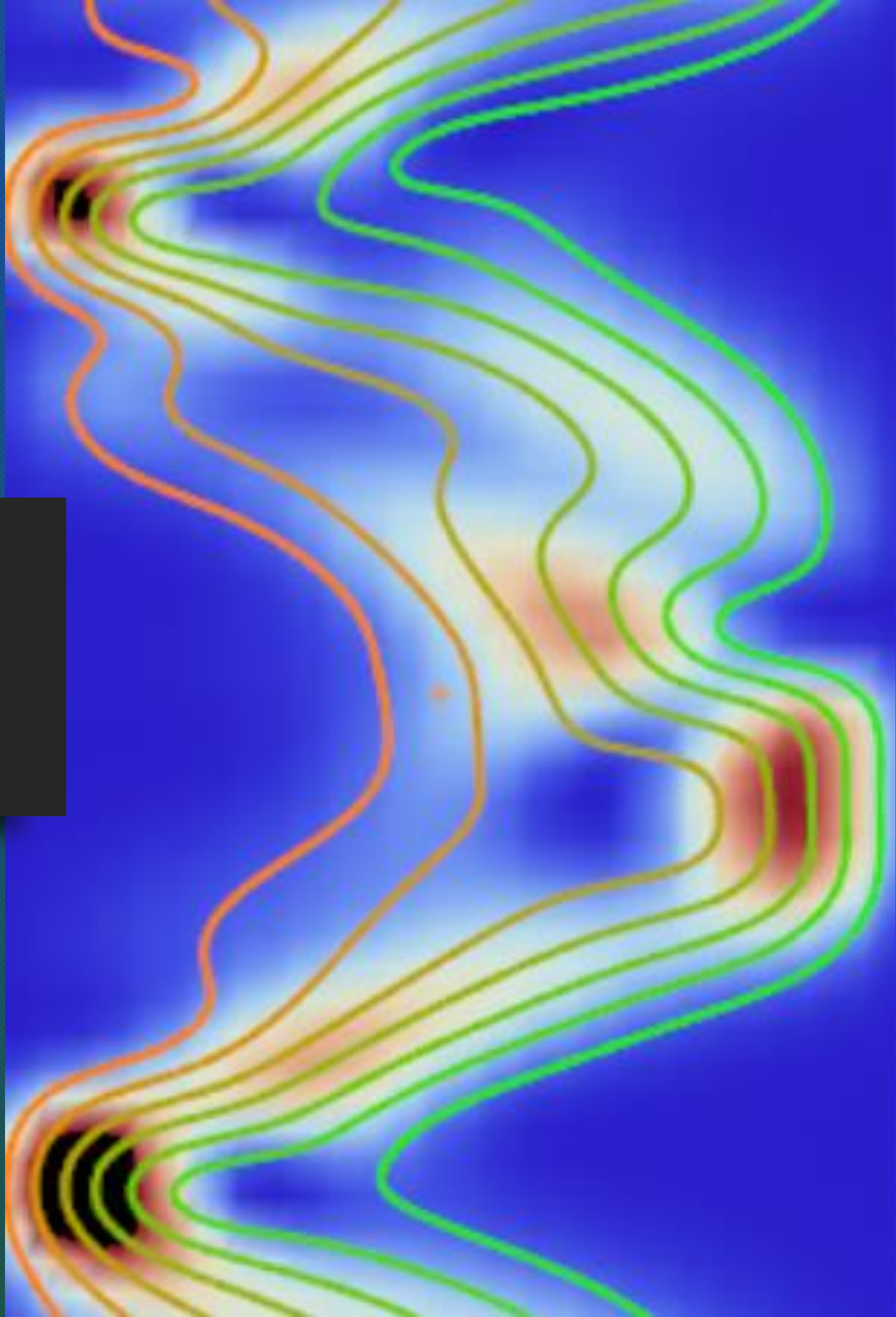
Tensor Analysis I

Differential and Integral Calculus with Tensors

MEG 324 SSG 321 Introduction to Continuum Mechanics

Instructors: OA Fakinlede & O Adewumi
www.oafak.com eds.s2pafrica.org oafak@unilag.edu.ng



- *"We do not fuss over smoothness assumptions: Functions and boundaries of regions are presumed to have continuity and differentiability properties sufficient to make meaningful underlying analysis…" Morton Gurtin, et al.*

Scope of Lecture

- The issues we shall cover in today's lecture are not hard to understand. They are fundamental to all tensor analysis.
- Be careful to note any area of difficulty. If you are specific, you can be assisted.
- We introduce the Gateaux differential as the solution to our inability to divide by tensors when we want to define a derivative

	TOPIC FOR WEEK 10
1	Reminder: Limit of a Quotient Differentiation in Scalar Domains 4-9
2	Important Consequences: Three examples 10-18
3	Tensor Arguments: The Problem. You cannot divide by a vector, tensor. How can you differentiate with respect to a tensor? 19-20
4	Solution: The Gateaux Generalization of our usual directional derivative; you can now differentiate with respect to an object of any dimension! 21-26

Differentiation & Large Objects

- We are already familiar with the techniques of differentiation of scalar-valued functions with respect to scalar arguments. These objects are defined in scalar domains. Here,

$$x, h \in \mathbb{R}$$

- The derivative, $f'(x)$, of the function, $f: \mathbb{R} \to \mathbb{R}$, is defined as,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

pose no problems as division by scalars is well-defined.

Large Objects: Scalar Arguments

- Things become more complex when we handle tensors.
  - A second-order tensor, contains nine scalars.
  - A vector – a first-order tensor → three scalar members.
- The complication does not arise from the size of the objects themselves.
  - Derivation of tensor objects with respect to scalar domains, with some adjustments, basically conforms to the same rules as the above derivation of scalars:
  - Division of a tensor by a scalar is accomplished by multiplying the tensor by the inverse of the scalar.
  - This operation is defined in all vector spaces to which our vectors and tensors belong.

- Consequently, the derivative of the tensor $\mathbf{T}(t)$, with respect to a scalar argument, such as time, for example, can be defined as,

$$\frac{d}{dt}\mathbf{T}(t) = \lim_{h\to 0}\frac{\mathbf{T}(t+h)-\mathbf{T}(t)}{h}$$
$$\equiv \lim_{h\to 0}\frac{1}{h}\big(\mathbf{T}(t+h)-\mathbf{T}(t)\big)$$

- The product of the scalar, $\frac{1}{h}$ and the difference of tensors is a tensor. Hence the derivative of a vector (or tensor) with respect to a scalar is a vector (tensor).

Large (Tensor) Objects: Scalar Arguments

www.oafak.com; eds.s2pafrica.org; oafak@unilag.edu.ng

Large Objects: Scalar Arguments

- If $\alpha(t) \in \mathbb{R}$, and tensor, $\mathbf{T}(t) \in \mathbb{L}$ are both functions of time $t \in \mathbb{R}$, we find,

$$\frac{d}{dt}(\alpha\mathbf{T}) = \lim_{h\to 0}\frac{\alpha(t+h)\mathbf{T}(t+h) - \alpha(t)\mathbf{T}(t)}{h}$$

$$= \lim_{h\to 0}\frac{\alpha(t+h)\mathbf{T}(t+h) - \alpha(t)\mathbf{T}(t+h) + \alpha(t)\mathbf{T}(t+h) - \alpha(t)\mathbf{T}(t)}{h}$$

$$= \lim_{h\to 0}\frac{\alpha(t+h)\mathbf{T}(t+h) - \alpha(t)\mathbf{T}(t+h)}{h} + \lim_{h\to 0}\frac{\alpha(t)\mathbf{T}(t+h) - \alpha(t)\mathbf{T}(t)}{h}$$

$$= \left(\lim_{h\to 0}\frac{\alpha(t+h) - \alpha(t)}{h}\right)\left(\lim_{h\to 0}\mathbf{T}(t+h)\right) + \alpha(t)\lim_{h\to 0}\frac{\mathbf{T}(t+h) - \mathbf{T}(t)}{h}$$

$$= \frac{d}{dt}(\alpha\mathbf{T}) = \alpha\frac{d\mathbf{T}}{dt} + \frac{d\alpha}{dt}\mathbf{T}$$

Large Objects: Scalar Arguments

- Proceeding in a similar fashion, for $\alpha(t) \in \mathbb{R}, \mathbf{u}(t), \mathbf{v}(t) \in \mathbb{E}$, and $\mathbf{S}(t), \mathbf{T}(t) \in \mathbb{L}$, all being functions of a scalar variable t, the results in the following table hold as expected.
- The Simple Rule, when obeyed, allows us to gain proficiency and transfer scalar knowledge to tensors:
- Don't be fooled by the symbols! They are overloaded. You are no longer in Real Scalar Space! You are in the Euclidean Vector Space. Rules are different!

Remember Where You Are!

Remember…

Expression	Note
$\frac{d}{dt}(\alpha\mathbf{u}) = \alpha\frac{d\mathbf{u}}{dt} + \frac{d\alpha}{dt}\mathbf{u}$	Each term on the RHS retains the commutative property of multiplication by a scalar.
$\frac{d}{dt}(\mathbf{u}\cdot\mathbf{v}) = \frac{d\mathbf{u}}{dt}\cdot\mathbf{v} + \mathbf{u}\cdot\frac{d\mathbf{v}}{dt}$	Slide 10.6 tells that the derivative of a vector is a vector. Each term on the RHS retains the commutative property of the scalar product
$\frac{d}{dt}(\mathbf{u}\times\mathbf{v}) = \frac{d\mathbf{u}}{dt}\times\mathbf{v} + \mathbf{u}\times\frac{d\mathbf{v}}{dt}$	Original product order must be maintained
$\frac{d}{dt}(\mathbf{u}\otimes\mathbf{v}) = \frac{d\mathbf{u}}{dt}\otimes\mathbf{v} + \mathbf{u}\otimes\frac{d\mathbf{v}}{dt}$	Original product order must be maintained
$\frac{d}{dt}(\mathbf{T}+\mathbf{S}) = \frac{d\mathbf{T}}{dt} + \frac{d\mathbf{S}}{dt}$	Sum of tensors
$\frac{d}{dt}\mathbf{TS} = \frac{d\mathbf{T}}{dt}\mathbf{S} + \mathbf{T}\frac{d\mathbf{S}}{dt}$	Product of tensors. Not commutative! Note that we must maintain the order of the product as shown. $\mathbf{T}\frac{d\mathbf{S}}{dt} \neq \frac{d\mathbf{S}}{dt}\mathbf{T}$
$\frac{d}{dt}\mathbf{T}:\mathbf{S} = \frac{d\mathbf{T}}{dt}:\mathbf{S} + \mathbf{T}:\frac{d\mathbf{S}}{dt}$	Scalar Product of tensors. Commutative; Order is not important $\mathbf{T}:\frac{d\mathbf{S}}{dt} = \frac{d\mathbf{S}}{dt}:\mathbf{T}$
$\frac{d}{dt}(\alpha\mathbf{T}) = \alpha\frac{d\mathbf{T}}{dt} + \frac{d\alpha}{dt}\mathbf{T}$	Order is not important in multiplication by a scalar.

Puzzle:

Given that, $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{L}$; what is wrong with $\frac{d}{dt}(\mathbf{A}:\mathbf{B}:\mathbf{C})$? Can you find, $\frac{d}{dt}(\mathbf{ABC})$? $\frac{d}{dt}(\mathbf{A}+\mathbf{B}+\mathbf{C})$? $\frac{d}{dt}(\mathbf{AB}^{-1})$?

Examples: One
Constancy of the Identity Tensor

$$\frac{d\mathbf{I}}{dt} = \mathbf{O}.$$

- From Slide 10.6, we recognize the fact that the derivative of the tensor with respect to a scalar must give a tensor. The value here is the annihilator or Zero tensor, $\mathbf{O}$.
- This fact that the Identity Tensor does not change, and has a Zero derivative, leads to important results.
- We look at some of these as our first example.

Constancy of the Identity: Inverses

- For any invertible tensor valued scalar function, $\mathbf{S}(t)$, we differentiate the equation, $\mathbf{S}^{-1}(t)\mathbf{S}(t) = \mathbf{I}$ to obtain,

$$\frac{d\mathbf{S}^{-1}}{dt}\mathbf{S} + \mathbf{S}^{-1}\frac{d\mathbf{S}}{dt} = \mathbf{O}$$

$$\Rightarrow \frac{d\mathbf{S}^{-1}}{dt} = -\mathbf{S}^{-1}\frac{d\mathbf{S}}{dt}\mathbf{S}^{-1}$$

… if we post-multiply both sides by $\mathbf{S}^{-1}$, the following important expression results for the derivative of the inverse tensor with respect to a scalar parameter, in terms of the derivative of the original tensor function:

$$\frac{d\mathbf{S}^{-1}}{dt} = -\mathbf{S}^{-1}\frac{d\mathbf{S}}{dt}\mathbf{S}^{-1}$$

Conversely,

$$\frac{d\mathbf{S}}{dt} = -\mathbf{S}\frac{d\mathbf{S}^{-1}}{dt}\mathbf{S}$$

Constancy of the Identity: Orthogonal Tensors

- An orthogonal tensor as well as its transpose can each be functions of a scalar parameter.

$$\mathbf{Q}(t)\mathbf{Q}^{\mathrm{T}}(t) = \mathbf{I}$$

- One consequence of this relationship is that the tensor valued function,

$$\boldsymbol{\Omega}(t) \equiv \frac{d\mathbf{Q}(t)}{dt}\mathbf{Q}^{\mathrm{T}}(t)$$

What is a skew tensor?

of the same scalar parameter must be skew.

- This is a consequence of differentiating the identity:

$\mathbf{Q}$ is orthogonal, therefore,

$$\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$$

$$\frac{d}{dt}\left(\mathbf{Q}\mathbf{Q}^{\mathrm{T}}\right) = \frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\frac{d\mathbf{Q}^{\mathrm{T}}}{dt} = \frac{d\mathbf{I}}{dt} = \mathbf{O}$$

Consequently,

$$\frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\frac{d\mathbf{Q}^{\mathrm{T}}}{dt} = -\left(\frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}}\right)^{\mathrm{T}}$$

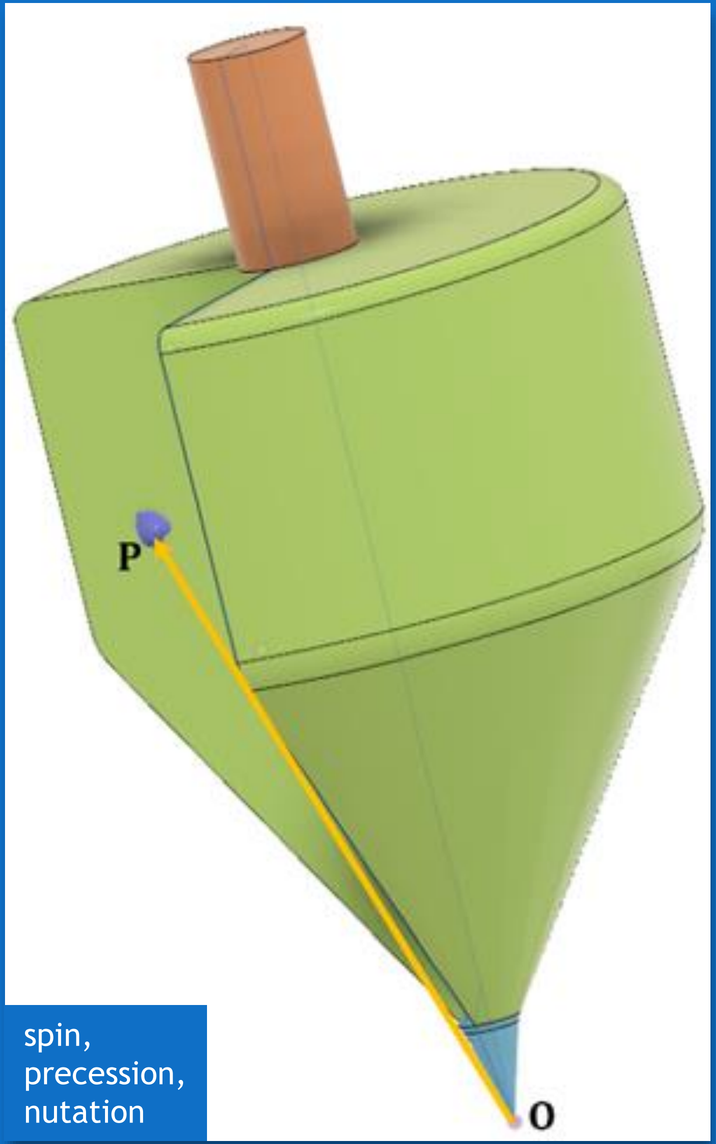
So we have that the tensor $\boldsymbol{\Omega} = \frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}}$ is negative of its own transpose, hence it is skew.

Constancy of the Identity: Angular Velocity

- Consider a rigid body fixed at one end $\mathbf{O}$ – for example, the spinning top shown. It is given a rotation $\mathbf{R}(t)$ from rest so that each point $\mathbf{P}$ is at a position vector $\mathbf{r}(t)$ at a time t, related to the original position $\mathbf{r}_o$ by the equation,
$$\mathbf{r}(t) = \mathbf{R}(t)\mathbf{r}_o$$

- We can find the velocity by differentiating the position vector,

$$\frac{d\mathbf{r}}{dt} = \frac{d\mathbf{R}}{dt}\mathbf{r}_o = \frac{d\mathbf{R}}{dt}\mathbf{R}^{-1}\mathbf{r}$$



spin, precession, nutation

- And the rotation is an orthogonal tensor, hence its inverse is its transpose, so that,

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{R}}{dt}\mathbf{R}^{\mathrm{T}}\mathbf{r} = \boldsymbol{\Omega}\mathbf{r}$$

- And, $\boldsymbol{\Omega}$ as we have seen above, is a skew tensor hence it is associated with an axial vector such that $(\boldsymbol{\omega}\,\times) = \boldsymbol{\Omega}$. From this fact we can see that every point in the body has an angular velocity, $\boldsymbol{\omega}$, such that,

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$$

- $\boldsymbol{\omega}$, defined by this expression, the axial vector of the $\frac{d\mathbf{R}}{dt}\mathbf{R}^{\mathrm{T}}$ where $\mathbf{R}(t)$ is the rotation function, is called the angular velocity.

Constancy of the Identity: Angular Velocity

Examples: Two Magnitude, Other Scalars

- We saw in the previous chapter that the tensor belongs to its own Euclidean vector space which is equipped with a scalar product
  - consequently, $\exists$ a scalar magnitude: $\forall\ \mathbf{A} \in \mathbb{L}$, $\|\mathbf{A}\| \equiv \sqrt{\mathbf{A}:\mathbf{A}} = \sqrt{\operatorname{tr}(\mathbf{A}\mathbf{A}^{\mathrm{T}})} = \sqrt{\operatorname{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{A})}$
- Consider the magnitude of a scalar (time, for example) dependent tensor,

$$\phi(t) = \sqrt{\mathbf{A}(t):\mathbf{A}(t)} = \|\mathbf{A}(t)\|$$

- so that, $\phi^2 = \mathbf{A}:\mathbf{A}.$

Differentiating this scalar equation, and remembering that the scalar operand here is just a product, we have,

$$\frac{d}{dt}\phi^2 = 2\phi\frac{d\phi}{dt} = \frac{d\mathbf{A}}{dt}:\mathbf{A} + \mathbf{A}:\frac{d\mathbf{A}}{dt} = 2\frac{d\mathbf{A}}{dt}:\mathbf{A}.$$

Why could we do this?

This simplifies to

$$\frac{d\phi}{dt} = \frac{d}{dt}\|\mathbf{A}(t)\| = \left(\frac{1}{\|\mathbf{A}(t)\|}\right)\frac{d\mathbf{A}(t)}{dt}:\mathbf{A}(t) = \frac{d\mathbf{A}(t)}{dt}:\frac{\mathbf{A}(t)}{\|\mathbf{A}(t)\|}$$

Examples: Three Tensor Invariants, The Trace

To obtain the derivative of the trace of a tensor, take the trace of the differentiated tensor.

- Trace is a linear operator. It follows immediately that

$$\frac{d}{dt}\,\mathrm{tr}\,\mathbf{A} = \mathrm{tr}\,\frac{d\mathbf{A}}{dt}$$

- To differentiate the trace of $\mathbf{A}(t)$, $t \in \mathbb{R}$, we select three linearly independent, constant $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$, we can write,

$$\frac{d}{dt} I_1(\mathbf{A}) = \frac{d}{dt}\mathrm{tr}\,\mathbf{A}$$

$$= \frac{d}{dt}\left(\frac{[\mathbf{Aa}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathbf{Ab}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \mathbf{Ac}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}\right)$$

$$= \frac{\left[\frac{d\mathbf{A}}{dt}\mathbf{a}, \mathbf{b}, \mathbf{c}\right] + \left[\mathbf{a}, \frac{d\mathbf{A}}{dt}\mathbf{b}, \mathbf{c}\right] + \left[\mathbf{a}, \mathbf{b}, \frac{d\mathbf{A}}{dt}\mathbf{c}\right]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}$$

$$= \mathrm{tr}\,\frac{d\mathbf{A}}{dt}$$

Posers:

- Is addition a linear operation? Derivative of a sum equals sum of derivatives?
- Is multiplication a linear operation? Derivative of a product, product of derivatives?

Examples: Three Tensor Invariants, Trace of the Cofactor

- The second invariant is NOT a linear scalar valued function of its tensor argument. However, we have the expression,

$$\mathbf{A}^{\mathrm{c}} = \mathbf{A}^{-\mathrm{T}} \det \mathbf{A}$$

$$\Rightarrow \operatorname{tr} \mathbf{A}^{\mathrm{c}} = \operatorname{tr}\left(\mathbf{A}^{-\mathrm{T}} \det \mathbf{A}\right)$$

- Differentiating with respect to t,

$$\frac{d}{dt} \operatorname{tr} \mathbf{A}^{\mathrm{c}} = \operatorname{tr} \frac{d}{dt}\left(\mathbf{A}^{-\mathrm{T}} \det \mathbf{A}\right)$$

- Not a very useful quantity.
- The *derivative of the third invariant* with respect to a scalar argument is of *momentous importance*.
- It is the basis of **Liouville's theorem** and is fundamental to the study of continuum flow in general.

Examples: Three Tensor Invariants, The Determinant

- The third invariant is not a linear function of its tensor argument.

$$I_3(\mathbf{A}) = \frac{[\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]} = \det \mathbf{A}$$

- so that, $[\mathbf{a}, \mathbf{b}, \mathbf{c}] \det \mathbf{A} = [\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}]$. Differentiating, we have,

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] \frac{d}{dt} \det \mathbf{A}$$

$$= \left[\frac{d\mathbf{A}}{dt}\mathbf{a}, \mathbf{Ab}, \mathbf{Ac}\right] + \left[\mathbf{Aa}, \frac{d\mathbf{A}}{dt}\mathbf{b}, \mathbf{Ac}\right] + \left[\mathbf{Aa}, \mathbf{Ab}, \frac{d\mathbf{A}}{dt}\mathbf{c}\right]$$

$$= \left[\frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}\right] + \left[\mathbf{Aa}, \frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\mathbf{Ab}, \mathbf{Ac}\right]$$

$$+ \left[\mathbf{Aa}, \mathbf{Ab}, \frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\mathbf{Ac}\right]$$

$$= \operatorname{tr}\left(\frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\right)[\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}]$$

So that, $\frac{d}{dt}\det \mathbf{A} = \operatorname{tr}\left(\frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}\right)\det \mathbf{A}$. A momentous theorem – Liouville's Theorem

Vector & Tensor Arguments

- When the domain of differentiation itself is a made up of large objects, the task of differentiation becomes more demanding. Such problems are standard in Continuum Mechanics. Examples:
  - Strain Energy function is a scalar, yet we can obtain the strains from it by differentiating with respect to the stress.
    - We are dealing there with the differentiation of a scalar function of a tensor: stress.

Why is this a tensor?

  - Velocity Gradient. Here, we are differentiating a vector field defined on the Euclidean point space, $\mathcal{E}$, with respect to the position vector of the points in $\mathcal{E}$.
- In these and several other derivatives of interest, the domains are no longer in the real scalar space.

Vector & Tensor Arguments !ERROR!

- When we are in a vector domain,

$$\mathbf{x}, \mathbf{h} \in \mathbb{E}$$

- The derivative, $\mathbf{F}'(\mathbf{x})$, of the function, $\mathbf{F}: \mathbb{E} \to \mathbb{E}$, is not properly defined as,

$$\mathbf{F}'(\mathbf{x}) = \lim_{\mathbf{h} \to \mathbf{o}} \frac{\mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x})}{\mathbf{h}}$$

creates several problems. For example, (1) division by vectors is not defined, and (2) there are many ways $\mathbf{h} \to \mathbf{o}$ can be achieved. Similar problems arise when the argument is a tensor: $\mathbf{H} \to \mathbf{O}$.

Vector & Tensor Arguments

- The approach to this challenge is twofold:
  - Recognize that the vectors and tensors live in their respective Euclidean **VECTOR** spaces where the concept of length is already defined.
  - Use the above to extend the concept of directional derivative to include the derivative of any object from a given Euclidean space with respect to objects from another.
- Such a generalization is in the Gateaux differential. Consider a map,

$$\mathbf{F}: \mathbb{V} \to \mathbb{W}$$

- This maps from the domain $\mathbb{V}$ to $\mathbb{W}$ both of which are Euclidean vector spaces. The concepts of limit and continuity carries naturally from the real space to any Euclidean vector space.

Vector & Tensor Arguments

- Let $\mathbf{v}_0 \in \mathbb{V}$ and $\mathbf{w}_0 \in \mathbb{W}$, as usual we can say that the limit
$$\lim_{\mathbf{v} \to \mathbf{v}_0} \mathbf{F}(\mathbf{v}) = \mathbf{w}_0$$
- if for any pre-assigned real number $\epsilon > 0$, no matter how small, we can always find a real number $\delta > 0$ such that $|\mathbf{F}(\mathbf{v}) - \mathbf{w}_0| \leq \epsilon$ whenever $|\mathbf{v} - \mathbf{v}_0| < \delta$. The function is said to be continuous at $\mathbf{v}_0$ if $\mathbf{F}(\mathbf{v}_0)$ exists and $\mathbf{F}(\mathbf{v}_0) = \mathbf{w}_0$

- Specifically, for $\alpha \in \mathbb{R}$ let this map be:

$$D\mathbf{F}(\mathbf{x},\mathbf{h}) \equiv \lim_{\alpha\to 0}\frac{\mathbf{F}(\mathbf{x}+\alpha\mathbf{h})-\mathbf{F}(\mathbf{x})}{\alpha} = \frac{d}{d\alpha}\mathbf{F}(\mathbf{x}+\alpha\mathbf{h})\bigg|_{\alpha=0}$$

- We focus attention on the second variable $\mathbf{h}$ while we allow the dependency on $\mathbf{x}$ to be as general as possible. We shall show that while the above function can be any given function of $\mathbf{x}$ (linear or nonlinear), the above map is always linear in $\mathbf{h}$ irrespective of what kind of Euclidean space we are mapping from or into. It is called the ***Gateaux Differential***.

The Gateaux Differential

Proof of Second Equation Term

- The second equation above is not obvious. It can be shown by remembering the scalar formula for derivative and treat

$$\phi(\alpha) \equiv \mathbf{F}(\mathbf{x} + \alpha\mathbf{h})$$

as a scalar function. We do that here as follows:

$$\frac{d\phi(\alpha)}{d\alpha} = \lim_{\Delta\alpha \to 0} \frac{\phi(\alpha + \Delta\alpha) - \phi(\alpha)}{\Delta\alpha}$$

Let $\phi(\alpha) \equiv \mathbf{F}(\mathbf{x} + \alpha\mathbf{h})$. Substituting, we have,

$$\frac{d\phi(\alpha)}{d\alpha} = \lim_{\Delta\alpha \to 0} \frac{\mathbf{F}(\mathbf{x} + (\alpha + \Delta\alpha)\mathbf{h}) - \mathbf{F}(\mathbf{x} + \alpha\mathbf{h})}{\Delta\alpha}$$

so that,

$$\left.\frac{d}{d\alpha}\mathbf{F}(\mathbf{x} + \alpha\mathbf{h})\right|_{\alpha=0} = \lim_{\Delta\alpha \to 0} \frac{\mathbf{F}(\mathbf{x} + \Delta\alpha\mathbf{h}) - \mathbf{F}(\mathbf{x})}{\Delta\alpha}$$

$$= \lim_{\beta \to 0} \frac{\mathbf{F}(\mathbf{x} + \beta\mathbf{h}) - \mathbf{F}(\mathbf{x})}{\beta}$$

$$= D\mathbf{F}(\mathbf{x}, \mathbf{h}).$$

Real functions in Real Domains.

- Let us make the Gateaux differential a little more familiar in real space in two steps: First, we move to the real space and allow $h \to dx$ and we obtain,

$$DF(x, dx) = \lim_{\alpha \to 0} \frac{F(x + \alpha dx) - F(x)}{\alpha} = \frac{d}{d\alpha} F(x + \alpha dx) \bigg|_{\alpha = 0}$$

- And let $\alpha dx \to \Delta x$, the middle term becomes,

$$\lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} dx = \frac{dF}{dx} dx$$

- from which it is obvious that the Gateaux derivative is a generalization of the well-known differential from elementary calculus. The Gateaux differential helps to compute a local linear approximation of any function (linear or nonlinear).

Linearity

- Gateaux differential is linear in its second argument, i.e., for $a \in \mathbb{R}$,
$$D\mathbf{F}(\mathbf{x}, a\mathbf{h}) = aD\mathbf{F}(\mathbf{x}, \mathbf{h})$$
- Furthermore,

$$\begin{aligned} D\mathbf{F}(\mathbf{x}, \mathbf{g} + \mathbf{h}) &= \lim_{\alpha \to 0} \frac{\mathbf{F}\big(\mathbf{x} + \alpha(\mathbf{g} + \mathbf{h})\big) - \mathbf{F}(\mathbf{x})}{\alpha} \\ &= \lim_{\alpha \to 0} \frac{\mathbf{F}\big(\mathbf{x} + \alpha(\mathbf{g} + \mathbf{h})\big) - \mathbf{F}(\mathbf{x} + \alpha\mathbf{g}) + \mathbf{F}(\mathbf{x} + \alpha\mathbf{g}) - \mathbf{F}(\mathbf{x})}{\alpha} \\ &= \lim_{\alpha \to 0} \frac{\mathbf{F}(\mathbf{y} + \alpha\mathbf{h}) - \mathbf{F}(\mathbf{y})}{\alpha} + \lim_{\alpha \to 0} \frac{\mathbf{F}(\mathbf{x} + \alpha\mathbf{g}) - \mathbf{F}(\mathbf{x})}{\alpha} \\ &= D\mathbf{F}(\mathbf{x}, \mathbf{h}) + D\mathbf{F}(\mathbf{x}, \mathbf{g}) \end{aligned}$$

as the variable $\mathbf{y} \equiv \mathbf{x} + \alpha\mathbf{g} \to \mathbf{x}$ as $\alpha \to 0$; For $a, b \in \mathbb{R}$, using similar arguments, we can also show that,

$$D\mathbf{F}(\mathbf{x}, a\mathbf{g} + b\mathbf{h}) = aD\mathbf{F}(\mathbf{x}, \mathbf{g}) + bD\mathbf{F}(\mathbf{x}, \mathbf{h})$$

Points to Note:

- The Gateaux differential is not unique to the point of evaluation.
  - Rather, at each point $\mathbf{x}$ there is a Gateaux differential for each "vector" $\mathbf{h}$. If the domain is a vector space, then we have a Gateaux differential for each of the infinitely many directions at each point. In two of more dimensions, there are infinitely many Gateaux differentials at each point!
  - $\mathbf{h}$ may not even be a vector, but second- or higher-order tensor.
  - It does not matter, as the tensors themselves are in a Euclidean space that define magnitude and direction as a result of the embedded inner product.
- The Gateaux differential is a one-dimensional calculation along a specified direction $\mathbf{h}$. Because it's one-dimensional, you can use ordinary one-dimensional calculus to compute it. Product rules and other constructs for the differentiation in real domains apply.