Analytical Dynamics: Lagrange's Equation and its Application – A Brief Introduction

D. S. Stutts, Ph.D.

Associate Professor of Mechanical Engineering
Missouri University of Science and Technology
Rolla, MO 65409-0050

January 24, 2011*

*Corrected typo in Equation (6) – thanks, Shantanu!

1 The Calculus of Variations

The calculus of variations is an extensive subject, and there are many fine references which present a detailed development of the subject – *see Bibliography.* The purpose of this addendum is do provide a brief background in the theory behind Lagrange's Equations. Fortunately, complete understanding of this theory is not *absolutely necessary* to use Lagrange's equations, but a *basic* understanding of variational principles can greatly increase your mechanical modeling skills.

1.1 Extremum of an Integral – The Euler-Lagrange Equation

Given the Integral of a functional (a function of functions) of the form

$$I(\epsilon) = \int_{t_1}^{t_2} F(U, \dot{U}, t)dt, \tag{1}$$

where t_1, and t_2 are arbitrary, ϵ is a small positive, real constant, and U and $\dot{U}$ are given by

$$U(t) = u(t) + \epsilon\eta(t), \text{ and } \dot{U}(t) = \dot{u}(t) + \epsilon\dot{\eta}(t). \tag{2}$$

The functions U, and u may be thought of as describing the possible positions of a dynamical system between the two instants in time, t_1, and t_2, where $u(t)$ represents the position when the integral described by Equation (1) is stationary, i.e. where it is an extremum, and $U(t)$ is $u(t)$ plus a *variation* $\epsilon\eta(t)$. The function $U(t)$ does not by definition render (1) stationary because we shall assume $\eta(t)$ is *independent* of $u(t)$, and we will assume that a *unique* function renders (1) an extremum. The reasons for these assumptions will become clear below. The important point so far is that have not made any restrictive statements about $I(\epsilon)$ other than it is an integral of a functional of the functions $U(t)$ and $\dot{U}(t)$. We will now specify that the functions $u(t)$, and $\eta(t)$ are of class C^2. That is, they possess continuous second derivatives with respect to t. Further, let us stipulate that $\eta(t)$ must vanish at $t = t_1$, and $t = t_2$. In other words, $u(t)$ and $U(t)$ coincide at the end points of the interval [t_1, t_2], where t_1, and t_2 are *arbitrary.*

Now that we have the stage more or less set up, lets see what rules the functional $F(U, \dot{U}, t)$ must obey to render (1) extreme. We have, by definition, that the function $u(t)$ renders I stationary, hence, we know this occurs when $U(t) = u(t)$, or $\epsilon = 0$. this situation is depicted in Figure 1.

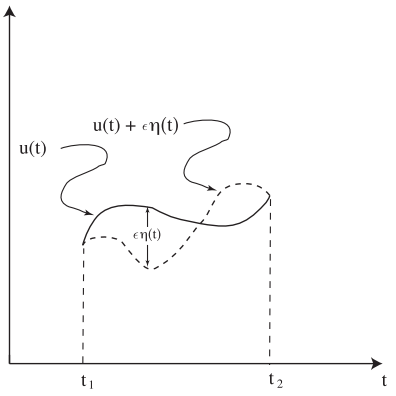

Figure 1. Relationship between extremizing function $u(t)$, and variation $\epsilon\eta(t)$.

Thus, assuming that t_1, and t_2 are not functions of ϵ, we set the first derivative of $I(\epsilon)$ equal to zero.

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \int_{t_1}^{t_2} \frac{dF}{d\epsilon}(U, \dot{U}, t)dt = 0. \tag{3}$$

However,

$$\frac{dF}{d\epsilon}(U,\dot{U},t) = \frac{\partial F}{\partial U}\frac{\partial U}{\partial \epsilon} + \frac{\partial F}{\partial \dot{U}}\frac{\partial \dot{U}}{\partial \epsilon}, \tag{4}$$

so substituting Equation (4) into Equation (3), and setting $\epsilon = 0$, we have

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \int_{t_1}^{t_2}\left(\frac{\partial F}{\partial u}\eta + \frac{\partial F}{\partial \dot{u}}\dot{\eta}\right)dt = 0. \tag{5}$$

Integration of Equation (5) by parts yields:

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \int_{t_1}^{t_2}\eta(t)\left(\frac{\partial F}{\partial u} - \frac{d}{dt}\frac{\partial F}{\partial \dot{u}}\right)dt + \left.\frac{\partial F}{\partial \dot{u}}\eta(t)\right|_{t_1}^{t_2} = 0. \tag{6}$$

The last term in Equation (6) vanishes because of the stipulation $\eta(t_1) = \eta(t_2) = 0$, which leaves

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \int_{t_1}^{t_2}\eta(t)\left(\frac{\partial F}{\partial u} - \frac{d}{dt}\frac{\partial F}{\partial \dot{u}}\right)dt = 0. \tag{7}$$

By the fundamental theorem of the calculus of variations [1], since $\eta(t)$ is arbitrary except at the end points t_1, and t_2, we must have, in general

$$\left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \frac{\partial F}{\partial u} - \frac{d}{dt}\frac{\partial F}{\partial \dot{u}} = 0. \tag{8}$$

Equation (8) is known as the Euler-Lagrange equation. It specifies the conditions on the functional F to extremize the integral $I(\epsilon)$ given by Equation (1). By extremize, we mean that $I(\epsilon)$ may be (1) maximum, (2) minimum, or (3) an inflection point – i.e. neither maximum, nor minimum. In fact, there is no guarantee of the existence of a *global* extremum; the integral may be only locally extreme for *small* values of ϵ. The determination of the nature of the stationary condition of $I(\epsilon)$ for the general case is beyond the scope of this document. Let it suffice to say that for *every* case considered in this class, $I(\epsilon)$ will be *globally minimum* when $\epsilon = 0$.

Equation (5) is often written

$$\delta I = \left.\frac{dI(\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \epsilon\int_{t_1}^{t_2}\left(\frac{\partial F}{\partial u}\eta + \frac{\partial F}{\partial \dot{u}}\dot{\eta}\right)dt = \int_{t_1}^{t_2}\left(\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial \dot{u}}\delta\dot{u}\right)dt = 0, \tag{9}$$

where $\delta u = \eta$ is the *variation* of u, and [2],

$$\frac{d}{dt}\delta u = \frac{d}{dt}(\epsilon\eta) = \epsilon\dot{\eta} = \delta\frac{du}{dt} = \delta\dot{u}. \tag{10}$$

Using Equation (10), and integrating Equation (9) by parts, we obtain

$$\delta I = \int_{t_1}^{t_2}\left(\frac{\partial F}{\partial u} - \frac{d}{dt}\frac{\partial F}{\partial \dot{u}}\right)\delta u dt, \tag{11}$$

with the stipulation, as before, that $\delta u(t_1) = \delta u(t_2) = 0$.

2 Hamilton's Principle

Hamilton's principal is, perhaps, the most important result in the calculus of variations. We derived the Euler-Lagrange equation for a single variable, u, but we will now shift our attention to a system N particles of mass m_i each. Hence, we may obtain N equations of the form

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i, \tag{12}$$

where the **bold** font indicates a *vector* quantity, and $\mathbf{F}_i$ denotes the *total* force on the i^{th} particle. D'Alembert's principle may be stated by rewriting Equation (12) as

$$m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i = 0 \tag{13}$$

Taking the *dot product* of each of the Equations (13) with the variation in position $\delta\mathbf{r}$, and summing the result over all N particles, yields

$$\sum_{i=1}^{N} (m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i) \cdot \delta\mathbf{r} = 0. \tag{14}$$

We note that the sum of the *virtual work* done by the applied forces over the virtual displacements is given by

$$\delta W = \sum_{i=1}^{N} \mathbf{F}_i \cdot \delta\mathbf{r}. \tag{15}$$

Next, we note that

$$\sum_{i=1}^{N} m_i \ddot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i = \sum_{i=1}^{N} m_i \left[\frac{d}{dt} (\dot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i) - \delta \left(\frac{1}{2} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \right) \right] = \sum_{i=1}^{N} m_i \frac{d}{dt} (\dot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i) - \delta T, \tag{16}$$

where δT is the variation of the kinetic energy. Hence, Equation (16) may be written,

$$\delta T + \delta W = \sum_{i=1}^{N} m_i \frac{d}{dt} (\dot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i). \tag{17}$$

In a manner similar to that shown in Figure 1, and in view of Equation (10) the possible dynamical paths of each particle may be represented as shown in Figure 2, where the *varied* dynamical path may be thought to occur *atemporally*.

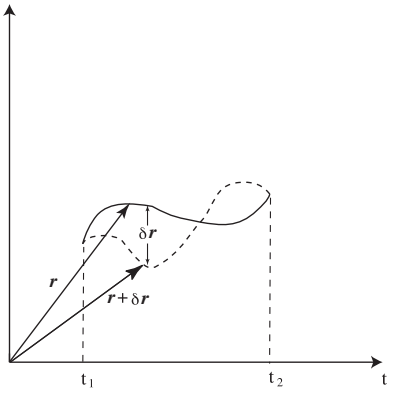

Figure 2. Possible dynamical paths for a particle between two arbitrary instants in time.

Since we again have that $\delta\mathbf{r}(t_1) = \delta\mathbf{r}(t_2) = 0$, we may multiply Equation (17) by dt, and and integrate between the two arbitrary times t_1, and t_2 to obtain

$$\int_{t_1}^{t_2} (\delta T + \delta W)\, dt = \sum_{i=1}^{N} m_i \left(\dot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i\right)\Big|_{t_1}^{t_2} = 0. \tag{18}$$

If δW can be expressed as the variation of the potential energy, $-\delta V$[1], Equation (18) may be written

$$\int_{t_1}^{t_2} (\delta T - \delta V)\, dt = 0. \tag{19}$$

Introducing the Lagrange function, $L = T - V$, Equation (19) becomes

$$\delta \int_{t_1}^{t_2} L dt = 0. \tag{20}$$

Equation (20) is the mathematical statement of *Hamilton's principal.* Hamilton's principal may be defined in words as follows.

Definition 1 *The actual path a body takes in configuration space renders the value of the definite integral* $I = \int_{t_1}^{t_2} L dt$ *stationary with respect to all arbitrary variations of the path between two instants* t_1, *and* t_2 *provided that the path variations vanish at* t_1, *and* t_2 *[2].*

In all physical systems, the stationary value of I will be a minimum.

2.1 Generalized Coordinates

Implicit in the definition of Hamilton's principle is that the system will move along a dynamical path consistent with the system constraints – i.e. along a *permissible* path. *Generalized coordinates* render the dynamical path explicitly permissible by describing it using the *minimum* number of independent coordinates. Thus, the i^{th} system position may be described as a function of the N generalized coordinates, and (in general) time, t, as follows:

$$\mathbf{r}_i = \mathbf{r}_i(q_1, q_2, q_3, \cdots, q_n, t). \tag{21}$$

Hence, the variation of the i^{th} position, which occurs atemporally, may be expressed as

$$\delta\mathbf{r}_i = \sum_{j=1}^{N} \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j. \tag{22}$$

Consider the following example.

Example 1 *Given the system shown in Figure 3, determine the virtual work* δW *done by the force* $\mathbf{F} = 4\mathbf{i} + 3\mathbf{j}$, *where* $\mathbf{i}$, *and* $\mathbf{j}$ *are the unit vectors in the x and y directions, over a virtual displacement* $\delta\mathbf{r}$ *consistent with the constraints:* $x = r\cos\theta$, *and* $y = r\sin\theta$. *Use the generalized coordinates,* $q_1 = r$, *and* $q_2 = \theta$.

[1]The negative sign on δV is chosen to reflect that conservative forces may always be written as the negative gradient of the potential energy: $\mathbf{F}_{ci} = -\boldsymbol{\nabla} V$. Where, for example, in Cartesian coordinates: $\boldsymbol{\nabla} = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}$.

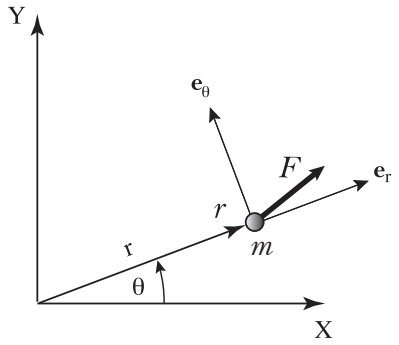

Figure 3. Particle of mass m moving under influence of force $\mathbf{F}$.

Solution:

We have a choice to use either the x, y-coordinate system, in which the force $\mathbf{F}$ is described, or the r, θ-coordinate system, which trivially describes the motion of the particle in terms of the unit vectors, $\mathbf{e}_r$, and $\mathbf{e}_\theta$. Let us choose the former choice to better illustrate the use of generalized coordinates. Writing the location of the mass in terms of x, and y, we have

$$\mathbf{r} = r\cos\theta\mathbf{i} + r\sin\theta\mathbf{j}. \tag{23}$$

Applying Equation (22), to Equation (23), we obtain

$$\delta\mathbf{r}_i = \sum_{j=1}^{N} \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j = (\cos\theta\mathbf{i} + \sin\theta\mathbf{j})\,\delta r + (-r\sin\theta\mathbf{i} + r\cos\theta\mathbf{j})\,\delta\theta. \tag{24}$$

The virtual work is given by $\delta W = \mathbf{F}\cdot\delta\mathbf{r}$, hence, we obtain

$$\delta W = (4\mathbf{i} + 3\mathbf{j})\cdot\left[(\cos\theta\mathbf{i} + \sin\theta\mathbf{j})\,\delta r + (-r\sin\theta\mathbf{i} + r\cos\theta\mathbf{j})\,\delta\theta\right], \tag{25}$$

or,

$$\delta W = 4\left(\cos\theta\delta r - r\sin\theta\delta\theta\right) + 3\left(\sin\theta\delta r + r\cos\theta\delta\theta\right). \tag{26}$$

or

$$\delta W = (4\cos\theta + 3\sin\theta)\delta r + (3\cos\theta - 4\sin\theta)r\delta\theta \tag{27}$$

Careful examination of Equation (27) reveals that the coefficients of the δr, and $r\delta\theta$ terms represent the the components of the applied force $\mathbf{F}$ in the $\mathbf{e}_r$, and $\mathbf{e}_\theta$ directions respectively.

3 Lagrange's Equations of Motion

Writing the position and velocity of each particle in the system as a function of the *generalized coordinates* q_i, and their derivatives with respect to time $\dot{q}_i$, we have that $L = L(q_1, q_2, \cdots, q_n, \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n)$. Hence, following the procedure detailed in Section 1, but replacing the functional F with L, and u with q_i, we obtain

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0. \tag{28}$$

Equation (28) is the Lagrange equation for systems where the virtual work may be expressed as a variation of a potential function, V. In the frequent cases where this is not the case, the so-called *extended Hamilton's principle* must be used.

3.1 Lagrange's Equations Via The Extended Hamilton's Principle

If the virtual work is not derivable from a potential function, then we must begin with equation (18). The kinetic energy is given by

$$T = \frac{1}{2}\sum_{k=1}^{N} m_k \dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_k, \tag{29}$$

and the virtual work may be computed in the manner of Equation (15). Care must be taken to account for p forces, and N generalized coordinates. Hence, the variation of the j^{th} position is given by

$$\delta \mathbf{r}_j = \sum_{k=1}^{N} \frac{\partial \mathbf{r}_j}{\partial q_k}\delta q_k \tag{30}$$

Hence, the virtual work done by p forces acting over N generalized coordinates is given by

$$\delta W = \sum_{j=1}^{p} \mathbf{F}_j \cdot \delta \mathbf{r}_j = \sum_{j=1}^{p} \mathbf{F}_j \cdot \sum_{k=1}^{N} \frac{\partial \mathbf{r}_j}{\partial q_k}\delta q_k. \tag{31}$$

Switching the order of summation, we have

$$\delta W = \sum_{k=1}^{N} \left(\sum_{j=1}^{p} \mathbf{F}_j \cdot \frac{\partial \mathbf{r}_j}{\partial q_k} \right) \delta q_k = \sum_{k=1}^{N} Q_k \delta q_k, \tag{32}$$

where

$$Q_k = \sum_{k=1}^{p} F_j \cdot \frac{\partial \mathbf{r}_j}{\partial q_k}, \text{ for } k = 1, 2, 3, \cdots, N. \tag{33}$$

Substitution of Equation (32) into Equation (18) yields

$$\int_{t_1}^{t_2} \sum_{k=1}^{N} \left[\frac{\partial T}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) + Q_k \right] \delta q_k dt = 0. \tag{34}$$

Since the δq_k are arbitrary between t_1, and t_2, we have

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} = Q_k. \tag{35}$$

If some of the forces are derivable from a potential function V, we may divide the virtual work up into *conservative* virtual work, which is done by those forces derivable from a potential function, and *non--conservative* virtual work done by those which are not derivable from a potential function V. Thus we have,

$$\delta W = \delta W^c + \delta W^{nc} = -\delta V + \sum_{k=1}^{N} Q_k^{nc} \delta q_k. \tag{36}$$

Substitution of Equation (36) into Equation (18) yields

$$\int_{t_1}^{t_2} \delta \left(T - V \right) dt + \int_{t_1}^{t_2} \sum_{k=1}^{N} Q_k^{nc} \delta q_k dt = 0, \tag{37}$$

or, from the definition of the Lagrangian,

$$\int_{t_1}^{t_2} \left(\delta L + \sum_{k=1}^{N} Q_k^{nc} \delta q_k \right) dt. \tag{38}$$

Applying Equation (8), we obtain Lagrange's equation in its most familiar form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = Q_k^{nc}. \tag{39}$$

3.2 Rayleigh's Dissipation function

An important case where a nonconservative force may be derived from a potential function is that of the viscous damping force. The potential function for viscous forces is called the *Rayleigh dissipation function* after Lord Rayleigh. Presented here without derivation, the Rayleigh dissipation function for a single linear viscous damper is given by

$$D = \frac{1}{2}c\dot{x}^2, \tag{40}$$

where c is the damping constant, and x is the displacement from inertial ground. In a system where there are multiple degrees of freedom, and several dampers between the mass particles, the velocity difference between the ends of the dampers must be accounted for. For example, in a two degree of freedom system, with one set of springs and dampers attached to ground, and another set between the two masses, we have

$$D = \frac{1}{2}\left[c_1\dot{x}_1^2 + c_2(\dot{x}_2 - \dot{x}_1)^2\right], \tag{41}$$

where c_1, and c_2 are the damping constants, and $\dot{x}_1$, and $\dot{x}_2$ are the velocities of the two masses.

In general, Equation 39 may be modified to include the Rayleigh dissipation function, and will assume the form:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} + \frac{\partial D}{\partial \dot{q}_k} = Q_k^{nc}. \tag{42}$$

3.3 Kinematic Requirements of Lagrange's Equation

Lagrangian dynamics, as described thus far, provides a very powerful means to determine the equations of motion for complicated discrete (finite degree of freedom) systems. However, there are two primary kinematic requirements which must be achieved before the determination of the potential functions, and subsequent application of Lagrange's equation.

1. *Coordinate choice:*

   (a) The choice of coordinates must be independent and orthogonal. While it is possible to use non--orthogonal coordinates, the additional complexity incurred is not worth the effort in discrete models. Examples of orthogonal coordinate choices include: Cartesian – x, y,z, cylindrical – r, θ, z, and spherical – r, θ, and ϕ.

   (b) The coordinates must locate the body with respect to an *inertial reference frame.* An inertial reference frame is simply one which is not accelerating.

2. *Translational and rotational energy:*

   In rigid bodies, both the translational and rotational kinetic energy must be accounted for. Three cases exist:

   (a) *Pure rotation* – An object which is in pure rotation has at least one point or line which has zero translational velocity. In this case, all of the kinetic energy is rotational, so only the rotational kinetic potential function need be accounted for.

   (b) *Pure translation* – An object is said to be in pure translation if it has no rotation. In this case only the translational kinetic potential function need be accounted for, so only the velocity of the center of mass is needed.

(c) *Translation and rotation* – A body which is both translating and rotating exhibits no stationary points as does a body in pure rotation. However, a translating and rotating body can exhibit *instantaneous centers of rotation* which have zero velocity with respect to an inertial reference for an instant. An important example of this case is rolling without slipping. The point of contact between a wheel and the ground has zero velocity, so the kinetic energy may be considered to be purely rotational as long as the inertia with respect to the instant center is used.

In the general case of rotation and translation, the *velocity of the center of mass* is used for the translational kinetic potential, and the *angular velocity about the mass center* is used to determine the rotational kinetic potential. Hence, the inertia about the mass center is used in this case. This approach may be used for *all* cases since *all motion* may be broken up into rotation about the center of mass, and translation of the center of mass.

As an example, consider a slender rod of mass, m, and length L used as a pendulum. The kinetic energy may be found in one case, using the instant center of rotation approach, to be $T = T_{rot} = \frac{1}{2}(\frac{1}{3}mL^2)\dot{\theta}^2$, were the pivot point it the reference for the moment of inertia, and in the general case, $T = T_{rot} + T_{trans}$, where the mass center is used. In later case, $T_{rot} = \frac{1}{2}(\frac{1}{12}mL^2)\dot{\theta}^2$, and $T_{trans} = \frac{1}{2}mv_{CM}^2 = \frac{1}{2}m(\frac{L}{2}\dot{\theta})^2$. Hence, adding the rotational and translational kinetic potentials yields the same result as obtained by using the instant center of rotation approach.

3.4 Lagrange Equation Examples

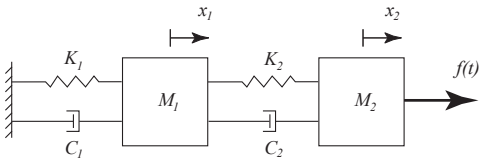

Figure 4. Two degree of freedom example.

Consider the system shown in Figure 4. The equations of motion may be easily found using Equation 42. In this case $q_1 = x_1$, and $q_2 = x_2$. First, we must find the potential functions.

Kinetic Energy:

$$T = \frac{1}{2}\left(M_1\dot{x}_1^2 + M_2\dot{x}_2^2\right) \tag{43}$$

Potential Energy:

$$V = \frac{1}{2}\left[K_1x_1^2 + K_2\left(x_2 - x_1\right)^2\right] \tag{44}$$

Rayleigh's Dissipation Function:

$$D = \frac{1}{2}\left[C_1\dot{x}_1^2 + C_2\left(\dot{x}_2 - \dot{x}_1\right)^2\right] \tag{45}$$

Generalized Force:

$$Q_2 = f(t) \tag{46}$$

Substitution of the potentials and the generalized force into Equation 42 yields the system equations of motion.

$$M_1\ddot{x}_1 + (C_1 + C_2)\dot{x}_1 - C_2\dot{x}_2 + (K_1 + K_2)x_1 - K_2x_2 = 0 \tag{47}$$

$$M_2\ddot{x}_2 - C_2\dot{x}_1 + C_2\dot{x}_2 - K_2x_1 + K_2x_2 = f(t). \tag{48}$$

4 Constrained Maxima and Minima

There are two principal advantages in using the analytical approach to determining the equation of motion for a dynamical system:

1. only positions and velocities need be determined. The resulting accelerations are determined automatically. This often means a considerable savings in computation.
2. All work-less constraint forces are automatically eliminated from the calculations. In the Newtonian approach, all contact forces applied to the body in question must be accounted for, and ultimately determined in order to solve the resulting equations of motion. In a kinematic chain, this requirement leads to significant effort.

The question arises, however: what if we need to know the constraint forces in a system? This is often the case in engineering design and analysis. Fortunately, Lagrange developed an elegant means to solve constrained problems of extremum in general which yields only the constraint forces of interest in dynamical systems.

4.1 Lagrange's Multipliers

Consider a function $u = f(x_1, x_2, \ldots, x_n)$, having at least two continuous derivatives with respect to the independent variables, to be extremized. Obviously, the function, u, must be at least of quadratic polynomial order, or it cannot exhibit an extremum! A necessary condition for extremum is that the total differential of the function u vanishes [3];

$$du = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \ldots, + \frac{\partial f}{\partial x_n}dx_n = 0. \tag{49}$$

Assuming that the variables, x_i, are independent, it follows that the sufficient condition for extremum is

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 0, \\ \frac{\partial f}{\partial x_2} &= 0, \\ &\vdots \\ \frac{\partial f}{\partial x_n} &= 0. \end{aligned} \tag{50}$$

Next, consider the case where some of the independent variables are related by constraints. It is easy to show that Equation (48) is still valid. Consider the function, $u = f(x, y, z)$, where $z = z(x, y)$, is related to x, and y through a constraint equation of the form

$$\phi(x, y, z) = 0. \tag{51}$$

Considering the variables x, and y, as the independent variables, the necessary conditions for extremum are

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0, \\ \frac{\partial u}{\partial y} &= \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0. \end{aligned} \tag{52}$$

Hence, the total differential becomes

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right) = 0. \tag{53}$$

Since

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy, \tag{54}$$

it follows that

$$\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 0. \tag{55}$$

Hence, Equation (48) is still valid even when there are constraint relations between the independent variables. Next, consider the total differential of the constraint given by Equation (50),

$$d\phi = \frac{\partial \phi}{\partial x}dx + \frac{\partial \phi}{\partial y}dy + \frac{\partial \phi}{\partial z}dz = 0. \tag{56}$$

Multiplying Equation (55) by some undetermined multiplier, λ, and adding it to Equation (54) yields

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z}\right)dz = 0. \tag{57}$$

The multiplier, λ, may be chosen so that

$$\begin{aligned} \frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y} &= 0 \\ \frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z} &= 0 \\ \phi(x, y, z) &= 0, \end{aligned} \tag{58}$$

so that the necessary condition for an extremum of $u = f(x, y, z)$ is satisfied.

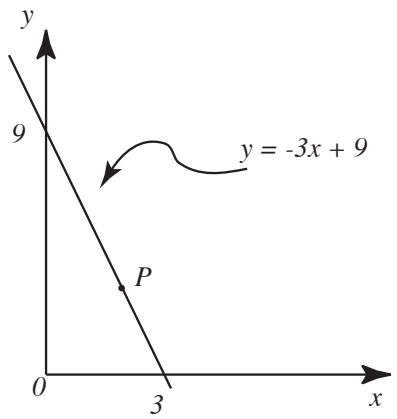

Figure 5. Analytical geometry example.

As an example of the application of Lagrangian multipliers, consider the problem of finding the coordinates of the nearest point to the origin, P, on a specified line [3]. The function to be extremized is the squared distance to the point given by

$$f(x,y) = r^2 = x^2 + y^2, \tag{59}$$

subject to the constraint

$$\phi(x,y) = y + 3x - 9 = 0. \tag{60}$$

We note here that $f(x,y)$ is of class C^2, whereas $x+y$, which is also a measure of the distance from the origin to the point P, is not. Alternatively, we could also use the Euclidian norm, $r = \sqrt{x^2+y^2}$, as the distance measure to be minimized, but the computations are slightly more complex. Applying Equations (57), we have

$$2x + 3\lambda = 0 \tag{61}$$

$$2y + \lambda = 0 \tag{62}$$

$$y + 3x - 9 = 0. \tag{63}$$

From Equation (62), we have $\lambda = -2y$. Substitution into (61) yields $y = \frac{1}{3}x$. Substitution of this result into (63) yields $x = \frac{27}{10}$, and $y = \frac{9}{10}$. Hence, the point $P = (\frac{27}{10}, \frac{9}{10})$ is the nearest point to the origin on the line given by $y = -3x + 9$. That this is so may be easily verified by taking the dot product of the vector from the origin to P with that in the same direction as the line to demonstrate that they are indeed perpendicular:

$$(\frac{27}{10}, \frac{9}{10}) \cdot (3, -9) = 0. \tag{64}$$

A mathematical shorthand may be employed to include the constraints and the function to be extremized in a single "augmented" function given by

$$f^* = f(x_1, x_2, \ldots, x_n) + \lambda\phi(x_1, x_2, \ldots, x_n), \tag{65}$$

where $\phi(x_1, x_2, \ldots, x_n)$ is the constraint function. Next, taking the total differential of Equation (64) while considering the Lagrange multiplier constant, we arrive at

$$\begin{aligned}
\frac{\partial f}{\partial x_1} + \lambda\frac{\partial \phi}{\partial x_1} &= 0 \\
\frac{\partial f}{\partial x_2} + \lambda\frac{\partial \phi}{\partial x_2} &= 0 \\
&\vdots \\
\frac{\partial f}{\partial x_n} + \lambda\frac{\partial \phi}{\partial x_n} &= 0 \\
\phi(x_1, x_2, \ldots, x_n) &= 0,
\end{aligned} \tag{66}$$

In the case where there are n variables related by m constraints, we must define m Lagrange multipliers. Hence, Equation (64) becomes

$$f^* = f(x_1, x_2, \ldots, x_n) + \sum_{j=1}^{m} \lambda_j \phi_j(x_1, x_2, \ldots, x_n), \tag{67}$$

and we define $n + m$ equations of the form

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial \phi_j}{\partial x_i} = 0, \text{ for } i = 1, 2, \ldots, n, \tag{68}$$

and

$$\phi_j(x_1, x_2, \ldots, x_n) = 0, \text{ for } j = 1, 2, \ldots, m. \tag{69}$$

4.2 Application of Lagrange Multipliers to Compute Equilibrium Reaction Forces

Next, we consider the application of Lagrange multipliers to determine the static reaction forces at equilibrium.

Theorem 1 *The total work done by the forces acting on a body in equilibrium during a reversible virtual displacement consistent with the system constraints is zero.*

The above theorem, stated here without proof, seems reasonable since equilibrium implies the absence of explicit time dependent forces. Consider the case for a conservative system where

$$\delta W = -\delta V \tag{70}$$

At equilibrium, we have that

$$\delta W = -\delta V = 0 \tag{71}$$

Equation (70) implies the well know fact that the potential energy is minimum at a stable equilibrium. Consider the potential energy to be a function of n coordinates, $x_1, x_2, \ldots, x_n$, such that

$$\delta V = \frac{\partial V}{\partial x_1}\delta x_1 + \frac{\partial V}{\partial x_2}\delta x_2 + \ldots + \frac{\partial V}{\partial x_n}\delta x_n = 0. \tag{72}$$

Next, we consider the same system to be subject to a constraint of the form

$$\phi(x_1, x_2, \ldots, x_n) = 0. \tag{73}$$

Taking the total variation of the constraint, we have

$$\delta\phi = \frac{\partial \phi}{\partial x_1}\delta x_1 + \frac{\partial \phi}{\partial x_2}\delta x_2 + \ldots + \frac{\partial \phi}{\partial x_n}\delta x_n = 0. \tag{74}$$

Multiplying Equation (73) by an unknown Lagrange multiplier, and *subtracting*[2] it from Equation (71) we have

$$\left(\frac{\partial V}{\partial x_1} - \lambda\frac{\partial \phi}{\partial x_1}\right)\delta x_1 + \left(\frac{\partial V}{\partial x_2} - \lambda\frac{\partial \phi}{\partial x_2}\right)\delta x_2 + \ldots + \left(\frac{\partial V}{\partial x_n} - \lambda\frac{\partial \phi}{\partial x_n}\right)\delta x_n = 0. \tag{75}$$

Equation (74) is analogous to Equation (56), and again, a system of $n+1$ equations analogous to Equation (57) results. As before, mathematical shorthand may be used to augment the potential energy, so that it assumes the form

$$V^* = V - \lambda\phi. \tag{76}$$

Then, the variation is taken as usual, but the Lagrange multiplier, λ, is assumed to be constant. Consider the following simple example:

Given the simple pendulum shown in Figure 6, determine the reaction force on the pivot, F_p, as well as the conditions of equilibrium using the Lagrange multiplier method.

Solution: Taking the pivot point as datum, we find that the potential energy is given by

$$V = -mgr\cos\theta,$$

[2]The variation of the constraint is subtracted, since it represents virtual work done by constraint forces *which always oppose motion*. Where Newton's second law is applied, such forces would carry a negative sign on the opposite side of the equation of motion from acceleration.

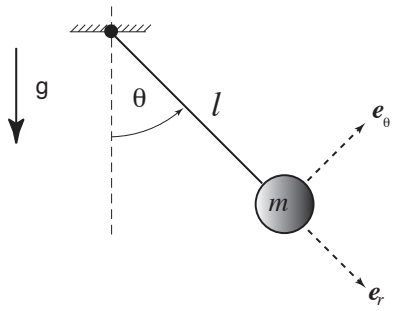

Figure 6. Simple pendulum example.

subject to the constraint

$$\phi = r - l = 0.$$

The augmented potential energy is given by

$$V^* = V - \lambda\phi = -mgr\cos\theta - \lambda(r - l)$$

Taking the variation of V^*, we have

$$\delta V^* = -mg\cos\bar{\theta}\delta r + mgr\sin\bar{\theta}\delta\theta - \lambda\delta r = 0$$

Since the variations are independent, we must have

$$\begin{aligned} -\lambda - mg\cos\bar{\theta} &= 0 \\ \sin\bar{\theta} &= 0, \end{aligned}$$

where $\bar{\theta}$ is the value of θ at equilibrium. The second equation merely implies the obvious fact that at stable equilibrium, $\bar{\theta} = 0$. The first equation implies $\lambda = -mg\cos\bar{\theta} = -mg$ at equilibrium. In this case, λ is the force of constraint acting on the pendulum mass. Hence, the reaction force at the pivot is $F_p = -\lambda = mg$ as expected. We note that this force is in the r-direction because the constraint was in the r-direction.

5 Bibliography

1. Weinstock, Robert, "Calculus of Variations with Applications to Physics and Engineering," 1974, *Dover Publications*, New York.

2. Meirovitch, Leonard, "Methods of Analytical Dynamics," 1970, *McGraw-Hill*, New York.

3. Sokolnikoff, I. S., "Advanced Calculus," 1939, *McGraw-Hill*, New York.