TENSOR CALCULUS

LONG CHEN

ABSTRACT. We give rules on tensor calculation.

CONTENTS

1. TENSORS

What is a tensor? While a matrix can be considered a 2nd order tensor, a 2nd order tensor is essentially an equivalent class of matrices. Mathematically speaking, a tensor is a multi-linear map, whereas a matrix represents a bilinear map in a specific coordinate system. Different coordinates yield different representations, but the underlying mapping remains unchanged.

1.1. **Definition.** Let V be an n-dimensional vector space. The linear functionals from V to $\mathbb{R}$ form another vector space, denoted V^*, called the dual space of V. When dealing with finite-dimensional vector spaces, we can identify V with V^* and with $\mathbb{R}^n$ after choosing a basis. Therefore for $v^* \in V^*$ and $v \in V$, the function $v^*(v)$ can be also denoted by $v^* \cdot v$ or (v^*, v) using the inner product of vectors. However, it is important to note that elements in V and V^* are distinct entities. To clarify this distinction, we use column vectors to represent elements in V and row vectors to represent elements in V^*.

Date: March 25, 2024.

Let us move to two vector spaces (V, W) and a bilinear function $f(\cdot,\cdot) : (V, W) \to \mathbb{R}$. The set of all such bilinear functions forms a vector space. To represent this space, we can choose bases for V and W, allowing us to identify a bilinear function as a matrix.

More formally, we introduce the tensor product of dual spaces, denoted $V^* \otimes W^*$. For any $v^* \in V^*$ and $w^* \in W^*$, we define $v^* \otimes w^*$ as a bilinear map $V \times W \to \mathbb{R}$ such that:

$$(v^* \otimes w^*)(v, w) := v^*(v) \cdot w^*(w) = (v^*, v) \cdot (w^*, w).$$

By identifying V as the dual of V^*, i.e., $V = (V^*)^*$, we can also define the tensor product $V \otimes W$.

Suppose $\{e^i, 1 \le i \le m\}$ is a basis of V^* and $\{e^j, 1 \le j \le n\}$ is a basis of W^*. Then the set $\{e^i \otimes e^j\}$ forms a basis of $V^* \otimes W^*$. For example, given a matrix $A = (a_{ij})_{m\times n}$, it can be expanded as $\sum_{ij} a_{ij} e^i \otimes e^j$, where $e^i = (0, \dots, 1, \dots, 0)$ represents the classical orthonormal basis of $\mathbb{R}^n$. It is crucial to note that the same tensor $v^* \otimes w^*$ may have various matrix representations, depending on the choice of bases.

Given two linear spaces V_1, V_2 and assume they are subspaces of a larger linear space V. We have the following operations of these two spaces

- $V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$;
- $V_1 \oplus V_2 = V_1 + V_2$ when $V_1 \cap V_2 = \{0\}$;
- $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$;
- $V_1 \otimes V_2 = \{v_1 \otimes v_2 : v_1 \in V_1, v_2 \in V_2\}$.

The tensor product $v_1 \otimes v_2$ is a bilinear mapping on the dual space $V_1^* \times V_2^*$. A natural product topology can be defined for $V_1 \times V_2$ component-wise.

The relation of dimensions are

- $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) \le \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2)$;
- $\dim(V_1 \otimes V_2) = \dim(V_1) \times \dim(V_2)$.

We emphasize that the sum $V_1 + V_2$ may not enlarge the space. For example, when $V_1 \subset V_2$ (a line on a plane), $V_1 + V_2 = V_2$.

How about trilinear and in general multilinear functions? Matrix contains only two directions. We need a new object, *tensor*, for multilinear functions. The tensor product of two vector space can be extended to accommodate multilinear functions of various orders.

Consider a (r, s) type tensor space, denoted V^r_s, defined as:

$$V^r_s := \underbrace{V \otimes \ldots \otimes V}_{r} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{s}.$$

An element in V^r_s takes the form:

$$x = x^{i_1,\dots,i_r}_{k_1,\dots,k_s} e_{i_1} \otimes \ldots \otimes e_{i_r} \otimes e^{k_1} \otimes \ldots \otimes e^{k_s}, \tag{1}$$

where a basis (e_i) is given for V and (e^i) for V^*. Tensors in this space represent multilinear functions, and their representations may differ under changes in bases.

Now, let us explore how these representations are related when the basis is changed. Suppose we have another basis $(\hat{e}_i)$ and a transformation matrix (α_i^j). This implies:

$$\hat{e}_i = \alpha_i^j e_j, \quad \hat{e}^i = \beta^i_j e^j, \quad (\alpha_i^j)(\beta_i^j) = I,$$

where Einstein summation convention is used: when an index appears as both a subscript and a superscript in an expression, it implies summation over all possible values for that index.

For a (r, s) type tensor

$$
\begin{aligned}
x &= \hat{x}^{i_1,\dots,i_r}_{k_1,\dots,k_s}\hat{e}_{i_1}\otimes\dots\otimes\hat{e}_{i_r}\otimes\hat{e}^{k_1}\otimes\dots\otimes\hat{e}^{k_s}\\
&= \hat{x}^{i_1,\dots,i_r}_{k_1,\dots,k_s}\alpha^{j_1}_{i_1}\dots\alpha^{j_r}_{i_r}\beta^{k_1}_{l_1}\dots\beta^{k_s}_{l_s}e_{j_1}\otimes\dots\otimes e_{j_r}\otimes e^{l_1}\otimes\dots\otimes e^{l_s}\\
&= x^{j_1,\dots,j_r}_{l_1,\dots,l_s}e_{j_1}\otimes\dots\otimes e_{j_r}\otimes e^{l_1}\otimes\dots\otimes e^{l_s}.
\end{aligned}
$$

Therefore

$$x^{j_1,\dots,j_r}_{l_1,\dots,l_s} = \hat{x}^{i_1,\dots,i_r}_{k_1,\dots,k_s}\alpha^{j_1}_{i_1}\dots\alpha^{j_r}_{i_r}\beta^{k_1}_{l_1}\dots\beta^{k_s}_{l_s}$$

or equivalently

$$\hat{x}^{j_1,\dots,j_r}_{l_1,\dots,l_s} = x^{i_1,\dots,i_r}_{k_1,\dots,k_s}\beta^{j_1}_{i_1}\dots\beta^{j_r}_{i_r}\alpha^{k_1}_{l_1}\dots\alpha^{k_s}_{l_s}$$

Two types of tensors play an important role: symmetric and skew-symmetric. Riemannian metric is symmetric and positive definite 2-tensor. Stress and strain in linear elasticity are symmetric tensors. Differential forms are skew-symmetric tensors.

1.2. Symmetric tensor.

We consider the strain ε and stress $\boldsymbol{\sigma}$ tensor used in linear elasticity. Both are 2nd order symmetric tensors.

Two different Cartesian coordinate systems can be related to each other by a rigid motion. Since the choice of origin for these systems does not influence the definition of $\boldsymbol{\sigma}$, we assume that the new coordinate system is derived by rotating the original one around its origin, that is, $\hat{\boldsymbol{x}} = \boldsymbol{Q}\boldsymbol{x}$, where $\boldsymbol{Q}$ is a unitary matrix.

We can also view $\boldsymbol{\sigma}$ as a linear mapping of vectors: $\boldsymbol{\sigma} : \boldsymbol{n} \to \boldsymbol{t} = \boldsymbol{\sigma}\boldsymbol{n}$. By applying $\boldsymbol{Q}$ to both sides, we obtain

$$\hat{\boldsymbol{t}} = \boldsymbol{Q}\boldsymbol{t} = \boldsymbol{Q}\boldsymbol{\sigma}\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{n} = \boldsymbol{Q}\boldsymbol{\sigma}\boldsymbol{Q}^{\intercal}\hat{\boldsymbol{n}},$$

which implies

$$\hat{\boldsymbol{\sigma}} = \boldsymbol{Q}\boldsymbol{\sigma}\boldsymbol{Q}^{\intercal}. \tag{2}$$

Hence, we can view the stress as a linear mapping between linear spaces (of vectors). The stress matrix is simply one representation of this mapping in a particular coordinate system. Different coordinates result in different representations related by (2), but the linear mapping remains consistent.

For a given linear operator T, an eigenvector v and its corresponding eigenvalue λ are defined as $Tv = \lambda v$. By the nature of eigenvalues, which depend only on the linear structure and not on the representation, the eigenvalues of $\boldsymbol{\sigma}$, and their combinations, e.g., $\operatorname{tr}(\boldsymbol{\sigma}) = \lambda_1 + \lambda_2 + \lambda_3$, $\det(\boldsymbol{\sigma}) = \lambda_1\lambda_2\lambda_3$, and $\operatorname{tr}(\boldsymbol{\varepsilon}) = \operatorname{div}\boldsymbol{u}$, remain invariant with the change of coordinates.

1.3. Differential forms.

2. Change of Coordinates

Recall that $\Phi : \Omega_R \to \Phi(\Omega_R)$ denotes the configuration mapping, where Ω_R represents the reference domain. The body force $\boldsymbol{f}$ is a vector function composed of 3-forms. Thus,

$$\int_V \boldsymbol{f}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{V_R} \boldsymbol{f}_R(\boldsymbol{x}_R)\,\mathrm{d}\boldsymbol{x}_R,$$

and, given $\mathrm{d}\boldsymbol{x} = \det(D\Phi)\,\mathrm{d}\boldsymbol{x}_R$, its transformation obeys

$$\boldsymbol{f}(\boldsymbol{x}) = \det(D\Phi)^{-1}\boldsymbol{f}_R(\boldsymbol{x}_R).$$

3. STRUCTURE OF THE MATRIX SPACE

Let $\mathbb{M}$ be the linear space of $d \times d$ matrices. The symmetric subspace is denoted by $\mathbb{S}$ and the anti-symmetric one by $\mathbb{K}$.

3.1. Matrix-vector and matrix-matrix products.

For a matrix $\boldsymbol{A}$, we can express it as a stack of row vectors $\boldsymbol{a}_i$ or column vectors $\boldsymbol{a}^j$, where $i, j = 1, \dots, d$. As an example, for $d = 3$, we have:

$$\boldsymbol{A} = \begin{pmatrix} - & \boldsymbol{a}_1 & - \\ - & \boldsymbol{a}_2 & - \\ - & \boldsymbol{a}_3 & - \end{pmatrix} = \begin{pmatrix} | & | & | \\ \boldsymbol{a}^1 & \boldsymbol{a}^2 & \boldsymbol{a}^3 \\ | & | & | \end{pmatrix}.$$

We define the Frobenius inner product in $\mathbb{M}$ as

$$(\boldsymbol{A}, \boldsymbol{B})_F = \boldsymbol{A} : \boldsymbol{B} = \boldsymbol{A} \circ \boldsymbol{B} := \sum_{ij} a_{ij} b_{ij} = \sum_i \boldsymbol{a}_i \cdot \boldsymbol{b}_i = \sum_j \boldsymbol{a}^j \cdot \boldsymbol{b}^j,$$

which leads to the Frobenius norm $\|\cdot\|_F$ of a matrix. Here, $\circ$ denotes the Hadamard product (the entry-wise product), and in MATLAB, this operation is `.*`. We extend this to the sum of the cross product of column vectors as

$$\boldsymbol{A}(\cdot\times)\boldsymbol{B} = \sum_{i=1}^{d} \boldsymbol{a}^i \times \boldsymbol{b}^i.$$

The matrix-vector product $\boldsymbol{A}\boldsymbol{b}$ can be interpreted as the linear combination of column vectors $\sum_i b_i \boldsymbol{a}^i$ or as the inner product of $\boldsymbol{b}$ with the row vectors $\boldsymbol{a}_i$ of $\boldsymbol{A}$. We define

$$\boldsymbol{A} \cdot \boldsymbol{b} := \boldsymbol{A}\boldsymbol{b} = \sum_i b_i \boldsymbol{a}^i = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{b} \\ \boldsymbol{a}_2 \cdot \boldsymbol{b} \\ \boldsymbol{a}_3 \cdot \boldsymbol{b} \end{pmatrix},$$

and

$$\boldsymbol{A} \times \boldsymbol{b} := \begin{pmatrix} \boldsymbol{a}_1 \times \boldsymbol{b} \\ \boldsymbol{a}_2 \times \boldsymbol{b} \\ \boldsymbol{a}_3 \times \boldsymbol{b} \end{pmatrix}.$$

When the vector is to the right of the matrix, the operation is defined row-wise:

$$\text{row-wise} \qquad \boldsymbol{A} \cdot \boldsymbol{b}, \quad \boldsymbol{A} \times \boldsymbol{b},$$

We also define the dot product and the cross product from the left

$$\text{column-wise} \qquad \boldsymbol{b} \cdot \boldsymbol{A}, \quad \boldsymbol{b} \times \boldsymbol{A},$$

which is applied column-wise to the matrix $\boldsymbol{A}$. Specifically,

$$\boldsymbol{b} \cdot \boldsymbol{A} = \boldsymbol{b}^{\intercal}\boldsymbol{A} = (\boldsymbol{b} \cdot \boldsymbol{a}^1 \quad \boldsymbol{b} \cdot \boldsymbol{a}^2 \quad \boldsymbol{b} \cdot \boldsymbol{a}^3)$$

and

$$\boldsymbol{b} \times \boldsymbol{A} = \begin{pmatrix} | & | & | \\ \boldsymbol{b} \times \boldsymbol{a}^1 & \boldsymbol{b} \times \boldsymbol{a}^2 & \boldsymbol{b} \times \boldsymbol{a}^3 \\ | & | & | \end{pmatrix}.$$

For clarity, we use the same notation $\boldsymbol{b}$ for both row and column vectors.

The order in which row and column products are performed does not affect the outcome, leading to the associative rule for triple products:

$$\boldsymbol{b} \times \boldsymbol{A} \times \boldsymbol{c} := (\boldsymbol{b} \times \boldsymbol{A}) \times \boldsymbol{c} = \boldsymbol{b} \times (\boldsymbol{A} \times \boldsymbol{c}).$$

Similar rules apply for $\boldsymbol{b} \cdot \boldsymbol{A} \cdot \boldsymbol{c}$ and $\boldsymbol{b} \cdot \boldsymbol{A} \times \boldsymbol{c}$, allowing us to omit parentheses without ambiguity.

Another advantage is the ability to handle the transpose of products easily. For the transpose of the product of two entities, we transpose each one, reverse their order, and add a negative sign if it involves the cross product. For instance, $(\boldsymbol{b}\times\boldsymbol{A})^{\intercal} = -\boldsymbol{A}^{\intercal}\times\boldsymbol{b}^{\intercal}$.

For two column vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, the tensor product $\boldsymbol{u}\otimes\boldsymbol{v} := \boldsymbol{u}\boldsymbol{v}^{\intercal}$ results in a matrix, also known as the dyadic product $\boldsymbol{u}\boldsymbol{v} := \boldsymbol{u}\boldsymbol{v}^{\intercal}$, using a more concise notation (omitting one $^{\intercal}$). The interaction of $\boldsymbol{u}\boldsymbol{v}$ with another vector $\boldsymbol{x}$ in row-wise and column-wise products affects the adjacent vector:

$$\boldsymbol{x}\cdot(\boldsymbol{u}\boldsymbol{v}) = (\boldsymbol{x}\cdot\boldsymbol{u})\boldsymbol{v}^{\intercal}, \quad (\boldsymbol{u}\boldsymbol{v})\cdot\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{v}\cdot\boldsymbol{x}),$$
$$\boldsymbol{x}\times(\boldsymbol{u}\boldsymbol{v}) = (\boldsymbol{x}\times\boldsymbol{u})\boldsymbol{v}, \quad (\boldsymbol{u}\boldsymbol{v})\times\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{v}\times\boldsymbol{x}).$$

3.2. **Trace.** For a $d\times d$ matrix $\boldsymbol{A}$, the trace is the sum of its diagonal entries:

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{d} a_{ii} = \sum_{i=1}^{d} \lambda_i(\boldsymbol{A}).$$

The latter equality indicates that the trace equals the sum of the eigenvalues, counted with their multiplicities. This relationship can be easily proved using the characteristic polynomial of $\boldsymbol{A}$.

Thus, if two matrices are similar, meaning $\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{C}$, then $\operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{B})$. Since $\boldsymbol{A}\boldsymbol{B}$ and $\boldsymbol{B}\boldsymbol{A}$ have the same spectrum (aside from potential differences in zero eigenvalues due to differing sizes), it follows that

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}\boldsymbol{A}).$$

From this, we deduce that the trace remains invariant under cyclic permutations of the product of matrices. For symmetric matrices, this invariance extends to any permutation of the matrices involved.

However, the trace operation does not distribute over the product of matrices in a straightforward manner, except in the case of the tensor product:

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{B}) \neq \operatorname{tr}(\boldsymbol{A})\operatorname{tr}(\boldsymbol{B}), \quad \text{but } \operatorname{tr}(\boldsymbol{A}\otimes\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A})\operatorname{tr}(\boldsymbol{B}).$$

The Frobenius inner product in $\mathbb{M}$, using the trace operator, is expressed as

$$(\boldsymbol{A},\boldsymbol{B})_F = \boldsymbol{A}:\boldsymbol{B} = \operatorname{tr}(\boldsymbol{B}^{\intercal}\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}^{\intercal}\boldsymbol{B}).$$

3.3. **An orthogonal decomposition.** Consider the exact sequence

$$0 \to \mathbb{T} \hookrightarrow \mathbb{M} \xrightarrow{\operatorname{tr}} \mathbb{R}, \tag{3}$$

where $\mathbb{T}$ denotes the subspace of traceless (or trace-free) matrices. The sequence (4) is exact because $\ker(\operatorname{tr}) = \mathbb{T}$ by definition. We define a right inverse of tr as:

$$\boldsymbol{I}/d : \mathbb{R}\to\mathbb{M}, \quad p \to p\boldsymbol{I}/d.$$

Here the scaling d^{-1} is introduced as a normalization since $(\boldsymbol{I},\boldsymbol{I})_F = d$. The map $\boldsymbol{I}/d$ serves as the right inverse of the trace operator $\operatorname{tr}:\mathbb{M}\to\mathbb{R}$, ensuring that $\operatorname{tr}\circ(\boldsymbol{I}/d) = \operatorname{tr}(\boldsymbol{I})/d = \operatorname{id}$. This leads to the Helmholtz decomposition from (4)

$$\mathbb{M} = \mathbb{T} \oplus^{\perp_F} \mathbb{R}\boldsymbol{I}, \tag{4}$$

where $\oplus^{\perp_F}$ indicates that the decomposition is orthogonal under the $(\cdot,\cdot)_F$ inner product.

By reversing the composition of tr and $\boldsymbol{I}/d$, we obtain the orthogonal projection to the subspace $\mathbb{R}\boldsymbol{I}$ in the F-product. We define $P_{\mathbb{R}}:\mathbb{M}\to\mathbb{R}\boldsymbol{I}$ through the operation $P_{\mathbb{R}} = \boldsymbol{I}/d\circ\operatorname{tr}$. The projection property is confirmed as

$$(P_{\mathbb{R}}\boldsymbol{\sigma}, p\boldsymbol{I})_F = \operatorname{tr}(\boldsymbol{\sigma})p = (\boldsymbol{\sigma}, p\boldsymbol{I})_F, \quad \forall p\in\mathbb{R}. \tag{5}$$

The orthogonal complement $(I - P_{\mathbb{R}})\boldsymbol{\sigma} \in \mathbb{T}$, representing the orthogonal projection onto $\mathbb{T}$, is referred to as the deviation of $\boldsymbol{\sigma}$ and denoted by $\boldsymbol{\sigma}^D$.

We can thus summarize this orthogonal decomposition with respect to $(\cdot,\cdot)_F$ as

$$\boldsymbol{\sigma} = (I - P_{\mathbb{R}})\boldsymbol{\sigma} + P_{\mathbb{R}}\boldsymbol{\sigma} = \boldsymbol{\sigma}^D + \operatorname{tr}(\boldsymbol{\sigma})\boldsymbol{I}_d/d.$$

In this decomposition, the volumetric stress tensor $P_{\mathbb{R}}\boldsymbol{\sigma}$ primarily affects the volume of the stressed body, whereas the stress deviator tensor $\boldsymbol{\sigma}^D$ influences its shape.

3.4. **Skew-symmetric matrices and the cross product.** When $d = 3$, we can establish an isomorphism between $\mathbb{R}^3$ and the space $\mathbb{K}$ of anti-symmetric matrices. This is achieved by defining the mapping $[\cdot]_\times : \mathbb{R}^3 \to \mathbb{K}$ as

$$[\omega]_\times = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \quad \text{for any } \omega = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \in \mathbb{R}^3.$$

Its inverse map is denoted as $\operatorname{vskw} : \mathbb{K} \to \mathbb{R}^3$, which satisfies $[\operatorname{vskw}(Z)]_\times = Z$ for $Z \in \mathbb{K}$, and $\operatorname{vskw}([\omega]_\times) = \omega$ for $\omega \in \mathbb{R}^3$. The notation $[\cdot]_\times$, sometimes denoted by mskw, is favored here to highlight its connection to the cross product of vectors. For any two vectors in $\mathbb{R}^3$, the following identity holds

$$u \times v = [u]_\times v,$$

where the right side is a standard matrix-vector product. It is also shown that

$$[u \times v]_\times = [u]_\times [v]_\times - [v]_\times [u]_\times,$$

indicating that the commutator of skew-symmetric 3×3 matrices corresponds to the cross product of vectors. Therefore, $[\cdot]_\times$ maintains the Lie algebra structure between $\mathbb{R}^3$ and $\mathbb{K}$.

In the case of $d = 2$, where $\dim \mathbb{K} = 1$, we define $[\cdot]_\times : \mathbb{R} \to \mathbb{K}$ as

$$[\omega]_\times = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

One can embed $\mathbb{R}$ into $\mathbb{R}^3$ by considering the vector $(0, 0, \omega)^\intercal$, then apply $[\cdot]_\times$, and finally, by removing the zero row and columns, truncate the resulting matrix to a 2×2 matrix.

3.5. **Another orthogonal decomposition.** The decomposition

$$\mathbb{M} = \mathbb{S} \oplus^{\perp_F} \mathbb{K}$$

constitutes an orthogonal decomposition under the $(\cdot,\cdot)_F$ inner product. For any matrix $\boldsymbol{B} \in \mathbb{M}$, this decomposition can be expressed as

$$\boldsymbol{B} = \operatorname{sym}(\boldsymbol{B}) + \operatorname{skw}(\boldsymbol{B}) := \frac{1}{2}(\boldsymbol{B} + \boldsymbol{B}^\intercal) + \frac{1}{2}(\boldsymbol{B} - \boldsymbol{B}^\intercal).$$

The skew-symmetric component can also be represented in terms of the cross product:

$$\operatorname{skw}(\boldsymbol{B}) = \frac{1}{2}[\boldsymbol{I}(\cdot\times)\boldsymbol{B}]_\times. \tag{6}$$

By contrast, the trace of a matrix is derived from the inner product with the identity matrix:

$$\operatorname{tr}(\boldsymbol{B}) = \boldsymbol{I} : \boldsymbol{B}.$$

This delineation illustrates how every matrix in $\mathbb{M}$ can be uniquely decomposed into a symmetric part, which characterizes the stretching or compression aspects, and an anti-symmetric part, which characterizes the rotational aspects of the transformation represented by the matrix.

4. FORMULAE INVOLVING DIFFERENTIAL OPERATORS

In the context of function spaces $C^1(\Omega)$, $C^1(\Omega)\otimes\mathbb{R}^d$, and $C^1(\Omega)\otimes\mathbb{M}$, we explore the interplay between differential operators and matrix operations. The Hamilton operator $\nabla = (\partial_1, \partial_2, \partial_3)^\intercal$ is regarded as a column vector, facilitating our discussion primarily in three dimensions $d = 3$. For $d = 2$, a 2-D vector $(u_1(x_1, x_2), u_2(x_1, x_2))^\intercal$ is embedded into $\mathbb{R}^3$ as $(u_1(x_1, x_2), u_2(x_1, x_2), 0)^\intercal$ to maintain consistency in representation.

4.1. **Gradient and Symmetric Gradient.** For a scalar function $v \in C^1(\Omega)$, the gradient ∇v is represented as a column vector and Dv as a row vector

$$\nabla v = \begin{pmatrix} \partial_1 v \\ \partial_2 v \\ \partial_3 v \end{pmatrix}, \quad Dv = (\nabla v)^\intercal = (\partial_1 v, \partial_2 v, \partial_3 v).$$

For a vector function $\boldsymbol{u} = (u_1, u_2, u_3)^\intercal$, standard operations such as $\operatorname{curl}\boldsymbol{u} = \nabla\times\boldsymbol{u}$ and $\operatorname{div}\boldsymbol{u} = \nabla\cdot\boldsymbol{u}$ are applied to delineate various aspects of vector fields.

The gradient of a vector $\boldsymbol{u}$ manifests as a matrix

$$D\boldsymbol{u} = (\partial_j u_i) = \begin{pmatrix} Du_1 \\ Du_2 \\ Du_3 \end{pmatrix} = \boldsymbol{u}\nabla = (\nabla\boldsymbol{u}^\intercal)^\intercal,$$

where the representation leverages the dyadic product $\boldsymbol{u}\boldsymbol{v} := \boldsymbol{u}\boldsymbol{v}^\intercal$ for clear exposition.

The symmetric gradient of a vector function $\boldsymbol{u}$ is defined as

$$\nabla^s\boldsymbol{u} = \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\intercal) = \frac{1}{2}(\boldsymbol{u}\nabla + \nabla\boldsymbol{u}),$$

utilizing the dyadic product to underscore the symmetry within the operation. Within the realm of elasticity, $\nabla^s\boldsymbol{u}$ is often referred to by $\varepsilon(\boldsymbol{u})$ or $\operatorname{def}(\boldsymbol{u})$.

Using the cross product, we arrive at the identity

$$\operatorname{skw}(D\boldsymbol{u}) = \frac{1}{2}[\nabla\times\boldsymbol{u}]_\times. \tag{7}$$

This allows us to express the decomposition for the matrix $D\boldsymbol{u}$ as

$$D\boldsymbol{u} = \nabla^s\boldsymbol{u} + \frac{1}{2}[\nabla\times\boldsymbol{u}]_\times. \tag{8}$$

In 2D, treating $(u_1(x_1, x_2), u_2(x_1, x_2))^\intercal$ as $(u_1(x_1, x_2), u_2(x_1, x_2), 0)^\intercal$, the operation $\nabla\times\boldsymbol{u}$ simplifies to

$$\operatorname{rot}\boldsymbol{u} := \partial_1 u_2 - \partial_2 u_1,$$

and the decomposition (9) adjusts to

$$D\boldsymbol{u} = \nabla^s\boldsymbol{u} + \frac{1}{2}\operatorname{rot}\boldsymbol{u}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{9}$$

Given that $\ker(\operatorname{grad}) = c$ and $\ker(D) = \boldsymbol{c} \in \mathbb{R}^3$, we establish a relation between $\ker(\nabla^s)$ and the anti-symmetric space $\mathbb{K}$

$$\ker(\nabla^s) = \{\boldsymbol{Z}\boldsymbol{x} + \boldsymbol{c},\ \boldsymbol{Z}\in\mathbb{K}, \boldsymbol{c}\in\mathbb{R}^3\} = \{\boldsymbol{\omega}\times\boldsymbol{x} + \boldsymbol{c},\ \boldsymbol{\omega}, \boldsymbol{c}\in\mathbb{R}^3\}. \tag{10}$$

A basis for $\ker(\nabla^s)$ is readily derived from a basis of $\mathbb{R}^3$ and $\dim\ker(\nabla^s) = 6$ for $d = 3$. In $d = 2$, $\mathbb{K}\cong\mathbb{R}$, and the cross product $\omega\times\boldsymbol{x}$ equates to $\omega\boldsymbol{e}_1\times(\boldsymbol{x}; 0) = \omega(-y, x)^\intercal$, signifying a 90° counter-clockwise rotation. Thus $\ker(\nabla^s)$ can be determined by $\omega \in \mathbb{R}, \boldsymbol{c}\in\mathbb{R}^2$ and $\dim\ker(\nabla^s) = 3$ for $d = 2$.

The application of differential operators div and curl to a vector function $\boldsymbol{u}$ translates to operations on the matrix function $D\boldsymbol{u}$ as follows:

$$\operatorname{div}\boldsymbol{u} = \nabla\cdot\boldsymbol{u} = (\boldsymbol{I}\cdot\nabla)\cdot\boldsymbol{u} = \operatorname{tr}(D\boldsymbol{u}) = \operatorname{tr}(\nabla^s\boldsymbol{u}),$$
$$\operatorname{curl}\boldsymbol{u} = \nabla\times\boldsymbol{u} = (\boldsymbol{I}\times\nabla)\boldsymbol{u} = [\nabla]_\times\boldsymbol{u} = \operatorname{vskw}(\operatorname{skw}(D\boldsymbol{u})).$$

The following identity

$$2\operatorname{div}\nabla^s\boldsymbol{u} = \Delta\boldsymbol{u} + \operatorname{grad}\operatorname{div}\boldsymbol{u} \tag{11}$$

can be proved as follows: for $k = 1, 2, \ldots, d$

$$(\operatorname{div}2\nabla^s\boldsymbol{u})_k = \sum_{i=1}^{d}\partial_i(\partial_i u_k + \partial_k u_i) = \Delta\boldsymbol{u}_k + \partial_k(\operatorname{div}\boldsymbol{u}).$$

The differential operators ∇^s, grad, and div interrelate as:

$$\operatorname{tr}\nabla^s(\boldsymbol{u}) = \operatorname{div}\boldsymbol{u}, \quad \operatorname{div}(p\boldsymbol{I}/d) = \operatorname{grad}p,$$

which is demonstrable through straightforward calculation and illsutrated in the following diagram.

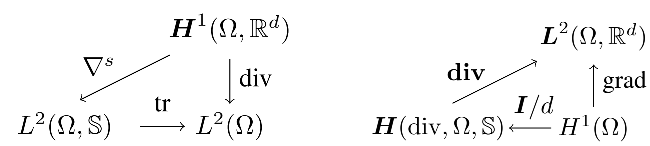

FIGURE 1. Relation of differential operators and trace operator

4.2. **Differentiation of matrix functions.** When engaging with matrix-vector operations using the Hamilton operator ∇, we encounter differentiation operations that can be classified as column-wise or row-wise. Column-wise differentiations include $\nabla\cdot\boldsymbol{A}$ and $\nabla\times\boldsymbol{A}$, while row-wise differentiations involve $\boldsymbol{A}\cdot\nabla$ and $\boldsymbol{A}\times\nabla$. To align with conventional notation, where differentiation is applied following the ∇ symbol, a more standardized representation:

$$\boldsymbol{A}\cdot\nabla := (\nabla\cdot\boldsymbol{A}^\intercal)^\intercal, \quad \boldsymbol{A}\times\nabla := -(\nabla\times\boldsymbol{A}^\intercal)^\intercal.$$

By repositioning the differential operator to the right, the notation is simplified, allowing formal application of the transpose rule for matrix-vector products.

In scholarly texts, differential operations on matrices, especially in the context of tensors, are typically applied in a row-wise manner. To distinguish this from the ∇ notation, we define the following operators in letters:

$$\operatorname{grad}\boldsymbol{u} := \boldsymbol{u}\nabla^\intercal = (\partial_j u_i) = (\nabla\boldsymbol{u})^\intercal,$$
$$\operatorname{curl}\boldsymbol{A} := -\boldsymbol{A}\times\nabla = (\nabla\times\boldsymbol{A}^\intercal)^\intercal,$$
$$\operatorname{div}\boldsymbol{A} := \boldsymbol{A}\cdot\nabla = (\nabla\cdot\boldsymbol{A}^\intercal)^\intercal.$$

It is important to recognize that the transpose operator $^\intercal$ plays a role for tensors, and notably, $\operatorname{grad}\boldsymbol{u}$ is distinct from $\nabla\boldsymbol{u}$, $\operatorname{curl}\boldsymbol{A}$ differs from both $\nabla\times\boldsymbol{A}$ and $\boldsymbol{A}\times\nabla$, and $\operatorname{div}\boldsymbol{A}$ is not equivalent to $\nabla\cdot\boldsymbol{A}$. For symmetric tensors, however, we observe that $\operatorname{div}\boldsymbol{A} = (\nabla\cdot\boldsymbol{A})^\intercal$ and $\operatorname{curl}\boldsymbol{A} = (\nabla\times\boldsymbol{A})^\intercal$.

4.3. **Integration by parts.** Integration by parts is a fundamental theorem in calculus that allows the transformation of integrals over a domain to integrals over its boundary.

The abstract form of integration by parts for tensors and vectors, assuming $L(\cdot)$ represents a linear operation, is expressed as:

$$\int_{\partial V} L(\boldsymbol{n}, \cdot)\,\mathrm{d}S = \int_V L(\nabla, \cdot)\,\mathrm{d}\boldsymbol{x}.$$

Here, the unit outward normal vector $\boldsymbol{n}$ on the boundary ∂V of a volume V is substituted by the Hamilton operator (or gradient operator) ∇. The component form of this equation,

$$\int_{\partial V} L(n_i, \cdot)\,\mathrm{d}S = \int_V L(\partial_i, \cdot)\,\mathrm{d}\boldsymbol{x},$$

proves more practical when dealing with mixed products of vectors.

To illustrate this concept with a simple example, consider verifying the following:

$$\int_{\partial V} \boldsymbol{\sigma}\boldsymbol{n}\,\mathrm{d}S = \int_V \boldsymbol{\sigma}\nabla\,\mathrm{d}\boldsymbol{x} = \int_V \operatorname{div}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{x},$$

where it is important to remember that the div operator is applied row-wise to the stress tensor $\boldsymbol{\sigma}$. This identity demonstrates how the force exerted by a stress field $\boldsymbol{\sigma}$ over the boundary ∂V of a volume V translates to the divergence of $\boldsymbol{\sigma}$ throughout the volume V. Such principles underlie the mathematical formulations of physical laws, including those governing fluid dynamics, elasticity, and electromagnetism.

4.4. **An example in linear elasticity.** This example shows how the symmetry of the stress tensor $\boldsymbol{\sigma}$ comes from the balance equations. Starting with the balance of forces and moments equations:

$$\int_V \boldsymbol{f}\,\mathrm{d}\boldsymbol{x} + \int_{\partial V} \boldsymbol{\sigma}\,\boldsymbol{n}\,\mathrm{d}S = 0 \quad \text{for all } V \subset \Omega, \tag{12}$$

$$\int_V \boldsymbol{f} \times \boldsymbol{x}\,\mathrm{d}\boldsymbol{x} + \int_{\partial V} (\boldsymbol{\sigma}\,\boldsymbol{n}) \times \boldsymbol{x}\,\mathrm{d}S = 0 \quad \text{for all } V \subset \Omega. \tag{13}$$

Using the integral by parts and let $V \to \{\boldsymbol{x}\}$, we get from (13) that $\operatorname{div}\boldsymbol{\sigma} + \boldsymbol{f} = 0$. Integration by parts implies

$$\int_{\partial V} (\boldsymbol{x} \times \boldsymbol{\sigma}) \cdot \boldsymbol{n}\,\mathrm{d}S = \int_V (\boldsymbol{x} \times \boldsymbol{\sigma}) \cdot \nabla\,\mathrm{d}\boldsymbol{x}.$$

By the product rule of differentiation,

$$\partial_i(\boldsymbol{x} \times \boldsymbol{\sigma}^i) = \boldsymbol{x} \times \partial_i\boldsymbol{\sigma}^i + \partial_i\boldsymbol{x} \times \boldsymbol{\sigma}^i.$$

As $\sum_i \partial_i\boldsymbol{\sigma}^i = \operatorname{div}\boldsymbol{\sigma}$,

$$\sum_i \boldsymbol{x} \times \partial_i\boldsymbol{\sigma}^i = \boldsymbol{x} \times (\operatorname{div}\boldsymbol{\sigma})$$

which cancel out with $\boldsymbol{x} \times \boldsymbol{f}$. For the second term, expand $\boldsymbol{x} = \sum x_i\boldsymbol{e}^i$ and thus $\partial_i\boldsymbol{x} = \boldsymbol{e}^i$. Consequently

$$\left[\sum_i \boldsymbol{e}^i \times \boldsymbol{\sigma}^i\right]_\times = [\boldsymbol{I}(\cdot\times)\boldsymbol{\sigma}]_\times = 2\operatorname{skw}(\boldsymbol{\sigma}).$$

Therefore $\sum_i \boldsymbol{e}^i \times \boldsymbol{\sigma}^i = 0$ implies $\operatorname{skw}(\boldsymbol{\sigma}) = 0$, i.e., $\boldsymbol{\sigma}$ is symmetric.

As an exercise, the reader is encouraged to prove that: for a symmetric matrix function $\boldsymbol{\sigma}$ and vector function $\boldsymbol{v}$

$$-\int_{\Omega} \operatorname{div} \boldsymbol{\sigma} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} = \int_{\Omega} \boldsymbol{\sigma} : \nabla^s \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} - \int_{\partial\Omega} (\boldsymbol{\sigma}\boldsymbol{n}) \cdot \boldsymbol{v} \, \mathrm{d}S. \tag{14}$$