A. Elementary Tensor Analysis

This appendix is intended to provide a survey of mathematical background needed for a modern development of continuum mechanics. The reader is expected to be familiar with some notions of vector spaces or matrix algebra. In the first part we shall review some basic notions of vector spaces and linear transformations. At the same time, elementary properties of tensors as well as tensor notations will be introduced.

A.1 Linear Algebra

We shall consider finite dimensional real vector spaces only. The field of real numbers is denoted by $I\!R$.

Definition. *A vector space V is a set equiped with two operations:*
1) $\boldsymbol{v}+\boldsymbol{u}\in V$, *called addition of* $\boldsymbol{v}$ *and* $\boldsymbol{u}$ *in* V,
2) $\alpha\boldsymbol{v}\in V$, *called scalar multiplication of* $\boldsymbol{v}\in V$ *by* $\alpha\in I\!R$,

which satisfy the following rules: for any $\boldsymbol{v},\boldsymbol{u},\boldsymbol{w}\in V$, *and any* $\alpha,\beta\in I\!R$,
1) $\boldsymbol{v}+\boldsymbol{u}=\boldsymbol{u}+\boldsymbol{v}$.
2) $\boldsymbol{v}+(\boldsymbol{u}+\boldsymbol{w})=(\boldsymbol{v}+\boldsymbol{u})+\boldsymbol{w}$.
3) *There exists a null vector* $0\in V$, *such that* $\boldsymbol{v}+0=\boldsymbol{v}$.
4) *For any* $\boldsymbol{v}\in V$, *there exist* $-\boldsymbol{v}\in V$, *such that* $\boldsymbol{v}+(-\boldsymbol{v})=0$.
5) $\alpha(\beta\boldsymbol{v})=(\alpha\beta)\boldsymbol{v}$.
6) $(\alpha+\beta)\boldsymbol{v}=\alpha\boldsymbol{v}+\beta\boldsymbol{v}$.
7) $\alpha(\boldsymbol{v}+\boldsymbol{u})=\alpha\boldsymbol{v}+\alpha\boldsymbol{u}$.
8) $1\boldsymbol{v}=\boldsymbol{v}$.

Definition. *A set of vectors* $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ *is said to be a basis of* V, *if*
1) *it is a linearly independent set, i.e., for any* $a_1,\cdots,a_n\in I\!R$, *if* $a_1\boldsymbol{v}_1+\cdots+a_n\boldsymbol{v}_n=0$ *then* $a_1=\cdots=a_n=0$.
2) *it spans the space* V, *i.e., for any* $\boldsymbol{u}\in V$, *the vector* $\boldsymbol{u}$ *can be expressed as a linear combination of* $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$.

Let $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ be a basis of V, then for any vector $\boldsymbol{u}\in V$, $\boldsymbol{u}$ can be expressed as

$$\boldsymbol{u}=\sum_{i=1}^{n}u^i\boldsymbol{e}_i,$$

where u^i, called the components of $\boldsymbol{u}$, are uniquely determined relative to the basis $\{\boldsymbol{e}_i\}$.

A vector space can have many different bases, but all of them will have the same number of elements. The number of elements in a basis is called the *dimension* of the space, in this case, we have $\dim V = n$.

A.1.1 Inner Product

We may think of a vector as a geometric object which has a length and pointing in a certain direction. To incorporate this notion we introduce an additional structure, inner product, into the vector space.

Definition. *An inner product is a map*

$$g : V \times V \to I\!R$$

with the following properties: For any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, *and* $\alpha \in I\!R$,
1) $g(\boldsymbol{u} + \alpha\boldsymbol{v}, \boldsymbol{w}) = g(\boldsymbol{u}, \boldsymbol{w}) + \alpha g(\boldsymbol{v}, \boldsymbol{w})$,
2) $g(\boldsymbol{u}, \boldsymbol{v}) = g(\boldsymbol{v}, \boldsymbol{u})$,
3) $g(\boldsymbol{u}, \boldsymbol{u}) > 0$, *if* $\boldsymbol{u} \neq 0$.

That is, an inner product is a positive-definite symmetric bilinear function on V. We call $g(\boldsymbol{u}, \boldsymbol{v})$ the inner product of $\boldsymbol{u}$ and $\boldsymbol{v}$. The vector space equipped with an inner product is called an *inner product space*. Hereafter, all vector spaces considered are always inner product spaces.

Notation. $g(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{v}$, if g is given and fixed.

Definition. *The norm of a vector* $\boldsymbol{v} \in V$ *is defined as*

$$|\boldsymbol{v}| = \sqrt{(\boldsymbol{v} \cdot \boldsymbol{v})}.$$

A vector space equipped with such a norm is called a Euclidean vector space.

The notion of angle between two vectors can be defined based on the following Schwarz inequality:

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq |\boldsymbol{u}|\,|\boldsymbol{v}|\,. \tag{A.1}$$

Definition. *For any non-zero* $\boldsymbol{u}, \boldsymbol{v} \in V$, *the angle between* $\boldsymbol{u}$ *and* $\boldsymbol{v}$, $\theta(\boldsymbol{u}, \boldsymbol{v}) \in [0, \pi]$, *is defined by*

$$\cos\theta(\boldsymbol{u}, \boldsymbol{v}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}|\,|\boldsymbol{v}|}.$$

The vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are said to be *orthogonal* if $\theta(\boldsymbol{u}, \boldsymbol{v}) = \pi/2$. Obviously, $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if and only if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

A vector $\boldsymbol{v}$ is called a *unit vector* if $|\boldsymbol{v}| = 1$. The *projection* of a vector $\boldsymbol{u}$ on the vector $\boldsymbol{v}$ can be defined as $|\boldsymbol{u}|\cos\theta(\boldsymbol{u}, \boldsymbol{v})$, or as $(\boldsymbol{u} \cdot \boldsymbol{e})$, where $\boldsymbol{e} = \boldsymbol{v}/\,|\boldsymbol{v}|$ is the unit vector in the direction of $\boldsymbol{v}$. The vector $(\boldsymbol{u} \cdot \boldsymbol{e})\,\boldsymbol{e}$ is called the projection vector of $\boldsymbol{u}$ in the direction of $\boldsymbol{v}$.

Let $\{\boldsymbol{e}_i,\ i = 1, \cdots, n\}$ be a basis of V. Denote the inner product of $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ by g_{ij},

$$g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j.$$

Clearly, g_{ij} is symmetric, $g_{ij} = g_{ji}$. Let $\boldsymbol{u} = u^i \boldsymbol{e}_i$, $\boldsymbol{v} = v^j \boldsymbol{e}_j$ be arbitrary vectors in V expressed in terms of the basis $\{\boldsymbol{e}_i\}$. Then

$$\begin{aligned}\boldsymbol{u} \cdot \boldsymbol{v} &= (u^i \boldsymbol{e}_i) \cdot (v^j \boldsymbol{e}_j) \\ &= u^i v^j (\boldsymbol{e}_i \cdot \boldsymbol{e}_j) = u^i v^j g_{ij},\end{aligned}$$

or

$$\boldsymbol{u} \cdot \boldsymbol{v} = g_{ij} u^i v^j. \qquad (A.2)$$

Here we have used the following summation convention.

Notation. (*Summation convention*) In the expression of a term, if an index is repeated once (and only once), a summation over the range of this index is assumed.

For example,

$$u^i \boldsymbol{e}_i = \sum_{i=1}^{n} u^i \boldsymbol{e}_i,$$

$$g_{ij} u^i v^j = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} u^i v^j.$$

Note that in these expressions, we purposely write the indices in two different levels so that the repeated summation indices are always one superindex and one subindex. The reason for doing so will become clear in the next section.

A.1.2 Dual Bases

Let $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ be a basis of V. There exists a non-zero vector orthogonal to the plane spanned by the $n-1$ vectors $\{\boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$, and if in addition the projection of this vector on $\boldsymbol{e}_1$ is prescribed, then this vector is uniquely determined. In this manner, for any given basis $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$, we can construct a set of vectors $\{\boldsymbol{e}^1, \cdots, \boldsymbol{e}^n\}$ such that

$$\boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i{}_j,$$

where $\delta^i{}_j$ is called the *Kronecker delta* defined by

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

From this construction, if $\boldsymbol{v} = v^i \boldsymbol{e}_i$ is a vector in V, then by taking the inner product with $\boldsymbol{e}^i$ we have

$$\boldsymbol{e}^i \cdot \boldsymbol{v} = \boldsymbol{e}^i \cdot (v^j \boldsymbol{e}_j) = v^j \delta^i{}_j = v^i.$$

Hence the i^{th} component of $\boldsymbol{v}$ relative to the basis $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ is its inner product with the vector $\boldsymbol{e}^i$. Therefore this set of vectors $\{\boldsymbol{e}^i\}$ associated with the basis $\{\boldsymbol{e}_i\}$ can be regarded as linear functions which map a vector to its components.[1]

We can easily show that this new set of vectors is a linearly independent set. Indeed, if for any linear combination $a_j \boldsymbol{e}^j = 0$, then it follows that $(a_j \boldsymbol{e}^j) \cdot \boldsymbol{e}_i = a_j \delta^j{}_i = a_i = 0$ for all i. Furthermore, it also spans the space V, for if $\boldsymbol{u} = u^i \boldsymbol{e}_i$ is a vector in V, then for any vector $\boldsymbol{v} = v^i \boldsymbol{e}_i$, from (A.2) and $v^j = \boldsymbol{e}^j \cdot \boldsymbol{v}$,

$$\boldsymbol{u} \cdot \boldsymbol{v} = g_{ij} u^i v^j = (g_{ij} u^i \boldsymbol{e}^j) \cdot \boldsymbol{v},$$

which implies that $\boldsymbol{u}$ can be expressed as $\boldsymbol{u} = u_i \boldsymbol{e}^i$ with

$$u_i = g_{ij} u^j.$$

Therefore we have proved that this new set of vectors $\{\boldsymbol{e}^i\}$ is also a basis of V.

Definition. *Let $\beta = \{\boldsymbol{e}_i\}$ and $\beta^* = \{\boldsymbol{e}^i\}$ be two bases of V related by the property*

$$\boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i{}_j.$$

They are said to be a pair of dual bases for V, or β^* is the dual basis of β.

The dual bases are uniquely determined from each other. For this reason, we have used the same notation for their elements except the different level of indices to distinguish them. Clearly, if $\boldsymbol{u}$ is a vector in V, then we can express $\boldsymbol{u}$ in terms of components in two different ways relative to the dual bases,

$$\boldsymbol{u} = u^i \boldsymbol{e}_i = u_j \boldsymbol{e}^j,$$

where we have also employed different level of component indices in order to be consistent with our summation convention, which sums over repeated indices in different levels. We call

u^i	the i^{th} *contravariant* component of $\boldsymbol{u}$,
u_j	the j^{th} *covariant* component of $\boldsymbol{u}$.

From the definition, it follows that

$$u^i = \boldsymbol{e}^i \cdot \boldsymbol{u}, \qquad u_j = \boldsymbol{e}_j \cdot \boldsymbol{u}, \tag{A.3}$$

and they are related by

$$u_i = g_{ij} u^j, \qquad u^i = g^{ij} u_j,$$

[1] In general, the space of all linear functions on V is called the dual space of V and denoted by V^*. In this note, for simplicity, we shall not distinguish vectors in V^* and V through the inner product.

where we have denoted

$$g^{ij} = \boldsymbol{e}^i \cdot \boldsymbol{e}^j.$$

The two operations

$$g_{ij} : u^j \mapsto u_i, \qquad g^{ij} : u_j \mapsto u^i,$$

enable us to *lower* and *raise* the component index. One can also show that

$$\boldsymbol{e}_j = g_{ij}\boldsymbol{e}^i, \qquad \boldsymbol{e}^i = g^{ij}\boldsymbol{e}_j.$$

Therefore, lowering or raising the index for dual bases can be made in the same manner. It is easy to verify that $[g^{ij}]$ is the inverse of the matrix $[g_{ij}]$, or

$$g^{ij}g_{jk} = \delta^i{}_k.$$

A basis $\{\boldsymbol{e}_i\}$ is called an *orthogonal basis* if all the elements of the basis are mutually orthogonal, i.e.,

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0 \quad \text{if} \quad i \neq j.$$

If in addition, $|\boldsymbol{e}_i| = 1$, for all i, it is called an *orthonormal basis.* Although in general, we carefully do our bookkeeping of super- and sub-indices, this becomes unnecessary if $\beta = \{\boldsymbol{e}_i\}$ is an orthonormal basis. Since then $g_{ij} = \delta_{ij}$, and

$$\boldsymbol{e}_i = g_{ij}\boldsymbol{e}^j = \delta_{ij}\boldsymbol{e}^j = \boldsymbol{e}^i.$$

That is, the basis β is identical to its dual basis β^*. Hence we do not have to distinguish contravariant and covariant components. In this case, we can write all the indices at the same level. for example,

$$\boldsymbol{v} = v_i\boldsymbol{e}_i.$$

Of course, according to our summation convention, we still sum over the repeated indices (now in the same level) in this situation.

Exercise A.1.1 Let $\beta' = \{\boldsymbol{e}_1 = (1,0), \boldsymbol{e}_2 = (2,1)\}$ be a basis of $\mathbb{R}^2$, and $\boldsymbol{v} = (1,-1)$ be a vector in $\mathbb{R}^2$.

1) Find the dual basis $\{\boldsymbol{e}^1, \boldsymbol{e}^2\}$ of β'.
2) Determine the matrix representations $[g_{ij}]$ and $[g^{ij}]$ relative to β'.
3) Determine the contravariant and covariant components of $\boldsymbol{v}$ relative to the bases and make a graphic representation of the results.

A.1.3 Tensor Product

The notion of matrix is related to linear functions on vector spaces. Let U and V be two vector spaces with inner product. A function $T : U \to V$, is called a *linear transformation* from U to V, if for any $\boldsymbol{u}, \boldsymbol{v} \in U$ and $\alpha \in I\!R$,

$$T(\boldsymbol{u} + \alpha \boldsymbol{v}) = T(\boldsymbol{u}) + \alpha T(\boldsymbol{v}).$$

Notation. $\mathcal{L}(U, V) = \{T : U \to V \mid T \text{ is linear}\}$.

If T and S are two linear transformations in $\mathcal{L}(U, V)$, we can define the addition $T + S$ and the scalar multiplication αT, as transformations in $\mathcal{L}(U, V)$, in the following manner, for all $\boldsymbol{v} \in U$,

$$\begin{aligned}(T + S)(\boldsymbol{v}) &= T(\boldsymbol{v}) + S(\boldsymbol{v}),\\ (\alpha T)(\boldsymbol{v}) &= \alpha T(\boldsymbol{v}).\end{aligned}$$

With these operations the set $\mathcal{L}(U, V)$ becomes a vector space.

Definition. *For any vectors $\boldsymbol{v} \in V$ and $\boldsymbol{u} \in U$, the tensor product of $\boldsymbol{v}$ and $\boldsymbol{u}$, denoted by $\boldsymbol{v} \otimes \boldsymbol{u}$, is defined as a linear transformation from U to V such that*

$$(\boldsymbol{v} \otimes \boldsymbol{u})(\boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v}, \tag{A.4}$$

for any $\boldsymbol{w} \in U$.

The tensor product of two vectors is a linear transformation. We call such a linear transformation a *simple tensor*. Of course, not every linear transformation can be obtained as a tensor product of two vectors. However, we can show that, indeed, it can always be expressed as a linear combination of simple tensors.

Proposition. *Let $\{\boldsymbol{e}_i\}$, $i = 1, \cdots, n$ and $\{\boldsymbol{d}_\alpha\}$, $\alpha = 1, \cdots, m$ be bases of V and U respectively. Then the set $\{\boldsymbol{e}_i \otimes \boldsymbol{d}_\alpha\}$, $i = 1, \cdots, n$, $\alpha = 1, \cdots, m$, forms a basis of $\mathcal{L}(U, V)$.*

Proof: Let $\{\boldsymbol{e}^i\}$ be the dual basis of $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{d}^\alpha\}$ the dual of $\{\boldsymbol{d}_\alpha\}$. If $a^{i\alpha}\boldsymbol{e}_i \otimes \boldsymbol{d}_\alpha = 0$, then

$$a^{i\alpha}(\boldsymbol{e}_i \otimes \boldsymbol{d}_\alpha)(\boldsymbol{d}^\beta) = a^{i\alpha}(\boldsymbol{d}_\alpha \cdot \boldsymbol{d}^\beta)\boldsymbol{e}_i = a^{i\alpha}\delta_\alpha^\beta \boldsymbol{e}_i = a^{i\beta}\boldsymbol{e}_i = 0,$$

which implies that $a^{i\beta} = 0$ since $\{\boldsymbol{e}_i\}$ is a basis. Therefore, $\{\boldsymbol{e}_i \otimes \boldsymbol{d}_\alpha\}$ is a linearly independent set. Moreover, for any $T \in \mathcal{L}(U, V)$, let

$$\boldsymbol{e}^i \cdot T(\boldsymbol{d}^\alpha) = T^{i\alpha}.$$

Then for any $\boldsymbol{v} \in V$ and any $\boldsymbol{u} \in U$,

$$\begin{aligned}\boldsymbol{v} \cdot T(\boldsymbol{u}) &= v_i \boldsymbol{e}^i \cdot T(u_\alpha \boldsymbol{d}^\alpha)\\ &= v_i u_\alpha \boldsymbol{e}^i \cdot T(\boldsymbol{d}^\alpha) = T^{i\alpha} v_i u_\alpha.\end{aligned}$$

On the other hand,

$$\begin{aligned}\boldsymbol{v}\cdot(\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha)(\boldsymbol{u}) &= v_j\boldsymbol{e}^j\cdot(\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha)(u_\beta\boldsymbol{d}^\beta)\\ &= v_ju_\beta(\boldsymbol{e}^j\cdot\boldsymbol{e}_i)(\boldsymbol{d}_\alpha\cdot\boldsymbol{d}^\beta)=v_iu_\alpha.\end{aligned}$$

Therefore, we have

$$\boldsymbol{v}\cdot T(\boldsymbol{u}) = T^{i\alpha}\boldsymbol{v}\cdot(\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha)(\boldsymbol{u}),$$

for any $\boldsymbol{v}$ and any $\boldsymbol{u}$, which leads to

$$T = T^{i\alpha}\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha.$$

That is, $\{\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha\}$ spans the space $\mathcal{L}(U,V)$. □

We may call $\mathcal{L}(U,V)$ the *tensor product space* of V and U and denote it by $V\otimes U$. Obviously, from this result, we have

$$\dim V\otimes U = (\dim V)(\dim U).$$

The basis $\{\boldsymbol{e}_i\otimes\boldsymbol{d}_\alpha\}$ is called a *product basis* of $V\otimes U$. Similarly, the sets $\{\boldsymbol{e}_i\otimes\boldsymbol{d}^\alpha\}$, $\{\boldsymbol{e}^i\otimes\boldsymbol{d}_\alpha\}$, and $\{\boldsymbol{e}^i\otimes\boldsymbol{d}^\alpha\}$ are also product bases of $V\otimes U$.

Notation. $V\otimes V=\mathcal{L}(V)=\mathcal{L}(V,V)$.

We shall call linear transformations in $\mathcal{L}(V)$ the *second order tensors*. Let $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{e}^j\}$ be dual bases of V, a second order tensor T then has different component forms relative to the different product bases.

$$\begin{aligned}T &= T^{ij}\boldsymbol{e}_i\otimes\boldsymbol{e}_j = T^i{}_j\boldsymbol{e}_i\otimes\boldsymbol{e}^j\\ &= T_i{}^j\boldsymbol{e}^i\otimes\boldsymbol{e}_j = T_{ij}\boldsymbol{e}^i\otimes\boldsymbol{e}^j,\end{aligned}$$

where the various components are given by

$$\begin{aligned}T^{ij} &= \boldsymbol{e}^i\cdot T\boldsymbol{e}^j, & T^i{}_j &= \boldsymbol{e}^i\cdot T\boldsymbol{e}_j,\\ T_{ij} &= \boldsymbol{e}_i\cdot T\boldsymbol{e}_j, & T_i{}^j &= \boldsymbol{e}_i\cdot T\boldsymbol{e}^j.\end{aligned}\tag{A.5}$$

They are related by

$$T^i{}_j = T^{ik}g_{kj} = g^{ik}T_{kj}, \quad etc., \tag{A.6}$$

with the operations of raising or lowering the indices discussed in the previous section.

The matrices $[T^{ij}]$, $[T^i{}_j]$, $[T_i{}^j]$, $[T_{ij}]$ are called the *matrix representations* of T relative to the corresponding product bases. Note that the first index refers to the row and the second index refers to the column of the matrix. It is important to distinguish the level as well as the position order of the component indices. In general $T^i{}_j\neq T_j{}^i$, therefore it may cause some confusions to write T^i_j with i and j at the same position one on top of the other. The relation (A.6) can be written in terms of matrix multiplication, in which the column of the first matrix is summed against the row of the second matrix,

$$[T^i{}_j] = [T^{ik}][g_{kj}] = [g^{ik}][T_{kj}].$$

These components are called the *associated components* of the second order tensor T. In classical tensor analysis, they are also called

T^{ij}	contravariant tensor of order 2,
T_{ij}	covariant tensor of order 2,
$T^i{}_j$, $T_i{}^j$	mixed tensor of order 2.

Note that if $S, T \in \mathcal{L}(V)$, then the *composition* $S \circ T$, defined as $S \circ T(\boldsymbol{v}) = S(T(\boldsymbol{v}))$ for all $\boldsymbol{v} \in V$, is also in $\mathcal{L}(V)$. The composition $S \circ T$ will be more conveniently denoted by ST. In terms of components and matrix operation, we have

$$[(ST)^i{}_j] = [S^i{}_k T^k_j] = [S^i{}_k][T^k_j].$$

Example A.1.1 The identity transformation, $1\boldsymbol{v} = \boldsymbol{v}$ for any $\boldsymbol{v}$ in V, has the components,

$$1 = \delta^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j = \delta_i{}^j \boldsymbol{e}^i \otimes \boldsymbol{e}_j = g^{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j = g_{ij} \boldsymbol{e}^i \otimes \boldsymbol{e}^j, \qquad (A.7)$$

since by (A.5), we have

$$1^i{}_j = \boldsymbol{e}^i \cdot 1\boldsymbol{e}_j = \boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i{}_j,$$
$$1_{ij} = \boldsymbol{e}_i \cdot 1\boldsymbol{e}_j = \boldsymbol{e}_i \cdot \boldsymbol{e}_j = g_{ij}.$$

Therefore, the Kronecker deltas are the mixed components of the identity tensor, while g_{ij} and g^{ij} are just its covariant and contravariant components. □

Example A.1.2 For $\boldsymbol{v} = v^i \boldsymbol{e}_i$ and $\boldsymbol{u} = u^i \boldsymbol{e}_i$ in V, their tensor product has the component form:

$$\boldsymbol{v} \otimes \boldsymbol{u} = v^i u^j \boldsymbol{e}_i \otimes \boldsymbol{e}_j.$$

Let $\boldsymbol{v} = (v_1, v_2)$ and $\boldsymbol{u} = (u_1, u_2)$ be two vectors in $I\!R^2$, then relative to the standard basis of $I\!R^2$, the matrix of $\boldsymbol{v} \otimes \boldsymbol{u}$ is given by

$$[(\boldsymbol{v} \otimes \boldsymbol{u})] = \begin{bmatrix} v_1 u_1 & v_1 u_2 \\ v_2 u_1 & v_2 u_2 \end{bmatrix}.$$

This product is sometimes referred to as *dyadic product* of vectors $\boldsymbol{v}$ and $\boldsymbol{u}$. □

In general, the tensor products $\boldsymbol{v} \otimes \boldsymbol{u}$ and $\boldsymbol{u} \otimes \boldsymbol{v}$ belong to two different spaces, namely $V \otimes U$ and $U \otimes V$ respectively. Even in the case $V = U$, by definition, $\boldsymbol{v} \otimes \boldsymbol{u}$ and $\boldsymbol{u} \otimes \boldsymbol{v}$ are different, i.e., the tensor product is not symmetric.

Definition. *For $A \in V \otimes U$, the transpose of A, denoted by A^T, is defined as a tensor in $U \otimes V$ such that*

$$\boldsymbol{v} \cdot A\boldsymbol{u} = \boldsymbol{u} \cdot A^T\boldsymbol{v}, \tag{A.8}$$

for any $\boldsymbol{v} \in V$ and any $\boldsymbol{u} \in U$.

Example A.1.3 For simple tensors, it follows that

$$(\boldsymbol{v} \otimes \boldsymbol{u})^T = \boldsymbol{u} \otimes \boldsymbol{v},$$

because for any $\boldsymbol{w}_1, \boldsymbol{w}_2 \in V$, we have

$$\begin{aligned} \boldsymbol{w}_1 \cdot (\boldsymbol{v} \otimes \boldsymbol{u})^T \boldsymbol{w}_2 &= \boldsymbol{w}_2 \cdot (\boldsymbol{v} \otimes \boldsymbol{u})\boldsymbol{w}_1 \\ &= (\boldsymbol{w}_2 \cdot \boldsymbol{v})(\boldsymbol{u} \cdot \boldsymbol{w}_1) = \boldsymbol{w}_1 \cdot (\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{w}_2. \end{aligned}$$

□

Example A.1.4 We have

$$A(\boldsymbol{u} \otimes \boldsymbol{v}) = A\boldsymbol{u} \otimes \boldsymbol{v}, \qquad (\boldsymbol{u} \otimes \boldsymbol{v})A = \boldsymbol{u} \otimes A^T\boldsymbol{v}.$$

Indeed, for any vector $\boldsymbol{w} \in V$, we obtain

$$A(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{w} = A\boldsymbol{u}(\boldsymbol{v} \cdot \boldsymbol{w}) = (A\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{w},$$

and

$$(\boldsymbol{u} \otimes \boldsymbol{v})A\boldsymbol{w} = \boldsymbol{u}(\boldsymbol{v} \cdot A\boldsymbol{w}) = \boldsymbol{u}(A^T\boldsymbol{v} \cdot \boldsymbol{w}) = (\boldsymbol{u} \otimes A^T\boldsymbol{v})\boldsymbol{w}.$$

□

If A is a second order tensor in $\mathcal{L}(V)$, then the components of the transpose A^T satisfy the following relations:

$$\begin{aligned} (A^T)^{ij} &= A^{ji}, & (A^T)_{ij} &= A_{ji}, \\ (A^T)^i{}_j &= A_j{}^i, & (A^T)_i{}^j &= A^j{}_i. \end{aligned} \tag{A.9}$$

We see from these relations that for contravariant or covariant tensors the matrix of A^T is simply the transpose of the matrix of A. However, from the second group of the relations in (A.9) for mixed tensors, this is not valid in general, since the matrix transpose of $[A^i{}_j]$, by changing rows and columns, is $[A^j{}_i]$, instead of $[A_j{}^i]$.

A tensor $S \in \mathcal{L}(V)$ is called *symmetric* if $S^T = S$, and is called *skew-symmetric* if $S^T = -S$. In other words, S is symmetric if $\boldsymbol{v} \cdot S\boldsymbol{u} = \boldsymbol{u} \cdot S\boldsymbol{v}$ and S is skew-symmetric if $\boldsymbol{v} \cdot S\boldsymbol{u} = -\boldsymbol{u} \cdot S\boldsymbol{v}$, for all $\boldsymbol{u}, \boldsymbol{v} \in V$.

Notation. $Sym(V) = \{S \in \mathcal{L}(V) \mid S^T = S\}$ and

$$Skw(V) = \{S \in \mathcal{L}(V) \mid S^T = -S\}.$$

Note that both $Sym(V)$ and $Skw(V)$ are subspaces of $\mathcal{L}(V)$. If $S \in Sym(V)$, then its components satisfy

$$S^{ij} = S^{ji}, \qquad S_{ij} = S_{ji},$$
$$S^i{}_j = S_j{}^i = g_{jk}\, g^{im} S^k{}_m.$$

In terms of matrix representation we have

$$[S^{ij}] = [S^{ij}]^T, \qquad [S_{ij}] = [S_{ij}]^T.$$

Note that although S is symmetric, the matrix $[S^i{}_j]$ is not symmetric in general,

$$[S^i{}_j] \neq [S^i{}_j]^T.$$

A second order tensor can also be regarded as a bilinear function in the following manner: For any $A \in \mathcal{L}(V)$, define the function on $V \times V$, also denoted by A,

$$A(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u} \cdot A\boldsymbol{v},$$

for any vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V. In particular, for simple tensors, we have

$$(\boldsymbol{u} \otimes \boldsymbol{v})(\boldsymbol{u}', \boldsymbol{v}') = \boldsymbol{u}' \cdot (\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{v}' = (\boldsymbol{u} \cdot \boldsymbol{u}')(\boldsymbol{v} \cdot \boldsymbol{v}').$$

By employing the notion of multilinear functions, we can generalize tensor products to higher orders. For example, we can define a tensor product of three vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ as a trilinear function on V by

$$(\boldsymbol{u} \otimes \boldsymbol{v} \otimes \boldsymbol{w})(\boldsymbol{u}', \boldsymbol{v}', \boldsymbol{w}') = (\boldsymbol{u} \cdot \boldsymbol{u}')(\boldsymbol{v} \cdot \boldsymbol{v}')(\boldsymbol{w} \cdot \boldsymbol{w}')$$

for any vectors $\boldsymbol{u}'$, $\boldsymbol{v}'$ and $\boldsymbol{w}'$ in V. One can show as before, that if $\{\boldsymbol{e}_i\}$ is a basis of V, then $\{\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k\}$ is a product basis for the space of all trilinear functions on V. We shall denote this space as $V \otimes V \otimes V$ and call it the space of third order tensors. If S is a third order tensor, then

$$S = S^{ijk} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k = S^i{}_{jk} \boldsymbol{e}_i \otimes \boldsymbol{e}^j \otimes \boldsymbol{e}^k = etc.$$

There are several different component forms relative to the different product bases. In a similar manner, tensor product of higher orders can be defined. We write

$$\overset{k}{\otimes} V = \overbrace{V \otimes \cdots \otimes V}^{k \text{ times}}$$

for tensors of order k. Clearly, $\dim \overset{k}{\otimes} V = (\dim V)^k$.

Exercise A.1.2 Let $\beta' = \{\boldsymbol{e}_1 = (1,0), \boldsymbol{e}_2 = (2,1)\}$ be a basis of $I\!R^2$, and $T \in \mathcal{L}(I\!R^2)$ be defined by

$$T(x_1, x_2) = (3x_1 + x_2, x_1 + 2x_2), \qquad \forall\, (x_1, x_2) \in I\!R^2. \tag{A.10}$$

1) Show that T is a symmetric transformation.
2) Determine the matrices of the associated components of T relative to β':

$$[T_{ij}],\ [T^{ij}],\ [T_i{}^j],\ [T^i{}_j].$$

Note that the last two matrices are not symmetric.

A.1.4 Transformation Rules for Components

The components of a tensor relative to a basis are uniquely determined and their values depend on the basis. Therefore, if we make a change of basis, they must change accordingly. In this section, we shall establish the transformation rules for components of tensors under a change of basis.

Consider a change of basis from $\beta = \{\boldsymbol{e}_i\}$ to another basis $\bar\beta = \{\bar{\boldsymbol{e}}_i\}$ given by

$$\bar{\boldsymbol{e}}_k = M_k{}^j\, \boldsymbol{e}_j. \tag{A.11}$$

We call $M_k{}^j$ the transformation matrix for the change of basis from β to $\bar\beta$. By the use of (A.3), we have

$$M_k{}^j = \bar{\boldsymbol{e}}_k \cdot \boldsymbol{e}^j,$$

from which we can also obtain the relation between the dual bases $\beta^* = \{\boldsymbol{e}^i\}$ and $\bar\beta^* = \{\bar{\boldsymbol{e}}^i\}$,

$$\boldsymbol{e}^j = M_k{}^j\, \bar{\boldsymbol{e}}^k. \tag{A.12}$$

The above two transformation relations (A.11) and (A.12) can be schematically represented by

$$\beta \quad \xrightarrow{\;M\;} \quad \bar\beta\ ,$$

$$\beta^* \xleftarrow{\;M^T\;} \bar\beta^*.$$

In other words, if M changes a basis β to another basis $\bar\beta$, their corresponding dual bases β^* and $\bar\beta^*$ are changed in the opposite direction through M^T.

The components of a vector transform in a similar manner. Indeed, let $\boldsymbol{v}$ be a vector in V, and

$$\begin{aligned}\boldsymbol{v} &= v^i\, \boldsymbol{e}_i = \bar v^i\, \bar{\boldsymbol{e}}_i \\ &= v_j\, \boldsymbol{e}^j = \bar v_j\, \bar{\boldsymbol{e}}^j.\end{aligned}$$

One can easily verify that the transformation rules for the components are

$$\bar{v}_k = M_k^{\ j}\, v_j, \qquad v^j = M_k^{\ j}\, \bar{v}^k, \tag{A.13}$$

which look exactly like the ones for the change of basis (A.11) and (A.12). In matrix notations, we have

$$[\bar{v}_k] = [M_k^{\ j}]\,[v_j], \qquad [v^j] = [M_k^{\ j}]^T [\bar{v}^k]$$

or schematically

$$[v_i] \xrightarrow{\ M\ } [\bar{v}_i],$$

$$[v^i] \xleftarrow{\ M^T\ } [\bar{v}^i].$$

That is, the covariant components transform in the same direction as the change of basis by M, while the contravariant components transform in the opposite direction by M^T. This is the reason why such components are called *co-* and *contra-*variant in classical tensor analysis, in which tensors are defined through their transformation properties.

For a second order tensor A in $\mathcal{L}(V)$,

$$\begin{aligned} A &= A_{ij}\, \boldsymbol{e}^i \otimes \boldsymbol{e}^j = \bar{A}_{ij}\, \bar{\boldsymbol{e}}^i \otimes \bar{\boldsymbol{e}}^j \\ &= A_i^{\ j}\, \boldsymbol{e}^i \otimes \boldsymbol{e}_j = \bar{A}_i^{\ j}\, \bar{\boldsymbol{e}}^i \otimes \bar{\boldsymbol{e}}_j. \end{aligned}$$

We have the following transformation rules,

$$\begin{aligned} \bar{A}_{ij} &= A_{mn}\, M_i^{\ m}\, M_j^{\ n}, \\ \bar{A}_i^{\ j} &= A_m^{\ n}\, M_i^{\ m}\, \overset{-1}{M}{}_n^{\ j}, \end{aligned} \tag{A.14}$$

where the matrix $[\,\overset{-1}{M}{}_i^{\ j}]$ is the inverse matrix of $[M_i^{\ j}]$. In matrix notations, the transformation rules can be written as

$$\begin{aligned} [\bar{A}_{ij}] &= [M_i^{\ m}]\,[A_{mn}]\,[M_j^{\ n}]^T, \\ [\bar{A}_i^{\ j}] &= [M_i^{\ m}]\,[A_m^{\ n}]\,[M_n^{\ j}]^{-1}. \end{aligned}$$

Transformation rules for other components and for tensors of higher orders are similar. The general rule can easily be obtained by composing the transformation rules for covariant and contravariant components as shown in (A.13) or (A.14).

Exercise A.1.3 Let $\beta = \{\boldsymbol{e}_1 = (1,0), \boldsymbol{e}_2 = (0,1)\}$ and $\bar{\beta} = \{\bar{\boldsymbol{e}}_1 = (1,0), \bar{\boldsymbol{e}}_2 = (2,1)\}$ be two bases of $I\!R^2$. Determine the transformation matrix of the change of basis from β to $\bar{\beta}$ and also the transformation matrix from $\bar{\beta}^*$ to β^*. Let $T \in \mathcal{L}(\boldsymbol{R}^2)$ be defined in (A.10). Determine the various components of T relative to the two different bases and verify the transformation rules.

Exercise A.1.4 For any two bases $\beta = \{\boldsymbol{e}_i\}$ and $\bar{\beta} = \{\bar{\boldsymbol{e}}_i\}$ of V, there exists a linear transformation $A \in \mathcal{L}(V)$ such that $\bar{\boldsymbol{e}}_k = A\boldsymbol{e}_k$. Show that the transformation matrix M for the change of basis from β to $\bar{\beta}$ is given by $M_k{}^j = \boldsymbol{e}^j \cdot A\boldsymbol{e}_k$, that is, $[M_k{}^j] = [A^j{}_k]$.

A.1.5 Determinant and Trace

In matrix algebra, the definition of determinant of a square matrix is based on the notion of permutation. Let $(1, \cdots, n)$ be an ordered set of natural numbers. A reordering of the elements in $(1, \cdots, n)$ is called a permutation. More precisely, a permutation is a one-to-one mapping $\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}$ resulting in the ordered set $(\sigma(1), \cdots, \sigma(n))$. There are exactly $n!$ permutation of $(1, \cdots, n)$. A permutation by exchanging order of two adjacent elements is called a transposition. It is known that any permutation can be obtained by merely subsequent transpositions, and although the number of such transpositions are not unique for a given permutation, the parity of this number is. Hence, a permutation is called even or odd according to parity of the number of transpositions in order to restore the permutation back to the natural order and one can define the sign of a permutation, denoted $\operatorname{sign}\sigma$, as $+1$ if σ is even and -1 if σ is odd.

Let $[M_{ij}]$ be a square matrix. The first index denotes the row and the second the column (it does not matter whether they are superindices or subindices). The determinant of the matrix can be calculated by

$$\det [M_{ij}] = \sum_{\sigma} (\operatorname{sign} \sigma)\, M_{\sigma(1)\,1} \cdots M_{\sigma(n)\,n}, \tag{A.15}$$

where the summation is taken over all permutations of $(1, \cdots, n)$.

On the other hand, since the matrix representation of a linear transformation depends on the choice of basis, the question arises of whether it is meaningful to define the determinant of a linear transformation as the determinant of its matrix representation. In the following, we shall see that the notion of determinant of linear transformation can be defined in a natural way, independent of the choice of basis and see how it is related to its matrix representations.

Definition. *Let V be a vector space of dimension n. A function $\omega : \overbrace{V \times \cdots \times V}^{n} \to I\!R$ is said to be an alternating n-linear form if it is n-linear and for all $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in V$,*

$$\omega(\boldsymbol{v}_{\sigma(1)}, \cdots, \boldsymbol{v}_{\sigma(n)}) = (\operatorname{sign} \sigma)\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n). \tag{A.16}$$

ω is called non-trivial if there exist $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$, such that $\omega(\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n) \neq 0$.

It is obvious that if ω is alternating then

$$\omega(\cdots, \boldsymbol{u}, \cdots, \boldsymbol{v}, \cdots) = 0, \quad \text{if } \boldsymbol{u} = \boldsymbol{v}. \tag{A.17}$$

More generally, if $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is linearly dependent then $\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = 0$. In other words, if $\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) \neq 0$ then $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is a linearly independent set, and since the number of vectors in this set equals $\dim V$, $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is also a basis of V.

Theorem. (uniqueness) *Let ω and ω' be two alternating n-linear forms and ω be non-trivial. Then there exists uniquely a $\lambda \in I\!R$, such that $\omega' = \lambda\omega$, i.e., $\forall\, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in V$,*

$$\omega'(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \lambda\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n).$$

Proof: Since ω is non-trivial, there exists a set of vectors, say $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$, such that $\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) \neq 0$, and hence it is a basis of V. Let λ be the number defined by

$$\lambda = \frac{\omega'(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)}{\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)}.$$

Suppose that $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in V$, and

$$\boldsymbol{v}_a = v_a^i \boldsymbol{e}_i, \qquad a = 1, \cdots, n.$$

Then using (A.16) and (A.17) one can easily obtain

$$\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \alpha\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n),$$
$$\omega'(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \alpha\omega'(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n),$$

where

$$\alpha = \sum_\sigma (\text{sign } \sigma) v_1^{\sigma(1)} \cdots v_n^{\sigma(n)}.$$

Therefore we have

$$\omega'(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \lambda\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n).$$

Moreover, this relation also shows that λ does not depend on the choice of basis. □

Let $T \in \mathcal{L}(V)$ be a linear transformation on V, and ω be a non-trivial alternating n-linear form on V. Define a map $T_\omega : V \times \cdots \times V \to I\!R$ by

$$T_\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \omega(T\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_n). \tag{A.18}$$

Clearly it is alternating and n-linear, hence by the uniqueness theorem, there exists a unique $\lambda \in I\!R$, such that

$$T_\omega = \lambda\omega.$$

We can easily see that the scalar λ so defined does not depend on the choice of ω. For if ω' is another non-trivial alternating n-linear form, then by the uniqueness theorem,

$$\omega' = \mu\omega, \qquad \mu \neq 0.$$

Therefore, we have

$$T_{\omega'} = \lambda'\omega' = \lambda'\mu\omega.$$

On the other hand, we have

$$\begin{aligned} T_{\omega'}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) &= \omega'(T\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_n) = \mu\omega(T\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_n) \\ &= \mu T_\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \mu\lambda\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n), \end{aligned}$$

which implies that

$$T_{\omega'} = \mu\lambda\omega.$$

Consequently, $\lambda = \lambda'$. Therefore, λ is uniquely determined by T alone and we can lay down the following definition.

Definition. *$T \in \mathcal{L}(V)$, the determinant of T, $\det T \in I\!R$, is defined by the following relation,*

$$(\det T)\,\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \omega(T\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_n), \tag{A.19}$$

for any non-trivial alternating n-linear form ω and for any $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in V$.

The function $\det : \mathcal{L}(V) \to I\!R$ has the following properties:

1) $\det \boldsymbol{u} \otimes \boldsymbol{v} = 0.$
2) $\det(\alpha 1) = \alpha^n.$ $\quad (A.20)$
3) $\det(ST) = (\det S)(\det T).$
4) $\det S^T = \det S.$

The first two properties are almost trivial. Here let us verify the property (3). By definition,

$$\begin{aligned} &\det(ST)\,\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) \\ &\quad = \omega(ST\boldsymbol{v}_1, \cdots, ST\boldsymbol{v}_n) = \omega(S(T\boldsymbol{v}_1), \cdots, S(T\boldsymbol{v}_n)) \\ &\quad = (\det S)\omega(T\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_n) = (\det S)(\det T)\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n). \end{aligned}$$

Since it holds for any $\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$, the relation (3) follows.

We can calculate the determinant of a linear transformation in term of its component matrix. Let $\{\boldsymbol{e}_i\}$ be a basis of V, and $T = T^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j$. Then by definition,

$$\begin{aligned} &(\det T)\,\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) \\ &\quad = \omega(T\boldsymbol{e}_1, \cdots, T\boldsymbol{e}_n) = \omega(T^{i_1}_{\ 1}\boldsymbol{e}_{i_1}, \cdots, T^{i_n}_{\ n}\boldsymbol{e}_{i_n}) \\ &\quad = \sum_\sigma (\text{sign }\sigma) T^{\sigma(1)}_{\ \ 1} \cdots T^{\sigma(n)}_{\ \ n}\, \omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n). \end{aligned}$$

Hence we obtain

$$\det T = \sum_{\sigma} (\text{sign } \sigma) T^{\sigma(1)}_{\ \ 1} \cdots T^{\sigma(n)}_{\ \ n},$$

which assures that

$$\det T = \det [\, T^i_{\ j}],$$

i.e., $\det T$ is equal to the determinant of the component matrix $[T^i_{\ j}]$ according to the definition (A.15). Similarly, one can show that it is also equal to determinant of $[T_i^{\ j}]$. Therefore we have

$$\begin{aligned} \det T &= \det[\, T^i_{\ j}] = \det[\, T_i^{\ j}] \\ &= \det[\, g^{ik} T_{kj}] = \det[\, g_{ik} T^{kj}]. \end{aligned}$$

Note that $\det T$ is not equal to $\det[\, T_{ij}]$ nor to $\det[\, T^{ij}]$ unless $\det[\, g_{ij}] = 1$.

Similar to the determinant, another scalar can be associated with a linear transformation. Let $T \in \mathcal{L}(V)$, and ω be a non-trivial alternating n-linear form. Define a map $\widetilde{T}_\omega : V \times \cdots \times V \to I\!R$ by

$$\widetilde{T}_\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \sum_{i=1}^{n} \omega(\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_i, \cdots, \boldsymbol{v}_n).$$

One can easily check that $\widetilde{T}_\omega$ is alternating and n-linear, hence by the uniqueness theorem, there exists a $\mu \in I\!R$, such that

$$\widetilde{T}_\omega = \mu\omega.$$

Moreover, μ does not depend on the choice of ω. Therefore, we can make the following definition.

Definition. *$T \in \mathcal{L}(V)$, the trace of T, $\operatorname{tr} T \in I\!R$, is defined by the following relation*

$$(\operatorname{tr} T)\, \omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) = \sum_{i=1}^{n} \omega(\boldsymbol{v}_1, \cdots, T\boldsymbol{v}_i, \cdots, \boldsymbol{v}_n), \qquad (A.21)$$

for any non-trivial alternating n-linear form ω and for any $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in V$.

The function $\operatorname{tr} : \mathcal{L}(V) \to I\!R$ has the following properties:

1) $\operatorname{tr}(\alpha S + T) = \alpha \ \operatorname{tr} S + \operatorname{tr} T$.
2) $\operatorname{tr} \mathit{1} = n$.
3) $\operatorname{tr}(\boldsymbol{v} \otimes \boldsymbol{u}) = \boldsymbol{v} \cdot \boldsymbol{u}$. (A.22)
4) $\operatorname{tr} S^T = \operatorname{tr} S$.
5) $\operatorname{tr}(ST) = \operatorname{tr}(TS)$.

The property (1) states that trace is a linear function on $\mathcal{L}(V)$. Here let us prove the property (3). Suppose that $\boldsymbol{v} = v^i \boldsymbol{e}_i$, then

$$\operatorname{tr}(\boldsymbol{v} \otimes \boldsymbol{u})\,\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) = \sum_{i=1}^{n} \omega(\boldsymbol{e}_1, \cdots, (\boldsymbol{v} \otimes \boldsymbol{u})\boldsymbol{e}_i, \cdots, \boldsymbol{e}_n)$$
$$= \sum_{i=1}^{n} (\boldsymbol{u} \cdot \boldsymbol{e}_i)\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{v}, \cdots, \boldsymbol{e}_n) = \sum_{i=1}^{n} (\boldsymbol{u} \cdot \boldsymbol{e}_i) v^i \omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n),$$

which implies that

$$\operatorname{tr}(\boldsymbol{v} \otimes \boldsymbol{u}) = \sum_{i=1}^{n} \boldsymbol{u} \cdot (v^i \boldsymbol{e}_i) = \boldsymbol{u} \cdot \boldsymbol{v}.$$

Hence (3) is proved.

In terms of components, let $T = T^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j = T_i{}^j \boldsymbol{e}^i \otimes \boldsymbol{e}_j$, then

$$\operatorname{tr} T = T^i{}_j \operatorname{tr}(\boldsymbol{e}_i \otimes \boldsymbol{e}^j) = T^j{}_j = T_j{}^j = g_{ij} T^{ij} = g^{ij} T_{ij}.$$

That is, $\operatorname{tr} T$ is equal to the sum of diagonal elements of the matrix $[T^i{}_j]$ or $[T_j{}^i]$, but in general is not equal to that of the matrix $[T_{ij}]$ or $[T^{ij}]$.

Example A.1.5 Show that $\det(\mathit{1} + \boldsymbol{u} \otimes \boldsymbol{v}) = 1 + \boldsymbol{u} \cdot \boldsymbol{v}$.
By definition, we have

$$\begin{aligned}
&\det(\mathit{1} + \boldsymbol{u} \otimes \boldsymbol{v})\,\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) \\
&\quad = \omega((\mathit{1} + \boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{e}_1, \cdots, (\mathit{1} + \boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{e}_n) \\
&\quad = \omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) + \sum_{i=1}^{n} \omega(\boldsymbol{e}_1, \cdots, (\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{e}_i, \cdots, \boldsymbol{e}_n) + \cdots \\
&\quad = \omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) + \operatorname{tr}(\boldsymbol{u} \otimes \boldsymbol{v})\,\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n),
\end{aligned}$$

where the dots represent terms involved with more than one factor of $(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{e}_i$ in ω. Since $(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{e}_i = (\boldsymbol{v} \cdot \boldsymbol{e}_i)\boldsymbol{u}$, which is a vector in the direction of $\boldsymbol{u}$ for any index i, those terms must all equal to zero because ω is an alternating form. □

The set of all non-trivial alternating n-linear forms on V consists of two disjoint classes. Two non-trivial alternating n-linear forms ω_1 and ω_2 are said to be equivalent if $\omega_1 = \lambda\omega_2$ for some $\lambda > 0$. Clearly, this is an equivalence relation which decomposes the set of non-trivial alternating n-linear forms into two equivalent classes. Each of these classes is called an *orientation* of

V. We call one of them, say Δ, the *positive orientation.* A basis $\{\boldsymbol{e}_i\}$ of V is called *positively oriented* if for any $\omega \in \Delta$,

$$\omega(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) > 0,$$

and $A \in \mathcal{L}(V)$ is said to be *orientation-preserving* if $A_\omega \in \Delta$, for any $\omega \in \Delta$. Here A_ω is defined by (A.18). Since $A_\omega = (\det A)\omega$, A preserves the orientation if and only if $\det A > 0$.

Let $\{\boldsymbol{e}_i\}$ and $\{\bar{\boldsymbol{e}}_i\}$ be two bases such that $A(\boldsymbol{e}_i) = \bar{\boldsymbol{e}}_i$. If $\det A > 0$ (or < 0), then $\{\boldsymbol{e}_i\}$ and $\{\bar{\boldsymbol{e}}_i\}$ are said to have the *same* (or the *opposite*) orientation.

Suppose that V is a three-dimensional vector space and let $\{\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3\}$ be an positively oriented orthonormal basis of V, then there exists a unique $e \in \Delta$, called the *volume element*, such that

$$e(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3) = 1.$$

Since $e \in \mathcal{L}(V \times V \times V, I\!R)$, it is a third order tensor and can be represented as

$$e = \varepsilon_{ijk}\, \boldsymbol{i}_i \otimes \boldsymbol{i}_j \otimes \boldsymbol{i}_k,$$

where $\varepsilon_{ijk} = e(\boldsymbol{i}_i, \boldsymbol{i}_j, \boldsymbol{i}_k)$ are the components of e relative to the basis $\{\boldsymbol{i}_k\}$. Obviously we have

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3), \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3), \\ 0 & \text{otherwise.} \end{cases}$$

One can easily check the following identities:

$$\begin{aligned} &\varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}, \\ &\varepsilon_{ijk}\varepsilon_{ijn} = 2\,\delta_{kn}, \\ &\varepsilon_{ijk}\varepsilon_{ijk} = 6, \end{aligned} \tag{A.23}$$

where δ_{mn} is the Kronecker delta.

Let $\{\boldsymbol{e}_k\}$ be a basis and $A \in \mathcal{L}(V)$ be a change of basis from $\{\boldsymbol{i}_k\}$ to $\{\boldsymbol{e}_k\}$, i.e., $A\,\boldsymbol{i}_k = \boldsymbol{e}_k$. Then the covariant components of the volume element relative to $\{\boldsymbol{e}_k\}$ are

$$e_{ijk} = e(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k), \quad e = e_{ijk}\, \boldsymbol{e}^i \otimes \boldsymbol{e}^j \otimes \boldsymbol{e}^k.$$

By (A.19), it follows that

$$e_{ijk} = (\det A)\varepsilon_{ijk}.$$

and also,

$$g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j = A\,\boldsymbol{i}_i \cdot A\,\boldsymbol{i}_j = (A^T A)_{ij},$$

which yields $g = (\det A)^2$, where $g = \det[g_{ij}]$. Therefore we have

$$e_{ijk} = \sqrt{g}\,\varepsilon_{ijk}, \tag{A.24}$$

if A preserves the orientation. Similarly, the contravariant components of the volume element are

$$e^{ijk} = e(\boldsymbol{e}^i, \boldsymbol{e}^j, \boldsymbol{e}^k), \quad e = e^{ijk}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k,$$

and

$$e^{ijk} = (\sqrt{g})^{-1}\varepsilon^{ijk},$$

where $\varepsilon^{ijk} = \varepsilon_{ijk}$. Moreover, the identities (A.23) can be written as

$$\begin{aligned} &e^{ijk}e_{imn} = \delta^j{}_m\delta^k{}_n - \delta^j{}_n\delta^k{}_m, \\ &e^{ijk}e_{ijn} = 2\,\delta^k{}_n, \\ &e^{ijk}e_{ijk} = 6. \end{aligned} \tag{A.25}$$

If $T \in \mathcal{L}(V)$ and $T = T^i{}_j\boldsymbol{e}_i \otimes \boldsymbol{e}^j$, then (A.19) leads to the following formula for the determinant of T,

$$e_{lmn}(\det T) = e_{ijk}T^i{}_lT^j{}_mT^k{}_n. \tag{A.26}$$

Multiplying e^{lmn} and using the last identity of (A.25), we obtain another formula for the determinant,

$$\det T = \frac{1}{6}\,e^{lmn}e_{ijk}T^i{}_l\,T^j{}_mT^k{}_n.$$

Exercise A.1.5 Consider the tensor defined by (A.10) in the previous exercise. Calculate $\det T$ and $\operatorname{tr} T$ by means of definition and also by the use of component matrices relative to β'.

A.1.6 Exterior Product and Vector Product

The usual vector product on a three-dimensional vector space can not be generalized directly to vector spaces in general. However it can be associated with the skew symmetric tensor product in a trivial manner.

Definition. *For any, $\boldsymbol{v}, \boldsymbol{u} \in V$, the exterior product of $\boldsymbol{v}$ and $\boldsymbol{u}$, denoted $\boldsymbol{v} \wedge \boldsymbol{u}$, is defined by*

$$\boldsymbol{v} \wedge \boldsymbol{u} = \boldsymbol{v} \otimes \boldsymbol{u} - \boldsymbol{u} \otimes \boldsymbol{v}.$$

It is obvious that the operation $\wedge : V \times V \longrightarrow V \otimes V$ is bilinear and skew-symmetric, i.e.,

$$\boldsymbol{v} \wedge \boldsymbol{u} = -\boldsymbol{u} \wedge \boldsymbol{v}.$$

The exterior product of two vectors $\boldsymbol{v} \wedge \boldsymbol{u}$ is a skew-symmetric tensor.

Suppose that $\{\boldsymbol{e}_i \otimes \boldsymbol{e}_j\}$, $i, j = 1, \cdots, n$ is a product basis of $V \otimes V$, then it is easy to verify that $\{\boldsymbol{e}_i \wedge \boldsymbol{e}_j\}$, $1 \le i < j \le n$ is a basis for $Skw(V)$. Therefore we have the following proposition.

Proposition. *If* $\dim V = n$, *then* $\dim Skw(V) = n(n-1)/2$. *In particular, if* $n = 3$, *then* $\dim Skw(V) = 3$.

Now suppose that V is an oriented Euclidean three-dimensional vector space. Since the space of skew-symmetric tensors is also three-dimensional we can define a map

$$\tau : Skw(V) \longrightarrow V$$

by the condition: for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$,

$$\tau(\boldsymbol{u} \wedge \boldsymbol{v}) \cdot \boldsymbol{w} = e(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}), \tag{A.27}$$

Here e is the volume element of V. This linear map, called the *duality map*, is one-to-one and onto and hence establishes a one-to-one correspondence between a skew-symmetric tensor and a vector. It is easy to verify that

$$\tau(\boldsymbol{e}_i \wedge \boldsymbol{e}_j) = e_{ijk}\boldsymbol{e}^k. \tag{A.28}$$

For an orthonormal basis $\{\boldsymbol{i}_k\}$ the duality map τ establishes the following correspondence,

$$\begin{aligned} \boldsymbol{i}_1 \wedge \boldsymbol{i}_2 &\longmapsto \boldsymbol{i}_3, \\ \boldsymbol{i}_2 \wedge \boldsymbol{i}_3 &\longmapsto \boldsymbol{i}_1, \\ \boldsymbol{i}_3 \wedge \boldsymbol{i}_1 &\longmapsto \boldsymbol{i}_2. \end{aligned}$$

For a skew-symmetric tensor W, let $\boldsymbol{w} = \tau(W)$ be the associated vector, which shall be denoted more conveniently by

$$\boldsymbol{w} = \langle W \rangle. \tag{A.29}$$

In component form, if

$$W = W^{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j, \qquad W^{ij} = -W^{ji},$$

or

$$W = \frac{1}{2} W^{ij}\boldsymbol{e}_i \wedge \boldsymbol{e}_j.$$

Then it follows from (A.28) that

$$\boldsymbol{w} = \frac{1}{2} e_{ijk} W^{ij} \boldsymbol{e}^k.$$

If the basis is orthonormal, it becomes

$$w_i = \frac{1}{2} \varepsilon_{ijk} W_{jk}, \qquad W_{ij} = \varepsilon_{ijk} w_k, \tag{A.30}$$

or in matrix form,

$$[W_{ij}] = \begin{bmatrix} 0 & w_3 & -w_2 \\ -w_3 & 0 & w_1 \\ w_2 & -w_1 & 0 \end{bmatrix}.$$

Remark. It is worthwhile to point out that the vector associated with a skew-symmetric tensor behaves differently from usual vectors under linear transformations. To see this, let $\boldsymbol{u}, \boldsymbol{v} \in V$, then for any $\boldsymbol{w} \in V$ and $Q \in \mathcal{L}(V)$, it follows from the definition that

$$\begin{aligned}\langle Q\boldsymbol{u} \wedge Q\boldsymbol{v}\rangle \cdot Q\boldsymbol{w} &= e(Q\boldsymbol{u}, Q\boldsymbol{v}, Q\boldsymbol{w}) \\ &= (\det Q)\, e(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (\det Q)\langle \boldsymbol{u} \wedge \boldsymbol{v}\rangle \cdot \boldsymbol{w},\end{aligned}$$

which implies that

$$\langle Q\boldsymbol{u} \wedge Q\boldsymbol{v}\rangle = (\det Q) Q\langle \boldsymbol{u} \wedge \boldsymbol{v}\rangle.$$

In other words, as the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ are transformed into $Q\boldsymbol{u}$, $Q\boldsymbol{v}$ respectively, the vector $\langle \boldsymbol{u} \wedge \boldsymbol{v}\rangle$ is transformed into $Q\langle \boldsymbol{u} \wedge \boldsymbol{v}\rangle$ only to within a scalar constant, or, into a vector which may point in one or the opposite sense of the same axial direction. For this reason, a vector associated with a skew-symmetric tensor is usually called an *axial vector*. □

The usual vector product, in the three-dimensional vector space, can now be defined from the exterior product in a similar manner.

Definition. *For any* $\boldsymbol{u}, \boldsymbol{v} \in V$, *the vector product* of $\boldsymbol{u}$ and $\boldsymbol{v}$, *denoted* $\boldsymbol{u}\times\boldsymbol{v}$, *is defined by*

$$\boldsymbol{u} \times \boldsymbol{v} = \langle \boldsymbol{u} \wedge \boldsymbol{v}\rangle. \qquad (A.31)$$

Clearly the operation $\times : V \times V \longrightarrow V$ is bilinear and skew-symmetric. In components (A.31) gives

$$\boldsymbol{u} \times \boldsymbol{v} = e_{ijk} u^j v^k \boldsymbol{e}^i,$$

If the basis is orthonormal, say $\{\boldsymbol{i}_k\}$, then it becomes

$$\boldsymbol{u} \times \boldsymbol{v} = \varepsilon_{ijk} u_j v_k \boldsymbol{i}_i,$$

which is the usual definition of the vector product.

The relations (A.27) and (A.31) imply that

$$e(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}.$$

This is usually called the *triple product* of $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$. For convenience, we shall also use the notation,

$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}.$$

With this notation, we can rewrite the definitions (A.19) and (A.21) of the determinant and the trace in the following form

$$\begin{aligned} \det A &= \frac{[A\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3]}{[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]}, \\ \operatorname{tr} A &= \frac{[A\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, A\boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{e}_2, A\boldsymbol{e}_3]}{[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]}. \end{aligned} \tag{A.32}$$

One may use the duality map to identify a skew-symmetric tensor with an axial vector, as well as the exterior product with the vector product. In other words, one may interpret the duality in either way in case no ambiguity would arise.

Exercise A.1.6 Verify the following relations, using index notations:
1) $W\boldsymbol{v} = -\boldsymbol{w} \times \boldsymbol{v}$.
2) $(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w} = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{w})\boldsymbol{u}$.
3) $|\boldsymbol{u} \times \boldsymbol{v}|^2 = |\boldsymbol{u}|^2|\boldsymbol{v}|^2 - |\boldsymbol{u} \cdot \boldsymbol{v}|^2$.
4) $|\boldsymbol{u} \times \boldsymbol{v}| = |\boldsymbol{u}|\,|\boldsymbol{v}| \sin\theta(\boldsymbol{u}, \boldsymbol{v})$.

A.1.7 Second Order Tensors

We shall review some of the important properties of linear transformations, i.e., the second order tensors, mostly without proofs in this section. The proofs can be found in most standard books in linear algebra.

First, let us introduce an inner product of two second order tensors. Let $A, B \in \mathcal{L}(V)$, we can define the *inner product* of A and B by

$$A \cdot B = \operatorname{tr}(AB^T),$$

which is obviously a bilinear, symmetric and positive-definite operation. We have

$$\mathit{1} \cdot A = \operatorname{tr} A,$$

where $\mathit{1}$ is the identity tensor, and for any $A, B, C \in \mathcal{L}(V)$,

$$AB \cdot C = B \cdot A^T C.$$

The *norm* of a tensor $A \in \mathcal{L}(V)$, can then be defined as

$$|A| = \sqrt{A \cdot A} = \sqrt{\operatorname{tr} AA^T}.$$

Note that if A_{ij} is the components of A relative to an orthonormal basis, then the norm of A is simply

$$|A| = (A_{11}^2 + A_{12}^2 + \cdots + A_{nn}^2)^{1/2}.$$

Now, suppose that $A \in \mathcal{L}(V)$ is one-to-one (therefore, onto), then there is a unique $A^{-1} \in \mathcal{L}(V)$, called the *inverse* of A, such that

$$AA^{-1} = A^{-1}A = \mathit{1}.$$

If A^{-1} exists, A is said to be *invertible* or *nonsingular*, otherwise, it is said to be *singular*. It can be proved that A is invertible if and only if $\det A \neq 0$, and for any nonsingular A and B,

$$(AB)^{-1} = B^{-1}A^{-1},$$
$$(A^{-1})^T = (A^T)^{-1} = A^{-T}.$$

Notation. $Inv(V) = \{F \in \mathcal{L}(V) \mid F \text{ is invertible}\}$.

Recall that a set G is called a *group* if it has the following properties:

1) If $A, B \in G$ then $AB \in G$.
2) If $A, B, C \in G$ then $A(BC) = (AB)C$.
3) There exists an identity element $\mathit{1} \in G$ such that $\mathit{1}A = A\mathit{1} = A$, for any $A \in G$.
4) For any $A \in G$, there exists $A^{-1} \in G$, such that $AA^{-1} = A^{-1}A = \mathit{1}$.

It is easy to verify that $Inv(V)$ forms a group under the operation of composition. It is usually known as the *general linear group* of V, denoted by $GL(V)$.

Definition. *$Q \in \mathcal{L}(V)$ is called an orthogonal transformation if it preserves the inner product of V. i.e., for all $\boldsymbol{u}, \boldsymbol{v} \in V$,*

$$Q\boldsymbol{u} \cdot Q\boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{v}.$$

Notation. $\mathcal{O}(V) = \{Q \in \mathcal{L}(V) \mid Q \text{ is orthogonal}\}$.

The set $\mathcal{O}(V)$ forms a group and is called the *orthogonal group* of V. Orthogonal transformations have the following properties:

1) $Q^T = Q^{-1}$.
2) $|\det Q| = 1$.
3) $|Q\boldsymbol{v}| = |\boldsymbol{v}|$.
4) $\theta(Q\boldsymbol{v}, Q\boldsymbol{u}) = \theta(\boldsymbol{v}, \boldsymbol{u})$.

The last two relations assert that orthogonal transformations also preserve norms and angles. An orthogonal transformation Q is said to be *proper* if $\det Q = 1$, and *improper* if $\det Q = -1$.

Notation. $\mathcal{O}^+(V) = \{Q \in \mathcal{O}(V) \mid \det Q = 1\}$.

The set $\mathcal{O}^+(V)$ also forms a group, called the *proper orthogonal group* of V. It is also called the *rotation group* since its elements are rotations. Note that the subset of $\mathcal{O}(V)$ with determinant equal to -1 does not form a group since it does not have an identity element.

Notation. $\mathcal{U}(V) = \{T \in \mathcal{L}(V) \mid |\det T| = 1\}$, and $SL(V) = \{T \in \mathcal{L}(V) \mid \det T = 1\}$.

Element of $\mathcal{U}(V)$ are called unimodular transformations and $\mathcal{U}(V)$ forms a group, called the *unimodular group* of V. $SL(V)$ also forms a group, called the *special linear group* of V. Clearly, we have the following relations:

$$\mathcal{O}^+(V) \subset \begin{matrix} SL(V) \\ \mathcal{O}(V) \end{matrix} \subset \mathcal{U}(V) \subset GL(V).$$

A.1.8 Some Theorems of Linear Algebra

We shall mention some important theorems of linear algebra relevant to the study of mechanics. They are all related to the concept of eigenvalues and eigenvectors.

Definition. *Let $A \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{R}$ is called an eigenvalue of A, if there exists a non-zero vector $\boldsymbol{v} \in V$, such that*

$$A\boldsymbol{v} = \lambda \boldsymbol{v}. \tag{A.33}$$

$\boldsymbol{v}$ is called the eigenvector of A associated with the eigenvalue λ.

It follows from the definition that λ is an eigenvalue if and only if

$$\det(A - \lambda 1) = 0. \tag{A.34}$$

The left-hand side of (A.34) is a polynomial of degree n in λ, where n is the dimension of V. We may write it in the form

$$(-\lambda)^n + I_1(-\lambda)^{n-1} + \cdots + I_{n-1}(-\lambda) + I_n = 0.$$

It is called the *characteristic equation* of A. Its real roots are the eigenvalues of A. The coefficients $I_1, \cdots, I_n$ are scalar functions of A and are called the *principal invariants* of A.

It can be shown that the characteristic equation is also satisfy by the tensor A itself. We have the following

Cayley–Hamilton Theorem. *A second order tensor $A \in \mathcal{L}(V)$ satisfies its own characteristic equation,*

$$(-A)^n + I_1(-A)^{n-1} + \cdots + I_{n-1}(-A) + I_n 1 = 0.$$

Example A.1.6 For $\dim V = 3$ and $A \in \mathcal{V}$, we have

$$\det(A - \lambda \mathit{1}) = -\lambda^3 + I_A \lambda^2 - \mathit{II}_A \lambda + \mathit{III}_A. \qquad (A.35)$$

The three principal invariants of A, more specifically denoted by I_A, II_A, and III_A can be obtained from the following relations:

$$I_A = \operatorname{tr} A, \qquad \mathit{II}_A = \operatorname{tr} A^{-1} \det A, \qquad \mathit{III}_A = \det A. \qquad (A.36)$$

Of course, the second relation is valid only when A is nonsingular.

Proof: From (A.32) we can write

$$\begin{aligned} \det(A - \lambda \mathit{1})[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] =\ & [(A - \lambda \mathit{1})\boldsymbol{e}_1, (A - \lambda \mathit{1})\boldsymbol{e}_2, (A - \lambda \mathit{1})\boldsymbol{e}_3] \\ =\ & -\lambda^3 [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] \\ & + \lambda^2([A\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, A\boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{e}_2, A\boldsymbol{e}_3]) \\ & - \lambda\,([\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3] + [A\boldsymbol{e}_1, \boldsymbol{e}_2, A\boldsymbol{e}_3] + [A\boldsymbol{e}_1, A\boldsymbol{e}_2, \boldsymbol{e}_3]) \\ & + [A\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3]. \end{aligned}$$

Comparing this with the right-hand side of (A.35), we obtain $(A.36)_{1,3}$ by the use of (A.32), as well as the following relation for the second invariant II_A,

$$\mathit{II}_A = \frac{[\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3] + [A\boldsymbol{e}_1, \boldsymbol{e}_2, A\boldsymbol{e}_3] + [A\boldsymbol{e}_1, A\boldsymbol{e}_2, \boldsymbol{e}_3]}{[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]}.$$

If $A \in \mathit{Inv}(V)$, then it implies the second relation of (A.36). In particular, if $\det A = 1$, we have $\mathit{II}_A = I_{A^{-1}}$. □

In general, the characteristic equation may not have real roots. However, it is known that if A is symmetric all the roots are real and there exists a basis of V consisting entirely of eigenvectors.

Spectral Theorem. *Let $S \in \mathit{Sym}(V)$, then there exists an orthonormal basis $\{\boldsymbol{e}_i\}$ of V, such that S can be written in the form*

$$S = \sum_{i=1}^{n} s_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i. \qquad (A.37)$$

Such a basis is called a *principal basis* for S. Relative to this basis, the component matrix of S is a diagonal matrix and the diagonal elements s_i are the eigenvalues of S associated with the eigenvectors $\boldsymbol{e}_i$ respectively. The eigenvalues s_i, $i = 1, \cdots, n$ may or may not be distinct.

Definition. *Let λ be an eigenvalue of $S \in \mathcal{L}(V)$. We call $V_\lambda = \{\boldsymbol{v} \in V \mid S\boldsymbol{v} = \lambda\boldsymbol{v}\}$ the characteristic space of S associated with λ.*

If S is a symmetric tensor and suppose that $\boldsymbol{v} \in V_\lambda$, $\boldsymbol{u} \in V_\mu$, where λ and μ are two distinct eigenvalues of S, then one can easily show that $\boldsymbol{v}\cdot\boldsymbol{u} = 0$, i.e. they are mutually orthogonal. Moreover, by the spectral theorem any vector $\boldsymbol{v}$ can be written in the form

$$\boldsymbol{v} = \sum_\lambda \boldsymbol{v}_\lambda, \qquad \boldsymbol{v}_\lambda \in V_\lambda, \tag{A.38}$$

where the summation is extended over all characteristic spaces of S.

Commutation Theorem. *Let $T \in \mathcal{L}(V)$ and $S \in Sym(V)$. Then*

$$ST = TS$$

if and only if T preserves all characteristic spaces of S. i.e., T maps each characteristic space of S into itself.

Proof: Suppose that S and T commute, and $S\boldsymbol{v} = \lambda\boldsymbol{v}$. Then

$$S(T\boldsymbol{v}) = T(S\boldsymbol{v}) = \lambda(T\boldsymbol{v}),$$

so that both $\boldsymbol{v}$ and $T\boldsymbol{v}$ belong to the characteristic space V_λ.

To prove the converse, since S is symmetric, for any $\boldsymbol{v} \in V$, let $\boldsymbol{v} = \sum_\lambda \boldsymbol{v}_\lambda$ be the decomposition relative to the characteristic spaces of S as given in (A.38). If T leaves each characteristic space V_λ invariant, then $T\boldsymbol{v}_\lambda \in V_\lambda$ and

$$S(T\boldsymbol{v}_\lambda) = \lambda(T\boldsymbol{v}_\lambda) = T(\lambda\boldsymbol{v}_\lambda) = T(S\boldsymbol{v}_\lambda).$$

Therefore, by (A.38), we have

$$ST\boldsymbol{v} = \sum_\lambda ST\boldsymbol{v}_\lambda = \sum_\lambda TS\boldsymbol{v}_\lambda = TS\boldsymbol{v},$$

which shows that $ST = TS$. □

There is only one subspace of V that is preserved by any rotation, namely V itself. Therefore, we have the following

Corollary. *A symmetric $S \in \mathcal{L}(V)$ commutes with every orthogonal transformation if and only if $S = \lambda 1$, for some $\lambda \in I\!R$.*

Definition. *$S \in \mathcal{L}(V)$ is said to be positive definite (positive semi-definite) if for any $\boldsymbol{v} \in V$ and $\boldsymbol{v} \neq 0$,*

$$\boldsymbol{v} \cdot S\boldsymbol{v} > 0 \ (\geq 0).$$

Similarly, S is said to be negative definite (negative semi-definite) if

$$\boldsymbol{v} \cdot S\boldsymbol{v} < 0 \ (\leq 0).$$

One can easily see that if S is symmetric, then it is positive definite if and only if all of its eigenvalues are positive. Consequently, for any symmetric positive definite transformation S, there is a *unique* symmetric positive definite transformation T such that $T^2 = S$ and the eigenvalues of T are the positive square roots of those of S associated with the same eigenvectors. We denote $T = \sqrt{S}$ and call T the *square root* of S. In other words, if S is expressed by (A.37) in terms of the principal basis, then

$$T = \sqrt{S} = \sum_{i=1}^{n} \sqrt{s_i}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_i.$$

Example A.1.7 Let $S \in \mathcal{L}(I\!R^2)$ be given by $S(x,y) = (3x+\sqrt{2}y, \sqrt{2}x+2y)$. Relative to the standard basis of $I\!R^2$, the matrix of S is

$$[S_{ij}] = \begin{bmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix},$$

which has the eigenvalues $s_1 = 4$ and $s_2 = 1$ and the corresponding principal basis $\boldsymbol{e}_1 = (\sqrt{2/3}, \sqrt{1/3})$ and $\boldsymbol{e}_2 = (-\sqrt{1/3}, \sqrt{2/3})$. Therefore, we have

$$T = \sqrt{S} = 2\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_2,$$

whose matrix relative to the standard basis becomes

$$[T_{ij}] = \frac{2}{3}\begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 & -\sqrt{2} \\ -\sqrt{2} & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 5 & \sqrt{2} \\ \sqrt{2} & 4 \end{bmatrix}.$$

One can easily verify that $[T_{ij}]^2 = [S_{ij}]$. □

Example A.1.8 Let S be a positive definite symmetric tensor in a two-dimensional space, then

$$\sqrt{S} = \frac{1}{b}(S + a\mathit{1}),$$

where $a = \sqrt{\det S}$ and $b = \sqrt{2a + \operatorname{tr} S}$.
Proof: Let $A = \sqrt{S}$. By the Cayley–Hamilton Theorem in the two-dimensional space, we have the identity

$$A^2 - (\operatorname{tr} A)A + (\det A)\mathit{1} = 0.$$

Since $A^2 = S$, if we let the eigenvalues of A be a_1 and a_2, then $\det S = a_1^2 a_2^2$ and $\operatorname{tr} S = a_1^2 + a_2^2$. Therefore

$$a = \sqrt{a_1^2 a_2^2} = a_1 a_2 = \det A,$$
$$b = \sqrt{2a_1a_2 + a_1^2 + a_2^2} = a_1 + a_2 = \operatorname{tr} A,$$

which together with the above identity prove the result. □

Polar Decomposition Theorem. *For any $F \in Inv(V)$, there exist symmetric positive definite transformations V and U and a orthogonal transformation R such that*

$$F = RU = VR.$$

Moveover, the transformations U, V and R are uniquely determined in the above decompositions.

Proof: We can easily verify that FF^T and F^TF are symmetric positive definite. Indeed for any $\boldsymbol{v} \neq 0$, we have

$$(\boldsymbol{v} \cdot F^T F \boldsymbol{v}) = (F\boldsymbol{v} \cdot F\boldsymbol{v}) > 0,$$

since F is nonsingular.

To prove the theorem, let us define

$$U = \sqrt{F^T F}, \quad R = FU^{-1}, \quad V = RUR^T. \tag{A.39}$$

By definition, U is symmetric positive definite and R is orthogonal since

$$\begin{aligned} RR^T &= FU^{-1}(FU^{-1})^T = FU^{-1}U^{-T}F^T \\ &= FU^{-2}F^T = F(F^TF)^{-1}F^T = 1. \end{aligned}$$

Moreover, from the definition (A.39) we also have

$$V^2 = RUR^T(RUR^T) = (RU)(RU)^T = FF^T.$$

Therefore, V is the square root of FF^T and hence is itself a symmetric positive definite transformation. Furthermore, the uniqueness follows from the definition of square root. □

The polar decomposition theorem, which decomposes a nonsingular transformation into a rotation and a positive definite tensor, is crucial in the development of continuum mechanics. The following decomposition of a tensor into its symmetric and skew-symmetric parts is also important in mechanics.

For any $T \in \mathcal{L}(V)$, let

$$A = \frac{1}{2}(T + T^T), \qquad B = \frac{1}{2}(T - T^T),$$

then

$$T = A + B, \qquad A \in Sym(V), \quad B \in Skw(V).$$

This is sometimes called the *Cartesian decomposition* of a tensor. Such a decomposition is also unique.

Exercise A.1.7 Let $A \in \mathcal{L}(V)$ be such that $(1 + A)$ is nonsingular. Verify that

1) $(1 + A)^{-1} = 1 - A(1 + A)^{-1}$.
2) $(1 + A)^{-1} = 1 - A + A^2 - \cdots + (-1)^n A^n + o(A^n)$ if $\lim_{|A| \to 0} \dfrac{o(A^n)}{|A|^n} = 0$.

Exercise A.1.8 Let $\boldsymbol{u}, \boldsymbol{v} \in V$. Show that if $1 + \boldsymbol{u} \cdot \boldsymbol{v} \neq 0$ then

$$(1 + \boldsymbol{u} \otimes \boldsymbol{v})^{-1} = 1 - \frac{\boldsymbol{u} \otimes \boldsymbol{v}}{1 + \boldsymbol{u} \cdot \boldsymbol{v}}.$$

Exercise A.1.9 For $\dim V = 3$, let $A \in \mathcal{L}(V)$ and $B = 1 + A$. Show that

$$\begin{aligned} I_B &= 3 + I_A, \\ II_B &= 3 + 2I_A + II_A, \\ III_B &= 1 + I_A + II_A + III_A, \end{aligned}$$

and if $a = \det B \neq 0$, verify that

$$(1 + A)^{-1} = \frac{1}{a}\Big((1 + I_A + II_A)1 - (1 + I_A)A + A^2\Big).$$

Exercise A.1.10 Prove the Cayley–Hamilton theorem for the special case that $A \in \mathcal{L}(V)$ is symmetric, by employing the spectral theorem.

Exercise A.1.11 Let $\beta = \{(1,0,0), (0,1,0), (0,0,1)\}$ be the standard basis of $I\!R^3$ and the matrix representation of $F \in \mathcal{L}(I\!R^3)$ relative to β be given by

$$F = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Suppose that $F = RU = VR$ be the polar decomposition of F. Find the matrix representation of U, V, and R relative to the standard basis β.

A.2 Tensor Calculus

In the second part of this appendix, we shall discuss some basic notions of calculus on Euclidean spaces: gradients and other differential operators of tensor functions.

A.2.1 Euclidean Point Space

Let $\mathcal{E}$ be a set of points and V be a Euclidean vector space of dimension n.

Definition. *$\mathcal{E}$ is called a Euclidean point space of dimension n, and V is called the translation space of $\mathcal{E}$, if for any pair of points $x, y \in \mathcal{E}$, there is a vector $\boldsymbol{v} \in V$, called the difference vector of x and y, written as*

$$\boldsymbol{v} = y - x, \tag{A.40}$$

with the following properties:
1) $\forall\, x \in \mathcal{E}, \quad x - x = 0 \in V$.
2) $\forall\, x \in \mathcal{E}, \forall\, \boldsymbol{v} \in V$, *there exists a unique point $y \in \mathcal{E}$, such that* (A.40) *is satisfied. We write $y = x + \boldsymbol{v}$.*
3) $\forall\, x, y, z \in \mathcal{E}, \quad (x - y) + (y - z) = (x - z)$.

Obviously, with (A.40) we can define the *distance* between x and y in $\mathcal{E}$, denoted $d(x, y)$, by

$$d(x, y) = |\boldsymbol{v}|\,,$$

or equivalently

$$d(x, y) = \sqrt{(x - y) \cdot (x - y)},$$

where the dot denotes the inner product on V.

Notation. $\mathcal{E}_x = \{\boldsymbol{v}_x = (x, \boldsymbol{v}) \mid \boldsymbol{v} = y - x,\ \forall\, y \in \mathcal{E}\}$.

$\mathcal{E}_x$ denotes the set of all difference vectors at x. It can be made into a Euclidean vector space in an obvious way, with the addition and scalar multiplication defined as

$$\begin{aligned} \boldsymbol{v}_x + \boldsymbol{u}_x &= (\boldsymbol{v} + \boldsymbol{u})_x, \\ \alpha\, \boldsymbol{v}_x &= (\alpha \boldsymbol{v})_x. \end{aligned}$$

We call $\mathcal{E}_x$ the *tangent space* of $\mathcal{E}$ at x.

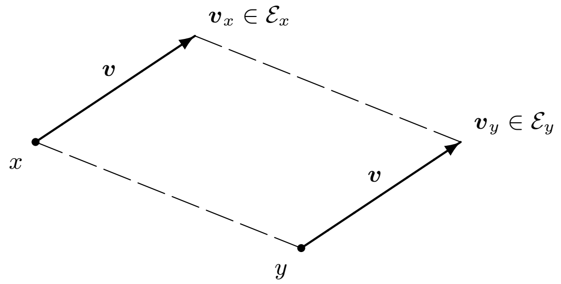


Fig. A.1. Parallel translation

Clearly $\mathcal{E}_x$ is a copy of V, i.e., it is isomorphic to V. In other words, for any $x \in \mathcal{E}$, the map $i_x : V \to \mathcal{E}_x$, called the *Euclidean parallelism*, taking $\boldsymbol{v}$ to $\boldsymbol{v}_x$ trivially establishes a one-to-one correspondence between $\mathcal{E}_x$ and V. The composite map

$$\tau_{xy} = i_y \circ i_x^{-1} : \mathcal{E}_x \longrightarrow \mathcal{E}_y$$

taking

$$\boldsymbol{v}_x = (x, \boldsymbol{v}) \longmapsto \boldsymbol{v}_y = (y, \boldsymbol{v})$$

defines the *parallel translation* of vectors at x to vectors at y (Fig. A.1).

Therefore, although $\mathcal{E}_x$ and $\mathcal{E}_y$ for $x \neq y$, are two different tangent spaces, they can be identified through V in an obvious manner,

$$\mathcal{E}_x \cong \mathcal{E}_y \cong V, \qquad \forall\, x, y \in \mathcal{E}.$$

In other words, $\boldsymbol{v}_x = (x, \boldsymbol{v}) \in \mathcal{E}_x$ and $\boldsymbol{u}_y = (y, \boldsymbol{u}) \in \mathcal{E}_y$ are regarded as the same vector if and only if $\boldsymbol{v} = \boldsymbol{u}$. In this manner, vectors at different tangent spaces can be added or subtracted as if they were in the same vector space.

A.2.2 Differentiation

Before we define the derivative of tensor functions on Euclidean space in general, let us recall the definition of derivative of a real-valued function of a real variable. Let $f : (a, b) \to I\!R$ be a function on the interval $(a, b) \subset I\!R$. The derivative of f at $t \in (a, b)$ is defined as

$$\frac{df(t)}{dt} = \lim_{h \to 0} \frac{1}{h} \Big(f(t + h) - f(t) \Big),$$

if the limit exists.

This definition can easily be extended to tensor-valued functions of a real variable. Let W be a space equipped with a norm (or a distance function). As examples, we have

$$\begin{aligned} &I\!R &&: \quad d(x,y) = |x-y|\,, \\ &\mathcal{E} &&: \quad d(x,y) = \sqrt{(x-y)\cdot(x-y)}, \\ &V &&: \quad |\boldsymbol{u}| = \sqrt{\boldsymbol{u}\cdot\boldsymbol{u}}, \\ &\mathcal{L}(V), Sym(V), Skw(V) &&: \quad |A| = \sqrt{\operatorname{tr} AA^T}. \end{aligned} \tag{A.41}$$

With a norm it makes sense to talk about limit and convergence in the space W.

Let $\boldsymbol{f} : (a,b) \to W$ be a function defined on an interval $(a,b) \subset I\!R$. The derivative of $\boldsymbol{f}$ at $t \in (a,b)$ is defined as

$$\frac{d\boldsymbol{f}(t)}{dt} = \lim_{h\to 0} \frac{1}{h}\Big(\boldsymbol{f}(t+h) - \boldsymbol{f}(t)\Big). \tag{A.42}$$

The derivative of $\boldsymbol{f}$ at t will also be denoted by $\dot{\boldsymbol{f}}(t)$. Obviously for any $t \in (a,b)$ we have $\dot{\boldsymbol{f}}(t) \in W$.

Note that if $\boldsymbol{f}$ is defined on a more general space, the expression on the right-hand side of the definition (A.42) may not make sense at all. However, we can rewrite the relation (A.42) in a different form.

For fixed t, let $D\boldsymbol{f}(t) : I\!R \to W$ be the linear transformation defined by

$$D\boldsymbol{f}(t)[h] = \dot{\boldsymbol{f}}(t)\, h.$$

Then (A.42) is equivalent to

$$\lim_{h\to 0} \frac{1}{|h|}\,|\boldsymbol{f}(t+h) - \boldsymbol{f}(t) - D\boldsymbol{f}(t)[h]| = 0.$$

In this form the definition of derivative can easily be generalized to other functions.

Tensor fields

Now we shall consider functions on a Euclidean point space $\mathcal{E}$. Let $\mathcal{D}$ be an open set in $\mathcal{E}$, and $\boldsymbol{f}$ be a tensor-valued function, $\boldsymbol{f} : \mathcal{D} \to W$. Such functions are usually called tensor fields, more specifically,

1) $W = I\!R$, $\boldsymbol{f}$ is called a scalar field on $\mathcal{D}$,

$$\boldsymbol{f} : x \in \mathcal{D} \longmapsto \boldsymbol{f}(x) \in I\!R.$$

2) $W = V$, $\boldsymbol{f}$ is called a vector field on $\mathcal{D}$,

$$\boldsymbol{f} : x \in \mathcal{D} \longmapsto \boldsymbol{f}(x) \in \mathcal{E}_x \cong V.$$

3) $W = \mathcal{L}(V)$, $\boldsymbol{f}$ is called a second order tensor field on $\mathcal{D}$,

$$\boldsymbol{f} : x \in \mathcal{D} \longmapsto \boldsymbol{f}(x) \in \mathcal{E}_x \otimes \mathcal{E}_x \cong \mathcal{L}(V).$$

4) $W = \mathcal{E}$, $\boldsymbol{f}$ is called a point field on $\mathcal{D}$ or a deformation of $\mathcal{D}$,

$$\boldsymbol{f} : x \in \mathcal{D} \longmapsto \boldsymbol{f}(x) \in \mathcal{E}.$$

Definition. *A function $\boldsymbol{f} : \mathcal{D} \to W$ is said to be differentiable at $x \in \mathcal{D} \subset \mathcal{E}$, if there exists a linear transformation $D\boldsymbol{f}(x) \in \mathcal{L}(V, W)$ at x, such that for any $\boldsymbol{v} \in V$,*

$$\lim_{|\boldsymbol{v}| \to 0} \frac{1}{|\boldsymbol{v}|} \left| \boldsymbol{f}(x + \boldsymbol{v}) - \boldsymbol{f}(x) - D\boldsymbol{f}(x)[\boldsymbol{v}] \right| = 0. \qquad (A.43)$$

The linear transformation $D\boldsymbol{f}(x)$ is uniquely determined by the above relation, and it is called the *gradient* (or *derivative*) of $\boldsymbol{f}$ at x, denoted by $\operatorname{grad} \boldsymbol{f}$, or $\nabla_x \boldsymbol{f}$, or simply $\nabla \boldsymbol{f}$. By definition, $\nabla \boldsymbol{f}(x)$ is a tensor in $W \otimes V$, or is a vector in V if $W = I\!R$.

The condition (A.43) is equivalent to

$$\boldsymbol{f}(x + \boldsymbol{v}) - \boldsymbol{f}(x) = \nabla \boldsymbol{f}(x)[\boldsymbol{v}] + o(\boldsymbol{v}),$$

where $o(\boldsymbol{v})$ is a quantity containing terms such that

$$\lim_{|\boldsymbol{v}| \to 0} \frac{o(\boldsymbol{v})}{|\boldsymbol{v}|} = 0.$$

Moreover, if we substitute $t\boldsymbol{v}$ for $\boldsymbol{v}$ for some fixed $\boldsymbol{v}$ in V, (A.43) is also equivalent to

$$\begin{aligned} \nabla \boldsymbol{f}(x)[\boldsymbol{v}] &= \lim_{t \to 0} \frac{1}{t} \Big(\boldsymbol{f}(x + t\boldsymbol{v}) - \boldsymbol{f}(x) \Big) \\ &= \frac{d}{dt} \boldsymbol{f}(x + t\boldsymbol{v}) \Big|_{t=0}, \end{aligned} \qquad (A.44)$$

The right-hand side of the above relation is usually known as the *directional derivative* of $\boldsymbol{f}$ relative to the vector $\boldsymbol{v}$. Note that for fixed x and $\boldsymbol{v}$, $\boldsymbol{f}(x + t\boldsymbol{v})$ is a tensor-valued function of a real variable and its derivative can easily determined from (A.42).

Functions on tensor spaces

Let W_1 and W_2 be two spaces on which a norm or a distance function is defined, such as the spaces mentioned in (A.41) and let $\mathcal{D} \subset W_1$ be an open subset. The gradient of tensor functions on $\mathcal{D}$ can be defined in a similar manner.

Definition. *A function* $\boldsymbol{F} : \mathcal{D} \to W_2$ *is said to be differentiable at* $X \in \mathcal{D} \subset W_1$, *if there exists a linear transformation* $D\boldsymbol{F}(X) \in \mathcal{L}(W_1, W_2)$ *at* X, *such that* $\forall\, Y \in \mathcal{D}$,

$$\lim_{|Y|\to 0} \frac{1}{|Y|}\, |\boldsymbol{F}(X+Y) - \boldsymbol{F}(X) - D\boldsymbol{F}(X)[Y]| = 0.$$

The linear transformation $D\boldsymbol{F}(X)$ is uniquely determined by the above relation, and it is called the *gradient* of $\boldsymbol{F}$ with respect to X, denoted by $\partial_X \boldsymbol{F}$. We have $\partial_X \boldsymbol{F} \in W_2 \otimes W_1$. The definition is equivalent to the condition: for any Y, we have

$$\boldsymbol{F}(X+Y) - \boldsymbol{F}(X) = \partial_X \boldsymbol{F}(X)[Y] + o(Y), \tag{A.45}$$

or

$$\partial_X \boldsymbol{F}(X)[Y] = \frac{d}{dt}\boldsymbol{F}(X+tY)\Big|_{t=0}. \tag{A.46}$$

For $\phi \in W_2 \otimes W_1$, and $Y \in W_1$, the notation $\phi[Y]$ used in the above relations is self-evident: for $\phi = K \otimes X$,

$$(K \otimes X)[Y] = (X \cdot Y)\, K, \qquad \forall\, K \in W_2,\ X, Y \in W_1.$$

Moreover, for all $\boldsymbol{v}, \boldsymbol{u} \in V$ and $A, S \in \mathcal{L}(V)$, we have

$$\begin{aligned}
&\boldsymbol{v}[\boldsymbol{u}] = \boldsymbol{v} \cdot \boldsymbol{u},\\
&A[\boldsymbol{u}] = A\boldsymbol{u},\\
&A[S] = A \cdot S = \operatorname{tr} AS^T,\\
&(\boldsymbol{v} \otimes \boldsymbol{u})[S] = \boldsymbol{v} \cdot S\boldsymbol{u}.
\end{aligned}$$

Gradients can easily be computed directly from the definition (A.45) or (A.46). We demonstrate this procedure with some examples.

Example A.2.1 Let $\phi : \mathcal{L}(V) \times V \to I\!R$ be defined by

$$\phi(A, \boldsymbol{v}) = \boldsymbol{v} \cdot A\boldsymbol{v}.$$

Then

$$\begin{aligned}
\phi(A, \boldsymbol{v}+\boldsymbol{u}) &= (\boldsymbol{v}+\boldsymbol{u}) \cdot A(\boldsymbol{v}+\boldsymbol{u})\\
&= \boldsymbol{v} \cdot A\boldsymbol{v} + \boldsymbol{v} \cdot A\boldsymbol{u} + \boldsymbol{u} \cdot A\boldsymbol{v} + \boldsymbol{u} \cdot A\boldsymbol{u}\\
&= \phi(A, \boldsymbol{v}) + \partial_{\boldsymbol{v}}\phi[\boldsymbol{u}] + o(\boldsymbol{u}),
\end{aligned}$$

so that

$$\begin{aligned}
\partial_{\boldsymbol{v}}\phi[\boldsymbol{u}] &= \boldsymbol{v} \cdot A\boldsymbol{u} + \boldsymbol{u} \cdot A\boldsymbol{v}\\
&= A^T\boldsymbol{v} \cdot \boldsymbol{u} + A\boldsymbol{v} \cdot \boldsymbol{u}\\
&= (A^T + A)\boldsymbol{v}[\boldsymbol{u}].
\end{aligned}$$

Therefore, we obtain

$$\partial_{\boldsymbol{v}}\phi = (A + A^T)\boldsymbol{v}.$$

Moreover, we have

$$\phi(A + S, \boldsymbol{v}) = \boldsymbol{v} \cdot (A + S)\boldsymbol{v} = \boldsymbol{v} \cdot A\boldsymbol{v} + \boldsymbol{v} \cdot S\boldsymbol{v},$$

which implies

$$\partial_A\phi[S] = \boldsymbol{v} \cdot S\boldsymbol{v} = (\boldsymbol{v} \otimes \boldsymbol{v})[S],$$

so that

$$\partial_A\phi = \boldsymbol{v} \otimes \boldsymbol{v}.$$

□

Example A.2.2 Let $\phi : \mathcal{L}(V) \to I\!R$ be defined for any fixed $\boldsymbol{u}, \boldsymbol{v} \in V$ by

$$\phi(A) = \boldsymbol{u} \cdot A\boldsymbol{v}.$$

From (A.46) we have

$$\partial_A\phi[S] = \frac{d}{dt}\Big(\boldsymbol{u} \cdot (A + tS)\boldsymbol{v}\Big)\Big|_{t=0} = \boldsymbol{u} \cdot S\boldsymbol{v} = (\boldsymbol{u} \otimes \boldsymbol{v})[S],$$

for all $S \in \mathcal{L}(V)$, and we obtain

$$\partial_A\phi = \boldsymbol{u} \otimes \boldsymbol{v}.$$

Now suppose that A is a symmetric tensor, hence the function ϕ is defined on the subspace $Sym(V)$ only,

$$\phi : Sym(V) \to I\!R,$$

and by definition $\partial_A\phi \in Sym(V)$ also. In this case, we have the same relation,

$$\partial_A\phi[S] = (\boldsymbol{u} \otimes \boldsymbol{v})[S],$$

but it holds only for all $S \in Sym(V)$. Therefore we conclude that

$$\partial_A\phi = \frac{1}{2}(\boldsymbol{u} \otimes \boldsymbol{v} + \boldsymbol{v} \otimes \boldsymbol{u}),$$

after symmetrization.

Similarly, if A is a skew-symmetric tensor, then $\partial_A\phi \in Skw(V)$ and the result must be skew-symmetrized,

$$\partial_A\phi = \frac{1}{2}(\boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u}).$$

□

Example A.2.3 We consider trace and determinant function. Since

$$\operatorname{tr}(A+S) = \operatorname{tr} A + \operatorname{tr} S = \operatorname{tr} A + \mathit{1} \cdot S,$$

so that trivially, the gradient of the trace is the identity transformation,

$$\partial_A(\operatorname{tr} A) = \mathit{1}. \tag{A.47}$$

For the gradient of the determinant, we have

$$(\partial_A \det A)[S] = \det(A+S) - \det(A) + o(S).$$

Let ω be a non-trivial alternating n-linear form, then

$$\begin{aligned}&\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)(\partial_A \det A)[S] \\ &\qquad = \omega((A+S)\boldsymbol{v}_1, \cdots, (A+S)\boldsymbol{v}_n) - \omega(A\boldsymbol{v}_1, \cdots, A\boldsymbol{v}_n) + o(S).\end{aligned}$$

By the linearity of ω, after throwing away all the higher order terms into $o(S)$, the right-hand side becomes

$$\begin{aligned}&= \sum_{i=1}^{n} \omega(A\boldsymbol{v}_1, \cdots, S\boldsymbol{v}_i, \cdots, A\boldsymbol{v}_n) + o(S) \\ &= \sum_{i=1}^{n} \omega(A\boldsymbol{v}_1, \cdots, AA^{-1}S\boldsymbol{v}_i, \cdots, A\boldsymbol{v}_n) + o(S) \\ &= (\det A) \sum_{i=1}^{n} \omega(\boldsymbol{v}_1, \cdots, A^{-1}S\boldsymbol{v}_i, \cdots, \boldsymbol{v}_n) + o(S) \\ &= (\det A)(\operatorname{tr} A^{-1}S)\omega(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) + o(S),\end{aligned}$$

by (A.21). Therefore, we have

$$(\partial_A \det A)[S] = (\det A)(\operatorname{tr} SA^{-1}) = (\det A)A^{-T}[S],$$

which implies the following formula,

$$\partial_A \det A = (\det A)A^{-T}. \tag{A.48}$$

□

In differential calculus, we frequently differentiate a composite function by the chain rule. This rule can be stated for composite tensor functions in general. Let W_1, W_2, W_3 be normed spaces of the type (A.41) and $\mathcal{D}_1 \subset W_1$, $\mathcal{D}_2 \subset W_2$ be open subsets, and let

$$\phi : \mathcal{D}_1 \to W_2, \qquad \psi : \mathcal{D}_2 \to W_3,$$

with $\phi(\mathcal{D}_1) \subset \mathcal{D}_2$. Then we have the following

Chain Rule. *Let ϕ be differentiable at $X \in \mathcal{D}_1$, and ψ be differentiable at $Y = \phi(X) \in \mathcal{D}_2$. Then the composition $\boldsymbol{f} = \psi \circ \phi$ is differentiable at X and*

$$D\boldsymbol{f}(X)[Z] = D\psi(\phi(X))[D\phi(X)[Z]], \tag{A.49}$$

for any $Z \in W_1$ or simply

$$D\boldsymbol{f}(X) = D\psi(Y) \circ D\phi(X).$$

Example A.2.4 If ϕ is a scalar-valued function of a vector variable, $\boldsymbol{g}(x)$ is a vector field on $\mathcal{E}$, and $\boldsymbol{h}(\boldsymbol{v})$ is a vector-valued function of a vector variable, then

$$\nabla \boldsymbol{h}(\boldsymbol{g}(x)) = \partial_{\boldsymbol{v}} \boldsymbol{h}\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)} (\nabla \boldsymbol{g}(x)),$$

$$\nabla \phi(\boldsymbol{g}(x)) = \big(\nabla \boldsymbol{g}(x)\big)^T \partial_{\boldsymbol{v}} \phi\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)}.$$

Let us verify the last one in the above formulae. For any $\boldsymbol{u} \in V$, from (A.49),

$$\nabla \phi(\boldsymbol{g}(x))[\boldsymbol{u}] = \partial_{\boldsymbol{v}} \phi\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)} [\nabla \boldsymbol{g}(x)[\boldsymbol{u}]] = \partial_{\boldsymbol{v}} \phi\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)} \cdot \big(\nabla \boldsymbol{g}(x)\big)\boldsymbol{u}$$

$$= \big(\nabla \boldsymbol{g}(x)\big)^T \partial_{\boldsymbol{v}} \phi\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)} \cdot \boldsymbol{u} = \big(\nabla \boldsymbol{g}(x)\big)^T \partial_{\boldsymbol{v}} \phi\Big|_{\boldsymbol{v}=\boldsymbol{g}(x)} [\boldsymbol{u}],$$

where in the third step we have used the definition of transpose (A.8). Note that $\nabla \boldsymbol{h}$, $\nabla \boldsymbol{g}$, and $\partial_{\boldsymbol{v}} \boldsymbol{h}$ are all second order tensors, while $\partial_{\boldsymbol{v}} \phi$ is a vector quantity. □

Another important result in differentiation is the product rule. For tensor functions in general, there are many different products available, for example, the product of a scalar and a vector, the inner product, the tensor product, the action of a tensor on a vector *etc.* These products have one property in common, namely, bilinearity. Therefore, in order to establish a product rule valid for all cases of interest, we consider the bilinear operation

$$\pi : W_1 \times W_2 \longrightarrow W_3$$

which assigns to each $\phi \in W_1$, $\psi \in W_2$, the product $\pi(\phi, \psi) \in W_3$. If ϕ, ψ are two functions,

$$\phi : \mathcal{D} \to W_1, \qquad \psi : \mathcal{D} \to W_2,$$

where $\mathcal{D}$ is an open subset of some normed space W, then the product $\boldsymbol{f} = \pi(\phi, \psi)$ is the function defined by

$$\begin{aligned} &\boldsymbol{f} : \mathcal{D} \longrightarrow W_3 \\ &\boldsymbol{f}(X) = \pi(\phi(X), \psi(X)), \qquad \forall X \in \mathcal{D}. \end{aligned}$$

We then have the following

Product Rule. *Suppose that ϕ and ψ are differentiable at $X \in \mathcal{D} \subset W$, then their product $\boldsymbol{f} = \pi(\phi, \psi)$ is differentiable at X and*

$$D\boldsymbol{f}(X)[V] = \pi(D\phi(X)[V], \psi(X)) + \pi(\phi(X), D\psi(X)[V]), \tag{A.50}$$

for all $V \in W$.

In other words, the derivative of the product $\pi(\phi, \psi)$ is the derivative of π holding ψ fixed plus the derivative of π holding ϕ fixed.

Example A.2.5 Let f be a scalar-valued, and $\boldsymbol{h}$, $\boldsymbol{q}$ be vector-valued functions on $\mathcal{D} \subset W$. For $W = I\!R$, we have

$$\begin{aligned} (f\boldsymbol{h})^{\cdot} &= \dot{f}\boldsymbol{h} + f\dot{\boldsymbol{h}}, \\ (\boldsymbol{q} \cdot \boldsymbol{h})^{\cdot} &= \dot{\boldsymbol{q}} \cdot \boldsymbol{h} + \boldsymbol{q} \cdot \dot{\boldsymbol{h}}. \end{aligned} \tag{A.51}$$

For $W = \mathcal{E}$, we have

$$\begin{aligned} \nabla(f\boldsymbol{h}) &= \boldsymbol{h} \otimes \nabla f + f\nabla\boldsymbol{h}, \\ \nabla(\boldsymbol{q} \cdot \boldsymbol{h}) &= (\nabla\boldsymbol{q})^T\boldsymbol{h} + (\nabla\boldsymbol{h})^T\boldsymbol{q}. \end{aligned} \tag{A.52}$$

For $W = V$, we have

$$\begin{aligned} \partial_{\boldsymbol{v}}(f\boldsymbol{h}) &= \boldsymbol{h} \otimes \partial_{\boldsymbol{v}} f + f\, \partial_{\boldsymbol{v}}\boldsymbol{h}, \\ \partial_{\boldsymbol{v}}(\boldsymbol{q} \cdot \boldsymbol{h}) &= (\partial_{\boldsymbol{v}}\boldsymbol{q})^T\boldsymbol{h} + (\partial_{\boldsymbol{v}}\boldsymbol{h})^T\boldsymbol{q}, \end{aligned} \tag{A.53}$$

Unlike the simple formulae in (A.51), the relations in (A.52) and (A.53) do not look like the familiar product rules, because they have to be consistent with our notation conventions.

Let us demonstrate the first relation of (A.52). By the product rule (A.50), for any $\boldsymbol{w} \in V$, we have

$$\begin{aligned} \nabla(f\boldsymbol{h})[\boldsymbol{w}] &= (\nabla f[\boldsymbol{w}])\boldsymbol{h} + f(\nabla\boldsymbol{h}[\boldsymbol{w}]) = (\nabla f \cdot \boldsymbol{w})\boldsymbol{h} + f(\nabla\boldsymbol{h})\boldsymbol{w} \\ &= (\boldsymbol{h} \otimes \nabla f)\boldsymbol{w} + f(\nabla\boldsymbol{h})\boldsymbol{w} = \big(\boldsymbol{h} \otimes \nabla f + f(\nabla\boldsymbol{h})\big)[\boldsymbol{w}], \end{aligned}$$

where in the third step we have used the definition (A.4). □

If $\boldsymbol{f} : \mathcal{D} \subset U \to W$ is differentiable and its derivative $D\boldsymbol{f}$ is continuous in $\mathcal{D}$, we say that $\boldsymbol{f}$ is of class C^1. The derivative is again a function, $D\boldsymbol{f} : \mathcal{D} \to W \otimes U$, for which we can talk about the differentiability and continuity. We say that $\boldsymbol{f}$ is of class C^2, if $D\boldsymbol{f}$ is of class C^1, and so forth. Frequently, we say a function is *smooth* to mean that it is of class C^k for some $k \geq 1$. We mention the following

Inverse Function Theorem. *Let $\mathcal{D} \subset W$ be an open subset and $\boldsymbol{f} : \mathcal{D} \to W$ be a one-to-one function of class $C^k (k \geq 1)$. Assume that the linear transformation $D\boldsymbol{f}(X) : W \to W$ is invertible at each $X \in \mathcal{D}$, then $\boldsymbol{f}^{-1}$ exists and is of class C^k.*

Example A.2.6 Let $\mathcal{D} \subset \mathcal{E}$ and $\phi : \mathcal{D} \to I\!R$ be of class C^2. Then the second gradient of ϕ is a symmetric tensor, that is, $\nabla(\nabla\phi) \in Sym(V)$.

Indeed, from the definition, we have

$$\nabla\phi(x + \boldsymbol{u}) - \nabla\phi(x) = \nabla(\nabla\phi)[\boldsymbol{u}] + o(\boldsymbol{u}).$$

Taking inner product with $\boldsymbol{v}$, we obtain

$$\nabla\phi(x + \boldsymbol{u})[\boldsymbol{v}] - \nabla\phi(x)[\boldsymbol{v}] = \boldsymbol{v} \cdot \nabla(\nabla\phi)\boldsymbol{u} + o(\boldsymbol{u}),$$

which implies that

$$\begin{aligned} \boldsymbol{v} \cdot \nabla(\nabla\phi)\boldsymbol{u} = \Big(&\phi(x + \boldsymbol{u} + \boldsymbol{v}) - \phi(x + \boldsymbol{u})\Big) \\ &- \Big(\phi(x + \boldsymbol{v}) - \phi(x)\Big) + o(\boldsymbol{u}) + o(\boldsymbol{v}). \end{aligned}$$

Since the right-hand side of the last relation is symmetric in $\boldsymbol{u}$ and $\boldsymbol{v}$, it follows that

$$\boldsymbol{v} \cdot \nabla(\nabla\phi)\boldsymbol{u} = \boldsymbol{u} \cdot \nabla(\nabla\phi)\boldsymbol{v},$$

which proves that the second gradient of ϕ is symmetric. □

Exercise A.2.1 Show that if $Q : I\!R \to \mathcal{O}(V)$ is differentiable, then $\dot{Q}Q^T$ is skew symmetric.

Exercise A.2.2 Let $\boldsymbol{h}(\boldsymbol{v}, A) = (\boldsymbol{v} \cdot A\boldsymbol{v})A^2\boldsymbol{v}$ be a vector function of a vector $\boldsymbol{v}$ and a second order tensor A. Compute $\partial_{\boldsymbol{v}}\boldsymbol{h}$ and $(\partial_A\boldsymbol{h})[S]$ for any $S \in \mathcal{L}(V)$.

Exercise A.2.3 If $A \in \mathcal{L}(V)$ is invertible, show that
1) $(\partial_A A^{-1})[S] = -A^{-1}SA^{-1}$, for any $S \in \mathcal{L}(V)$,
2) $\partial_A \operatorname{tr}(A^{-1}) = -(A^{-2})^T$.

Exercise A.2.4 Let A be a second order tensor. Show that
1) For any positive integer k,

$$\partial_A \operatorname{tr} A^k = k(A^{k-1})^T.$$

2) For principal invariants I_A, II_A, III_A,

$$\begin{aligned} &\partial_A I_A = 1, \\ &\partial_A II_A = (I_A 1 - A)^T, \\ &\partial_A III_A = (II_A 1 - I_A A + A^2)^T. \end{aligned} \tag{A.54}$$

Hint: Calculate $\partial_A \det(A + \lambda 1) = \partial_A(\lambda^3 + I_A\lambda^2 + II_A\lambda + III_A)$.

A.2.3 Coordinate System

Tensor functions can be expressed in terms of components relative to smooth fields of bases in the Euclidean point space $\mathcal{E}$ associated with a coordinate system.

Definition. *Let $\mathcal{D} \subset \mathcal{E}$ be an open set. A coordinate system on $\mathcal{D}$ is a smooth one-to-one mapping*

$$\psi : \mathcal{D} \longrightarrow U,$$

where U is an open set in $I\!R^n$, such that ψ^{-1} is also smooth.

Let $x \in \mathcal{D}$,

$$\psi : x \longmapsto (x^1, \cdots, x^n) = \psi(x).$$

$(x^1, \cdots, x^n)$ is called the (*curvilinear*) *coordinate* of x, and the functions

$$\begin{aligned} &\chi^i : \mathcal{D} \longrightarrow I\!R \\ &\chi^i(x) = x^i, \qquad i = 1, \cdots, n, \end{aligned} \tag{A.55}$$

are called the i^{th} *coordinate function* of ψ. For convenience, we call (x^i) a coordinate system on $\mathcal{D}$.

Let $\chi = \psi^{-1}$, then

$$x = \chi(x^1, \cdots, x^n). \tag{A.56}$$

For $x^1, \cdots, x^n$ fixed, the mapping

$$\begin{aligned} &\lambda_i : I\!R \longrightarrow \mathcal{D} \\ &\lambda_i(t) = \chi(x^1, \cdots, x^i + t, \cdots, x^n), \end{aligned} \tag{A.57}$$

is a curve in $\mathcal{D}$ passing through x at $t = 0$, called the i^{th} *coordinate curve* at x (Fig. A.2). We denote the tangent of this curve at x by $\boldsymbol{e}_i(x)$.

$$\boldsymbol{e}_i(x) = \dot{\lambda}_i(t)\Big|_{t=0} = \frac{\partial \chi}{\partial x^i}\Big|_{(x^1, \cdots, x^n)}. \tag{A.58}$$

Proposition. *The set $\{\boldsymbol{e}_i(x), \ \ i = 1, \cdots, n\}$ forms a basis for the tangent space $\mathcal{E}_x$.*

Proof: For any vector $\boldsymbol{v} \in \mathcal{E}_x$, we can define a curve through x by

$$\lambda(t) = x + t\boldsymbol{v}.$$

Let

$$\lambda(t) = \chi(\lambda^1(t), \cdots, \lambda^n(t)),$$

where $\lambda^i(t)$ are the coordinates of $\lambda(t)$ given by

$$\lambda^i(t) = \chi^i(x + t\boldsymbol{v}). \tag{A.59}$$

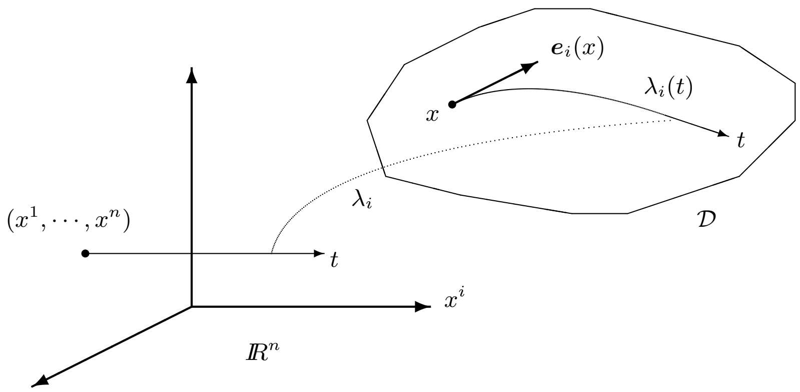


Fig. A.2. Coordinate curve

Then the tangent vector

$$\boldsymbol{v} = \dot{\lambda}(t)\Big|_{t=0} = \frac{\partial \chi}{\partial x^i}\Big|_x \frac{d\lambda^i}{dt}\Big|_{t=0} = \frac{d\lambda^i}{dt}\Big|_{t=0} \boldsymbol{e}_i(x),$$

by (A.58). In other words, $\{\boldsymbol{e}_i(x)\}$ spans the space $\mathcal{E}_x$. □

The set $\{\boldsymbol{e}_i(x)\}$ is a basis of $\mathcal{E}_x$ for each x. This field of bases is called the *natural basis* of the coordinate system (x^i) for V, the translation space of $\mathcal{E}$. The corresponding dual basis of this natural basis is denoted by $\{\boldsymbol{e}^i(x)\}$.

Combining (A.55) and (A.56), we have

$$x^i = \chi^i(\chi(x^1, \cdots, x^n)),$$

which implies

$$\frac{\partial x^i}{\partial x^j} = \delta^i_j = (\nabla \chi^i) \cdot \frac{\partial \chi}{\partial x^j} = (\nabla \chi^i) \cdot \boldsymbol{e}_j(x),$$

by (A.58). Therefore, the two natural bases of the coordinate system (x^i) are given by the following relations:

$$\boldsymbol{e}_i(x) = \frac{\partial \chi}{\partial x^i}\Big|_x, \qquad \boldsymbol{e}^i(x) = \nabla \chi^i(x). \tag{A.60}$$

The inner products,

$$g_{ij}(x) = \boldsymbol{e}_i(x) \cdot \boldsymbol{e}_j(x), \qquad g^{ij}(x) = \boldsymbol{e}^i(x) \cdot \boldsymbol{e}^j(x),$$

are called the *metric tensors* of the coordinate system.

Now let us consider change of coordinate systems. Let (x^i) and $(\bar{x}^i)$ be two coordinate systems on $\mathcal{D}$, and $\{\boldsymbol{e}_i(x)\}$, $\{\bar{\boldsymbol{e}}_i(x)\}$ be the corresponding natural bases. Suppose that the coordinate transformations are given by

$$x^i = x^i(\bar{x}^1, \cdots, \bar{x}^n),$$
$$\bar{x}^k = \bar{x}^k(x^1, \cdots, x^n).$$

Then by taking the gradients, one immediately obtain the change of the corresponding natural bases given by

$$\boldsymbol{e}^i(x) = \frac{\partial x^i}{\partial \bar{x}^k}\bar{\boldsymbol{e}}^k(x), \qquad \boldsymbol{e}_i(x) = \frac{\partial \bar{x}^k}{\partial x^i}\bar{\boldsymbol{e}}_k(x). \tag{A.61}$$

Comparing the change of bases considered in Sect. A.1.4, $[\partial x^i/\partial \bar{x}^k]$ plays the role of the transformation matrix $[M_k{}^i]$ in (A.12), and hence, the transformation rules (A.14) for the components of an arbitrary tensor in the change of coordinate system becomes

$$\bar{A}^i{}_j = A^k{}_l \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^l}{\partial \bar{x}^j}. \tag{A.62}$$

For other components of tensors in general, the transformation rules are similar.

Example A.2.7 Let us consider a deformation $\kappa : \mathcal{D} \to \mathcal{E}$,

$$\kappa(x) = \tilde{x}.$$

Let (x^i) be a coordinate system on $\mathcal{D}$, and $(\tilde{x}^\alpha)$ be a coordinate system on $\kappa(\mathcal{D})$,

$$x = \chi(x^1, \cdots, x^n), \qquad \tilde{x} = \tilde{\chi}(\tilde{x}^1, \cdots, \tilde{x}^n).$$

The deformation κ is usually expressed explicitly in the form,

$$\tilde{x}^\alpha = \kappa^\alpha(x^1, \cdots, x^n), \qquad \alpha = 1, \cdots, n. \tag{A.63}$$

Using the chain rule, we obtain, with $x^i = \chi^i(x)$,

$$\nabla\kappa(x) = \frac{\partial \tilde{\chi}}{\partial \tilde{x}^\alpha}\bigg|_{\tilde{x}} \frac{\partial \kappa^\alpha}{\partial x^i}\bigg|_x \nabla\chi^i(x),$$

which by (A.60) becomes

$$\nabla\kappa(x) = \frac{\partial \kappa^\alpha}{\partial x^i}\bigg|_x \tilde{\boldsymbol{e}}_\alpha(\kappa(x)) \otimes \boldsymbol{e}^i(x).$$

This is the component form of the deformation gradient $\nabla\kappa(x)$ in terms of two different coordinate systems (x^i) and $(\tilde{x}^\alpha)$. With respect to these two natural bases at two different points, namely, x and $\kappa(x)$, the components of the deformation gradient are just the partial derivatives of the deformation function (A.63), which can most easily be calculated. Other component forms of $\nabla\kappa$ can be obtained through the metric tensors and by the change of bases relative to the coordinate systems. □

A.2.4 Covariant Derivatives

We shall now consider the component form of the gradient of a tensor field in general relative to the natural basis of a coordinate system. Let (x^i) be a coordinate system on $\mathcal{D} \subset \mathcal{E}$, and $\{\boldsymbol{e}_i(x)\}$, $\{\boldsymbol{e}^i(x)\}$ be its natural bases.

To begin with, let us consider a scalar field, $f : \mathcal{D} \to I\!R$, the gradient of f is then a vector field. By (A.44), (A.57), and (A.58) we have

$$\begin{aligned}(\nabla f(x)) \cdot \boldsymbol{e}_i(x) &= \lim_{t\to 0} \frac{1}{t} \left(f(x + t\boldsymbol{e}_i) - f(x)\right) \\ &= \lim_{t\to 0} \frac{1}{t} \big(f(\chi(x^1, \cdots, x^i + t, \cdots, x^n)) - f(\chi(x^1, \cdots, x^n))\big) \\ &= \frac{\partial(f \circ \chi)}{\partial x^i}\Big|_{(x^1,\cdots,x^n)},\end{aligned}$$

which are the covariant components of ∇f.

Usually, we shall write $f(\chi(x^1, \cdots, x^n))$ as $f(x^1, \cdots, x^n)$ for simplicity. Therefore, the component form of the gradient of $f(x)$ becomes

$$\nabla f(x) = \frac{\partial f}{\partial x^i}\Big|_x \boldsymbol{e}^i(x). \tag{A.64}$$

In other words, for the gradient of a scalar field f, its covariant component relative to the natural basis, $(\nabla f)_i$, is *just* the partial derivative relative to the coordinate x^i.

Now let us consider the gradients of natural bases themselves. For each i fixed, $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{e}^i\}$ can be regarded as vector fields on $\mathcal{D}$,

$$\boldsymbol{e}_i : x \in \mathcal{D} \longmapsto \boldsymbol{e}_i(x) \in \mathcal{E}_x.$$

Let us denote the gradients of natural bases by

$$\begin{aligned}\Gamma_i(x) &= \nabla \boldsymbol{e}_i(x) \in \mathcal{E}_x \otimes \mathcal{E}_x, \\ \Gamma^i(x) &= \nabla \boldsymbol{e}^i(x) \in \mathcal{E}_x \otimes \mathcal{E}_x.\end{aligned} \tag{A.65}$$

We write

$$\Gamma_i = \Gamma_i{}^j{}_k \boldsymbol{e}_j \otimes \boldsymbol{e}^k, \qquad \Gamma^i = \Gamma^i{}_{jk} \boldsymbol{e}^j \otimes \boldsymbol{e}^k. \tag{A.66}$$

The components $\Gamma_i{}^j{}_k$ and $\Gamma^i{}_{jk}$ are called the *Christoffel symbols*. Note that $\Gamma_i{}^j{}_k$ and $\Gamma^i{}_{jk}$ are not the associated components of a third order tensor.

By taking the gradient of $(\boldsymbol{e}^i(x) \cdot \boldsymbol{e}_j(x))$, one can obtain the relation,

$$\Gamma_j{}^i{}_k = -\Gamma^i{}_{jk}. \tag{A.67}$$

Moreover, since $\Gamma^i = \nabla(\nabla\chi^i(x))$ by $(\mathrm{A}.60)_1$ and the second gradient is a symmetric tensor, we have the following symmetry conditions,

$$\Gamma^i{}_{jk} = \Gamma^i{}_{kj}, \qquad \Gamma_j{}^i{}_k = \Gamma_k{}^i{}_j. \tag{A.68}$$

Since both Christoffel symbols are related in such a simple manner, usually only one is in use, namely, $\Gamma_j{}^i{}_k$, and it is called the Christoffel symbol of the second kind in classical tensor analysis.

Now let us calculate the gradient of a vector field in terms of the coordinate system. Suppose that $\boldsymbol{v}(x)$ is a vector field and

$$\boldsymbol{v}(x) = v^i(x)\boldsymbol{e}_i(x) = v_i(x)\boldsymbol{e}^i(x).$$

Then by $(A.52)_1$, (A.64), (A.65), and (A.66), we have

$$\begin{aligned}\nabla\boldsymbol{v} &= \nabla(v^i\boldsymbol{e}_i)\\ &= \boldsymbol{e}_i \otimes \nabla v^i + v^i\nabla\boldsymbol{e}_i\\ &= \boldsymbol{e}_i \otimes \frac{\partial v^i}{\partial x^k}\boldsymbol{e}^k + v^i\Gamma_i{}^j{}_k\boldsymbol{e}_j \otimes \boldsymbol{e}^k\\ &= \Big(\frac{\partial v^j}{\partial x^k} + v^i\Gamma_i{}^j{}_k\Big)\boldsymbol{e}_j \otimes \boldsymbol{e}^k.\end{aligned}$$

Hence, the gradient of $\boldsymbol{v}(x)$ has the component form,

$$\nabla\boldsymbol{v} = v^j{}_{,k}\boldsymbol{e}_j \otimes \boldsymbol{e}^k,$$

where

$$v^j{}_{,k} = \frac{\partial v^j}{\partial x^k} + v^i\,\Gamma_i{}^j{}_k. \tag{A.69}$$

Similarly, we also have

$$\nabla\boldsymbol{v} = v_{j,k}\boldsymbol{e}^j \otimes \boldsymbol{e}^k,$$

where

$$v_{j,k} = \frac{\partial v_j}{\partial x^k} - v_i\,\Gamma_j{}^i{}_k. \tag{A.70}$$

Here the relation (A.67) has been used.

$v^j{}_{,k}$ and $v_{j,k}$ are the mixed and the covariant components of $\nabla\boldsymbol{v}$. The comma, stands for the operation called the *covariant derivative*, since it increases the covariant order of the components by one.

More generally, suppose that A is a second order tensor field, then ∇A is a third order tensor field which has the following component form,

$$\nabla A = A^i{}_{j,k}\,\boldsymbol{e}_i \otimes \boldsymbol{e}^j \otimes \boldsymbol{e}^k,$$

where

$$A^i{}_{j,k} = \frac{\partial A^i{}_j}{\partial x^k} + A^l{}_j \, \Gamma_l{}^i{}_k - A^i{}_l \, \Gamma_j{}^l{}_k. \tag{A.71}$$

Covariant derivatives of other components can easily be written down using the same recipes for covariant and contravariant components respectively.

We have seen in (A.7) that the components of the metric tensor, $g_{ij}(x)$ and $g^{ij}(x)$, are also the components of the identity tensor, therefore their covariant derivatives must vanish,

$$g_{ij,k} = 0, \qquad g^{ij}{}_{,k} = 0. \tag{A.72}$$

Consequently by (A.24), the covariant derivative of the volume tensor also vanish,

$$e_{ijk,l} = 0, \qquad e^{ijk}{}_{,l} = 0.$$

In other words, the components of the metric tensor and the volume tensor behave like constant tensors in covariant derivation although they are in general functions of x.

From $(A.72)_1$, we can derive a formula for the determination of the Christoffel symbols in terms of the metric tensor. By (A.71) we have

$$\frac{\partial g_{ij}}{\partial x^k} = g_{lj} \Gamma_i{}^l{}_k + g_{il} \Gamma_j{}^l{}_k.$$

Rotating the indices (i, j, k) of this relation, then adding two of the three resulting equations and subtracting the remaining one, we get

$$2 \, g_{lj} \Gamma_i{}^l{}_k = \Big(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ij}}{\partial x^k} - \frac{\partial g_{ik}}{\partial x^j} \Big).$$

Hence, we have the following formula:

$$\Gamma_i{}^j{}_k = \frac{1}{2} \, g^{jl} \Big(\frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l} \Big). \tag{A.73}$$

The Christoffel symbols are not components of a third order tensor. For two coordinate systems (x^i) and $(\bar{x}^i)$, they have the following transformation rules:

$$\bar{\Gamma}_i{}^j{}_k = \Gamma_r{}^s{}_t \frac{\partial x^r}{\partial \bar{x}^i} \frac{\partial \bar{x}^j}{\partial x^s} \frac{\partial x^t}{\partial \bar{x}^k} + \frac{\partial^2 x^r}{\partial \bar{x}^i \partial \bar{x}^k} \frac{\partial \bar{x}^j}{\partial x^r}.$$

A.2.5 Other Differential Operators

Divergence and curl of a vector field can be defined in the usual way and their definitions can be adopted also for tensor fields.

Definition. *The divergence of a vector field* $\boldsymbol{u}$ *is a scalar field defined by*

$$\operatorname{div} \boldsymbol{u} = \operatorname{tr}(\nabla \boldsymbol{u}). \tag{A.74}$$

In component form,

$$\operatorname{div} \boldsymbol{u} = u^i{}_{,i}.$$

Definition. *The curl (or rotation) of* $\boldsymbol{u}$ *is a vector field defined by*

$$\operatorname{curl} \boldsymbol{u} = \langle \nabla \boldsymbol{u}^T - \nabla \boldsymbol{u} \rangle.$$

In component form,

$$\operatorname{curl} \boldsymbol{u} = e^{ijk} u_{k,j} \boldsymbol{e}_i.$$

Here duality map defined in (A.29) is employed and according to (A.30) $\operatorname{curl} \boldsymbol{u}$ is the axial vector of the skew-symmetric part of the gradient of $(-2\boldsymbol{u})$. One can easily verify the following condition:

$$\boldsymbol{v} \cdot \operatorname{curl} \boldsymbol{u} = \operatorname{div}(\boldsymbol{u} \times \boldsymbol{v}),$$

for any constant vector field $\boldsymbol{v}$. This condition can be used as the definition for the curl operator. In a similar manner, we can define the divergence of a second order tensor in terms of the divergence of a vector.

Definition. *The divergence of a second order tensor field* S *is a vector field defined by the condition: for any constant vector field* $\boldsymbol{v}$,

$$\boldsymbol{v} \cdot \operatorname{div} S = \operatorname{div}(S^T \boldsymbol{v}). \tag{A.75}$$

In component form, we have

$$\operatorname{div} S = S^{ij}{}_{,j} \boldsymbol{e}_i.$$

Definition. *The Laplacian of a scalar (or vector) field* ϕ*, denoted by* $\nabla^2 \phi$*, is a scalar (or vector) field defined by*

$$\nabla^2 \phi = \operatorname{div}(\nabla \phi).$$

In component form, if ϕ is a scalar field,

$$\nabla^2 \phi = g^{jk} (\phi_{,j})_{,k} = g^{jk} \phi_{,jk}.$$

If $\phi = \boldsymbol{h}$ is a vector field,

$$\nabla^2 \boldsymbol{h} = g^{jk} h^i{}_{,jk} \boldsymbol{e}_i.$$

In the above expressions, the comma denotes the covariant derivative.

Example A.2.8 Let f and $\boldsymbol{u}$, $\boldsymbol{v}$ be scalar and vector fields respectively. Then we can show the following relations:

$$\begin{aligned} &\operatorname{div}(f\boldsymbol{u}) = \boldsymbol{u} \cdot \nabla f + f \ \operatorname{div} \boldsymbol{u}, \\ &\operatorname{div}(\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{v} \cdot \operatorname{curl} \boldsymbol{u} - \boldsymbol{u} \cdot \operatorname{curl} \boldsymbol{v}, \\ &\nabla^2(\boldsymbol{u} \cdot \boldsymbol{v}) = \nabla^2 \boldsymbol{u} \cdot \boldsymbol{v} + 2 \nabla \boldsymbol{u} \cdot \nabla \boldsymbol{v} + \boldsymbol{u} \cdot \nabla^2 \boldsymbol{v}. \end{aligned} \tag{A.76}$$

Let us verify the first relation.

$$\begin{aligned}\operatorname{div}(f\boldsymbol{u}) &= \operatorname{tr}(\nabla(f\boldsymbol{u}))\\ &= \operatorname{tr}\Big(\boldsymbol{u}\otimes\nabla f + f(\nabla\boldsymbol{u})\Big)\\ &= \operatorname{tr}(\boldsymbol{u}\otimes\nabla f) + f\ \operatorname{tr}(\nabla\boldsymbol{u}),\end{aligned}$$

which gives $(A.76)_1$. In this calculation, we have used the definition (A.74), the relation $(A.52)_1$, and the linearity of the trace operator.

Verification of the other relations in (A.76) may not be so straightforward in *direct notation.* And more annoyingly, these relations as well as the relations (A.52) and (A.53) are not easy to memorize. Nevertheless, if we express all of these relations in *index notation*, they all become trivially simple. Indeed, (A.76) may be written out straightly as:

$$\begin{aligned}&(fu^i)_{,i} = f_{,i}u^i + fu^i{}_{,i},\\ &(g^{il}e_{ljk}u^jv^k)_{,i} = g^{il}e_{ljk}u^j{}_{,i}v^k + g^{il}e_{ljk}u^jv^k{}_{,i},\\ &g^{jk}(u^iv_i)_{,jk} = g^{jk}(u^i{}_{,j}v_i + u^iv_{i,j})_{,k}\\ &\qquad\qquad = g^{jk}u^i{}_{,jk}v_i + 2\,g^{jk}u^i{}_{,k}v_{i,j} + g^{jk}u^iv_{i,jk}.\end{aligned}$$

which are merely the usual product rules of differentiating scalar functions and the symmetry of second gradient. The only difference here is that the comma denotes the covariant derivative instead of the usual partial derivative. □

Remark. From the observation made in the above example, the use of index notation is often encouraged, especially when complicated calculations are involved. In arbitrary curvilinear coordinate systems, contravariant and covariant indices must be carefully distinguished and the pair of repeated indices, for which the summation convention is applied, must always appear in different levels. An index can be raised or lowered to its proper level with the metric tensor g_{ij} or g^{ij}. Moreover, since the gradients of the metric tensor and the volume tensor vanish, in covariant differentiation, the metric tensor g_{ij} as well as the components of the volume element e_{ijk} can be treated as constants. Furthermore, if Cartesian coordinate system is used, there is no difference between contravariant and covariant components and hence all the indices can be written at the same level, and more conveniently, the covariant derivative becomes the partial derivative and $g_{ij} = \delta_{ij}$, $e_{ijk} = \varepsilon_{ijk}$ are constants.

It is important to note that given an expression in index notation, one can always turn it into an expression in direct notation or *vice versa.* Therefore, in handling calculations, the choice of using direct notation or index notation, or even using Cartesian index notation is totally up to one's taste and convenience.

We shall also mention some important theorems of integral calculus often used in mechanics.

Divergence Theorem. *Let $\mathcal{R}$ be a bounded regular region*[2] *in $\mathcal{E}$, and let $\phi : \mathcal{R} \to I\!R$, $\boldsymbol{h} : \mathcal{R} \to V$, $S : \mathcal{R} \to \mathcal{L}(V)$ be smooth fields. Then*

$$\int_{\partial\mathcal{R}} \phi \boldsymbol{n}\, da = \int_{\mathcal{R}} \nabla\phi\, dv,$$

$$\int_{\partial\mathcal{R}} \boldsymbol{v} \cdot \boldsymbol{n}\, da = \int_{\mathcal{R}} \operatorname{div} \boldsymbol{v}\, dv, \qquad (A.77)$$

$$\int_{\partial\mathcal{R}} S\boldsymbol{n}\, da = \int_{\mathcal{R}} \operatorname{div} S\, dv,$$

where $\boldsymbol{n}$ is the outward unit normal field on $\partial\mathcal{R}$.

Proof: The relations $(A.77)_{1,2}$ are well-known classical results. To show $(A.77)_3$, let $\boldsymbol{v}$ be an arbitrary constant vector. Then

$$\begin{aligned} \boldsymbol{v} \cdot \int_{\partial\mathcal{R}} S\boldsymbol{n}\, da &= \int_{\partial\mathcal{R}} \boldsymbol{v} \cdot S\boldsymbol{n}\, da = \int_{\partial\mathcal{R}} S^T\boldsymbol{v} \cdot \boldsymbol{n}\, da \\ &= \int_{\mathcal{R}} \operatorname{div}(S^T\boldsymbol{v})\, dv = \int_{\mathcal{R}} \boldsymbol{v} \cdot \operatorname{div} S\, dv \\ &= \boldsymbol{v} \cdot \int_{\mathcal{R}} \operatorname{div} S\, dv, \end{aligned}$$

where we have used $(A.77)_2$ and the definition (A.75). □

Proposition. *Let $\phi : \mathcal{D} \to W$ be a continuous function on an open set $\mathcal{D}$ in $\mathcal{E}$. If*

$$\int_{\mathcal{N}} \phi\, dv = 0,$$

for any $\mathcal{N} \subset \mathcal{D}$, then ϕ is identically zero in $\mathcal{D}$, i.e.,

$$\phi(x) = 0, \qquad \forall x \in \mathcal{D}.$$

Proof: Suppose that $\phi(x_\circ) \neq 0$ for some $x_\circ \in \mathcal{D}$, then since ϕ is continuous, there exists a small neighborhood $\mathcal{N} \subset \mathcal{D}$ containing $x_\circ$, such that $\phi(x) \neq 0$, $\forall x \in \mathcal{N}$. Therefore, by the mean value theorem of integral calculus,

$$\int_{\mathcal{N}} \phi\, dv = K\phi(\bar{x}) \neq 0,$$

[2] A regular region, roughly speaking, is a closed region with piecewise smooth boundary.

for some $\bar{x} \in \mathcal{N}$, where K denotes the volume of $\mathcal{N}$. This contradicts the hypothesis. □

This proposition and the divergence theorem enable us to deduce local field equations from the integral balance laws.

Exercise A.2.5 Let f, $\boldsymbol{u}, \boldsymbol{v}$, and S be smooth scalar, vector, and second order tensor fields. Verify the following identities:
1) $\operatorname{div}(S\boldsymbol{u}) = \boldsymbol{u} \cdot \operatorname{div} S^T + \operatorname{tr}(S\nabla\boldsymbol{u})$,
2) $\operatorname{div}(fS) = S\,\nabla f + f \operatorname{div} S$,
3) $\operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{v}) = (\nabla\boldsymbol{u})\boldsymbol{v} + \boldsymbol{u} \operatorname{div} \boldsymbol{v}$,
4) $\operatorname{div}(\nabla\boldsymbol{u})^T = \nabla(\operatorname{div}\boldsymbol{u})$.

Exercise A.2.6 Let f and $\boldsymbol{v}$ be smooth scalar and vector fields respectively. Show that
1) $\operatorname{curl}\nabla f = 0$,
2) $\operatorname{div}\operatorname{curl}\boldsymbol{v} = 0$,
3) If $\operatorname{div}\boldsymbol{v} = 0$ and $\operatorname{curl}\boldsymbol{v} = 0$, then $\nabla^2\boldsymbol{v} = 0$.

Exercise A.2.7 Let $\boldsymbol{v}$ and S be smooth vector and tensor field on a bound regular region $\mathcal{R}$ respectively. Show that
1) $\int_{\partial\mathcal{R}} \boldsymbol{v} \otimes \boldsymbol{n}\, da = \int_{\mathcal{R}} \nabla\boldsymbol{v}\, dv$,
2) $\int_{\partial\mathcal{R}} \boldsymbol{v} \otimes S\boldsymbol{n}\, da = \int_{\mathcal{R}} \Big(\boldsymbol{v} \otimes \operatorname{div} S + (\nabla\boldsymbol{v})S^T\Big)\, dv$.

A.2.6 Physical Components

Let (x^i) be a coordinate system on $\mathcal{E}$ and $\{\boldsymbol{e}_i(x)\}$ and $\{\boldsymbol{e}^i(x)\}$ be its natural bases. The system (x^i) is called an *orthogonal coordinate system* if the metric tensor

$$g_{ij}(x) = 0, \quad \text{for} \quad i \neq j, \ \ \forall x \in \mathcal{E}.$$

For an orthogonal coordinate system, we can define a field of orthonormal basis, denoted by $\{\boldsymbol{e}_{\langle i\rangle}(x)\}$, by normalizing the natural basis,

$$\boldsymbol{e}_{\langle i\rangle} = \frac{\boldsymbol{e}_i}{|\boldsymbol{e}_i|}. \qquad \text{(no sum)}$$

In this expression the summation notation is not invoked as indicated explicitly. Since

$$|\boldsymbol{e}_i| = \sqrt{\boldsymbol{e}_i \cdot \boldsymbol{e}_i} = \sqrt{g_{ii}}, \qquad \text{(no sum)}$$

therefore,

$$\begin{aligned} \boldsymbol{e}_{\langle i\rangle} &= \frac{\boldsymbol{e}_i}{\sqrt{g_{ii}}} = \frac{\boldsymbol{e}^i}{\sqrt{g^{ii}}} \\ &= \sqrt{g^{ii}}\, \boldsymbol{e}_i = \sqrt{g_{ii}}\, \boldsymbol{e}^i. \qquad \text{(no sum)} \end{aligned}$$

Here we have noted that normalization of the two dual natural bases of an orthogonal coordinate system gives rise to the same orthonormal basis.

The components of a tensor field relative to the orthonormal basis $\{\boldsymbol{e}_{\langle i\rangle}(x)\}$ are called the *physical components* in the coordinate system (x^i). For a vector field $\boldsymbol{v}$,

$$\boldsymbol{v} = v^i \boldsymbol{e}_i = v_i \boldsymbol{e}^i = v_{\langle i\rangle} \boldsymbol{e}_{\langle i\rangle}.$$

The physical components $v_{\langle i\rangle}$ are given by

$$v_{\langle i\rangle} = \sqrt{g_{ii}}\, v^i = \frac{v_i}{\sqrt{g_{ii}}}. \qquad \text{(no sum)} \qquad (A.78)$$

For a second order tensor field T,

$$\begin{aligned} T &= T^{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j = T_{ij} \boldsymbol{e}^i \otimes \boldsymbol{e}^j = T^i{}_j \boldsymbol{e}_i \otimes \boldsymbol{e}^j \\ &= T_{\langle ij\rangle} \boldsymbol{e}_{\langle i\rangle} \otimes \boldsymbol{e}_{\langle j\rangle}. \end{aligned}$$

The physical components $T_{\langle ij\rangle}$ are given by

$$T_{\langle ij\rangle} = \sqrt{g_{ii}}\sqrt{g_{jj}}\, T^{ij} = \frac{T_{ij}}{\sqrt{g_{ii}}\sqrt{g_{jj}}} = \frac{\sqrt{g_{ii}}}{\sqrt{g_{jj}}}\, T^i{}_j. \qquad \text{(no sum)} \qquad (A.79)$$

In particular, we have $g_{\langle ij\rangle} = \delta_{ij}$.

The advantage of using physical components is obvious in practical applications. Since the norms of the basis vectors of the natural basis in general vary from point to point in $\mathcal{E}$, hence it is inconvenient for the measurement of physical quantities relative to this basis.

A.2.7 Orthogonal Coordinate Systems

We now consider three orthogonal coordinate systems most commonly used: the Cartesian, the cylindrical, and the spherical coordinate systems and derive their basic characteristics.

a) Cartesian Coordinate System

Fix a point o in $\mathcal{E}$. Let $\{\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3\}$ be an orthonormal basis of V. For any $x \in \mathcal{E}$, then $x - o \in V$. We write

$$x - o = x_i\, \boldsymbol{i}_i.$$

Clearly, this defines a coordinate system

$$x \longmapsto (x_1, x_2, x_3)$$

with $\{\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3\}$ as its natural basis, which is of course independent of $x \in \mathcal{E}$. We call such a system a *Cartesian coordinate system.*

For a Cartesian coordinate system, we have

$$g_{ij}(x) = \delta_{ij}, \qquad \forall\, x \in \mathcal{E},$$

and hence by (A.73)

$$\Gamma_{j\ k}^{\ i}(x) = 0.$$

It is also a custom to write the basis $\{\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3\}$ as $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ and the coordinate (x_1, x_2, x_3) as (x, y, z) for a Cartesian coordinate system.

b) Cylindrical Coordinate System

The cylindrical coordinate system (r, θ, z) is defined as

$$x = \chi(r, \theta, z),$$

by the following coordinate transformation (see Fig. A.3 (a)),

$$\begin{aligned} x_1 &= r\,\cos\theta, \qquad && r > 0 \\ x_2 &= r\,\sin\theta, \qquad && 0 < \theta < 2\pi \\ x_3 &= z, \end{aligned} \tag{A.80}$$

where $x = (x_1, x_2, x_3)$ is the Cartesian coordinate system.

The natural bases are denoted by $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\}$ and $\{\boldsymbol{e}^r, \boldsymbol{e}^\theta, \boldsymbol{e}^z\}$. From (A.80) and $(A.60)_2$, we can determine the basis in terms of the Cartesian components.

$$\begin{aligned} \boldsymbol{e}_r &= \frac{\partial \chi}{\partial r} = \cos\theta\,\boldsymbol{i}_1 + \sin\theta\,\boldsymbol{i}_2, \\ \boldsymbol{e}_\theta &= \frac{\partial \chi}{\partial \theta} = -r\,\sin\theta\,\boldsymbol{i}_1 + r\,\cos\theta\,\boldsymbol{i}_2, \\ \boldsymbol{e}_z &= \frac{\partial \chi}{\partial z} = \boldsymbol{i}_3. \end{aligned}$$

Therefore, we obtain the matrix representations of the metric tensor in the cylindrical coordinate system,

$$[g_{ij}] = \begin{bmatrix} 1 & & \\ & r^2 & \\ & & 1 \end{bmatrix}, \qquad [g^{ij}] = \begin{bmatrix} 1 & & \\ & r^{-2} & \\ & & 1 \end{bmatrix},$$

and the Christoffel symbols by (A.73),

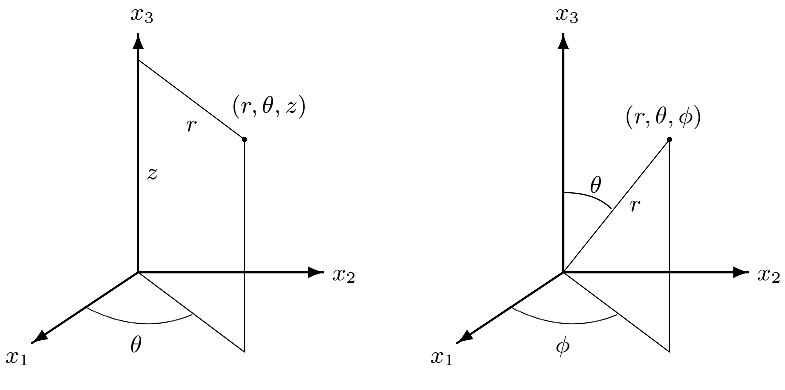
$$\begin{aligned} &\Gamma_{r\ \theta}^{\ \theta} = \Gamma_{\theta\ r}^{\ \theta} = \frac{1}{r}, \\ &\Gamma_{\theta\ \theta}^{\ r} = -r, \\ &\text{others } = 0. \end{aligned}$$

Moreover, we have

$$e_r = e^r, \quad e_\theta = r^2\, e^\theta, \quad e_z = e^z,$$

and

$$\begin{aligned} e_{\langle r\rangle} &= \cos\theta\, i_1 + \sin\theta\, i_2, \\ e_{\langle \theta\rangle} &= -\sin\theta\, i_1 + \cos\theta\, i_2, \\ e_{\langle z\rangle} &= i_3. \end{aligned}$$



(a) Cylindrical (b) Spherical

Fig. A.3. Coordinate systems

c) Spherical Coordinate System

The spherical coordinate system (r, θ, ϕ) is defined as

$$x = \chi(r, \theta, \phi),$$

by the following coordinate transformation (see Fig. A.3 (b)),

$$\begin{aligned} x_1 &= r\, \sin\theta\, \cos\phi, \qquad && r > 0 \\ x_2 &= r\, \sin\theta\, \sin\phi, && 0 < \theta < \pi \\ x_3 &= r\, \cos\theta, && 0 < \phi < 2\pi \end{aligned}$$

where $x = (x_1, x_2, x_3)$ is the Cartesian coordinate system.

The natural bases are denoted by $\{e_r, e_\theta, e_\phi\}$ and $\{e^r, e^\theta, e^\phi\}$. We have

$$\begin{aligned} e_r &= \sin\theta\, \cos\phi\, i_1 + \sin\theta\, \sin\phi\, i_2 + \cos\theta\, i_3, \\ e_\theta &= r\, \cos\theta\, \cos\phi\, i_1 + r\, \cos\theta\, \sin\phi\, i_2 - r\, \sin\theta\, i_3, \\ e_\phi &= -r\, \sin\theta\, \sin\phi\, i_1 + r\, \sin\theta\, \cos\phi\, i_2, \end{aligned}$$

and

$$\boldsymbol{e}_r = \boldsymbol{e}^r, \quad \boldsymbol{e}_\theta = r^2\, \boldsymbol{e}^\theta, \quad \boldsymbol{e}_\phi = r^2 \sin^2\theta\, \boldsymbol{e}^\phi.$$

The matrix representations of the metric tensor has the forms

$$[g_{ij}] = \begin{bmatrix} 1 & & \\ & r^2 & \\ & & r^2\sin^2\theta \end{bmatrix}, \qquad [g^{ij}] = \begin{bmatrix} 1 & & \\ & r^{-2} & \\ & & (r\,\sin\theta)^{-2} \end{bmatrix},$$

and the Christoffel symbols are

$$\begin{aligned}
&\Gamma_{r\ \theta}^{\ \theta} = \Gamma_{\theta\ r}^{\ \theta} = \Gamma_{r\ \phi}^{\ \phi} = \Gamma_{\phi\ r}^{\ \phi} = \frac{1}{r},\\
&\Gamma_{\theta\ \theta}^{\ r} = -r,\\
&\Gamma_{\phi\ \phi}^{\ r} = -r\,\sin^2\theta,\\
&\Gamma_{\theta\ \phi}^{\ \phi} = \Gamma_{\phi\ \theta}^{\ \phi} = \cot\theta,\\
&\Gamma_{\phi\ \phi}^{\ \theta} = -\sin\theta\,\cos\theta,\\
&\text{others } = 0,
\end{aligned}$$

Moreover, the orthonormal basis for the physical components are

$$\begin{aligned}
\boldsymbol{e}_{\langle r\rangle} &= \sin\theta\,\cos\phi\,\boldsymbol{i}_1 + \sin\theta\,\sin\phi\,\boldsymbol{i}_2 + \cos\theta\,\boldsymbol{i}_3,\\
\boldsymbol{e}_{\langle \theta\rangle} &= \cos\theta\,\cos\phi\,\boldsymbol{i}_1 + \cos\theta\,\sin\phi\,\boldsymbol{i}_2 - \sin\theta\,\boldsymbol{i}_3,\\
\boldsymbol{e}_{\langle \phi\rangle} &= -\sin\phi\,\boldsymbol{i}_1 + \cos\phi\,\boldsymbol{i}_2.
\end{aligned}$$

Remark. More frequently, we would like to express quantities in these coordinate systems in terms of their physical components. A simple way to do this is to derive the expressions first in terms of contravariant or covariant components and then convert them into physical components using relations like (A.78) and (A.79).

Example A.2.9 Let us calculate the Laplacian of a scalar field Φ in the spherical coordinate system. We have

$$\begin{aligned}
\Phi_{,j} &= \frac{\partial\Phi}{\partial x^j},\\
\Phi_{,jk} &= \frac{\partial^2\Phi}{\partial x^j\partial x^k} - \frac{\partial\Phi}{\partial x^i}\Gamma_{j\ k}^{\ i},
\end{aligned}$$

from which we obtain the following covariant components:

$$\begin{aligned}
\Phi_{,rr} &= \frac{\partial^2 \Phi}{\partial r^2}, \\
\Phi_{,\theta\theta} &= \frac{\partial^2 \Phi}{\partial \theta^2} - \frac{\partial \Phi}{\partial r}\Gamma_{\theta}{}^{r}{}_{\theta} = \frac{\partial^2 \Phi}{\partial \theta^2} + r\frac{\partial \Phi}{\partial r}, \\
\Phi_{,\phi\phi} &= \frac{\partial^2 \Phi}{\partial \phi^2} - \frac{\partial \Phi}{\partial r}\Gamma_{\phi}{}^{r}{}_{\phi} - \frac{\partial \Phi}{\partial \theta}\Gamma_{\phi}{}^{\theta}{}_{\phi} = \frac{\partial^2 \Phi}{\partial \phi^2} + r\sin^2\theta\frac{\partial \Phi}{\partial r} + \sin\theta\cos\theta\frac{\partial \Phi}{\partial \theta}.
\end{aligned}$$

We have $\Phi_{,rr} = \Phi_{,\langle rr\rangle}$, $\Phi_{,\theta\theta} = r^2\Phi_{,\langle\theta\theta\rangle}$, $\Phi_{,\phi\phi} = r^2\sin^2\theta\,\Phi_{,\langle\phi\phi\rangle}$ in terms of physical components. That is,

$$\begin{aligned}
\Phi_{,\langle rr\rangle} &= \frac{\partial^2 \Phi}{\partial r^2}, \\
\Phi_{,\langle\theta\theta\rangle} &= \frac{1}{r^2}\frac{\partial^2 \Phi}{\partial \theta^2} + \frac{1}{r}\frac{\partial \Phi}{\partial r}, \\
\Phi_{,\langle\phi\phi\rangle} &= \frac{1}{r^2\sin^2\theta}\frac{\partial^2 \Phi}{\partial \phi^2} + \frac{1}{r}\frac{\partial \Phi}{\partial r} + \frac{\cot\theta}{r^2}\frac{\partial \Phi}{\partial \theta}.
\end{aligned}$$

Therefore, the Laplacian $\nabla^2\Phi$, which is the sum $\Phi_{,\langle rr\rangle} + \Phi_{,\langle\theta\theta\rangle} + \Phi_{,\langle\phi\phi\rangle}$ in physical components, becomes

$$\nabla^2\Phi = \frac{\partial^2 \Phi}{\partial r^2} + \frac{2}{r}\frac{\partial \Phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \Phi}{\partial \theta^2} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 \Phi}{\partial \phi^2} + \frac{\cot\theta}{r^2}\frac{\partial \Phi}{\partial \theta}.$$

□

Example A.2.10 We give the physical components of the divergence of a symmetric tensor field T in the following coordinate systems:

a) Cartesian coordinate system (x, y, z):

$$\begin{aligned}
(\operatorname{div} T)_{\langle x\rangle} &= \frac{\partial T_{\langle xx\rangle}}{\partial x} + \frac{\partial T_{\langle xy\rangle}}{\partial y} + \frac{\partial T_{\langle xz\rangle}}{\partial z}, \\
(\operatorname{div} T)_{\langle y\rangle} &= \frac{\partial T_{\langle xy\rangle}}{\partial x} + \frac{\partial T_{\langle yy\rangle}}{\partial y} + \frac{\partial T_{\langle yz\rangle}}{\partial z}, \qquad (A.81)\\
(\operatorname{div} T)_{\langle z\rangle} &= \frac{\partial T_{\langle xz\rangle}}{\partial x} + \frac{\partial T_{\langle yz\rangle}}{\partial y} + \frac{\partial T_{\langle zz\rangle}}{\partial z}.
\end{aligned}$$

b) Cylindrical coordinate system (r, θ, z):

$$\begin{aligned}
(\operatorname{div} T)_{\langle r\rangle} &= \frac{\partial T_{\langle rr\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle r\theta\rangle}}{\partial \theta} + \frac{\partial T_{\langle rz\rangle}}{\partial z} + \frac{T_{\langle rr\rangle} - T_{\langle\theta\theta\rangle}}{r}, \\
(\operatorname{div} T)_{\langle\theta\rangle} &= \frac{\partial T_{\langle r\theta\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle\theta\theta\rangle}}{\partial \theta} + \frac{\partial T_{\langle\theta z\rangle}}{\partial z} + \frac{2}{r}T_{\langle r\theta\rangle}, \qquad (A.82)\\
(\operatorname{div} T)_{\langle z\rangle} &= \frac{\partial T_{\langle rz\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle\theta z\rangle}}{\partial \theta} + \frac{\partial T_{\langle zz\rangle}}{\partial z} + \frac{1}{r}T_{\langle rz\rangle}.
\end{aligned}$$

c) Spherical coordinate system (r, θ, ϕ):

$$
\begin{aligned}
(\operatorname{div} T)_{\langle r\rangle} &= \frac{\partial T_{\langle rr\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle r\theta\rangle}}{\partial \theta} + \frac{1}{r\,\sin\theta}\frac{\partial T_{\langle r\phi\rangle}}{\partial \phi} \\
&\quad + \frac{1}{r}\Big(2\,T_{\langle rr\rangle} - T_{\langle \theta\theta\rangle} - T_{\langle \phi\phi\rangle} + \cot\theta\, T_{\langle r\theta\rangle}\Big), \\
(\operatorname{div} T)_{\langle \theta\rangle} &= \frac{\partial T_{\langle r\theta\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle \theta\theta\rangle}}{\partial \theta} + \frac{1}{r\,\sin\theta}\frac{\partial T_{\langle \theta\phi\rangle}}{\partial \phi} \\
&\quad + \frac{1}{r}\Big(3\,T_{\langle r\theta\rangle} + \cot\theta\,(T_{\langle \theta\theta\rangle} - T_{\langle \phi\phi\rangle})\Big), \\
(\operatorname{div} T)_{\langle \phi\rangle} &= \frac{\partial T_{\langle r\phi\rangle}}{\partial r} + \frac{1}{r}\frac{\partial T_{\langle \theta\phi\rangle}}{\partial \theta} + \frac{1}{r\,\sin\theta}\frac{\partial T_{\langle \phi\phi\rangle}}{\partial \phi} \\
&\quad + \frac{1}{r}\Big(3\,T_{\langle r\phi\rangle} + 2\,\cot\theta\, T_{\langle \theta\phi\rangle}\Big).
\end{aligned}
\qquad (A.83)
$$

□

Exercise A.2.8 Let $\boldsymbol{u}$ be a vector field. Show that

1) in cylindrical coordinate system,

$$\operatorname{div}\boldsymbol{u} = \frac{\partial u_{\langle r\rangle}}{\partial r} + \frac{1}{r}\frac{\partial u_{\langle \theta\rangle}}{\partial \theta} + \frac{\partial u_{\langle z\rangle}}{\partial z} + \frac{1}{r}u_{\langle r\rangle};$$

2) in spherical coordinate system,

$$\operatorname{div}\boldsymbol{u} = \frac{\partial u_{\langle r\rangle}}{\partial r} + \frac{1}{r}\frac{\partial u_{\langle \theta\rangle}}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial u_{\langle \phi\rangle}}{\partial \phi} + \frac{2}{r}u_{\langle r\rangle} + \frac{\cot\theta}{r}u_{\langle \theta\rangle}.$$

Exercise A.2.9 Let $\boldsymbol{u}$ be a vector field and $E = \frac{1}{2}(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T)$. Express E in cylindrical and spherical coordinate systems,

1) relative to the natural basis,
2) in terms of physical components.

Exercise A.2.10 Let T be a symmetric tensor field. Compute $\operatorname{div} T$, in cylindrical and spherical coordinate systems,

1) relative to the natural basis,
2) in terms of physical components. (Verify (A.82) and (A.83)).

Exercise A.2.11 Let $\Phi : I\!R \to \mathcal{E}$ be a curve. Suppose that $\{\boldsymbol{e}_i(x)\}$ is the natural basis and $\phi^i(t)$ is the coordinate of $\Phi(t)$ in the coordinate system (x^i). Show that

1) $\dot{\Phi}(t) = \dot{\phi}^i(t)\,\boldsymbol{e}_i(\phi(t))$,
2) $\ddot{\Phi}(t) = \Big(\ddot{\phi}^i(t) + \dot{\phi}^j(t)\dot{\phi}^k(t)\Gamma_j{}^i{}_k(\phi(t)\Big)\,\boldsymbol{e}_i(\phi(t))$.