Notes on The Calculus of Variations

Charles Byrne (Charles_Byrne@uml.edu)
Department of Mathematical Sciences
University of Massachusetts at Lowell
Lowell, MA 01854, USA

April 2, 2009

1 Introduction

Typically, we have been concerned with maximizing or minimizing real-valued functions of one or several variables, possibly subject to constraints. In these notes, we consider another type of optimization problem, maximizing or minimizing *a function of functions*. The functions themselves we shall denote by simply $y = y(x)$, instead of the more common notation $y = f(x)$, and the function of functions will be denoted $J(y)$; in the calculus of variations, such functions of functions are called *functionals*. We then want to optimize $J(y)$ over a class of *admissible* functions $y(x)$. We shall focus on the case in which x is a single real variable, although there are situations in which the functions y are functions of several variables.

When we attempt to minimize a function $g(x_1, ..., x_N)$, we consider what happens to g when we perturb the values x_n to $x_n + \Delta x_n$. In order for $\mathbf{x} = (x_1, ..., x_N)$ to minimize g, it is necessary that

$$g(x_1 + \Delta x_1, ..., x_N + \Delta x_N) \geq g(x_1, ..., x_N),$$

for all perturbations $\Delta x_1, ..., \Delta x_N$. For differentiable g, this means that the gradient of g at $\mathbf{x}$ must be zero. In the calculus of variations, when we attempt to minimize $J(y)$, we need to consider what happens when we perturb the function y to a nearby *admissible* function, denoted $y + \Delta y$. In order for y to minimize $J(y)$, we need

$$J(y + \Delta y) \geq J(y),$$

for all Δy that make $y + \Delta y$ admissible. We end up with something analogous to a first derivative of J, which is then set to zero. The result is a differential equation, called the *Euler-Lagrange Equation*, which must be satisfied by the minimizing y.

2 Some Examples

In this section we present some of the more famous examples of problems from the calculus of variations.

2.1 The Shortest Distance

Among all the functions $y = y(x)$, defined for x in the interval $[0, 1]$, with $y(0) = 0$ and $y(1) = 1$, the straight-line function $y(x) = x$ has the shortest length. Assuming the functions are differentiable, the formula for the length of such curves is

$$J(y) = \int_0^1 \sqrt{1 + \Big(\frac{dy}{dx}\Big)^2} dx. \tag{2.1}$$

Therefore, we can say that the function $y(x) = x$ minimizes $J(y)$, over all such functions.

In this example, the functional $J(y)$ involves only the first derivative of $y = y(x)$ and has the form

$$J(y) = \int f(x, y(x), y'(x))dx, \tag{2.2}$$

where $f = f(u, v, w)$ is the function of three variables

$$f(u, v, w) = \sqrt{1 + w^2}. \tag{2.3}$$

In general, the functional $J(y)$ can come from almost any function $f(u, v, w)$. In fact, if higher derivatives of $y(x)$ are involved, the function f can be a function of more than three variables. In this chapter we shall confine our discussion to problems involving only the first derivative of $y(x)$.

2.2 The Brachistochrone Problem

Consider a frictionless wire connecting the two points $A = (0, 0)$ and $B = (1, 0)$. For notational convenience, we assume that the positive y-axis extends below the x-axis, instead of above it, so that the point B is below A.

A metal ball rolls from point A to point B under the influence of gravity. What shape should the wire take in order to make the travel time of the ball the smallest? This famous problem, known as the *Brachistochrone Problem*, was posed in 1696 by Johann Bernoulli. This event is viewed as marking the beginning of the calculus of variations.

The velocity of the ball along the curve is $v = \frac{ds}{dt}$, where s denotes the arc-length. Therefore,

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

Because the ball is falling under the influence of gravity only, the velocity it attains after falling from $(0,0)$ to (x,y) is the same as it would have attained had it fallen y units vertically; only the travel times are different. This is because the loss of potential energy is the same either way. The velocity attained after a vertical free fall of y units is $\sqrt{2gy}$. Therefore, we have

$$dt = \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}dx}{\sqrt{2gy}}.$$

The travel time from A to B is therefore

$$J(y) = \frac{1}{\sqrt{2g}}\int_0^1 \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\frac{1}{\sqrt{y}}dx. \tag{2.4}$$

For this example, the function $f(u,v,w)$ is

$$f(u,v,w) = \frac{\sqrt{1+w^2}}{\sqrt{v}}. \tag{2.5}$$

The interested reader should consider the connection between the solution to the Brachistochrone Problem and the curved path taken by a refracted ray of light in a medium with a variable speed of light.

2.3 Minimal Surface Area

Given a function $y = y(x)$ with $y(0) = 1$ and $y(1) = 0$, we imagine revolving this curve around the x-axis, to generate a surface of revolution. The functional $J(y)$ that we wish to minimize now is the surface area. Therefore, we have

$$J(y) = \int_0^1 y\sqrt{1 + y'(x)^2}dx. \tag{2.6}$$

Now the function $f(u,v,w)$ is

$$f(u,v,w) = v\sqrt{1+w^2}. \tag{2.7}$$

2.4 The Maximum Area

Among all curves of length L connecting the points $(0,0)$ and $(1,0)$, find the one for which the area A of the region bounded by the curve and the x-axis is maximized. The length of the curve is given by

$$L = \int_0^1 \sqrt{1 + y'(x)^2}dx, \tag{2.8}$$

and the area, assuming that $y(x) \geq 0$ for all x, is

$$A = \int_0^1 y(x)dx. \tag{2.9}$$

This problem is different from the previous ones, in that we seek to optimize a functional, subject to a second functional being held fixed. Such problems are called *problems with constraints.*

3 Comments on Notation

The functionals $J(y)$ that we shall consider in this chapter have the form

$$J(y) = \int f(x, y(x), y'(x))dx, \tag{3.1}$$

where $f = f(u, v, w)$ is some function of three real variables. It is common practice, in the calculus of variations literature, to speak of $f = f(x, y, y')$, rather than $f(u, v, w)$. Unfortunately, this leads to potentially confusing notation, such as when $\frac{\partial f}{\partial u}$ is written as $\frac{\partial f}{\partial x}$, which is not the same thing as the total derivative of $f(x, y(x), y'(x))$,

$$\frac{d}{dx} f(x, y(x), y'(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'(x) + \frac{\partial f}{\partial y'} y''(x). \tag{3.2}$$

Using the notation of this chapter, Equation (3.2) becomes

$$\frac{d}{dx} f(x, y(x), y'(x)) = \frac{\partial f}{\partial u}(x, y(x), y'(x)) +$$

$$\frac{\partial f}{\partial v}(x, y(x), y'(x)) y'(x) + \frac{\partial f}{\partial w}(x, y(x), y'(x)) y''(x). \tag{3.3}$$

The common notation forces us to view $f(x, y, y')$ both as a function of three unrelated variables, x, y, and y', and as $f(x, y(x), y'(x))$, a function of the single variable x.

For example, suppose that

$$f(u, v, w) = u^2 + v^3 + \sin w,$$

and

$$y(x) = 7x^2.$$

Then

$$f(x, y(x), y'(x)) = x^2 + (7x^2)^3 + \sin(14x), \tag{3.4}$$

$$\frac{\partial f}{\partial x}(x, y(x), y'(x)) = 2x, \tag{3.5}$$

and

$$\frac{d}{dx} f(x, y(x), y'(x)) = \frac{d}{dx}\Big(x^2 + (7x^2)^3 + \sin(14x)\Big)$$

$$= 2x + 3(7x^2)^2(14x) + 14\cos(14x). \tag{3.6}$$

4 The Euler-Lagrange Equation

In the problems we shall consider in this chapter, admissible functions are differentiable, with $y(x_1) = y_1$ and $y(x_2) = y_2$; that is, the graphs of the admissible functions pass through the end points (x_1, y_1) and (x_2, y_2). If $y = y(x)$ is one such function and $\eta(x)$ is a differentiable function with $\eta(x_1) = 0$ and $\eta(x_2) = 0$, then $y(x) + \epsilon\eta(x)$ is admissible, for all values of ϵ. For fixed admissible function $y = y(x)$, we define

$$J(\epsilon) = J(y(x) + \epsilon\eta(x)), \tag{4.1}$$

and force $J'(\epsilon) = 0$ at $\epsilon = 0$. The tricky part is calculating $J'(\epsilon)$.

Since $J(y(x) + \epsilon\eta(x))$ has the form

$$J(y(x) + \epsilon\eta(x)) = \int_{x_1}^{x_2} f(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x))dx, \tag{4.2}$$

we obtain $J'(\epsilon)$ by differentiating under the integral sign.

Omitting the arguments, we have

$$J'(\epsilon) = \int_{x_1}^{x_2} \frac{\partial f}{\partial v}\eta + \frac{\partial f}{\partial w}\eta' dx. \tag{4.3}$$

Using integration by parts and $\eta(x_1) = \eta(x_2) = 0$, we have

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial w}\eta' dx = -\int_{x_1}^{x_2} \frac{d}{dx}(\frac{\partial f}{\partial w})\eta dx. \tag{4.4}$$

Therefore, we have

$$J'(\epsilon) = \int_{x_1}^{x_2} \Big(\frac{\partial f}{\partial v} - \frac{d}{dx}(\frac{\partial f}{\partial w})\Big)\eta dx. \tag{4.5}$$

In order for $y = y(x)$ to be the optimal function, this integral must be zero for every appropriate choice of $\eta(x)$, when $\epsilon = 0$. It can be shown without too much trouble that this forces

$$\frac{\partial f}{\partial v} - \frac{d}{dx}(\frac{\partial f}{\partial w}) = 0. \tag{4.6}$$

Equation (4.6) is the *Euler-Lagrange Equation*.

For clarity, let us rewrite that Euler-Lagrange Equation using the arguments of the functions involved. Equation (4.6) is then

$$\frac{\partial f}{\partial v}(x, y(x), y'(x)) - \frac{d}{dx}\Big(\frac{\partial f}{\partial w}(x, y(x), y'(x))\Big) = 0. \tag{4.7}$$

5 Special Cases of the Euler-Lagrange Equation

The Euler-Lagrange Equation simplifies in certain special cases.

5.1 If f is independent of v

If the function $f(u, v, w)$ is independent of the variable v then the Euler-Lagrange Equation (4.7) becomes

$$\frac{\partial f}{\partial w}(x, y(x), y'(x)) = c, \tag{5.1}$$

for some constant c. If, in addition, the function $f(u, v, w)$ is a function of w alone, then so is $\frac{\partial f}{\partial w}$, from which we conclude from the Euler-Lagrange Equation that $y'(x)$ is constant.

5.2 If f is independent of u

Note that we can write

$$\frac{d}{dx}f(x, y(x), y'(x)) = \frac{\partial f}{\partial u}(x, y(x), y'(x)) + \frac{\partial f}{\partial v}(x, y(x), y'(x))y'(x) + \frac{\partial f}{\partial w}(x, y(x), y'(x))y''(x). \tag{5.2}$$

We also have

$$\frac{d}{dx}\Big(y'(x)\frac{\partial f}{\partial w}(x, y(x), y'(x))\Big) = y'(x)\frac{d}{dx}\Big(\frac{\partial f}{\partial w}(x, y(x), y'(x))\Big) + y''(x)\frac{\partial f}{\partial w}(x, y(x), y'(x)). \tag{5.3}$$

Subtracting Equation (5.3) from Equation (5.2), we get

$$\frac{d}{dx}\Big(f(x, y(x), y'(x)) - y'(x)\frac{\partial f}{\partial w}(x, y(x), y'(x))\Big) =$$

$$\frac{\partial f}{\partial u}(x, y(x), y'(x)) + y'(x)\Big(\frac{\partial f}{\partial v} - \frac{d}{dx}\frac{\partial f}{\partial w}\Big)(x, y(x), y'(x)). \tag{5.4}$$

Now, using the Euler-Lagrange Equation, we see that Equation (5.4) reduces to

$$\frac{d}{dx}\Big(f(x, y(x), y'(x)) - y'(x)\frac{\partial f}{\partial w}(x, y(x), y'(x))\Big) = \frac{\partial f}{\partial u}(x, y(x), y'(x)). \tag{5.5}$$

If it is the case that $\frac{\partial f}{\partial u} = 0$, then equation (5.5) leads to

$$f(x, y(x), y'(x)) - y'(x)\frac{\partial f}{\partial w}(x, y(x), y'(x)) = c, \tag{5.6}$$

for some constant c.

6 Using the Euler-Lagrange Equation

We derive and solve the Euler-Lagrange Equation for each of the examples presented previously.

6.1 The Shortest Distance

In this case, we have

$$f(u, v, w) = \sqrt{1 + w^2}, \tag{6.1}$$

so that

$$\frac{\partial f}{\partial v} = 0,$$

and

$$\frac{\partial f}{\partial u} = 0.$$

We conclude that $y'(x)$ is constant, so $y(x)$ is a straight line.

6.2 The Brachistochrone Problem

Equation (2.5) tells us that

$$f(u, v, w) = \frac{\sqrt{1 + w^2}}{\sqrt{v}}. \tag{6.2}$$

Then, since

$$\frac{\partial f}{\partial u} = 0,$$

and

$$\frac{\partial f}{\partial w} = \frac{w}{\sqrt{1+w^2}\sqrt{v}},$$

Equation (5.6) tells us that

$$\frac{\sqrt{1+y'(x)^2}}{\sqrt{y(x)}} - y'(x)\frac{y'(x)}{\sqrt{1+y'(x)^2}\sqrt{y(x)}} = c. \tag{6.3}$$

Equivalently, we have

$$\sqrt{y(x)}\sqrt{1+y'(x)^2} = \sqrt{a}. \tag{6.4}$$

Solving for $y'(x)$, we get

$$y'(x) = \sqrt{\frac{a-y(x)}{y(x)}}. \tag{6.5}$$

Separating variables and integrating, using the substitution

$$y = a\sin^2\theta = \frac{a}{2}(1-\cos 2\theta),$$

we obtain

$$x = 2a\int \sin^2\theta d\theta = \frac{a}{2}(2\theta - \sin 2\theta) + k. \tag{6.6}$$

From this, we learn that the minimizing curve is a *cycloid*, that is, the path a point on a circle traces as the circle rolls.

There is an interesting connection, discussed by Simmons in [1], between the brachistochrone problem and the refraction of light rays. Imagine a ray of light passing from the point $A = (0, a)$, with $a > 0$, to the point $B = (c, b)$, with $c > 0$ and $b < 0$. Suppose that the speed of light is v_1 above the x-axis, and $v_2 < v_1$ below the x-axis. The path consists of two straight lines, meeting at the point $(0, x)$. The total time for the journey is then

$$T(x) = \frac{\sqrt{a^2+x^2}}{v_1} + \frac{\sqrt{b^2+(c-x)^2}}{v_2}.$$

Fermat's Principle of Least Time says that the (apparent) path taken by the light ray will be the one for which x minimizes $T(x)$. From calculus, it follows that

$$\frac{x}{v_1\sqrt{a^2+x^2}} = \frac{c-x}{v_2\sqrt{b^2+(c-x)^2}},$$

and from geometry, we get *Snell's Law*:

$$\frac{\sin\alpha_1}{v_1} = \frac{\sin\alpha_2}{v_2},$$

where α_1 and α_2 denote the angles between the upper and lower parts of the path and the vertical, respectively.

Imagine now a stratified medium consisting of many horizontal layers, each with its own speed of light. The path taken by the light would be such that $\frac{\sin\alpha}{v}$ remains constant as the ray passes from one layer to the next. In the limit of infinitely many infinitely thin layers, the path taken by the light would satisfy the equation $\frac{\sin\alpha}{v} =$ constant, with

$$\sin\alpha = \frac{1}{\sqrt{1+y'(x)^2}}.$$

As we have already seen, the velocity attained by the rolling ball is $v = \sqrt{2gy}$, so the equation to be satisfied by the path $y(x)$ is

$$\sqrt{2gy(x)}\sqrt{1+y'(x)^2} = \text{constant},$$

which is what we obtained from the Euler-Lagrange Equation.

6.3 Minimizing the Surface Area

For the problem of minimizing the surface area of a surface of revolution, the function $f(u,v,w)$ is

$$f(u,v,w) = v\sqrt{1+w^2}. \tag{6.7}$$

Once again, $\frac{\partial f}{\partial u} = 0$, so we have

$$\frac{y(x)y'(x)^2}{\sqrt{1+y'(x)^2}} - y(x)\sqrt{1+y'(x)^2} = c. \tag{6.8}$$

It follows that

$$y(x) = b\cosh\frac{x-a}{b}, \tag{6.9}$$

for appropriate a and b.

7 Problems with Constraints

We turn now to the problem of optimizing one functional, subject to a second functional being held constant. The basic technique is similar to ordinary optimization subject to constraints: we use Lagrange multipliers. We begin with a classic example.

7.1 The Isoperimetric Problem

A classic problem in the calculus of variations is the *Isoperimetric Problem*: find the curve of a fixed length that encloses the largest area. For concreteness, suppose the curve connects the two points $(0,0)$ and $(1,0)$ and is the graph of a function $y(x)$. The problem then is to maximize the area integral

$$\int_0^1 y(x)dx, \tag{7.1}$$

subject to the perimeter being held fixed, that is,

$$\int_0^1 \sqrt{1+y'(x)^2}dx = P. \tag{7.2}$$

With

$$f(x, y(x), y'(x)) = y(x) + \lambda\sqrt{1+y'(x)^2},$$

the Euler-Lagrange Equation becomes

$$\frac{d}{dx}\Big(\frac{\lambda y'(x)}{\sqrt{1+y'(x)^2}}\Big) - 1 = 0, \tag{7.3}$$

or

$$\frac{y'(x)}{\sqrt{1+y'(x)^2}} = \frac{x-a}{\lambda}. \tag{7.4}$$

Using the substitution $t = \frac{x-a}{\lambda}$ and integrating, we find that

$$(x-a)^2 + (y-b)^2 = \lambda^2, \tag{7.5}$$

which is the equation of a circle. So the optimal function $y(x)$ is a portion of a circle.

What happens if the assigned perimeter P is greater than $\frac{\pi}{2}$, the length of the semicircle connecting $(0,0)$ and $(1,0)$? In this case, the desired curve is not the graph of a function of x, but a parameterized curve of the form $(x(t), y(t))$, for, say, t in the interval $[0,1]$. Now we have one independent variable, t, but two dependent ones, x and y. We need a generalization of the Euler-Lagrange Equation to the multivariate case.

8 The Multivariate Case

Suppose that the integral to be optimized is

$$J(x,y) = \int_a^b f(t, x(t), x'(t), y(t), y'(t))dt, \tag{8.1}$$

where $f(u, v, w, s, r)$ is a real-valued function of five variables. In such cases, the Euler-Lagrange Equation is replaced by the two equations

$$\frac{d}{dt}\Big(\frac{\partial f}{\partial w}\Big) - \frac{\partial f}{\partial v} = 0,$$

$$\frac{d}{dx}\Big(\frac{\partial f}{\partial r}\Big) - \frac{\partial f}{\partial s} = 0. \tag{8.2}$$

We apply this now to the problem of maximum area for a fixed perimeter.

We know from Green's Theorem in two dimensions that the area A enclosed by a curve C is given by the integral

$$A = \frac{1}{2}\oint_C (xdy - ydx) = \frac{1}{2}\int_0^1 (x(t)y'(t) - y(t)x'(t))dt. \tag{8.3}$$

The perimeter P of the curve is

$$P = \int_0^1 \sqrt{x'(t)^2 + y'(t)^2}dt. \tag{8.4}$$

So the problem is to maximize the integral in Equation (8.3), subject to the integral in Equation (8.4) being held constant.

The problem is solved by using a Lagrange multiplier. We write

$$J(x, y) = \int_0^1 \Big(x(t)y'(t) - y(t)x'(t) + \lambda\sqrt{x'(t)^2 + y'(t)^2}\Big)dt. \tag{8.5}$$

The generalized Euler-Lagrange Equations are

$$\frac{d}{dt}\Big(\frac{1}{2}x(t) + \frac{\lambda y'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}\Big) + \frac{1}{2}x'(t) = 0, \tag{8.6}$$

and

$$\frac{d}{dt}\Big(-\frac{1}{2}y(t) + \frac{\lambda x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}\Big) - \frac{1}{2}y'(t) = 0. \tag{8.7}$$

It follows that

$$y(t) + \frac{\lambda x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} = c, \tag{8.8}$$

and

$$x(t) + \frac{\lambda y'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} = d. \tag{8.9}$$

Therefore,

$$(x - d)^2 + (y - c)^2 = \lambda^2. \tag{8.10}$$

The optimal curve is then a portion of a circle.

9 Finite Constraints

Suppose that we want to minimize the functional

$$J(y) = \int_a^b f(x, y(x), y'(x))dx,$$

subject to the constraint

$$g(x, y(x)) = 0.$$

Such a problem is said to be one of *finite constraints*. In this section we illustrate this type of problem by considering the geodesic problem.

9.1 The Geodesic Problem

The space curve $(x(t), y(t), z(t))$, defined for $a \leq t \leq b$, lies on the surface described by $G(x, y, z) = 0$ if $G(x(t), y(t), z(t)) = 0$ for all t in $[a, b]$. The *geodesic problem* is to find the curve of shortest length lying on the surface and connecting points $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$. The functional to be minimized is the arc length

$$J = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}dt, \tag{9.11}$$

where $\dot{x} = \frac{dx}{dt}$.

We assume that the equation $G(x, y, z) = 0$ can be rewritten as

$$z = g(x, y),$$

that is, we assume that we can solve for the variable z, and that the function g has continuous second partial derivatives. We may not be able to do this for the entire surface, as the equation of a sphere $G(x, y, z) = x^2 + y^2 + z^2 - r^2 = 0$ illustrates, but we can usually solve for z, or one of the other variables, on part of the surface, as, for example, on the upper or lower hemisphere.

We then have

$$\dot{z} = g_x\dot{x} + g_y\dot{y} = g_x(x(t), y(t))\dot{x}(t) + g_y(x(t), y(t))\dot{y}(t), \tag{9.12}$$

where $g_x = \frac{\partial g}{\partial x}$.

Lemma 9.1 *We have*

$$\frac{\partial \dot{z}}{\partial x} = \frac{d}{dt}(g_x).$$

Proof: From Equation (9.12) we have

$$\frac{\partial \dot{z}}{\partial x} = \frac{\partial}{\partial x}(g_x\dot{x} + g_y\dot{y}) = g_{xx}\dot{x} + g_{yx}\dot{y}.$$

We also have

$$\frac{d}{dt}(g_x) = \frac{d}{dt}(g_x(x(t), y(t)) = g_{xx}\dot{x} + g_{xy}\dot{y}.$$

Since $g_{xy} = g_{yx}$, the assertion of the lemma follows. ∎

From the Lemma we have both

$$\frac{\partial \dot{z}}{\partial x} = \frac{d}{dt}(g_x), \tag{9.13}$$

and

$$\frac{\partial \dot{z}}{\partial y} = \frac{d}{dt}(g_y). \tag{9.14}$$

Substituting for z in Equation (9.11), we see that the problem is now to minimize the functional

$$J = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + (g_x\dot{x} + g_y\dot{y})^2}dt, \tag{9.15}$$

which we write as

$$J = \int_a^b F(x, \dot{x}, y, \dot{y})dt. \tag{9.16}$$

Note that the only place where x and y occur is in the g_x and g_y terms.

The Euler-Lagrange Equations are then

$$\frac{\partial F}{\partial x} - \frac{d}{dt}(\frac{\partial F}{\partial \dot{x}}) = 0, \tag{9.17}$$

and

$$\frac{\partial F}{\partial y} - \frac{d}{dt}(\frac{\partial F}{\partial \dot{y}}) = 0. \tag{9.18}$$

Using

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial \dot{z}}\frac{\partial(g_x\dot{x} + g_y\dot{y})}{\partial x}$$

$$= \frac{\partial f}{\partial \dot{z}}\frac{\partial}{\partial x}(\frac{dg}{dt}) = \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial x}$$

and

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial y},$$

we can rewrite the Euler-Lagrange Equations as

$$\frac{d}{dt}(\frac{\partial f}{\partial \dot{x}}) + g_x \frac{d}{dt}(\frac{\partial f}{\partial \dot{z}}) = 0, \tag{9.19}$$

and

$$\frac{d}{dt}(\frac{\partial f}{\partial \dot{y}}) + g_y \frac{d}{dt}(\frac{\partial f}{\partial \dot{z}}) = 0. \tag{9.20}$$

To see why this is the case, we reason as follows. First

$$\begin{aligned}\frac{\partial F}{\partial \dot{x}} &= \frac{\partial f}{\partial \dot{x}} + \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial \dot{x}} \\ &= \frac{\partial f}{\partial \dot{x}} + \frac{\partial f}{\partial \dot{z}} g_x,\end{aligned}$$

so that

$$\begin{aligned}\frac{d}{dt}(\frac{\partial F}{\partial \dot{x}}) &= \frac{d}{dt}(\frac{\partial f}{\partial \dot{x}}) + \frac{d}{dt}(\frac{\partial f}{\partial \dot{z}} g_x) \\ &= \frac{d}{dt}(\frac{\partial f}{\partial \dot{x}}) + \frac{d}{dt}(\frac{\partial f}{\partial \dot{z}}) g_x + \frac{\partial f}{\partial \dot{z}}\frac{d}{dt}(g_x) \\ &= \frac{d}{dt}(\frac{\partial f}{\partial \dot{x}}) + \frac{d}{dt}(\frac{\partial f}{\partial \dot{z}}) g_x + \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial \dot{x}}.\end{aligned}$$

Let the function $\lambda(t)$ be defined by

$$\frac{d}{dt}(\frac{\partial f}{\partial \dot{z}}) = \lambda(t) G_z,$$

and note that

$$g_x = -\frac{G_x}{G_z},$$

and

$$g_y = -\frac{G_y}{G_z}.$$

Then the Euler-Lagrange Equations become

$$\frac{d}{dt}(\frac{\partial f}{\partial \dot{x}}) = \lambda(t) G_x, \tag{9.21}$$

and

$$\frac{d}{dt}(\frac{\partial f}{\partial \dot{y}}) = \lambda(t) G_y. \tag{9.22}$$

Eliminating $\lambda(t)$ and extending the result to include z as well, we have

$$\frac{\frac{d}{dt}(\frac{\partial f}{\partial \dot{x}})}{G_x} = \frac{\frac{d}{dt}(\frac{\partial f}{\partial \dot{y}})}{G_y} = \frac{\frac{d}{dt}(\frac{\partial f}{\partial \dot{z}})}{G_z}. \tag{9.23}$$

Notice that we could obtain the same result by calculating the Euler-Lagrange Equation for the functional

$$\int_a^b f(\dot{x}, \dot{y}, \dot{z}) + \lambda(t) G(x(t), y(t), z(t)) dt. \tag{9.24}$$

9.2 An Example

Let the surface be a sphere, with equation

$$0 = G(x, y, z) = x^2 + y^2 + z^2 - r^2.$$

Then Equation (9.23) becomes

$$\frac{f\ddot{x} - \dot{x}\dot{f}}{2xf^2} = \frac{f\ddot{y} - \dot{y}\dot{f}}{2yf^2} = \frac{f\ddot{z} - \dot{z}\dot{f}}{2zf^2}.$$

We can rewrite these equations as

$$\frac{\ddot{x}y - x\ddot{y}}{\dot{x}y - x\dot{y}} = \frac{y\ddot{z} - z\ddot{y}}{y\dot{z} - z\dot{y}} = \frac{\dot{f}}{f}.$$

The numerators are the derivatives, with respect to t, of the denominators, which leads to

$$\log|x\dot{y} - y\dot{x}| = \log|y\dot{z} - z\dot{y}| + c_1.$$

Therefore,

$$x\dot{y} - y\dot{x} = c_1(y\dot{z} - z\dot{y}).$$

Rewriting, we obtain

$$\frac{\dot{x} + c_1\dot{z}}{x + c_1 z} = \frac{\dot{y}}{y},$$

or

$$x + c_1 z = c_2 y,$$

which is a plane through the origin. The geodesics on the sphere are great circles, that is, the intersection of the sphere with a plane through the origin.

10 Exercises

10.1 *Suppose that the cycloid in the brachistochrone problem connects the starting point* $(0, 0)$ *with the point* $(\pi a, -2a)$*, where* $a > 0$*. Show that the time required for the ball to reach the point* $(\pi a, -2a)$ *is* $\pi\sqrt{\frac{a}{g}}$.

10.2 *Show that, for the situation in the previous exercise, the time required for the ball to reach* $(\pi a, -2a)$ *is again* $\pi\sqrt{\frac{a}{g}}$*, if the ball begins rolling at any intermediate point along the cycloid. This is the tautochrone property of the cycloid.*

References

[1] Simmons, G. (1972) *Differential Equations, with Applications and Historical Notes*. New York: McGraw-Hill.