1 NOTATIONS AND CARTESIAN TENSORS

In this chapter we will present the set of notations that will be used and we will find it convenient to use both *matrix* and *tensor notation* depending on the particular application. As *index notation* is an integrated part of tensor algebra, the advantage of this notation will be illustrated and we will then provide an elementary discussion of the concept of tensors and why they appear in a naturally manner when formulating physical relations.

1.1 Matrix notation

In general, a matrix consists of a collection of certain quantities which are termed the *components* of the matrix. The components are ordered in rows and columns and if the number of rows or columns is equal to one, the matrix is *one-dimensional*; otherwise it is *two-dimensional*. A treatment of matrix algebra can be found in many textbooks, for instance Ayres (1962), Eves (1980) or Strang (1980). The intention here is not to provide a resume of matrix algebra, but simply to present sufficient information of the notation used.

A column matrix is denoted by a bold-face, usually lower-case letter, for instance

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}; \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}; \qquad \boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} \tag{1.1}$$

where c_1, c_2, c_3 and c_4 are the components of the matrix $\boldsymbol{c}$. The *dimension* of a matrix is given by the number of rows and columns, i.e. the *column matrix* $\boldsymbol{c}$ of (1.1) has the dimension 4x1. The *transpose* $\boldsymbol{a}^T$ of $\boldsymbol{a}$ is given by the *row matrix*

$$\boldsymbol{a}^T = [a_1 \;\; a_2 \;\; a_3]$$

The *length* of $\boldsymbol{a}$ or $\boldsymbol{a}^T$ is denoted by $|\boldsymbol{a}|$ and we have

$$|\boldsymbol{a}| = (a_1^2 + a_2^2 + a_3^2)^{1/2}$$

The *scalar product* of two column matrices $\boldsymbol{a}$ and $\boldsymbol{b}$ having the same dimensions is defined according to

$$\boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{b}^T\boldsymbol{a} = [b_1 \quad b_2 \quad b_3] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = b_1a_1 + b_2a_2 + b_3a_3 \tag{1.2}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ in the present case are given by (1.1). Therefore, the length $|\boldsymbol{a}|$ of $\boldsymbol{a}$ can be written as

$$|\boldsymbol{a}| = (\boldsymbol{a}^T\boldsymbol{a})^{1/2}$$

A two-dimensional matrix is denoted by a bold-face, usually an upper-case letter, for instance

$$\boldsymbol{B} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \quad ; \qquad \boldsymbol{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \\ C_{31} & C_{32} \end{bmatrix} \tag{1.3}$$

where $\boldsymbol{B}$ is termed a *square matrix* since the number of rows and columns is equal. The transpose $\boldsymbol{B}^T$ of $\boldsymbol{B}$ is obtained by interchanging rows and columns in $\boldsymbol{B}$, i.e.

$$\boldsymbol{B}^T = \begin{bmatrix} B_{11} & B_{21} & B_{31} \\ B_{12} & B_{22} & B_{32} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$

and the matrix $\boldsymbol{B}$ is *symmetric* if $\boldsymbol{B} = \boldsymbol{B}^T$. The unit matrix $\boldsymbol{I}$ is defined

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.4}$$

A *zero matrix* is defined as a matrix where all components are zero. Examples are

$$\boldsymbol{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad ; \quad \boldsymbol{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We note that the *inverse* $\boldsymbol{B}^{-1}$ of a square matrix $\boldsymbol{B}$ is defined by

$$\boldsymbol{B}^{-1}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{B}^{-1} = \boldsymbol{I} \tag{1.5}$$

and that $\boldsymbol{B}^{-1}$ exists if the *determinant* $\det \boldsymbol{B}$ of $\boldsymbol{B}$ is different from zero. If $\det \boldsymbol{B} \neq 0$, then $\boldsymbol{B}$ is *nonsingular*; otherwise it is *singular*. For matrices having the correct dimension the *matrix product* $\boldsymbol{AB}$ exists and we recall that

$$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T \quad ; \qquad (\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

and for two square matrices we have

$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \ \det \boldsymbol{B} \tag{1.6}$$

For a square matrix $\boldsymbol{A}$, consider the quantity $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$, which is a number; this quantity is called a *quadratic form.* If

$$\boxed{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} > 0 \qquad \textit{for all } \boldsymbol{x} \neq 0} \tag{1.7}$$

then the matrix $\boldsymbol{A}$ is said to be *positive definite*. It is recalled that

$$\boxed{\textit{If } \boldsymbol{A} \textit{ is positive definite then } \det \boldsymbol{A} \neq 0}$$

We also mention that a matrix $\boldsymbol{A}$ is called *positive semi-definite* if

$$\boxed{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \geq 0 \qquad \textit{for all } \boldsymbol{x} \neq 0}$$

1.2 Cartesian coordinate system

Whenever a coordinate system is employed in the following, we will for simplicity only make use of the standard *orthogonal, rectangular and right-handed coordinate system* shown in Fig. 1.1. The word rectangular signifies that the coordinate axes are straight orthogonal lines. For reasons that will be unfolded in a moment we label the coordinate axes by x_1, x_2 and x_3 instead of the usual notation of x, y and z.

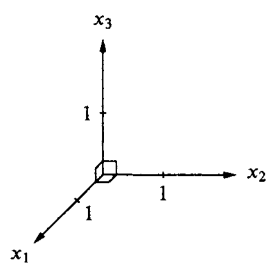

Figure 1.1: Cartesian coordinate system.

In order to maintain the standard definition of distance between two points in this coordinate system, the unit length along all the coordinate axes is equal to the unit length scale. Such a coordinate system is termed a *Cartesian coordinate system* in recognition of the French philosopher and mathematician Descartes (1596-1650), whose Latin name is Cartesius and who introduced the concept of a coordinate system. It is obvious that a certain set of coordinates, i.e. a certain set of x_1-, x_2- and x_3-values defines uniquely the position of a point in the coordinate system.

1.3 Index notation

Index notation is often used in tensor algebra and it is therefore often termed as *tensor notation*. Index notation implies that complicated expressions can be written in a very compact fashion that emphasizes the physical content of these expressions and greatly facilitates mathematical manipulations.

The coordinate axes x_1, x_2 and x_3 in Fig. 1.1 can be written more briefly as x_i, where the index i takes the values $i = 1, 2$ and 3. The column matrix $\boldsymbol{a}$ given by (1.1) can then be written as $[a_i]$ where the brackets [ ] around a_i emphasize that we in the present case interpret the quantity a_i as a matrix. Therefore

$$\boldsymbol{a} = [a_i] = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

where, again, the index i takes the values 1, 2 and 3. In what follows, Latin indices, unless otherwise specified, assume the values 1, 2 and 3; on the other hand, Greek indices will extend over a range to be specified in each case. If reference is made to a_i we refer to the entire quantity given by a_i whereas a specific component of a_i like the one given by, for instance, $i = 2$ is referred to as a_2.

An important convention in index notation is the so-called *summation convention*, which states that if an index is repeated twice then a summation over this index is implied. As an example, the product $b_i a_i$, where the index i is repeated twice, means

$$b_i a_i = b_1 a_1 + b_2 a_2 + b_3 a_3$$

and a comparison with (1.2) shows that $\boldsymbol{b}^T \boldsymbol{a} = b_i a_i$. It is also a convention in index notation that an index cannot be repeated more than twice. If it is repeated twice, it is called a *dummy index* and if it is not repeated, it is called a *free index*, i.e

$$\boxed{\begin{array}{l} \textit{index} = \begin{cases} \textit{free if it appears once} \\ \textit{dummy if it appears twice} \end{cases} \\ \textit{An index can only be free or dummy} \end{array}}$$

It is obvious that the specific letter used for a dummy index is immaterial and we have, for instance, $b_i a_i = b_k a_k$. However, for a free index the specific letter used is of extreme importance. It should also be noted that whereas the position of a quantity in a matrix expression is of significance - we have for example $\boldsymbol{b}^T \boldsymbol{a} \neq \boldsymbol{a} \boldsymbol{b}^T$ - this is not the case in index notation where, for instance, $b_i a_i = a_i b_i = \boldsymbol{a}^T \boldsymbol{b} = \boldsymbol{b}^T \boldsymbol{a}$.

It is also possible to work with quantities having two indices and it is evident that the matrix $\boldsymbol{B}$ given by (1.3) can be written as

$$\boldsymbol{B} = [B_{ij}] = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

where the brackets [ ] around B_{ij} again emphasize that in the present case we interpret the quantity B_{ij} as a matrix.

Using the summation convention, it follows that the inhomogeneous equation system $\boldsymbol{Bx} = \boldsymbol{a}$ can be written as $B_{ij}x_j = a_i$ and that

$$B_{ii} = B_{11} + B_{22} + B_{33}$$

From the rules defined, it follows that each term in an expression must possess the same number of free indices, i.e. whereas $B_{ij}x_j = a_i$ is a valid expression, the formulations $B_{ij}x_j = C$ and $B_{ij}x_j = A_{ij}$ are invalid. The operation, where two free indices are made equal to each other, so that a dummy index arises, is called *contraction.* As an example, contraction of A_{ij} gives A_{ii}.

The *Kronecker delta* δ_{ij} plays an essential role in index notation and tensor algebra and it is defined as

$$\boxed{\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}} \tag{1.8}$$

i.e. it is equal to the unit matrix $\boldsymbol{I}$ given by (1.4). Using the summation convention it follows that

$$B_{ij}\delta_{jk} = B_{ik} \tag{1.9}$$

This result follows from the fact that δ_{ij} only contributes with the value of unity provided that j and k take the same value. Alternatively, the trivial use of the summation convention yields

$$B_{ij}\delta_{jk} = B_{i1}\delta_{1k} + B_{i2}\delta_{2k} + B_{i3}\delta_{3k}$$

and an evaluation of this relation for each i- and k-value results in expression (1.9). Another example of the use of Kronecker's delta arises from the matrix equation $\boldsymbol{AB} = \boldsymbol{I}$. In index notation this is written as

$$A_{ik}B_{kj} = \delta_{ij}$$

which shows that A_{ik} is the inverse of B_{ik}, cf. (1.5). A final important illustration of the use of Kronecker's delta is the expression

$$\frac{\partial a_i}{\partial a_j} = \delta_{ij}$$

This identity follows from the fact that $\partial a_i/\partial a_j$ is zero if $i \neq j$ and unity if i and j are equal.

In accordance with the matrix notation, it follows that the quantity M_{ij} is *symmetric* if

$$M_{ij} = M_{ji}$$

Moreover, a quantity N_{ij} is termed *anti-symmetric* - or *skew-symmetric* - if

$$N_{ij} = -N_{ji}$$

holds. This implies that all diagonal terms in N_{ij} are equal to zero.

Suppose we have an arbitrary quantity P_{ij}. It is always possible to write P_{ij} according to

$$\boxed{P_{ij} = P^s_{ij} + P^a_{ij}} \tag{1.10}$$

where the symmetric part P^s_{ij} of P_{ij} is defined by

$$\boxed{P^s_{ij} = \frac{1}{2}(P_{ij} + P_{ji})} \tag{1.11}$$

and the anti-symmetric part P^a_{ij} of P_{ij} is defined by

$$\boxed{P^a_{ij} = \frac{1}{2}(P_{ij} - P_{ji})} \tag{1.12}$$

It is easy to prove that this decomposition of P_{ij} is unique. To show this, assume that

$$P_{ij} = \tilde{P}^s_{ij} + \tilde{P}^a_{ij} \tag{1.13}$$

where $\tilde{P}^s_{ij}$ is symmetric and $\tilde{P}^a_{ij}$ is anti-symmetric and possibly different from P^s_{ij} and P^a_{ij} respectively. Interchanging i and j in (1.13) and using that $\tilde{P}^s_{ij}$ is symmetric and $\tilde{P}^a_{ij}$ is anti-symmetric, we obtain

$$P_{ji} = \tilde{P}^s_{ij} - \tilde{P}^a_{ij} \tag{1.14}$$

Addition and subtraction of (1.13) and (1.14) give

$$\tilde{P}^s_{ij} = \frac{1}{2}(P_{ij} + P_{ji})\ ; \qquad \tilde{P}^a_{ij} = \frac{1}{2}(P_{ij} - P_{ji})$$

A comparison with (1.11) and (1.12) shows that $\tilde{P}^s_{ij} = P^s_{ij}$ and $\tilde{P}^a_{ij} = P^a_{ij}$, i.e. the decomposition (1.10) of P_{ij} into a symmetric and a anti-symmetric part is unique.

A problem often encountered is the multiplication of a symmetric quantity A^s_{ij} with a quantity B_{ij} not necessarily symmetric. It turns out that

$$\boxed{A^s_{ij}B_{ij} = A^s_{ij}B^s_{ij}}$$

This result follows if we can prove that $A^s_{ij}B^a_{ij} = 0$. To show this we have

$$\begin{aligned} A^s_{ij}B^a_{ij} &= A^s_{ji}B^a_{ji} && \text{(redefinition of dummy indices)} \\ &= A^s_{ij}B^a_{ji} && \text{(symmetry of } A^s_{ij}\text{)} \\ &= -A^s_{ij}B^a_{ij} && \text{(anti-symmetry of } B^a_{ij}\text{)} \end{aligned}$$

which implies that $A^s_{ij}B^a_{ij} = 0$.

A so-called *comma convention* is also used in index notation. It states that whenever a quantity is differentiated with respect to the coordinates x_i, we use a comma to indicate this differentiation. Examples are

$$\boxed{\frac{\partial f}{\partial x_i} = f_{,i}\,; \qquad \frac{\partial a_i}{\partial x_j} = a_{i,j}}$$

We finally observe that in matrix notation we are restricted to working with one- and two-dimensional arrays. This is not the case in index notation where, for instance, the quantity e_{ijk} exists and comprises $3 \times 3 \times 3 = 27$ components. Likewise, the quantity D_{ijkl} exists and comprises $3 \times 3 \times 3 \times 3 = 81$ components.

1.4 Change of coordinate system

Before we present the concept of a tensor, we may first note that the essential issue of a tensor is that it behaves in a certain manner when a coordinate change is performed. Let us therefore first discuss coordinate changes between Cartesian coordinate systems.

Such a coordinate change can only occur in form of a *translation* and/or a *rotation* of the original coordinate system. Letting the old coordinates be denoted by x_i and the new coordinates by x'_i, Fig. 1.2 illustrates possible coordinate changes.

Consider first a translation of the coordinate system, Fig. 1.2a). It is obvious that we have

$$x_i = x'_i + c_i$$

i.e.

$$x'_i = x_i - c_i \qquad \text{or} \qquad \boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{c} \tag{1.15}$$

where c_i is constant and contains information of the translation of the old origin to the new origin. To be specific, c_i is the coordinates of the new origin measured

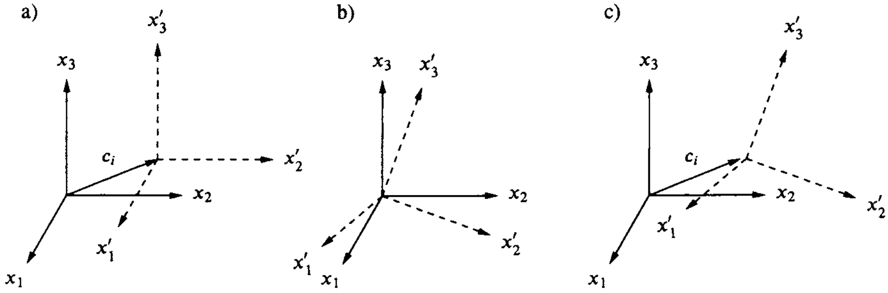

Figure 1.2: Coordinate changes; a) translation; b) rotation; c) translation and rotation.

in the old x_i-coordinate system. That is, if the coordinates x_i to a given fixed point are known in the old coordinate system, (1.15) provides the corresponding coordinates to that point in the new coordinate system.

Consider next a rotation of the coordinate system, Fig. 1.2b). In this case we expect x_1' to be a linear function of x_1, x_2 and x_3, i.e.

$$x_1' = A_{11}x_1 + A_{12}x_2 + A_{13}x_3$$

where A_{11}, A_{12} and A_{13} are certain constants. Likewise, for x_2' and x_3' we expect that

$$\begin{aligned} x_2' &= A_{21}x_1 + A_{22}x_2 + A_{23}x_3 \\ x_3' &= A_{31}x_1 + A_{32}x_2 + A_{33}x_3 \end{aligned}$$

Using index and matrix notation these expressions can be combined into

$$x_i' = A_{ij}x_j \qquad \text{or} \qquad \boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x} \tag{1.16}$$

where the *transformation matrix* $\boldsymbol{A}$ for a given rotation is constant, i.e. independent of the coordinates.

From (1.15) and (1.16) it follows that the most general coordinate transformation, which comprises a translation and a rotation, can be written as

$$\boxed{x_i' = A_{ij}(x_j - c_j) \qquad or \qquad \boldsymbol{x}' = \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{c})} \tag{1.17}$$

If $A_{ij} = \delta_{ij}$, this expression corresponds to a translation only and if $c_i = 0$ we only have a rotation of the coordinate system; finally, if $x_j = c_j$ then $x_i' = 0$ as expected. Since A_{ij} and c_i are constant quantities, it follows from (1.17) that

$$dx_i' = A_{ij}dx_j \tag{1.18}$$

The transformation matrix A_{ij} turns out to have a remarkable property. To see this, consider a point having the coordinates x_j and a neighboring point having the coordinates $x_j + dx_j$. According to Pythagoras' theorem, the distance ds between these two points is given by

$$ds^2 = (x_j + dx_j - x_j)(x_j + dx_j - x_j)$$

i.e.

$$ds^2 = dx_j dx_j = dx_j dx_k \delta_{kj} \tag{1.19}$$

The distance between two fixed points must be the same irrespective of the Cartesian coordinate system used. Therefore, in another coordinate system we have

$$ds^2 = dx'_i dx'_i = A_{ij} dx_j A_{ik} dx_k \tag{1.20}$$

where (1.18) has been used. Subtraction of (1.19) and (1.20) results in

$$(\delta_{jk} - A_{ij} A_{ik}) dx_j dx_k = 0$$

As this expression holds for arbitrary dx_j-values, we conclude that $A_{ij} A_{ik} = \delta_{jk}$, i.e.

$$\boxed{A_{ki} A_{kj} = \delta_{ij}} \tag{1.21}$$

In matrix notation, (1.21) reads

$$\boldsymbol{A}^T \boldsymbol{A} = \boldsymbol{I} \tag{1.22}$$

From (1.22) and (1.5), we conclude that the transformation matrix $\boldsymbol{A}$ is *orthogonal*, i.e.

$$\boxed{\boldsymbol{A}^T = \boldsymbol{A}^{-1}} \tag{1.23}$$

From (1.23) follows that

$$\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I} \tag{1.24}$$

which in index notation takes the form

$$\boxed{A_{ik} A_{jk} = \delta_{ij}} \tag{1.25}$$

The similarity in structure when compared with (1.21) should be noticed. However, we notice the different positions of the dummy index in (1.21) and (1.25) and we emphasize that the transformation matrix A_{ij} is, in general, unsymmetric.

Since $\det \boldsymbol{A} = \det \boldsymbol{A}^T$, we find from (1.24) and (1.6) that

$$\det(\boldsymbol{A}\boldsymbol{A}^T) = \det \boldsymbol{A} \cdot \det \boldsymbol{A}^T = (\det \boldsymbol{A})^2 = \det \boldsymbol{I} = 1$$

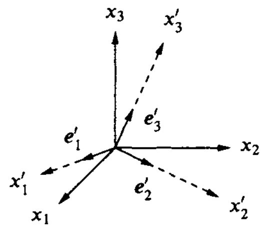

Figure 1.3: Rotation of coordinate system.

i.e. $\det \boldsymbol{A} = \pm 1$. For no rotation of the coordinate system, we have $\boldsymbol{A} = \boldsymbol{I}$, i.e. $\det \boldsymbol{A} = 1$ and of continuity reasons, we conclude that

$$\boxed{\det \boldsymbol{A} = 1} \tag{1.26}$$

holds for all right-handed coordinate systems.

Let us see, how in practice we may determine the components of the transformation matrix. For this purpose, consider the rotation of the coordinate system as shown in Fig. 1.3. Let $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$ denote unit vectors along the x_1'-, x_2'- and x_3'-axis respectively. The components of these unit vectors in the old x_i-system are given by

$$\boldsymbol{e}_1' = \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \end{bmatrix} \quad ; \quad \boldsymbol{e}_2' = \begin{bmatrix} e_{21} \\ e_{22} \\ e_{23} \end{bmatrix} \quad ; \quad \boldsymbol{e}_3' = \begin{bmatrix} e_{31} \\ e_{32} \\ e_{33} \end{bmatrix}$$

where the first index refers to the axis in the x_i'-system and the second index to the component measured in the x_i-system. According to (1.17), we have

$$\boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{x}' \tag{1.27}$$

It turns out that

$$\boxed{\boldsymbol{A}^T = \left[\ \boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\ \right] = \begin{bmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{bmatrix}} \tag{1.28}$$

To check this expression for $\boldsymbol{A}^T$, let us recall our definitions of $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$. Then, setting $\boldsymbol{x}' = [1\ \ 0\ \ 0]^T$ in (1.27) gives with (1.28) that $\boldsymbol{x} = \boldsymbol{e}_1'$, whereas $\boldsymbol{x}' = [0\ \ 1\ \ 0]^T$ gives $\boldsymbol{x} = \boldsymbol{e}_2'$ and $\boldsymbol{x}' = [0\ \ 0\ \ 1]^T$ provides $\boldsymbol{x} = \boldsymbol{e}_3'$. These results are in accordance with our definitions of $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$ and therefore prove the correctness of (1.28).

Considering $\boldsymbol{e}_1'$ we observe that the component e_{11} is $\cos\theta_{11}$, where θ_{11} is the angle between the x_1'-axis and the x_1-axis. Likewise, component $e_{12} = \cos\theta_{12}$ where θ_{12} is the angle between the x_1'-axis and the x_2-axis. Finally, component $e_{13} = \cos\theta_{13}$ where θ_{13} is the angle between the x_1'-axis and the x_3-axis. Similar interpretations hold for the components of $\boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$.

1.5 Cartesian tensors

Now we will present an elementary discussion of the concept of tensors and why they appear naturally when formulating physical relations. As we only use Cartesian coordinate systems, no difference exists between so-called covariant and contravariant tensors and therefore, by a tensor we always mean a *Cartesian tensor*. This simplifies the concept of a tensor significantly and for our purpose, there is no need to work with general tensors. For a discussion of general tensor analysis the reader is referred, for instance, to Malvern (1969), Segel (1987), Sokolnikoff (1951) or Spain (1965).

As previously mentioned, the essential issue of a tensor is that it behaves in a certain manner when a change of coordinate system is performed. We shall now establish this relation.

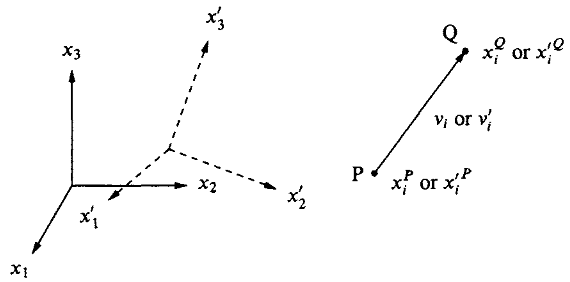

Figure 1.4: Vector from P to Q

We define a *vector* in the usual manner as a quantity having a length and a direction. In Fig. 1.4, the two fixed points P and Q have the coordinates x_i^P and x_i^Q in the old coordinate system and the coordinates $x_i'^P$ and $x_i'^Q$ in the new x_i'-coordinate system. The components of the vector v_i from P to Q in the old x_i-system are then given by

$$v_i = x_i^Q - x_i^P \tag{1.29}$$

where v_1, v_2 and v_3 are the components of the vector in the x_1-, x_2- and x_3-direction respectively. Likewise, the components of the vector v_i' from P to Q measured in the new x_i'-system are given by

$$v_i' = x_i'^Q - x_i'^P \tag{1.30}$$

where v_1', v_2' and v_3' are the components of the vector in the x_1'-, x_2'- and x_3'-direction respectively. From (1.17) we have

$$x_i'^Q = A_{ij}x_j^Q - c_i \ ; \qquad x_i'^P = A_{ij}x_j^P - c_i$$

Insertion into (1.30) and recognition of (1.29) result in

$$\boxed{v_i' = A_{ij}v_j \qquad or \qquad \boldsymbol{v}' = \boldsymbol{A}\boldsymbol{v}} \tag{1.31}$$

We have now established the important relation that shows how the components of a vector changes if a coordinate transformation is made. Here we have derived (1.31) from the usual definition of a vector, but we will now define a *quantity* v_i *as a vector if it transforms according to* (1.31). A vector is also called a *first-order tensor*, where first order refers to the fact that v_i only possesses one index. Now we have an indication of the statement expressed previously that tensors are quantities which behave in a certain manner when a coordinate change is performed.

It is of extreme importance that whereas any quantity containing three pieces of information can be written in the index form b_i, this does not make b_i a vector i.e. a first-order tensor as b_i will not, in general, transform according to (1.31). As an example, assume that $\boldsymbol{a}$ is a vector and consider the quantity $b_i = (|\boldsymbol{a}|, \theta_1, \theta_2)$ where $|\boldsymbol{a}|$ = the length of $\boldsymbol{a}$, θ_1 = the angle between $\boldsymbol{a}$ and the x_1-axis and θ_2 = the angle between $\boldsymbol{a}$ and the x_2-axis. In this case, b_i is certainly not a vector, since each of the components of b_i maintains its value irrespectively of the coordinate system, i.e. b_i does not fulfill the transformation rule (1.31). It is now apparent why we have chosen to use the name column matrix for **a** given by (1.1). Even though a vector a_i can be written in the same manner, the column matrix $\boldsymbol{a}$ is not necessary a vector.

Using (1.21) it is straightforward to invert (1.31) because multiplication by A_{ik} gives $v_k = A_{ik}v_i'$, i.e.

$$v_i = A_{ji}v_j' \qquad \text{or} \qquad \boldsymbol{v} = \boldsymbol{A}^T\boldsymbol{v}' \tag{1.32}$$

As indicated below, it is easy to show formally that velocity and acceleration vectors indeed are vectors.

Consider a specific particle of a body. This particle is described by its coordinates, which are functions of time, i.e. $x_i = x_i(t)$ where t is the time. The velocity components v_i are then defined by

$$v_i = \dot{x}_i \tag{1.33}$$

where a dot denotes the derivative with respect to time and v_1, v_2 and v_3 are the components of v_i in the x_1-, x_2- and x_3-direction respectively. Likewise, in a new coordinate system the velocity v_i' is defined by

$$v_i' = \dot{x}_i' \tag{1.34}$$

Differentiating (1.17) with respect to time and noting that A_{ij} and c_i are constants it follows directly from the definitions (1.33) and (1.34) that v_i transforms according to (1.31); i.e. the velocity v_i is indeed a vector. Differentiating (1.31) with respect to time and assuming that v_i is the velocity vector it appears that also the acceleration vector is, in fact, a vector.

As a force vector is defined as a quantity having a length and direction it follows in complete analogy with (1.29) and (1.30), which lead to (1.31), that a force vector is, in fact, a vector.

We have already touched upon quantities containing one piece of information and which take the same value irrespectively of the coordinate system. Such a quantity b is called a *scalar*, an *invariant* or a *zero-order* tensor and it transforms according to

$$\boxed{b' = b}$$

i.e. it takes the same value in the old coordinate system x_i and in the new coordinate system x'_i. A specific example of a scalar is the distance between two fixed points in the space as used in (1.19) and (1.20).

We have dwelt on the fact that tensors are quantities which transform in a particular manner when coordinate changes are made. It is now timely to ask why tensors are of relevance for our present purpose. The reason for this is of extraordinary importance, because it turns out that the relations of physics are conveniently expressed in terms of tensors. To illustrate this important aspect we write Newton's second law for a particle in the old coordinate system x_i according to

$$F_i = ma_i \tag{1.35}$$

where F_i is the force vector, m is the mass and a_i the acceleration vector. The vectors F_i and a_i are interpreted in the usual way that, for instance, F_2 is the component of F_i in the x_2-direction. When writing (1.35), we did not specify our coordinate system in any manner so in another coordinate system x'_i, we expect that Newton's second law takes the form $F'_i = m'a'_i$, i.e.

$$F'_i = ma'_i \tag{1.36}$$

where it has been assumed that the mass m is an invariant, i.e. independent of the coordinate system. As F_i and a_i are vectors, they transform according to (1.31) i.e. we have

$$F'_j = A_{ji}F_i \quad ; \qquad a'_j = A_{ji}a_i$$

Multiply (1.35) by A_{ji} and use the expressions above to obtain

$$F'_j = ma'_j$$

which is precisely the form stipulated in (1.36). It appears that irrespectively of the coordinate system, we write Newton's law in the same form, either (1.35) or (1.36), and this is possible only because F_i and a_i, in fact, are vectors, i.e. first-order tensors and because the mass m is an invariant, i.e. a zero-order tensor. Therefore, the occurrence of vectors and scalars in physical relations is a result of the fact that we expect physical laws to be independent of the particular coordinate system we choose to work with.

Above we illustrated that if a quantity like b_i appears in a physical relation, we expect it to be a vector. Let us pursue the argument above and assume that we have a physical relation which in the x_i-coordinate system states that

$$b_i = B_{ij}c_j \tag{1.37}$$

where b_i and c_i are assumed to be vectors and B_{ij} some quantity. When writing (1.37) we did not specify our coordinate system in any manner, so we expect that in another coordinate system x'_i the same physical relation is expressed through

$$b'_i = B'_{ij}c'_j \tag{1.38}$$

or

$$b'_k = B'_{kl}c'_l \tag{1.39}$$

Multiply (1.37) by A_{ki} and use (1.31) to obtain

$$b'_k = A_{ki}B_{ij}c_j$$

Transformation of c_j according to (1.32) yields

$$b'_k = A_{ki}B_{ij}A_{lj}c'_l \tag{1.40}$$

Subtraction of (1.39) and (1.40) provides

$$(B'_{kl} - A_{ki}B_{ij}A_{lj})c'_l = 0$$

This expression should hold for arbitrary c'_l-values and B_{ij} must therefore transform according to

$$\boxed{B'_{kl} = A_{ki}B_{ij}A_{lj} \qquad or \qquad B' = ABA^T} \tag{1.41}$$

We have found that if it is allowable to write a physical relation as (1.37) in one coordinate system and as (1.38) in another coordinate system - and we certainly expect this to be acceptable - then the quantity B_{ij} must transform according to (1.41). A quantity B_{ij}, which transforms according to (1.41), is defined to be a *second-order tensor*. It is obvious that whereas any square matrix containing 3x3 components can be written in index notation as B_{ij}, this does not make B_{ij} a second-order tensor. Only those B_{ij}-quantities, which transform according to (1.41), are second-order tensors.

We started with (1.37) where b_i and c_i were assumed to be vectors and B_{ij} some quantity. We then concluded that B_{ij} must be a second-order tensor, which transforms according to (1.41). This conclusion is an example of the use of the so-called *quotient theorem*.

Using (1.21) it is easy to invert (1.41). Multiplication of (1.41) by A_{km} gives

$$A_{km}B'_{kl} = B_{mj}A_{lj}$$

and multiplication by A_{ln} yields

$$B_{mn} = A_{km}B'_{kl}A_{ln}$$

which can be written as

$$B_{kl} = A_{ik}B'_{ij}A_{jl} \qquad \text{or} \qquad \boldsymbol{B} = \boldsymbol{A}^T\boldsymbol{B}'\boldsymbol{A} \tag{1.42}$$

Let us finally consider the following physical relation expressed in the x_i-coordinate system by

$$B_{ij} = D_{ijkl}M_{kl}$$

where B_{ij} and M_{kl} are assumed to be second-order tensors. In the x'_i-coordinate system we expect the relation

$$B'_{ij} = D'_{ijkl}M'_{kl}$$

If this is true, then by arguments like before, it is easily shown that the quantity D_{ijkl} must transform according to

$$\boxed{D'_{ijkl} = A_{im}A_{jn}D_{mnpq}A_{kp}A_{lq}} \tag{1.43}$$

Such a quantity is defined as a *fourth-order tensor* . Using (1.21) it follows in a straightforward manner that

$$D_{ijkl} = A_{mi}A_{nj}D'_{mnpq}A_{pk}A_{ql}$$

At this point it should be evident why tensors appear naturally when formulating physical relations. Thus it is not surprising that, in the following chapters, we will encounter a variety of tensors which characterize different physical phenomena. Indeed, if we start with tensors in a certain expression and then, after certain manipulations are left with quantities that are not tensors, we have a clear warning that something dubious has probably sneaked into our considerations. Although it is not a necessity to use a tensor formulation when dealing with physical phenomena, we have seen strong indications why it is very convenient to do so. Matrix formulations are often used instead of tensors, the main reason being that matrices are convenient when it comes to numerical computations. Often, tensors are used to derive the general relations governing the specific problem investigated and hereafter a corresponding matrix formulation is obtained from the tensor formulation in order to facilitate later numerical calculations. Later we will see applications of that approach.

To represent tensors both a component representation, i.e. index notation, as well as so-called *index free notation* can be used, cf. Gurtin (1981). It turns out that the index representation of tensors as presented above is very attractive when manipulating various expressions. The index free notation of tensors - also called *direct notation* - makes no reference to the base vectors of the coordinate system in question and thereby it makes no reference to components. The index free notation is therefore advantageous when representing general concepts. As an example, consider a tensor, which in index notation is written as A_{ij}. We may also adopt the matrix notation for the same quantity and we then denote it by $\boldsymbol{A}$. If an index free formulation is adopted for the tensor in question, the standard notation would also be $\boldsymbol{A}$. It appears that the same notation $\boldsymbol{A}$ is used in the literature both for the matrix notation and for the free index notation of tensors. Since matrices are not necessarily tensors, this implies a certain ambiguity in the interpretation. To avoid this problem, we will, in the present text, avoid the use of index free notation and all boldface letters, like $\boldsymbol{A}$, shall be viewed as matrices.

1.6 Example of tensors - Isotropic tensors

We have previously defined an invariant as a quantity that takes the same value in all coordinate systems. It is of considerable interest that if a quantity is known to be a tensor, it is then possible to establish various invariants from this tensor. As an example, assume that B_{ij} is a second-order tensor. As B_{ij} then obeys the transformation rule (1.41), we obtain by contraction and use of (1.21) that

$$B'_{kk} = A_{ki}B_{ij}A_{kj} = \delta_{ij}B_{ij} = B_{ii}$$

i.e. the quantity B_{ii} is an invariant. We shall later see that other invariants may be obtained from a second-order tensor.

We have previously introduced Kronecker's δ_{ij}, cf. (1.8), and we may ask whether δ_{ij} is a second-order tensor. When defining δ_{ij} no reference was made to any coordinate system, i.e. we must have that $\delta'_{ij} = \delta_{ij}$. With (1.25), it then follows that

$$\delta'_{ij} = \delta_{ij} = A_{ik}A_{jk} = A_{ik}\delta_{kl}A_{jl}$$

It appears that (1.41) is fulfilled, which means that δ_{ij} is a second-order tensor. In fact, it is called an *isotropic tensor*, because it takes the same value irrespectively of the coordinate system.

Let us now derive the most general isotropic second-order tensor. From (1.41) and imposing that B_{ij} is isotropic, we have

$$B_{kl} = A_{ki}B_{ij}A_{lj} \tag{1.44}$$

Multiplication by A_{km} and using that $A_{km}A_{ki} = \delta_{mi}$ gives

$$A_{km}B_{kl} = B_{mj}A_{lj} \qquad \text{i.e} \qquad A_{kj}\delta_{jm}B_{kl} = B_{mj}A_{kj}\delta_{kl}$$

We then obtain

$$(\delta_{mj}B_{kl} - B_{mj}\delta_{kl})A_{kj} = 0$$

Since the transformation matrix A_{kj} is arbitrary, it follows that the expression within the parenthesis must be equal to zero, i.e. $\delta_{mj}B_{kl} - B_{mj}\delta_{kl} = 0$. To evaluate this expression, choose $j = m = 1$; it then follows that $B_{kl} - B_{11}\delta_{kl} = 0$. Likewise, for $j = m = 2$ we obtain $B_{kl} - B_{22}\delta_{kl} = 0$, whereas $j = m = 3$ gives $B_{kl} - B_{33}\delta_{kl} = 0$. We then conclude that

The most general isotropic second-order tensor is given by $k\delta_{ij}$ where k is an arbitrary invariant

(1.45)

We will next demonstrate that differentiation, multiplication, addition and subtraction of tensors lead to quantities that also are tensors.

Suppose that f is an invariant, then in applications we will often encounter the quantity $\partial f/\partial x_i = f_{,i}$; this quantity is called the *gradient* of f and we shall prove that it is a vector, i.e. a first-order tensor. From (1.17) and since A_{ij} and c_i are constant quantities, we obtain

$$\frac{\partial x_i'}{\partial x_k} = A_{ik} \tag{1.46}$$

We then obtain

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_j'}\frac{\partial x_j'}{\partial x_i} = A_{ji}\frac{\partial f}{\partial x_j'}$$

and a comparison with (1.32) proves that $\partial f/\partial x_i = f_{,i}$ is a vector, i.e. a first-order tensor.

Likewise, considering the quantity $\partial f/\partial B_{ij}$ where f is an invariant and B_{ij} is a second-order tensor, we shall prove that $\partial f/\partial B_{ij}$ is a second-order tensor. We have

$$\frac{\partial f}{\partial B_{ij}} = \frac{\partial f}{\partial B_{kl}'}\frac{\partial B_{kl}'}{\partial B_{ij}}$$

As B_{ij} is a second-order tensor, (1.41) implies that

$$\frac{\partial B_{kl}'}{\partial B_{ij}} = A_{km}\frac{\partial B_{mn}}{\partial B_{ij}}A_{ln} = A_{km}\delta_{mi}\delta_{nj}A_{ln} = A_{ki}A_{lj}$$

i.e.

$$\frac{\partial f}{\partial B_{ij}} = A_{ki}\frac{\partial f}{\partial B_{kl}'}A_{lj}$$

A comparison with (1.41) proves that $\partial f/\partial B_{ij}$ is a second- order tensor.

Consider now the quantity $\partial v_i/\partial x_j = v_{i,j}$ where v_i is a vector. According to (1.32) we have

$$\frac{\partial v_i}{\partial x_j} = A_{ki}\frac{\partial v'_k}{\partial x_j} = A_{ki}\frac{\partial v'_k}{\partial x'_l}\frac{\partial x'_l}{\partial x_j}$$

and use of (1.46) implies that

$$\frac{\partial v_i}{\partial x_j} = A_{ki}\frac{\partial v'_k}{\partial x'_l}A_{lj} \tag{1.47}$$

A comparison with (1.41) proves that $\partial v_i/\partial x_j = v_{i,j}$ is a second-order tensor.

For the often encountered quantity $\partial B_{ij}/\partial x_j$ where B_{ij} is a second-order tensor, we find according to (1.42) that

$$\frac{\partial B_{ij}}{\partial x_j} = A_{ki}\frac{\partial B'_{kl}}{\partial x_j}A_{lj} = A_{ki}\frac{\partial B'_{kl}}{\partial x'_m}\frac{\partial x'_m}{\partial x_j}A_{lj}$$

and use of (1.46) implies that

$$\frac{\partial B_{ij}}{\partial x_j} = A_{ki}\frac{\partial B'_{kl}}{\partial x'_m}A_{mj}A_{lj}$$

From expression (1.25) we conclude that

$$\frac{\partial B_{ij}}{\partial x_j} = A_{ki}\frac{\partial B'_{kl}}{\partial x'_l}$$

and a comparison with (1.32) reveals that $\partial B_{ij}/\partial x_j = B_{ij,j}$ is a vector, i.e. a first-order tensor.

Finally, consider the quantity $a_i b_j$ where a_i and b_j are assumed to be vectors. As a_i and b_j are vectors, it follows from (1.32) that

$$a_i b_j = A_{ki}a'_k b'_l A_{lj} \tag{1.48}$$

which, according to (1.42), shows that $a_i b_j$ is a second-order tensor. Occasionally, the product $a_i b_j$ is called a *dyad* and the sum of dyads is termed a *dyadic*. Analogous with the derivation of (1.48), it follows directly that products of tensors result in the creation of new tensors. As an example, the product $a_i B_{ij}$ is a vector, if a_i and B_{ij} is a vector and a second-order tensor respectively. Moreover, it follows directly that the sum or difference of two tensors results in a new tensor.

Because of the transformation properties, a tensor is known completely in all coordinate systems if it is known in one of them. In particular, if all the components vanish in one system, they vanish in all. This seemingly trivial statement is helpful in various mathematical proofs. As an example, it will be shown later that static equilibrium requires

$$\sigma_{ij,j} + b_i = 0$$

where σ_{ij} is the stress tensor and b_i the body forces per unit volume. Defining the quantity D_i as $D_i = \sigma_{ij,j} + b_i$, we recognize that D_i is a first-order tensor, i.e. a vector; equilibrium is then expressed by $D_i = 0$. Therefore, if the vector $D_i = 0$ in one coordinate system it is zero in all coordinate systems. That is, if the body is in equilibrium in one coordinate system, it is unnecessary to reinvestigate equilibrium in any other coordinate system.

After these preliminary remarks about notations and tensors and as the reader should now be familiar with index notation, it is timely to proceed with what is our main interest: *the behavior of materials*. The first thing of relevance is the ability to describe the deformation of the material in a proper manner and this is the subject of the next chapter.