Levi-Civita Symbol and Vector Formula

xjunkong

December 2016

1 Introduction

A great discovery solves a great problem but there is a grain of discovery in the solution of any problem, Pólya and Conway [2004]. If you could master one or two great skills in vector analysis, we will never be afraid of any complicated vector or tensor equations. One I can see is the Levi-Civita and the other is the splitting of magnitude and direction of any vector (but tensor can have multiple magnitude and direction pairs).

2 Tensor

What is a tensor? From WIKI, a tensor is an algebraic object that describes a (multilinear) relationship between sets of algebraic objects related to a vector space.

1. A fancy word for a matrix
2. A container which can house data in n dimensions
3. A linear transformation from vectors to vectors
4. An ordered set of numbers that transform in a particular way upon a change of basis
5. The other names for "dyads" and "dyadics".

3 Vector

No coordinates or basis vectors are actually needed for the laws of physics. The derivations are at most expressible without the aid of any coordinates or bases. That means that the components are secondary and they only exist after one has chosen a set of basis vectors. Components (we claim) are an impediment to a clear and deep understanding of the laws of physics. The coordinate-free, component-free description is deeper, and once one becomes accustomed to it much more clear and understandable.

4 Vector Analysis

This part is extracted mostly from Callen [2002] and with some important ones being derived and explained. The following conventions are used in this appendix and throughout the book:

1. t is the scalar quantity, often referring to time passed from the start $t = 0$;
2. $\mathbf{x}$ is the vector from the origin to the point (x, y, z);
3. f, g, ϕ, ψ are scalar functions of $\mathbf{x}$, t;
4. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are vector functions of $\mathbf{x}$, t;
5. $A = |\mathbf{A}| \equiv \sqrt{\mathbf{A} \cdot \mathbf{A}}$ is the magnitude or length of the vector $\mathbf{A}$;
6. $\hat{\mathbf{A}} \equiv \mathbf{A}/|\mathbf{A}|$ is a unit vector in the $\mathbf{A}$ direction;
7. $\mathbf{T}$, $\mathbf{W}$, $\mathbf{AB}$, etc., are dyad (second rank tensor) functions of $\mathbf{x}$, t that will be called simply tensors;
8. $\mathbf{I}$ is the identity tensor or unit dyad;
9. $\mathbf{T}^\top$ is the transpose of tensor $\mathbf{T}$ (interchange of indices of the tensor elements), a tensor;
10. $\mathrm{tr}(\mathbf{T})$ is the trace of the tensor $\mathbf{T}$ (sum of its diagonal elements), a scalar;
11. $\det(\mathbf{T}) \equiv \|\mathbf{T}\|$ is the determinant of the tensor $\mathbf{T}$ (determinant of the matrix of tensor elements), a scalar.

4.1 D.1 Vector Algebra

Basic algebraic relations:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \qquad \text{commutative addition} \tag{1}$$
$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C} \qquad \text{associative addition} \tag{2}$$
$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}) \qquad \text{difference} \tag{3}$$
$$f\mathbf{A} = \mathbf{A}f \qquad \text{commutative scalar multiplication} \tag{4}$$
$$(f + g)\mathbf{A} = f\mathbf{A} + g\mathbf{A} \qquad \text{distributive scalar multiplication} \tag{5}$$
$$f(\mathbf{A} + \mathbf{B}) = f\mathbf{A} + f\mathbf{B} \qquad \text{distributive scalar multiplication} \tag{6}$$
$$f(g\mathbf{A}) = (fg)\mathbf{A} \qquad \text{associative scalar multiplication} \tag{7}$$

Dot product:

$$\mathbf{A} \cdot \mathbf{B} = 0 \quad \text{implies } (\mathbf{A} = 0 \text{ or } \mathbf{B} = 0) \text{ or } \mathbf{A} \perp \mathbf{B} \tag{8}$$
$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} \quad \text{commutative dot product} \tag{9}$$
$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \quad \text{distributive dot product} \tag{10}$$
$$(f\mathbf{A}) \cdot (g\mathbf{B}) = fg(\mathbf{A} \cdot \mathbf{B}) \quad \text{associative scalar, dot product} \tag{11}$$

Cross product:

$$\mathbf{A} \times \mathbf{B} = 0 \quad \text{implies } (\mathbf{A} = 0 \text{ or } \mathbf{B} = 0) \text{ or } \mathbf{A} \parallel \mathbf{B} \tag{12}$$
$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}, \quad \mathbf{A} \times \mathbf{A} = 0, \text{ anti-commutative cross product} \tag{13}$$
$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C} \quad \text{distributive cross product} \tag{14}$$
$$(f\mathbf{A}) \times (g\mathbf{B}) = fg(\mathbf{A} \times \mathbf{B}) \quad \text{associative scalar, cross product} \tag{15}$$

Scalar relations:

$$\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{A} \times \mathbf{B} \cdot \mathbf{C} = (\mathbf{C} \times \mathbf{A}) \cdot \mathbf{B} \quad \text{dot-cross product} \tag{16}$$
$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}) \tag{17}$$
$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) + (\mathbf{B} \times \mathbf{C}) \cdot (\mathbf{A} \times \mathbf{D}) + (\mathbf{C} \times \mathbf{A}) \cdot (\mathbf{B} \times \mathbf{D}) = 0 \tag{18}$$

Vector relations:

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \qquad \text{bac-cab rule} \\ &= (\mathbf{C} \times \mathbf{B}) \times \mathbf{A} = \mathbf{A} \cdot (\mathbf{CB} - \mathbf{BC}) \end{aligned} \tag{19}$$
$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = 0 \tag{20}$$
$$(\mathbf{A} \cdot \mathbf{B})\mathbf{C} = \mathbf{A} \cdot (\mathbf{BC}), \quad \text{associative dot product} \tag{21}$$
$$\begin{aligned} (\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) &= \mathbf{C}(\mathbf{A} \times \mathbf{B} \cdot \mathbf{D}) - \mathbf{D}(\mathbf{A} \times \mathbf{B} \cdot \mathbf{C}) \\ &= \mathbf{B}(\mathbf{C} \times \mathbf{D} \cdot \mathbf{A}) - \mathbf{A}(\mathbf{C} \times \mathbf{D} \cdot \mathbf{B}) \end{aligned} \tag{22}$$

Projection of a vector $\mathbf{A}$ in directions relative to a vector $\mathbf{B}$:

$$\mathbf{A} = \mathbf{A}_{\parallel}(\mathbf{B}/B) + \mathbf{A}_{\perp} = \mathbf{A}_{\parallel}\hat{\mathbf{b}} + \mathbf{A}_{\perp} \tag{23}$$
$$\hat{\mathbf{b}} \equiv \mathbf{B}/B, \quad \text{unit vector in } \mathbf{B} \text{ direction} \tag{24}$$
$$\mathbf{A}_{\parallel} \equiv \mathbf{B} \cdot \mathbf{A}/B = \hat{\mathbf{b}} \cdot \mathbf{A}, \quad \text{component of } \mathbf{A} \text{ along } \mathbf{B} \tag{25}$$
$$\begin{aligned} \mathbf{A}_{\perp} &\equiv -\mathbf{B} \times (\mathbf{B} \times \mathbf{A})/\mathbf{B}^2, \quad \text{component of } \mathbf{A} \text{ perpendicular to } \mathbf{B} \\ &= -\hat{\mathbf{b}} \times (\hat{\mathbf{b}} \times \mathbf{A}) \end{aligned} \tag{26}$$

4.2 D.2 Tensor Algebra

Scalar relations:

$$\mathbf{I} : \mathbf{AB} \equiv (\mathbf{I} \cdot \mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot \mathbf{B} \tag{27}$$
$$\mathbf{AB} : \mathbf{CD} \equiv \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})\mathbf{D} = (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \cdot \mathbf{D}) = \mathbf{D} \cdot \mathbf{AB} \cdot \mathbf{C} = \mathbf{B} \cdot \mathbf{CD} \cdot \mathbf{A} \tag{28}$$
$$\mathbf{I} : \mathbf{T} = \mathrm{tr}(\mathbf{T}), \mathbf{T} : \mathbf{T} \equiv |\mathbf{T}|2 \tag{29}$$
$$\mathbf{T} : \mathbf{AB} = (\mathbf{T} \cdot \mathbf{A}) \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{T} \cdot \mathbf{A} \tag{30}$$
$$\mathbf{AB} : \mathbf{T} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{T}) = \mathbf{B} \cdot \mathbf{T} \cdot \mathbf{A} \tag{31}$$
$$\mathbf{B} \times \mathbf{T} : \mathbf{W} = -(\mathbf{T} \cdot \mathbf{W})\mathbf{T} : \mathbf{B} \times \mathbf{I} \tag{32}$$

Vector relations:

$$\mathbf{I}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{I} = \mathbf{A} \tag{33}$$
$$\mathbf{A}\cdot\mathbf{T}^\top = \mathbf{T}\cdot\mathbf{A}, \mathbf{T}^\top\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{T} \tag{34}$$
$$\mathbf{A}\cdot(\mathbf{CB}-\mathbf{BC}) = \mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}),\ \text{bac - cab rule} \tag{35}$$
$$(\mathbf{A}\times\mathbf{C})\cdot\mathbf{T} = \mathbf{A}\cdot(\mathbf{C}\times\mathbf{T}) = -\mathbf{C}\cdot(\mathbf{A}\times\mathbf{T}) \tag{36}$$
$$\mathbf{T}\cdot(\mathbf{A}\times\mathbf{C}) = (\mathbf{T}\times\mathbf{A})\cdot\mathbf{C} = -(\mathbf{T}\times\mathbf{C})\cdot\mathbf{A} \tag{37}$$
$$\mathbf{A}\cdot(\mathbf{T}\times\mathbf{C}) = (\mathbf{A}\cdot\mathbf{T})\times\mathbf{C} = -\mathbf{C}\times(\mathbf{A}\cdot\mathbf{T}) \tag{38}$$
$$(\mathbf{A}\times\mathbf{T})\cdot\mathbf{C} = \mathbf{A}\times(\mathbf{T}\cdot\mathbf{C}) = -(\mathbf{T}\cdot\mathbf{C})\times\mathbf{A} \tag{39}$$
$$\mathbf{A}\cdot(\mathbf{T}\times\mathbf{C}) - \mathbf{C}\cdot(\mathbf{T}\times\mathbf{A}) = [\mathbf{I}\,\mathrm{tr}(\mathbf{T}) - \mathbf{T}]\cdot(\mathbf{A}\times\mathbf{C}) \tag{40}$$
$$(\mathbf{A}\times\mathbf{T})\cdot\mathbf{C} - (\mathbf{C}\times\mathbf{T})\cdot\mathbf{A} = (\mathbf{A}\times\mathbf{C})\cdot[\mathbf{I}\,\mathrm{tr}(\mathbf{T}) - \mathbf{T}] \tag{41}$$

Tensor relations:

$$\mathbf{I}\cdot\mathbf{AB} = (\mathbf{I}\cdot\mathbf{A})\mathbf{B} = \mathbf{AB}, \mathbf{AB}\cdot\mathbf{I} = \mathbf{A}(\mathbf{B}\cdot\mathbf{I}) = \mathbf{AB} \tag{42}$$
$$\mathbf{I}\times\mathbf{A} = \mathbf{I}\times\mathbf{A} \tag{43}$$
$$\mathbf{A}\times(\mathbf{BC}) = (\mathbf{A}\times\mathbf{B})\mathbf{C}, (\mathbf{AB})\times\mathbf{C} = \mathbf{A}(\mathbf{B}\times\mathbf{C}) \tag{44}$$
$$(\mathbf{A}\times\mathbf{B})\times\mathbf{I} = \mathbf{I}\times(\mathbf{A}\times\mathbf{B}) = \mathbf{BA} - \mathbf{AB} \tag{45}$$
$$(\mathbf{A}\times\mathbf{T})\mathbf{T} = -\mathbf{T}^\top\times\mathbf{A}, (\mathbf{T}\times\mathbf{A})\mathbf{T} = -\mathbf{A}\times\mathbf{T}^\top \tag{46}$$
$$(\mathbf{A}\times\mathbf{T}) - (\mathbf{A}\times\mathbf{T})\mathbf{T} = \mathbf{I}\times[\mathbf{A}\,\mathrm{tr}(\mathbf{T}) - \mathbf{T}\cdot\mathbf{A}] \tag{47}$$
$$(\mathbf{T}\times\mathbf{A}) - (\mathbf{T}\times\mathbf{A})\mathbf{T} = \mathbf{I}\times[\mathbf{A}\,\mathrm{tr}(\mathbf{T}) - \mathbf{A}\cdot\mathbf{T}] \tag{48}$$
$$\mathbf{T}_S = \frac{1}{2}(\mathbf{T}+\mathbf{T}^\top), \quad \text{symmetric part of tensor } \mathbf{T} \tag{49}$$
$$\mathbf{T}_A = \frac{1}{2}(\mathbf{T}-\mathbf{T}^\top), \quad \text{anti-symmetric part of tensor } \mathbf{T} \tag{50}$$
$$\begin{aligned}\mathbf{B}\times\mathbf{T}_S\times\mathbf{B} &= B^2\mathbf{T}_S - (\mathbf{BB}\cdot\mathbf{T}_S + \mathbf{T}_S\cdot\mathbf{BB}) \\ &\quad - (\mathbf{I}B^2 - \mathbf{BB})(\mathbf{IB}^2 - \mathbf{BB})\cdot\mathbf{T}_S/B^2 - \mathbf{BB}(\mathbf{BB}\cdot\mathbf{T}_S)/B^2\end{aligned} \tag{51}$$

4.3 D.3 Derivatives

The outer product of two vectors,

$$\nabla\cdot\left(\mathbf{BA}^\top\right) = \mathbf{A}\left(\nabla\cdot\mathbf{B}\right) + \left(\mathbf{B}\cdot\nabla\right)\mathbf{A} \tag{52}$$

Note that $\mathbf{B}\cdot\nabla()$ is the directional derivative of something in the direction of $\mathbf{B}$ multiplied (scaled) by $\mathbf{B}$'s magnitude. Therefore the divergence of $\mathbf{BA}^\top$ can be expressed in two directions of $\mathbf{A}$ (scaled by divergence of $\mathbf{B}$) and $\mathbf{B}$ (its directional derivative of $\mathbf{A}$ scaled by the magnitude of $\mathbf{B}$)

Temporal derivatives:

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} \quad \text{is a vector tangent to the curve defined by } \mathbf{A}(t) \tag{53}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(f\mathbf{A}) = \frac{\mathrm{d}f}{\mathrm{d}t}\mathbf{A} + f\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} \tag{54}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}+\mathbf{B}) = \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} + \frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \tag{55}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}\cdot\mathbf{B}) = \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\cdot\mathbf{B} + \mathbf{A}\cdot\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \tag{56}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}\times\mathbf{B}) = \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\times\mathbf{A} + \mathbf{A}\times\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \tag{57}$$

Definitions of partial derivatives in space ($\nabla \equiv \partial/\partial\mathbf{x} = del$ or ∇ is the differential vector operator):

$$\nabla f \equiv \frac{\partial f}{\partial \mathbf{x}}, \quad \text{gradient of scalar function } f\text{, a vector — vector in the direction and measure of the greatest rate of spatial change of } f \tag{58}$$

$$\nabla\cdot\mathbf{A} \equiv \frac{\partial}{\partial\mathbf{x}}\cdot\mathbf{A}, \quad \text{divergence of vector function } \mathbf{A}\text{, a scalar — divergence } (\nabla\cdot\mathbf{A} > 0) \text{ or convergence } (\nabla\cdot\mathbf{A} < 0) \text{ of } \mathbf{A} \text{ lines} \tag{59}$$

$$\nabla\times\mathbf{A} \equiv \frac{\partial}{\partial\mathbf{x}}\times\mathbf{A}, \quad \text{curl (or rotation) of vector function } \mathbf{A}\text{, a vector — vorticity of } \mathbf{A} \text{ lines} \tag{60}$$

$$\nabla^2 f \equiv \nabla\cdot\nabla f, \quad \text{del square or Laplacian (divergence of gradient) — derivative of scalar function } f\text{, a scalar also written as } \Delta f\text{, 3D measure of curvature of } f\text{, is larger where } \nabla^2 f < 0 \text{ and smaller where } \nabla^2 f > 0) \tag{61}$$

$$\nabla^2\mathbf{A} \equiv (\nabla\cdot\nabla)\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla\times(\nabla\times\mathbf{A}), \quad \text{Laplacian derivative of vector function } \mathbf{A}\text{, a vector. Mnemonic: } \textbf{GDMCC}\text{:Grad-Div Minus Curl-Curl} \tag{62}$$

For the general vector coordinate $\mathbf{x} \equiv x\hat{\mathbf{e}}_x + y\hat{\mathbf{e}}_y + z\hat{\mathbf{e}}_z$ and $|\mathbf{x}| \equiv \sqrt{x^2+y^2+z^2}$:

$$\nabla\cdot\mathbf{x} = 3 \qquad \nabla\cdot(\mathbf{x}/|\mathbf{x}|) = 2/|\mathbf{x}| \tag{63}$$

$$\nabla\times\mathbf{x} = 0 \qquad \nabla\times(\mathbf{x}/|\mathbf{x}|) = 0 \tag{64}$$

$$\nabla|\mathbf{x}| = \mathbf{x}/|\mathbf{x}| \qquad \nabla(1/|\mathbf{x}|) = -\mathbf{x}/|\mathbf{x}|^3 \tag{65}$$

$$\nabla\mathbf{x} = \mathbf{I} \tag{66}$$

$$(\mathbf{A}\cdot\nabla)\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) = [\mathbf{A} - (\mathbf{x}\cdot\mathbf{A})\mathbf{x}/|\mathbf{x}|^2]/|\mathbf{x}| \qquad \equiv \mathbf{A}_\perp/|\mathbf{x}| \tag{67}$$

$$\nabla^2(1/|\mathbf{x}|) \equiv \nabla\cdot\nabla(1/|\mathbf{x}|) = -\nabla\cdot(\mathbf{x}/|\mathbf{x}|^3) \qquad = -4\pi\delta(\mathbf{x}) \tag{68}$$

First derivatives with scalar functions:

$$\nabla(f+g) = \nabla f + \nabla g \tag{69}$$
$$\nabla(fg) = (\nabla f)g + f\nabla g = \nabla(gf) \tag{70}$$
$$\nabla(f\mathbf{A}) = (\nabla f)\mathbf{A} + f\nabla\mathbf{A} \tag{71}$$
$$\nabla \cdot f\mathbf{A} = \nabla f \cdot \mathbf{A} + f\nabla \cdot \mathbf{A} \tag{72}$$
$$\nabla \times f\mathbf{A} = \nabla f \times \mathbf{A} + f\nabla \times \mathbf{A} \tag{73}$$
$$\nabla \cdot f\mathbf{T} = \nabla f \cdot \mathbf{T} + f\nabla \cdot \mathbf{T} \tag{74}$$
$$\nabla \times f\mathbf{T} = \nabla f \times \mathbf{T} + f\nabla \times \mathbf{T}. \tag{75}$$

First derivative scalar relations:

$$\nabla \cdot (\mathbf{A}+\mathbf{B}) = \nabla \cdot \mathbf{A} + \nabla \cdot \mathbf{B} \tag{76}$$
$$\nabla \cdot (\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot\nabla\times\mathbf{A} - \mathbf{A}\cdot\nabla\times\mathbf{B} \tag{77}$$
$$\begin{aligned}(\mathbf{B}\cdot\nabla)(\mathbf{A}\cdot\mathbf{C}) &= \mathbf{C}\cdot(\mathbf{B}\cdot\nabla)\mathbf{A} + \mathbf{A}\cdot(\mathbf{B}\cdot\nabla)\mathbf{C} \\ &\equiv \mathbf{CB} : \nabla\mathbf{A} + \mathbf{AB} : \nabla\mathbf{C}\end{aligned} \tag{78}$$
$$\mathbf{A}\cdot\nabla\mathbf{B}\cdot\mathbf{C} - \mathbf{C}\cdot\nabla\mathbf{B}\cdot\mathbf{A} \equiv (\mathbf{CA}-\mathbf{AC}) : \nabla\mathbf{B} = (\mathbf{A}\times\mathbf{C})\cdot\nabla\times\mathbf{B} \tag{79}$$
$$\begin{aligned}2\mathbf{A}\cdot\nabla\mathbf{B}\cdot\mathbf{C} \equiv 2\mathbf{CA} : \nabla\mathbf{B} = {}& \mathbf{A}\cdot\nabla(\mathbf{B}\cdot\mathbf{C}) + \mathbf{C}\cdot\nabla(\mathbf{B}\cdot\mathbf{A}) \\ & - \mathbf{B}\cdot\nabla(\mathbf{A}\cdot\mathbf{C}) + (\mathbf{B}\times\mathbf{C})\cdot(\nabla\times\mathbf{A}) \\ & + (\mathbf{B}\times\mathbf{A})\cdot(\nabla\times\mathbf{C}) + (\mathbf{A}\times\mathbf{C})\cdot(\nabla\times\mathbf{B})\end{aligned} \tag{80}$$
$$\mathbf{I} : \nabla\mathbf{B} = \nabla\cdot\mathbf{B} \tag{81}$$
$$\mathbf{A}\times\mathbf{I} : \nabla\mathbf{B} = \mathbf{A}\cdot\nabla\times\mathbf{B} \tag{82}$$
$$\mathbf{A}\cdot\nabla\cdot\mathbf{T} = \nabla\cdot(\mathbf{A}\cdot\mathbf{T}) - \nabla\mathbf{A} : \mathbf{T} = \nabla\cdot(\mathbf{A}\cdot\mathbf{T}) - \mathbf{T} : \nabla\mathbf{A} \tag{83}$$

First derivative vector relations:

$$\nabla \times (\mathbf{A} + \mathbf{B}) = \nabla \times \mathbf{A} + \nabla \times \mathbf{B} \tag{84}$$

$$\begin{aligned}\nabla(\mathbf{A} \cdot \mathbf{B}) &= \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} \\ &= (\nabla \mathbf{A}) \cdot \mathbf{B} + (\nabla \mathbf{B}) \cdot \mathbf{A}\end{aligned} \tag{85}$$

$$\nabla(\mathbf{B}^2/2) \equiv \nabla(\mathbf{B} \cdot \mathbf{B}/2) = \mathbf{B} \times (\nabla \times \mathbf{B}) + (\mathbf{B} \cdot \nabla)\mathbf{B} = (\nabla \mathbf{B}) \cdot \mathbf{B} \tag{86}$$

$$(\mathbf{B} \cdot \nabla)(\mathbf{A} \times \mathbf{C}) = (\mathbf{B} \cdot \nabla)\mathbf{A} \times \mathbf{C} + \mathbf{A} \times (\mathbf{B} \cdot \nabla)\mathbf{C} \tag{87}$$

$$\nabla \cdot \mathbf{A}\mathbf{B} = (\nabla \cdot \mathbf{A})\mathbf{B} + (\mathbf{A} \cdot \nabla)\mathbf{B} = (\nabla \cdot \mathbf{A})\mathbf{B} + \mathbf{A} \cdot (\nabla \mathbf{B}) \tag{88}$$

$$\nabla \cdot \mathbf{I} = 0 \tag{89}$$

$$\nabla \cdot (\mathbf{I} \times \mathbf{A}) = \nabla \times \mathbf{A} \tag{90}$$

$$\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - \mathbf{A} \cdot (\nabla \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} \tag{91}$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A}) + (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} = \nabla \cdot (\mathbf{B}\mathbf{A} - \mathbf{A}\mathbf{B}) \tag{92}$$

$$\mathbf{A} \cdot \nabla \mathbf{B} \times \mathbf{C} + \mathbf{C} \times \nabla \mathbf{B} \cdot \mathbf{A} = \mathbf{C} \times [\mathbf{A} \times (\nabla \times \mathbf{B})] \tag{93}$$

$$\mathbf{A} \cdot \nabla \mathbf{B} \times \mathbf{C} - \mathbf{C} \cdot \nabla \mathbf{B} \times \mathbf{A} = [(\nabla \cdot \mathbf{B})\mathbf{I} - \nabla \mathbf{B}] \cdot (\mathbf{A} \times \mathbf{C}) \tag{94}$$

$$\mathbf{A} \times \nabla \mathbf{B} \cdot \mathbf{C} - \mathbf{C} \times \nabla \mathbf{B} \cdot \mathbf{A} = (\mathbf{A} \times \mathbf{C}) \cdot [(\nabla \cdot \mathbf{B})\mathbf{I} - \nabla \mathbf{B}] \tag{95}$$

$$\mathbf{I} \cdot \nabla \mathbf{B} = \nabla \mathbf{B}, \nabla \mathbf{B} \cdot \mathbf{I} = \nabla \mathbf{B} \tag{96}$$

$$\nabla \times \mathbf{A}\mathbf{B} = (\nabla \times \mathbf{A})\mathbf{B} - \mathbf{A} \times \nabla \mathbf{B} \tag{97}$$

$$\nabla(\mathbf{A} \times \mathbf{B}) = \nabla \mathbf{A} \times \mathbf{B} - \nabla \mathbf{B} \times \mathbf{A} \tag{98}$$

$$\begin{aligned}&\mathbf{A} \times \nabla \mathbf{B} + \nabla \mathbf{B} \times \mathbf{A} \\ &\qquad = \mathbf{I} \times [(\nabla \cdot \mathbf{B})\mathbf{A} - (\nabla \mathbf{B}) \cdot \mathbf{A}] + [\mathbf{A} \cdot (\nabla \times \mathbf{B})]\mathbf{I} - \mathbf{A}(\nabla \times \mathbf{B}) \\ &\qquad = \mathbf{I} \times [(\nabla \cdot \mathbf{B})\mathbf{A} - \mathbf{A} \cdot (\nabla \mathbf{B})] + [\mathbf{A} \cdot (\nabla \times \mathbf{B})]\mathbf{I} - (\nabla \times \mathbf{B})\mathbf{A}\end{aligned} \tag{99}$$

$$\nabla \mathbf{B} \times \mathbf{A} + (\mathbf{A} \times \nabla \mathbf{B})\mathbf{T} = [\mathbf{A} \cdot (\nabla \times \mathbf{B})]\mathbf{I} - \mathbf{A}(\nabla \times \mathbf{B}) \tag{100}$$

$$\mathbf{A} \times \nabla \mathbf{B} + (\nabla \mathbf{B} \times \mathbf{A})\mathbf{T} = [\mathbf{A} \cdot (\nabla \times \mathbf{B})]\mathbf{I} - (\nabla \times \mathbf{B})\mathbf{A} \tag{101}$$

$$\mathbf{A} \times \nabla \mathbf{B} - (\mathbf{A} \times \nabla \mathbf{B})\mathbf{T} = \mathbf{I} \times [(\nabla \cdot \mathbf{B})\mathbf{A} - (\nabla \mathbf{B}) \cdot \mathbf{A}] \tag{102}$$

$$\nabla \mathbf{B} \times \mathbf{A} - (\nabla \mathbf{B} \times \mathbf{A})\mathbf{T} = [(\nabla \cdot \mathbf{B})\mathbf{A} - \mathbf{A} \cdot (\nabla \mathbf{B})] \times \mathbf{I} \tag{103}$$

Second derivative relations:

$$\nabla \cdot \nabla f \equiv \nabla^2 f \tag{104}$$

$$\nabla \times \nabla f = 0 \tag{105}$$

$$\nabla \cdot \nabla f \times \nabla g = 0 \tag{106}$$

$$\nabla \cdot \nabla \mathbf{A} \equiv \nabla^2 \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}) \tag{107}$$

$$\nabla \cdot \nabla \times \mathbf{A} = 0 \tag{108}$$

$$\nabla \cdot (\mathbf{B} \cdot \nabla)\mathbf{A} = (\mathbf{B} \cdot \nabla)(\nabla \cdot \mathbf{A}) - (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{B}) \tag{109}$$

$$\nabla \times [(\mathbf{A} \cdot \nabla)\mathbf{A}] = (\mathbf{A} \cdot \nabla)(\nabla \times \mathbf{A}) + (\nabla \cdot \mathbf{A})(\nabla \times \mathbf{A}) - [(\nabla \times \mathbf{A}) \cdot \nabla]\mathbf{A} \tag{110}$$

Derivatives of projections of $\mathbf{A}$ in $\mathbf{B}$ direction [$\hat{\mathbf{b}} \equiv \mathbf{B}/B$, $\mathbf{A} = A_\parallel \hat{\mathbf{b}} +$

$\mathbf{A}_\perp,\ A_\parallel \equiv \hat{\mathbf{b}}\cdot\mathbf{A},\ \mathbf{A}_\perp \equiv -\hat{\mathbf{b}}\times(\hat{\mathbf{b}}\times\mathbf{A}),\ (\hat{\mathbf{b}}\cdot\nabla)\hat{\mathbf{b}} = -\hat{\mathbf{b}}\times(\nabla\times\hat{\mathbf{b}})\equiv\boldsymbol{\kappa}]$:

$$\nabla\cdot\mathbf{A} = (A_\parallel/\mathbf{B})(\nabla\cdot\mathbf{B}) + (\mathbf{B}\cdot\nabla)(A_\parallel/B) + \nabla\cdot\mathbf{A}_\perp \tag{111}$$

$$\nabla\cdot\mathbf{A}_\perp = -\mathbf{A}_\perp\cdot[\nabla\ln(B) + (\hat{\mathbf{b}}\cdot\nabla)\hat{\mathbf{b}}] - (1/\mathbf{B})\hat{\mathbf{b}}\cdot\nabla\times(\mathbf{B}\times\mathbf{A}) \tag{112}$$

$$\begin{aligned}\hat{\mathbf{b}}\cdot\nabla\mathbf{A}\cdot\hat{\mathbf{b}} &\equiv \hat{\mathbf{b}}\hat{\mathbf{b}}:\nabla\mathbf{A} = (\hat{\mathbf{b}}\cdot\nabla)\mathbf{A}_\parallel - \mathbf{A}_\perp\cdot(\hat{\mathbf{b}}\cdot\nabla)\hat{\mathbf{b}}\\ &= \mathbf{A}\cdot\nabla\ln(B) - (1/\mathbf{B})\hat{\mathbf{b}}\cdot\nabla\times(\mathbf{B}\times\mathbf{A}) + \nabla\cdot\mathbf{A} - (\mathbf{A}_\parallel/\mathbf{B})(\nabla\cdot\mathbf{B})\end{aligned} \tag{113}$$

$$\text{For } \mathbf{A}_\perp = (1/B^2)\mathbf{B}\times\nabla f,\ \hat{\mathbf{b}}\cdot\nabla\times(\mathbf{B}\times\mathbf{A}_\perp) = (\hat{\mathbf{b}}\cdot\nabla f)(\hat{\mathbf{b}}\cdot\nabla\times\hat{\mathbf{b}}) \tag{114}$$

4.4 D.4 Integrals

For a volume V enclosed by a closed, continuous surface S with differential volume element d^3x and differential surface element $\mathrm{d}S \equiv \hat{\mathbf{n}}\mathrm{d}S$ where $\hat{\mathbf{n}}$ is the unit normal outward from the volume V , for well-behaved functions f, g,$\mathbf{A}$, $\mathbf{B}$ and $\mathbf{T}$:

$$\int_V \mathrm{d}^3x\nabla f = \oiint_S \mathrm{d}\mathbf{S}f, \tag{115}$$

$$\int_V \mathrm{d}^3x\nabla\cdot\mathbf{A} = \oiint_S \mathrm{d}\mathbf{S}\cdot\mathbf{A}, \quad \text{Divergence or Gauss' theorem,} \tag{116}$$

$$\int_V \mathrm{d}^3x\nabla\cdot\mathbf{T} = \oiint_S \mathrm{d}\mathbf{S}\cdot\mathbf{T}, \tag{117}$$

$$\int_V \mathrm{d}^3x\nabla\times\mathbf{A} = \oiint_S \mathrm{d}\mathbf{S}\times\mathbf{A}, \tag{118}$$

$$\int_V \mathrm{d}^3x f\nabla 2g = \int_V \mathrm{d}^3x\nabla f\cdot\nabla g + \oiint_S \mathrm{d}\mathbf{S}\cdot f\nabla g, \quad \text{Green's 1st identity} \tag{119}$$

$$\int_V \mathrm{d}^3x(f\nabla^2 g - g\nabla^2 f) = \oiint_S \mathrm{d}\mathbf{S}\cdot(f\nabla g - g\nabla f), \text{ Green's 2nd identity} \tag{120}$$

$$\begin{aligned}&\int_V \mathrm{d}^3x[\mathbf{A}\cdot\nabla\times(\nabla\times\mathbf{B}) - \mathbf{B}\cdot\nabla\times(\nabla\times\mathbf{A})]\\ &\quad = \oiint \mathrm{d}S\cdot[\mathbf{B}\times(\nabla\times\mathbf{A}) - \mathbf{A}\times(\nabla\times\mathbf{B})],\\ &\qquad\qquad \text{vector form of Green's second identity.}\end{aligned} \tag{121}$$

The gradient, divergence and curl partial differential operators can be defined using integral relations in the limit of small surfaces ΔS encompassing small volumes ΔV, as follows:

$$\nabla f \equiv \lim_{\Delta V\to 0} \oiint_{\Delta S} \mathrm{d}Sf \qquad \text{gradient,} \tag{122}$$

$$\nabla\cdot\mathbf{A} \equiv \lim_{\Delta V\to 0} \oiint_{\Delta S} \mathrm{d}S\cdot\mathbf{A} \qquad \text{divergence,} \tag{123}$$

$$\nabla\times\mathbf{A} \equiv \lim_{\Delta V\to 0} \oiint_{\Delta S} \mathrm{d}S\times\mathbf{A} \qquad \text{curl.} \tag{124}$$

For S representing an open surface bounded by a closed, continuous contour $\mathbf{C}$ with line element $\mathrm{d}\boldsymbol{\ell}$ which is defined to be positive when the right-hand-rule sense of the line integral around $\mathbf{C}$ points in the $\mathrm{d}S$ direction:

$$\iint_S \mathrm{d}S \times \nabla f = \oint_{\mathbf{C}} \mathrm{d}\boldsymbol{\ell} f, \tag{125}$$

$$\iint_S \mathrm{d}S \cdot \nabla \times \mathbf{A} = \oint_{\mathbf{C}} \mathrm{d}\boldsymbol{\ell} \cdot \mathbf{A}, \qquad \text{Stokes' theorem,} \tag{126}$$

$$\iint_S (\mathrm{d}S \times \nabla) \times \mathbf{A} = \oint_{\mathbf{C}} \mathrm{d}\boldsymbol{\ell} \times \mathbf{A}, \tag{127}$$

$$\iint_S \mathrm{d}S \cdot (\nabla f \times \nabla g) = \oint_{\mathbf{C}} \mathrm{d}\boldsymbol{\ell} \cdot f \nabla g = \oint_C f \mathrm{d}g = -\oint_C g \mathrm{d}f, \qquad \text{Green's theorem.} \tag{128}$$

The appropriate differential line element $\mathrm{d}\boldsymbol{\ell}$, surface area $\mathrm{d}\mathbf{S}$, and volume d^3x can be defined in terms of any three differential line elements $\mathrm{d}\boldsymbol{\ell}(i)$, $i = 1, 2, 3$ that are linearly independent, i.e., $\llbracket \mathrm{d}\boldsymbol{\ell}(1)\ \mathrm{d}\boldsymbol{\ell}(2)\ \mathrm{d}\boldsymbol{\ell}(3) \rrbracket = \mathrm{d}\boldsymbol{\ell}(1) \cdot \mathrm{d}\boldsymbol{\ell}(2) \times \mathrm{d}\boldsymbol{\ell}(3) \neq 0$ by

5 Levi-Civita

5.1 Introduction

For the whole university life, it was a shame not being able to fully understand the Levi-Civita symbol ε_{ijk} The Levi-Civita symbol is a tensor of rank three and is defined by

$$\varepsilon_{ijk} = \varepsilon^{ijk} = \begin{cases} 0, & \text{if any two indices equal} \\ 1, & \text{if } (i,j,k) \text{ is even permutation of } (1,2,3) \\ -1, & \text{if } (i,j,k) \text{ is odd permutation of } (1,2,3) \end{cases} \tag{129}$$

The Levi-Civita symbol ε_{ijk} is anti-symmetric on each pair of indices. The Levi-Civita symbol can be expressed as the determinant, or mixed triple product of the unit vectors $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ of a normalised and direct orthogonal frame of reference.

$$\varepsilon_{ijk} = \det(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j, \hat{\mathbf{e}}_k) = \hat{\mathbf{e}}_i \cdot (\hat{\mathbf{e}}_j \times \hat{\mathbf{e}}_k) \tag{130}$$

The cross product can be defined by analogy as follows

$$\mathbf{a} \times \mathbf{b} = \det(\hat{\mathbf{e}}_i, \mathbf{a}_j, \mathbf{b}_k) = \det \begin{vmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \varepsilon_{ijk} \hat{\mathbf{e}}_i a_j b_k \tag{131}$$

The determinant of a 3×3 square matrix $\mathbf{A} = [a_{ij}]$ can be written

$$\det(\mathbf{A}) = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} a_{1i} a_{2j} a_{3k} \tag{132}$$

5.2 Properties

- The Levi-Civita tensor ε_{ijk} has $C_3^1 C_3^1 C_3^1 = 3 \times 3 \times 3 = 27$ components.
- $C_3^1(C_3^1 C_2^1 + 1) = 21$ components with repeated indices are equal to 0.
- 3 components with even permutation of indices $(1, 2, 3)$ are equal to 1.
- 3 components with old permutation of indices $(1, 2, 3)$ are equal to -1.

5.3 Identities

The product of two Levi-Civita symbols can be expressed as a function of the Kronecker's symbol[1] $\delta_{ij} = \delta_i^j = \delta_j^i = 1$ iff $i = j$, otherwise zero. The properties are easily observed by expressing ε_{ijk} in a determinant form. This also can be generalized for $n > 3$. Let's look at the 3×3 identity matrix consisting of the three unit column vectors $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$.

$$\left[\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\ \hat{\mathbf{e}}_3\right] = \left[\begin{array}{c:c:c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \tag{133}$$

Obviously, the unit columns are orthonormal,

$$\hat{\mathbf{e}}_i^{\mathrm{T}} \hat{\mathbf{e}}_j = \delta_{ij} = \delta_i^j = \delta_j^i, \quad i, j = 1, 2, 3, \tag{134}$$

where δ_{ij} is the Kronecker delta.

Consider determinants consisting of three columns selected out of the three unit columns. Then by the properties of determinants:

$$\left|\hat{\mathbf{e}}_i\ \hat{\mathbf{e}}_j\ \hat{\mathbf{e}}_k\right| = 0 \quad \text{if} \quad i = j,\ i = k, \text{or}\ \ j = k. \tag{135}$$

Further,

$$\left|\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\ \hat{\mathbf{e}}_3\right| = \left|\hat{\mathbf{e}}_2\ \hat{\mathbf{e}}_3\ \hat{\mathbf{e}}_1\right| = \left|\hat{\mathbf{e}}_3\ \hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_2\right| = 1 \tag{136}$$

$$\left|\hat{\mathbf{e}}_2\ \hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_3\right| = \left|\hat{\mathbf{e}}_3\ \hat{\mathbf{e}}_2\ \hat{\mathbf{e}}_1\right| = \left|\hat{\mathbf{e}}_1\ \hat{\mathbf{e}}_3\ \hat{\mathbf{e}}_2\right| = -1. \tag{137}$$

Hence

$$\varepsilon_{ijk} = \left|\hat{\mathbf{e}}_i\ \hat{\mathbf{e}}_j\ \hat{\mathbf{e}}_k\right|. \tag{138}$$

Introduce two 3×3 unit matrices $\mathbf{A}$ and $\mathbf{B}$ as short-hand notations:

$$\det(\mathbf{A}) := \left|\hat{\mathbf{e}}_i\ \hat{\mathbf{e}}_j\ \hat{\mathbf{e}}_k\right|, \qquad \det(\mathbf{B}) := \left|\hat{\mathbf{e}}_\ell\ \hat{\mathbf{e}}_m\ \hat{\mathbf{e}}_n\right|. \tag{139}$$

Use $\varepsilon_{ijk}\varepsilon_{\ell mn} = \det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{A}^{\mathrm{T}})\det(\mathbf{B}) = \det(\mathbf{A}^{\mathrm{T}}\mathbf{B})$. The determinant of the product of two matrices can be written as:

$$\varepsilon_{ijk}\varepsilon^{lmn} = \det\left(\begin{bmatrix} \hat{\mathbf{e}}_i^{\mathrm{T}} \\ \hat{\mathbf{e}}_j^{\mathrm{T}} \\ \hat{\mathbf{e}}_k^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{e}}^\ell & \hat{\mathbf{e}}^m & \hat{\mathbf{e}}^n \end{bmatrix}\right) = \begin{vmatrix} \delta_i^\ell & \delta_i^m & \delta_i^n \\ \delta_j^\ell & \delta_j^m & \delta_j^n \\ \delta_k^\ell & \delta_k^m & \delta_k^n \end{vmatrix} \tag{140}$$

$$= \delta_i^\ell(\delta_j^m\delta_k^n - \delta_j^n\delta_k^m) - \delta_i^m(\delta_j^\ell\delta_k^n - \delta_j^n\delta_k^\ell) + \delta_i^n(\delta_j^\ell\delta_k^m - \delta_j^m\delta_k^\ell) \tag{141}$$

[1] we will not make difference between upper or lower index notation

If $i = \ell$ in (141), then

$$\begin{aligned}\varepsilon_{ijk}\varepsilon^{imn} &= \delta_i^i(\delta_j^m\delta_k^n - \delta_j^n\delta_k^m) - \delta_i^m(\delta_j^i\delta_k^n - \delta_j^n\delta_k^i) + \delta_i^n(\delta_j^i\delta_k^m - \delta_j^m\delta_k^i)\\ &= 3(\delta_j^m\delta_k^n - \delta_j^n\delta_k^m) - \delta_j^m\delta_k^n + \delta_k^m\delta_j^n + \delta_j^n\delta_k^m - \delta_j^m\delta_k^n\\ &= \delta_j^m\delta_k^n - \delta_j^n\delta_k^m \end{aligned} \tag{142}$$

It is seen that the two equal indices lead to the total index number (or dimension of the tensor) reduced by 2. If further let $j = m$ in (142), then

$$\begin{aligned}\varepsilon_{ijk}\varepsilon^{ijn} &= \delta_j^j\delta_k^n - \delta_j^n\delta_k^j\\ &= 3\delta_k^n - \delta_k^n\\ &= 2\delta_k^n \end{aligned} \tag{143}$$

From (143), the fact $\varepsilon_{ijk}\varepsilon^{ijk} = 2 \times 3 = 6$ is obtained[2].

In fact, repetition of any two indices will follow the same rule of contraction. To see the flexibility of the contraction, let $n = k$ in (141)

$$\mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{bmatrix}\hat{\mathbf{e}}_i^{\mathrm{T}}\\ \hat{\mathbf{e}}_j^{\mathrm{T}}\\ \hat{\mathbf{e}}_k^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{e}}_\ell & \hat{\mathbf{e}}_m & \hat{\mathbf{e}}_k\end{bmatrix} = \begin{bmatrix}\delta_{i\ell} & \delta_{im} & \delta_{ik}\\ \delta_{j\ell} & \delta_{jm} & \delta_{jk}\\ \delta_{k\ell} & \delta_{km} & \delta_{kk}\end{bmatrix} = \begin{bmatrix}\delta_{i\ell} & \delta_{im} & 0\\ \delta_{j\ell} & \delta_{jm} & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{144}$$

The zeros in the third row/column is due to the independence of unit basis $(\hat{\mathbf{e}}_i\ \hat{\mathbf{e}}_j\ \hat{\mathbf{e}}_k)$ with $i \neq j \neq k$ and $(\hat{\mathbf{e}}_\ell\ \hat{\mathbf{e}}_m\ \hat{\mathbf{e}}_k)$ with $\ell \neq m \neq k$, if otherwise, $\varepsilon_{ijk} = 0$. Therefore the determinant is

$$\det(\mathbf{A}^{\mathrm{T}}\mathbf{B}) = \begin{vmatrix}\delta_{i\ell} & \delta_{im} & 0\\ \delta_{j\ell} & \delta_{jm} & 0\\ 0 & 0 & 1\end{vmatrix} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}. \tag{145}$$

A generalization of the property to arbitrary n is clear now:

$$\sum_{k=1}^{n} \varepsilon_{i_1\ i_2\ \dots\ i_{n-1}\ k}\varepsilon_{j_1\ j_2\ \dots\ j_{n-1}\ k} = \begin{vmatrix}\delta_{i_1j_1} & \delta_{i_1j_2} & \delta_{i_1j_3} & \cdots & \delta_{i_1j_{n-1}}\\ \delta_{i_2j_1} & \delta_{i_2j_2} & \cdots & \cdots & \delta_{i_2j_{n-1}}\\ \cdots & & & & \cdots\\ \delta_{i_{n-1}j_1} & \delta_{i_{n-1}j_2} & \cdots & \cdots & \delta_{i_{n-1}j_{n-1}}\end{vmatrix}. \tag{146}$$

The second property of the Levi-Civita symbol follows from

$$\begin{bmatrix}\hat{\mathbf{e}}_i^{\mathrm{T}}\\ \hat{\mathbf{e}}_p^{\mathrm{T}}\\ \hat{\mathbf{e}}_q^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{e}}_j & \hat{\mathbf{e}}_p & \hat{\mathbf{e}}_q\end{bmatrix} = \begin{bmatrix}\delta_{ij} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{147}$$

The determinant of the last matrix is equal to δ_{ij}. The same holds for p and q interchanged. In the case of general n the sum is over $(n-1)!$ permutations [note that $(3-1)! = 2$]. The final property contains a summation over six (3!) non-zero terms; each term is the determinant of the identity matrix, which is unity.

[2]Einstein Summation Convention: If an index appears (exactly) twice, then it is summed over and appears only on one side of an equation. A single index (called a free index) appears once on each side of the equation.

6 Skew-symmetric matrix

In mathematics, particularly in linear algebra, a skew-symmetric (or antisymmetric or antimetric matrix is a square matrix whose transpose equals its additive inverse $\mathbf{A}^\top = -\mathbf{A}$.

- The sum of two skew-symmetric matrices is skew-symmetric.
- A scalar multiple of a skew-symmetric matrix is skew-symmetric.
- The elements on the diagonal of a skew-symmetric matrix are zero, and therefore its trace equals zero.
- If $\mathbf{A}$ is a real skew-symmetric matrix and λ is a real eigenvalue, then $\lambda = 0$, i.e. the nonzero eigenvalues of a skew-symmetric matrix are non-real.
- If $\mathbf{A}$ is a real skew-symmetric matrix, then $\mathbb{E} + \mathbf{A}$ is invertible, where $\mathbb{E}$ is the identity matrix.
- If $\mathbf{A}$ is a skew-symmetric matrix then $\mathbf{A}^2$ is a symmetric negative semi-definite matrix $\langle \mathbf{x}, \mathbf{A}^2\mathbf{x}\rangle = \mathbf{x}^\top \mathbf{A}^2 \mathbf{x} = -(\mathbf{A}\mathbf{x})^\top \mathbf{A}\mathbf{x} = -\|\mathbf{A}\mathbf{x}\|_2^2 \leq 0$.
- Levi-Civita symbol ε_{ijk} is the completely skew-symmetric third-order tensor.
- The space of $n \times n$ skew-symmetric matrices has dimension of a vector space $\frac{1}{2}n(n-1)$.
- If $\mathbf{A}$ is a skew-symmetric matrix then $[\mathbf{A}]^2 = \mathbf{A}\mathbf{A}^\top - |\mathbf{A}|^2 \mathbb{E}$
- The real matrix $\mathbf{A}$ is skew-symmetric iff $\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle = (\mathbf{A}\mathbf{x})^\top \mathbf{y} = -\langle \mathbf{x}, \mathbf{A}\mathbf{y}\rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

To compute moment of inertia of a mass around an axis, the perpendicular vector from the mass to the axis is needed. If the axis L is defined by the unit vector $\hat{\mathbf{k}}$ through the reference point $\mathbf{R}$, then the perpendicular vector from the line L to the point $\mathbf{r}$ is given by

$$\delta\mathbf{r}_i^\perp = \delta\mathbf{r}_i - \left(\hat{\mathbf{k}} \cdot \delta\mathbf{r}_i\right)\hat{\mathbf{k}} = \left(\mathbb{E} - \hat{\mathbf{k}}\hat{\mathbf{k}}^\top\right)\delta\mathbf{r}_i \tag{148}$$

where $[\mathbb{E}]$ is the 3×3 identity matrix, and $\hat{\mathbf{k}}\hat{\mathbf{k}}^\mathrm{T}$ is the outer product matrix formed from the unit vector $\hat{\mathbf{k}}$ along the axis line L. The matrix $\mathbb{E} - \hat{\mathbf{k}}\hat{\mathbf{k}}^\mathrm{T}$ in this equation subtracts the $\hat{\mathbf{k}}-$parallel component from $\delta\mathbf{r}_i = \mathbf{r} - \mathbf{R}$ and then the perpendicular (to the line L) component is obtained.

The previous sections show that in computing the moment of inertia matrix this operator yields a similar operator using the components of the vector $\delta\mathbf{r}_i$ that is

$$\left[\mathbb{E}\,|\delta\mathbf{r}|^2 - \delta\mathbf{r}\delta\mathbf{r}^\top\right] \tag{149}$$

It is helpful to keep the following identities in mind to compare the equations that define the inertia tensor and the inertia matrix.

$$[(\mathbf{a}\times\mathbf{b})\times\mathbb{I}]\cdot\mathbf{r} = [\mathbb{I}\times(\mathbf{a}\times\mathbf{b})]\cdot\mathbf{r} = (\mathbf{a}\times\mathbf{b})\times\mathbf{r} = (\mathbf{ba}-\mathbf{ab})\cdot\mathbf{r} \tag{150}$$

where $\mathbb{I} = \hat{\mathbf{i}}\hat{\mathbf{i}}+\hat{\mathbf{j}}\hat{\mathbf{j}}+\hat{\mathbf{k}}\hat{\mathbf{k}}$ is the idemfactor (versor), which is a complete dyadic, see GIBBS and WILSON [1901]. Therefore

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{r} = (\mathbf{ba}-\mathbf{ab})\cdot\mathbf{r} \tag{151}$$

$$\mathbf{r}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{r}\cdot(\mathbf{ba}-\mathbf{ab}) \tag{152}$$

Let $[\mathbf{R}]$ be the skew symmetric matrix associated with the position vector $\mathbf{x} = (x, y, z)$ and a rotation vector $\boldsymbol{\omega}$, then the triple product, (also called Lagrange's formula) can be used to express unit moment of inertia matrix

$$-[\mathbf{R}]^2\boldsymbol{\omega} = -[\mathbf{R}]\left([\mathbf{R}]\boldsymbol{\omega}\right) = -\mathbf{x}\times(\mathbf{x}\times\boldsymbol{\omega}) = \mathbf{x}\times(\boldsymbol{\omega}\times\mathbf{x}) \tag{153}$$

$$= -\begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}^2 \boldsymbol{\omega} = \begin{bmatrix} y^2+z^2 & -xy & -xz \\ -yx & x^2+z^2 & -yz \\ -zx & -zy & x^2+y^2 \end{bmatrix}\boldsymbol{\omega} \tag{154}$$

The squared skew-symmetric matrix $[\mathbf{R}]^2$ can be proved by using the Levi-Civita symbol

$$[\mathbf{R}]^2(\mathbf{x}) = [\mathbf{S}](\mathbf{x})[\mathbf{S}](\mathbf{x}) = -[\mathbf{S}](\mathbf{x})[\mathbf{S}]^\top(\mathbf{x}) \quad (\text{due to } [\mathbf{S}]^\top = -[\mathbf{S}]) \tag{155}$$

$$= [-\epsilon_{ikm}x_m(-\epsilon_{kjn}x_n)] = [\epsilon_{ikm}\epsilon_{kjn}x_m x_n] \tag{156}$$

$$= [(\delta_{mj}\delta_{in} - \delta_{mn}\delta_{ij})x_m x_n] \tag{157}$$

$$= \left[x_i x_j - \delta_{ij}\sum_{n=1}^{3} x_n^2\right] = \mathbf{x}\mathbf{x}^\top - \mathbf{x}^\top\mathbf{x}[\mathbb{I}] \tag{158}$$

where $\mathbf{x}^\top$ is the transpose of the vector $\mathbf{x}$ and $[\mathbb{I}]$ is the 3×3 identity matrix so as to avoid confusion with the inertia matrix.

This can be viewed as another way of computing the perpendicular distance from an axis to a point, because the matrix formed by the outer product $[\mathbf{R}\mathbf{R}^\top]$ yields the identify

$$-[\mathbf{R}]^2 = |\mathbf{R}|^2\,\hat{\mathbf{e}} - \mathbf{R}\mathbf{R}^\top \tag{159}$$

$$= \begin{bmatrix} x^2+y^2+z^2 & 0 & 0 \\ 0 & x^2+y^2+z^2 & 0 \\ 0 & 0 & x^2+y^2+z^2 \end{bmatrix} - \begin{bmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{bmatrix} \tag{160}$$

Also notice, that

$$|\mathbf{R}|^2 = \mathbf{R}\cdot\mathbf{R} = \operatorname{tr}\left(\mathbf{R}\mathbf{R}^\top\right) \tag{161}$$

where tr denotes the sum of the diagonal elements of the outer product matrix, known as its trace.

Inertia matrix in different reference frames

The use of the inertia matrix in Newton's second law assumes its components are computed relative to axes parallel to the inertial frame and not relative to a body-fixed reference frame. This means that as the body moves the components of the inertia matrix change with time. In contrast, the components of the inertia matrix measured in a body-fixed frame are constant.

6.1 Vector calculus identities

$$
\begin{aligned}
\nabla\cdot(\mathbf{A}\times\mathbf{B}) &= (\nabla\times\mathbf{A})\cdot\mathbf{B}-\mathbf{A}\cdot(\nabla\times\mathbf{B}) && (162)\\
\nabla\times(\mathbf{A}\times\mathbf{B}) &= \mathbf{A}(\nabla\cdot\mathbf{B})-\mathbf{B}(\nabla\cdot\mathbf{A})+(\mathbf{B}\cdot\nabla)\mathbf{A}-(\mathbf{A}\cdot\nabla)\mathbf{B} && (163)\\
&= (\nabla\cdot\mathbf{B}+\mathbf{B}\cdot\nabla)\mathbf{A}-(\nabla\cdot\mathbf{A}+\mathbf{A}\cdot\nabla)\mathbf{B} && (164)\\
&= \nabla\cdot\left(\mathbf{B}\mathbf{A}^\top\right)-\nabla\cdot\left(\mathbf{A}\mathbf{B}^\top\right) && (165)\\
&= \nabla\cdot\left(\mathbf{B}\mathbf{A}^\top-\mathbf{A}\mathbf{B}^\top\right) && (166)\\
\mathbf{A}\times(\nabla\times\mathbf{B}) &= \nabla_{\mathbf{B}}(\mathbf{A}\cdot\mathbf{B})-(\mathbf{A}\cdot\nabla)\mathbf{B} && (167)\\
&= \mathbf{A}\cdot\mathbf{J}_{\mathbf{B}}^\top-(\mathbf{A}\cdot\nabla)\mathbf{B}=\mathbf{A}\cdot\nabla\mathbf{B}-(\mathbf{A}\cdot\nabla)\mathbf{B} && (168)\\
&= \mathbf{A}\cdot\left(\mathbf{J}_{\mathbf{B}}^\top-\mathbf{J}_{\mathbf{B}}\right) && (169)\\
(\mathbf{A}\times\nabla)\times\mathbf{B} &= \mathbf{A}\cdot\nabla\mathbf{B}-\mathbf{A}(\nabla\cdot\mathbf{B}) && (170)\\
&= \mathbf{A}\times(\nabla\times\mathbf{B})+(\mathbf{A}\cdot\nabla)\mathbf{B}-\mathbf{A}(\nabla\cdot\mathbf{B}) && (171)\\
\nabla\cdot\left(\mathbf{b}\mathbf{a}^\mathsf{T}\right) &= \mathbf{a}\,(\nabla\cdot\mathbf{b})+(\mathbf{b}\cdot\nabla)\,\mathbf{a}. && (172)
\end{aligned}
$$

$$
\begin{aligned}
\nabla(\mathbf{A}\cdot\mathbf{B}) &= \underbrace{(\mathbf{A}\cdot\nabla)\mathbf{B}+\mathbf{A}\times(\nabla\times\mathbf{B})}_{=\mathbf{A}\cdot\nabla\mathbf{B}}+\underbrace{(\mathbf{B}\cdot\nabla)\mathbf{A}+\mathbf{B}\times(\nabla\times\mathbf{A})}_{=\mathbf{B}\cdot\nabla\mathbf{A}} && (173)\\
&= \mathbf{A}\cdot\mathbf{J}_{\mathbf{B}}^\top+\mathbf{B}\cdot\mathbf{J}_{\mathbf{A}}^\top=\mathbf{A}\cdot\nabla\mathbf{B}+\mathbf{B}\cdot\nabla\mathbf{A} && (174)
\end{aligned}
$$

where $\mathbf{J}_{\mathbf{A}}^\top=\nabla\mathbf{A}=(\partial A_j/\partial x_i)_{ij}=(\partial_i A_j)$ denotes the transposed Jacobian matrix of the vector field $\mathbf{A}=(A_1,\ldots,A_n)$, and in the last expression the $\cdot$ operations are understood not to act on the ∇ directions (different definition in Wiki), but to inner-product of the vector immediately after ∇. Alternatively, using Feynman subscript notation,

$$\nabla(\mathbf{A}\cdot\mathbf{B})=\nabla_{\mathbf{A}}(\mathbf{A}\cdot\mathbf{B})+\nabla_{\mathbf{B}}(\mathbf{A}\cdot\mathbf{B}). \tag{175}$$

Note the difference:

$$\mathbf{A}\cdot\nabla\mathbf{B} = \nabla_{\mathbf{B}}(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{J}_{\mathbf{B}}^{\top} = A_j\left(\frac{\partial B_j}{\partial x_i}\right) = \sum_j A_j\partial_i B_j \tag{176}$$

$$(\mathbf{A}\cdot\nabla)\mathbf{B} = \left(A_j\frac{\partial}{\partial x_j}\right)B_i = \mathbf{A}\cdot\mathbf{J}_{\mathbf{B}} = A_j\left(\frac{\partial B_i}{\partial x_j}\right) = \sum_j A_j\partial_j B_i \tag{177}$$

$$\frac{1}{2}\nabla\left(\mathbf{A}\cdot\mathbf{A}\right) = \mathbf{A}\cdot\mathbf{J}_{\mathbf{A}}^{\top} = \mathbf{A}\cdot\nabla\mathbf{A} = (\mathbf{A}\cdot\nabla)\mathbf{A} + \mathbf{A}\times(\nabla\times\mathbf{A}). \tag{178}$$

7 Extension of Stokes' Theorem and its Applications

7.1 General Stokes' Theorem

Stokes' theorem (full name: Kelvin–Stokes theorem after Lord Kelvin and George Stokes), is the fundamental theorem for curls or the curl theorem in vector calculus on $\mathbb{R}^3$. Let Σ be a smooth oriented surface in $\mathbb{R}^3$ with boundary $\partial\Sigma$ and $\hat{\mathbf{n}}$ the unit outward normal. If a vector field $\mathbf{A} = (P(x,y,z), Q(x,y,z), R(x,y,z))$ is defined and has continuous first order partial derivatives in a region containing Σ, then

$$\iint_{\Sigma}(\nabla\times\mathbf{A})\cdot\mathrm{d}\mathbf{s} = \oint_{\partial\Sigma}\mathbf{A}\cdot\mathrm{d}\boldsymbol{\ell} \tag{179}$$

More explicitly, the equality says that

$$\iint_{\Sigma}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}s_x+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}s_y+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}s_z=\oint_{\partial\Sigma}P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z \tag{180}$$

where the infinitesimal area components for the surface integral are $\mathrm{d}s_x = \mathrm{d}\mathbf{s}\cdot\hat{\mathbf{n}}_x = \mathrm{d}y\mathrm{d}z$, $\mathrm{d}s_y = \mathrm{d}\mathbf{s}\cdot\hat{\mathbf{n}}_y = \mathrm{d}x\mathrm{d}z$ and $\mathrm{d}s_z = \mathrm{d}\mathbf{s}\cdot\hat{\mathbf{n}}_z = \mathrm{d}x\mathrm{d}y$. By applying the Gauss' theorem,

$$\iiint_V(\nabla\cdot\nabla\times\mathbf{A})\mathrm{d}V = \iint_{\Sigma}(\nabla\times\mathbf{A})\cdot\mathrm{d}\mathbf{s} = \oint_{\partial\Sigma}\mathbf{A}\cdot\mathrm{d}\boldsymbol{\ell} \tag{181}$$

A definition of $\nabla\times\mathbf{A}$ which is independent of the axes $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ may be obtained by applying Stokes's theorem to an infinitesimal plane area. Consider a point P. Pass a plane through P and draw in it, concentric with P, a small circle of area $\mathrm{d}\mathbf{s}$.

$$(\nabla\times\mathbf{A})\cdot\mathrm{d}\mathbf{s} = \oint\mathbf{A}\cdot\mathrm{d}\boldsymbol{\ell} \tag{182}$$

When $\mathrm{d}\mathbf{s}$ has the same direction as $\nabla\times\mathbf{A}$ the value of the line integral will be a maximum, for the cosine of the angle between $\nabla\times\mathbf{A}$ and $\mathrm{d}\mathbf{s}$ will be equal to unity. So $\mathrm{d}\mathbf{s}/|\mathrm{d}\mathbf{s}|$ can be used to represent the direction of $\nabla \times \mathbf{A}$ but the magnitude $|\mathrm{d}\mathbf{s}|$ has to be divided from both sides in (182). For this value of $\mathrm{d}\mathbf{s}$ goes to zero,

$$\nabla \times \mathbf{A} = \lim_{\mathrm{d}\mathbf{s}\to 0}\left[\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}\mathbf{s}\cdot\mathrm{d}\mathbf{s}}\oint \mathbf{A}\cdot \mathrm{d}\boldsymbol{\ell}\right] \tag{183}$$

Hence the curl $\nabla \times \mathbf{A}$ of a vector function $\mathbf{A}$ has at each point of space the direction of the normal to that plane in which the line integral of $\mathbf{A}$ taken about a small circle concentric with the point in question is a maximum. The magnitude of the curl at the point is equal to the magnitude of that line integral of maximum value divided by the area of the circle about which it is taken.

7.2 The Different Forms of Stokes' Theorem and Variants

The following is to derive several different variants of the Stokes's Theory from the original (179) or (180). In the derivations, the notations are defined here unless specified otherwise. Let $\mathbf{r}$ be the position vector $(x\hat{\mathbf{i}}, y\hat{\mathbf{j}}, z\hat{\mathbf{k}})$, $\mathbf{c}$ an auxiliary constant vector and f, ϕ scalar functions. For compact expressions, some special notations are introduced, for instance $\mathbf{r_c}$ is short for $\mathbf{r}\times\mathbf{c}$ and the up-down brackets around the three vectors $\llbracket\mathbf{a}\ \mathbf{b}\ \mathbf{c}\rrbracket$ stands for cyclic permutation equivalence of the scalar triple product.

$$\begin{aligned}
\llbracket\mathbf{a}\ \mathbf{b}\ \mathbf{c}\rrbracket &= \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) && = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) \\
&= (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} && = (\mathbf{b}\times\mathbf{c})\cdot\mathbf{a} = (\mathbf{c}\times\mathbf{a})\cdot\mathbf{b} \qquad (184)\\
&= -\llbracket\mathbf{a}\ \mathbf{c}\ \mathbf{b}\rrbracket && = -\llbracket\mathbf{b}\ \mathbf{a}\ \mathbf{c}\rrbracket = -\llbracket\mathbf{c}\ \mathbf{b}\ \mathbf{a}\rrbracket \qquad (185)
\end{aligned}$$

7.2.1 Original but in different equivalents

The original form (181) with $\mathbf{A} = \mathbf{b}$ can be written by the cyclic permutation notation as,

$$\oint_\ell \mathbf{b}\cdot\mathrm{d}\boldsymbol{\ell} = \iint_s (\nabla\times\mathbf{b})\cdot\mathrm{d}\mathbf{s} = \iiint_V (\nabla\cdot\nabla\times\mathbf{b})\mathrm{d}V \tag{186}$$

$$= \iint_s [\mathbf{n}\cdot(\nabla\times\mathbf{b})]\mathrm{d}s \tag{187}$$

$$= \iint_s \llbracket\mathbf{n}\ \nabla_{\mathbf{b}}\ \mathbf{b}\rrbracket\mathrm{d}s \tag{188}$$

7.2.2 If $\mathbf{A} = f\mathbf{c}$, a scalar function multiplied by constant vector

In this case, the scalar function f can dependent on space (can also on time), but the vector $\mathbf{c}$ is independent of space but no restriction on time.

$$\oint_\ell f\mathbf{c}\cdot\mathrm{d}\boldsymbol{\ell} = \mathbf{c}\cdot\oint_\ell f\mathrm{d}\boldsymbol{\ell} = \iint_s (\nabla\times f\mathbf{c})\cdot\mathrm{d}\mathbf{s} = \iint_s (\nabla f\times\mathbf{c})\cdot\mathrm{d}\mathbf{s} = \mathbf{c}\cdot\iint_s \mathbf{n}\times\nabla f\mathrm{d}s$$

$$\oint_\ell f\mathrm{d}\boldsymbol{\ell} = \iint_s \mathrm{d}\mathbf{s}\times\nabla f = \iint_s \mathbf{n}\times\nabla f\mathrm{d}s \tag{189}$$

7.2.3 If $\mathbf{A} = \nabla(f\phi)$, a gradient of the product of two scalar functions

Then we have

$$\mathbf{A} = \nabla(f\phi) = f\nabla\phi + \phi\nabla f \tag{190}$$

$$\oint_\ell \nabla(f\phi)\cdot \mathrm{d}\boldsymbol{\ell} = \iint_s \nabla\times\nabla(f\phi)\cdot\mathrm{d}\mathbf{s} = 0 \tag{191}$$

$$\begin{aligned}\oint_\ell f\nabla\phi\cdot\mathrm{d}\boldsymbol{\ell} = \oint_\ell f\mathrm{d}\phi = -\oint_\ell \phi\mathrm{d}f \qquad &(192)\\ = \iint_s \nabla f\times\nabla\phi\cdot\mathrm{d}\mathbf{s} \qquad &(193)\\ = \iint_s [\![\nabla f\ \nabla\phi\ \mathbf{n}]\!]\mathrm{d}s \qquad &(194)\end{aligned}$$

7.2.4 If $\mathbf{A} = f\mathbf{a}$, a scalar function scaled vector

In this case, both the scalar function f and the vector $\mathbf{a}$ are functions of space, as contrast to Section7.2.2.

$$\begin{aligned}\oint_\ell f\mathbf{a}\cdot\mathrm{d}\boldsymbol{\ell} = \iint_s(\nabla\times f\mathbf{a})\cdot\mathrm{d}\mathbf{s} = \iint_s(\nabla f\times\mathbf{a} + f\nabla\times\mathbf{a})\cdot\mathrm{d}\mathbf{s} \qquad &(195)\\ = \iint_s [\![\nabla f\ \mathbf{a}\ \mathbf{n}]\!] + [\![f\nabla_{\mathbf{a}}\ \mathbf{a}\ \mathbf{n}]\!]\mathrm{d}s \qquad &(196)\end{aligned}$$

7.2.5 If $\mathbf{A} = \mathbf{a}\times\mathbf{c}$, a cross product of scalar function scaled vector and a constant vector

Note $\mathbf{c}$ denotes a constant vector, i.e., independent on space.

$$\begin{aligned}\oint_\ell(\mathbf{a}\times\mathbf{c})\cdot\mathrm{d}\boldsymbol{\ell} = \mathbf{c}\cdot\oint_\ell \mathrm{d}\boldsymbol{\ell}\times\mathbf{a} \qquad &(197)\\ = \iint_s[\nabla\times(\mathbf{a}\times\mathbf{c})]\cdot\mathbf{n}\mathrm{d}s \qquad &(198)\\ = \mathbf{c}\cdot\iint_s(\mathbf{n}\times\nabla)\times\mathbf{a}\mathrm{d}s \qquad &(199)\\ = \mathbf{c}\cdot\iint_s \varepsilon_{ijk}(\mathbf{n}\times\nabla)_j a_k\mathrm{d}s \qquad &(200)\\ = \mathbf{c}\cdot\iint_s \varepsilon_{ijk}\varepsilon_{jlm}n_\ell\partial_m a_k\mathrm{d}s \qquad &(201)\\ = \mathbf{c}\cdot\iint_s n_k\partial_i a_k - n_i\partial_k a_k\mathrm{d}s \qquad &(202)\end{aligned}$$

$$\begin{aligned}\oint_\ell \mathrm{d}\boldsymbol{\ell}\times\mathbf{a} = \iint_s(\mathbf{n}\times\nabla)\times\mathbf{a}\mathrm{d}s = \iint_s \mathbf{n}\cdot(\nabla_{\mathbf{a}}\mathbf{a}) - (\nabla\cdot\mathbf{a})\mathbf{n}\mathrm{d}s \qquad &(203)\\ = \iint_s \mathbf{n}\times(\nabla\times\mathbf{a}) + (\mathbf{n}\cdot\nabla)\mathbf{a} - (\nabla\cdot\mathbf{a})\mathbf{n}\mathrm{d}s \qquad &(204)\end{aligned}$$

7.2.6 If $\mathbf{A} = f\mathbf{a} \times \mathbf{c}$, a cross product of scalar function scaled vector and a constant vector

$$\oint_\ell (f\mathbf{a} \times \mathbf{c}) \cdot \mathrm{d}\boldsymbol{\ell} = \mathbf{c} \cdot \oint_\ell \mathrm{d}\boldsymbol{\ell} \times f\mathbf{a} \tag{205}$$

$$= \mathbf{c} \cdot \iint_s (\mathbf{n} \times \nabla) \times f\mathbf{a}\mathrm{d}s \tag{206}$$

$$= \mathbf{c} \cdot \iint_s \varepsilon_{ijk}(\mathbf{n} \times \nabla)_j (fa_k)\mathrm{d}s \tag{207}$$

$$= \mathbf{c} \cdot \iint_s \varepsilon_{ijk}\varepsilon_{jlm} n_\ell \partial_m (fa_k)\mathrm{d}s \tag{208}$$

$$= \mathbf{c} \cdot \iint_s n_k f \partial_i a_k + n_k a_k \partial_i f - n_i f \partial_k a_k - n_i a_k \partial_k f \mathrm{d}s \tag{209}$$

$$\oint_\ell \mathrm{d}\boldsymbol{\ell} \times f\mathbf{a} = \iint_s (\mathbf{n} \times \nabla) \times f\mathbf{a}\mathrm{d}s \tag{210}$$

$$= \iint_s f\mathbf{n} \times (\nabla \times \mathbf{a}) + f(\mathbf{n} \cdot \nabla)\mathbf{a} + (\mathbf{n} \cdot \mathbf{a})\nabla f - (\nabla f \cdot \mathbf{a})\mathbf{n} - (\nabla \cdot \mathbf{a}) f\mathbf{n}\mathrm{d}s \tag{211}$$

7.2.7 If $\mathbf{A} = f\nabla\phi \times \mathbf{c}$, a cross product of scalar function scaled gradient and a constant vector

Here, we introduce a potential function ϕ which satisfies Laplace equation, i.e. $\nabla^2\phi = 0$, therefore

$$\oint_\ell (f\nabla\phi \times \mathbf{c}) \cdot \mathrm{d}\boldsymbol{\ell} = \mathbf{c} \cdot \oint_\ell \mathrm{d}\boldsymbol{\ell} \times f\nabla\phi \tag{212}$$

neglecting the inner product by the constant vector $\mathbf{c}$, we obtain

$$\oint_\ell \mathrm{d}\boldsymbol{\ell} \times f\nabla\phi = \iint_s \left[f\frac{\partial}{\partial n}\nabla\phi + \frac{\partial\phi}{\partial n}\nabla f - (\nabla f \cdot \nabla\phi)\mathbf{n} \right] \mathrm{d}s \tag{213}$$

7.2.8 If $\mathbf{A} = \mathbf{a} \times (\mathbf{r} \times \mathbf{c})$, a vector triple product

Note $\mathbf{c}$ denotes a constant vector.

$$\oint_\ell \mathbf{a} \times (\mathbf{r} \times \mathbf{c}) \cdot \mathrm{d}\boldsymbol{\ell} = \oint_\ell (\mathrm{d}\boldsymbol{\ell} \times \mathbf{a}) \cdot (\mathbf{r} \times \mathbf{c}) = \oint_\ell [(\mathrm{d}\boldsymbol{\ell} \times \mathbf{a}) \times \mathbf{r}] \cdot \mathbf{c} \tag{214}$$

$$= \iint_s \{\nabla \times [\mathbf{a} \times (\mathbf{r} \times \mathbf{c})]\} \cdot \mathbf{n}\mathrm{d}s \xRightarrow{\mathbf{r_c} = \mathbf{r} \times \mathbf{c}} \iint_s \{\nabla \times [\mathbf{a} \times \mathbf{r_c})]\} \cdot \mathbf{n}\mathrm{d}s \tag{215}$$

$$= \iint_s \{(\mathbf{r_c} \cdot \nabla)\mathbf{a} - \mathbf{r_c}\nabla \cdot \mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{r_c} + \mathbf{a}\nabla \cdot \mathbf{r_c}\} \cdot \mathbf{n}\mathrm{d}s \tag{216}$$

$$= \iint_s \left\{ \underbrace{(\mathbf{r_c} \cdot \nabla)\mathbf{a}}_{(1)} \underbrace{-(\mathbf{a} \cdot \nabla)\mathbf{r_c}}_{(2)} + \underbrace{\mathbf{a}\nabla \cdot \mathbf{r_c}}_{(3)} \underbrace{-\mathbf{r_c}\nabla \cdot \mathbf{a}}_{(4)} \right\} \cdot \mathbf{n}\mathrm{d}s \tag{217}$$

$$
\begin{aligned}
\textcircled{1}\cdot\mathbf{n} &= \{[(\mathbf{r}\times\mathbf{c})\cdot\nabla]\mathbf{a}\}\cdot\mathbf{n} && (218)\\
&= \llbracket \nabla\times\mathbf{a}\ \mathbf{r}\times\mathbf{c}\ \mathbf{n} \rrbracket + \llbracket (\mathbf{n}\cdot\nabla)\mathbf{a}\ \mathbf{r}\ \mathbf{c} \rrbracket && (219)\\
&= \llbracket \mathbf{n}\times(\nabla\times\mathbf{a})\ \mathbf{r}\ \mathbf{c} \rrbracket + \llbracket (\mathbf{n}\cdot\nabla)\mathbf{a}\ \mathbf{r}\ \mathbf{c} \rrbracket && (220)\\
&= \Big\{ \underline{[\mathbf{n}\times(\nabla\times\mathbf{a})]\times\mathbf{r}} + [(\mathbf{n}\cdot\nabla)\mathbf{a}]\times\mathbf{r} \Big\}\cdot\mathbf{c} && (221)\\
&= \Big\{ \underline{(\mathbf{r}\cdot\mathbf{n})\nabla\times\mathbf{a} - [\mathbf{r}\cdot(\nabla\times\mathbf{a})]\mathbf{n}} + [(\mathbf{n}\cdot\nabla)\mathbf{a}]\times\mathbf{r} \Big\}\cdot\mathbf{c} && (222)\\
\textcircled{2}\cdot\mathbf{n} &= -\{(\mathbf{a}\cdot\nabla)(\mathbf{r}\times\mathbf{c})\}\cdot\mathbf{n} && (223)\\
&= -a_m(\partial_m\varepsilon_{ijk}r_jc_k)n_i = -\varepsilon_{ijk}a_mn_i(c_k\partial_mr_j + r_j\partial_mc_k) && (224)\\
&= -\varepsilon_{ijk}a_jc_kn_i && (225)\\
&= -\llbracket \mathbf{a}\ \mathbf{c}\ \mathbf{n} \rrbracket = (\mathbf{a}\times\mathbf{n})\cdot\mathbf{c} && (226)\\
\textcircled{3}\cdot\mathbf{n} &= \{\mathbf{a}\nabla\cdot(\mathbf{r}\times\mathbf{c})\}\cdot\mathbf{n} && (227)\\
&= a_m[\varepsilon_{ijk}(c_k\partial_ir_j + r_j\partial_ic_k)]n_m && (228)\\
&= \mathbf{a}[\mathbf{c}\cdot(\nabla\times\mathbf{r}) - \mathbf{r}\cdot(\nabla\times\mathbf{c})]\cdot\mathbf{n} && (229)\\
&= 0 && (230)\\
\textcircled{4}\cdot\mathbf{n} &= (\nabla\cdot\mathbf{a})(\mathbf{r}\times\mathbf{c})\cdot\mathbf{n} && (231)\\
&= (\nabla\cdot\mathbf{a})\llbracket \mathbf{r}\ \mathbf{n}\ \mathbf{c} \rrbracket && (232)\\
&= \{(\nabla\cdot\mathbf{a})(\mathbf{r}\times\mathbf{n})\}\cdot\mathbf{c} && (233)
\end{aligned}
$$

$$
\begin{aligned}
\int_\ell (\mathrm{d}\boldsymbol{\ell}\times\mathbf{a})\times\mathbf{r} = \iint_s \{ & \underbrace{(\mathbf{r}\cdot\mathbf{n})\nabla\times\mathbf{a} - [\mathbf{r}\cdot(\nabla\times\mathbf{a})]\mathbf{n}}_{\mathbf{n}\times(\nabla\times\mathbf{a})\times\mathbf{r}} + [(\mathbf{n}\cdot\nabla)\mathbf{a}]\times\mathbf{r} + \\
& \mathbf{a}\times\mathbf{n} + (\nabla\cdot\mathbf{a})(\mathbf{r}\times\mathbf{n})\}\mathrm{d}s \qquad (234)
\end{aligned}
$$

Using above derivations, it is found that some literature contains incorrect use of Stokes formula. For instance the popularly cited X-B [2007] where the middle field reformulation for 2nd order wave loads used incorrectly Stokes's formulae, the corrected one might be shown below.

$$
\mathbf{r}\times[\mathbf{n}\cdot(\nabla\mathbf{a})] = \mathbf{r}\times[\mathbf{n}\times(\nabla\times\mathbf{a}) + (\mathbf{n}\cdot\nabla)\mathbf{a}] \qquad (235)
$$

$$
\begin{aligned}
\int_\ell (\mathrm{d}\boldsymbol{\ell}\times\mathbf{a})\times\mathbf{r} = \iint_s & \{(\nabla\cdot\mathbf{a})(\mathbf{r}\times\mathbf{n}) + \mathbf{a}\times\mathbf{n} - \mathbf{r}\times[(\mathbf{n}\cdot\nabla)\mathbf{a}] \\
& -[\mathbf{r}\cdot(\nabla\times\mathbf{a})]\mathbf{n} + (\mathbf{r}\cdot\mathbf{n})\nabla\times\mathbf{a}\}\,\mathrm{d}s \qquad (236)
\end{aligned}
$$

$$\oint_c \mathbf{t} \times \mathbf{n} f \mathrm{d}s = \iint_S (\mathbf{n} \times \nabla) \times \mathbf{n} f \mathrm{d}s \tag{237}$$

$$= \iint_S \varepsilon_{ijk} (\mathbf{n} \times \nabla)_j (f n_k) \mathrm{d}s \tag{238}$$

$$= \iint_S \varepsilon_{ijk} \varepsilon_{jlm} n_\ell \partial_m (f n_k) \mathrm{d}s \tag{239}$$

$$= \iint_S n_k f \partial_i n_k + n_k n_k \partial_i f - n_i f \partial_k n_k - n_i n_k \partial_k f \mathrm{d}s \tag{240}$$

$$= \iint_S f \mathbf{n} \times (\nabla \times \mathbf{n}) + f(\mathbf{n} \cdot \nabla) \mathbf{n} + (\mathbf{n} \cdot \mathbf{n}) \nabla f - (\nabla f \cdot \mathbf{n}) \mathbf{n} - (\nabla \cdot \mathbf{n}) f \mathbf{n} \mathrm{d}s \tag{241}$$

7.3 Common Formulas

Given that $\mathbf{r} = x_i \hat{\mathbf{e}}_i$ is the position vector, $\mathbf{a}$ is a constant vector, and $f(r)$ is a scalar function of $r = |\mathbf{r}|$, show that $\nabla \cdot \mathbf{r} = 3$, $\nabla \times \mathbf{r} = \mathbf{0}$, $\nabla(\mathbf{a} \cdot \mathbf{r}) = \mathbf{a}$,

$$\nabla f = \frac{\mathrm{d}f}{\mathrm{d}r} \nabla r = \frac{\mathrm{d}f}{\mathrm{d}r} \frac{\mathbf{r}}{r} = \frac{\mathrm{d}f}{\mathrm{d}r} \hat{\mathbf{r}} \tag{242}$$

If $f = r = |\mathbf{r}|$, then

$$\nabla r = \frac{\partial r}{\partial x_i} \hat{\mathbf{e}}_i = \frac{x_i \hat{\mathbf{e}}_i}{r} = \frac{\mathbf{r}}{r} = \hat{\mathbf{r}} \tag{243}$$

$$\nabla \frac{1}{r} = -\frac{\nabla r}{r^2} = -\frac{\mathbf{r}}{r^3} \tag{244}$$

$$\nabla \frac{1}{r^n} = -\frac{n \nabla r}{r^{n+1}} = -\frac{n \mathbf{r}}{r^{n+2}} \tag{245}$$

$$\nabla r^n = n r^{n-1} \nabla r = n r^{n-1} \hat{\mathbf{r}} \tag{246}$$

If $f = \mathbf{r}$, then

$$\nabla \mathbf{r} = \sum_{i=1}^{3} \nabla (x_i \hat{\mathbf{e}}_i) \tag{247}$$

$$= \sum_{i=1}^{3} \Big((\nabla x_i) \cdot \hat{\mathbf{e}}_i + x_i \nabla \hat{\mathbf{e}}_i) \Big) \qquad \text{(unit vector is a constant)} \tag{248}$$

$$= \sum_{i=1}^{3} \frac{\mathbf{r}}{r} \frac{\partial x_i}{\partial r} \cdot \hat{\mathbf{e}}_i \qquad \text{(Substitute } f = x_i \text{ in 242)} \tag{249}$$

$$= \sum_{i=1}^{3} \frac{\mathbf{r}}{r} \frac{\cdot \hat{\mathbf{e}}_i}{\frac{\partial r}{\partial x_i}} \tag{250}$$

$$= \sum_{i=1}^{3} \frac{\mathbf{r}}{r} \frac{\cdot \hat{\mathbf{e}}_i}{\frac{x_i}{r}} \tag{251}$$

$$= \sum_{i=1}^{3} \frac{\mathbf{r} \cdot \hat{\mathbf{e}}_i}{x_i} \tag{252}$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{x_j \hat{\mathbf{e}}_j \cdot \hat{\mathbf{e}}_i}{x_i} \tag{253}$$

$$= \mathbb{E}_{3\times 3} \tag{254}$$

$$\nabla \cdot \mathbf{r} = \frac{\hat{\mathbf{e}}_i \partial}{\partial x_i} \cdot x_j \hat{\mathbf{e}}_j = \hat{\mathbf{e}}_i \partial_i \cdot x_j \hat{\mathbf{e}}_j = 3 \tag{255}$$

$$= \frac{3}{r^2} - \mathbf{r} \cdot \frac{2\mathbf{r}}{r^4} = \frac{1}{r^2} = \frac{1}{r^2} \tag{256}$$

$$\nabla \cdot \frac{\hat{\mathbf{r}}}{r} = \nabla \cdot \frac{\mathbf{r}}{r^2} = \frac{\mathbf{1}}{r^2} \nabla \cdot \mathbf{r} + \mathbf{r} \cdot \nabla \frac{1}{r^2} \tag{257}$$

$$\nabla \cdot \hat{\mathbf{r}} = \frac{2}{r} \tag{258}$$

$$\nabla \cdot \frac{\hat{\mathbf{r}}}{r^2} = \nabla \cdot \frac{\mathbf{r}}{r^3} = 4\pi \delta^3 (\mathbf{r}^3) \tag{259}$$

$$\nabla \cdot \frac{\hat{\mathbf{r}}}{r^n} = \nabla \cdot \frac{\mathbf{r}}{r^{n+1}} = \frac{2-n}{r^{n+1}} \tag{260}$$

$$\nabla \times \hat{\mathbf{r}} r^n = r^n \nabla \times \hat{\mathbf{r}} + \nabla r^n \times \hat{\mathbf{r}} \tag{261}$$

$$= r^n \varepsilon_{ijk} \partial_j \hat{\mathbf{e}}_k + n r^{n-1} \nabla r \times \hat{\mathbf{r}} \tag{262}$$

$$= r^n \varepsilon_{ijk} \delta_{jk} + n r^{n-1} \hat{\mathbf{r}} \times \hat{\mathbf{r}} = 0 \tag{263}$$

$$
\begin{aligned}
\nabla(\mathbf{a}\cdot\mathbf{r}) &= \partial_i(a_j r_j) = a_j\partial_i r_j + r_j\partial_i a_j \\
&= \nabla_r(\mathbf{r}\cdot\mathbf{a}) + \cancelto{0}{\nabla_a(\mathbf{a}\cdot\mathbf{r})} \qquad (\mathbf{a}\text{ is constant}) && (264)\\
&= \mathbb{E}_{3\times3}\cdot\mathbf{a} && (265)\\
&= \mathbf{a} && (266)
\end{aligned}
$$

First, consider the triple product,

$$
\begin{aligned}
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= a_i[\mathbf{b}\times\mathbf{c}]_i = \varepsilon_{ijk}a_i b_j c_k \\
&= b_j\varepsilon_{jki}c_k a_i = b_j[\mathbf{c}\times\mathbf{a}]_j \\
&= c_k\varepsilon_{kij}a_i b_j = c_k[\mathbf{a}\times\mathbf{b}]_k \\
&= \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) && (267)
\end{aligned}
$$

Next, a double cross product can be expanded as

$$
\begin{aligned}
[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i &= \varepsilon_{ijk}a_j[\mathbf{b}\times\mathbf{c}]_k = \varepsilon_{ijk}a_j\varepsilon_{klm}b_\ell c_m \\
&= \varepsilon_{ijk}\varepsilon_{klm}a_j b_\ell c_m \\
&= (\delta_i^\ell\delta_j^m - \delta_i^m\delta_j^\ell)a_j b_\ell c_m \\
&= \delta_i^\ell\delta_j^m a_j b_\ell c_m - \delta_i^m\delta_j^\ell a_j b_\ell c_m \\
&= (\delta_i^\ell b_\ell)(\delta_j^m a_j c_m - (\delta_i^m c_m)(\delta_j^\ell a_j b_\ell) \\
&= b_i(a_j c_j) - c_i(a_j b_j) \\
&= [\mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})]_i && (268)
\end{aligned}
$$

The expansion of the curl of a cross product as

$$
\begin{aligned}
[\nabla\times(\mathbf{a}\times\mathbf{b})]_i &= \varepsilon_{ijk}\partial_j[\mathbf{a}\times\mathbf{b}]_k = \varepsilon_{ijk}\partial_j(\varepsilon_{klm}a_\ell b_m) \\
&= \varepsilon_{ijk}\varepsilon_{klm}(b_m\partial_j a_\ell + a_\ell\partial_j b_m) \\
&= (\delta_i^\ell\delta_j^m - \delta_i^m\delta_j^\ell)(b_m\partial_j a_\ell + a_\ell\partial_j b_m) \\
&= \delta_i^\ell\delta_j^m b_m\partial_j a_\ell + \delta_i^\ell\delta_j^m a_\ell\partial_j b_m - \delta_i^m\delta_j^\ell b_m\partial_j a_\ell - \delta_i^m\delta_j^\ell a_\ell\partial_j b_m \\
&= b_j\partial_j a_i + a_i\partial_j b_j - b_i\partial_j a_j - a_j\partial_j b_i \\
&= [(\mathbf{b}\cdot\nabla)\mathbf{a}]_i + [\mathbf{a}(\nabla\cdot\mathbf{b})]_i - [\mathbf{b}(\nabla\cdot\mathbf{a})]_i - [(\mathbf{a}\cdot\nabla)\mathbf{b}]_i \\
&= [(\mathbf{b}\cdot\nabla)\mathbf{a} + \mathbf{a}(\nabla\cdot\mathbf{b}) - \mathbf{b}(\nabla\cdot\mathbf{a}) - (\mathbf{a}\cdot\nabla)\mathbf{b}]_i && (269)
\end{aligned}
$$

The expansion of the gradient of a dot product can be

$$
\begin{aligned}
\nabla(\mathbf{a}\cdot\mathbf{b}) &= \partial_i(a_j b_j) = a_j\partial_i b_j + b_j\partial_i a_j \\
&= \nabla_b(\mathbf{b}\cdot\mathbf{a}) + \nabla_a(\mathbf{a}\cdot\mathbf{b}) \qquad \text{(Feynman notation)} && (270)\\
&= \nabla(\mathbf{b}\cdot\mathbf{a})_a + \nabla(\mathbf{a}\cdot\mathbf{b})_b \qquad \text{(GIBBS and WILSON [1901])} && (271)\\
&= a_j(\partial_i b_j - \partial_j b_i) + b_j(\partial_i a_j - \partial_j a_i) + a_j\partial_j b_i + b_j\partial_j a_i \\
&= \mathbf{a}\times(\nabla\times\mathbf{b}) + \mathbf{b}\times(\nabla\times\mathbf{a}) + (\mathbf{a}\cdot\nabla)\mathbf{b} + (\mathbf{b}\cdot\nabla)\mathbf{a} && (272)
\end{aligned}
$$

The following derivation is one alternative by double cross products.

$$
\begin{aligned}
[\mathbf{a} \times (\nabla \times \mathbf{b})]_i &= \varepsilon_{ijk} a_j [\nabla \times \mathbf{b}]_k = \varepsilon_{ijk} a_j \varepsilon_{klm} \partial_\ell b_m = \varepsilon_{ijk} \varepsilon_{klm} a_j \partial_\ell b_m \\
&= (\delta_i^\ell \delta_j^m - \delta_i^m \delta_j^\ell) a_j \partial_\ell b_m \\
&= \delta_i^\ell \delta_j^m a_j \partial_\ell b_m - \delta_i^m \delta_j^\ell a_j \partial_\ell b_m \\
&= a_j \partial_i b_j - a_j \partial_j b_i
\end{aligned} \tag{273}
$$

$$
\begin{aligned}
[\mathbf{b} \times (\nabla \times \mathbf{a})]_i &= \varepsilon_{ijk} b_j [\nabla \times \mathbf{a}]_k = \varepsilon_{ijk} b_j \varepsilon_{klm} \partial_\ell a_m = \varepsilon_{ijk} \varepsilon_{klm} b_j \partial_\ell a_m \\
&= (\delta_i^\ell \delta_j^m - \delta_i^m \delta_j^\ell) b_j \partial_\ell a_m \\
&= \delta_i^\ell \delta_j^m b_j \partial_\ell a_m - \delta_i^m \delta_j^\ell a_j \partial_\ell a_m \\
&= b_j \partial_i a_j - b_j \partial_j a_i
\end{aligned} \tag{274}
$$

The sum of the above two equations eqs.(273)+(274) can be seen as

$$
\begin{aligned}
[\mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a})]_i &= a_j \partial_i b_j + b_j \partial_i a_j - a_j \partial_j b_i - b_j \partial_j a_i \\
&= \partial_i (a_j b_j) - a_j \partial_j b_i - b_j \partial_j a_i \\
&= [\nabla(\mathbf{a} \cdot \mathbf{b}) - (\mathbf{a} \cdot \nabla)\mathbf{b} - (\mathbf{b} \cdot \nabla)\mathbf{a}]_i
\end{aligned} \tag{275}
$$

Rearrangement of (275) will result in (272). This is a useful identity in the derivation where material derivative is involved, such as Navier-Stokes or vorticity equations. A special case if $\mathbf{v} = \mathbf{a} = \mathbf{b}$ will be

$$
\mathbf{v} \cdot \nabla \mathbf{v} = \frac{1}{2} \nabla (\mathbf{v} \cdot \mathbf{v}) - \mathbf{v} \times (\nabla \times \mathbf{v}). \tag{276}
$$

7.3.1 Variant of Green's Theorem

The divergence theorem is applied to the vector $\mathbf{P} \times (\nabla \times \mathbf{Q})$, giving

$$
\iiint_v \nabla \cdot [\mathbf{P} \times (\nabla \times \mathbf{Q})] \mathrm{d}v = \oiint_s [\mathbf{P} \times (\nabla \times \mathbf{Q})] \cdot \mathbf{n} \mathrm{d}s \tag{277}
$$

where n is a unit normal vector directed outward from s. Upon expanding the integrand of the volume integral a vector analog of Green's first identity is obtained,

$$
\iiint_v \{(\nabla \times \mathbf{P}) \cdot (\nabla \times \mathbf{Q}) - \mathbf{P} \cdot [\nabla \times (\nabla \times \mathbf{Q})]\} \mathrm{d}v = \oiint_s [\mathbf{P} \times (\nabla \times \mathbf{Q})] \cdot \mathbf{n} \mathrm{d}s \tag{278}
$$

The vector analog of Green's second identity ("Green's theorem") is obtained by reversing the roles of P and Q in (278) and subtracting one expression from the other.

$$
\iiint_v \{\mathbf{Q} \cdot [\nabla \times (\nabla \times \mathbf{P})] - \mathbf{P} \cdot [\nabla \times (\nabla \times \mathbf{Q})] \mathrm{d}v = \oiint_s [\mathbf{P} \times (\nabla \times \mathbf{Q}) - \mathbf{Q} \times (\nabla \times \mathbf{P})] \cdot \mathbf{n} \mathrm{d}s \tag{279}
$$

The first identity (278) is the logical basis for uniqueness proofs in connection with vector fields. It will be noted that if one place $\mathbf{P} = \mathbf{Q} = \mathbf{E}$, (278) proves to be identical with Poynting's theorem.

A novel way to deal with vector differentiation will be just copied from Feynman Lectures on Physics. Look at the divergence of a curl of two vectors.

$$\nabla \cdot (\mathbf{B} \times \hat{\mathbf{e}}) \tag{280}$$

Rather than working out all the components of (280), you will be shown a trick that is very useful for this kind of problem. It is a trick that allows you to use all the rules of vector algebra on expressions with the ∇ operator, without getting into trouble. The trick is to throw out - for a while at least - the rule of the calculus notation about what the derivative operator works on. You see, ordinarily, the order of terms is used for two separate purposes. One is for calculus: $f(d/dx)g$ is not the same as $g(d/dx)f$; and the other is for vectors: $\mathbf{a}\times\mathbf{b}$ is different from $\mathbf{b}\times\mathbf{a}$. We can, if we want, choose to abandon momentarily the calculus rule. Instead of saying that a derivative operates on everything to the right, we make a new rule that doesn't depend on the order in which terms are written down. Then we can juggle terms around without worrying.

Here is our new convention: we show, by a subscript, what a differential operator works on; the order has no meaning. Suppose we let the operator D stand for $\partial/\partial x$. Then D_f means that only the derivative of the variable quantity f is taken. Then

$$D_f f = \frac{\partial}{\partial x} f. \tag{281}$$

But if we have $D_f fg$, it means

$$D_f fg = (\frac{\partial f}{\partial x})g. \tag{282}$$

But notice now that according to our new rule, $fD_f g$ means the same thing. We can write the same thing any which way:

$$D_f fg = gD_f f = fD_f g = fgD_f. \tag{283}$$

You see, the D_f can even come after everything. (It's surprising that such a handy notation is never taught in books on mathematics or physics.)

7.4 Inner product space

The notation $\langle x|y\rangle$ which is commonly used in physics will be assumed to be antilinear (conjugate-linear) map in the *first* argument while $\langle x,\, y\rangle$ which is commonly used in mathematics, will be assumed to be antilinear its the *second* argument. The relationship is $\langle x,\, y\rangle = \langle y\,|\,x\rangle$ for all $x, y \in H$. The real part of any inner product (no matter which argument is antilinear and no matter if it is real or complex) is a symmetric bilinear map that for any $x, y \in H$

$$\begin{aligned}\Re(x,y) &:= \Re\langle y|ix\rangle = \Re\langle ix, y\rangle \qquad (284)\\ &= \frac{1}{2}\left(\langle y, ix\rangle + \langle ix, y\rangle\right)\\ &= \frac{1}{2i}\left(-\langle iy|ix\rangle + \langle ix|iy\rangle\right) = \frac{1}{2i}\left(\langle iy, ix\rangle - \langle ix, iy\rangle\right)\\ &= \frac{1}{2i}\left(-i\langle iy|x\rangle - i\langle x, iy\rangle\right) = \frac{1}{2i}\left(-i\langle iy, x\rangle - i\langle x, iy\rangle\right)\\ &= -\frac{1}{2}\left(\langle iy|x\rangle + \langle x|iy\rangle\right) = \frac{1}{2i}\left(\langle x|y\rangle - \langle y|x\rangle\right)\\ &= -\Re\left\langle x \mid iy\right\rangle = \Im\left\langle x \mid y\right\rangle = \Im\left\langle y, x\right\rangle \qquad (285)\end{aligned}$$

Thus $\Re\left\langle ix \mid y\right\rangle = -\Re\left\langle x \mid iy\right\rangle$ which says that to move a factor of $i = \sqrt{-1}$ to the other argument, introduce a negative sign. This is due to the fact that there is one negating operation of anitilinear mapping. So if $z = a + bi$, the following will be obviously applicable

$$\Re(z) = a = \frac{1}{2}(z + z^*) = \Im(iz) \qquad (286)$$

$$\Im(z) = b = \frac{1}{2i}(z - z^*) = \Re(-iz). \qquad (287)$$

Equivalently, the properties of inner product, i.e., conjugate symmetric or Hermitian $\langle \alpha x|\beta y\rangle = \alpha^*\beta\langle x|y\rangle$ can be used to derive the Polarization identity that is to express the inner product of any two vectors in terms of the norm of a normed vector space.

$$\langle x|y\rangle + \langle y|x\rangle = 2\Re\langle x|y\rangle \qquad (288)$$

$$\langle x|y\rangle - \langle y|x\rangle = 2\Im\langle x|y\rangle \qquad (289)$$

Firstly, it is almost trivial to know that the inner product (or called norm, distance) of the sum of x and $\pm y$ can be written as

$$\begin{aligned}\|x\pm y\|^2 &= \|x\|^2 + \|y\|^2 \pm \langle x|y\rangle \pm \langle y|x\rangle \qquad (290)\\ &= \|x\|^2 + \|y\|^2 \pm 2\Re\left\langle x \mid y\right\rangle \text{ for all vectors } x, y. \qquad (291)\end{aligned}$$

And by adding and subtracting of the two equations with $\pm$ signs, respectively, it is straightforward to see

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \qquad (292)$$

$$\|x+y\|^2 - \|x-y\|^2 = 4\Re\langle x|y\rangle. \qquad (293)$$

Let us look at this configuration

$$
\begin{aligned}
&-j\|x+jy\|^2+j\|x-jy\|^2\\
&=j\Big(\langle x-jy|x-jy\rangle-\langle x+jy|x+jy\rangle\Big)\\
&=j\Big(\langle x|x\rangle-\langle x|jy\rangle-\langle jy|x\rangle+\langle jy|jy\rangle-\langle x|x\rangle-\langle x|jy\rangle-\langle jy|x\rangle-\langle jy|jy\rangle\Big)\\
&=-2j\Big(\langle x|jy\rangle+\langle jy|x\rangle\Big)\\
&=-4j\Re\langle x|jy\rangle \qquad (294)
\end{aligned}
$$

So the Polarization identity can be seen

$$
\begin{aligned}
4\langle x|y\rangle &= \|x+y\|^2-\|x-y\|^2-j\|x+jy\|^2+j\|x-jy\|^2 &(295)\\
&=4\Big(\Re\langle x|y\rangle+j\Im\langle x|y\rangle\Big) &(296)\\
&=4\Big(\Re\langle x|y\rangle-j\Re\langle x|jy\rangle\Big) &(297)\\
&=4\Big(\Re\langle x|y\rangle+j\Re\langle jx|y\rangle\Big) &(298)
\end{aligned}
$$

by using the relationship $\Im\langle x|y\rangle=\Re\langle jx|y\rangle=-\Re\langle x|jy\rangle$.

8 Proof of Helmholtz Decomposition

In vector analysis, we know that any vector $\mathbf{v}$ can be decomposed into one longitudinal component $\mathbf{v}_\parallel$ parallel to a unit vector $\hat{\mathbf{e}}$ and the other transverse component $\mathbf{v}_\perp$ perpendicular to the given vector $\hat{\mathbf{e}}$ such that

$$
\begin{aligned}
\mathbf{v} &= \mathbf{v}_\parallel+\mathbf{v}_\perp\\
&=(\mathbf{v}\cdot\hat{\mathbf{e}})\hat{\mathbf{e}}+[\mathbf{v}-(\mathbf{v}\cdot\hat{\mathbf{e}})\hat{\mathbf{e}}]\\
&=(\mathbf{v}\cdot\hat{\mathbf{e}})\hat{\mathbf{e}}-\hat{\mathbf{e}}\times(\hat{\mathbf{e}}\times\mathbf{v}) &(299)\\
&=(\mathbf{v}\cdot\hat{\mathbf{e}})\hat{\mathbf{e}}+(\mathbb{E}-\hat{\mathbf{e}}\hat{\mathbf{e}})\cdot\mathbf{v} &(300)
\end{aligned}
$$

where the longitudinal component $\mathbf{v}_\parallel=(\mathbf{v}\cdot\hat{\mathbf{e}})\hat{\mathbf{e}}$ and $\mathbf{v}_\perp$ is the other transverse component as shown in Figure 1.

By similarity, Helmholtz's Theorem states that any vector field can be written as the sum of a gradient and a curl. The decomposition is formulated in $\mathbf{r}\subseteq\mathbb{R}^3$-space. By a Fourier transform, the decomposition may be formulated in **k**-space. This is advantageous because differentiations in **r**-space become multiplications in **k**-space. We will show that divergence in **r**-space becomes an inner product in **k**-space and a curl becomes a cross product. Thus, we define the mutually inverse Fourier transforms,

$$\tilde{\mathfrak{F}}(\mathbf{k})=\frac{1}{\sqrt{(2\pi)^3}}\int_V e^{-i\mathbf{k}\cdot\mathbf{r}}\mathbf{F}(\mathbf{r})d^3\mathbf{r} \qquad (301)$$

$$\mathbf{F}(\mathbf{r})=\frac{1}{\sqrt{(2\pi)^3}}\int_V e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{\mathfrak{F}}(\mathbf{k})d^3\mathbf{k} \qquad (302)$$

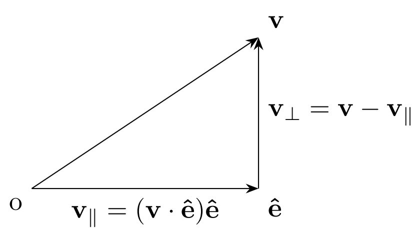

Figure 1: Vector Decomposition.

An arbitrary vector field in **k**-space can be decomposed in components parallel and perpendicular to **k**,

$$\tilde{\mathfrak{F}}_\parallel(\mathbf{k}) \equiv \big(\hat{\mathbf{k}}\cdot\tilde{\mathfrak{F}}(\mathbf{k})\big)\hat{\mathbf{k}}, \qquad \text{with} \qquad \hat{\mathbf{k}} \equiv \frac{\mathbf{k}}{|\mathbf{k}|}, \tag{303}$$

$$\tilde{\mathfrak{F}}_\perp(\mathbf{k}) \equiv \tilde{\mathfrak{F}}(\mathbf{k}) - \tilde{\mathfrak{F}}_\parallel(\mathbf{k}) = \hat{\mathbf{k}}\times\tilde{\mathfrak{F}}(\mathbf{k})\times\hat{\mathbf{k}} = -\hat{\mathbf{k}}\times(\hat{\mathbf{k}}\times\tilde{\mathfrak{F}}(\mathbf{k})) \tag{304}$$

so that

$$\tilde{\mathfrak{F}}(\mathbf{k}) = \tilde{\mathfrak{F}}_\perp(\mathbf{k}) + \tilde{\mathfrak{F}}_\parallel(\mathbf{k}). \tag{305}$$

Clearly,

$$\mathbf{k}\cdot\tilde{\mathfrak{F}}_\perp(\mathbf{k}) = 0 \qquad \text{and} \qquad \mathbf{k}\times\tilde{\mathfrak{F}}_\parallel(\mathbf{k}) = 0. \tag{306}$$

Transforming back, we get

$$\mathbf{F}_\perp(\mathbf{r}) \equiv \frac{1}{\sqrt{(2\pi)^3}}\int_V e^{i\mathbf{k}\cdot\mathbf{r}}\,\tilde{\mathfrak{F}}_\perp(\mathbf{k})\,d^3\mathbf{k}, \qquad \mathbf{F}_\parallel(\mathbf{r}) \equiv \frac{1}{\sqrt{(2\pi)^3}}\int_V e^{i\mathbf{k}\cdot\mathbf{r}}\,\tilde{\mathfrak{F}}_\parallel(\mathbf{k})\,d^3\mathbf{k}, \tag{307}$$

which satisfy the properties

$$\boldsymbol{\nabla}\cdot\mathbf{F}_\perp(\mathbf{r}) = \frac{i}{\sqrt{(2\pi)^3}}\int_V e^{i\mathbf{k}\cdot\mathbf{r}}\,\mathbf{k}\cdot\tilde{\mathfrak{F}}_\perp(\mathbf{k})\,d^3\mathbf{k} = 0 \tag{308}$$

$$\boldsymbol{\nabla}\times\mathbf{F}_\parallel(\mathbf{r}) = \frac{i}{\sqrt{(2\pi)^3}}\int_V e^{i\mathbf{k}\cdot\mathbf{r}}\,\mathbf{k}\times\tilde{\mathfrak{F}}_\parallel(\mathbf{k})\,d^3\mathbf{k} = 0. \tag{309}$$

Hence we have found the required decomposition.

9 Matrix Calculus

9.1 Gradient of Linear Function

Consider a linear function of the form

$$f(\mathbf{x}) = \mathbf{a}^{\mathbf{x}} \tag{310}$$

where $\mathbf{a}$ and $\mathbf{x}$ are n-dimensional vectors. We have already known that

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}, \tag{311}$$

and that the gradient of a scalar function is defined as a column vector with its components corresponding to the partial derivatives. Let's step into the inside of a vector (or matrix) to derive the gradient of a linear function (here obtained from vector inner product).

$$\begin{aligned}
\frac{\partial}{\partial x_k} f(\mathbf{x}) &= \frac{\partial}{\partial x_k}\left(\begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) && (312) \\
&= \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \frac{\partial}{\partial x_k}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) && (313) \\
&= \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \left(\begin{bmatrix} \delta_{ik} \\ \vdots \\ \delta_{nk} \end{bmatrix}\right) && (314) \\
&= a_k && (315)
\end{aligned}$$

We then assemble all the resulting derivatives of components for $k = 1 \dots n$.

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} f(\mathbf{x}) = \begin{bmatrix} \vdots \\ \frac{\partial}{\partial x_k} \\ \vdots \end{bmatrix} \mathbf{a}^{\mathbf{x}} = \begin{bmatrix} \vdots \\ a_k \\ \vdots \end{bmatrix} = \mathbf{a} \tag{316}$$

Note: the gradient definition shall not be confused by the definition of the derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ w.r.t the column vector $\mathbf{x}$. If y is a scalar function of independent scalar variable x, then the differential $\mathrm{d}y$ of y is related to $\mathrm{d}x$ by the formula

$$\mathrm{d}y = \frac{dy}{dx}\mathrm{d}x \tag{317}$$

where $\mathrm{d}y/\mathrm{d}x$ denotes the derivative of y with respect to x. The definition can be still applicable when the independent variable has higher dimensions, like vector or matrix. For instance, the differential of the scalar function $f : \mathbb{R}^n \to \mathbb{R}$ is still a scalar

$$\mathrm{d}f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} \mathrm{d}\mathbf{x} \tag{318}$$

which is also consistent with the result of a column vector d$\mathbf{x}$ left-multiplied by a row vector. Since $f(\mathbf{x}) = \mathbf{a}^{\mathbf{x}}$ is a scalar function, the result of differentiation w.r.t. $\mathbf{x}$ must be a row vector. In this way

$$\frac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}\mathbf{a}^{\mathbf{x}}}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}^{\mathbf{a}}}{\mathrm{d}\mathbf{x}} = \mathbf{a}^{=}\nabla^{f(\mathbf{x})} \tag{319}$$

Therefore, we can conclude that the derivative with respect to a column vector is ended up with a row vector. The full version of chain rule for the vector function can be defined as

$$\frac{\mathrm{d}f(g,h)}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}(g(\mathbf{x})^T)}{\mathrm{d}\mathbf{x}} \frac{\partial f(g,h)}{\partial g} + \frac{\mathrm{d}(h(\mathbf{x})^T)}{\mathrm{d}\mathbf{x}} \frac{\partial f(g,h)}{\partial h} \tag{320}$$

where $f, g, h : \mathbb{R}^n \to \mathbb{R}$.

9.2 Gradient of Quadratic Function

Consider a quadratic function of the form

$$f(\mathbf{x}) = \mathbf{x}^{\mathbf{A}}\mathbf{x} \tag{321}$$

where $\mathbf{x}$ is a n-size vector and $\mathbf{A}$ is a $n \times n$ matrix. We can utilize the chain-rule to derive the gradient in matrix notation in a straight-forward way.

$$\frac{\partial}{\partial x_k} f(\mathbf{x}) = \frac{\partial}{\partial x_k} \left(\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{322}$$

$$= \left(\frac{\partial}{\partial x_k} \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \right) \mathbf{A} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \mathbf{A} \frac{\partial}{\partial x_k} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{323}$$

$$= \begin{bmatrix} \delta_{1k} & \dots & \delta_{nk} \end{bmatrix} \mathbf{A} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \mathbf{A} \begin{bmatrix} \delta_{ik} \\ \vdots \\ \delta_{nk} \end{bmatrix} \tag{324}$$

$$= a_{ki}x_i + x_i a_{ik} \tag{325}$$

Therefore, after assembling all the components, we have

$$\nabla^{\mathrm{T}} f(\mathbf{x}) = (\mathbf{A}^{+}\mathbf{A})\mathbf{x} \tag{326}$$

9.3 Hessian Matrix

If we take the double gradients, i.e., the gradient of the gradient, we obtain the so-called Hessian matrix, here donated by $\nabla\nabla$ in stead of ∇^2. However, not only a few mathematicians like to use ∇^2, which becomes to bewilder many engineers due to the fact that ∇^2 stands for the Laplacian were already deeply ingrained in the education. Especially, when some writers do not differentiate the scalar notations with vector/matrix ones. So to avoid any confusion, the Laplacian is donated by $\nabla^2 = \nabla \cdot \nabla$, the divergent of the gradient and is a scalar, while the Hessian Matrix is the gradient of the gradient, therefore explicitly expressed as $\nabla\nabla$. From the word-width point of view, ∇^2 can save spaces for $\nabla \cdot \nabla$, but not for $\nabla\nabla$.

If we take the gradient of (326) by the use of (316), it readily arrives at

$$\begin{aligned}
\nabla\nabla f(\mathbf{x}) &= \nabla\nabla \left(\mathbf{x}^{\mathbf{A}}\mathbf{x}\right) && (327)\\
&= \nabla \left[(\mathbf{A}^{+}\mathbf{A})\mathbf{x}\right] && (328)\\
&= \left(\mathbf{A}^{+}\mathbf{A}\right) && (329)\\
&= \mathbf{A} + \mathbf{A} && (330)
\end{aligned}$$

9.4 Least Square Application

Given $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{y} \in \mathbb{R}^{m\times p}$, we form a matrix equation in $\mathbf{x} \in \mathbb{R}^{n\times p}$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{331}$$

If we want to find the least-squares solution, then we minimize $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_F$, where the norm $\|\cdot\|_F$ is taken as the Frobenius-norm[3] which is equivalent to

[3]Frobenius-norm is one of the oldest and entrywise/elementwise norms. It is defined as L_2-norm, i.e., the square root of the sum of the squares of all the matrix entries. Let $\mathbf{A} = [a_{ij}]$ be an $m \times n$ matrix with rows $\mathbf{a}_1^T, \ldots, \mathbf{a}_m^T$. That is

$$\mathbf{A} = \begin{bmatrix} — & \mathbf{a}_1^T & — \\ — & \mathbf{a}_2^T & — \\ & \vdots & \\ — & \mathbf{a}_m^T & — \end{bmatrix}$$

The Frobenius norm is

$$\begin{aligned}
\|\mathbf{A}\|_F^2 &= \sqrt{\langle \mathbf{A}, \mathbf{A}\rangle} = \sqrt{\sum_{i=1}^{m} \mathbf{a}_i^T \mathbf{a}_i} = \sqrt{\sum_{i=1}^{m} \|\mathbf{a}_i\|_2^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \\
&= \sqrt{\mathrm{Tr}(\mathbf{A}\mathbf{A}^T)} = \sqrt{\mathrm{Tr}\left(\begin{bmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \mathbf{a}_1^T\mathbf{a}_2 & \ldots & \mathbf{a}_1^T\mathbf{a}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m^T\mathbf{a}_1 & \mathbf{a}_m^T\mathbf{a}_2 & \ldots & \mathbf{a}_m^T\mathbf{a}_m \end{bmatrix}.\right)} \qquad (332)
\end{aligned}$$

Frobenius norm is also called spectrum norm. Other relevant norms, for instance L_1-norm

$$\|\mathbf{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|$$

least squares norm for matrices. However, if we do not want the norm of $\mathbf{x}$ to become too large, then we minimize the following objective function.

$$\begin{aligned} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_F^2+\lambda\|\mathbf{x}\|_F^2 &= \mathrm{tr}((\mathbf{A}\ \mathbf{x}-\mathbf{y})^T(\mathbf{A}\mathbf{x}-\mathbf{y}))+\lambda\mathrm{tr}(\mathbf{x}^T\mathbf{x}) \\ &= \mathrm{tr}(\mathbf{x}^T(\mathbf{A}^T\mathbf{A}\lambda\mathbf{I}_n)\mathbf{x}-\mathbf{y}^T\mathbf{A}\ \mathbf{x}\ -\mathbf{x}^T\mathbf{A}^T\mathbf{y}\mathbf{y}^T\mathbf{y}) \end{aligned} \tag{333}$$

where $\lambda > 0$. Differentiating with respect to $\mathbf{x}$, we obtain

$$2(\mathbf{A}^T\mathbf{A}+\lambda\mathbf{I}_n)\mathbf{x}-2\mathbf{A}^T\mathbf{y} \tag{334}$$

Finding where the derivative vanishes, we obtain the matrix equation

$$(\mathbf{A}^T\mathbf{A}\lambda\mathbf{I}_n)\mathbf{x}=\mathbf{A}^T\mathbf{y} \tag{335}$$

If $\lambda > 0$, then $\mathbf{A}^T\mathbf{A}\lambda\mathbf{I}_n$ is always invertible. Hence, the unique minimizer is

$$\hat{\mathbf{x}} := (\mathbf{A}^T\mathbf{A}\lambda\mathbf{I}_n)^{-1}\mathbf{A}^T\mathbf{y} \tag{336}$$

9.5 Differential of Vector Function

In general, the derivative of a function $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^m$ at a point $p\in\mathbb{R}^n$, if it exists, is the unique linear transformation $Df(p)\in L(\mathbb{R}^n,\mathbb{R}^m)$ such that

$$\lim_{h\to 0}\frac{\|\mathbf{f}(p+h)-\mathbf{f}(p)-\mathrm{d}\mathbf{f}(p)h\|}{\|h\|}=0; \tag{337}$$

the matrix of $\mathrm{d}\mathbf{f}(p)$ with respect to the standard orthonormal bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, called the Jacobian matrix of $\mathbf{f}$ at p, therefore lies in $M_{m\times n}(\mathbb{R})$. The (i,j) entry of the Jacobian at $\mathbf{x}$ is

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}) \tag{338}$$

From Wiki, suppose $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^m$ is a function such that each f_i $(i=1\ldots m)$ of its first-order partial derivatives exist on $\mathbb{R}^n$ independent argument. This function takes a point $\mathbf{x}\in\mathbb{R}^n$ as input and produces the vector $\mathbf{f}(\mathbf{x})\in\mathbb{R}^m$ as output. Then the Jacobian matrix of $\mathbf{f}$ is defined to be an $m\times n$ matrix, denoted by $\mathbf{J}$, whose (i,j)th entry is $\mathbf{J}_{ij}=\frac{\partial f_i}{\partial x_j}$, or explicitly

$$\mathbf{J}=\begin{bmatrix}\frac{\partial\mathbf{f}}{\partial x_1} & \cdots & \frac{\partial\mathbf{f}}{\partial x_n}\end{bmatrix}=\begin{bmatrix}\nabla^{\mathrm{T}}f_1\\ \vdots\\ \nabla^{\mathrm{T}}f_m\end{bmatrix}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n}\end{bmatrix} \tag{339}$$

---

and L_∞-norm

$$\|\mathbf{A}\|_\infty=\max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|$$

are referred as column-sum norm and row-sum norm, respectively.

where $\nabla^{\mathrm{T}} f_i$ is the transpose (row vector) of the gradient of the i component, i.e., the i^{th} row of the Jacobian is then given by

$$\left[\frac{\partial f_i}{\partial x_1}, \frac{\partial f_i}{\partial x_2}, \ldots, \frac{\partial f_i}{\partial x_n}\right](\mathbf{x}) = \nabla^{\mathrm{T}} f_i(\mathbf{x}). \tag{340}$$

The j^{th} column of the Jacobian contains the partial derivatives of the m components $(f_1 \ldots f_m)$ of $\mathbf{f}$ with respect to the j^{th} variable x_j. There is no standard notation for such a column, and $\mathbf{J_f}(\mathbf{x})\mathbf{e}_j$ or $\partial_j \mathbf{f}(\mathbf{x})$ are candidate notations without confusion. Compactly, $f_{i,j}(\mathbf{x})$ can be used for the partial derivative (338).

9.6 Matrix Derivatives

Under the above link the extensive differential method for vector and matrix are covered with the following notations. **Notation**

1. x, $\mathbf{x}$ and $\mathbf{X}$ denote *scalar*, **vector** and **MATRIX** respectively.

2. The letters toward the beginning of the alphabet-such as $(a, b, c, \cdots)$ denote constants, and those toward the end of the alphabet-such as $(\cdots, x, y, z)$ denote variables.

3. $\mathbf{X}^T$ denotes matrix transpose, $\operatorname{tr}(\mathbf{X})$ is the trace, $|\mathbf{X}|$ is the determinant, and $\operatorname{adj}(\mathbf{X})$ is the adjugate matrix.

4. $\otimes$ is the Kronecker product and $\circ$ is the Hadamard product[4].

5. Here we use *numerator layout* while the online tool Matrix Calculus seems to use *mixed layout*. Please refer to Wiki - Matrix Calculus - Layout Conventions for the detailed layout definitions, and keep in mind that different layouts lead to different results.

References

George Pólya and John Horton Conway. *How to solve it*. Princeton science library. Princeton University Press, expanded princeton science library ed. edition, 2004. ISBN 0-691-11966-X ; 978-0-691-11966-3. Includes bibliographical references.

G Callen. Fundamentals of plasma physics. appendix d. vector analysis. *https://drive.google.com/file/d/1hLQsL6A8GVQgDYfXfZ95GBZ36R-nFIja*, 2002.

[4]Hadamard product is also known as the element-wise product, entrywise product or Schur product. For two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimension $m \times n$, the Hadamard product $A \circ B$ (or $A \odot B$) is a matrix of the same dimension as the operands, with elements given by

$$(A \circ B)_{ij} = (A \odot B)_{ij} = (A)_{ij}(B)_{ij}.$$

The Hadamard product is undefined for matrices of different dimensions.

J. WILLARD GIBBS and E. BIDWELL WILSON. *Vector Analysis; a text-book for the use of students of mathematics and physics*. New York, C. Scribner's Sons, New York, 1901.

Chen X-B. Middle-field formulation for the computation of wave-drift loads. *J Eng Math 59, 61–82*, 2007.