# 1.15 Tensor Calculus 2: Tensor Functions

## 1.15.1 Vector-valued functions of a vector

Consider a vector-valued function of a vector

$$\mathbf{a} = \mathbf{a}(\mathbf{b}), \qquad a_i = a_i(b_j)$$

This is a function of three independent variables $b_1, b_2, b_3$, and there are nine partial derivatives $\partial a_i / \partial b_j$. The partial derivative of the vector **a** with respect to **b** is defined to be a second-order tensor with these partial derivatives as its components:

$$\frac{\partial \mathbf{a}(\mathbf{b})}{\partial \mathbf{b}} \equiv \frac{\partial a_i}{\partial b_j} \mathbf{e}_i \otimes \mathbf{e}_j \tag{1.15.1}$$

It follows from this that

$$\frac{\partial \mathbf{a}}{\partial \mathbf{b}} = \left( \frac{\partial \mathbf{b}}{\partial \mathbf{a}} \right)^{-1} \quad \text{or} \quad \frac{\partial \mathbf{a}}{\partial \mathbf{b}} \frac{\partial \mathbf{b}}{\partial \mathbf{a}} = \mathbf{I}, \quad \frac{\partial a_i}{\partial b_m} \frac{\partial b_m}{\partial a_j} = \delta_{ij} \tag{1.15.2}$$

To show this, with $a_i = a_i(b_j), b_i = b_i(a_j)$, note that the differential can be written as

$$da_1 = \frac{\partial a_1}{\partial b_j} db_j = \frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_i} da_i = da_1 \left( \frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_1} \right) + da_2 \left( \frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_2} \right) + da_3 \left( \frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_3} \right)$$

Since $da_1, da_2, da_3$ are independent, one may set $da_2 = da_3 = 0$, so that

$$\frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_1} = 1$$

Similarly, the terms inside the other brackets are zero and, in this way, one finds Eqn. 1.15.2.

## 1.15.2 Scalar-valued functions of a tensor

Consider a scalar valued function of a (second-order) tensor

$$\phi = \phi(\mathbf{T}), \qquad \mathbf{T} = T_{ij} \mathbf{e}_i \otimes \mathbf{e}_j .$$

This is a function of nine independent variables, $\phi = \phi(T_{ij})$, so there are nine different partial derivatives:

$$\frac{\partial \phi}{\partial T_{11}}, \frac{\partial \phi}{\partial T_{12}}, \frac{\partial \phi}{\partial T_{13}}, \frac{\partial \phi}{\partial T_{21}}, \frac{\partial \phi}{\partial T_{22}}, \frac{\partial \phi}{\partial T_{23}}, \frac{\partial \phi}{\partial T_{31}}, \frac{\partial \phi}{\partial T_{32}}, \frac{\partial \phi}{\partial T_{33}}$$

The partial derivative of $\phi$ with respect to **T** is defined to be a second-order tensor with these partial derivatives as its components:

$$\boxed{\frac{\partial \phi}{\partial \mathbf{T}} \equiv \frac{\partial \phi}{\partial T_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j} \quad \textbf{Partial Derivative with respect to a Tensor} \qquad (1.15.3)$$

The quantity $\partial \phi(\mathbf{T}) / \partial \mathbf{T}$ is also called the gradient of $\phi$ with respect to **T**.

Thus differentiation with respect to a second-order tensor raises the order by 2. This agrees with the idea of the gradient of a scalar field where differentiation with respect to a vector raises the order by 1.

### Derivatives of the Trace and Invariants

Consider now the trace: the derivative of $\mathrm{tr}\mathbf{A}$, with respect to **A** can be evaluated as follows:

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{A}} \mathrm{tr}\mathbf{A} &= \frac{\partial A_{11}}{\partial \mathbf{A}} + \frac{\partial A_{22}}{\partial \mathbf{A}} + \frac{\partial A_{33}}{\partial \mathbf{A}} \\
&= \frac{\partial A_{11}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j + \frac{\partial A_{22}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j + \frac{\partial A_{33}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j \\
&= \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 \\
&= \mathbf{I}
\end{aligned} \qquad (1.15.4)$$

Similarly, one finds that {▲Problem 1}

$$\boxed{\begin{array}{ccc}
\dfrac{\partial(\mathrm{tr}\mathbf{A})}{\partial \mathbf{A}} = \mathbf{I} & \dfrac{\partial(\mathrm{tr}\mathbf{A}^2)}{\partial \mathbf{A}} = 2\mathbf{A}^{\mathrm{T}} & \dfrac{\partial(\mathrm{tr}\mathbf{A}^3)}{\partial \mathbf{A}} = 3(\mathbf{A}^2)^{\mathrm{T}} \\
\dfrac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial \mathbf{A}} = 2(\mathrm{tr}\mathbf{A})\mathbf{I} & \dfrac{\partial((\mathrm{tr}\mathbf{A})^3)}{\partial \mathbf{A}} = 3(\mathrm{tr}\mathbf{A})^2\mathbf{I} &
\end{array}} \qquad (1.15.5)$$

**Derivatives of Trace Functions**

From these and 1.11.17, one can evaluate the derivatives of the invariants {▲Problem 2}:

$$\boxed{\begin{aligned}
\frac{\partial \mathrm{I}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathbf{I} \\
\frac{\partial \mathrm{II}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathrm{I}_{\mathbf{A}}\mathbf{I} - \mathbf{A}^{\mathrm{T}} \\
\frac{\partial \mathrm{III}_{\mathbf{A}}}{\partial \mathbf{A}} &= (\mathbf{A}^{\mathrm{T}})^2 - \mathrm{I}_{\mathbf{A}}\mathbf{A}^{\mathrm{T}} + \mathrm{II}_{\mathbf{A}}\mathbf{I} = \mathrm{III}_{\mathbf{A}}\mathbf{A}^{-\mathrm{T}}
\end{aligned}} \quad \textbf{Derivatives of the Invariants} \qquad (1.15.6)$$

### Derivative of the Determinant

An important relation is

$$\frac{\partial}{\partial \mathbf{A}}(\det \mathbf{A}) = (\det \mathbf{A})\mathbf{A}^{-\mathrm{T}} \tag{1.15.7}$$

which follows directly from 1.15.6c.

### Other Relations

The total differential can be written as

$$\begin{aligned} d\phi &= \frac{\partial \phi}{\partial T_{11}} dT_{11} + \frac{\partial \phi}{\partial T_{12}} dT_{12} + \frac{\partial \phi}{\partial T_{13}} dT_{13} + \cdots \\ &\equiv \frac{\partial \phi}{\partial \mathbf{T}} : d\mathbf{T} \end{aligned} \tag{1.15.8}$$

This total differential gives an approximation to the total increment in $\phi$ when the increments of the independent variables $T_{11}, \cdots$ are small.

The second partial derivative is defined similarly:

$$\frac{\partial^2 \phi}{\partial \mathbf{T} \partial \mathbf{T}} \equiv \frac{\partial^2 \phi}{\partial T_{ij} \partial T_{pq}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_p \otimes \mathbf{e}_q, \tag{1.15.9}$$

the result being in this case a fourth-order tensor.

Consider a scalar-valued function of a tensor, $\phi(\mathbf{A})$, but now suppose that the components of **A** depend upon some scalar parameter *t*: $\phi = \phi(\mathbf{A}(t))$. By means of the chain rule of differentiation,

$$\dot{\phi} = \frac{\partial \phi}{\partial A_{ij}} \frac{dA_{ij}}{dt} \tag{1.15.10}$$

which in symbolic notation reads (see Eqn. 1.10.10e)

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial \mathbf{A}} : \frac{d\mathbf{A}}{dt} = \mathrm{tr}\left[\left(\frac{\partial \phi}{\partial \mathbf{A}}\right)^{\mathrm{T}} \frac{d\mathbf{A}}{dt}\right] \tag{1.15.11}$$

### Identities for Scalar-valued functions of Symmetric Tensor Functions

Let **C** be a symmetric tensor, $\mathbf{C} = \mathbf{C}^{\mathrm{T}}$. Then the partial derivative of $\phi = \phi(\mathbf{C}(\mathbf{T}))$ with respect to **T** can be written as {▲Problem 3}

(1) $\dfrac{\partial \phi}{\partial \mathbf{T}} = 2\mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}}$ for $\mathbf{C} = \mathbf{T}^{\mathrm{T}}\mathbf{T}$

(2) $\dfrac{\partial \phi}{\partial \mathbf{T}} = 2\dfrac{\partial \phi}{\partial \mathbf{T}}\mathbf{C}$ for $\mathbf{C} = \mathbf{T}\mathbf{T}^{\mathrm{T}}$ (1.15.12)

(3) $\dfrac{\partial \phi}{\partial \mathbf{T}} = 2\mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}} = 2\dfrac{\partial \phi}{\partial \mathbf{C}}\mathbf{T} = \mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}} + \dfrac{\partial \phi}{\partial \mathbf{C}}\mathbf{T}$ for $\mathbf{C} = \mathbf{T}\mathbf{T}$ and symmetric $\mathbf{T}$

### Scalar-valued functions of a Symmetric Tensor

Consider the expression

$$\mathbf{B} = \frac{\partial \phi(\mathbf{A})}{\partial \mathbf{A}} \qquad B_{ij} = \frac{\partial \phi(A_{ij})}{\partial A_{ij}} \tag{1.15.13}$$

If **A** is a symmetric tensor, there are a number of ways to consider this expression: two possibilities are that $\phi$ can be considered to be

(i) a symmetric function of the 9 variables $A_{ij}$

(ii) a function of 6 independent variables: $\phi = \phi(A_{11}, \overline{A}_{12}, \overline{A}_{13}, A_{22}, \overline{A}_{23}, A_{33})$ where

$$\begin{aligned}
\overline{A}_{12} &= \frac{1}{2}(A_{12} + A_{21}) = A_{12} = A_{21} \\
\overline{A}_{13} &= \frac{1}{2}(A_{13} + A_{31}) = A_{13} = A_{31} \\
\overline{A}_{23} &= \frac{1}{2}(A_{23} + A_{32}) = A_{23} = A_{32}
\end{aligned}$$

Looking at (i) and writing $\phi = \phi(A_{11}, A_{12}(\overline{A}_{12}), \cdots, A_{21}(\overline{A}_{12}), \cdots)$, one has, for example,

$$\frac{\partial \phi}{\partial \overline{A}_{12}} = \frac{\partial \phi}{\partial A_{12}}\frac{\partial A_{12}}{\partial \overline{A}_{12}} + \frac{\partial \phi}{\partial A_{21}}\frac{\partial A_{21}}{\partial \overline{A}_{12}} = \frac{\partial \phi}{\partial A_{12}} + \frac{\partial \phi}{\partial A_{21}} = 2\frac{\partial \phi}{\partial A_{12}},$$

the last equality following from the fact that $\phi$ is a symmetrical function of the $A_{ij}$.

Thus, depending on how the scalar function is presented, one could write

(i) $B_{11} = \dfrac{\partial \phi}{\partial A_{11}}, \quad B_{12} = \dfrac{\partial \phi}{\partial A_{12}}, \quad B_{13} = \dfrac{\partial \phi}{\partial A_{13}}, \quad$ etc.

(ii) $B_{11} = \dfrac{\partial \phi}{\partial A_{11}}, \quad B_{12} = \dfrac{1}{2}\dfrac{\partial \phi}{\partial \overline{A}_{12}}, \quad B_{13} = \dfrac{1}{2}\dfrac{\partial \phi}{\partial \overline{A}_{13}}, \quad$ etc.

### 1.15.3 Tensor-valued functions of a tensor

The derivative of a (second-order) tensor **A** with respect to another tensor **B** is defined as

$$\frac{\partial \mathbf{A}}{\partial \mathbf{B}} \equiv \frac{\partial A_{ij}}{\partial B_{pq}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_p \otimes \mathbf{e}_q \tag{1.15.14}$$

and forms therefore a fourth-order tensor. The total differential $d\mathbf{A}$ can in this case be written as

$$d\mathbf{A} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}} : d\mathbf{B} \tag{1.15.15}$$

Consider now

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \frac{\partial A_{ij}}{\partial A_{kl}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l$$

The components of the tensor are independent, so

$$\frac{\partial A_{11}}{\partial A_{11}} = 1, \quad \frac{\partial A_{11}}{\partial A_{12}} = 0, \quad \cdots \quad \text{etc.} \qquad \boxed{\frac{\partial A_{mn}}{\partial A_{pq}} = \delta_{mp}\delta_{nq}} \tag{1.15.16}$$

and so

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{I}, \tag{1.15.17}$$

the fourth-order identity tensor of Eqn. 1.12.4.

#### Example

Consider the scalar-valued function $\phi$ of the tensor **A** and vector **v** (the "dot" can be omitted from the following and similar expressions),

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

The gradient of $\phi$ with respect to **v** is

$$\frac{\partial \phi}{\partial \mathbf{v}} = \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \cdot \mathbf{A}\mathbf{v} + \mathbf{v} \cdot \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{v}} = \mathbf{A}\mathbf{v} + \mathbf{v}\mathbf{A} = \left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\mathbf{v}$$

On the other hand, the gradient of $\phi$ with respect to **A** is

$$\frac{\partial \phi}{\partial \mathbf{A}} = \mathbf{v} \cdot \frac{\partial \mathbf{A}}{\partial \mathbf{A}} \mathbf{v} = \mathbf{v} \cdot \mathbf{I}\mathbf{v} = \mathbf{v} \otimes \mathbf{v}$$

■

Consider now the derivative of the inverse, $\partial \mathbf{A}^{-1} / \partial \mathbf{A}$. One can differentiate $\mathbf{A}^{-1}\mathbf{A} = \mathbf{0}$ using the product rule to arrive at

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} \mathbf{A} = -\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \mathbf{A}}$$

One needs to be careful with derivatives because of the position of the indices in 1.15.14); it looks like a post-operation of both sides with the inverse leads to $\partial \mathbf{A}^{-1} / \partial \mathbf{A} = -\mathbf{A}^{-1}(\partial \mathbf{A} / \partial \mathbf{A})\mathbf{A}^{-1} = -A_{ik}^{-1} A_{jl}^{-1} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l$. However, this is not correct (unless **A** is symmetric). Using the index notation (there is no clear symbolic notation), one has

$$\begin{aligned}
&\frac{\partial A_{im}^{-1}}{\partial A_{kl}} A_{mj} = -A_{im}^{-1} \frac{\partial A_{mj}}{\partial A_{kl}} \qquad (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l) \\
\to\ &\frac{\partial A_{im}^{-1}}{\partial A_{kl}} A_{mj} A_{jn}^{-1} = -A_{im}^{-1} \frac{\partial A_{mj}}{\partial A_{kl}} A_{jn}^{-1} \\
\to\ &\frac{\partial A_{im}^{-1}}{\partial A_{kl}} \delta_{mn} = -A_{im}^{-1} \delta_{mk} \delta_{jl} A_{jn}^{-1} \\
\to\ &\frac{\partial A_{ij}^{-1}}{\partial A_{kl}} = -A_{ik}^{-1} A_{lj}^{-1} \qquad (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l)
\end{aligned} \tag{1.15.18}$$

■

## 1.15.4 The Directional Derivative

The directional derivative was introduced in §1.6.11. The ideas introduced there can be extended to tensors. For example, the directional derivative of the trace of a tensor **A**, in the direction of a tensor **T**, is

$$\partial_{\mathbf{A}}(\mathrm{tr}\mathbf{A})[\mathbf{T}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathrm{tr}(\mathbf{A} + \varepsilon\mathbf{T}) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (\mathrm{tr}\mathbf{A} + \varepsilon\mathrm{tr}\mathbf{T}) = \mathrm{tr}\mathbf{T} \tag{1.15.19}$$

As a further example, consider the scalar function $\phi(\mathbf{A}) = \mathbf{u} \cdot \mathbf{A}\mathbf{v}$, where **u** and **v** are constant vectors. Then

$$\partial_{\mathbf{A}}\phi(\mathbf{A}, \mathbf{u}, \mathbf{v})[\mathbf{T}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} [\mathbf{u} \cdot (\mathbf{A} + \varepsilon\mathbf{T})\mathbf{v}] = \mathbf{u} \cdot \mathbf{T}\mathbf{v} \tag{1.15.20}$$

Also, the gradient of $\phi$ with respect to **A** is

$$\frac{\partial \phi}{\partial \mathbf{A}} = \frac{\partial}{\partial \mathbf{A}}(\mathbf{u} \cdot \mathbf{A}\mathbf{v}) = \mathbf{u} \otimes \mathbf{v} \tag{1.15.21}$$

and it can be seen that this is an example of the more general relation

$$\partial_{\mathbf{A}} \phi[\mathbf{T}] = \frac{\partial \phi}{\partial \mathbf{A}} : \mathbf{T} \tag{1.15.22}$$

which is analogous to 1.6.41. Indeed,

$$\begin{aligned} \partial_{\mathbf{x}} \phi[\mathbf{w}] &= \frac{\partial \phi}{\partial \mathbf{x}} \cdot \mathbf{w} \\ \partial_{\mathbf{A}} \phi[\mathbf{T}] &= \frac{\partial \phi}{\partial \mathbf{A}} : \mathbf{T} \\ \partial_{\mathbf{u}} \mathbf{v}[\mathbf{w}] &= \frac{\partial \mathbf{v}}{\partial \mathbf{u}} \mathbf{w} \end{aligned} \tag{1.15.23}$$

### Example (the Directional Derivative of the Determinant)

It was shown in §1.6.11 that the directional derivative of the determinant of the $2 \times 2$ matrix **A**, in the direction of a second matrix **T**, is

$$\partial_{\mathbf{A}} (\det \mathbf{A})[\mathbf{T}] = A_{11}T_{22} + A_{22}T_{11} - A_{12}T_{21} - A_{21}T_{12}$$

This can be seen to be equal to $\det \mathbf{A} \left( \mathbf{A}^{-\mathrm{T}} : \mathbf{T} \right)$, which will now be proved more generally for tensors **A** and **T**:

$$\begin{aligned} \partial_{\mathbf{A}} (\det \mathbf{A})[\mathbf{T}] &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det(\mathbf{A} + \varepsilon \mathbf{T}) \\ &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det\left[\mathbf{A}\left(\mathbf{I} + \varepsilon \mathbf{A}^{-1}\mathbf{T}\right)\right] \\ &= \det \mathbf{A} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det\left(\mathbf{I} + \varepsilon \mathbf{A}^{-1}\mathbf{T}\right) \end{aligned}$$

The last line here follows from (1.10.16a). Now the characteristic equation for a tensor **B** is given by (1.11.4, 1.11.5),

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = 0 = \det(\mathbf{B} - \lambda \mathbf{I})$$

where $\lambda_i$ are the three eigenvalues of **B**. Thus, setting $\lambda = -1$ and $\mathbf{B} = \varepsilon \mathbf{A}^{-1}\mathbf{T}$,

$$
\begin{aligned}
\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] &= \det \mathbf{A} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(1+\lambda_1|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_2|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_3|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(1+\varepsilon\lambda_1|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_2|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_3|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A} \left(\lambda_1|_{\mathbf{A}^{-1}\mathbf{T}}+\lambda_2|_{\mathbf{A}^{-1}\mathbf{T}}+\lambda_3|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A}\ \mathrm{tr}\left(\mathbf{A}^{-1}\mathbf{T}\right)
\end{aligned}
$$

and, from (1.10.10e),

$$
\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] = \det \mathbf{A}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right) \tag{1.15.24}
$$

■

### Example (the Directional Derivative of a vector function)

Consider the $n$ homogeneous algebraic equations $\mathbf{f}(\mathbf{x}) = \mathbf{o}$:

$$
\begin{gathered}
f_1(x_1, x_2, \cdots, x_n) = 0 \\
f_2(x_1, x_2, \cdots, x_n) = 0 \\
\cdots \\
f_n(x_1, x_2, \cdots, x_n) = 0
\end{gathered}
$$

The directional derivative of $\mathbf{f}$ in the direction of some vector $\mathbf{u}$ is

$$
\begin{aligned}
\partial_{\mathbf{x}} \mathbf{f}(\mathbf{x})[\mathbf{u}] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{f}(\mathbf{z}(\varepsilon)) \qquad (\mathbf{z} = \mathbf{x} + \varepsilon\mathbf{u}) \\
&= \left(\frac{\partial \mathbf{f}(\mathbf{z})}{\partial \mathbf{z}} \frac{d\mathbf{z}}{d\varepsilon}\right)_{\varepsilon=0} \\
&= \mathbf{K}\mathbf{u}
\end{aligned} \tag{1.15.25}
$$

where $\mathbf{K}$, called the **tangent matrix** of the system, is

$$
\mathbf{K} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & & \partial f_2/\partial x_n \\ \vdots & & & \vdots \\ \partial f_n/\partial x_1 & & \cdots & \partial f_n/\partial x_n \end{bmatrix}, \quad \partial_{\mathbf{x}} \mathbf{f}[\mathbf{u}] = (\mathrm{grad}\mathbf{f})\mathbf{u}
$$

which can be compared to (1.15.23c).

■

### Properties of the Directional Derivative

The directional derivative is a linear operator and so one can apply the usual product rule. For example, consider the directional derivative of $\mathbf{A}^{-1}$ in the direction of $\mathbf{T}$:

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (\mathbf{A} + \varepsilon\mathbf{T})^{-1}$$

To evaluate this, note that $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\mathbf{A}\right)[\mathbf{T}] = \partial_{\mathbf{A}}(\mathbf{I})[\mathbf{T}] = \mathbf{0}$, since **I** is independent of **A**. The product rule then gives $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}]\mathbf{A} = -\mathbf{A}^{-1}\partial_{\mathbf{A}}(\mathbf{A})[\mathbf{T}]$, so that

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = -\mathbf{A}^{-1}\partial_{\mathbf{A}}\mathbf{A}[\mathbf{T}]\mathbf{A}^{-1} = -\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1} \tag{1.15.26}$$

Another important property of the directional derivative is the **chain rule**, which can be applied when the function is of the form $\mathbf{f}(\mathbf{x}) = \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}))$. To derive this rule, consider (see §1.6.11)

$$\mathbf{f}(\mathbf{x}+\mathbf{u}) \approx \mathbf{f}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}], \tag{1.15.27}$$

where terms of order $o(\mathbf{u})$ have been neglected, i.e.

$$\lim_{|\mathbf{u}|\to 0} \frac{o(\mathbf{u})}{|\mathbf{u}|} = 0.$$

The left-hand side of the previous expression can also be written as

$$\begin{aligned}\hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}+\mathbf{u})) &\approx \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]) \\ &\approx \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x})) + \partial_{\mathbf{B}}\hat{\mathbf{f}}(\mathbf{B})[\partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]]\end{aligned}$$

Comparing these expressions, one arrives at the chain rule,

$$\boxed{\partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}] = \partial_{\mathbf{B}}\hat{\mathbf{f}}(\mathbf{B})[\partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]]} \qquad \textbf{Chain Rule} \tag{1.15.28}$$

As an application of this rule, consider the directional derivative of $\det\mathbf{A}^{-1}$ in the direction **T**; here, **f** is $\det\mathbf{A}^{-1}$ and $\hat{\mathbf{f}} = \hat{\mathbf{f}}(\mathbf{B}(\mathbf{A}))$. Let $\mathbf{B} = \mathbf{A}^{-1}$ and $\hat{\mathbf{f}} = \det\mathbf{B}$. Then, from Eqns. 1.15.24, 1.15.25, 1.10.3h, f,

$$\begin{aligned}\partial_{\mathbf{A}}\left(\det\mathbf{A}^{-1}\right)[\mathbf{T}] &= \partial_{\mathbf{B}}(\det\mathbf{B})[\partial_{\mathbf{A}}\mathbf{A}^{-1}[\mathbf{T}]] \\ &= (\det\mathbf{B})\left(\mathbf{B}^{-\mathrm{T}} : \left(-\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{\mathrm{T}} : \left(\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right)\end{aligned} \tag{1.15.29}$$

## 1.15.5 Formal Treatment of Tensor Calculus

Following on from §1.6.12 and §1.14.6, a scalar function $f : V^2 \to R$ is **differentiable** at $\mathbf{A} \in V^2$ if there exists a second order tensor $Df(\mathbf{A}) \in V^2$ such that

$$f(\mathbf{A}+\mathbf{H}) = f(\mathbf{A}) + Df(\mathbf{A}):\mathbf{H} + o(\|\mathbf{H}\|) \quad \text{for all} \quad \mathbf{H} \in V^2 \tag{1.15.30}$$

In that case, the tensor $Df(\mathbf{A})$ is called the **derivative** of *f* at **A**. It follows from this that $Df(\mathbf{A})$ is that tensor for which

$$\partial_{\mathbf{A}} f[\mathbf{B}] = Df(\mathbf{A}):\mathbf{B} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} f(\mathbf{A}+\varepsilon\mathbf{B}) \quad \text{for all} \quad \mathbf{B} \in V^2 \tag{1.15.31}$$

For example, from 1.15.24,

$$\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] = \det \mathbf{A}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right) = \left(\det \mathbf{A}\ \mathbf{A}^{-\mathrm{T}}\right):\mathbf{T} \tag{1.15.32}$$

from which it follows, from 1.15.31, that

$$\frac{\partial}{\partial \mathbf{A}} \det \mathbf{A} = \det \mathbf{A}\ \mathbf{A}^{-\mathrm{T}} \tag{1.15.33}$$

which is 1.15.7.

Similarly, a tensor-valued function $\mathbf{T}: V^2 \to V^2$ is **differentiable** at $\mathbf{A} \in V^2$ if there exists a fourth order tensor $D\mathbf{T}(\mathbf{A}) \in V^4$ such that

$$\mathbf{T}(\mathbf{A}+\mathbf{H}) = \mathbf{T}(\mathbf{A}) + D\mathbf{T}(\mathbf{A})\mathbf{H} + o(\|\mathbf{H}\|) \quad \text{for all} \quad \mathbf{H} \in V^2 \tag{1.15.34}$$

In that case, the tensor $D\mathbf{T}(\mathbf{A})$ is called the **derivative** of **T** at **A**. It follows from this that $D\mathbf{T}(\mathbf{A})$ is that tensor for which

$$\partial_{\mathbf{A}} \mathbf{T}[\mathbf{B}] = D\mathbf{T}(\mathbf{A}):\mathbf{B} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{T}(\mathbf{A}+\varepsilon\mathbf{B}) \quad \text{for all} \quad \mathbf{B} \in V^2 \tag{1.15.35}$$

## 1.15.6 Problems

1. Evaluate the derivatives (use the chain rule for the last two of these)
$$\frac{\partial(\mathrm{tr}\mathbf{A}^2)}{\partial \mathbf{A}}, \quad \frac{\partial(\mathrm{tr}\mathbf{A}^3)}{\partial \mathbf{A}}, \quad \frac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial \mathbf{A}}, \quad \frac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial \mathbf{A}}$$
2. Derive the derivatives of the invariants, Eqn. 1.15.5. [Hint: use the Cayley-Hamilton theorem, Eqn. 1.11.15, to express the derivative of the third invariant in terms of the third invariant.]
3. (a) Consider the scalar valued function $\phi = \phi(\mathbf{C}(\mathbf{F}))$, where $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. Use the chain rule
$$\frac{\partial \phi}{\partial \mathbf{F}} = \frac{\partial \phi}{\partial C_{mn}} \frac{\partial C_{mn}}{\partial F_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j$$
to show that

$$\frac{\partial \phi}{\partial \mathbf{F}} = 2\mathbf{F}\frac{\partial \phi}{\partial \mathbf{C}}, \qquad \frac{\partial \phi}{\partial F_{ij}} = 2F_{ik}\frac{\partial \phi}{\partial C_{kj}}$$

(b) Show also that

$$\frac{\partial \phi}{\partial \mathbf{U}} = 2\mathbf{U}\frac{\partial \phi}{\partial \mathbf{C}} = 2\frac{\partial \phi}{\partial \mathbf{C}}\mathbf{U}$$

for $\mathbf{C} = \mathbf{U}\mathbf{U}$ with $\mathbf{U}$ symmetric.

[Hint: for (a), use the index notation: first evaluate $\partial C_{mn} / \partial F_{ij}$ using the product rule, then evaluate $\partial \phi / \partial F_{ij}$ using the fact that **C** is symmetric.]

4. Show that

   (a) $\dfrac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{B} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$, (b) $\dfrac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{A} \otimes \mathbf{A}^{-1} = -\mathbf{A}^{-1} \otimes \mathbf{A}^{-1}$

5. Show that

$$\frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial \mathbf{A}} : \mathbf{B} = \mathbf{B}^{\mathrm{T}}$$

6. By writing the norm of a tensor $|\mathbf{A}|$, 1.10.14, where **A** is symmetric, in terms of the trace (see 1.10.10), show that

$$\frac{\partial |\mathbf{A}|}{\partial \mathbf{A}} = \frac{\mathbf{A}}{|\mathbf{A}|}$$

7. Evaluate

   (i) $\partial_{\mathbf{A}}\left(\mathbf{A}^2\right)[\mathbf{T}]$

   (ii) $\partial_{\mathbf{A}}\left(\mathrm{tr}\mathbf{A}^2\right)[\mathbf{T}]$ (see 1.10.10e)

8. Derive 1.15.29 by using the definition of the directional derivative and the relation 1.15.7, $\partial(\det \mathbf{A}) / \partial \mathbf{A} = (\det \mathbf{A})\mathbf{A}^{-\mathrm{T}}$.