SUPPLEMENTS TO MATHEMATICAL METHODS IN SCIENCE AND ENGINEERING
http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470041420.html

CHAPTER 17: VARIATIONAL ANALYSIS

I. Solutions or Hints to Selected Problems:

1. **(Newton's bucket experiment)** A bucket half-filled with water is rotated about its axis. Find the shape of the surface of the water after the equilibrium is set.

   Solution:

   In *Principia* (1689) Newton describes a simple experiment with a bucket half-filled with water and suspended with a rope from a fixed point in space. In this experiment, first the rope is twisted tightly and then after the water has settled and its surface becomes flat, the rope is released. At first, the bucket spins rapidly with the water remaining at rest with its surface flat. Eventually, the friction between the water and the bucket communicates the motion of the bucket to the water and the water begins to rotate also. As the water begins to rotate, it also rises along the sides of the bucket. Slowly, the relative motion between the bucket and the water ceases and the surface of the water assumes a concave shape. Finally, the rope unwinds completely and begins to twist in the other direction, thus slowing and eventually stopping the bucket. Shortly after the bucket has stopped, the water continues its rotation with its surface being concave. The question is; what causes this concave shape of the surface of the water?

   At first, the bucket is spinning but the water is at rest and its surface is flat. Eventually, when there is no relative motion between the bucket and the water, the surface is concave. Finally, when the water is spinning but the bucket is at rest, the surface is still concave. From these it is clear that the relative rotation of the water and the bucket is not what determines the shape of the surface.

   The crucial question is; what is spinning and with respect to what? Let us try to understand the shape of the surface in terms of interactions. Since the bucket, the water and the rest of the universe are on the average neutral, electromagnetic forces can not be the reason. The gravitational interaction between the bucket and the water is surely

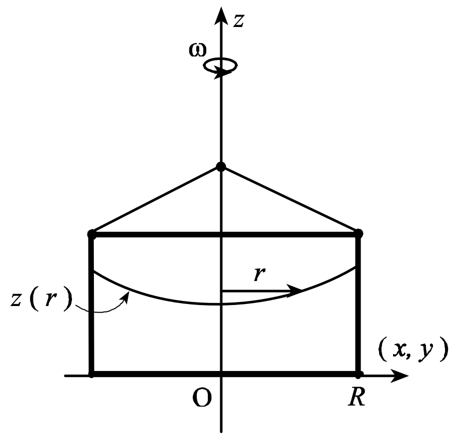

Fig. 0.1 Newton's bucket experiment.

negligible, hence it can not be the reason either. According to Newton's theory, gravity is a scalar interaction, thus the gravitational force between two masses is independent of their velocity and thus depends only on their separation. In this regard, Newton could not have used the gravitational interaction of the water with the rest of the universe also. This lead Newton reluctantly to explain the concave shape as due to rotation with respect to **absolute space**. In other words, the surface of the water is flat when the water is not rotating with respect to absolute space and when there is rotation with respect to absolute space it is concave.

A satisfactory solution comes only with the Einstein's general theory of relativity, where the gravitational force between two masses depends not just on their separation but also on their relative velocity. This is analogous to Maxwell's theory, where the electromagnetic forces are described by a vector potential, hence the force between two charged particles has a velocity dependent part aside from the usual coulomb force. However, in the case of general theory of relativity gravity is described by a tensor potential, the metric tensor, thus the gravitational force is much more complicated than it is in a vector interaction. In this context, not just the shape of the surface of the water in the Newton's bucket experiment, but all fictitious forces in Newton's theory, in principle, can be explained as the gravitational interaction of matter with the mean matter distribution in the universe.

Let us now find the equation of the concave shape that the surface of the water assumes. For simplicity, we assume a cylindrical container (Fig. 0.1) with the radius R and rotating with uniform angular velocity

ω about its axis. We determine the surface height, $z(r)$, of the water by minimizing the potential energy. For a given mass element of the water, we can write the infinitesimal potential energy as

$$dE = (\rho g z - \frac{1}{2}\rho\omega^2 r^2)dv, \tag{0.1}$$

where ρ is the uniform density of the water and g is the acceleration of gravity. We now write the functional, $I[z(r)]$, that needs to be minimized for $z(r)$ as

$$\begin{aligned} I[z(r)] &= \int\int\int_V dE & (0.2)\\ &= \int_0^{2\pi}\int_0^R\int_0^{z(r)} (\rho g z - \frac{1}{2}\rho\omega^2 r^2) r \; dz \; dr \; d\theta & (0.3)\\ &= \pi\rho\int_0^R (g z^2 - \omega^2 r^2 z) r dr. & (0.4)\end{aligned}$$

Important:

Note that the integrand in the above functional does not involve any derivatives of $z(r)$, hence the boundary conditions

$$z(0) \text{ and } z(R), \tag{0.5}$$

are not needed in the derivation of the Euler Equation (17.14), which becomes:

$$\frac{\partial\left[(g z^2 - \omega^2 r^2 z) r\right]}{dz} = 0, \tag{0.6}$$

thus yielding the surface of revolution, $z(r)$, representing the free surface of the water as

$$z(r) = \frac{\omega^2 r^2}{2g}. \tag{0.7}$$

One final remark that needs to be made is that this result is only partially true, since we have not defined the optimization problem correctly. For a proper description of the problem we have to take into account the fact that water is incompressible, that is, its volume is fixed. Now the functional in Equation (0.4) has to be extremized subject to the constraint

$$\begin{aligned} J[z(r)] &= \int\int\int_V dv & (0.8)\\ &= \int_0^{2\pi}\int_0^R\int_0^{z(r)} r \; dz \; dr \; d\theta & (0.9)\\ &= 2\pi\int_0^R z(r) dr & (0.10)\\ &= V_0, & (0.11)\end{aligned}$$

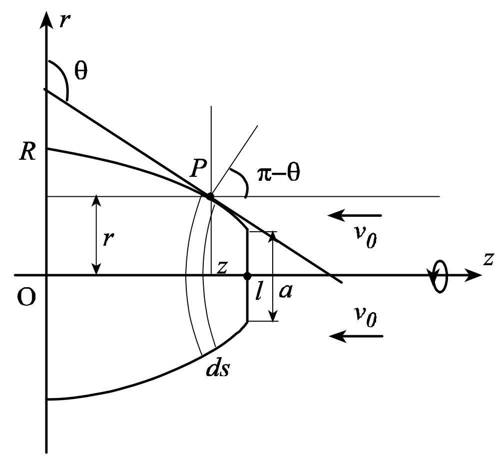

Fig. 0.2 Drag force on a surface of revolution.

thus making the problem one of isoperimetric type and can be solved by using the method discussed in Section 17.6.

2. Consider an axially symmetric object moving in a perfect incompressible fluid with constant velocity. Assuming that at any point on the surface the drag force per unit area is proportional to the normal component of the velocity, find the shape that minimizes the drag force on the object.

Solution:

Since the object is axially symmetric, we consider the surface of revolution shown in Figure 0.2, where θ is the angle that the tangent at point P makes with the plane perpendicular to the z-axis. Since the normal component of the velocity at P is

$$v_\perp = v_0 \cos(\pi - \theta) \tag{0.12}$$

$$= -v_0 \cos\theta, \tag{0.13}$$

we write the drag force on the infinitesimal strip with the area (Fig. 0.2)

$$2\pi r \; ds \tag{0.14}$$

and projected along the z-axis, as

$$\alpha \left(v_0^2 \cos^2\theta\right) \cos\theta \; 2\pi r \; ds, \tag{0.15}$$

where α is the drag coefficient. Since $ds = dr/\cos\theta$, the total drag on the body is the integral

$$J = 2\pi\alpha v_0^2 \int_0^R r \cos^2\theta \; dr. \tag{0.16}$$

Using the definition of the surface of revolution, $z(r)$, we can write

$$\frac{dz}{dr} = \tan\theta, \tag{0.17}$$

hence

$$\cos\theta = \frac{1}{[1+z'^2]^{1/2}}. \tag{0.18}$$

Now the functional to be minimized for $z(r)$ becomes

$$J[z(r)] = 2\pi\alpha v_0^2 \int_0^R \frac{r\ dr}{1+z'^2}, \tag{0.19}$$

which yields the Euler equation

$$\frac{rz'}{[1+z'^2]^2} = c_0, \tag{0.20}$$

where c_0 is an integration constant. Note that since the integrand does not depend on z explicitly, we have written the first integral [Eq. (17.14)] immediately. For the solution we call $z' = p$ and solve the above equation for r to write

$$r = \frac{c_0}{p}(1+p^2)^2, \tag{0.21}$$

which when differentiated gives

$$dr = c_0(-\frac{1}{p^2} + 2 + 3p^2)dp. \tag{0.22}$$

Using

$$\frac{dz}{dr} = p, \tag{0.23}$$

we also obtain

$$\int dz = \int p\ dr \tag{0.24}$$

$$= c_0 \int p(-\frac{1}{p^2} + 2 + 3p^2)dp \tag{0.25}$$

or

$$z = c_0\left(-\ln p + p^2 + \frac{3p^4}{4}\right) + c_1. \tag{0.26}$$

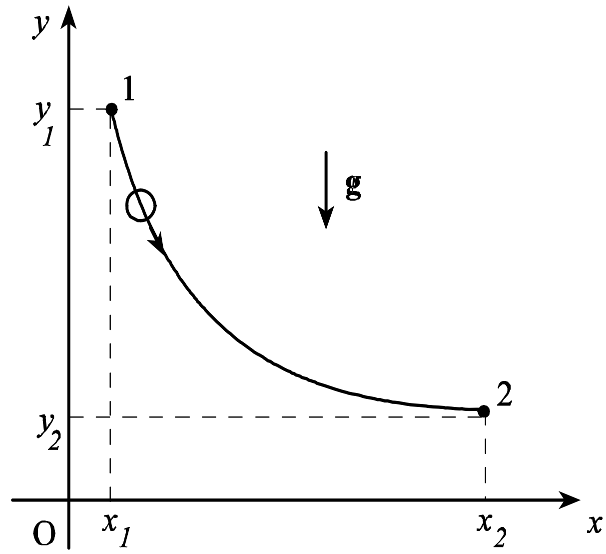

Fig. 0.3 The brachistochrone problem.

Equations (0.21) and (0.26) represent the parametric expression of the needed surface of revolution. To determine the integration constants we can use the values

$$z(a/2) = z_1, \tag{0.27}$$
$$z(R) = z_2. \tag{0.28}$$

For a complete treatment of this problem, which was originally discussed by Newton and which still has engineering interest, see Bryson and Ho (1969).

3. **(Problem 17.11: The brachistocrone problem)** Find the shape of the curve joining two points, along which a particle, initially at rest, falls freely under the influence of gravity from the higher point to the lower point in the least amount of time.

Solution:

Velocity of the bead as it falls freely along the wire is

$$v = \frac{ds}{dt}, \tag{0.29}$$

where ds is the arclength

$$ds = \sqrt{dx^2 + dy^2}. \tag{0.30}$$

We write the total time, T, to descend from point 1 to 2 (Fig 0.3) as

$$T = \int_1^2 \frac{ds}{v} \tag{0.31}$$

$$= \int_1^2 \frac{1}{v}\sqrt{1+y'^2}\ dx, \tag{0.32}$$

where $y' = dy/dx$. Since the velocity changes along the wire, we still need to express $v(x)$ as a function of $y(x)$ and possibly x. For this we use the expression of the conservation of energy, which we write as

$$\left[mgy + \frac{1}{2}mv^2\right]_1 = \left[mgy + \frac{1}{2}mv^2\right]_2 = E, \tag{0.33}$$

where E is the conserved total energy of the bead. Since the bead is initially at rest:

$$v_1 = 0, \tag{0.34}$$

we use Equation (0.33) to write

$$v = \sqrt{2g\left(\frac{E}{mg} - y\right)}. \tag{0.35}$$

Using this we finally obtain the functional

$$T[y(x)] = \int_1^2 \sqrt{\frac{1+y'^2}{2g(E'-y)}}\ dx, \tag{0.36}$$

where we wrote $E' = E/mg$. Note that E' is always greater than y for $x \in (1.2)$ in Equation (0.36). We now define a new variable as

$$z = (E' - y) \tag{0.37}$$

to write

$$T[z(x)] = \int_1^2 \sqrt{\frac{1+z'^2}{2gz}}\ dx. \tag{0.38}$$

Since the integrand does not depend on the independent variable, we can write the first integral as

$$\frac{1}{\sqrt{z[1+z'^2]}} = C_0, \tag{0.39}$$

which can be solved for z' as

$$z' = \sqrt{\frac{1-C_0^2 z}{C_0 z}}, \tag{0.40}$$

or as

$$C_1 + \int_0^u \sqrt{\frac{u}{1-u}}\, du = C_0^2 x, \tag{0.41}$$

where we have defined

$$u = C_0^2 z. \tag{0.42}$$

To evaluate the integral we substitute

$$u = \sin^2\theta \tag{0.43}$$

to write

$$C_1 + \int_0^\theta 2\sin^2\theta\, d\theta = C_1 + \int_0^\theta (1-\cos 2\theta)\, d\theta = C_0^2 x, \tag{0.44}$$

which yields

$$x = A + B(2\theta - \sin 2\theta), \tag{0.45}$$

where

$$A = \frac{C_1}{C_0^2},\ B = \frac{1}{2C_0^2}. \tag{0.46}$$

Since from Equations (0.42) and (0.43), z is also related to θ through the relation

$$z = B(1-\cos 2\theta), \tag{0.47}$$

we obtain the parametric expression of the curve as

$$z = B(1-\cos 2\theta), \tag{0.48}$$
$$x = A + B(2\theta - \sin 2\theta). \tag{0.49}$$

4. **(Rayleigh-Ritz method: First-order)** Consider the differential equation

$$y'' + \lambda a(x) y(x) = 0,\ y(0) = y(1) = 0, \tag{0.50}$$

which could represent the vibrations of a rod with nonuniform cross-section given by $a(x)$. By choosing a suitable trial function, estimate the lowest eigenvalue for $a(x) = x$.

Solution:

Using the trial functions

$$y(x) = \sin \pi x \tag{0.51}$$

and

$$y(x) = x(1-x) \tag{0.52}$$

we can estimate the lowest eigenvalue, λ_0, as

$$\lambda_0 \le \frac{\int_0^1 |y'(x)|^2 \, dx}{\int_0^1 a(x) \, |y(x)|^2 \, dx}, \tag{0.53}$$

which yields the values

$$\lambda_0 \le 19.74 \tag{0.54}$$

and

$$\lambda_0 \le 20.0, \tag{0.55}$$

for $y(x) = \sin \pi x$ and $y(x) = x(1-x)$, respectively.

One can show that for

$$a(x) = \alpha + \beta x, \tag{0.56}$$

Equation (0.50) can be reduced to Bessel's equation, where for $\alpha = 0$, $\beta = 1$, the exact lowest eigenvalue is given as

$$\lambda_0 = 18.956. \tag{0.57}$$

We can improve our approximation by choosing the trial function as

$$y(x) = \sin \pi x + c \sin 2\pi x. \tag{0.58}$$

This leads to the inequality

$$\lambda_0 \le \frac{2\pi^2(1+4c^2)}{1+c^2-64c/9\pi^2}. \tag{0.59}$$

Minimizing the right-hand side gives

$$c = -0.11386 \tag{0.60}$$

and the improved estimate becomes

$$\lambda_0 \le 18.961. \tag{0.61}$$

5. **(Rayleigh-Ritz method: Second-order)** For the previous problem find an upper bound to the second-order eigenvalue.

   Solution:

In page 539 of Bayin (2006) we have concluded Example 17.9 by saying that the method we use to estimate the lowest eigenvalue can also be used for the higher-order eigenvalues, granted that the trial function is chosen orthogonal to the lower eigenfunctions. In the previous problem we have estimated the lowest eigenvalue via the test function

$$y_0 = \sin \pi x - 0.11386 \sin 2\pi x. \tag{0.62}$$

For the second-order trial function we use

$$y_1 = \sin \pi x + d \sin 2\pi x, \tag{0.63}$$

where d is determined such that y_0 and y_1 are orthogonal. A simple calculation yields

$$d = \frac{9\pi^2(-0.11386) - 32}{32(-0.11386) - 9\pi^2} \tag{0.64}$$

$$= 2.1957, \tag{0.65}$$

which gives the estimate

$$\lambda_1 \leq 94.45. \tag{0.66}$$

An exact calculation in terms of Bessel functions gives

$$\lambda_1 = 81.89. \tag{0.67}$$

Note and also show that the estimates for λ_0 and λ_1 are both upper bounds to the exact eigenvalues.

6. If $y(x)$ extremizes $J[y(x)]$, then regardless of the prescribed end conditions, show that the first variation must vanish:

$$\delta J[y(x)] = 0. \tag{0.68}$$

Solution:

Using the variational notation we write the variation of the functional

$$J[y(x)] = \int_1^2 F(y, y', x)dx \tag{0.69}$$

as

$$J[y(x) + \delta y] - J[y(x)] = \delta J + \delta^2 J + \delta^3 J + \cdots, \tag{0.70}$$

where the second variation is given as

$$\delta^2 J = \frac{1}{2!} \int_1^2 \left[F_{yy}\delta y^2 + 2F_{yy'}\delta y \delta y' + F_{y'y'}\delta y'^2 \right] dx. \tag{0.71}$$

Since $\delta y = \varepsilon\eta(x)$, $\delta y' = \varepsilon\eta'(x)$, etc. [Eq. (17.4)], where ε is a small parameter, $\delta^2 J$ is smaller in magnitude by at least by one power of ε than δJ and so are the higher order variations. On the other hand, $\delta J[\delta y]$ can be written as

$$\delta J[\delta y] = \varepsilon \int_1^2 [F_y \eta(x) + F_{y'}\eta'(x)]\, dx, \tag{0.72}$$

which can be made to be positive or negative for the positive or the negative choices of the small parameter ε, respectively, hence $\delta J[y(x)]$ must vanish for any $y(x)$ that extremizes the functional in Equation (0.69).

II. Optimum Control Theory

Let us now discuss a slightly different problem, where we have to produce a certain amount, say by weight, of goods to meet a certain order at time $t = T$. The problem is to determine the best strategy to follow so that our cost is minimum. One obvious strategy is to produce at a constant rate determined by the amount of goods to be delivered at time T. To see whether this actually minimizes our cost or not, let us formulate this as a variational problem. We first let $x(t)$ be the total amount of goods accumulated at $t \geq 0$, hence its derivative, $x'(t)$, gives the production rate. For the cost there are mainly two sources, one of which is the production cost per unit item, c_P, which can be taken as proportional to the production rate:

$$c_P = k_1 x'(t), \tag{0.73}$$

Naturally, producing faster while maintaining the quality of the product increases the cost per item. Besides, producing the goods faster will increase our inventory unnecessarily before the delivery time, thus increasing the holding cost, c_H, which is defined as the cost per unit item per unit time. As a first approximation, we can take c_H to be proportional to $x(t)$:

$$c_H = k_2 x(t). \tag{0.74}$$

We can now write the total cost of production over the time interval

$$(t, t + \Delta t) \tag{0.75}$$

as

$$\begin{aligned} \delta J &= c_P \delta x + c_H \delta t && (0.76)\\ &= [c_P x'(t) + c_H]\, \delta t && (0.77)\\ &= \left[k_1 x'(t)^2 + k_2 x(t)\right] \delta t. && (0.78) \end{aligned}$$

We also assume that production starts at $t = 0$ with zero inventory, $x(0) = x_0 = 0$, and we need $x(T) = x_T$, where x_T is the amount of goods to be delivered at $t = T$. We can know write the total cost of the entire process as the functional

$$J[x(t)] = \int_0^T \left[k_1 x'(t)^2 + k_2 x(t)\right] dt. \tag{0.79}$$

The problem is to find a production strategy, $x(x)$, that minimizes the functional $J[x(t)]$, subject to the initial conditions

$$x(0) = 0 \text{ and } x(T) = x_T. \tag{0.80}$$

For this problem an acceptable solution should also satisfy the conditions

$$x(t) \geq 0 \text{ and } x'(t) \geq 0. \tag{0.81}$$

Solution of the unconstrained problem with the given initial equations [Eq. (0.80)] is

$$x(t) = \left(x_T - \frac{k_2}{4k_1}T^2\right)\frac{t}{T} + \frac{k_2}{4k_1}t^2. \tag{0.82}$$

The uniform rate of production,

$$x(t) = \frac{x_T t}{T}, \tag{0.83}$$

even though satisfies the end conditions [Eq. (0.80)] and the inequalities in Equation (0.81), does not minimize $J[x(t)]$ for $k_2 \neq 0$. Besides, for realistic problems due to finite capacity we also have an upper and a lower bound for the production rate, hence we also need to satisfy the inequalities

$$x'_M \geq x'(t) \geq x'_m \geq 0, \tag{0.84}$$

where x'_M and x'_m represent the possible maximum and the minimum production rates, respectively. The unconstrained solution is valid only for the times that the inequalities in Equations (0.81) and (0.84) are satisfied. Variational problems with constraints on $x(t)$ and/or $x'(t)$, expressed either as equalities, or inequalities, are handled by the **optimal control theory**, which is a derivative of the variational analysis. In the minimum cost production schedule, to obtain the desired result we need to control the production rate, $x'(t)$, hence the optimal control theory is needed to determine the correct strategy.

Basic Optimum Control Theory: Dynamics versus Controlled Dynamics

In physics a **dynamical system** is described by the second law of Newton as

$$\overrightarrow{F} = m\overrightarrow{a}, \tag{0.85}$$

where $\overrightarrow{F}$ represents the net force acting on the mass, m, and $\overrightarrow{a}$ is the acceleration. For example, for the one dimensional motion of a mass falling in uniform gravity, g, under the influence of a restoring force, $-kx$, and a friction force, $-\mu\dot{x}$, the second law of Newton becomes

$$m\ddot{x} = -\mu\dot{x} - kx - mg, \tag{0.86}$$

where k and μ are constants. With the appropriate initial conditions, $x(0)$ and $\dot{x}(0)$, we can solve this differential equation to find the position, $x(t)$, at a later time. If we also attach a thrust mechanism that allows us to apply force, $f(t)$, to the mass m, then we can control its dynamics so that it arrives at a specific point at a specific time and with a predetermined velocity. Equation (0.85) is now written as

$$\ddot{x} = -\frac{\mu}{m}\dot{x} - \frac{k}{m}x - g + \frac{f(t)}{m}. \tag{0.87}$$

We now define two new variables, y_1 and y_2, that define the **state of the system**:

$$y_1(t) = x(t),\ y_2(t) = \dot{x}(t), \tag{0.88}$$

and introduce u_1 and u_2, called the **control variables** or parameters as

$$u_1(t) = 0,\ u_2(t) = \frac{f(t)}{m}. \tag{0.89}$$

We can write them as the column vectors

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \text{ and } \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{0.90}$$

Controlled dynamics of this system is now governed by the differential equation

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{u}, t), \tag{0.91}$$

where $\mathbf{f}(\mathbf{y}, \mathbf{u}, t)$ is given as

$$\mathbf{f}(\mathbf{y}, \mathbf{u}, t) = \begin{pmatrix} y_2 \\ -\frac{k}{m}y_1 - \frac{\mu}{m}y_2 - g + u_2 \end{pmatrix} \tag{0.92}$$

$$= \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\mu}{m} \end{pmatrix}\mathbf{y} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{u} + \begin{pmatrix} 0 \\ -g \end{pmatrix}. \tag{0.93}$$

We introduce the matrices

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\mu}{m} \end{pmatrix}, \tag{0.94}$$

$$\mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{0.95}$$

$$\mathbf{f}_0 = \begin{pmatrix} 0 \\ -g \end{pmatrix}. \tag{0.96}$$

to write the above equation as

$$\mathbf{f}(\mathbf{y}, \mathbf{u}, t) = \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} + \mathbf{f}_0. \tag{0.97}$$

Note that Equation (0.91) gives two differential equations:

$$\dot{y}_1 = y_2 \tag{0.98}$$

and

$$\dot{y}_2 = -\frac{k}{m}y_1 - \frac{\mu}{m}y_2 - g + u_2, \tag{0.99}$$

to be solved simultaneously, which are coupled and linear. However, in general they are nonlinear and can not be decoupled. For a realistic solution of the fuel-optimal horizontal motion of a rocket problem, one also has to consider the loss of mass due to thrusting (Geering, 2007).

General Statement of a Controlled Dynamics Problem

A general optimal control problem involves the following features:

I) State variables and Controls:

State of the system is described by the state variable, $\mathbf{y}$, written as the column $(n \times 1)$ vector

$$\mathbf{y}(t) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \tag{0.100}$$

while all the admissable controls are described by the $(m \times 1)$ column vector

$$\mathbf{u}(t) = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}. \tag{0.101}$$

II) Vector differential equation of state:

Dynamical evolution of the system is described by the ordinary differential equation

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{u}, t), \tag{0.102}$$

also called the **equation of state**, where $\mathbf{f}(\mathbf{y}, \mathbf{u}, t)$ is a known $(n \times 1)$ continuously differentiable column vector, with the usual **initial condition**

$$\mathbf{y}(t_0) = \mathbf{y}_0. \tag{0.103}$$

Depending on the problem, the terminal state, $\mathbf{y}(T)$ is either fixed or left free.

III) Constraints:

(a) Some constraints on the controls, which are of the form

$$\mathbf{u}(t_1) = \mathbf{u}_1, \tag{0.104}$$

for some t_1 in the time domain.

(b) Some constraints on the controls in the form of inequalities, such as

$$\mathbf{u}_m \leq \mathbf{u} \leq \mathbf{u}_M. \tag{0.105}$$

(c) Some constraints on the state variables, which are either expressed as equality:

$$\Phi(\mathbf{y},t) = 0 \tag{0.106}$$

or as inequality:

$$\Theta(\mathbf{y},t) \geq 0. \tag{0.107}$$

(d) One could also have constraints mixing the state variables and the controls and expressed in various forms.

IV) Solution:

For a given choice of an admissable control, $\mathbf{u}(t)$, we solve the initial value problem [Eq. (0.102)] for $\mathbf{y}(t)$. In other cases we seek for an admissable $\mathbf{u}(t)$ that steers $\mathbf{y}(t)$ to a target value $\mathbf{y}(T)$ at some terminal time T. In optimal control problems, we look for the admissable control variables, $\mathbf{u}(t)$, such that the functional

$$J[u(t)] = \int_{t_0}^{T} F(\mathbf{y}, \mathbf{u}, t)dt + \Psi(T, \mathbf{y}(T)), \tag{0.108}$$

where $F(\mathbf{y}(t), \mathbf{u}(t), t)$ and $\Psi(T, \mathbf{y}(T))$ are known functions, is minimized or maximized. Note that $F(\mathbf{y}, \mathbf{u}, t)$ in Equation (0.108) is different from $f(\mathbf{y}, \mathbf{u}, t)$ in Equation (0.102). In certain type of problems we look for the maximum of $J[u(t)]$, where it is called the **payoff functional**, while $F(\mathbf{y}, \mathbf{u}, t)$ is the **running payoff** and $\Psi(T, \mathbf{y}(T))$ is called the **terminal payoff**. In certain other problems, the minimum of $J[u(t)]$ is desired, where it is called the cost functional.

Connection With Variational Analysis

There is a definite connection between optimal control theory and variational analysis. If we set $\dot{x} = u$ in the action

$$J[x(t)] = \int_1^2 \pounds(x, \dot{x}, t)dt \tag{0.109}$$

and write

$$J[x(t)] = \int_1^2 \pounds(x, u, t)dt, \tag{0.110}$$

and take the constraint as the entire real axis for u, we transform a variational problem to an optimal control one with

$$\dot{x} = u \tag{0.111}$$

representing the equation of state [Eq. (0.102)]. Similarly, if we solve the equation of state [Eq. (0.102)] for u in terms of $\dot{y}, y$ and t, and substitute the result into the payoff functional in Equation (0.110), we can convert an optimal control problem into a variational problem.

However, it should be kept in mind that there is a philosophical difference between the two approaches. In Lagrangian mechanics nature does the optimization and hence controls the entire process. All we can do is to adjust the initial conditions. For example, when firing a cannon ball controlling the system through initial conditions helps to achieve a simple goal like having the ball drop to a specific point. However, if we are sending astronauts to the moon, to assure that they land on the moon safely we have to steer the process all the way. Among other things, we have to assure that the fuel is used efficiently, we have to make sure that the accelerations involved and the cabin conditions stay within certain limits and we have to assure that the rocket lands softly on the moon with enough fuel left to return. Optimum control theory basically allows us to develop the most advantageous strategy to achieve the desired result through some control variables that we build into the system, like the thrust system. In optimal control theory, we are basically *steering* the system to achieve a certain goal.

Controllability of a System

A major concern in optimal control theory is the controllability of a given system. Landing a rocket safely on the moon is a difficult problem, but if we insist on landing it at a specific point at a specific time, that is, if we also fix the terminal state, it becomes a much more difficult problem. In general, it is not clear that a system can be steered from an initial state to a predetermined final state with an admissable choice of the control variables. To demonstrate some of the basic ideas, we confine ourselves to linear systems where the equation of state can be written as

$$\dot{\mathbf{y}} = F(\mathbf{y}, \mathbf{u}, t) \tag{0.112}$$

$$= \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u}. \tag{0.113}$$

Here, $\mathbf{A}$ and $\mathbf{B}$ are $(n \times n)$ and $(n \times m)$ matrices, respectively. To simplify the matter further, consider time-invariant (autonomous) systems. For such systems the $\mathbf{A}$ and $\mathbf{B}$ matrices are constant matrices, hence the controllability of such a system does not depend on the initial time. Let us now consider that $\mathbf{A}$ has a complete set of eigenvectors and let $\mathbf{M}$ be the matrix, columns of which are composed of the eigenvectors of $\mathbf{A}$. We also define the column vector $\mathbf{z}$ as

$$\mathbf{z} = \mathbf{M}^{-1}\mathbf{y}, \tag{0.114}$$

and write Equation (0.114) as

$$\dot{\mathbf{z}} = \mathbf{M}^{-1}\dot{\mathbf{y}} \tag{0.115}$$

$$= \mathbf{M}^{-1}\left[\mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u}\right] \tag{0.116}$$

$$= \mathbf{M}^{-1}\mathbf{A}\mathbf{M}\mathbf{M}^{-1}\mathbf{y} + M^{-1}\mathbf{B}\mathbf{u} \tag{0.117}$$

$$= \left(\mathbf{M}^{-1}\mathbf{A}\mathbf{M}\right) M^{-1}\mathbf{y} + M^{-1}\mathbf{B}\mathbf{u} \tag{0.118}$$

$$= \left(\mathbf{M}^{-1}\mathbf{A}\mathbf{M}\right)\mathbf{z} + \left(M^{-1}\mathbf{B}\right)\mathbf{u}, \tag{0.119}$$

where $\mathbf{M}^{-1}\mathbf{A}\mathbf{M}$ is a diagonal $(n \times n)$ matrix, λ, with its diagonal terms being the eigenvalues:

$$\lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \tag{0.120}$$

From here it is seen that if the matrix $\mathbf{M}^{-1}\mathbf{B}$ has a zero row, say, the kth row, then the kth component of $\mathbf{z}$ satisfies

$$\dot{z}_k = \lambda_k z_k. \tag{0.121}$$

That is, $z_k(t)$ is determined entirely by the initial conditions at t_0. In general, for a linear autonomous system, if $\mathbf{A}$ has a complete set of eigenfunctions, a necessary and sufficient condition for its controllability is that $\mathbf{M}^{-1}\mathbf{B}$ has no zero rows. For linear autonomous systems, where the constant matrix $\mathbf{A}$ does not necessarily has a complete set of eigenvectors, then the following theorem is more useful:

Theorem 0.1: A linear autonomous system is controllable, if and only if the $(n \times nm)$ matrix

$$\mathbf{C} = [\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \ldots, \mathbf{A}^{n-1}\mathbf{B}] \tag{0.122}$$

is of rank n. Proof of this theorem can be found in Wan (1995).

There exists a set of necessary conditions that the optimal solution of an optimal control theory problem should satisfy. This set of conditions is called the *Pontryagin's Minimum Principle*, which can also be used to solve several optimal control problems. For a formulation of the optimal control problem via the Pontryagin's minimum principal see Geering (2007).

Example 0.1. An Inventory Control Model:

A firm has an inventory of y_1 amount (by weight) of goods produced at the rate of $u_1 = u(t)$. If the rate of sales, which could be taken from the past records, is y_2, we can write the rate of change of the inventory as

$$\dot{y}_1 = u - y_2. \tag{0.123}$$

It is natural to think that the firm will make a bigger effort to sell when the inventory increases, hence we can take $\dot{y}_2$ as proportional to the inventory:

$$\dot{y}_2 = \alpha^2 y_1, \tag{0.124}$$

where α is real and positive. If C_P is the price per unit sale, c_p is the cost per unit produced, and c_h is the holding cost per unit item per unit time, we can write the total revenue over a period of T as the integral

$$J[\mathbf{u}] = \int_0^T F dt \tag{0.125}$$

$$= \int_0^T [C_P y_2 - c_p u - c_h y_1]\, dt. \tag{0.126}$$

Note that we have only one control variable, hence we take $u_2 = 0$ in this problem. We now look for the control variable, $u(t)$, that maximizes the revenue, $J[\mathbf{u}]$, subject to the initial conditions:

$$y_1(0) = y_{10}, \tag{0.127}$$

$$y_2(0) = y_{20}. \tag{0.128}$$

We now write the two conditions [Eqs. (0.123) and (0.124)] as

$$\dot{y}_1 - u + y_2 = 0, \tag{0.129}$$

$$\dot{y}_2 - \alpha^2 y_1 = 0 \tag{0.130}$$

and incorporate them into the problem through two Lagrange multipliers, $\lambda_1(t)$, $\lambda_2(t)$, by defining a new *Lagrangian*, H, as

$$H = F - \lambda_1[\dot{y}_1 - u + y_2] - \lambda_2[\dot{y}_2 - \alpha^2 y_1] \tag{0.131}$$

and consider the variation of

$$\begin{aligned} I[\mathbf{u}] &= \int_0^T H dt \\ &= \int_0^T \left[F - \lambda_1\left(\dot{y}_1 - u + y_2\right) - \lambda_2(\dot{y}_2 - \alpha^2 y_1)\right] dt \\ &= -\left[\lambda_1 y_1 + \lambda_2 y_2\right]_0^T \\ &+ \int_0^T \left[(C_P y_2 - c_p u - c_h y_1) + (\dot{\lambda}_1 + \alpha^2 \lambda_2) y_1 + (\dot{\lambda}_2 - \lambda_1) y_2 + \lambda_1 u\right] dt. \end{aligned} \tag{0.132}$$

Note that the stationary values of $J[\mathbf{u}]$ are also the stationary values of $I[\mathbf{u}]$ (Wan, 1995, pg. 345). However, we also have to take into account that in any realistic business the rate of production is always limited, that is,

$$0 \leq u_m \leq u \leq u_M, \tag{0.133}$$

where u_m and u_M represent the possible minimum and the maximum production rates possible. In this regard, we can not insist on the optimal strategy to be a stationary value of $J[\mathbf{u}]$. We can at most ask for $\delta I[\mathbf{u}]$ be nonincreasing, that is, $\delta I[\mathbf{u}] \leq 0$, for a maximum of $J[\mathbf{u}]$:

$$\begin{aligned}\delta I[\mathbf{u}] = &-[\lambda_1(T)\delta y_1(T) + \lambda_2(T)\delta y_2(T)] \\ &+ \int_0^T \left[(\dot{\lambda}_1 + \alpha^2\lambda_2 - c_h)\delta y_1 + (\dot{\lambda}_2 - \lambda_1 + C_P)\delta y_2 + (\lambda_1 - c_p)\delta u\right] dt \leq 0.\end{aligned} \tag{0.134}$$

Since we have fixed the initial conditions [Eqs. (0.127) and (0.128)], we have taken

$$\delta y_1(0) = \delta y_2(0) = 0. \tag{0.135}$$

For simplicity, we also choose the Lagrange multipliers such that

$$\lambda_1(T) = 0, \tag{0.136}$$

$$\lambda_2(T) = 0, \tag{0.137}$$

$$\dot{\lambda}_1(t) + \alpha^2\lambda_2(t) - c_h = 0, \tag{0.138}$$

$$\dot{\lambda}_2(t) - \lambda_1(t) + C_P = 0. \tag{0.139}$$

The first two terms eliminate the surface term in Equation (0.134), which is needed, since we are not given the terminal values $y_1(T)$ and $y_2(T)$, and the last two equations are needed to avoid the need for a relation between $\delta y_1$and δy_2 in the integrand, thus reducing Equation (0.134) to

$$\delta I[\mathbf{u}] = \int_0^T (\lambda_1 - c_p)\delta u dt \leq 0. \tag{0.140}$$

The two coupled linear equations for $\lambda_1(t)$ and $\lambda_2(t)$ [Eqs. (0.138) and (0.139)] can be solved immediately. After incorporating the end conditions [Eqs. (0.136) and (0.137)] we obtain

$$\lambda_1(t) = C_P\{1 - \cos(\alpha[T-t])\} - \frac{c_h}{\alpha}\sin(\alpha[T-t]), \tag{0.141}$$

$$\lambda_2(t) = \frac{c_h}{\alpha^2}\{1 - \cos(\alpha[T-t])\} + \frac{C_p}{\alpha}\sin(\alpha[T-t]). \tag{0.142}$$

With $\lambda_1(t)$ determined as in Equation (0.141), we can not in general have

$$\lambda_1(t) - c_p = 0, \tag{0.143}$$

obviously not when c_p is a constant, hence we can not use $\delta I = 0$. We now turn to $I[u]$ in Equation (0.132) and substitute the expressions found for $\lambda_1(t)$ and $\lambda_2(t)$ to get

$$I[\mathbf{u}] = -[\lambda_1(0)y_{10} + \lambda_2(0)y_{20}] + \int_0^T (\lambda_1 - c_p)u\ dt. \tag{0.144}$$

For a maximum of $J[u]$ we need to pick the largest possible value of u that makes the integral a maximum. In other words, we need

$$u(t) = \begin{cases} u_M & \text{when} \quad (\lambda_1 - c_p) > 0, \\ \\ u_m & \text{when} \quad (\lambda_1 - c_p) < 0. \end{cases} . \tag{0.145}$$

We now check the inequality in Equation (0.140). Since $\lambda_1(T) = 0$, we have

$$\lambda_1(T) - c_p = -c_p < 0, \tag{0.146}$$

hence it is also satisfied. Optimum control models, where the control variables alternate from two extreme values are called the **bang-bang** models. For the limitations of this simple control model see Wan (1995). What is depicted here is a very brief introduction to the interesting field of optimal control theory. For the interested reader who wants to explore this subject further, we recomment the following books and sites.

IV. References and Useful Sites

Books:

Bryson, A., and Y.C. Ho, *Applied Optimal Control Theory*, Ginn, Lexington, MA, 1969.

Gamkrelidze, R.V., *Principles of Optimal Control Theory*, Plenum, New York, 1978.

Geering, H.P., *Optimal Control With Engineering Applications*, Springer, Berlin, 2007.

Griffel, D.E., *Applied Functional Analysis*, Ellis Horwood Ltd., New York, 1988.

Wan, F.Y.M., *Introduction To the Calculus of Variations and Its Applications*, Chapman &Hall, NY, 1995.

Weinstock, R., *Calculus of Variations*, Dover, New York, 1974.

Useful Links:

Calculus of Variations:

`http://en.wikipedia.org/wiki/Calculus_of_variations`,

`http://mathworld.wolfram.com/CalculusofVariations.html`.

Newton's Bucket:

`http://demonstrations.wolfram.com/NewtonsRotatingBucketExperiment/`,

`http://en.wikipedia.org/wiki/Bucket_argument`.

Inventory Control Models:

`http://demonstrations.wolfram.com/ASimpleInventoryControlModel/`.

Optimum Control Theory:

`http://en.wikipedia.org/wiki/Optimal_control`,

`http://math.berkeley.edu/~evans/control.course.pdf`,

`http://mathworld.wolfram.com/OperationsResearch.html`.

Selçuk Bayin (March, 2009)