to the polar coordinates (r,φ) is $(\mathbf{e}_r,\mathbf{e}_\varphi)$ and not $(\hat{\mathbf{n}}_r,\hat{\mathbf{n}}_\varphi)$. Now, we can expand an arbitrary vector $\mathbf{A}$ using one or another basis

$$\mathbf{A} = \tilde{A}^r\,\mathbf{e}_r + \tilde{A}^\varphi\,\mathbf{e}_\varphi = A^r\hat{\mathbf{n}}_r + A^\varphi\hat{\mathbf{n}}_\varphi\,, \tag{4.16}$$

where

$$A^r = \tilde{A}^r \qquad \text{and} \qquad A^\varphi = r\,\tilde{A}^\varphi. \tag{4.17}$$

In what follows we shall always follow this pattern: if the coordinate basis (like $\mathbf{e}_r,\mathbf{e}_\varphi$) is not normalized, we shall mark the components of the vector in this basis by tilde. The simpler notations without tilde are always reserved for the components of the vector in the normalized basis.

Why do we need to give this preference to the normalized basis? The reason is that the normalized basis is much more useful for the applications, e.g. in physics. Let us remember that the physical quantities usually have the specific dimensionality. Using the normalized basis means that all components of the vector, even if we are using the curvilinear coordinates, have the same dimension. At the same time, the dimensions of the distinct vector components in a non-normalized basis may be different. For example, if we consider the vector of velocity $\mathbf{v}$, its magnitude is measured in centimeters per seconds. Using our notations, in the normalized basis both v^r and v^φ are also measured in centimeters per seconds. But, in the original coordinate basis $(\mathbf{e}_r, \mathbf{e}_\varphi)$ the component $\tilde{v}^\varphi$ is measured in $1/sec$, while the unit for $\tilde{v}^r$ is cm/sec. As we see, in this coordinate basis the different components of the same vector have different dimensionalities, while in the normalized basis they are equal. It is not necessary to be physicist to understand which basis is more useful. However, we can not disregard completely the coordinate basis, because it enables us to perform practical transformations from one coordinate system to another in a general form. Of course, for the orthogonal basis one can easily perform the transformations via rotations, but for the more complicated cases (say, elliptic or hyperbolic coordinates) we are forced to use the coordinate basis. Below we shall organize all the calculations in such a way that first one has to perform the tranformation between the coordinate basises and then make an extra transformation to the normalized basis.

Exercise 2. For the polar coordinates on the $2D$ plane find metric tensor by tensor transformation of the metric in Cartesian coordinates.

Solution: In cartesian coordinates

$$g_{ab} = \delta_{ab}\,, \quad \text{that is} \quad g_{xx} = 1 = g_{yy}\,, \qquad g_{xy} = g_{yx} = 0\,.$$

Let us apply the tensor transformation rule:

$$g_{\varphi\varphi} = \frac{\partial x}{\partial \varphi}\frac{\partial x}{\partial \varphi}\, g_{xx} + \frac{\partial x}{\partial \varphi}\frac{\partial y}{\partial \varphi}\, g_{xy} + \frac{\partial y}{\partial \varphi}\frac{\partial x}{\partial \varphi}\, g_{yx} + \frac{\partial y}{\partial \varphi}\frac{\partial y}{\partial \varphi}\, g_{yy} =$$

$$= \left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2 = r^2 \sin^2\varphi + r^2\cos^2\varphi = r^2\,.$$

Similarly we find, using $g_{xy} = g_{yx} = 0$,

$$g_{\varphi r} = \frac{\partial x}{\partial \varphi}\frac{\partial x}{\partial r}\, g_{xx} + \frac{\partial y}{\partial \varphi}\frac{\partial y}{\partial r}\, g_{yy} = (-r\sin\varphi)\cdot\cos\varphi + r\cos\varphi\cdot\sin\varphi = 0 = g_{r\varphi}$$

and

$$g_{rr} = \frac{\partial x}{\partial r}\frac{\partial x}{\partial r}\, g_{xx} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial r}\, g_{yy} = \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 = \cos^2\varphi + \sin^2\varphi = 1\,.$$

After all, the metric is

$$\begin{pmatrix} g_{rr} & g_{r\varphi} \\ g_{\varphi r} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix},$$

that is exactly the same which we obtained earlier in (4.13).

This derivation of the metric has a simple geometric interpretation. Consider two points which have infinitesimally close φ and r (see the Figure 4.1). The distance between these two points dl is defined by the relation

$$ds^2 = dx^2 + dy^2 = g_{ab}\, dX^a\, dX^b = g_{ij}\, dx^i\, dx^j = dr^2 + r^2 d\varphi^2 = g_{rr} dr dr + g_{\varphi\varphi} d\varphi d\varphi + 2g_{\varphi r} d\varphi dr\,.$$

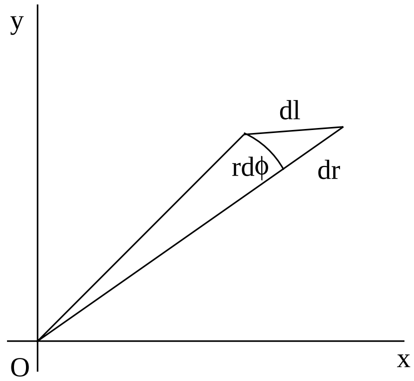

Figure 4.1: Line element in $2D$ polar coordinates, corresponding to the infinitesimal $d\phi$ and dr.

Hence, the tensor form of the transformation of the metric correspond to the coordinate-independent distance between two infinitesimally close points.

Let us consider the motion of the particle on the plane in polar coordinates. The position of the particle in the instant of time t is given by the radius-vector $\mathbf{r} = \mathbf{r}(t)$, its velocity by $\mathbf{v} = \dot{\mathbf{r}}$ and acceleration by $\mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{r}}$. Our purpose is to find the components of these three vectors in polar coordinates. We can work directly in the normalized basis. Using (4.12), we arrive at

$$\hat{\mathbf{n}}_r = \hat{\mathbf{i}} \cos\varphi + \hat{\mathbf{j}} \sin\varphi \,, \qquad \hat{\mathbf{n}}_\varphi = -\hat{\mathbf{i}} \sin\varphi + \hat{\mathbf{j}} \cos\varphi \,. \tag{4.18}$$

As we have already mentioned, the transformation from one orthonormal basis $(\hat{\mathbf{i}}, \hat{\mathbf{j}})$ to another one $(\hat{\mathbf{n}}_r, \hat{\mathbf{n}}_\varphi)$ is rotation, and the angle of rotation φ depends on time. Taking the first and second derivatives, we obtain

$$\dot{\hat{\mathbf{n}}}_r = \dot{\varphi}\,(-\hat{\mathbf{i}} \sin\varphi + \hat{\mathbf{j}} \cos\varphi) = \dot{\varphi}\,\hat{\mathbf{n}}_\varphi \,,$$

$$\dot{\hat{\mathbf{n}}}_\varphi = \dot{\varphi}\,(-\hat{\mathbf{i}} \cos\varphi - \hat{\mathbf{j}} \sin\varphi) = -\dot{\varphi}\,\hat{\mathbf{n}}_r \tag{4.19}$$

and

$$\ddot{\hat{\mathbf{n}}}_r = \ddot{\varphi}\,\hat{\mathbf{n}}_\varphi - \dot{\varphi}^2\,\hat{\mathbf{n}}_r \,, \qquad \ddot{\hat{\mathbf{n}}}_\varphi = -\ddot{\varphi}\,\hat{\mathbf{n}}_r - \dot{\varphi}^2\,\hat{\mathbf{n}}_\varphi \,, \tag{4.20}$$

where $\dot{\varphi}$ and $\ddot{\varphi}$ are angular velocity and acceleration.

The above formulas enable us to derive the velocity and acceleration of the particle. In the Cartesian coordinates, of course,

$$\mathbf{r} = x\,\hat{\mathbf{i}} + y\,\hat{\mathbf{j}} \,, \qquad \mathbf{v} = \dot{x}\,\hat{\mathbf{i}} + \dot{y}\,\hat{\mathbf{j}} \,, \qquad \mathbf{a} = \ddot{x}\,\hat{\mathbf{i}} + \ddot{y}\,\hat{\mathbf{j}} \,. \tag{4.21}$$

In polar coordinates, using the formulas (4.20), we get

$$\mathbf{r} = r\,\hat{\mathbf{n}}_r \,, \qquad \mathbf{v} = \dot{r}\,\hat{\mathbf{n}}_r + r\,\dot{\varphi}\,\hat{\mathbf{n}}_\varphi \,, \qquad \mathbf{a} = (\ddot{r} - r\,\dot{\varphi}^2)\,\hat{\mathbf{n}}_r + (r\ddot{\varphi} + 2\dot{r}\,\dot{\varphi})\,\hat{\mathbf{n}}_\varphi \,. \tag{4.22}$$

Exercise 3. Suggest physical interpretation of the last two formulas.

4.3 Cylindric and spherical coordinates

Def. 1 Cylindric coordinates (r, φ, z) in $3D$ are defined by the relations

$$x = r\cos\varphi\,, \qquad y = r\sin\varphi\,, \qquad z = z\,, \tag{4.23}$$

$$\text{where} \qquad 0 \leq r < \infty\,, \qquad 0 \leq \varphi < 2\pi \quad \text{and} \quad -\infty < z < \infty\,.$$

Exercise 4. (a) Derive explicit expressions for the basic vectors for the case of cylindric coordinates in $3D$; (b) Using the results of the point (a), calculate all metric components. (c) Using the results of the point (a), construct the normalized non-coordinate basis. (d) Compare all previous results to the ones for the polar coordinates in $2D$.

Answers:

$$\mathbf{e}_r = \hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi\,, \qquad \mathbf{e}_\varphi = -\hat{\mathbf{i}}\,r\sin\varphi + \hat{\mathbf{j}}\,r\cos\varphi\,, \qquad \mathbf{e}_z = \hat{\mathbf{k}}\,. \tag{4.24}$$

$$g_{ij} = \begin{pmatrix} g_{rr} & g_{r\varphi} & g_{rz} \\ g_{\varphi r} & g_{\varphi\varphi} & g_{\varphi z} \\ g_{zr} & g_{z\varphi} & g_{zz} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,. \tag{4.25}$$

$$\hat{\mathbf{n}}_r = \hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi\,, \qquad \hat{\mathbf{n}}_\varphi = -\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi\,, \qquad \hat{\mathbf{n}}_z = \hat{\mathbf{k}}\,. \tag{4.26}$$

Def. 2 Spherical coordinates are defined by the relations

$$x = r\cos\varphi\sin\chi\,, \qquad y = r\sin\varphi\sin\chi\,, \qquad z = r\cos\chi\,, \tag{4.27}$$

$$\text{where} \qquad 0 \leq r < \infty\,, \qquad 0 \leq \varphi < 2\pi \quad \text{and} \quad 0 \leq \chi \leq \pi\,.$$

Exercise 5. Repeat the whole program of the Exercise 4 for the case of spherical coordinates.

Answers:

$$\mathbf{e}_r = \hat{\mathbf{i}}\cos\varphi\sin\chi + \hat{\mathbf{j}}\sin\varphi\sin\chi + \hat{\mathbf{k}}\cos\chi\,,$$
$$\mathbf{e}_\varphi = -\hat{\mathbf{i}}\,r\sin\varphi\sin\chi + \hat{\mathbf{j}}\,r\cos\varphi\sin\chi\,,$$
$$\mathbf{e}_\chi = \hat{\mathbf{i}}\,r\cos\varphi\cos\chi + \hat{\mathbf{j}}\,r\sin\varphi\cos\chi - \hat{\mathbf{k}}\,r\sin\chi\,,$$

$$g_{ij} = \begin{pmatrix} g_{rr} & g_{r\varphi} & g_{r\chi} \\ g_{\varphi r} & g_{\varphi\varphi} & g_{\varphi\chi} \\ g_{\chi r} & g_{\chi\varphi} & g_{\chi\chi} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2\sin\chi & 0 \\ 0 & 0 & r^2 \end{pmatrix}\,. \tag{4.28}$$

$$\begin{aligned} \hat{\mathbf{n}}_r &= \cos\varphi\sin\chi\,\hat{\mathbf{i}} + \sin\varphi\sin\chi\,\hat{\mathbf{j}} + \cos\chi\,\hat{\mathbf{k}}\,, \\ \hat{\mathbf{n}}_\varphi &= -\sin\varphi\,\hat{\mathbf{i}} + \cos\varphi\,\hat{\mathbf{j}}\,, \\ \hat{\mathbf{n}}_\chi &= \cos\varphi\cos\chi\,\hat{\mathbf{i}} + \sin\varphi\cos\chi\,\hat{\mathbf{j}} - \cos\chi\,\hat{\mathbf{k}}\,. \end{aligned} \tag{4.29}$$

Exercise 6. Repeat the calculations of Exercise 2 for the case of cylindric and spherical coordinates.

Exercise 7. Derive the expressions for the velocity and acceleration in $3D$ (analogs of the equations (4.19), (4.20), (4.22)), for the case of the cylindric coordinates.

Hint. In Cartesian coordinates in $3D$ the corresponding expressions have the form

$$\mathbf{r} = x\,\hat{\mathbf{i}} + y\,\hat{\mathbf{j}} + z\,\hat{\mathbf{k}}\,, \qquad \mathbf{v} = \dot{x}\,\hat{\mathbf{i}} + \dot{y}\,\hat{\mathbf{j}} + \dot{z}\,\hat{\mathbf{k}}\,, \qquad \mathbf{a} = \ddot{x}\,\hat{\mathbf{i}} + \ddot{y}\,\hat{\mathbf{j}} + \ddot{z}\,\hat{\mathbf{k}}\,. \tag{4.30}$$

Let us derive the expressions for the velocity and acceleration in $3D$ (analogs of the equations (4.19), (4.20), (4.22)), for the case of the spherical coordinates. This is a cumbersome calculation, and we shall present it in details. The starting point is the formulas (4.29). Now we need to derive the first and second time derivatives for the vectors of this orthonormal basis. A simple calculus gives the following result for the first derivatives:

$$\dot{\mathbf{n}}_r = \dot{\varphi}\,\sin\chi\Big(-\,\hat{\mathbf{i}}\sin\varphi\, + \hat{\mathbf{j}}\,\cos\varphi\Big) + \dot{\chi}\Big(\hat{\mathbf{i}}\,\cos\varphi\,\cos\chi + \hat{\mathbf{j}}\,\sin\varphi\,\cos\chi - \hat{\mathbf{k}}\,\sin\chi\Big)\,,$$

$$\dot{\mathbf{n}}_\varphi = -\,\dot{\varphi}\,\Big(\hat{\mathbf{i}}\,\cos\varphi + \hat{\mathbf{j}}\,\sin\varphi\Big)\,, \tag{4.31}$$

$$\dot{\mathbf{n}}_\chi = \dot{\varphi}\,\cos\chi\,\Big(-\,\hat{\mathbf{i}}\,\sin\varphi + \hat{\mathbf{j}}\,\cos\varphi\Big) + \dot{\chi}\,\Big(-\,\hat{\mathbf{i}}\,\sin\chi\,\cos\varphi + \hat{\mathbf{j}}\,\sin\chi\,\sin\varphi - \hat{\mathbf{k}}\,\cos\chi\Big)\,,$$

where, as before,

$$\dot{\mathbf{n}}_r = \frac{d\,\hat{\mathbf{n}}_r}{dt}\,, \qquad \dot{\mathbf{n}}_\varphi = \frac{d\,\hat{\mathbf{n}}_\varphi}{dt}\,, \qquad \dot{\mathbf{n}}_\chi = \frac{d\,\hat{\mathbf{n}}_\chi}{dt}\,.$$

It is easy to understand that this is not what we really need, because the derivatives (4.31) are given in the constant Cartesian basis and not in the $\hat{\mathbf{n}}_r$, $\hat{\mathbf{n}}_\varphi$, $\hat{\mathbf{n}}_\chi$ basis. Therefore we need to perform an inverse transformation from the basis $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ to the basis $\hat{\mathbf{n}}_r$, $\hat{\mathbf{n}}_\varphi$, $\hat{\mathbf{n}}_\chi$. The general transformation formula for the coordinate basis looks like

$$\mathbf{e}'_i = \frac{\partial x^k}{\partial x'^i}\,\mathbf{e}_k$$

and when we apply it we get

$$\mathbf{i} = \tfrac{\partial r}{\partial x}\mathbf{e}_r + \tfrac{\partial \varphi}{\partial x}\mathbf{e}_\varphi + \tfrac{\partial \chi}{\partial x}\mathbf{e}_\chi\,, \quad \mathbf{j} = \tfrac{\partial r}{\partial y}\mathbf{e}_r + \tfrac{\partial \varphi}{\partial y}\mathbf{e}_\varphi + \tfrac{\partial \chi}{\partial y}\mathbf{e}_\chi\,, \quad \mathbf{k} = \tfrac{\partial r}{\partial z}\mathbf{e}_r + \tfrac{\partial \varphi}{\partial z}\mathbf{e}_\varphi + \tfrac{\partial \chi}{\partial z}\mathbf{e}_\chi\,. \tag{4.32}$$

It is clear that we need all nine partial derivatives which appear in the last relations. These derivatives can be found from the observation that

$$\begin{pmatrix} x'_r & x'_\varphi & x'_\chi \\ y'_r & y'_\varphi & y'_\chi \\ z'_r & z'_\varphi & z'_\chi \end{pmatrix} \times \begin{pmatrix} r'_x & r'_y & r'_z \\ \varphi'_x & \varphi'_y & \varphi'_z \\ \chi'_x & \chi'_y & \chi'_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,. \tag{4.33}$$

Inverting the first matrix, we obtain (a simple method of inverting this matrix is discussed in

Chapter 7)

$$r'_x = \sin\chi\cos\varphi\,, \qquad \varphi'_x = -\frac{\sin\varphi}{r\,\sin\chi}\,, \qquad \chi'_x = \frac{\cos\varphi\,\cos\chi}{r}$$

$$r'_y = \sin\chi\sin\varphi\,, \qquad \varphi'_y = \frac{\cos\varphi}{r\,\sin\chi}\,, \qquad \chi'_y = \frac{\sin\varphi\,\cos\chi}{r}\,,$$

$$r'_z = \cos\chi\,, \qquad \varphi'_z = 0 \qquad \chi'_z = -\frac{\sin\chi}{r}\,. \tag{4.34}$$

After replacing (4.34) into (4.32) we arrive at the relations

$$\begin{aligned} \hat{\mathbf{i}} &= \hat{\mathbf{n}}_r \cos\varphi\sin\chi - \hat{\mathbf{n}}_\varphi \sin\varphi + \hat{\mathbf{n}}_\chi \cos\varphi\cos\chi\,, \\ \hat{\mathbf{j}} &= \hat{\mathbf{n}}_r \sin\varphi\sin\chi + \hat{\mathbf{n}}_\varphi \cos\varphi + \hat{\mathbf{n}}_\chi \sin\varphi\cos\chi\,, \\ \hat{\mathbf{k}} &= \hat{\mathbf{n}}_r \cos\chi - \hat{\mathbf{n}}_\chi \sin\chi\,. \end{aligned} \tag{4.35}$$

inverse to (4.29). Using this relation, we can derive the first derivative of the vectors $\hat{\mathbf{n}}_r$, $\hat{\mathbf{n}}_\varphi$, $\hat{\mathbf{n}}_\chi$ in a proper form

$$\begin{aligned} \dot{\hat{\mathbf{n}}}_r &= \dot\varphi\,\sin\chi\,\hat{\mathbf{n}}_\varphi + \dot\chi\,\hat{\mathbf{n}}_\chi\,, \\ \dot{\hat{\mathbf{n}}}_\varphi &= -\dot\varphi\left(\sin\chi\,\hat{\mathbf{n}}_r + \cos\chi\,\hat{\mathbf{n}}_\chi\right), \\ \dot{\hat{\mathbf{n}}}_\chi &= \dot\varphi\,\cos\chi\,\hat{\mathbf{n}}_\varphi - \dot\chi\,\hat{\mathbf{n}}_r\,. \end{aligned} \tag{4.36}$$

Now we are in a position to derive the particle velocity

$$\mathbf{v} = \dot{\mathbf{r}} = \dot r\,\hat{\mathbf{n}}_r + r\dot\chi\,\hat{\mathbf{n}}_\chi + r\dot\varphi\,\sin\chi\,\hat{\mathbf{n}}_\varphi \tag{4.37}$$

and acceleration

$$\begin{aligned} \mathbf{a} = \ddot{\mathbf{r}} = \left(\ddot r - r\dot\chi^2 - r\dot\varphi^2\,\sin^2\chi\right)\hat{\mathbf{n}}_r + \left(2r\dot\varphi\dot\chi\cos\chi + 2\dot\varphi\dot r\sin\chi + r\ddot\varphi\sin\chi\right)\hat{\mathbf{n}}_\varphi + \\ + \left(2\dot r\dot\chi + r\ddot\chi - r\dot\varphi^2\sin\chi\cos\chi\right)\hat{\mathbf{n}}_\chi\,. \end{aligned} \tag{4.38}$$

Exercise 8. Find the metric for the case of the hyperbolic coordinates in $2D$

$$x = r\cdot\cosh\varphi\,, \qquad y = r\cdot\sinh\varphi\,. \tag{4.39}$$

Reminder. The hyperbolic functions are defined as follows:

$$\cosh\varphi = \frac{1}{2}\,\left(e^{\varphi} + e^{-\varphi}\right)\,; \qquad \sinh\varphi = \frac{1}{2}\,\left(e^{\varphi} - e^{-\varphi}\right)\,.$$

Properties:

$$\cosh^2\varphi - \sinh^2\varphi = 1\,; \qquad \cosh'\varphi = \sinh\varphi\,; \qquad \sinh'\varphi = \cosh\varphi\,.$$

Answer:

$$g_{rr} = \cosh 2\varphi\,, \quad g_{\varphi\varphi} = r^2 \cosh 2\varphi\,, \quad g_{r\varphi} = r \sinh 2\varphi\,.$$

Exercise 9. Find the components ε_{123} and ε^{123} for the cases of cylindric and spherical coordinates.

Hint. It is very easy Exercise. You only need to derive the metric determinants

$$g = \det \|g_{ij}\| \quad \text{and} \quad \det \left\|g^{ij}\right\| = 1/g$$

for both cases and use the formulas (3.30) and (3.33).

Chapter 5

Derivatives of tensors, covariant derivatives

In the previous Chapters we learned how to use the tensor transformation rule. Consider a tensor in some (maybe curvilinear) coordinates $\{x^i\}$. If one knows the components of the tensor in these coordinates and the relation between them and some new coordinates $x'^i = x'^i\left(x^j\right)$, it is sufficient to derive the components of the tensor in new coordinates. As we already know, the tensor transformation rule means that the tensor components transform while the tensor itself corresponds to a (geometric or physical) quantity which is coordinate-independent. When we change the coordinate system, our point of view changes and hence we may see something different despite the object is the very same in both cases.

Why do we need to have a control over the change of coordinates? One of the reasons is that the physicallaws must be formulated in such a way that they remain correct in any coordinates.

However, here we meet a serios problem. The physical lows include not only the quantities themselves, but also, quite often, their derivatives. How can we treat derivatives in the tensor framework?

The two important questions are:

1) Is the partial derivative of a tensor always producing a tensor?

2) If not, how one can define a "correct" derivative of a tensor, so that differentiation produces a new tensor?

Let us start from a scalar field φ. Consider its partial derivative using the chain rule

$$\partial_i \varphi \,=\, \varphi_{,i} \,=\, \frac{\partial \varphi}{\partial x^i}\,. \tag{5.1}$$

Notice that here we used three different notations for the same partial derivative. In the transformed coordinates $x'^i = x'^i\left(x^j\right)$ we obtain, using the chain rule and $\varphi'(x') = \varphi(x)$

$$\partial_{i'} \varphi' \,=\, \frac{\partial \varphi'}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i}\,\frac{\partial \varphi}{\partial x^j} = \frac{\partial x^j}{\partial x'^i}\,\partial_j \varphi\,.$$

The last formula shows that the partial derivative of a scalar field $\varphi_i \,=\, \partial_i \varphi$ transforms as a covariant vector field. Of course, there is no need to modify the partial derivative in this case.

Def. 1. The vector (5.1) is called gradient of a scalar field $\varphi\,(x)$, and is denoted

$$\mathrm{grad}\,\varphi = \mathbf{e}^i\,\varphi_{,i} \qquad \text{or} \qquad \nabla\varphi = \,\mathrm{grad}\,\varphi\,. \tag{5.2}$$

In the last formula we have introduced a new object: vector differential operator ∇

$$\nabla = \mathbf{e}^i \frac{\partial}{\partial x^i} \,,$$

which is also called the Hamilton operator. When acting on scalar, ∇ produces a covariant vector which is called gradient. But, as we shall see later on, the operator ∇ may also act on vectors or tensors.

The next step is to consider a partial derivative of a covariant vector $b_i\,(x)$. This partial derivative $\partial_j b_i$ looks like a second rank $(0,2)$-type tensor. But it is easy to see that it is not a tensor! Let us make a corresponding transformation

$$\partial_{j'} b_{i'} = \frac{\partial}{\partial x'^j}\, b_{i'} = \frac{\partial}{\partial x'^j}\left(\frac{\partial x^k}{\partial x'^i}\, b_k\right) = \frac{\partial x^k}{\partial x'^i}\,\frac{\partial x^l}{\partial x'^j}\cdot \partial_l b_k + \frac{\partial^2 x^k}{\partial x'^j \partial x'^i}\cdot b_k \,. \tag{5.3}$$

Let us comment on a technical subtlety in the calculation (5.3). The parenthesis contains a product of the two expressions which are functions of different coordinates x'^j and x^i. When the partial derivative $\partial/\partial x'^j$ acts on the function of the coordinates x^i (like $b_k = b_k(x)$), it must be applied following the chain rule

$$\frac{\partial}{\partial x'^j} = \frac{\partial x^l}{\partial x'^j}\,\frac{\partial}{\partial x^l}\,.$$

The relation (5.3) provides a very important information. Since the last term in the *r.h.s.* of this equation is nonzero, derivative of a covariant vector is not a tensor. Indeed, the last term in (5.3) may be equal to zero and then the partial derivative of a vector is tensor. This is the case for the global coordinate transformation, then the matrix $\partial x^k/\partial x'^i$ is constant. Then its derivatives are zeros and the tensor rule in (5.3) really holds. But, for the general case of the curvilinear coordinates (e.g. polar in $2D$ or spherical in $3D$) the formula (5.3) shows the non-tensor nature of the transformation.

Exercise 1. Derive formulas, similar to (5.3), for the cases of
i) Contravariant vector $a^i\,(x)$;
ii) Mixed tensor $T_i^j\,(x)$.

And so, we need to construct such derivative of a tensor which should be also tensor. For this end we shall use the following strategy. First of all we choose a name for this new derivative: it will be called **covariant**. After that the problem is solved through the following definition:

Def. 1. The covariant derivative ∇_i satisfies the following two conditions:

i) Tensor transformation after acting to any tensor;

ii) In a Cartesian coordinates $\{X^a\}$ the covariant derivative coincides with the usual partial derivative

$$\nabla_a = \partial_a = \frac{\partial}{\partial X^a}\,.$$

As we shall see in a moment, these two conditions fix the form of the covariant derivative of a tensor in a unique way.

The simplest way of obtaining an explicit expression for the covariant derivative is to apply what may be called a **covariant continuation**. Let us consider an arbitrary tensor, say mixed (1, 1)-type one W_i^j. In order to define its covariant derivative, we perform the following steps:

1) Transform it to the Cartesian coordinates X^a

$$W_b^a = \frac{\partial X^a}{\partial x^i}\frac{\partial x^j}{\partial X^b}\,W_j^i\,.$$

2) Take partial derivative with respect to the Cartesian coordinates X^c

$$\partial_c W_b^a \;=\; \frac{\partial}{\partial X^c}\Big(\frac{\partial X^a}{\partial x^i}\frac{\partial x^j}{\partial X^b}\,W_j^i\Big)\,.$$

3) Transform this derivative back to the original coordinates x^i using the tensor rule.

$$\nabla_k W_m^l = \frac{\partial X^c}{\partial x^k}\frac{\partial x^l}{\partial X^a}\frac{\partial X^b}{\partial x^m}\,\partial_c W_b^a\,.$$

It is easy to see that according to this procedure the conditions i) and ii) are satisfied automatically. Also, by construction, the covariant derivative follows the Leibnitz rule for the product of two tensors A and B

$$\nabla_i\,(A\cdot B) \;=\; \nabla_i A\cdot B \,+\, A\cdot\nabla_i B\,. \tag{5.4}$$

Let us make an explicit calculation for the covariant vector T_i. Following the above scheme of covariant continuation, we obtain

$$\nabla_i T_j = \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\,(\partial_a T_b) \;=\; \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\cdot\frac{\partial}{\partial X^a}\left(\frac{\partial x^k}{\partial X^b}T_k\right) =$$

$$= \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial x^k}{\partial X^b}\,\frac{\partial x^l}{\partial X^a}\frac{\partial T_k}{\partial x^l} \;+\; \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\,T_k\,\frac{\partial^2 x^k}{\partial X^a\,\partial X^b} \;=$$

$$= \partial_i T_j \,-\, \Gamma^k_{ji}\,T_k\,, \qquad \text{where} \qquad \Gamma^k_{ji} = -\frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial^2 x^k}{\partial X^a\,\partial X^b}\,. \tag{5.5}$$

Oe can obtain another representation for the symbol Γ^l_{ki}. Let us start from the general

Theorem 1. Suppose the elements of the matrix $\wedge$ depend on the parameter κ and $\wedge(\kappa)$ is differentiable and invertible matrix for $\kappa\in(a,b)$. Then, within the region (a,b), the inverse matrix $\wedge^{-1}(\kappa)$ is also differentiable and its derivative equals to

$$\frac{d\wedge^{-1}}{d\kappa} = -\,\wedge^{-1}\,\frac{\partial\wedge}{\partial\kappa}\,\wedge^{-1}\;. \tag{5.6}$$

Proof. The differentiability of the inverse matrix can be easily proved in the same way as for the inverse function in analysis. Taking derivative of $\wedge\cdot\wedge^{-1} = I$ we obtain

$$\frac{d\wedge}{d\kappa}\,\wedge^{-1} + \wedge\,\frac{\partial\wedge^{-1}}{\partial\kappa} = 0\,.$$

After multiplying this equation by $\wedge^{-1}$ from the left, we arrive at (5.6).

Consider

$$\wedge^b_k = \frac{\partial X^b}{\partial x^k}\,, \qquad \text{then the inverse matrix is} \qquad \left(\wedge^{-1}\right)^k_a = \frac{\partial x^k}{\partial X^a}\,.$$

Using (5.6) with x^i playing the role of parameter κ, we arrive at

$$\frac{\partial^2 X^b}{\partial x^i \partial x^k} = \frac{\partial}{\partial x^i}\wedge^b_k = -\,\wedge^b_l\,\frac{\partial\left(\wedge^{-1}\right)^l_a}{\partial x^i}\wedge^a_k =$$

$$= -\frac{\partial X^b}{\partial x^l}\left(\frac{\partial X^c}{\partial x^i}\,\frac{\partial}{\partial X^c}\,\frac{\partial x^l}{\partial X^a}\right)\frac{\partial X^a}{\partial x^k} = -\frac{\partial X^b}{\partial x^l}\,\frac{\partial X^a}{\partial x^k}\,\frac{\partial X^c}{\partial x^i}\,\frac{\partial^2 x^l}{\partial X^c \partial X^a}\,.$$

If applying this equality to (5.5), the second equivalent form of the symbol Γ^i_{ki} emerges

$$\Gamma^i_{ki} = \frac{\partial x^j}{\partial X^b}\,\frac{\partial^2 X^b}{\partial x^i \partial x^k} = -\frac{\partial^2 x^j}{\partial X^b \partial X^a}\,\frac{\partial X^a}{\partial x^k}\,\frac{\partial X^b}{\partial x^i}\,. \tag{5.7}$$

Exercise 2. Check, by making direct calculation and using (5.7) that

$$1) \qquad \nabla_i S^j = \partial_i S^j + \Gamma^j_{ki} S^k\,. \tag{5.8}$$

Exercise 3. Repeat the same calculation for the case of a mixed tensor and show that

$$2) \qquad \nabla_i W^j_k = \partial_i W^j_k + \Gamma^j_{li} W^l_k - \Gamma^l_{ki} W^j_l\,. \tag{5.9}$$

Observation. It is important that the expressions (5.8) and (5.9) are consistent with (5.7). In order to see this, we notice that the contraction $T_i S^i$ is a scalar, since T_i and S^j are co- and contravariant vectors. Using (5.4), we obtain

$$\nabla_j\left(T_i S^i\right) = T_i \cdot \nabla_j S^i + \nabla_j T_i \cdot S^i\,.$$

Let us assume that $\nabla_j S^i = \partial_j S^i + \tilde{\Gamma}^i_{kj} S^k$. For a while we take $\tilde{\Gamma}^i_{kj}$ as an independent quantity, but later one we shall see that it coincides with Γ^i_{kj}. Then

$$\partial_j\left(T_i S^i\right) = \nabla_j\left(T_i S^i\right) = T_i \nabla_j S^i + \nabla_j T_i \cdot S^i =$$

$$= T_i\,\partial_j S^i \,+\, T_i\,\tilde{\Gamma}^i_{kj}\,S^k \,+\, \partial_j T_i \cdot S^i \,-\, T_k\,\Gamma^k_{ij}\,S^i\,.$$

Changing the names of the umbral indices, we rewrite this equality in the form

$$\partial_j\left(T_i S^i\right) = S^i \cdot \partial_j T_i + T_i \cdot \partial_j S^i \;=\; T_i \partial_j S^i + \partial_j T_i \cdot S^i + T_i S^k\left(\tilde{\Gamma}^i_{kj} - \Gamma^i_{kj}\right)\,.$$

It is easy to see that the unique consistent option is to take $\tilde{\Gamma}^i_{kj} = \Gamma^i_{kj}$. Similar consideration may be used for any tensor. For example, in order to obtain the covariant derivative for the mixed $(1,1)$-tensor, one can use the fact that $T_i S^k W^i_k$ is a scalar, etc. This method has a serious advantage, because using it we can formulate the general rule for the covariant derivative of an arbitrary tensor

$$\nabla_i T^{j_1 j_2 \dots}{}_{k_1 k_2 \dots} \;=\; \partial_i T^{j_1 j_2 \dots}{}_{k_1 k_2 \dots} + \Gamma^{j_1}_{li}\,T^{l\,j_2 \dots}{}_{k_1 k_2 \dots} + \Gamma^{j_2}_{li}\,T^{j_1 l \dots}{}_{k_1 k_2 \dots} + \,\dots$$

$$-\Gamma^l_{k_2 i}\, T^{j_1 j_2 \dots}{}_{k_1 l \dots} - \Gamma^l_{k_1 i}\, T^{j_1 j_2 \dots}{}_{l\, k_2 \dots} - \dots\,. \tag{5.10}$$

As we already know, the covariant derivative differs from the partial derivative by the terms linear in the coefficients Γ^i_{jk}. Therefore, these coefficients play an important role. The first observation is that the transformation rule for Γ^i_{jk} must be non-tensor. In the equation (5.10), the non-tensor transformation for Γ^i_{jk} cancels with the non-tensor transformation for the partial derivative and, in sum, they give a tensor transformation for the covariant derivative.

Exercise 4. Derive, starting from each of the two representations (5.7), the transformation rule for the symbol Γ^i_{jk} from the coordinates x^i to other coordinates x'^j.

Result:

$$\Gamma^{i'}_{j'k'} = \frac{\partial x^{i'}}{\partial x^l}\,\frac{\partial x^m}{\partial x^{j'}}\,\frac{\partial x^n}{\partial x^{k'}}\,\Gamma^l_{mn} + \frac{\partial x^{i'}}{\partial x^r}\,\frac{\partial^2 x^r}{\partial x^{j'}\,\partial x^{k'}}\,. \tag{5.11}$$

The next observation is that the symbol Γ^i_{jk} is zero in the Cartesian coordinates X^a, where covariant derivative is nothing but the partial derivative. Taking X^a as a second coordinates $x^{i'}$, we arrive back at (5.7).

Finally, all the representations for Γ^i_{jk} which we have formulated, are not useful for practical calculations, and we need to construct one more. At that moment, we shall give the special name to this symbol.

Def. 2. Γ^i_{jk} is called Cristoffel symbol or affine connection.

Let us remark that for a while we are considering only the flat space, that is the space which admits global Cartesian coordinates and (equivalently) gloval orthonormal basis. In this special case Cristoffel symbol and affine connection are the same thing. But there are other, more general geometries which we do not consider in this Chapter. For this geometries Cristoffel symbol and affine connection may be different objects. In these cases it is customary to denote Cristoffel symbol as $\left\{ {i \atop jk} \right\}$ and keep notation $\Gamma^{\,i}_{jk}$ for the affine connection.

The most useful form of the affine connection Γ^i_{jk} is expressed via the metric tensor. In order to obtain this form of Γ^i_{jk} we remember that in the flat space there is a special orthonormal basis X^a in which the metric has the form

$$g_{ab} \equiv \delta_{ab}\,, \qquad \text{and hence} \qquad \partial_c\, g_{ab} \equiv \nabla_c\, g_{ab} \equiv 0\,.$$

Therefore, in any other coordinate system

$$\nabla_i\, g_{jk} = \frac{\partial X^c}{\partial x^i}\,\frac{\partial X^b}{\partial x^j}\,\frac{\partial X^a}{\partial x^k}\,\partial_c g_{ab} = 0\,. \tag{5.12}$$

If we apply to (5.12) the explicit form of the covariant derivative (5.10), we arrive at the equation

$$\nabla_i g_{jk} = \partial_i g_{jk} - \Gamma^l_{ji}\, g_{lk} - \Gamma^l_{ki}\, g_{lj} = 0\,. \tag{5.13}$$

Making permutations of indices we get

$$\begin{array}{ll} \partial_i g_{jk} = \Gamma^l_{ji}\, g_{lk} + \Gamma^l_{ki}\, g_{lj} & (i) \\ \partial_j g_{ik} = \Gamma^l_{ij}\, g_{lk} + \Gamma^l_{kj}\, g_{il} & (ii) \\ \partial_k g_{ij} = \Gamma^l_{ik}\, g_{lj} + \Gamma^l_{jk}\, g_{li} & (iii) \end{array} .$$

Taking linear combination $(i) + (ii) - (iii)$ we arrive at the relation

$$2\,\Gamma^l_{ij}\, g_{lk} \;=\; \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\,.$$

Contracting both parts with g^{km} (rememeber that $g^{km}\, g_{ml} = \delta^k_l$) we arrive at

$$\Gamma^k_{ij} \;=\; \frac{1}{2}\, g^{kl}\, \left(\partial_i\, g_{jl} + \partial_j\, g_{il} - \partial_l\, g_{ij}\right)\,. \tag{5.14}$$

The last formula is the most useful for practical calculation of Γ^i_{jk}. As an example, let us calculate Γ^i_{jk} for the polar coordinates $x = r\cos\varphi, \quad y = r\sin\varphi$ in $2D$. We start with the very detailed example

$$\Gamma^\varphi_{\varphi\varphi} = \frac{1}{2} g^{\varphi\varphi}\left(\partial_\varphi g_{\varphi\varphi} + \partial_\varphi g_{\varphi\varphi} - \partial_\varphi g_{\varphi\varphi}\right) + \frac{1}{2}\, g^{\varphi r}\left(\partial_\varphi g_{r\varphi} + \partial_\varphi g_{r\varphi} - \partial_r g_{\varphi\varphi}\right)\,.$$

The second term here is zero, since $g^{\varphi r} = 0$. Also the inverse of the diagonal metric $g_{ij} = \mathrm{diag}\,\big(r^2, 1\big)$ is indeed $g^{ij} = \mathrm{diag}\,\big(1/r^2,\, 1\big)$, such that $g^{\varphi\varphi} = 1/r^2\,, \quad g^{rr} = 1$. Then, since $g_{\varphi\varphi}$ does not depend on φ, we have $\Gamma^\varphi_{\varphi\varphi} = 0$.

$$\Gamma^\varphi_{rr} = \frac{1}{2} g^{\varphi\varphi}\left(2\partial_r g_{\varphi r} - \partial_\varphi g_{rr}\right) = 0\,,$$

$$\Gamma^r_{r\varphi} = \frac{1}{2} g^{rr}\left(\partial_r g_{\varphi r} + \partial_\varphi g_{rr} - \partial_r g_{\varphi r}\right) = 0\,,$$

$$\Gamma^\varphi_{r\varphi} = \frac{1}{2} g^{\varphi\varphi}\left(\partial_r g_{\varphi\varphi} + \partial_\varphi g_{r\varphi} - \partial_\varphi g_{r\varphi}\right) = \frac{1}{2r^2}\partial_r r^2 = \frac{1}{r}\,,$$

$$\Gamma^r_{\varphi\varphi} = \frac{1}{2} g^{rr}\left(\partial_\varphi g_{r\varphi} + \partial_\varphi g_{r\varphi} - \partial_r g_{\varphi\varphi}\right) = \frac{1}{2}\left(-\partial_r r^2\right) = -r\,,$$

$$\Gamma^r_{rr} = \frac{1}{2}\, g^{rr}\left(\partial_r g_{rr}\right) = 0\,, \tag{5.15}$$

After all, only $\Gamma^\varphi_{r\varphi} = \frac{1}{r}$ and $\Gamma^r_{\varphi\varphi} = -r$ are non-zero.

As an application of these formulas, we can consider the derivation of the Laplace operator acting on scalar and vector fields in $2D$ in polar coordinates. For any kind of field, the Laplace operator can be defines as

$$\Delta = g^{ij}\nabla_i\nabla_j\,. \tag{5.16}$$

In the Cartesian coordinates this operator has the well-known universal form

$$\Delta \;=\; \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \;=\; g^{ij}\,\nabla_i\nabla_j \;=\; \nabla^i\nabla_i\,.$$

What we want to know is how the operator Δ looks in an arbitrary coordinates, that is why we constructed it as covariant scalar operator expressed in terms of covariant derivatives. In the case of a scalar field Ψ, we have

$$\Delta\Psi = g^{ij}\nabla_i\nabla_j\Psi = g^{ij}\nabla_i\partial_j\Psi\,. \tag{5.17}$$

In (5.17), the second covariant derivative acts on the vector $\partial_i\Psi$, hence

$$\Delta\Psi \;=\; g^{ij}\left(\partial_i\partial_j\Psi - \Gamma^k_{ij}\partial_k\right)\,\Psi = \left(g^{ij}\partial_i\partial_j - g^{ij}\Gamma^k_{ij}\partial_k\right)\,\Psi\,. \tag{5.18}$$

Similarly, for the vector field $A^i\,(x)$ we obtain

$$\Delta A^i = g^{jk}\nabla_j\nabla_k A^i = g^{jk}\left[\partial_j\left(\nabla_k A^i\right) - \Gamma^l_{kj}\nabla_l A^i + \Gamma^i_{lj}\nabla_k A^l\right] =$$

$$= g^{jk}\left[\partial_j\left(\partial_k A^i + \Gamma^i_{lk}A^l\right) - \Gamma^l_{kj}\left(\partial_l A^i + \Gamma^i_{ml}A^k\right) + \Gamma^i_{lj}\left(\partial_k A^l + \Gamma^l_{mk}A^m\right)\right] =$$

$$= g^{jk}\left[\partial_j\partial_k A^i + \Gamma^i_{lk}\partial_j A^l + A^l\partial_j\Gamma^i_{lk} - \Gamma^l_{kj}\partial_l A^i + \Gamma^i_{lj}\partial_k A^l - \Gamma^l_{kj}\Gamma^i_{ml}A^m + \Gamma^i_{lj}\Gamma^l_{mk}A^m\right]\,. \tag{5.19}$$

Observation 1. The Laplace operators acting on vectors and scalars look quite different in curvilinear coordinates. They have the same form only in the Cartesian coordinates.

Observation 2. The formulas (5.17) and (5.19) are quite general, they hold for any dimension D and for an arbitrary choice of coordinates.

Let us perform the explcit calculations for the Laplace operator acting on scalars. The relevant non-zero contractions of Cristoffel symbol are $g^{ij}\Gamma^k_{ij} = g^{\varphi\varphi}\Gamma^k_{\varphi\varphi} + g^{rr}\Gamma^k_{rr}$, so using (5.15) we can see that only one trace

$$g^{\varphi\varphi}\Gamma^r_{\varphi\varphi} = -\frac{1}{r} \qquad \text{is non-zero}\,.$$

Then,

$$\Delta\Psi = \left[g^{\varphi\varphi}\frac{\partial^2}{\partial\varphi^2} + g^{rr}\frac{\partial^2}{\partial r^2} - \left(-\frac{1}{r}\right)\frac{\partial}{\partial r}\right]\Psi = \left[\frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2} + \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right]\Psi$$

and finally we obtain the well-known formula

$$\Delta\Psi \;=\; \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Psi}{\partial r}\right) \;+\; \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\varphi^2}\,. \tag{5.20}$$

Def. 3. For vector field $\mathbf{A}(x) = A^i\,\mathbf{e}_i$ the scalar quantity $\nabla_i A^i$ is called divergence of $\mathbf{A}$. The operation of taking divergence of the vector can be seen as a contraction of the vector with the Hamilton operator (5.2), so we have (exactly as in the case of a gradient) two distinct notations for divergence

$$\mathrm{div}\,\mathbf{A} = \nabla\mathbf{A} = \nabla_i A^i = \frac{\partial A^i}{\partial x^i} + \Gamma^i_{ji}A^j\,.$$

Exercises:

1) * Finish calculation of $\Delta A^i\,(x)$ in polar coordinates.

2) Write general expression for the Laplace operator acting on the covariant vector $\Delta B_i(x)$, similar to (5.19). Discuss the relation between two formulas and the formula (5.17).

3) Write general expression for $\nabla_i A^i$ and derive it in the polar coordinates (the result can be checked by comparison with similar calculations in cylindric and spherical coordinates in $3D$ in Chapter 7).

4) The commutator of two covariant derivatives is defined as

$$[\nabla_i, \nabla_j] \;=\; \nabla_i \nabla_j - \nabla_j \nabla_i \quad . \tag{5.21}$$

Prove, without **any** calculations, that for $\forall$ tensor T the commutator of two covariant derivatives is zero

$$[\nabla_i, \nabla_j]\, T^{k_1 \dots k_s}{}_{l_1 \dots l_t} \;=\; 0\,.$$

Hint. Use the existence of global Cartesian coordinates in the flat space and the tensor form of the transformations from one coordinates to another.

Observation. Of course, the same commutator is *non-zero* if we consider a more general geometry, which does not admit a global orthonormal basis. For example, after reading the Chapter 8 one can check that this commutator is non-zero for the geometry on a curved $2D$ surface in a $3D$ space, e.g. for the sphere.

5) Verify that for the scalar field Ψ the following relation takes place:

$$\Delta\Psi = \text{div}\,(\,\text{grad}\,\Psi)\,.$$

Hint. Use the fact that $\text{grad}\,\Psi = \nabla\Psi = \mathbf{e}^i\partial_i\Psi$ is a covariant vector. Before taking div, which is defined as an operation over the contravariant vector, one has to *rise* the index: $\nabla^i\Psi = g^{ij}\nabla_j\Psi$. Also, one has to use the metricity condition $\nabla_i g_{jk} = 0$.

6) Derive $\text{grad}\,(\,\text{div}\,\mathbf{A})$ for an arbitrary vector field $\mathbf{A}$ in polar coordinates.

7) Prove the following relation between the contraction of the Christoffel symbol Γ^k_{ij} and the derivative of the metric determinant $g = \det\|g_{\mu\nu}\|$:

$$\Gamma^j_{ij} \;=\; \frac{1}{2g}\,\frac{\partial g}{\partial x^i} \;=\; \partial_i \ln\sqrt{g}\,. \tag{5.22}$$

Hint. Use the formula (3.22).

8) Use the relation (5.22) for the derivation of the $\text{div}\,\mathbf{A}$ in polar coordinates. Also, calculate $\text{div}\,\mathbf{A}$ for elliptic coordinates

$$x = ar\cos\varphi\,, \qquad y = br\sin\varphi \tag{5.23}$$

and hyperbolic coordinates

$$x = \rho\cosh\chi\,, \qquad y = \rho\sinh\chi\,. \tag{5.24}$$

9) Prove the relation

$$g^{ij}\,\Gamma^k_{ij} \;=\; -\frac{1}{\sqrt{g}}\,\partial_i\left(\sqrt{g}\,g^{ik}\right) \tag{5.25}$$

and use it for the derivation of $\Delta\Psi$ in polar, elliptic and hyperbolic coordinates.

Chapter 6

Grad, div, rot and relations between them

6.1 Basic definitions and relations

Here we consider some important operations over vector and scalar fields, and relations between them. In this Chapter, all the consideration will be restricted to the special case of Cartesian coordinates, but the results can be easily generalized to any other coordinates by using the tensor transformation rule. As far as we are dealing with vectors and tensors, we know how to transform then.

Caution: When transforming the relations into corvilinear coordinates, one has to take care to use only the Levi-Civita tensor ε_{ijk} and not the maximal antisymmetric symbol E_{ijk}. The two object are indeed related as $\varepsilon_{ijk} = \sqrt{g}\, E_{ijk}$, where $E_{123} = 1$ in any coordinate system.

We already know three of four relevant differential operators:

$$\begin{aligned}
\operatorname{div} \mathbf{V} &= \partial_a V^a = \nabla \mathbf{V} && \text{divergence} \;; \\
\operatorname{grad} \Psi &= \hat{\mathbf{n}}^a \partial_a \Psi = \nabla \Psi && \text{gradient} \;; \\
\Delta &= g^{ab} \partial_a \partial_b && \text{Laplace operator}\,,
\end{aligned} \tag{6.1}$$

where

$$\nabla = \hat{\mathbf{n}}^a \partial_a = \hat{\mathbf{i}}\,\frac{\partial}{\partial x} + \hat{\mathbf{j}}\,\frac{\partial}{\partial y} + \hat{\mathbf{k}}\,\frac{\partial}{\partial z}$$

is the Hamilton operator. Of course, the vector operator ∇ can be formulated in $\forall$ other coordinate system using the covariant derivative

$$\nabla \;=\; \mathbf{e}^i \nabla_i \;=\; g^{ij}\, \mathbf{e}_j\, \nabla_i \tag{6.2}$$

acting on the corresponding (scalar, vector or tensor) fields. Let us define the fourth relevant differential operator, which is called **rotor**, or **curl**, or **rotation**. The rotor operator acts on vectors. Contrary to grad and div, rotor can be defined only in $3D$.

Def. 1. The rotor of the vector $\mathbf{V}$ is defined as

$$\operatorname{rot} \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ V_x & V_y & V_z \end{vmatrix} = \begin{vmatrix} \hat{\mathbf{n}}_1 & \hat{\mathbf{n}}_2 & \hat{\mathbf{n}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ V_1 & V_2 & V_3 \end{vmatrix} \;=\; E^{abc}\, \hat{\mathbf{n}}_a \partial_b V_c\,. \tag{6.3}$$

Other useful notation for the rot $\mathbf{V}$ uses its representation as the vector product of ∇ and $\mathbf{V}$

$$\text{rot}\ \mathbf{V} = \nabla \times \mathbf{V}\,.$$

One can easily prove the following important relations:

$$1) \qquad \text{rot grad}\ \Psi \equiv 0\,. \tag{6.4}$$

In the component notations the *l.h.s.* of this equality looks like

$$E^{abc}\,\hat{\mathbf{n}}_a \partial_b \left(\text{grad}\ \Psi \right)_c = E^{abc}\,\hat{\mathbf{n}}_a \partial_b \partial_c\,\Psi\,.$$

Indeed, this is zero because of the contraction of the symmetric symbol $\partial_b \partial_c = \partial_c \partial_b$ with the antisymmetric one E^{abc} (see **Exercise 3.11**).

$$2) \qquad \text{div rot}\ \mathbf{V} \equiv 0\,. \tag{6.5}$$

In the component notations the *l.h.s.* of this equality looks like

$$\partial_a\, E^{abc}\, \partial_b\, V_c \;=\; E^{abc}\, \partial_a\, \partial_b\, V_c\,,$$

that is zero because of the very same reason: contraction of the symmetric $\partial_b \partial_c$ with the antisymmetric E^{abc} symbols.

Exercises: Prove the following identities (Notice - in the exercises 11 and 12 we use the square brackets for the commutators $[\nabla,\, A] = \nabla\, A - A\, \nabla$, while in other cases the same bracket denotes vector product of two vectors):

1) $\text{grad}\ (\varphi\Psi) = \varphi\, \text{grad}\, \Psi + \Psi\, \text{grad}\, \varphi\,;$

2) $\text{div}\ (\varphi\mathbf{A}) = \varphi\, \text{div}\, \mathbf{A} + (\mathbf{A} \cdot \text{grad}\,)\, \varphi = \varphi \nabla \mathbf{A} + (\mathbf{A} \cdot \nabla)\, \varphi\,;$

3) $\text{rot}\ (\varphi\mathbf{A}) = \varphi\, \text{rot}\, \mathbf{A} - [\mathbf{A}, \nabla]\, \varphi\,;$

4) $\text{div}\ [\mathbf{A}, \mathbf{B}] = \mathbf{B} \cdot\ \text{rot}\, \mathbf{A} - \mathbf{A} \cdot\ \text{rot}\, \mathbf{B}\,;$

5) $\text{rot}\ [\mathbf{A}, \mathbf{B}] = \mathbf{A}\, \text{div}\, \mathbf{B} - \mathbf{B}\, \text{div}\, \mathbf{A} + (\mathbf{B}, \nabla)\, \mathbf{A} - (\mathbf{A}, \nabla)\, \mathbf{B}\,;$

6) $\text{grad}\ (\mathbf{A} \cdot \mathbf{B}) = [\mathbf{A},\, \text{rot}\, \mathbf{B}] + [\mathbf{B},\, \text{rot}\, \mathbf{A}] + (\mathbf{B} \cdot \nabla)\, \mathbf{A} + (\mathbf{A}, \nabla)\, \mathbf{B}\,;$

7) $(\mathbf{C}, \nabla)\, (\mathbf{A} \cdot \mathbf{B}) = (\mathbf{A}, (\mathbf{C}, \nabla)\, \mathbf{B}) + (\mathbf{B}, (\mathbf{C}, \nabla)\, \mathbf{A})\ ;$

8) $(\mathbf{C} \cdot \nabla)\, [\mathbf{A}, \mathbf{B}] = [\mathbf{A}, (\mathbf{C}, \nabla)\, \mathbf{B}] - [\mathbf{B}, (\mathbf{C}, \nabla)\, \mathbf{A}]\ ;$

9) $(\nabla, \mathbf{A})\, \mathbf{B} = (\mathbf{A}, \nabla)\, \mathbf{B} + \mathbf{B}\, \text{div}\, \mathbf{A}\,;$

10) $[\mathbf{A}, \mathbf{B}] \cdot\ \text{rot}\, \mathbf{C} = (\mathbf{B}, (\mathbf{A}, \nabla)\, \mathbf{C}) - (\mathbf{A}, (\mathbf{B}, \nabla)\, \mathbf{C})\ ;$

11) $[[\mathbf{A}, \nabla]\,, \mathbf{B}] = (\mathbf{A}, \nabla)\, \mathbf{B} + [\mathbf{A},\, \text{rot}\, \mathbf{B}] - \mathbf{A} \cdot\ \text{div}\, \mathbf{B}\,;$

12) $[[\nabla, \mathbf{A}]\,, \mathbf{B}] = \mathbf{A} \cdot\ \text{div}\, \mathbf{B} - (\mathbf{A}, \nabla)\, \mathbf{B} - [\mathbf{A},\, \text{rot}\, \mathbf{B}] - (\mathbf{B},\, \text{rot}\, \mathbf{A})\ .$

6.2 On the classification of differentiable vector fields

Let us introduce a new notion concerning the classification of differentiable vector fields.

Def. 2. Consider a differentiable vector $\mathbf{C}(\mathbf{r})$. It is called **potential** vector field, if it can be presented as a gradient of some scalar field $\Psi(\mathbf{r})$ (called potential)

$$\mathbf{C}(\mathbf{r}) \,=\, \text{grad}\,\Psi(\mathbf{r})\,. \tag{6.6}$$

Examples of the potential vector field can be easily found in physics. For instance, the potential force $\mathbf{F}$ acting on a particle is defined as $\mathbf{F} = -\,\text{grad}\,U(\mathbf{r})$, where $U(\mathbf{r})$ is the potential energy of the particle.

Def. 3. A differentiable vector field $\mathbf{B}(\mathbf{r})$ is called **solenoidal**, if it can be presented as a rotor of some vector field $\mathbf{A}(\mathbf{r})$

$$\mathbf{B}(\mathbf{r}) \,=\, \text{rot}\,\mathbf{A}(\mathbf{r})\,. \tag{6.7}$$

The most known physical example of the solenoidal vector field is the magnetic field $\mathbf{B}$ which is derived from the vector potential $\mathbf{A}$ exactly through (6.7).

There is an important theorem which is sometimes called *Fundamental Theorem of Vector Analysis.*

Theorem. Suppose $\mathbf{V}(\mathbf{r})$ is a smooth vector field, defined in the whole $3D$ space, which falls sufficiently fast at infinity. Then $\mathbf{C}(\mathbf{r})$ has unique (up to a gauge transformation) represenation as a sum

$$\mathbf{V} \,=\, \mathbf{C} \,+\, \mathbf{B}\,, \tag{6.8}$$

where $\mathbf{C}$ and $\mathbf{B}$ are potential and solenoidal fields corespondingly.

Proof. We shall accept, without proof (which can be found in many courses of Mathematical Physics, e.g. [11]), the following Theorem: The Laplace equation with a given boundary conditions at infinity has unique solution. In particular, the Laplace equation with a zero boundary conditions at infinity has unique zero solution.

Our first purpose is to prove that the separation of the vector field into the solenoidal and potential part is possible. Let us suppose

$$\mathbf{V} \,=\, \text{grad}\,\Psi \,+\, \text{rot}\,\mathbf{A} \tag{6.9}$$

and take divergence. Using (6.5), after some calculations, we arrive at the Poisson equation for Ψ

$$\Delta\Psi \,=\, \text{div}\,\mathbf{V}\,. \tag{6.10}$$

Since the solution of this equation is (up to a constant) unique, we now meet a single equation for $\mathbf{A}$. Let us take rot from both parts of the equation (6.9). Using the identity (6.4), we obtain

$$\text{rot}\,\mathbf{V} \,=\, \text{rot}\,(\,\text{rot}\,\mathbf{A}) \,=\, \text{grad}\,(\,\text{div}\,\mathbf{A}) \,-\, \Delta\mathbf{A}\,. \tag{6.11}$$

Let us notice that making the transformation

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \text{grad } f(\mathbf{r})$$

(in physics this is called gauge transformation) we can always provide that $\text{div}\,\mathbf{A}' \equiv 0$. At the same time, due to (6.4), this transformation does not affect the $\text{rot}\ \mathbf{A}$ and its contribution to $\mathbf{V}(\mathbf{r})$. Therefore, we do not need to distinguish $\mathbf{A}'$ and $\mathbf{A}$ and can simply suppose that $\text{grad}\ (\text{div}\,\mathbf{A}) \equiv 0$. Then, the equation (6.8) has a unique solution with respect to the vectors $\mathbf{B}$ and $\mathbf{C}$, and the proof is complete. Any other choice of the function $f(\mathbf{r})$ leads to a different equation for the remaining part of $\mathbf{A}$, but for any *given choice* of $f(\mathbf{r})$ the solution is also unique. The proof can be easily generalized to the case when the boundary conditions are fixed not at infinity but at some finite closed surface. The reader can find more detailed treatment of this and similar theorems in the book [4].

Chapter 7

Grad, div, rot and Δ in cylindric and spherical coordinates

The purpose of this Chapter is to calculate the operators grad , rot , div and Δ in cylindric and spherical coordinates. The method of calculations can be easily applied to more complicated cases. This may include derivation of other differential operators, acting on tensors; also derivation in other, more complicated (in particular, non-orthogonal) coordinates and to the case of dimension $D \neq 3$. In part, our considerations will repeat those we have already performed above, in Chapter 5, but we shall always make calculations in a slightly different manner, expecting that student can additionally benefit from the comparison.

Let us start with the **cylindric coordinates**

$$x = r\cos\varphi\,, \quad y = r\sin\varphi\,, \quad z = z\,. \tag{7.1}$$

In part, we can use here the results for the polar coordinates in $2D$, obtained in Chapter 5. In particular, the metric in cylindric coordinates is

$$g_{ij} = \text{diag}\,(g_{rr}, g_{\varphi\varphi}, g_{zz}) = \text{diag}\,\big(1,\, r^2,\, 1\big)\,. \tag{7.2}$$

Also, the inverse metric is $g^{ij} = \text{diag}\,\big(1,\, 1/r^2,\, 1\big)$. Only two components of the connection Γ^i_{jk} are different from zero

$$\Gamma^{\varphi}_{r\varphi} = \frac{1}{r} \quad \text{and} \quad \Gamma^{r}_{\varphi\varphi} = -r\,. \tag{7.3}$$

Let us start with the transformation of basic vectors. As always, local transformations of basic vectors and coordinates are performed by the use of the mutually inverse matrices

$$dx'^i = \frac{\partial x'^i}{\partial x^j}\,dx^j \quad \Longleftrightarrow \quad \mathbf{e}'_i = \frac{\partial x^j}{\partial x'^i}\,\mathbf{e}_j\,.$$

We can present the transformation in the form

$$\begin{pmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix} = \frac{D\,(r,\varphi,z)}{D\,(x,y,z)} \times \begin{pmatrix} \mathbf{e}_r \\ \mathbf{e}_\varphi \\ \mathbf{e}_z \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{e}_r \\ \mathbf{e}_\varphi \\ \mathbf{e}_z \end{pmatrix} = \frac{D\,(x,y,z)}{D\,(r,\varphi,z)} \times \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}. \tag{7.4}$$

It is easy to see that in the z-sector the Jacobian is trivial, so one has to concentrate attention to the xy-plane. Direct calculations give us (compare to the section 4.3)

$$\left(\frac{\partial X^a}{\partial x^i}\right) = \frac{D\,(x,y,z)}{D\,(r,\varphi,z)} = \begin{pmatrix} \partial_r x & \partial_r y & \partial_r z \\ \partial_\varphi x & \partial_\varphi y & \partial_\varphi z \\ \partial_z x & \partial_z y & \partial_z z \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -r\sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.5}$$

This is equivalent to the linear transformation

$$\mathbf{e}_r = \hat{\mathbf{n}}_x \cos\varphi + \hat{\mathbf{n}}_y \cdot \sin\varphi$$
$$\mathbf{e}_\varphi = -r\hat{\mathbf{n}}_x \sin\varphi + r\hat{\mathbf{n}}_y \cos\varphi$$
$$\mathbf{e}_z = \hat{\mathbf{n}}_z \,.$$

As in the $2D$ case, we introduce the normalized basic vectors

$$\hat{\mathbf{n}}_r = \mathbf{e}_r\,, \quad \hat{\mathbf{n}}_\varphi = \frac{1}{r}\,\mathbf{e}_\varphi\,, \quad \hat{\mathbf{n}}_z = \mathbf{e}_z\,.$$

It is easy to see that the transformation

$$\begin{pmatrix} \hat{\mathbf{n}}_r \\ \hat{\mathbf{n}}_\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \hat{\mathbf{n}}_x \\ \hat{\mathbf{n}}_y \end{pmatrix}$$

is nothing but rotation. Of course this is due to

$$\hat{\mathbf{n}}_r \perp \hat{\mathbf{n}}_\varphi \quad \text{and} \quad |\hat{\mathbf{n}}_r| = |\hat{\mathbf{n}}_\varphi| = 1\,.$$

As before, for any vector we denote

$$\mathbf{A} = \tilde{A}^r\mathbf{e}_r + \tilde{A}^\varphi\mathbf{e}_\varphi + \tilde{A}^z\mathbf{e}_z \;=\; A^r\hat{\mathbf{n}}_r + A^\varphi\hat{\mathbf{n}}_\varphi + A^z\hat{\mathbf{n}}_z\,,$$

where $\tilde{A}^r = A^r, \quad r\tilde{A}^\varphi = A^\varphi, \quad \tilde{A}^z = A^z\,.$

Now we calculate divergence of the vector in cylindric coordinates

$$\operatorname{div}\mathbf{A} \;=\; \nabla\mathbf{A} = \nabla_i\tilde{A}^i \;=\; \partial_i\tilde{A}^i + \Gamma^i_{ji}\tilde{A}^j \;=\; \partial_r\tilde{A}^r + \partial_\varphi\tilde{A}^\varphi + \frac{1}{r}\tilde{A}^r + \partial_z\tilde{A}^z \;=$$

$$= \;\frac{1}{r}\frac{\partial}{\partial r}\,(rA^r) + \frac{1}{r}\frac{\partial}{\partial\varphi}A^\varphi + \frac{\partial}{\partial z}\,A^z\,. \tag{7.6}$$

Let us calculate gradient of a scalar field Ψ in cylindric coordinates. By definition,

$$\nabla\Psi = \mathbf{e}^i\partial_i\Psi = \Big(\hat{\mathbf{i}}\,\partial_x + \hat{\mathbf{j}}\,\partial_y + \hat{\mathbf{k}}\,\partial_z\Big)\,\Psi\,, \tag{7.7}$$

because for the orthonormal basis the conjugated basis coisides with the original one.

By making the very same calculations as in the $2D$ case, we arrive at the expression for the gradient in cylindric coordinates

$$\nabla\Psi = \left(\mathbf{e}^r\,\frac{\partial}{\partial r} + \mathbf{e}^\varphi\,\frac{\partial}{\partial\varphi} + \mathbf{e}^z\,\frac{\partial}{\partial z}\right)\Psi\,.$$

Since

$$\mathbf{e}^r = \hat{\mathbf{n}}_r\,, \quad \mathbf{e}^\varphi = \frac{1}{r}\,\hat{\mathbf{n}}_\varphi\,, \quad \mathbf{e}^z = \hat{\mathbf{n}}_z\,,$$

we obtain the gradient in the orthonormal basis

$$\nabla\Psi = \left(\hat{\mathbf{n}}_r\,\frac{\partial}{\partial r} + \hat{\mathbf{n}}_\varphi\,\frac{1}{r}\,\frac{\partial}{\partial\varphi} + \hat{\mathbf{n}}_z\,\frac{\partial}{\partial z}\right)\Psi\,. \tag{7.8}$$

Exercise 1. Obtain the same formula by making rotation of the basis in the expression (7.7).

Let us consider the derivation of rotation of the vector $\mathbf{V}$ in the cylindric coordinates. First we remark that in the formula (7.6) for $\nabla_i A^i$ we have considered the divergence of the contravariant vector. The same may be maintained in the calculation or the $\text{rot}\,\mathbf{V}$, because only $\left\{\tilde{V}^r, \tilde{V}^\varphi, \tilde{V}^z\right\}$ components correspond to the basis (and to the change of coordiantes) which we are considering. On the other hand, after we expand the $\text{rot}\,\mathbf{V}$ in the normalized basis, there is no difference between co and contravariant components.

In the Cartesian coordinates

$$\text{rot}\,\mathbf{V} = E^{abc}\hat{\mathbf{n}}_a \partial_b V_c = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \partial_x & \partial_y & \partial_z \\ V_x & V_y & V_z \end{vmatrix} \tag{7.9}$$

Before starting the calculation in cylindric coordinates, let us learn to write the covariant and normalized components of the vector $\mathbf{V}$ in this coordinates. It is easy to see that

$$\tilde{V}_r = g_{rr}\tilde{V}^r = \tilde{V}^r \quad , \quad \tilde{V}_\varphi = g_{\varphi\varphi}\tilde{V}^\varphi = r^2\tilde{V}^\varphi \quad , \quad \tilde{V}_z = g_{zz}\tilde{V}^z = \tilde{V}^z \tag{7.10}$$

and furthermore

$$V_r = V^r = \tilde{V}_r \quad , \quad V_\varphi = V^\varphi = r\,\tilde{V}^\varphi \quad , \quad V_z = \tilde{V}^z = \tilde{V}^z\,. \tag{7.11}$$

It is worth noticing that the same relations hold also for the components of the vector $\text{rot}\,\mathbf{V}$.

It is easy to see that the Christoffel symbol here is not relevant, because

$$\text{rot}\,\mathbf{V} = \varepsilon^{ijk}\mathbf{e}_i \nabla_j V_k = \varepsilon^{ijk}\mathbf{e}_i\left(\partial_j V_k - \Gamma^l_{kj}V_l\right) = \varepsilon^{ijk}\mathbf{e}_i\partial_j A_k\,, \tag{7.12}$$

where we used an obvious relation $\varepsilon^{ijk}\Gamma^l_{kj} = 0$.

First we perform the calculations in the basis $\{\mathbf{e}_r, \mathbf{e}_\varphi, \mathbf{e}_z\}$ and only then pass to the normalized one

$$\hat{\mathbf{n}}_r = \mathbf{e}_r \quad , \quad \hat{\mathbf{n}}_\varphi = \frac{1}{r}\mathbf{e}_\varphi \quad , \quad \hat{\mathbf{n}}_z = \mathbf{e}_z\,.$$

Since $g_{ij} = diag\left(1, r^2, 1\right)$, the determinant is $g = \det\,(g_{ij}) = r^2$.

Then, starting calculations in the coordinate basis $\mathbf{e}_r$, $\mathbf{e}_\varphi$, $\mathbf{e}_z$ and then transorming the result into the normalized basis, we obtain

$$(\,\text{rot}\,\mathbf{V})^z = \frac{1}{\sqrt{r^2}}E^{zjk}\partial_j\tilde{V}_k = \frac{1}{r}\left(\partial_r\tilde{V}_\varphi - \partial_\varphi\tilde{V}_r\right) =$$

$$= \frac{1}{r}\,\partial_r\left[\frac{1}{r}\cdot r^2 V^\varphi\right] - \frac{1}{r}\,\partial_\varphi V^r \;=\; \frac{1}{r}\frac{\partial}{\partial r}\,(r\cdot V^\varphi) - \frac{1}{r}\frac{\partial}{\partial\varphi}V^r\,. \tag{7.13}$$

Furthermore, making similar calculations for other components of the rotation, we obtain

$$(\,\text{rot}\,\mathbf{V})^r = \frac{1}{r}E^{rjk}\partial_j\tilde{V}_k = \frac{1}{r}\left(\partial_\varphi\tilde{V}_z - \partial_z\tilde{V}_\varphi\right) = \frac{1}{r}\frac{\partial V^z}{\partial\varphi} - \frac{1}{r}\frac{\partial}{\partial z}\,(rV^\varphi) = \frac{1}{r}\frac{\partial V^z}{\partial\varphi} - \frac{\partial V^\varphi}{\partial z} \tag{7.14}$$

and

$$(\,\mathrm{rot}\,\mathbf{V})^{\varphi} = r \cdot \frac{1}{r}\, E^{\varphi jk} \partial_j \tilde{V}_k = \partial_z \tilde{V}_r - \partial_r \tilde{V}_z \,, \tag{7.15}$$

where the first factor of r appears because we want the expansion of $\mathrm{rot}\,\mathbf{V}$ in the normalized basis $\hat{\mathbf{n}}_i$ rather than in the coordinate one $\mathbf{e}_i$.

Remark: All components of $\mathrm{rot}\,\mathbf{A}$ which were derived above exactly correspond to the ones written in standard textbooks on electrodynamics and vector analizis, where the result is usually achieved via the Stokes theorem.

Finally, in order to complete the list of formulas concerning the cilindric coordinates, we write down the result for the Laplace operator (see Chapter 5, where it was calculated for the polar coordinates in $2D$).

$$\Delta\Psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r \frac{\partial \Psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \Psi}{\partial \varphi^2} + \frac{\partial^2 \Psi}{\partial z^2}\,.$$

Let us now consider the **spherical coordinates.**

$$\begin{cases} x = r\cos\varphi\sin\theta\,, & 0 \leq \theta \leq \pi \\ y = r\sin\varphi\sin\theta\,, & 0 \leq \varphi \leq 2\pi \\ z = r\cos\theta\,, & 0 \leq r \leq +\infty \end{cases}$$

The Jacobian matrix has the form

$$\left(\frac{\partial X^a}{\partial x^i}\right) = \frac{D\,(x,y,z)}{D\,(r,\varphi,\theta)} =$$

$$= \begin{pmatrix} \cos\varphi\sin\theta & \sin\varphi\sin\theta & \cos\theta \\ -r\sin\varphi\sin\theta & r\cos\varphi\sin\theta & 0 \\ r\cos\varphi\cos\theta & r\sin\varphi\cos\theta & -r\sin\theta \end{pmatrix} = \begin{pmatrix} x_{,r} & y_{,r} & z_{,r} \\ x_{,\varphi} & y_{,\varphi} & z_{,\varphi} \\ x_{,\theta} & y_{,\theta} & z_{,\theta} \end{pmatrix} = \mathcal{U}\,, \tag{7.16}$$

where $x_{,r} = \frac{\partial x}{\partial r}$ and so on. The matrix $\mathcal{U}$ is not orthogonal, but one can easily check that

$$\mathcal{U}\cdot\mathcal{U}^T = \mathrm{diag}\,\left(1,\ r^2\sin^2\theta,\ r^2\right) \tag{7.17}$$

where the transposed matrix can be, of course, easily calculated

$$\mathcal{U}^T = \begin{pmatrix} \cos\varphi\sin\theta & -r\sin\varphi\sin\theta & r\cos\varphi\cos\theta \\ \sin\varphi\sin\theta & r\cos\varphi\sin\theta & r\sin\varphi\cos\theta \\ \cos\theta & 0 & -r\sin\theta \end{pmatrix} \tag{7.18}$$

Therefore we can easily transform this matrix to the unitary form by multiplying it to the corresponding diagonal matrix. Hence, the inverse matrix has the form

$$\mathcal{U}^{-1} = \frac{D\,(r,\varphi,\theta)}{D\,(x,y,z)} = \mathcal{U}^T\cdot\mathrm{diag}\,\left(1,\ \frac{1}{r^2\sin^2\theta},\ \frac{1}{r^2}\right) = \begin{pmatrix} \cos\varphi\sin\theta & \frac{-\sin\varphi}{r\,\sin\theta} & \frac{\cos\varphi\cos\theta}{r} \\ \sin\varphi\sin\theta & \frac{\cos\varphi}{r\,\sin\theta} & \frac{\sin\varphi\cos\theta}{r} \\ \cos\theta & 0 & -\frac{\sin\theta}{r} \end{pmatrix}, \tag{7.19}$$

Exercise 1. Check $\mathcal{U} \cdot \mathcal{U}^{-1} = \hat{1}$, and compare the above method of deriving the inverse matrix with the standard one which has been used in section 4. Try to describe the situations when we can use the present, more economic, method of inverting the matrix. Is it possible to do so for the transformation to the non-orthogonal basis?

Exercise 2. Check that $g_{rr} = 1$, $g_{\varphi\varphi} = r^2 \sin^2\theta$, $g_{\theta\theta} = r^2$ and that the other component of the metric equal zero.

Hint. Use $g_{ij} = \frac{\partial x^a}{\partial x^i}\frac{\partial x^b}{\partial x^j} \cdot \delta_{ab}$ and matrix $\mathcal{U} = \frac{D(x,y,z)}{D(r,\varphi,\theta)}$.

Exercise 3. Check that $g^{ij} = diag\left(1, \frac{1}{r^2\sin^2\theta}, \frac{1}{r^2}\right)$. Try to solve this problem in several distinct ways.

Observation. The metric g_{ij} is non-degenerate everywhere except the poles $\theta = 0$ and $\theta = \pi$.

The next step is to calculate the Cristoffel symbol in spherical coordinates. The general expression has the form (5.14)

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}\right) .$$

A useful technical observation is that, since the metric g^{ij} is diagonal, in the sum over index l only the terms with $l = i$ are non-zero. The expressions for the components are

1) $\Gamma^r_{rr} = \frac{1}{2} g^{rr} \left(\partial_r g_{rr} + \partial_r g_{rr} - \partial_r g_{rr}\right) = 0\,;$

2) $\Gamma^r_{\varphi\varphi} = \frac{1}{2} g^{rr} \left(2\partial_\varphi g_{r\varphi} - \partial_r g_{\varphi\varphi}\right) = -\frac{1}{2} g^{rr} \partial_r g_{\varphi\varphi} = -\frac{1}{2}\frac{\partial}{\partial r} \cdot \left(r^2\sin^2\theta\right) = -r\sin^2\theta\,;$

3) $\Gamma^r_{\varphi\theta} = \frac{1}{2} g^{rr} \left(\partial_\varphi g_{\theta r} + \partial_\theta g_{\varphi r} - \partial_r g_{\theta\varphi}\right) = 0\,;$

4) $\Gamma^r_{\theta\theta} = \frac{1}{2} g^{rr} \left(2\partial_\theta g_{\theta r} - \partial_r g_{\theta\theta}\right) = -\frac{1}{2} g^{rr} \partial_r g_{\theta\theta} = -\frac{1}{2}\frac{\partial}{\partial r} r^2 = -r\,;$

5) $\Gamma^\varphi_{\varphi\varphi} = \frac{1}{2} g^{\varphi\varphi} \left(2\partial_\varphi g_{\varphi\varphi} - \partial_\varphi g_{\varphi\varphi}\right) = 0\,;$

6) $\Gamma^\varphi_{\varphi r} = \frac{1}{2} g^{\varphi\varphi} \left(\partial_\varphi g_{r\varphi} + \partial_r g_{\varphi\varphi} - \partial_\varphi g_{r\varphi}\right) = \frac{1}{2r^2\sin^2\theta} \cdot 2r\sin^2\theta = \frac{1}{r}\,;$

7) $\Gamma^\varphi_{rr} = \frac{1}{2} g^{\varphi\varphi} \left(2\partial_r g_{r\varphi} - \partial_\varphi g_{rr}\right) = 0\,;$

8) $\Gamma^\varphi_{\varphi\theta} = \frac{1}{2} g^{\varphi\varphi} \left(\partial_\varphi g_{\theta\varphi} + \partial_\theta g_{\varphi\varphi} - \partial_\varphi g_{\varphi\theta}\right) = \frac{2r^2\sin\theta\cos\theta}{2r^2\sin^2\theta} = \cot\theta\,;$

9) $\Gamma^\varphi_{r\theta} = \frac{1}{2} g^{\varphi\varphi} \left(\partial_r g_{\theta\varphi} + \partial_\theta g_{r\varphi} - \partial_\varphi g_{r\theta}\right) = 0\,;$

10) $\Gamma^\varphi_{\theta\theta} = \frac{1}{2} g^{\varphi\varphi} \left(2\partial_\theta g_{\varphi\theta} - \partial_\varphi g_{\theta\theta}\right) = 0\,;$

11) $\Gamma^r_{r\varphi} = \frac{1}{2} g^{rr} \left(\partial_r g_{\varphi r} + \partial_\varphi g_{rr} - \partial_r g_{r\varphi}\right) = 0\,;$

12) $\Gamma^r_{r\theta} = \frac{1}{2} g^{rr} \left(\partial_r g_{\theta r} + \partial_\theta g_{rr} - \partial_r g_{\theta r} \right) = 0\,;$

13) $\Gamma^\theta_{\theta\theta} = \frac{1}{2} g^{\theta\theta} \left(\partial_\theta g_{\theta\theta} \right) = 0\,;$

14) $\Gamma^\theta_{\varphi\varphi} = \frac{1}{2} g^{\theta\theta} \left(2\partial_\varphi g_{\varphi\theta} - \partial_\theta g_{\varphi\varphi} \right) = \frac{1}{2r^2} \left[-r^2 \frac{\partial \sin^2\theta}{\partial\theta} \right] = -\frac{1}{2} \sin 2\theta\,;$

15) $\Gamma^\theta_{rr} = \frac{1}{2} g^{\theta\theta} \left(2\partial_r g_{r\theta} - \partial_\theta g_{rr} \right) = 0\,;$

16) $\Gamma^\theta_{\theta\varphi} = \frac{1}{2} g^{\theta\theta} \left(\partial_\theta g_{\varphi\theta} + \partial_\varphi g_{\theta\theta} - \partial_\theta g_{\theta\varphi} \right) = 0\,;$

17) $\Gamma^\theta_{\theta r} = \frac{1}{2} g^{\theta\theta} \left(\partial_\theta g_{r\theta} + \partial_r g_{\theta\theta} - \partial_\theta g_{r\theta} \right) = \frac{1}{2}\frac{1}{r^2} \cdot \frac{\partial}{\partial r} \cdot r^2 = \frac{1}{r}\,;$

18) $\Gamma^\theta_{\varphi r} = \frac{1}{2} g^{\theta\theta} \left(\partial_\varphi g_{r\theta} + \partial_r g_{\varphi\theta} - \partial_\theta g_{\varphi r} \right) = 0\,.$

As we can see, the nonzero coefficients are only

$$\Gamma^\varphi_{\varphi\theta} = \cot\theta\,, \qquad \Gamma^\varphi_{r\varphi} = \frac{1}{r}\,, \qquad \Gamma^r_{\varphi\varphi} = -r\sin^2\theta\,,$$

$$\Gamma^r_{\theta\theta} = -r\,, \qquad \Gamma^\theta_{\varphi\varphi} = -\frac{1}{2}\sin^2\theta\,, \qquad \Gamma^\theta_{\theta r} = \frac{1}{r}\,.$$

Now we can easily derive $\operatorname{grad}\Psi$, $\operatorname{div}\mathbf{A}$, $\operatorname{rot}\mathbf{A}$ and $\Delta\Psi$ in spherical coordinates. According to our practice, we shall express all results in the normalized orthogonal basis $\hat{\mathbf{n}}_r, \hat{\mathbf{n}}_\varphi, \hat{\mathbf{n}}_\theta$. The relations between the components of the normalized basis and the corresponding elements of the coordinate basis are

$$\hat{\mathbf{n}}_r = \mathbf{e}_r \quad , \qquad \hat{\mathbf{n}}_\varphi = \frac{\mathbf{e}_\varphi}{r\sin\theta} \quad , \qquad \hat{\mathbf{n}}_\theta = \frac{\mathbf{e}_\theta}{r}\,. \tag{7.20}$$

Then,

$$\operatorname{grad}\Psi\,(r,\varphi,\theta) = \mathbf{e}^i \partial_i \Psi = g^{ij}\mathbf{e}_j \partial_i \Psi = g^{rr}\,\mathbf{e}_r\,\frac{\partial\Psi}{\partial r} + g^{\varphi\varphi}\,\mathbf{e}_\varphi\,\frac{\partial\Psi}{\partial\varphi} + g^{\theta\theta}\,\mathbf{e}_\theta\,\frac{\partial\Psi}{\partial\theta} =$$

$$= \mathbf{e}_r \frac{\partial\Psi}{\partial r} + \frac{\mathbf{e}_\varphi}{r^2\sin^2\theta}\,\frac{\partial\Psi}{\partial\varphi} + \frac{\mathbf{e}_\theta}{r^2}\,\frac{\partial\Psi}{\partial\theta} \;=\; \hat{\mathbf{n}}_r\,\frac{\partial\Psi}{\partial r} + \frac{\hat{\mathbf{n}}_\varphi}{r\sin\theta}\,\frac{\partial\Psi}{\partial\varphi} + \frac{\hat{\mathbf{n}}_\varphi}{r}\,\frac{\partial\Psi}{\partial\theta}\,.$$

Hence, we obtained the following components of the gradient of the scalar field Ψ in spherical coordinates:

$$(\operatorname{grad}\Psi)_r = \frac{\partial\Psi}{\partial r} \quad , \qquad (\operatorname{grad}\Psi)_\varphi = \frac{1}{r\sin\theta}\,\frac{\partial\Psi}{\partial\varphi} \quad , \qquad (\operatorname{grad}\Psi)_\theta = \frac{1}{r}\,\frac{\partial\Psi}{\partial\theta}\,. \tag{7.21}$$

Consider the derivation of divergence in spherical coordinates. Using the components of the Christoffel symbol derived above, we obtain

$$\operatorname{div}\mathbf{A} = \nabla_i \tilde{A}^i = \partial_i \tilde{A}^i + \Gamma^i_{ji}\tilde{A}^j = \partial_i \tilde{A}^i + \cot\theta\cdot\tilde{A}^\theta + \frac{2}{r}\tilde{A}^r =$$

$$= \frac{\partial\tilde{A}^r}{\partial r} + \frac{\partial\tilde{A}^\theta}{\partial\theta} + \frac{\partial\tilde{A}^\varphi}{\partial\varphi} + \cot\theta\cdot\tilde{A}^\theta + \frac{2\tilde{A}^r}{r}\,.$$

Now we have to rewrite this using the normalized basis $\hat{\mathbf{n}}_i$. The relevant relations are

$$\mathbf{A} = \tilde{A}^i \mathbf{e}_i = A^i \hat{\mathbf{n}}_i \,,$$

where

$$\tilde{A}^r = A^r \quad , \qquad \tilde{A}^\varphi = \frac{1}{r\sin\theta} A^\varphi \quad , \qquad \tilde{A}^\theta = \frac{1}{r} A^\theta \,.$$

Then

$$\operatorname{div}\mathbf{A} = \frac{\partial A^r}{\partial r} + \frac{1}{r}\frac{\partial A^\theta}{\partial \theta} + \frac{\cot\theta}{r} A^\theta + \frac{2A^r}{r} + \frac{1}{r\sin\theta}\frac{\partial A^\varphi}{\partial\varphi} =$$
$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A^r\right) + \frac{1}{r\sin\theta}\frac{\partial A^\varphi}{\partial \varphi} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(A^\theta\cdot\sin\theta\right) \,.$$

Now we can derive $\Delta\Psi$. There are two possibilities for making this calculation: to use the relation $\Delta\Psi = \operatorname{div}\left(\operatorname{grad}\Psi\right)$ or derive directly using covariant derivatives. Let us follow the second (a bit more economic) option, and also use the diagonal form of the metric tensor

$$\Delta\Psi = g^{ij}\nabla_i\nabla_j\Psi \;=\; g^{ij}\left(\partial_i\partial_j - \Gamma^k_{ij}\partial_k\right)\Psi \;=\; g^{rr}\frac{\partial^2\Psi}{\partial r^2} + g^{\varphi\varphi}\frac{\partial^2\Psi}{\partial\varphi^2} + g^{\theta\theta}\frac{\partial^2\Psi}{\partial\theta^2} - g^{ij}\Gamma^k_{ij}\frac{\partial\Psi}{\partial x^k} \,.$$

Since we already have all components of the Christoffel symbol, it is easy to obtain

$$g^{ij}\Gamma^r_{ij} = g^{\varphi\varphi}\Gamma^r_{\varphi\varphi} + g^{\theta\theta}\Gamma^r_{\theta\theta} = -\frac{2}{r} \quad \text{and} \quad g^{ij}\Gamma^\theta_{ij} = \Gamma^\theta_{\varphi\varphi}g^{\varphi\varphi} = -\frac{1}{r^2}\,\operatorname{cotan}\theta$$

while $\Gamma^\varphi_{ij}\, g^{ij} = 0$. After this we immediately arrive at

$$\Delta\Psi = \frac{\partial^2\Psi}{\partial r^2} + \frac{1}{r^2\,\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2} + \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\theta^2} + \frac{2}{r}\frac{\partial\Psi}{\partial r} + \frac{1}{r^2}\,\operatorname{cotan}\theta\,\frac{\partial\Psi}{\partial\theta} =$$
$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right) \;+\; \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial\Psi}{\partial\theta}\right) \;+\; \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2} \,. \tag{7.22}$$

In order to calculate

$$\operatorname{rot}\mathbf{A} = \varepsilon^{ijk}\mathbf{e}_i\nabla_j\tilde{A}_k \tag{7.23}$$

one has to use the formula $\varepsilon^{ijk} = E^{ijk}/\sqrt{g}$, where $E^{123} = E^{r\theta\varphi} = 1$. The determinant is of course $g = r^4\sin^2\theta$, therefore $\sqrt{g} = r^2\sin\theta$. Also, due to the identify $\varepsilon^{ijk}\Gamma^l_{jk} = 0$, we can use partial derivatives instead of the covariant ones.

$$\operatorname{rot}\mathbf{A} = \frac{1}{\sqrt{g}}\,\varepsilon^{ijk}\hat{\mathbf{n}}_i\partial_j\tilde{A}_k \,, \tag{7.24}$$

then substitute $\tilde{A}_k = g_{kl}\tilde{A}^l$ and finally

$$\tilde{A}_r = A^r \,, \quad \tilde{A}_\varphi = r\sin\theta\cdot A^\varphi \,, \quad \tilde{A}_\theta = r\cdot A^\theta \,.$$

Using $E^{r\theta\varphi} = 1$, we obtain

$$(\operatorname{rot}\mathbf{A})^r = \frac{1}{r^2\sin\theta}\left(\partial_\theta\tilde{A}_\varphi - \partial_\varphi\tilde{A}_\theta\right) \;=\; \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}\left(A^\varphi\cdot\sin\theta\right) - \frac{\partial A^\theta}{\partial\varphi}\right] \,. \tag{7.25}$$

This component does not need to be renormalized because of the property $\mathbf{e}_r = \hat{\mathbf{n}}_r$. For other components the situation is different, and they have to be normalized after the calculations. In particular, the φ-component must be multiplied by $r\sin\theta$ according to (7.20). Thus,

$$(\,\mathrm{rot}\,\mathbf{A})^\varphi = r\sin\theta\cdot\frac{E^{\varphi ij}}{r^2\,\sin\theta}\,\partial_i\tilde{A}_j \;=\; \frac{1}{r}\,\Big[\partial_r\tilde{A}_\theta - \partial_\theta\tilde{A}_r\Big] = \frac{1}{r}\,\frac{\partial}{\partial r}\,\Big(rA^\theta\Big) - \frac{1}{r}\,\frac{\partial A^r}{\partial\theta}\,. \tag{7.26}$$

Similarly,

$$(\,\mathrm{rot}\,\mathbf{A})^\theta = r\cdot\frac{1}{r^2\sin\theta}E^{\theta ij}\partial_i\tilde{A}_j = \frac{1}{r\sin\theta}\,\Big(\partial_\varphi\tilde{A}_r - \partial_r\tilde{A}_\varphi\Big) =$$

$$= \frac{1}{r\sin\theta}\,\left(\frac{\partial A^r}{\partial\varphi} - \frac{\partial}{\partial r}\,\left[r\sin\theta A^\varphi\right]\right) = \frac{1}{r\sin\theta}\,\frac{\partial A^r}{\partial\varphi} - \frac{1}{r}\frac{\partial}{\partial r}\,\left(rA^\varphi\right)\,. \tag{7.27}$$

And so, we obtained $\mathrm{grad}\,\Psi$, $\mathrm{div}\,\mathbf{A}$, $\mathrm{rot}\,\mathbf{A}$ and $\Delta\Psi$ in both cylindrical and spherical coordinates. Let us present a list of results in the normalized basis

I. Cylindric coordinates.

$$(\nabla\Psi)_r = \frac{\partial\Psi}{\partial r}\,,\qquad (\nabla\Psi)_\varphi = \frac{1}{r}\frac{\partial\Psi}{\partial\varphi}\,,\qquad (\nabla\Psi)_z = \frac{\partial\Psi}{\partial z}\,,$$

$$\mathrm{div}\,\mathbf{V} = \frac{1}{r}\frac{\partial}{\partial r}\,(rV^r) + \frac{1}{r}\frac{\partial V^r}{\partial\varphi} + \frac{\partial V^z}{\partial z}\,,$$

$$(\,\mathrm{rot}\,\mathbf{V})^z = \frac{1}{r}\,\frac{\partial\,(rV^\varphi)}{\partial r} - \frac{1}{r}\,\frac{\partial V^r}{\partial_\varphi}\,,\qquad (\,\mathrm{rot}\,\mathbf{V})^r = \frac{1}{r}\,\frac{\partial V^z}{\partial\varphi} - \frac{\partial V^\varphi}{\partial z}\,,$$

$$(\,\mathrm{rot}\,\mathbf{V})^\varphi \;=\; \frac{\partial V^r}{\partial z} - \frac{\partial V^z}{\partial r}\,,$$

$$\Delta\Psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\varphi^2} + \frac{\partial^2\Psi}{\partial z^2}\,.$$

II. Spherical coordinates.

$$(\nabla\Psi)_r = \frac{\partial\Psi}{\partial r}\,,\qquad (\nabla\Psi)_\varphi = \frac{1}{r\sin\theta}\,\frac{\partial\Psi}{\partial\varphi}\,,\qquad (\nabla\Psi)_\theta = \frac{1}{r}\,\frac{\partial\Psi}{\partial\theta}\,,$$

$$(\nabla\mathbf{V}) = \frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2V^r\right) + \frac{1}{r\sin\theta}\,\frac{\partial V^\varphi}{\partial\varphi} + \frac{1}{r\sin\theta}\,\frac{\partial V^\varphi}{\partial\theta}\,\Big(V^\theta\,\sin\theta\Big)\,,$$

$$\Delta\Psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\Psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\,\frac{\partial^2\Psi}{\partial\varphi^2}\,,$$

$$[\nabla\times\mathbf{V}]^r = \frac{1}{r\sin\theta}\,\left[\frac{\partial}{\partial\theta}\,(V^\varphi\cdot\sin\theta) - \frac{\partial V^\theta}{\partial\varphi}\right]\,,$$

$$[\nabla\times\mathbf{V}]^\varphi = \frac{1}{r}\,\frac{\partial}{\partial r}\,\Big(rV^\theta\Big) - \frac{1}{r}\,\frac{\partial V^r}{\partial\theta}\,,\qquad [\nabla\times\mathbf{V}]^\theta = \frac{1}{r\sin\theta}\,\frac{\partial V^r}{\partial\varphi} - \frac{1}{r}\,\frac{\partial}{\partial r}\,(rV^\varphi)\,.$$

Exercise 4. Use the relation (5.25) from Chapter 5 as a starting point for the derivation of the $\Delta\Psi$ in both cylindric and spherical coordinates. Compare the complexity of this calculation with the one presented above.

Chapter 8

Integrals over D-dimensional space.
Curvilinear, surface and volume integrals.

In this Chapter we define and discuss the volume integrals in D-dimensional space, curvilinear and surface integrals of the first and second type.

8.1 Brief reminder concerning $1D$ integrals and D-dimensional volumes

Let us start from a brief reminder of the notion of a definite (Riemann) integral in $1D$ and the notion of the area of a surface in $2D$.

Def. 1. Consider a function $f(x)$, which is defined and restricted on $[a, b]$. In order to construct the integral, one has to perform the division of the interval by inserting n intermediate points

$$a = x_0 < x_1 < ... < x_i < x_{i+1} < ... < x_n = b \tag{8.1}$$

and also choose arbitrary points $\xi_i \in [x_i, x_{i+1}]$ inside each interval $\Delta x_i = x_{i+1} - x_i$. The next step is to construct the integral sum

$$\sigma = \sum_{i=0}^{n-1} f(\xi_i)\, \Delta x_i\,. \tag{8.2}$$

For a fixed division (8.1) and fixed choice of the intermediate points ξ_i, the integral sum σ is just a number. But, since we have a freedom of changing (8.1) and ξ_i, this sum is indeed a function of many variables x_i and ξ_i. Let us define a procedure of adding new points x_i and correspondingly ξ_i. These points must be added in such a way that the maximal length $\lambda = \max\{\Delta x_i\}$ should decrease after inserting a new point into the division (8.1). Still we have a freedom to insert these points in a different ways and hence we have different procedures and different sequences of the integral sums. Within each procedure we meet a numerical sequence $\sigma = \sigma_n$. When $\lambda \to 0$, of course $n \to \infty$. If the limit

$$I = \lim_{\lambda \to 0} \sigma \tag{8.3}$$

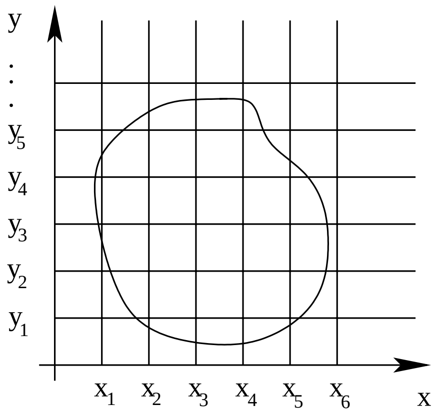

Figure 8.1: To the definition of the area.

is finite and universal, in sense it does not depend on the choice of the points x_i and ξ_i, it is called definite (or Riemann) integral of the function $f(x)$ over the interval $[a, b]$

$$I = \int_a^b f(x)dx\,. \tag{8.4}$$

The following Theorem plays a prominent role in the integral calculus:

Theorem 1. For f(x) continuous on $[a, b]$ the integral (8.4) exists. The same is true if f(x) is continuous everywhere except a finite number of points on $[a, b]$.

Observation. The integral possesses many well-known properties, which we will not review here. One can consult the textbooks on Real Analysis for this purpose.

Def. 2. Consider a restricted region $(S) \subset R^2$, The restricted region means that (S) is a subset of a finite circle of radius $C > 0$

$$x^2 + y^2 \leq C\,.$$

In order to define the area S of the figure (S), we shall cover it by the lattice $x_i = i/n$ and $y_j = j/n$ (see the picture) with the size of the cell $\frac{1}{n}$ and its area $\frac{1}{n^2}$. The individual cell may be denoted as

$$\Delta_{ij} = \{x, y \,|\, x_i < x < x_{i+1}\,,\; y_j < y < y_{j+1}\}$$

Suppose the number of the cells Δ_{ij} which are situated inside (S) is N_{in} and the number of cells which have common points with (S) is N_{out}. One can define an internal area of the figure (S) as $S_{in}(n)$ and an external area of the figure (S) as $S_{out}(n)$, where

$$S_{in}(n) = \frac{1}{n^2}\,N_{in} \quad \text{and} \quad S_{out}(n) = \frac{1}{n^2}\,N_{out}\,.$$

Taking the limit $n \to \infty$ we meet

$$\lim_{n\to\infty} S_{in}(n) = S_{int} \quad \text{and} \quad \lim_{n\to\infty} S_{out} = S_{ext}\,,$$

where S_{int} and S_{ext} are called internal and external areas of the figure (S). In case they coincide $S_{int} = S_{ext} = S$, we shall define the number S to be the area of the figure (S).

Remark 1. Below we consider only such (S) for which S exists. In fact, the existence of the area is guaranteed for any figure (S) with the continuous boundary (∂S).

Remark 2. In the same manner one can define the volume of the figure (V) in a space of an arbitrary dimension D (in particular for $D=3$). It is important (V) to be a restricted region in R^D

$$\exists L > 0\,, \qquad (V) \subset \mathcal{O}(0,L) = \{(x_1, x_2, ..., x_D)\,|\, x_1^2 + x_2^2 + ... + x_D^2 \leq L^2\} \subset R^D\,.$$

8.2 Volume integrals in curvilinear coordinates

The first problem for us is to formulate the D-dimensional integral in an arbitrary curvilinear coordinates $x^i = x^i(X^a)$. For this end we have to use the metric tensor and the Levi-Chivita tensor in the curvilinear coordinates.

Let us start from the metric tensor and its geometric meaning. As we already learned in the Chapter 2, the square of the distance between any two points X^a and Y^a, in a globally defined basis, is given by (2.7)

$$s_{xy}^2 = \sum_{a=1}^{D} (X^a - Y^a)^2 = g_{ij}\left(x^i - y^j\right)\left(x^j - y^j\right)$$

where x^i and y^j are the coordinates of the two points in an arbitrary global coordinates. For the local change of coordinates

$$x^i = x^i\left(X^1, X^2, ..., X^D\right)$$

the similar formula holds for the infinitesimal distances

$$ds^2 \;=\; g_{ij}\, dx^i\, dx^j\,. \tag{8.5}$$

Therefore, if we have a curve in D-dimensional space $x^i = x^i(\tau)$ (where τ is an arbitrary monotonic parameter along the curve), then the length of the curve between the points A with the coordinates $x^i(a)$ and B with the coordinates $x^i(b)$ is

$$s_{AB} \;=\; \int_{(AB)} ds \;=\; \int_a^b d\tau\,\sqrt{g_{ij}\,\frac{dx^i}{d\tau}\,\frac{dx^j}{d\tau}}\,. \tag{8.6}$$

We can make a conclusion the direct geometric sense of a metric is that it defines a distance between two infinitesimally close points and also the length of the finite curve in the D-dimensional space, in an arbitrary curvilinear coordinates.

Exercise 1. Consider the formula (8.6) for the case of Cartesian coordinates, and also for a global non-orthogonal basis. Discuss the difference between these two cases.

Now we are in a position to consider the volume integration in the Euclidean D-dimensional space. The definition of the integral in D-dimensional space

$$I \;=\; \int_{(V)} f(X^1, ..., X^D)\, dV \tag{8.7}$$

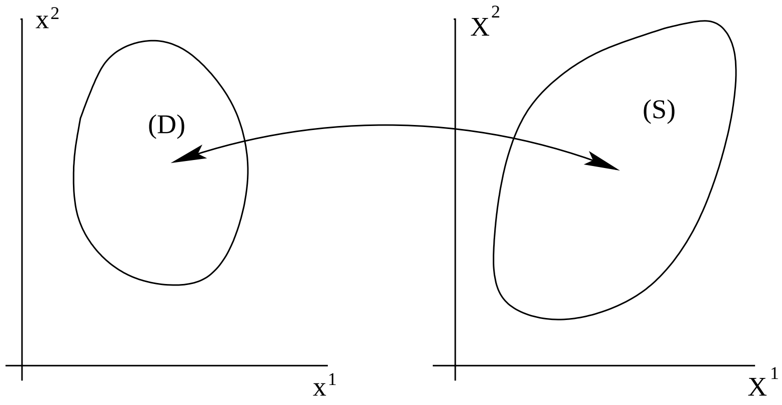

Figure 8.2: Illustration of the mapping $(G) \rightarrow (S)$ in $2D$.

implies that we can define the volume of the figure (V) bounded by the $(D\text{-}1)$-dimensional continuous surface (∂V). After this one has to follow the procedure similar to that for the $1D$ Riemann integral: divide the figure $(V) = \bigcup_i (V_i)$ such that each sub-figure (V_i) should have a well-defined volume V_i, for each sub-figure we have to choose points $M_i \in (V_i)$ and define[1] $\lambda = \max\{\text{diam}\,(V_i)\}$. The integral is defined as a universal limit of the integral sum

$$I = \lim_{\lambda\to 0} \sigma\,, \qquad \text{where} \qquad \sigma = \sum_i f\,(\,M_i\,)\;V_i\,. \tag{8.8}$$

Our purpose it to rewrite the integral (8.7) using the curvilinear coordinates x^i. In other words, we have to perform a change of variables $x^i = x^i\,(X^a)$ in this integral. First of all, the form of the figure (V) may change. In fact, we are performing the mapping of the figure (V) in the space with by the coordinates X^a into another figure (D), which lies in the space with the coordinates x^i. It is always assumed that this mapping is invertible everywhere except, at most, some finite number of isolated singular points. After the change of variables, the limits of integrations over the variables x^i correspond to the figure (D), not to the original figure (V).

The next step is to replace the function by the composite function of new variables

$$f(X^a) \;\rightarrow\; f(X^a(x^i)\,) \;=\; f(\,X^1(x^i),\,X^2(x^i),\,...\,,\,X^D(x^i)\,)\,.$$

Looking at the expression for the integral sum in (8.8) it is clear that the transformation of the function $f\,(\,M_i\,)$ may be very important. Of course, we can consider the integral in the case when the function is not necessary a scalar. It may be, e.g., a component of some tensor. But, in this case, if we consider curvilinear coordinates, the result of the integration may be an object which transforms in a non-regular way. In order to avoid the possible problems, let us suppose that the function f is a scalar. In this case the transformation of the function f does not pose a problem. Taking the additivity of the volume into account, we arrive at the necessity to transform the infinitesimal volume element

$$dV = dX^1 dX^2 \quad ... \quad dX^D \tag{8.9}$$

[1]Reminder: $\text{diam}\,(V_i)$ is a maximal (supremum) of the distances between the two points of the figure (V_i).

into the curvilinear coordinates x^i. Thus, the transformation of the volume element is the last step in transforming the integral. When transforming the volume element, we have to provide it to be a scalar, for otherwise the volume V will be different in different coordinate systems. In order to rewrite (8.9) in an invariant form, let us perform transformations

$$dX^1\, dX^2 \dots dX^D \;=\; \frac{1}{D!} \left| E_{a_1 a_2 \dots a_D} \right| dX^{a_1}\, dX^{a_2} \dots dX^{a_D} \;=$$

$$=\; \frac{1}{D!} \left| \varepsilon_{a_1 a_2 \dots a_D} \right| dX^{a_1}\, dX^{a_2} \dots dX^{a_D}\,. \tag{8.10}$$

Due to the modulo, all the non-zero terms in (8.10) give identical contributions. Since the number of such terms is $D!$, the two factorials cancel and we have a first identity. In the second equality we have used the coincidence of the Levi-Chivita tensor $\varepsilon_{a_1 a_2 \dots a_D}$ and the coordinate-independent tensor density $E_{a_1 a_2 \dots a_D}$ in the Cartesian coordinates. The last form of the volume element (8.10) is remarkable, because it is explicitly seen that dV is indeed a scalar. Indeed, we meet here the contraction of the tensor[2]. $\varepsilon_{a_1 \dots a_D}$ with the infinitesimal vectors dX^{a_k}. Since dV is an infinitesimal scalar, it can be easily transformed into other coordinates

$$dV \;=\; \frac{1}{D!} \left| \varepsilon_{i_1 i_2 \dots i_D} \right| dx^{i_1} \dots dx^{i_D} \;=\; \sqrt{g}\, \cdot dx^1 dx^2 \dots dx^D\,. \tag{8.11}$$

In the last step we have used the same identical transformation as in first equality in (8.10).

The quantity

$$J \;=\; \sqrt{g} = \sqrt{\det \|g_{ij}\|} \;=\; \left| \det \left(\frac{\partial X^a}{\partial x^i} \right) \right| \;=\; \frac{D(X^a)}{D(x^i)} \tag{8.12}$$

is nothing but the well-known Jacobian of the coordinate transformation. The modulo eliminates the possible change of sign which may come from a "wrong" choice of the order of the new coordinates. For example, any permutation of two coordinates, e.g., $x^1 \leftrightarrow x^2$ changes the sign of the determinant $\det\left(\frac{\partial X^a}{\partial x^i}\right)$, but the infinitesimal volume element dV should be always positive, as $J = \sqrt{g}$ always is.

8.3 Curvilinear integrals

Until now we have considered integrals over flat spaces, and now start to consider integrals over curves and curved surfaces.

Consider first a curvilinear integral of the 1-st type. Suppose we have a curve (L), defined by the vector function of the continuous parameter t:

$$\mathbf{r} = \mathbf{r}(t) \quad \text{where} \quad a \le t \le b\,. \tag{8.13}$$

It is supposed that when the parameter t grows up from $t = a$ to $t = b$, the point $\mathbf{r}$ moves, monotonically, from the initial point $\mathbf{r}(a)$ to the final one $\mathbf{r}(b)$. In particular, we suppose that in all points of the curve

$$dl^2 \;=\; g_{ij}\, \dot{x}^i\, \dot{x}^j \;>0\,.$$

[2] Verify, as an **Exercise 2**, that dx^i really transform as a contravariant vector components.

where the overdot stands for the derivative with respect to the parameter t. As we already learned above, the length of the curve is given by the integral

$$L = \int_{(L)} dl = \int_a^b \sqrt{g_{ij}\, \dot{x}^i\, \dot{x}^j}\; dt\,.$$

Def. 2. Consider a continuous function $f(x^i) = f(x^1, x^2, ..., x^D)$ defined along the curve (8.13). We define the **curvilinear integral of the 1-st type**

$$I_1 = \int_{(L)} f\,(\mathbf{r})\, dl \tag{8.14}$$

in a manner quite similar to the definition of the Riemann integral. Namely, one has to divide the interval

$$a = t_0 < t_1 < ... < t_i < t_{i+1} < ... < t_n = b \tag{8.15}$$

and also choose some $\xi_i \in [t_i,\, t_{i+1}]$ inside each interval $\Delta t_i = t_{i+1} - t_i$. After that we have to establish the length of each particular curve

$$\Delta\, l_i = \int_{t_i}^{t_{i+1}} \sqrt{g_{ij}\; \dot{x}^i\, \dot{x}^j}\; dt \tag{8.16}$$

and construct the integral sum

$$\sigma = \sum_{i=0}^{n-1} f\,(\,x(\xi_i)\,)\; \Delta\, l_i\,. \tag{8.17}$$

It is easy to see that this sum is also an integral sum for the Riemann integral

$$\int_a^b \sqrt{g_{ij}\, \dot{x}^i\, \dot{x}^j}\;\, f(\,x^i(t)\,)\, dt\,. \tag{8.18}$$

Introducing the maximal length $\lambda = \max\,\{\Delta l_i\}$, and taking the limit (through the same procedure which was described in the first section) $\lambda \to 0$, we meet a limit

$$I = \lim_{\lambda\to 0} \sigma\,. \tag{8.19}$$

If this limit is finite and does not depend on the choice of the points t_i and ξ_i, it is called curvilinear integral of the 1-st type (8.14).

The existence of the integral is guaranteed for the smooth curve and continuous function in the integrand. Also, the procedure above provides a method of explicit calculation of the curvilinear integral of the 1-st type through the formula (8.18).

The properties of the curvilinear integral of the first type include

$$\text{additivity} \quad \int_{AB} + \int_{BC} = \int_{AC} \qquad \text{and symmetry} \quad \int_{AB} = \int_{BA}\,.$$

Both properties can be easily established from (8.18).

Observation. In order to provide the coordinate-independence of the curvilinear integral of the 1-st type, we have to choose the function $f(x^i)$ to be a scalar field. Of course, the integral may be also defined for the non-scalar quantity. But, in this case the integral may be coordinate dependent and its geometrical (or physical) sense may be unclear. For example, if the function $f(x^i)$ is a component of a vector, and if we consider only global transformations of the basis, it is easy to see that the integral will also transform as a component of a vector. But, if we allow the local transformations of a basis, the integral may transform in a non-regular way and is geometric interpretation becomes unclear. The situation is equivalent to the one for the volume integrals and to the one for all other integrals we shall discuss below.

Let us notice that the scalar law of transformation for the curvilinear integral may be also achieved in the case when the integrated function is a vector. The corresponding construction is, of course, different from the curvilinear integral of the 1-st type.

Def. 3. The **curvilinear integral of the second type** is a scalar which results from the integration of the vector field along the curve. Consider a vector field $\mathbf{A}(\mathbf{r})$ defined along the curve (L). Let us consider, for the sake of simplicity, the $3D$ case and keep in mind an obvious possibility to generalize the construction for an arbitrary D. One can construct the following differential scalar:

$$\mathbf{A} \cdot d\mathbf{r} \;=\; A_x d_x + A_y d_y + A_z d_z = A_i dx^i \,. \tag{8.20}$$

If the curve is parametrized by the continuous monotonic parameter t, the last expression can be presented as

$$A_i dx^i \;=\; \mathbf{A} \cdot d\mathbf{r} \;=\; (\, A_x\, \dot{x} + A_y\, \dot{y} + A_z\, \dot{z}\,)\; dt \;=\; A_i\, \dot{x}^i\, dt\,. \tag{8.21}$$

This expression can be integrated along the curve (L) to give

$$\int_{(L)} \mathbf{A} d\mathbf{r} \;=\; \int_{(L)} A_x dx + A_y dy + A_z dz \;=\; \int_{(L)} A_i dx^i\,. \tag{8.22}$$

Each of these three integrals can be defined as a limit of the corresponding integral sum in the Cartesian coordinates. For instance,

$$\int_{(L)} A_x dx \;=\; \lim_{\lambda \to 0} \sigma_x \tag{8.23}$$

where

$$\sigma_x \;=\; \sum_{i=0}^{n-1} A_x\,(\,\mathbf{r}(\xi_i)\,) \cdot \Delta x_i\,. \tag{8.24}$$

Here we used the previous notations from Eq. (8.15) and below, and also introduced a new one

$$\Delta x_i \;=\; x\,(t_{i+1}) - x\,(t_i)\;.$$

The limit of the integral sum is taken exactly as in the previous case of the curvilinear integral of the 1-st type. The scalar property of the whole integral is guaranteed by the scalar transformation rule for each term in the integral sum. It is easy to establish the following relation between the curvilinear integral of the 2-nd type, the curvilinear integral of the 1-st type and the Riemann integral:

$$\int_{(L)} A_x dx + A_y dy + A_z dz \;=\; \int_{(L)} A_i \,\frac{dx^i}{dl}\, dl \;=\; \int_a^b A_i \,\frac{dx^i}{dt}\, dt\,, \tag{8.25}$$

where l is a natural parameter along the curve

$$l \;=\; \int_a^t \sqrt{g_{ij}\;\dot{x}^i\,\dot{x}^j}\;dt'\,, \qquad 0 \leq l \leq L\,.$$

Let us notice that, in the case of the Cartesian coordinates $X^a = (x, y, z))$, the derivatives in the second integral in (8.25) are nothing but the cosines of the angles between the tangent vector $d\mathbf{r}/dl$ along the curve and the basic vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$.

$$\cos\alpha \;=\; \frac{d\mathbf{r}\cdot\hat{\mathbf{i}}}{dl} \;=\; \frac{dx}{dl}\,, \qquad \cos\beta \;=\; \frac{d\mathbf{r}\cdot\hat{\mathbf{j}}}{dl} \;=\; \frac{dy}{dl}\,, \qquad \cos\gamma \;=\; \frac{d\mathbf{r}\cdot\hat{\mathbf{k}}}{dl} \;=\; \frac{dz}{dl}\,. \tag{8.26}$$

Of course, $\cos^2\alpha \,+\, \cos^2\beta \,+\, \cos^2\gamma \,=\, 1.$

Now we can rewrite Eq (8.25) in the following useful forms

$$\int_{(L)} \mathbf{A}d\mathbf{r} \;=\;=\; \int_0^L \,(\,A_x\;\cos\alpha \;+\; A_y\;\cos\beta \;+\; A_z\;\cos\gamma\,)\;dl \;=$$

$$=\; \int_a^b \,(\,A_x\;\cos\alpha \;+\; A_y\;\cos\beta \;+\; A_z\;\cos\gamma\,)\;\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}\;dt\,, \tag{8.27}$$

where we used $dl^2 = g_{ab}\dot{X}^a\dot{X}^b\,dt^2 \;=\; (\dot{x}^2+\dot{y}^2+\dot{z}^2)\,dt^2$.

The main properties of the curvilinear integral of the second type are

additivity $\displaystyle\int_{AB} + \int_{BC} = \int_{AC}$ and antisymmetry $\displaystyle\int_{AB} = -\int_{BA}$.

Exercises.[3]

1) Derive the following curvilinear integrals of the first type:

1-i. $\displaystyle\int_{(C)} (x+y)\,dl\,,$

where C is a contour of the triangle with the vertices $O(0,0)$, $A(1,0)$, $B(0,1)$.

1-ii. $\displaystyle\int_{(C)} (x^2+y^2)\,dl\,,$ where C is a curve

$$x = a(\cos t \,+\, t\,\sin t)\,, \quad y = a(\sin t \,-\, t\,\cos t)\,, \quad \text{where} \quad 0 \leq t \leq 2\pi\,.$$

[3] Most of the exercises of this Chapter are taken from the famous book [10].

1-iii. $\int\limits_{(C)} x\,dl$, where C is a part of the logarithmic spiral

$$r \,=\, a\,e^{k\varphi}\,, \quad \text{which is situated inside the circle} \quad r \leq a\,.$$

1-iv. $\int\limits_{(C)} (x^2+y^2+z^2)\,dl$, where C is a part of the spiral line

$$x = a\cos t\,, \quad y = a\sin t\,, \quad z = bt \quad (0 \leq t \leq 2\pi)\,.$$

1-v. $\int\limits_{(C)} x^2\,dl$, where C is a circumference

$$x^2+y^2+z^2 = a^2 \qquad x+y+z = 0\,.$$

2) Derive the following curvilinear integrals of the second type:

2-i. $\int\limits_{(OA)} xdy \pm ydx\,,$

where $A(1,2)$ and (OA) are the following curves:

(a) Straight line.

(b) Part of parabola with Oy axis.

(c) Polygonal line OBA where $B(0,2)$.

Discuss the result, and in particular try to explain why in the case of the positive sign all three cases gave the same integral.

2-ii. $\int\limits_{(C)} (x^2-2xy)dx \,+\, (y^2-2xy)dy$, where (C) is a parabola $y=x^2$, $-1 \leq x \leq 1$.

2-iii. $\int\limits_{(C)} (x+y)dx \,+\, (x-y)dy\,,$ where (C) is an ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ in the counterclockwise direction.

8.4 $2D$ Surface integrals in a $3D$ space

Consider the integrals over the $2D$ surface (S) in the $3D$ space $(S) \subset R^3$. Suppose the surface is defined by three smooth functions

$$x = x(u,v)\,, \qquad y = y(u,v)\,, \qquad z = z(u,v)\,, \tag{8.28}$$

where u and v are called internal coordinates on the surface. One can regard (8.28) as a mapping of a figure (G) on the plane with the coordinates $u,\,v$ into the $3D$ space with the coordinates $x,\,y,\,z$.

Example. The sphere (S) of the constant radius R can be described by internal coordinates (angles) φ, θ. Then the relation (8.28) looks like

$$x = R\cos\varphi\sin\theta\,, \qquad y = R\sin\varphi\sin\theta\,, \qquad z = R\cos\theta\,, \tag{8.29}$$

where φ and θ play the role of internal coordinates u and v. The values of these coordinates are taken from the rectangle

$$(G) \,=\, \{(\varphi,\,\theta)\,|\,0 \leq \varphi < 2\pi\,,\; 0 \leq \theta < \pi\}\,,$$

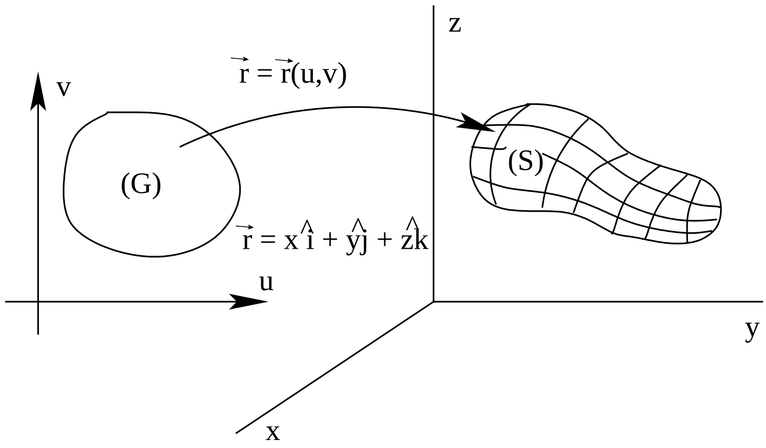

Figure 8.3: Picture of a mapping of (G) into (S).

ad the parametric description of the sphere (8.29) is nothing but the mapping $(G) \to (S)$. One can express all geometric quantities on the surface of the sphere using the internal coordinates φ and θ, without addressing the coordinates of the external space x, y, z.

In order to establish the geometry of the surface, let us start by considering the infinitesimal lime element dl linking two points of the surface. We can suppose that these two infinitesimally close points belong to the same smooth curve $u = u(t)$, $v = v(t)$ situated on the surface. Then the line element is

$$dl^2 \,=\, dx^2 + dy^2 + dz^2\,. \tag{8.30}$$

where

$$dx \,=\, \left(\frac{\partial x}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial x}{\partial v}\frac{\partial v}{\partial t}\right) dt \,=\, x'_u du + x'_v dv\,,$$

$$dy \,=\, \left(\frac{\partial y}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial t}\right) dt \,=\, y'_u du + y'_v dv\,,$$

$$dz \,=\, \left(\frac{\partial z}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial t}\right) dt \,=\, z'_u du + z'_v dv\,. \tag{8.31}$$

Using the last formulas, we obtain

$$dl^2 \,=\, \left(x'_u du + x'_v dv\right)^2 + \left(y'_u du + y'_v dv\right)^2 + \left(z'_u du + z'_v dv\right)^2 =$$

$$= \left({x'_u}^2 + {y'_u}^2 + {z'_u}^2\right) du^2 + 2\left(x'_u \cdot x'_v + y'_u \cdot y'_v + z'_u \cdot z'_v\right) dudv + \left({x'_v}^2 + {y'_v}^2 + {z'_v}^2\right) dv^2 \,=$$

$$= g_{uu} du^2 + 2 g_{uv} dudv + g_{vv} dv^2\,. \tag{8.32}$$

where

$$g_{uu} = {x'_u}^2 + {y'_u}^2 + {z'_u}^2\,, \qquad g_{uv} = x'_u x'_j + y'_u y'_v + z'_u z'_v\,, \qquad g_{vv} = {x'_v}^2 + {y'_v}^2 + {z'_v}^2\,. \tag{8.33}$$

are the elements of the so-called induced metric on the surface. As we can see, the expression $dl^2 = g_{ij}dx^i dx^j$ is valid for the curved surfaces as well as for the curvilinear coordinates in a flat D-dimensional spaces. Let us consider this relation from another point of view and introduce two basis vectors on the surface in such a way that the scalar products of these vectors equal to the corresponding metric components. The tangent plane at any point is composed by linear combinations of the two vectors

$$\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u} \qquad \text{and} \qquad \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}\,. \tag{8.34}$$

Below we suppose that at any point of the surface $\mathbf{r}_u \times \mathbf{r}_v \neq 0$. This means that the lines of constant coordinates u and v are not parallel or, equivalently, that the internal coordinates are not degenerate.

For the smooth surface the infinitesimal displacement may be considered identical to the infinitesimal displacement over the tangent plane. Therefore, for such a displacement we can write

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\,du + \frac{\partial \mathbf{r}}{\partial v}\,dv \tag{8.35}$$

and hence (here we use notations $\mathbf{a}^2$ and $(\mathbf{a})^2$ for the square of the vector $= \mathbf{a}\cdot\mathbf{a}$)

$$dl^2 \,=\, (d\mathbf{r})^2 \,=\, \left(\frac{\partial \mathbf{r}}{\partial u}\right)^2 du^2 + 2\,\frac{\partial \mathbf{r}}{\partial u}\,\frac{\partial \mathbf{r}}{\partial v}\,dudv + \left(\frac{\partial \mathbf{r}}{\partial v}\right)^2 dv^2\,, \tag{8.36}$$

that is an alternative form of Eq. (8.32). Thus, we see that the tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ may be safely interpreted as a basic vectors on the surface, in the sense that the components of the induced metric are nothing but the scalar products of these two vectors

$$g_{uu} = \mathbf{r}_u \cdot \mathbf{r}_u\,, \qquad g_{uv} = \mathbf{r}_u \cdot \mathbf{r}_v\,, \qquad g_{vv} = \mathbf{r}_v \cdot \mathbf{r}_v\,. \tag{8.37}$$

Exercise 2. (i) Derive the tangent vectors $\mathbf{r}_\varphi$ and $\mathbf{r}_\theta$ for the coordinates φ, θ on the sphere; (ii) Check that for these two vectors are orthogonal $\mathbf{r}_\varphi \cdot \mathbf{r}_\theta = 0$; (iii) Using the tangent vectors, derive the induced metric on the sphere. Check this calculation using the formulas (8.33); (iv) Compare the induced metric on the sphere and the $3D$ metric in the spherical coordinates. Invent the procedure how to obtain the induced metric on the sphere from the $3D$ metric in the spherical coordinates. (v) * Analyze whether the internal metric on any surface may be obtained through a similar procedure.

Exercise 3. Repeat the points (i) and (ii) of the previous exercise for the surface of the torus. The equation of the torus in cylindric coordinates is $(r-b)^2 + z^2 = a^2$, where $b > a$. Try to find angular variables (internal coordinates) which are most useful and check that in this coordinates the metric is diagonal. Give geometric interpretation.

Let us now calculate the infinitesimal element of the area of the surface, corresponding to the infinitesimal surface element $dudv$ on the figure (G) in the uv-plane. After the mapping, on the tangent plane we have a parallelogram spanned on the vectors $\mathbf{r}_u du$ and $\mathbf{r}_v dv$. According to the

Analytic Geometry, the area of this parallelogram equals a modulo of the vector product of the two vectors

$$|d\mathbf{A}| \,=\, |\,\mathbf{r}_u du \,\times\, \mathbf{r}_v dv| \,=\, |\mathbf{r}_u \times \mathbf{r}_v|\; du \cdot dv\,. \tag{8.38}$$

In order to evaluate this product, it is useful to take a square of the vector (here α is the angle between $\mathbf{r}_u$ and $\mathbf{r}_v$)

$$(d\mathbf{A})^2 \,=\, |\mathbf{r}_u \times \mathbf{r}_v|^2\, du^2 dv^2 \,=\, r_u^2 r_v^2 \sin^2\alpha \cdot du^2 dv^2 \,=$$

$$= du^2 dv^2 \left\{ r_u^2 r_v^2 \left(1 - \cos^2\alpha\right)\right\} = du^2 dv^2 \left\{ \mathbf{r}^2 \cdot \mathbf{r}^2 - (\mathbf{r}_u \cdot \mathbf{r}_v)^2 \right\} =$$

$$= \left(g_{uu} \cdot g_{vv} - g_{uv}^2\right) du^2 dv^2 \,=\, g\, du^2 dv^2\,, \tag{8.39}$$

where g is the determinant of the induced metric. Once again, we arrived at conventional formula $dA = \sqrt{g}\, dudv$, where

$$g \,=\, \begin{vmatrix} g_{uu} & g_{uv} \\ g_{vu} & g_{vv} \end{vmatrix}$$

is metric determinant. Finally, the area S of the whole surface (S) may be defined as an integral

$$S \,=\, \int\!\!\int_{(G)} \sqrt{g}\, du\, dv\,. \tag{8.40}$$

Geometrically, this definition corresponds to a triangulation of the curved surface with the consequent limit, when the size of all triangles simultaneously tend to zero.

Exercise 4 Find g and $\sqrt{g}$ for the sphere, using the induced metric in the (φ, θ) coordinates. Apply this result to calculate the area of the sphere. Repeat the same for the surface of the torus (see Exercise 3).

In a more general situation, we have n-dimensional surface embedded into the D-dimensional space R^D, with $D > n$. Introducing internal coordinates u^i on the surface, we obtain its parametric equation of the surface in the form

$$x^\mu = x^\mu\left(u^i\right)\,, \tag{8.41}$$

where

$$\begin{cases} i = 1, ..., n \\ \mu = 1, ..., D \end{cases}.$$

The vector in R^D may be presented in the form $\bar{r} = r^\mu \bar{e}_\mu$, where $\bar{e}_\mu$ are the corresponding basis vectors. The tangent vectors are given by the partial derivatives $\bar{r}_i = \partial\bar{r}/\partial u^i$. Introducing a curve $u^i = u^i(t)$ on the surface

$$dx^\mu = \frac{\partial x^\mu}{\partial u^i}\,\frac{du^i}{dt}\,dt = \frac{\partial x^\mu}{\partial u^i}\,du^i\,. \tag{8.42}$$

we arrive at the expression for the distance between the two infinitesimally close points

$$ds^2 \,=\, g_{\mu\nu} dx^\mu dx^\nu \,=\, \frac{\partial x^\mu}{\partial u^i}\,\frac{\partial x^\nu}{\partial u^i}\, g_{\mu\nu}\, du^i du^j\,. \tag{8.43}$$

Therefore, in this general case we meet the same relation as for the usual coordinate transformation

$$g_{ij}\,(u^u)\;=\;\frac{\partial x^\mu}{\partial u^i}\,\frac{\partial x^\nu}{\partial u^j}\;g_{\mu\nu}\,(x^\alpha)\;. \tag{8.44}$$

The condition of the non-degeneracy is that the induced metric must be non-degenerate

$$\text{rank}\;(g_{ij}) = n\,.$$

Sometimes (e.g. for the coordinates on the sphere at the pole points $\theta = 0,\,\pi$) the metric may be degenerate in some isolated points. The volume element of the area (surface element) in the n-dimensional case also looks as

$$dS \;=\; J\,du^1 du^2 ... du^n\,, \qquad \text{where} \qquad J \;=\; \sqrt{g}\,. \tag{8.45}$$

Hence, there is no big difference between two-dimensional and n-dimensional surfaces. For the sake of simplicity we shall concentrate, below, on the two-dimensional case. Now we are in a position to define the surface integral of the first type. Suppose we have a surface (S) defined as a mapping (8.28) of the closed finite figure (G) in a uv-plane, and a function $f(u,v)$ on this surface. Instead one can take a function $g(x,y,z)$ which generates a function on the surface through the relation

$$f(u,v) \;=\; g\,(\,x(u,v),\;y(u,v),\;z(u,v)\,)\;. \tag{8.46}$$

Def. 4. The surface integral of the first type is defined as follows: Let us divide (S) into particular sub-surfaces (S_i) such that each of them has a well-defined area

$$S_i \;=\; \int_{G_i} \sqrt{g}\,du\,dv\,. \tag{8.47}$$

Of course, one has to perform the division of (S) into (S_i) such that the intersections of the two sub-figures have zero area. On each of the particular surfaces (S_i) we choose a point $M_i\,(\xi_i,\eta_i)$ and construct an integral sum

$$\sigma \;=\; \sum_{i=1}^{n} f\,(\,M_i\,)\cdot S_i\,, \tag{8.48}$$

where $f\,(M_i) = f\,(\xi_i,\,\eta_i)$. The next step is to define $\lambda = \max\{\,\text{diam}\,(S_i)\,\}$. If the limit

$$\mathcal{I}_1 \;=\; \lim_{\lambda\to 0}\sigma \tag{8.49}$$

is finite and does not depend on the choice of (S_i) and on the choice of M_i, it is called the surface integral of the first type

$$\mathcal{I}_1 \;=\; \int\!\!\int_{(S)} f(u,v)\,dS \;=\; \int\!\!\int_{(S)} g(x,y,z)\,dS\,. \tag{8.50}$$

¿From the construction it is clear that this integral can be indeed calculated as a double integral over the figure (G) in the uv-plane

$$\mathcal{I}_1 \;=\; \int\!\!\int_{(S)} f(u,v)\,dS \;=\; \int\!\!\int_{(G)} f(u,v)\,\sqrt{g}\;dudv\,. \tag{8.51}$$

Figure 8.4: Examples of two-side and one-side surfaces

Remark. The surface integral (8.51) is, by construction, a scalar, if the function $f(u,v)$ is a scalar. Therefore, it can be calculated in any coordinates (u',v'). Of course, when we change the internal coordinates on the surface to (u',v'), the following aspects must be changed:

1) Form of the area $(G) \rightarrow (G')$;

2) the surface element $\sqrt{g} \rightarrow \sqrt{g'}$;

3) the form of the integrand $f(u,v) = f'(u',v') = f\left(u\left(u',v'\right), v\left(u',v'\right)\right)$. In this case the surface integral of the first type is a coordinate-independent geometric object.

Let us now consider move complicated type of integral: surface integral of the second type. The relation between the two types of surface integrals is similar to that between the two types of curvilinear integrals, which we have considered in the previous section. First of all, we need to introduce the notion of an oriented surface. Indeed, for the smooth surface with independent tangent vectors $\mathbf{r}_u = \partial\mathbf{r}/\partial u$ and $\mathbf{r}_v = \partial\mathbf{r}/\partial v$ one has a freedom to choose which of them is the first and which is the second. This defines the orientation for their vector product, and corresponds to the choice of orientation for the surface. In a small vicinity of a point on the surface with the coordinates u, v there are two opposite directions associated with the normalized normal vectors

$$\pm\,\hat{\mathbf{n}} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}\,. \tag{8.52}$$

Anyone of these directions could be taken as positive.

Exercise 5. Using the formula (8.52), derive the normal vector for the sphere and torus (see Exercises 2,3). Give geometrical interpretation. Discuss, to which extent the components of the normal vector depend on the choice of internal coordinates.

The situation changes if we consider the whole surface. One can distinguish two types of surfaces: 1) oriented, or two-side surface, like the one of the cylinder or sphere and 2) non-oriented (one-side) surface, like the Mobius surface (see the Figure 8.4).

Below we consider only smooth, parametrically defined two-side oriented surfaces. If

$$|\mathbf{r}_u \times \mathbf{r}_v| \neq 0$$

in all points on the surface, then the normal vector $\hat{\mathbf{n}}$ never changes its sign. We shall choose, for definiteness, the positive sign in (8.52). According to this formula, the vector $\hat{\mathbf{n}}$ is normal to the surface. It is worth noticing, that if we consider the infinitesimal piece of surface which corresponds to the $dudv$ area on the corresponding plane (remember that the figure (G) is the inverse image of the "real" surface (S) on the plane with the coordinates u, v), then the area of the corresponding parallelogram on (S) is $dS = |\mathbf{r}_u \times \mathbf{r}_v| \, dudv$. This parallelogram may be characterized by its area dS and by the normal vector $\hat{\mathbf{n}}$ (which may be, of course, different in the distinct points of the surface). Using these two quantities, we can construct a new vector $d\mathbf{S} = \hat{\mathbf{n}} \cdot dS = [\mathbf{r}_u, \mathbf{r}_v] \, dudv$. This vector quantity may be regarded as a vector element of the (oriented) surface. This vector contains information about both area of the infinitesimal parallelogram (it is equivalent to the knowledge of the Jacobian $\sqrt{g}$) and about its orientation in the external $3D$ space.

For the calculational purposes it is useful to introduce the components of $\hat{\mathbf{n}}$, using the same cosines which we already considered for the curvilinear 2-nd type integral.

$$\cos\alpha = \hat{\mathbf{n}} \cdot \hat{\mathbf{i}}, \qquad \cos\beta = \hat{\mathbf{n}} \cdot \hat{\mathbf{j}}, \qquad \cos\gamma = \hat{\mathbf{n}} \cdot \hat{\mathbf{k}}. \tag{8.53}$$

Now we are able to formulate the desired definition.

Def. 4. Consider a continuous vector field $\mathbf{A}(\mathbf{r})$ defined on the surface (S). The surface integral of the second type

$$\int\int_{(S)} \mathbf{A} \, d\mathbf{S} = \int\int_{(S)} \mathbf{A} \cdot \mathbf{n} \, dS \tag{8.54}$$

is a universal (if it exists, of course) limit of the integral sum

$$\sigma = \sum_{i=1}^{n} \mathbf{A}_i \cdot \hat{\mathbf{n}}_i \, S_i \,, \tag{8.55}$$

where both vectors must be taken in the same point $M_i \in (S_i)$: $\mathbf{A}_i = \mathbf{A}(M_i)$ and $\hat{\mathbf{n}}_i = \hat{\mathbf{n}}(M_i)$. Other notations here are the same as in the case of the surface integral of the first type, that is

$$(S) = \bigcup_{i=1}^{n} (S_i) \,, \qquad \lambda = \max \{\text{diam } (S_i)\}$$

and the integral corresponds to the limit $\lambda \to 0$.

By construction, the surface integral of the second type equals to the following double integral over the figure (G) in the uv-plane:

$$\int\int_{(S)} \mathbf{A} \, d\mathbf{S} = \int\int_{(G)} (A_x \cos\alpha + A_y \cos\beta + A_z \cos\gamma) \, \sqrt{g} \, dudv \,. \tag{8.56}$$

Remark. One can prove that the surface integral exists if $\mathbf{A}(u, v)$ and $\mathbf{r}(u, v)$ are smooth functions and the surface (S) has finite area. The sign of the integral changes if we change the orientation of the surface to the opposite one $\hat{\mathbf{n}} \to -\hat{\mathbf{n}}$.

Exercise 6. Calculate the following surface integrals:

$$\mathbf{1)} \qquad \int\int_{(S)} z^2 \, dS \,.$$

(S) is a part of the surface of the cone $x = r \cos\varphi \sin\alpha$, $y = r \sin\alpha \sin\theta$, $z = r \cos\alpha$, where

$$\alpha = const \,, \quad 0 < \alpha < \pi/2 \,, \quad 0 < r < a \,, \quad 0 < \varphi < 2\pi \,.$$

$$\mathbf{2)} \qquad \iint_{(S)} (xz + zy + yx)\, dS\,.$$

(S) is a part of the conic surface $z^2 = x^2 + y^2$, cut by the surface $x^2 + y^2 = 2ax$.

Exercise 7. Calculate the mass of the hemisphere

$$x^2 + y^2 + z^2 = a^2\,, \qquad z > 0\,.$$

if its density at any point $M(x, y, z)$ equals z/a.

Exercise 8. Calculate the moment of inertia with respect to the axis z of the homogeneous spherical shell of the radius a and the surface density ρ_0.

Exercise 9. Consider the spherical shell of the radius R which is homogeneously charged such that the total charge of the sphere is Q. Find the electric potential at an arbitrary point outside and inside the sphere.

Hint. Remember that, since the electric potential is a scalar, the result does not depend on the coordinate system you use. Hence, it is recommended to chose the coordinate system which is the most useful, e.g. settle the center of coordinates at the center of the sphere and put the point of interest at the z axis.

Exercise 10. Derive the following surface integrals of the second type:

$$\mathbf{1)} \qquad \oint\!\!\oint_{(S)} (xdydz + ydxdz + zdxdy)\,,$$

where (S) is an external side of the sphere $x^2 + y^2 + z^2 = a^2$.

$$\mathbf{2)} \qquad \oint\!\!\oint_{(S)} (x^2dydz + y^2dxdz + z^2dydx)\,.$$

where (S) is an external side of the sphere $(x-a)^2 + (y-b)^2 + (z-c)^2 = a^2$.

Try to explain the results using qualitative considerations.

Chapter 9

Theorems of Green, Stokes and Gauss

In this Chapter we shall obtain several important relations between diferent types of integrals.

9.1 Integral theorems

Let us first remember that any well-defined multiple integral can be usually calculated by reducing to the consequent ordinary definite integrals. Consider, for example, the double integral over the region $(S) \subset R^2$. Assume (S) is restricted by the lines

$$\begin{cases} x = a \\ x = b \end{cases}, \quad \begin{cases} y = \alpha(x) \\ y = \beta(x) \end{cases}, \quad \text{where} \quad \alpha(x)\,, \quad \beta(x) \quad \text{are continuous functions}\,. \tag{9.1}$$

Furthermore, we may suppose that the integrand function $f(x, y)$ is continuous. Then

$$\int\int_{(S)} f(x,y)dS = \int_a^b dx \int_{\alpha(x)}^{\beta(x)} dy \cdot f(x,y)\,. \tag{9.2}$$

In a more general case, when the figure (S) can not be described by (9.1), this figure should be divided into the (S_i) sub-figures of the type (9.1). After that one can calculate the double integral using the formula ((9.2)) and use additivity of the double integral

$$\int\int_{(S)} = \sum_i \int\int_{(S_i)}\,.$$

Def. 1. We define the positively oriented contour (∂S^+) restricting the figure (S) on the plane such that moving along (∂S^+) in the positive direction one always has the internal part of (S) on the left.

Observation. This definition can be easily generalized to the case of the $2D$ surface (S) in the $3D$ space. For this end one has to fix orientation $\hat{vecn}$ of (S) and then the **Def. 1** applies directly, by looking from the end of the normal vector $\hat{vecn}$.

Now we are in a position to prove the following important statement:

Theorem. (Green's formula). For the surface (S) dividable to finite number of pieces (S_i) of the form (9.1), and for the functions $P(x, y)$, $Q(x, y)$ which are continuous on (S) together

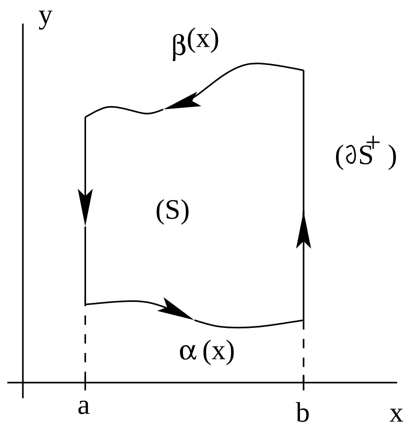

Figure 9.1: Illustration to the demonstration (9.4) of the Green formula.

with their first derivatives

$$P'_x\,, \quad P'_y\,, \quad Q'_x\,, \quad Q'_y\,,$$

the following relation between the double and second type curvilinear integrals holds:

$$\int\!\!\int_{(S)} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy \;=\; \oint_{(\partial S^+)} P\,dx \,+\, Q\,dy\,. \tag{9.3}$$

Proof: Consider the region (S) on the figure, and take

$$\int\!\!\int_{(S)} - \frac{\partial P}{\partial y}\, dxdy \;=\; - \int_a^b dx \int_{\alpha(x)}^{\beta(x)} \frac{\partial P}{\partial y}\, dy$$

Using the Newton-Leibnitz formula we transform this into

$$\begin{aligned} & - \int_a^b \left[P\left(x, \beta(x)\right) - P\left(x, \alpha(x)\right) \right] dx = \\ & = \; - \int_{(23)} Pdx + \int_{(14)} Pdx = \int_{(32)} Pdx + \int_{(14)} Pdx = \oint_{(14321)} Pdx \;=\; \oint_{(\partial S^+)} Pdx\,. \end{aligned} \tag{9.4}$$

In the case the figure (S) does not have the form presented at the Figure, it can be cut into finite number of part which have a given form.

Exercise 5. Prove the Q-part of the Green formula.

Hint. Consider the region restricted by

$$\left\{ \begin{array}{l} y = a \\ y = b \end{array} \right. \qquad \text{and} \qquad \left\{ \begin{array}{l} x = \alpha(y) \\ x = \beta(y) \end{array} \right. .$$

Observation. One can use the Green formula to calculate the area of the surface. For that one can take $Q = x$, $P = 0$ or $Q = 0$, $P = -y$ or $Q = x/2$, $P = -y/2$ etc. Then

$$S = \oint_{(\partial S^+)} xdy \;=\; - \oint_{(\partial S^+)} ydx \;=\; \frac{1}{2} \oint_{(\partial S^+)} xdy - ydx\,. \tag{9.5}$$

Exercise 6. Calculate, using the above formula, the areas of a square and of a circle.

Let us now consider a more general statement. Consider the oriented region of the smooth surface (S) and the right-oriented contour (∂S^+) which bounds this region. We assume that $x^i = x^i(u,v)$, where $x^i = (x,y,z)$, so that the point $(u,v) \in (G)$, $(G) \to (S)$ due to the mapping for the figures and

$$(\partial G^+) \to (\partial S^+)$$

is a corresponding mapping for their boundaries.

Stokes's Theorem. For any continuous vector function $\mathbf{F}(\mathbf{r})$ with continuous partial derivatives $\partial_i F_j$, the following relation holds:

$$\int\int_{(S)} \mathrm{rot}\,\mathbf{F}\cdot d\mathbf{S} \,=\, \oint_{(\partial S^+)} \mathbf{F}\,d\mathbf{r}\,. \tag{9.6}$$

The last integral in (9.6) is called a circulation of the vector field $\mathbf{F}$ over the closed path (∂S^+).

Proof. Let us, as before, denote

$$\mathbf{F}(u,v) = \mathbf{F}(x(u,v),\, y(u,v),\, z(u,v))\,.$$

It proves useful to introduce the third coordinate w such that even outside the surface (S) we could write

$$\mathbf{F} \,=\, \mathbf{F}(x^i) \,=\, \mathbf{F}(\,x^i(u,v,w)\,)\,.$$

The surface (S) corresponds to the value $w = 0$. The operation of introducing the third coordinate may be regarded as a continuation of the coordinate system $\{u,v\}$ to the region of space outside the initial surface[1].

After we introduce the third coordinate, we can also implement a new basis vector $\mathbf{e}_w$. One can always provide that at any point of the space

$$\mathbf{F} = \mathbf{F}(u,v,w) = \,F^u\mathbf{e}_u + F^v\mathbf{e}_v + F^w\mathbf{e}_w\,, \qquad \text{where} \qquad \mathbf{e}_w \perp \mathbf{e}_u \qquad \text{and} \qquad \mathbf{e}_w \perp \mathbf{e}_v\,. \tag{9.7}$$

The components of the vectors $\mathbf{F}$, $\mathrm{rot}\,\mathbf{F}$, $d\mathbf{S}$ and $d\mathbf{r}$ transform in a standard way, such that the products

$$\mathrm{rot}\,\mathbf{F}\cdot d\mathbf{S} \,=\, \varepsilon^{ijk}\,(dS_i)\,\partial_j F_k \qquad \text{and} \qquad \mathbf{F}\cdot d\mathbf{r} = F_i dx^i$$

are scalars with respect to the coordinate transformation from (x,y,z) to (u,v,w). Therefore, the relation (9.6) can be proved in any particular coordinate system. If being correct in one coordinate system, it is correct for any other coordinates too. In what follows we shall use the coordinates (u,v,w). Then

$$\oint_{(\partial S^+)} \mathbf{F}d\mathbf{r} = \oint_{(\partial G^+)} \mathbf{F}d\mathbf{r} = \oint_{(\partial G^+)} F_u du + F_v dv\,, \tag{9.8}$$

[1] As an example we may take internal coordinates φ, θ on the sphere. We can introduce one more coordinate r and the new coordinates φ, θ, r will be, of course, the spherical ones, which cover all the $3D$ space. The only one difference is that the new coordinate r, in this example, is not zero on the surface $r = R$. But this can be easily corrected by the redefinition $\rho = r - R$, where $-R < \rho < \infty$.

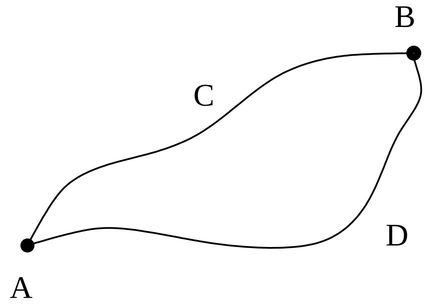

Figure 9.2: Two possible curves of integration between the points A and B.

because the contour (∂G^+) lies in the (u, v) plane.

Furthermore, since the figure (G) belongs to the (u, v) plane, the oriented area vector $d\mathbf{G}$ is parallel to the $\mathbf{e}_w$ axis. Therefore, due to the Eq. (9.7),

$$(d\mathbf{G})_u = (d\mathbf{G}_v) = 0\,. \tag{9.9}$$

Then

$$\int\!\!\int_{(S)} \text{rot}\,\mathbf{F}\cdot d\mathbf{S} \;=\; \int\!\!\int_{(G)} \text{rot}\,\mathbf{F}\cdot d\mathbf{G} \;=\; \int_{(G)} (\,\text{rot}\,\mathbf{F})_w\; dG^w \;=$$

$$= \int\!\!\int_{(G)} \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) dudv\,. \tag{9.10}$$

It is easy to see that the remaining relation

$$\oint_{(\partial G^+)} F_u du + F_v dv \;=\; \int\!\!\int_{(G)} \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) dudv \tag{9.11}$$

is nothing but the Green formula which has been already proved in the previous Theorem. The proof is complete.

There is an important consequence of the Stokes Theorem.

Theorem. For the $\mathbf{F}(\mathbf{r}) = \text{grad}\, U(\mathbf{r})$, where U is some smooth function, the curvilinear integral between two points A and B doesn't depend on the path (AB) and is simply

$$\oint_{(AB)} \mathbf{F}\, d\mathbf{r} \;=\; U(B) - U(A)\,.$$

Proof. Consider some closed simple path (Γ) which can be always regarded as (∂S^+) for some oriented surface (S). Then, according to the Stokes theorem

$$\oint_{(\Gamma)} \mathbf{F}\, d\mathbf{r} \;=\; \int\!\!\int_{(S)} \text{rot}\,\mathbf{F}\cdot d\mathbf{S}\,. \tag{9.12}$$

If $\mathbf{F} = \text{grad}\, U$, then $\text{rot}\,\mathbf{F} = 0$ and $\oint_{(\Gamma)} \mathbf{F} d\mathbf{r} = 0$.

Let us now consider two different paths between the points A and B: (ACB) and (ADB). As we have just seen, $\oint_{(ABCDA)} \mathbf{F} d\mathbf{r} = 0$. But this also means that

$$\int_{(ACB)} \mathbf{F}\, d\mathbf{r} - \int_{(ADB)} \mathbf{F}\, d\mathbf{r} \,=\, 0\,, \tag{9.13}$$

that completes the proof.

Observation 1. For the $2D$ case this result can be seen just from the Green's formula.

Observation 2. A useful criterion of $\mathbf{F}$ being $\operatorname{grad} U$ is

$$\frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x}\,.$$

and the same for any couple of partial derivatives. The reason is that

$$\frac{\partial F_z}{\partial x} = \frac{\partial^2 U}{\partial x \partial z} = \frac{\partial^2 U}{\partial z \partial x} = \frac{\partial F_x}{\partial z}\,,$$

provided the second derivatives are continuous functions. In total, there are three such relations

$$\frac{\partial F_z}{\partial x} = \frac{\partial F_x}{\partial z}\,, \qquad \frac{\partial F_y}{\partial x} = \frac{\partial F_x}{\partial y}\,, \qquad \frac{\partial F_z}{\partial y} = \frac{\partial F_y}{\partial z}\,. \tag{9.14}$$

The last of the Theorems we are going to prove in this section is

Gauss-Ostrogradsky Theorem. Consider a $3D$ figure $(V) \subset R^3$ and also define (∂V^+) to be the externally oriented boundary of (V). Consider a vector field $\mathbf{E}(\mathbf{r})$ defined on (V) and on its boundary (∂V^+) and suppose that the components of this vector field are continuous, as are their partial derivatives

$$\frac{\partial E^x}{\partial x}\,, \qquad \frac{\partial E^y}{\partial y}\,, \qquad \frac{\partial E^z}{\partial z}\,.$$

Then these components satisfy the following integral relation

$$\oint\!\!\oint_{(\partial V^+)} \mathbf{E} \cdot d\mathbf{S} \,=\, \int\!\!\int\!\!\int_{(V)} \operatorname{div} \mathbf{E}\; dV\,. \tag{9.15}$$

Proof. Consider $\int\int\int_{(V)} \partial_z E_z dV$. Let's suppose, similar to that as we did in the proof of the Green's formula, that (V) is restricted by the surfaces $z = \alpha(x,y)$, $z = \beta(x,y)$ and by some closed cylindric surface with the projection (∂S) (boundary of the figure (S)) on the xy plane. Then

$$\int\!\!\int\!\!\int_{(V)} \frac{\partial E_z}{\partial z} dV \,=\, \int\!\!\int_{(S)} dS \int_{\alpha(x,y)}^{\beta(x,y)} dz \frac{\partial E_z(x,y,z)}{\partial z} \,=$$

$$= \int\!\!\int_{(S)} dx\, dy\, \Big[- E_z\,(x,\, y,\, \alpha(x,y)) + E_z\,(x,\, y,\, \beta(x,y)) \Big] \,=\, \oint\!\!\oint_{(\partial V^+)} E_z dS^z\,.$$

In the same way one can prove the other three relations, that completes the proof.

9.2 Div, grad and rot from new point of view

Using the Stokes and Gauss-Ostrogradsky theorems, one can give more geometric definitions of divergence and rotation of the vector. Suppose we want to know the projection of $\mathrm{rot}\,\mathbf{F}$ on the direction of the unit vector $\hat{\mathbf{n}}$. Then we have to take some infinitesimal surface vector $d\mathbf{S} = dS\,\hat{\mathbf{n}}$. In this case, due to the continuity of all components of the vector $\mathrm{rot}\,\mathbf{F}$, we have

$$\hat{\mathbf{n}} \cdot \mathrm{rot}\,\mathbf{F} \;=\; \lim_{dS\to 0} \frac{d}{dS} \oint_{(\partial S^+)} \mathbf{F} \cdot d\mathbf{r} \qquad \text{where} \quad \left(\partial S^+\right) \quad \text{is a border of (S)}\,.$$

The last relation provides a necessary food for the geometric intuition of what $\mathrm{rot}\,\mathbf{F}$ is.

In the case of $\mathrm{div}\,\mathbf{A}$ one meets the following relation

$$\mathrm{div}\,\mathbf{A} \;=\; \lim_{V\to 0} \frac{1}{V} \oiint_{(\partial V^+)} \mathbf{A}\, d\mathbf{S}\,,$$

where the limit $V \to 0$ must be taken in such a way that also $\lambda \to 0$, where $\lambda = \mathrm{diam}\,(V)$. This formula indicates to the relation between $\mathrm{div}\,\mathbf{A}$ at some point and the presence of the source for this field at this point. If such source is absent, $\mathrm{div}\,\mathbf{A} = 0$.

The last two formulas make the geometric sense of div and rot operators explicit. In order to complete the story, let us explain the geometric sense of the remaining differential operator grad. For a scalar field $\varphi(\mathbf{r})$, the $\mathrm{grad}\,\varphi(\mathbf{r})$ is the vector field which satisfies the following conditions:

1) It is directed according to the maximal growth of φ.

2) It has the absolute value equal to the maximal derivative of $\varphi(\mathbf{r})$ with respect to the natural parameter along the corresponding curve, in all possible directions.

Exercise 1. Using the Stokes formula, derive the curvilinear integrals:

$$\mathbf{1)} \qquad \oint_{(C)} (ydx + zdy + zdz)\,,$$

where (C) is a circumference $x^2 + y^2 + z^2 = a^2$, $x + y + z = 0$, which is run counterclockwise if looking from the positive side of the axis Ox;

$$\mathbf{2)} \qquad \oint_{(C)} (x^2 - yz)dx + (y^2 - xz)dy + (z^2 - xy)dz\,,$$

where (C) is a part of the spiral line $x = a\,\cos\varphi,\; y = a\,\sin\varphi,\; z = \frac{h}{2\pi}\,\varphi$;

$$\mathbf{3)} \qquad \oint_{(C)} \begin{vmatrix} dx & dy & dz \\ \cos\alpha & \cos\beta & \cos\gamma \\ x & y & z \end{vmatrix}\,,$$

where the contour (C) is situated in a plane $x\cos\alpha + y\cos\beta + z\cos\gamma - p = 0$ and restricts the figure with the area S.

Exercise 2. Using the Gauss-Ostrogradsky formula, derive the following surface integrals:

$$\mathbf{1)} \qquad \oiint_{(S)} (x^2 dydz + y^2 dxdz + z^2 dxdy)\,,$$

where (S) is an external surface of the cube $0 \le x \le a,\ 0 \le y \le a,\ 0 \le z \le a$;

$$\textbf{2)} \qquad \int\!\!\int_{(S)} (x^2 \cos\alpha + y^2 \cos\beta + z^2 \cos\gamma)\, dS\,,$$

where (S) is a part of the conic surface $x^2 + y^2 = z^2,\ ,\ 0 \le z \le h$ and $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are the components of the normal vector to the surface.

Exercise 3. Using the Gauss-Ostrogradsky formula, prove the following statement. Assume (S) is a smooth surface which is a boundary of a finite $3D$ body (V), $u(x,y,z)$ is a smooth function on $(V) \cup (S)$ and

$$\Delta\, u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\,.$$

Then the following relations hold:

$$\textbf{1)} \qquad \oint\!\!\oint_{(S)} \frac{\partial u}{\partial n}\, dS = \int\!\!\int\!\!\int_{(V)} \Delta\, u\, dxdydz\,;$$

$$\textbf{2)} \qquad \oint\!\!\oint_{(S)} u\, \frac{\partial u}{\partial n}\, dS = \int\!\!\int\!\!\int_{(V)} \Big[\Big(\frac{\partial u}{\partial x}\Big)^2 + \Big(\frac{\partial u}{\partial y}\Big)^2 + \Big(\frac{\partial u}{\partial y}\Big)^2 + u\, \Delta\, u \Big]\, dxdydz\,.$$

where (S) is a part of the conic surface $x^2 + y^2 = z^2,\ ,\ 0 \le z \le h$ and $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are the components of the normal vector to the surface.

Bibliography

[1] A.T. Fomenko, S.P. Novikov and B. A. Dubrovin, *Modern Geometry-Methods and Applications, Part I: The Geometry of Surfaces, Transformation Groups, and Fields*, (Springer, 1992).

[2] A.J. McConnell, Applications of Tensor Analysis. (Dover, 1957).

[3] I.S. Sokolnikoff, *Tensor Analysis: Theory and Applications to Geometry and Mechanics of Continua*, (John Wiley and Sons, N.Y. 1968).

[4] A. I. Borisenko and I. E. Tarapov, *Vector and Tensor Analysis with Applications*, (Dover, N.Y. 1968 - translation from III Russian edition).

[5] B. F. Schutz, *Geometrical Methods of Mathematical Physics*, (Cambridge University Press, 1982).

[6] A. Z. Petrov, *Einstein Spaces*, (Pergamon, 1969).

[7] S. Weinberg, *Gravitation and Cosmology*, (John Wiley and Sons. Inc., 1972).

[8] V.V. Batygin and I. N. Toptygin, *Problems in Electrodynamics*, (Academic Press, 1980).

[9] A.P. Lightman, W.H. Press, R.H. Price, S.A. Teukolsky, *Problem book in Relativity and Gravitation,* (Princeton University Press, 1975).

[10] B. P. Demidovich, *5000 Problemas de Analisis Matematico*, (Paraninfo, 1992, in Spanish), also *Collection of Problems and Exercises on Mathematical Analysis*, (Mifril, 1995, In Russian).

[11] V.S. Vladimirov, *Equations of Mathematical Physics*, (Imported Pubn., 1985).

[12] J. J. Sakurai, *Modern Quantum Mechanics*, (Addison-Wesley Pub Co, 1994).

[13] L.D. Landau and E.M. Lifshits, *Mechanics - Course of Theoretical Physics, Volume 1*, (Butterworth-Heinemann, 1976).

[14] L.D. Landau and E.M. Lifshits, *The Classical Theory of Fields - Course of Theoretical Physics, Volume 2*, (Butterworth-Heinemann, 1987).

[15] H. Goldstein, *Classical Mechanics*, (Prentice Hall, 2002).

[16] G.B. Arfken and H.J. Weber, *Mathematical Methods for Physicists*, (Academic Press, 1995).

[17] P.M. Morse e H. Fishbach, *Methods of Theoretical Physics*, (McGraw-Hill, 1953).