

HIGHER DERIVATIVES AND THE INVERSE DERIVATIVE OF A TENSOR-VALUED FUNCTION OF A TENSOR

BY

ANDREW N. NORRIS

Mechanical and Aerospace Engineering, Rutgers University, Piscataway NJ 08854-8058, USA

Abstract.

The n^{th} derivative of a tensor valued function of a tensor is defined by a finite number of coefficients each with closed form expression.

1. Introduction. We consider tensor functions on symmetric second order tensors, $\text{Sym} \rightarrow \text{Sym}$, defined by a scalar function $f(x)$ of a single variable according to

$$f(\mathbf{A}) = \sum_{i=1}^d f(\alpha_i) \mathbf{A}_i, \quad (1.1)$$

where d is the eigen-index of $\mathbf{A}$. The tensor $\mathbf{A} \in \text{Sym}$ is arbitrary with spectral decomposition

$$\mathbf{A} = \sum_{i=1}^d \alpha_i \mathbf{A}_i, \quad \mathbf{A}_i \mathbf{A}_j = \begin{cases} \mathbf{A}_i & i = j, \\ 0, & i \neq j, \end{cases} \quad \sum_{i=1}^d \mathbf{A}_i = \mathbf{I}. \quad (1.2)$$

The specific case of tensors acting on 3-dimensional vectors, $d \leq 3$, is discussed in this paper, although the results can be readily generalized.

Derivatives of $f(\mathbf{A})$ are defined by the expansion

$$\begin{aligned} f(\mathbf{A} + \mathbf{X}) = & f(\mathbf{A}) + \nabla f(\mathbf{A}) \mathbf{X} + \frac{1}{2} \nabla^{(2)} f(\mathbf{A}) : \mathbf{X} \mathbf{X} + \frac{1}{3!} \nabla^{(3)} f(\mathbf{A}) : \mathbf{X} \mathbf{X} \mathbf{X} \\ & + \dots + \frac{1}{n!} \nabla^{(n)} f(\mathbf{A}) : \underbrace{\mathbf{X} \mathbf{X} \dots \mathbf{X}}_{n \text{ terms}} + \dots \end{aligned} \quad (1.3)$$

The n^{th} derivative $\nabla^{(n)} f(\mathbf{A})$ is a tensor of order $2(n+1)$ which contracts n -times with the second order tensor $\mathbf{X}$ to produce a second order tensor. The first derivative, or gradient, is

$$\nabla f(\mathbf{A}) = \sum_{i,j=1}^d f_{ij}^{(1)} \mathbf{A}_i \boxtimes \mathbf{A}_j, \quad f_{ij}^{(1)} = \begin{cases} f'(\alpha_i), & i = j, \\ \frac{f(\alpha_i) - f(\alpha_j)}{\alpha_i - \alpha_j}, & i \neq j, \end{cases} \quad (1.4)$$

E-mail address: norris@rutgers.edu

where $\boxtimes$ denotes the outer tensor product, defined in Section 2. The identity (1.4) is well known and has appeared in various formats. The first explicit presentation I am aware of is due to Ogden [10] who defines a fourth order tensor $\mathcal{L}^1 = \partial f(\mathbf{A})/\partial \mathbf{A}$ (with slight change in notation). Ogden gives the components $\mathcal{L}_{ijkl}^1$ in terms of the eigenvectors $\mathbf{a}_i$ of $\mathbf{A}$, i.e. $\mathcal{L}^1 = \mathcal{L}_{ijkl}^1 \mathbf{a}_i \otimes \mathbf{a}_j \otimes \mathbf{a}_k \otimes \mathbf{a}_l$. These coefficients are related to those of (1.4) by the fact that $\mathbf{A}_i = \mathbf{a}_i \otimes \mathbf{a}_i$ (no sum) when $d = 3$ and hence $\mathcal{L}_{iii}^1 = f_{ii}^{(1)}$, $\mathcal{L}_{ijij}^1 = f_{ij}^{(1)}$. The fundamental result (1.4) was also derived by Carlson and Hoger [1], although the present notation is based on [14].

Ogden [10] (Section 3.4) also presented the second derivative. In the present notation it is

$$\begin{aligned} \frac{1}{2} \nabla^{(2)} f(\mathbf{A}) &= \sum_{i,j,k=1}^d f_{ijk}^{(2)} \mathbf{A}_i \boxtimes \mathbf{A}_j \boxtimes \mathbf{A}_k, \\ f_{iii}^{(2)} &= \frac{1}{2} f''(\alpha_i), \\ f_{iij}^{(2)} &= \frac{f(\alpha_j) - f(\alpha_i) - (\alpha_j - \alpha_i) f'(\alpha_i)}{(\alpha_j - \alpha_i)^2}, \\ f_{ijk}^{(2)} &= \frac{[f(\alpha_j) - f(\alpha_i)](\alpha_i + \alpha_j - 2\alpha_k) - [f(\alpha_i) + f(\alpha_j) - 2f(\alpha_k)](\alpha_j - \alpha_i)}{2(\alpha_i - \alpha_j)(\alpha_j - \alpha_k)(\alpha_k - \alpha_i)}. \end{aligned} \quad (1.5)$$

Ogden's expressions (1.4) and (1.5) are special cases of the more general formulae for derivatives of isotropic tensor functions derived by Chadwick and Ogden [2].

It is interesting to note that the tensor gradient function (first derivative) involves finite differences of the function f as well as its derivative evaluated at the eigenvalues of $\mathbf{A}$. Similarly, the second derivative contains second order finite differences. The general result derived here shows that the coefficients for the n^{th} derivative are related to an interpolating polynomial.

Our main result is the following:

THEOREM 1.1. The n^{th} derivative of the tensor function $f(\mathbf{A})$ is given by

$$\frac{1}{n!} \nabla^{(n)} f(\mathbf{A}) = \sum_{i_1, i_2, \dots, i_{n+1}=1}^d f_{i_1, i_2, \dots, i_{n+1}}^{(n)} \mathbf{A}_{i_1} \boxtimes \mathbf{A}_{i_2} \boxtimes \dots \boxtimes \mathbf{A}_{i_{n+1}}. \quad (1.6)$$

The coefficients $f_{i_1, i_2, \dots, i_{n+1}}^{(n)}$ are unaltered under all permutations of the $n+1$ indices. The $(n+2)(n+3)/2$ distinct coefficients can be classified into¹ $\lfloor \frac{(n+4)^2+4}{12} \rfloor$ expressions $f_{ijk}^{\nu_i, \nu_j, \nu_k}$ where ν_i, ν_j, ν_k , are the number of occurrences of distinct indices i, j, k respectively, with $i \neq j \neq k \neq i$ and $\nu_i + \nu_j + \nu_k = n+1$. The coefficient is

$$f_{ijk}^{\nu_i, \nu_j, \nu_k} = \sum_{\substack{l=i,j,k \\ \nu_l > 0}} \frac{1}{(\nu_l - 1)!} \frac{d^{\nu_l-1}}{dx^{\nu_l-1}} \frac{f(x)}{\prod_{\substack{m=i,j,k \\ m \neq l}} (x - \alpha_m)^{\nu_m}} \Big|_{x=\alpha_l}. \quad (1.7)$$

Alternatively, the coefficient can be found from the unique interpolating polynomial $P(x)$ of degree n that fits the data at the three points $x = \{\alpha_i, \alpha_j, \alpha_k\}$ defined by the $n+1$

¹ $\lfloor x \rfloor$ is the floor function.

values $f^{(I)}(\alpha_i)$, $f^{(J)}(\alpha_j)$, and $f^{(K)}(\alpha_k)$ for $0 \leq I \leq \nu_i - 1$, $0 \leq J \leq \nu_j - 1$, $0 \leq K \leq \nu_k - 1$ where $f^{(l)}(x)$ is the l^{th} derivative. Let $P(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_0$, then

$$f_{ijk}^{\nu_i, \nu_j, \nu_k} = p_n. \quad (1.8)$$

The first few expressions for the coefficients are

$$f_{ijk}^{0,0,n+1} = \frac{1}{n!} f^{(n)}(\alpha_k), \quad (1.9a)$$

$$f_{ijk}^{0,1,n} = \frac{1}{(\alpha_j - \alpha_k)^{n-1}} \left[f(\alpha_j) - \sum_{l=0}^{n-1} \frac{1}{l!} (\alpha_j - \alpha_k)^l f^{(l)}(\alpha_k) \right], \quad (1.9b)$$

$$f_{ijk}^{0,2,n-1} = \frac{1}{(\alpha_j - \alpha_k)^{n-1}} \left[\sum_{l=0}^{n-1} \frac{(n-l)}{l!} (\alpha_j - \alpha_k)^l f^{(l)}(\alpha_k) + (\alpha_j - \alpha_k) f'(\alpha_j) - n f(\alpha_j) \right], \quad (1.9c)$$

$$f_{ijk}^{1,1,n-1} = \frac{1}{(\alpha_i - \alpha_j)} \left\{ \frac{f(\alpha_i)}{(\alpha_i - \alpha_k)^{n-1}} - \frac{f(\alpha_j)}{(\alpha_j - \alpha_k)^{n-1}} - \sum_{l=0}^{n-2} \frac{1}{l!} f^{(l)}(\alpha_k) \left[\frac{1}{(\alpha_i - \alpha_k)^{n-1-l}} - \frac{1}{(\alpha_j - \alpha_k)^{n-1-l}} \right] \right\}, \quad (1.9d)$$

$$f_{ijk}^{1,2,n-2} = \frac{1}{(\alpha_i - \alpha_j)^2} \left\{ \frac{f(\alpha_i)}{(\alpha_i - \alpha_k)^{n-2}} - \frac{f(\alpha_j)}{(\alpha_j - \alpha_k)^{n-2}} \left[1 + (n-2) \frac{(\alpha_j - \alpha_i)}{(\alpha_j - \alpha_k)} \right] + \frac{(\alpha_j - \alpha_i)}{(\alpha_j - \alpha_k)^{n-2}} f'(\alpha_j) - \sum_{l=0}^{n-3} \frac{1}{l!} f^{(l)}(\alpha_k) \left[\frac{1}{(\alpha_i - \alpha_k)^{n-2-l}} - \frac{1}{(\alpha_j - \alpha_k)^{n-2-l}} \right] - \frac{1}{(\alpha_j - \alpha_k)^{n-2-l}} \left[1 + (n-l-2) \frac{(\alpha_j - \alpha_i)}{(\alpha_j - \alpha_k)} \right] \right\}. \quad (1.9e)$$

These are sufficient to determine all derivatives up to and including the fourth. Thus, the gradient tensor ($n = 1$) requires the two expressions $f_{ijk}^{0,0,2}$ and $f_{ijk}^{0,1,1}$ evident from (1.4); the second derivative ($n = 2$) requires $f_{ijk}^{0,0,3}$, $f_{ijk}^{0,1,2}$ and $f_{ijk}^{1,1,1}$, which may be read off from (1.5); the third derivative ($n = 3$) involves four distinct formulas for $f_{ijk}^{0,0,4}$, $f_{ijk}^{0,1,3}$, $f_{ijk}^{0,2,2}$ and $f_{ijk}^{1,1,2}$; and the fourth derivative ($n = 4$) requires the five expressions $f_{ijk}^{0,0,5}$, $f_{ijk}^{0,1,4}$, $f_{ijk}^{0,2,3}$, $f_{ijk}^{1,1,3}$ and $f_{ijk}^{1,2,2}$. Note that (1.9b) and (1.9c) reduce to (1.9a) in the limit as $\alpha_i \rightarrow \alpha_k$. Similarly, (1.9d) and (1.9e) reduce to (1.9c) in the limit.

The main objective of this paper is to prove Theorem 1.1. Some new results concerning the properties of the fourth order gradient tensor ∇f are also presented and the inverse fourth order tensor $\nabla^{-1} f$ is introduced. Both the gradient and its inverse are discussed with application to strain measure functions [6, 12]. The proof of Theorem 1.1 begins with a new derivation of the well known expression for the gradient ∇f . The essential structure of the second and higher order derivatives is shown to depend on a general algebraic identity. This identity also reveals the appearance of the characteristic finite difference terms. The proof is completed by making connections with contour integrals and with interpolation polynomials. The results are presented in terms of Kronecker products of tensors which makes the expressions more transparent.

The paper is laid out as follows. Notation is introduced in Section 2 followed by the derivation of Theorem 1.1 in Section 3. The inverse gradient tensor is introduced and its properties discussed in Section 4.

2. Notation and preliminaries. We consider second order tensors acting on vectors in a three dimensional inner product space, $\mathbf{x} \rightarrow \mathbf{A}\mathbf{x}$ with transpose $\mathbf{A}^t$ such that $\mathbf{y} \cdot \mathbf{A}\mathbf{x} = \mathbf{x} \cdot \mathbf{A}^t \mathbf{y}$. Spaces of symmetric and skew-symmetric tensors are distinguished, $\text{Lin} = \text{Sym} \oplus \text{Skw}$ where $\mathbf{A} \in \text{Sym}$ (Skw) iff $\mathbf{A}^t = \mathbf{A}$ ($\mathbf{A}^t = -\mathbf{A}$). Products $\mathbf{A}\mathbf{B} \in \text{Lin}$ are defined by $\mathbf{y} \cdot \mathbf{A}\mathbf{B}\mathbf{x} = (\mathbf{A}^t \mathbf{y}) \cdot \mathbf{B}\mathbf{x}$.

Psym is the space of positive definite second order tensors. Functions of a symmetric tensor can be phrased in terms of its spectral decomposition (1.1) where $\mathbf{A}_i \in \text{Psym}$ and the distinct eigenvalues α_i , $i = 1 \dots, d \leq 3$ are real numbers. The single function $f(\mathbf{A})$ is a special case of isotropic tensor functions of the form $\mathbf{T}(\mathbf{A}) = \sum_{i=1}^3 f_i(\alpha_1, \alpha_2, \alpha_3) \mathbf{A}_i$, involving three functions of three variables. Chadwick and Ogden [2] derived first and second derivatives for more general situation (see [3] for a recent overview). Our interest here is strictly limited to one function f as in (1.1).

$\mathbb{Lin}$ is the space of fourth order tensors acting on Lin . The square tensor product or Kronecker product $\mathbf{A} \boxtimes \mathbf{B}$, $\text{Lin} \times \text{Lin} \rightarrow \mathbb{Lin}$, is defined in the usual manner as [11]

$$(\mathbf{A} \boxtimes \mathbf{B})\mathbf{X} = \mathbf{A}\mathbf{X}\mathbf{B}^t, \quad \forall \mathbf{X} \in \text{Lin}. \quad (2.1)$$

The following property of $\boxtimes$ will be used extensively,

$$(\mathbf{A} \boxtimes \mathbf{B})(\mathbf{X} \boxtimes \mathbf{Y}) = (\mathbf{A}\mathbf{X}) \boxtimes (\mathbf{B}\mathbf{Y}). \quad (2.2)$$

The generalization of (2.1) is

$$\underbrace{(\mathbf{A} \boxtimes \mathbf{B} \dots \boxtimes \mathbf{C})}_{n+1} : \underbrace{\mathbf{X}\mathbf{Y} \dots \mathbf{Z}}_n = \mathbf{A}\mathbf{X}\mathbf{B}^t \mathbf{Y} \dots \mathbf{Z}\mathbf{C}^t. \quad (2.3)$$

Basic properties of the Kronecker product can be found in e.g. [9].

The gradient of a tensor function $f(\mathbf{A})$ is a fourth order tensor $\nabla f \in \mathbb{Lin}$ defined by

$$\nabla f(\mathbf{A}) \mathbf{X} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [f(\mathbf{A} + \epsilon \mathbf{X}) - f(\mathbf{A})]. \quad (2.4)$$

Higher order derivatives are defined recursively in accordance with (1.3),

$$\nabla^{(n)} f(\mathbf{A}) \mathbf{X} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\nabla^{(n-1)} f(\mathbf{A} + \epsilon \mathbf{X}) - \nabla^{(n-1)} f(\mathbf{A})], \quad n \geq 2. \quad (2.5)$$

We note some properties of the tensor gradient. First, the tensor gradient of the product of two functions is

$$\nabla[f(\mathbf{A})g(\mathbf{A})] = (\mathbf{I} \boxtimes g(\mathbf{A})) \nabla f(\mathbf{A}) + (f(\mathbf{A}) \boxtimes \mathbf{I}) \nabla g(\mathbf{A}). \quad (2.6)$$

For example, let $g = 1/f \equiv f^{-1}$, then using $\nabla[f(\mathbf{A})g(\mathbf{A})] = 0$ yields

$$\nabla f^{-1}(\mathbf{A}) = -[f^{-1}(\mathbf{A}) \boxtimes f^{-1}(\mathbf{A})] \nabla f(\mathbf{A}). \quad (2.7)$$

The composition of two functions is $f \circ g(\mathbf{A}) = f(g(\mathbf{A}))$ has gradient

$$\nabla f \circ g(\mathbf{A}) = \nabla f(g(\mathbf{A})) \nabla g(\mathbf{A}). \quad (2.8)$$

Only symmetric tensor arguments are considered in this paper. The extension of the gradient tensor to non-symmetric tensor arguments is discussed by Itskov [7].

3. Proof of Theorem 1.1. The proof has several stages, starting with consideration of f in the form of a power series. This allows us to identify certain properties of the derivatives that are subsequently generalized to arbitrary f . Implicit in this approach is the idea that a function of a single variable can be approximated to any desired level of precision by a polynomial. The use of polynomials allows us to see the structure of the higher order derivatives, not only the first or second.

3.1. *A power series.* For the moment assume

$$f(x) = x^m, \quad (3.1)$$

then $f(\mathbf{A} + \mathbf{X})$ is

$$\begin{aligned} (\mathbf{A} + \mathbf{X})^m &= (\mathbf{A} + \mathbf{X})(\mathbf{A} + \mathbf{X})(\dots)(\mathbf{A} + \mathbf{X}) \\ &= \mathbf{A}^m + (\mathbf{A}^{m-1}\mathbf{X} + \mathbf{A}^{m-2}\mathbf{X}\mathbf{A} + \dots + \mathbf{X}\mathbf{A}^{m-1}) \\ &\quad + (\mathbf{A}^{m-2}\mathbf{X}^2 + \mathbf{A}^{m-3}\mathbf{X}\mathbf{A}\mathbf{X} + \dots + \mathbf{X}^2\mathbf{A}^{m-2}) + \dots \end{aligned} \quad (3.2)$$

The $\mathcal{O}(\mathbf{X})$ and $\mathcal{O}(\mathbf{X}^2)$ contributions contain m and $m(m-1)/2$ separate terms, respectively. The term of $\mathcal{O}(\mathbf{X}^k)$ has $\binom{m}{k}$ elements. A more suggestive and useful form of (3.2) is apparent by rewriting it as

$$\begin{aligned} (\mathbf{A} + \mathbf{X})^m - \mathbf{A}^m &= \sum_{k=1}^m \mathbf{A}^{m-k} \mathbf{X} \mathbf{A}^{k-1} + \sum_{k=1}^{m-1} \mathbf{A}^{m-1-k} \mathbf{X} \sum_{l=1}^k \mathbf{A}^{l-1} \mathbf{X} \mathbf{A}^{k-l} + \dots \\ &= \sum_{k=1}^m (\mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1}) \mathbf{X} + \sum_{k=1}^{m-1} \sum_{l=1}^k \mathbf{A}^{m-1-k} \boxtimes \mathbf{A}^{l-1} \boxtimes \mathbf{A}^{k-l} : \mathbf{X} \mathbf{X} \\ &\quad + \sum_{\substack{i_1, i_2, \dots, i_{n+1}=0 \\ i_1 + \dots + i_{n+1} = m-n}} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} x_{n+1}^{i_{n+1}} \mathbf{A}^{i_1} \boxtimes \mathbf{A}^{i_2} \boxtimes \dots \boxtimes \mathbf{A}^{i_n} \boxtimes \mathbf{A}^{i_{n+1}} \\ &\quad : \underbrace{\mathbf{X} \mathbf{X} \dots \mathbf{X}}_n + \dots \end{aligned} \quad (3.3)$$

This permits us to consider the derivatives in sequence. Our objective is to understand the term of $\mathcal{O}(\mathbf{X}^n)$, but it is easier and more instructive to begin with the lowest, $n = 1$.

3.2. *The first derivative.* Let us first focus on the $\mathcal{O}(\mathbf{X})$ term. Consider the identity

$$\begin{aligned} (\mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}) \sum_{k=1}^m \mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1} &= \sum_{k=1}^m (\mathbf{A}^{m+1-k} \boxtimes \mathbf{A}^{k-1} - \mathbf{A}^{m-k} \boxtimes \mathbf{A}^k) \\ &= \mathbf{A}^m \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}^m. \end{aligned} \quad (3.4)$$

Then

$$(\mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}) [(\mathbf{A} + \mathbf{X})^m - \mathbf{A}^m] = (\mathbf{A}^m \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}^m) \mathbf{X} + \mathcal{O}(\mathbf{X}^2), \quad (3.5)$$

which may be generalized to power series functions

$$f(x) = \sum_{m=0}^M c_m x^m, \quad (3.6)$$

as

$$(\mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}) [f(\mathbf{A} + \mathbf{X}) - f(\mathbf{A})] = [f(\mathbf{A}) \boxtimes \mathbf{I} - \mathbf{I} \boxtimes f(\mathbf{A})] \mathbf{X} + \mathcal{O}(\mathbf{X}^2). \quad (3.7)$$

Equation (3.7) implies

LEMMA 3.1. The tensor gradient of any differentiable function satisfies

$$(\mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}) \nabla f(\mathbf{A}) = f(\mathbf{A}) \boxtimes \mathbf{I} - \mathbf{I} \boxtimes f(\mathbf{A}). \quad (3.8)$$

We will return to the form of this equation later when the inverse gradient tensor is discussed. At this stage we note an isomorphism between the tensorial identity (3.4) and an algebraic equation in two variables. Let x and y represent the first and second occurrence of $\mathbf{A}$ in $\mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1} \rightarrow x^{m-k} y^{k-1}$. Then eq. (3.4) is equivalent to

$$\begin{aligned} (x-y) \sum_{k=1}^m x^{m-k} y^{k-1} &= (x-y)(x^{m-1} + x^{m-2}y + x^{m-3}y^2 + \dots + y^{m-1}) \\ &= x^m - y^m. \end{aligned} \quad (3.9)$$

This is the well known factorization of $x^m - y^m$. The analogous relation for the second derivative of f is derived next.

3.3. *The second derivative.* The algebraic equality analogous to (3.9) for the $\mathcal{O}(\mathbf{X}^2)$ terms involves three variables, say x , y and z , and is

$$(x-y)(y-z)(z-x) \sum_{k=1}^{m-1} \sum_{l=1}^k x^{m-1-k} y^{l-1} z^{k-l} = (z-y)x^m + (x-z)y^m + (y-x)z^m. \quad (3.10)$$

As proof of this statement note that the tensorial equivalent of (3.10) is

$$\begin{aligned} &(\mathbf{A} \boxtimes \mathbf{I} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A} \boxtimes \mathbf{I})(\mathbf{I} \boxtimes \mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{I} \boxtimes \mathbf{A})(\mathbf{I} \boxtimes \mathbf{I} \boxtimes \mathbf{A} - \mathbf{A} \boxtimes \mathbf{I} \boxtimes \mathbf{I}) \frac{1}{2} \nabla^{(2)} f(\mathbf{A}) \\ &= f(\mathbf{A}) \boxtimes (\mathbf{I} \boxtimes \mathbf{A} - \mathbf{A} \boxtimes \mathbf{I}) + (\mathbf{A} \boxtimes f(\mathbf{A}) \boxtimes \mathbf{I} - \mathbf{I} \boxtimes f(\mathbf{A}) \boxtimes \mathbf{A}) + (\mathbf{I} \boxtimes \mathbf{A} - \mathbf{A} \boxtimes \mathbf{I}) \boxtimes f(\mathbf{A}). \end{aligned} \quad (3.11)$$

It may be checked, using (3.3), that this indeed holds.

The algebraic identities for the $n = 1$ and $n = 2$ derivatives are, respectively,

$$\sum_{k=1}^m x^{m-k} y^{k-1} = \frac{x^m}{x-y} + \frac{y^m}{y-x}, \quad (3.12a)$$

$$\sum_{k=1}^{m-1} \sum_{l=1}^k x^{m-1-k} y^{l-1} z^{k-l} = \frac{x^m}{(x-y)(x-z)} + \frac{y^m}{(y-x)(y-z)} + \frac{z^m}{(z-x)(z-y)}. \quad (3.12b)$$

This form of the identity will prove more useful in the general case of the n^{th} derivative, considered next.

3.4. *Higher derivatives.* Referring to eq. (3.3) it is clear that the term of $O(\mathbf{X}^n)$ involves $\binom{m}{k}$ products of the form

$$\mathbf{A}^{i_1} \mathbf{X} \mathbf{A}^{i_2} \mathbf{X} \dots \mathbf{X} \mathbf{A}^{i_n} \mathbf{X} \mathbf{A}^{i_{n+1}}, \quad \text{with } i_1 + i_2 + \dots + i_{n+1} = m - n. \quad (3.13)$$

The analogous algebraic quantity has a closed form expression similar to that found for the first and second derivatives in (3.12). It is given by the following general identity among $n + 1$ independent variables:

LEMMA 3.2.

$$\sum_{\substack{i_1, i_2, \dots, i_{n+1}=0 \\ i_1 + i_2 + \dots + i_{n+1} = m - n}}^{m-n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} x_{n+1}^{i_{n+1}} = \sum_{i=1}^{n+1} \frac{x_i^m}{\prod_{\substack{j=1 \\ j \neq i}}^{n+1} (x_i - x_j)}. \quad (3.14)$$

As proof², consider the identity

$$\frac{1}{(1 - x_1 t)(1 - x_2 t) \dots (1 - x_{n+1} t)} = \frac{1}{t^n} \sum_{i=1}^{n+1} \frac{1}{(1 - x_i t)} \frac{1}{\prod_{\substack{j=1 \\ j \neq i}}^{n+1} (x_i - x_j)}. \quad (3.15)$$

The right hand side is simply the expansion of the left in partial fractions. Equation (3.14) follows by comparing the coefficients of t^{m-n} on either side of (3.15).

Motivated by the expansion (3.3), consider the following ansatz for the derivatives,

$$\nabla f(\mathbf{A}) = \sum_{i,j=1}^d f_{ij}^{(1)} \mathbf{A}_i \boxtimes \mathbf{A}_j, \quad (3.16a)$$

$$\frac{1}{2} \nabla^{(2)} f(\mathbf{A}) = \sum_{i,j,k=1}^d f_{ijk}^{(2)} \mathbf{A}_i \boxtimes \mathbf{A}_j \boxtimes \mathbf{A}_k, \quad (3.16b)$$

$$\frac{1}{n!} \nabla^{(n)} f(\mathbf{A}) = \sum_{\substack{i_1, i_2, \dots, \\ \dots, i_n, i_{n+1}=1}}^d f_{i_1, i_2, \dots, i_{n+1}}^{(n)} \mathbf{A}_{i_1} \boxtimes \mathbf{A}_{i_2} \boxtimes \dots \boxtimes \mathbf{A}_{i_{n+1}}. \quad (3.16c)$$

The key feature is the use of the spectral basis of $\mathbf{A}$.

The tensorial equations (3.8) and (3.11), or their algebraic versions (3.9) and (3.10), imply

$$(\alpha_i - \alpha_j) f_{ij}^{(1)} = f(\alpha_i) - f(\alpha_j), \quad (3.17a)$$

$$(\alpha_i - \alpha_j)(\alpha_j - \alpha_k)(\alpha_k - \alpha_i) f_{ijk}^{(2)} = (\alpha_k - \alpha_j) f(\alpha_i) + (\alpha_i - \alpha_k) f(\alpha_j) + (\alpha_j - \alpha_i) f(\alpha_k), \quad (3.17b)$$

²Thanks to Doron Zeilberger for the proof.

These may be expressed in the obvious alternate form,

$$f_{ij}^{(1)} = \frac{f(\alpha_i) - f(\alpha_j)}{\alpha_i - \alpha_j}, \quad i \neq j, \quad (3.18a)$$

$$f_{ijk}^{(2)} = \frac{f(\alpha_i)}{(\alpha_i - \alpha_j)(\alpha_i - \alpha_k)} + \frac{f(\alpha_j)}{(\alpha_j - \alpha_k)(\alpha_j - \alpha_i)} + \frac{f(\alpha_k)}{(\alpha_k - \alpha_i)(\alpha_k - \alpha_j)}, \quad i \neq j \neq k \neq i. \quad (3.18b)$$

These relations do not, however, provide all the coefficients required to define the derivatives according to (3.16a) and (3.16b). Specifically, the coefficients $f_{ii}^{(1)}$, $f_{ikk}^{(2)}$ and $f_{kkk}^{(2)}$ are not defined, but may be obtained from (3.17) by taking the limits $\alpha_j \rightarrow \alpha_i$ etc. and using l'Hopital's rule. The result is $f_{ii}^{(1)} = f'(\alpha_i)$ in agreement with (1.4), and

$$f_{iik}^{(2)} = \begin{cases} \frac{f'(\alpha_i)}{\alpha_i - \alpha_k} - \frac{f(\alpha_i) - f(\alpha_k)}{(\alpha_i - \alpha_k)^2}, & i \neq k, \\ \frac{1}{2}f''(\alpha_i), & i = k, \end{cases} \quad (3.19)$$

in agreement with Theorem 1.1 and with (1.5) from [10].

How does this generalize to the n^{th} derivative, and is it correct to take the limits as described to find the “missing” coefficients? In partial answer to the first question, note that Lemma 3.2 applied to the function f represented as a power series, (3.6), suggests the equivalence

$$\sum_{i=1}^{n+1} \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^{n+1} (x_i - x_j)} \rightarrow f_{i_1, i_2, \dots, i_{n+1}}^{(n)}, \quad n \geq 3, \quad (3.20)$$

where the x_i and x_j are replaced by α_i and α_j . However, no matter what combination of indices are taken for the coefficients on the right hand side, there will be pairs of indices with the same values, e.g. $f_{ijkk}^{(3)}$. In this sense the cases $n = 1$ and $n = 2$ are special because it is possible to have indices that are distinct, which enables us to write the identities (3.18). This does not apply for $n \geq 3$. The indicial overlap must be taken into account in order to arrive at the correct form of (3.20). Not surprisingly, the correct result is equivalent to taking the appropriate limit of the left member of (3.20). Consider the third order derivative, for instance. The coefficient $f_{ijkk}^{(3)}$ corresponds to $x_1 = \alpha_i$, $x_2 = \alpha_j$, $x_3 = \alpha_k$ in the limit as $x_4 \rightarrow \alpha_k$. That is,

$$\begin{aligned} f_{ijkk}^{(3)} &= \lim_{x \rightarrow \alpha_k} \left[\frac{f(\alpha_i)}{(\alpha_i - \alpha_j)(\alpha_i - \alpha_k)(\alpha_i - x)} + \frac{f(\alpha_j)}{(\alpha_j - \alpha_k)(\alpha_j - x)(\alpha_j - \alpha_i)} \right. \\ &\quad \left. + \frac{f(\alpha_k)}{(\alpha_k - \alpha_i)(\alpha_k - \alpha_j)(\alpha_k - x)} + \frac{f(x)}{(x - \alpha_i)(x - \alpha_j)(x - \alpha_k)} \right] \\ &= \frac{1}{(\alpha_i - \alpha_j)} \left[\frac{f(\alpha_i) - f(\alpha_k)}{(\alpha_i - \alpha_k)^2} - \frac{f(\alpha_j) - f(\alpha_k)}{(\alpha_j - \alpha_k)^2} \right] + \frac{f'(\alpha_k)}{(\alpha_i - \alpha_k)(\alpha_j - \alpha_k)}, \end{aligned} \quad (3.21)$$

in agreement with (1.9d) for $n = 3$. The remaining coefficients $f_{ikkk}^{(3)}$, $f_{iikk}^{(3)}$ and $f_{kkkk}^{(3)}$ can be obtained by taking further limits. This method of evaluating the coefficients, while correct, is particularly tedious as n becomes larger and more limits must be taken.

Let us therefore consider the form of $f_{i_1, i_2, \dots, i_{n+1}}^{(n)}$. First, it is clear that the coefficient is unchanged under any permutation of the indices $i_1, i_2, \dots, i_{n+1}$. In other words, it is

totally symmetric in the $n + 1$ indices. Secondly, the indices assume only three different values: 1, 2 and 3. Since there are $n + 1$ indices then apart from the first and second derivatives $f_{ij}^{(1)}$ ($n = 1$) and $f_{ijk}^{(1)}$ ($n = 2$) there will always be coincident indices present in the coefficient $f_{i_1, i_2, \dots, i_{n+1}}^{(n)}$. These are indices with the same value whether that value is 1, 2 or 3. For $n \geq 3$ there are at least $n - 1$ coincident indices. Suppose the index k , distinct from i and j , occurs $\nu_k > 1$ times in $\{i_1, i_2, \dots, i_{n+1}\}$. Similarly, the indices i and j each occur $\nu_i > 0$ and $\nu_j > 0$ times. This suggests the change of notation

$$f_{i_1, i_2, \dots, i_{n+1}}^{(n)} \rightarrow f_{ijk}^{\nu_i, \nu_j, \nu_k}, \quad \text{with } \nu_i + \nu_j + \nu_k = n + 1. \quad (3.22)$$

We are now ready to describe two approaches for calculating these coefficients.

3.5. *A closed form expression.* We first note that the coefficient $f_{i_1, i_2, \dots, i_{n+1}}^{(n)}$ of (3.16c) can be evaluated as a contour integral. This follows from (3.20) and the integral identity

$$\sum_{i=1}^{n+1} \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^{n+1} (x_i - x_j)} = \frac{1}{2\pi i} \int_C \frac{dz f(z)}{\prod_{l=1}^{n+1} (z - x_l)}, \quad (3.23)$$

where C is any contour in the complex plane that encloses the set of points $\{x_1, x_2, \dots, x_{n+1}\}$. The integrand in (3.23) must be modified to account for multiple occurrences of the indices, in the notation of (3.22). We obtain the integral identity

$$f_{ijk}^{\nu_i, \nu_j, \nu_k} = \frac{1}{2\pi i} \int_C \frac{dz f(z)}{(z - \alpha_i)^{\nu_i} (z - \alpha_j)^{\nu_j} (z - \alpha_k)^{\nu_k}}, \quad (3.24)$$

which can be evaluated by residues, with the final result (1.7).

3.6. *Interpolation interpretation.* The following result is key:

LEMMA 3.3. If $\{x_1, x_2, \dots, x_{n+1}\}$ are $n + 1$ distinct points then

$$\sum_{i=1}^{n+1} \frac{f(x_i)}{\prod_{\substack{k=1 \\ k \neq i}}^{n+1} (x_i - x_k)} = \frac{1}{n!} \frac{d^n}{dx^n} P(x), \quad (3.25)$$

where $P(x)$ is the Lagrange interpolating polynomial of the $n + 1$ function values $f(x_i)$, $i = 1, 2, \dots, n + 1$.

The proof follows from the explicit form of the interpolating polynomial of degree n ,

$$P(x) = \sum_{i=1}^{n+1} f(x_i) L_i(x), \quad (3.26)$$

where $L_i(x)$, $i = 1, 2, \dots, n + 1$, are the Lagrange polynomials satisfying $L_i(x_j) = \delta_{ij}$,

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n+1} (x - x_j) \bigg/ \prod_{\substack{k=1 \\ k \neq i}}^{n+1} (x_i - x_k). \quad (3.27)$$

Differentiating $P(x)$ of (3.26) yields the identity (3.25).

This connection between the coefficients and polynomial interpolation provides an alternative means to calculate the coefficient $f_{ijk}^{\nu_i, \nu_j, \nu_k}$. The interpolation problem that must be solved is not the straightforward one of Lemma 3.3 with a single datum at each of $n + 1$ distinct points. Some of the points can have more than one piece of data. Specifically, associated with $x = \alpha_l$ where $l = i, j$ and k , will be the ν_l data comprising $f(\alpha_l), f'(\alpha_l), \dots, f^{(\nu_l-1)}(\alpha_l)$. In this way there is a well defined and unique interpolating polynomial of degree x^n for each $f_{ijk}^{\nu_i, \nu_j, \nu_k}$ with

$$f_{ijk}^{\nu_i, \nu_j, \nu_k} = \frac{1}{n!} \frac{d^n}{dx^n} P(x). \quad (3.28)$$

The interpolation problem that must be solved does not fit into the standard category of Lagrange interpolating polynomials, since some or many of the points are coincident. However, the procedure for generating the polynomial is not difficult. Appendix A illustrates the method by example, for the case of f known at three points along with all derivatives at one of these points up to $f^{(n-2)}$.

3.7. The number of independent coefficients. The final detail is a count of the number $N(n)$ of independent expressions of the form $f_{ijk}^{\nu_i, \nu_j, \nu_k}$ required for the n^{th} derivative. This is the number of elements in the set of integer triples $\{I, J, K \in \mathbb{Z} : 0 \leq I \leq J \leq K, I + J + K = n + 1\}$. That is,

$$N(n) = \dim \{I, J \in \mathbb{Z} : 0 \leq I \leq J \leq n + 1 - I - J\}. \quad (3.29)$$

This defines a region $A(n)$ on the integer grid of $\{I, J\}$ bounded by the three straight lines $L_1 : I = 0$, $L_2 : I - J = 0$ and $L_3 : I + 2J = n + 1$. $N(n)$ is the number of points in $A(n)$, including on its boundary. Enumeration yields

$$N(n) = \lfloor \frac{(n+4)^2 + 4}{12} \rfloor \quad (3.30)$$

where the floor function $\lfloor x \rfloor$ is the integral part of x .

This completes the proof of Theorem 1.1. Note that the two methods discussed for calculating the coefficients, based on the contour integral and the interpolation problem, are both related to divided differences [4].

4. The inverse gradient.

4.1. Definition. The inverse tensor function $\nabla^{-1}f(\mathbf{A}) \in \mathbb{L}\text{in}$ is defined by $(\nabla f)\nabla^{-1}f = (\nabla^{-1}f)\nabla f = \mathbb{I}$. As an example application let f^{inv} be the inverse of f , such that $f \circ f^{\text{inv}}(\mathbf{A}) = f^{\text{inv}} \circ f(\mathbf{A}) = \mathbf{A}$. It follows from (2.8) that

$$\nabla f^{\text{inv}}(\mathbf{A}) = \nabla^{-1}f(f^{\text{inv}}(\mathbf{A})). \quad (4.1)$$

It is clear from (1.4) that the definition of $\nabla^{-1}f(\mathbf{A})$ is problematic if $f'(\alpha_i)$ vanishes. This possibility is precluded by restricting consideration to strictly monotonic *strain measure functions* acting on positive definite tensors.

4.2. *Application to strain measure functions.* The function f is a *strain measure* [6, 12] if it is a smooth function $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ which satisfies

$$f(1) = 0, \quad f'(1) = 1, \quad f' > 0. \quad (4.2)$$

For the remainder of the paper we restrict attention to strain measure functions. Examples will be presented for the Seth-Hill strain measure functions,

$$f^{[m]}(x) = m^{-1}(x^m - 1). \quad (4.3)$$

The inverse f^{inv} is well defined since strain measure functions are one-to-one. Hence, the inverse gradient is relevant to calculating the gradient of the inverse function, via the identity (2.8).

Strain measure functions act on positive definite strain tensors $\mathbf{S} \in \text{Psym}$, with

$$\mathbf{S} = \sum_{i=1}^d \lambda_i \mathbf{S}_i, \quad \mathbf{I} = \sum_{i=1}^d \mathbf{S}_i, \quad \lambda_i > 0, \quad \mathbf{S}_i \mathbf{S}_j = \mathbf{S}_j \mathbf{S}_i = \delta_{ij} \mathbf{S}_i, \quad (4.4)$$

where $d \leq 3$ is now the eigen-index of $\mathbf{S}$.

Accordingly, we extend the concept of positive definiteness to fourth order tensors. It is necessary to define an inner product on Lin , which is done in the usual manner as $\mathbf{A} \cdot \mathbf{B} = \text{tr}(\mathbf{A}\mathbf{B}^t)$. Then Lin is the space of fourth order tensors acting on Lin , $\mathbf{X} \rightarrow \mathbb{A}\mathbf{X}$ with transpose $\mathbb{A}^t$ such that $\mathbf{Y} \cdot \mathbb{A}\mathbf{X} = \mathbf{X} \cdot \mathbb{A}^t \mathbf{Y}$ for all $\mathbf{X}, \mathbf{Y} \in \text{Lin}$. The vector space may be decomposed $\text{Lin} = \text{Sym} \oplus \text{Skw}$ where Sym and Skw denote the spaces of symmetric ($\mathbb{A}^t = \mathbb{A}$) and skew-symmetric ($\mathbb{A}^t = -\mathbb{A}$) tensors, respectively. Any $\mathbb{A} \in \text{Lin}$ can be uniquely partitioned into symmetric and skew parts: $\mathbb{A} = \mathbb{A}^{(+)} + \mathbb{A}^{(-)}$, where $\mathbb{A}^{(\pm)} = (\mathbb{A} \pm \mathbb{A}^t)/2$. The identity $\mathbb{I}$ satisfies $\mathbb{I}\mathbf{X} = \mathbf{X}$ for all $\mathbf{X} \in \text{Lin}$. The product $\mathbb{A}\mathbb{B} \in \text{Lin}$ is defined by $\mathbf{Y} \cdot \mathbb{A}\mathbb{B}\mathbf{X} = (\mathbb{A}^t \mathbf{Y}) \cdot \mathbb{B}\mathbf{X}$. Psym is the space of positive definite fourth order tensors: $\mathbb{A} \in \text{Psym}$ iff $\mathbf{X} \cdot \mathbb{A}\mathbf{X} > 0$, for all nonzero $\mathbf{X} \in \text{Sym}$.

The spectral form of $\mathbb{A} \in \text{Psym}$ is

$$\mathbb{A} = \sum_{I=1}^D a_I \mathbb{A}_I, \quad \mathbb{I} = \sum_{I=1}^D \mathbb{A}_I, \quad \mathbb{A}_I \mathbb{A}_J = \begin{cases} \mathbb{A}_I & I = J, \\ 0, & I \neq J, \end{cases} \quad (4.5)$$

where D is the eigen-index, $a_I > 0$ and $\mathbb{A}_I \in \text{Psym}$. Functions of fourth order tensors can therefore be defined according to

$$f(\mathbb{A}) = \sum_{I=1}^D f(a_I) \mathbb{A}_I. \quad (4.6)$$

For any $\mathbf{A} \in \text{Psym}$ and its eigen-tensors $\mathbf{A}_i$, $i = 1, \dots, d$ of (1.2), define the associated set of $D = \frac{1}{2}d(d+1)$ basis tensors for Sym by

$$\mathbb{A}_I = \begin{cases} \mathbf{A}_I \boxtimes \mathbf{A}_I, & I = 1, \dots, d, \\ \mathbf{A}_i \boxtimes \mathbf{A}_j + \mathbf{A}_j \boxtimes \mathbf{A}_i, & I = d+1, \dots, D = \frac{1}{2}d(d+1). \end{cases} \quad (4.7)$$

Here $I = d+1, \dots, D$ corresponds to distinct pairs (i, j) with $i < j$. Thus, $D = 6$ for $d = 3$, while $D = 3$ for $d = 2$ and $D = 1$ in the trivial case of $d = 1$. It may be readily

checked that $\mathbb{A}_I \mathbb{A}_J = \mathbb{A}_I \delta_{IJ}$. The identity $\mathbb{I} = \mathbf{I} \boxtimes \mathbf{I}$ implies the partition of unity

$$\mathbb{I} = \sum_{I=1}^D \mathbb{A}_I.$$

In particular,

LEMMA 4.1. The gradient of a strain measure function and its inverse are positive definite fourth order tensors, i.e. $\nabla f(\mathbf{A}), \nabla^{-1} f(\mathbf{A}) \in \mathbb{P}\text{sym}$, and

$$(\nabla f) \nabla^{-1} f = (\nabla^{-1} f) \nabla f = \mathbb{I}. \quad (4.8)$$

The spectral forms are

$$\nabla f(\mathbf{A}) = \sum_{I=1}^D f_I \mathbb{A}_I, \quad \nabla^{-1} f(\mathbf{A}) = \sum_{I=1}^D f_I^{-1} \mathbb{A}_I, \quad (4.9a)$$

with

$$f_I(\mathbf{A}) = \begin{cases} f'(\alpha_I), & I = 1, \dots, d, \\ \frac{f(\alpha_i) - f(\alpha_j)}{\alpha_i - \alpha_j}, & I = d+1, \dots, D. \end{cases} \quad (4.9b)$$

The proof is evident from the spectral decompositions. The positivity of f_I , $I = 1, \dots, D$ is a consequence of $\alpha_i > 0$, $i = 1, \dots, d$ combined with the monotonicity of f and the mean value theorem. The positive definite nature of $\nabla f(\mathbf{A})$ implies that it has a unique positive definite inverse.

The gradient function and its inverse has an interesting alternative representation for integer and fractional values of m , respectively:

LEMMA 4.2. For integer values of $m \neq 0$,

$$\nabla f^{[m]}(\mathbf{A}) = \frac{1}{|m|} \begin{cases} \sum_{k=1}^m \mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1}, & m > 0, \\ \sum_{k=m+1}^0 \mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1}, & m < 0. \end{cases} \quad (4.10a)$$

$$\nabla^{-1} f^{(\frac{1}{m})}(\mathbf{A}) = \frac{1}{|m|} \begin{cases} \sum_{k=1}^m \mathbf{A}^{1-\frac{k}{m}} \boxtimes \mathbf{A}^{\frac{k}{m}-\frac{1}{m}}, & m > 0, \\ \sum_{k=m+1}^0 \mathbf{A}^{1-\frac{k}{m}} \boxtimes \mathbf{A}^{\frac{k}{m}-\frac{1}{m}}, & m < 0. \end{cases} \quad (4.10b)$$

The proof uses the expansion of eq. (4.9) with $f_I \rightarrow f_I^{[m]}$ where

$$f_I^{[m]}(\mathbf{A}) = \begin{cases} \alpha_I^{m-1}, & I = 1, \dots, d \\ \frac{1}{m} \left(\frac{\alpha_i^m - \alpha_j^m}{\alpha_i - \alpha_j} \right), & I = d+1, \dots, D. \end{cases} \quad (4.11)$$

Consider the case of positive m . Then using the spectral form for $\mathbf{A}$

$$\begin{aligned}
\frac{1}{m} \sum_{k=1}^m \mathbf{A}^{m-k} \boxtimes \mathbf{A}^{k-1} &= \frac{1}{m} \sum_{i,j=1}^d \mathbf{A}_i \boxtimes \mathbf{A}_j \sum_{k=1}^m \alpha_i^{m-k} \alpha_j^{k-1} \\
&= \sum_{i=1}^d \alpha_i^{m-1} \mathbf{A}_i \boxtimes \mathbf{A}_i + \frac{1}{m} \sum_{\substack{i,j=1 \\ i \neq j}}^d \mathbf{A}_i \boxtimes \mathbf{A}_j \alpha_i^m \alpha_j^{-1} \sum_{k=1}^m \left(\frac{\alpha_j}{\alpha_i}\right)^k \\
&= \sum_{I=1}^d f_I^{[m]} \mathbb{A}_I + \frac{1}{m} \sum_{\substack{i,j=1 \\ i \neq j}}^d \mathbf{A}_i \boxtimes \mathbf{A}_j \left(\frac{\alpha_i^m - \alpha_j^m}{\alpha_i - \alpha_j}\right) \\
&= \nabla f^{[m]}(\mathbf{A}),
\end{aligned} \tag{4.12}$$

where the identity $x + x^2 + \dots + x^m = x(1 - x^m)/(1 - x)$ has been used. The proof for negative m is similar. The results for $\nabla^{-1} f^{(\frac{1}{m})}(\mathbf{A})$ are a consequence of the identity

$$f_I^{(\frac{1}{m})}(\mathbf{A}) = 1/f_I^{[m]}(\mathbf{A}^{\frac{1}{m}}), \quad m \neq 0. \tag{4.13}$$

Some examples of interest, first for integer values of m :

$$\nabla f^{(1)}(\mathbf{A}) = \mathbf{I} \boxtimes \mathbf{I} = \mathbb{I}, \tag{4.14a}$$

$$\nabla f^{(-1)}(\mathbf{A}) = \mathbf{A}^{-1} \boxtimes \mathbf{A}^{-1}, \tag{4.14b}$$

$$\nabla f^{(2)}(\mathbf{A}) = \frac{1}{2}(\mathbf{I} \boxtimes \mathbf{A} + \mathbf{A} \boxtimes \mathbf{I}), \tag{4.14c}$$

$$\nabla f^{(-2)}(\mathbf{A}) = \frac{1}{2}(\mathbf{A}^{-2} \boxtimes \mathbf{A}^{-1} + \mathbf{A}^{-1} \boxtimes \mathbf{A}^{-2}). \tag{4.14d}$$

The inverse tensors

$$\nabla^{-1} f^{(1)}(\mathbf{A}) = \mathbb{I}, \tag{4.15a}$$

$$\nabla^{-1} f^{(-1)}(\mathbf{A}) = \mathbf{A} \boxtimes \mathbf{A}, \tag{4.15b}$$

follow by observation, as do the general relations

$$\nabla f^{(-m)}(\mathbf{A}) = \mathbf{A}^{-m} \boxtimes \mathbf{A}^{-m} \nabla f^{[m]}(\mathbf{A}), \tag{4.16a}$$

$$\nabla^{-1} f^{(-m)}(\mathbf{A}) = \mathbf{A}^m \boxtimes \mathbf{A}^m \nabla^{-1} f^{[m]}(\mathbf{A}), \tag{4.16b}$$

which express tensors in terms of their counterparts of opposite sign.

$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A} = \mathbf{C}$ is an important and common equation in mechanics [13, 11]. The solution may be written simply and succinctly as $\mathbf{X} = (\mathbf{A} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \mathbf{A})^{-1} \mathbf{C}$ where [8] (eq. (22))

$$(\mathbf{A} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \mathbf{A})^{-1} = \sum_{i,j=1}^d (\alpha_i + \alpha_j)^{-1} \mathbf{A}_i \boxtimes \mathbf{A}_j. \tag{4.17}$$

In the current notation this may be expressed $\mathbf{X} = \frac{1}{2} \nabla^{-1} f^{(2)}(\mathbf{A}) \mathbf{C}$, see eq. (4.14c). In fact, the identity (4.10a) implies that the solution to

$$\mathbf{A}^{m-1} \mathbf{X} + \mathbf{A}^{m-2} \mathbf{X} \mathbf{A} + \mathbf{A}^{m-3} \mathbf{X} \mathbf{A}^2 + \dots + \mathbf{X} \mathbf{A}^{m-1} = \mathbf{C}, \tag{4.18}$$

is

$$\mathbf{X} = \frac{1}{m} \nabla^{-1} f^{[m]}(\mathbf{A}) \mathbf{C} = \sum_{i,j=1}^d \left(\frac{\alpha_i - \alpha_j}{\alpha_i^m - \alpha_j^m} \right) \mathbf{A}_i \boxtimes \mathbf{A}_j \mathbf{C}. \quad (4.19)$$

The solution to $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A} = \mathbf{C}$ can be written in a form that does not require the spectral split, e.g. $\mathbf{X} = [2(I_1 I_2 - I_3)]^{-1} [(I_1^2 - I_2)\mathbf{C} + \mathbf{A}\mathbf{C}\mathbf{A} - (\mathbf{A}^2\mathbf{C} + \mathbf{C}\mathbf{A}^2) + I_1 I_3 \mathbf{A}^{-1} \mathbf{C} \mathbf{A}^{-1}]$, [11] (eq. (27)) where I_1, I_2, I_3 are the invariants of $\mathbf{A}$. Similar expressions, although far more complicated, can be generated for the solution to (4.18) and related equations, but we leave such matters aside.

Examples of fractional powers include the well known case of $\frac{1}{2}$ [11], for which (4.10b) gives

$$\nabla^{-1} f^{(\frac{1}{2})}(\mathbf{A}) = \frac{1}{2} (\sqrt{\mathbf{A}} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \sqrt{\mathbf{A}}) \Rightarrow \nabla \sqrt{\mathbf{A}} = (\sqrt{\mathbf{A}} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \sqrt{\mathbf{A}})^{-1}. \quad (4.20)$$

Similarly,

$$\nabla^{-1} f^{(-\frac{1}{2})}(\mathbf{A}) = \frac{1}{2} (\sqrt{\mathbf{A}} \boxtimes \mathbf{A} + \mathbf{A} \boxtimes \sqrt{\mathbf{A}}) \Rightarrow \nabla \mathbf{A}^{-1/2} = -(\sqrt{\mathbf{A}} \boxtimes \mathbf{A} + \mathbf{A} \boxtimes \sqrt{\mathbf{A}})^{-1}, \quad (4.21)$$

and

$$\nabla^{-1} f^{(\frac{1}{3})}(\mathbf{A}) = \frac{1}{3} (\mathbf{A}^{2/3} \boxtimes \mathbf{I} + \mathbf{A}^{1/3} \boxtimes \mathbf{A}^{1/3} + \mathbf{I} \boxtimes \mathbf{A}^{2/3}), \quad (4.22a)$$

$$\nabla^{-1} f^{(-\frac{1}{3})}(\mathbf{A}) = \frac{1}{3} (\mathbf{A}^{1/3} \boxtimes \mathbf{A} + \mathbf{A}^{2/3} \boxtimes \mathbf{A}^{2/3} + \mathbf{A} \boxtimes \mathbf{A}^{1/3}), \quad (4.22b)$$

imply expressions for $\nabla \mathbf{A}^{\pm 1/3}$, etc.

Note that the case of $m = 0$, corresponding to

$$\nabla f^{(0)}(\mathbf{A}) = \nabla \ln(\mathbf{A}), \quad (4.23)$$

is specifically excluded in Lemma 3.3. Although there is no expression for $\nabla \ln(\mathbf{A})$ analogous to (4.10a), its inverse can be represented as follows:

LEMMA 4.3. The tensor $\nabla^{-1} \ln(\mathbf{A}) = \nabla^{-1} f^{(0)}(\mathbf{A})$, which is the inverse of $\nabla \ln(\mathbf{A})$, is

$$\nabla^{-1} \ln(\mathbf{A}) = \int_0^1 dx \mathbf{A}^x \boxtimes \mathbf{A}^{1-x}, \quad \text{for any } \mathbf{A} \in \text{Psym}. \quad (4.24)$$

The proof follows from (4.10b) by taking the limit of $m \rightarrow \infty$. An alternative and direct proof can be obtained using the spectral form of $\mathbf{A}$,

$$\begin{aligned} \int_0^1 dx \mathbf{A}^x \boxtimes \mathbf{A}^{1-x} &= \sum_{i,j=1}^d \mathbf{A}_i \boxtimes \mathbf{A}_j \int_0^1 dx \alpha_i^x \alpha_j^{1-x} \\ &= \sum_{i,j=1}^d \frac{\alpha_i - \alpha_j}{\ln \alpha_i - \ln \alpha_j} \mathbf{A}_i \boxtimes \mathbf{A}_j, \end{aligned} \quad (4.25)$$

where the ratio becomes α_i if $i = j$. The right member of (4.25) is obviously the inverse of $\nabla \ln(\mathbf{A})$ from eq. (1.4).

4.3. *The tensors $\mathbb{J}$, $\mathbb{J}^*$, $\mathbb{K}$, and the solution to $AX - XA = Y$.* Consider the equation

$$\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{A} = \mathbf{Y}, \quad \mathbf{A} \in \text{Psym}, \quad (4.26)$$

for the unknown $\mathbf{X}$ in terms of $\mathbf{Y}$ which is either symmetric or skew, and $\mathbf{X}$ is of the opposite parity [5]. The equation can be written

$$\mathbb{J}(\mathbf{A})\mathbf{X} = \mathbf{Y}, \quad (4.27)$$

where

$$\mathbb{J}(\mathbf{A}) \equiv \mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A}, \quad \mathbf{A} \in \text{Psym}. \quad (4.28)$$

We will only consider $\mathbb{J}(\mathbf{A})$ for symmetric $\mathbf{A}$, implying $\mathbb{J} \in \text{Sym}$ and $\mathbb{J}$ maps $\text{Sym} \rightarrow \text{Skw}$ and $\text{Skw} \rightarrow \text{Sym}$. Therefore, $\mathbb{J}$ does not possess eigenvalues, eigenvectors or an inverse in the usual sense. However, it is possible to define the pseudo-inverse, or equivalently the Moore-Penrose inverse, $\mathbb{J}^*$, which satisfies

$$\begin{aligned} \mathbb{J}\mathbb{J}^*\mathbb{J} &= \mathbb{J}, \\ \mathbb{J}^*\mathbb{J}\mathbb{J}^* &= \mathbb{J}^*. \end{aligned} \quad (4.29)$$

Further understanding comes from consideration of the spectral form

$$\mathbb{J}(\mathbf{A}) = \sum_{i,j=1}^d (\alpha_i - \alpha_j) \mathbf{A}_i \boxtimes \mathbf{A}_j. \quad (4.30)$$

This representation implies that $\mathbf{A}_i$ are null vectors, i.e. $\mathbb{J}(\mathbf{A})\mathbf{A}_i = 0$, $i = 1, \dots, d$. Noting that

$$\mathbb{J}(\mathbf{A})[\mathbf{A}_i \boxtimes \mathbf{A}_j \pm \mathbf{A}_j \boxtimes \mathbf{A}_i] = (\alpha_i - \alpha_j)[\mathbf{A}_i \boxtimes \mathbf{A}_j \mp \mathbf{A}_j \boxtimes \mathbf{A}_i], \quad (4.31)$$

gives

$$\mathbb{J}^*(\mathbf{A})[\mathbf{A}_i \boxtimes \mathbf{A}_j \pm \mathbf{A}_j \boxtimes \mathbf{A}_i] = (\alpha_i - \alpha_j)^{-1}[\mathbf{A}_i \boxtimes \mathbf{A}_j \mp \mathbf{A}_j \boxtimes \mathbf{A}_i], \quad \text{for } i \neq j \text{ only.} \quad (4.32)$$

The caveat $i \neq j$ is crucial, and indicates that the non-null parts of $\mathbb{J}$ and its pseudo-inverse define maps between three dimensional subspaces of Sym and Skw . In particular,

$$\mathbb{J}^*(\mathbf{A}) = (\mathbf{A} \boxtimes \mathbf{I} - \mathbf{I} \boxtimes \mathbf{A})^* = \sum_{\substack{i,j=1 \\ i \neq j}}^d (\alpha_i - \alpha_j)^{-1} \mathbf{A}_i \boxtimes \mathbf{A}_j. \quad (4.33)$$

This formula for the pseudo-inverse clearly satisfies (4.29). It is interesting to compare the form of (4.33) with (4.17).

Note that

$$\mathbb{J}^*\mathbb{J} = \mathbb{J}\mathbb{J}^* = \mathbb{I} - \sum_{I=1}^d \mathbb{A}_I = \sum_{I=d+1}^D \mathbb{A}_I. \quad (4.34)$$

Define $\mathbb{K}(\mathbf{A})$, $\mathbf{A} \in \text{Psym}$, and its pseudoinverse,

$$\mathbb{K}(\mathbf{A}) = \sum_{i=1}^d \alpha_i \mathbb{A}_i, \quad \mathbb{K}^*(\mathbf{A}) = \sum_{i=1}^d \alpha_i^{-1} \mathbb{A}_i = \mathbb{K}(\mathbf{A}^{-1}), \quad (4.35)$$

then

$$\mathbb{J}(\mathbf{A})\mathbb{J}^*(\mathbf{A}) + \mathbb{K}(\mathbf{A})\mathbb{K}^*(\mathbf{A}) = \mathbb{I} \quad \text{for any } \mathbf{A} \in \text{Psym}. \quad (4.36)$$

The tensors $\mathbb{J}$ and $\mathbb{K}$ allow us to write $\nabla f(\mathbf{A})$ and its inverse in a succinct manner as, respectively,

$$\nabla f(\mathbf{A}) = \mathbb{K}(f'(\mathbf{A})) + \mathbb{J}^*(\mathbf{A})\mathbb{J}(f(\mathbf{A})), \quad (4.37a)$$

$$\nabla^{-1}f(\mathbf{A}) = \mathbb{K}^*(f'(\mathbf{A})) + \mathbb{J}(\mathbf{A})\mathbb{J}^*(f(\mathbf{A})). \quad (4.37b)$$

Appendix.

A. Solving the interpolation problems. Consider three points, say x_1, x_2 and x_3 . Given $f(x_1)$, $f(x_2)$ and $f^{(l)}(x_3)$, $l = 0, 1, 2, \dots, n-2$ find the interpolating polynomial $P(x)$ of degree n . For simplicity, but no lack of generality, take $x_3 = 0$, and consider the ansatz

$$P(x) = \sum_{l=0}^n p_l x^l, \quad p_l = \frac{1}{l!} f^{(l)}(0), \quad l = 1, 2, \dots, n-2. \quad (A.1)$$

This satisfies all $n-1$ conditions at $x = 0$, and the conditions at x_1 and x_2 are met if $P(x_1) = f(x_1)$ and $P(x_2) = f(x_2)$. These imply a pair of linear equations in p_{n-1} and p_n which are easily solved. Only the coefficient p_n is required as this determines the n^{th} derivative of $P(x)$,

$$p_n = \frac{1}{x_1 - x_2} \left[\frac{f(x_1)}{x_1^{n-1}} - \frac{f(x_2)}{x_2^{n-1}} - \sum_{l=0}^{n-2} \frac{f^{(l)}(0)}{l!} \left(\frac{x_1^l}{x_1^{n-1}} - \frac{x_2^l}{x_2^{n-1}} \right) \right]. \quad (A.2)$$

This implies eq. (1.9d).

Consider the case where the values of the function and its first $n-3$ derivatives are given at one point, say $x = 0$, along with $f(x_1)$, $f(x_2)$ and $f'(x_2)$. Using the same ansatz as (A.1) with the three unknowns p_{n-2} , p_{n-1} and p_n , implies three simultaneous equations. The solution for p_n is

$$p_n = \frac{1}{(x_1 - x_2)^2} \left[\frac{f(x_1)}{x_1^{n-2}} - \frac{f(x_2)}{x_2^{n-2}} [1 - (n-2)(\frac{x_1}{x_2} - 1)] + \frac{f'(x_2)}{x_2^{n-2}} (x_2 - x_1) - \sum_{l=0}^{n-3} \frac{f^{(l)}(0)}{l!} \left(\frac{x_1^l}{x_1^{n-2}} - \frac{x_2^l}{x_2^{n-2}} [1 - (n-2-l)(\frac{x_1}{x_2} - 1)] \right) \right]. \quad (A.3)$$

This yields the expression (1.9e).

Acknowledgment. Thanks to Doron Zeilberger for being true to his motto: “Who you gonna call?”.

REFERENCES

- [1] D. E. Carlson and A. Hoger, *The derivative of a tensor-valued function of a tensor*, Q. Appl. Math. **44** (1986), 409–423.
- [2] P. Chadwick and R. W. Ogden, *A theorem of tensor calculus and its application to isotropic elasticity*, Arch. Rat. Mech. Anal. **44** (1971), no. 1, 54–68.
- [3] Y-C. Chen and G-S. Dui, *The derivative of isotropic tensor functions, elastic moduli and stress rate: I. Eigenvalue formulation*, Math. Mech. Solids **9** (2004), no. 5, 493–511.
- [4] W. F. Donoghue, Jr., *Monotone matrix functions and analytic continuation*, Springer-Verlag, New York, 1974.

- [5] Guan-Suo Dui, *Some basis-free formulae for the time rate and conjugate stress of logarithmic strain tensor*, J. Elasticity **83** (2006), no. 2, 113–151.
- [6] R. Hill, *Aspects of invariance in solid mechanics*, Adv. Appl. Mech. **18** (1978), 1–75.
- [7] M. Itskov and N. Aksel, *A closed-form representation for the derivative of non-symmetric tensor power series*, Int. J. Solids Struct. **39** (2002), no. 24, 5963–5978.
- [8] C. S. Jog, *Derivatives of the stretch, rotation and exponential tensors in n -dimensional vector spaces*, J. Elasticity **82** (2006), no. 2, 175–192.
- [9] O. Kintzel and Y. Bascédilar, *Fourth-order tensors - tensor differentiation with applications to continuum mechanics. I: Classical tensor analysis*, Z. Angew. Math. Mech. **86** (2006), no. 4, 291–311.
- [10] R. W. Ogden, *Non-linear elastic deformations*, Ellis Horwood, 1984.
- [11] L. Rosati, *A novel approach to the solution of the tensor equation $AX+XA=H$* , Int. J. Solids Struct. **37** (2000), no. 25, 3457–3477.
- [12] M. Scheidler, *Time rates of generalized strain tensors. Part I: Component formulas*, Mech. Materials **11** (1991), 199–210.
- [13] ———, *The tensor equation $AX+XA=F(A,H)$, with applications to kinematics of continua*, J. Elasticity **36** (1994), no. 2, 117–153.
- [14] H. Xiao, O. T. Bruhns, and A. Meyers, *Strain rates and material spins*, J. Elasticity **52** (1998), no. 1, 1–41.