Introduction to Calculus of Variations

February 15, 2013

> For since the fabric of the universe is most perfect and the work of a most wise creator, nothing at all takes place in the universe in which some rule of the maximum or minimum does not appear.
> – Leonhard Euler

1 Differential vs variational statements

Most of problems in science and engineering can be formulated in two completely different yet equivalent statements: a differential statement and a variational statement. A differential statement includes a set of governing differential equations established inside the domain and a set of boundary conditions to be satisfied along the boundaries, while a variational statement is to find stationary conditions for an integral with unknown functions contained in the integrand. The variational statement is attractive primarily for their beauty of form and spirit due to the fact that the extremum principle is a unique aspect of nature. In comparison to differential statements, the variational statements are advantageous in the following aspects:

- Variational statements usually have clear physical meanings, and they are invariant to coordinate systems.
- In some problems, variational statements provide more realistic description than differential statements, for example, the description of concentrated loads.
- Variational statements provide a powerful approximation or numerical method for solving the problem whose solution might be very difficult to find from the differential statement. As an example, the popular finite element method is based on variational statements. Sometimes, variational statements can also provide upper bounds or lower bounds for the exact solution although it is difficult to find the exact solution.
- Using a variational statement, it is usually more systematically than conventional ways to derive the governing differential equations and corresponding boundary conditions.

The mathematics for variational statements is the calculus of variations, which is a natural extension of the conventional calculus for finding stationary points for continuous

functions, which can be generally classified as unconstrained stationary problems and constrained stationary problems.

2 Unconstrained stationary problems

Consider a continuous function $F(u_1)$ defined in a domain. The necessary conditions for this function to be stationary is $\frac{\mathrm{d}F}{\mathrm{d}u_1} = 0$. This condition corresponds to a horizontal tangent to the curve, as illustrated in Figure 1. To determine whether the stationary point is a minimum, a maximum or a saddle point, we need to evaluate the sign of the second derivative: $\frac{\mathrm{d}^2F}{\mathrm{d}u_1^2} > 0$ for a minimum point; $\frac{\mathrm{d}^2F}{\mathrm{d}u_1^2} < 0$ for a maximum point; and $\frac{\mathrm{d}^2F}{\mathrm{d}u_1^2} = 0$ for a saddle point.

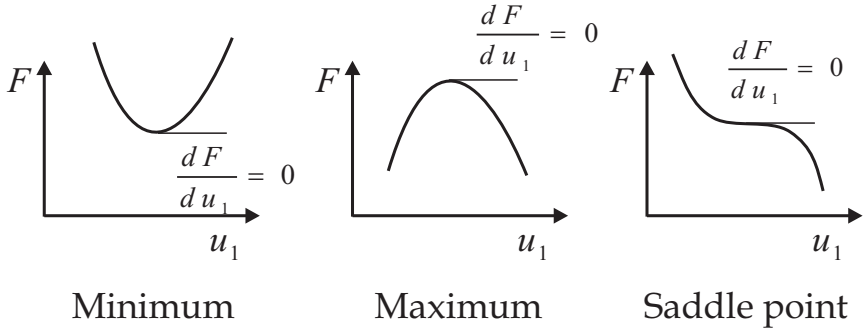

Figure 1: Stationary points of a function.

When the function is of n variables, $F = F(u_1, u_2, \ldots, u_n)$. The stationary points of this function are defined as those for which

$$\frac{\partial F}{\partial u_i} = 0, \quad i = 1, 2, \ldots, n. \tag{1}$$

To determine whether a stationary point of the multivariable function is a minimum, a maximum, or a saddle point, it is necessary to consider the second partial derivatives of the function. If

$$\sum_{i,j=1,n} \frac{\partial^2 F}{\partial u_i \partial u_j}\, \mathrm{d}u_i\, \mathrm{d}u_j > 0 \tag{2}$$

at a stationary point for all increments $\mathrm{d}u_i$ and $\mathrm{d}u_j$, the function presents a minimum point. If, on the other hand, the same quantity is negative for all $\mathrm{d}u_i$ and $\mathrm{d}u_j$, the function presents a maximum point. Finally, if the same quantity can be positive or negative depending on the choice of the increments, the function presents a saddle point.

2.1 Example

Example 1: Find the path of light, starting from a given point $P(x_1, y_1)$ and arriving at point $Q(x_0, y_0)$ where medium 1 and medium 2 are different homogeneous media in which

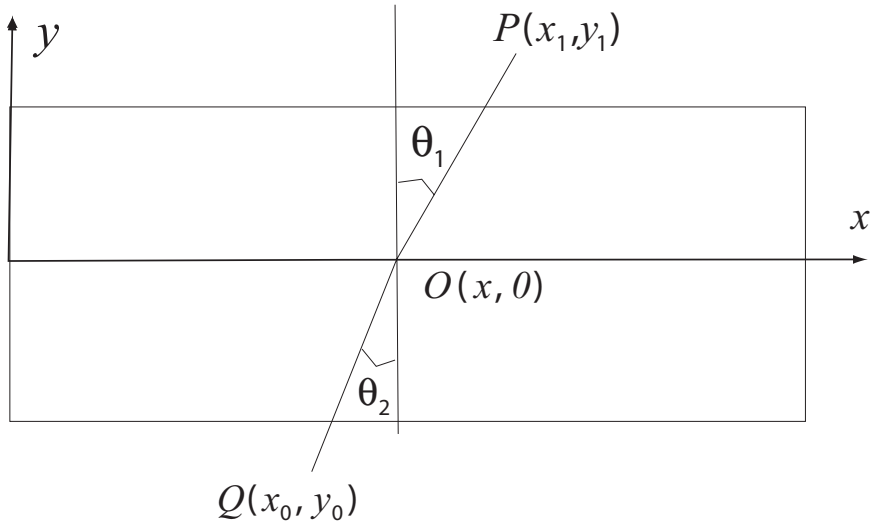

Figure 2: Fermat's principle.

the velocities of light are c_1 and c_2, respectively. See Figure 2 for an illustration.

Solution: According to the principle of Fermat, the time elapsed in the passage of light between two fixed points is minimum with respect to possible paths connecting two points. Suppose the interface is chosen to be x axis and the light intersects the interface of the medium at point $O(x, 0)$; see Figure 2. The total time T for the light to travel through the two media is

$$T = \frac{\sqrt{(x_1 - x)^2 + y_1^2}}{c_1} + \frac{\sqrt{(x - x_0)^2 + y_0^2}}{c_2}.$$

The equation $\frac{\mathrm{d}T}{\mathrm{d}x} = 0$ leads to

$$\frac{x_1 - x}{c_1\sqrt{(x_1 - x)^2 + y_1^2}} = \frac{x - x_0}{c_2\sqrt{(x - x_0)^2 + y_0^2}}$$

which is equivalent to

$$\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_2}{c_2}$$

the so-called Snell's law.

3 Constrained stationary problems - the Lagrange multiplier method

Consider once more the problem of determining a stationary point of a function of several variables $F = F(u_1, u_2, \ldots, u_n)$. However, in this case, the variables are not independent, rather they are subjected to a constraint of the form

$$f(u_1, u_2, \ldots, u_n) = 0. \tag{3}$$

Conceptually one could use the constraint to express one variable, say u_n, in term of the others. Then, u_n is eliminated from F to obtain a function of $n-1$ independent variables $F = F(u_1, u_2, \ldots, u_{n-1})$, a problem identical to that treated in the previous section. In many practical situations it might be cumbersome, or even impossible, to completely eliminate one variable of the problem.

This elimination process can be avoided by using an alternative and systematic approach: the Lagrange multiplier method. At a stationary point, the differentiation of F must vanish

$$\mathrm{d}F = \frac{\partial F}{\partial u_1}\,\mathrm{d}u_1 + \frac{\partial F}{\partial u_2}\,\mathrm{d}u_2 + \ldots + \frac{\partial F}{\partial u_n}\,\mathrm{d}u_n = 0. \tag{4}$$

However this does not imply $\partial F/\partial u_i = 0$ for $i = 1, 2, \ldots, n$ because we cannot chose the differentiations $\mathrm{d}u_i$ arbitrarily as they must satisfy the constraint in Eq. (3). The relation among the differentiations $\mathrm{d}u_i$ can be explicitly written by taking a differentiation of the constraint in Eq. (3) to find

$$\mathrm{d}f = \frac{\partial f}{\partial u_1}\,\mathrm{d}u_1 + \frac{\partial f}{\partial u_2}\,\mathrm{d}u_2 + \ldots + \frac{\partial f}{\partial u_n}\,\mathrm{d}u_n = 0. \tag{5}$$

A linear combination of Eqs. (4) and (5) yields

$$\frac{\partial F}{\partial u_1}\,\mathrm{d}u_1 + \ldots + \frac{\partial F}{\partial u_n}\,\mathrm{d}u_n + \lambda\left[\frac{\partial f}{\partial u_1}\,\mathrm{d}u_1 + \ldots + \frac{\partial f}{\partial u_n}\,\mathrm{d}u_n\right] = 0, \tag{6}$$

where λ is an arbitrary function of the variables $u_1, u_2, \ldots, u_n$ and it is called the *Lagrange multiplier*. Regrouping the various terms then leads to

$$\sum_{i=1}^{n}\left[\frac{\partial F}{\partial u_i} + \lambda\,\frac{\partial f}{\partial u_i}\right]\,\mathrm{d}u_i = 0. \tag{7}$$

In principle, one could now express $\mathrm{d}u_n$ in term of the $n-1$ other differentiations $\mathrm{d}u_i$ to be left with $n-1$ independent, arbitrary differentiations. To avoid this cumbersome step, the arbitrary Lagrange multiplier can be chosen such that

$$\frac{\partial F}{\partial u_n} + \lambda\,\frac{\partial f}{\partial u_n} = 0. \tag{8}$$

With this choice, the last term of the sum in Eq. (7) vanishes for all $\mathrm{d}u_n$. Note if $\frac{\partial f}{\partial u_n} = 0$, the Lagrange multiplier becomes undetermined. We should choose another term $\frac{\partial F}{\partial u_i} + \lambda\,\frac{\partial f}{\partial u_i}$ which has $\frac{\partial f}{\partial u_i} \neq 0$. Here we assumed that $\frac{\partial f}{\partial u_n} \neq 0$ without loss of generality. The choice of Lagrange multiplier in such a way effectively enables us to eliminate $\mathrm{d}u_n$. Hence, there is no need to express this differentiation in terms of the $n-1$ others which can now be treated as independent, arbitrary quantities, implying

$$\frac{\partial F}{\partial u_i} + \lambda\,\frac{\partial f}{\partial u_i} = 0, \quad i = 1, 2, \ldots, n-1. \tag{9}$$

These equations in Eqs. (8) and (9) are actually those necessary conditions for function $F^* = F + \lambda f$ to be stationary which are

$$\frac{\partial F^*}{\partial u_i} = \frac{\partial F}{\partial u_i} + \lambda\frac{\partial f}{\partial u_i} + f\frac{\partial \lambda}{\partial u_i} = \frac{\partial F}{\partial u_i} + \lambda\frac{\partial f}{\partial u_i} = 0, \quad i = 1, 2, \ldots, n, \tag{10}$$

where the second equality is due to the constraint, Eq. (3), and

$$\frac{\partial F^*}{\partial \lambda} = f = 0. \tag{11}$$

These $n+1$ equations including Eqs. (10) and (11) can be used to solve for the $n+1$ variables including u_i $(i = 1, 2, \ldots, n)$ and the Lagrange multiplier λ, a total of $n+1$ unknown variables.

In summary, the original, constrained stationary problem can be replaced by an *unconstrained stationary problem* for the function F^*. It is important to note that the initial problem involves n variables u_i and one constraint. The final problem involves $n+1$ variables u_i for $i = 1, 2, \ldots, n$ and the Lagrange multiplier λ but no constraint. And also it is pointed out that although the stationary conditions of F and F^* remain the same, the extreme behavior (maximum, minimum, or saddle) might be different. In other words a solution corresponds to the minimum of F does not necessarily correspond to minimum of F^*.

The Lagrange multiplier method provides us a systematic way to find the stationary point of the original function. Whether this point is a maximum, minimum or a saddle point can be obtained from the second derivative of the original function F or the physics of the problem itself. For example, if the function can only have maximum value under some constraints, and you find two stationary points using the Lagrange multiplier method, you can plug these two stationary points to find out which one gives a bigger value for the function, which corresponds to the maximum value of the original function.

If there are multiple constraints, we can follow a similar procedure to introduce multiple Lagrange multipliers to transform an original, constrained stationary problem to an unconstrained stationary problem.

3.1 Example

Example 2: Design a box with given surface area so that the volume is maximized. We know that the box has no cover along one of the xy planes and the given surface area is $A = 2xz + 2yz + xy = C$.

Solution: Mathematically, it means we need to find the maximum of the volume $V = xyz$ under the constraint $2xz + 2yz + xy = C$. Let us solve this constrained stationary value problem using the conventional approach first. From the constraint, we can express x in terms of y, z as

$$x = \frac{C - 2yz}{2z + y}$$

Then the volume can be expressed as

$$V = \frac{C - 2yz}{2z + y} yz$$

The maximum volume will be reached if $y = 2z = \sqrt{C/3}$, which satisfies the two conditions required for V to be stationary: $\frac{\partial V}{\partial y} = \frac{\partial V}{\partial z} = 0$.

Next, let us try to solve the same problem using the Lagrange multiplier method. Let us construct the following function

$$V^* = xyz + \lambda(2xz + 2yz + xy - C)$$

For this function to be stationary, the following equations should be satisfied

$$\begin{aligned}
\frac{\partial V^*}{\partial x} &= yz + \lambda(2z + y) = 0 \\
\frac{\partial V^*}{\partial y} &= xz + \lambda(2z + x) = 0 \\
\frac{\partial V^*}{\partial z} &= xy + \lambda(2x + 2y) = 0 \\
\frac{\partial V^*}{\partial \lambda} &= (2xz + 2yz + xy - C) = 0
\end{aligned}$$

The only meaningful solution of the above equations is

$$x = y = 2z = \sqrt{\frac{C}{3}}, \quad \lambda = -\frac{1}{4}\sqrt{\frac{C}{3}}$$

The corresponding volume is

$$V = \frac{C^{\frac{3}{2}}}{6\sqrt{3}}.$$

According to the nature of problem, we know that maximum volume value exists. And we only find one stationary point and this stationary point must correspond to the maximum volume.

Indeed we have obtained the same result using both approaches.

4 Stationary values of functionals

We have just discussed the stationary values of continuous functions which has a specific functional form in terms of the unknown variables. Another type of problems frequently in engineering can be formulated as stationary value problem of some functionals with the arguments are unknown functions which themselves are functions of some variables. A functional is usually expressed as

$$I = I[y(x)] \tag{12}$$

showing the relationship between the functional I, the unknown function $y(x)$ and the fundamental variable x. Usually the stationary value of the functional I corresponds to a specific form of $y(x)$. Sometimes, we also simply write $I[y]$ to indicate the functional relation of I in terms of y which itself is a function of x.

First, we will use some examples to show the relevancy of this type of problems. Then we will extend the techniques we have learned in dealing with continuous functions in previous sections to deal with this type of problems.

4.1 Some Sample Problems

In this section, we list a few examples to help the readers appreciate this type of problems because such problems are not commonly covered in the mathematics courses we have studied.

Example 3: Find a curve connecting two points (x_0, y_0) and (x_1, y_1) in a two-dimensional domain (x, y) so that the distance measured along the curve is the shortest.
Solution: Denoting the arc length along the curve as s, a differential increment can be expressed as $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \sqrt{1 + y'^2}\ \mathrm{d}x$. The distance along the curve is

$$L = \int_{x_0}^{x_1} \sqrt{1 + y'^2}\ \mathrm{d}x \tag{13}$$

Then the question of the problem can be stated as finding a specific functional form of $y(x)$ so that L will be minimum among all the possible choices of curves.

Example 4 (Brachistochrone problem): Find the path between two points, $A(0, 0)$ and $B(a, b)$, in a vertical, frictionless plane along which a mass point under the influence of gravity will slide from A to B in the shortest possible time.
Solution: Denoting s as the arc length, we can express the speed as

$$v = \frac{\mathrm{d}s}{\mathrm{d}t}$$

Thus, the time to travel from A to be B is given by

$$t = \int_A^B \frac{\mathrm{d}s}{v}$$

where the arc length s can be obtained by $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \sqrt{1 + y'^2}\ \mathrm{d}x$, and the speed v can be obtained by conservation of energy such that

$$\frac{1}{2}mv^2 = mgy$$

giving $v = \sqrt{2gy}$. Hence the problem asks us to find the function $y(x)$ so that the time, described using the following functional of $y(x)$

$$t = \int_0^a \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}\ \mathrm{d}x \tag{14}$$

will be minimum among all the possible choices of $y(x)$.

4.2 Attacked by the conventional calculus

Let us deal with the problem of finding unknown function $y(x)$ that makes the following integral stationary

$$I[y] = \int_{x_0}^{x_1} [p(x)(y')^2 + q(x)y^2 + 2f(x)y]\ \mathrm{d}x \tag{15}$$

where the notation $()'$ is used to indicate a derivative with respect to x. y is an unknown function of x and is subjected to boundary conditions

$$y(x_0) = y_0, \quad y(x_1) = y_1, \tag{16}$$

where p, q, and f are given continuous functions in $x_0 \leq x \leq x_1$. Values of the integral in Eq. (15) depends on the function $y(x)$. Hence, $I[y]$ is called a *functional* with argument $y(x)$ because $I[y]$ is a function of another function, $y(x)$.

The statement of the problem implies the existence of a solution $y(x)$ that results in a stationary value of I among all the possible choices of $y(x)$ satisfying the boundary conditions in Eq. (16). Furthermore, we assume that $y(x)$ is continuous and differentiable as needed in the problem. To show that this function corresponds to the stationary value of $I[y]$ using what we have learned in conventional calculus, we introduce the following trial function $\bar{y}$

$$\bar{y}(x) = y(x) + \alpha\eta(x) \tag{17}$$

where α is a parameter and $\eta(x)$ is an arbitrary, continuous function that vanishes at the boundaries such that $\eta(x_0) = \eta(x_1) = 0$. In other words, $\bar{y}(x)$ is a possible choice of $y(x)$. When $\alpha = 0$, the trial function $\bar{y}$ becomes the solution $y(x)$. If we can show that $I[\bar{y}]$ is stationary when $\alpha = 0$, we can confirm that $y(x)$ corresponds to the stationary value of I. Because both $y(x)$ and $\eta(x)$ are continuous and differentiable, we have

$$\bar{y}'(x) = y'(x) + \alpha\eta'(x) \tag{18}$$

Substituting Eq. (17) and Eq. (18) into Eq. (15), we have

$$I[\bar{y}(x)] = I[y(x) + \alpha\eta(x)] = I[y(x)] + \delta I + \frac{1}{2}\delta^2 I = I[y(x)] + \alpha I_1 + \frac{1}{2}\alpha^2 I_2 \tag{19}$$

with

$$\delta I = 2\alpha \int_{x_0}^{x_1} [py'\eta' + qy\eta + f\eta] \, \mathrm{d}x \equiv \alpha I_1 \tag{20}$$

and

$$\delta^2 I = 2\alpha^2 \int_{x_0}^{x_1} [p\eta'^2 + q\eta^2] \, \mathrm{d}x \equiv \alpha^2 I_2 \tag{21}$$

δI and $\delta^2 I$ are called the *first variation* and the *second variation* of the functional $I[y]$, respectively.

Observe the following:

$$\frac{\partial I[\bar{y}]}{\partial \alpha} = \frac{\partial I[\bar{y}]}{\partial \bar{y}}\frac{\partial \bar{y}}{\partial \alpha} + \frac{\partial I[\bar{y}]}{\partial \bar{y}'}\frac{\partial \bar{y}'}{\partial \alpha} = 2\int_{x_0}^{x_1} [(qy + f)\eta + py'\eta'] \, \mathrm{d}x = I_1 \tag{22}$$

$$\frac{\partial^2 I[\bar{y}]}{\partial \alpha^2} = 2\frac{\partial I_1}{\partial \alpha} = 2\int_{x_0}^{x_1} [q\eta^2 + p\eta'^2] \, \mathrm{d}x = I_2 \tag{23}$$

For $I[\bar{y}(x)]$, a second-order polynomial of α, to reach a stationary value at $\alpha = 0$ (so that $y(x)$ corresponds to the stationary value of $I[\bar{y}]$), we require $I_1 = 0$ (in other words the first variation of the functional δI vanishes) , which corresponds to the condition in Eq. (1)

for ordinary functions. The stationary value of $I[\bar{y}]$ at $\alpha = 0$, $I[y]$, becomes a minimum if $I_2 > 0$ (in other words $\delta^2 I > 0$), a maximum if $\delta^2 I < 0$ (in other words $\delta^2 I < 0$), a saddle point if $\delta^2 I = 0$ (in other words $\delta^2 I = 0$). Clearly for the given functional in Eq. (15), if we have $p > 0$ and $q > 0$, the stationary point will be a minimum.

Integrating the first term of δI by parts, we obtain

$$\delta I = 2\alpha \left(py'\eta|_{x_0}^{x_1} + \int_{x_0}^{x_1} [-(py')' + qy + f]\eta \, \mathrm{d}x \right) = 2\alpha \int_{x_0}^{x_1} [-(py')' + qy + f]\eta \, \mathrm{d}x = 0 \quad (24)$$

The first term vanishes because η vanishes at the boundaries. To proceed further, we need to use the following **fundamental lemma** of the calculus of variations:

> **Lemma**: If the relation $\int_{x_0}^{x_1} \rho(x)\eta(x) \, \mathrm{d}x = 0$, where $\rho(x)$ is a continuous and differentiable function of x, holds for all continuous functions $\eta(x)$ that vanishes on the boundaries, it follows that $\rho(x) = 0$ identically.

This lemma can be trivially proved by contradiction. Suppose at some point x^*, $\rho(x^*) \neq 0$, then $\rho(x)$ is not zero and does not change its sign in a sufficiently small neighborhood, V, close to x^*, because $\rho(x)$ is a continuous and differentiable function. If we choose $\eta(x)$ to be of the same sign as $\rho(x)$ in V and zero outside V, we have

$$\int_{x_0}^{x_1} \rho(x)\eta(x) \, \mathrm{d}x > 0 \quad (25)$$

because the integrand $\rho(x)\eta(x)$ does not change sign inside V. We arrive at a contradiction, which means the Lemma must hold.

Since Eq. (24) holds for any continuous and function η, the coefficient of η in the integrand must vanish such that

$$-(py')' + qy + f = 0. \quad (26)$$

This equation is called the *Euler-Lagrange equation* of the functional in Eq. (15). The solution to the stationary problem of the functional in Eq. (15) can be obtained by solving the Euler-Lagrange equation, Eq. (26), along with the boundary conditions, Eq. (16). It is noted that η and η' in Eq. (20) are not independent continuous function but are related through $\frac{\mathrm{d}\eta}{\mathrm{d}x} = \eta'$. This is the very reason that the integration by parts is used in Eq. (24) so that we can use the fundamental lemma to obtain the Euler-Lagrange equation in Eq. (26).

Following the same line of logic, we can prove the following two extensions of the fundamental lemma.

- **Extension I**
  If we have for a continuous and differentiable function $\rho(x)$,

  $$\int_{x_0}^{x_1} \rho(x)\eta(x) \, \mathrm{d}x + B_0\eta(x_0) + B_1\eta(x_1) = 0 \quad (27)$$

  holds for any arbitrary continuous function $\eta(x)$, then

  $$\rho(x) = 0 \quad B_0 = 0 \quad B_1 = 0.$$

Indeed, if we set $\eta(x_0) = \eta(x_1) = 0$, then from the fundamental lemma, we know $\rho(x) = 0$. Then Eq. (27) is reduced to be

$$B_0\eta(x_0) + B_1\eta(x_1) = 0 \tag{28}$$

for arbitrary $\eta(x_0)$ and $\eta(x_1)$, where one can easily conclude that $B_0 = B_1 = 0$.

- **Extension II**
  If we have multiple continuous and differentiable functions $\rho_1(x), \rho_2(x), \ldots, \rho_n(x)$, and
  $$\int_{x_0}^{x_1} \left[\rho_1(x)\eta_1(x) + \rho_2(x)\eta_2(x) + \ldots + \rho_n(x)\eta_n(x)\right] \, \mathrm{d}x = 0 \tag{29}$$
  for any choice of arbitrary continuous functions $\eta_1(x), \eta_2(x), \ldots, \eta_n(x)$, then
  $$\rho_1(x) = 0 \quad \rho_2(x) = 0 \quad \ldots \quad \rho_n(x) = 0$$
  This statement can be proved consecutively by showing $\rho_i(x) = 0$ for $i = 1, \ldots, n$ by choosing $\eta_i(x)$ to be arbitrary and all the other arbitrary functions to be zero.

4.2.1 Example

Example 5: Find a function $y(x)$ that minimizes the functional

$$I[y] = \int_0^{\pi} \left[(y')^2 + y^2 + 2xy\right] \, \mathrm{d}x \tag{30}$$

and satisfy the conditions $y(0) = 1$ and $y(\pi) = 0$.

Comparing to Eq. (15), we see that $p = q = 1, f(x) = x$. Hence the corresponding Euler-Lagrange equation is

$$y'' - y - x = 0,$$

The same equation can be obtained using Eq. (49) as derived later. This Euler-Lagrange equation has the following solution

$$y = c_1 \cosh x + c_2 \sinh x - x,$$

where c_1 and c_2 can be determined from the corresponding boundary conditions as

$$c_1 = 1, \quad c_2 = \frac{\pi - \cosh \pi}{\sinh \pi}.$$

Hence the final solution, the function $y(x)$ that minimizes the functional $I[y]$, is

$$y = \cosh x + \frac{\pi - \cosh \pi}{\sinh \pi} \sinh x - x. \tag{31}$$

In fact, the condition that the first variation vanishes only ensures that the solved function $y(x)$ corresponds to the stationary value of the functional. To verify whether it corresponds

to a minimum, maximum or saddle point of the functional, we still need to evaluate the second variation of the functional, which can be easily shown to be (see Eq (21))

$$\delta^2 I = \alpha^2 \int_0^\pi \left[(\delta y')^2 + (\delta y)^2 \right] \, \mathrm{d}x \tag{32}$$

Clearly $\delta^2 I$ is always greater than zero. Hence the function we have obtained in Eq. (31) corresponds to the minimum value of the functional. It will also be interesting to find out what will be the value. Simply substituting the solution in Eq. (31) into Eq. (30) and carrying out the integration, we find out that

$$I = 2 - \pi - \frac{\pi^3}{3} + (1 + \pi^2) \coth \pi - \frac{2\pi}{\sinh \pi} \approx -1.1108$$

To partially verify that all the other choices of the function y will result in a bigger number for I. Let us choose

$$y = \frac{\pi - x}{\pi}$$

Note that the possible choice of the function y should satisfy the boundary conditions. Carrying out the integration, we can find out that

$$I = \frac{1}{\pi} + \frac{\pi}{3} + \frac{\pi^2}{3} \approx 4.65538$$

which is clearly larger than the value of I based on the function y in Eq. (31) solved following the principle of the calculus of variations.

4.3 Calculus of variations

To deal with stationary problems of functionals with general integrands, we need to know how to manipulate variations. Consider a function $f(x)$, $a < x < b$. Change the function at any (or every point) $x = x_0$ to $\bar{f}(x)$, while holding $x = x_0$ fixed. Form the difference

$$\bar{f}(x_0) - f(x_0) \equiv \delta f(x_0) \tag{33}$$

which is called the variation of f. δf is a continuous and differentiable, but otherwise arbitrary function. In other words, δf is a virtual change that brings the function $f(x)$ to a new, arbitrary function $\bar{f}(x)$. If $f(x)$ is prescribed with a given value at certain point, such as the boundary, the corresponding variation at that point should be equal to zero. As shown in Figure 3, $f(x)$ has prescribed values at both boundary points, we than have $\delta f(a) = \delta f(b) = 0$. The variation δf is equal to zero at both ends.

The only and important difference between the calculus of variations and conventional calculus exists between an increment $\mathrm{d}f$ of a function $f(x)$ and a variation δf of the same function, as depicted in Figure 4. $\mathrm{d}f$ is an infinitesimal change in $f(x)$ resulting from an infinitesimal change $\mathrm{d}x$ in the independent variable; $\mathrm{d}f/\mathrm{d}x$ represents the tangent at the point. On the other hand, δf is an arbitrary virtual change that brings $f(x)$ to $\bar{f}(x)$. In taking the variation, the functional form is varied and the fundamental variable x is held fixed, while in taking the differentiation, the fundamental variable is varied and the

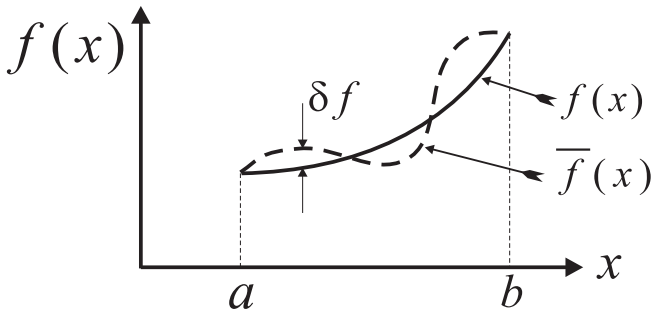

Figure 3: The concept of variation of a function.

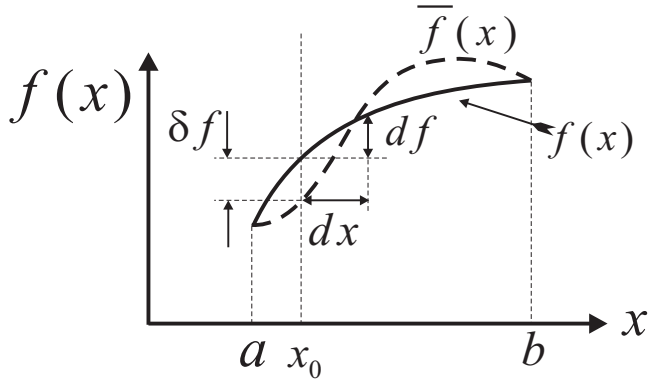

Figure 4: The difference between an increment $\mathrm{d}f$ and a variation δf.

functional form is held fixed. Clearly, the two quantities, $\mathrm{d}f$ and δf, are unrelated. As shown in Figure 4, $\mathrm{d}f$ is positive whereas δf is negative at x_0.

Although the concepts associated with the notation $\mathrm{d}f$ and δf are clearly distinct, the manipulations of these two symbols are quite similar. For instance, the order of application of the two operations can be interchanged. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}x}(\delta f) = \frac{\mathrm{d}}{\mathrm{d}x}(\bar{f} - f) = \frac{\mathrm{d}\bar{f}}{\mathrm{d}x} - \frac{\mathrm{d}f}{\mathrm{d}x} = \delta(\frac{\mathrm{d}f}{\mathrm{d}x}). \tag{34}$$

Similarly, the order of the integration and variation operations commute

$$\delta \int_a^b F \, \mathrm{d}x = \int_a^b \bar{F} \, \mathrm{d}x - \int_a^b F \, \mathrm{d}x = \int_a^b (\bar{F} - F) \, \mathrm{d}x = \int_a^b \delta F \, \mathrm{d}x. \tag{35}$$

The following rules of operation for the variation symbol can be easily verified by its definition

$$\begin{aligned} \delta(F_1 + F_2) &= \delta F_1 + \delta F_2 \\ \delta(F_1 F_2) &= \delta F_1 F_2 + F_1 \delta F_2 \\ \delta(F_1/F_2) &= (\delta F_1 F_2 - F_1 \delta F_2)/F_2^2 \\ \delta(F^n) &= n F^{n-1} \delta F \end{aligned} \tag{36}$$

which are exactly the same as the operations defined for the differentiation symbol d.

Now, let us rework the the simple problem in previous section attacked using the conventional calculus using the concept of variation of a function. Clearly, we can have the following change of variables

$$\delta y = \bar{y} - y = \alpha\eta \tag{37}$$

and

$$\delta y' = \bar{y}' - y' = \alpha\eta' \tag{38}$$

Using the variation symbol, we can rewrite the first variation in Eq. (20) as

$$\begin{aligned} \delta I &= 2\int_{x_0}^{x_1} [py'\delta y' + qy\delta y + f\delta y]\ \mathrm{d}x = 2\left[\int_{x_0}^{x_1} py'\ \mathrm{d}(\delta y) + \int_{x_0}^{x_1} (qy+f)\delta y\ \mathrm{d}x\right] \\ &= 2py'\delta y|_{x_0}^{x_1} + \int_{x_0}^{x_1} [-(py')' + qy + f]\,\delta y\ \mathrm{d}x \end{aligned} \tag{39}$$

According to Eq. (16), δy will vanish at the boundaries and it is otherwise arbitrary. The necessary condition for $I[y]$ to be stationary is $\delta I = 0$, which will give us the same Euler-Lagrange equation as in Eq. (26) using the fundamental lemma of calculus of variations.

Next, let us consider the stationary problem of the following definite integral

$$I = \int_a^b F(y,\ y',\ x)\ \mathrm{d}x, \tag{40}$$

where the integrand of the functional is a generic function of x (the integration variable), y (an unknown function of x), and y' (the derivative of y with respect to x). The unknown function is prescribed at the boundaries. In Eq. (40), the dependence of F on y, y', and x is given, and after integration, the coordinate x disappears. Hence, the value of I depends on the choice of the unknown function $y(x)$. Our purpose is to determine the unknown function $y(x)$ so that I will be stationary. Since there are an infinite number of functional forms for y defined between a and b, I is equivalent to a function of an infinite number of variables in some sense.

The stationarity of I, Eq. (40), requires that the variation of the functional vanishes, that is

$$\delta I = \delta\int_a^b F(y,\ y',\ x)\ \mathrm{d}x = \int_a^b \delta F(y,\ y',\ x)\ \mathrm{d}x = 0. \tag{41}$$

According to the definition of variation, we have

$$\delta F(y,\ y',\ x) = F(y+\delta y,\ y'+\delta y',\ x) - F(y,\ y',\ x) = \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' + \epsilon_1 \tag{42}$$

where ϵ_1 are terms with orders higher than $\sqrt{(\delta y)^2 + (\delta y')^2}$. Similarly to what we define the first derivative in calculus, we neglect these higher-order terms to define our first variation of a functional as

$$\delta F(y,\ y',\ x) \equiv \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' \tag{43}$$

The difference of the variation of the unknown function, Eq. (33), and the variation of a known function of unknown functions, Eq. (45), is timely noted here. According to our knowledge in calculus, we have

$$\mathrm{d}F(y,\ y',\ x) \equiv \frac{\partial F}{\partial y}\mathrm{d}y + \frac{\partial F}{\partial y'}\mathrm{d}y' + \frac{\partial F}{\partial x}\mathrm{d}x \tag{44}$$

The main difference is that in variation, the fundamental variable is held fixed and only the unknown functions are varying.

In general, if a functional $F(x, y, z, u_1, u_2, \ldots, u_n)$ is a function of fundamental variables x, y, z and multiple unknown functions, $u_1, u_2, \ldots, u_n$, we define the "first variation in F", noted δF, as

$$\delta F = \frac{\partial F}{\partial u_1}\ \delta u_1 + \frac{\partial F}{\partial u_2}\ \delta u_2 + \ldots + \frac{\partial F}{\partial u_n}\ \delta u_n. \tag{45}$$

Note $u_1, u_2, \ldots, u_n$ are not necessarily independent of each other. Again, it is repeated that the partial derivatives with respect to fundamental variables x, y, z does not contribute to the definition of the variation of the functional.

With the help of Eq. (43) (or in more general form in Eq. (45)), we can rewrite the stationary condition of the functional I in Eq. (41) as

$$\delta I = \int_a^b \left[\frac{\partial F}{\partial y}\ \delta y + \frac{\partial F}{\partial y'}\ \delta y'\right]\ \mathrm{d}x = 0. \tag{46}$$

It is noted that when calculating the variation of the functional F, only y and y' can be varied and x is hold fixed.

Integration by parts is now applied to the second term in the square bracket

$$\int_a^b \frac{\partial F}{\partial y'}\ \delta\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\ \mathrm{d}x = \int_a^b \frac{\partial F}{\partial y'}\ \frac{\mathrm{d}}{\mathrm{d}x}(\delta y)\ \mathrm{d}x = -\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\ \delta y\ \mathrm{d}x + \ \frac{\partial F}{\partial y'}\ \delta y|_a^b. \tag{47}$$

The boundary term vanishes since $\delta y(a) = \delta y(b) = 0$, and the stationarity condition then becomes

$$\delta I = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right]\delta y\ \ \mathrm{d}x = 0. \tag{48}$$

Because the integral must go to zero for *all arbitrary variations* δy, according to the fundamental lemma of calculus of variations, the bracketed term must vanish, which yields

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0. \tag{49}$$

Here again, the above reasoning can be reversed. Starting from Eq. (49), and performing the integration by parts in the reversed order implies $\delta I = 0$. In summary, *the necessary and sufficient condition for the definite integral to be at a stationary point is that Eq. (49)* is satisfied. This differential equation is called the *Euler-Lagrange equation* of the problem.

If F does not depend on x explicitly, we have the following

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y'' - y''\frac{\partial F}{\partial y'} - y'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) \tag{50}$$

after canceling the middle two terms and taking advantage of the Euler-Lagrange equation in Eq. (49), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'\frac{\partial F}{\partial y'}\right) = 0 \quad \text{or} \quad F - y'\frac{\partial F}{\partial y'} = \text{constant } c. \tag{51}$$

Usually if the functional F does not depend on x explicitly, Eq. (51) can be considered as a simpler version of the Euler-Lagrange equation in Eq. (49).

4.3.1 Example

Example 6: Example 3 (continued)
From Section 4.1, we know that the curve yields the minimum distance between two points (x_0, y_0) and (x_1, y_1) can be posed as finding the minimum of the following functional

$$I[y] = \int_{x_0}^{x_1} \sqrt{1 + y'^2}\, \mathrm{d}x$$

According to calculus of variations, the necessary and sufficient condition for I to be stationary is for $y(x)$ to satisfy the Euler-Lagrange equation in Eq. (49), which is

$$\frac{y''}{(1 + y'^2)^{\frac{3}{2}}} = 0$$

which implies $y'' = 0$, which further implies $y' = a$ with a as a constant. Thus we have

$$y = ax + b$$

which means the curve corresponds to the minimum distance between two points is a straight line with the two constants a and b determined from the end points.

As $\sqrt{1 + y'^2}$ is not an explicit function of x, we can use Eq. (51) to obtain the governing equation which is

$$\frac{1}{\sqrt{1 + y'^2}} = c$$

which implies

$$y' = \pm\sqrt{\frac{1}{c^2} - 1}$$

We are free to let $a = \pm\sqrt{\frac{1}{c^2} - 1}$ as it is a constant to be determined from the boundary conditions. Thus, we obtain the same straight line which is expected.

Example 7: Brachistochrone problem (continued)
From Section 4.1, we know that the Brachistochrone problem can be posed as finding the minimum of the following functional

$$I = \int_0^a \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}\, \mathrm{d}x.$$

Because the integrand does not depend on x explicitly, it is simpler for us to use Eq. (51) instead of Eq. (49) to obtain the governing differential equation. For this particular problem, we have

$$\frac{1}{\sqrt{2gy}\sqrt{1+y'^2}} = c.$$

Squaring both sides and rearranging slightly result in

$$y(1+y'^2) = \frac{1}{2gc^2} \equiv c_1.$$

which can be rewritten as

$$y' = \frac{\sqrt{k-y}}{\sqrt{y}}$$

Although strictly speaking, $y' = \pm\frac{\sqrt{k-y}}{\sqrt{y}}$, we consider the positive option only which will not affect the solution. Using $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$, we can rewrite the above equation as

$$dx = \frac{\sqrt{y}\mathrm{d}y}{\sqrt{k-y}}$$

Integrating both sides, we have

$$x = \int \frac{\sqrt{y}}{\sqrt{k-y}}\mathrm{d}y + c_1$$

Denoting $y = k\sin^2\theta$, we have

$$x = \int 2k\sin^2\theta \mathrm{d}\theta + c_1 = k(\theta - \frac{1}{2}\sin 2\theta) + c_1$$

The solution is a cycloid. As the curve starts from $(0,0)$, we conclude $c_1 = 0$ and the starting angle $\theta_s = 0$. Thus the solution is

$$x = k(\theta - \frac{1}{2}\sin 2\theta)$$

with $a = k(\theta - \frac{1}{2}\sin 2\theta)$, and $b = k\sin^2\theta$ to determine k and the ending angle θ_e. For example, if the ending point is (2,2), we can solve $k = 2.29167$ and $\theta_e = 1.20601$, and the curve is plotted in Figure 5.

4.4 Boundary conditions

Not all the stationary problems of functionals have prescribed boundary conditions for unknown functions. We also need to deal with problems with unknown functions free to vary along the boundaries. Let us consider the following functional

$$I[y] = \int_{x_0}^{x_1} \left[p(x)(y')^2 + q(x)y^2 + 2f(x)y\right]\ \mathrm{d}x + h_1y^2(x_1) + h_0y^2(x_0),$$

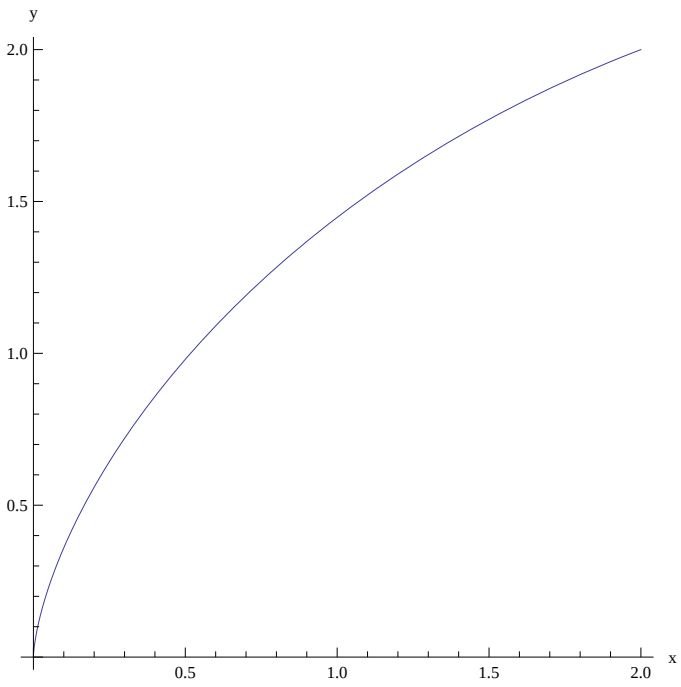

Figure 5: Solution of the Brachistochrone problem

where h_0 and h_1 are given constants. The unknown function $y(x)$ is not prescribed at the boundaries. The notation $y(x_0), y(x_1)$ indicate the evaluate of the unknown function $y(x)$ evaluated at x_0, x_1, respectively. The condition for the functional to be stationary is

$$\begin{aligned}\delta I =& 2\int_{x_0}^{x_1} \left[-(py')' + qy + f\right]\delta y \, \mathrm{d}x \\ &+ 2py'\delta y|_{x_0}^{x_1} + 2h_1 y(x_1)\delta y(x_1) + 2h_0 y(x_0)\delta y(x_0) = 0\end{aligned} \tag{52}$$

We know that δy is arbitrary in the domain $x_0 \le x \le x_1$, and its coefficient must vanish in the domain, giving us an Euler-Lagrange equation same as Eq. (26). Hence the free boundary variational problem is governed by the same Euler-Lagrange equation, Eq. (26). With the Euler-Lagrange equation being satisfied, we are left with the boundary terms in Eq. (53), which implies

$$2\left[py'(x_1) + h_1 y(x_1)\right]\delta y(x_1) + 2\left[-py'(x_0) + h_0 y(x_0)\right]\delta y(x_0) = 0 \tag{53}$$

Since δy is an arbitrary unknown function, $\delta y(x_0)$ and $\delta y(x_1)$ are arbitrary at the boundaries, we can conclude

$$p(x_1)y'(x_1) + h_1 y(x_1) = 0 \tag{54}$$

$$-p(x_0)y'(x_0) + h_0 y(x_0) = 0 \tag{55}$$

These two equations are the so-called *natural boundary conditions*. The boundary conditions in Eq. (16) are called the *geometric boundary conditions*. In general, if the functional

contains the derivative of the unknown function up to the mth order, then all the boundary conditions expressed in terms of the unknown function and its derivative up to the $(m-1)$th order are called the geometric boundary conditions and all the boundary conditions expressed in terms of the unknown function and any of its derivatives higher than $(m-1)$th order are called the natural boundary conditions. An advantage of calculus of variations is that the boundary conditions can be directly identified from the boundary terms of the first variation. For example, in Eq. (53), if there exist geometry boundary conditions at either $x = x_0$ or $x = x_1$, the corresponding term will be automatically equal to zero as the corresponding variation of the unknown function, $\delta y(x_0) or \delta y(x_1)$, is equal to zero. Otherwise, the coefficient in front of the variation of the unknown function should be equal to zero.

Thus, the free boundary variational problem should be solve using the Euler-Lagrange equation, Eq. (26), along with the natural boundary conditions, Eq. (55). Since both the natural boundary conditions and the Euler-Lagrange equation are derived in an equal base, they together are called natural conditions. It is clear that the natural conditions can be directly drawn from the Extension I of the fundamental lemma.

4.4.1 Example

Example 8: A bar of extension stiffness EA has a length L. It is subject to a distributed tensile load $p(x)$ and fixed at $x = 0$. There is a spring with spring stiffness k attached to the end at $x = L$. Find the governing differential equation and corresponding boundary conditions for this problem.

The tension behavior of this bar is governed by so-called principle of minimum total potential energy. Using $u(x)$ denoting the tensile displacement of the bar, we have

$$u(0) = 0 \tag{56}$$

as a boundary condition at $x = 0$.

We can compute the total potential energy as

$$I[u(x)] = \int_0^L \frac{1}{2} EAu'^2 \mathrm{d}x + \frac{1}{2} ku(L)^2 - \int_0^L p(x)u(x)\mathrm{d}x$$

where the first integral denotes the strain energy stored due to extension of the bar, the second term denotes the energy stored in the spring due to compression of the spring, and the last term denotes the potential due to the work done by applied loads. To minimize the potential energy, we require that the first variation of the total potential energy vanishes, yielding

$$\begin{aligned} 0 = \delta I &= \int_0^L EAu'\delta u' \mathrm{d}x + ku(L)\delta u(L) - \int_0^L p(x)\delta u(x)\mathrm{d}x \\ &= \int_0^L \left[p - (EAu')'\right] \delta u \mathrm{d}x + \left[EAu' + ku(L)\right] \delta u(L) \end{aligned} \tag{57}$$

Here the boundary term $-EAu'\delta u(0)$ is dropped as it is zero because $\delta u(0) = 0$ due to the fact that the bar is fixed at $x = 0$. From Eq. (58), we can conclude

$$p - (EAu')' = 0 \tag{58}$$

as the governing differential equation and

$$EAu' + ku(L) = 0 \tag{59}$$

as the boundary condition at $x = L$.

Then the problem can be solved using the governing differential equation in Eq. (58) along with the two boundary conditions in Eq. (56) and (59).

4.5 Subsidiary conditions

So far, we have discussed variational problem of functionals with or without prescribed boundary conditions for the unknown functions. In engineering applications, we may also meet variational problems subjected to additional constraint conditions, or so-called subsidiary conditions.

Let us consider the variational problem of the following functional

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

with the boundary conditions

$$y(x_0) = y_0 \text{ and } y(x_1) = y_1 \tag{60}$$

and a subsidiary condition

$$\int_{x_0}^{x_1} G(y,\ y',\ x)\ \mathrm{d}x = \text{constant} = C \tag{61}$$

Using an argument similar to that of the section on the Lagrange multiplier method, the stationary conditions for this variational problem can be obtained from the following variational statement

$$\delta I^* = 0 \ \text{ with } \ I^* = \int_{x_0}^{x_1} F(y,\ y',\ x)\ \mathrm{d}x + \lambda \left(\int_{x_0}^{x_1} G(y,\ y',\ x)\ \mathrm{d}x - C \right),$$

with the boundary conditions $y(x_0) = y_0$ and $y(x_1) = y_1$. λ is an arbitrary constant acting as the Lagrange multiplier. Carrying out the variation, we obtain

$$\begin{aligned} \delta I^* = &\int_{x_0}^{x_1} \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) + \lambda \left[\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right) \right] \right\} \delta y\ \mathrm{d}x \\ &+ \delta\lambda \left(\int_{x_0}^{x_1} G(y,\ y',\ x)\ \mathrm{d}x - C \right) = 0 \end{aligned} \tag{62}$$

Since δy is arbitrary in the domain, the corresponding Euler-Lagrange equation is

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) + \lambda \left[\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right) \right] = 0 \tag{63}$$

We also know that $\delta\lambda$ is arbitrary, which gives us the subsidiary condition, Eq. (61). Thus we can solve the constrained stationary value problem using the Euler-Lagrange equation

in Eq. (63) along with the subsidiary condition in Eq. (61) and the boundary conditions in Eqs. (60).

In what we just discussed, the subsidiary condition, Eq. (61), is given in terms of a definite integral. In some situations the subsidiary condition might be given in terms of differential equations. Consider to find the stationary conditions for the following functional

$$I[y] = \int_{x_0}^{x_1} F(x, y, y') \, \mathrm{d}x,$$

with the boundary conditions

$$y(x_0) = y_0, \; y(x_1) = y_1$$

and a subsidiary condition

$$G(x, y, y') = 0 \tag{64}$$

In this case, the Lagrange multiplier method still works, but the multiplier λ will be a function of the fundamental variable x. The new functional to be considered is

$$J[y, \lambda] = \int_{x_0}^{x_1} F(x, y, y') \, \mathrm{d}x + \int_{x_0}^{x_1} \lambda G(x, y, y') \, \mathrm{d}x$$

The first variation of this functional becomes

$$\delta J = \int_{x_0}^{x_1} \left\{ \left[\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'} \right)' + \lambda \frac{\partial G}{\partial y} - \left(\lambda \frac{\partial G}{\partial y'} \right)' \right] \delta y + \delta \lambda G(x, y, y') \right\} \, \mathrm{d}x$$

Here $\delta y = 0$ at $x = x_0$ and $x = x_1$ due to prescribed boundary conditions which have already been taken advantage of. The Euler-Lagrange equations are

$$\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'} \right)' + \lambda \frac{\partial G}{\partial y} - \left(\lambda \frac{\partial G}{\partial y'} \right)' = 0$$

and

$$G(x, y, y') = 0$$

which is the same as the constraint equation.

The reason why λ is a function of x is obvious. The subsidiary condition in Eq. (64) is defined at any point of x in the domain, which can be considered as infinite many point constraints. We need infinite many Lagrange multipliers, which can be mathematically replaced by a continuous function of the domain.

4.5.1 Example

Example 9: A uniform power line with length C and density ρ is hanging between two points (x_0, y_0) and (x_1, y_1). With the gravitational force acting in the y direction, find the shape of the power line in equilibrium.

The equilibrium state is determined from the minimum potential energy condition. The potential energy is

$$I[y] = \rho \int_{x_0}^{x_1} y \sqrt{1 + y'^2} \, \mathrm{d}x.$$

The length of the power line is constant, which implies

$$J = \int_{x_0}^{x_1} \sqrt{1 + y'^2}\ \mathrm{d}x - C = 0 \tag{65}$$

The boundary conditions are

$$y(x_0) = y_0 \qquad y(x_1) = y_1 \tag{66}$$

If we construct a new functional

$$I^* = I + \lambda J = \int_{x_0}^{x_1} \left[\rho y\sqrt{1 + y'^2} + \lambda\left(\sqrt{1 + y'^2} - \frac{C}{x_1 - x_0}\right)\right]\ \mathrm{d}x$$

we can easily obtain the corresponding Euler-Lagrange equation following Eq. (51) as

$$\frac{\rho y + \lambda}{\sqrt{1 + y'^2}} = c_1$$

Then

$$\int \frac{\mathrm{d}y}{\sqrt{[(\rho y + \lambda)/c_1]^2 - 1}} = \int \mathrm{d}x$$

which implies the following solution

$$\frac{\rho x}{c_1} + c_2 = \cosh^{-1}\frac{\rho y + \lambda}{c_1}.$$

The constants c_1, c_2, and λ are determined through the conditions in Eq. (65) and Eq. (66).

4.6 Functionals of higher derivatives and multiple functions

The above procedure can be easily generalized to obtain stationary conditions for definite integrals with higher derivatives. For example, considering the stationary problem of the following functional

$$I[y] = \int_{x_0}^{x_1} F(x, y, y', y'', \ldots, y^{(n)})\ \mathrm{d}x$$

The first variation is

$$\delta I[y] = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' + \ldots + \frac{\partial F}{\partial y^{(n)}}\delta y^{(n)}\right]\ \mathrm{d}x$$

By carrying out integration by parts repeatedly, we can eliminate all the derivatives of y from the integral, transforming it into the form

$$\delta I[y] = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'}\right)' + \left(\frac{\partial F}{\partial y''}\right)'' - \ldots + (-1)^n\left(\frac{\partial F}{\partial y^{(n)}}\right)^{(n)}\right]\delta y\ \mathrm{d}x$$

The Euler-Lagrange equation becomes

$$\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'}\right)' + \left(\frac{\partial F}{\partial y''}\right)'' - \ldots + (-1)^n \left(\frac{\partial F}{\partial y^{(n)}}\right)^{(n)} = 0$$

The boundary terms resulting from the integration by parts vanish when $y, y', y'', \ldots, y^{(n-1)}$ are prescribed at the boundaries $x = x_0$ and $x = x_1$ so that

$$\delta y = 0, \delta y' = 0, \ldots, \delta y^{(n-1)} = 0$$

at the boundaries or the coefficients of these variations at the boundaries are zero.

A functional could also be a function of several functions. If we want to find the stationary value of the following functional

$$I[y, z] = \int_{x_0}^{x_1} F(x, y, z, y', z') \, \mathrm{d}x,$$

where y, z are unknown functions of x, the first variation must be zero, that is

$$\begin{aligned}\delta I &= \int_{x_0}^{x_1} \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y' + \frac{\partial F}{\partial z}\delta z + \frac{\partial F}{\partial z'}\delta z' \, \mathrm{d}x \\ &= \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'}\right)'\right] \delta y \, \mathrm{d}x + \frac{\partial F}{\partial y'}\delta y|_{x_0}^{x_1} + \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial z} - \left(\frac{\partial F}{\partial z'}\right)'\right] \delta z \, \mathrm{d}x + \frac{\partial F}{\partial z'}\delta z|_{x_0}^{x_1}\end{aligned}$$

The corresponding Euler-Lagrange equation is

$$\frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial y'}\right)' = 0, \quad \frac{\partial F}{\partial z} - \left(\frac{\partial F}{\partial z'}\right)' = 0$$

If y, z are prescribed at the boundaries, the boundary related terms in δI will vanishes. If y, z are free to vary at the boundaries, the coefficients in front of δy and δz must be zero, which are the natural boundary conditions of this variational problem.

Another extension is that there could be multiple fundamental variables in the functional. Let us consider, for example, the problem of finding an extremum of the double integral

$$I[u] = \int\int_G F(x, y, u, u_{,x}, u_{,y}) \, \mathrm{d}x \, \mathrm{d}y$$

over a given region of a two-dimensional domain G by determining a suitable function u, which is a unknown function of the fundamental variables x and y and takes prescribed values on the boundary Γ. Here $u_{,x} \equiv \frac{\partial u}{\partial x}$ and $u_{,y} \equiv \frac{\partial u}{\partial y}$. The first variation must be zero, that is

$$\delta I = \int\int_G (\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_{,x}}\delta u_{,x} + \frac{\partial F}{\partial u_{,y}}\delta u_{,y}) \, \mathrm{d}x \, \mathrm{d}y$$

We know that

$$\int\int_G \frac{\partial F}{\partial u_{,x}}\delta u_{,x} \, \mathrm{d}x \, \mathrm{d}y = \int\int_G \frac{\partial(\frac{\partial F}{\partial u_{,x}}\delta u)}{\partial x} \, \mathrm{d}x \, \mathrm{d}y - \int\int_G \frac{\partial}{\partial x}(\frac{\partial F}{\partial u_{,x}})\delta u \, \mathrm{d}x \, \mathrm{d}y$$

Using the gradient theorem and considering $\frac{\partial F}{\partial u_{,x}}\delta u$ altogether as a quantity, we obtain

$$\int\int_G \frac{\partial F}{\partial u_{,x}}\delta u_{,x}\ \mathrm{d}x\ \mathrm{d}y = \int_\Gamma \frac{\partial F}{\partial u_{,x}}\delta u n_x\ \mathrm{d}s - \int\int_G \frac{\partial}{\partial x}(\frac{\partial F}{\partial u_{,x}})\delta u\ \mathrm{d}x\ \mathrm{d}y,$$

where n_x is the x component of the outward normal to the boundary line element $\mathrm{d}s$. The Euler-Lagrange equation is obtained as

$$\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u_{,x}}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial u_{,y}}\right) = 0$$

If u is prescribed along the boundary, then $\delta u = 0$. If not, then the coefficient in front of δu along the boundary should be zero, which is

$$\frac{\partial F}{\partial u_{,x}}n_x + \frac{\partial F}{\partial u_{,y}}n_y = 0.$$

It is straightforward to extend what we have learned to functionals of multiple functions and their higher derivatives, and the unknown functions could be multivariable. Such generalization will be illustrated in our mechanics applications.

4.7 Example

Example 10: Minimize the mechanical potential energy of a beam with deflection y under applied force $f(x)$,

$$I[y] = \int_0^l \left[\frac{1}{2}EI(y'')^2 - fy\right]\ \mathrm{d}x$$

with EI as the bending stiffness of the beam.

The first variation becomes

$$\begin{aligned}\delta I =& \int_0^l EIy''\delta y''\ \mathrm{d}x - \int_0^l f\delta y\ \mathrm{d}x = EIy''\delta y'|_0^l - \int_0^l (EIy'')'\delta y'\ \mathrm{d}x - \int_0^l f\delta y\ \mathrm{d}x \\ =& EIy''\delta y'|_0^l - (EIy'')'\delta y|_0^l + \int_0^l \left[(EIy'')'' - f\right]\delta y\ \mathrm{d}x\end{aligned}$$

The Euler-Lagrange equation is the famous beam deflection equation $(EIy'')'' = f$. If the end of the beam at $x = 0$ is clamped, we have both y and y' prescribed. If this end is simply supported, we know that $y = 0$ and y' is arbitrary (free to rotate), which means the coefficient in front of $\delta y'(0)$ should vanishes, *i.e.*, $EIy''(0) = 0$, which further implies that the bending moment vanishes at this end. If this end is free, then both y and y' are unknown at this end and the coefficients in front of their variations should be zero, *i.e.*, $EIy''(0) = 0$ and $(EIy''(0))' = 0$, which implies that both the bending moment and the shear force are zero. The boundary conditions at $x = l$ can be obtained in the same manner.

Example 11: Minimize the following functional

$$I[y,z] = \int_{x_0}^{x_1} (y^2 - z^2)\mathrm{d}x$$

under the derivative constraint of

$$y' - y + z = 0$$

We construct a new functional as

$$I^* = \int_{x_0}^{x_1} \left[y^2 - z^2 + \lambda(x)(y' - y + z)\right] \mathrm{d}x$$

The corresponding Euler-Lagrange equations are

$$2y - \lambda + \lambda' = 0$$

$$-2z + \lambda = 0$$

and

$$y' - y + z = 0$$

Solutions for $y(x), z(x), \lambda(x)$ can be found from these three differential equations along with boundary conditions.

Example 12: Convert a second-order system into a first-order system. Consider minimizing the following functional

$$I[y] = \int_{x_0}^{x_1} F(x, y, y', y'')\mathrm{d}x$$

with prescribed boundary conditions $y(x_0), y(x_1), y'(x_0), y'(x_1)$. We can introduce a new function

$$z(x) = y'(x)$$

so that the original variational statement becomes minimizing the following functional

$$I[y,z] = \int_{x_0}^{x_1} F(x, y, z, z')\mathrm{d}x$$

subject to a constraint

$$z - y' = 0$$

The corresponding Euler-Lagrange equations can be found using the Lagrange multiplier method. Constructing the following functional

$$I[y,z,\lambda] = \int_{x_0}^{x_1} F(x, y, z, z') + \lambda(x)(z - y')\mathrm{d}x$$

The Euler-Lagrange equations are

$$\frac{\partial F}{\partial y} + \frac{\partial \lambda}{\partial x} = 0 \tag{67}$$

$$\frac{\partial F}{\partial z} + \lambda - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial z'} = 0 \tag{68}$$

$$z - y' = 0 \tag{69}$$

From Eq. (68), we have

$$\lambda = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial z'} - \frac{\partial F}{\partial z}$$

Differentiating both sides, we have

$$\frac{\partial \lambda}{\partial x} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\frac{\partial F}{\partial z'} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial z}$$

Substituting this equation and Eq. (69) into the first Euler-Lagrange equation, we have

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\frac{\partial F}{\partial y''} = 0$$

which is the same as minimizing the original functional containing second derivatives of y directly. There are situations that the first-order system can be more conveniently solved.

Note, with introducing the constraint, the original functional can also be expressed as

$$I[y, z] = \int_{x_0}^{x_1} F(x, y, y', z')\mathrm{d}x$$

The readers are encouraged to check that one can also use this functional to convert the original second-order system into a first-order system. The Lagrange multiplier will be different but the corresponding Euler-Lagrange equations can be used to reproduce the original second-order Euler-Lagrange equations.

5 Transform from differential statements to variational statements

We have already seen how a variational statement, Eq. (41), can be transformed to an equivalent differential statement, Eq. (49). For problems which can be clearly identified as stationary problems of functionals of unknown functions, it is easier to formulate the variational statement. However situations may arise that we only know the differential statement of a problem, we want to find an equivalent variational statement for a certain purpose. Consider the following differential statement

$$L(u) + f = 0 \quad \text{in} \quad \Omega \tag{70}$$

$$B(u) - g = 0 \quad \text{on} \quad \Gamma \tag{71}$$

where L is a differential operator, B is an operator which could be differential or algebraic, Ω is the defined domain, and Γ is the corresponding boundary. The equivalent variational statement can be written as

$$\int_\Omega \delta u[L(u)+f]\ \mathrm{d}\Omega - \int_\Gamma \delta u[B(u)-g]\ \mathrm{d}\Gamma = 0 \tag{72}$$

If L is a linear and self-adjoint operator, which is true for most engineering problems, we can perform integration by parts to achieve the following variational statement

$$\delta\Pi(u) = 0 \tag{73}$$

where $\Pi(u)$ is a functional of the function u.

5.1 Example 13

Find the corresponding variational statement for the two-dimensional steady-state heat conduction problem for a homogeneous material which is governed by the following differential statement

$$k(\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2})+Q=0 \ \text{ in } \ \Omega \tag{74}$$

$$\phi-\bar{\phi}=0 \ \text{ on } \ \Gamma_\phi \tag{75}$$

$$k\frac{\partial\phi}{\partial n}-\bar{q}=0 \quad \text{on } \ \Gamma_q \tag{76}$$

where ϕ is the temperature field, k is the heat conduction coefficient, n is the outward normal of the boundary curve, Q is the density of internal heat source, $\bar{\phi}$ and $\bar{q}$ are given temperature or heat flux on the boundary Γ_ϕ and Γ_q, respectively. Let us restrict that $Q, \bar{\phi}$, and $\bar{q}$ are not functions of ϕ. According to Eq. (72), we have

$$\int_\Omega \delta\phi[k(\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2})+Q]\ \mathrm{d}\Omega - \int_{\Gamma_q}\delta\phi(k\frac{\partial\phi}{\partial n}-\bar{q})\ \mathrm{d}\Gamma = 0$$

by taking advantage of the fact that $\delta\phi = 0$ on Γ_ϕ. After integrating by parts, we obtain

$$\begin{aligned}&\int_\Omega\left(-k\frac{\partial\delta\phi}{\partial x}\frac{\partial\phi}{\partial x}-k\frac{\partial\delta\phi}{\partial y}\frac{\partial\phi}{\partial y}+\delta\phi Q\right)\ \mathrm{d}\Omega+\\ &\int_{\Gamma_\phi+\Gamma_q} k\delta\phi\left(\frac{\partial\phi}{\partial x}n_x+\frac{\partial\phi}{\partial x}n_x\right)\ \mathrm{d}\Gamma-\int_{\Gamma_q}\delta\phi(k\frac{\partial\phi}{\partial n}-\bar{q})\ \mathrm{d}\Gamma=0\end{aligned} \tag{77}$$

Recognizing $\frac{\partial\phi}{\partial x}n_x+\frac{\partial\phi}{\partial y}n_y=\frac{\partial\phi}{\partial n}$ and $\delta\phi=0$ along $\Gamma_\phi=0$, we can simplify the above formula as

$$\int_\Omega\left(-k\frac{\partial\delta\phi}{\partial x}\frac{\partial\phi}{\partial x}-k\frac{\partial\delta\phi}{\partial y}\frac{\partial\phi}{\partial y}+\delta\phi Q\right)\ \mathrm{d}\Omega+\int_{\Gamma_q}\delta\phi\bar{q}\ \mathrm{d}\Gamma=0 \tag{78}$$

which implies

$$\delta\Pi(\phi)=0 \ \text{ with } \ \Pi(\phi)=\int_\Omega[\frac{1}{2}k(\frac{\partial\phi}{\partial x})^2+\frac{1}{2}k(\frac{\partial\phi}{\partial y})^2-\phi Q]\ \mathrm{d}\Omega-\int_{\Gamma_q}\phi\bar{q}\ \mathrm{d}\Gamma.$$

6 The Ritz method

Only a very limited set of Euler-Lagrange equations have obtainable exact solutions. Most governing differential equations in engineering applications in general and mechanics in particular are very difficult, if not impossible, to obtain. We need to seek approximations to the exact solution as alternatives. The Ritz method provides us an approach to find a solution (either exact or approximate) by directly operating the corresponding variational statement. Two concepts become necessary: *complete* system of functions and *admissible* functions. If any continuously differentiable function, $y(x)$, is uniformly approximated by $y_n(x)$, a linear combination of finite number of $w_1, w_2, \ldots, w_n$, *i.e.*,

$$y_n \approx \sum_{k=1}^{n} a_k w_k,$$

the system of $w_1, w_2, \ldots, w_n$ is called a complete system. Examples of the complete system are polynomials and trigonometric functions. For a function to be admissible, it should satisfy the geometric boundary conditions and continuously differentiable up to $m-1$ order, with m as the highest-order derivative of the unknown function in the functional.

The basic idea of the Ritz method is to find the "best" solution from a set of trial solutions which are admissible and selected from a complete system of functions. The general solution procedure using the Ritz method for a given variational statement can be summarized as the following:

1. Select a set of trial functions, $w_1, w_2, \ldots, w_n$, which are admissible and from a complete system of functions;

2. Form the linear combination of the trial functions as the approximation of the unknown function such that $y_n = \sum_{k=1}^{n} a_k w_k$ with a_k as unknown constant coefficients;

3. Substitute y_n as y into the functional and perform the integration to evaluate the functional to be a function of a_i, *i.e.*, $I[y] = I(a_1, a_2, \ldots, a_n)$;

4. Obtain a system of equations $\frac{\partial I}{\partial a_i} = 0$ by carrying out a set of partial derivatives $\frac{\partial I}{\partial a_i}, i = 1, 2, \ldots, n$;

5. Solve for the unknown coefficients a_k, and an approximation to y is found as $y \approx \sum_{k=1}^{n} a_k w_k$.

The accuracy of approximation can be improved by including more terms in the trial function. If the trial functions contain the exact solution, the Ritz method will help us find the exact solution. It is emphasized here that one should start from the beginning terms of the complete function series to add more terms in the complete function system for accuracy improvements. Only if we know that some terms will not contribute, we can skip these terms. For example, for the complete function system of polynomials, we can select functions from $1, x, x^2, x^3, \ldots$. If we know that our solution will be symmetric with respect to x, we can choose from $1, x^2, x^4, \ldots$.

Usually, we consider the Ritz method as a direct method to solve the stationary problems of functionals and solving the Euler-Lagrange equations as an indirect method. A

hybrid of the direct and indirect method called the *Kantorovich method* takes a slightly different path from that of the Ritz method we just discussed. Consider a functional that has the form

$$I[y] = \int_t \int_x F(x, t, y) \, \mathrm{d}x \, \mathrm{d}t$$

with the boundary conditions $y(x_1, t) = y_1(t), y(x_2, t) = y_2(t)$, and $y(x, t_1) = y_3(x)$. Suppose the trial function is of the form

$$y(x, t) = g_p(x) + \sum_{j=1}^{n} f_j(t) g_j^H(x)$$

where $g_p(x)$ satisfies the boundary conditions and $g_j^H(x)$ satisfies the homogeneous boundary conditions. At this moment, no condition is prescribed for $f_j(t)$. When this trial function is substituted into the functional, a new reduced-degree-freedom functional

$$I[f] = \int_t F\left(f_j(t)\right) \, \mathrm{d}t$$

from which we can derive the corresponding Euler-Lagrange equations to solve for $f_j(t)$. In many cases, the Kantorovich method yields a better solution than the straight Ritz method because part of the procedure is analytical.

6.1 Examples

Example 14
Using the Ritz method to solve the following differential equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + u + x = 0 \quad \text{for} \quad 0 \le x \le 1$$

along with the boundary condition $u(0) = u(1) = 0$ approximately.
Solution: It is easy to verify that the given differential statement is equivalent to the following variational statement

$$\delta I[u] = 0 \quad \text{with} \quad I[u] = \int_0^1 [-\frac{1}{2}(\frac{\mathrm{d}u}{\mathrm{d}x})^2 + \frac{1}{2}u^2 + ux] \, \mathrm{d}x.$$

Let us select the trial function as

$$u_1 = a_1 x(1 - x)$$

which satisfies the given boundary conditions. Substituting the trial function into the functional, and after simplification, we obtain the following:

$$I = -\frac{1}{2}\frac{3}{10}a_1^2 + \frac{1}{12}a_1.$$

From $\delta I = 0$, we obtain $a_1 = \frac{5}{18}$. Hence, we find an approximate solution

$$u \approx u_1 = \frac{5}{18}x(1 - x).$$

It can be easily verified that u_1 does not satisfy the governing differential equation.

If we choose another trial function as

$$u_2 = a_1 \sin x + a_2 x$$

satisfying the boundary condition at $x = 0$. To satisfy the boundary condition at $x = 1$,

$$a_1 \sin 1 + a_2 = 0$$

we need $a_2 = -a_1 \sin 1$, which means the trial solution should be modified to be

$$u_2 = a_1(\sin x + x \sin 1).$$

Substituting the trial function into the functional, and after simplification, we obtain the following:

$$I = -\frac{1}{2}a_1^2 \sin 1(\frac{2}{3}\sin 1 - \cos 1) + a_1(\frac{2}{3}\sin 1 - \cos 1).$$

From $\delta I = 0$, we obtain $a_1 = \frac{1}{\sin 1}$. Hence, we find another approximate solution

$$u \approx u_2 = \frac{\sin x}{\sin 1} - x.$$

In fact, we can check that u_2 satisfy the governing differential equation identically, which means we have found the exact solution because our set of trial functions incidentally contains the exact solution. Figure 6 shows the difference between different solutions for the same problem, where u_3 is obtained by the trial function $u_3 = a_1 x(1-x) + a_2 x^2(1-x^2)$. Interested readers are encouraged to obtain this solution themselves. The difference between u_3 and u_2 (the exact solution) is smaller than that between u_1 and u_2, which makes perfect sense because u_3 has more terms in the trial solution.

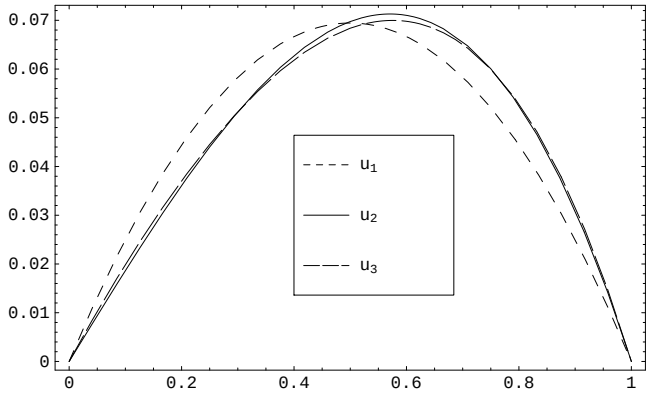

Figure 6: Approximations found by the Ritz method.

Example 15
A two-dimensional domain Ω defined as $x \in [0, \pi]$ and $y \in [0, 1]$. Solve the following partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \ \text{ in } \ \Omega$$

along with the boundary conditions $u|_{x=0} = u|_{x=\pi} = 0$, $u|_{y=0} = 0$, and $u|_{y=1} = \sin x$.

Solution: It is not trivial to find the solution for this partial differential equation directly, however according to Section 5.1, we know this differential equation corresponds to the stationary value of the following functional

$$I[u] = \frac{1}{2} \int_0^1 \int_0^\pi [(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2] \, \mathrm{d}x \, \mathrm{d}y$$

Assume the trial function to be $u(x, y) = f(y) \sin x$. Notice this trial function satisfies the boundary conditions in x. Substituting this function into the functional and integrating over x, we can reduce $I[u]$ to be

$$I[f] = \frac{1}{4}\pi \int_0^1 \left[f^2 + (f')^2\right] \, \mathrm{d}y$$

The Euler-Lagrange equation of this reduced functional is an ordinary differential equation in terms of f

$$f - f'' = 0.$$

As for the boundary conditions for this ordinary differential equation, we know $u|_{y=0} = 0$, which implies $f(0) = 0$, and $u|_{y=1} = \sin x$, which implies $f(1) = 1$. The solution is then

$$f(y) = \frac{\sinh(y)}{\sinh 1}$$

Hence the complete solution can be written as

$$u(x, y) = \frac{\sinh(y)}{\sinh 1} \sin x.$$

which can be easily verified to be the exact solution of the original partial differential equations with the prescribed boundary conditions.

7 The Galerkin method

There is another method for finding approximate solutions for differential equations. Although it is not directly related with the calculus of variations, it is of the same spirit and can be derived from the stationary problem of a definite integral.

Let us consider the stationary value problem that we have discussed in Section 4.2. If we use the Ritz method to solve this problem by assuming

$$y = \frac{x - x_0}{x_1 - x_0} y_1 + \frac{x - x_1}{x_0 - x_1} y_0 + a_1 w_1 + a_2 w_2 + \ldots + a_n w_n. \tag{79}$$

with w_1, w_2 taking from a complete system and satisfying the homogenous boundary conditions. The Ritz method yields

$$\frac{\partial I}{\partial a_i} = 2\int_0^l [py'\frac{\partial y'}{\partial a_i} + qy\frac{\partial y}{\partial a_i} + f\frac{\partial y}{\partial a_i}]\ \mathrm{d}x = 2\int_0^l (py'w_i' + qyw_i + fw_i)\ \mathrm{d}x = 0$$

for $i = 1, 2, \ldots, n$. After integration by parts, the above equation can be written as

$$2\int_0^l [(py')' + qy + f]w_i\ \mathrm{d}x = 0 \tag{80}$$

It is interesting to note that the integrand in the above equation is the multiplication of the Euler-Lagrange equation by the fundamental function. The n equations in Eq. (80) can be used to determine the corresponding n unknowns $a_1, a_2, \ldots, a_n$. The method using Eq. (80) to find an approximation solution, Eq. (79), is called the Galerkin method for the differential equation $(py')' + qy + f = 0$.

In general, if we have a problem with a given differential equation $L[u] = 0$ defined over a two-dimensional domain G and boundary conditions $u(s) = \psi(s), M[u] = m(s)$, where L, M are differential operators, then we can find an approximate solution for this problem using the Galerkin method. We choose the fundamental functions to satisfy

$$w_n(s) = 0 \qquad M[w_n] = 0$$

and

$$u_0(s) = \psi(s), M[u_0] = m(s)$$

then we can use the Galerkin method to seek an approximation solution in the following form

$$u = u_0 + a_1w_1 + a_2w_2 + \ldots + a_nw_n,$$

with the unknown coefficients determined from

$$\int\int_G L[u]w_i\ \mathrm{d}A = 0 \quad \text{for} \quad i = 1, 2, \ldots, n \tag{81}$$

where $\mathrm{d}A$ is the surface element of G in which the differential equation is defined.

Although the Galerkin method can be derived from the Ritz method for stationary problems of definite integrals, the Galerkin method can be used to solve a differential statement approximately, even if a corresponding variational statement does not exist. Sometimes the formulation in Eq. (81) is called the weak form of the differential statement. From the above derivation, we can also easily concluded that whenever a corresponding variational statement exists for the differential statement, the Galerkin method will yield the same results as the Ritz method. However, it is noted that the trial solution should satisfy all the boundary conditions including the geometry boundary conditions and the natural boundary conditions while the trial solution for Ritz method needs only satisfying the geometry boundary conditions.

7.1 Example

Example 16 Using the Galerkin method to solve the following differential equation

$$u'' + u = 2x \quad \text{for} \quad 0 \le x \le 1$$

along with the boundary condition $u(0) = u(1) = 0$ for an approximate unknown function $u(x)$.

Solution:

Let us use polynomials to find our approximation. Let us try the following function

$$u_1(x) = a_1 x(1-x)$$

Clearly this function satisfies the boundary conditions. To determine the unknown constant a_1, we need to form the following integral according to Eq. (81)

$$\int_0^1 (u'' + u - 2x)x(1-x)\mathrm{d}x = 0$$

Note to carry out the integration, we need to substitute the trial function, which results in

$$-\frac{1}{6} - \frac{3a_1}{10} = 0 \to a_1 = -\frac{5}{9}$$

Hence, we have find an approximate solution as

$$u_1(x) = -5/9x(1-x)$$

If we are not satisfied with this approximation, we can try the following function

$$u_2(x) = a_1 x(1-x) + a_2 x^2(1-x)$$

Again this trial function satisfies the boundary conditions. To determine the unknown constants a_1 and a_2, we need to form the following integrations according to Eq. (81)

$$\int_0^1 (u'' + u - 2x)x(1-x)\mathrm{d}x = 0 \qquad \int_0^1 (u'' + u - 2x)x^2(1-x)\mathrm{d}x = 0$$

which result in

$$-\frac{1}{6} - \frac{3a_1}{10} - \frac{3a_2}{20} = 0 \qquad -\frac{1}{10} - \frac{3a_1}{20} - \frac{13a_2}{105} = 0$$

The unknown constants determined from these two equations are $a_1 = -\frac{142}{369}$ and $a_2 = -\frac{14}{41}$. Hence, we have find another approximate solution as

$$u_2(x) = -\frac{142}{369}x(1-x) - \frac{14}{41}x^2(1-x)$$

The given boundary value problem is simple and can be solved exactly with the solution as

$$u(x) = 2(x - \frac{\sin x}{\sin 1})$$

To illustrate our approximations, we also plot the approximate solution u_1 and u_2 obtained using the Galerkin method along with the exact solution in Figure 7. As shown in the plot, u_2, although has a different form from the exact solution, agrees with the exact solution numerically very well.

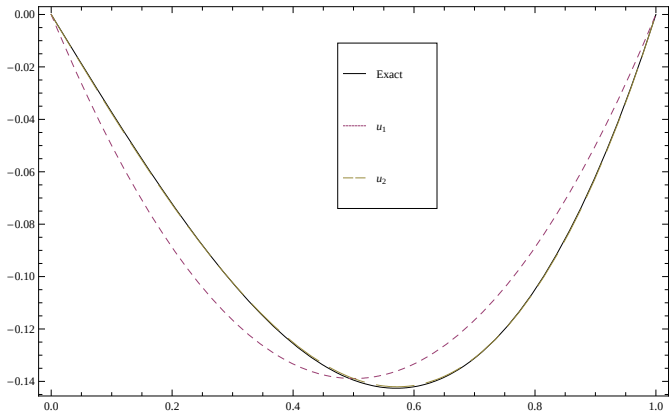

Figure 7: Approximations found by the Galerkin method.

8 Variational-asymptotic method

In many problems of mechanics and physics, there are small parameters such as the thickness of a plate or a shell, the dimension of the cross-section of a beam, or the wavelength and frequency of an oscillating body. Such problems can be properly investigated using asymptotic methods. Analogous to differential statements and variational statements for formulating a problem, there are two types of asymptotic methods to analyze problems involving small parameters: formal asymptotic method and the variational-asymptotic method (VAM). The formal asymptotic method expands the field variables into an asymptotic series and then perform an order analysis of the differential statement of the problem to solve the field variables according to different orders. It is clear that for the problems which allow variational formulations, there should exist a direct variational approach based on direct asymptotic analysis of the corresponding functionals and automatically taking into account the variational structure of the equations and those properties of the solutions. The method of asymptotic analysis of functionals, which is called the variational-asymptotic method by its author Prof. Berdichevsky, allows one to consider the stationary value problem for functions or functional involving small parameters from a common point of view. For this very reason, we do not differentiate the use of function and functional although their meanings become self-evident for specific cases. It is a very useful method to simplify the procedure of finding the stationary points for a functional depending on one or more small parameters. It is applicable to problems that can be posed in terms of seeking the stationary points of a functional with some inherent small parameters. It is especially the right tool for construction of accurate models for composite structures because their behavior is governed by a variational statement, and there are some small geometrical parameters characterizing these structures, such as the dimension of the cross-section of a beam.

There are a number of advantages of the variational-asymptotic method (VAM) compared to the widely used formal asymptotic method (FAM). First, VAM is much simpler than FAM as VAM only deals with one functional while FAM deals with a system of differential equations. Hence, dropping a small term in the functional is equivalent to

neglecting such quantities in several differential equations simultaneously, which are not always easy to recognize as small ones. Second, the approximate equations obtained using VAM is always variational, which is not necessarily the case with FAM. Thus it is easier to prove the correctness of approximate equations obtained by VAM, as well as to modify them if needed. The variational natural of the equations obtained using VAM is also very convenient if one wants to solve the approximate equations using a numerical technique such as the finite element method.

8.1 Basics of asymptotic analysis

To understand VAM, we need first to know some basics of asymptotic analysis. The asymptotic analysis is a mathematical approach providing an alternative for us to solve the problem if the exact solution is very difficult to obtain.

Let us use a simple example to understand the use of the asymptotic series intuitively. For example, a student wants to find the numerical value of $\sin 39^\circ$ but his calculator is not powerful enough to evaluate trigonometric functions although he knows $\sin 30^\circ = \frac{1}{2}$ and $\cos 30^\circ = \frac{\sqrt{3}}{2}$. He can find the value approximately using the Taylor series technique. We can expand $\sin\theta$ around a value θ_0 as

$$\sin\theta = \sin\theta_0 + (\theta-\theta_0)\cos\theta_0 - \frac{1}{2}(\theta-\theta_0)^2\sin\theta_0 - \frac{1}{6}(\theta-\theta_0)^3\cos\theta_0 + \dots \tag{82}$$

Let $\theta_0 = 30^\circ$ and $\theta = 39^\circ$, we can approximate $\sin 39^\circ$ using the first two terms of the Taylor series as

$$\sin 39^\circ = \sin 30^\circ + (39-30)\frac{\pi}{180}\cos 30^\circ = \frac{1}{2} + \frac{\sqrt{3}\pi}{40} = 0.636035 \tag{83}$$

If we calculate the value up to the first three terms, we have

$$\begin{aligned}\sin 39^\circ &= \sin 30^\circ + (39-30)\frac{\pi}{180}\cos 30^\circ - \frac{1}{2}\left[(39-30)\frac{\pi}{180}\right]^2\sin\theta_0 \\ &= \frac{1}{2} + \frac{\sqrt{3}\pi}{40} - \frac{\pi^2}{1600} = 0.629866\end{aligned} \tag{84}$$

Of course, with an advanced calculator, you will know the exact value of $\sin 39^\circ$ is 0.62932. It is remarkable that using the first three terms of the Taylor series relying on very basic algebraic skills (plus, minus, and times), we already get the first three digits correct with an error less than 0.1%. A requirement for this to work is that $\theta - \theta^0$ must be smaller than one in radians, that is a small number. Otherwise, the higher powers of $\theta - \theta^0$ will introduce significant errors and the first few terms of the Taylor series will not be able to approximate the exact value of the original function. If $\theta - \theta^0$ is small in radians, the Taylor series is in fact an asymptotic series of $\sin\theta$, or the asymptotic expansion of $\sin\theta$.

VAM is also based on the idea of neglecting the small terms. However, to apply it in a systematic way, one has to learn how to recognize small terms. To this end, we need to learn some terminology frequently used in asymptotic analysis: $O, o, \sim$. Suppose $f(x)$ and $g(x)$ are continuous functions defined on some domain and possess limits as $x \to x_0$ in the domain. Then we can define the following shorthand notation for the relative properties of these functions in the limit $x \to x_0$.

- $f(x) = O\left(g(x)\right)$ as $x \to x_0$ if $|f(x)| \leq K|g(x)|$ in the neighborhood of x_0 with K denoting a constant. We say that $f(x)$ is asymptotically bounded by $g(x)$ in magnitude as $x \to x_0$ or $f(x)$ is of the order of $g(x)$.

- $f(x) = o\left(g(x)\right)$ as $x \to x_0$ if $|f(x)| \leq \epsilon|g(x)|$ in the neighborhood of x_0 for all positive value ϵ. We say that $f(x)$ is asymptotically smaller than $g(x)$.

- $f(x) \sim g(x)$ as $x \to x_0$ if $f(x) = g(x) + o\left(g(x)\right)$ in the neighborhood of x_0. We say that $f(x)$ is asymptotically equal to $g(x)$.

To correctly recognize small terms in a functional, we not only need to know the asymptotic order of the functions, but also often need to know the asymptotic order of their derivatives. We need to introduce the notion of the characteristic length. Consider a function $f(x)$ defined for $x \in [a, b]$ and sufficiently smooth in this domain. We denote the amplitude of change of $f(x)$ on $[a, b]$ as the maximum difference of the function evaluated at any two points in the domain, *i.e.*

$$\bar{f} = \max_{x_1, x_2 \in [a,b]} |f(x_1) - f(x_2)| \tag{85}$$

Then for a sufficiently small number l, the following inequality holds:

$$\left|\frac{\mathrm{d}f}{\mathrm{d}x}\right| \leq \frac{\bar{f}}{l} \tag{86}$$

The largest constant l satisfying the above inequality is termed the characteristic length of function $f(x)$ in its own definition domain. For example if we need to estimate higher derivatives, then the corresponding terms are included in the definition of l, and the characteristic length is the largest constant satisfying the following inequalities

$$\left|\frac{\mathrm{d}f}{\mathrm{d}x}\right| \leq \frac{\bar{f}}{l}, \qquad \left|\frac{\mathrm{d}^2 f}{\mathrm{d}^2 x}\right| \leq \frac{\bar{f}}{l^2}, \qquad \ldots, \qquad \left|\frac{\mathrm{d}^k f}{\mathrm{d}^k x}\right| \leq \frac{\bar{f}}{l^k} \tag{87}$$

where k is the highest derivative we want to estimate the asymptotic order. If the derivatives of the function change significantly on the domain, we need to use the local characteristic length $l(x)$ for different segments of the domain or even for each point.

This definition of characteristic length can be easily generalized to functions of multiple variables.

8.2 Operating rules of the variational-asymptotic method

The variational-asymptotic method is of heuristic character. It is formulated as a set of rules, along with their applications illustrated using examples.

Let a functional $I(u, \varepsilon)$ depending on a small parameter ε be given at some set M of elements u. For a beam-like structure, the variable u represents the 3D displacement field and ε is the aspect ratio of the cross-section with respect to the length along the beam axis. Assuming that the functional $I(u, \varepsilon)$ has a stationary point, denoted by $\breve{u}$, it is clear that $\breve{u}$ is a function of the small parameter ε. We further assume that $\breve{u}$ approaches its

asymptotic limit u_0 as $\varepsilon \to 0$, which is often called the zeroth-order approximation of $\breve{u}$. For real problems, $\breve{u}$ might be very difficult to obtain, we are asking whether it is possible to construct a functional based on the information given by $I(u, \varepsilon)$ so that its stationary value is u_0? How can we construct refined functionals, the stationary points of which are the approximations of $\breve{u}$ with given accuracy in terms of the small parameter ε?

It is clear that answers to these questions are related to the asymptotic analysis of the functional $I(u, \varepsilon)$. Let us start with the functional $I_0(u)$ by neglecting terms that are the small in asymptotic sense in $I(u, \varepsilon)$, *i.e.*, $I_0(u) = I(u, 0)$. The following cases might be encountered:

- Case 1: $I_0(u)$ has isolated stationary points
- Case 2: $I_0(u)$ has non-isolated stationary points
- Case 3: $I_0(u)$ does not have stationary points
- Case 4: $I_0(u)$ is meaningless, *i.e.*, the functional $I(u, \varepsilon)$ is not defined for $\varepsilon = 0$.

8.2.1 Case 1

If $I_0(u)$ has isolated stationary points $\breve{u}_0$, one can expect that these stationary points are the first approximations of the stationary points of the original functional $I(u, \varepsilon)$. The second asymptotic term can be obtained in the following way. Let us express $u = \breve{u}_0 + u'$, where $u' \to 0$ for $\varepsilon \to 0$. Substituting this expression into the original functional $I(u, \varepsilon)$, we will have $I(\breve{u}_0 + u, \varepsilon)$. Keeping the leading terms containing u', we can obtain a functional $I_1(u', \varepsilon)$. The stationary points with respect of u' of this functional are the next approximation of the asymptotic series of the stationary points u of the original functional $I(u, \varepsilon)$. Next, we are going to use an example to show how this works.

Example 12 Find the zeroth-order approximation of the stationary points of the function of one variable u

$$f(u, \varepsilon) = u^2 + u^3 + 2\varepsilon u + \varepsilon u^2 + \varepsilon^2 u$$

with ε as a small parameter.

According to VAM, we obtain the zeroth-order approximation of the function as $f_0(u) = f(u, 0) = u^2 + u^3$, which has two stationary points $u = 0$ and $u = -\frac{2}{3}$. For the original function $f(u, \varepsilon)$, the stationary points are

$$u = \frac{1}{3}(-1 - \varepsilon \pm \sqrt{1 - 4\varepsilon - 2\varepsilon^2})$$

Expanding these two stationary points asymptotically in terms of ε, we have

$$u = \begin{cases} -\dfrac{2}{3} + \dfrac{\varepsilon}{3} + \varepsilon^2 + o(\varepsilon^2) \\ 0 - \varepsilon - \varepsilon^2 + o(\varepsilon^2) \end{cases}$$

It is clear that the stationary points we find from $f_0(u)$ are the zeroth-order approximation of the stationary points of $f(u, \varepsilon)$.

To find the next approximation, we need to start with the two stationary points we just found in the zeroth-order approximation. Consider we need to find our next approximation in the neighborhood of $-\frac{2}{3}$, which is one of our $\check{u}_0$. Setting $u = -\frac{2}{3} + u'$, we obtain the following function

$$f(-\frac{2}{3} + u', \varepsilon) = -u'^2 + \frac{2u'\varepsilon}{3} + \underline{u'^3 + u'^2\varepsilon + u'\varepsilon^2} + \underline{\underline{\frac{4}{27} - \frac{8\varepsilon}{9}}}$$

The double underlined terms are additive constants which will not affect the stationary points and can be simply dropped. The underlined terms are much smaller than those not underlined terms. That is

$$|u'^3| \ll |u'^2| \qquad |u'^2\varepsilon| \ll |u'^2| \qquad |u'\varepsilon^2| \ll |\frac{2u'\varepsilon}{3}|$$

in view of the fact that both u' and ε are small. Keeping the leading terms with respect to u' in the function $f(-\frac{2}{3} + u', \varepsilon)$, we arrive at the following function

$$f_1(u', \varepsilon) = -u'^2 + \frac{2u'\varepsilon}{3}$$

It is stationary when $u' = \frac{1}{3}\varepsilon$. Note that the asymptotic order of u' is not assumed a priori, but is determined as the stationary point of the function $f_1(u', \varepsilon)$. Hence, we have obtained the first-order approximation of the stationary point in the neighborhood of $-\frac{2}{3}$ as

$$\check{u}_\varepsilon = -\frac{2}{3} + \frac{1}{3}\varepsilon + o(\varepsilon)$$

The first-order approximation in the neighborhood of 0, which is also one of our $\check{u}_0$ can be obtained analogously. Setting $u = 0 + u' = u'$, we obtain the following function

$$f(u', \varepsilon) = u'^2 + 2\varepsilon u' + \underline{u'^3 + \varepsilon u'^2 + \varepsilon^2 u'}$$

The underlined terms are much smaller than those not underlined terms. That is

$$|u'^3| \ll |u'^2| \qquad |u'^2\varepsilon| \ll |u'^2| \qquad |u'\varepsilon^2| \ll |2u'\varepsilon|$$

in view of the fact that both u' and ε are small. Keeping the leading terms with respect to u' in the function $f(u', \varepsilon)$, we arrive at the following function

$$f_1(u', \varepsilon) = u^2 + 2u'\varepsilon$$

It is stationary when $u' = -\varepsilon$. Hence, we have obtained the first-order approximation of the stationary point in the neighborhood of 0 such that

$$\check{u}_\varepsilon = 0 - \varepsilon + o(\varepsilon)$$

Till now, we have exactly reproduced that the first two terms of the asymptotic expansion of the exact solution. We can continue this process to find higher-order approximations. For example, for the stationary point of $f(u, \varepsilon)$ in the neighborhood of 0, we set $u =$

$0 - \varepsilon + u'' = -\varepsilon + u''$, with $u'' = o(\varepsilon)$. Keeping only the leading terms with respect to u'' in the function

$$f(-\varepsilon + u'', \varepsilon) = u''^2 + 2u''\varepsilon^2 + \underline{u''^3 - 2u''^2\varepsilon} - \underline{\underline{(\varepsilon^2 + \varepsilon^3)}}$$

we get the following functional

$$f_2(u'', \varepsilon) = u''^2 + 2u''\varepsilon^2$$

because we know

$$|u''^3| \ll |u''^2| \qquad |2u''^2\varepsilon| \ll |u''^2|$$

in view of the fact that both u' and ε are small. The stationary point of $f_2(u'', \varepsilon)$ is $u'' = -\varepsilon^2$. Then we have obtained the second-order approximation of the stationary point as

$$\check{u}_\varepsilon = 0 - \varepsilon - \varepsilon^2 + o(\varepsilon^2)$$

Note that the terms that were insignificant in obtaining the leading terms of one approximation may become important in determining the leading terms of the function for the next approximation. For example, $2u''^2\varepsilon$ is neglected in the function for the first-order approximation, yet it becomes essential for the second-order approximation.

Similarly we can construct the second-order approximation near the neighborhood of $-2/3$. The approximate stationary values and the exact values in the neighborhood of $-2/3$ and 0 are plotted with respect to the small parameter ε in Figure 8 and Figure 9, respectively. As shown in these two plots, when ε is very small, say smaller than 0.05, both the first-order approximation and the second-order approximation estimate the exact values very well. However, when ε gets bigger, say smaller than 0.1 yet greater than 0.05, only the second-order approximation gives an accurate estimation. When ε gets even bigger, we need find higher-order approximations. Fortunately, in most engineering problems, when we consider one parameter, say the aspect ratio of a beam or a plate, is small, usually it is smaller than 0.1, which intuitively implies that the second-order approximation usually can provide a good approximation for engineering problems involving small parameters.

The main issue in the asymptotic analysis is to recognize the leading terms and the negligible terms. Usually, this is the most important and most difficult point of the asymptotic analysis. It was relatively easy to identify such terms in the example we just considered. To determine which terms are negligible, we need to consider the following two conditions:

- For two terms $A(u, \varepsilon)$ and $B(u, \varepsilon)$ in the functional $I(u, \varepsilon)$, if

$$\lim_{\varepsilon \to 0} \max_{u \in M} \left| \frac{B(u, \varepsilon)}{A(u, \varepsilon)} \right| = 0 \tag{88}$$

  then $B(u, \varepsilon)$ is negligible in comparison to $A(u, \varepsilon)$ for all stationary points. Such terms are called globally secondary.

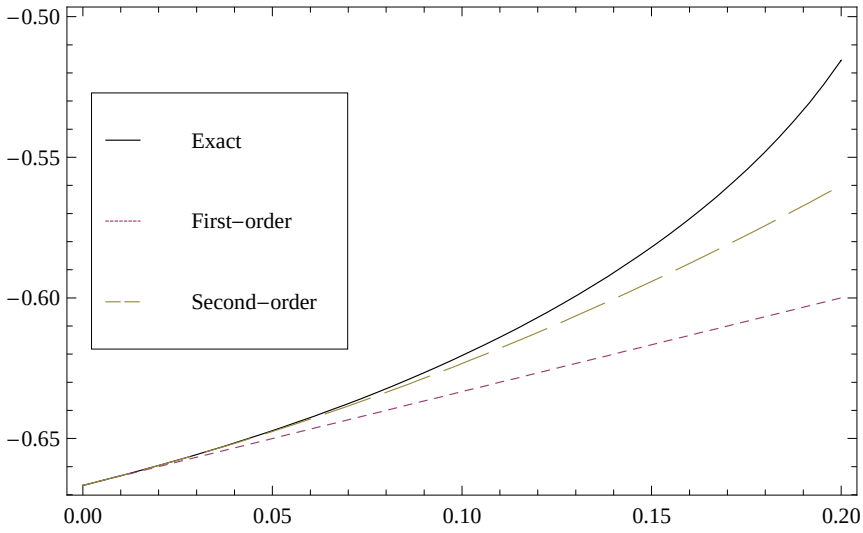

Figure 8: Comparison of the stationary points in the neighborhood of $-2/3$.

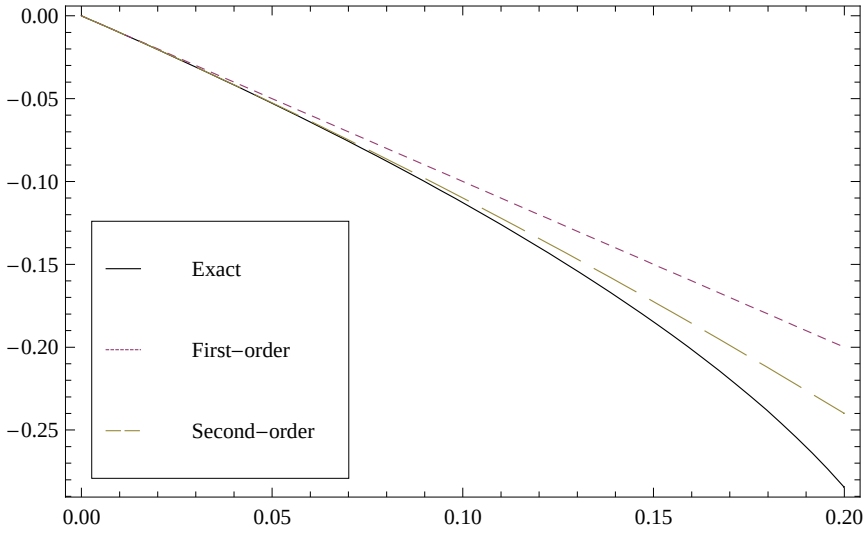

Figure 9: Comparison of the stationary points in the neighborhood of 0.

- Let $\breve{u}_\varepsilon \to 0$ for $\varepsilon \to 0$, and for any sequence $\{u_n\}$ converging to $u = 0$. If

$$\lim_{\substack{n\to\infty \\ \varepsilon\to 0}} \left| \frac{B(u,\varepsilon)}{A(u,\varepsilon)} \right| = 0 \tag{89}$$

  then $B(u,\varepsilon)$ is negligible in comparison to $A(u,\varepsilon)$ for the stationary point $\breve{u}_\varepsilon$. Such terms are called locally secondary.

In Example 1, the term εu^2 is globally secondary with respect to u^2, the term $\varepsilon^2 u$ is globally secondary with respect to $2\varepsilon u$ while u^3 is locally secondary with respect to u^2 in the neighborhood of the point $u = 0$. However, along with these terms, the functional may have some terms which cannot be easily identified as "small" because the corresponding "large" terms are not present in the functional. Consider, for example, the function

$$f(u,\varepsilon) = u^3 + \varepsilon u^2 + \varepsilon^3 u \tag{90}$$

with ε as a small parameter. The function $f_0(u) = f(u,0) = u^3$, and one may expect that all stationary points of the function $f(u,\varepsilon)$ are in the vicinity of $u = 0$. When we proceed to the next approximation by setting $u = 0 + \breve{u}_\varepsilon$, we will have

$$f(\breve{u}_\varepsilon,\varepsilon) = \breve{u}_\varepsilon^3 + \varepsilon\breve{u}_\varepsilon^2 + \varepsilon^3\breve{u}_\varepsilon \tag{91}$$

It seems we have a difficulty here to determine which terms are negligible because we don't know the relative order of $\breve{u}_\varepsilon$ and ε. If, for example, $\breve{u}_\varepsilon \sim \varepsilon$, then we can neglect the last term. If $\breve{u}_\varepsilon \sim \varepsilon^2$, then we can neglect the first term and keep the last two terms. However, such information is not known a priori. Hence we call such terms as doubtful terms because we cannot determine whether they are negligible or not. All the terms in Eq. (91) are doubtful.

A recipe for studying a functional with doubtful terms is suggested as follows. We first drop some doubtful terms in the functional $I(u,\varepsilon)$ to obtain a modified functional $\tilde{I}(u,\varepsilon)$. Then we find the stationary points of the functional $\tilde{I}$ and evaluate the doubtful terms at these stationary points. If the doubtful terms are asymptotically smaller than the kept ones, then $\tilde{I}$ provides the leading asymptotics of $I(u,\varepsilon)$. Otherwise, the doubtful terms should be kept. This recipe can be illustrated using the following example.

Example 13 Using the variational-asymptotic method to investigate the stationary points of the function in Eq. (90) with ε as a small parameter.

First, the stationary point of $f_0(u) = f(u,0)$ is $u = 0$. Then we set $u = 0 + \breve{u}_\varepsilon$, we obtain the function in Eq. (91). There is no clear indication that any one of the terms is small compared to others.

Let us drop the last term first. We obtain the function

$$\tilde{f} = \breve{u}_\varepsilon^3 + \varepsilon\breve{u}_\varepsilon^2$$

the stationary points of which are $\breve{u}_\varepsilon = 0$ and $\breve{u}_\varepsilon = -\frac{2}{3}\varepsilon$. At $\breve{u}_\varepsilon = 0$, it is not clear whether or not the last term can be disregarded. While at $\breve{u}_\varepsilon = -\frac{2}{3}\varepsilon$, the last term is negligible because the first two terms are of the order of ε^3 and the last term is of the order of ε^4.

Let us now discard the second term, we obtain the function

$$\tilde{f} = \breve{u}_\varepsilon^3 + \varepsilon^3 \breve{u}_\varepsilon$$

which does not have any stationary points if $\varepsilon > 0$, while for $\varepsilon \leq 0$, the stationary points are $\pm\sqrt{-\frac{1}{3}\varepsilon^3}$. The first and third terms are on the order of $\varepsilon^{\frac{9}{2}}$, which are small compared to the second one which is of the order of ε^4. Therefore, the second term cannot be discarded.

Finally, let us try to drop the first term. We obtain the function

$$\tilde{f} = \varepsilon \breve{u}_\varepsilon^2 + \varepsilon^3 \breve{u}_\varepsilon$$

which has one stationary point at $\breve{u}_\varepsilon = -\frac{1}{2}\varepsilon^2$. the first term is of the order of ε^6 and is much smaller than the two kept ones which are of the order of ε^5. Therefore, discarding the first term makes sense.

Therefore, in the neighborhood of zero we found two stationary points, $\breve{u}_\varepsilon = -\frac{2}{3}\varepsilon + o(\varepsilon)$ and $\breve{u}_\varepsilon = -\frac{1}{2}\varepsilon^2 + o(\varepsilon^2)$. Since the function in Eq. (90) has only two real stationary points, we have find the first-approximations of these two stationary points.

Fortunately, we can obtain the exact stationary points for the original function which are

$$u = \frac{1}{3}(-\varepsilon \pm \sqrt{\varepsilon^2 - 3\varepsilon^3})$$

Expanding these two stationary points asymptotically in terms of ε, we have

$$u = \begin{cases} -\dfrac{2\varepsilon}{3} + \dfrac{\varepsilon^2}{2} + o(\varepsilon^2) \\ -\dfrac{\varepsilon^2}{2} + o(\varepsilon^2) \end{cases}$$

Clearly, using VAM, we have correctly reproduced the first terms of both asymptotic series of the exact stationary points.

8.2.2 Case 2

In certain problems, we could meet situations that the functional $I_0(u)$ may have non-isolated stationary points. A typical finite-dimensional case can be shown for the function of two variables of the form

$$f(x, y, \varepsilon) = f_0(x) + \varepsilon g(x, y)$$

where x and y are defined in the real space, *i.e.*, $x \in (-\infty, \infty)$ and $y \in (-\infty, \infty)$. Let us assume that the function $f_0(x)$ has some stationary points $\breve{x}_0$. If the presumably small term $\varepsilon g(x, y)$ is dropped, we get the function $f_0(x)$ which has in the (x, y)-plane continuum of stationary points $(\breve{x}_0, y)$ for $y \in (-\infty, \infty)$. Clearly, the stationary points form a continuous line along the y direction.

Before formulating general rules for dealing with such cases, let us first consider two examples.

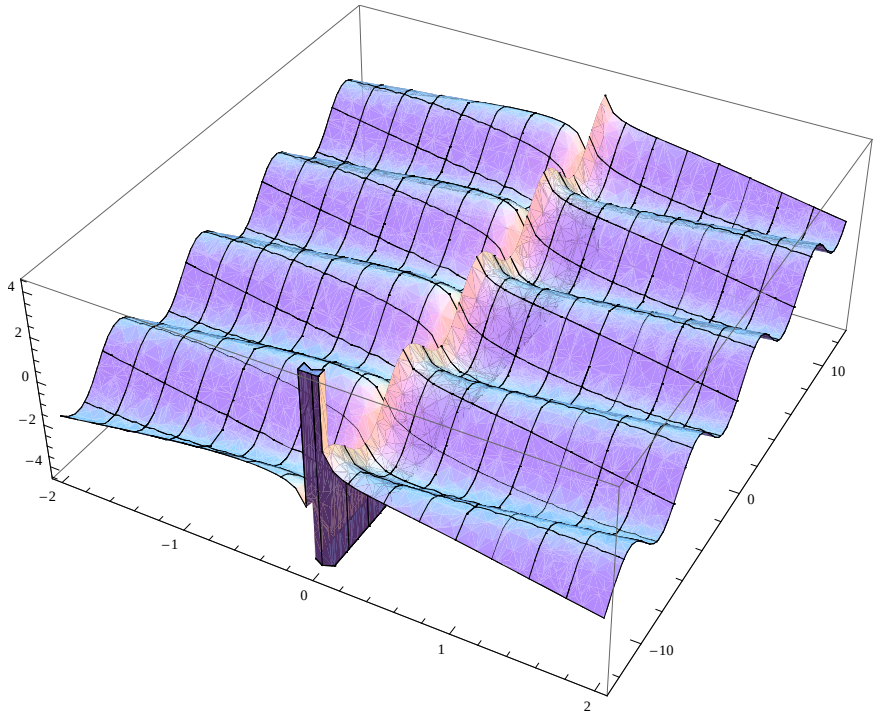

Figure 10: Sketch of the function $f(x, y, \varepsilon) = \cos(x - y) + \varepsilon \left(\frac{1}{x} + y\right)$ with $\varepsilon = 0.1$.

Example 14. Let us find the asymptotic expansion of the stationary points of the function.

$$f(x, y, \varepsilon) = \cos(x - y) + \varepsilon \left(\frac{1}{x} + y\right)$$

with $x \in (-\infty, \infty)$ and $y \in (-\infty, \infty)$.

Note that this function is not defined when $x = 0$. If $\left(\frac{1}{x} + y\right)$ is bounded for $\varepsilon \to 0$, then the second term is negligible and $f_0(x, y) = \cos(x - y)$. The set of stationary points of the function f_0 is $x = y + k\pi$ with $x \in (-\infty, \infty)$ and $y \in (-\infty, \infty)$ and $k = 0, \pm 1, \pm 2, \ldots$. Let us denote this set by M_0. The function of $f(x, y, \varepsilon)$ on M_0 becomes a function of one variable y (and a number k) such that

$$f(y, \varepsilon) = \cos(k\pi) + \varepsilon \left(\frac{1}{y + k\pi} + y\right)$$

The stationary points of f on M_0 with respect to y can be obtained through $\frac{\partial f(y,\varepsilon)}{\partial y} = \varepsilon \left(1 - \frac{1}{(k\pi + y)^2}\right) = 0$ as $y = \pm 1 - k\pi$. It should be expected that the stationary points of function $f(x, y, \varepsilon)$ converge to $(\pm 1, \pm 1 - \pi k)$ for $\varepsilon \to 0$. The validity of this proposition can be checked directly. Let us perform a change of variable $y \to z : y = x + z$, and the function f will become a sum of functions which depends only on x and z, such that

$$f(x, z, \varepsilon) = \cos z + \varepsilon z + \varepsilon \left(\frac{1}{x} + x\right)$$

The stationary points of this function can be obtained by solving the following stationary conditions

$$\frac{\partial f}{\partial x} = -\frac{1}{x^2} + 1 = 0 \qquad \frac{\partial f}{\partial z} = -\sin z + \varepsilon = 0$$

which are $x = \pm 1$ and $z = \sin^{-1}(\varepsilon)$. These points clearly converge to $(\pm 1, \pm 1 - \pi k)$ in terms of x, y for $\varepsilon \to 0$. The stationary behavior of this function for $\varepsilon = 0.1$ can also be observed from Figure 10. Note the graphics along $x = 0$ is not real as it is not defined along this line.

The described procedure of finding the stationary points made without some precautions can provide wrong results. This is illustrated in the following example.

Example 15 Let us find the stationary points of the following function

$$f(x, y, \varepsilon) = x^2 - 2x + 4\varepsilon(x-1)y + \varepsilon^2 y^2 + 2\varepsilon^2 y$$

The set M_0 of the stationary points of the function $f_0(x, y) = x^2 - 2x$ is the line $(1, y)$. The function $f(x, y, \varepsilon)$ is a function of y on M_0: $f = -1 + \varepsilon^2(y^2 + 2y)$. This function has a stationary point at $y = -1$. By analogy with the previous example we could assume that the stationary point of the function $f(x, y, \varepsilon)$ tends to the point $(1, -1)$ for $\varepsilon \to 0$. This, however, is not true. Indeed, the exact solution for the stationary conditions for the function $f(x, y, \varepsilon)$ are

$$\frac{\partial f}{\partial x} = x - 1 + 2\varepsilon y = 0 \qquad \frac{\partial f}{\partial y} = 2\varepsilon(x-1) + \varepsilon y + \varepsilon = 0$$

which can be solved to find the stationary points as $\check{x}_\varepsilon = 1 - \frac{2\varepsilon}{3}$, $\check{y}_\varepsilon = \frac{1}{3}$. Why did we get the wrong answer using the variational-asymptotic method?

Let us fix some point of the set M_0 (it is defined by assigning a value to y). The search for the stationary point will be carried out by first finding the stationary point with respect to x for a fixed y (the corresponding stationary value of f becomes a function of y only), and then finding the stationary point over all y. Since we suspect that the stationary point is in the neighborhood of set M_0, it is sufficient to consider not all values of x but only those $x = 1 + x'$ for which x' are small. Substituting it into the original function, we have

$$f(1 + x', y, \varepsilon) = (1 + x')^2 - 2(1 + x') + 4\varepsilon x' y + \varepsilon^2 y^2 + 2\varepsilon^2 y = x'^2 - 1 + 4\varepsilon x' y + \varepsilon^2 y^2 + 2\varepsilon^2 y$$

Keeping only the leading terms with respect to x' (y is kept fixed) in the function $f(1 + x', y, \varepsilon)$, we obtain the function $x'^2 + 4\varepsilon x' y$. Consequently, $x' = -2\varepsilon y$. So, we have to consider the function $f(1 - 2\varepsilon y, y, \varepsilon)$, which is

$$f(1 - 2\varepsilon y, y, \varepsilon) = -1 - 3\varepsilon^2 y^2 + 2\varepsilon^2 y$$

but not $f(1, y, \varepsilon)$ on M_0. This yields $\check{y}_0 = \frac{1}{3}$, in compliance with the correct result. It follows from the above consideration that the stationary point is a saddle point - the function f is maximum with respect to y and minimum with respect to x

The difference between Example 14 and Example 15 is the following. In Example 3, let us fix y and denote the stationary point of the function $f_0(x, y)$ for fixed y by x_0

$(x_0 = y + \pi k)$. Let x' be the first correction to x_0. It is easy to check that the difference between functions $f(x_0 + x', y, \varepsilon)$ and $f(x_0, y, \varepsilon)$ are asymptotically smaller than those terms kept in $f(x_0, y, \varepsilon)$, which is clearly shown as follows

$$\begin{aligned} f(x_0 + x', y, \varepsilon) = \cos(k\pi) + \varepsilon \left(\frac{1}{y + k\pi} + y \right) \underline{-\frac{1}{2}\cos(k\pi)x'^2 + o(x'^2)} \\ + \underline{\varepsilon \left(-\frac{x'}{(y + \pi k)^2} + \frac{x'^2}{(y + \pi k)^3} + o(x'^2) \right)} \end{aligned} \tag{92}$$

The leading terms of $f(x_0 + x', y, \varepsilon)$ are the same as $f(x_0, y, \varepsilon)$ as the underlined terms are negligible in view of the fact that x' is asymptotically smaller than x_0.

However, it is not the same with Example 15. In this example, the difference between functions $f(x_0 + x', y, \varepsilon)$ and $f(x_0, y, \varepsilon)$ are not all negligible in comparison to $f(x_0, y, \varepsilon)$ as demonstrated below

$$f(x_0 + x', y, \varepsilon) = -1 + 2y\varepsilon^2 + y^2\varepsilon^2 + \underline{x'^2 + 4x'y\varepsilon} = f(x_0, y, \varepsilon) + \underline{x'^2 + 4x'y\varepsilon} \tag{93}$$

It is not clear whether the underlined terms are smaller that those not underlined terms as we don't know the relative order of x'. If x' is of the order of ε, then the first underlined term should be kept. This implies one cannot use $f(x_0, y, \varepsilon)$ to seek an approximation for y as not all the leading terms of the asymptotical expanding of the original functional are included. Instead, we should use $f(x_0 + x', y, \varepsilon)$ to solve x' first and then seek the approximation for y. Of course, we need to check whether all the leading terms are included in $f(x_0 + x', y, \varepsilon)$, which requires us to verify whether the difference between $f(x_0 + x' + x'', y, \varepsilon)$ and $f(x_0 + x', y, \varepsilon)$ is small compared to $f(x_0 + x', y, \varepsilon)$.

Now we are ready to formulate the general rules for searching the stationary points.

Denote M_0 a set of stationary points of $I_0(u)$ and any element $u \in M$ can be presented as $u = u_0 + u'$ with $u_0 \in M_0$, $u' \in M'$. Fixing u_0 and considering u' to be a small term, we keep the leading terms with respect to u' in the functional $I(u_0 + u', \varepsilon)$. We obtain a functional $I_1(u_0, u', \varepsilon)$. Let us find the stationary points of I_1 with respect to u'. which will depend on u_0.

Suppose that u' is uniquely determined by u_0, $u' = u'(u_0, \varepsilon)$. Representing u as $u = u_0 + u'(u_0, \varepsilon) + u''$, with $u_0 \in M_0$, $u'' \in M''$ and $u'' = o(u')$. Keeping the leading terms in the functional with respect to u'', we find u''. Consider the functionals $I(u_0, \varepsilon)$ and $I(u_0 + u'(u_0, \varepsilon), \varepsilon)$ on M_0. If their stationary points do not differ significantly, then we can expect that the leading approximation of the stationary points of the initial functional are the stationary points of the functional $I(u_0, \varepsilon)$ on M_0. If the stationary points of the functionals $I(u_0, \varepsilon)$ and $I(u_0 + u'(u_0, \varepsilon), \varepsilon)$ differ significantly, but the stationary points of the functionals $I(u_0 + u'(u_0, \varepsilon), \varepsilon)$ and $I(u_0 + u'(u_0, \varepsilon), u''(u_0, \varepsilon), \varepsilon)$ do not, then the leading approximation is the stationary points of the functional $I(u_0 + u'(u_0, \varepsilon), \varepsilon)$. Note the pitfall we have encountered in Example 15. If the stationary points of the functionals $I(u_0 + u'(u_0, \varepsilon), \varepsilon)$ and $I(u_0 + u'(u_0, \varepsilon), u''(u_0, \varepsilon), \varepsilon)$ differ greatly, then the next approximation should be considered. After the leading term of the asymptotic expansion is found, the next approximation can be constructed following the same procedure.

The application of this scheme is sometimes hindered by the following obstacle: u' may be not uniquely determined by u_0, and for fixed u_0, runs through some set M_1. Let us

fix any element u_0 of M_0 and represent u in the form $u = u_0 + u_1$ where u_1 is less than u_0 in the asymptotic sense. Keeping in $I(u_0 + u_1, \varepsilon)$ only the leading terms containing u_1 and the leading interaction terms between u_1 and u_0, we obtain another functional $I_1(u_0, u_1, \varepsilon)$. If this functional has a unique stationary point, then u_1 is determined by u_0. If not, the same procedure has to be iterated. Usually, after a finite number of steps, no additional variables appear and the expansion $u = u_0 + u_1 + u_2 + \ldots$ can be written as $u = v + v_1(v, \varepsilon) + v_2(v, \varepsilon)$, where v is an element of some set N, and the subsequent terms of $v_1, v_2, \ldots$ are uniquely determined by v. If the stationary points of the functional $I(v + w(v, \varepsilon), \varepsilon)$ and the functional $I(v, \varepsilon)$ on N are close, then the stationary points of $I(v, \varepsilon)$ are the leading terms of the asymptotic expansion of the stationary points of the initial functional. If the stationary points of the functional $I(v + w(v, \varepsilon), \varepsilon)$ significantly differ from the stationary points of the functional $I(v, \varepsilon)$ but not significantly differ from the stationary points of the functional $I(v + w(v, \varepsilon) + w_1(v, \varepsilon), \varepsilon)$, the leading term is given by the stationary points of the functional $I(v + w(v, \varepsilon), \varepsilon)$. Otherwise, the next approximations should be considered.

Usually, in order to estimate the order of the difference of the two stationary points, it is sufficient to compare the values of the functionals at those points: if the values of the functionals do not differ significantly, their stationary points are in close proximity to each other.

Now, let us move on to the construction of the next approximation. In physical applications, they are usually referred as refined theories. Refined theories are needed in the cases when the zeroth-order approximation is too crude, which can happen if the actual values of the small parameter ε are not sufficiently small. The method of constructing the refined theories consists of two steps. First, a functional is derived which allows one to obtain the corrections to the solution of the next order of magnitude. Second, this refined functional is extrapolated to all, even not small, values of ε.

Usually, to construct a refined theory which takes into account the corrections on the order of $\epsilon_1(\varepsilon)$, one should keep in the functional all the terms on the order of $\epsilon_1(\varepsilon)$ compared to unity. We illustrate this statement by a minimization problem in linear space M of quadratic functional $I(u, \varepsilon)$:

$$I(u, \varepsilon) = \mathcal{E}(u, u, \varepsilon) - L(u, \varepsilon) \tag{94}$$

Here $\mathcal{E}(u, v, \varepsilon)$ is a symmetric bilinear functional of u, v and $L(u, \varepsilon)$ a linear functional. It is assumed that these two functionals can be written in the following series with respect to the small parameter ε.

$$\begin{aligned} \mathcal{E}(u, v, \varepsilon) &= \mathcal{E}_0(u, v) + \epsilon_1(\varepsilon)\mathcal{E}_1(u, v) + \epsilon_2(\varepsilon)\mathcal{E}_2(u, v) + \ldots \\ L(u, \varepsilon) &= L_0(u, \varepsilon) + \epsilon_1(h)L_1(u, \varepsilon) + \epsilon_2(h)L_2(u, \varepsilon) + \ldots \end{aligned} \tag{95}$$

where $\epsilon_2(\varepsilon)$ is smaller in the asymptotic sense than $\epsilon_1(\varepsilon)$ which can be denoted as $\epsilon_2 = o(\epsilon_1)$. $\mathcal{E}_0(u, u)$ is a positive definite functional that is

$$\mathcal{E}_0(u, u) > 0 \quad \text{if} \quad u \neq 0 \tag{96}$$

The first approximation u_0 to the minimum point of $I(u, \varepsilon)$ is the minimum point of $I_0 = \mathcal{E}_0(u, u) - L_0(u)$. The Euler-Lagrange equation for u_0 is

$$2\mathcal{E}_0(u_0, \overline{u}) - L_0(\overline{u}) = 0 \tag{97}$$

where $\overline{u}$ is an arbitrary element of M.

Let us represent u in the form $u = u_0 + u^*$. Keeping in $I(u, a)$ the leading terms containing u^* and the leading interaction terms containing u^* and u_0 and using the equation

$$2\mathcal{E}_0(u_0, u^*) - L_0(u^*) = 0 \tag{98}$$

which follows from Eq. (97), we have the following functional to solve for the minimum point u^*:

$$I^*(u_0, u^*, a) = \mathcal{E}_0(u^*, u^*) + \epsilon_1(\varepsilon)\mathcal{E}_1(u_0, u^*) - \epsilon_1(\varepsilon)L_1(u^*) \tag{99}$$

After substitution of $u^* = \epsilon_1(\varepsilon)w$, we obtain the following functional

$$I_1(u_0, w, a) = \mathcal{E}_0(w, w) + \mathcal{E}_1(u_0, w) - L_1(w) \tag{100}$$

which does not depend on the small parameter ε. Therefore $u^* \sim \epsilon_1(\varepsilon)$, and for constructing a solution with accuracy up to the order of $\epsilon_1(\varepsilon)$, all terms of the order of $\epsilon_1(\varepsilon)$ of the functional should be kept. Carefully observing Eq. (99), one can conclude that if the goal is to obtain a functional that is asymptotically correct up to the order of $\epsilon_1(a)$, it is unnecessary to solve for u^* because the leading terms associated with u^* are of the order of $\epsilon_1(a)^2$.

8.2.3 Case 3 & 4

So far, in all cases the functional $I_0(u)$ had the stationary points. It remains to discuss the cases when the functional $I_0(u)$ does not have stationary points or it is meaningless. Such cases are not exotic. For example, the function of one variable

$$f(u, \varepsilon) = u + \varepsilon u^2 + \sin \varepsilon u \tag{101}$$

is such a case, because $f_0(u)$ does not have stationary points. Another example is the minimization problem for the functional

$$I(u, \varepsilon) = \int_{\varepsilon}^{\infty} \frac{1}{r}(u^2 + u_{,r}^2)\mathrm{d}r \tag{102}$$

on the set of functions $u(r)$ which are equal to 1 for $r = \varepsilon$. This variational problem arises in the modeling of incompressible flow along an axis-symmetric thin body. The functional is meaningless for $\varepsilon = 0$. The boundary layer problems do usually also belong to this case.

In general, there are no recipes for these cases, except, probably, for the following one: an attempt should be made to reduce these problems to the problems of the type considered earlier by means of changing the required functions or some other transformations. For example, for the function in Eq. (101), the substitution: $\varepsilon u = v, \varepsilon f = g$, yields the function

$$g(v, \varepsilon) = v + v^2 + \varepsilon \sin v$$

which may be studied by the variational-asymptotic method which has been discussed.

The fields to which one can apply the VAM is still growing. However, there is no doubt about applying this method to structural mechanics. In any case, the result obtained by the VAM converges asymptotically to the exact solutions.

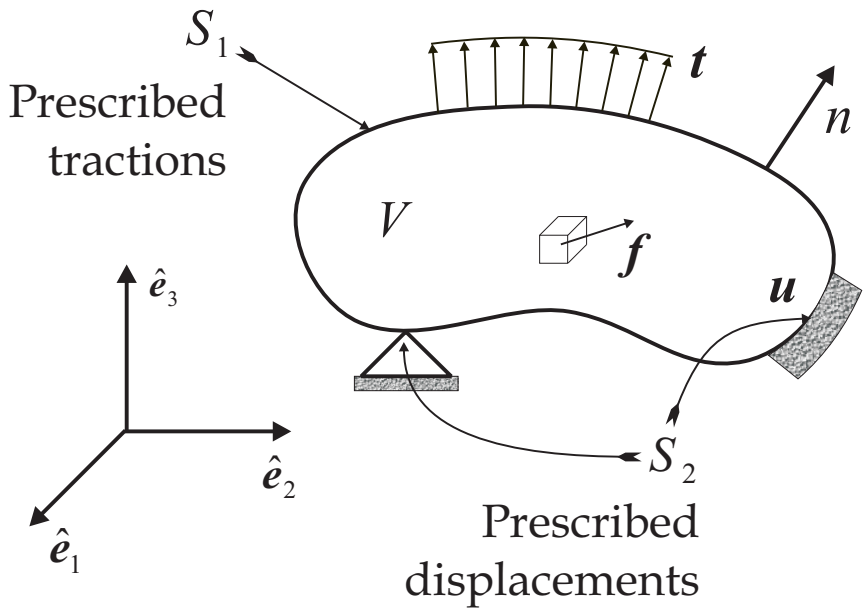

Figure 11: General elasticity problem.

9 The differential statement of elasticity problems

Consider a general elasticity problem consisting of an elastic body of arbitrary shape subjected to surface tractions and body forces as well as geometric boundary conditions such as prescribed displacements at a point or over a portion of its outer surface, as depicted in Figure 11. The volume of the body is denoted $\mathcal{V}$ and its outer surface $\mathcal{S}$. The outward normal to S is the unit vector $\hat{\boldsymbol{n}}$. $\mathcal{S}_1$ and $\mathcal{S}_2$ denote the portions of the outer surface where prescribed tractions $\boldsymbol{t}$ and prescribed displacements $\boldsymbol{u}$ are applied, respectively. At a point of the outer surface, either tractions or displacements can be prescribed, but it is impossible to prescribe both together. However it is possible to prescribe a relationship between the displacements and the tractions applied along the boundary, for example, a spring type support. This type of boundary conditions can in general be considered as a traction boundary condition. Consequently, we consider $\mathcal{S}_1$ and $\mathcal{S}_2$ share no common points and $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$. Note that a point of the outer surface that is traction free belongs to $\mathcal{S}_1$, since zero tractions $\boldsymbol{t} = 0$ are prescribed at that point. Body forces $\boldsymbol{f}$ might also be applied over the entire volume of the body. Gravity forces are a typical example of body forces, but such forces can also arise as a result of electric or magnetic fields.

The equations we have just derived in continuum mechanics can be used to form a differential statement for the linear elasticity theory which can be solved to find the displacement, strain, and stress fields at all points in $\mathcal{V}$. For linear elasticity theory, we implicitly assume that the deformation gradients are small and the strains are infinitesimal. These equations can be broken into three groups:

1. The *equations of equilibrium*,

$$\sigma_{ji,j} + f_i = 0 \tag{103}$$

are the most fundamental equations. They were derived from the principle of conservation of linear momentum stating that the time rate change of linear momentum of a body is equal to the summation of forces acting on that body.

2. The *strain-displacement equations*,

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{104}$$

merely define the strain components that are used for the characterization of the deformation of the body at a point. The strain-displacement relationships were derived from purely geometric considerations. We have also shown that the strains cannot be given arbitrarily, and they have to satisfy the compatibility equations,

$$\varepsilon_{ij,kl} + \varepsilon_{kl,ij} - \varepsilon_{jl,ik} - \varepsilon_{ik,jl} = 0 \tag{105}$$

3. The *constitutive equations* relate the stress and strain components. They consist of a mathematical idealization of the experimentally observed behavior of materials. The isotropic, linear elastic material behavior described in

$$\sigma_{ij} = \lambda\varepsilon_{\alpha\alpha}\delta_{ij} + 2\mu\varepsilon_{ij} \tag{106}$$

is a commonly used, but highly idealized constitutive law. Many materials can present one or more of the following features: anisotropy, plasticity, viscoelasticity, or creep, to name just a few commonly observed material behaviors.

Of course, these equations must be supplemented with sufficient boundary conditions and initial conditions which are used to describe the interaction between the body in question and the outside environment, such as the stress boundary conditions,

$$\bar{t}_j = n_i\sigma_{ij} \tag{107}$$

on $\mathcal{S}_1$ and prescribed displacement on $\mathcal{S}_2$.

In summary, the complete solution of a problem of linear elasticity involves the following three key elements:

1. A kinetically admissible stress field, which is defined as any stress field satisfying the equations of equilibrium, Eqs. (103), at all points in $\mathcal{V}$ and the stress boundary conditions, Eqs. (107), at all points on $\mathcal{S}_1$.

2. A kinematically admissible displacement field and the corresponding compatible strain field. A displacement field $\boldsymbol{u}(x_1, x_2, x_3)$ is said to be **kinematically admissible** if it is continuous and differentiable at all points in $\mathcal{V}$ and satisfies the prescribed geometry along the boundary at all points on $\mathcal{S}_2$. A strain field $\boldsymbol{\varepsilon}(x_1, x_2, x_3)$ is said to be **compatible** if it is derived from a kinematically admissible displacement field through the strain-displacement relationships, Eq. (104), or if the strain field satisfies the compatibility equations, Eq. (105).

3. A constitutive model satisfied by the stress field and the strain field at all points in $\mathcal{V}$.

It can be proved that the solution to any linear elasticity problem, which have the aforementioned features, is unique.

As far as the actual formulation is concerned, not all the equations have to enter the formulation at the same time. It is clear that we have 15 unknowns including $\sigma_{ij}, \varepsilon_{ij}$, and u_i for any linear elasticity problem, while there are a total of 21 equations including 3 equations of equilibrium in Eq. (103), 6 strain-displacement relations in Eq. (104), 6 compatibility equations in Eq. (105), and 6 equations in the constitutive model in Eq. (106). In fact, the strain-displacement relationships in Eq. (104) cannot enter the formulation with the compatibility equations at the same time because both conditions implying a compatible strain field.

One can use the 15 equations in Eq. (103), Eq. (104), and Eq. (106) to solve for $\sigma_{ij}, \varepsilon_{ij}$, and u_i. To achieve this, one can develop so-called a displacement formulation to choose the displacements u_i as the main unknowns and eliminate ε_{ij} using Eq. (104), and eliminate σ_{ij} using Eq. (106) to reformulate the three equations of equilibrium, Eq. (103), in terms of displacements u_i. After the displacement field has been obtained, one can calculate the strain field from Eq. (104), and then the stress field from Eq. (106).

One can also use the 15 equations in Eq. (103), Eq. (106), and Eq. (105) to solve for σ_{ij} and ε_{ij}. Recall that the 6 equations in Eq. (105) in fact only represent three mathematically independent equations. Then use Eq. (104) to back calculate the displacement field, u_i.

10 Variational principles in mechanics

It is clear that we have to manipulate a group of equations systematically and purposely to obtain a solution to satisfy the differential statement of Elasticity. Only a few exact solutions exist for simple Elasticity problems. It is very difficult, if not impossible, to obtain the solution for complex Elasticity problems. In this section, we will present an alternate formalism, variational principles, for the solution of Elasticity problems. Variational principles is the basis for powerful numerical techniques, such as the finite element method, that are routinely used to obtain approximate solutions to complex Elasticity problems. Although there are numerous variational principles in mechanics, we will focus on two commonly used principles in Solid Mechanics including the Principle of Virtual Work (PVW) and the Principle of Minimum Total Potential Energy (PMTPE).

10.1 The principle of virtual work

Consider an elastic body that is in equilibrium under applied body forces and surface tractions, as depicted in Figure 11. This implies that the stress field is statically admissible, *i.e.*, the equilibrium equations, Eq. (103), are satisfied at all points in $\mathcal{V}$ and the stress boundary conditions, Eqs. (107), at all points on $\mathcal{S}_1$. The following statement is now constructed

$$-\int_{\mathcal{V}} (\sigma_{ji,j} + f_i)\delta u_i \mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} (n_j\sigma_{ji} - t_i)\delta u_i \mathrm{d}\mathcal{S} = 0. \tag{108}$$

This statement was constructed in the following manner. Each of the three equilibrium equations was multiplied by an arbitrary, virtual change in displacement, then integrated over the range of validity of the equation, $\mathcal{V}$. Similarly, each of the three stress boundary conditions was multiplied by an arbitrary, virtual change in displacement, then integrated over the range of validity of the equation, $\mathcal{S}_1$. Since the stress field is statically admissible, each term in parenthesis is zero, and multiplication by an arbitrary quantity results in a zero product. Each of the two integrals then vanishes, as does their sum. The minus sign in front the first integral is added due to traditional reasons.

Next, integration by parts is performed. For the first term of the volume integral this operation is

$$-\int_{\mathcal{V}} \sigma_{ji,j}\delta u_i \mathrm{d}\mathcal{V} = \int_{\mathcal{V}} \sigma_{ji}\delta u_{i,j}\mathrm{d}\mathcal{V} - \int_{\mathcal{S}} n_j\sigma_{ji}\ \delta u_i\ \mathrm{d}\mathcal{S}, \tag{109}$$

If we choose the virtual displacements to be kinematically admissible, which implies $\delta u_i = 0$ on $\mathcal{S}_2$. Expression (108) now simplifies to

$$\int_{\mathcal{V}} \sigma_{ji}\delta u_{i,j}\mathrm{d}\mathcal{V} = \int_{\mathcal{V}} f_i\delta u_i\mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} t_i\delta u_i\mathrm{d}\mathcal{S}. \tag{110}$$

If we also define a *virtual, compatible strain field* as

$$\delta\varepsilon_{ij} = \frac{1}{2}(\delta u_{i,j} + \delta u_{j,i}) \tag{111}$$

Then the expression (110) can be further rewritten as

$$\int_{\mathcal{V}} \sigma_{ij}\delta\varepsilon_{ij}\mathrm{d}\mathcal{V} = \int_{\mathcal{V}} f_i\delta u_i\mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} t_i\delta u_i\mathrm{d}\mathcal{S}. \tag{112}$$

where symmetry of the stress tensor has been used. The term on the left hand side of this expression can be interpreted as the virtual work done by the internal stresses whereas that on the right corresponds to the work done by the externally applied loads. Eq. (112) can thus be interpreted as follows

> **Principle of Virtual Work** *A body is in equilibrium if the virtual work done by the internal stresses equals the virtual work done by the externally applied loads, for all arbitrary kinematically admissible virtual displacement fields and corresponding compatible strain fields.*

It has been shown thus far that if a stress field is statically admissible, the principle of virtual work must hold. It can also be shown that if the principle of virtual work holds true, then the stress field must be statically admissible. Indeed, this principle implies Eq. (110), which in turns implies Eq. (108) by reversing the integration by parts process. Finally, the equations of equilibrium and stress boundary conditions are recovered because Eq. (108) must hold for all arbitrary, kinematically admissible virtual displacement fields. In summary, the conditions of equilibrium, Eqs. (103) and (107), and the principle of virtual work are two entirely equivalent statements. Because the principle of virtual work is solely a statement of equilibrium, it is always true, no matter whether the material is linear or nonlinear, elastic or not elastic. However, for the solution of specific Elasticity

problems, it must be complemented with an appropriate constitutive model. For example, for Linear Elasticity problems, we need to use the generalized Hooke's law to relate the stress field and the strain field.

For the convenience of computation, we can adopt the engineering notation to rewrite the principle of virtual work, Eq. (112), in a matrix form as

$$\int_{\mathcal{V}} \sigma^T \delta\varepsilon \mathrm{d}\mathcal{V} = \int_{\mathcal{V}} f^T \delta u \mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} t^T \delta u \mathrm{d}\mathcal{S}. \tag{113}$$

with $\sigma = \lfloor \sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{23}\ \sigma_{13}\ \sigma_{12} \rfloor^T$, $\varepsilon = \lfloor \varepsilon_{11}\ \varepsilon_{22}\ \varepsilon_{33}\ 2\varepsilon_{23}\ 2\varepsilon_{13}\ 2\varepsilon_{12} \rfloor^T$, $f = \lfloor f_1\ f_2\ f_3 \rfloor^T$, $t = \lfloor t_1\ t_2\ t_3 \rfloor^T$, and $u = \lfloor u_1\ u_2\ u_3 \rfloor^T$.

10.2 The principle of minimum total potential energy

The principle of virtual work should be augmented with an appropriate constitutive model to characterize the material. For hyperelastic materials, there exists a scalar state function, *strain energy density function*, such that

$$\sigma_{ij} = \frac{\partial W}{\partial \varepsilon_{ij}} \tag{114}$$

Introducing this constitutive model into the left hand side of the principle of virtual work, Eq. (112), leads to

$$\int_{\mathcal{V}} \sigma_{ij} \delta\varepsilon_{ij} \mathrm{d}\mathcal{V} = \int_{\mathcal{V}} \frac{\partial W}{\partial \varepsilon_{ij}} \delta\varepsilon_{ij} \mathrm{d}\mathcal{V} = \int_{\mathcal{V}} \delta W \mathrm{d}\mathcal{V} = \delta \int_{\mathcal{V}} W \mathrm{d}\mathcal{V} = \delta A(\epsilon_{ij}) = \delta A(u_i) \tag{115}$$

where the chain rule for derivatives was used. The strain energy density and the *total strain energy* of the body, $A = \int_{\mathcal{V}} W \ \mathrm{d}\mathcal{V}$, can be expressed in terms of the displacement field u using the strain displacement relationships because the principle of virtual work requires a compatible strain field. The principle of virtual work now writes

$$\delta A(u) = \int_{\mathcal{V}} f_i \delta u_i \mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} t_i \delta u_i \mathrm{d}\mathcal{S}. \tag{116}$$

Next, if the body forces and surface tractions can be derived from potential functions

$$f_i = -\frac{\partial \phi}{\partial u_i}; \quad t_i = -\frac{\partial \psi}{\partial u_i}, \tag{117}$$

where ϕ is the *potential of the body forces*, and ψ the *potential of the surface tractions*. For instance, the potential of fixed surface tractions is simply $\psi = -t_i u_i$, whereas the potential of the body forces associated with a gravitational acceleration field g_i is $\phi = -m g_i u_i$, where m is the mass per unit volume of the body. It is important to note that potential functions do not exist for all types of forces. For example, potential functions do not always exist for forces that depend on the orientation of the body, such as follower forces or aerodynamic forces, for instance.

The right hand side Eq. (116) now becomes

$$\begin{aligned}\int_{\mathcal{V}} f_i \delta u_i \mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} t_i \delta u_i \mathrm{d}\mathcal{S} &= -\int_{\mathcal{V}} \frac{\partial \phi}{\partial u_i} \delta u_i \mathrm{d}\mathcal{V} - \int_{\mathcal{S}_1} \frac{\partial \psi}{\partial u_i} \delta u_i \mathrm{d}\mathcal{S} = -\int_{\mathcal{V}} \delta\phi(u) \mathrm{d}\mathcal{V} - \int_{\mathcal{S}_1} \delta\psi(u) \mathrm{d}\mathcal{S} \\ &= -\delta \int_{\mathcal{V}} \phi(u) \mathrm{d}\mathcal{V} - \delta \int_{\mathcal{S}_1} \psi(u) \mathrm{d}\mathcal{S} = -\delta\Phi(u)\end{aligned} \tag{118}$$

where $\Phi(u) = \int_{\mathcal{V}} \phi(u) \mathrm{d}\mathcal{V} + \int_{\mathcal{S}_1} \psi(u) \mathrm{d}\mathcal{S}$ is the total potential of the applied loads. Introducing this result into Eq. (116) leads to

$$\delta A(u) = -\delta\Phi(u), \text{ or } \quad \delta\left(A(u) + \Phi(u)\right) = 0. \tag{119}$$

The *total potential energy* of the body is now defined as

$$\Pi(u) = A(u) + \Phi(u), \tag{120}$$

and it follows that

$$\delta\Pi(u) = 0. \tag{121}$$

This statement can be interpreted as follows

> **Principle of Minimum Total Potential Energy** *Among all kinematically admissible displacements fields, the actual displacement field that corresponds to the equilibrium configuration of the body makes* Π *an absolute minimum.*

Eq. (121) only proves that at equilibrium, the total potential energy presents a stationary point. As discussed previously, the sign of the second variation $\delta^2\Pi$ will determine whether the stationary point actually is a minimum. The first variation in Π is

$$\delta\Pi(u) = \int_{\mathcal{V}} \frac{\partial W}{\partial \varepsilon_{ij}} \delta\varepsilon_{ij} \mathrm{d}\mathcal{V} - \int_{\mathcal{V}} b_i \delta u_i \mathrm{d}\mathcal{V} - \int_{\mathcal{S}_1} t_i \delta u_i \mathrm{d}\mathcal{S}, \tag{122}$$

and its second variation is then

$$\delta^2\Pi(u) = \int_{\mathcal{V}} \frac{\partial^2 W}{\partial \varepsilon_{ij} \partial \varepsilon_{kl}} \, \delta\varepsilon_{ij} \, \delta\varepsilon_{kl} \mathrm{d}\mathcal{V}. \tag{123}$$

Based on physical grounds, the strain energy density function must be a positive-definite function of the strain components, which implies $\frac{\partial^2 W}{\partial \varepsilon_{ij} \partial \varepsilon_{kl}} \, \delta\varepsilon_{ij} \, \delta\varepsilon_{kl} \geq 0$ for all $\delta\varepsilon_{ij}$. Indeed, if the strain energy function were not positive-definite, strain states would exist that generate a negative strain energy, *i.e.*, the elastic body would generate energy under deformation, a situation that is physically impossible. It follows that $\delta^2\Pi \geq 0$, and hence, Π presents an absolute minimum at its stationary point.

If the principle of minimum total potential energy holds true, the total potential energy must present a stationary point, implying Eq. (119). In turn, this equation implies the principle of virtual work, in which the stresses were expressed in terms of the strains using the constitutive model of the form of Eq. (114), and strains were themselves expressed in

terms of displacements using the strain-displacement relationships. In summary, the principle of minimum total potential energy implies the equations of equilibrium of the problem expressed in terms of the displacement field. Such equations are the Euler-Lagrange equations corresponding to the stationarity condition for the total potential energy. The principle of minimum total potential energy implies the principle of virtual work, but the principle of virtual work only implies the principle of minimum total potential energy under restrictive assumptions on existence of a strain energy density function and of potentials of the body forces and surface tractions.

11 Problems

1. For a functional $I[y] = \int_0^1 (y'^2 + 12xy)\ \mathrm{d}x$ along with $y(0) = 0, y(1) = 0$, find the function y which corresponds to the stationary value of I. Is the stationary value a minimum, maximum, or saddle point? Choose at least two other functions of y which satisfies the boundary conditions to evaluate the functional to partially verify your conclusion.

2. Find the number n_i of atoms in energy level E_i that minimizes the free energy $G = E - TS$ with $E = n_1E_1 + n_2E_2 + \ldots$ as the internal energy, $S = k \log \frac{n!}{n_1!n_2!n_3!\ldots}$ as the entropy, $n = n_1 + n_2 + \ldots$ as the total number of atoms. E_i is the energy level of n_i number of atoms. T is the absolute temperature, and k is the Boltzman constant. T, E_i, k, n are given constants, n and n_i are large enough to assume $\log n_i! = n_i \log(n_i - 1)$.

3. Find the shortest distance from the origin of coordinates to the hyperbola, $x^2 + 8xy + 7y^2 = 225$.

4. The ground state of energy of an electron with mass m confined in a box with sides a, b, and c is expressed as

$$E = \frac{h^2}{8m}(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}).$$

   When the volume of the box is kept constant as V_0, find the sides of the box to give a minimum value of E.

5. Using two approaches to find the stationary points of function $f(x, y) = 2x^2 - 2xy + y^2 + 18x + 6y$ subject to the constraint: $x = y$.

6. Using two approaches to find the maximum area of a rectangular with a given perimeter L.

7. Find the minimum surface of revolution of the curve $y = y(x)$ rotating about the x axis, where $y(0) = 1$ and $y(1) = 1$ at the ends of the curve. Plot this curve. (Hint: if you have difficulty to solve the constants analytically, you can solve it numerically.)

8. On surface $2x^2 - y^2 - z^2 = 0$, provide a variational statement for the equations of the smooth curve having minimum length connecting point $A(0, 0, 0)$ and point $B(1, 1, 1)$.

9. Sketch the stationary curve of the functional $I[y] = \int_{-1}^{1} \sqrt{y(1 + y'^2)}\, \mathrm{d}x$ with boundary conditions $y(-1) = 1$ and $y(1) = 1$.

10. Among all the curves of given length L connecting the two points $(x_0, 0)$ and $(x_1, 0)$ on the x axis, find the one which maximizes the area enclosed by the curve and the x axis.

11. The differential statement of one-dimensional heat transfer with ϕ as the unknown function is

$$\frac{d^2\phi}{dx^2} + Q(x) = 0 \quad \text{for} \quad (0 \le x \le L) \tag{124}$$

along with the boundary conditions $\phi = \phi_0$ at $x = 0$ and $\phi = \phi_0$ at $x = L$. What is the corresponding variational statement of the same problem? If

$$Q(x) = \begin{cases} \phi_0/L^2 & (0 \le x \le L/2) \\ 0.1\phi_0/L^2 & (L/2 < x \le L) \end{cases} \tag{125}$$

then:

   (a) Solve the differential statement of this problem exactly;

   (b) Use the Ritz method to solve this problem approximately by using trial functions as $\phi = \sum_{i=1}^{n} a_i x^{i-1}$. Obtain the approximate solutions for n=2, 3, 4, and 5, respectively;

   (c) Normalize the solution using $\eta = x/L$ and $\bar{\phi} = \frac{\phi}{\phi_0}$. Plot the approximate solution and exact solution for $\bar{\phi}$ with respect to η in the same figure.

   (d) Use the Galerkin method to solve this problem approximately by using trial functions as $\phi = \sum_{i=1}^{n} a_i x^{i-1}$. Obtain the approximate solutions for n=2, 3, 4, and 5, respectively. And verify that the solutions are the same as obtained using the Ritz method.

12. Minimize the functional defined in $0 \le x \le a$ and $0 \le y \le b$,

$$I[\psi] = \int\int_R [(\frac{\partial\psi}{\partial x})^2 + (\frac{\partial\psi}{\partial y})^2 - 2g\psi]\, \mathrm{d}x\, \mathrm{d}y$$

with ψ vanishes along the boundaries at $x = 0$, $x = a$, $y = 0$, and $y = b$, and g is given by

$$g = \sum_{m,n=1}^{\infty} C_{mn} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b}.$$

13. Use the Ritz method to solve the stationary problem of $I[y] = \int_0^1 [y'^2 + y^2 + 2xy]\, \mathrm{d}x$ with $y = 0$ at $x = 0$ and 1. Compare your approximate solution with the exact solution and verify that your approximate solution will converge to the exact solution as more terms are included in the trial function used for the Ritz method.

14. Find the Euler-Lagrange equation and the (natural) boundary conditions to minimize the functional

$$I[y] = \int_0^l (y'^2 + y^2 + 2xy)\ \mathrm{d}x + l[y(0) - a] + l[y(l) - b],$$

where l, a, and b are given constants.

15. Find the natural conditions to minimize the following functional, or so-called Lagrangian,

$$I[x, y, z] = \int_{t_0}^{t_l} [\frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \frac{1}{2}k(x^2 + y^2 + z^2)\ \mathrm{d}t,$$

where m is the mass of the particle and k is the spring constant.

16. Find an approximate solution to minimize the functional

$$I[\psi] = \int\int_R [(\frac{\partial \psi}{\partial x})^2 + (\frac{\partial \psi}{\partial y})^2]\ \mathrm{d}x\ \mathrm{d}y$$

with the subsidiary condition

$$H[\psi] = \int\int_R \psi^2\ \mathrm{d}x\ \mathrm{d}y = 1$$

and $\psi = 0$ at the boundary of R, where R is a rectangular domain $0 \le x \le a$ and $0 \le x \le b$. (Hint: one can assume $\psi = \sum_{m,n=1}^{\infty} a_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}$)

17. Show that the Euler-Lagrange equation for

$$I[u] = \int\int_G (\frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2} - (\frac{\partial^2 u}{\partial x \partial y})^2)\ \mathrm{d}x\ \mathrm{d}y$$

is identically zero. Does this functional have stationary values when boundary values of derivatives of u are prescribed?

18. Determine whether the following functionals have stationary points. If yes, under which condition? If no, why?

    (a) $I[y] = \int_{x_0}^{x_1} (y - x^3 y')\ \mathrm{d}x$.

    (b) $I[y] = \int_0^1 (2xy + x^2 y')\ \mathrm{d}x$ with $y(0) = 0, y(1) = 1$.

    (c) $I[y] = \int_0^{\frac{\pi}{2}} \left[x \sin y + \left(\frac{x^2}{2} \cos y\right) y'\right]\ \mathrm{d}x$, with $y(0) = 0, y(\frac{\pi}{2}) = \frac{\pi}{2}$.

    (d) $I[y] = \int_0^{\frac{\pi}{2}} (y^2 + y' \sin 2x)\ \mathrm{d}x$ with $y(0) = k, y(\frac{\pi}{2}) = -1$.

19. Find the curve which corresponds to the stationary value of the following functionals.

    (a) $I[y] = \int_a^b (x^2 + 16y^2 - y''^2)\ \mathrm{d}x$.

(b) $I[y] = \int_0^1 (1 + x + y''^2)\ \mathrm{d}x$, with $y(0) = 0, y(1) = 1, y'(0) = y'(1) = 1$.

(c) $I[y] = \int_0^{\frac{\pi}{2}} (y''^2 - 2y'^2 + y^2 - x^2)\ \mathrm{d}x$, with $y(0) = y'(0) = 0, y(\frac{\pi}{2}) = 1, y'(\frac{\pi}{2}) = \frac{\pi}{2}$.

(d) $I[y, z] = \int_{x_0}^{x_1} (2yz - 2y^2 + y'^2 - z'^2)\ \mathrm{d}x$.

(e) $I[y, z] = \int_0^1 (y'z' + y'^2 + z'^2)\ \mathrm{d}x$ with $y(0) = z(0) = 0, y(1) = z(1) = 1$.

20. The potential energy of a circular plate with radius R under axisymmetric distributed load $q(r)$, with $r \in [0, R]$, can be express in terms of its deflection $w(r)$ as

$$I[w] = D\pi \int_0^R (rw''^2 + \frac{w'^2}{r} + 2\mu w'w'' - \frac{2q}{D} rw)\ \mathrm{d}r$$

where D, μ are elastic constants. We know that the equilibrium of the plate corresponds to the minimum value of the potential energy. Show that $w(r)$ must satisfy the following equation of equilibrium.

$$rw'''' + 2w''' - \frac{w''}{r} + \frac{w'}{r^2} = \frac{qr}{D}.$$

21. Find the Euler-Lagrange equation for the following functional

$$J[u(x, y, z)] = \int_G \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 + 2uf(x, y, z) \right]\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z.$$

with $f(x, y, z)$ as a given known function.

22. Express the functional $J[u(x, y)] = \int_D f(x, y, u, u_{,x}, u_{,y})\ \mathrm{d}x\ \mathrm{d}y$ in a polar coordinate system and find the corresponding Euler-Lagrange equation in the polar coordinate system. Show that the same Euler-Lagrange equation can be obtained through a direct transformation of the Euler-Lagrange equation in the Cartesian coordinate system.

23. Using the Galerkin method to find approximate solutions for the following differential equation

$$u'' - u + 1 = 0 \quad \text{for} \quad 0 \le x \le 1$$

along with the boundary condition $u'(0) = u(1) = 0$ for an approximate solution for the unknown function $u(x)$. You need to find at least two approximate solutions and you also need to compare with the exact solution.

24. The Saint Venant torsion problem is governed by the following partial differential equation

$$\frac{\partial^2 \Phi}{\partial x_2^2} + \frac{\partial^2 \Phi}{\partial x_3^2} = -2G\kappa_1.$$

Φ is the so-called Prandtl stress function, G is a constant denoting shear modulus, κ_1 is a constant denoting the twist rate. For simply connected domain (*i.e.*, no holes in the domain), $\Phi = 0$ along the boundary curve. Solve this problem approximately

using the Kantorovich method for a rectangular domain with $x_2 \in [-a, a]$ and $x_3 \in [-b, b]$. The exact solution of this problem is

$$\Phi = G\kappa_1(b^2 - x_3^2) - \frac{32G\kappa_1 b^2}{\pi^3}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^3}\frac{\cosh\lambda_n x_2}{\cosh\lambda_n a}\cos\lambda_n x_3$$

with $\lambda_n = \frac{2n+1}{2}\frac{\pi}{b}$. Compare the approximate solution with the exact solution for $\frac{\Phi}{G\kappa_1 a^2}$ by using $b = a/5$ and normalizing the length parameters with respect to a. What will be your suggestion to refine this approximation? Hint: you can use a trial function in the form of $\Phi = f(x_2)(b^2 - x_3^2)$.

25. Find the third-order approximation of the stationary points for the first example in the Section of variational asymptotic method. Plot the the exact solution along with the first-order approximation, the second-order approximation, and the third-order approximation.

26. Using the variational asymptotic method to find an approximate solution for the stationary point for function

$$f(u, \varepsilon) = u + u^2 + \varepsilon \sin u$$

asymptotically correct up to the second-order. Compare the approximate solutio with exact solution for $\varepsilon = 0.1$ and $\varepsilon = 0.2$.

27. A state of plane strain relative to the (x_1, x_2) plane has a strain energy density function given by

$$W = \frac{1}{2}b_{11}\varepsilon_{11}^2 + b_{22}\varepsilon_{22}^2 + b_{33}\varepsilon_{12}^2 + 2b_{12}\varepsilon_{11}\varepsilon_{22} + 2b_{13}\varepsilon_{11}\varepsilon_{12} + 2b_{23}\varepsilon_{22}\varepsilon_{12}$$

where the b_{ij} are elastic coefficients. Use the Principle of Minimum Total Potential Energy to derive the equations of equilibrium in terms of the displacements $u_1(x_1, x_2)$ and $u_2(x_1, x_2)$ including the effects of body forces.

28. Carry out the variation of the total potential energy to obtain the corresponding Euler-Lagrange equations and boundary conditions. Verify that these equations and boundary conditions are equivalent to the equilibrium and stress boundary conditions of the elasticity problem.