<u>Variational Calculus</u> deals with functionals, i.e. functions of functions, in the form of a definite integral whose integrand contains a function $u(x)$ that is yet to be determined, such as:

$$I[u(x)] = \int_{x_0}^{x_1} F[x, u(x), u'(x)]\,dx$$

and it is utilized when searching for the function $u(x)$ for which the functional is an extremum.

e.g. Fermat's principle of optics states that the path of light $y(x)$ between two points is the one requires the minimum travel time:

$$T[y(x)] = \int_{t_0}^{t_1} dt = \int_{x_0}^{x_1} \frac{ds}{v(x,y)} = \int_{x_0}^{x_1} \frac{1}{v(x,y)} \underbrace{\sqrt{1+(y')^2}\,dx}_{ds}$$

where $v(x,y)$ is the speed of light in the medium.

As a typical variational calculus problem, consider the functional

$$I[u(x)] = \int_{x_0}^{x_1} F[x, u(x), u'(x)]\,dx$$

where $u(x)$ is smooth (diffrentible) family of curves passing through the specified end points $u(x_0) = u_0$ and $u(x_1) = u_1$ which are classified as Dirichlet or <u>essential</u> boundary conditions.

Among these functions (curves) we search for the function $u(x)$ for which the functional $I[u(x)]$ is an <u>extremum</u> or $u(x)$ is a <u>stationary function</u> for $I[u(x)]$.

For this purpose, let us introduce neighborhood functions $\hat{u}(x)$ that is "close" to the actual, yet unknown, stationary function $u(x)$ defined as

$$\hat{u}(x) = u(x) + \left.\tfrac{\partial u}{\partial \varepsilon}\right|_{\varepsilon=0} \varepsilon + O(\varepsilon^2)$$

where ε is a small parameter.

The two functions $u(x)$ and $\hat{u}(x)$ are close in the sense of some norm, the $O(\varepsilon) \equiv \left.\tfrac{\partial u}{\partial \varepsilon}\right|_{\varepsilon=0} \varepsilon$ term is called the "variation of u"

$$\delta u \equiv \hat{u}(x) - u(x) = \left.\tfrac{\partial u}{\partial \varepsilon}\right|_{\varepsilon=0} \varepsilon \equiv \varepsilon \eta(x)$$

where $\eta(x)$ denote suitable functions with $\eta(x_0) = \eta(x_1) = 0$.
The functional (Frechet) derivative of $I[u(x)]$ is then defined by

$$\begin{aligned}
\left.\tfrac{dI}{d\varepsilon}\right|_{\varepsilon=0} &= \lim_{\varepsilon\to 0} \tfrac{I[u+\varepsilon\eta]-I[u]}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \tfrac{1}{\varepsilon} \int_{x_0}^{x_1} \{F[x, u+\varepsilon\eta, u'+\varepsilon\eta'] - F[x,u,u']\}\,dx \\
&= \lim_{\varepsilon\to 0} \tfrac{1}{\varepsilon} \int_{x_0}^{x_1} \{\varepsilon\eta \tfrac{\partial F}{\partial u} + \varepsilon \tfrac{d\eta}{dx} \tfrac{\partial F}{\partial u'} + O(\varepsilon^2)\}\,dx = \int_{x_0}^{x_1} \{\eta \tfrac{\partial F}{\partial u} + \tfrac{\partial \eta}{\partial x} \tfrac{\partial F}{\partial u'}\}\,dx \\
&= \underbrace{\left[\eta \tfrac{\partial F}{\partial u'}\right]_{x_0}^{x_1}}_{0} + \int_{x_0}^{x_1} \eta \{\tfrac{\partial F}{\partial u} - \tfrac{d}{dx} \tfrac{\partial F}{\partial u'}\}\,dx
\end{aligned}$$

Stationarity condition results in the Euler-Lagrange equations,

$$\left.\tfrac{dI}{d\varepsilon}\right|_{\varepsilon=0} = 0 \quad \Rightarrow \quad \tfrac{\partial F}{\partial u} - \tfrac{d}{dx} \tfrac{\partial F}{\partial u'} = 0$$

as a necessary condition for an extremum by the Fundamental Lemma.

<u>Fundamental Lemma</u> of the Calculus of Variations
If $g(x)$ is a continuous function in $x_0 \le x \le x_1$ and if

$$\int_{x_0}^{x_1} g(x)h(x)\,dx = 0$$

where $h(x)$ is an arbitrary function in the same interval with $h(x_0) = h(x_1) = 0$, then $g(x) = 0$ at every point in the interval.

Euler-Lagrange equation may be regarded as an equivalent differential form of the variational form $I[u(x)]$. An equivalent form is

$$\left(\tfrac{\partial^2 F}{\partial u'^2}\right) \tfrac{d^2 u}{dx^2} + \left(\tfrac{\partial^2 F}{\partial u' \partial u}\right) \tfrac{du}{dx} + \left(\tfrac{\partial^2 F}{\partial u' \partial x} - \tfrac{\partial F}{\partial u}\right) = 0$$

the 2nd order ODE for $u(x)$. It may be linear or nonlinear depending on the nature of the integrand $F = F[x, u, u']$.

e.g. Verify that the stationary function $u(x)$ for the functional

$$I[u] = \int_0^1 [(xu')^2 + 2u^2]\,dx \quad \text{subject to} \quad u(0) = 0 \quad \text{and} \quad u(1) = 2$$

satisfies the ODE: $x^2 u'' + 2xu' - 2u = 0$. Find $u(x)$.

e.g. Verify that the problem of finding a plane curve $u = u(x)$ with the shortest length connecting the two points (x_0, u_0) and (x_1, u_1) is formulated as the functional

$$I[u] = \int_{x_0}^{x_1} ds = \int_{x_0}^{x_1} \sqrt{1+(u')^2}\, dx$$

and leads to the Euler-Lagrange equation: $u' = \text{constant}$. Find $u(x)$.

e.g. Verify that the problem of finding a minimum surface area by revolving the plane curve $u = u(x)$ around the x-axis can be formulated as the functional

$$I[u] = \iint_A dA = \int_{x_1}^{x_2} 2\pi u(x) ds = \int_{x_0}^{x_1} 2\pi u(x)\sqrt{1+(u')^2}\, dx$$

and leads to the Euler-Lagrange equation: $uu'' - (u')^2 = 1$. Find $u(x)$ by change of variables from (x,u) to (u,p) using $u' = p$ and thus $u'' = p\frac{dp}{du}$.

Natural Boundary Conditions arise when the values of the stationary function $u = u(x)$ are not specified at the endpoints x_0 and x_1. Consider the total variation of $F[x,u,u']$ by $\delta F = \left(\frac{\partial F}{\partial x}\right)\delta x + \left(\frac{\partial F}{\partial u}\right)\delta u + \left(\frac{\partial F}{\partial u'}\right)\delta u'$ where $\delta u = \hat{u}(x) - u(x) \equiv \varepsilon\eta(x)$.

Since x is an independent variable and so it does not change, $\delta x = 0$, and since

$$\delta u' = \delta\left(\frac{du}{dx}\right) = \frac{d}{dx}\left(\delta u\right) \quad \text{or} \quad \delta u' = \varepsilon\eta'(x) = \frac{d}{dx}\left(\varepsilon\eta\right) = \frac{d}{dx}\left(\delta u\right),$$

we have $\delta F = \left(\frac{\partial F}{\partial u}\right)\delta u + \left(\frac{\partial F}{\partial u'}\right)(\delta u)'$. Now, the extremization of the functional

$$I[u] = \int_{x_0}^{x_1} F[x,u,u']\, dx$$

leads to

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

Now, for u specified at the boundaries $\delta u\big|_{x_0} = \delta u\big|_{x_1} = 0$, otherwise $\frac{\partial F}{\partial u'}\big|_{x_0} = \frac{\partial F}{\partial u'}\big|_{x_1} = 0$ required for the vanishing boundary terms to have left with Euler-Lagrange eqn. These later conditions which are not dictated externally (essentially) are called natural (or Neumann or nonessential) boundary conditions.

e.g. Verify that the extremization of the functional

$$I[u] = \frac{1}{2}\int_0^{\frac{\pi}{2}} \left\{-(u')^2 + u^2 + 2xu\right\} dx$$

where $u(0)$ and $u(\frac{\pi}{2})$ are not specified leads to the natural b.c. $u'(0)$ and $u'(\frac{\pi}{2})$. Find $u(x)$.

In addition to having b.c. not specified at the boundaries, the b.c. may be added to the functional to be included in the extremum problem. This is called dual functional problem.

e.g. Verify that the extremization of the functional

$$I[u] = \frac{1}{2}[u(1)]^2 + \frac{1}{2}\int_0^1 \left\{(u')^2\right\} dx$$

where $u(0) = 1$ and $u(1)$ is not specified but to be minimized along with the definite integral, leads to

$$0 = \delta I = u(1)\,\delta u\big|_{x=1} + \int_0^1 \left\{u'\,\delta u'\right\} dx$$

with the natural b.c. $u'(1) + u(1) = 0$ and the ODE $u'' = 0$. Find $u(x)$.

Variable Endpoints arise when the functional

$$I[u] = \int_{x_0}^{x_1} F[x,u,u']\, dx$$

is to be extremalized when the endpoints x_0 and x_1 vary along some functions $f(x)$ and $g(x)$, thus $u_0 = u(x_0) = f(x_0)$ and $u_1 = u(x_1) = g(x_1)$. Now

$$\delta I = I[u + \delta u] - I[u]$$

becomes

$$\begin{aligned}
\delta I &= \int_{x_0+\delta x_0}^{x_1+\delta x_1} F[x, u+\delta u, u'+\delta u']dx - \int_{x_0}^{x_1} F[x,u,u']dx \\
&= \int_{x_0}^{x_1} F[x, u+\delta u, u'+\delta u']dx - \int_{x_0}^{x_1} F[x,u,u']dx \\
&+ \int_{x_1}^{x_1+\delta x_1} F[x, u+\delta u, u'+\delta u']dx - \int_{x_0}^{x_0+\delta x_0} F[x, u+\delta u, u'+\delta u']dx
\end{aligned}.$$

and further

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

The typical argument used is that

$$\int_{x_1}^{x_1+\delta x_1} F\left[x,u+\delta u,u'+\delta u'\right]dx \sim \int_{x_1}^{x_1+\delta x_1}\left\{F\left[x,u,u'\right]+\left(\tfrac{\partial F}{\partial u}\right)\delta u+\left(\tfrac{\partial F}{\partial u'}\right)\delta u'\right\}dx$$

$$\sim \left[F\left[x,u,u'\right]+\left(\tfrac{\partial F}{\partial u}\right)\delta u+\left(\tfrac{\partial F}{\partial u'}\right)\delta u'\right]_{x_1}\delta x_1 \sim \left[F\right]_{x_1}\delta x_1$$

where quadratic terms, such as $\delta u\,\delta x$ and $\delta u'\delta x$, are ignored.

Also at a typical boundary, say, at x_0 where $u(x_0)=f(x_0)$, the variation is

$$u(x_0+\delta x_0)+\varepsilon\eta(x_0+\delta x_0)=f(x_0+\delta x_0).$$

It can also be written as

$$u(x_0+\delta x_0)-u(x_0)+\varepsilon\eta(x_0+\delta x_0)=f(x_0+\delta x_0)-f(x_0)$$

that becomes

$$u'\big|_{x_0}\delta x_0+\varepsilon\,\eta\big|_{x_0}=f'\big|_{x_0}\delta x_0 \quad\Rightarrow\quad \delta u_0\equiv\varepsilon\,\eta\big|_{x_0}=(f'-u')\big|_{x_0}\delta x_0$$

where again quadratic terms, such as $\varepsilon\delta x$, are ignored.

These arguments lead to

$$\delta I=\int_{x_0}^{x_1}\left\{\tfrac{\partial F}{\partial u}-\tfrac{d}{dx}\left(\tfrac{\partial F}{\partial u'}\right)\right\}\delta u\,dx+\left[\left(\tfrac{\partial F}{\partial u'}\right)(g'-u')+F\right]_{x_1}\delta x_1-\left[\left(\tfrac{\partial F}{\partial u'}\right)(f'-u')+F\right]_{x_0}\delta x_0.$$

For extremum, we have the Euler-Lagrange eqn. $\frac{\partial F}{\partial u}-\frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right)=0$ and the transversality conditions

$$\left[\left(\tfrac{\partial F}{\partial u'}\right)(f'-u')+F\right]_{x_0}=0 \quad\text{and}\quad \left[\left(\tfrac{\partial F}{\partial u'}\right)(g'-u')+F\right]_{x_1}=0.$$

e.g. Recall previously that the shortest distance problem $F\left[x,u,u'\right]=\sqrt{1+(u')^2}$ led to $u'=\text{constant}$. Now let the endpoints vary along $f(x)=x-3$ and $g(x)=\exp(x)$, respectively. Write down the transversality conditions and find $u(x)$.

Higher-Order Dervatives can also be handled similarly as follows: Consider

$$I\left[u\right]=\int_{x_0}^{x_1}F\left[x,u,u',u''\right]dx.$$

It leads to

$$0=\delta I=\int_{x_0}^{x_1}\left\{\left(\tfrac{\partial F}{\partial u}\right)\delta u+\left(\tfrac{\partial F}{\partial u'}\right)(\delta u)'+\left(\tfrac{\partial F}{\partial u''}\right)(\delta u)''\right\}dx$$

and further to the Euler-Lagrange equation

$$\tfrac{\partial F}{\partial u}-\tfrac{d}{dx}\left(\tfrac{\partial F}{\partial u'}\right)+\tfrac{d^2}{dx^2}\left(\tfrac{\partial F}{\partial u''}\right)=0.$$

The essential boundary conditions $\delta u\big|_{x_i}=0$ and $\delta u'\big|_{x_i}=0$ and/or the natural boundary conditions $\frac{d}{dx}\left(\frac{\partial F}{\partial u''}\right)-\frac{\partial F}{\partial u'}\Big|_{x_i}=0$ and $\frac{\partial F}{\partial u''}\Big|_{x_i}=0$ may be required for $i=0,1$.

e.g. Verify the derivations above.

Two Independent Variables may also arise in variational formulations:

$$I\left[u\right]=\iint_A F\left[x,y,u,u_x,u_y\right]dxdy.$$

It leads to

$$0=\delta I=\iint_A\left\{\left(\tfrac{\partial F}{\partial u}\right)\delta u+\left(\tfrac{\partial F}{\partial u_x}\right)\delta u_x+\left(\tfrac{\partial F}{\partial u_y}\right)\delta u_y\right\}dxdy$$

where $\delta u\equiv\hat{u}(x,y)-u(x,y)=\varepsilon\eta(x,y)$. In particular, using the argument

$$\tfrac{\partial F}{\partial u_x}\delta u_x=\tfrac{\partial F}{\partial u_x}\delta\left(\tfrac{\partial u}{\partial x}\right)=\tfrac{\partial F}{\partial u_x}\tfrac{\partial}{\partial x}\left(\delta u\right)=\tfrac{\partial}{\partial x}\left(\tfrac{\partial F}{\partial u_x}\delta u\right)-\tfrac{\partial}{\partial x}\left(\tfrac{\partial F}{\partial u_x}\right)\delta u$$

and similarly for δu_y term, it reduces to

$$0=\iint_A\left\{\left(\tfrac{\partial F}{\partial u}-\tfrac{\partial}{\partial x}\left(\tfrac{\partial F}{\partial u_x}\right)-\tfrac{\partial}{\partial y}\left(\tfrac{\partial F}{\partial u_y}\right)\right)\delta u+\nabla\cdot\left[\tfrac{\partial F}{\partial u_x}\delta u \quad \tfrac{\partial F}{\partial u_y}\delta u\right]\right\}dxdy$$

$$=\iint_A\left\{\left(\tfrac{\partial F}{\partial u}-\tfrac{\partial}{\partial x}\left(\tfrac{\partial F}{\partial u_x}\right)-\tfrac{\partial}{\partial y}\left(\tfrac{\partial F}{\partial u_y}\right)\right)\delta u\right\}dxdy+\oint_C \mathbf{n}\cdot\left[\tfrac{\partial F}{\partial u_x}\delta u \quad \tfrac{\partial F}{\partial u_y}\delta u\right]ds$$

using planar divergence theorem

$$\iint_A \nabla\cdot\mathbf{v}\,dA=\oint_C \mathbf{n}\cdot\mathbf{v}\,ds$$

where $\mathbf{n}$ denotes the outward normal vector on curve C bounding the planar region A and s denotes the arclength parameter.

For given parametric representation of C: $\mathbf{R}(s)=\left[x(s) \quad y(s)\right]$, the normal vector is $\mathbf{n}(s)=\left[y'(s) \quad -x'(s)\right]$, while tangent vector is $\mathbf{t}(s)=\left[x'(s) \quad y'(s)\right]$, so that $\mathbf{n}\cdot\mathbf{t}=0$, the boundary term becomes

$$\oint_C \mathbf{n}\cdot\left[\tfrac{\partial F}{\partial u_x}\delta u \quad \tfrac{\partial F}{\partial u_y}\delta u\right]ds=\oint_C\left[\tfrac{\partial F}{\partial u_x}\tfrac{dy}{ds}-\tfrac{\partial F}{\partial u_y}\tfrac{dx}{ds}\right]\delta u\, ds.$$

The extremum of I is then obtained when Euler-Lagrange equation

$$\tfrac{\partial F}{\partial u}-\tfrac{\partial}{\partial x}\left(\tfrac{\partial F}{\partial u_x}\right)-\tfrac{\partial}{\partial y}\left(\tfrac{\partial F}{\partial u_y}\right)=0$$

is satisfied subject to the essential boundary conditions such as $u(x,y)=u_0(x,y)$ on C implying $\delta u=0$, or natural boundary conditions

$$\tfrac{\partial F}{\partial u_x}\tfrac{dy}{ds}-\tfrac{\partial F}{\partial u_y}\tfrac{dx}{ds}=0 \quad \text{on} \quad C.$$

e.g. Construct the Euler-Lagrange equation together with the complete set of boundary coditions for the extremization of the functional

$$I\left[u\right]=\iint_A \tfrac{1}{2}\left[u_x^2+u_xu_y+u^2+2xu\right]dxdy$$

subject to $u(0,y)=3y$ and $u(x,0)=2x$ for the unit square region $A=\left\{(x,y)\,\middle|\,0\le x,y\le 1\right\}$.

Constrained Functionals arises when a stationary function is to be determined in the presence of a constraint. This constraint may be an integral, a differential or an algebraic expression. Consider the functional

$$I\left[u\right]=\int_{x_0}^{x_1}F\left[x,u,u'\right]dx.$$

subject to a constraint expressed as a definite integral

$$\int_{x_0}^{x_1}G\left[x,u,u'\right]dx=K$$

where K is a specified constant. Lagrange multiplier approach may be utilized as follows: Let λ be an arbitrary constant so that the constraint can be attached to the functional and extremalized

$$0=\delta\left\{\int_{x_0}^{x_1}F\left[x,u,u'\right]dx+\lambda\left(\int_{x_0}^{x_1}G\left[x,u,u'\right]dx-K\right)\right\}=\delta\left\{\int_{x_0}^{x_1}\tilde{F}\left[x,u,u'\right]dx\right\}$$

where $\tilde{F}\left[\lambda,x,u,u'\right]=F\left[x,u,u'\right]+\lambda G\left[x,u,u'\right]$. This leads to

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

that together with the essential boundary conditions, say, $u(x_0)=u_0$ and $u(x_1)=u_1$ determines λ and the stationary function $u(x)$.

e.g. Consider the problem of constructing the curve $u(x)$ passing through $u(-1)=u(1)=0$ that has minimum length and encloses an area $A=\frac{\pi}{2}$ with the x-axis. Thus,

$$I\left[u\right]=\int_{-1}^{1}ds=\int_{-1}^{1}\sqrt{1+\left(u'\right)^2}\,dx \quad \text{subject to} \quad \int_{-1}^{1}u(x)\,dx=\tfrac{\pi}{2}.$$

Show that it leads to “circular arc” as the stationary function.

e.g. Consider now the functional

$$I\left[u\right]=\int_{x_0}^{x_1}\left[q(x)u^2-p(x)\left(u'\right)^2\right]dx \quad \text{subject to} \quad \int_{x_0}^{x_1}\omega(x)u^2dx=1$$

where p,q,ω are known functions. Show that the extremalization of I leads to

$$\tfrac{d}{dx}\left(p(x)\tfrac{du}{dx}\right)+\left(q(x)+\lambda\omega(x)\right)u=0.$$

Thus the functional has the stationary function as solution to the Sturm-Liouville equation and the Lagrange multiplier λ is an eigenvalue. The constraint is interpreted as the weighted norm of $u(x)$ be equal to unity. The natural boundary conditions for this problem are

$$\left.\tfrac{\partial\tilde{F}}{\partial u'}\right|_{x_i}=0 \quad\Rightarrow\quad \left.p(x)u'\right|_{x_i}=0$$

for $i=0,1$, where $\tilde{F}\left[\lambda,x,u,u'\right]=q(x)u^2-p(x)\left(u'\right)^2+\lambda\omega(x)u^2$.

When the constraint is algebraic $\phi(u,v)=0$ or differential $\phi(u,v,u',v')=0$ in, say, two dependent variable case, the augmented functional is formed as

$$\tilde{F}=F\left[x,u,v,u',v'\right]+\lambda(x)\phi.$$

The Lagrange multiplier is now a function of the independent variable x and apply at each point x in the domain.

This is justified because the algebraic or differential constraints applies locally at each point in contrast to global integral consraints that apply over the entire domain as a whole.

The extremalization now leads to

$$0=\int_{x_0}^{x_1}\left\{\left[\tfrac{\partial\tilde{F}}{\partial u}-\tfrac{d}{dx}\left(\tfrac{\partial\tilde{F}}{\partial u'}\right)\right]\delta u+\left[\tfrac{\partial\tilde{F}}{\partial v}-\tfrac{d}{dx}\left(\tfrac{\partial\tilde{F}}{\partial v'}\right)\right]\delta v\right\}dx$$

and the Euler-Lagrange equations.

e.g. Find the stationary function extremalizing the functional,

$$I\left[u\right]=\int_0^1\left\{\tfrac{1}{2}\left(\left(u'\right)^2+\left(v'\right)^2\right)+uv\right\}dx$$

subject to $\phi=v'-u=0$ and $u(0)=u(1)=v(0)=0,\ v(1)=1.$

Variational Form of a known differential equation (as Euler-Lagrange equation) may be obtained by reversing the procedure of variational calculus.

Consider, for example, the Laplace equation (Dirichlet problem)

$$\Im[u] \equiv \nabla^2 u = \tfrac{\partial^2 u}{\partial x^2} + \tfrac{\partial^2 u}{\partial y^2} = 0 \quad \text{in D subject to} \quad u = u_0 \quad \text{on} \quad \partial D.$$

In order to obtain a variational formulation, let

$$0 = \iint_D \Im[u]\, \delta u \, dD = \iint_D \left\{ \tfrac{\partial^2 u}{\partial x^2} + \tfrac{\partial^2 u}{\partial y^2} \right\} \delta u \, dD.$$

Consider a rectangular domain $D = \{(x,y) \mid x_0 \le x \le x_1,\ y_0 \le y \le y_1\}$ for convenience and integrate by parts, first in x,

$$\int \left\{ \int_{x_0}^{x_1} \tfrac{\partial^2 u}{\partial x^2} \delta u \, dx \right\} dy = \int \left\{ \left[u_x \delta u \right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \tfrac{\partial u}{\partial x} \delta u_x \, dx \right\} dy = -\tfrac{1}{2} \delta \left\{ \iint_D \left(\tfrac{\partial u}{\partial x} \right)^2 dxdy \right\}$$

then in y, to get the variational form corresponding to the Dirichlet problem

$$0 = \delta I[u] = \delta \left\{ \iint_D \left[\left(\tfrac{\partial u}{\partial x} \right)^2 + \left(\tfrac{\partial u}{\partial y} \right)^2 \right] dD \right\}.$$

It indicates that the diffusion process $\frac{\partial u}{\partial t} = c^2 \Im[u]$ as modelled by the Dirichlet problem is a smoothing process in which overall magnitudes of the gradients are minimized.

This procedure, in the language of Finite Element Method, for example, is called converting the strong (differential) form to the weak (integral – variational) form.

The inverse problem of determining the variational form from the differential equation is always possible for linear, self-adjoint differential operators. The resulting equivalent variational form often provides valuable insight into the physical processes that the differential equation models as in the Dirichlet problem modeling the diffusion process.

Approximate Solution methods may be sought when resulting Euler-Lagrange differential equations do not allow a closed form solution. The approach may be summarized as follows:

1. Solve the Euler-Lagrange differential equations using numerical methods, such as spectral, or finite-element methods, where they are based on weighted residual methods (Galerkin approach).

2. Solve the integral variational form approximately using the Rayleigh-Ritz method.

Rayleigh-Ritz method directly applies to the variational form rather than the resulting differential form. Consider the variational problem

$$0 = \delta I = \int_{x_0}^{x_1} \delta F[x, u, u', \ldots]\, dx$$

where the stationary function $u(x)$ is to be determined. It is approximated by a linear combination of known basis (trial) functions $\{\phi_j(x)\}_{j=0}^{n}$ as follows:

$$u(x) \approx \tilde{u}(x) = \sum_{j=0}^{n} c_j \phi_j(x)$$

If the endpoints are specified as $u(x_0) = u_0$ and $u(x_1) = u_1$, the first basis function $\phi_0(x)$ is designed to satisfy the boundary conditions while the remaining basis functions are to satisfy the homogeneous boundary conditions, such that

$$\phi_0(x_i) = u_i \quad \text{and} \quad \phi_1(x_i) = \phi_2(x_i) = \ldots = \phi_n(x_i) = 0.$$

Substituting the trial function into the functional produces the variational problem

$$\delta \int_{x_0}^{x_1} \tilde{F}[x, c_1, c_2, \ldots, c_n]\, dx = 0$$

where the constants c_j's ($c_0 = 1$) vary to render the functional stationary.

e.g. Apply Rayleigh-Ritz method to the DE: $u'' - u = -x$ with $u(0) = u(L) = 0$ to get:

$$0 = \int_0^L \{u'' - u + x\} \delta u dx = \int_0^L \{-u'\delta u' - u\delta u + x\delta u\} dx \ \Rightarrow\ \delta \left\{ \int_0^L \left[(u')^2 + u^2 - 2xu \right] dx \right\} = 0.$$

Let $\phi_0(x) = 0$ and $\phi_1(x) = x(L-x)$, $\phi_2(x) = x^2(L-x)$,..., $\phi_n(x) = x^n(L-x)$.

- One-term approximation $\tilde{u}(x) = c_1\phi_1(x)$ yields

$$\delta\left\{ \tfrac{11}{30} c_1^2 - \tfrac{1}{6} c_1 \right\} = 0 \quad \Rightarrow \quad \left\{ \tfrac{11}{15} c_1 - \tfrac{1}{6} \right\} \delta c_1 = 0 \quad \Rightarrow \quad c_1 = \tfrac{5}{22}.$$

- Show that two-term approximation $\tilde{u}(x) = c_1\phi_1(x) + c_2\phi_2(x)$ yields $c_1 = \frac{69}{473}$, $c_2 = \frac{7}{43}$.

Recall that for self-adjoint DE, i.e. DE in the Sturm-Liouville form

$$\Im[u] \equiv \tfrac{d}{dx}\left(p(x)\tfrac{du}{dx}\right) + q(x)u = f(x)$$

proper variational formulation exists.

In other words,

$$\int_{x_0}^{x_1}\left(\Im[u]-f\right)\delta u dx = 0 \quad \Leftrightarrow \quad \delta\left\{\int_{x_0}^{x_1} F dx\right\} = 0$$

the two forms are equivalent. Now introducing the approximate solution before converting into variational form leads to

$$\int_{x_0}^{x_1}\left(\Im[\tilde{u}]-f\right)\tfrac{\partial\tilde{u}}{\partial c_i}\delta c_i dx = 0 \quad \Rightarrow \quad \int_{x_0}^{x_1}\left(\Im[\tilde{u}]-f\right)\phi_i dx = 0$$

for $i = 1,2,\ldots,n$.

This is the projection of the residual $\left(\Im[\tilde{u}]-f\right)$ onto the space of admissible (test) functions known as Galerkin method. In fact, Galerkin method is more general and it does not require the DE to be able to be written in a proper variational form.

Consider now the differential eigenvalue problem $\Im[u] = -\lambda\omega(x)u(x)$ where $\Im[\cdot]$ is of Sturm-Liouville form. The variational form

$$\int_{x_0}^{x_1}\Im[u]\delta u dx = -\lambda\int_{x_0}^{x_1}\omega(x)u(x)\delta u dx \quad \Leftrightarrow \quad \lambda[u] = -\int_{x_0}^{x_1}\Im[u]\delta u dx \Big/ \int_{x_0}^{x_1}\omega(x)u(x)\delta u dx$$

leads to the egenvalue as a functional, therefore, minimizing the functional $\lambda[u]$ will reult in the smallest eigenvalue.

e.g. Consider the Sturm-Liouville problem: $u'' + \lambda u = 0$ with $u(0) = u(1) = 0$.

$$\lambda[u] = -\int_0^1 u''\delta u dx \Big/ \int_1^1 u(x)\delta u dx.$$

Let $\tilde{u}(x) = c_1 x(1-x)$ so that $\tilde{u}''(x) = -2c_1$ and $\delta\tilde{u} = x(1-x)\delta c_1$. Then

$$\lambda = -\int_0^1 2c_1 x(1-x)\delta c_1 dx \Big/ \int_1^1 c_1 x^2(1-x)^2\delta c_1 dx = 10 \approx \pi^2 = 9.8696\ldots \quad \left(\lambda_n = n^2\pi^2\right).$$

e.g. Consider vibrating square membrane problem: $u_{xx} + u_{yy} = \frac{1}{c^2}u_{tt}$, $0 < x, y < 1$. Consider harmonic moion that $u(x,y,t) = \exp(i\omega t)\phi(x,y)$, then

$$\phi_{xx} + \phi_{yy} = -\tfrac{\omega^2}{c^2}\phi \quad \Rightarrow \quad \Im[\phi] = -\lambda\phi \quad \Rightarrow \quad \lambda[\phi] = -\iint_D \Im[\phi]\delta\phi dD \Big/ \iint_D \phi\delta\phi dD.$$

where $\lambda = \omega^2/c^2$. Consider clamped boundaries and so $\tilde{\phi}(x,y) = c_1 x(1-x)y(1-y)$.

Show that $\lambda = \frac{1}{45}c_1\delta c_1 \Big/ \frac{1}{900}c_1\delta c_1 = 20 \approx 2\pi^2$.

<u>Appendix A</u> The point x_0 is a stationary (or critical point) of $f(x)$ if

$$\tfrac{df}{dx} = 0 \quad \text{at} \quad x = x_0 \quad \text{or} \quad df = \tfrac{df}{dx}dx = 0$$

where df is called total differential of $f(x)$. The following possibilities exist

1. If $\frac{d^2f}{dx^2} < 0$ at x_0, then f has a local maximum at x_0.
2. If $\frac{d^2f}{dx^2} > 0$ at x_0, then f has a local minimum at x_0.
3. If $\frac{d^2f}{dx^2} = 0$ at x_0, then it is inconclusive, e.g. $f(x) = x^3$.

In 2D, for an extremum of $f(x,y)$ to occur at (x_0, y_0) it is necessary that

$$df = \tfrac{df}{dx}dx + \tfrac{df}{dy}dy = 0 \quad \Rightarrow \quad \tfrac{df}{dx} = \tfrac{df}{dy} = 0 \quad \text{at} \quad (x_0, y_0)$$

and the sufficiency comes from

1. $f_{xx}f_{yy} - f_{xy}^2 > 0$ and $f_{xx} < 0$ for a local maximum at (x_0, y_0).
2. $f_{xx}f_{yy} - f_{xy}^2 > 0$ and $f_{xx} > 0$ for a local minimum at (x_0, y_0).
3. $f_{xx}f_{yy} - f_{xy}^2 = 0$, for a saddle point at (x_0, y_0).
4. $f_{xx}f_{yy} - f_{xy}^2 = 0$, inconclusive.

<u>Appendix B</u> For an extremum of a function $f(x,y,z)$ at (x_0, y_0, z_0) it is necessary that

$$df = f_x dx + f_y dy + f_z dz = 0 \quad \Rightarrow \quad f_x = f_y = f_z = 0 \quad \text{at} \quad (x_0, y_0, z_0)$$

because the differentials dx, dy, dz vary independently.

If the extremum of $f(x,y,z)$ is sought subject to a constraint $g(x,y,z) = c$, then they are no longer independent

$$dg = g_x dx + g_y dy + g_z dz = 0 \quad \Rightarrow \quad dz = -\left(g_x dx + g_y dy\right)\big/ g_z \quad \text{(say)}$$

provided that $g_z \neq 0$. This then implies

$$df = \left(f_x - \tfrac{f_z}{g_z}g_x\right)dx + \left(f_y - \tfrac{f_z}{g_z}g_y\right)dy = 0 \quad \Rightarrow \quad f_x + \lambda g_x = f_y + \lambda g_y = 0$$

where $\lambda = -\frac{f_z}{g_z}$ and thus $f_z + \lambda g_z = 0$.

This can also be obtained by defining $\tilde{f} = f + \lambda(g - c)$ so that extremalization yields

$$d\tilde{f} = \tilde{f}_x dx + \tilde{f}_y dy + \tilde{f}_z dz = 0 \quad \Rightarrow \quad \left(f_x + \lambda g_x\right) = \left(f_y + \lambda g_y\right) = \left(f_z + \lambda g_z\right) = 0$$

together with $g = c$. Here, λ is called **<u>Lagrange Multiplier</u>**.

Appendix C

Integration by parts plays an important part in the calculus of variations.

In 1D: $$\int_{x_0}^{x_1} p\,dq = \left[pq\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} q\,dp$$

that follows from $d(pq) = p\,dq + q\,dp$.

In 2D or higher: $$\iint_A \psi \nabla^2 \phi \, dA = \int_{\partial A} \psi \left(\nabla\phi \cdot \mathbf{n}\right) dS - \iint_A \nabla\psi \cdot \nabla\phi \, dA$$

that follows from the divergence theorem.

$$\iint_A \nabla \cdot \mathbf{v} \, dA = \int_{\partial A} \left(\mathbf{v} \cdot \mathbf{n}\right) dS$$

where $\mathbf{n}$ is the outward normal vector to the boundary ∂A of A, that is a closed curve (2D) or a closed surface (3D) and from the vector differential calculus identity

$$\nabla \cdot \mathbf{v} = \nabla\psi \cdot \nabla\phi + \psi \nabla^2 \phi$$

when $\mathbf{v} = \psi \nabla \phi$ where ψ and ϕ are arbitrary scalar functions.

Appendix D The adjoint $\Im^*$ of a differential operator $\Im$ satisfies

$$\left(v, \Im[u]\right)_\omega = \left(\Im^*[v], u\right)_\omega$$

where u and v are arbitrary functions and $(\cdot,\cdot)_\omega : H \times H \to \Re$ denotes weighted inner product in a function space H, such as

$$\left(u, v\right)_\omega = \int_a^b u(x) v(x)\, \omega(x) dx\,.$$

Consider a 2nd order linear differential operator

$$\Im[u] = \tfrac{1}{\omega(x)}\left(a_2(x)u'' + a_1(x)u' + a_0(x)u\right) \quad \text{for} \quad a < x < b\,.$$

The construction of the adjoint operator yields

$$\left(v, \Im[u]\right)_\omega = \int_a^b \tfrac{1}{\omega}\left(a_2 u'' + a_1 u' + a_0 u\right) v\, \omega dx = \int_a^b \underbrace{\tfrac{1}{\omega}\left(\left(a_2 v\right)'' - \left(a_1 v\right)' + a_0 v\right)}_{\Im^*[v]} u\, \omega dx\,.$$

after integration by parts for homogeneous boundary conditions.

It can be shown that for $\Im$ to be self-adjoint $\Im = \Im^*$, it is required that $a_1 = a_2{}'$, thus

$$\Im = \tfrac{1}{\omega(x)}\left(\tfrac{d}{dx}\left(a_2 \tfrac{d}{dx}\right) + a_0\right)$$

is the form of a self-adjoint differential operator called Sturm-Liouville operator.