Contents

Chapter 1

Inroduction

1.1 Continuum theory

Matter is formed of molecules which in turn consist of atoms and sub-atomic particles. Thus matter is not continuous. However, there are many aspects of everyday experience regarding the behaviors of materials, such as the deflection of a structure under loads, the rate of discharge of water in a pipe under a pressure gradient or the drag force experienced by a body moving in the air etc., which can be described and predicted with theories that pay no attention to the molecular structure of materials. The theory which aims at describing relationships between gross phenomena, neglecting the structure of material on a smaller scale, is known as continuum theory. The continuum theory regards matter as indefinitely divisible. Thus, within this theory, one accepts the idea of an infinitesimal volume of materials referred to as a particle in the continuum, and in every neighborhood of a particle there are always neighbor particles. Whether the continuum theory is justified or not depends on the given situation; for example, while the continuum approach adequately describes the behavior of real materials in many circumstances, it does not yield results that are in accord with experimental observations in the propagation of waves of extremely small wavelength. On the other hand, a rarefied gas may be adequately described by a continuum in certain circumstances. At any case, it is misleading to justify the continuum approach on the basis of the number of molecules in a given volume. After all, an infinitesimal volume in the limit contains no molecules at all. Neither is it necessary to infer that quantities occurring in continuum theory must be interpreted as certain particular statistical averages. In fact, it has been known that the same continuum equation can be arrived at by different hypothesis about the molecular structure and definitions of gross variables. While molecular-statistical theory, whenever available, does enhance the understanding of the continuum theory, the point to be made is simply that whether the continuum theory is justified in a given situation is a matter of experimental test, not of philosophy. Suffice it to say that more than a hundred years of experience have justified such a theory in a wide variety of situations.

1.2 Continuum Mechanics

The analysis of the kinematic and mechanical behavior of materials modeled on the continuum assumption is what we know as continuum mechanics. There are two main themes into which the topics of continuum mechanics are divided. In the first, emphasis is on the derivation of fundamental equations which are valid for all continuous media. These equations are based upon universal laws of physics such as the conservation of mass and the principles of energy and momentum. In the second, attention focuses on the development of constitutive equations which characterize the behavior of specific idealized materials, of which the perfectly elastic solid and the viscous fluid are the best known examples. These equations provide the focal points for studies in elasticity, plasticity, viscoelasticity, and fluid mechanics. Mathematically, the fundamental equations of continuum mechanics mentioned above may be developed in two separate but essentially equivalent formulations. One, the integral or global form, derives from the basic principles applied to a finite volume of the material. The other, a differential or field approach, leads to equations which result from the basic principles applied to a very small (infinitesimal) element of volume. In practice, it is often useful and convenient to deduce the field equations from their global counterparts. In addition to the fundamental assumption of material continuity, we impose two further restrictions on the bodies considered in this text. First, we require them to be homogeneous, that is, to possess identical mechanical properties at all locations. And second, we consider in general only those materials which are isotropic and which thereby have identical physical properties in every direction at a given point. Anisotropic bodies will be mentioned only briefly. The continuum assumption is the only one needed in the derivation of the general (field) equations. Those of homogeneity and isotropy enter into the theory with the introduction of constitutive equations. Boundary value problems in technical disciplines rooted in continuum theory are formulated in terms of the basic field equations together with the appropriate constitutive equations and relevant boundary conditions. Linear elasticity and classical fluid mechanics are the best known of these disciplines, and an abbreviated discussion of these in the context of continuum mechanics is given in this course.

1.3 Essential mathematics

1.3.1 Scalars, Vectors, and Tensors

LEARNING a discipline's language is the first step a student takes towards becoming competent in that discipline. The language of continuum mechanics is the algebra and calculus of tensors. Here, tensors is the generic name for those mathematical entities which are used to represent the important physical quantities of continuum mechanics. Only that category of tensors known as Cartesian tensors is used in this text, and definitions of these will be given in the pages that follow. The tensor equations used to develop the fundamental theory of continuum mechanics may be written in either of two distinct notations the symbolic notation or the indices notation. We shall make use of both notations, employing whichever is more convenient for the derivation or analysis at hand, but taking care to establish the interrelationships between the two. As it happens, a considerable variety of physical and geometrical quantities have important roles in continuum mechanics, and fortunately, each of these may be represented by some form of tensor. For example, such quantities as density and temperature may be specified completely by giving their magnitude, that is, by stating a numerical value. These quantities are represented mathematically by scalars, which are referred to as zeroth-order tensors. It should be emphasized that scalars are not constants, but may actually be functions of position or time or both. Also, the exact numerical value of a scalar will depend upon the units in which it is expressed. Thus the temperature at a certain location may be given by either 68^0F or 20^0C. As a general rule, lowercase Greek letters in italic print such as a, A,), etc. will be used as symbols for scalars in both the indicial and symbolic notations. Several physical quantities of mechanics such as force and velocity require not only an assignment of magnitude, but also a specification of direction for their complete characterization. As a trivial example, a $20 - N$ force acting vertically at a point is substantially different than a $20 - N$ force acting horizontally at the point. Quantities possessing such directional properties are represented by vectors, which are first-order tensors.

Geometrically, vectors are displayed as arrows, having a definite length (the magnitude), a specified orientation (the direction), and also a sense of action as indicated by the head and the tail of the arrow. Certain quantities in mechanics which are not truly vectors are also portrayed by arrows, for example, finite rotations. Consequently, in addition to the magnitude and direction characterization, the complete definition of a vector requires the further statement that "vectors add (and subtract) in accordance with the triangle rule by which the arrow representing the vector sum of two vectors extends from the tail of the first component arrow to the head of the second when the component arrows are arranged 'head-to-tail.' " Although vectors are independent of any particular coordinate system, it is often useful to define a vector in terms of its coordinate components, and in this respect it is necessary to reference the vector to an appropriate set of axes. In view of our restriction to Cartesian tensors, we limit ourselves to consideration of Cartesian coordinate systems for designating the components of a vector. A significant number of physical quantities having important status in continuum mechanics require mathematical entities of higher order than vectors for their representation in the hierarchy of tensors. As we shall see, among the best known of these are the stress tensor and the strain tensors. These particular tensors are second-order tensors, said to have a rank of two. Third-order and fourth-order tensors are not uncommon in continuum mechanics, but they are not nearly as plentiful as second-order tensors. Henceforth, the unqualified use of the word tensor in this text means second-order tensor. With only a few exceptions, primarily those which represent the stress and strain tensors, we shall denote second-order tensors by upper-case Latin letters in bold-faced print, a typical example being the tensor T. Tensors, like vectors, are independent of any coordinate system, but just as with vectors, when we wish to specify a tensor by its components we are obliged to refer to a suitable set of reference axes. The precise definitions of tensors of various order will be given subsequently in terms of the transformation properties of their components between two related sets of Cartesian coordinate

1.3.2 Conventions

As mentioned in the introduction, all laws of continuum mechanics must be formulated in terms of quantities that are independent of coordinates. It is the purpose of this chapter to introduce such mathematical entities. We shall begin by introducing a short-hand notation - the indicial notation, which will be followed by the concept of tensors introduced as a linear transformation. The basic field operations needed for continuum formulations and their representations in curvilinear coordinates are presented after that.

Summation Convent

Consider the sum

$$s = a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 + \cdots + a_n \cdot x_n$$

We can write the above equation in a compact form by using the summation sign:

$$s = \sum_{i=1}^{n} a_i \cdot x_i, \quad \text{or} \quad s = \sum_{m=1}^{n} a_m \cdot x_m.$$

The index i or m in these equations is a dummy index in the sense that the sum is independent of the letter used.

2. Consider the linear transformation of coordinates. Let a point P have coordinates
$(x_1, x_2, \ldots, x_n)$. The new $x_1', x_2', \ldots, x_n'$ coordinates of P may be expressed according the next

equations

$$x_i^{'} = \sum_{j=1}^{n} l_{ij} \cdot x_j, \tag{1.1}$$

The suffix j appears **twice in the sums on** the right-hand sides and sum over all possible values of j. This situation occurs so frequently that it is convenient adopt convention which often avoids the necessity of writing summation sing.

- *SUMMATION CONVENTION. Whenever an index is repeated once, it is a dummy index indicating a summation with the index running through the integers 1,2, ..., n.*

This convention is known as Einstein's summation convention. Using the summation convention equation (1.1) becomes simply

$$x_i^{'} = l_{ij} \cdot x_j, \tag{1.2}$$

For example:

$$a_{ii} = a_{mm} = a_{11} + a_{22} + a_{33}$$

$$a_i e_i = a_1 e_1 + a_2 e_2 + a_3 e_3.$$

The summation convention obviously can be used to express a double sum, a triple sum. etc.

$$\sum_{i=1}^{3} \cdot \sum_{j=1}^{3} a_{ij} x_i x_j \quad \Longleftrightarrow \quad a_{ij} x_i x_j$$

When the summation convention is in use, care must be taken to avoid using any suffix more than twice in the same equation.

Free indices

Consider the following system of three equation

$$\begin{array}{rl} x_1^{'} = & a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ x_2^{'} = & a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ x_3^{'} = & a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{array} . \tag{1.3}$$

Using the summation convention (1.3) can be written as

$$\begin{array}{rl} x_1^{'} = & a_{1m}x_m \\ x_2^{'} = & a_{2m}x_m \\ x_3^{'} = & a_{3m}x_m \end{array} .$$

Which can be shortened into

$$x_i^{'} = \quad a_{im}x_m, \qquad i = 1, 2, 3. \tag{1.4}$$

An index which appears only once in each term of an equation such as the index in eq. (1.4) is called a *"free index"*.

A further example is given by

$$e_m^{'} = Q_{mi} e_i, \qquad m = 1, 2, 3$$

$$e_l^{'} = Q_{li} e_i, \qquad l = 1, 2, 3.$$

$a_i = b_j$ is a meaningless equation,
$T_{ij} = A_{im} A_{jm} \qquad i, j = 1, 2, 3.$
But equation such as $T_{ij} = T_{ik}$ have no meaning.

Kronecker Delta

Kronecker delta is defined by

$$\delta_{ij} = \begin{cases} 0, & \text{when } i \neq j, \\ 1 & \text{when } i = j, \end{cases} \tag{1.5}$$

That is,

$$\delta_{11} = \delta_{22} = \delta_{33} = 1,$$
$$\delta_{12} = \delta_{13} = \delta_{21} = \delta_{23} = \delta_{31} = \delta_{32} = 0.$$

In other words, the matrix of the Kronecker delta is the identity matrix, i.e.

$$[\delta_{ij}] = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We also will use the Kronecker delta with upper, down, and mixed suffix (upper and down)

$$\delta^{i\cdot}_{\cdot j} = \delta^{\cdot j}_{i\cdot} = \delta^{ij}$$

$$\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3,$$
$$\delta_{1m}a_m = \delta_{11}a_1 + \delta_{12}a_2 + \delta_{13}a_3 = a_1$$
$$\delta_{2m}a_m = \delta_{21}a_1 + \delta_{22}a_2 + \delta_{23}a_3 = a_2$$
$$\delta_{im}T_{mj} = \delta_{i1}T_{1j} + \delta_{i2}T_{2j} + \delta_{i3}T_{3j} = T_{ij}$$
$$\delta_{im}\delta_{mj} = \delta_{ij}.$$

Manipulation with the Indicial Notation

Substitution

If $a_i = U_{im}b_m$ and $b_i = V_{im}e_m$. In order to substitute the b_i's into a_i's we first change the free index from i to m and dummy index m to some other letter, say n so that

$$b_m = V_{mn}e_n \qquad \Longrightarrow \qquad a_i = U_{im}V_{mn}e_n.$$

Multiplication

If $p = a_m b_m$ and $q = c_m d_m$ then $pq = a_m b_m c_n d_n$.

It is important to note that $pq \neq a_m b_m c_m d_m$.

Example 1. If $\overline{a} = a_i\overline{e_i}$ and $\overline{b} = b_i\overline{e_i}$ then the dot product of vectors $\overline{a}$ and $\overline{b}$ is

$$\overrightarrow{a} \cdot \overrightarrow{b} = (a_i\overrightarrow{e_i}) \cdot (b_j\overrightarrow{e_j}) = a_i b_j(\overrightarrow{e_i} \cdot \overrightarrow{e_j}).$$

In particular, if $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}$ are unit vectors perpendicular to one another, then $\overrightarrow{e_i} \cdot \overrightarrow{e_j} = \delta_{ij}$, and $\overrightarrow{a} \cdot \overrightarrow{b} = a_i b_j \delta_{ij} = a_i b_i = a_j b_j = a_1 b_1 + a_2 b_2 + a_3 b_3$.

Factoring

If $T_{ij}n_j - \lambda n_i = 0$ we can write $n_i = \delta_{ij}n_j$ and we will have

$$T_{ij}n_j - \lambda\delta_{ij}n_j = 0 \qquad \Longrightarrow \qquad (T_{ij} - \lambda\delta_{ij})n_j = 0.$$

Contraction

The operation of identifying two indices and so summing on them is known as contraction. For example, T_{ii} is contraction of T_{ij}

$$T_{ii} = T_{11} + T_{22} + T_{33}.$$

If $\quad T_{ij} = \lambda\Theta\delta_{ij} + 2\mu E_{ij} \quad$, then $T_{ii} = \lambda\Theta\delta_{ii} + 2\mu E_{ii} = 3\lambda\Theta + 2\mu E_{ii}$.

1.3.3 Linear algebra

Basic definitions and examples

We consider a vector space V over the field R, where R is a set of real numbers.

The operation of adding two vectors in **V**.

First, note that the result is again a vector in **V**.

$$\forall \overrightarrow{v}, \overrightarrow{w} \in \mathbf{V} \Longrightarrow \overrightarrow{v} + \overrightarrow{w} \in \mathbf{V} \quad \text{closure under addition} \tag{1.6}$$

$$(\overrightarrow{v} + \overrightarrow{w}) + \overrightarrow{z} = \overrightarrow{v} + (\overrightarrow{w} + \overrightarrow{z}) \text{ associative property} \tag{1.7}$$

$$\overrightarrow{v} + \overrightarrow{w} = \overrightarrow{w} + \overrightarrow{v} \quad \text{commutative property} \tag{1.8}$$

There is the vector $\overrightarrow{0} = (0, 0, \dots, 0)$ with the property

$$\overrightarrow{v} + \overrightarrow{0} = \overrightarrow{0} + \overrightarrow{v} = \overrightarrow{v} \tag{1.9}$$

The vector $\overrightarrow{0}$ is called a zero vector (an additive identity).

Furthermore, given any vector $\overrightarrow{v} \in R^n$, there is another vector $(-\overrightarrow{v})$, so that

$$\overrightarrow{v} + (-\overrightarrow{v}) = (-\overrightarrow{v}) + \overrightarrow{v} = \overrightarrow{0} \tag{1.10}$$

The vector $(-\overrightarrow{v})$ is called an additive inverse for $\overrightarrow{v}$.

Any set V with an operation + satisfying the five properties given (1.6-1.9) is called an **Abelian group.**

Thus **V** is an abelian group under the operation of vector addition.

There is another operation in **V** , and that is the operation of multiplying a real number (scalar) times a vector to get a new vector. Given $a \in R$, $\overrightarrow{v} \in \mathbf{V}$, then

$$a\overrightarrow{v} \in R^n - \text{closure under scalar multiplication.} \tag{1.11}$$

This scalar multiplication obeys some properties:

$$a(b\overrightarrow{v}) = b(a\overrightarrow{v}) \quad \text{associative property} \tag{1.12}$$

$$a(\overrightarrow{v} + \overrightarrow{w}) = a\overrightarrow{v} + a\overrightarrow{w} \quad \text{distributive property 1} \tag{1.13}$$

$$(a + b)\overrightarrow{v} = a\overrightarrow{v} + b\overrightarrow{v} \quad \text{distributive property 2} \tag{1.14}$$

$$1 \cdot \overrightarrow{v} = \overrightarrow{v} \forall \overrightarrow{v}, \tag{1.15}$$

where 1 is the identity of scalar multiplication.

Definition 1.1. *A set* **V** *that forms an Abelian group under + and has an operation of multiplication on real number a by an element* $\overrightarrow{v} \in \mathbf{V}$ *to get another element* $a\overrightarrow{v} \in \mathbf{V}$ *satisfying the properties (1.12)-(1.15) is called a vector space (over the real numbers).*

Basis

Definition 1.2. *(linearly dependent) A set of $n-$th vectors $\overrightarrow{a_1}, \overrightarrow{a_2}, \ldots, \overrightarrow{a_n}$ is said to be linearly dependent if and only if there are $n-$ real numbers $\beta_1, \beta_2, \ldots, \beta_n$ not all zero such that $\beta_i \overrightarrow{a_i} = 0$. Otherwise it is called linearly independent.*

A vector space $\mathbf{V}$ is said to be of finite dimension or to be n-dimensional, written as $\dim \mathbf{V} = n$, if there exists a set of linearly independent vectors $\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_n}$, which spans $\mathbf{V}$.

Definition 1.3. *(basis) Any set of linearly independent vectors $\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_n}$ of n-dimensional space is called a BASIS.*

Confirmation of this definition is the foolowing theorem.

Theorem 1.1.

Let $\mathbf{V}$ be n-dimensional vector space. Then
(i) any set of $n+1$ or more vectors is linearly dependent;
(ii) any linearly independent set of vectors can be a part of a basis, i.e., can be extended to a basis;
(iii) any linearly independent set of vectors composes a basis.

Theorem 1.2. *The vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ in a vector space V are linearly independent if $c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} + \ldots + c_n\overrightarrow{v_n} = 0$ if and only if all $c_i = 0$.*

Example 2. $(2,3,5), (1,4,2), (1,-1,3) \in R^3$.

We form the linear combination

$$c_1(2,3,5) + c_2(1,4,2) + c_3(1,-1,3) = (0,0,0).$$

$$\begin{array}{ll} c_1 \cdot 2 + c_2 \cdot 1 + c_3 \cdot 1 & = 0 \\ c_1 \cdot 3 + c_2 \cdot 4 + c_3 \cdot -1 & = 0 \\ c_1 \cdot 5 + c_2 \cdot 2 + c_3 \cdot 3 & = 0 \end{array} \Longrightarrow \begin{pmatrix} 2 & 1 & 1 \\ 3 & 4 & -1 \\ 5 & 2 & 3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or $A\overrightarrow{c} = \overrightarrow{0}$, but $\det(A) = 0$ we can find $c = (-1,1,1)$ such that -1$\cdot(2,3,5) + 1 \cdot (1,4,2) + 1 \cdot (1,-1,3) = (0,0,0)$.

This vectors are linearly dependent.

(Algorithm for Independence). If $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ is a collection of vectors in R^n,then they are independent iff the matrix equation $A\overrightarrow{e} = \overrightarrow{0}$ has only trivial solution, where A is the matrix whose $j-$th column vector is $\overrightarrow{v_j}$.

Example 3. $p_1 = 2 - 3 \cdot x^2, \quad p_2 = 1 + 2x - x^2, p_3 = 1 + x + x^2$

To check whether they are independent, we write $c_1p_1 + c_2p_2 + c_3p_3$ and solve for (c_1, c_2, c_3)

$c_1(2 - 3x^2) + c_2(1 + 2x - x^2) + c_3(1 + x + x^2) = 0$.

Collecting terms gives

$(2c_1 + c_2 + c_3) + (2c_2 + c_3)x + (-3c_1 - c_2 + c_3)x^2 = 0$

we have the matrix equation $A\overrightarrow{c} = \overrightarrow{0}$

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ -3 & -1 & 1 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 9/4 \end{pmatrix}; det(A) \neq 0$$

So that the only solution is $\overrightarrow{c} = 0$. Thus the three polynomials are independent.

Example 4. (basis) The four matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Form a basis for all 2 by 2 matrices

Example 5. The polynomials $1, x, \cdots, x^n$ form a basis for the vector space $P^n(R,R)$ of polynomials of degree less than or equal to n.

Exercise 1.
For each part determine whether the given vectors are linearly independent
(a) $(1,2,3),(4,5,6),(7,8,9)$
(b) $(2,1,4),(3,9,2),(3,6,2),(4,7,2)$
(c) $(1,1,1),(1,1,0),(1,0,0)$
(d) $(1,5,2,7),(3,2,1,5),(3,2,1,1),(2,6,2,1)$

We denote basis as $\{e_i\}_1^n$ or if the value n is known from text we will denote basis simply as $\{e_i\}$.

For instance: basis vectors $\overrightarrow{e_1} = (1,0,\dots,0)$, $\overrightarrow{e_1} = (1,0,\dots,0)$, $\overrightarrow{e_1} = (1,0,\dots,0),\dots,\overrightarrow{e_1} = (1,0,\dots,0)$ are form the basis of the $n-$dimensional vector space.

Remark 1.5. *If we use notation Z_i^j for matrix then upper index j is a number of row, and the lower index i is a number of column.*

COVARIANT and CONTRAVARIANT coordinates

Let $\{e_i\}_1^n$ be a basis in $\mathbf{V}$. We define the set of vectors $\{e^i\}_1^n$ which satisfies the property:

$$(e^i, e_j) = \delta_j^i \ (i,j = 1,2,\dots,n)$$

Exercise 2. *Prove that $\{e^i\}_1^n$ is a basis.*

The basis $\{e^i\}_1^n$, which corresponds to the basis $\{e_i\}_1^n$ is called a co–basis.

Exercise 3. *Prove that if $\{e_i\}$ is an orthogonal basis, then the co-basis $\{e^i\}$ is also orthogonal.*

If $e^i = e_i, \ \ i = 1,2,...,n$ (if the basis $\{e_i\}$ coincides with its cobasis $\{e^i\}$) in this case the basis $\{e_i\}$ is orthonormal basis (orthobasis).

Remark 1.5. *An orthonormal basis we will call an orthobasis.*

Every element of vector spce $p \in V$ has a unique expression as linear combination

$$p = p^i e_i, \qquad p = p_i e^i.$$

Definition 1.4. *The coordinates p_i are called covariant cooordinates of p with respect to the cobasis $\{e^i\}$, the coordinates p^i are called contravariant components of p with respect to the basis $\{e_i\}$.*

Exercise 4.
(a) Show that the vectors $(1,1,1),(1,1,0),(1,0,0)$ are basis of R^3.
(b) Find the cobasis corresponding to a given basis.
(c) let $p = (1,0,1)$ be a vector of R^3. Find covariant and contravariant components of p with respect to the basis and cobasis.

Let $\{e_i\}$ is the "old" basis and $\{e_i^{'}\}$ is the "new" one. We can write

$$e_i^{'} = A_i^j e_j; \qquad e^{'i} = \overline{A_j^i} e^j$$

Definition 1.5. *A matrix $(A) = (A_i^j)$ is called a transition matrix from the basis $\{e_i^{'}\}$ to the basis $\{e_i\}$.*

Theorem 1.3. *The matrix $\overline{A} = (\overline{A_j^i})$ is an inverse matrix of A ($\overline{A} = A^{-1}$).*

Proof. We have $e_i^{'} = A_i^j e_j$ and $e^{'i} = \overline{A_j^i} e^j$. We have to show $(\overline{A})$=(A^{-1}) or
$\overline{A_j^i} \cdot A_l^j = \delta_l^i$,
$e^{'i} = \overline{A_j^i} e^j$

$$e^{'i} \cdot e_l^{'} = \overline{A_j^i} e^j \cdot e_l^{'} = \overline{A_j^i} e^j \cdot A_l^m e_m = \overline{A_j^i} A_l^m e^j \cdot e_m = \overline{A_j^i} A_l^m \delta_m^j = \overline{A_j^i} A_l^j = \delta_l^i$$

Example 6. Let $\{e_i^{'}\}_1^3$ be $e_1^{'} = (1,-1,3),\ e_2^{'} = (2,1,0),\ e_3^{'} = (1,1,1)$ and $\{e_i\}_1^3$ is standard basis of R^3 $e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)$.

Then A_i^j is easy to find, since we can write

$$\begin{aligned} e_1^{'} &= 1 \cdot e_1 - 1 \cdot e_2 + 3 \cdot e_3 \\ e_2^{'} &= 2 \cdot e_1 + 1 \cdot e_2 + 0 \cdot e_3 \\ e_3^{'} &= 1 \cdot e_1 + 1 \cdot e_2 + 1 \cdot e_3 \end{aligned} \Longrightarrow (A_i^j) = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Example 7. Consider an arbitrary vector p having components $p_i^{'}$ and $p^{'i}$ in the basis $\{e_i^{'}\}$ and $\{e^{'i}\}$ and p_i and p^i in the basis $\{e_i\}$ and $\{e^i\}$. Show that

$$p_i^{'} = A_i^j p_j, \quad p^{'i} = \overline{A_j^i} p^j$$

Solution.

$p = p_i^{'} e^{'i} = p_i^{'} \overline{A_j^i} e^j$,

$p = p_j e^j$,

we can see that $p_j = p_i^{'} \overline{A_j^i} = \overline{A_j^i} p_i^{'}$,

$A_l^j p_j = A_l^j \overline{A_j^i} p_i^{'} = \delta_l^i p_i^{'} = p_l^{'}, \Longrightarrow p_l^{'} = A_l^j p_j$.

We can use the same way to proof second equality.

Exercise 5. *Prove that* $p^{'i} = \overline{A_j^i} p^j$.

Linear Transformations

Definition 1.6. *A linear transformation* $L : R^n \longrightarrow R^m$ *is a function satisfying the property:* $\forall \lambda, \mu \in R$ *and* $a, b \in R^n$ *there is*

$$L\langle \lambda a + \mu b \rangle = \lambda L\langle a \rangle + \mu L\langle b \rangle$$

Example 8. A function $L : R^n \longrightarrow R^m$ is given by $L\langle \overrightarrow{v} \rangle = A\overrightarrow{v}$ (where A is a m by n matrix and $\overrightarrow{v}$ is being considered as a column vector) satisfies this property and so it is a linear transformation.

A rotation by an angle θ corresponds to the multiplication by the matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

A set of all linear transformations $L : R^n \longrightarrow R^m$ is Euclidean space. We will use a notation $\mathcal{L}(R^n, R^m)$ for this set.

In the case $\mathcal{L}(R^n, R^n)$ it will be used the notation $\mathcal{L}(R^n)$.

Definition 1.7. *A linear transformation* $L : V \longrightarrow W$ *is called an* **isomorphism** *if it is one-to-one transformation* V **on** W. *The vector spaces* V *and* W *are called ISOMORPHIC.*

Proposition.

If V and W are finite dimensional vector spaces, then they are isomorphic iff they have the same dimension.

Linear transformation and matrices

There is a close relation between linear transformations $\mathcal{L}(R^n, R^m)$ and m by n matrices.

Let we have a linear mapping $L : R^n \longrightarrow R^m$ and $\{e_i\}_1^n \in R^n$, $\{f_j\}_1^m \in R^m$ be bases of R^n and R^m. respectively. Then

$$L\langle e_j\rangle = L_j^i f_i, \quad (j = 1, 2, ..., n) \tag{1.16}$$

Definition 1.8. *A matrix $L = (L_j^i)$ whose entries L_j^i satisfy (1.16) is called the matrix that represents L with respect to the bases $\{e_i\}$ and $\{f_j\}$ (or matrix of the linear mapping L with respect to the bases $\{e_i\}$ and $\{f_j\}$).*

If $\{e_i'\} \in R^n$ and $\{f_j'\} \in R^m$ are other bases. Let $(L)' = (L_j'^i)$ is the matrix of the same linear transformation L . Then we have

$$(L)(A) = (B)(L)',$$

where $(A) = (A_i^r)$ and $(B) = (B_j^s)$ are the transition matrices

$$e_i' = A_i^r e_r; \quad f_j' = B_j^s f_s.$$

In fact, we have

$$L\langle e_r\rangle = L_r^s f_s, \quad L\langle e_i'\rangle = L_i'^j f_j',$$

$$L\langle A_i^r e_r\rangle = L\langle e_i'\rangle = L_i'^j f_j' = L_i'^j B_j^s f_s,$$

$$L\langle A_i^r e_r\rangle = A_i^r L\langle e_r\rangle = A_i^r L_r^s f_s,$$

and we can see that

$$A_i^r L_r^s f_s = L_i'^j B_j^s f_s \quad \text{or } (L)(A) = (B)(L)'.$$

COMPOSITION of linear transformations $K \circ L$ where $L : R^n \longrightarrow R^m$ is defined by the rule

$$K \circ L\, \langle a\rangle = K\, \langle L\, \langle a\rangle\rangle$$

If $(L) = (L_i^j)$ and $(K) = (K_j^s)$ are matrices of L and K with respect to corresponding bases, then $(K \circ L) = (K_j^s L_i^j)$ with respect to the same bases.

The case $m = n$. We say that a linear transformation $L \in \mathcal{L}(R^n)$ is non-singular if $\det(L) \neq 0$.

Let a linear map $L : R^n \longrightarrow R^n$ be non-singular, then there is the inverse mapping L^{-1}.

Definition 1.9. *Let $L : R^n \longrightarrow R^n$ be a linear transformation. An eigenvalue of L is a number $\lambda \in R$ such that there is a nonzero vector $\overrightarrow{e} \in R^n$ with $L\langle \overrightarrow{e}\rangle = \lambda \overrightarrow{e}$. Such vector $\overrightarrow{e}$ is called an eigenvector of L associated with this eigenvalue.*

How can the eigenvalues and eigenvectors for L be found? In many cases these can be found via geometric descriptions of L.

Example 9. Consider the reflection to the plane $x + y + z = 0$ in R^3. Any vector in the plane is mapped to itself. Thus vectors in the plane are eigenvectors for the eigenvalue 1. The basis $(1, -1, 0)$, $(1, 0, -1)$ of the plane gives two eigenvectors for the eigenvalue 1. The normal vector $(1, 1, 1)$ to the plane is transformed to its negative via reflection, so it is an eigenvector for the eigenvalue -1.

We now show that the eigenvalues of L are the same as the eigenvalues of the corresponding matrix.

We define the characteristic polynomial of L to be the characteristic polynomial of a matrix (L), which represents L.

Firstly we prove, that if λ is eigenvalue of map L then λ is eigenvalue of matrix (L). If λ is an eigenvalue of L with corresponding eigenvector $\overrightarrow{v}$ and we choose $\overrightarrow{v}$ to be a first vector in the basis of R^n. Then the matrix (L), which represents L with respect to this basis has in the first column and in the first row only one element λ, because $L\langle\overrightarrow{v}\rangle = \lambda\overrightarrow{v}$. It means that λ is one of the eigenvalues of the matrix (L).

Conversely, if λ is an eigenvalue of matrix (L), it means that there is vector–column $\overrightarrow{w} \in R^n$ with property $L_i^j w_j = \lambda w_i$. Let us consider vector $\overrightarrow{w} = w_i e^i$, where $\{e^i\}$ is a basis in R^n. Then

$$L\langle\overrightarrow{w}\rangle = L\left\langle w_j e^j\right\rangle = w_j L\left\langle e^j\right\rangle = w_j L_i^j e^i = L_i^j w_j e^i = \lambda w_i e^i = \lambda\overrightarrow{w}.$$

Therefore, eigenvalues of L are precisely the roots of the **Characteristic polynomial** of L

$$\det(L-\lambda I) = 0$$

where I is the unit matrix. For the case $n=3$ a characteristic polynomial has the form

$$\lambda^3 - J_1\lambda^2 + J_2\lambda - J_3 = 0$$

where J_k are invariants of the transformation L ((L) is a matrix of the mapping L).

These invariants have special names

trace L: $J_1 = tr(L) = L_1^1 + L_2^2 + L_3^3 = L_i^i$ (or $sp(L)$ from German *"Spur"*).

the second invariant:

$$J_2 = J_2(L) = \begin{vmatrix} L_1^1 & L_2^1 \\ L_1^2 & L_2^2 \end{vmatrix} + \begin{vmatrix} L_1^1 & L_3^1 \\ L_1^3 & L_3^3 \end{vmatrix} + \begin{vmatrix} L_2^2 & L_3^2 \\ L_2^3 & L_3^3 \end{vmatrix}.$$

$$J_2(L) = \frac{1}{2}(tr^2 L - tr\, L^2) = \frac{1}{2}(L_i^i L_j^j - L_j^i L_i^j)$$

the third invariant:

$$J_3(L) = \det(L) = \det\begin{vmatrix} L_1^1 & L_2^1 & L_3^1 \\ L_1^2 & L_2^2 & L_3^2 \\ L_1^3 & L_2^3 & L_3^3 \end{vmatrix}$$

$$J_3(L) = \frac{1}{6}\left[tr^3(L) - 3tr\,(L)\,tr\,(L)^2 + 2\,tr\,(L)^3\right]$$

We note that, in terms of eigenvalues of L, which are roots of the characteristic equation, the invariants take the simplest form

$$J_1 = \lambda_1 + \lambda_2 + \lambda_3,\ J_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1,$$
$$J_3 = \lambda_1\lambda_2\lambda_3.$$

The Hamilton-Kelly identity is

$$L^3 - J_1 L^2 + J_2 L - J_3 I = 0$$

Let $L \in \mathcal{L}(R^n)$ be a linear transformation and $(L) = (L_j^i)$ be a representation of L in some basis. Note that the numbers J_i and $\det((L))$ do not depend on the basis. Therefore one can define the functions from the set of linear transformations $\mathcal{L}(R^n)$ into R:

$$tr\ L = L_i^i, \quad \det L = \det\,((L)).$$

Some properties of the invariants

$$tr(K+L) = tr(K) + tr(L),$$
$$\det(K \circ L) = \det(K)\cdot\det(L)$$

The case $m = 1$.

A linear transformation $\Phi : R^n \longrightarrow R$ is called a linear form defined on the space R^n. The space $\mathcal{L}(R^n; R)$ is called **conjugate** to the space R^n and we use the notation R^{n*}.

There exists a metrical isomorphism $\kappa : R^n \leftrightarrow R^{n*}$, which determines a one-to-one correspondence between vectors $b \in R^n$ and linear forms: a linear form is given by the formula

$$\kappa(b) < a >= a \cdot b$$

Theorem 1.4. *For any linear form $\Phi : R^n \longrightarrow R$ there exists unique vector b_Φ such that*

$$\Phi < a >= b_\Phi \cdot a$$

Bilinear forms

A function $\Phi : R^n \times R^n \longrightarrow R$ is called a bilinear form over R^n space if for all $\lambda, \mu \in R$ and $a, b, c \in R^n$

$$\Phi \langle \lambda a + \mu b, c \rangle = \lambda \Phi \langle a, c \rangle + \mu \Phi \langle b, c \rangle$$

$$\Phi \langle a, \lambda b + \mu c \rangle = \lambda \Phi \langle a, b \rangle + \mu \Phi \langle a, c \rangle$$

If $\Phi \langle a, b \rangle = \Phi \langle b, a \rangle$, then this form is called a symmetric form.

A bilinear form g determined by the formula

$$g \langle a, b \rangle = a \cdot b$$

is called a fundamental form of R^n space.

Example 10. If $L \in \mathcal{L}(R^n)$, then the function $\Phi \langle a, b \rangle = a \cdot L \langle b \rangle$ is a bilinear form.

Theorem 1.5. *For any bilinear form Φ over R^n there exists unique linear transformation $L \in \mathcal{L}(R^n)$ such that*

$$\Phi < a, b >= a \cdot L < b > .$$

Exercise 6. *Prove the theorem.*

The theorem establishes isomorphism between the set of bilinear forms Φ over R^n and the space of linear transformations $L \in \mathcal{L}(R^n)$. For instance, the fundamental form g corresponds to the identical transformation.

Let us consider a linear transformation $L \in \mathcal{L}(R^n)$. Φ. The form Φ^* defined by the formula

$$\Phi^* < a, b >= b \cdot L < a >$$

is a bilinear form. According to the theorem there exists the linear transformation L^* such that

$$\Phi^* < a, b >= a \cdot L^* < b > .$$

Hence,

$$a \cdot L^* < b >= \Phi^* < a, b >= b \cdot L < a > .$$

Therefore, for any linear transformation $L \in \mathcal{L}(R^n)$ there exists unique linear transformation $L^* \in \mathcal{L}(R^n)$ that

$$a \cdot L^* < b >= b \cdot L < a > .$$

The linear transformation L^* is called conjugate to a linear transformation L.

In the same orthonormal basis $\{e_i\}$ the matrix of (L^*) is a transpose matrix of the matrix (L).

Properties of *:

$$(L+N)^* = L^* + N^*, \quad (L \circ N)^* = N^* \circ L^*, \quad tr(L) = tr(L^*), \quad tr(L \circ N) = tr(N \circ L),$$

$$tr(K^* \circ L) = tr(K \circ L^*) = \sum_{i,j=1}^{n} K_i^j L_i^j,$$

where $(K) = (K_j^s)$ and $(L) = (L_i^j)$.

Definition 1.10. *A trace of composition of K and L^* (or K^* and L) is called a* **dot product** *(or a scalar product) of the transformations K and L. We use the notation*

$$K : L = tr(K^* \circ L) = tr(K \circ L^*).$$

Note that

$$K : L = L : K$$

Symmetry

Definition 1.11. *A linear transformation L is called a symmetric linear transformation if $L^* = L$ and antisymmetric (or skew-symmetric) if $L^* = -L$.*

The set of symmetrical linear transformations we denote by $\mathcal{L}_s(R^n)$.

If L is a linear mapping, then we can write $L = L_s + L_a$, where

$$L_s = \frac{1}{2}(L + L^*), \quad L_a = \frac{1}{2}(L - L^*).$$

Here L_s is symmetric and L_a is antisymmetric. This representation of L through symmetric and antisymmetric transformations is unique. If $\mathcal{L}_s(R^n)$ is the set of symmetric transformations and $\mathcal{L}_a(R^n)$ is the set of antisymmetric transformations, then they are linear subspaces of $\mathcal{L}(R^n)$ and

$$\mathcal{L}(R^n) = \mathcal{L}_a(R^n) \oplus \mathcal{L}_s(R^n)$$

• All eigenvalues of symmetric linear transformations are real. Non-zero eigenvalues of anti-symmetric linear transformations are purely imaginary.

• If L is a symmetric linear transformation, then there is an orthobasis in R^n for which the matrix (L) is diagonal.

Let an antisymmetric linear transformation $A \in \mathcal{L}_a(R^3)$ have the following matrix with respect to some orthonormal basis $\{e_i\}$

$$(A) = \begin{bmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{bmatrix}.$$

One can obtain the vector $\overrightarrow{a} = a^i \overrightarrow{e_i}$. Assume that $\{e_i'\}$ is another orthonormal basis $\{e_i\}$ and the matrix representation of the antisymmetric transformation A in this basis is $(A)'$. For this matrix one finds the vector $\overrightarrow{a}' = a'^i \overrightarrow{e_i'}$.

Exercise 7. *Prove that $\overrightarrow{a}' = \overrightarrow{a}$.*

Hint. The transition matrix T from an orthonormal basis to another orthonormal basis is orthogonal:

$$T^{-1} = T^*.$$

Therefore we constructed one-to-one mapping from $\mathcal{L}_a(R^3)$ onto R^3. For this mapping one can determine the inverse linear transformation $E : R^3 \longrightarrow \mathcal{L}_a(R^3)$, which acts according to

$$E\left\langle \overrightarrow{a} \right\rangle = A.$$

Exercise 8. *Prove that E is an isomorphism.*

Note that

$$(E < a >) < b >= -(E < b >) < a >= (E^* < b >) < a > .$$

Definition 1.12. *A* **cross** *or* **vector** *product of two vectors $\overline{u}$ and $\overline{v}$, denoted by $\overline{u} \times \overline{v}$ is a vector quantity. A direction of $\overline{u} \times \overline{v}$ is the direction of the extended thumb when the fingers of the right hand are closed from $\overline{u}$ to $\overline{v}$ (origins coinciding) through the smallest possible angle θ $(0 \leq \theta \leq \pi)$. The magnitude of $\overline{u} \times \overline{v}$ is defined by*

$$\mid \overline{u} \times \overline{v} \mid = uv \sin\theta$$

(The magnitude of $\overline{u} \times \overline{v}$ is the area of the parallelogram determined by $\overline{u}$ and $\overline{v}$).

Definition 1.13. *A vector product of a vector $\overline{a} = (a_1, a_2, a_3)$ with a vector $\overline{b} = (b_1, b_2, b_3)$ defined as a new vector*

$$\overline{a} \times \overline{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1),$$

where a_i, b_i are components with respect to an orthobasis.

Definition 1.14. *One can verify that*

$$E\left\langle \overrightarrow{a} \right\rangle \left\langle \overrightarrow{b} \right\rangle = \overrightarrow{a} \times \overrightarrow{b}$$

Definition 1.15. *A vector product of $\overline{a} = a_i\overline{e}^i, \overline{b} = b_i\overline{e}^i$ in an orthonormal basis $\{\overline{e}^i\}$*

$$\overline{a} \times \overline{b} = \begin{bmatrix} \overline{e}^1 & \overline{e}^2 & \overline{e}^3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

It is easy to verify the following rules governing these products

$$\overline{u} \times \overline{v} = -\overline{v} \times \overline{u};$$

$$\overline{u} \times (\overline{v} + \overline{w}) = \overline{u} \times \overline{v} + \overline{u} \times \overline{w};$$

$$\overline{u} \times \overline{u} = 0;$$

$\overline{u} \times \overline{v} = 0 \Leftrightarrow$ $\overline{u}$ is either parallel to $\overline{v}$ or $\overline{u} = 0$ or $\overline{v} = 0$.

The product $(\overline{u} \times \overline{v}) \cdot \overline{w}$ is called the box or triple scalar product.

Its geometrical interpretation depends on its sign:

(a) If $(\overline{u} \times \overline{v}) \cdot \overline{w} > 0$, then $(\overline{u} \times \overline{v}) \cdot \overline{w}$ is the volume of the parallelepiped defined by $\overline{u}$, $\overline{v}$, $\overline{w}$. In this case we say that $\overline{u}, \overline{v}, \overline{w}$ form a right-handed triad of vectors.

(b) If $(\overline{u} \times \overline{v}) \cdot \overline{w} < 0$, then $(\overline{u} \times \overline{v}) \cdot \overline{w}$ is the negative of the volume of the parallelepiped whose sides are $\overline{u}$, $\overline{v}$, $\overline{w}$. In this case $\overline{u}$, $\overline{v}$, $\overline{w}$ form a left-handed triad.

(c) If $(\overline{u} \times \overline{v}) \cdot \overline{w} = 0$, the vectors $\overline{u}$, $\overline{v}$, $\overline{w}$ are coplanar (parallel to the same plane).

It follows from the geometrical interpretation that

$$(\overline{u} \times \overline{v}) \cdot \overline{w} = \overline{u} \cdot (\overline{v} \times \overline{w}).$$

The product $(\overline{u} \times \overline{v}) \times \overline{w}$ is called a triple vector product. In general, the triple vector product is not associative:

$(\overline{u} \times \overline{v}) \times \overline{w} \neq \overline{u} \times (\overline{v} \times \overline{w})$.

The vector product $(\overline{u} \times \overline{v}) \times \overline{w}$ of three or more vectors can be reduced to simpler products with the aid of the identity

$$(\overline{a} \times \overline{b}) \times \overline{c} = \overline{b}(\overline{a} \cdot \overline{c}) - \overline{c}(\overline{a} \cdot \overline{b}).$$

This is known as "*back cab*" rule.

Example 11. Let A be any symmetric $n \times n$ matrix. Then we can define a symmetric bilinear form $\mathcal{B}_A : R^n \times R^n \longrightarrow R$

$$\mathcal{B}_A(\overrightarrow{x}, \overrightarrow{y}) = \overrightarrow{x}^* A \overrightarrow{y}$$

Here $\overrightarrow{x}$ and $\overrightarrow{y}$ are considered as column–vectors.

Symmetry is derived from the fact that $A = A^*$:

$$\mathcal{B}_A(\overrightarrow{y}, \overrightarrow{x}) = \overrightarrow{y}^* A \overrightarrow{x} = \overrightarrow{y}^* A^* \overrightarrow{x} = \overrightarrow{x}^* A \overrightarrow{y} = \mathcal{B}_A(\overrightarrow{x}, \overrightarrow{y}).$$

Definition 1.16. *A linear transformation $O \in \mathcal{L}(R^n)$ is called an orthogonal transformation if it satisfies*

$$O \circ O^* = I.$$

For the orthogonal transformations it is fair the formulae

$$O^* = O^{-1}, \ \det([O]) = \pm 1.$$

Any composition of two orthogonal transformations is orthogonal.

Definition 1.17. *A linear transformation $\overline{A}$ is called an equivalent transformation to A if there exists an orthogonal transformation O such that*

$$\overline{A} = O \circ A \circ O^*$$

Let a function $f : \mathcal{L}_s(R^n) \to \mathcal{L}_s(R^n)$ be a function, which maps a linear symmetric transformation into a linear symmetric transformation.

Definition 1.18. *A mapping $f : \mathcal{L}_s(R^n) \to \mathcal{L}_s(R^n)$ is called an invariant mapping with respect to orthogonal transformations if for any orthogonal transformation O there is*

$$f(O \circ A \circ O^*) = O \circ f(A) \circ O^*.$$

In continuum mechanics such functions are called isotropic functions.

Theorem 1.6. *All continuous isotropic functions $f : \mathcal{L}_s(R^n) \to \mathcal{L}_s(R^n)$ have the representation*

$$f(A) = \sum_{k=0}^{n-1} \varphi_k A^k,$$

where the coefficients φ_k are scalar functions only invariants of a linear transformation A.

Proof. We will prove this theorem for $n = 3$.

We choose orthogonal basis $\{e_i\}$ and we consider all linear transformations in this basis. Let A be a symmetrical linear transformation. Because the matrix A is a symmetrical matrix, there exists an orthogonal matrix O such that

$$\overline{A} = O \circ A \circ O^* = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix},$$

where a_i $(i = 1, 2, 3)$ are eigenvalues of the matrix A.

We show that the matrix $\overline{B} = O \circ f(A) \circ O^* = f(\overline{A})$ is also a diagonal matrix.

In fact, let the matrix $\overline{B}$ has the representation

$$\overline{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

We take the orthogonal matrix

$$O_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then by virtue of that O_1O is an orthogonal matrix, we obtain

$$O_1\overline{B}O_1^* = O_1(Of(A)O^*)O_1^* = f(O_1OA(O_1\,O)^*) = f(O_1\overline{A}O_1^*).$$

But $O_1\overline{A}O_1^* = \overline{A}$, and

$$O_1\overline{B}O_1^* = \begin{pmatrix} b_{11} & -b_{12} & -b_{13} \\ -b_{21} & b_{22} & b_{23} \\ -b_{31} & b_{32} & b_{33} \end{pmatrix} = f(\overline{A}) = \overline{B}.$$

We obtained $b_{12} = b_{13} = 0$. Similarly, we find that $b_{23} = b_{32} = 0$ and $b_{ii} = f_i(a_1, a_2, a_3)$. If we take the orthogonal transformation

$$O_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then

$$O_2\overline{B}O_2^* = \begin{pmatrix} b_{22} & 0 & 0 \\ 0 & b_{11} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} = f(O_2\overline{A}O_2^*) = f\begin{pmatrix} a_2 & 0 & 0 \\ 0 & a_1 & 0 \\ 0 & 0 & a_3 \end{pmatrix}.$$

This means

$$f_1(a_2, a_1, a_3) = f_2(a_1, a_2, a_3),\;\; f_2(a_2, a_1, a_3) = f_1(a_1, a_2, a_3),\;\; f_3(a_2, a_1, a_3) = f_3(a_1, a_2, a_3).$$

We will consider some cases.

1^0. Assume that $(a_1-a_2)(a_2-a_3)(a_3-a_1)\neq 0$. Let us consider the linear system of equations

$$b_{ii}=\alpha+\beta a_i+\gamma a_i^2 \ (i=1,2,3)$$

with respect to α,β,γ. The determinant of this system is (Gramma's determinant)

$$\Delta=\det\begin{pmatrix}1 & a_1 & a_1^2\\ 1 & a_2 & a_2^2\\ 1 & a_3 & a_3^2\end{pmatrix}=(a_1-a_2)(a_2-a_3)(a_3-a_1)\neq 0.$$

Because $\Delta\neq 0$, then

$$\alpha=\frac{1}{\Delta}\det\begin{pmatrix}b_{11} & a_1 & a_1^2\\ b_{22} & a_2 & a_2^2\\ b_{33} & a_3 & a_3^2\end{pmatrix},\beta=\frac{1}{\Delta}\det\begin{pmatrix}1 & b_{11} & a_1^2\\ 1 & b_{22} & a_2^2\\ 1 & b_{33} & a_3^2\end{pmatrix},\gamma=\frac{1}{\Delta}\det\begin{pmatrix}1 & a_1 & b_{11}\\ 1 & a_2 & b_{22}\\ 1 & a_3 & b_{33}\end{pmatrix}.$$

The functions $\alpha=\alpha(a_1,a_2,a_3),\beta=\beta(a_1,a_2,a_3),\gamma=\gamma(a_1,a_2,a_3)$ have property that they are unaltered by interchanges of pairs of a_1,a_2,a_3. In fact, for example, $\Delta'=-\Delta$,

$$\alpha(a_2,a_1,a_3)=\frac{1}{\Delta'}\det\begin{pmatrix}f_1(a_2,a_1,a_3) & a_2 & a_2^2\\ f_2(a_2,a_1,a_3) & a_1 & a_1^2\\ f_3(a_2,a_1,a_3) & a_3 & a_3^2\end{pmatrix}=-\frac{1}{\Delta}\det\begin{pmatrix}f_2(a_1,a_2,a_3) & a_2 & a_2^2\\ f_1(a_1,a_2,a_3) & a_1 & a_1^2\\ f_3(a_1,a_2,a_3) & a_3 & a_3^2\end{pmatrix}=\alpha(a_1,a_2,a_3).$$

Definition 1.19. *Any function $f(x_1,x_2,x_3)$ whose value is unaltered by interchanges of pairs of x_1,x_2,x_3 is said to be symmetrical in x_1,x_2,x_3.*

Theorem 1.7. *A continuous symmetrical function can be expressed as a function of the invariants $f(x_1,x_2,x_3)=g(J_1,J_2,J_3)$, where $J_1=x_1+x_2+x_3,\ J_2=x_1x_2+x_2x_3+x_3x_1,\ J_3=x_1x_2x_3$.*

Hence,

$$\alpha=\alpha(J_1(A),J_2(A),J_3(A)),\ \beta=\beta(J_1(A),J_2(A),J_3(A)),\ \gamma=\gamma(J_1(A),J_2(A),J_3(A))$$

and $\overline{B}=\alpha I+\beta\overline{A}+\gamma\overline{A}^2$. But $\overline{B}=O\circ B\circ O^*$, thus

$$B=O^*\circ\overline{B}\circ O=\alpha I+\beta O^*\circ\overline{A}\circ O+\gamma O^*\circ\overline{A}^2\circ O=\alpha I+\beta A+\gamma A^2.$$

2^0. Let $a_1=a_2\neq a_3$, then $f_2(a_1,a_2,a_3)=f_1(a_1,a_2,a_3)$ and we consider the equations

$$b_{ii}=\alpha+\beta a_i+\gamma 0 \ (i=1,3)$$

By virtue of $a_1\neq a_3$, we have

$$\alpha=\frac{1}{a_3-a_1}\det\begin{pmatrix}b_{11} & a_1\\ b_{33} & a_3\end{pmatrix},\ \beta=\frac{b_{33}-b_{11}}{a_3-a_1}.$$

The functions $\alpha=\alpha(a_1,a_2,a_3),\beta=\beta(a_1,a_2,a_3)$ are again symmetrical functions of a_1,a_2,a_3, and they are functions of invariants

$$\alpha=\alpha(J_1(A),J_2(A),J_3(A)),\ \beta=\beta(J_1(A),J_2(A),J_3(A)).$$

In this case we have $\overline{B}=\alpha I+\beta\overline{A}$. Hence, we get

$$B=O^*\circ\overline{B}\circ O=\alpha I+\beta O^*\circ\overline{A}\circ O=\alpha I+\beta A.$$

3^0. Let $a_1=a_2=a_3$, then $b_{11}=b_{22}=b_{33}$ and we consider the equation

$$b_{11}=\alpha.$$

The function α is a symmetrical function of a_1,a_2,a_3, therefore $\alpha=\alpha(J_1(A),J_2(A),J_3(A))$ and

$$B=O^*\circ\overline{B}\circ O=\alpha I.$$

Multilinear transformations

Definition 1.20. *A transformation* $u : \underbrace{R^n \times \ldots \times R^n} \longrightarrow R^m$ *is said to be r–multilinear (or simply multilinear) if it is linear in each variable.*

For multilinear transformation $u : \underbrace{R^n \times \ldots \times R^n} \longrightarrow R^m$ we write

$$(a_1, a_2, \ldots, a_r) \longrightarrow u \langle a_1, a_2, \ldots, a_r \rangle \in R^m,$$

where $a_i \in R^n$.

A linear transformation ($r = 1$) is a particular case of multilinear transformation. Bilinear forms ($r = 2;\ m = 1$) are also particular cases of multilinear transformations.

Definition 1.21. *A multilinear transformation with values in* R ($m = 1$) *is called a multilinear form defined in* R^n.

1.4 Tensors

Definition 1.22. *A multilinear (r-linear) form* Φ *defined in* R^n *is called a* **tensor**. *The number* $r > 0$ *is called an order of the tensor* Φ.

Let $\{e_i\}$ be a basis and $\{e^i\}$ be the cobasis in R^n. The numbers $\Phi \langle a_1, a_2, \ldots, a_r \rangle$ are called components of the tensor Φ of order r in the basis $\{e_i\}$ or $\{e^i\}$, where the vectors $a_s, (s = 1, \ldots, r)$ are from the basis or cobasis. If all a_s belong to the basis $\{e_i\}$, then the components of the tensor Φ are called covariant components:

$$\Phi_{i_1, i_2, \ldots, i_r} = \Phi \langle e_{i_1}, e_{i_2}, \ldots, e_{i_r} \rangle .$$

If all a_s belong to cobasis $\{e^i\}$

$$\Phi^{k_1, k_2, \ldots, k_r} = \Phi \left\langle e^{k_1}, e^{k_2}, \ldots, e^{k_r} \right\rangle ,$$

then they are contravariant components of Φ. All other types of components are called mixed components of Φ. Mixed components of Φ are called p times covariant and q times contravariant if we use p vectors from the basis and q vectors from the cobasis.

A tensor of order r has only 2^r different types of components. The number of components of each type is equal to n^r.

Change of tensor components

The formulae for changing tensor components are obtained from the properties of linear forms. Let A and $\overline{A}$ be transition matrices

$$e'_i = A^j_i e_j; \quad e'^i = \overline{A}^i_j e^j.$$

Covariant components of a tensor are changed by the rule

$$\Phi_{i_1, i_2, \ldots, i_r} \Longrightarrow \Phi'_{i_1, i_2, \ldots, i_r} = \Phi_{j_1, j_2, \ldots, j_r} A^{j_1}_{i_1} \ldots A^{j_r}_{i_r}.$$

Contravariant components are changed by the rule

$$\Phi^{k_1, \ldots, k_r} \Longrightarrow \Phi'^{k_1, \ldots, k_r} = \Phi^{i_1, i_2, \ldots, i_r} \overline{A}^{k_1}_{i_1} \ldots \overline{A}^{k_r}_{i_r}.$$

Changed mixed components ("new") are a product of "old" components and the elements of A for all covariant indices (upper indices) and elements of the matrix $\overline{A}$ for all contravariant indices (lower indices), for example:

$$\Phi'^{i_1,i_2,\dots i_\alpha}_{i_{\alpha+1},\dots,i_r} = \Phi^{j_1,j_2,\dots,j_\alpha}_{j_{\alpha+1},\dots,j_r} A^{i_1}_{j_1} \dots A^{i_\alpha}_{j_\alpha} \overline{A}^{j_{\alpha+1}}_{i_{\alpha+1}} \dots \overline{A}^{j_r}_{i_r}.$$

This rule shows us that if components of a tensor are known in one basis, then they are known in any basis. It means that tensor is determined by the set of its coordinates in some basis.

This property can be considered as a definition of a tensor.

Definition 1.23. *A tensor of order r is a set of n^r numbers, which obey to the rule of changing components as defined above.*

Particular cases.

A tensor of zeroth order has only one invariant scalar component and it is usually referred to a **scalar.**

A tensor of first order can be considered as a vector. As was proven there is isomorphism between linear forms over R^n and the vector space R^n.

A tensor of second order has four different types of components. Each of its has n^r components. As it has been proven, there is a one-to-one correspondence between bilinear forms and linear transformations

$$\Phi < a, b >= a \cdot L < b > .$$

Therefore there is an isomorphism between second order tensors and linear transformations.

Fundamental tensor

Definition 1.24. *A tensor of second order g determined by the fundamental form $g\langle a, b\rangle = a \cdot b$ is called a fundamental tensor.*

It is obvious that $(g^{ij}) = (g_{ij})^{-1}$. Another essential property of the fundamental tensor is its symmetry

$$g_{ij} = g_{ji}, \quad g^{ij} = g^{ji}.$$

It can be shown that

$$e^i = g^{ij} e_j, \quad e_j = g_{ij} e^i.$$

Proof.

Let $e^i = T^{ij} e_j$, then

$$g^{ik} = g\left\langle e^i, e^k\right\rangle = e^i \cdot e^k = T^{ij} e_j \cdot e^k = T^{ij} \delta^k_j = T^{ik}.$$

Exercise 9. *Prove that $e_j = g_{ij} e^i$.*

In the Euclidean space R^n covariant and contravariant components of a tensor are not completely independent. Each set of covariant components gives raise to a set of contravariant components and conversely. In the terminology usually applied in tensor analysis g_{ik}, g^{ik} are used: to lower and to raise indices, respectively.

For example: ($r = 1$, vector).

If $\overrightarrow{a} = a_i e^i = a^j e_j$, then

$$a_i e^i = a_i(g^{ij} e_j) = a_i g^{ij} e_j = a^j e_j.$$

Hence,

$$a^j = a_i g^{ij} \quad a_i = a^k g_{ik}.$$

If $r = 3$ (third order tensor), then

$$\Phi^{..k}_{ij.} = \Phi_{ijs} g^{ks}, \quad \Phi^{..k}_{ij.} = \Phi^{pqk} g_{ip} g_{jq}.$$

Raising and lowering indices

The effect of multiplying tensor components by g^{ik} (or g_{ik}) and summing w.r.t. k is raising (or lowering) indices.

Exercise 10. *Prove that:*
(a) $T^{i\cdot}_{\cdot j} = g^{ik}T_{kj}$, *(b)* $T^{ij} = g^{ik}T^{\cdot j}_{k\cdot}$,
(b) $T_{ij} = g_{ik}T^{k\cdot}_{\cdot j}$, *(c)* $T_{ij} = ?\ T^{kl}$.

Exercise 11. *Show that the following relations for tensor components are equivalent:*
(a) $S_{ij} = S_{ji} \Leftrightarrow S^{kl} = S^{lk} \Leftrightarrow S^{m\cdot}_{\cdot n} = S^{\cdot m}_{n\cdot}$,
(b) $a_{ij} = -a_{ji} \Leftrightarrow a^{kl} = -a^{lk} \Leftrightarrow a^{m\cdot}_{\cdot n} = -a^{\cdot m}_{n\cdot}$.

Exercise 12. *If T is a second order tensor and $T_{ij}v^iv^j = 0$ for all vectors v, then $T_{ij} = -T_{ji}$.*

1.4.1 Fundamental operations with tensor coordinates

1. Summation. A summation of two or more multilinear forms is a multilinear form. In coordinates, for example, for third order tensors A and B, the coordinates of the sum are

$$C^{\cdot mp}_{q\cdot\cdot} = A^{\cdot mp}_{q\cdot\cdot} + B^{\cdot mp}_{q\cdot\cdot}.$$

Converserly, it is easy to prove that a sum of coordinates of two (or more) tensors of the same type composes a tensor.

2. Subtraction (similar).

3. Outer Multiplication. A product of tensor components A and B is a tensor whose order is a sum of the order of the given tensors and its coordinates are defined by

$$A^{\cdot pr}_{q\cdot\cdot}B^{m\cdot}_{\cdot s} = C^{\cdot prm\cdot}_{q\cdots s}$$

The tensor C is called an outer product of A and B (C is also called a Kronekker product).

Note that not every tensor can be written as a product of two tensors of lower orders.

Contraction

If one contravariant and one covariant indices of tensor components are the same, then it means summation w.r.t. this index. The result of the summation is a tensor of order on two less than the original tensor. This process of summation is called a contraction. For example: let $\Phi^{mpr\cdot\cdot}_{\cdot\cdot\cdot qs}$ be a tensor of order 5, then we obtain $\Psi^{mp\cdot}_{\cdot\cdot q} = \Phi^{mpr\cdot\cdot}_{\cdot\cdot\cdot qr}$ are components of a tensor of third order. Setting $p = q$ one obtains $C^m = \Psi^{mp\cdot}_{\cdot\cdot p}$ components of a first order tensor. Coordinateless representation of the contraction can be given by the following:

$$\Psi\left\langle a, b, c\right\rangle = \Phi\left\langle a, b, e^i, c, e_i\right\rangle = \Phi\left\langle a, b, e_i, c, e^i\right\rangle.$$

Note that this process does not depend on a basis.

Inner multiplication

By the process of outer multiplication of two tensors followed by a contraction we obtain a new tensor called an **inner product** of the given tensors. This process is called an **inner multiplication**.

For example, for given tensors with components $A^{mp\cdot}_{\cdot\cdot q}$ and $B^{r\cdot\cdot}_{\cdot st}$ the outer product has the components $A^{mp\cdot}_{\cdot\cdot q}B^{r\cdot\cdot}_{\cdot st}$. Letting $q = r$ we get the inner product $A^{mp\cdot}_{\cdot\cdot r}B^{r\cdot\cdot}_{\cdot st}$. Letting $q = r$ and $p = s$, we obtain another inner product $A^{mp\cdot}_{\cdot\cdot r}B^{r\cdot\cdot}_{\cdot pt}$. Inner and outer multiplication of tensors is commutative and associative.

Quotient law

Let B be a mathematical or physical quantity, which is represented in a basis $\{e_i\}$ by an order set of n^2 scalars B_{ij}. Assume that if the basis $\{e_i\}$ is changed, then the scalars B_{ij} are changed such way that for all vectors $\overrightarrow{v} = v^\alpha e_\alpha$ the scalars $u_i = B_{ij}v^j$ are components of the vector $\overrightarrow{u}$. Then B is a second order tensor.

Proof. Let us consider

$$u_i = B_{ij}v^j, \ (i = 1, \dots, n),$$

where v^j and u_i are contravariant and covariant components of vectors. It means that in the new system of coordinates $\{e_i'\}$ we have

$$u_p^{'} = u_i A_p^i = B_{il} v^l A_p^i$$

and

$$u_p^{'} = B_{pk}^{'} v^{'k} = B_{pk}^{'} \overline{A}_l^k v^l,$$

where A and $\overline{A}$ are transition matrices. Hence,

$$(B_{il}A_p^i - B_{pk}^{'}\overline{A}_l^k)v^l = 0$$

for any vector v. Therefore

$$B_{il}A_p^i - B_{pk}^{'}\overline{A}_l^k = 0$$

Thus, after multipling on A_j^l we get

$$B_{pj}^{'} = B_{il}A_p^i A_j^l.$$

It means that B_{il} are covariant tensor components: they are changed according to the law required for second order tensor components.

The result just proven extends to tensors of higher order and is usually called the **quotient rule**: suppose that it is not known whether a quantity X is a tensor or not. If an inner product of X with an arbitrary tensor is a tensor, then X is also a tensor.

1.4.2 Dyadic Product of two vectors ($\otimes$)

A dyadic product of vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ denoted by $\overrightarrow{a} \otimes \overrightarrow{b}$ is defined as a transformation, which transforms an arbitrary vector $\overrightarrow{c}$ according to the rule:

$$(\overrightarrow{a} \otimes \overrightarrow{b}) < \overline{c} >= \overrightarrow{a}(\overrightarrow{b}, \overline{c})$$

From the definition one obtains that dyad is a linear transformation

$$\overline{a} \otimes \overline{b} < c\overline{u} + \beta\overline{v} >= c\overline{a} \otimes \overline{b} < \overline{u} > + \beta\overline{a} \otimes \overline{b} < \overline{v} > .$$

A pair $\overrightarrow{a} \otimes \overrightarrow{b}$ is called a dyad or direct product of vectors $\overrightarrow{a}$ and $\overrightarrow{b}$. The tensor, which corresponds to the dyad is also called a dyad. Because for the dyadic tensor

$$\Phi < c, d >= (c, (a \times b) < d >) = (c, a)(b, d),$$

then coordinates of the dayadic tensor are

$$\Phi_{ij} = a_i b_j,$$

where $a = a_i e^i$, $b = b_i e^i$.

Example 12. Let $\overline{e}$ be an arbitrary unit vector. We can form the dyad $\overline{e} \otimes \overline{e}$. The transformation $\overline{e} \otimes \overline{e}$ maps any vector $\overline{v}$ into its vector projection on $\overline{e}$: $\overline{e}(\overline{e \cdot v})$.

The linear transformation $P = \overline{e} \otimes \overline{e}$ corresponds to a second-order tensor that can be called a projection tensor.

1.4.3 The permutation symbol

Definition 1.25. *The third order tensor given by the formula*

$$\varepsilon\left\langle \overline{a}, \overline{b}, \overline{c} \right\rangle = \overline{a} \cdot (\overline{b} \times \overline{c})$$

is called a **permutation tensor**.

By means of covariant components ε_{ijl} and contravariant components ε^{ijl} of the permutation tensor we can express the cross product of basis and cobasis vectors:

$$\overline{e}_i \times \overline{e}_j = \varepsilon_{ijl}\overline{e}^l;\ \overline{e}^i \times \overline{e}^j = \varepsilon^{ijl}\overline{e}_l.$$

For any vectors $\overline{a}, \overline{b}$ there is

$$\overline{a} \times \overline{b} = \varepsilon_{ijl} a^j b^l \overline{e}^i = \varepsilon^{ijl} a_j b_l \overline{e}_i.$$

According to the properties of triple scalar product all components ε_{ijl} can be expressed through one component, for example, $\varepsilon_{123} = \varepsilon$.

A basis $\{e_i\}$ is called right-handed if $\varepsilon_{123} = \varepsilon > 0$ and it is called left-handed if $\varepsilon_{123} = \varepsilon < 0$. Note that

$$\varepsilon_{ijk} = \begin{cases} +\varepsilon, & \text{if } (i,j,k) \text{ is an } \textbf{even} \text{ permutation of } (1,2,3) \\ -\varepsilon, & \text{if } (i,j,k) \text{ is an } \textbf{odd} \text{ permutation of } (1,2,3) \\ 0, & \text{if two or more indices are equal} \end{cases}$$

Hence,

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = -\varepsilon_{321} = -\varepsilon_{213} = -\varepsilon_{132}.$$

Example 13. Prove that

$$[\overline{A} \cdot (\overline{B} \times \overline{C})][\overline{a} \cdot (\overline{b} \times \overline{c})] = \det \begin{bmatrix} \overline{A}\cdot\overline{a} & \overline{A}\cdot\overline{b} & \overline{A}\cdot\overline{c} \\ \overline{B}\cdot\overline{a} & \overline{B}\cdot\overline{b} & \overline{B}\cdot\overline{c} \\ \overline{C}\cdot\overline{a} & \overline{B}\cdot\overline{b} & \overline{C}\cdot\overline{c} \end{bmatrix}$$

Proof.

Let $\overline{A} = A^m\overline{e}_m;\ \overline{B} = B^n\overline{e}_n, \overline{C} = C^k\overline{e}_k$ and $\{\overline{e}_i\}$ be an orthobasis, then

$$\overline{A} \cdot (\overline{B} \times \overline{C}) = A^m\overline{e}_m B^n\overline{e}_n \times C^k\overline{e}_k = A^m B^n C^k \overline{e}_m \cdot (\overline{e}_n \times \overline{e}_k) = A^m B^n C^k \varepsilon_{mnk}.$$

According to the definition we have

$$\varepsilon_{123} \det \begin{bmatrix} A^1 & A^2 & A^3 \\ B^1 & B^2 & B^3 \\ C^1 & C^2 & C^3 \end{bmatrix} = A^m B^n C^k \varepsilon_{mnk}$$

$$[\overline{A} \cdot (\overline{B} \times \overline{C})][\overline{a} \cdot (\overline{b} \times \overline{c})] = \det \begin{bmatrix} A^1 & A^2 & A^3 \\ B^1 & B^2 & B^3 \\ C^1 & C^2 & C^3 \end{bmatrix} \det \begin{bmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{bmatrix} =$$

$$\det \begin{bmatrix} A^1 & A^2 & A^3 \\ B^1 & B^2 & B^3 \\ C^1 & C^2 & C^3 \end{bmatrix} \det \begin{bmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{bmatrix}^T =$$

$$\det\left(\begin{bmatrix} A^1 & A^2 & A^3 \\ B^1 & B^2 & B^3 \\ C^1 & C^2 & C^3 \end{bmatrix} \cdot \begin{bmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{bmatrix}^T \right) = \det \begin{bmatrix} \overline{A}\cdot\overline{a} & \overline{A}\cdot\overline{b} & \overline{A}\cdot\overline{c} \\ \overline{B}\cdot\overline{a} & \overline{B}\cdot\overline{b} & \overline{B}\cdot\overline{c} \\ \overline{C}\cdot\overline{a} & \overline{B}\cdot\overline{b} & \overline{C}\cdot\overline{c} \end{bmatrix}.$$

Example 14. $\varepsilon_{123} = \varepsilon = \pm\sqrt{g},\ g = \det\|g_{ij}\|$.
Proof.

$$\varepsilon_{123}\varepsilon_{123} = [\overline{e}_1\cdot(\overline{e}_2\times\overline{e}_3)][\overline{e}_1\cdot(\overline{e}_2\times\overline{e}_3)] = det\begin{bmatrix}\overline{e}_1\cdot\overline{e}_1 & \overline{e}_1\cdot\overline{e}_2 & \overline{e}_1\cdot\overline{e}_3\\ \overline{e}_2\cdot\overline{e}_1 & \overline{e}_2\cdot\overline{e}_2 & \overline{e}_2\cdot\overline{e}_3\\ \overline{e}_3\cdot\overline{e}_1 & \overline{e}_3\cdot\overline{e}_2 & \overline{e}_3\cdot\overline{e}_3\end{bmatrix} = \det\begin{bmatrix}g_{11} & g_{12} & g_{13}\\ g_{21} & g_{22} & g_{23}\\ g_{31} & g_{32} & g_{33}\end{bmatrix}$$

The same for $\varepsilon^{123} = \pm\frac{1}{\sqrt{g}}$. In fact,

$$\varepsilon^{123}\varepsilon^{123} = [\overline{e}^1\cdot(\overline{e}^2\times\overline{e}^3)][\overline{e}^1\cdot(\overline{e}^2\times\overline{e}^3)] = det\begin{bmatrix}\overline{e}^1\cdot\overline{e}^1 & \overline{e}^1\cdot\overline{e}^2 & \overline{e}^1\cdot\overline{e}^3\\ \overline{e}^2\cdot\overline{e}^1 & \overline{e}^2\cdot\overline{e}^2 & \overline{e}^2\cdot\overline{e}^3\\ \overline{e}^3\cdot\overline{e}^1 & \overline{e}^3\cdot\overline{e}^2 & \overline{e}^3\cdot\overline{e}^3\end{bmatrix} = \det\begin{bmatrix}g^{11} & g^{12} & g^{13}\\ g^{21} & g^{22} & g^{23}\\ g^{31} & g^{32} & g^{33}\end{bmatrix} =$$

$$\det||g^{ij}|| = 1/g,\quad \varepsilon^{123} = \pm\frac{1}{\sqrt{g}}.$$

Example 15. $\varepsilon_{ijk}\varepsilon^{prk} = \delta_i^p\delta_j^r - \delta_i^r\delta_j^p$.
Let us prove that

$$\varepsilon_{ijk}\varepsilon^{prq} = [\overline{e}_i\cdot(\overline{e}_j\times\overline{e}_k)][\overline{e}^p\cdot(\overline{e}^r\times\overline{e}^q)] = det\begin{bmatrix}\overline{e}_i\cdot\overline{e}^p & \overline{e}_i\cdot\overline{e}^r & \overline{e}_i\cdot\overline{e}^q\\ \overline{e}_j\cdot\overline{e}^p & \overline{e}_j\cdot\overline{e}^r & \overline{e}_j\cdot\overline{e}^q\\ \overline{e}_k\cdot\overline{e}^p & \overline{e}_k\cdot\overline{e}^r & \overline{e}_k\cdot\overline{e}^q\end{bmatrix} = \det\begin{bmatrix}\delta_i^p & \delta_i^r & \delta_i^q\\ \delta_j^p & \delta_j^r & \delta_j^q\\ \delta_k^p & \delta_k^r & \delta_k^q\end{bmatrix}$$

Example 16. Prove that $\overline{a}\times(\overline{b}\times\overline{c}) = \overline{b}(\overline{a}\cdot\overline{c}) - \overline{c}(\overline{a}\cdot\overline{b})$.

Proof. Let $\overline{a} = a^i\overline{e}_i,\ \overline{b} = b_i\overline{e}^i,\ \overline{c} = c_i\overline{e}^i$, where $\{\overline{e}_i\}$ and $\{\overline{e}^i\}$ are basis and cobasis. We suppose that $\overline{b}\times\overline{c} = \overline{d}$. Then

$$\overline{a}\times\overline{d} = \varepsilon_{ijl}a^jd^l\overline{e}^i,\quad \overline{d} = \overline{b}\times\overline{c} = \varepsilon^{\alpha\beta\gamma}b_\beta c_\gamma\overline{e}_\alpha,$$
$$d^l = \overline{d}\cdot\overline{e}^l = \varepsilon^{\alpha\beta\gamma}b_\beta c_\gamma\overline{e}_\alpha\cdot\overline{e}^l = \varepsilon^{\alpha\beta\gamma}b_\beta c_\gamma\delta_\alpha^l = \varepsilon^{l\beta\gamma}b_\beta c_\gamma.$$
$$\overline{a}\times(\overline{b}\times\overline{c}) = \varepsilon_{ijl}\varepsilon^{l\beta\gamma}a^jb_\beta c_\gamma\overline{e}^i = \varepsilon_{ijl}\varepsilon^{\beta\gamma l}a^jb_\beta c_\gamma\overline{e}^i =$$
$$= (\delta_i^\beta\delta_j^\gamma - \delta_i^\gamma\delta_j^\beta)a^jb_\beta c_\gamma\overline{e}^i = \delta_i^\beta\delta_j^\gamma a^jb_\beta c_\gamma\overline{e}^i - \delta_i^\gamma\delta_j^\beta a^jb_\beta c_\gamma\overline{e}^i =$$
$$= a^\gamma b_\beta c_\gamma\overline{e}^\beta - a^\beta b_\beta c_\gamma\overline{e}^\gamma = b_\beta\overline{e}^\beta a^\gamma c_\gamma - c_\gamma\overline{e}^\gamma a^\beta b_\beta = \overline{b}(\overline{a}\cdot\overline{c}) - \overline{c}(\overline{a}\cdot\overline{b}).$$

Exercise 13. *Evaluate* $\varepsilon_{ijk}\varepsilon^{ijk}$.

Exercise 14. *Write* $(\overline{a}\times\overline{b})\cdot(\overline{c}\times\overline{d})$ *in component form.*

Exercise 15. *Prove that* $\overline{A}\cdot(\overline{B}\times\overline{C}) = \overline{B}\cdot(\overline{C}\times\overline{A}) = \overline{C}\cdot(\overline{A}\times\overline{B})$.

Exercise 16. *Prove that*
(a) $\overline{A}\cdot(\overline{B}\times\overline{C}) = (\overline{A}\times\overline{B})\cdot\overline{C}$;
(b) $\overline{A}\cdot(\overline{A}\times\overline{C}) = 0$.

1.5 Tensor calculus

Let $\Omega\subset R^n$ be open set in R^n. A tensor field is a function over Ω that associates any point of Ω a tensor. In this section we study an elementary calculus of tensor and vector fields.

Definition 1.26. *A transformation* $u:\Omega\longrightarrow R^m$ *is continuous at the point* $x\in\Omega$ *if*

$$\lim_{|h|_{R^n}\to 0}|u(x+h) - u(x)|_{R^m} = 0$$

Definition 1.27. *A transformation* $u:\Omega\to R^m$ *is continuous on the set* Ω *if* u *is continuous at any point* $x\in\Omega$.

Differentiation

Let $u : \Omega \to R^m$ be continuous on Ω.

Definition 1.28. *A transformation u is called differentiable at the point $x \in \Omega$ if there exists a linear transformation $L : R^n \to R^m$ such that*

$$\lim_{|h|\to 0} \left|h^{-1}\right| |u(x+h) - u(x) - L\langle h\rangle| = 0$$

Definition 1.29. *A transformation u is called differentiable on Ω, if it is differentiable at any point $x \in \Omega$.*

Definition 1.30. *A linear mapping L, defined above is called a derivative of the transformation u at the point x and it is denoted by $\dfrac{\partial u}{\partial x}(x)$.*

Definition 1.31. *A new transformation $\dfrac{\partial u}{\partial x} : \Omega \to \mathcal{L}\left(R^n; R^m\right)$ given by the formula $x \to \dfrac{\partial u}{\partial x}\langle x\rangle$ is called a derivative transformation of u.*

A transformation u is called continuously differentiable on Ω if the derivative transformation $\dfrac{\partial u}{\partial x}$ is continuous on Ω. Here the symbol $\dfrac{\partial u}{\partial x}$ is used only as notation for the "new" linear transformation

$$\frac{\partial u}{\partial x} : R^n \to \mathcal{L}(R^n, R^m) \text{ or } \frac{\partial u}{\partial x} \in \mathcal{L}\left(R^n; R^m\right)$$

If $u : R^n \to R^m$ and $v : R^m \to R^p$ are differentiable, then their composition $v \circ u : R^n \to R^p$ is also differentiable. And we have

$$\frac{\partial}{\partial x}(v \circ u) = \frac{\partial v}{\partial x} \circ \frac{\partial u}{\partial x}$$

The case $n = 1$.

In this case $u : R(t) \to R^m$ is usually called a vector–function of the variable t. If $u = u^j f_j$, and $\{f_j\}_1^m \subset R^m$ is a basis then $\frac{\partial u}{\partial t} = \frac{\partial u^j}{\partial t} f_j$.

General case can be reduced to this simple one. Let $u : \Omega \to R^m$ be a differentiable vector field with $\Omega \subset R^n$. One can consider the new transformation $v : R(t) \to R^m$, where $v(t) = u(x + ta)$ and $a \in R^n$ is an arbitrary vector (this vector is called a test vector). We have

$$\frac{\partial u}{\partial x}(x)\langle a\rangle = \lim_{t\to 0}\frac{u(x+ta) - u(x)}{t} = \lim_{t\to 0}\frac{v(t) - v(0)}{t} = \frac{\partial v}{\partial x}(0).$$

Problems.

1. Expand $\dfrac{d}{dt}\left[\overline{u}(t) \times \overline{v}(t) \cdot \overline{w}(t)\right]$.

Solution.

$$\frac{d\overline{u}}{dt} \times \overline{v}(t) \cdot \overline{w}(t) + \overline{u}(t) \times \frac{d\overline{v}}{dt} \cdot \overline{w}(t) + \overline{u}(t) \times \overline{v}(t) \cdot \frac{d\overline{w}}{dt}.$$

2. Show that if $\overline{e}(t)$ has a constant magnitude, then $\dfrac{d\overline{e}}{dt}$ is either zero or perpendicular to $\overline{e}$.

Solution

$$\overline{e} \cdot \overline{e} = \left|\overline{e}^2\right| = const, \quad \frac{d\overline{e} \cdot \overline{e}}{dt} = 2\overline{e} \cdot \frac{d\overline{e}}{dt} = 0.$$

3. Let $u = \sin t e_1 + e^t e_2 + e_3$. Find $\dfrac{d\overline{u}}{dt}$.

Solution.

$$\frac{d\overline{u}}{dt} = \cos t\, e_1 + e^t\, e_2.$$

Let $m = 1$. A transformation $\varphi : R^n \to R$ $(m = 1)$ is simply a scalar function. We use a special symbol for the derivative transformation (derivative transformation is a linear form)

$$\frac{\partial \varphi}{\partial x} = \nabla \varphi$$

The transformation $\nabla\varphi$ is called a gradient of φ. The operator ∇ is also known as a NABLA or *del operator*. The del operator ∇ represents a linear function that maps a tensor field into another tensor field. For instance, the gradient of scalar function $f(x_1, x_2)$ maps a scalar field into a vector field $\nabla f(x_1, x_2)$.

For the linear form $\nabla\varphi$ there exists an unique vector b such that

$$\nabla\varphi < a >= a \cdot b.$$

The vector b is also denoted by $\nabla\varphi$.

1.5.1 A coordinate representation

Let $m = 1$. If $\{e_i\}$ is a basis in R^n, and $x = x^i e_i$, then

$$\nabla\varphi < e_i >= \lim_{t\to 0} \frac{\varphi(x + te_i) - \varphi(x)}{t} = \frac{\partial \varphi}{\partial x^i}.$$

Therefore, the vector b is $b = \frac{\partial \varphi}{\partial x^i} e_i$, or

$$b = (\frac{\partial \varphi}{\partial x^1}, \frac{\partial \varphi}{\partial x^2}, \cdots, \frac{\partial \varphi}{\partial x^n}).$$

Let $n = 1$. If $\{f_j\}$ is a basis in R^m, then $u = u^j f_j$. The derivative is calculated by component–wise

$$\frac{\partial u}{\partial x} = \lim_{t\to 0} \frac{u^j(x + t) - u^j(x)}{t} f_j = \frac{\partial u^j}{\partial x} f_j.$$

In the general case $x = x^i e_i$, $u = u^j f_j$ the matrix of linear transformation $(\frac{\partial u}{\partial x})$ has the representation

$$(\frac{\partial u}{\partial x}) = (\frac{\partial u^j}{\partial x^i}),$$

where $\frac{\partial u^j}{\partial x^i}$ are usual partial derivatives of the function $u^j(x^1, \ldots, x^n)$. The matrix $\left(\frac{\partial u^j}{\partial x^i}\right)$ is called the Jacobi matrix of the transformation u. A transformation $u : \Omega \to R^m$ is continuously differentiable on Ω iff all partial derivatives $\partial u^j / \partial x^i$ exist and they are continuous functions on the set Ω.

Example 17. A directional derivative of f in the direction $\overline{e}$ is

$$\frac{df}{ds} = \overline{e} \cdot \nabla f.$$

Example 18. One can use the gradient to derive the equations of the normal line and tangent plane to a surface $f(\overline{x}) = const$ at the point (a_1, a_2, a_3). If $\overline{r}$ is a point on the normal line, then ∇f is parallel to the normal:

$$\overline{r} - \overline{a} = t \cdot \nabla f$$

or in coordinate form

$$x_i - a_i = t\frac{\partial f}{\partial x_i}.$$

If $\overline{r}$ is a point on the tangent plane, then ∇f is perpendicular to the plane

$$\overline{(r - \overline{a})} \cdot \nabla f = 0,$$

or

$$(x_i - a_i)\frac{\partial f}{\partial x_i}(a_1, a_2, a_3) = 0.$$

Exercise 17. *Find an equation of the normal line and tangent plane to the surface* $x_3 = 1 + x_1 x_2$ *at the point* $(1, 1, 2)$.

Exercise 18. *Find equations of straight lines normal and tangent to the curve* $f(x_1, x_2) =$ *const at the point* (a_1, a_2).

In the case $m = n$ a linear transformation $\frac{\partial u}{\partial x}$ belongs to $\mathcal{L}(R^n)$. A matrix representation of this transformation is a square $n \times n$ matrix. Two invariants of this linear transformation have special names.

The trace $tr(\frac{\partial u}{\partial x})$ is called a **divergence**. The determinant $\det(\frac{\partial u}{\partial x})$ is called the *Jacobian* of the transformation u. Thus,

$$div\, u = tr(\frac{\partial u}{\partial x}) = \frac{\partial u^i}{\partial x^i}, \quad \left|\frac{\partial u}{\partial x}\right| = \det(\frac{\partial u}{\partial x}) = \left|\frac{\partial u^j}{\partial x^i}\right|.$$

1.5.2 Divergence of a tensor

In continuum mechanics we will use a new vector operation, it is named a *divergence of a tensor*.

Let $P : \Omega \to \mathcal{L}(R^n)$ be a differentiable transformation and $\overline{a} \in R^n$ be an arbitrary test vector and $x \in \Omega$. We can determine a transformation $u_a(x) : R^n \to R^n$ by

$$u_a(x) = P^*(x) \langle \overline{a} \rangle .$$

This is a differentiable function. We construct the divergence of this transformation (as above)

$$div(u_a(x)) = tr(\frac{\partial u_a}{\partial x}) = tr(\frac{\partial}{\partial x} P^*(x) \langle \overline{a} \rangle)$$

In the left hand side we have a scalar in the right hand side we have a linear form. There is an unique vector p such that

$$tr(\frac{\partial}{\partial x} P^*(x) \langle \overline{a} \rangle) = \overline{a} \cdot p.$$

The vector p is called a divergence of the tensor P and it is denoted by $p = div\, P$.

Example 19. Let

$$P(x, y) = \left[\begin{array}{cc} x & xy \\ y & x^2 \end{array} \right],$$

be a matrix representation of a linear transformation P in the basis $\overline{e}_1 = (1, 0), \overline{e}_2 = (0, 1)$. Find $div\, P$.

Solution. Note that

$$div\, P = (div\, P)^i \cdot \overline{e}_i = ((div\, P)^1, (div\, P)^2).$$

Because $\{\overline{e}_1, \overline{e}_2\}$ is an orthonormal basis

$$P^*(x,y) = P^T = \begin{bmatrix} x & y \\ xy & x^2 \end{bmatrix}.$$

Hence,

$$P^*(x,y)\,\langle \overline{e}_1 \rangle = \begin{bmatrix} x & y \\ xy & x^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ xy \end{bmatrix}.$$

Therefore,

$$P^*(x,y)\,\langle \overline{e}_2 \rangle = \begin{bmatrix} x & y \\ xy & x^2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} y \\ x^2 \end{bmatrix},$$

$$(div\,P)^1 = div(\,P^*(x)\,\langle \overline{e}_1 \rangle) = \frac{\partial}{\partial x} x + \frac{\partial}{\partial y} xy = 1 + x,$$

$$(div\,P)^2 = div\,(P^*(x)\,\langle \overline{e}_2 \rangle) = \frac{\partial}{\partial x} y + \frac{\partial}{\partial y} x^2 = 0.$$

For the case $m = n = 3$ a *rotor* (or *curl*) of a vector field u is defined by the formula

$$rot\ u = E^{-1} \left\langle \frac{\partial u}{\partial x} - (\frac{\partial u}{\partial x})^* \right\rangle.$$

Hence, a coordinate representation of the curl in an orthonormal basis is defined by:

$$(\frac{\partial u}{\partial x} - (\frac{\partial u}{\partial x})^*) = \begin{bmatrix} u^1_{,x_1} & u^1_{,x_2} & u^1_{,x_3} \\ u^2_{,x_1} & u^2_{,x_2} & u^2_{,x_3} \\ u^3_{,x_1} & u^3_{,x_2} & u^3_{,x_3} \end{bmatrix} - \begin{bmatrix} u^1_{,x_1} & u^2_{,x_1} & u^3_{,x_1} \\ u^1_{,x_2} & u^2_{,x_2} & u^3_{,x_2} \\ u^1_{,x_3} & u^2_{,x_3} & u^3_{,x_3} \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & (u^1_{x_2} - u^2_{x_1}) & (u^1_{x_3} - u^3_{x_1}) \\ (u^2_{x_1} - u^1_{x_2}) & 0 & (u^2_{x_3} - u^3_{x_2}) \\ (u^3_{x_1} - u^1_{x_3}) & (u^3_{x_2} - u^2_{x_3}) & 0 \end{bmatrix}.$$

We remind that in an orthonormal basis

$$a = E^{-1}\,\langle A_a \rangle = E^{-1} \begin{bmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{bmatrix} = \begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix}.$$

Thus,

$$rot\ u = E^{-1} \left\langle \frac{\partial u}{\partial x} - \left(\frac{\partial u}{\partial x} \right)^* \right\rangle = \left(\frac{\partial u^3}{\partial x_2} - \frac{\partial u^2}{\partial x_3}, \frac{\partial u^1}{\partial x_3} - \frac{\partial u^3}{\partial x_1}, \frac{\partial u^2}{\partial x_1} - \frac{\partial u^1}{\partial x_2} \right)$$

or another representation is

$$\nabla \times \overline{u} = \left(\frac{\partial}{\partial x} \overline{i} + \frac{\partial}{\partial y} \overline{j} + \frac{\partial}{\partial z} \overline{k} \right) \times \left(u_1 \overline{i} + u_2 \overline{j} + u_3 \overline{k} \right) = \det \begin{bmatrix} \overline{i} & \overline{j} & \overline{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_1 & u_2 & u_3 \end{bmatrix} =$$

$$\left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z} \right) \overline{i} + \left(\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial y} \right) \overline{j} + \left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right) \overline{k}$$

1.5.3 Second derivative

If derivative transformation $\dfrac{\partial u}{\partial x} : \Omega \to \mathcal{L}(R^n; R^m)$ is differentiable, then the function u is called twice differentiable. The derivative transformation $\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)$ is called a second derivative or derivative transformation of second order. We use a notation $\frac{\partial^2 u}{\partial x^2}$. A value of second derivative at the point x is called a second derivative of mapping u at a point x. A second derivative of mapping is a bilinear transformation from $R^n \times R^n$ into R^m. The bilinear mapping $\frac{\partial^2 u}{\partial x^2} : R^n \times R^n \to R^m$ is defined by the formula of successive differentiating

$$\frac{\partial^2 u}{\partial x^2}\langle a, b\rangle = \frac{\partial}{\partial x}(\frac{\partial u}{\partial x}\langle b\rangle)\langle a\rangle$$

Note that if a function u is twice continuously differentiable, then the second derivative is symmetric

$$\frac{\partial^2 u}{\partial x^2}\langle a, b\rangle = \frac{\partial^2 u}{\partial x^2}\langle b, a\rangle$$

1.5.4 High order derivatives

High order derivatives can be defined by the induction

$$\frac{\partial^k u}{\partial x^k} = \frac{\partial}{\partial x}(\frac{\partial^{k-1} u}{\partial x^{k-1}}).$$

A derivative $\frac{\partial^k u}{\partial x^k}$ at a point $x \in \Omega$ of a k-times continuously differentiable mapping is a symmetric k–linear transformation $\underbrace{R^n \times R^n \times \cdots \times R^n}_{k} \to R^m$. In the general case one can write down

$$\frac{\partial^p}{\partial x^p}\left(\frac{\partial^q u}{\partial x^q}\langle b_1, \ldots, b_q\rangle\right)\langle a_1, \ldots, a_p\rangle = \frac{\partial^{p+q} u}{\partial x^{p+q}}\langle a_1, \ldots, a_p, b_1, \ldots, b_q\rangle.$$

1.6 Curvilinear coordinate systems

1.6.1 Coordinate system

For the sake of simplicity we consider here coordinates in three–dimensional Euclidean space R^3. All facts and definitions are valid in arbitrary space R^n.

A set of triples (K^1, K^2, K^3) , where K^1, K^2, K^3 are real numbers, is called an arithmetic space A^3. For all elements of A^3 the operations of summation, subtraction and multiplication by a scalar, and dot product are defined by usual way. For example,

$$(K^1, K^2, K^3) \cdot (L^1, L^2, L^3) = K^1 L^1 + K^2 L^2 + K^3 L^3.$$

This space of triples is an Euclidean space.

Let $\Omega \subset R^3(\overline{x})$ be an open set. A one–to–one and reciprocal continuously differentiable mapping $K : \Omega \to A^3$ is called a coordinate system. This mapping is defined by the formula

$$\overline{x} \to K(\overline{x}) = (K^1(\overline{x}), K^2(\overline{x}), K^3(\overline{x}))$$

The values of functions $K^i(\overline{x})$ are called coordinates (curvilinear coordinates) of the point $\overline{x}$.

For any fixed point x_0 the equation $K^i(\overline{x}) = K^i(\overline{x}_0)$ determines a coordinate surface $\Pi_i \subset R^3(\overline{x})$. This coordinate surface passes through the point $\overline{x}_0$.

Any pair Π_i, Π_j of these surfaces is intersected along the curves

$$l_1 = \Pi_2 \cap \Pi_3, \quad l_2 = \Pi_3 \cap \Pi_1 \quad l_3 = \Pi_1 \cap \Pi_2,$$

which are called coordinate curves (or curvilinear axes of coordinates). Along the l_i–line only the coordinate K^i is changed; two other coordinates are constants.

1.6.2 Coordinate basis

Let a point $\overline{x} \in \Omega$ be fixed. At this point there are three vectors

$$\overline{e}_i = \frac{\partial \overline{x}}{\partial K^i}, \quad (i = 1, 2, 3),$$

which form a basis in R^3. This basis is called a coordinate basis of the coordinate system K at the point $\overline{x}$. A vector $\overline{e}_i$ is a tangent vector to the coordinate line l_i.

Similarly, one can define other three vectors at the same point $\overline{x} \in \Omega$

$$\overline{e}^i = \frac{\partial K^i}{\partial \overline{x}} = \nabla K^i, \quad (i = 1, 2, 3).$$

These vectors form a cobasis that corresponds to the basis $\{\overline{e}_i\}$. The basis $\{\overline{e}^i\}$ is called a coordinate cobasis of the coordinate system K at the point $\overline{x}$. The vectors $\overline{e}^i$ are normal to the coordinate surfaces Π_i. If $K' : \Omega \to A^3$ is a "new" coordinate system, then the trasformation $K' \circ K^{-1} : A^3 \to A^3$ is defined. This transformation acts by the rule

$$(K^1, K^2, K^3) \to (K'^1, K'^2, K'^3) = (K' \circ K^{-1})(K^1, K^2, K^3)$$

and an inverse mapping $K \circ (K')^{-1}$ is determined, too. Then, "new" basis and cobasis are related with "old" bases by usual formulae

$$\overline{e}'_i = A_i^j \overline{e}_j; \ \overline{e}'^i = \overline{A}_j^i \overline{e}^j,$$

where

$$A_i^j = \frac{\partial K^j}{\partial K'^i}; \ \overline{A}_j^i = \frac{\partial K'^i}{\partial K^j}; \quad (i, j = 1, 2, 3)$$

1.6.3 Orthogonal coordinate systems

A coordinate system is called **orthogonal** (at a point or on a set) if its basis is orthogonal

$$\overline{e}_i \cdot \overline{e}_j = 0, \quad \overline{e}^i \cdot \overline{e}^j = 0, \ (i \neq j)$$

(at the point or on the set). Coordinates of the fundamental tensor g with respect to an orthogonal coordinate system are

$$(g_{ij}) = \begin{bmatrix} h_1 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & h_3 \end{bmatrix},$$

where

$$h_i = g_{ii} = |\overline{e}_i|^2 = \left| \frac{\partial \overline{x}}{\partial K^i} \right|^2, \ i = 1, 2, 3.$$

1.6.4 Christoffel's symbols

Vectors of a coordinate basis $\{\overline{e}_i\}$ depend on $\overline{x}$ or on $(K^1, K^2, K^3) = K(\overline{x})$. Derivatives of the basis vectors can be represented in terms of the basis $\{\overline{e}_i\}$

$$\frac{\partial \overline{e}_i}{\partial K^j} = \Gamma^s_{ij}\overline{e}_s,$$

where the coefficients Γ^s_{ij} are called Christoffel's symbols of second order. Note that the Christoffel's symbols are not components of any tensor. These symbols are symmetric with respect to lower indices

$$\Gamma^s_{ij} = \Gamma^s_{ji}.$$

A dual formula for the representation of the Christoffel symbols is

$$\frac{\partial \overline{e}^i}{\partial K^j} = -\Gamma^i_{js}\overline{e}^s.$$

The Christoffel symbols are related to the derivatives of the fundamental tensor

$$\Gamma^l_{ij} = \frac{1}{2}\left(\frac{\partial g_{is}}{\partial K^j} + \frac{\partial g_{js}}{\partial K^i} - \frac{\partial g_{ij}}{\partial K^s}\right) g^{ls}.$$

Let us prove these identities. In the strength of the identities

$$g_{ij} = \overline{e}_i \cdot \overline{e_j}, \quad g_{is} = \overline{e}_i \cdot \overline{e}_s, \quad g_{js} = \overline{e}_j \cdot \overline{e}_s,$$

$$\frac{\partial g_{ij}}{\partial K^s} = \frac{\partial \overline{e}_i}{\partial K^s} \cdot \overline{e_j} + \overline{e_i} \cdot \frac{\partial \overline{e}_j}{\partial K^s} = \Gamma^p_{is}\overline{e}_p \cdot \overline{e_j} + \overline{e_i} \cdot \Gamma^p_{js}\overline{e}_p$$

one can obtain

$$\frac{\partial g_{is}}{\partial K^j} = \frac{\partial \overline{e}_i}{\partial K^j} \cdot \overline{e}_s + \overline{e}_i \cdot \frac{\partial \overline{e}_s}{\partial K^j} = \Gamma^p_{ij}\overline{e}_p \cdot \overline{e}_s + \overline{e}_i \cdot \Gamma^p_{sj}\overline{e}_p,$$

$$\frac{\partial g_{js}}{\partial K^i} = \frac{\partial \overline{e}_j}{\partial K^i} \cdot \overline{e}_s + \overline{e}_j \cdot \frac{\partial \overline{e}_s}{\partial K^i} = \Gamma^p_{ji}\overline{e}_p \cdot \overline{e}_s + \overline{e}_j \cdot \Gamma^p_{si}\overline{e}_p.$$

Hence,

$$\frac{\partial g_{is}}{\partial K^j} + \frac{\partial g_{js}}{\partial K^i} - \frac{\partial g_{ij}}{\partial K^s} = \Gamma^p_{ij}\overline{e}_p \cdot \overline{e}_s + \overline{e}_i \cdot \Gamma^p_{sj}\overline{e}_p + \Gamma^p_{ji}\overline{e}_p \cdot \overline{e}_s + \overline{e}_j \cdot \Gamma^p_{si}\overline{e}_p - \Gamma^p_{is}\overline{e}_p \cdot \overline{e_j} - \overline{e_i} \cdot \Gamma^p_{js}\overline{e}_p =$$

$$= 2\Gamma^p_{ji}\overline{e}_p \cdot \overline{e}_s = 2\Gamma^p_{ji}g_{ps}.$$

Thus, after multiplying on g^{sl} one has the proof.

Note that

$$\Gamma^s_{is} = \frac{1}{\sqrt{|g|}}\frac{\partial\sqrt{|g|}}{\partial K^i}, \quad (i = 1, 2, 3),$$

where $|g| = \det(g_{ij})$. This follows from $|g| = [e_1 \cdot (e_2 \times e_3)]^2$ and

$$\frac{\partial |g|}{\partial K^i} = 2[e_1 \cdot (e_2 \times e_3)][\frac{\partial e_1}{\partial K^i} \cdot (e_2 \times e_3) + e_1 \cdot (\frac{\partial e_2}{\partial K^i} \times e_3) + e_1 \cdot (e_2 \times \frac{\partial e_3}{\partial K^i})] =$$
$$= 2\varepsilon^2_{123}(\Gamma^1_{1i} + \Gamma^2_{2i} + \Gamma^3_{3i}).$$

1.6.5 Derivative tensor

Let Φ be a tensor of order r, which is dependent on $\overline{x} \in \Omega \subset R^3$ and $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r$ be a set of test vectors. Hence, there is the function $\Phi \left\langle \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r \right\rangle : \Omega \to R$ defined by the formula

$$\overline{x} \to \Phi(\overline{x}) \left\langle \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r \right\rangle .$$

The derivative of this map is a linear transformation. This means that

$$\Phi' \left\langle \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r, \overline{b} \right\rangle = \frac{\partial}{\partial \overline{x}} (\Phi(\overline{x}) \left\langle \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r \right\rangle) \left\langle \overline{b} \right\rangle$$

is a $(r+1)$-linear form.

Therefore, Φ' is a tensor of $(r+1)$ order. The tensor Φ' is called a derivative tensor of the tensor Φ. Coordinates of the tensor $\Phi' \left\langle \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_r, \overline{b} \right\rangle$ which correspond to basis vectors $\overline{b} = e_i$ are called covariant derivatives of the components of the tensor Φ. We use the notation

$$\Phi'_{i_1 \ldots i_r l} = \Phi_{i_1 \ldots i_r, l}.$$

Here a one covariant index, which is written after the comma, is added.

1.6.6 Covariant derivative

Covariant derivatives are expressed in terms of partial derivatives with respect to corresponding coordinates, Christoffel's symbols and components of a tensor. The simplest are covariant derivatives of a scalar, which coincide with usual partial derivatives

$$\Phi_{,i} = \frac{\partial \Phi}{\partial K^i}.$$

Let us consider first a tensor field of first order. Taking the derivative one has

$$\begin{gathered}
\frac{\partial}{\partial K^l} \Phi_i(\overline{x}) = \frac{\partial}{\partial K^l} \left(\Phi(\overline{x}) < \overline{e_i} > \right) = \Phi'(\overline{x}) < \overline{e_i}, \frac{\partial \overline{x}}{\partial K^l} > + \Phi(\overline{x}) < \frac{\partial \overline{e_i}}{\partial K^l} >= \\
= \Phi'(\overline{x}) < \overline{e_i}, \overline{e_l} > + \Phi(\overline{x}) < \Gamma^s_{il} \overline{e}_s >= (\Phi'(\overline{x}))_{il} + \Gamma^s_{il} (\Phi(\overline{x}))_s = \\
= (\Phi')_{i,l}(\overline{x}) + \Gamma^s_{il} \Phi_s(\overline{x}).
\end{gathered}$$

Thus,

$$\Phi_{i,l} = (\Phi')_{i,l} = \frac{\partial}{\partial K^l} \Phi_i - \Gamma^s_{il} \Phi_s.$$

Similar for contravariant coordinates of a vector

$$\begin{gathered}
\frac{\partial}{\partial K^l} \Phi^i(\overline{x}) = \frac{\partial}{\partial K^l} \left(\Phi(\overline{x}) < \overline{e^i} > \right) = \Phi'(\overline{x}) < \overline{e^i}, \frac{\partial \overline{x}}{\partial K^l} > + \Phi(\overline{x}) < \frac{\partial \overline{e^i}}{\partial K^l} >= \\
= \Phi'(\overline{x}) < \overline{e^i}, \overline{e_l} > + \Phi(\overline{x}) < -\Gamma^i_{sl} \overline{e^s} >= (\Phi'(\overline{x}))^i_{.l} - \Gamma^i_{sl} (\Phi(\overline{x}))^s = \\
= (\Phi')^i_{,l}(\overline{x}) - \Gamma^i_{sl} \Phi^s(\overline{x})
\end{gathered}$$

or

$$\Phi^i_{.,l} = (\Phi')^i_{,l} = \frac{\partial}{\partial K^l} \Phi^i + \Gamma^i_{sl} \Phi^s.$$

For a second order tensor

$$\Phi_{ij,l} = \frac{\partial \Phi_{ij}}{\partial K^l} - \Gamma^s_{li} \Phi_{sj} - \Gamma^s_{lj} \Phi_{is}, \quad \Phi^{ij}_{,l} = \frac{\partial \Phi^{ij}}{\partial K^l} + \Gamma^i_{ls} \Phi^{sj} + \Gamma^j_{ls} \Phi^{is},$$

$$\Phi^{\cdot j}_{i.,l} = \frac{\partial \Phi^{\cdot j}_{i.}}{\partial K^l} - \Gamma^s_{li}\Phi^{\cdot j}_{s.} + \Gamma^j_{ls}\Phi^{\cdot s}_{i.}, \quad \Phi^{j\cdot}_{.i,l} = \frac{\partial \Phi^{j\cdot}_{.i}}{\partial K^l} - \Gamma^s_{li}\Phi^{j\cdot}_{.s} + \Gamma^j_{ls}\Phi^{s\cdot}_{.i},$$

In the general case for covariant derivatives of a tensor Φ of order r one obtains

$$\Phi_{i_1\dots i_r,l} = \frac{\partial \Phi_{i_1\dots i_r}}{\partial K^l} - \sum_{\sigma=1}^{r} \Gamma^s_{li_\sigma}\Phi_{i_1\dots s\dots i_r},$$

where the index s is placed instead of the index i_σ. Similarly, for coordinates $\Phi^{j_1\dots j_r}_{\dots\dots\ ,l} = \Phi' < \overline{e^{j_1}}, \dots, \overline{e^{j_r}}, \overline{e_l} >$

$$\Phi^{j_1\dots j_r}_{\dots\dots\ ,l} = \frac{\partial \Phi^{j_1\dots j_r}}{\partial K^l} + \sum_{\sigma=1}^{r} \Gamma^{j_\sigma}_{ls}\Phi^{j_1\dots s\dots j_r},$$

where the index s is placed instead of the index j_σ.

Covariant differentiation of sum and products of tensor components are got by usual laws of differentiation of product and summation.

Special interest of covariant derivatives is the covariant derivatives of the fundamental tensor

$$g_{ij,l} = g^{ij}_{\dots,l} = 0.$$

This results are known as **Ricci's lemma**. A practical importance of this lies in the fact that the metric tensor acts as a constant when the operation of covariant differentiation is applied to it. Therefore, it can be moved freely in covariant differentiation. This property gives us a simple way to calculate covariant derivatives of mixed and contravariant components of an arbitrary tensor, if we know all covariant derivatives of covariant components of this tensor.

Proof (**Ricci's lemma**).

The proof is obtained from the following calculations

$$g_{ij,l} = \frac{\partial g_{ij}}{\partial K^l} - \Gamma^s_{li}g_{sj} - \Gamma^s_{lj}g_{is} =$$

$$\frac{\partial g_{ij}}{\partial K^l} - \frac{1}{2}g^{sm}\left(\frac{\partial g_{lm}}{\partial K^i} + \frac{\partial g_{im}}{\partial K^l} - \frac{\partial g_{li}}{\partial K^m}\right)g_{sj} - \frac{1}{2}g^{sm}\left(\frac{\partial g_{lm}}{\partial K^j} + \frac{\partial g_{jm}}{\partial K^l} - \frac{\partial g_{lj}}{\partial K^m}\right)g_{is} =$$

$$= \frac{\partial g_{ij}}{\partial K^l} - \frac{1}{2}\delta^m_j\left(\frac{\partial g_{lm}}{\partial K^i} + \frac{\partial g_{im}}{\partial K^l} - \frac{\partial g_{li}}{\partial K^m}\right) - \frac{1}{2}\delta^m_i\left(\frac{\partial g_{lm}}{\partial K^j} + \frac{\partial g_{jm}}{\partial K^l} - \frac{\partial g_{lj}}{\partial K^m}\right) =$$

$$= \frac{\partial g_{ij}}{\partial K^l} - \frac{1}{2}\left(\frac{\partial g_{lj}}{\partial K^i} + \frac{\partial g_{ij}}{\partial K^l} - \frac{\partial g_{li}}{\partial K^j}\right) - \frac{1}{2}\left(\frac{\partial g_{li}}{\partial K^j} + \frac{\partial g_{ji}}{\partial K^l} - \frac{\partial g_{lj}}{\partial K^i}\right) = 0.$$

Let us consider some properties of covariant derivatives.

(a) Derivative of a vector:

$$\frac{\partial \overline{\Phi}}{\partial K^j} = \frac{\partial(\Phi_i\overline{e}^i)}{\partial K^j} = \frac{\partial \Phi_i}{\partial K^j}\overline{e}^i + \Phi_i\frac{\partial \overline{e}^i}{\partial K^j} = \frac{\partial \Phi_i}{\partial K^j}\overline{e}^i + \Phi_i(-1)\Gamma^i_{js}\overline{e}^s =$$

$$\left(\frac{\partial \Phi_i}{\partial K^j} - \Gamma^i_{js}\Phi_s\right)\overline{e}^i = \Phi_{i,j}\overline{e}^i.$$

(b) Covariant derivative of a scalar product has the same form as usual partial derivative. In fact, let $B = \Phi^i\Psi_i$ be a scalar product of two differentiable vectors $\overline{\Phi}$ and $\overline{\Psi}$:

$$B_{,j} = \frac{\partial B}{\partial K^j} = \frac{\partial(\Phi^i\Psi_i)}{\partial K^j} = \frac{\partial \Phi^i}{\partial K^j}\Psi_i + \Phi^i\frac{\partial \Psi_i}{\partial K^j} = \frac{\partial \Phi^i}{\partial K^j}\Psi_i + \Phi^i(\Psi_{i,j} + \Gamma^s_{ji}\Psi_s) =$$

$$= \left(\frac{\partial \Phi^i}{\partial K^j} + \Gamma^i_{sj}\Phi^s\right)\Psi_i + \Phi^i\Psi_{i,j} = \Phi^i\Psi_{i,j} + \Phi^i_{,j}\Psi_i.$$

Here we used

$$\Phi^i_{,j} = \frac{\partial \Phi^i}{\partial K^j} + \Gamma^i_{js}\Phi^s, \;\; \Psi_{i,j} = \frac{\partial \Psi_i}{\partial K^j} - \Gamma^s_{ij}\Psi_s$$

Therefore,

$$B_{,j} = \frac{\partial B}{\partial K^j} = \Phi^i \Psi_{i,j} + \Phi^i_{,j}\Psi_i.$$

1.6.7 Differential Operators

In this section we study some differential operators that are expressed in terms of covariant derivatives. We use these formulae for covariant representation of differential equations of continuum mechanics.

Gradient of a function

Definition 1.32. *The vector*

$$\nabla F = \frac{\partial F}{\partial \overline{x}} = \frac{\partial F}{\partial K^i}\, e^i$$

is called a gradient of the function F. Covariant components of the gradient are:

$$(\nabla F)_i = F_{,i} = \frac{\partial F}{\partial K^i}.$$

Derivative of a vector

Let $\overline{v}$ be a vector. A derivative of the vector $\overline{v}$ is a linear transformation $\dfrac{\partial \overline{v}}{\partial \overline{x}} \in \mathcal{L}(R^n)$. This linear transformation corresponds to the second order tensor, which is denoted by the same symbol:

$$\frac{\partial \overline{v}}{\partial \overline{x}} \left\langle \overline{a},\, \overline{b} \right\rangle = \overline{a} \cdot \frac{\partial \overline{v}}{\partial \overline{x}} \left\langle \overline{b} \right\rangle$$

Covariant coordinates of the $\dfrac{\partial \overline{v}}{\partial \overline{x}}$ are

$$\left(\frac{\partial \overline{v}}{\partial \overline{x}}\right)_{ij} = e_i \left(\frac{\partial \overline{v}}{\partial \overline{x}}\right) < e_j >= e_i (\lim_{t\to 0} \frac{\overline{v}(x+te_j)-\overline{v}(x)}{t}) = e_i \frac{\partial \overline{v}}{\partial K^j} = \frac{\partial v_i}{\partial K^j} - \Gamma^s_{ij} v_s = v_{i,j}.$$

Divergence of a vector

First we remind definition of trace of linear operator L, $tr(L)$, where $L \in \mathcal{L}(R^n)$. If $\{e_i\}$ is a basis in R^n then $tr(L) = L^i_i$, where $L^j_i = e^j \cdot L < e_i >$ and $L < e_i >= L^j_i e_j$. Note that $tr(L)$ is a contraction of the tensor corresponding to the linear transformation L.

The divergence of a vector $\overline{v}$ is a scalar

$$div\, \overline{v} = tr\left(\frac{\partial \overline{v}}{\partial \overline{x}}\right) = e^i (\frac{\partial \overline{v}}{\partial \overline{x}}) < e_i > .$$

The divergence of a vector v can be expressed in terms of covariant derivatives of the vector $\overline{v}$

$$div\, \overline{v} = v^i_{,i} = \frac{\partial v^i}{\partial K^i} + \Gamma^i_{is} v^s = \frac{\partial v^i}{\partial K^i} + \frac{v^i}{\sqrt{|g|}} \frac{\partial \sqrt{|g|}}{\partial K^i} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial K^i}\left(\sqrt{|g|} v^i\right).$$

The Laplace operator of a function

The scalar

$$\triangle F = div\ (\nabla F) = ((\nabla F)^i)_{,i} = (g^{is}(\nabla F)_s)_{,i} =$$

$$= \left(g^{is} \frac{\partial F}{\partial K^s} \right)_{,i} = g^{is} (\frac{\partial F}{\partial K^s})_{,i} = g^{is} [\frac{\partial^2 F}{\partial K^s \partial K^i} - \Gamma^{\alpha}_{is} \frac{\partial F}{\partial K^\alpha}]$$

is called the Laplace operator of the function F.

If one denotes $\overline{v} = \nabla F$, then $v^i = g^{is} \dfrac{\partial F}{\partial K^s}$, therefore

$$\triangle F = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial K^i} \left(\sqrt{|g|} g^{is} \frac{\partial F}{\partial K^s} \right)$$

Curl of a vector

A *Curl* (or rotor) of a vector $\overline{v}$ is the vector

$$curl\, \overline{v} = E^{-1} \left\langle \left(\frac{\partial \overline{v}}{\partial \overline{x}} \right) - \left(\frac{\partial \overline{v}}{\partial \overline{x}} \right)^* \right\rangle$$

For obtaining contravariant components of this vector note that

$$E < curl\, \overline{v} >= \frac{\partial \overline{v}}{\partial \overline{x}} - \left(\frac{\partial \overline{v}}{\partial \overline{x}} \right)^*.$$

Then

$$e_j((E < curl\, \overline{v} >) < e_s >) = e_j((curl\, \overline{v})^l e_l \times e_s) = \varepsilon_{jls} (curl\, \overline{v})^l$$

and

$$e_j((E < curl\, \overline{v} >) < e_s >) = (\frac{\partial \overline{v}}{\partial \overline{x}})_{js} - (\frac{\partial \overline{v}}{\partial \overline{x}})_{sj} = \frac{\partial v_j}{\partial K^s} - \frac{\partial v_s}{\partial K^j}.$$

Hence,

$$\varepsilon_{123} \left(curl\, \overline{v} \right)^1 = \frac{\partial v_3}{\partial K^2} - \frac{\partial v_2}{\partial K^3}, \quad \varepsilon_{123} \left(curl\, \overline{v} \right)^2 = \frac{\partial v_1}{\partial K^3} - \frac{\partial v_3}{\partial K^1},$$

$$\varepsilon_{123} \left(curl\, \overline{v} \right)^3 = \frac{\partial v_2}{\partial K^1} - \frac{\partial v_1}{\partial K^2}.$$

If one uses a right-handed basis, then $\varepsilon_{123} = \sqrt{|g|}$ and

$$\left(curl\, \overline{v} \right)^1 = \frac{1}{\sqrt{|g|}} \left(\frac{\partial v_3}{\partial K^2} - \frac{\partial v_2}{\partial K^3} \right), \quad \left(curl\, \overline{v} \right)^2 = \frac{1}{\sqrt{|g|}} \left(\frac{\partial v_1}{\partial K^3} - \frac{\partial v_3}{\partial K^1} \right),$$

$$\left(curl\, \overline{v} \right)^3 = \frac{1}{\sqrt{|g|}} \left(\frac{\partial v_2}{\partial K^1} - \frac{\partial v_1}{\partial K^2} \right).$$

Divergence of a tensor

Let a be a test vector, then

$$a \cdot div\, P = div\ (P^* \langle a \rangle) = tr\left(\frac{\partial}{\partial x} P^* \langle a \rangle\right) = (P^* \langle a \rangle)^j_{,j} = \frac{\partial (P^* < a >)^j}{\partial K^j} + \Gamma^j_{j\alpha}(P^* < a >)^\alpha .$$

Because $(P^* < a >)^j = e^j P^* < a >= aP < e^j >$, and $P < e^j >= P^{sj} e_s$, then

$$a \cdot div\, P = a\left(\frac{\partial P^{sj}}{\partial K^j} e_s + P^{sj}\frac{\partial e_s}{\partial K^j} + \Gamma^j_{j\alpha} P^{s\alpha} e_s\right) =$$
$$(a \cdot e_s)\left(\tfrac{\partial P^{sj}}{\partial K^j} + \Gamma^j_{j\alpha}P^{s\alpha} + \Gamma^s_{\alpha j}P^{\alpha j}\right) = (a \cdot e_s)P^{sj}_{,j}$$

or

$$a_s(div\, P)^s = a_s P^{sj}_{,j}.$$

This means that

$$(div\, P)^s = P^{sj}_{,j}.$$

Another representation is the following

$$(div\, P)^s = \ P^{sj}{}_{,j} = div\,(\overline{P}^s) + \Gamma^s_{j\alpha}P^{j\alpha},$$

where $\overline{P}^s = (P^{s1}, P^{s2}, P^{s3})$ and

$$div\,(\overline{P}^s) = \frac{\partial P^{sj}}{\partial K^j} + \Gamma^j_{j\alpha}P^{s\alpha}.$$

The Laplace operator of a vector

The vector

$$\triangle \overline{v} = div\left(\frac{\partial \overline{v}}{\partial \overline{x}}\right)$$

is called the Laplace operator of the vector $\overline{v}$.

Contravariant components of this vector are

$$(\triangle \overline{v})^l = \left(\frac{\partial \overline{v}}{\partial \overline{x}}\right)^{li}_{,i} = g^{ij}\left(\left(\frac{\partial \overline{v}}{\partial \overline{x}}\right)^{l.}_{.j}\right)_{,i} = g^{ij}(v^l_{,i})_{,j} = g^{ij}[\frac{\partial v^l_{,i}}{\partial K^j} - \Gamma^s_{ji}v^l_{,s} + \Gamma^l_{js}v^s_{,i}] =$$

$$= g^{ij}[(\frac{\partial^2 v^l}{\partial K^j \partial K^i} + \frac{\partial \Gamma^l_{is}}{\partial K^j}v^s + \Gamma^l_{is}\frac{\partial v^s}{\partial K^j}) - \Gamma^s_{ji}(\frac{\partial v^l}{\partial K^s} + \Gamma^l_{s\alpha}v^\alpha) + \Gamma^l_{js}(\frac{\partial v^s}{\partial K^i} + \Gamma^s_{i\alpha}v^\alpha)].$$

After regrouping one has

$$g^{ij}(\frac{\partial^2 v^l}{\partial K^j \partial K^i} - \Gamma^s_{ji}\frac{\partial v^l}{\partial K^s}) + g^{ij}(\frac{\partial \Gamma^l_{is}}{\partial K^j} - \Gamma^\alpha_{ji}\Gamma^l_{\alpha s} + \Gamma^l_{j\alpha}\Gamma^\alpha_{is})v^s + 2g^{ij}\Gamma^l_{is}\frac{\partial v^s}{\partial K^j} =$$

$$= g^{ij}(\triangle v^l) + 2g^{ij}\Gamma^l_{is}\frac{\partial v^s}{\partial K^j} + g^{ij}(\frac{\partial \Gamma^l_{is}}{\partial K^j} - \Gamma^\alpha_{ji}\Gamma^l_{\alpha s} + \Gamma^l_{j\alpha}\Gamma^\alpha_{is})v^s.$$

Acceleration

The vector

$$\frac{d\overline{v}}{dt} = \frac{\partial \overline{v}}{\partial t} + \frac{\partial \overline{v}}{\partial \overline{x}} \langle \overline{v} \rangle$$

is called an acceleration of $\overline{v}$.

Covariant and contravariant components of the acceleration are

$$\left(\frac{d\overline{v}}{dt}\right)_i = \frac{\partial v_i}{\partial t} + v^s v_{i,s} = \frac{\partial v_i}{\partial t} + v^s \frac{\partial v_i}{\partial K^s} - \Gamma^j_{is} v^s v_j,$$

$$\left(\frac{d\overline{v}}{dt}\right)^i = \frac{\partial v^i}{\partial t} + v^s v^i_{\cdot,s} = \frac{\partial v^i}{\partial t} + v^s \frac{\partial v^i}{\partial K^s} + \Gamma^i_{js} v^j v^s,$$

Physical Components of Tensors

If vectors of coordinate bases and cobases are not normed, then components of tensors have different numerical values in different bases even if directions of basis vectors coincides. For using specific physical values it is inconvenient. To correct this there are physical components of tensors.

Definition 1.33. *Numerical values of tensor components divided by the length of corresponding basis or cobasis vectors, which define these components are called physical components of the tensor.*

Let us consider curlinear coordinate system

$$\overline{e}_i = \frac{\partial \overline{x}}{\partial K^i}, \quad \overline{e}^i = \frac{\partial K^i}{\partial \overline{x}} = \nabla K^i.$$

The lengths of the basis and cobasis vectors are

$$|\overline{e}_i|^2 = \overline{e}_i \cdot \overline{e}_i = g_{ii}, \quad |\overline{e}^i|^2 = \overline{e}^i \cdot \overline{e}^i = g^{ii}.$$

The vectors $\overline{e}^1_i = \dfrac{\overline{e}_i}{|\overline{e}_i|}$ have the unite length. If $a_i = ae_i$ are covariant coordinates of a vector a, then the physical components are

$$\widetilde{a}_i = \frac{a_i}{|\overline{e}_i|}$$

For covariant components of a tensor one has

$$L_{ij} = L\left\langle \overline{e}_i, \overline{e}_j \right\rangle = |e_i||e_j| L\left\langle \overline{e}^1_i, \overline{e}^1_j \right\rangle = |e_i||e_j| \widetilde{L}_{ij}$$

or

$$\widetilde{L}_{ij} = \frac{L_{ij}}{(|\overline{e}_i|\,|\overline{e}_j|)}$$

In orthonormal system of coordinates for any tensor its physical components of all types, having the same indices, are equal.

1.7 Special Coordinate Systems

Let $\{e_i\}$ be a fixed right-handed orthonormal basis in R^3. We suppose that a point (vector) $\overline{x}$ has components $\overline{x} = (x, y, z)$ in this basis. Curvilinear coordinates are introduced by transformation $K : R^3 \to A^3$, which acts according to the formulae

$$K(\overline{x}) = (K^1(x, y, z), K^2(x, y, z), K^3(x, y, z)).$$

1.7.1 Rectangular Cartesian Coordinates

Let us consider a transformation K:

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

The inverse transformation K^{-1} is

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

The coordinate surfaces $x = const$, $y = const$ or $z = const$ are planes, which are parallel to yz plane, zx plane or xy plane, respectively. The coordinate curves are straight lines, which are parallel to the x–axis, y–axis or z–axis, respectively. They are orthogonal. The basis $\overline{e}_i = \frac{\partial \overline{x}}{\partial K^i}$ coincides with the cobasis $\overline{e}^i = \frac{\partial K^i}{\partial \overline{x}}$:

$$\overline{e}_1 = \overline{e}^1,\ \overline{e}_2 = \overline{e}^2,\ \overline{e}_3 = \overline{e}^3.$$

The fundamental tensor is an identical tensor

$$(g_{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = (g^{ij});\ |g| = 1.$$

All Christoffel's symbols are equal to zero. The permutational tensor has the value $\varepsilon = 1$. The coordinate system x, y, z is a right-handle coordinate system. Physical components of a vector $\overline{v} = (v_x, v_y, v_z)$ coincide with tensor components

$$(v_1, v_2, v_3) = (v_x, v_y, v_z) = (v^1, v^2, v^3).$$

Physical components of a second order tensor P coincide with corresponding tensor components:

$$(P) = \begin{bmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{bmatrix} = (P_{ij}) = (P^{ij}).$$

The gradient of a function F is

$$\nabla F = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right).$$

The matrix of covariant derivatives of a vector $\overline{v}$ is (i is a number of row)

$$(v^i_{,j}) = \left(\frac{\partial v^i}{\partial x^j} \right) = \begin{bmatrix} \frac{\partial v_x}{\partial x} & \frac{\partial v_x}{\partial y} & \frac{\partial v_x}{\partial z} \\ \frac{\partial v_y}{\partial x} & \frac{\partial v_y}{\partial y} & \frac{\partial v_y}{\partial z} \\ \frac{\partial v_z}{\partial x} & \frac{\partial v_z}{\partial y} & \frac{\partial v_z}{\partial z} \end{bmatrix}.$$

The divergence of a vector $\overline{v}$ is

$$div\, v = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}.$$

The Laplace operator of a function F is

$$\triangle F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2}.$$

The rotor $\overline{\omega}$ of a vector $\overline{v}$ ($\overline{\omega} = rot\,\overline{v}$) has components

$$\omega^1 = \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z},\ \omega^2 = \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x},\ \omega^3 = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}.$$

The divergence of a tensor P is

$$(div\,P)^i = div\,(\overline{P}^i),$$

$$\begin{aligned}\overline{P}^1 &= (P_{xx}, P_{xy}, P_{xz}),\\ \overline{P}^2 &= (P_{yx}, P_{yy}, P_{yz}),\\ \overline{P}^3 &= (P_{zx}, P_{zy}, P_{zz}).\end{aligned}$$

The Laplace operator of a vector $\overline{v}$ is

$$(\triangle\overline{v})^1 = \triangle(v_x),\ (\triangle\overline{v})^2 = \triangle(v_y),\ (\triangle\overline{v})^3 = \triangle(v_z).$$

The acceleration is

$$\left(\frac{d\overline{v}}{dt}\right)^1 = \mathcal{D}(v_x),\ \left(\frac{d\overline{v}}{dt}\right)^2 = \mathcal{D}(v_y),\ \left(\frac{d\overline{v}}{dt}\right)^3 = \mathcal{D}(v_z),$$

where

$$\mathcal{D}(f) = \frac{\partial f}{\partial t} + v_x\frac{\partial f}{\partial x} + v_y\frac{\partial f}{\partial y} + v_z\frac{\partial f}{\partial z}.$$

1.7.2 The cylindrical coordinate system

For the cylindrical coordinate system the transformation K is

$$K^1 = r = \sqrt{x^2+y^2},\ K^2 = \varphi = arctan\,\frac{y}{x},\ K^3 = z.$$

The inverse mapping K^{-1} is

$$x = r\,\cos\varphi,\ y = r\,\sin\varphi,\ z = z,\ (0 \le \varphi \le 2\pi).$$

Coordinate surfaces $r = const > 0$ are circular cylinders (coaxial to z–axis); $\varphi = const$ are half-planes passing through z–axis and $z = const$ are planes perpendicular to z–axis. Coordinate curves are: l_1 (intersection of $\varphi = const$ and $z = const$) are straight rays going from z–axis and perpendicular to it; l_2 (intersection of $r = const$ and $z = const$) are circles (these circles lie on the planes, which are perpendicular to z–axis with a center in the z–axis); l_3 (intersection of $r = const$ and $\varphi = const$) are straight lines that are parallel to z–axis.

The basis and cobasis of the cylindrical coordinate system are orthogonal and consist of the vectors

$$\overline{e}_1 = \frac{\partial\overline{x}}{\partial K^1} = \left(\frac{\partial x}{\partial r}, \frac{\partial y}{\partial r}, \frac{\partial z}{\partial r}\right) = (\cos\,\varphi, \sin\,\varphi, 0),$$

$$\overline{e}_2 = \frac{\partial\overline{x}}{\partial K^2} = \left(\frac{\partial x}{\partial\varphi}, \frac{\partial y}{\partial\varphi}, \frac{\partial z}{\partial\varphi}\right) = r\,(-\sin\,\varphi, \cos\varphi, 0),$$

$$\overline{e}_3 = \frac{\partial\overline{x}}{\partial K^3} = \left(\frac{\partial x}{\partial z}, \frac{\partial y}{\partial z}, \frac{\partial z}{\partial z}\right) = (0, 0, 1),$$

$$\overline{e}^1 = (\frac{\partial K^1}{\partial x}, \frac{\partial K^1}{\partial y}, \frac{\partial K^1}{\partial z}) = (\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}, 0) = (\cos\,\varphi, \sin\,\varphi, 0),$$

$$\overline{e}^2 = (\frac{\partial K^2}{\partial x}, \frac{\partial K^2}{\partial y}, \frac{\partial K^2}{\partial z}) = (-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0) = \frac{1}{r}(-\sin\varphi, \cos\varphi, 0),$$

$$\overline{e}^3 = (\frac{\partial K^3}{\partial x}, \frac{\partial K^3}{\partial y}, \frac{\partial K^3}{\partial z}) = (0,0,1).$$

Hence,

$$\overline{e}_1 = (\cos\varphi, \sin\varphi, 0),\ \overline{e}_2 = r\,(-\sin\varphi, \cos\varphi, 0),\ \overline{e}_3 = (0,0,1),$$

$$\overline{e}^1 = (\cos\varphi, \sin\varphi, 0),\ \overline{e}^2 = \frac{1}{r}\,(-\sin\varphi, \cos\varphi, 0),\ \overline{e}^3 = (0,0,1).$$

The fundamental tensor is

$$(g_{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ (g^{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{bmatrix},\ |g| = r^2.$$

Because the Christoffel's symbols are

$$\Gamma^l_{ij} = \frac{1}{2} g^{ls}\left(\frac{\partial g_{is}}{\partial K^j} + \frac{\partial g_{js}}{\partial K^i} - \frac{\partial g_{ij}}{\partial K^s}\right),$$

then for the cylindrical coordinate system

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r},\ \ \Gamma^1_{22} = -r$$

and all others are equal to zero. For example,

$$\Gamma^2_{12} = \frac{1}{2} g^{2s}\left(\frac{\partial g_{1s}}{\partial K^2} + \frac{\partial g_{2s}}{\partial K^1} - \frac{\partial g_{12}}{\partial K^s}\right) = \frac{1}{2} g^{22}\left(\frac{\partial g_{12}}{\partial K^2} + \frac{\partial g_{22}}{\partial K^1} - \frac{\partial g_{12}}{\partial K^2}\right) = \frac{1}{2}\frac{1}{r^2}2r = \frac{1}{r}.$$

For the permutation tensor is $\varepsilon_{123} = r$ with

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = -\varepsilon_{321} = -\varepsilon_{213} = -\varepsilon_{132}.$$

Since cylindrical coordinate system is orthogonal, then all physical components of any type concide. Let (v_r, v_φ, v_z) be physical components of a vector $\overline{v}$, then the tensor components of the vector $\overline{v}$ are

$$(v^1, v^2, v^3) = (v_r, \frac{v_\varphi}{r}, v_z),\ \ (v_1, v_2, v_3) = (v_r, r v_\varphi, v_z)$$

We remind that physical components of a vector are related with covariant components by the formulae:

$$\widetilde{v}_2 = \frac{v_2}{|e_2|} = \frac{v_2}{r},\ v_2 = r\widetilde{v}_2 = r v_\varphi.$$

Let $P^{ij} = P\left\langle \overline{e}^i, \overline{e}^j \right\rangle$ are contravariant components of a second order tensor P. The physical components $\widetilde{P}^{ij}$ are

$$\widetilde{P}^{ij} = \frac{P^{ij}}{(|\overline{e}^i|\,|\overline{e}^j|)}.$$

For the cylindrical coordinate system, for example,

$$\widetilde{P}^{21} = \frac{P^{21}}{(|\overline{e}^2|\,|\overline{e}^1|)} = r\,P^{21},\ P^{21} = \frac{1}{r}\widetilde{P}^{21} = \frac{1}{r}P_{\varphi r}.$$

Hence, if the physical components of a tensor P are

$$(P) = \begin{bmatrix} P_{rr} & P_{r\varphi} & P_{rz} \\ P_{\varphi r} & P_{\varphi\varphi} & P_{\varphi z} \\ P_{zr} & P_{z\varphi} & P_{zz} \end{bmatrix},$$

then contravariant componets of the tensor P are

$$(P^{ij}) = \begin{bmatrix} P_{rr} & \frac{1}{r}P_{r\varphi} & P_{rz} \\ \frac{1}{r}P_{\varphi r} & \frac{1}{r^2}P_{\varphi\varphi} & \frac{1}{r}P_{\varphi z} \\ P_{zr} & \frac{1}{r}P_{z\varphi} & P_{zz} \end{bmatrix}.$$

The coordinates of the gradient of a function F is

$$(\nabla F)_1 = (\nabla F)^1 = \frac{\partial F}{\partial r}, \;\; (\nabla F)_2 = \frac{\partial F}{\partial \varphi}, (\nabla F)^2 = \frac{1}{r^2}\frac{\partial F}{\partial \varphi},$$

$$(\nabla F)_3 = (\nabla F)^3 = \frac{\partial F}{\partial z}.$$

A matrix of covariant derivatives

$$\Phi^i_{,j} = \frac{\partial \Phi^i}{\partial K^j} + \Gamma^i_{js}\Phi^s;$$

is (here i is the number of a row)

$$\left(v^i_{,j}\right) = \begin{bmatrix} \frac{\partial v_r}{\partial r} & \frac{\partial v_r}{\partial \varphi} - v_\varphi & \frac{\partial v_r}{\partial z} \\ \frac{1}{r}\frac{\partial v_\varphi}{\partial r} & \frac{1}{r}\frac{\partial v_\varphi}{\partial \varphi} + \frac{v_r}{r} & \frac{1}{r}\frac{\partial v_\varphi}{\partial z} \\ \frac{\partial v_z}{\partial r} & \frac{\partial v_z}{\partial \varphi} & \frac{\partial v_z}{\partial z} \end{bmatrix}.$$

Because

$$div\,\overline{v} = \frac{1}{\sqrt{g}}\frac{\partial}{\partial K^i}\left(\sqrt{g}v^i\right),$$

the divergence of a vector $\overline{v}$ can be expressed as follows

$$div\,\overline{v} = \frac{1}{r}\frac{\partial (r\,v_r)}{\partial r} + \frac{1}{r}\frac{\partial v_\varphi}{\partial \varphi} + \frac{\partial v_z}{\partial z}.$$

Similar for the Laplace operator of a function F

$$\triangle F = \frac{1}{\sqrt{g}}\frac{\partial}{\partial K^i}\left(\sqrt{g}g^{is}\frac{\partial F}{\partial K^s}\right),$$

hence,

$$\triangle F = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 F}{\partial \varphi^2} + \frac{\partial^2 F}{\partial z^2}.$$

For the rotor $\overline{\omega} = rot\,\overline{v}$ of a vector $\overline{v}$ one has

$$\omega^1 = \frac{1}{r}\frac{\partial v_z}{\partial \varphi} - \frac{\partial v_\varphi}{\partial z}, \; \omega^2 = \frac{1}{r}\frac{\partial v_r}{\partial z} - \frac{1}{r}\frac{\partial v_z}{\partial r}, \; \omega^3 = \frac{1}{r}\frac{\partial (r v_\varphi)}{\partial r} - \frac{1}{r}\frac{\partial v_r}{\partial \varphi}.$$

Here we used the representation for contravariant components of the vector $rot(\overline{v})$:

$$(rot\,\overline{v})^l = \varepsilon^{ijl}v_{j,i} = \varepsilon^{ijl}\frac{\partial v_i}{\partial K^j}.$$

The divergence of a tensor P is a vector with contravariant components:

$$(div\, P)^i = P^{is}_{,s} = div\,(\overline{P}^i) + \Gamma^i_{js} P^{js}.$$

Hence,

$$(div\, P)^1 = div\,(\overline{P}^1) - \frac{1}{r} P_{\varphi\varphi},\ \overline{P}^1 = (P_{rr}, \frac{1}{r} P_{r\varphi}, P_{rz}),$$

$$(div\, P)^2 = div\,(\overline{P}^2) + \frac{1}{r^2}(P_{r\varphi} + P_{\varphi r}),\ \overline{P}^2 = (\frac{1}{r} P_{\varphi r}, \frac{1}{r^2} P_{\varphi\varphi}, \frac{1}{r} P_{\varphi z}),$$

$$(div\, P)^3 = div\,(\overline{P}^3),\ \overline{P}^3 = (P_{zr}, \frac{1}{r} P_{z\varphi}, P_{zz}).$$

Therefore,

$$(div\, P)^1 = div\,(\overline{P}^1) - \frac{1}{r} P_{\varphi\varphi} = \frac{1}{r}\frac{\partial}{\partial r}(r P_{rr}) + \frac{1}{r}\frac{\partial}{\partial \varphi}(P_{r\varphi}) + \frac{\partial}{\partial z}(P_{rz}) - \frac{1}{r} P_{\varphi\varphi};$$

$$(div\, P)^2 = div\,(\overline{P}^2) + \frac{1}{r^2}(P_{r\varphi} + P_{\varphi r}) =$$

$$= \frac{1}{r}\frac{\partial}{\partial r}(P_{\varphi r}) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}(P_{r\varphi}) + \frac{1}{r}\frac{\partial}{\partial z}(P_{\varphi z}) + \frac{1}{r^2}(P_{r\varphi} + P_{\varphi r});$$

$$(div\, P)^3 = div\,(\overline{P}^3) = \frac{1}{r}\frac{\partial}{\partial r}(r P_{zr}) + \frac{1}{r}\frac{\partial}{\partial \varphi}(P_{z\varphi}) + \frac{\partial}{\partial z}(P_{zz}).$$

The Laplace operator of a vector $\overline{v}$ is the vector with contravariant components:

$$\triangle\,(\overline{v})^1 = \triangle\,(v_r) - \frac{2}{r^2}\frac{\partial v_\varphi}{\partial \varphi} - \frac{v_r}{r^2},$$

$$\triangle\,(\overline{v})^2 = \triangle\left(\frac{v_\varphi}{r}\right) + \frac{2}{r}\frac{\partial}{\partial r}\left(\frac{v_\varphi}{r}\right) + \frac{2}{r^3}\frac{\partial v_r}{\partial \varphi},$$

$$\triangle\,(\overline{v})^3 = \triangle\,(v_z).$$

Here we used

$$\triangle\,(\overline{v})^l = \triangle\left(v^l\right) + 2\, g^{ij}\Gamma^l_{is}\frac{\partial v^s}{\partial K^j} + g^{ij}\left(\frac{\partial \Gamma^l_{ip}}{\partial K^j} + \Gamma^s_{ip}\Gamma^l_{js} - \Gamma^s_{ij}\Gamma^l_{ps}\right) v^p.$$

The acceleration has the components

$$\left(\frac{d\overline{v}}{dt}\right)^1 = \mathcal{D}(v_r) - \frac{v_\varphi^2}{r};\ \left(\frac{d\overline{v}}{dt}\right)^2 = \frac{1}{r}\mathcal{D}(v_\varphi) + \frac{v_r v_\varphi}{r^2};\ \left(\frac{d\overline{v}}{dt}\right)^3 = \mathcal{D}(v_z),$$

where

$$\mathcal{D}(f) = \frac{\partial f}{\partial t} + v_r\frac{\partial f}{\partial r} + \frac{v_\varphi}{r}\frac{\partial f}{\partial \varphi} + v_z\frac{\partial f}{\partial z}.$$

1.7.3 Spherical Coordinate system

For a spherical coordinate system the transformation K is

$$K^1 = r = \sqrt{x^2+y^2+z^2},\ K^2 = \theta = arctan\,\frac{\sqrt{x^2+y^2}}{z},\ K^3 = \varphi = arctan\,\frac{y}{x}.$$

The inverse mapping K^{-1} is

$$x = r\,\sin\,\theta\cos\varphi,\ y = r\,\sin\,\theta\sin\varphi,\ z = r\,\cos\,\theta,\ (0 \le \varphi \le 2\pi,\ 0 \le \theta \le \pi).$$

Coordinate surfaces are: $r = const > 0$ are spheres having centers at the origin, $\theta = const$ are half–cones having vertex at the origin, $\varphi = const$ are half–planes passing through z–axis. Coordinate curves are: l_1 (intersection of $\theta = const$ and $\varphi = const$) are straight rays emanating from the origin, l_2 (intersection of $r = const$ and $\varphi = const$) are semi–circles having centers at the origin and which diameters lie on z–axis, l_3 (intersection of $r = const$ and $\theta = const$) are circles, which lie on planes perpendicular to the z–axis.

The spherical coordinate system is orthogonal with the bais and cobasis vectors

$$\overline{e}_1 = (\sin\,\theta\cos\varphi\,,\sin\,\theta\sin\,\varphi,\cos\,\theta) = \overline{e}^1,$$

$$\overline{e}_2 = r\,(\cos\,\theta\,\cos\,\varphi,\cos\,\theta\,\sin\,\varphi,-\sin\,\theta),\ \overline{e}^2 = \frac{1}{r^2}\overline{e}_2,$$

$$\overline{e}_3 = r\sin\,\theta\,(-\sin\,\varphi,\cos\,\varphi,0),\ \overline{e}^3 = \frac{1}{r^2\sin^2\theta}\overline{e}_3.$$

The fundamental tensor is

$$(g_{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{bmatrix},\ (g^{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{bmatrix},\ |g| = r^4\sin^2\,\theta.$$

The Christoffel's symbols are (we write down only nonvanishing symbols)

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r},\ \Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r}\ \ \Gamma^1_{22} = -r,$$

$$\Gamma^3_{23} = \Gamma^3_{32} = \cot\,\theta,\ \Gamma^1_{33} = -r\,\sin^2\,\theta,\ \Gamma^2_{33} = -\sin\,\theta\,\cos\,\theta.$$

For example,

$$\Gamma^1_{33} = \frac{1}{2}g^{1s}\left(\frac{\partial g_{3s}}{\partial K^3} + \frac{\partial g_{3s}}{\partial K^3} - \frac{\partial g_{33}}{\partial K^s}\right) = \frac{1}{2}g^{11}\left(\frac{\partial g_{31}}{\partial K^3} + \frac{\partial g_{31}}{\partial K^3} - \frac{\partial g_{33}}{\partial K^1}\right) =$$

$$= \frac{1}{2}\left(-\frac{\partial r^2\sin^2\theta}{\partial r}\right) = -\frac{1}{2}2r\,\sin^2\theta = -r\,\sin^2\,\theta.$$

Spherical coordinate system is right–handle coordinate system, hence $\varepsilon_{123} = \sqrt{|g|} = r^2\,\sin\,\theta$.

Let $\overline{v} = (v_r, v_\theta, v_\varphi)$ be physical components of a vector $\overline{v}$, then tensor components of the vector $\overline{v}$ are

$$(v^1, v^2, v^3) = (v_r, \frac{v_\theta}{r}, \frac{v_\varphi}{r\,\sin\,\theta}),\quad (v_1, v_2, v_3) = (v_r, rv_\theta, r\,\sin\,\theta\,v_\varphi).$$

For example, $|\overline{e}_3| = \sqrt{(\overline{e}_3\cdot\overline{e}_3)} = \sqrt{r^2\sin^2\,\theta\,(\sin^2\,\varphi+\cos^2\,\varphi)} = r\sin\,\theta$ and therefore,

$$v_\varphi = \frac{v_3}{|\overline{e}_3|} = \frac{v_3}{r\sin\,\theta},\quad \Rightarrow\ v_3 = r\,\sin\,\theta\,v_\varphi.$$

Let a second order tensor P has the physical components:

$$(P)=\begin{bmatrix} P_{rr} & P_{r\theta} & P_{r\varphi} \\ P_{\theta r} & P_{\theta\theta} & P_{\theta\varphi} \\ P_{\varphi r} & P_{\varphi\theta} & P_{\varphi\varphi} \end{bmatrix}.$$

Then contravariant components of P are

$$(P^{ij})=\begin{bmatrix} P_{rr} & \frac{1}{r}P_{r\theta} & \frac{1}{r\,\sin\,\theta}P_{r\varphi} \\ \frac{1}{r}P_{\theta r} & \frac{1}{r^2}P_{\theta\theta} & \frac{1}{r^2\,\sin\,\theta}P_{\theta\varphi} \\ \frac{1}{r\,\sin\,\theta}P_{\varphi r} & \frac{1}{r^2\,\sin\,\theta}P_{\varphi\theta} & \frac{1}{r^2\,\sin^2\,\theta}P_{\varphi\varphi} \end{bmatrix}.$$

The gradient of a function F is

$$(\nabla F)_1=(\nabla F)^1=\frac{\partial F}{\partial r},\quad (\nabla F)_2=\frac{\partial F}{\partial \theta},\quad (\nabla F)^2=\frac{1}{r^2}\frac{\partial F}{\partial \theta},$$

$$(\nabla F)_3=\frac{\partial F}{\partial \varphi},\quad (\nabla F)^3=\frac{1}{r^2\,\sin^2\,\theta}\frac{\partial F}{\partial \varphi}.$$

The matrix of covariant derivatives is (here i is a number of row)

$$\left(v^i_{,j}\right)=\begin{bmatrix} \frac{\partial v_r}{\partial r} & \frac{\partial v_r}{\partial \theta}-v_\theta & \frac{\partial v_r}{\partial \varphi}-\sin\theta v_\varphi \\ \frac{1}{r}\frac{\partial v_\theta}{\partial r} & \frac{1}{r}\frac{\partial v_\theta}{\partial \theta}+\frac{v_r}{r} & \frac{1}{r}\frac{\partial v_\theta}{\partial \varphi}-\frac{\cos\theta}{r}v_\varphi \\ \frac{1}{r\sin\theta}\frac{\partial v_\varphi}{\partial r} & \frac{1}{r\sin\theta}\frac{\partial v_\varphi}{\partial \theta} & \frac{1}{r\sin\theta}\frac{\partial v_\varphi}{\partial \varphi}+\frac{v_r}{r}+\frac{\cot\theta}{r}v_\theta \end{bmatrix}.$$

In the strength of

$$div\,\overline{v}=\frac{1}{\sqrt{g}}\frac{\partial}{\partial K^i}\left(\sqrt{g}v^i\right),$$

the divergence of a vector $\overline{v}$ is

$$div\,\overline{v}=\frac{1}{r^2}\frac{\partial(r^2\,v_r)}{\partial r}+\frac{1}{r\sin\theta}\frac{\partial\sin\theta\,v_\theta}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial v_\varphi}{\partial\varphi},$$

The Laplace operator of a function is

$$\triangle F=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial F}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 F}{\partial\varphi^2}.$$

We used here

$$\triangle F=\frac{1}{\sqrt{g}}\frac{\partial}{\partial K^i}\left(\sqrt{g}g^{is}\frac{\partial F}{\partial K^s}\right).$$

The rotor $\overline{\omega}=rot\,\overline{v}$ of a vector $\overline{v}$ is

$$\omega^1=\frac{1}{r\sin\theta}\left(\frac{\partial(\sin\,\theta\,v_\varphi)}{\partial\theta}-\frac{\partial v_\theta}{\partial\varphi}\right),\quad \omega^2=\frac{1}{r^2\sin\theta}\left(\frac{\partial v_r}{\partial\varphi}-\sin\theta\frac{\partial(rv_\varphi)}{\partial r}\right),$$

$$\omega^3=\frac{1}{r^2\sin\theta}\left(\frac{\partial(rv_\theta)}{\partial r}-\frac{\partial v_r}{\partial\theta}\right).$$

We used the representation

$$(rot\,\overline{v})^l=\varepsilon^{ijl}v_{j,i}=\varepsilon^{ijl}\frac{\partial v_i}{\partial K^j}.$$

For example,

$$\omega^1 = \varepsilon^{ij1}\frac{\partial v_j}{\partial K^i} = \varepsilon^{231}\frac{\partial v_3}{\partial K^2} + \varepsilon^{321}\frac{\partial v_2}{\partial K^3} = \varepsilon^{123}\left(\frac{\partial v_3}{\partial \theta} - \frac{\partial v_2}{\partial \varphi}\right) =$$
$$= \varepsilon^{123}\left(\frac{\partial(r\sin\,\theta\, v_\varphi)}{\partial\theta} - \frac{\partial r v_\theta}{\partial\varphi}\right) = r\,\varepsilon^{123}\left(\frac{\partial(\sin\,\theta\, v_\varphi)}{\partial\theta} - \frac{\partial v_\theta}{\partial\varphi}\right) =$$
$$= \frac{1}{r\sin\theta}\left(\frac{\partial(\sin\,\theta\, v_\varphi)}{\partial\theta} - \frac{\partial v_\theta}{\partial\varphi}\right)$$

where $\varepsilon^{123} = \dfrac{1}{\varepsilon_{123}} = \dfrac{1}{|g|} = \dfrac{1}{r^2\sin\theta}$.

The divergence of a tensor P is the vector with the contravariant coordinates

$$(div\,P)^i = P^{is}_{,s} = div\,(\overline{P}^i) + \Gamma^i_{js}P^{js}.$$

Thus, in spherical coordinate system

$$(div\,P)^1 = div\,(\overline{P}^1) - \frac{1}{r}(P_{\theta\theta} + P_{\varphi\varphi}),$$
$$(div\,P)^2 = div\,(\overline{P}^2) + \frac{1}{r}(P_{r\theta} + P_{\theta r}) - \frac{\cot\theta}{r^2}P_{\varphi\varphi},$$
$$(div\,P)^3 = div\,(\overline{P}^3) + \frac{1}{r^2\sin\theta}(P_{r\varphi} + P_{\varphi r}) + \frac{\cot\theta}{r^2\sin\theta}(P_{\theta\varphi} + P_{\varphi\theta}),$$

where

$$\overline{P}^1 = (P_{rr}, \frac{1}{r}P_{r\theta}, \frac{1}{r\sin\theta}P_{r\varphi}),\ \overline{P}^2 = (\frac{1}{r}P_{\theta r}, \frac{1}{r^2}P_{\theta\theta}, \frac{1}{r^2\sin\theta}P_{\theta\varphi}),$$
$$\overline{P}^3 = (\frac{1}{r\sin\theta}P_{\varphi r}, \frac{1}{r^2\sin\theta}P_{\varphi\theta}, \frac{1}{r^2\sin^2\theta}P_{\varphi\varphi}).$$

The Laplace operator of a vector $\overline{v}$ is

$$\triangle\,(\overline{v})^1 = \triangle\,(v_r) - \frac{2}{r}\frac{\partial v_\theta}{\partial\theta} - \frac{2}{r^2\sin\theta}\frac{\partial v_\varphi}{\partial\varphi} - \frac{2v_r}{r^2} - \frac{2\cot\theta}{r^2}v_\theta,$$
$$\triangle\,(\overline{v})^2 = \triangle\left(\frac{v_\theta}{r}\right) + \frac{2}{r^2}\frac{\partial v_\theta}{\partial r} + \frac{2}{r^3}\frac{\partial v_r}{\partial\theta} - \frac{2\cot\theta}{r^3\sin\theta}\frac{\partial v_\varphi}{\partial\varphi} - \frac{v_\theta}{r^3\sin^2\theta},$$
$$\triangle\,(\overline{v})^3 = \triangle\left(\frac{v_\varphi}{r\sin\theta}\right) + \frac{2}{r\sin\theta}\frac{\partial}{\partial r}\left(\frac{v_\varphi}{r}\right) + \frac{2\cot\theta}{r^3}\frac{\partial}{\partial\theta}\left(\frac{v_\varphi}{\sin\theta}\right) +$$
$$+\frac{2}{r^3\sin^2\theta}\frac{\partial v_r}{\partial\varphi} + \frac{2\cot\theta}{r^3\sin^2\theta}\frac{\partial v_\theta}{\partial\varphi}.$$

Here we used

$$\triangle\,(\overline{v})^l = \triangle\left(v^l\right) + 2\,g^{ij}\Gamma^l_{is}\frac{\partial v^s}{\partial K^j} + g^{ij}\left(\frac{\partial\Gamma^l_{ip}}{\partial K^j} + \Gamma^s_{ip}\Gamma^l_{js} - \Gamma^s_{ij}\Gamma^l_{ps}\right)v^p.$$

The acceleration of a vector $\overline{v}$ has components

$$\left(\frac{d\overline{v}}{dt}\right)^1 = \mathcal{D}(v_r) - \frac{v_\varphi^2 + v_\theta^2}{r},\quad \left(\frac{d\overline{v}}{dt}\right)^2 = \frac{1}{r}\mathcal{D}(v_\theta) + \frac{v_r v_\theta - \cot\theta\, v_\varphi^2}{r^2},$$
$$\left(\frac{d\overline{v}}{dt}\right)^3 = \frac{1}{r\sin\theta}\mathcal{D}(v_\varphi) + \frac{v_r v_\varphi + \cot\theta\, v_\varphi v_\theta}{r^2\sin\theta},$$

where

$$\mathcal{D}(f) = \frac{\partial f}{\partial t} + v_r\frac{\partial f}{\partial r} + \frac{v_\theta}{r}\frac{\partial f}{\partial\theta} + \frac{v_\varphi}{r\sin\theta}\frac{\partial f}{\partial\varphi}.$$

Chapter 2

Mathematical modelling

2.1 Mathematical background

Let $\Omega \subset R^n$ be a bounded open set in R^n.

2.1.1 Volume.

Definition. We say that a set $\omega \subset R^n$ has a piecewise smooth boundary $\partial\omega$, if it is possible to represent the set $\partial\omega$ in the form of union of a finite number of subsets σ_j (smooth pieces) with the following property: for every σ_j there is a point $x_0 \in \sigma_j$ and basis $\{e_i\} \subset R^n$ (in which $x = x_i e_i$) that

$$\sigma_j = \{x \in R^n : \ x_n = \psi^j(x_1, x_2, ..., x_{n-1}), \ (x_1, x_2, ..., x_{n-1}) \in \sigma_j^n \subset R^{n-1}\},$$

where $\psi^j : \sigma_j^n \to R$ is a continuously differentiable mapping (σ_j^n is called a projection of σ_j on R^{n-1}).

A bounded set $\omega \subset R^n$ with a piecewise smooth boundary $\partial\omega$ is called a volume.

Examples (of volume in R^3): cube, ball

$$x_1^2 + x_2^2 + x_3^2 = k^2$$

tetrahedron

$$\sigma_1 : x_1 = 0$$

$$\sigma_2 : x_2 = 0; \quad \sigma_3 : x_2 = 0$$

$$\sigma_4 : x_1 + x_2 + x_3 = 1$$

segment of cylinder

$$\sigma_1 : x_1^2 + x_2^2 = 1$$

$$\sigma_2 : x_3 = 0$$

$$\sigma_3 : x_3 = 1$$

2.1.2 Additive functions of a set

Let $\{\omega\}$be a collection of all volumes $\omega \subset \Omega \subset R^n$. A mapping $\Phi : \{\omega\} \to R^m$ is called an additive mapping, if for any disjoint volumes $\omega_1, \omega_2 \subset \{\omega\}$ it is satisfired the equality

$$\Phi(\omega_1 \cup \omega_2) = \Phi(\omega_1) + \Phi(\omega_2).$$

Let $x \in \Omega$ and a set $\square_\alpha(x)$ be open cube (hypercube) in R^n:

$$\square_\alpha(x) = \{y \in R^n : ||y - x||_\infty < \alpha\}$$

Definition. An additive mapping $\Phi : \{\omega\} \to R^m$ is called continuous at the point x if $\lim_{\alpha\to 0} \Phi(\square_\alpha(x)) = 0$.

If a mapping Φ is continuous at any point $x \in \Omega$, then it is called a continuous mapping in Ω.

An additive mapping Φ is called differentiable at a point $x \in \Omega$, if there is a vector $u \in R^m$ that

$$\lim_{\alpha\to 0} ||(2\alpha)^{-n}\Phi(\square_\alpha(x)) - u|| = 0.$$

The vector $u \in R^m$ is called a derivative of the mapping Φ and we write $u = \Phi'$.

A mapping Φ is called differentiable in Ω, if it is differentiable at every point $x \in \Omega$. In this case the mapping $u : \Omega \to R^m$ appears acting by the formula $u(x) = \Phi'(x)$.

2.1.3 Integral.

Let $u : \Omega \to R^m$ be a continuous function. Then we have the following main theorem of integral theory.

There exists only one continuously differentiable in Ω additive function of set $\Phi_u : \{\omega\} \to R^m$ for which $\Phi'_u(x) = u(x)$ for all $x \in \Omega$. This function Φ_u is called a primitive function of u. Its values for any volume $\omega \subset \Omega$ are integrals of the function $u : \Omega \to R^m$ on volume ω:

$$\Phi_u(\omega) = \int_\omega u \, d\omega.$$

2.1.4 Surface measure and integral

Let a smooth hypersurface γ with a projection γ_1 on R^{n-1} be given by the equation

$$x_n = \psi(x_1, x_2, ..., x_{n-1}), \quad (x_1, x_2, ..., x_{n-1}) \in \gamma_1,$$

where $\psi : \gamma_1 \to R$ is a continuously differentiable function on γ_1. Let a subset $\sigma \subset \gamma$ be open with respect to γ and it has a projection σ_1 on R^{n-1}. Then σ_1 is open set (in R^{n-1}).

Definition. The number

$$\mu_1(\sigma) = \int_{\sigma_1} \sqrt{1 + |\nabla\psi|^2} \, d\sigma_1$$

is called a surface Lebesque measure of the set σ.

With the help of this measure an integral for a continuous function $\varphi : \gamma \to R^m$ is defined. We consider additive continuously differentiable function of sets of R^{n-1}, which has the derivative

$$\varphi\sqrt{1 + |\nabla\psi|^2}$$

We denote the primitive of this function by $\int_{\sigma_1} \varphi\sqrt{1 + |\nabla\psi|^2} \, d\sigma_1$. And we define

$$\int_\sigma \varphi \, d\sigma = \int_{\sigma_1} \varphi\sqrt{1 + |\nabla\psi|^2} \, d\sigma_1$$

If γ is piecewise smooth surface, which consists of a finite number N of disjoint smooth pieces γ_i, such that $\gamma = \cup_{i=1}^N \gamma_i$, then $\int_\sigma \varphi \, d\sigma = \sum_{i=1}^N \int_{\gamma_i} \varphi \, d\sigma$.

2.1.5 Gauss-Ostrogradskii theorem

Let $\Omega \subset R^n$ be an open set and a function $u : \Omega \to R^n$ be a continuously differentiable on Ω. We consider such volumes $\omega \subset \Omega$ that the closure $\bar{\omega} \subset \Omega$ ($\bar{\omega}$ is a closure of ω in R^n). Let $\partial\omega$ be a piecewise smooth boundary of ω:

$$\partial\omega = \cup_{i=1}^N \gamma_i.$$

We define a positive side of $\partial\omega$ by the following way. Let $x_n = \psi^i(x_1, x_2, ..., x_{n-1})$ be a parametrization of the surface γ_i. If there exists such $\delta > 0$ that for all λ ($0 < \lambda < \delta$) the points $(x_1, x_2, ..., x_{n-1}, \lambda + \psi(x_1, x_2, ..., x_{n-1})) \notin \omega$, then the unit vector

$$n = \frac{1}{\sqrt{1 + |\nabla\psi|^2}}(\frac{\partial\psi^i}{\partial x_1}, \frac{\partial\psi^i}{\partial x_2}, ..., \frac{\partial\psi^i}{\partial x_{n-1}}, 1)$$

defines a positive side of $\partial\omega$ and the vector n is called a positive or outward drawn unit normal. Otherwise, if $(x_1, x_2, ..., x_{n-1}, \lambda + \psi(x_1, x_2, ..., x_{n-1})) \in \omega$, then the outward drawn unit normal vector is the vector

$$n = \frac{1}{\sqrt{1 + |\nabla\psi|^2}}(\frac{\partial\psi^i}{\partial x_1}, \frac{\partial\psi^i}{\partial x_2}, ..., \frac{\partial\psi^i}{\partial x_{n-1}}, -1)$$

The Gauss-Ostrogradskii theorem asserts that the equality (Gauss-Ostrogradskii formula):

$$\int_\omega div(u)\, d\omega = \int_{\partial\omega} (nu)\, d\sigma.$$

is fair.

Also it is fair the generalization of this formula on the case of continuously differentiable mapping $P : \Omega \to L(R^n)$, i.e. for a tensor of the second rank:

$$\int_\omega div(P)\, d\omega = \int_{\partial\omega} P < n > \, d\sigma.$$

Proof. Let a be a test vector. Then we have the chain of equalities

$$a \int_\omega div(P)\, d\omega = \int_\omega div(P^* < a >)\, d\omega = \int_{\partial\omega} nP^* < a > \, d\sigma = .$$

$$= \int_{\partial\omega} aP < n > \, d\sigma = a \int_{\partial\omega} P < n > \, d\sigma.$$

Lemma. If a function h is continuous on $\Omega \subset R^3$ (Ω is open set in R^3) and if for any volume $\omega \subset \Omega$ we have the equality $\int_\omega h\, d\omega = 0$, then $h = 0$ on Ω.

Rule. If we have a one-to-one continuously differentiable function $x : \Omega_0 \to \Omega$, where $\Omega_0 \subset R^n$, $\Omega \subset R^n$, and if $\omega = x(\omega_0)$, where $\omega_0 \subset \Omega_0$, then

$$\int_\omega f\, d\omega = \int_{\omega_0} fJ\, d\omega_0.$$

Here J is the Jacobian.

2.1.6 Linear transformations.

Let us consider transformations $L(R^n)$.

Definition. Linear transformation $A : R^n \to R^n$ is called symmetrical transformation if

$$A^* = A.$$

Linear transformation B is called transpose one of A if for all vectors $a, b \in R^n$ we have

$$(a, Ab) = (Ba, b).$$

In matrix representation it means that

$$[B] = [A]^*$$

A set of symmetric linear transformations we denote $L_s(R^n)$.

Definition. Linear transformation $O \in L(R^n)$ is called orthogonal transformation if it satisfies to

$$O \circ O^* = I.$$

For the orthogonal transformations it is fair formulae

$$O^* = O^{-1},\ \det([O]) = \pm 1.$$

Composition of two orthogonal transformations is orthogonal transformation.

Definition. Linear transformation $\overline{A}$ is called equivalent transformation to A if there exists orthogonal transformation O that

$$\overline{A} = O \circ A \circ O^*$$

Let function $f : L_s(R^n) \to L_s(R^n)$ be function which mappings linear symmetrical transformation into linear symmetrical transformation.

Definition. A mapping $f : L_s(R^n) \to L_s(R^n)$ is called an invariant mapping with respect to orthogonal transformations if for each orthogonal transformation O we have

$$f(O \circ A \circ O^*) = O \circ f(A) \circ O^*$$

In continuum mechanics such functions are called isotropic functions.

Theorem. All continuous isotropic functions $f : L_s(R^n) \to L_s(R^n)$ have representation

$$f(A) = \sum_{k=0}^{n-1} \varphi_k A^k,$$

where coefficients φ_k are scalar functions only invariants of linear transformation A.

Proof. We will prove this theorem for $n = 3$.

We choose orthogonal basis $\{e_i\}$ and we consider all linear transformations in this basis. Let A be a symmetrical linear transformation. Because matrix A is a symmetrical matrix then there exists an orthogonal matrix O such that

$$\overline{A} = O \circ A \circ O^* = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix},$$

where a_i $(i = 1, 2, 3)$ are eigenvalues of matrix A.

We show that matrix $\overline{B} = O \circ f(A) \circ O^* = f(\overline{A})$ is a diagonal matrix, too.

Really, let matrix $\overline{B}$ have representation

$$\overline{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

and we take the orthogonal matrix

$$O_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then by virtue of $O_1 O$ is orthogonal matrix, we get

$$O_1 \overline{B} O_1^* = O_1(Of(A)O^*)O_1^* = f(O_1 O A (O_1^* O)^*) = f(O_1 \overline{A} O_1^*).$$

But $O_1 \overline{A} O_1^* = \overline{A}$, and

$$O_1 \overline{B} O_1^* = \begin{pmatrix} b_{11} & -b_{12} & -b_{13} \\ -b_{21} & b_{22} & b_{23} \\ -b_{31} & b_{32} & b_{33} \end{pmatrix} = f(\overline{A}) = \overline{B}.$$

We receive that $b_{12} = b_{13} = 0$. Similarly, we find that $b_{23} = b_{32} = 0$ and $b_{ii} = f_i(a_1, a_2, a_3)$. We take orthogonal transformation

$$O_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then

$$O_2 \overline{B} O_2^* = \begin{pmatrix} b_{22} & 0 & 0 \\ 0 & b_{11} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} = f(O_2 \overline{A} O_2^*) = f \begin{pmatrix} a_2 & 0 & 0 \\ 0 & a_1 & 0 \\ 0 & 0 & a_3 \end{pmatrix},$$

it means, that

$$f_1(a_2, a_1, a_3) = f_2(a_1, a_2, a_3),\ f_2(a_2, a_1, a_3) = f_1(a_1, a_2, a_3),\ f_3(a_2, a_1, a_3) = f_3(a_1, a_2, a_3).$$

We will consider some cases.

1^0. Assume that $(a_1 - a_2)(a_2 - a_3)(a_3 - a_1) \neq 0$. Let us consider linear system of equations

$$b_{ii} = \alpha + \beta a_i + \gamma a_i^2 \ (i = 1, 2, 3)$$

with respect to α, β, γ. Determinant of this system is (Gramma's determinant)

$$\Delta = \det \begin{pmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{pmatrix} = (a_1 - a_2)(a_2 - a_3)(a_3 - a_1) \neq 0.$$

Because $\Delta \neq 0$, then we get

$$\alpha = \frac{1}{\Delta} \det \begin{pmatrix} b_{11} & a_1 & a_1^2 \\ b_{22} & a_2 & a_2^2 \\ b_{33} & a_3 & a_3^2 \end{pmatrix}, \beta = \frac{1}{\Delta} \det \begin{pmatrix} 1 & b_{11} & a_1^2 \\ 1 & b_{22} & a_2^2 \\ 1 & b_{33} & a_3^2 \end{pmatrix}, \gamma = \frac{1}{\Delta} \det \begin{pmatrix} 1 & a_1 & b_{11} \\ 1 & a_2 & b_{22} \\ 1 & a_3 & b_{33} \end{pmatrix}.$$

Functions $\alpha = \alpha(a_1, a_2, a_3), \beta = \beta(a_1, a_2, a_3), \gamma = \gamma(a_1, a_2, a_3)$ have property that they are unaltered by interchanges of pairs of a_1, a_2, a_3. Really, for example, $\Delta' = -\Delta$,

$$\alpha(a_2, a_1, a_3) = \frac{1}{\Delta'} \det \begin{pmatrix} f_1(a_2, a_1, a_3) & a_2 & a_2^2 \\ f_2(a_2, a_1, a_3) & a_1 & a_1^2 \\ f_3(a_2, a_1, a_3) & a_3 & a_3^2 \end{pmatrix} = -\frac{1}{\Delta} \det \begin{pmatrix} f_2(a_1, a_2, a_3) & a_2 & a_2^2 \\ f_1(a_1, a_2, a_3) & a_1 & a_1^2 \\ f_3(a_1, a_2, a_3) & a_3 & a_3^2 \end{pmatrix} = \alpha(a_1, a_2, a_3).$$

Definition. Any function $f(x_1, x_2, x_3)$ whose value is unaltered by interchanges of pairs of x_1, x_2, x_3 is said to be symmetrical in x_1, x_2, x_3.

Theorem. Continuous symmetrical function can be expressed as a function of the invariants $f(x_1, x_2, x_3) = g(J_1, J_2, J_3)$, where $J_1 = x_1 + x_2 + x_3,\ J_2 = x_1x_2 + x_2x_3 + x_3x_1$.

Then we receive that

$$\alpha = \alpha(J_1(A), J_2(A), J_3(A)),\ \beta = \beta(J_1(A), J_2(A), J_3(A)),\ \gamma = \gamma(J_1(A), J_2(A), J_3(A))$$

and $\overline{B} = \alpha I + \beta\overline{A} + \gamma\overline{A}^2$. But $\overline{B} = O \circ B \circ O^*$, then

$$B = O^* \circ \overline{B} \circ O = \alpha I + \beta O^* \circ \overline{A} \circ O + \gamma O^* \circ \overline{A}^2 \circ O = \alpha I + \beta A + \gamma A^2.$$

2^0. Now let $a_1 = a_2 \neq a_3$, then $f_2(a_1, a_2, a_3) = f_1(a_1, a_2, a_3)$ and we consider equations

$$b_{ii} = \alpha + \beta a_i + \gamma 0 \ (i = 1, 3)$$

By virtue of $a_1 \neq a_3$, we have

$$\alpha = \frac{1}{a_3 - a_1} \det \begin{pmatrix} b_{11} & a_1 \\ b_{33} & a_3 \end{pmatrix},\ \beta = \frac{b_{33} - b_{11}}{a_3 - a_1}.$$

Functions $\alpha = \alpha(a_1, a_2, a_3), \beta = \beta(a_1, a_2, a_3)$ are again symmetrical functions of a_1, a_2, a_3, and they are functions of invariants

$$\alpha = \alpha(J_1(A), J_2(A), J_3(A)),\ \beta = \beta(J_1(A), J_2(A), J_3(A)).$$

In this case we have $\overline{B} = \alpha I + \beta\overline{A}$, then we get

$$B = O^* \circ \overline{B} \circ O = \alpha I + \beta O^* \circ \overline{A} \circ O = \alpha I + \beta A.$$

3^0. Let $a_1 = a_2 = a_3$, then we have $b_{11} = b_{22} = b_{33}$ and we consider equation

$$b_{11} = \alpha$$

Then α is a symmetrical function of a_1, a_2, a_3. We have $\alpha = \alpha(J_1(A), J_2(A), J_3(A))$ and

$$B = O^* \circ \overline{B} \circ O = \alpha I.$$

2.2 Differential equations

2.2.1 Cauchy problem.

Let $u : \Omega \times \tau \to R^n$ be a function, where Ω is an open set $\Omega \subset R^n(x)$ and $\tau \subset R^1(t)$ is an open interval. We study the problem. Find a differentiable function $x : \tau \to R^n$ such that for all $t \in \tau$

$$\frac{dx}{dt} = u(x, t). \tag{2.1}$$

We require that the function $x(t)$ satisfies the initial condition

$$x(t_0) = \xi \tag{2.2}$$

for some point $t = t_0 \in \tau$ and $\xi \in \Omega$.

Definition The problem of finding a solution $x(t)$ of equation (2.1) and satisfying condition (2.2):

$$\frac{dx}{dt} = u(x,t), \; x(t_0) = \xi \tag{2.3}$$

is called a Cauchy problem.

We define the following condition on a function $u : \Omega \times \tau \to R^n$.

Condition D. (a) A function $u : \Omega \times \tau \to R^n$ is continuous on $\Omega \times \tau \subset R^{n+1}$

(b) For all fixed $t \in \tau$ the function $u_t : \Omega \to R^n$ is continuously differentiable on Ω. Here $u_t(x) = u(x,t)$.

Theorem. Let a function $u(x,t)$ ($u : \Omega \times \tau \to R^n$) satisfies the condition D. There exists one and only one solution $x = x(\xi, t)$ of the Cauchy problem (2.3), defined on some interval $\tau_1 \subset \tau$, which contains the point $t_0 \in \tau_1$. For all fixed $t \in \tau_1$ this solution is continuously differentiable as the function $x : \xi \to x(\xi, t)$ and the derivative $\frac{\partial x}{\partial \xi}$ satisfies the perturbation equation

$$\frac{d}{dt}(\frac{\partial x}{\partial \xi}) = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \xi}, \quad \frac{\partial x}{\partial \xi}(t_0) = I.$$

Corollary (Euler's formula). Let a function $u(x,t)$ satisfies the condition D. The formula

$$\frac{dJ}{dt} = J\, div(u)$$

is correct.

Proof. Let $L = \frac{\partial x}{\partial \xi}$, $N = \frac{\partial u}{\partial x}$ be matrices with elements (L_{ij}) and (N_{ij}), and $J = \det(L)$. By virtue of the theorem we have

$$\frac{\partial L_{ij}}{\partial t} = N_{i\alpha} L_{\alpha j}.$$

We wish to prove that

$$\frac{\partial}{\partial t} J = J\; tr(N). \tag{2.4}$$

Let A_{ij} be a cofactor of L_{ij}, then

$$J\delta_{ij} = L_{i\alpha} A_{j\alpha},$$

where δ_{ij} is the Kroneker's symbol. In particular, if $i = j$, then (the symbol i is fixed)

$$J = L_{i\alpha} A_{i\alpha} \tag{2.5}$$

Because $A_{i\alpha}$ does not depend on L_{ij} for all α, then it is encountered only one time in the sum (2.5), it means that

$$A_{ij} = \frac{\partial}{\partial L_{ij}} J$$

Differentiating the determinant J with respect to time t and using (2.4), we obtain

$$\frac{\partial}{\partial t} J = \frac{\partial J}{\partial L_{\alpha\beta}} \frac{\partial L_{\alpha\beta}}{\partial t} = A_{\alpha\beta}(N_{\alpha\gamma} L_{\gamma\beta}) = J\delta_{\alpha\gamma} N_{\alpha\gamma} = J\, N_{\alpha\alpha} = J\, tr(N).$$

Remark. If one consider J as a function of the independent variables (x, t), then $J(\xi, t) = J(x(\xi, t), t)$ and
$$\frac{dJ}{dt} \equiv \frac{\partial J}{\partial t} + \frac{\partial J}{\partial x_\alpha}\frac{dx_\alpha}{dt}.$$
In this case the Euler's formula has the representation
$$\frac{dJ}{dt} = J\,\frac{\partial u_\alpha}{\partial x_\alpha}.$$

2.3 Subject and method of continuum mechanics

1. The subject of study in continuum mechanics is physical bodies (physical continuum), having properties of *continuous media* and *internal mobility.*

A physical continuum is a medium field with a continuous matter such that every part of the medium, however small, is itself a continuum and entirely filled with the matter. A property of internal mobility (or deformation) consists in translating separated parts of the physical continuum with respect to each other, but the external form is invariant.

Strictly saying, by virtue of atom-molecular construction of any matter, we have no such physical bodies. When we say about physical continuum we suppose that the property of continuous matter is approximately fulfilled. It means that we regard the process in which the character scale of molecular process is less than minimum scale of study interaction with media. These scales are distinguished for different conditions. For example, average distance between particles (molecules) of air near of the Earth is $l \sim 10^{-6}cm$, in atmosphere on the height of $60km$ it is $l \sim 10^{-3}cm$ and in cosmos it is $l \sim 1cm$. If one consider that lower bound of length L on which processes are studied in these mediums is equal to $10^{-1}cm$, $10^{2}cm$ and $10^{5}cm$, then for all those cases we have $l/L \sim 10^{-5}$. Therefore cosmos media we can regard as physical continuum in the same meaning as we assume for the air near the Earth.

That is the continuum hypothesis implies that a very small volume will contain a large number of molecules. For example, $V = 1cm^3$ of air contains $N = 2.687 \times 10^{19}$ molecules under normal conditions (Avogadro hypothesis). Thus, in a cube $0.001cm$ on a side, there are 2.687×10^{10} molecules–which is a large number. We are not interested in the properties of each molecule at some point x but rather in the average over a large number of molecules in the neighborhood of the point x. Mathematically, the association of averaged values of properties at a point x also gives rise to a continuum of points and numbers. In summary, the continuum hypothesis implies the postulate: "Matter is continuously distributed throughout the region under consideration with a large number of molecules even in macroscopically small volumes."

Schematically continuum media is separated on gas, liquid and solid. It is a conditional separation. In this consideration we use some statistical aspects of molecular motions. For example, in gases, the molecules are far apart having an average separation between the molecules of the order of $3.5 \times 10^{-7}cm$. The cohesive forces between the molecules are weak. The molecules randomly collide and exchange their momentum, heat, and other properties and thus give rise to viscosity, thermal conductivity, etc. These effects, though molecular in origin, are considered the physical properties of the continuum itself. In liquids, the separation between the molecules is much smaller and the cohesive forces between a molecule and its neighbors are quite strong. Again, the averaged molecular properties resulting from these cohesive forces are taken as the properties of the medium. While air and water are treated through the same continuum hypothesis, the effects of their motions are different due to the differences in their molecular properties, e.g. viscosity, thermal conductivity, etc.

2. Continuum mechanics describes the global behavior of gases, liquids or solids under the influence of external disturbances.

The concept of a physical continuum makes the powerful methods of calculus available for the study of nonuniform distributions of physical variables and provides an easily visualized physical model that closely approximates observations of matter in the large.

The problems of continuum mechanics are multiform. Continuum mechanics is a foundation for the understanding of many aspects of the applied sciences and engineering. It is a subject of enormous interest in numerous fields like biology, biomedicine, geophysics, meteorology, physical chemistry, plasma physics and almost all branches of engineering.

Continuum mechanics is separated on experimental–physical and theoretical parts. We will consider only theoretical continuum mechanics.

A method of theoretical continuum mechanics consists of constructing a mathematical model of behavior of continuous media. A mathematical model is a system of relationships (equations and inequalities) between values, which characterize different properties of media. Usually they are differential (finite) equations. To these equations the initial and boundary data are supplemented. Mathematical model has to have a property of correctness. It means that the solutions of involving in it equations have to exist, to be unique and stable. For some models there is no strict prove of a correctness, in this case one has to use the criteria of practice. Physical experiments serve as tests for the validity of a theoretical model.

After constructing a mathematical model we produce purely mathematical methods for studying it. For this goal we use analytical and numerical methods. By virtue of difficulty of solving of continuum mechanics equations the methods of simplifications of them have a wide spreading.

3. For the better understanding of physical foundations of construction of mathematical model of continuum mechanics at first we consider a molecular ("microscopic") description.

Let some volume V of continuum media contains N molecules μ_i $(i = 1, 2, ..., N)$ with coordinates of position x and mass m_i. A motion $x_i(t)$ of molecule μ_i obeys to the second Newton's law

$$m_i \frac{d^2 x_i}{dt} = f_i,\ x_i(t_0) = x_{i0},\ \frac{dx_i}{dt}(t_0) = v_{i0},\ (i = 1, 2, ..., N),$$

where f_i is a force, which acts on the molecule μ_i. A solution of these equations defines position and velocity of molecule μ_i at any moment of time t. If we have been able to solve these equations we could answer on any question about behavior of media in the volume V. However this way is impossible, because the number N is very large and we don't exactly know the forces f_i. Therefore in continuum mechanics we adopt a macroscopic viewpoint: we ignore all the fine details of the molecular or atomic structure and, for the purpose of study, we replace the microscopic medium with a hypothetical continuum in which the basic values are replaced by average values.

To distinguish the continuum or macroscopic model from a microscopic one a concept of the mean free path plays a fundamental role. This concept can be defined as an average distance a molecule travels between successive collissions with other molecules. The ratio of the mean free path λ to the characteristic length L of the physical boundaries of interest, called the Knudsen number $K_n = \lambda/L$, may be used to determine the separating line between the macroscopic and microscopic models. Based on the Knudsen number the motion regimes are grouped as:

(a) continuum $(Kn < 0.1)$;

(b) rarefied gas $(0.1 < Kn < 5)$;

(c) free molecular flow $(Kn > 5)$.

Regimes (a) and (b) belong to macroscopic models. All these regimes are encoutered in real life.

Two macroscopic theories are the most widespread: molecular-kinetic theory and phenomenological one.

In molecular-kinetic theory all average values are described with the help of theoretical-probability approach. The mathematical model has the form of the Boltzmann equation. We study the phenomenological theory.

4. A basis of the phenomenological theory forms a representation that each point of a body V has density, velocity and other mechanical values. These values are defined as limits of some average values, which are formed by the following way.

Let molecules μ_i $(i = 1, 2, ..., N)$ from volume V have mass m_i, velocity v_i and internal energy U_i. With the help of these values one calculates macrocharacteristics of the volume V: $M = \sum_{i=1}^{N} m_i$ is the mass, $K = \sum_{i=1}^{N} m_i v_i$ is the impuls, $E = \sum_{i=1}^{N}(m_i \frac{v_i^2}{2} + U_i)$ is the full energy. Then $\rho_* = M/V$ is the average density, $v = K/M$ is the average velocity, $U_* = U/V$ is the average energy. Here $U = \sum_{i=1}^{N}(m_i \frac{(v_i - v_*)^2}{2} + U_i)$. The macroscopic characteristics of the volume V can be expressed by means of the average values:

$$M = \rho_* V,\ K = \rho_* v_* V,\ E = (\frac{1}{2}\rho_* v_*^2 + U_*)V.$$

The hypothesis of the physical continuum allows to give to the point x the "limit" values of average ones, for example, $\rho = \lim\ \rho_*,\ v = \lim\ v_*$ where the volume V vanishies that $x \in V$. A mathematical model has the form of conservation laws of changing macroscopic characteristics with respect to time.

We will construct the phenomenological theory of continuum mechanics as the theory of some mathematical structure. This mathematical structure is based on some system of axioms.

2.4 Basic definitions and axioms

Definition 2.1. *Continuous medium is a part of changing with time physical space. It means that continuous media is a part of Euclidean three dimensional space R^3, and the time is independent from events. We consider non-relativistic Newtonian approach, i.e. the time is absolute.*

Axiom 1 (axiom of space-time). Continuous medium is a subset of three dimensional Euclidean affine space. The time is absolute.

An Euclidean-affine space is a curvature-free space in which a set of rectangular Cartesian coordinates can always be introduced on a global scale. It is a linear three dimensional space over field of real numbers R. In this space the origin point O is fixed. Open connected sets $\Omega \in R^3$ is regarded as a positions (configurations) of continuous medium.

Set $\Omega \in R^3$ is called a material domain (or media) if an additive positive function of sets $M(\omega)$ is defined on it, which is named by mass. It is supposed that for any (nonempty) volume $\omega \subset \Omega$ its mass is $M(\omega) > 0$. The additiveness of mass means that if $\omega_1 \subset \Omega, \omega_2 \subset \Omega$ and $\omega_1 \cap \omega_2 = 0$, then $M(\omega_1 \cup \omega_2) = M(\omega_1) + M(\omega_2)$.

Besides mass we determine another additive function of set, which we call internal energy and we denote it by E_i.

Definition 2.2. *Media Ω is called material continuum, if functions M and E_i are differentiable on Ω and their densities (volume densities) are bounded.*

Volume density of mass is denoted by ρ and it is called a density of media (or simply density). Volume density of energy is denoted by ρU and U is called specific internal energy (internal energy per unit mass). The following formula

$$M(\omega) = \int_{\omega} \rho d\omega, E_i(\omega) = \int_{\omega} \rho U d\omega$$

determine the connection between additive function of set ω and its volume density.

Axiom 2 (axiom of material continuum). Continuous medium is material continuum.

The transition of continuous medium from position Ω_1 into position Ω_2 is called its motion. A motion of continuous medium depends of time t, which is changed in some interval $\tau = (a, b) \in R$. Position of medium at the moment of time t is denoted by Ω_t. We fix a moment of time $t_0 \in \tau$ and for all $t \in \tau$ we consider one-parametrical family of movements γ_t from position Ω_{t_0} onto Ω_t. It means that we have mapping $\gamma : \Omega_{t_0} \times \tau \longrightarrow \Omega_t$. We denote $\gamma_t(\xi) = \gamma(\xi, t)$ for all $\xi \in \Omega_{t_0}$ and $\gamma_\xi(t) = \gamma(\xi, t)$ for all $t \in \tau$ or we will write $\gamma_t : \Omega_{t_0} \longrightarrow \Omega_t, \gamma_\xi : \tau \longrightarrow \Omega_t$. A Set $\{x \in R^3 : x = \gamma_\xi(t), t \in \tau\}$ is called a trajectory of point $\xi \in \Omega_{t_0}$.

Axiom 3 (axiom of movement). For all $t \in \tau$ there exist the movement γ_t of continuous medium from position Ω_{t_0} onto position Ω_t and mapping $\gamma_t : \Omega_{t_0} \longrightarrow \Omega_t$ is homeomorphism; for all point $\xi \in \Omega_{t_0}$ the mapping $\gamma_\xi : \tau \longrightarrow \Omega_t$ is continuous and piecewise continuously differentiable function on τ.

Remark.(definition of Homeomorphism)
Homeomorphism From Wikipedia, the free encyclopedia. (Redirected from Homeomorphic)
This word should not be confused with homomorphism.

In topology, two geometrical objects (or "spaces") are called homeomorphic if, roughly speaking, the first can be deformed into the second by stretching and bending; cutting is also allowed, but only if the two parts are later glued back together along exactly the same cut. For example, a square and a circle are homeomorphic. A hollow sphere containing a smaller solid ball is homeomorphic to a hollow cube with a solid cube outside of it. If two objects are homeomorphic, one can find a continuous function which maps points from the first object to corresponding points of the second object, and vice versa. Such a function is called a homeomorphism; intuitively, it maps points in the first object that are "close together" to points in the second object that are close together, and points in the first object that are not close together to points in the second object that are not close together. Topology is the study of those properties of objects that do not change when homeomorphisms are applied.

For a formal definition, suppose X and Y are topological spaces, and f is a function from X to Y. Then f is a homeomorphism iff all the following hold:

1. f is a bijection, 2. f is continuous, 3. the inverse function f^{-1} is continuous.

If there exists a homeomorphism $f : X \to Y$, then Y is said to be homeomorphic to X (or to be a homeomorph of X). In this case, Y is also homeomorphic to X, since f^{-1} is a homeomorphism, and we say that X and Y belong to the same homeomorphism class.

For example, the unit circle S^1 and the unit square in R^2 are homeomorphic. The open interval $(-1, 1)$ is homeomorphic to the real numbers R. The product space $S^1 \times S^1$ and the two-dimensional torus are homeomorphic.

The third requirement, that f^{-1} be continuous, is essential. Consider for instance the function $f : [0, 2\pi) \to S^1$ defined by $f(\varphi) = (cos(\varphi), sin(\varphi))$. This function is bijective and continuous, but not a homeomorphism.

If two spaces are homeomorphic then they have exactly the same topological properties. For example, if one of them is compact, then the other is as well; if one of them is connected, then the other is as well; if one of them is Hausdorff, then the other is as well; their homology groups will coincide. Note however that this does not extend to properties defined via a metric; there are metric spaces which are homeomorphic even though one of them is complete and the other is not.

Homeomorphisms are the isomorphisms in the category of all topological spaces. As such, the composition of two homeomorphisms is again a homeomorphism, and the set of all homeomorphisms $X \to X$ forms a group.

Informal discussion

The intuitive criterion of stretching, bending, cutting and gluing back together takes a certain amount of practice to apply correctly–it may not be obvious from the description above that deforming a line segment to a point is impermissible, for instance. It is thus important to realize that it is the formal definition given above that counts.

This characterization of a homeomorphism often leads to confusion with the concept of homotopy, which is actually defined as a continuous deformation, but from one function to another, rather than one space to another. In the case of a homeomorphism, envisioning a continuous deformation is a mental tool for keeping track of which points on space X correspond to which points on Y – one just follows them as X deforms. In the case of homotopy, the continuous deformation from one map to the other is of the essence, and it is also less restrictive, since none of the maps involved need to be one-to-one or onto. Homotopy does lead to a relation on spaces: homotopy equivalence.

There is a name for the kind of deformation involved in visualizing a homeomorphism. It is (except when cutting and regluing are required) an isotopy between the identity map on X and the homeomorphism from X to Y.

This axiom allows to postulate a point of continuous medium. A material point (or particle) of continuous medium is called a point $x = \gamma(\xi, t) \in \Omega_t$, which is obtained as a result of movement of fixed point $\xi \in \Omega_{t_0}$. Every particle describes in R^3 the trajectory of this point.

Definition 2.3. *A set of points which consists of the same particles for all $t \in \tau$ is called a material volume ω_t. By virtue of axiom 3 for all $\xi \in \Omega_{t_0}$ and all (except maybe finite number) values $t \in \tau$ there exists a derivative*

$$\frac{\partial}{\partial t}\gamma(\xi, t).$$

Definition 2.4. *A derivative $\frac{\partial}{\partial t}\gamma(\xi, t)$ is called a velocity of point $\xi \in \Omega_{t_0}$ and it is denoted by*

$$v = \frac{\partial}{\partial t}\gamma(\xi, t).$$

Let F be either a scalar or vector or tensor function of position x and time t, representing some physical property of the movement. There are two ways of description of field F given on the moving continuous medium. The first one is called by Eulerian way. It consists of in the giving of value of field F on the position Ω_t as a function of $x \in R^3$ and time $t \in \tau$, i.e. it has a value $F(x, t)$.

A second way is called Lagrangian one. In this case given field is considered as a function of each particle $\xi \in \Omega_{t_0}$ at the moment of time $t \in \tau$. Let it be $F(\xi, t)$. The functions $F(x, t)$ and $F(\xi, t)$ are connected by identity

$$F(\xi, t) = F(\gamma(\xi, t), t). \tag{2.6}$$

There are two possible time derivatives: $\frac{\partial F(\xi,t)}{\partial t}$ and $\frac{\partial F(x,t)}{\partial t}$

A value $\dfrac{\partial F(x, t)}{\partial t}$ is the rate of change of field F measured by an observer stationed at the fixed point $x \in \Omega_t$ and it is a local time variation of F.

On the other hand, $\dfrac{\partial F(\xi, t)}{\partial t}$ is a rate of change of $F(\xi, t)$ measured by an observer moving with the particle. The differentiation (2.6) with respect to t gives

$$\frac{\partial F(\xi, t)}{\partial t} = \frac{\partial F(x, t)}{\partial t} + \frac{\partial F(x, t)}{\partial x} < v > .$$

A value $\dfrac{\partial F(x, t)}{\partial t} + \dfrac{\partial F(x, t)}{\partial x} < v >$ is called a total derivative (material or substational derivative, or the derivative following the motion) and it is denoted by symbol $\dfrac{dF(x, t)}{dx}$. So, for any smooth field $F = F(x, t)$ its derivative is given by formula

$$\frac{dF(x, t)}{dt} \equiv \frac{\partial F(x, t)}{\partial t} + \frac{\partial F(x, t)}{\partial x} < v > .$$

In particular, if $F = x = \gamma(\xi, t)$ and we obtain the formula for the definition of velocity

$$\frac{dx}{dt} = \frac{\partial \gamma(\xi, t)}{\partial t} = v.$$

Definition 2.5. *Coordinates (ξ, t) are called material or Lagrangian coordinates and (x, t) are called spatial or Eulerian coordiates.*

A difference between these descriptions is essential. For example, if field of a vector of velocity is known in Lagrange description, i.e. we have a vector–function $v(\xi, t)$, then we can find a trajectories of particles (and then it means we can find a movement of continuous medium)

$$x = \xi + \int_{t_0}^{t} v(\xi, t')dt'$$

And if we know a field v in Eulerian description (it means that we have $v = v(x, t)$, then the same problem of determination of trajectories gives us the Cauchy problem for the system of ordinary differential equations

$$\frac{dx}{dt} = v(x, t), x(\xi, t_0) = \xi.$$

In spite of the simplicity of first problem a Lagrangian description is convenient not always. In particular, the main differential equations of continuum mechanics have simpler form in Eulerian description.

In Eulerian description a map $\gamma : \Omega_{t_0} \times \tau \longrightarrow \Omega_t$ is obtained as a solution of Cauchy problem (2). If vector–function $v(x, t)$ is continuously differentiable then for such solution there exists Jacobian $J = \det(\frac{\partial x}{\partial \xi})$. For the Jacobian we have a kinematic formula, known as Euler's formula:

$$\frac{dJ}{dt} = J div(v).$$

In addition to the main numerical characteristics of the material media (mass, energy) there are the following additive functions of the set $\omega \subset \Omega$:

(i) linear momentum:

$$K(\omega) = \int_{\omega} \rho v d\omega$$

(ii) angular momentum:

$$H(\omega) = \int_{\omega} \rho(x \times v) d\omega$$

(iii) kinetic energy:

$$E_k(\omega) = \frac{1}{2} \int_{\omega} \rho v^2 d\omega$$

(iv) total energy: $E(\omega) = E_k(\omega) + E_i(\omega)$.

The changes of these magnitudes under movement are the result of force and energetic actions on the volume ω. These actions are realized with the help of new magnitudes: resultant force $F(\omega)$, resultant moment $\mathrm{G}(\omega)$ and power $N(\omega)$.

If we take these magnitudes for a fixed moving material volume $\omega_t \subset \Omega_t$ then they will be only functions of time t. A following axiom determine the relations between them.

Axiom 4 (balance, poise). For arbitrary moving material volume $\omega_t \subset \Omega_t$ and in any time $t \in \tau$ we have

$$\frac{d}{dt} M(\omega_t) = 0, \ \frac{d}{dt} K(\omega_t) = F(\omega_t), \ \frac{d}{dt} H(\omega_t) = G(\omega_t), \ \frac{d}{dt} E(\omega_t) = N(\omega_t).$$

Sometimes this axiom is called a "hardening" principle, because these equalities are fulfilled for movement of rigid body.

2.5 Material volume and surface

A very useful in continuum mechanics is notation of a "closed system" or a "material volume". A material volume is an arbitrary collection of particles enclosed by a surface also formed of particles. All points of the material volume, including the points of its boundary, move with the local continuum velocity. A material volume moves with the flow and deforms in shape as the flow progresses, with the stipulation that no mass ever fluxes in or out of the volume, viz., both the volume and its boundary are always composed of the same particles. We shall denote a material volume by ω_t and its surface by $\partial\omega_t$. Note that the use of material volume in continuum mechanics is in the form of a thought experiment in which one isolates a parcel of material volume out of the flow field and gives it a hypothetical surface. This helps formulate the conservation laws for continuum mechanics in a straightforward manner.

To obtain the necessary conditions for a bounding surface of a material volume, first note that the material volume is enclosed by a surface formed of particles. This surface is called the material bounding surface. Since there cannot be a transfer of material across a material bounding surface, the particles forming the inside surface of the material volume can never become the particles forming the outside surface of the same material volume. Consequently, it qualifies as a material surface since it is always composed of the same material points.

Let $f(x,t) = 0$ be the equation of a bounding surface $\partial\omega_t$ enclosing the material volume $\partial\omega$. Let n be the unit external normal to $\partial\omega_t$, then:

$$n = \frac{\nabla f}{|\nabla f|}$$

To obtain the necessary conditions for $\partial\omega_t$ to be the bounding surface of a material volume, we follow Kelvin who states that "...to express the fact that every particle of fluid remains on the same side of the surface, or that there is no flux across it, we must find the normal motion of the surface ..." Let v_n be the velocity of any point normal to $\partial\omega_t$; then the equation of the surface at time $t + \delta t$ is

$$f(x + \delta x, t + \delta t) = 0$$

where

$$\delta x = n\, v_n \delta t$$

Using Taylor's expansion, we have:

$$\frac{\partial f}{\partial t} + v_n(n\ \nabla f) = 0 \tag{2.7}$$

If the bounding surface is a material surface, then the velocity at any point normal to the surface must be equal to the normal continuum velocity, i.e., $v_n = v \cdot n$. Using Equation (2.7), we get:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + v_n(n\ \nabla f) = 0$$

Thus, for a bounding surface to be a material surface, the surface $f(x,t) = 0$ must satisfy this equation.

Remark (another proof).

$$f(x,t) = 0,\ \ 0 = f(x + \delta x, t + \delta t) = f(x,t) + \frac{\partial f}{\partial t}\delta t + \nabla f\ \delta x + ...$$

$$= \delta t(\frac{\partial f}{\partial t} + \nabla f\ \frac{\delta x}{\delta t}) + ... = \delta t(\frac{\partial f}{\partial t} + \nabla f\ v) + ...$$

$$f(\gamma(\xi,t),t) = 0,\ \ \frac{df}{dt} = f_t + v\nabla f = f_t + vn|\nabla f|.$$

2.6 Relation between elemental volumes

We shall denote any arbitrary closed volume in the Ω_{t_0} by ω_{t_0} so that it is a volume "frozen" in time with surface $\partial\omega_{t_0}$. Let $F(x,t)$ be a physical property per unit volume and let it be continuously differentiable in Ω_t. The amount of this property in a material volume ω_t at the time $t > t_0$ is

$$\int_{\omega_t} F(x,t)\, d\omega = \int_{\omega_{t_0}} F^0(\xi,t) J\, d\omega_0$$

The rate of change of value $\int_{\omega_t} F(x,t)\, d\omega$ as the volume moves with the movement is (we assume that movement $\gamma : \Omega_{t_0} \to \Omega_t$ is continuously differentiable one):

$$\frac{d}{dt}\int_{\omega_t} F(x,t)\, d\omega = \frac{d}{dt}(\int_{\omega_{t_0}} F^0(\xi,t) J\, d\omega_0) = \int_{\omega_{t_0}} \frac{\partial}{\partial t}(F^0(\xi,t)J)\, d\omega_0 =$$

$$= \int_{\omega_{t_0}} (J\frac{\partial F^0(\xi,t)}{\partial t} + F^0(\xi,t)\frac{\partial J}{\partial t})d\omega_0$$

By virtue of Euler's formula, we get

$$\frac{d}{dt}\int_{\omega_t} F(x,t)\, d\omega = \int_{\omega_{t_0}} (J\frac{\partial F^0(\xi,t)}{\partial t} + F^0(\xi,t)J(div\ v))d\omega_0$$

Reverting to the material volume ω_t, we obtain

$$\frac{d}{dt}\int_{\omega_t} F(x,t)\, d\omega = \int_{\omega_t} (\frac{dF}{dt} + F(div\ v))d\omega$$

Using the definition of $\frac{d}{dt}$, we also have:

$$\frac{d}{dt}\int_{\omega_t} F(x,t)\, d\omega = \int_{\omega_t} (\frac{\partial F}{\partial t} + v\nabla F + F\ div\ v)d\omega = \int_{\omega_t} (\frac{\partial F}{\partial t} + div\ (Fv))d\omega$$

We use Gauss-Ostrogradskii divergence theorem

$$\frac{d}{dt}\int_{\omega_t} F(x,t)\, d\omega = \int_{\omega_t} \frac{\partial F}{\partial t}d\omega + \int_{\partial\omega_t} Fvn\, d\sigma,$$

where n is the outward drawn unit normal to $\partial\omega_t$.

We received the Reynolds' transport theorem, which states that the rate of change of the total property $F(x,t)$ contained in material volume is equal to the volume integral of the instantaneous changes of $F(x,t)$ while keeping ω_t fixed at a given time, plus the surface integral of the rate of spreading of $F(x,t)$ to the adjoining region due to the surface velocity v.

2.7 Kinetics of movement

For further we have to specify the right sides in the formulae of Axiom 4. At first we define a concept of resultant forces. We will consider two types of forces which act on a material volume ω:

(a) body forces,
(b) surface forces.

The body forces are forces of an extensive character acting on the bulk portions of the continuous medium and arise due to some external cause. Examples of the external causes are (a) the force of gravity, (b) forces of electric and magnetic origin acting on a continuous medium carrying charged particles, etc.. The body force is proportional to the volume of a continuous medium and therefore it is expressed as a force per unit volume.

Definition 2.6. *An additive vector-function F_e having a density (body force per unit mass) is called an external body force. If we denote the body force per unit mass by the symbol $f(x,t)$, so that the body force per unit volume will be ρf. Therefore, the external body force acting on the volume ω is given by formula*

$$F_e(\omega) = \int_\omega \rho f \, d\omega.$$

And the moment of external body force acting on any material volume ω is defined by the formula

$$G_e(\omega) = \int_\omega \rho (x \times f) \, d\omega.$$

The surface forces are forces of an intensive or local nature. The surface forces arise due to mechanical interaction between contiguous portions of a continuous medium. To explain the phenomena from a continuum point of view, we consider two adjacent portions of continuous medium separated by an imaginary surface drawn in the medium. At the separating surface there exists a direct mechanical contact between the medium particles on the two sides of the surface, thus, giving rise to forces of action and reaction. If the continuous medium on one side is imagined to have been replaced by the force system which it has produced, then at each point of the imaginary surface there will be a force vector.

Internal surface force acts on a volume ω only through its surface $\partial\omega$. In order to define it we consider cross section Σ of Ω by some plane dividing Ω on two parts Ω_1 and Ω_2.

Definition 2.7. *Additive vector-function F_i of sets $\sigma \subset \Sigma$ is called an internal surface force acting through a cross section Σ from the side Ω_2 on the Ω_1.*

Axiom 5 (of internal surface forces). An internal surface force is defined for any cross section Σ of Ω and it has a density (surface) on Σ.

Remark. This axiom is named a Cauchy stress principle and it asserts the existence and differentiability of this force.

Let n be a local outward drawn unit normal vector of Σ directed in the side of Ω_2 (positive side of Ω_1). The density of internal surface force we denote by p_n.

Definition 2.8. *A vector p_n is called stress vector of surface forces acting on Ω_1 through the area with the normal n.*

And for the $\sigma \subset \Sigma$ the force, which acts on part Ω_1 from the side of part Ω_2 through an area σ is equal to

$$F_i(\sigma) = \int_\sigma p_n \, d\sigma.$$

Definition 2.9. *The value*

$$F_i(\omega) = \int_{\partial\omega} p_n \, d\sigma.$$

is called an internal surface force acting on volume $\omega \subset \Omega$ from the side of Ω. Here n is positive outward drawn unit normal vector to the surface volume ω.

The value

$$G_i(\omega) = \int_{\partial\omega} (x \times p_n) \, d\sigma$$

is called a moment of internal surface force, acting on the volume ω.

Axiom 6 (of forces and moments). The (main) resultant force and resultant moment, acting on any material volume $\omega \subset \Omega$ are given by formulae:

$$F(\omega) = F_i(\omega) + F_e(\omega) = \int_{\partial\omega} p_n \, d\sigma + \int_\omega \rho f \, d\omega,$$

$$G(\omega) = G_i(\omega) + G_e(\omega) = \int_{\partial\omega} (x \times p_n) \, d\sigma + \int_\omega \rho(x \times f) \, d\omega.$$

Remark. This axiom asserts that only such moments and forces act on the volume $\omega \subset \Omega$ (we have no other forces and moments).

We have to determine a power $N(\omega)$.

Definition 2.10. *The values*

$$N_i(\omega) = \int_{\partial\omega} v p_n \, d\sigma, \; N_e(\omega) = \int_\omega \rho v f \, d\omega$$

are called powers developing by internal surface forces and external body forces.

By the similar way as for an internal surface force we define a heat output.

Definition 2.11. *An additive scalar function Q of sets $\sigma \subset \Sigma$ is called a heat output through area Σ from the part Ω_2 into Ω_1.*

Axiom 7 (of heat output). A heat output is defined for any cross section Σ of Ω and it has a density (surface density) on Σ. The surface density of heat output is denoted by q_n and the value

$$Q(\sigma) = \int_\sigma q_n \, d\sigma$$

gives the heat output from the Ω_2 into Ω_1 through the area $\sigma \subset \Sigma$.

Definition 2.12. *The value*

$$Q(\omega) = \int_{\partial\omega} q_n \, d\sigma$$

is called a heat output into volume $\omega \subset \Omega$ from the domain $\Omega \backslash \bar{\omega}$. Here n is a positive outward drawn unit normal vector to the volume surface $\partial\omega$.

Axiom 8 (of energy transfer). The power $N(\omega)$ getting by any volume $\omega \subset \Omega$ is equal to

$$N(\omega) = N_i(\omega) + N_e(\omega) + Q(\omega) = \int_{\partial\omega} v p_n \, d\sigma + \int_\omega \rho v f \, d\omega + \int_{\partial\omega} q_n \, d\sigma.$$

Remark. This axiom fixes an assumption that we have no any mechanisms of the receiving of energy by a volume $\omega \subset \Omega$.

Resume.

We can summarize the previous axioms and definitions as the following classical **mathematical model** of moving continuous media.

Mathematical model 1 (integral conservation laws).

In the moving continuous media for any moving volume $\omega_t \subset \Omega_t$ and any moment of time $t \in \tau$ the following equalities are hold:

$$\frac{d}{dt}(\int_{\omega_t} \rho \, d\omega) = 0,$$

$$\frac{d}{dt}(\int_{\omega_t} \rho v \, d\omega) = \int_{\partial\omega_t} p_n \, d\sigma + \int_{\omega_t} \rho f \, d\omega, \tag{2.8}$$

$$\frac{d}{dt}\Big(\int_{\omega_t} \rho(x \times v)\, d\omega\Big) = \int_{\partial\omega_t} (x \times p_n)\, d\sigma + \int_{\omega_t} \rho(x \times f)\, d\omega,$$

$$\frac{d}{dt}\Big(\int_{\omega_t} \rho(\frac{v^2}{2} + U)\, d\omega\Big) = \int_{\partial\omega_t} v p_n\, d\sigma + \int_{\omega_t} \rho v f\, d\omega + \int_{\partial\omega_t} q_n\, d\sigma.$$

Every of these equalities is called by "conservation law" of the corresponding mechanical value: conservation law of mass, conservation law of linear momentum, conservation law of angular momentum, conservation law of energy.

Definition 2.13. *(main definition of continuum mechanics). A moving continuous media is an object satisfying the axioms A_1-A_8. A mathematical model consists of four conservation laws.*

2.8 Continuous motion

The main functions (magnitudes) related to a moving continuous medium: a density ρ, a specific internal energy U, a velocity v, a stress p_n, with a normal vector n, a density of heat output q_n, and a density of external body forces f, we will consider (further) in Eulerian description. It means that these functions are functions of (x, t) in a domain $W \subset R^4(x, t)$. The magnitudes p_n and q_n depend on unit vector $n \in R^3$ (point of unit sphere S_1) and therefore they are given on the product $W \times S_1$.

At first we study a class of movements of continuous media where the main magnitudes are sufficiently smooth functions.

Definition. A movement of continuous media is called continuous in a domain W if the functions ρ, U, v, p_n, q_n are continuously differentiable functions in W, the functions p_n, q_n are continuous in $W \times S_1$, and the function f is continuous in W.

Let us consider the derivative

$$\frac{d}{dt}\int_{\omega_t} \rho F\, d\omega$$

where the vector-function $F(x, t)$ is continuously differentiable and a movement of continuous media is continuous. In order to calculate this value we do a transition to the Lagrange's system of coordinates $x = \gamma(x, t)$. The integral has the form

$$\int_{\omega_t} \rho(x,t)F(x,t)\, d\omega = \int_{\omega_{t_0}} \rho(\xi,t)F(\xi,t)J(\xi,t)\, d\omega_0$$

By virtue of the theorem of Real analysis we can replace an integral and derivative and in the strength of Euler's formula, we obtain

$$\frac{\partial}{\partial t}(\rho(\xi,t)F(\xi,t)J(\xi,t)) = \frac{d}{dt}(\rho(x,t)F(x,t)J(x,t)) = J[\frac{d}{dt}(\rho(x,t)F(x,t)) + \rho(x,t)F(x,t)\, div_x(v)]$$

Hence,

$$\frac{d}{dt}\int_{\omega_t} \rho(x,t)F(x,t)\, d\omega = \int_{\omega_t} \big(\frac{d}{dt}(\rho(x,t)F(x,t)) + \rho(x,t)F(x,t)\, div_x(v)]\big)\, d\omega$$

For example, if $F = 1$, then

$$\frac{d}{dt}\int_{\omega_t} \rho(x,t)\, d\omega = \int_{\omega_t} \big(\frac{d}{dt}\rho(x,t) + \rho(x,t)\, div_x(v)\big)\, d\omega = 0.$$

Because ω_t is an arbitrary volume then by virtue of the lemma

$$\frac{d}{dt}\rho + \rho\, div_x(v) = 0.$$

This equation is called a continuity equation. It is equivalent to the mass conservation law on class of continuous motions.

The continuity equation allows simplifying

$$\frac{d}{dt}\int_{\omega_t} \rho F\, d\omega = \int_{\omega_t} \rho \frac{dF}{dt}\, d\omega$$

2.9 Conservation of linear momentum.

By virtue of formula (2.8) the equation of linear momentum takes the form

$$\int_{\partial\omega_t} p_n\, d\sigma = \int_{\omega_t} \varphi\, d\omega$$

where $\varphi = \rho(\frac{dv}{dt} - f)$ is continuous in Ω_t.

Theorem. There exists a tensor field of second order P in W that for all $(x,t) \in W$

$$p_n(x,t) = P(x,t) < n >$$

Proof. At a fixed point $(x,t) \in W$, (in fixed time t point $x \in \Omega_t$) the vector p_n is a continuous function $p : n \to p_n$ (we write $p_n = p(n)$). We have to prove that this function is linear. Let $\{e^i\}$ be an orthonormal basis in R^3. Then $n = n_\alpha e^\alpha$. For the proof of the theorem it is enough to prove that

$$p_n = n_\alpha p_{e^\alpha},$$

because we can determine the tensor

$$P < n >= n_\alpha p_{e^\alpha}$$

At first we prove that $p_{-n} = -p_n$ or

$$p(-n) = -p(n).$$

Let $B_\varepsilon(x) \subset \Omega_t$ be ball at the point $x \in \Omega_t$ with radius ε and Σ be plane with normal vector n and passing through the point x. The ball $B_\varepsilon(x)$ is divided by this plane on two half-balls ω_1 and $\omega_2 : B_\varepsilon(x) = \omega_1 \cup \omega_2$ with cross-section $\sigma_\varepsilon = \omega_1 \cap \omega_2$. We have the equalities

$$\int_{\partial\omega_1} p_n\, d\sigma = \int_{\omega_1} \varphi\, d\omega,\ \int_{\partial\omega_2} p_n\, d\sigma = \int_{\omega_2} \varphi\, d\omega,\ \int_{\partial B_\varepsilon} p_n\, d\sigma = \int_{B_\varepsilon} \varphi\, d\omega.$$

Because

$$\int_{B_\varepsilon} \varphi\, d\omega = \int_{\omega_1} \varphi\, d\omega + \int_{\omega_2} \varphi\, d\omega,$$

then we get

$$\int_{\sigma_\varepsilon} (p_n + p_{-n})\, d\sigma = 0.$$

By virtue of continuity of the function p and because ε is an arbitrary small parameter, we obtain

$$p_n + p_{-n} = 0.$$

Let a fixed vector $n^0 = n^0_\alpha e^\alpha$ be with $n^0_\alpha \neq 0$. We consider the infinitesimal tetrahedron Δ_ε, whose vertex is at x with its three faces σ_{ε_i} being parallel to the coordinate planes:

$$h_1 = \frac{\varepsilon}{n^0_1},\ h_2 = \frac{\varepsilon}{n^0_2},\ h_3 = \frac{\varepsilon}{n^0_3}$$

Denoting σ_ε slanted face of this tetrahedron with outward drawn normal n^0, we have $\sigma_{\varepsilon_i} = |n^0_i|\sigma_\varepsilon$. We take height ε such that $\Delta_\varepsilon \subset \Omega_t$. Then

$$\partial\Delta_\varepsilon = \sigma_\varepsilon \cup \sigma_{\varepsilon_1} \cup \sigma_{\varepsilon_2} \cup \sigma_{\varepsilon_3}$$

and

$$\int_{\partial\Delta_\varepsilon} p_n\, d\sigma = \int_{\sigma_\varepsilon} p(n^0)\, d\sigma + \sum_{i=1}^{3} \int_{\sigma_{\varepsilon_i}} p_n\, d\sigma = \int_{\Delta_\varepsilon} \varphi\, d\omega.$$

Because on the face σ_{ε_i} we have $n = \pm e^i$ (we take $+$ if $n^0_i < 0$ and $-$if $n^0_i > 0$) we find

$$\int_{\sigma_\varepsilon} p(n^0)\, d\sigma + \sum_{i=1}^{3} \int_{\sigma_{\varepsilon_i}} p(\mp e^i)\, d\sigma = \int_{\Delta_\varepsilon} \varphi\, d\omega.$$

By virtue of continuity of the function $p(n)$ and the Mean value theorem we have

$$\sigma_\varepsilon\ p_{x_0}(n^0) + \sum_{i=1}^{3} \sigma_{\varepsilon_i}\ p_{x_i}(\mp e^i) = \varphi(z)\ \varepsilon\sigma_\varepsilon$$

where $x_0 \in \sigma_\varepsilon,\ x_i \in \sigma_{\varepsilon_i},\ z \in \Delta_\varepsilon$. Then

$$p_{x_0}(n^0) - \sum_{i=1}^{3} n^0_i\ p_{x_i}(e^i) = \varepsilon\varphi(z).$$

If $\varepsilon \to 0$, then $x_0 \to x, x_i \to x$. In the strength of continuity of the functions p and φ, we have

$$p_x(n^0) - \sum_{i=1}^{3} n^0_i\ p_x(e^i) = 0.$$

If one or two coordinates of the vector n^0 are equal to zero, then in the strength of continuity of $p(n)$ with respect to n we receive the proof of the theorem for all vectors n.

Definition. The tensor P is called a stress tensor.

Now let us consider the conservation law of linear momentum. Using Gauss-Ostrogradskii theorem, we have

$$\int_{\partial\omega_t} p_n\, d\sigma = \int_{\partial\omega_t} P<n>\ d\sigma = \int_{\omega_t} div(P)\, d\omega.$$

Hence, the equation of linear momentum is reduced to

$$\int_{\omega_t} (\rho\frac{dv}{dt} - div(P) - \rho f)\, d\omega = 0.$$

For the continuous motion the integrand function is continuous. Since this equation is valid for any volume ω_t, then we get the differential form of the conservation law of linear momentum

$$\rho\frac{dv}{dt} = div(P) + \rho f.$$

2.10 Conservation of angular momentum.

Let us consider the linear transformation $E : R^3 \to L(R^3)$, which is defined by the formula

$$E(a) = \begin{pmatrix} 0 & -a^3 & a^2 \\ a^3 & 0 & -a^1 \\ -a^2 & a^1 & 0 \end{pmatrix},$$

where $a = a_i e^i$ is a vector, $\{e^i\}$ is an orthonormal basis. We have

$$E(a) < b >= a \times b = -E(b) < a > .$$

If we use Gauss-Ostrogradskii theorem and $x \times P < n >= (E(x) \circ P) < n >$ we get

$$\int_{\partial\omega_t} x \times P < n > \; d\sigma = \int_{\partial\omega_t} (E(x) \circ P) < n > \; d\omega = \int_{\omega_t} div(P \circ E(x)) \, d\omega.$$

The angular momentum equation is reduced to

$$\int_{\omega_t} (\rho(x \times \frac{dv}{dt}) - div(E(x) \circ P) - \rho x \times f) \, d\omega = 0$$

or the differential form has the form

$$\rho(x \times \frac{dv}{dt}) = div(E(x) \circ P) + \rho x \times f.$$

If we substitute in this equation the value $\rho\frac{dv}{dt}$ being found from the linear momentum equation we obtain

$$div(E(x) \circ P) = x \times div(P). \tag{2.9}$$

Theorem. Equation (2.9) is fulfilled if and only if the second order tensor P is a symmetric tensor, i.e., $P^* = P$.

Proof. The following chain of equalities are fair with an arbitrary constant vector (test vector) a

$$a \; div(E(x) \circ P) = div((E(x) \circ P)^* < a >) = -div(P^* \circ E(x) < a >) =$$

$$= -tr(\frac{\partial}{\partial x}(P^* \circ E(x)) < a >) = -E(x) < a > div(P) - tr(P^* \circ \frac{\partial}{\partial x}(E(x) < a >)) =$$

$$= a(x \times div(P)) + tr(P^* \circ \frac{\partial}{\partial x}(E(a) < x >)) = a(x \times div(P)) + tr(P^* \circ E(a)).$$

Hence,

$$tr(P^* \circ E(a)) = 0$$

for arbitrary test-vector a. Because

$$tr(P^* \circ E(a)) = tr((P^* \circ E(a))^*) = tr(E(a)^* \circ P) = -tr(E(a) \circ P) = -tr(P \circ E(a))$$

we have

$$2tr(P^* \circ E(a)) = tr(P^* \circ E(a)) - tr(P \circ E(a)) = 0$$

or we can rewrite

$$tr((P^* - P) \circ E(a)) = 0.$$

For the tensor $(P^* - P)$ there exists a vector c such that

$$E(c) = P^* - P.$$

Therefore

$$tr(E(c) \circ E(a)) = -2ca = 0.$$

By virtue of arbitrariness of the vector a the vector c is equal to 0, it means that $E(c) = 0$ or

$$P^* - P = 0.$$

Conversely, let $P^* = P$, then $tr(P^* \circ E(a)) = 0$ and therefore (2.9).

This theorem means that for any continuous motion a conservation law of angular momentum is equivalent to symmetry of the stress tensor P.

2.11 Conservation law of energy.

For a continuous motion the conservation law of energy is reduced to the equation

$$\int_{\partial\omega_t} q_n \, d\sigma = \int_{\omega_t} \psi \, d\omega$$

where $\psi = \rho \frac{d}{dt}(\frac{1}{2}v^2 + U) - div(P < v >) - \rho v f$. This equation can be used for a proof of existence of a heat output rate vector (heat flux).

Theorem. For a continuous motion in W, there exists a vector field q in W that for all $(x,t) \in W$

$$q_n(x,t) = -q(x,t)\, n$$

Proof is the same as in the theorem of existence of the tensor P.

Exercise. Prove this theorem.

Definition. Vector q is called a heat output rate vector (or heat flux).

Introducing the heat output rate vector allows transforming a surface integral into the volume integral:

$$\int_{\partial\omega_t} q_n \, d\sigma = -\int_{\partial\omega_t} qn \, d\sigma = -\int_{\omega_t} div(q) \, d\omega.$$

The conservation law of energy becomes

$$\int_{\omega_t} (\psi + div(q)) \, d\omega = 0.$$

By virtue of continuity of motion and arbitrariness of a material volume ω_t, we obtain $\psi + div(q) = 0$ or

$$\rho \frac{d}{dt}(\frac{1}{2}v^2 + U) = div(P < v >) + \rho v f - div(q).$$

We simplify this equation by using the relationships

$$P : \frac{\partial v}{\partial x} = \frac{1}{2} P : (\frac{\partial v}{\partial x} + (\frac{\partial v}{\partial x})^*), \ \frac{d}{dt}(\frac{1}{2}v^2) = 2v\frac{d}{dt}v, \ div(P < v >) = v \, div(P) + P : D,$$

where D is a rate-of-strain tensor

$$2D = \frac{\partial v}{\partial x} + (\frac{\partial v}{\partial x})^*$$

Substituting $\rho\frac{dv}{dt}$ from the differential form of the linear momentum conservation law (equation of motion) we find

$$\rho\frac{dU}{dt} = P : D - div(q).$$

This equation is called an energy equation (or heat flux equation).

Resume.

For arbitrary continuous motion of continuous media described by model M_1, there exist continuously differentiable fields of symmetric stress tensor P and a vector of heat output rate with which the integral conservation laws are equivalent to the system of differential equations

$$\frac{d\rho}{dt} + \rho\, div(v) = 0,$$

$$\rho\frac{dv}{dt} = div(P) + \rho f,$$

$$\rho\frac{dU}{dt} = P : D - div(q).$$

This system of partial differential equations is called a model of continuous motion of continuum mechanics. If we assume that the body force is prescribed, then the model M_2 consists of five independent (scalar) equations involving fourteen unknown variables, namely, ρ, v, P, U, q. A model is called "closed" if a number of unknown variables is equal to the number of equations in the model. And so, we have a problem of closing the model M_2. This problem has to be solved on the basis of an additional information about continuous media.

2.12 Invariants of stress tensor

Let us consider eigenvalues λ and eigenvectors l of the linear transformation P :

$$(P - \lambda I) < l >= 0.$$

Because $l \neq 0$ then we receive characteristic equation (secular equation)

$$\det(P - \lambda I) = -\lambda^3 + J_1\lambda^2 - J_2\lambda + J_3,$$

where $J_k(P)$ are invariants of stress tensor:

$$J_1(P) = tr(P) = \sigma_1 + \sigma_2 + \sigma_3,$$

$$J_2(P) = \frac{1}{2}[(tr(P))^2 - tr(P^2)] = \sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_1\sigma_3 - \tau_{12}^2 - \tau_{23}^2 - \tau_{13}^2,$$

$$J_3(P) = \det t(P).$$

By virtue of symmetry of stress tensor ($P^* = P$) a characteristic equation always has three real roots: $\sigma_I, \sigma_{II}, \sigma_{III}$.

Definition. The roots of characteristic equation are called principal stresses. Directions of eigenvectors corresponding to these eigenvalues are called principal axes.

For the symmetric matrix we always can take orthonormal basis which consists of the eigenvectors. This basis $\{e_i\}$ is called principal basis of tensor P. In this basis the stress tensor is a diagonal tensor with principal stress values on the main diagonal:

$$P = \begin{pmatrix} \sigma_I & 0 & 0 \\ 0 & \sigma_{II} & 0 \\ 0 & 0 & \sigma_{III} \end{pmatrix}.$$

In the terms of the principal stresses, the stress invariants may be written:

$$J_1(P) = \sigma_I + \sigma_{II} + \sigma_{III},$$

$$J_2(P) = \sigma_I\sigma_{II} + \sigma_{II}\sigma_{III} + \sigma_I\sigma_{III},$$

$$J_3(P) = \sigma_I\sigma_{II}\sigma_{III}.$$

Example. The components of the stress tensor at x are given in MPa with respect to axes x_1, x_2, x_3 by the matrix

$$[P] = \begin{pmatrix} 57 & 0 & 24 \\ 0 & 50 & 0 \\ 24 & 0 & 43 \end{pmatrix}$$

Determine the principal stresses and the principal stress directions at x.

Solution. For the given stress tensor, secular equation takes the form of the determinant

$$\det \begin{pmatrix} 57-\lambda & 0 & 24 \\ 0 & 50-\lambda & 0 \\ 24 & 0 & 43-\lambda \end{pmatrix} = 0$$

which, upon cofactor expansion about the first row, results in the equation

$$(57-\lambda)(50-\lambda)(43-\lambda) - (24)^2(50-\lambda) = 0$$

or in its readily factored form

$$(50-\lambda)(\lambda-25)(\lambda-75) = 0$$

Hence, the principal stress values are $\sigma_I = 25,\ \sigma_{II} = 50,\ \sigma_{III} = 75$. Note that we confirm that the first stress invariant,

$$J_1 = 57 + 50 + 43 = 25 + 50 + 75 = 150$$

To determine the principal directions, we first consider $\sigma_I = 25$, for which Eq 3.6-3 provides three equations for the direction cosines of the principal direction $l^1 = (l_1^1, l_2^1, l_3^1)$ of $\sigma_I = 25$, namely,

$$32l_1^1 + 24l_3^1 = 0,\ 25l_2^1 = 0,\ 24l_1^1 + 18l_3^1 = 0.$$

Obviously, $l_2^1 = 0$ from the second of these equations, and from the other two, $l_3^1 = -4l_1^1/3$ so that, from the normalizing condition $l_\alpha^1 l_\alpha^1 = 1$, we see that $(l_1^1)^2 = 9/25$, which gives $l_1^1 = \pm 3/5$ and $l_3^1 = \mp 4/5$.

Next, for $\sigma_{II} = 50$, Eq 3.6-3 gives,

$$77l_1^2 + 24l_3^2 = 0,\ 24l_1^2 - 7l_3^2 = 0,$$

which are satisfied only by $l_1^2 = l_3^2 = 0$. Then from $l_\alpha^2, l_\alpha^2 = 1, l_2^2 = \pm 1$. Finally, for σ_{III}=75, Eq 3.6-3 gives

$$-18l_1^3 + 24l_3^3 = 0,\ -25l_2^3 = 0,\ 24l_1^3 - 32l_3^3 = 0.$$

Here from the second equation $l_2^3 = 0$ and from either of the other two equations $4l_3^3 = 3l_1^3$, so that from $l_\alpha^3 l_\alpha^3 = 1$ we have $l_1^3 = \pm 4/5$ and $l_3^3 = \pm 3/5$.

2.13 Special cases of the stress state.

2.13.1 Octahedral stress

The stress when $n = \pm\frac{1}{\sqrt{3}}(1,1,1)$ in the principal axes is called octahedral stress. For this unit normal we have

$$p_n = nP < n >= \frac{1}{3}(\sigma_I + \sigma_{II} + \sigma_{III}) = \frac{1}{3}J_1(P),$$

$$(p_\tau^{oct})^2 \equiv p_\tau^2 = (P < n >)^2 - (p_n)^2 = \frac{1}{3}[(\sigma_I^2 + \sigma_{II}^2 + \sigma_{III}^2) - (\sigma_I + \sigma_{II} + \sigma_{III})^2] =,$$

$$= \frac{1}{9}[(\sigma_I - \sigma_{II})^2 + (\sigma_{II} - \sigma_{III})^2 + (\sigma_{III} - \sigma_I)^2] = \frac{2}{9}[J_1^2(P) - 3J_2(P)].$$

Octahedral stress plays a prominent role in the plastic theory.

2.13.2 The uniaxial state of stress.

The uniaxial state of stress is that for which only one principal stress component is different than zero (for example, $\sigma_{II} = \sigma_{III} = 0$. Then in the principal axes:

$$[P] = \begin{pmatrix} \sigma_I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; J_1(P) = \sigma_I, \; J_2(P) = J_3(P) = 0.$$

2.13.3 Simple shear (pure shear).

The state of simple shear is called such stress state when $\sigma_I + \sigma_{II} = 0, \; \sigma_{III} = 0$. Then normal stress on the plane with unit normal vector $n = \frac{1}{\sqrt{2}}(1, \pm 1, 0)$ in the principal axes is equal to zero ($p_n = 0$). On these planes we have a maximum of shear stress called octahedral stress. For this unit normal we have the principal axes:

$$p_\tau = \frac{\sigma_I}{\sqrt{2}}(1, \pm 1, 0).$$

2.13.4 Spherical stress state.

The stress state with $\sigma_I = \sigma_{II} = \sigma_{III}$ is called spherical stress state. For the spherical stress state for all planes. For this stress state we have in the principal axes:

$$p_n = \sigma_I, \; p_\tau = 0.$$

2.13.5 Plane stress.

In plane stress, all stress components are parallel to one plane. If we choose basis vectors $\{e_i\}$ such that direction of vector e_3 coincides with third principal axis and $\sigma_{III} = 0$, then the matrix representation of stress tensor is:

$$[P] = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and the stress vector lies on the plane with basis $\{e_1, e_2\}$. The characteristic equation for the plane stress state is

$$\lambda^3 - J_1(P)\lambda^2 + J_2(P)\lambda = \lambda(\lambda^2 - J_1(P)\lambda + J_2(P)) = 0.$$

2.13.6 The deviatoric stress tensor.

The stress state at a place in a continuum body can be decomposed into components called the spherical and deviator stress tensors. If we denote a value

$$p = \frac{1}{3}\sigma_{ii} = \frac{1}{3}J_1(P),$$

(which is called mean normal stress) then

$$P' = P - pI$$

is called deviator stress tensor. For it the first invariant $J_1(P') = 0$. Others invariants are

$$J_2(P') = \sigma_{kk}^2 - J_2(P) = \frac{3}{2}(p_\tau^{oct})^2,$$

$$J_3(P') = J_3(P) - \frac{1}{3}\sigma_{kk}J_2(P) + \frac{2}{3}\sigma_{kk}^3.$$

2.14 Deformation

In continuum mechanics material bodies consisting of particles (material points) are studied. Any change of a material volume $\Omega_{t_0} \to \Omega_t$ is the result of a displacement of the points of the body. The change in size or shape, or possibly both is called a deformation. A movement of material volume is described by the homeomorphism $\gamma_t : \Omega_{t_0} \to \Omega_t$ (see Axiom 3). Let us suppose that γ_t and $(\gamma_t)^{-1}$ are continuously differentiable in the neighborhood of some particle $\xi \in \Omega_{t_0}$.

Definition 2.14. *The vector*

$$w = x - \xi = \gamma(\xi, t) - \xi$$

is called a displacement vector of the particle ξ.

Let us consider two neighbor particles of the body situated at the point ξ and $\xi + d\xi$ in the initial undeformed configuration Ω_{t_0}. Under the displacement these particles move to the positions $x = \gamma_t(\xi)$ and $x + dx$ in the deformed configuration Ω_t. For the value of dx one has

$$dx = T < d\xi >,$$

where the tensor $T = \dfrac{\partial x}{\partial \xi}$ is called a deformation gradient tensor or simply the deformation gradient. The tensor T characterizes a local deformation at ξ. For describing motions we use several measures of deformation. First let us consider the measure based on the change during the deformation the following magnitude:

$$(dx)^2 - (d\xi)^2 = (T < d\xi >)(T < d\xi >) - (d\xi)^2 = d\xi((T^* \circ T - I) < d\xi >)$$

The symmetric tensor

$$2\widehat{\varepsilon} = T^* \circ T - I$$

is called a Lagrangean finite strain tensor. The difference $(dx)^2 - (d\xi)^2$ may also be developed in terms of spatial (Eulerian) variables

$$(dx)^2 - (d\xi)^2 = (dx)^2 - (T^{-1} < dx >)(T^{-1} < dx >) = dx\,(I - (T^{*-1} \circ T^{-1}) < dx >),$$

where the symmetric tensor

$$2\varepsilon = I - T^{*-1} \circ T^{-1} = 2T^{*-1} \circ \widehat{\varepsilon} \circ T^{-1}$$

is called the Eulerian finite strain tensor.

Another measure of the deformation is a change of the angle between two elements $d\xi$ and $d_1\xi$:

$$dx \cdot d_1x = T < d\xi > T < d_1\xi >= d\xi\,(T^* \circ T) < d_1\xi >= d\xi\,(2\widehat{\varepsilon} + I) < d_1\xi >$$

or

$$d\xi \cdot d_1\xi = T^{-1} < dx > \cdot T^{-1} < d_1x >= dx \cdot ((T^{*-1} \circ T^{-1}) < d_1x >) = dx \cdot ((I - 2\varepsilon) < d_1x >).$$

These finite strain tensors may also be expressed in terms of the displacement gradient $\frac{\partial w}{\partial \xi}$:

$$2\widehat{\varepsilon} = (I + \frac{\partial w}{\partial \xi})^*(I + \frac{\partial w}{\partial \xi}) - I = (\frac{\partial w}{\partial \xi})^* + \frac{\partial w}{\partial \xi} + (\frac{\partial w}{\partial \xi})^* \circ \frac{\partial w}{\partial \xi},$$

$$2\varepsilon = 2T^{*-1} \circ \widehat{\varepsilon} \circ T^{-1} = T^{*-1}(T^* \circ (\frac{\partial w}{\partial x})^* + \frac{\partial w}{\partial x} \circ T + T^* \circ (\frac{\partial w}{\partial x})^* \frac{\partial w}{\partial x} \circ T)T^{-1}.$$

Since

$$T = I + \frac{\partial w}{\partial \xi}, \ \frac{\partial w}{\partial \xi} = \frac{\partial w}{\partial x} \circ T,$$

then

$$T^{-1} = I - \frac{\partial w}{\partial x}$$

and

$$\begin{aligned} 2\varepsilon &= (\tfrac{\partial w}{\partial x})^* T^{-1} + T^{*-1}\tfrac{\partial w}{\partial x} + (\tfrac{\partial w}{\partial x})^* \tfrac{\partial w}{\partial x} = \\ &= (\tfrac{\partial w}{\partial x})^*(I - \tfrac{\partial w}{\partial x}) + (I - \tfrac{\partial w}{\partial x})^* \tfrac{\partial w}{\partial x} + (\tfrac{\partial w}{\partial x})^* \tfrac{\partial w}{\partial x} = \\ &= (\tfrac{\partial w}{\partial x})^* + \tfrac{\partial w}{\partial x} - (\tfrac{\partial w}{\partial x})^* \tfrac{\partial w}{\partial x}. \end{aligned}$$

The strain tensors in terms of Lagrangean or Eulerian components are symmetric second order tensors. Moreover eigenvalues of the tensor $T^* \circ T$ are positive. In fact, let a vector $l \neq 0$ be an eigenvector of the tensor $T^* \circ T$, which corresponds to an eigenvalue λ:

$$T^* \circ T < l >= \lambda\, l.$$

Hence,

$$\lambda(l, l) = l(T^* \circ T < l >) = (T < l >)^2 > 0.$$

The principal axes of $T^* \circ T$ and $\widehat{\varepsilon}$ coincide, and their eigenvalues, denoted by λ_i and ε_i, are related with the formulae

$$\lambda_i = 1 + 2\varepsilon_i, \ (i = 1, 2, 3).$$

Therefore the eigenvalues of a strain tensor satisfy the conditions

$$\varepsilon_i > -\frac{1}{2}.$$

The principal axes of tensor $\widehat{\varepsilon}$ are called principal axes of deformation, and eigenvalues ε_i are called principal strains.

If a strain tensor is equal to zero, then the length of the element $d\xi$ is unchanged, and the angle between any two elements $d\xi$ and $d_1\xi$ will also be unchanged. Thus in the absence of strain only a rigid body displacement can occur.

Theorem. In any continuum media (for continuous motion) density ρ only depends on invariants of a Lagrangean finite strain tensor

$$\rho = \frac{\rho_0}{\sqrt{1 + 2J_1(\widehat{\varepsilon}) + 4J_2(\widehat{\varepsilon}) + 8J_3(\widehat{\varepsilon})}}$$

where $\rho_0 = \rho(x, t_0)$.

Proof. From the continuity equation one has

$$\frac{d\rho}{dt} = -\rho div(v).$$

by virtue of the Euler's formula we obtain

$$\frac{d\rho J}{dt} = J\frac{d\rho}{dt} + \rho\frac{dJ}{dt} = 0.$$

It means that

$$J\rho = \rho_0(\xi).$$

Because $J = \det(T)$ and $T^* \circ T = 2\widehat{\varepsilon} + I$, then

$$\det(T^* \circ T) = \det(T^*)\det(T) = \det(T)^2 = \det(2\widehat{\varepsilon} + I) = 8\det(\widehat{\varepsilon} + \frac{1}{2}I) =$$

$$= 8\det(\widehat{\varepsilon} - \lambda I)_{|\lambda=-\frac{1}{2}} = 8(-\lambda^3 + J_1(\widehat{\varepsilon})\lambda^2 - J_2(\widehat{\varepsilon})\lambda + J_3(\widehat{\varepsilon}))_{|\lambda=-\frac{1}{2}} = 1 + 2J_1(\widehat{\varepsilon}) + 4J_2(\widehat{\varepsilon}) + 8J_3(\widehat{\varepsilon}).$$

or

$$J = \sqrt{1 + 2J_1(\widehat{\varepsilon}) + 4J_2(\widehat{\varepsilon}) + 8J_3(\widehat{\varepsilon})}.$$

2.15 Elements of mathematical thermodynamics

Thermodynamics studies relations between the heat energy and other kinds of energies and gives rules of reciprocal conversion of one kind of energy into another. For example, if a body is heated, then strains and stresses are developed. Conversely, if a body is strained rapidly, then heat is generated inside the body.

The main notion of thermodynamics is a notion of a physical body state. Phenomenological description of the state is given with the help of various functions called state variables. For example, introduced before the density ρ (or specific volume $\tau = 1/\rho$), the internal energy U are parameters of the state of continuous medium. Also absolute temperature $\theta > 0$, specific entropy η and pressure p are basic state variables.

Let $z = (z_1, z_2, ..., z_k)$ be a set of main state variables of continuous medium: other state variables are functions of these variables. Such medium is called k–parameter medium. In the space of state variables a change of variables from one state $z^{(1)}$ to another state $z^{(2)}$ are separated out. These changes are called processes. If for a process from $z^{(1)}$ into $z^{(2)}$ there exists, process from $z^{(2)}$ into $z^{(1)}$, then such process is called reversible, otherwise it is called irreversible. The state variables are obeyed the first and the second laws of thermodynamics.

The first law of thermodynamics may be stated as follows: the time rate of change of kinetic energy and internal energy in a body is equal to the rate of work done on the body plus the changes in all other energies, such as heat, magnetic, electrical, and chemical, per unit time or

$$\frac{d}{dt}(E_k + E_i) = (N_i + N_e) + Q + \sum_\alpha N_\alpha.$$

In absence of the energies N_α other than those due to mechanical power $(N_i + N_e)$ and heat Q, we write the conservation law of energy.

The second law of thermodynamics is based on the concept of entropy associated with irreversible thermodynamic processes. The entropy is regarded as a measure of change of energy dissipation with respect to temperature.

We define an entropy S as an additive continuously differentiable function

$$S = \int_\omega \rho\eta \, d\omega,$$

where η is an entropy density per unit mass. Furthermore, the total entropy production B is defined by:

$$B = \frac{dS}{dt} + \int_{\partial\omega} (\frac{q}{\theta}) n \, d\sigma - \int_\omega \rho(\frac{h}{\theta}) \, d\omega \geq 0. \tag{2.10}$$

This expression is referred as the second law of thermodynamics in continuum mechanics, which states that the total entropy production is always greater than or equal to zero. This is also known as the Clausius-Duhem inequality. We may rewrite equation (2.10) as

$$\int_\omega [\rho \frac{d\eta}{dt} + \nabla(\frac{q}{\theta}) - \rho(\frac{h}{\theta})] \, d\omega \geq 0.$$

or

$$\rho\theta \frac{d\eta}{dt} + \nabla \cdot q - \frac{1}{\theta} q \cdot \nabla\theta - \rho h \geq 0,$$

which is a local form of the Clausius-Duhem inequality. The value

$$\Phi = \rho\theta \frac{d\eta}{dt} - \rho h + div(q)$$

is called an internal dissipation. Therefore, the second thermodynamics law states

$$\Phi - \frac{1}{\theta} q \cdot \nabla\theta \geq 0. \tag{2.11}$$

The product $(-\theta\eta)$ is the irreversible heat energy due to entropy as related to temperature, with the negative sign indicating that compressive reaction results from thermal expansion (temperature rise) in a restrained body. The sum of internal energy U and irreversible heat energy $(-\theta\eta)$ is known as a Helmholtz free energy $F = U - \theta\eta$. Substituting it into the energy equation, one obtains

$$\rho \frac{dF}{dt} = P : D - \rho\eta \frac{d\theta}{dt} - \Phi.$$

Remark. For any irreversible process we count $\Phi > 0$, whereas $\Phi = 0$ is for a reversible process.

Remark. Here we added the additional term $\int_\omega \rho h \, d\omega$, in which h is the heat supply per unit mass. Before this consideration we studied (see Axiom 7) that $(h = 0)$

$$Q = \int_{\partial\omega} q_n \, d\sigma.$$

Axiom 9 (thermodynamics axiom). For continuous medium are fair the first and the second laws of thermodynamics.

Because condition (2.11) takes place for irreversible and reversible processes, then we must choose heat flux such that

$$-\frac{1}{\theta} q \cdot \nabla\theta \geq 0.$$

For example, the Fourier heat conduction law states that

$$q = -\kappa \nabla \theta,$$

where κ is a new state variable, which is called a coefficient of heat conductivity.

Axiom 10 (Fourier's axiom). A heat flux is proportional to gradient of temperature.

A coefficient of heat conductivity κ is always positive. In the models of continuum mechanics it is considered as a known function of other state variables. Therefore, the energy equation has the form

$$\rho \frac{dU}{dt} = P : D + div(\kappa \nabla \theta).$$

2.15.1 Ideal continuous media

For "ideal" continuous media it is supposed that the stress tensor P is a spherical tensor

$$P = -pI,$$

here p is called a pressure. State variables of "ideal" continuous medium are defined by five state variables

$$\rho = \frac{1}{\tau}, U, \theta, \eta, p.$$

Because the stress tensor is a spherical tensor, then for any reversible process one has

$$\theta d\eta = dU + pd\tau. \tag{2.12}$$

This equality is called the main thermodynamics identity.

Let us consider two–parameter ideal continuous medium. Between five state variables from equation (2.12) one can obtain two equations. For example, if one take τ and η as the main state variables, then other variables are functions of τ and η:

$$U = U(\tau, \eta),\ \theta = \theta(\tau, \eta),\ p = p(\tau, \eta).$$

Substituting these functions into the main thermodynamical identity, we get

$$\theta d\eta = U_\tau d\tau + U_\eta d\eta + pd\tau.$$

In this identity $d\tau$ and $d\eta$ are arbitrary. Hence,

$$\theta = U_\eta(\tau, \eta),\ p = -U_\tau(\tau, \eta).$$

For full description of thermodynamical state it is enough to have the function $U = U(\tau, \eta)$. Such functions are called constitutive equations. As usual these equations are obtained from the physical experiments.

In continuum mechanics very often we use the following constitutive equations:

(a) the internal energy U is a function of the state variables τ and η: $U = U(\tau, \eta)$;

(b) the heat content (enthalpy) $i = U + p\tau$ is a function of the state variables p and η: $U + p\tau = i(p, \eta)$;

(c) the free energy $F = U - \theta\eta$ is a function of the state variables τ and θ: $U - \theta\eta = F(\tau, \theta)$;

(d) the thermodynamical potential $\psi = U - \theta\eta + p\tau$ is a function of the state variables p and θ: $U - \theta\eta + p\tau = \psi(p, \theta)$.

Remark. The main thermodynamical identity has another representation

$$\theta\, \dot{\eta} = \dot{U} + p\, \dot{\tau}.$$

Exercises. Find two equations for the cases (a)–(d).
Answers.
(b) $\theta = i_\eta(p,\eta),\ \tau = i_p(p,\eta)$;
(c) $\eta = -F_\theta(\tau,\theta),\ p = -F_\tau(\tau,\theta)$;
(d) $\tau = \psi_p(p,\theta),\ \eta = -\psi_\theta(p,\theta)$.

2.16 Fluids

Definition. Fluid or gas is such continuous media in which the stress tensor P is a function of rate-of-strain tensor D. Also the stress tensor depends on the thermodynamic state variables $\Pi = (\rho, U, \theta, \eta, p)$ and coordinates x and time t

$$P = F(D, \Pi, x, t).$$

Concretization of the function F is formulated in the Stokes axioms.

Axiom (Stokes). For fluids and gases are valid.

1^0. The form of F does not depend either on position in space or on time (medium is homogeneous):

$$P = F(D, \Pi).$$

2^0. The function $F(D,\Pi)$ is a continuous symmetric function of rate-of-strain tensor D (medium is isotropic).

3^0. If $D = 0$, then medium is "ideal":

$$F(0, \Pi) = -pI$$

By virtue of the theorem and the property 2^0 one obtains

$$P = \alpha I + \beta D + \gamma D^2,$$

where the coefficients α, β, γ are functions of invariants of the tensor D and also thermodynamic state variables.

Axiom (state). Fluids and gases are two-parameter media for which the main thermodynamic identity

$$\theta d\eta = dU + pd\tau$$

takes place. The coefficient of heat conductivity κ is a known function of the state variables. For example, $\kappa = \kappa(\rho, \eta)$.

For a closure of system of equations, which describe a motion of fluids and gases one has to know the functions

$$\alpha = \alpha(J(D), \rho, \eta),\ \beta = \beta(J(D), \rho, \eta),\ \gamma = \gamma(J(D), \rho, \eta),$$

where $J(D) = (J_1(D), J_2(D), J_3(D))$. We obtain a closed model of fluid or gas

$$\frac{d\rho}{dt} + \rho div(v) = 0,$$

$$\rho\frac{dv}{dt} = div(P) + \rho f,$$

$$\rho\frac{dU}{dt} = P : D + div(\kappa\nabla\theta),$$

where $P = \alpha I + \beta D + \gamma D^2$. From the Stokes axioms we only have

$$\alpha(0, \rho, \eta) = -p,$$

the others we have to obtain from experimental data or we have to take additional assumptions.

Concretization of this model requires huge experimental data. Therefore this model is not applied in practice.

Besides noted postulates, Stokes further assumed that P is a linear function of D.

Axiom. The function $F(D, \Pi)$ is a linear function of D.

From this axiom we have $\gamma = 0$ and β does not depend on invariants of a rate-of-strain tensor D. Because only invariant $J_1(D)$ is a linearly dependent, then the function is a linear function of $J_1(D)$:

$$\alpha = -p + \lambda J_1(D),\ \beta = 2\mu.$$

The scalar invariants λ and μ depend on the thermodynamic state variables and they are called the first and the second coefficients of viscosity, respectively. Hence, the stress tensor is

$$P = -pI + \lambda div(v) + 2\mu D.$$

Sometimes it is useful to define the quantity

$$\mu' = \lambda + \frac{2}{3}\mu,$$

which is called a bulk coefficient of viscosity.

After substituting the representation of the stress tensor P into equations of fluid, one has

$$div(P) = -\nabla p + \nabla(\lambda div(v)) + div(2\mu D),$$

$$P : D = -p div(v) + \Phi,$$

where $\Phi = \lambda(div(v))^2 + 2\mu D : D = (\lambda + \frac{2}{3}\mu)(div(v))^2 + 2\mu D' : D'$, and D' is a deviatoric of strain-of-rate tensor:

$$D' = D - \frac{1}{3}div(v)I.$$

The function Φ is called a dissipation function. The second law of thermodynamics gives

$$\Phi + \frac{\kappa}{\theta}(\nabla\theta)^2 \geq 0.$$

Because Φ does not depend on $\nabla\theta$, we have to require that $\Phi \geq 0$.

Let us consider the conditions under which Φ is always nonegative, i.e., $\Phi \geq 0$. We rewrite the value of the function Φ in the principal axes of the rate-of-strain tensor:

$$D = O \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} O^*,$$

where O is an orthogonal transformation. In this matrix representation we have

$$\Phi = \lambda(J_1(D))^2 + 2\mu O_{ik} d_k O_{lk} O_{i\beta} d_\beta O_{l\beta} = \lambda(J_1(D))^2 + 2\mu \delta_{k\beta} d_k \delta_{k\beta} d_\beta =$$

$$= \lambda(d_1 + d_2 + d_3)^2 + 2\mu(d_1^2 + d_2^2 + d_3^2)$$

$$= \frac{1}{3}[(3\lambda + 2\mu)(d_1 + d_2 + d_3)^2 + 2\mu((d_1 - d_2)^2 + (d_2 - d_3)^2 + (d_3 - d_1)^2)].$$

If $\Phi \geq 0$ for all tensors D, then we obtain

$$3\lambda + 2\mu \geq 0,\ \mu \geq 0.$$

For real fluids these inequalities must always be satisfied.

Assume that η and ρ (or τ) are the main thermodynamical state variables, then from the main thermodynamical identity we have

$$\rho\frac{dU}{dt} = \rho\theta\frac{d\eta}{dt} - pdiv(v)$$

and then the energy equation has the form

$$\rho\theta\frac{d\eta}{dt} = div(\kappa\nabla\theta) + \Phi.$$

If λ, μ, κ and U are known functions of two parameters ρ and η, then we obtain the closed model of fluid dynamics

$$\frac{d\rho}{dt} + \rho div(v) = 0,\ \rho\frac{dv}{dt} = -\nabla p + \nabla(\lambda div(v)) + div(2\mu D) + \rho f, \tag{2.13}$$

$$\rho\theta\frac{d\eta}{dt} = div(\kappa\nabla\theta) + \Phi.$$

$$\theta = \frac{\partial U}{\partial\eta},\ p = \rho^2\frac{\partial U}{\partial\rho}.$$

This model is called a model of a viscous compressible gas.

2.16.1 Partial models of fluids and gases.

Viscous incompressible flow equations

Definition. Fluid flow for which the density remains constant are called incompressible flows.

For the incompressible flows

$$\frac{d\rho}{dt} = 0.$$

From the continuity equation we have

$$div(v) = 0.$$

Axiom (incompressibility). For incompressible flows ($\rho = const$), the second coefficient of viscosity is constant and does not depend on temperature.

In this case we find

$$div(2\mu D) = 2\mu div(D) = \mu(div(\frac{\partial v}{\partial x}) + div((\frac{\partial v}{\partial x})^*)).$$

Let us simplify the expressions $div(\frac{\partial v}{\partial x})$ and $div((\frac{\partial v}{\partial x})^*)$:

$$a \cdot div((\frac{\partial v}{\partial x})^*) = tr(\frac{\partial}{\partial x}(\frac{\partial v}{\partial x} < a >)) = \frac{\partial}{\partial x_\alpha}(\frac{\partial v_\alpha}{\partial x_\beta}a_\beta) = a_\beta\frac{\partial^2 v_\alpha}{\partial x_\alpha \partial x_\beta} =$$

$$= a_\beta\frac{\partial}{\partial x_\beta}(\frac{\partial v_\alpha}{\partial x_\alpha}) = a_\beta(\frac{\partial}{\partial x_\beta}div(v)) = a \cdot \nabla div(v)$$

$$a \cdot div(\frac{\partial v}{\partial x}) = tr(\frac{\partial}{\partial x}((\frac{\partial v}{\partial x})^* < a >)) = tr(\frac{\partial}{\partial x}(\frac{\partial a_\alpha v_\alpha}{\partial x})) = \sum_\beta \frac{\partial^2}{\partial x_\beta^2}(v_\alpha a_\alpha) =$$

$$= a_\alpha \sum_\beta \frac{\partial^2 v_\alpha}{\partial x_\beta^2} = a \cdot \Delta v,$$

or

$$div(2\mu D) = \mu(\nabla div(v) + \Delta v).$$

Therefore, we have

$$div(v) = 0, \tag{2.14}$$
$$\frac{dv}{dt} + \frac{1}{\rho}\nabla p = \nu \Delta v + f,$$

where $\nu = \mu/\rho$ is constant (it is called a kinematic viscosity). In this system of equations we have unknown only v and ρ. These equations are called the Navier-Stokes equations.

It is remarkable that thermodynamic does not participate in this model. The energy equation

$$\frac{d\theta}{dt} = \frac{1}{\rho c_v} div(\kappa \nabla \theta) + \Phi',$$

can be solved afterwards. Here

$$\Phi' = \frac{2\nu}{c_v} D' : D',$$

and c_v is a known function of temperature.

Axiom (of ideal fluid). For ideal fluids the coefficient of kinematic viscosity is equal to zero $\nu = 0$.

In this case the system of equations (2.14) becomes simpler

$$div(v) = 0,$$

$$\frac{dv}{dt} + \frac{1}{\rho}\nabla p = f.$$

This system is called the Euler equations. For the Euler equation the energy equation has the form ($\Phi' = 0$):

$$\frac{d\theta}{dt} = \frac{1}{\rho c_v(\theta)} div(\kappa \nabla \theta).$$

If κ and c_v are constant functions, then we obtain the classical heat equation ($k = \kappa/(\rho c_v)$)

$$\frac{d\theta}{dt} = k \Delta \theta.$$

Equations of ideal gas (Euler's equations)

Equations of inviscid (nonviscous) flow, originally derived by Euler, can directly be obtained from (2.13) by setting λ, μ and κ equal to zero.

Axiom (of an ideal gas). For an ideal gas

$$\lambda = 0, \ \mu = 0, \ \kappa = 0.$$

In this case the system of gas dynamics equations is

$$\frac{d\rho}{dt} + \rho div(v) = 0, \ \rho \frac{dv}{dt} + \nabla p = \rho f,$$

$$\frac{d\eta}{dt} = 0,\ p = p(\rho, \eta).$$

Flow in which p can be expressed as a function of ρ is said to be barotropic. For barotropic flows the system of equations is separated on two parts

$$\frac{d\rho}{dt} + \rho div(v) = 0,\ \frac{dv}{dt} + \nabla F(\rho) = f,$$

and

$$\frac{d\eta}{dt} = 0,\ p = p(\rho),$$

where $F(\rho) = \int \frac{dp}{\rho}$. This system of equations is called a system of a barortopic gas.

2.17 Elastic solids.

Definition. Continuous media in which the stress tensor P is a function of the finite Lagrangian strain tensor $\widehat{\varepsilon}$ are called solids.

Definition. In elastic solids thermodynamical processes are reversible ones. Thermodynamics state variables are θ and finite Lagrangian strain tensor $\widehat{\varepsilon}$. The others state parameters are connected with $\widehat{\varepsilon}$ and θ by additional axioms.

Because in elastic solids thermodynamical processes are reversible ones than

$$\Phi \equiv \rho\theta\eta^{\cdot} + div(q) = 0.$$

Axiom (of reversible process). In elastic solids there is equality

$$\rho\theta\eta^{\cdot} + div(q) = 0.$$

Then the equation of conservation law we can rewrite as following

$$\rho\frac{dU}{dt} - P : D + div_x(q) = \rho\frac{dU}{dt} - \widehat{P} : \frac{\partial\widehat{\varepsilon}}{\partial t} + div_x(q),$$

where $P = T \circ \widehat{P} \circ T^*$ and because

$$P : D = \widehat{P} : \frac{\partial\widehat{\varepsilon}}{\partial t}$$

Really:

$$2\frac{\partial\widehat{\varepsilon}}{\partial t} = \frac{\partial T^*}{\partial t} \circ T + T^* \circ \frac{\partial T}{\partial t} = (\frac{\partial v}{\partial \xi})^* \circ T + T^* \circ \frac{\partial v}{\partial \xi} = \frac{\partial T^*}{\partial t} \circ T + T^* \circ \frac{\partial T}{\partial t} =$$

$$= (\frac{\partial v}{\partial x} \circ T)^* \circ T + T^* \circ (\frac{\partial v}{\partial x} \circ T) = 2T^* \circ D \circ T,$$

Then

$$\rho\frac{\partial F}{\partial t} = \widehat{P} : \frac{\partial\widehat{\varepsilon}}{\partial t} - \rho\eta\frac{\partial\theta}{\partial t} - \Phi.$$

with Helmholtz free energy $F = U - \theta\eta$.

Axiom. Free energy F of elastic solid is a function only θ and $\widehat{\varepsilon}$. Free energy $F = F(\theta, \widehat{\varepsilon})$ is isotropic scalar function of $\widehat{\varepsilon}$. After substitution in the energy equation we get

$$\rho\frac{\partial F}{\partial t} = \rho\frac{\partial F}{\partial\widehat{\varepsilon}} : \frac{\partial\widehat{\varepsilon}}{\partial t} + \frac{\partial F}{\partial\theta}\frac{\partial\theta}{\partial t} = \rho\frac{\partial F}{\partial t} = \widehat{P} : \frac{\partial\widehat{\varepsilon}}{\partial t} - \rho\eta\frac{\partial\theta}{\partial t}.$$

If we compare left and right side we have

$$\rho\frac{\partial F}{\partial \widehat{\varepsilon}} = \widehat{P},\ \frac{\partial F}{\partial \theta} = -\eta.$$

Lemma. Stress tensor $\widehat{P}$ is isotropic tensor function of $\widehat{\varepsilon}$.

Proof. Let F' be derivative $\frac{\partial F}{\partial \widehat{\varepsilon}}$. We have to show that for arbitrary orthogonal transformation O we get

$$F'(O\widehat{\varepsilon}O^*) = OF'(\widehat{\varepsilon})O^*.$$

Let $\overline{\varepsilon} = O\widehat{\varepsilon}O^*$, then, because $F(O\widehat{\varepsilon}O^*) = F(\widehat{\varepsilon})$ and (by definition of derivative)

$$F(\widehat{\varepsilon}+A) - F(\widehat{\varepsilon}) \cong F'(\widehat{\varepsilon}) : A,\ F(\overline{\varepsilon}+OAO^*) - F(\overline{\varepsilon}) \cong F'(\overline{\varepsilon}) : (OAO^*)$$

$$F'(\widehat{\varepsilon}) : A = F'(\overline{\varepsilon}) : (OAO^*)$$

$$B : (OAO^*) = tr(B^*OAO^*) = tr(OO^*B^*OAO^*) = tr(O^*B^*OA) = (O^*BO) : A.$$

By virtue of arbitrariness of A we get

$$F'(\widehat{\varepsilon}) = O^*F'(\overline{\varepsilon})O$$

or

$$OF'(\widehat{\varepsilon})O^* = F'(O\widehat{\varepsilon}O^*)$$

Now we are ready to rewrite all equations. We will use following relationships

$$div_x(P) = div_\xi(T\widehat{P}) - T\widehat{P}T^* < div_\xi(T^{*-1}) >$$

$$div_x(q) = div_\xi(T^{-1} < q >) - q\cdot div_\xi(T^{*-1}),\ \ \nabla_x\theta = T^{*-1} < \nabla_\xi\theta > .$$

Proof. 1^0.

$$a\cdot div_x(P) = tr(\frac{\partial}{\partial x}P < a >) = tr((\frac{\partial}{\partial \xi}P < a >)T^{-1}) = tr(T^{-1}(\frac{\partial}{\partial \xi}P < a >)) =$$

$$= tr(\frac{\partial}{\partial \xi}(T^{-1}P < a >)) - tr(\frac{\partial T^{-1} < b >}{\partial \xi}_{|b=P<a>}) =$$

$$a\cdot div_\xi(T^{-1}P)^* - div_\xi(T^{*-1})\cdot P < a >= a\cdot(div_\xi(PT^{*-1}) - P < div_\xi(T^{*-1}) >=$$

$$= a\cdot(div_\xi(T\widehat{P}) - T\widehat{P}T^* < div_\xi(T^{*-1}) >).$$

2^0.

$$div_x(q) = tr(\frac{\partial q}{\partial x}) = tr(\frac{\partial q}{\partial \xi}T^{-1}) = tr(T^{-1}\frac{\partial q}{\partial \xi}) = tr(\frac{\partial}{\partial \xi}(T^{-1} < q >)) - tr(\frac{\partial T^{-1} < b >}{\partial \xi}_{|b=q}) =$$

$$= div_\xi(T^{-1} < q >) - q\cdot div_\xi(T^{*-1}).$$

3^0.

$$a\cdot\nabla_x\theta = a\cdot(\frac{\partial\theta}{\partial x}) = \frac{\partial\theta}{\partial\xi}\cdot T^{-1} < a >= a\cdot T^{*-1} < \nabla_\xi\theta > .$$

Let us consider $x = \gamma(\xi,t) = w(\xi,t) + \xi$ then

$$T = \frac{\partial w}{\partial \xi} + I,\ 2\widehat{\varepsilon} = \frac{\partial w}{\partial \xi} + (\frac{\partial w}{\partial \xi})^* + (\frac{\partial w}{\partial \xi})^*\frac{\partial w}{\partial \xi}$$

Now we have

$$\rho\frac{\partial^2 w}{\partial t^2} = div_\xi(T\widehat{P}) - T\widehat{P}T^* < div_\xi(T^{*-1}) > +\rho f.$$

$$\rho\theta(\frac{\partial^2 F}{\partial\widehat{\varepsilon}\partial\theta} : \frac{\partial\widehat{\varepsilon}}{\partial t} + \frac{\partial^2 F}{\partial\theta^2}\frac{\partial\theta}{\partial t}) = -div_\xi(T^{-1} < q >) + q\cdot div_\xi(T^{*-1}),$$

where $q = -\kappa T^{*-1} < \nabla_\xi\theta >,\ \widehat{P} = \alpha_* I + \beta_*\widehat{\varepsilon} + \gamma_*\widehat{\varepsilon}^2,\ \rho = \rho_0(1 + 2J_1(\widehat{\varepsilon}) + 4J_2(\widehat{\varepsilon}) + 8J_3(\widehat{\varepsilon}))^{-1/2}$.

Axiom. There exist "natural" state of elastic solid in which

$$P = 0,\ \widehat{\varepsilon} = 0,\ \theta = \theta_0(= const).$$

2.17.1 Linear theory of elasticity.

Assumptions:

(a) (geometrically linear solids) :
small: w and all derivatives of w are small: they have the same order.
(b) (physically linear solids):
$\widehat{P}$ is linear and homogeneously depends of $\widehat{\varepsilon}$ and $(\theta - \theta_0)$.
(c) (physically linear solids):
T and $\frac{\widetilde{\theta}}{\theta_0}$ have the same order of smallness.
Some manipulations:

$$2\widehat{\varepsilon} = \frac{\partial w}{\partial\xi} + (\frac{\partial w}{\partial\xi})^* + (\frac{\partial w}{\partial\xi})^*\frac{\partial w}{\partial\xi} \approx \frac{\partial w}{\partial\xi} + (\frac{\partial w}{\partial\xi})^*$$

$$\widehat{\varepsilon} \approx \varepsilon,\ \widehat{P} \approx P,$$

$$\alpha_* = \alpha_*(\widetilde{\theta}, J_1(\widehat{\varepsilon}), J_2(\widehat{\varepsilon}), J_3(\widehat{\varepsilon})) \approx \alpha_*^0 + \alpha_*^1\widetilde{\theta} + \alpha_*^2 J_1(\widehat{\varepsilon}) + \alpha_*^3 J_2(\widehat{\varepsilon}) + \alpha_*^4 J_3(\widehat{\varepsilon})$$

Because

$$P = 0,\ \widehat{\varepsilon} = 0,\ \theta = \theta_0(= const)$$

then $\alpha_*^0 = 0$and

$$P = (-\sigma\widetilde{\theta} + \lambda J_1(\widehat{\varepsilon}))I + 2\mu\widehat{\varepsilon},$$

where σ, λ, μ are constants.

$$\rho = \rho_0(1 - J_1(\widehat{\varepsilon})),$$

$$\rho_0\frac{\partial^2 w}{\partial t^2} = -\sigma\nabla_\xi\widetilde{\theta} + \lambda\nabla_\xi J_1(\varepsilon) + \mu div_\xi(2\varepsilon) + \rho_0 f.$$

$$\rho_0 c_0\frac{\partial\widetilde{\theta}}{\partial t} = \kappa_0\Delta_\xi\widetilde{\theta} - \sigma\frac{\partial J_1(\varepsilon)}{\partial t}.$$

But

$$J_1(\varepsilon) = div_\xi(w),\ div_\xi(2\varepsilon) = div_\xi(\frac{\partial w}{\partial\xi} + (\frac{\partial w}{\partial\xi})^*) = \Delta_\xi w + \nabla_\xi(div_\xi(w)),$$

because

$$div_\xi(\frac{\partial w}{\partial\xi}) = \Delta_\xi w,\ div_\xi((\frac{\partial w}{\partial\xi})^*) = \nabla_\xi(div_\xi(w)).$$

2.18 Shock relations

It is well-known experimental fact that in a real continuum media there exist surfaces, formed of the material particles of the material volume, across which steeply high gradients of pressure, density, temperature and velocity occur. In this case of real continuum media the transition from one side to the other side occurs through a thin layer of material, called shock layer.

To establish the subject of shock waves from the first principle we start from the equations of the mathematical model of moving continuous media (integral conservation laws).

In the moving continuous media for any moving volume $\omega_t \subset \Omega_t$ and any moment of time $t \in \tau$ following equalities are fair:

$$\frac{d}{dt}\left(\int_{\omega_t} \rho \, d\omega\right) = 0, \tag{2.15}$$

$$\frac{d}{dt}\left(\int_{\omega_t} \rho v \, d\omega\right) = \int_{\partial\omega_t} p_n \, d\sigma + \int_{\omega} \rho f \, d\omega,$$

$$\frac{d}{dt}\left(\int_{\omega_t} \rho(x \times v) \, d\omega\right) = \int_{\partial\omega_t} (x \times p_n) \, d\sigma + \int_{\omega} \rho(x \times f) \, d\omega,$$

$$\frac{d}{dt}\left(\int_{\omega_t} \rho\left(\frac{v^2}{2} + U\right) d\omega\right) = \int_{\partial\omega_t} v p_n \, d\sigma + \int_{\omega_t} \rho v f \, d\omega + \int_{\partial\omega_t} q_n \, d\sigma.$$

All these equations can be written as concretization of the equation

$$\frac{d}{dt}\int_{\omega_t} F \, d\omega = \int_{\partial\omega_t} \varphi \cdot n \, d\sigma + \int_{\omega_t} \chi \, d\omega,$$

where F, φ and χ are some functions.

For the continuous motion this equation can be reduced to

$$\int_{\omega_t} \left(\frac{\partial F}{\partial t} + div(Fv - \varphi)\right) d\omega = \int_{\omega_t} \chi \, d\omega,$$

Let us consider bounded domain $G \subset R^4(x,t)$ with piecewise smooth boundary $\Gamma = \partial G$ (four dimensional volume), for each point of which time t belongs to the same interval $\tau = (t_1, t_2)$. We take integral of the equation with respect to time t from t_1 to t_2:

$$\int_{t_1}^{t_2} \left(\int_{\omega_t} \left(\frac{\partial F}{\partial t} + div(Fv - \varphi)\right) d\omega\right) dt = \int_{t_1}^{t_2} \left(\int_{\omega_t} \chi \, d\omega\right) dt.$$

Because of the theorem of mathematical analysis

$$\int_{t_1}^{t_2} \int_{\omega_t} H \, d\omega = \int_G H \, dG$$

for function $H(x,t)$. Let $\mathbf{l}$ be unit normal vector of direction t in R^4, then

$$\frac{\partial F}{\partial t} + div(Fv - \varphi) = Div(F\mathbf{l} + Fv - \varphi) \tag{2.16}$$

where Div is a divergence in the space R^4. Therefore we can use the Gauss-Ostrogradskii theorem

$$\int_G \left(\frac{\partial F}{\partial t} + div(Fv - \varphi)\right) dG = \int_\Gamma (F\mathbf{l} + Fv - \varphi) \cdot \nu \, d\Gamma,$$

where ν is unit outward normal vector to Γ in the space R^4. Then equation (2.16) has representation

$$\int_\Gamma (F\mathbf{1} + Fv - \varphi) \cdot \nu \, d\Gamma = \int_G H \, dG. \tag{2.17}$$

In the case of continuous motion equation (2.17) is equivalent to (2.16) for any material volume. But we can use equation (2.17) for more general motions than the continuous motions, because equation (2.17) has sense for more general functions.

Definition. Motion of continuous media is called generalized motion if the functions ρ, U, P, v, q are bounded measurable functions of the independent variables (x, t) and for them integral relations (2.17) are satisfied for any four dimensional volume $G \subset R^4$.

Class of generalized motions are difficult for analysis. It has not even studied for more simple models of continuum mechanics. We consider one very important subclass of generalized motions: class motions with strong discontinuity.

Let motion be considered in the domain $W \subset R^4$ and this domain is divided by some smooth surface

$$\Pi = \{(x,t) \in R^4 | \; f(x,t) = 0\}$$

on two domains W_1 and W_2.

Definition. Generalized motion of continuous media is called a motion with strong discontinuity if in each domains W_1 and W_2 the functions ρ, U, P, v, q have continuous limit values on the surface Π. If these values are different for W_1 and W_2, then a cross-section

$$\Pi_t = \{x \in R^3 | \; f(x,t) = 0\}$$

of hypersurface Π is called a surface of strong discontinuity.

By virtue of this definition the functions ρ, U, P, v, q have discontinuity of the first order (finite jump) on the surface surface Π. In every point on Π they have two values: $\rho_1, U_1, P_1, v_1, q_1$ are limit values from the domain W_1 and $\rho_2, U_2, P_2, v_2, q_2$ are limit values from the domain W_2. We show that these sets of values can not be arbitrary: they are connected by some relations, which are called equations of the strong discontinuity.

Now we start obtaining these equations.

Let a point $x \in \Pi_t$ and $x + nH(\Delta t) \in \Pi_{t+\Delta t}$. Here n is a normal unit vector to Π_t.

Definition. A limit

$$D_n = \lim_{\Delta t \to 0} \frac{H(\Delta t)}{\Delta t}$$

is called a velocity of replacement of surface Π_t in the direction of normal n.

We note that the normal vector ν has the same direction as a vector $(f_t, \nabla_x f)$ and $\nabla_x f = (\nabla_x f \cdot n)n$. Connection of D_n and four dimensional normal vector ν is obtained from the following formulae

$$f(x,t) = 0,\ f(x + nH(\Delta t), t + \Delta t) = 0.$$

Therefore, after expanding the function $f(x + nH(\Delta t), t + \Delta t)$ into Taylor series, we have

$$0 = f(x + nH(\Delta t), t + \Delta t) = f(x,t) + f_t\Delta t + (\nabla_x f)nH(\Delta t) + O((\Delta t)^2).$$

Dividing the last equation on Δt and tending $\Delta t \to 0$ we obtain

$$f_t + (\nabla_x f)nD_n = 0.$$

It means that vectors ν and $\mathbf{l} + D_n n$ are orthogonal.

On the hypersurface Π we select small domain σ with smooth boundary γ and we construct close hypersurface $\Gamma = \sigma_1 + \sigma_2 + \sigma_3$. Here σ_3 is a side surface of the cylinder G with directrix ν, σ_1 and σ_2 are parts of cylinder which are "parallel" to the surface σ located on a small distance h from the surface σ (see fig.2). In equation (2.17) the integral will consist of the three integrals: with respect to surfaces σ_1, σ_2 and σ_3. Then we tend $h \to 0$. Because a measure of the surface σ_3 tends to 0 and integrand is bounded, then integral

$$\int_{\sigma_3} (F\mathbf{l} + Fv - \varphi) \cdot \nu \, d\Gamma \to 0.$$

Integrals with respect to σ_1 and σ_2 tend to integrals with respect to σ but with different directions of the normal vectors $\nu_1 = -\nu_2$. Therefore we have

$$\int_{\sigma} [F\mathbf{l} + Fv - \varphi] \cdot \nu \, d\Gamma = 0.$$

Here [ ] is a symbol of jump: $[a] = a_2 - a_1$, where a_1 and a_2 are limit values of the function a from different sides of the surface σ. By virtue of arbitrariness of the surface σ we obtain

$$[F\mathbf{l} + Fv - \varphi] \cdot \nu = 0.$$

or

$$[Ff_t + (Fv - \varphi)\nabla_x f] = 0.$$

Substituting $f_t = -(\nabla_x f \cdot n)D_n$ and $\nabla_x f = (\nabla_x f \cdot n)n$ we can rewrite as

$$(\nabla_x f \cdot n)[F(v_n - D_n) - \varphi_n] = 0,$$

where $v_n = v \cdot n$ and $\varphi_n = \varphi \cdot n$.

Concretization of this abstract equation to (2.15) gives

$$[\rho(v_n - D_n)] = 0, \quad [\rho(v_n - D_n)v - P < n >] = 0, \quad [\rho(v_n - D_n)(\frac{v^2}{2} + U) - vP < n > -q_n] = 0.$$

It is more convenient to rewrite these equations by introducing

$$v = v_n n + v_\tau, \; P < n >= p_{nn} n + p_{n\tau}, \; v_n' = v_n - D_n.$$

Then we have

$$[\rho v_n'] = 0,$$

$$\rho v_n'[v_n'] = [p_{nn}], \; \rho v_n'[v_\tau] = [p_{n\tau}],$$

$$\rho v_n'[\frac{1}{2}((v_n')^2 + v_\tau^2) + U - \frac{1}{\rho}p_{nn}] = [v_\tau p_{n\tau} + q_n].$$

From these equations we can see that if $v_n' = 0$, then the first equations is satisfied.

Definition. A strong discontinuity with $v_n' = 0$ or

$$v_n = D_n$$

is called a contact discontinuity.

Equations of contact discontinuity are

$$v_n = D_n, \; [p_n] = 0, \; [v_\tau]p_{n\tau} + [q_n] = 0.$$

Another typical case of a strong discontinuity is for "ideal" media in which a stress tensor $P = -pI$ is a spherical and $q = 0$.

Definition. A strong discontinuity with $v_n' \neq 0$ is called a shock wave.

Equations of a shock wave are

$$[\rho v_n'] = 0, \; [p + \rho(v_n')^2] = 0, \; [v_\tau] = 0, \; [\frac{1}{2}(v_n')^2 + U + \frac{p}{\rho}] = 0.$$