2 Differential vector calculus

2.1 Linear algebra

Focus now on 3D, and adopt convention that position vector $\mathbf{r} = (x, y, z) \equiv (x_1, x_2, x_3) \in \mathbb{R}^3$ to describe equations pertaining to physical applications.

Notation: The Cartesian (unit) basis vectors in $\mathbb{R}^3$ are $\hat{\mathbf{x}} = (1, 0, 0) \equiv \mathbf{e}_1$, $\hat{\mathbf{y}} = (0, 1, 0) \equiv \mathbf{e}_2$ and $\hat{\mathbf{z}} = (0, 0, 1) \equiv \mathbf{e}_3$ such that $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} \equiv x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$.

Also use $r = |\mathbf{r}|$ is the length of the vector.

Defn: The *dot product* of two vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ is defined

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 \equiv \sum_{j=1}^{3} u_jv_j$$

Notation: *Einstein summation convention*: Drop the $\sum_{j=1}^{3}$ in the above on the understanding that repeated suffices imply summation. I.e.

$$\mathbf{u} \cdot \mathbf{v} = u_jv_j$$

E.g. $r = |\mathbf{r}| = \sqrt{\mathbf{r} \cdot \mathbf{r}} = \sqrt{x_i^2}$.

Defn: The *Kronecker delta* symbol δ_{ij} is defined to be

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

E.g. 1) $x_i = \delta_{ij}x_j$ since this is $\sum_{j=1}^{3} \delta_{ij}x_j$

E.g. 2) $\delta_{ii} = 3$.

E.g. 3) $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

E.g. 4) $\mathbf{x} = x_j\mathbf{e}_j$ and taking dot product with $\mathbf{e}_i$ gives $x_i = \mathbf{x} \cdot \mathbf{e}_i$.

Defn: The *cross product* of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, vector given by

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \equiv (u_2v_3 - v_2u_3)\mathbf{e}_1 + (u_3v_1 - v_3u_1)\mathbf{e}_2 + (u_1v_2 - v_1u_2)\mathbf{e}_3. \tag{7}$$

Note: Defintion implies antisymmetry: $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$.

Defn: The *Levi-Civita tensor* (or antisymmetric tensor) is defined by

1. $\epsilon_{123} = 1$

2. $\epsilon_{ijk} = 0$ if any repeated suffices. E.g. $\epsilon_{113} = 0$.

3. Interchanging suffices implies reversal of sign. E.g. $\epsilon_{ijk} = -\epsilon_{jik}$.

Implies ϵ_{ijk} are invariant under cyclic rotation of suffices. Thus $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$, and all 21 others are zero.

Note: (Very nice) cross product can be written as

$$[\mathbf{u} \times \mathbf{v}]_i = \epsilon_{ijk} u_j v_k. \tag{8}$$

there is a *double sum* on the right hand side, by the summation convention. The defintion of ϵ_{ijk} guarantees the antisymmetry of the cross product (check !).

Proposition:

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{9}$$

Proof: Exercise !

Example: Prove the *vector triple product* relation

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}.$$

Proof:

$$\begin{aligned}
[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk} a_j [\mathbf{b} \times \mathbf{c}]_k \\
&= \epsilon_{ijk} a_j \epsilon_{klm} b_l c_m \\
&= \epsilon_{kij}\epsilon_{klm} a_j b_l c_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) a_j b_l c_m \\
&= a_j c_j b_i - a_j b_j c_i = (\mathbf{a} \cdot \mathbf{c}) b_i - (\mathbf{a} \cdot \mathbf{b}) c_i
\end{aligned}$$

True for $i = 1, 2, 3$, so result is proved.

2.2 Scalar and vector fields

Defn: Conventional language:

A *scalar field* on $\mathbb{R}^3$ is a function $f : \mathbb{R}^3 \to \mathbb{R}$.
A *vector field* on $\mathbb{R}^3$ is a map $\mathbf{v} : \mathbb{R}^3 \to \mathbb{R}^3$.

Scalar and vector fields defined in $\mathbb{R}^3$ are of particular importance for physical applications.

E.g.s:

- (Scalar fields) Temperature $T(\mathbf{r})$; mass density $\rho(\mathbf{r})$ for a fluid or gas; electric charge density $q(\mathbf{r})$.
- (Vector fields) Velocity $\mathbf{v}(\mathbf{r})$ of a fluid or gas; electric and magnetic fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{B}(\mathbf{r})$, displacement fields in elastic solid $\mathbf{u}(\mathbf{r})$.

There are three basic first-order differential operations in vector calculus.

2.3 Gradient (grad)

Defn: The *gradient* of a scalar field f, denoted $\boldsymbol{\nabla} f$, is the vector field given by

$$\boldsymbol{\nabla} f(\mathbf{r}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$$

(I.e. the gradient maps scalar to vector fields.)

E.g. 1)

$$\boldsymbol{\nabla} \tan^{-1}\left(\frac{y}{x}\right) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right) = \frac{1}{x^2+y^2}(-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}).$$

E.g. 2) Recall $r = \sqrt{x^2+y^2+x^2}$

$$\boldsymbol{\nabla} r = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) = \frac{(x,y,z)}{r} = \frac{\mathbf{r}}{r}$$

$\mathbf{r}/r$ is the unit vector from the origin to the point $\mathbf{r}$; we often denote this as $\hat{\mathbf{r}}$.

Note: $[\boldsymbol{\nabla} r]_i = x_i/r$.

E.g. 3) If $f(\mathbf{r}) = g(r)$ (i.e. a function depends only on the distance from the origin) then

$$[\boldsymbol{\nabla} g(r)]_i = \frac{\partial g(r)}{\partial x_i} = \frac{dg(r)}{dr}\frac{\partial r}{\partial x_i} \equiv g'(r)[\boldsymbol{\nabla} r]_i = g'(r)\left(\frac{\mathbf{r}}{r}\right)_i$$

and thus $\boldsymbol{\nabla} g(r) = g'(r)\hat{\mathbf{r}}$ (c.f. potentials, central forces in Mech 1).

Recall from Calculus 1, two important interpretations of the gradient:

2.3.1 Interpretation of the gradient

Provided $\boldsymbol{\nabla} f$ is nonzero, the gradient points in the direction in which f changes most rapidly.

Proof: let $\mathbf{v}$ be s.t. $|\mathbf{v}| = 1$. Then rate of change of f in direction $\mathbf{v}$ is the directional derivative (see (1)) $D_{\mathbf{v}} f(\mathbf{r}) = \mathbf{v} \cdot \boldsymbol{\nabla} f = |\boldsymbol{\nabla} f| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\boldsymbol{\nabla} f$. Maximised when $\theta = 0$. I.e. when $\mathbf{v}$ in direction $\boldsymbol{\nabla} f$.

2.3.2 Another interpretation of the gradient

The gradient of f is perpendicular to the level surfaces of f.

(A level surface S is defined by values of $\mathbf{r}$ s.t. $f(\mathbf{r}) = C$, a constant.)

Proof: Let $\mathbf{c}(t)$ lie in S. Then $f(\mathbf{c}(t)) = C$, for all t. The chain rule yields

$$0 = \frac{d}{dt} f(\mathbf{c}(t)) = \boldsymbol{\nabla} f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$$

and since $\mathbf{c}'(t)$ is parallel to S at $\mathbf{c}(t)$, we have our result.

2.3.3 Examples

E.g. 1) The temperature T in a room is a function of 3D position (x, y, z):

$$T(\mathbf{r}) = \frac{e^x \sin(\pi y)}{1 + z^2}$$

If you stand at the point $(1, 1, 1)$ in which direction will the room get coolest fastest ?

A:

$$\boldsymbol{\nabla} T = \left(\frac{e^x \sin(\pi y)}{1 + z^2}, \frac{\pi e^x \cos(\pi y)}{1 + z^2}, -\frac{2z e^x \sin(\pi y)}{(1 + z^2)^2} \right)$$

and at $(x, y, z) = (1, 1, 1)$, $\nabla T = (\frac{1}{2}e, 0, -\frac{1}{2}e)$. So a vector pointing in the direction where temperature gets coolest is $(-1, 0, 1)$.

E.g. 2) Take $f(x, y) = x^2 + 2y^2$. Then $\boldsymbol{\nabla} f = (2x, 4y)$.

Then: e.g. (i) ∇f evaulated at $(x, y) = (1, 1)$ is $(2, 4)$ and so the steepest ascent of f at $(1, 1)$ is in direction $\tan^{-1}(2)$ w.r.t. x axis and gradient in that direction is $|\nabla f| = 2\sqrt{5}$;

or e.g. (ii) $D_{(1,0)} f = (2x, 0)$ and this equals 0 if $x = 0$ meaning the gradient pointing in the +ve x-direction of f along $x = 0$ is zero.

E.g. 3) You are on a bicycle about to climb Clifton Vale. The gradient of this road is 20%. You cannot cycle up such steep gradients, but you can manage 10% gradients. At what angle to the road do you need to zig-zag to cut the gradient down to 10% ?

A: 20% gradient is one-in-five and, with x pointing up the road, the height of the hill is given by $f(x, y) = \frac{1}{5}x$. So $\nabla f = (\frac{1}{5}, 0)$. Let us cycle in a direction θ w.r.t. to the x-axis. Then the unit vector pointing in this direction is

$$\mathbf{u} = (\cos\theta, \sin\theta)$$

and the directional derivative $D_{\mathbf{u}} f$ gives the gradient in the direction $\mathbf{u}$. We want this to be 10% or one-in-ten so

$$\tfrac{1}{10} = D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f = (\cos\theta, \sin\theta) \cdot (\tfrac{1}{5}, 0) = \tfrac{1}{5}\cos\theta$$

and this means $\theta = 60°$.

2.4 Divergence (Div)

Defn: The *divergence* of a vector field $\mathbf{v}(\mathbf{r})$, denoted $\boldsymbol{\nabla} \cdot \mathbf{v}$, is the scalar field given by

$$\boldsymbol{\nabla} \cdot \mathbf{v} = \frac{\partial}{\partial x_j} v_j(\mathbf{r}) \equiv \partial_j v_j.$$

2.4.1 Interpretation of divergence

Harder without physical setting, but broadly it measures the expansion (positive divergence) or contraction of a field at a point. It measures flux (rate of transport of a field) entering a point.

2.4.2 Examples

1) $\mathbf{v}(\mathbf{r}) = (xyz, xyz, xyz)$ then

$$\boldsymbol{\nabla} \cdot \mathbf{v} = yz + zx + xy$$

2)

$$\boldsymbol{\nabla} \cdot \mathbf{r} = \frac{\partial x_j}{\partial x_j} = \delta_{jj} = 3.$$

3)

$$\boldsymbol{\nabla} \cdot (\mathbf{a} \times \mathbf{r}) = \frac{\partial}{\partial x_i} \epsilon_{ijk} a_j x_k = \epsilon_{ijk} a_j \frac{\partial x_k}{\partial x_i} = \epsilon_{ijk} a_j \delta_{ik} = \epsilon_{iji} a_j = 0$$

since $\epsilon_{iji} = 0$.

2.5 Curl

Defn: The *curl* of a vector field $\mathbf{v}(\mathbf{r})$, denoted $\boldsymbol{\nabla} \times \mathbf{v}$, is the vector field (i.e. in $\mathbb{R}^3$) given by

$$\boldsymbol{\nabla} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial_x & \partial_y & \partial_z \\ v_1 & v_2 & v_3 \end{vmatrix} \equiv \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}, \frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}, \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right).$$

Alternatively,

$$[\boldsymbol{\nabla} \times \mathbf{v}]_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} v_k \tag{10}$$

as for cross products.

2.5.1 Interpretation of Curl

Again harder without physical setting, but broadly it measures the rotation or circulation of a vector field (because it needs direction) at a point.

2.5.2 Examples

1) Let $\mathbf{v}(\mathbf{r}) = (y^2, x^2, y^2)$. Then

$$\boldsymbol{\nabla} \times \mathbf{v} = (2y, 0, 2(x - y))$$

2)

$$[\boldsymbol{\nabla} \times \mathbf{r}]_i = \epsilon_{ijk} \partial_j x_k = \epsilon_{ijk} \delta_{jk} = \epsilon_{ijj} = 0,$$

so $\boldsymbol{\nabla} \times \mathbf{r} = 0$.

2.6 Second-order differential operations

Schematically, grad, div and curl act as follows:

$$\begin{aligned}\text{grad: scalar fields} &\rightarrow \text{vector fields}\\ \text{div: vector fields} &\rightarrow \text{scalar fields}\\ \text{curl: vector fields} &\rightarrow \text{vector fields}\end{aligned}$$

The operations of grad, div and curl can be combined. Thus, the following combination of operations make sense:

$$\begin{aligned}\text{curl grad: scalar fields} &\rightarrow \text{vector fields}\\ \text{div grad: scalar fields} &\rightarrow \text{scalar fields}\\ \text{grad div: vector fields} &\rightarrow \text{vector fields}\\ \text{div curl: vector fields} &\rightarrow \text{scalar fields}\\ \text{curl curl: vector fields} &\rightarrow \text{vector fields}\end{aligned}$$

2.6.1 Null Identities

1) For scalar fields f

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) = 0.$$

Proof: We have that

$$[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f)]_i = \epsilon_{ijk}\partial_j\partial_k f = -\epsilon_{ikj}\partial_j\partial_k f = -\epsilon_{ikj}\partial_k\partial_j f = -[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f)]_i.$$

Thus since the expression equals its own negative, it must vanish.

2) For vector fields, $\mathbf{v}$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = 0$$

Proof: We have that

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = \partial_i\epsilon_{ijk}\partial_j v_k = \epsilon_{ijk}\partial_i\partial_j v_k,$$

which must vanish for the same reason.

The remaining combinations of grad, div and curl are related to a second-order differential operator called the Laplacian...

2.7 The Laplacian

Defn: The *Laplacian* of a scalar field $f(\mathbf{r})$, denoted $\nabla^2 f$ (or $\triangle f$), is the scalar field given by

$$\triangle f = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f(\mathbf{r}) = \partial_i^2 f \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) f.$$

The *Laplacian* of a vector field $\mathbf{v}(\mathbf{r})$, is

$$\triangle \mathbf{v} = (\triangle v_1, \triangle v_2, \triangle v_3)\,.$$

Example: For a vector field $\mathbf{v}(\mathbf{r})$,

$$\triangle \mathbf{v} = -\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v}) + \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v}).$$

Proof: We consider the ith component of the 1st RHS term:

$$\begin{aligned}
[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v})]_i &= \epsilon_{ijk}\partial_j(\boldsymbol{\nabla} \times \mathbf{v})_k \\
&= \epsilon_{ijk}\partial_j \epsilon_{klm}\partial_l v_m \\
&= \epsilon_{kij}\epsilon_{klm}\partial_j\partial_l v_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l v_m \\
&= \partial_i\partial_m v_m - \partial_j\partial_j v_i = [\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v}) - \triangle \mathbf{v}]_i
\end{aligned}$$

which shows that all components agree with our claim.

2.8 Curvilinear coordinate systems

All differential operators defined above were expressed in Cartesian coordinates. For many practical problems more natural to express problems in coordinates aligned with principal features of the problem. E.g. Polars are appropriate for circular domains.

Q: How do we recast the differential operators in a differential coordinate system ?

2.8.1 Coordinate transformations

Defn: *Curvilinear coordinates* are defined by a smooth function $\mathbf{r} : \mathbb{R}^3 \to \mathbb{R}^3$ which maps a point $\mathbf{q} = (q_1, q_2, q_3)$ in one coordinate system to a point in Cartesian space: $\mathbf{r} \equiv (x, y, z) = \mathbf{r}(\mathbf{q}) = \mathbf{r}(q_1, q_2, q_3)$. I.e.

$$x = x(q_1, q_2, q_3), \quad y = y(q_1, q_2, q_3), \quad z = z(q_1, q_2, q_3)$$

The inverse map, if it exists (see later) is

$$q_1 = q_1(x, y, z), \quad q_2 = q_2(x, y, z), \quad q_3 = q_3(x, y, z).$$

E.g. In 2D, if $q_1 = r$ and $q_2 = \theta$ then $x = x(r, \theta) = r\cos\theta$ and $y = y(r, \theta) = r\sin\theta$. The inverse map is $r = r(x, y) = \sqrt{x^2 + y^2}$, $\theta = \theta(x, y) = \tan^{-1}(y/x)$.

Defn: The surfaces $q_i = const$ are called *coordinate surfaces*. The space curves formed by their intersection in pairs are called the *coordinate curves*. The *coordinate axes* are determined by the tangents to the coordinate curves at the intersection of three surfaces. They are not, in general, fixed directions in space.

In Cartesians the standard basis can be written

$$\hat{\mathbf{x}} = \frac{\partial \mathbf{r}}{\partial x}, \quad \hat{\mathbf{y}} = \frac{\partial \mathbf{r}}{\partial y}, \quad \hat{\mathbf{z}} = \frac{\partial \mathbf{r}}{\partial z}$$

and a point P in space is written $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$.