

THE DERIVATIVE OF A TENSOR-VALUED FUNCTION OF A TENSOR*

BY

DONALD E. CARLSON (*Institute for Mathematics and its Applications, University of Minnesota and
University of Illinois at Urbana-Champaign*)

AND

ANNE HOGER (*Carnegie-Mellon University*)

1. Introduction.¹ Tensor-valued functions of tensors appear frequently in continuum physics, and it would often be useful to have expressions for their derivatives. For example, one may have a tensor-valued function of a tensor that in turn depends on a real parameter, such as time, and want the total time rate of change; i.e., suppose $\mathbf{X}: \mathbb{R} \rightarrow \text{Sym}(V_n)$, $\mathbf{F}: \text{Sym}(V_n) \rightarrow \text{Sym}(V_n)$ and consider the derivative of the composite function $\mathbf{G} = \mathbf{F} \circ \mathbf{X}$. If \mathbf{X} and \mathbf{F} are differentiable, then $\dot{\mathbf{G}}(t) = D\mathbf{F}(\mathbf{X}(t))[\dot{\mathbf{X}}(t)]$. However, in the past, the derivative $D\mathbf{F}$ has generally been unknown, and researchers have been forced to approach $\dot{\mathbf{G}}$ without recourse to the chain rule (e.g., see Guo [2] for the case $\mathbf{F}(\mathbf{X}) = \sqrt{\mathbf{X}}$ or Gurtin and Spear [3] for $\mathbf{F}(\mathbf{X}) = \ln \mathbf{X}$). Of course, the derivative of a tensor-valued function of a tensor is a topic of interest in its own right.

When \mathbf{F} is sufficiently simple, such as a polynomial, $D\mathbf{F}$ can be calculated directly from the definition of the derivative (see Gurtin [1, p. 22]). Sometimes $D\mathbf{F}$ can be obtained indirectly from the defining properties of \mathbf{F} (e.g., Hoger and Carlson [4] found $D\sqrt{\mathbf{X}}$ by differentiating $\sqrt{\mathbf{X}} \sqrt{\mathbf{X}} = \mathbf{X}$). The purpose of the present study is to provide $D\mathbf{F}$ in the general case.

We define the function \mathbf{F} through the spectral decomposition of its argument in the following manner. Given $f: \mathbb{R} \rightarrow \mathbb{R}$, then $\mathbf{F}: \text{Sym}(V_n) \rightarrow \text{Sym}(V_n)$ is defined by

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^n f(x_i) \mathbf{e}_i(\mathbf{X}) \otimes \mathbf{e}_i(\mathbf{X}), \quad (1.1)^2$$

* Received November 19, 1984.

¹ Throughout, we use the notation of Gurtin [1] or obvious variants of it. One exception is that our \mathbf{F} is not the deformation gradient of continuum mechanics, but rather is a function on $\text{Sym}(V_n)$ into $\text{Sym}(V_n)$. V_n is a normed n -dimensional vector space, usually $n = 2$ or 3 .

² This is the definition most commonly employed in continuum mechanics; see Rinehart [5] for comparison of various approaches. Often the domain of f is an open interval in \mathbb{R} , and the domain of \mathbf{F} is an open subset of $\text{Sym}(V_n)$. All of our results apply to this situation as well.

where $\{x_1, x_2, \dots, x_n\}$ is the spectrum of \mathbf{X} and $\{\mathbf{e}_1(\mathbf{X}), \mathbf{e}_2(\mathbf{X}), \dots, \mathbf{e}_n(\mathbf{X})\}$ is an orthonormal basis of eigenvectors of \mathbf{X} . The nonuniqueness of $\{\mathbf{e}_i(\mathbf{X})\}$ when \mathbf{X} has repeated eigenvalues has been a major source of difficulty in obtaining the derivative of \mathbf{F} . We circumvent this by replacing (1.1) with the equivalent representation

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^p f(x_i) \mathbf{E}_i(\mathbf{X}), \quad (1.2)$$

where $\{x_1, x_2, \dots, x_p\}$ is the set of *distinct* eigenvalues of \mathbf{X} and $\mathbf{E}_i(\mathbf{X})$ is the eigenprojection corresponding to x_i (i.e., the orthogonal projection operator on the null space of $\mathbf{X} - x_i \mathbf{I}$).

The advantages of (1.2) over (1.1) are that the eigenprojections are unique and there are formulas for them in terms of \mathbf{X} and its eigenvalues. Indeed,

$$\mathbf{E}_i(\mathbf{X}) = \begin{cases} \prod_{\substack{j=1 \\ j \neq i}}^p \frac{\mathbf{X} - x_j \mathbf{I}}{x_i - x_j}, & p > 1, \\ \mathbf{I}, & p = 1 \end{cases} \quad (1.3)$$

(see, e.g., Hoffman and Kunze [6] or Halmos [7]). Of course, the integer p depends on \mathbf{X} , which means that the representation (1.2) for $\mathbf{F}(\mathbf{X})$ (as well as the spectral decomposition of \mathbf{X} itself) changes with the number of distinct eigenvalues of \mathbf{X} . Nevertheless, we shall see that (1.2) defines a differentiable function if f has enough derivatives.

Important properties of the eigenprojections that will be used later are

$$\mathbf{E}_i(\mathbf{X}) \mathbf{E}_j(\mathbf{X}) = \begin{cases} \mathbf{E}_i(\mathbf{X}), & i = j, \\ \mathbf{0}, & i \neq j, \end{cases} \quad (1.4)$$

$$\sum_{i=1}^p \mathbf{E}_i(\mathbf{X}) = \mathbf{I}. \quad (1.5)$$

Our approach is to apply to (1.2) the well-known result that *if for each \mathbf{T} in a basis for $\text{Sym}(V_n)$ the directional derivative $d\mathbf{F}(\mathbf{X} + \alpha\mathbf{T})/d\alpha|_{\alpha=0}$ exists in a neighborhood of \mathbf{X} and is continuous at \mathbf{X} , then \mathbf{F} is differentiable at \mathbf{X} and its derivative is given by*

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \left. \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \right|_{\alpha=0}.$$

It is interesting to note that even though the eigenvalues and eigenprojections are not differentiable at arguments with repeated eigenvalues, the function defined by (1.2) is differentiable for a sufficiently smooth f .

This program is carried out directly in a self-contained, elementary manner for $n = 2$ in Sec. 2. The three-dimensional case is worked out in Sec. 3 first for an \mathbf{X} with three distinct eigenvalues, and then the results are extended to the case of repeated eigenvalues with the help of a new lemma due to Ball [8]. The two-dimensional case also could be handled with Ball's lemma; however, we feel that the added insight afforded by the completely elementary treatment in Sec. 2 warrants the extra effort. The n -dimensional case is considered briefly in Sec. 4, and a formula for the derivative at an \mathbf{X} with n distinct eigenvalues is given.

Of course, the expressions for derivatives found here agree with those few already known for polynomials and other simple functions, but now more complicated functions

such as the tensor logarithm can be differentiated. This is central to the determination in [9] of the relationship between the time rate of change of the logarithmic strain tensor and the stretching tensor.

2. The two-dimensional case. In this section, we start with the formulas for the eigenvalues obtained by solving the characteristic equation and then consider the directional derivatives of the eigenvalues, the eigenprojections, and the function F defined by (1.2). The analysis falls naturally into cases determined by the multiplicities of the eigenvalues.

2.1 *The tensor* $Y(\alpha) = X + \alpha T$. In preparation for the calculation of the directional derivatives of functions of X , we consider the linear perturbation

$$Y(\alpha) = X + \alpha T, \quad \alpha \in \mathbb{R},$$

where $X, T \in \text{Sym}(V_2)$ are arbitrarily fixed.

The principal invariants of $Y(\alpha)$ are

$$\left. \begin{aligned} I_1(\alpha) &= \text{tr } X + \alpha \text{tr } T, \\ I_2(\alpha) &= \det X + \alpha[(\text{tr } X)(\text{tr } T) - \text{tr}(XT)] + \alpha^2 \det T. \end{aligned} \right\} \quad (2.1.2)$$

The ordered eigenvalues of $Y(\alpha)$, $y_1(\alpha) \leq y_2(\alpha)$, are the roots of the characteristic equation

$$y_i(\alpha)^2 - I_1(\alpha)y_i(\alpha) + I_2(\alpha) = 0.$$

Thus,

$$y_1(\alpha) = \frac{1}{2}I_1(\alpha) - \frac{1}{2}\sqrt{d(\alpha)}, \quad y_2(\alpha) = \frac{1}{2}I_1(\alpha) + \frac{1}{2}\sqrt{d(\alpha)}, \quad (2.1.3)$$

where the discriminant

$$\begin{aligned} d(\alpha) &= I_1(\alpha)^2 - 4I_2(\alpha) \\ &= (\text{tr } X)^2 - 4 \det X + \alpha 2[2 \text{tr}(XT) - (\text{tr } X)(\text{tr } T)] + \alpha^2[(\text{tr } T)^2 - 4 \det T] \\ &= (x_2 - x_1)^2 + \alpha 2[2 \text{tr}(XT) - (\text{tr } X)(\text{tr } T)] + \alpha^2(t_2 - t_1)^2 \end{aligned} \quad (2.1.4)$$

is necessarily nonnegative since $Y(\alpha) \in \text{Sym}(V_2)$. Here, $x_1 \leq x_2$ and $t_1 \leq t_2$ are the eigenvalues of X and T , respectively. We note that

$$y_i(0) = x_i, \quad (2.1.5)$$

which reflects the fact that the eigenvalues of X depend continuously on X .

Since $y_2(\alpha) - y_1(\alpha) = \sqrt{d(\alpha)}$, it follows from (2.1.4) that the eigenprojections of $Y(\alpha)$, $E_i(Y(\alpha)) \equiv E_i(\alpha)$, fall into one of three categories.

Case 1 ($x_1 < x_2$). Here $d(\alpha) > 0$ for sufficiently small α , and by (1.3)

$$E_1(\alpha) = \frac{Y(\alpha) - y_2(\alpha)I}{y_1(\alpha) - y_2(\alpha)}, \quad E_2(\alpha) = \frac{Y(\alpha) - y_1(\alpha)I}{y_2(\alpha) - y_1(\alpha)}. \quad (2.1.6)$$

At $\alpha = 0$, these become

$$E_1(0) = \frac{X - x_2I}{x_1 - x_2} = E_1(X), \quad E_2(0) = \frac{X - x_1I}{x_2 - x_1} = E_2(X). \quad (2.1.7)$$

Case 2 ($x_1 = x_2 = x$, $t_1 < t_2$). Here $d(\alpha) = \alpha^2(t_2 - t_1)^2$, and

$$E_1(\alpha) = \frac{Y(\alpha) - y_2(\alpha)I}{|\alpha|(t_2 - t_1)}, \quad E_2(\alpha) = \frac{Y(\alpha) - y_1(\alpha)I}{|\alpha|(t_1 - t_2)}. \quad (2.1.8)$$

In this case, the eigenprojections are indeterminate at $\alpha = 0$; and since $|\alpha|$ is not differentiable at $\alpha = 0$, its presence in the denominator prohibits the use of l'Hospital's rule. However, we can calculate the following one-sided limits.

$$\left. \begin{aligned} \lim_{\alpha \rightarrow 0^+} \mathbf{E}_1(\alpha) &= \lim_{\alpha \rightarrow 0^-} \mathbf{E}_2(\alpha) = \frac{\mathbf{T} - t_2 \mathbf{I}}{t_1 - t_2} = \mathbf{E}_1(\mathbf{T}), \\ \lim_{\alpha \rightarrow 0^-} \mathbf{E}_1(\alpha) &= \lim_{\alpha \rightarrow 0^+} \mathbf{E}_2(\alpha) = \frac{\mathbf{T} - t_1 \mathbf{I}}{t_2 - t_1} = \mathbf{E}_2(\mathbf{T}). \end{aligned} \right\} \tag{2.1.9}$$

Case 3 ($x_1 = x_2 = x, t_1 = t_2 = t$). Here $d(\alpha) = 0$, and

$$\mathbf{E}_1(\alpha) = \mathbf{I}. \tag{2.1.10}$$

Remark. Since the eigenprojections are discontinuous at an \mathbf{X} with equal eigenvalues, they clearly are not differentiable there.

2.2 *The directional derivatives of the eigenvalues.* The derivatives of the eigenvalues of $\mathbf{Y}(\alpha)$ with respect to α are readily calculated from (2.1.3) together with (2.1.2) and (2.1.4). Due to the presence of $\sqrt{d(\alpha)}$ in (2.1.3), the analysis is partitioned into the three cases of the previous subsection.

Case 1 ($x_1 < x_2$). Equations (2.1.2)–(2.1.4) and (2.1.7) imply

$$\left. \frac{d}{d\alpha} y_i(\alpha) \right|_{\alpha=0} = \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]. \tag{2.2.1}$$

Since the directional derivative is continuous in \mathbf{X} for each \mathbf{T} , we have differentiability in this case; and

$$Dx_i(\mathbf{X})[\mathbf{T}] = \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]. \tag{2.2.2}$$

Remark. (2.2.2) is a well-known result (see, e.g., Kato [10, p. 79]). We have included the above derivation to keep the exposition self-contained. See Ball [8, Lemma 5.1] for an alternative elementary treatment.

Case 2 ($x_1 = x_2 = x, t_1 < t_2$). Here $\sqrt{d(\alpha)} = |\alpha|(t_2 - t_1)$, and only one-sided derivatives exist:

$$\left. \begin{aligned} \frac{d^+}{d\alpha} y_1(\alpha) &= t_1, & \frac{d^-}{d\alpha} y_1(\alpha) &= t_2, \\ \frac{d^+}{d\alpha} y_2(\alpha) &= t_2, & \frac{d^-}{d\alpha} y_2(\alpha) &= t_1. \end{aligned} \right\} \tag{2.2.3}$$

Thus, the eigenvalues do not even have directional derivatives in Case 2. The nonexistence of $dy_i/d\alpha$ reflects the fact that we have insisted on the ordering $y_1(\alpha) \leq y_2(\alpha)$ (cf. the remark below).

Case 3 ($x_1 = x_2 = x, t_1 = t_2 = t$). Here $d(\alpha) = 0$, and

$$\frac{d}{d\alpha} y_i(\alpha) = t. \tag{2.2.4}$$

Remark. Of course, we have capitalized on the fact that for $n = 2$ there are simple formulas for the eigenvalues. For $n = 3, 4$ the formulas are not simple, and for $n > 4$ no such formulas exist. For any n if $dy_i(\alpha)/d\alpha$ is known to exist, one can proceed (as we do in Subsec. 3.1 and Sec. 4) by differentiating the characteristic equation. There is no difficulty as long as the eigenvalues are distinct, but when they coalesce the coefficient of $dy_i(\alpha)/d\alpha$ vanishes. In the latter situation, l'Hospital's rule can be used or the characteristic equation can be differentiated again. In any case, the differentiability of the $y_i(\alpha)$ with

respect to α is a crucial matter. Rellich [11, pp. 44–45] (see also Kato [10, p. 122]) has shown that the *unordered* $y_i(\alpha)$ may be chosen so as to be smooth in α .

2.3 *The directional derivatives of the eigenprojections.* Since

$$\mathbf{E}_1(\mathbf{X}) + \mathbf{E}_2(\mathbf{X}) = \mathbf{I}, \quad (2.3.1)$$

it suffices to consider only $\mathbf{E}_1(\mathbf{X})$.

Case 1 ($x_1 < x_2$). Equation (2.1.6) together with (2.1.1), (2.1.5), (2.1.7), (2.2.1), and (2.3.1) yields

$$\begin{aligned} \left. \frac{d}{d\alpha} \mathbf{E}_1(\alpha) \right|_{\alpha=0} &= - \left. \frac{d}{d\alpha} \mathbf{E}_2(\alpha) \right|_{\alpha=0} \\ &= \frac{1}{(x_1 - x_2)} \{ \mathbf{T} - \text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}]\mathbf{E}_1(\mathbf{X}) - \text{tr}[\mathbf{E}_2(\mathbf{X})\mathbf{T}]\mathbf{E}_2(\mathbf{X}) \}. \end{aligned} \quad (2.3.2)$$

By continuity, the eigenprojections are differentiable in this case, and

$$D\mathbf{E}_i(\mathbf{X})[\mathbf{T}] = \left. \frac{d}{d\alpha} \mathbf{E}_i(\alpha) \right|_{\alpha=0}. \quad (2.3.3)$$

Equation (2.3.2) can be cast into a more useful form with the aid of Rivlin's [12] identity

$$\text{tr}(\mathbf{A}\mathbf{B})\mathbf{I} = \mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A} - (\text{tr } \mathbf{A})\mathbf{B} - (\text{tr } \mathbf{B})\mathbf{A} + (\text{tr } \mathbf{A})(\text{tr } \mathbf{B})\mathbf{I}.$$

We take $\mathbf{A} = \mathbf{E}_1(\mathbf{X})$, $\mathbf{B} = \mathbf{T}$, note from (2.1.7) that $\text{tr } \mathbf{E}_1(\mathbf{X}) = 1$, and make the substitution $\mathbf{I} = \mathbf{E}_1(\mathbf{X}) + \mathbf{E}_2(\mathbf{X})$ on the right-hand side to get

$$\text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}]\mathbf{I} = \mathbf{E}_1(\mathbf{X})\mathbf{T} + \mathbf{T}\mathbf{E}_1(\mathbf{X}) - \mathbf{T} + (\text{tr } \mathbf{T})\mathbf{E}_2(\mathbf{X}).$$

Multiplication on the right by $\mathbf{E}_1(\mathbf{X})$ and use of the properties (1.4) yields

$$\text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}]\mathbf{E}_1(\mathbf{X}) = \mathbf{E}_1(\mathbf{X})\mathbf{T}\mathbf{E}_1(\mathbf{X}),$$

which with the aid of (2.1.7) becomes

$$\text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}]\mathbf{E}_1(\mathbf{X}) = \frac{1}{(x_1 - x_2)^2} [\mathbf{X}\mathbf{T}\mathbf{X} - x_2(\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) + x_2^2\mathbf{T}]. \quad (2.3.4)$$

A similar result holds for $\mathbf{E}_2(\mathbf{X})$, and consequently (2.3.2) can be written as

$$\begin{aligned} \left. \frac{d}{d\alpha} \mathbf{E}_1(\alpha) \right|_{\alpha=0} &= - \left. \frac{d}{d\alpha} \mathbf{E}_2(\alpha) \right|_{\alpha=0} \\ &= \frac{1}{(x_1 - x_2)^3} [(\text{tr } \mathbf{X})(\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) - 2\mathbf{X}\mathbf{T}\mathbf{X} - 2(\det \mathbf{X})\mathbf{T}]. \end{aligned} \quad (2.3.5)$$

Case 2 ($x_1 = x_2 = x$, $t_1 < t_2$). The occurrence of $|\alpha|$ in (2.1.8) restricts us to the calculation of one-sided derivatives. However, with (2.1.1), (2.1.5), (2.2.3), and l'Hospital's rule, we get

$$\left. \frac{d^+}{d\alpha} \mathbf{E}_1(\alpha) \right|_{\alpha=0} = \left. \frac{d^-}{d\alpha} \mathbf{E}_1(\alpha) \right|_{\alpha=0} = \mathbf{0};$$

thus,

$$\left. \frac{d}{d\alpha} \mathbf{E}_i(\alpha) \right|_{\alpha=0} = \mathbf{0}. \quad (2.3.6)$$

Case 3 ($x_1 = x_2 = x$, $t_1 = t_2 = t$). It follows from (2.1.10) that

$$\left. \frac{d}{d\alpha} \mathbf{E}_i(\alpha) \right|_{\alpha=0} = \mathbf{0}. \quad (2.3.7)$$

Remark. Even though the eigenprojections have directional derivatives in all directions at an \mathbf{X} with $x_1 = x_2$, they are not differentiable there; comparison of (2.3.6) and (2.3.7) with (2.3.5) shows that the directional derivatives are not continuous at such an \mathbf{X} (cf. the remark at the end of Subsection 2.1).

2.4 *The directional derivative of F.* Given $f: \mathbb{R} \rightarrow \mathbb{R}$, we define $\mathbf{F}: \text{Sym}(V_2) \rightarrow \text{Sym}(V_2)$ by

$$\mathbf{F}(\mathbf{X}) = \begin{cases} f(x_1)\mathbf{E}_1(\mathbf{X}) + f(x_2)\mathbf{E}_2(\mathbf{X}), & x_1 \neq x_2, \\ f(x)\mathbf{I}, & x_1 = x_2 = x. \end{cases} \quad (2.4.1)$$

We assume that f has as many derivatives as needed and denote them by primes in the sequel.

Case 1 ($x_1 < x_2$). Since

$$\mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) = \mathbf{F}(\mathbf{Y}(\alpha)) = f(y_1(\alpha))\mathbf{E}_1(\alpha) + f(y_2(\alpha))\mathbf{E}_2(\alpha),$$

(2.1.5), (2.1.7), (2.2.1), (2.3.4), and (2.3.5) yield

$$\begin{aligned} \left. \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \right|_{\alpha=0} &= \frac{1}{(x_1 - x_2)^3} \left[\{(x_1 - x_2)[f'(x_1) + f'(x_2)] \right. \\ &\quad - 2[f(x_1) - f(x_2)]\} \mathbf{X}\mathbf{T}\mathbf{X} + \{-(x_1 - x_2)[x_2 f'(x_1) + x_1 f'(x_2)] \\ &\quad + (x_1 + x_2)[f(x_1) - f(x_2)]\} (\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) \\ &\quad + \{(x_1 - x_2)[x_2^2 f'(x_1) - x_1^2 f'(x_2)] \\ &\quad \left. - 2x_1 x_2 [f(x_1) - f(x_2)]\} \mathbf{T} \right]. \end{aligned} \quad (2.4.2)$$

As expected in this case, there is continuity in \mathbf{X} for each \mathbf{T} ; so \mathbf{F} is differentiable, and

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \left. \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \right|_{\alpha=0}. \quad (2.4.3)$$

Case 2 ($x_1 = x_2 = x$, $t_1 < t_2$). Again

$$\mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) = f(y_1(\alpha))\mathbf{E}_1(\alpha) + f(y_2(\alpha))\mathbf{E}_2(\alpha),$$

but because of the difficulties with $dy_i(\alpha)/d\alpha$ and $\mathbf{E}_i(\alpha)$ here, we can calculate only one-sided derivatives. By (2.1.5), (2.1.9), (2.2.3), and (2.3.7),

$$\lim_{\alpha \rightarrow 0^+} \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) = \lim_{\alpha \rightarrow 0^-} \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) = f'(x)[t_1 \mathbf{E}_1(\mathbf{T}) + t_2 \mathbf{E}_2(\mathbf{T})].$$

Hence,

$$\left. \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \right|_{\alpha=0} = f'(x)\mathbf{T}. \quad (2.4.4)$$

It is interesting to note how the discontinuities of $\mathbf{E}_i(\alpha)$ and $dy_i(\alpha)/d\alpha$ combine here to give a continuous result for any choice of f .

Case 3 ($x_1 = x_2 = x$, $t_1 = t_2 = t$). As was noted in Subsec. 2.1, $d(\alpha) = 0$ here; so it follows from (2.1.2) and (2.1.3) that $y_1(\alpha) = y_2(\alpha) = \frac{1}{2} \text{tr } \mathbf{X} + \frac{1}{2} \alpha \text{tr } \mathbf{T}$. Consequently,

$$\mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) = f(x + \alpha t)\mathbf{I},$$

and

$$\left. \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \right|_{\alpha=0} = f'(x)\mathbf{T}.$$

Thus, \mathbf{F} has a directional derivative in all cases, and in Cases 2 and 3 it is given by the same formula. Continuity will be investigated in the next subsection.

2.5 *The derivative of \mathbf{F} .* In the previous subsection, we showed that the function $\mathbf{F}: \text{Sym}(V_2) \rightarrow \text{Sym}(V_2)$ which is defined by (2.4.1), has a directional derivative, $d\mathbf{F}(\mathbf{X} + \alpha\mathbf{T})/d\alpha|_{\alpha=0}$, everywhere and in every direction. When the eigenvalues of \mathbf{X} are distinct, it is given by (2.4.2); and when the eigenvalues of \mathbf{X} are equal, it is given by (2.4.4). In this subsection, we will show that for each $\mathbf{T} \in \text{Sym}(V_2)$, $d\mathbf{F}(\mathbf{X} + \alpha\mathbf{T})/d\alpha|_{\alpha=0}$ is continuous at every $\mathbf{X} \in \text{Sym}(V_2)$ and thereby conclude that \mathbf{F} is differentiable on $\text{Sym}(V_2)$.

Obviously, we have continuity at any \mathbf{X} with distinct eigenvalues, so what remains to be shown is that for $d\mathbf{F}(\mathbf{X} + \alpha\mathbf{T})/d\alpha|_{\alpha=0}$ given by (2.4.2),

$$\lim_{\mathbf{X} \rightarrow x\mathbf{I}} \frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \Big|_{\alpha=0} = f'(x)\mathbf{T} \tag{2.5.1}$$

for any $x \in \mathbb{R}$.

Equation (2.4.2) is of the form

$$\frac{d}{d\alpha} \mathbf{F}(\mathbf{X} + \alpha\mathbf{T}) \Big|_{\alpha=0} = a(x_1, x_2)\mathbf{X}\mathbf{T}\mathbf{X} + b(x_1, x_2)(\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) + c(x_1, x_2)\mathbf{T}. \tag{2.5.2}$$

There is never any difficulty with the tensor factors; indeed,

$$\mathbf{X}\mathbf{T}\mathbf{X} \rightarrow x^2\mathbf{T} \quad \text{and} \quad (\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) \rightarrow 2x\mathbf{T} \quad \text{as} \quad \mathbf{X} \rightarrow x\mathbf{I}. \tag{2.5.3}$$

However, the scalar coefficients are not well defined, and their limits require more study.

Consider

$$a(x_1, x_2) = \frac{1}{(x_1 - x_2)^3} \{ (x_1 - x_2)[f'(x_1) + f'(x_2)] - 2[f(x_1) - f(x_2)] \}.$$

For any $x_1 < x_2$,

$$a(x_1, x_2) = \frac{1}{2d^3} \left\{ d \left[f' \left(\xi - \frac{d}{\sqrt{2}} \right) + f' \left(\xi + \frac{d}{\sqrt{2}} \right) \right] + \sqrt{2} \left[f \left(\xi - \frac{d}{\sqrt{2}} \right) - f \left(\xi + \frac{d}{\sqrt{2}} \right) \right] \right\},$$

where d is the distance of the point (x_1, x_2) from the line $x_1 = x_2$, and (ξ, ξ) is the intersection of the line $x_1 = x_2$ with the line normal to it which passes through (x_1, x_2) . Using the proof of l'Hospital's rule, it is not difficult for one to show that if f is four times continuously differentiable, then

$$\lim_{d \rightarrow 0} a(x_1, x_2) = \frac{1}{6} f'''(\xi) \quad \text{uniformly w.r.t. } \xi.$$

Obviously,

$$\lim_{\xi \rightarrow x} f'''(\xi) = f'''(x),$$

and therefore

$$\lim_{(x_1, x_2) \rightarrow (x, x)} a(x_1, x_2) = \frac{1}{6} f'''(x). \tag{2.5.4}$$

In the same manner, we obtain

$$\left. \begin{aligned} \lim_{(x_1, x_2) \rightarrow (x, x)} b(x_1, x_2) &= \frac{1}{6} [3f'''(x) - xf''''(x)], \\ \lim_{(x_1, x_2) \rightarrow (x, x)} c(x_1, x_2) &= \frac{1}{6} [6f'(x) - 6xf''(x) + x^2f'''(x)]. \end{aligned} \right\} \quad (2.5.5)$$

Equations (2.5.2)–(2.5.5) imply (2.5.1). Hence, we have established the following

THEOREM. Given $f: \mathbb{R} \rightarrow \mathbb{R}$, define $\mathbf{F}: \text{Sym}(V_2) \rightarrow \text{Sym}(V_2)$ by

$$\mathbf{F}(\mathbf{X}) = \begin{cases} f(x_1)\mathbf{E}_1(\mathbf{X}) + f(x_2)\mathbf{E}_2(\mathbf{X}), & x_1 \neq x_2, \\ f(x)\mathbf{I}, & x_1 = x_2 = x. \end{cases}$$

Then, if f is four times continuously differentiable on \mathbb{R} , \mathbf{F} is continuously differentiable on $\text{Sym}(V_2)$ and

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \begin{cases} \frac{1}{(x_1 - x_2)^3} \left[\left\{ (x_1 - x_2)[f'(x_1) + f'(x_2)] - 2[f(x_1) - f(x_2)] \right\} \mathbf{X}\mathbf{T}\mathbf{X} \right. \\ \quad \left. + \left\{ -(x_1 - x_2)[x_2f'(x_1) + x_1f'(x_2)] \right. \right. \\ \quad \left. \left. + (x_1 + x_2)[f(x_1) - f(x_2)] \right\} (\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) \right. \\ \quad \left. + \left\{ (x_1 - x_2)[x_2^2f'(x_1) + x_1^2f'(x_2)] \right. \right. \\ \quad \left. \left. - 2x_1x_2[f(x_1) - f(x_2)] \right\} \mathbf{T} \right], & x_1 \neq x_2, \\ f'(x)\mathbf{T}, & x_1 = x_2 = x. \end{cases} \quad (2.5.6)$$

For some choices of f (such as $f(x) = x^2$ or even $f(x) = \sqrt{x}$), the two formulas in (2.5.6) admit a common expression for both $x_1 \neq x_2$ and $x_1 = x_2$. In other cases (such as $f(x) = \ln x$), they do not.

3. The three-dimensional case. One could proceed here as we did in the two-dimensional case. However, it is more economical to use Ball's lemma to extend the formula for the derivative derived for three distinct eigenvalues to the case of repeated eigenvalues. The notation parallels that used in Sec. 2.

3.1 Three distinct eigenvalues. For $\mathbf{X}, \mathbf{T} \in \text{Sym}(V_3)$, the characteristic equation of $\mathbf{Y}(\alpha) = \mathbf{X} + \alpha\mathbf{T}$ is

$$y_i(\alpha)^3 - I_1(\mathbf{Y}(\alpha))y_i(\alpha)^2 + I_2(\mathbf{Y}(\alpha))y_i(\alpha) - I_3(\mathbf{Y}(\alpha)) = 0, \quad (3.1.1)$$

where the coefficients are the principal invariants

$$I_1(\mathbf{A}) = \text{tr } \mathbf{A}, \quad I_2(\mathbf{A}) = \frac{1}{2} [(\text{tr } \mathbf{A})^2 - \text{tr}(\mathbf{A}^2)], \quad I_3(\mathbf{A}) = \det \mathbf{A}. \quad (3.1.2)$$

As noted in the remark following (2.2.4), the unordered $y_i(\alpha)$ may be chosen so as to be smooth functions of α . Consequently, we can differentiate (3.1.1).

The derivatives of the invariants are found by Truesdell and Noll [13, p. 26] to be³

$$\left. \begin{aligned} DI_1(\mathbf{A})[\mathbf{C}] &= \text{tr } \mathbf{C}, & DI_2(\mathbf{A})[\mathbf{C}] &= -\text{tr}\{[\mathbf{A} - I_1(\mathbf{A})\mathbf{I}]\mathbf{C}\}, \\ DI_3(\mathbf{A})[\mathbf{C}] &= \text{tr}\{[\mathbf{A}^2 - I_1(\mathbf{A})\mathbf{A} + I_2(\mathbf{A})\mathbf{I}]\mathbf{C}\}. \end{aligned} \right\} \quad (3.1.3)$$

Equations (3.1.1)–(3.1.3) together with

$$y_i(0) = x_i, \quad I_i(\mathbf{Y}(0)) = I_i(\mathbf{X})$$

yield

$$\frac{d}{d\alpha} y_1(\alpha) \Big|_{\alpha=0} = \text{tr} \left\{ \frac{(\mathbf{X} - x_2\mathbf{I})(\mathbf{X} - x_3\mathbf{I})}{(x_1 - x_2)(x_1 - x_3)} \mathbf{T} \right\} = \text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}].$$

In general,

$$Dx_i(\mathbf{X})[\mathbf{T}] = \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]. \quad (3.1.4)^4$$

Next we consider the derivatives of the eigenprojections. From (1.3),

$$\mathbf{E}_1(\mathbf{X}) = \frac{(\mathbf{X} - x_2\mathbf{I})(\mathbf{X} - x_3\mathbf{I})}{(x_1 - x_2)(x_1 - x_3)};$$

and, therefore, with (3.1.4) and

$$\mathbf{X} = x_1\mathbf{E}_1(\mathbf{X}) + x_2\mathbf{E}_2(\mathbf{X}) + x_3\mathbf{E}_3(\mathbf{X}),$$

we find that

$$\begin{aligned} D\mathbf{E}_1(\mathbf{X})[\mathbf{T}] &= \frac{1}{(x_1 - x_2)(x_1 - x_3)} [\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X} - (x_2 + x_3)\mathbf{T} \\ &\quad - [(x_1 - x_2) + (x_1 - x_3)] \text{tr}[\mathbf{E}_1(\mathbf{X})\mathbf{T}]\mathbf{E}_1(\mathbf{X}) \\ &\quad - (x_2 - x_3) \{ \text{tr}[\mathbf{E}_2(\mathbf{X})\mathbf{T}]\mathbf{E}_2(\mathbf{X}) + \text{tr}[\mathbf{E}_3(\mathbf{X})\mathbf{T}]\mathbf{E}_3(\mathbf{X}) \}]. \end{aligned} \quad (3.1.5)$$

The derivatives of \mathbf{E}_2 and \mathbf{E}_3 may be obtained from (3.1.5) by the cyclic permutations $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ of the subscripts. Equation (3.1.5) can be simplified with the aid of Rivlin's [12] identity

$$\begin{aligned} \mathbf{A}\mathbf{B}\mathbf{A} + \mathbf{A}^2\mathbf{B} + \mathbf{B}\mathbf{A}^2 - (\text{tr } \mathbf{A})(\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}) - (\text{tr } \mathbf{B})\mathbf{A}^2 \\ - [\text{tr}(\mathbf{A}\mathbf{B}) - (\text{tr } \mathbf{A})(\text{tr } \mathbf{B})]\mathbf{A} + \frac{1}{2} [(\text{tr } \mathbf{A})^2 - \text{tr}(\mathbf{A}^2)]\mathbf{B} \\ - \left\{ \text{tr}(\mathbf{A}^2\mathbf{B}) - (\text{tr } \mathbf{A})\text{tr}(\mathbf{A}\mathbf{B}) + \frac{1}{2}(\text{tr } \mathbf{B}) [(\text{tr } \mathbf{A})^2 - \text{tr}(\mathbf{A}^2)] \right\} \mathbf{I} = \mathbf{0}. \end{aligned}$$

Take $\mathbf{A} = \mathbf{E}_i(\mathbf{X})$, $\mathbf{B} = \mathbf{T}$, and use the properties (1.4) and $\text{tr } \mathbf{E}_i(\mathbf{X}) = 1$ to get

$$\text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]\mathbf{E}_i(\mathbf{X}) = \mathbf{E}_i(\mathbf{X})\mathbf{T}\mathbf{E}_i(\mathbf{X}). \quad (3.1.6)$$

Finally, we turn to the derivative of the function $\mathbf{F}: \text{Sym}(V_3) \rightarrow \text{Sym}(V_3)$ defined by

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^3 f(x_i)\mathbf{E}_i(\mathbf{X}).$$

³ The derivation in [13] rests on the assumption that \mathbf{A} be invertible. This is not essential for the final result; each principal invariant can be expressed as an easily differentiated function of traces of powers (cf. [14]).

⁴ Cf. the remark following (2.2.2).

The differentiability of x_i and \mathbf{E}_i have already been established in the present case; consequently, if f is differentiable,

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \sum_{i=1}^3 \{ f'(x_i) Dx_i(\mathbf{X})[\mathbf{T}]\mathbf{E}_i(\mathbf{X}) + f(x_i) D\mathbf{E}_i(\mathbf{X})[\mathbf{T}] \}.$$

Then with (3.1.4)–(3.1.6) and (1.3), we obtain

$$\begin{aligned} D\mathbf{F}(\mathbf{X})[\mathbf{T}] &= (\mathcal{F}_1 + \mathcal{F}_2 + \mathcal{F}_3)\mathbf{X}^2\mathbf{T}\mathbf{X}^2 \\ &\quad - [(x_1 + x_2)\mathcal{F}_3 + (x_2 + x_3)\mathcal{F}_1 + (x_3 + x_1)\mathcal{F}_2](\mathbf{X}^2\mathbf{T}\mathbf{X} + \mathbf{X}\mathbf{T}\mathbf{X}^2) \\ &\quad + (x_1x_2\mathcal{F}_3 + x_2x_3\mathcal{F}_1 + x_3x_1\mathcal{F}_2)(\mathbf{X}^2\mathbf{T} + \mathbf{T}\mathbf{X}^2) \\ &\quad + [(x_1 + x_2)^2\mathcal{F}_3 + (x_2 + x_3)^2\mathcal{F}_1 + (x_3 + x_1)^2\mathcal{F}_2]\mathbf{X}\mathbf{T}\mathbf{X} \\ &\quad + [-x_1x_2(x_1 + x_2)\mathcal{F}_3 - x_2x_3(x_2 + x_3)\mathcal{F}_1 - x_3x_1(x_3 + x_1)\mathcal{F}_2 \\ &\quad + \mathcal{f}_1 + \mathcal{f}_2 + \mathcal{f}_3](\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) \\ &\quad + [(x_1x_2)^2\mathcal{F}_3 + (x_2x_3)^2\mathcal{F}_1 + (x_3x_1)^2\mathcal{F}_2 \\ &\quad - (x_1 + x_2)\mathcal{f}_3 - (x_2 + x_3)\mathcal{f}_1 - (x_3 + x_1)\mathcal{f}_2]\mathbf{T}, \end{aligned} \tag{3.1.7}$$

where

$$\left. \begin{aligned} \mathcal{f}_i &= \frac{f(x_i)}{(x_i - x_j)(x_i - x_k)}, \\ \mathcal{F}_i &= \frac{1}{(x_i - x_j)^2(x_i - x_k)^2} \left[f'(x_i) - (x_i - x_j)(\mathcal{f}_i + \mathcal{f}_k) - (x_i - x_k)(\mathcal{f}_i + \mathcal{f}_j) \right] \end{aligned} \right\} \tag{3.1.8}$$

($j \neq i \neq k \neq j$).

3.2 Ball's lemma. According to the results of the previous subsection, \mathbf{F} is differentiable on $\text{Sym}(V_3)$, except possibly at those tensors with repeated eigenvalues. Recently, Ball [8, Proposition 2.2] provided a technical lemma which is exactly what is needed here to establish differentiability where the eigenvalues are repeated without calculating the individual cases separately as was done in Sec. 2.

The essential ingredient of the lemma is the notion of a sparse set. A set $\mathcal{X} \subset \mathbb{R}^m$ is *sparse* if given any $z \in \mathcal{X}$ and any nonzero $\zeta \in \mathbb{R}^m$ there exist sequences $\{z_j\} \rightarrow z$, $\{\zeta_j\} \rightarrow \zeta$ and a number $\epsilon > 0$ such that for each $j = 1, 2, \dots$ the line segment $\{z_j + t\zeta_j; t \in [0, \epsilon]\}$ intersects \mathcal{X} at most countably often. Then a variant of Ball's lemma suitable for our purposes may be stated as follows.

Let $\mathcal{U} \subset \mathbb{R}^n$ be open and let $\mathcal{X} \subset \mathbb{R}^n$ be closed and sparse. Let $g: \mathcal{U} \rightarrow \mathbb{R}^l$ be continuous on \mathcal{U} and continuously differentiable on $\mathcal{U} - \mathcal{X}$. If for each $y \in \mathcal{U} \cap \mathcal{X}$, the limit

$$\lim_{\substack{x \rightarrow y \\ x \in \mathcal{U} - \mathcal{X}}} Dg(x)$$

exists, then g is continuously differentiable on \mathcal{U} .

In our case, we identify $\text{Sym}(V_3)$ with \mathbb{R}^6 and take $\mathcal{U} = \text{Sym}(V_3)$. The exceptional set \mathcal{X} is the set of tensors in $\text{Sym}(V_3)$ with repeated eigenvalues. Ball [8, Sec. 5] has shown this set to be closed and sparse.

3.3 *The continuity of F.* Combining (1.2) and (1.3), we see that the function $\mathbf{F}: \text{Sym}(V_3) \rightarrow \text{Sym}(V_3)$ corresponding to $f: \mathbb{R} \rightarrow \mathbb{R}$ can be expressed as

$$\mathbf{F}(\mathbf{X}) = \begin{cases} \frac{1}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)} \left\{ -[(x_1 - x_2)f(x_3) + (x_2 - x_3)f(x_1) + (x_3 - x_1)f(x_2)]\mathbf{X}^2 + [(x_1^2 - x_2^2)f(x_3) + (x_2^2 - x_3^2)f(x_1) + (x_3^2 - x_1^2)f(x_2)]\mathbf{X} - [(x_1 - x_2)x_1x_2f(x_3) + (x_2 - x_3)x_2x_3f(x_1) + (x_3 - x_1)x_3x_1f(x_2)]\mathbf{I} \right\}, & x_1 \neq x_2 \neq x_3 \neq x_1, & (3.3.1) \\ \frac{1}{x_1 - x} \{ [f(x_1) - f(x)]\mathbf{X} - [xf(x_1) - x_1f(x)]\mathbf{I} \}, & x_1 \neq x_2 = x_3 = x, & (3.3.2) \\ f(x)\mathbf{I}, & x_1 = x_2 = x_3 = x. & (3.3.3) \end{cases}$$

The purpose of the present subsection is to show that the function \mathbf{F} so defined is continuous on $\text{Sym}(V_3)$. Since the demonstration parallels the one in Subsection 2.5, not all of the details will be presented.

Clearly, \mathbf{F} is continuous at any \mathbf{X} with three distinct eigenvalues; and at such an \mathbf{X} , $\mathbf{F}(\mathbf{X})$ has the form

$$\mathbf{F}(\mathbf{X}) = a(x_1, x_2, x_3)\mathbf{X}^2 + b(x_1, x_2, x_3)\mathbf{X} + c(x_1, x_2, x_3)\mathbf{I}. \tag{3.3.4}$$

It remains to investigate the limit of this formula as the eigenvalues of \mathbf{X} coalesce.

Double coalescence. Consider $\mathbf{X} \rightarrow \bar{\mathbf{X}}$, where \mathbf{X} and $\bar{\mathbf{X}}$ have eigenvalues $x_1 \neq x_2 \neq x_3 \neq x_1$ and $\bar{x}_1 \neq \bar{x}_2 = \bar{x}_3 = x$, respectively. By (1.5), the spectral decomposition of $\bar{\mathbf{X}}$ can be written as

$$\bar{\mathbf{X}} = \bar{x}_1\mathbf{E}_1(\bar{\mathbf{X}}) + x[\mathbf{I} - \mathbf{E}_1(\bar{\mathbf{X}})];$$

and, therefore,

$$\bar{\mathbf{X}}^2 = \bar{x}_1^2\mathbf{E}_1(\bar{\mathbf{X}}) + x^2[\mathbf{I} - \mathbf{E}_1(\bar{\mathbf{X}})],$$

which with (1.3) becomes

$$\bar{\mathbf{X}}^2 = \frac{1}{\bar{x}_1 - x} [(\bar{x}_1^2 - x^2)\bar{\mathbf{X}} - (x\bar{x}_1^2 - \bar{x}_1x^2)\mathbf{I}]. \tag{3.3.5}$$

The scalar coefficients in (3.3.4) can be handled as follows. First let $(x_1, x_2, x_3) \rightarrow (x_1, \xi, \xi)$ along the line normal to the $x_2 = x_3$ plane. Then

$$\left. \begin{aligned} a(x_1, x_2, x_3) &\rightarrow \frac{1}{(x_1 - \xi)^2} [f(x_1) - f(\xi) - (x_1 - \xi)f'(\xi)], \\ b(x_1, x_2, x_3) &\rightarrow \frac{1}{(x_1 - \xi)^2} \{ -2\xi[f(x_1) - f(\xi)] + (x_1^2 - \xi^2)f'(\xi) \}, \\ c(x_1, x_2, x_3) &\rightarrow \frac{1}{(x_1 - \xi)^2} [\xi^2f(x_1) + (x_1 - 2\xi)x_1f(\xi) - \xi(x_1 - \xi)x_1f'(\xi)] \end{aligned} \right\} \tag{3.3.6}$$

uniformly w.r.t. x_1 and ξ . Moreover, the right-hand members in (3.3.6) are continuous in

x_1 and $\xi(x_1 \neq \xi)$. With this observation, we see that the limit of the right-hand side of (3.3.4) as $\mathbf{X} \rightarrow \bar{\mathbf{X}}$ is in agreement with (3.3.2).

Triple coalescence. Consider $\mathbf{X} \rightarrow x\mathbf{I}$. Suppose to begin with that \mathbf{X} has eigenvalues $x_1 \neq x_2 \neq x_3 \neq x_1$. To investigate the limits of the scalar coefficients in (3.3.4), we first let $(x_1, x_2, x_3) \rightarrow (\xi, \xi, \xi)$ along the line normal to the line $x_1 = x_2 = x_3$. Then

$$\begin{aligned} a(x_1, x_2, x_3) &\rightarrow \frac{-1}{2(d_1 - d_2)(d_2 - d_3)(d_3 - d_1)} \left[(d_1 - d_2)d_3^2 \right. \\ &\quad \left. + (d_2 - d_3)d_1^2 + (d_3 - d_1)d_2^2 \right] f''(\xi), \\ b(x_1, x_2, x_3) &\rightarrow \frac{1}{(d_1 - d_2)(d_2 - d_3)(d_3 - d_1)} \left\{ [(d_1^2 - d_2^2)d_3 \right. \\ &\quad \left. + (d_2^2 - d_3^2)d_1 + (d_3^2 - d_1^2)d_2] f'(\xi) \right. \\ &\quad \left. + [(d_1 - d_2)d_3^2 + (d_2 - d_3)d_1^2 + (d_3 - d_1)d_2^2] \xi f''(\xi) \right\}, \quad (3.3.7)^5 \\ c(x_1, x_2, x_3) &\rightarrow \frac{-1}{(d_1 - d_2)(d_2 - d_3)(d_3 - d_1)} \left\{ [(d_1 - d_2)d_1d_2 \right. \\ &\quad \left. + (d_2 - d_3)d_2d_3 + (d_3 - d_1)d_3d_1] f(\xi) \right. \\ &\quad \left. + [(d_1^2 - d_2^2)d_3 + (d_2^2 - d_3^2)d_1 + (d_3^2 - d_1^2)d_2] \xi f'(\xi) \right. \\ &\quad \left. + \frac{1}{2} [(d_1 - d_2)d_3^2 + (d_2 - d_3)d_1^2 + (d_3 - d_1)d_2^2] \xi^2 f''(\xi) \right\} \end{aligned}$$

uniformly w.r.t. ξ . Here, the d_i are the components of the unit vector directed from (ξ, ξ, ξ) toward (x_1, x_2, x_3) . Since the right-hand members of (3.3.7) are continuous in ξ and since $\mathbf{X}^2 \rightarrow x^2\mathbf{I}$, we see that the limit of the right-hand side of (3.3.4) as $\mathbf{X} \rightarrow x\mathbf{I}$ agrees with (3.3.3).

If \mathbf{X} has two distinct eigenvalues, say $x_1 \neq x_2 = x_3 = \hat{x}$, then by (3.3.2) $\mathbf{F}(\mathbf{X})$ has the form

$$\mathbf{F}(\mathbf{X}) = r(x_1, \hat{x})\mathbf{X} + s(x_1, \hat{x})\mathbf{I}.$$

The limits of the functions r and s as $\mathbf{X} \rightarrow x\mathbf{I}$ can be worked out by the method used in Subsec. 2.5. The result is $r(x_1, \hat{x}) \rightarrow f'(x)$, $s(x_1, \hat{x}) \rightarrow f(x) - xf'(x)$; consequently, $\mathbf{F}(\mathbf{X}) \rightarrow f(x)\mathbf{I}$, which is in agreement with (3.3.3).

Thus, we have the following

THEOREM. If $f: \mathbb{R} \rightarrow \mathbb{R}$ is three times continuously differentiable, the function $\mathbf{F}: \text{Sym}(V_3) \rightarrow \text{Sym}(V_3)$ defined by (3.3.1)–(3.3.3)⁶ is continuous.

3.4 Extension to repeated eigenvalues. Finally, in order to invoke Ball's lemma, we must show that the expression obtained for $D\mathbf{F}(\mathbf{X})[\mathbf{T}]$ in Sec. 3.1 for three distinct eigenvalues has continuous extensions to the two cases of repeated eigenvalues. The limits can be calculated exactly as in the previous subsection, and the results are summarized in the

⁵ If the product $(d_1 - d_2)(d_2 - d_3)(d_3 - d_1) = 0$, then at least one of the factors must vanish. Suppose $d_2 - d_3 = 0$. Then $x_2 - \xi = x_3 - \xi$, which contradicts our starting assumption.

⁶ Alternatively, by (1.2) and (1.3) for $n = 3$.

following

THEOREM. Given $f: \mathbb{R} \rightarrow \mathbb{R}$, define $\mathbf{F}: \text{Sym}(V_3) \rightarrow \text{Sym}(V_3)$ by

$$\mathbf{F}(\mathbf{X}) = \begin{cases} f(x_1)\mathbf{E}_1(\mathbf{X}) + f(x_2)\mathbf{E}_2(\mathbf{X}) + f(x_3)\mathbf{E}_3(\mathbf{X}), & x_1 \neq x_2 \neq x_3 \neq x_1, \\ f(x_1)\mathbf{E}_1(\mathbf{X}) + f(x)[\mathbf{I} - \mathbf{E}_1(\mathbf{X})], & x_1 \neq x_2 = x_3 = x,^7 \\ f(x)\mathbf{I}, & x_1 = x_2 = x_3 = x. \end{cases}$$

Then, if f is seven times continuously differentiable on \mathbb{R} , \mathbf{F} is continuously differentiable on $\text{Sym}(V_3)$ and

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \left\{ \begin{aligned} & (\mathcal{F}_1 + \mathcal{F}_2 + \mathcal{F}_3)\mathbf{X}^2\mathbf{T}\mathbf{X}^2 - [(x_1 + x_2)\mathcal{F}_3 + (x_2 + x_3)\mathcal{F}_1 \\ & \quad + (x_3 + x_1)\mathcal{F}_2](\mathbf{X}^2\mathbf{T}\mathbf{X} + \mathbf{X}\mathbf{T}\mathbf{X}^2) \\ & \quad + (x_1x_2\mathcal{F}_3 + x_2x_3\mathcal{F}_1 + x_3x_1\mathcal{F}_2)(\mathbf{X}^2\mathbf{T} + \mathbf{T}\mathbf{X}^2) \\ & \quad + [(x_1 + x_2)^2\mathcal{F}_3 + (x_2 + x_3)^2\mathcal{F}_1 \\ & \quad + (x_3 + x_1)^2\mathcal{F}_2]\mathbf{X}\mathbf{T}\mathbf{X} \\ & \quad + [-x_1x_2(x_1 + x_2)\mathcal{F}_3 - x_2x_3(x_2 + x_3)\mathcal{F}_1 \\ & \quad - x_3x_1(x_3 + x_1)\mathcal{F}_2 + \not{f}_1 + \not{f}_2 + \not{f}_3](\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X}) \\ & \quad + [(x_1x_2)^2\mathcal{F}_3 + (x_2x_3)^2\mathcal{F}_1 \\ & \quad + (x_3x_1)^2\mathcal{F}_2 - (x_1 + x_2)\not{f}_3 - (x_2 + x_3)\not{f}_1 \\ & \quad - (x_3 + x_1)\not{f}_2]\mathbf{T}, \quad x_1 \neq x_2 \neq x_3 \neq x_1, \\ & \quad \frac{1}{(x_1 - x)^3} \{ [(x_1 - x)[f'(x_1) + f'(x)] \\ & \quad - 2[f(x_1) - f(x)]\}\mathbf{X}\mathbf{T}\mathbf{X} \\ & \quad + \{ -(x_1 - x)[xf'(x_1) + x_1f'(x)] \\ & \quad + (x_1 + x)[f(x_1) - f(x)]\}\mathbf{X}\mathbf{T} + \mathbf{T}\mathbf{X} \\ & \quad + \{(x_1 - x)[x^2f'(x_1) + x_1^2f(x)] \\ & \quad - 2x_1x[f(x_1) - f(x)]\}\mathbf{T}, \quad x_1 \neq x_2 = x_3 = x, \\ & \quad f'(x)\mathbf{T}, \quad x_1 = x_2 = x_3 = x, \end{aligned} \right. \tag{3.4.1}$$

where \not{f}_i and \mathcal{F}_i are given by (3.1.8).

Remark. Note that the expression for $D\mathbf{F}(\mathbf{X})[\mathbf{T}]$ when the underlying vector space has dimension $n = 3$ and \mathbf{X} has two distinct eigenvalues is of the same form as the one appropriate to $n = 2$ and \mathbf{X} having two distinct eigenvalues. There is similar agreement between the $n = 2$ and $n = 3$ cases when \mathbf{X} has one distinct eigenvalue in each case. We conjecture that, in general, the formula for $D\mathbf{F}(\mathbf{X})[\mathbf{T}]$ depends only on the number of distinct eigenvalues of \mathbf{X} and is independent of the dimension of the underlying space.

⁷ Here (1.5) has been used to write $\mathbf{E}_2(\mathbf{X}) = \mathbf{I} - \mathbf{E}_1(\mathbf{X})$.

4. The n -dimensional case for n distinct eigenvalues. Here we sketch the extensions of the results of Subsec. 3.1 to the n -dimensional case. We do not consider repeated eigenvalues; however, for any choice of n and any particular type of coalescence, the methods of Subsec. 3.4 could be applied.

For $\mathbf{X}, \mathbf{T} \in \text{Sym}(V_n)$ the characteristic equation of $\mathbf{Y}(\alpha) = \mathbf{X} + \alpha\mathbf{T}$ is

$$\sum_{k=0}^n (-1)^k I_k(\mathbf{Y}(\alpha)) y_i(\alpha)^{n-k} = 0, \tag{4.1}$$

where $I_0 \equiv 1$. Since⁸

$$DI_k(\mathbf{X})[\mathbf{T}] = \text{tr} \left[\mathbf{T} \sum_{j=0}^{k-1} I_{k-1-j}(\mathbf{X})(-\mathbf{X})^j \right], \quad k = 1, 2, \dots, n,$$

differentiation of (4.1) yields⁹

$$\Delta_i \left. \frac{d}{d\alpha} y_i(\alpha) \right|_{\alpha=0} + N_i = 0, \tag{4.2}$$

where

$$\Delta_i = \sum_{k=0}^{n-1} (-1)^k (n-k) I_k(\mathbf{X}) x_i^{n-1-k}, \tag{4.3}$$

$$N_i = \text{tr}(\mathbf{N}_i \mathbf{T}), \tag{4.4}$$

$$\mathbf{N}_i = \sum_{k=1}^n (-1)^k x_i^{n-k} \left[\sum_{j=0}^{k-1} I_{k-1-j}(\mathbf{X})(-\mathbf{X})^j \right]. \tag{4.5}$$

By expanding

$$(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n),$$

it is not difficult for one to see on comparison with (4.3) that

$$\Delta_i = (x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n). \tag{4.6}$$

Similarly, expansion of

$$(\mathbf{X} - x_1 \mathbf{I}) \cdots (\mathbf{X} - x_{i-1} \mathbf{I})(\mathbf{X} - x_{i+1} \mathbf{I}) \cdots (\mathbf{X} - x_n \mathbf{I})$$

and use of the characteristic equation of \mathbf{X} in (4.5) leads to

$$\mathbf{N}_i = -(\mathbf{X} - x_1 \mathbf{I}) \cdots (\mathbf{X} - x_{i-1} \mathbf{I})(\mathbf{X} - x_{i+1} \mathbf{I}) \cdots (\mathbf{X} - x_n \mathbf{I}). \tag{4.7}$$

Equations (4.2), (4.4), (4.6), (4.7), and (1.3) imply

$$Dx_i(\mathbf{X})[\mathbf{T}] = \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]. \tag{4.8}^{10}$$

Next we consider the derivatives of the eigenprojections. It is convenient here to write (1.3) in the form

$$\mathbf{E}_i(\mathbf{X}) = \prod_{j=1}^n \mathbf{P}_{ij}(\mathbf{X}), \tag{4.9}$$

⁸ Truesdell and Noll [13, p. 26]. Cf. the first footnote in Subsection 3.1.

⁹ Cf. the remark following (2.2.5).

¹⁰ Cf. the remark following (2.2.2).

where

$$P_{ij}(\mathbf{X}) = \begin{cases} \frac{\mathbf{X} - x_j \mathbf{I}}{x_i - x_j}, & i \neq j, \\ \mathbf{I}, & i = j. \end{cases} \tag{4.10}$$

By (4.8) and (4.10),

$$DP_{ij}(\mathbf{X})[\mathbf{T}] = \begin{cases} \frac{1}{x_i - x_j} [\mathbf{T} - \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]\mathbf{I} - \text{tr}\{[\mathbf{E}_i(\mathbf{X}) - \mathbf{E}_j(\mathbf{X})]\mathbf{T}\}P_{ij}(\mathbf{X})], & i \neq j, \\ \mathbf{0}, & i = j. \end{cases} \tag{4.11}$$

With (4.11), differentiation of (4.9) yields

$$DE_i(\mathbf{X})[\mathbf{T}] = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{x_i - x_j} [P_{i1}(\mathbf{X}) \cdots P_{i(j-1)}(\mathbf{X}) \{ \mathbf{T} - \text{tr}[\mathbf{E}_i(\mathbf{X})\mathbf{T}]\mathbf{I} \} P_{i(j+1)}(\mathbf{X}) \cdots P_{in}(\mathbf{X}) - \text{tr}\{[\mathbf{E}_i(\mathbf{X}) - \mathbf{E}_j(\mathbf{X})]\mathbf{T}\}E_i(\mathbf{X})]. \tag{4.12}$$

Finally, if f is differentiable, then the derivative of the function $\mathbf{F}: \text{Sym}(V_n) \rightarrow \text{Sym}(V_n)$ defined by

$$\mathbf{F}(\mathbf{X}) = \sum_{i=1}^n f(x_i)E_i(\mathbf{X})$$

is

$$D\mathbf{F}(\mathbf{X})[\mathbf{T}] = \sum_{i=1}^n \{ f'(x_i)Dx_i(\mathbf{X})[\mathbf{T}]E_i(\mathbf{X}) + f(x_i)DE_i(\mathbf{X})[\mathbf{T}] \},$$

where $Dx_i(\mathbf{X})[\mathbf{T}]$ and $DE_i(\mathbf{X})[\mathbf{T}]$ are given by (4.8) and (4.12), respectively.

REFERENCES

- [1] M. E. Gurtin, *An introduction to continuum mechanics*, Academic Press, New York, 1981
- [2] Z.-h. Guo, *Rates of stretch tensors*, J. Elasticity **14**, 263–267 (1984)
- [3] M. E. Gurtin and K. Spear, *On the relationship between the logarithmic strain rate and the stretching tensor*, Internat. J. Solids Struct. **19**, 437–444 (1983)
- [4] A. Hoger and D. E. Carlson, *On the derivative of the square root of a tensor and Guo’s rate theorems*, J. Elasticity **14**, 329–336 (1984)
- [5] R. F. Rinehart, *The equivalence of definitions of a matrix function*. Amer. Math. Monthly **62**, 395–414 (1955)
- [6] K. Hoffman and R. Kunze, *Linear algebra*, Prentice-Hall, Englewood Cliffs, 1961
- [7] P. R. Halmos, *Finite-dimensional vector spaces*, 2nd Ed., D. Van Nostrand, Princeton, 1958
- [8] J. M. Ball, *Differentiability properties of symmetric and isotropic functions*, Duke Math. J. **51**, 699–728 (1984)
- [9] A. Hoger, *The time derivative of logarithmic strain*, Internat. J. Solids Struct. (to appear)
- [10] T. Kato, *Perturbation theory for linear operators*, 2nd Ed., Springer, Berlin-Heidelberg-New York, 1976
- [11] F. Rellich, *Perturbation theory of eigenvalue problems*, Gordon and Breach, New York, 1969
- [12] R. S. Rivlin, *Further remarks on the stress-deformation relations for isotropic materials*, J. Rat. Mech. Anal. **4**, 681–702 (1955)
- [13] C. Truesdell and W. Noll, *The non-linear field theories of mechanics*, Handbuch der Physik III/3, Springer, Berlin-Heidelberg-New York, 1965
- [14] D. E. Carlson and A. Hoger, *On the derivatives of the principal invariants of a second-order tensor*, J. Elasticity (to appear)