5. The Gradient Operator

A brief review is provided here for the gradient operator $\vec{\nabla}$ in both Cartesian and orthogonal non-Cartesian coordinate systems.

Sections in this Chapter:

5.01 Gradient, Divergence, Curl and Laplacian (Cartesian)

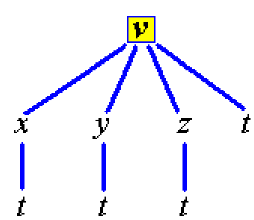
Let z be a function of two independent variables (x, y), so that $z = f(x, y)$.
The function $z = f(x, y)$ defines a surface in $\mathbb{R}^3$.
At any point (x, y) in the x-y plane, the direction in which one must travel in order to experience the greatest possible rate of increase in z at that point is the direction of the **gradient vector**,

$$\vec{\nabla} f = \frac{\partial f}{\partial x}\hat{\mathbf{i}} + \frac{\partial f}{\partial y}\hat{\mathbf{j}}$$

The magnitude of the gradient vector is that greatest possible rate of increase in z at that point. The gradient vector is not constant everywhere, unless the surface is a plane. (The symbol $\vec{\nabla}$ is usually pronounced "del").

The concept of the gradient vector can be extended to functions of any number of variables. If $u = f(x, y, z, t)$, then $\vec{\nabla} f = \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} & \dfrac{\partial f}{\partial t} \end{bmatrix}^{\mathrm{T}}$.

If $\mathbf{v}$ is a function of position $\mathbf{r}$ and time t, while position is in turn a function of time, then by the chain rule of differentiation,



$$\frac{d\vec{\mathbf{v}}}{dt} = \frac{\partial \vec{\mathbf{v}}}{\partial x}\frac{dx}{dt} + \frac{\partial \vec{\mathbf{v}}}{\partial y}\frac{dy}{dt} + \frac{\partial \vec{\mathbf{v}}}{\partial z}\frac{dz}{dt} + \frac{\partial \vec{\mathbf{v}}}{\partial t} = \left(\frac{d\vec{\mathbf{r}}}{dt}\bullet\vec{\nabla}\right)\vec{\mathbf{v}} + \frac{\partial \vec{\mathbf{v}}}{\partial t}$$

$$\Rightarrow \frac{d\vec{\mathbf{v}}}{dt} = \left(\vec{\mathbf{v}}\bullet\vec{\nabla}\right)\vec{\mathbf{v}} + \frac{\partial \vec{\mathbf{v}}}{\partial t}$$

which is of use in the study of fluid dynamics.

The gradient operator can also be applied to vectors via the scalar (dot) and vector (cross) products:

The **divergence** of a vector field $\mathbf{F}(x, y, z)$ is

$$\boxed{\operatorname{div}\vec{\mathbf{F}} = \vec{\nabla}\bullet\vec{\mathbf{F}} = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix}^{\mathrm{T}} \bullet \begin{bmatrix} F_1 & F_2 & F_3 \end{bmatrix}^{\mathrm{T}} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}}$$

A region free of sources and sinks will have zero divergence:
the total flux into any region is balanced by the total flux out from that region.

The **curl** of a vector field $\mathbf{F}(x, y, z)$ is

$$\boxed{\operatorname{curl}\vec{\mathbf{F}} = \vec{\nabla}\times\vec{\mathbf{F}} = \begin{vmatrix} \hat{\mathbf{i}} & \frac{\partial}{\partial x} & F_1 \\ \hat{\mathbf{j}} & \frac{\partial}{\partial y} & F_2 \\ \hat{\mathbf{k}} & \frac{\partial}{\partial z} & F_3 \end{vmatrix} = \begin{bmatrix} \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \\ \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \\ \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \end{bmatrix}}$$

In an irrotational field, curl $\vec{\mathbf{F}} = \vec{\mathbf{0}}$.
Whenever $\vec{\mathbf{F}} = \vec{\nabla}\phi$ for some twice differentiable potential function ϕ, curl $\vec{\mathbf{F}} = \vec{\mathbf{0}}$ or

$$\boxed{\operatorname{curl}(\operatorname{grad}\phi) \equiv \vec{\nabla}\times\vec{\nabla}\phi \equiv \vec{\mathbf{0}}}$$

<u>Proof</u>:

$$\vec{\mathbf{F}} = \vec{\nabla}\phi = \begin{bmatrix} F_1 & F_2 & F_3 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{bmatrix}^{\mathrm{T}}$$

$$\Rightarrow \quad \operatorname{curl}\vec{\nabla}\phi = \begin{vmatrix} \hat{\mathbf{i}} & \frac{\partial}{\partial x} & \frac{\partial \phi}{\partial x} \\ \hat{\mathbf{j}} & \frac{\partial}{\partial y} & \frac{\partial \phi}{\partial y} \\ \hat{\mathbf{k}} & \frac{\partial}{\partial z} & \frac{\partial \phi}{\partial z} \end{vmatrix} = \begin{bmatrix} \frac{\partial^2 \phi}{\partial y\,\partial z} - \frac{\partial^2 \phi}{\partial z\,\partial y} \\ \frac{\partial^2 \phi}{\partial z\,\partial x} - \frac{\partial^2 \phi}{\partial x\,\partial z} \\ \frac{\partial^2 \phi}{\partial x\,\partial y} - \frac{\partial^2 \phi}{\partial y\,\partial x} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Among many identities involving the gradient operator is

$$\boxed{\operatorname{div}\left(\operatorname{curl}\vec{\mathbf{F}}\right) \equiv \vec{\boldsymbol{\nabla}}\bullet\vec{\boldsymbol{\nabla}}\times\vec{\mathbf{F}} \equiv 0}$$

for all twice-differentiable vector functions $\vec{\mathbf{F}}$

Proof:

$$\operatorname{div}\operatorname{curl}\vec{\mathbf{F}} = \frac{\partial}{\partial x}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

$$= \cancel{\frac{\partial^2 F_3}{\partial x\,\partial y}} - \bcancel{\frac{\partial^2 F_2}{\partial x\,\partial z}} + \xcancel{\frac{\partial^2 F_1}{\partial y\,\partial z}} - \cancel{\frac{\partial^2 F_3}{\partial y\,\partial x}} + \bcancel{\frac{\partial^2 F_2}{\partial z\,\partial x}} - \xcancel{\frac{\partial^2 F_1}{\partial z\,\partial y}} \equiv 0$$

The divergence of the gradient of a scalar function is the **Laplacian**:

$$\boxed{\operatorname{div}\left(\operatorname{grad} f\right) \equiv \vec{\boldsymbol{\nabla}}\bullet\vec{\boldsymbol{\nabla}} f \equiv \nabla^2 f \equiv \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}}$$

for all twice-differentiable scalar functions f.

In orthogonal non-Cartesian coordinate systems, the expressions for the gradient operator are not as simple.

5.02 Differentiation in Orthogonal Curvilinear Coordinate Systems

For any orthogonal curvilinear coordinate system (u_1, u_2, u_3) in $\mathbb{R}^3$,
the unit tangent vectors along the curvilinear axes are $\hat{\mathbf{e}}_i = \hat{\mathbf{T}}_i = \dfrac{1}{h_i}\dfrac{\partial \vec{\mathbf{r}}}{\partial u_i}$,

where the scale factors $\boxed{h_i = \left|\dfrac{\partial \vec{\mathbf{r}}}{\partial u_i}\right|}$.

The displacement vector $\vec{\mathbf{r}}$ can then be written as $\vec{\mathbf{r}} = u_1\hat{\mathbf{e}}_1 + u_2\hat{\mathbf{e}}_2 + u_3\hat{\mathbf{e}}_3$,
where the unit vectors $\hat{\mathbf{e}}_i$ form an **orthonormal basis** for $\mathbb{R}^3$.

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$$

The differential displacement vector **dr** is (by the Chain Rule)

$$\mathbf{d}\vec{\mathbf{r}} = \frac{\partial \vec{\mathbf{r}}}{\partial u_1}du_1 + \frac{\partial \vec{\mathbf{r}}}{\partial u_2}du_2 + \frac{\partial \vec{\mathbf{r}}}{\partial u_3}du_3 = h_1\,du_1\,\hat{\mathbf{e}}_1 + h_2\,du_2\,\hat{\mathbf{e}}_2 + h_3\,du_3\,\hat{\mathbf{e}}_3$$

and the differential arc length ds is given by

$$ds^2 = \mathbf{d}\vec{\mathbf{r}} \bullet \mathbf{d}\vec{\mathbf{r}} = \left(h_1\,du_1\right)^2 + \left(h_2\,du_2\right)^2 + \left(h_3\,du_3\right)^2$$

The element of volume dV is

$$dV = h_1h_2h_3\,du_1du_2du_3 = \underbrace{\left|\frac{\partial(x,y,z)}{\partial(u_1,u_2,u_3)}\right|}_{\text{Jacobian}} du_1du_2du_3$$

$$= \begin{vmatrix} \dfrac{\partial x}{\partial u_1} & \dfrac{\partial y}{\partial u_1} & \dfrac{\partial z}{\partial u_1} \\ \dfrac{\partial x}{\partial u_2} & \dfrac{\partial y}{\partial u_2} & \dfrac{\partial z}{\partial u_2} \\ \dfrac{\partial x}{\partial u_3} & \dfrac{\partial y}{\partial u_3} & \dfrac{\partial z}{\partial u_3} \end{vmatrix} du_1du_2du_3$$

<u>Example 5.02.1</u>: Find the scale factor h_θ for the spherical polar coordinate system $(x, y, z) = (r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)$:

$$\frac{\partial \vec{\mathbf{r}}}{\partial \theta} = \begin{bmatrix} \dfrac{\partial x}{\partial \theta} & \dfrac{\partial y}{\partial \theta} & \dfrac{\partial z}{\partial \theta} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \end{bmatrix}^{\mathrm{T}}$$

$$\Rightarrow\ h_\theta = \left|\frac{\partial \vec{\mathbf{r}}}{\partial \theta}\right| = \sqrt{r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta}$$

$$= \sqrt{r^2\cos^2\theta\,\cancel{\left(\cos^2\phi + \sin^2\phi\right)} + r^2\sin^2\theta} = \sqrt{r^2\left(\cos^2\theta + \sin^2\theta\right)} = r$$

5.03 Summary Table for the Gradient Operator

Gradient operator $$\vec{\boldsymbol{\nabla}} = \frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial}{\partial u_1} + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial}{\partial u_2} + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial}{\partial u_3}$$

Gradient $$\vec{\boldsymbol{\nabla}} V = \frac{\hat{\mathbf{e}}_1}{h_1}\frac{\partial V}{\partial u_1} + \frac{\hat{\mathbf{e}}_2}{h_2}\frac{\partial V}{\partial u_2} + \frac{\hat{\mathbf{e}}_3}{h_3}\frac{\partial V}{\partial u_3}$$

Divergence $$\vec{\boldsymbol{\nabla}} \bullet \vec{\mathbf{F}} = \frac{1}{h_1 h_2 h_3}\left(\frac{\partial(h_2 h_3 F_1)}{\partial u_1} + \frac{\partial(h_3 h_1 F_2)}{\partial u_2} + \frac{\partial(h_1 h_2 F_3)}{\partial u_3}\right)$$

Curl $$\vec{\boldsymbol{\nabla}} \times \vec{\mathbf{F}} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & \dfrac{\partial}{\partial u_1} & h_1 F_1 \\ h_2 \hat{\mathbf{e}}_2 & \dfrac{\partial}{\partial u_2} & h_2 F_2 \\ h_3 \hat{\mathbf{e}}_3 & \dfrac{\partial}{\partial u_3} & h_3 F_3 \end{vmatrix}$$

Laplacian $$\boldsymbol{\nabla}^2 V = \frac{1}{h_1 h_2 h_3}\left(\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial V}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial V}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial V}{\partial u_3}\right)\right)$$

<u>Scale factors</u>:

Cartesian: $h_x = h_y = h_z = 1$.

Cylindrical polar: $h_\rho = h_z = 1\,,\ h_\phi = \rho$.

Spherical polar: $h_r = 1\,,\ \ h_\theta = r\,,\ \ h_\phi = r\sin\theta$.

<u>Example 5.03.1</u>: The Laplacian of V in spherical polars is

$$\nabla^2 V = \frac{1}{r^2\sin\theta}\left(\frac{\partial}{\partial r}\left(r^2\sin\theta\,\frac{\partial V}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial V}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\phi}\right)\right)$$

or $$\nabla^2 V = \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2 V}{\partial\theta^2} + \cot\theta\,\frac{\partial V}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2}$$

Example 5.03.2

A potential function $V(\vec{\mathbf{r}})$ is spherically symmetric, (that is, its value depends only on the distance r from the origin), due solely to a point source at the origin. There are no other sources or sinks anywhere in $\mathbb{R}^3$. Deduce the functional form of $V(\vec{\mathbf{r}})$.

$V(\vec{\mathbf{r}})$ is spherically symmetric $\Rightarrow$ $V(r, \theta, \phi) = f(r)$

In any regions not containing any sources of the vector field, the divergence of the vector field $\vec{\mathbf{F}} = \vec{\nabla} V$ (and therefore the Laplacian of the associated potential function V) must be zero. Therefore, for all $r \neq 0$, $\operatorname{div} \vec{\mathbf{F}} = \vec{\nabla} \bullet \vec{\nabla} V = \nabla^2 V = 0$

But

$$\nabla^2 V = \frac{1}{r^2 \sin\theta} \left(\frac{\partial}{\partial r} \left(r^2 \sin\theta \frac{\partial V}{\partial r} \right) + \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial V}{\partial \theta} \right) + \frac{\partial}{\partial \phi} \left(\frac{1}{\sin\theta} \frac{\partial V}{\partial \phi} \right) \right)$$

$$\Rightarrow \quad \nabla^2 V = \frac{1}{r^2 \cancel{\sin\theta}} \left(\frac{d}{dr} \left(r^2 \cancel{\sin\theta} \frac{dV}{dr} \right) + 0 + 0 \right) = 0$$

$$\Rightarrow \quad \frac{d}{dr} \left(r^2 \frac{dV}{dr} \right) = 0 \quad \Rightarrow \quad r^2 \frac{dV}{dr} = B \quad \Rightarrow \quad \frac{dV}{dr} = B r^{-2}$$

$\Rightarrow \quad V = \dfrac{B r^{-1}}{-1} + A$, where A, B are arbitrary constants of integration.

Therefore the potential function must be of the form

$$\boxed{V(r, \theta, \phi) = A - \frac{B}{r}}$$

This is the standard form of the potential function associated with a force that obeys the inverse square law $F \propto \dfrac{1}{r^2}$.

5.04 Derivatives of Basis Vectors

Cartesian: $\frac{d}{dt}\hat{\mathbf{i}} = \frac{d}{dt}\hat{\mathbf{j}} = \frac{d}{dt}\hat{\mathbf{k}} = \vec{\mathbf{0}}$

$$\vec{\mathbf{r}} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$$

$$\Rightarrow \quad \vec{\mathbf{v}} = \dot{x}\hat{\mathbf{i}} + \dot{y}\hat{\mathbf{j}} + \dot{z}\hat{\mathbf{k}}$$

Cylindrical Polar Coordinates:

$x = \rho\cos\phi\,,\quad y = \rho\sin\phi\,,\quad z = z$

$$\frac{d}{dt}\hat{\boldsymbol{\rho}} = \frac{d\phi}{dt}\hat{\boldsymbol{\phi}}$$

$$\frac{d}{dt}\hat{\boldsymbol{\phi}} = -\frac{d\phi}{dt}\hat{\boldsymbol{\rho}}$$

$$\frac{d}{dt}\hat{\mathbf{k}} = \vec{\mathbf{0}}$$

$$\mathbf{r} = \rho\,\hat{\boldsymbol{\rho}} + z\,\hat{\mathbf{k}}$$

$$\Rightarrow \quad \vec{\mathbf{v}} = \dot{\rho}\,\hat{\boldsymbol{\rho}} + \rho\dot{\phi}\,\hat{\boldsymbol{\phi}} + \dot{z}\,\hat{\mathbf{k}}$$

[radial and transverse components of $\vec{\mathbf{v}}$]

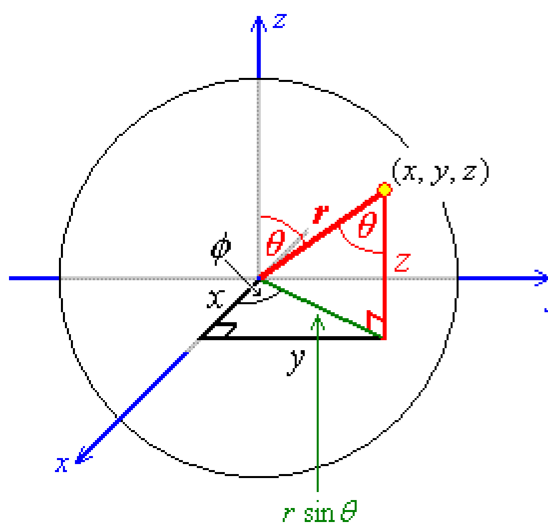
Spherical Polar Coordinates.

The "declination" angle θ is the angle between the positive z axis and the radius vector $\vec{\mathbf{r}}$. $0 \le \theta \le \pi$.

The "azimuth" angle ϕ is the angle on the x-y plane, measured anticlockwise from the positive x axis, of the shadow of the radius vector. $0 \le \phi < 2\pi$.

$$z = r\cos\theta\,.$$

The shadow of the radius vector on the x-y plane has length $r\sin\theta$.



It then follows that

$$x = r\sin\theta\cos\phi \quad \text{and} \quad y = r\sin\theta\sin\phi\,.$$

$$\frac{d}{dt}\hat{\mathbf{r}} = \frac{d\theta}{dt}\hat{\boldsymbol{\theta}} + \frac{d\phi}{dt}\sin\theta\,\hat{\boldsymbol{\phi}}$$

$$\frac{d}{dt}\hat{\boldsymbol{\theta}} = -\frac{d\theta}{dt}\hat{\mathbf{r}} + \frac{d\phi}{dt}\cos\theta\,\hat{\boldsymbol{\phi}}$$

$$\frac{d}{dt}\hat{\boldsymbol{\phi}} = -\frac{d\phi}{dt}\left(\sin\theta\,\hat{\mathbf{r}} + \cos\theta\,\hat{\boldsymbol{\theta}}\right)$$

$$\vec{\mathbf{r}} = r\,\hat{\mathbf{r}}$$

$$\Rightarrow \quad \vec{\mathbf{v}} = \dot{r}\,\hat{\mathbf{r}} + r\dot{\theta}\,\hat{\boldsymbol{\theta}} + r\dot{\phi}\sin\theta\,\hat{\boldsymbol{\phi}}$$

Example 5.04.1

Find the velocity and acceleration in cylindrical polar coordinates for a particle travelling along the helix $x = 3\cos 2t$, $y = 3\sin 2t$, $z = t$.

Cylindrical polar coordinates: $x = \rho\cos\phi$, $y = \rho\sin\phi$, $z = z$

$$\Rightarrow \quad \rho^2 = x^2 + y^2, \quad \tan\phi = \frac{y}{x}$$

$$\rho^2 = 9\cos^2 2t + 9\sin^2 2t = 9 \quad \Rightarrow \quad \rho = 3 \quad \Rightarrow \quad \dot{\rho} = 0$$

$$\tan\phi = \frac{3\sin 2t}{3\cos 2t} = \tan 2t \quad \Rightarrow \quad \phi = 2t \quad \Rightarrow \quad \dot{\phi} = 2$$

$$z = t \quad \Rightarrow \quad \dot{z} = 1$$

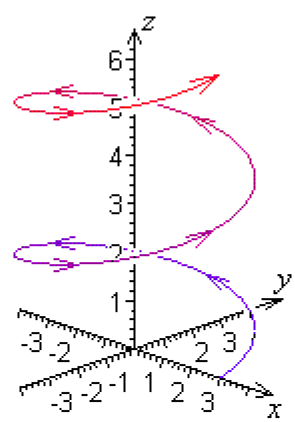
$$\Rightarrow \quad \vec{\mathbf{r}} = 3\hat{\boldsymbol{\rho}} + z\hat{\mathbf{k}}$$

$$\Rightarrow \quad \vec{\mathbf{v}} = \frac{d\vec{\mathbf{r}}}{dt} = \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\dot{\phi}\hat{\boldsymbol{\phi}} + \dot{z}\hat{\mathbf{k}} = 0\hat{\boldsymbol{\rho}} + 3\times 2\hat{\boldsymbol{\phi}} + 1\hat{\mathbf{k}} = \underline{\underline{6\hat{\boldsymbol{\phi}} + \hat{\mathbf{k}}}}$$

[The velocity has no radial component – the helix remains the same distance from the z axis at all times.]

$$\vec{\mathbf{a}} = \frac{d\vec{\mathbf{v}}}{dt} = 6\dot{\hat{\boldsymbol{\phi}}} + \dot{\hat{\mathbf{k}}} = -6\dot{\phi}\hat{\boldsymbol{\rho}} + \vec{\mathbf{0}} = \underline{\underline{-12\hat{\boldsymbol{\rho}}}}$$

[The acceleration vector points directly at the z axis at all times.]



Other examples are in the problem sets.

END OF CHAPTER 5

6. Calculus of Variations

The method of calculus of variations involves finding the path between two points that provides the minimum (or maximum) value of integrals of the form

$$\int_a^b F(x, y, y')\, dx$$

Sections in this Chapter:

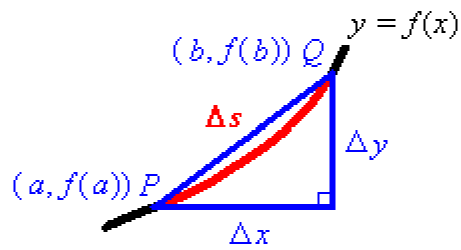
Sections for reference; **not** examinable:

6.01 Introduction

Example 6.01.1

To find the shortest path, (the **geodesic**), between two points, we need to find an expression for the arc length along a path between the two points.

Consider a pair of nearby points.
The element of arc length Δs is approximately the hypotenuse of the triangle.



$$(\Delta s)^2 \approx (\Delta x)^2 + (\Delta y)^2$$

$$\Rightarrow \frac{(\Delta s)^2}{(\Delta x)^2} \approx \frac{(\Delta x)^2}{(\Delta x)^2} + \frac{(\Delta y)^2}{(\Delta x)^2}$$

In the limit as the two points approach each other and $\Delta x \to 0$, we obtain

$$\left(\frac{ds}{dx}\right)^2 = 1 + \left(\frac{dy}{dx}\right)^2$$

$$\Rightarrow \left(\frac{ds}{dx}\right) = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

The arc length s between any two points $x = a$ and $x = b$ along any path C in $\mathbb{R}^2$ is the line integral

$$s = \int_C ds = \int_C \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx = \int_C \sqrt{1 + \left(f'(x)\right)^2}\, dx \quad \text{where } C \text{ is the path } y = f(x)$$

The geodesic will be the path C for which the line integral for s attains its minimum value. Of course, in a flat space such as $\mathbb{R}^2$, that geodesic is just the straight line between the two points.

6.02 Theory

We wish to find the curve $y(x)$ which passes through the points (x_0, y_0) and (x_1, y_1) and which minimizes the integral

$$I = \int_{x_0}^{x_1} F\big(x, y(x), y'(x)\big)\,dx$$

Consider the one parameter family of curves $y(x) = u(x) + \alpha\,\eta(x)$, where α is a real parameter, $\eta(x)$ is an arbitrary function except for the requirement $\eta(x_0) = \eta(x_1) = 0$ and $u(x)$ represents the (as yet unknown) solution.

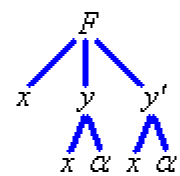
Every member of this family of curves passes through the points (x_0, y_0) and (x_1, y_1). For any member of the family,

$$I(\alpha) = \int_{x_0}^{x_1} F\big(x, u(x) + \alpha\,\eta(x), u'(x) + \alpha\,\eta'(x)\big)\,dx$$

we know that $y(x) = u(x)$ minimizes I.

Therefore the minimum for I occurs when $\alpha = 0$, so that $\left.\dfrac{dI}{d\alpha}\right|_{\alpha=0} = 0$.

Carrying out a Leibnitz differentiation of the integral $I(\alpha)$,



$$\frac{dI}{d\alpha} = 0 - 0 + \int_{x_0}^{x_1} \frac{\partial}{\partial \alpha} F\big(x, u(x) + \alpha\,\eta(x), u'(x) + \alpha\,\eta'(x)\big)\,dx$$

$$= \int_{x_0}^{x_1} \left[0 + \frac{\partial F}{\partial y}\frac{\partial}{\partial \alpha}\big(u(x) + \alpha\,\eta(x)\big) + \frac{\partial F}{\partial y'}\frac{\partial}{\partial \alpha}\big(u'(x) + \alpha\,\eta'(x)\big)\right] dx$$

At the minimum $\alpha = 0$, so that $y(x) = u(x)$ and $y'(x) = u'(x)$. Therefore

$$0 = \int_{x_0}^{x_1} \left[\eta(x)\frac{\partial F}{\partial u} + \eta'(x)\frac{\partial F}{\partial u'} \right] dx$$

Also note, by the product rule of differentiation, that

$$\frac{d}{dx}\left(\eta(x)\frac{\partial F}{\partial u'}\right) = \eta'(x)\frac{\partial F}{\partial u'} + \eta(x)\frac{d}{dx}\left(\frac{\partial F}{\partial u}\right)$$

Therefore the integral can be written as

$$0 = \int_{x_0}^{x_1} \left[\eta(x)\frac{\partial F}{\partial u} + \frac{d}{dx}\left(\eta(x)\frac{\partial F}{\partial u'}\right) - \eta(x)\frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right] dx$$

$$0 = \int_{x_0}^{x_1} \eta(x)\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right] dx + \int_{x_0}^{x_1} \frac{d}{dx}\left(\eta(x)\frac{\partial F}{\partial u'}\right) dx$$

$$0 = \int_{x_0}^{x_1} \eta(x)\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right] dx + \left[\eta(x)\frac{\partial F}{\partial u'}\right]_{x_0}^{x_1}$$

But $\eta(x_0) = \eta(x_1) = 0$

Therefore the minimizing curve $u(x)$ satisfies

$$\int_{x_0}^{x_1} \eta(x)\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)\right] dx = 0$$

But $\eta(x)$ is an arbitrary function of x, which leads to

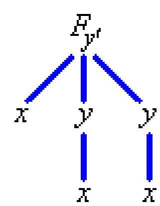
$$\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) = 0$$

Thus, if $y = f(x)$ is a path that minimizes the integral $\int_a^b F(x, y, y')\, dx$, then $y = f(x)$ and $F(x, y, y')$ must satisfy the **Euler equation for extremals**

$$\boxed{\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0}$$

Euler's equation requires the assumption that $F(x, y, y')$ has continuous second derivatives in all three of its variables and that all members of the family $y(x) = u(x) + \alpha\, \eta(x)$ have continuous second derivatives.

Expansion of Euler's Equation:



$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\big(x, y(x), y'(x)\big)\right) - \frac{\partial F}{\partial y}\big(x, y(x), y'(x)\big) = 0$$

$$\Rightarrow \quad \frac{\partial^2 F}{\partial x\, \partial y'} + y'(x)\frac{\partial^2 F}{\partial y\, \partial y'} + y''(x)\frac{\partial^2 F}{\partial y'^2} - \frac{\partial F}{\partial y} = 0$$

or

$$\boxed{y'' F_{y'y'} + y' F_{yy'} + \left(F_{xy'} - F_y\right) = 0}$$

Note: Leibnitz differentiation of $I(z) = \int_{f(z)}^{g(z)} F(x, z)\, dx$ with respect to z is:

$$\frac{dI}{dz} = g'(z) F\big(g(z), z\big) - f'(z) F\big(f(z), z\big) + \int_{f(z)}^{g(z)} \frac{\partial}{\partial z} F(x, z)\, dx$$

A special case of this is $\frac{d}{dx}\int_a^x f(t)\, dt = f(x)$.
