Chapter 3

Euler-Lagrange Equations

In this chapter, we apply the results of Sections 2.4 through 2.7 to some problems from the calculus of variations. We shall see that solutions of these problems must satisfy differential equations (called Euler-Lagrange equations). We begin with the so-called C^2-theory in which it is assumed that the integrand f and the admissible functions y have continuous second-order derivatives. After developing this theory and investigating some examples, we will develop an analogous C^1-theory in which the integrand and admissible functions are assumed only to have continuous first-order derivatives.

3.1 C²-Theory

Let $a, b, A, B \in \mathbb{R}$ with $a < b$ and $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be given. We assume that f has continuous second-order partial derivatives. Let $\mathfrak{X} = C^2[a, b]$ and define $\mathscr{Y} \subset \mathfrak{X}$

$$\mathscr{Y} := \left\{ y \in C^2[a, b] \,:\, y(a) = A \text{ and } y(b) = B \right\}.$$

(It is readily verified that $C^2[a, b]$ is a real linear space and is a subspace of $C^1[a, b]$.) We wish to minimize the functional $J : \mathscr{Y} \to \mathbb{R}$ defined by

$$J(y) := \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

Let $y_* \in \mathscr{Y}$ be given and suppose that J attains a minimum at $y_* \in \mathscr{Y}$. We want to carry out the following three steps:

(1) Identify the class of admissible variations $\mathscr{V}_{y_*}$.

(2) Determine $\delta J(y_*; v)$ for each $v \in \mathscr{V}_{y_*}$.

(3) Figure out what can be deduced from the condition $\delta J(y_*; v) = 0$ for all $v \in \mathscr{V}_{y_*}$.

For step (1), we apply Lemma 2.3. We take $\mathcal{W} = \mathbb{R}$, let $w_1 = A$, $w_2 = B$, and define $L_1, L_2 : \mathfrak{X} \to \mathcal{W}$ by

$$L_1(z) = z(a) \quad L_2(z) = z(b) \quad \text{for all } z \in \mathfrak{X}.$$

It is clear that L_1 and L_2 are linear. We conclude from Lemma 2.3 that for each $y \in \mathscr{Y}$ the class of admissible variations at y is (see Figure 3.1)

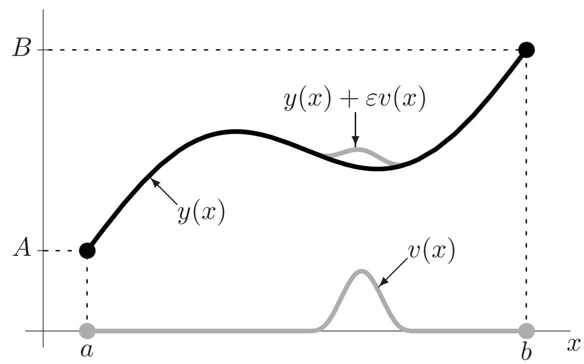

Figure 3.1: Example of a $v \in \mathscr{V}_y$ and $y + \varepsilon v$

$$\mathscr{V}_y = \left\{ v \in C^2[a,b] \,:\, v(a) = v(b) = 0 \right\},$$

and that the interval I in the definition of admissible variation can be taken to be $\mathbb{R}$. Since $\mathscr{V}_y$ does not depend upon y, we drop the subscript and simply denote the admissible variations by $\mathscr{V}$.

For step (2), we use Theorem 2.4. For each $y \in \mathscr{Y}$ and $v \in \mathscr{V}$, we know that $\delta J(y; v)$ exists and is given by

$$\delta J(y;v) = \int_a^b \left\{ f_{,2}(x, y(x), y'(x))v(x) + f_{,3}(x, y(x), y'(x))v'(x) \right\} dx.$$

We now carry out step (3). Since y_* is a minimizer for J over $\mathscr{Y}$, Theorem 2.2 tells us that

$$\delta J(y_*;v) = \int_a^b \left\{ f_{,2}(x, y_*(x), y_*'(x))v(x) + f_{,3}(x, y_*(x), y_*'(x))v'(x) \right\} dx = 0 \quad \text{for all } v \in \mathscr{V}. \tag{3.1}$$

To analyze (3.1), let us define $F, G \in C^1[a,b]$ by

$$F(x) := f_{,2}(x, y_*(x), y_*'(x)) \quad \text{and} \quad G(x) := f_{,3}(x, y_*(x), y_*'(x)) \quad \text{for all } x \in [a,b], \tag{3.2}$$

so we can rewrite (3.1) as

$$\int_a^b \Big\{F(x)v(x) + G(x)v'(x)\Big\}\,dx = 0 \quad \text{for all } v \in \mathscr{V}. \tag{3.3}$$

Since $G \in C^1[a,b]$, we may integrate the second term by parts to obtain

$$\begin{aligned}\int_a^b G(x)v'(x)\,dx &= [G(x)v(x)]\Big|_a^b - \int_a^b G'(x)v(x)\,dx \\ &= G(b)v(b) - G(a)v(a) - \int_a^b G'(x)v(x)\,dx \quad \text{for all } v \in \mathscr{V}. \end{aligned} \tag{3.4}$$

Using the fact that $v(a) = v(b) = 0$ for every $v \in \mathscr{V}$ we find that

$$\int_a^b G(x)v'(x)\,dx = -\int_a^b G'(x)v(x)\,dx \quad \text{for all } v \in \mathscr{V}.$$

Thus (3.3) becomes

$$\int_a^b \{F(x)v(x) - G'(x)v(x)\}\,dx = \int_a^b \{F(x) - G'(x)\}\,v(x)\,dx = 0 \quad \text{for all } v \in \mathscr{V}.$$

Recalling the definitions (3.2), we conclude that if y_* minimizes J on $\mathscr{Y}$ then we must have

$$\int_a^b \left\{f_{,2}(x, y_*(x), y_*'(x)) - \frac{d}{dx}\,[f_{,3}(x, y_*(x), y_*'(x))]\right\} v(x)\,dx = 0 \quad \text{for all } v \in \mathscr{V}. \tag{3.5}$$

The following lemma due to Lagrange will permit us to draw a very powerful conclusion from (3.5).

Lemma 3.1 (Lagrange) *Let $g \in C[a,b]$ be given, and assume that $\int_a^b g(x)v(x)\,dx = 0$ for all $v \in C^2[a,b]$ satisfying $v(a) = v(b) = 0$. Then $g(x) = 0$ for all $x \in [a,b]$.*

Applying this lemma to (3.5) yields the following condition on y_*:

$$f_{,2}(x, y_*(x), y_*'(x)) - \frac{d}{dx}\,[f_{,3}(x, y_*(x), y_*'(x))] = 0 \quad \text{for all } x \in [a,b]. \tag{E-L$_1$}$$

The equation $(\text{E-L})_1$ is called the first Euler-Lagrange equation or simply the Euler-Lagrange equation for J. If y_* is a minimizer (or a maximizer) for J on $\mathscr{Y}$ then it must satisfy this equation.

Before proving Lagrange's lemma, let us examine the Euler-Lagrange equation for some examples.

3.1.1 Example 3.1.1 (cf. Example 1.2)

Consider the functional $J : \mathscr{Y} \to \mathbb{R}$ defined by

$$J(y) := \int_0^1 \left[y(x)^2 + y'(x)^2 \right] dx \quad \text{for all } y \in \mathscr{Y},$$

where

$$\mathscr{Y} := \left\{ y \in C^2[0,1] \,:\, y(0) = 0 \text{ and } y(1) = 1 \right\}.$$

This is the same minimization problem as in Section 1.2, except that the admissible functions are assumed to belong to $C^2[0,1]$ rather than to $C^1[0,1]$. If we define $f : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x,y,z) := y^2 + z^2 \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R} \times \mathbb{R},$$

then J has the form

$$J(y) = \int_0^1 f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

To find the Euler-Lagrange equation for this functional, we need the partial derivatives of f with respect to its second and third arguments. We easily find

$$f_{,2}(x,y,z) = 2y \quad \text{and} \quad f_{,3}(x,y,z) = 2z \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R} \times \mathbb{R}.$$

Therefore

$$f_{,2}(x, y(x), y'(x)) = 2y(x) \quad \text{and} \quad f_{,3}(x, y(x), y'(x)) = 2y'(x) \quad \text{for all } y \in \mathscr{Y},$$

and the first Euler-Lagrange equation for J is

$$2y(x) - \frac{d}{dx}\left[2y'(x)\right] = 0 \quad \text{for all } x \in [0,1]. \tag{E-L$_1$}$$

Thus a minimizer for J on $\mathscr{Y}$ must satisfy

$$\begin{cases} y''(x) - y(x) = 0; \\ y(0) = 0 \text{ and } y(1) = 1. \end{cases} \tag{3.6}$$

Let us solve the boundary value problem (3.6). The Euler-Lagrange equation is a second-order homogeneous linear differential equation with constant coefficients. The characteristic equation is $r^2 - 1 = 0$ which has roots $r = \pm 1$. It follows that the solution to (3.6) has the form

$$y(x) = C_1 e^x + C_2 e^{-x}$$

for some constants C_1 and C_2. It only remains to find C_1 and C_2 so that the boundary conditions are satisfied. We have

$$y(0) = 0 \Rightarrow C_2 = -C_1;$$

and

$$y(1) = 1 \Rightarrow C_1 e + C_2 e^{-1} = 1 \Rightarrow C_1 = \frac{1}{e - e^{-1}} \text{ and } C_2 = \frac{-1}{e - e^{-1}}.$$

Thus the only solution to (3.6) is given by

$$y(x) = \frac{e^x - e^{-x}}{e - e^{-1}}.$$

What we know at this point is that if there is a minimizer, it is given by the formula above. It turns out that this function is indeed a minimizer. We shall establish this last claim in Section 5.7.1.

3.1.2 Example 3.1.2

For this example, we take

$$\mathscr{Y} := \left\{ y \in C^2[1, 32] \,:\, y(1) = 1 \text{ and } y(32) = 2 \right\}$$

and define $J : \mathscr{Y} \to \mathbb{R}$ by

$$J(y) := \int_1^{32} x^2 y'(x)^6 \, dx \quad \text{for all } y \in \mathscr{Y}.$$

The integrand for J is the function $f : [1, 32] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x, y, z) = x^2 z^6 \quad \text{for all } (x, y, z) \in [1, 32] \times \mathbb{R} \times \mathbb{R}.$$

To write $(\text{E-L})_1$ for J, we observe that

$$f_{,2}(x, y, z) = 0 \quad \text{and} \quad f_{,3}(x, y, z) = 6x^2 z^5 \quad \text{for all } (x, y, z) \in [1, 32] \times \mathbb{R} \times \mathbb{R}.$$

Therefore, the first Euler-Lagrange equation for this example is

$$\frac{d}{dx}\left[x^2 y'(x)^5\right] = 0 \quad \text{for all } x \in [1, 32], \tag{E-L$_1$}$$

so a minimizer for J on $\mathscr{Y}$ must satisfy

$$\begin{cases} \dfrac{d}{dx}\left[x^2 y'(x)^5\right] = 0; \\ y(1) = 1 \text{ and } y(32) = 2. \end{cases} \tag{3.7}$$

To solve (3.7), observe that we can integrate the Euler-Lagrange equation. This yields

$$x^2 y'(x)^5 = C_1 \Rightarrow y'(x) = C_2 x^{-\frac{2}{5}} \Rightarrow y(x) = C_3 x^{\frac{3}{5}} + C_4.$$

(Note that we have made the substitutions C_3 for $\frac{5}{3}C_2$ and C_2 for $C_1^{\frac{1}{5}}$). We now use the boundary conditions to find values for C_3 and C_4:

$$y(1) = 1 \Rightarrow C_3 + C_4 = 1 \Rightarrow C_3 = 1 - C_4;$$

and

$$y(32) = 2 \Rightarrow 8C_3 + C_4 = 2 \Rightarrow C_4 = \frac{6}{7} \Rightarrow C_3 = \frac{1}{7}.$$

The only solution to (3.7) is therefore

$$y(x) = \frac{1}{7}x^{\frac{3}{5}} + \frac{6}{7}.$$

We shall show in Section 5.7.1 that the solution found above is, in fact, a minimizer.

3.2 Proof of Lagrange's Lemma

In this section, we provide a proof of Lagrange' lemma. For ease of reference, we restatte the lemma here in a self-contained form.

Lemma 3.1 (Lagrange) *Let $a, b \in \mathbb{R}$ with $a < b$ and $g \in C[a,b]$ be given and put*

$$\mathscr{V} := \left\{ v \in C^2[a,b] \,:\, v(a) = v(b) = 0 \right\}.$$

Assume that $\displaystyle\int_a^b g(x)v(x)\,dx = 0$ *for all* $v \in \mathscr{V}$. *Then* $g(x) = 0$ *for all* $x \in [a,b]$.

Our proof of this lemma relies on the construction of a function from $\mathscr{V}$ having special properties.

Lemma 3.2 *Let $a, b, \alpha, \beta \in \mathbb{R}$ with $a < \alpha < \beta < b$ be given. Define the function $v_* : [a,b] \to \mathbb{R}$ by (see Figure 3.2)*

$$v_*(x) := \begin{cases} 0, & a \le x < \alpha; \\ (x-\alpha)^4(x-\beta)^4, & \alpha \le x \le \beta; \\ 0, & \beta < x \le b. \end{cases}$$

Then $v_* \in C^2[a,b]$ and $v_*(a) = v_*(b) = 0$.

Proof of Lemma 3.2. It is clear that $v_* \in C[a,b]$ and $v_*(a) = v_*(b) = 0$. So we need only show that $v_*', v_*'' \in C[a,b]$.

First we show that v_*' is continuous on $[a,b]$. We easily find that v_* is differentiable at each point in $[a,\alpha) \cup (\alpha,\beta) \cup (\beta,b]$ and

$$v_*'(x) = 0 \quad \text{for all } x \in [a,\alpha) \cup (\beta,b]$$

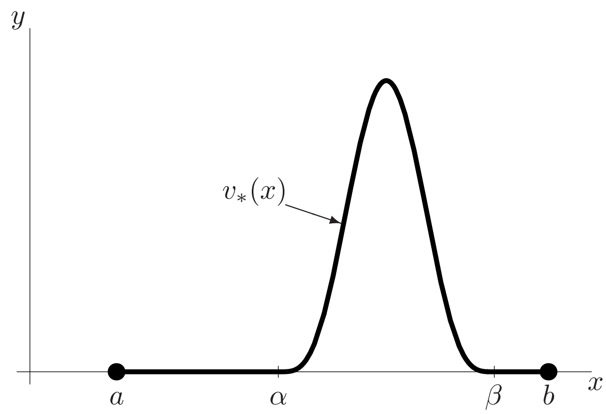

Figure 3.2: Graph of v_*

and

$$v_*'(x) = 4(x-\alpha)^4(x-\beta)^3 + 4(x-\alpha)^3(x-\beta)^4 \quad \text{for all } x \in (\alpha, \beta).$$

We see that v_*' is continuous at each point in $[a,\alpha) \cup (\alpha,\beta) \cup (\beta, b]$, and it only remains to show that v_*' is continuous at α and β. Examining the limit as x tends to α from the left and right, we find

$$\lim_{x\to\alpha^-} v_*'(x) = 0 = \lim_{x\to\alpha^+} \left[4(x-\alpha)^4(x-\beta)^3 + 4(x-\alpha)^3(x-\beta)^4\right] = \lim_{x\to\alpha^+} v_*'(x).$$

It follows that v_* is differentiable at α and v_*' is continuous at α. It is similarly established that v_*' is continuous at β. Thus $v_*' \in C[a,b]$.

Now we prove that $v_*'' \in C[a,b]$. We have

$$v_*''(x) = 0 \quad \text{for all } x \in [a,\alpha) \cup (\beta, b]$$

and

$$\begin{aligned} v_*''(x) = 12(x-\alpha)^4(x-\beta)^2 \\ + 32(x-\alpha)^3(x-\beta)^3 + 12(x-\alpha)^2(x-\beta)^4 \quad \text{for all } x \in (\alpha,\beta). \end{aligned}$$

Obviously v_*'' is continuous on $[a,\alpha) \cup (\alpha,\beta) \cup (\beta,b]$. As with the first derivative of v_*, an examination of the one-sided limits of v_*'' at α and β shows that the second derivative of v_* is continuous at α and β. It follows that $v_*'' \in \mathscr{C}[a,b]$.

Therefore $v_* \in \mathscr{C}^2[a,b]$, as claimed. □

Now we turn to the proof of Lagrange's Lemma.

Proof of Lemma 3.1. We will prove the contrapositive assertion. To this end, suppose that g is not identically zero on the interval $[a,b]$. We want to show that there exists a $v \in \mathscr{V}$ such that $\int_a^b g(x)v(x)\,dx \neq 0$. (This will prove the lemma.)

Since g is not identically zero, there must exist an $x_0 \in (a, b)$ such that $g(x_0) \neq 0$. Without loss of generality, we assume that $g(x_0) > 0$. By the continuity of g, we may choose $\delta > 0$ so that $a < x_0 - \delta < x_0 < x_0 + \delta < b$ and $g(x) > 0$ for every $x \in (x_0 - \delta, x_0 + \delta)$ (see Figure 3.3). Define $v_* : [a, b] \to \mathbb{R}$ by

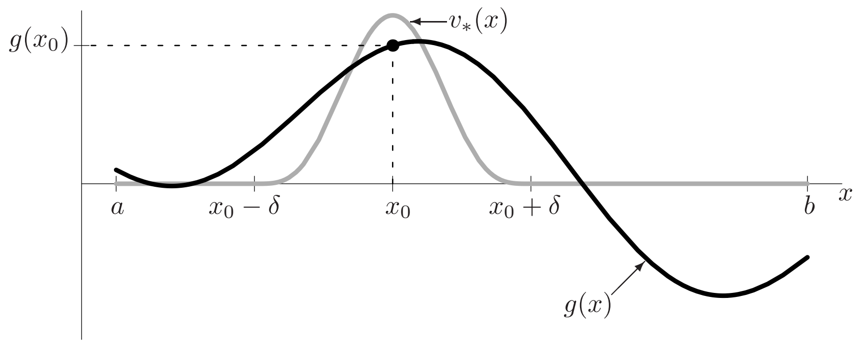

Figure 3.3: Example of a $v_* \in \mathscr{V}$

$$v_*(x) := \begin{cases} 0, & a \leq x < x_0 - \delta; \\ (x - x_0 + \delta)^4 (x - x_0 - \delta)^4, & x_0 \leq x \leq x_0 + \delta; \\ 0, & x_0 + \delta < x \leq b. \end{cases}$$

By Lemma 3.2, we know that $v_* \in \mathscr{V}$. Observe that we have $g(x)v_*(x) > 0$ for every $x \in (x_0 - \delta, x_0 + \delta)$ and $g(x)v_*(x) = 0$ elsewhere on [a,b]. It follows that $\int_a^b g(x)v(x)\,dx > 0$ and the proof is complete. □

3.3 Problems with Free Endpoints

So far we have focused on problems where the values of the admissible functions are prescribed at both endpoints. For these problems, we have found that a minimizer (or maximizer) must satisfy the Euler-Lagrange equation as well as the prescribed boundary conditions. Now we wish to determine what conditions must be satisfied by a minimizer (or maximizer) when the values of the admissible functions y are prescribed at only one end, or when the values of y are not prescribed at either end. An endpoint at which the value of y is not prescribed is called a *free end*. As we shall see, in problems with free ends, the Euler-Lagrange equations still holds and at each free end a *natural boundary condition* must hold.

We begin with some notation. Let $a, b, A, B \in \mathbb{R}$ with $a < b$ be given. Put

$$\mathfrak{X} := C^2[a, b].$$

Let $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a given function with continuous second-order partial derivatives. Define the functional $J : \mathfrak{X} \to \mathbb{R}$ by

$$J(y) := \int_a^b f(x, y(x), y'(x))\,dx \quad \text{for all } y \in \mathfrak{X}.$$

We now specify a variety of classes of admissible functions and indicate the corresponding classes of admissible variations (which can easily be determined using Lemma 2.3). Let

$$\begin{aligned}
\mathrm{Y}_{a,b} &:= \left\{y \in C^2[a, b] \,:\, y(a) = A \text{ and } y(b) = B\right\};\\
\mathrm{Y}_a &:= \left\{y \in C^2[a, b] \,:\, y(a) = A\right\};\\
\mathrm{Y}_b &:= \left\{y \in C^2[a, b] \,:\, y(b) = B\right\};\\
\mathrm{Y} &:= C^2[a, b];\\
\mathrm{V}_{a,b} &:= \left\{v \in C^2[a, b] \,:\, v(a) = v(b) = 0\right\};\\
\mathrm{V}_a &:= \left\{v \in C^2[a, b] \,:\, v(a) = 0\right\};\\
\mathrm{V}_b &:= \left\{v \in C^2[a, b] \,:\, v(b) = 0\right\};\\
\mathrm{V} &:= C^2[a, b].
\end{aligned}$$

Notice that $\mathrm{Y}_{a,b}$ is the usual class of admissible functions when both endpoints are pinned, and earlier we found that $\mathrm{V}_{a,b}$ is the class of admissible variations for each $y \in \mathrm{Y}_{a,b}$. A function in the class Y_a only has its value at a prescribed; it is not required to have any particular value at the endpoint b, and we say that the functions in Y_a have a free endpoint at b. Since a function $y \in \mathrm{Y}_a$ has no prescribed value at b, an admissible variation v for y need not satisfy $v(b) = 0$. It is straightforward to check that V_a, as defined above, is the class of admissible variations for every $y \in \mathrm{V}_a$. Analogously Y_b is the class of admissible functions with a free endpoint at a, and V_b is the class of admissible variations for every $y \in \mathrm{Y}_b$. The set Y is the class of admissible functions for a problem with both endpoints free. The class of admissible variations for each $y \in \mathrm{Y}$ is V (which is all of $C^2[a, b]$).

For the sake of definiteness, we focus on the problem of minimizing J when both endpoints are free. We therefore suppose that J attains a minimum over Y at $y_* \in \mathrm{Y}$. We wish to find some useful conditions that y_* must satisfy.

By Theorem 2.2, for each $v \in \mathrm{V}$ either $\delta J(y_*; v) = 0$ or $\delta J(y_*; v)$ does not exist. Using Theorem 2.4 we find that for each $v \in \mathrm{V}$, the Gâteaux variation of J at y_* in the direction v exists and

$$\delta J(y_*; v) = \int_a^b \left\{f_{,2}(x, y_*(x), y_*'(x))v(x) + f_{,3}(x, y_*(x), y_*'(x))v'(x)\right\} dx = 0 \quad \text{for all } v \in \mathrm{V}. \tag{3.8}$$

Since f has continuous second partial derivatives and $y_* \in C^2[a, b]$, we may define $F, G \in C^1[a,b]$ by

$$F(x) := f_{,2}(x, y_*(x), y_*'(x)) \text{ and } G(x) := f_{,3}(x, y_*(x), y_*'(x)) \quad \text{for all } x \in [a,b]. \tag{3.9}$$

Thus (3.8) becomes

$$\int_a^b \Big\{ F(x)v(x) + G(x)v'(x) \Big\}\, dx = 0 \quad \text{for all } v \in \mathrm{V}. \tag{3.10}$$

As in Section 3.1, we observe that $G \in C^1[a,b]$ and we may therefore integrate the second term in the integrand in (3.10) by parts. Doing so, we obtain

$$\begin{aligned} \int_a^b G(x)v'(x)\, dx &= [G(x)v(x)]\Big|_a^b - \int_a^b G'(x)v(x)\, dx \\ &= G(b)v(b) - G(a)v(a) - \int_a^b G'(x)v(x)\, dx \quad \text{for all } v \in \mathrm{V}. \end{aligned} \tag{3.11}$$

For problems with both ends pinned, the first two terms on the right-hand side of (3.11) were equal to 0. Now, however, since $v \in \mathrm{V}$ tells us nothing about the values of $v(a)$ or $v(b)$, we must carry these two terms along with us. Using (3.11) in (3.10), we find that

$$G(b)v(b) - G(a)v(a) + \int_a^b \{F(x) - G'(x)\}\, v(x)\, dx = 0 \quad \text{for all } v \in \mathrm{V}. \tag{3.12}$$

We now make a key observation: since $\mathrm{V}_{a,b} \subset \mathrm{V}$, if (3.12) holds for all $v \in \mathrm{V}$, then it must hold for all $v \in \mathrm{V}_{a,b}$. Therefore we deduce from (3.12) that

$$G(b)v(b) - G(a)v(a) + \int_a^b \{F(x) - G'(x)\}\, v(x)\, dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}. \tag{3.13}$$

Now for each $v \in \mathrm{V}_{a,b}$, we have $v(a) = v(b) = 0$. Substituting this into (3.13), we find

$$\int_a^b \left\{ F(x) - \frac{d}{dx}[G(x)] \right\} v(x)\, dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}.$$

This is exactly the same condition we found in Section 3.1, so Lagrange's Lemma (Lemma 3.1) allows us to conclude that

$$F(x) - G'(x) = 0 \quad \text{for all } x \in [a,b]. \tag{3.14}$$

With (3.14) in hand, we return to (3.12). We now have

$$G(b)v(b) - G(a)v(a) = 0 \quad \text{for all } v \in \mathrm{V}, \tag{3.15}$$

since (3.14) tells us that the integral term must be zero. We will argue next that (3.15) implies $G(a) = G(b) = 0$. We first choose $v_a \in \mathrm{V}$ such that $v_a(a) = 1$ and $v_a(b) = 0$. For example, we may choose the linear function

$$v_a(x) = \frac{b-x}{b-a}.$$

Since $v_a \in \mathrm{V}$, we have from (3.15) that

$$G(b)v_a(b) - G(a)v_a(a) = -G(a) = 0 \Rightarrow G(a) = 0. \tag{3.16}$$

Similarly, we choose $v_b \in \mathrm{V}$ such that $v_b(b) = 1$ and $v_b(a) = 0$, and with this choice we conclude that

$$G(b)v_b(b) - G(a)v_b(a) = 0 \Rightarrow G(b) = 0. \tag{3.17}$$

Thus $G(a) = G(b) = 0$.

So we have found that (3.10) implies the following:

(1) $F(x) - G'(x) = 0 \quad \text{for all } x \in [a,b]$;

(2) $G(a) = 0$;

(3) $G(b) = 0$.

Recalling the definitions (3.9), we conclude that if y_* minimizes the functional J over Y, then we must have

$$f_{,2}(x, y_*(x), y_*'(x)) - \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x))\right] = 0 \quad \text{for all } x \in [a,b]; \qquad (\text{E-L})_1$$

$$f_{,3}(a, y_*(a), y_*'(a)) = 0; \qquad (\text{NBC})_a$$

$$f_{,3}(b, y_*(b), y_*'(b)) = 0. \qquad (\text{NBC})_b$$

The equation $(\text{E-L})_1$ is the Euler-Lagrange equation we found in Section 3.1. The conditions $(\text{NBC})_a$ and $(\text{NBC})_b$ are called the natural boundary conditions at a and b respectively.

We have shown that if y_* minimizes J over the class Y, then y_* must satisfy $(\text{E-L})_1$, $(\text{NBC})_a$ and $(\text{NBC})_b$. If instead y_* were a minimizer for J over the class Y_a, then y_* is pinned at a and there is no natural boundary condition to be satisfied at a while there is a natural boundary condition $(\text{NBC})_b$ at b. An analogous conclusion may be made if y_* minimizes J over Y_b. To summarize this section, we list the following:

(no free ends) If y_* minimizes J over $\mathrm{Y}_{a,b}$, then y_* must satisfy $(\text{E-L})_1$, $y_*(a) = A$ and $y_*(b) = B$.

(free end at a) If y_* minimizes J over Y_b, then y_* must satisfy $(\text{E-L})_1$, $(\text{NBC})_a$ and $y_*(b) = B$.

(free end at b) If y_* minimizes J over Y_a, then y_* must satisfy $(\text{E-L})_1$, $y_*(a) = A$ and $(\text{NBC})_b$.

(both ends free) If y_* minimizes J over Y, then y_* must satisfy $(\text{E-L})_1$, $(\text{NBC})_a$ and $(\text{NBC})_b$.

3.3.1 Example 3.3.1

We illustrate our results from the previous section with an example (compare to Examples 1.2 and 3.1.1). Set

$$\mathscr{Y} := \left\{ y \in C^2[0,1] \, : \, y(1) = 1 \right\}.$$

The functions in $\mathscr{Y}$ have a free endpoint at 0. We consider the functional $J : \mathscr{Y} \to \mathbb{R}$ defined by

$$J(y) := \int_0^1 \left[y(x)^2 + y'(x)^2 \right] dx \quad \text{for all } y \in \mathscr{Y}.$$

If we define $f : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x,y,z) := y^2 + z^2 \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R} \times \mathbb{R},$$

then J has the form

$$J(y) = \int_0^1 f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathscr{Y}.$$

To write out the Euler-Lagrange equation and natural boundary condition for this problem, we compute the partial derivatives

$$f_{,2}(x,y,z) = 2y \text{ and } f_{,3}(x,y,z) = 2z \quad \text{for all } (x,y,z) \in [0,1] \times \mathbb{R} \times \mathbb{R}.$$

The first Euler-Lagrange equation may thus be written as

$$2y(x) - 2y''(x) = 0 \quad \text{for all } x \in [0,1], \tag{E-L$_1$}$$

and the natural boundary condition at 0 is

$$f_{,3}(0, y(0), y'(0)) = 2y'(0) = 0. \tag{NBC$_0$}$$

Hence, a minimizer for J over $\mathscr{Y}$ must satisfy

$$\begin{cases} y''(x) - y(x) = 0; \\ y'(0) = 0 \text{ and } y(1) = 1. \end{cases} \tag{3.18}$$

We will now solve (3.18). In Section 3.1.1, we found that the general solution to the differential equation above has the form

$$y(x) = C_1 e^x + C_2 e^{-x}.$$

Therefore, it only remains to find C_1 and C_2 such that $y'(0) = 0$ and $y(1) = 1$. We have

$$y'(0) = 0 \Rightarrow C_1 - C_2 = 0 \Rightarrow C_1 = C_2;$$

and

$$y(1) = 1 \Rightarrow C_1 e + C_2 e^{-1} = 1 \Rightarrow C_1 = \frac{1}{e + e^{-1}} \text{ and } C_2 = \frac{1}{e + e^{-1}}.$$

Hence the only solution to (3.18) is

$$y(x) = \frac{e^x + e^{-x}}{e + e^{-1}}.$$

3.4 The Second Euler-Lagrange Equation

This section contains a very useful alternative to the first Euler-Lagrange equation.

Theorem 3.1 *Let $a, b \in \mathbb{R}$ with $a < b$ be given, and suppose that the function $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has continuous second partial derivatives. Let $y \in C^2[a,b]$ be given and assume that y satisfies*

$$f_{,2}(x, y(x), y'(x)) = \frac{d}{dx}\left[f_{,3}(x, y(x), y'(x))\right] \quad \text{for all } x \in [a,b]. \tag{E-L$_1$}$$

Then there exists $c \in \mathbb{R}$ such that

$$f(x, y(x), y'(x)) - y'(x) f_{,3}(x, y(x), y'(x)) = c + \int_a^x f_{,1}(t, y(t), y'(t))\, dt \quad \text{for all } x \in [a,b]. \tag{E-L$_2$}$$

Equation (E-L)$_2$ is called the second Euler-Lagrange equation. In contrast with the first Euler-Lagrange equation, the second does not involve second-order derivatives of y. However, the second equation is not necessary simpler than the first, due to the presence of the integral term.

The proof of Theorem 3.1 is not too difficult and is left as an exercise. The idea is to show that (E-L)$_2$ holds if and only if

$$y'(x)\left\{f_{,2}(x, y(x), y'(x)) - \frac{d}{dx}\left[f_3(x, y(x), y'(x))\right]\right\} = 0 \quad \text{for all } x \in [a,b].$$

This indicates that a function y might satisfy $(\text{E-L})_2$ without satisfying $(\text{E-L})_1$. In particular, constant functions always satisfy $(\text{E-L})_2$.

In some problems, the first Euler-Lagrange equation is more convenient than the second, while in other problems the second Euler-Lagrange equation is preferable to the first. In certain situations, it is helpful to use both equations. Therefore it is prudent to keep both $(\text{E-L})_1$ and $(\text{E-L})_2$ in mind when tackling a problem.

3.4.1 An Important Special Case of $(\text{E-L})_2$

If the integrand f has no explicit x-dependence, i.e, if $f_{,1} \equiv 0$, then the second Euler-Lagrange equation reduces to a first-order differential equation (involving an undetermined constant) because the integral term vanishes. In particular, if

$$f(x, y, z) = F(y, z) \quad \text{for all } x \in [a, b], y \in \mathbb{R}, z \in \mathbb{R},$$

where $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a given function with continuous second-order partial derivatives, then the second Euler-Lagrange equations takes the form

$$F(y(x), y'(x)) - y'(x) F_{,2}(y(x), y'(x)) = c, \tag{E-L$_2$}$$

for some $c \in \mathbb{R}$.

3.4.2 The Brachistochrone Problem

Let $b, B > 0$ be given. For the brachistochrone problem (Section 1.4.2), we found that the functional to be minimized was

$$J(y) = \int_0^b \sqrt{\frac{1 + y'(x)^2}{y(x)}}\, dx.$$

The integrand for J is the function $f : [0, b] \times (0, \infty) \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x, y, z) := \frac{(1 + z^2)^{\frac{1}{2}}}{y^{\frac{1}{2}}} \quad \text{for all } (x, y, z) \in [0, b] \times (0, \infty) \times \mathbb{R}.$$

We are interested in minimizing J over a class of admissible functions satisfying $y(0) = 0$ and $y(b) = B$. The theory that we have developed so far does not apply to this problem because the integrand is not well behaved when $y = 0$ and we are forcing y to vanish at the left endpoint of the interval. Nevertheless, it is instructive to ignore this technical difficulty for the moment and formally determine the Euler-Lagrange equations. To write the first Euler-Lagrange equation, we compute

$$f_{,2}(x, y, z) = -\frac{(1 + z^2)^{\frac{1}{2}}}{2y^{\frac{3}{2}}} \quad \text{for all } (x, y, z) \in [0, b] \times (0, \infty) \times \mathbb{R}$$

and

$$f_{,3}(x,y,z) = \frac{z}{y^{\frac{1}{2}}(1+z^2)^{\frac{1}{2}}} \quad \text{for all } (x,y,z) \in [0,b] \times (0,\infty) \times \mathbb{R},$$

which yields

$$-\frac{(1+y'(x)^2)^{\frac{1}{2}}}{2y(x)^{\frac{3}{2}}} = \frac{d}{dx}\left[\frac{y'(x)}{y(x)^{\frac{1}{2}}(1+y'(x)^2)^{\frac{1}{2}}}\right] \quad \text{for all } x \in [0,b]. \qquad \text{(E-L)}_1$$

It looks as though it would be very difficult to find solutions to this equation.

Since f has no explicit x-dependence, i.e., since $f_{,1}(x,y,z) = 0$ for all $(x,y,z) \in [0,b] \times (0,\infty) \times \mathbb{R}$, the second Euler-Lagrange equation becomes

$$\frac{(1+y'(x)^2)^{\frac{1}{2}}}{y(x)^{\frac{1}{2}}} - \frac{y'(x)^2}{y(x)^{\frac{1}{2}}(1+y'(x)^2)^{\frac{1}{2}}} = c \quad \text{for all } x \in [0,b], \qquad \text{(E-L)}_2$$

for some $c \in \mathbb{R}$. After some simplification, we find that y satisfies (E-L)_2 if and only if it satisfies

$$y(x)(1+y'(x)^2) = \frac{1}{c^2} \quad \text{for all } x \in [0,b], \tag{3.19}$$

for some $c \in \mathbb{R}\backslash\{0\}$. We observe that if y satisfies (3.19) for some $c \in \mathbb{R}\backslash\{0\}$, then $y'(x)^2$ must be large whenever $y(x) > 0$ is small. So if $y \in \mathscr{Y}$ satisfies (E-L)_2 and $y(0) = 0$, we should expect the graph for y to have a vertical tangent at 0.

3.5 Transit Time for a Boat

Suppose that we wish to steer a boat from a given location on the bank of a river with current to a given location on the opposite bank in such a way that the transit time is minimized. For simplicity, we assume the following:

(1) the river banks are straight and parallel;

(2) the current in the river always runs parallel to the river banks;

(3) the strength of the current depends only upon the distance from the initial river bank;

(4) the speed of our boat relative to the water is constant;

(5) the speed of the current is always less than the speed of the boat.

To set up the problem, we choose our coordinate system so that one of the river banks coincides with the y-axis, the other bank is represented by the line $x = b$ with $b > 0$, and our boat's initial position is the origin. We let $Q = (b, B)$ be the destination. We assume that the river flows in the positive y-direction. The

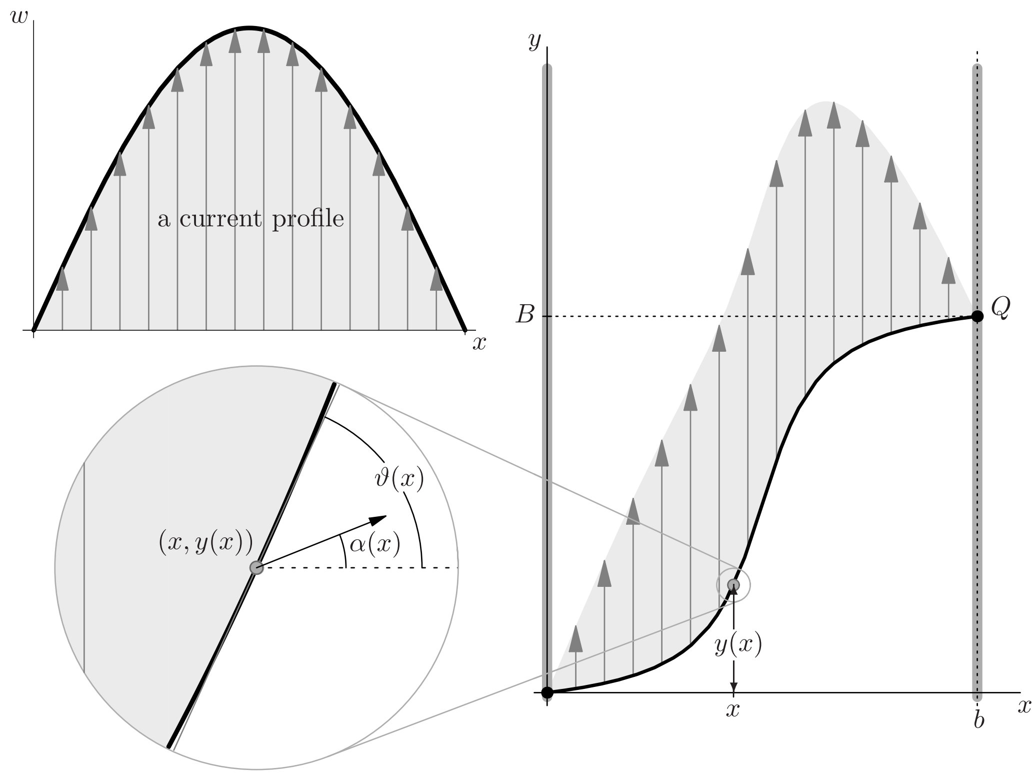

Figure 3.4: Setup for finding the transit time for a boat

speed of the boat relative to the water is denoted by ω, time is denoted by t, and T is the total time to cross the river. We assume that the path of the boat can be represented as the graph of some function $y \in C^1[0,b]$.

Let $\alpha : [0,b] \to (-\frac{\pi}{2}, \frac{\pi}{2})$ be the angle (measured off the x-axis) in which the boat is steered, let $\theta : [0,b] \to (-\frac{\pi}{2}, \frac{\pi}{2})$ be the angle at which the boat is traveling, and let $w : [0,b] \to [0,\omega)$ be the velocity of the current (we are incorporating assumption (5) here). We assume that $w \in C[0,b]$; we will impose additional smoothness assumptions on w later. If we are x-units from the initial river bank, the velocity of the current is $w(x)$, and we are steering the boat in the direction $\alpha(x)$ at the speed ω relative to the water. Though we are steering in the direction $\alpha(x)$, the boat is actually moving in the direction $\theta(x)$ because of the river's current.

If the velocity of the boat (relative to land) in the x-direction is $\frac{dx}{dt}$, then the total time required to traverse the river is

$$T = \int_0^b \frac{1}{\frac{dx}{dt}}\, dx. \tag{3.20}$$

Since the current in the river is assumed to be orthogonal to the x-axis, the current contributes nothing to the boat's velocity in the x-direction, and this velocity is given by

$$\frac{dx}{dt} = \omega \cos \alpha(x).$$

Thus (3.20) may be rewritten as

$$T = \frac{1}{\omega} \int_0^b \sec \alpha(x)\, dx. \tag{3.21}$$

We want to express $\sec \alpha$ in terms of the path of the boat y and the velocity of the current w. When the boat is at a point $(x, y(x))$, the direction in which the boat is traveling is $\theta(x)$. So the slope of the tangent line to the graph of y at x is

$$\begin{aligned} y'(x) = \tan \theta(x) &= \frac{\text{boat's speed in the } y\text{-direction}}{\text{boat's speed in the } x\text{-direction}} \\ &= \frac{\omega \sin \alpha(x) + w(x)}{\omega \cos \alpha(x)} = \frac{\sin \alpha(x) + \frac{w(x)}{\omega}}{\cos \alpha(x)} \\ &= \tan \alpha(x) + \frac{w(x)}{\omega} \sec \alpha(x). \end{aligned} \tag{3.22}$$

For convenience, we define the normalized current $e : [0, b] \to [0, 1)$ by

$$e(x) := \frac{w(x)}{\omega}.$$

Solving for $\tan \alpha$ in (3.22), we find

$$\tan \alpha(x) = y'(x) - e(x) \sec \alpha(x). \tag{3.23}$$

In order to eliminate $\tan \alpha$ from (3.23), we square both sides of (3.23) and use the trigonometric identity $\tan^2 \alpha = \sec^2 \alpha - 1$ to obtain

$$[y'(x) - e(x) \sec \alpha(x)]^2 = \tan^2 \alpha(x) = \sec^2 \alpha(x) - 1.$$

Collecting terms to one side, we find

$$[1 - e(x)^2] \sec^2 \alpha(x) + 2y'(x)e(x) \sec \alpha(x) - y'(x)^2 - 1 = 0.$$

Utilizing the quadratic formula yields

$$\sec \alpha(x) = \frac{-y'(x)e(x) \pm \sqrt{y'(x)^2 + 1 - e(x)^2}}{1 - e(x)^2}. \tag{3.24}$$

At this point, we need to understand what to do with the $\pm$ sign in (3.24). Since $\alpha(x) \in (\frac{-\pi}{2}, \frac{\pi}{2})$,we know that $\sec \alpha(x) > 0$ for every $x \in [0, b]$. Therefore, we want the positive root in (3.24). Upon examining (3.24), we see that this root is given by

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

Having found an expression for $\sec \alpha$ in terms of y and w, we substitute it into (3.21) and find that

$$T = \frac{1}{\omega} \int_0^b \frac{\sqrt{y'(x)^2 + 1 - e(x)^2} - y'(x)e(x)}{1 - e(x)^2}\, dx. \tag{3.25}$$

This is the total transit time for the boat when it is steered along the graph of y across the river.

The optimal path across the river is that which minimizes the expression for T given in (3.25). Our problem, therefore, is to minimize the functional $J : \mathscr{Y} \to \mathbb{R}$ given by

$$J(y) := \int_0^b \frac{\sqrt{1 - e(x)^2 + y'(x)^2} - e(x)y'(x)}{1 - e(x)^2}\, dx \quad \text{for all } y \in \mathscr{Y},$$

where

$$\mathscr{Y} := \left\{ y \in C^2[0, b] \,:\, y(0) = 0 \text{ and } y(b) = B \right\}.$$

Notice that if the river has no current, then $w(x) = 0$ which implies $e(x) = 0$ for every $x \in [0, b]$. In such a situation, equation (3.25) collapses to

$$T = \frac{1}{\omega} \int_0^b \sqrt{1 + y'(x)^2}\, dx.$$

This formula says that the transit time is simply the length of the path (i.e., the total distance traveled) divided by the speed. The minimizer in this case is just the straight line joining the origin to the point (b, B), and the transit time for the boat is the distance between the origin and (b, B) divided by the speed of the boat. This is exactly what one should expect to occur when there is no current in the river.

Returning to the more general problem, let us try to find solutions to one of the Euler-Lagrange equations for J. The integrand $f : [0, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for J is given by

$$f(x, y, z) := \frac{\sqrt{1 - e(x)^2 + z^2}}{1 - e(x)^2} - \frac{e(x)z}{1 - e(x)^2} \quad \text{for all } (x, y, z) \in [0, b] \times \mathbb{R} \times \mathbb{R}.$$

In order to ensure that f has continuous second-order partial derivatives, we assume that $w \in C^2[0, b]$ so that $e \in C^2[0, b]$. Computing the partial derivatives of f we find that

$$f_{,2}(x, y, z) = 0 \quad \text{for all } (x, y, z) \in [0, b] \times \mathbb{R} \times \mathbb{R}$$

and

$$f_{,3}(x,y,z) = \frac{1}{1-e(x)^2}\left(\frac{z}{\sqrt{1-e(x)^2+z^2}} - e(x)\right) \quad \text{for all } (x,y,z) \in [0,b]\times\mathbb{R}\times\mathbb{R},$$

so the first Euler-Lagrange equation becomes

$$\frac{d}{dx}\left[\frac{1}{1-e(x)^2}\left(\frac{y'(x)}{\sqrt{1-e(x)^2+y'(x)^2}} - e(x)\right)\right] = 0. \qquad \text{(E-L)}_1$$

Thus y is a solution to (E-L)_1 if and only if

$$\frac{1}{1-e(x)^2}\left(\frac{y'(x)}{\sqrt{1-e(x)^2+y'(x)^2}} - e(x)\right) = \beta$$

for some $\beta \in \mathbb{R}$. Multiplying through by $1-e(x)^2$, we obtain

$$\beta(1-e(x)^2) = \frac{y'(x)}{\sqrt{1-e(x)^2+y'(x)^2}} - e(x). \tag{3.26}$$

Let us define $\gamma : [0,b]\times\mathbb{R} \to \mathbb{R}$ by

$$\gamma(x,\beta) := e(x) + \beta(1-e(x)^2) \quad \text{for all } x \in [0,b], \beta \in \mathbb{R},$$

so that (3.26) becomes

$$\frac{y'(x)}{\sqrt{1-e(x)^2+y'(x)^2}} = \gamma(x,\beta) \quad \text{for all } x \in [0,b], \beta \in \mathbb{R}. \tag{3.27}$$

It follows from (3.27) that $y'(x)$ and $\gamma(x,\beta)$ always have the same sign. Observe also that (3.27) implies

$$|\gamma(x,\beta)| < 1 \quad \text{for all } x \in [0,b], \beta \in \mathbb{R}.$$

Squaring (3.27) and rearranging terms gives

$$y'(x)^2 = \frac{\gamma(x,\beta)^2(1-e(x)^2)}{1-\gamma(x,\beta)^2}. \tag{3.28}$$

Using the fact that $y'(x)$ and $\gamma(x,\beta)$ have the same sign, we can take square roots in (3.28) to obtain

$$y'(x) = G(x,\beta),$$

where

$$G(x,\beta) := \frac{\gamma(x,\beta)\sqrt{1-e(x)^2}}{\sqrt{1-\gamma(x,\beta)^2}}. \tag{3.29}$$

The fundamental theorem of calculus gives

$$y(x) = C + \int_0^x G(t,\beta)\,dt \quad \text{for all } x \in [0,b].$$

Now $y(0) = 0$ implies $C = 0$ and consequently

$$y(b) = B \Rightarrow \int_0^b G(t, \beta)\, dt = B.$$

The solution to $(\text{E-L})_1$ satisfying the boundary conditions is therefore given by

$$y(x) = \int_0^x G(t, \beta)\, dt \quad \text{for all } x \in [0, b],$$

where G is defined by (3.29) and $\beta \in \mathbb{R}$ is chosen to satisfy

$$\int_0^b G(t, \beta)\, dt = B.$$

3.6 C^1-Theory

We will now investigate conditions that must be satisfied by minimizers (or maximizers) under weaker smoothness assumptions on the integrand and on the admissible functions. In particular we will now assume that f and y have continuous first-order derivatives, but we do not make any assumptions regarding second-order derivatives. By examining the argument presented in Sections 3.1 and 3.3, we see that the main obstacle is that the integration by parts that was performed in (3.4) is no longer valid under the relaxed assumptions of the C^1-theory. The key idea for overcoming this obstacle is to integrate the other term in (3.1) by parts. An appropriate analogue of Lagrange's Lemma will lead us again to the first Euler-Lagrange equation. As in Section 3.3, we will deal with several types of problems simultaneously.

Let $a, b, A, B \in \mathbb{R}$ with $a < b$ be given. Put

$$\mathfrak{X} := C^1[a, b].$$

Let $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a given function with continuous first-order partial derivatives. Consider the functional $J : \mathfrak{X} \to \mathbb{R}$ defined by

$$J(y) := \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathfrak{X}.$$

We define several types of classes of admissible functions and the corresponding admissible variations. Put

$$\begin{aligned}
\mathrm{Y}_{a,b} &:= \left\{y \in C^1[a,b] \;:\; y(a) = A \text{ and } y(b) = B\right\};\\
\mathrm{Y}_a &:= \left\{y \in C^1[a,b] \;:\; y(a) = A\right\};\\
\mathrm{Y}_b &:= \left\{y \in C^1[a,b] \;:\; y(b) = B\right\};\\
\mathrm{Y} &:= C^1[a,b];\\
\mathrm{V}_{a,b} &:= \left\{v \in C^1[a,b] \;:\; v(a) = v(b) = 0\right\};\\
\mathrm{V}_a &:= \left\{v \in C^1[a,b] \;:\; v(a) = 0\right\};\\
\mathrm{V}_b &:= \left\{v \in C^1[a,b] \;:\; v(b) = 0\right\};\\
\mathrm{V} &:= C^1[a,b].
\end{aligned}$$

These are essentially the same classes given in Section 3.3 – the only difference being that they are no longer subsets of $\mathscr{C}^2[a,b]$ but subsets of $\mathscr{C}^1[a,b]$ instead.

For now, we focus on the problem of minimizing J over $\mathrm{Y}_{a,b}$ – this is the problem with both endpoints fixed. Later we will discuss problems with free endpoints. Let $y_* \in \mathrm{Y}_{a,b}$ be given and suppose that J attains a minimum at y_*

For each $v \in \mathrm{V}_{a,b}$, Theorem 2.4 tells us that the Gâteaux variation of J at y_* in the direction v exists and gives us a formula for $\delta J(y; v)$. Moreover, by Theorem 2.2, $\delta J(y_*; v)$ must be zero for each $v \in \mathrm{V}$ Therefore, we have

$$\delta J(y_*; v) = \int_a^b \left\{ f_{,2}(x, y_*(x), y_*'(x))v(x) + f_{,3}(x, y_*(x), y_*'(x))v'(x) \right\} dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}. \tag{3.30}$$

Define $F, G : [a,b] \to \mathbb{R}$ by

$$F(x) := f_{,2}(x, y_*(x), y_*'(x)) \text{ and } G(x) := f_{,3}(x, y_*(x), y_*'(x)) \quad \text{for all } x \in [a,b]. \tag{3.31}$$

and notice that $F, G \in \mathscr{C}[a,b]$ by virtue of our smoothness assumptions on f and y. We may now rewrite (3.30) as

$$\int_a^b \left\{ F(x)v(x) + G(x)v'(x) \right\} dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}. \tag{3.32}$$

In Section 3.1, we knew that G was continuously differentiable and we were able to integrate the second term in (3.32) by parts. Now, instead, we know only that F and G are continuous. To proceed from (3.32), we will integrate the first term by parts.

This idea is due to du Bois-Reymond. Define $H \in \mathscr{C}^1[a,b]$ by

$$H(x) := \int_a^x F(t)\, dt \quad \text{for all } x \in [a,b]. \tag{3.33}$$

Observe that $H' = F$, so (3.32) becomes

$$\int_a^b \Big\{ H'(x)v(x) + G(x)v'(x) \Big\}\, dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}.$$

Integrating the term involving H by parts yields

$$[H(x)v(x)]\Big|_a^b - \int_a^b H(x)v'(x)\, dx + \int_a^b G(x)v'(x)\, dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}. \tag{3.34}$$

Since $v \in \mathrm{V}_{a,b}$ implies $v(a) = v(b) = 0$, the first term in (3.34) is zero. Thus

$$\int_a^b \Big\{ G(x) - H(x) \Big\} v'(x)\, dx = 0 \quad \text{for all } v \in \mathrm{V}_{a,b}. \tag{3.35}$$

At this point, we need to use a lemma of du Bois-Reymond.

Lemma 3.3 (du Bois-Reymond) *Let $a, b \in \mathbb{R}$ with $a < b$ and $w \in C[a,b]$ be given. Put*

$$\mathscr{V} := \left\{ v \in \mathscr{C}^1[a,b] \mid v(a) = v(b) = 0 \right\}.$$

Assume that $\int_a^b w(x)v'(x)\, dx = 0$ for all $v \in \mathscr{V}$. Then there exists a $c \in \mathbb{R}$ such that $w(x) = c$ for all $x \in [a,b]$.

Proof. As in the proof of Lagrange's Lemma, the key idea involves a judicious choice of $v \in \mathscr{V}$. Notice that the only possible c that can work is the average value of w over the interval $[a,b]$.

Put

$$c_* := \frac{1}{b-a} \int_a^b w(t)\, dt.$$

We will show that $w(x) = c_*$ at each $x \in [a,b]$. It suffices to establish

$$\int_a^b [w(x) - c_*]^2\, dx = 0. \tag{3.36}$$

Since $\int_a^b w(x)v'(x)\, dx = 0$ for every $v \in \mathscr{V}$, we see that

$$\begin{aligned} \int_a^b [w(x) - c_*]v'(x)\, dx &= \int_a^b w(x)v'(x)\, dx - c_*[v(x)]\Big|_a^b \\ &= \int_a^b w(x)v'(x)\, dx = 0 \quad \text{for all } v \in \mathscr{V}. \end{aligned}$$

To obtain (3.36), we would like to choose $v \in \mathscr{V}$ such that $v'(x) = w(x) - c_*$ at each $x \in [a,b]$. Let us define $v_* \in C^1[a,b]$ by

$$v_*(x) := \int_a^x [w(t) - c_*]\, dt.$$

Notice that $v_*'(x) = w(x) - c_*$ at each $x \in [a,b]$, so our proof will be complete if we show that $v_* \in \mathscr{V}$. In other words, it only remains to show $v_*(a) = v_*(b) = 0$. By its definition, we have $v_*(a) = 0$ and

$$\begin{aligned} v_*(b) &= \int_a^b [w(t) - c_*]\, dx = \int_a^b w(t)\, dt - \int_a^b c_*\, dt = \int_a^b w(t) - (b-a)c_* \\ &= \int_a^b w(t)\, dt - \int_a^b w(t)\, dt = 0. \end{aligned}$$

Thus we find $v_* \in \mathscr{V}$ and therefore (3.36) must hold. It follows that $w(x) = c_*$ at each $x \in [a,b]$ and this proves the lemma. □

With du Bois-Reymond's Lemma in hand, we now return to (3.35). It follows that there exists some $c \in \mathbb{R}$ such that

$$G(x) - H(x) = c \Rightarrow G(x) = H(x) + c \quad \text{for all } x \in [a,b].$$

By definition (3.33), we see that $H + c$ is continuously differentiable. Hence G is continuously differentiable, and we have

$$G'(x) = H'(x) = F(x) \quad \text{for all } x \in [a,b]. \tag{3.37}$$

Now by definition (3.31), we have that the function $x \mapsto f_{,3}(x, y(x), y'(x))$ is continuously differentiable on $[a,b]$ and

$$f_{,2}(x, y_*(x), y_*'(x)) = \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x))\right] \quad \text{for all } x \in [a,b] \tag{E-L$_1$}$$

For future reference, we restate our conclusion in (3.37) as a lemma.

Lemma 3.4 *Let $a, b \in \mathbb{R}$ with $a < b$ and $F, G \in C[a,b]$ be given. Put*

$$\mathscr{V} := \left\{ v \in C^1[a,b] \,:\, v(a) = v(b) = 0 \right\}.$$

Assume that

$$\int_a^b \Big\{ F(x)v(x) + G(x)v'(x) \Big\}\, dx = 0 \quad \text{for all } v \in \mathscr{V}.$$

Then, we have $G \in C^1[a,b]$ and $G'(x) = F(x)$ at each $x \in [a,b]$.

3.6.1 Problems with Free Endpoints

Continuing our development of the $\mathscr{C}^1$-theory for problems in the calculus of variations, we state a corresponding version of Lemma 3.4 for the case when $\mathscr{V}$ is replaced by V_a, V_b or V.

Lemma 3.5 *Let $a, b \in \mathbb{R}$ with $a < b$ and $F, G \in C[a, b]$ be given.*

(i) If $\int_a^b \{F(x)v(x) + G(x)v'(x)\} = 0$ for every $v \in \mathrm{V}_a$, then $G(b) = 0$, $G \in C^1[a, b]$ and $G'(x) = F(x)$ at each $x \in [a, b]$.

(ii) If $\int_a^b \{F(x)v(x) + G(x)v'(x)\} = 0$ for every $v \in \mathrm{V}_b$, then $G(a) = 0$, $G \in \mathscr{C}^1[a, b]$ and $G'(x) = F(x)$ at each $x \in [a, b]$.

(iii) If $\int_a^b \{F(x)v(x) + G(x)v'(x)\} = 0$ for every $v \in \mathrm{V}$, then $G(a) = G(b) = 0$, $G \in C^1[a, b]$ and $G'(x) = F(x)$ at each $x \in [a, b]$.

Proof. We only prove (i), the proof for the others being similar (see Section 3.3).

Therefore, we are assuming that

$$\int_a^b \Big\{F(x)v(x) + G(x)v'(x)\Big\} = 0 \quad \text{for all } v \in \mathrm{V}_a. \tag{3.38}$$

Since $\mathrm{V}_{a,b} \subset \mathrm{V}_a$, we may use Lemma 3.4 to conclude that $G \in C^1[a, b]$ and $G'(x) = F(x)$ at each $x \in [a, b]$.

Once we know that $G \in \mathscr{C}^1[a, b]$, our argument can follow that given in Section 3.3. We integrate the second term in (3.38) by parts and write

$$\begin{aligned}\int_a^b \Big\{F(x)v(x) + G(x)v'(x)\Big\}\, dx = &[G(x)v(x)]\Big|_a^b + \int_a^b \Big\{F(x) - \frac{d}{dx}[G(x)]\Big\} v(x)\, dx \\ =&G(b)v(b) - G(a)v(a) \\ =&-G(a)v(a) = 0 \quad \text{for all } v \in \mathrm{V}_a.\end{aligned}$$

By choosing a $v \in \mathrm{V}_a$ such that $v(a) = 1$, we deduce that $G(a) = 0$.

So $G(a) = 0$ and $G \in C^1[a, b]$ with $G'(x) = F(x)$ at each $x \in [a, b]$. The lemma is proved. □

Recall that $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is assumed to have continuous first partial derivatives, and our functional was given by

$$J(y) = \int_a^b f(x, y(x), y'(x))\, dx \quad \text{for all } y \in \mathfrak{X},$$

where $\mathfrak{X} = C^1[a,b]$. The first Euler-Lagrange equation for J is

$$f_{,2}(x, y_*(x), y_*'(x)) = \frac{d}{dx}\left[f_{,3}(x, y_*(x), y_*'(x))\right] \quad \text{for all } x \in [a,b], \qquad \text{(E-L)}_1$$

and the natural boundary conditions at a and b, respectively, are

$$f_{,3}(a, y_*(a), y_*'(a)) = 0 \qquad \text{(NBC)}_a$$

and

$$f_{,3}(b, y_*(b), y_*'(b)) = 0. \qquad \text{(NBC)}_b$$

Using Lemma 3.5, we deduce the following conclusions:

(both ends pinned) If y_* minimizes J over $\mathrm{Y}_{a,b}$,
then y_* must satisfy (E-L)_1, $y_*(a) = A$ and $y_*(b) = B$.

(free end at a) If y_* minimizes J over Y_b,
then y_* must satisfy (E-L)_1, (NBC)_a and $y_*(b) = B$.

(free end at b) If y_* minimizes J over Y_a,
then y_* must satisfy (E-L)_1, $y_*(a) = A$ and (NBC)_b.

(both ends free) If y_* minimizes J over Y,
then y_* must satisfy (E-L)_1, (NBC)_a and (NBC)_b.

We remark that if y_* minimizes (or maximizes) J over one of the sets $\mathrm{Y}_{a,b}$, Y_a, Y_b or Y, then there exists a $c \in \mathbb{R}$ such that

$$f(x, y_*(x), y_*'(x)) - y_*'(x) f_{,3}(x, y_*(x), y_*'(x)) = c + \int_a^x f_{,1}(t, y_*(t), y_*'(t))\, dt \quad \text{for all } x \in [a,b], \qquad \text{(E-L)}_2$$

but the proof is completely elementary.

3.6.2 Example 3.6.2 (cf. Examples 1.2, 3.1.1 and 3.3.1)

Let us consider an example. Put

$$\mathscr{Y} := \left\{ y \in C^1[0,1] \,:\, y(0) = 0 \text{ and } y(1) = 1 \right\}.$$

We want to minimize the functional $J : \mathscr{Y} \to \mathbb{R}$ given by

$$J(y) := \int_0^1 \left[y(x)^2 + y'(x)^2 \right] dx \quad \text{for all } y \in \mathscr{Y}$$

over the class $\mathscr{Y}$. This functional was also used in Sections 1.2, 3.1.1 and 3.3.1.

By our results in the previous section, if J attains a minimum at $y \in \mathscr{Y}$, then $2y'$ must be continuously differentiable and y must satisfy

$$2y(x) = \frac{d}{dx}\left[2y'(x)\right] \quad \text{for all } x \in [0,1]. \tag{E-L$_1$}$$

Therefore, a minimizer for J over $\mathscr{Y}$ must satisfy

$$\begin{cases} y''(x) = y(x); \\ y(0) = 0 \text{ and } y(1) = 1. \end{cases} \tag{3.39}$$

Notice that from Lemma 3.4 one concludes that $y \in C^2[0,1]$. In fact, if y is a minimizer for J, then y must be in $C^\infty[0,1]$: since y satisfies (3.39) and $y \in C^2[0,1]$, we find $y'' \in C^2[0,1]$, and thus $y \in C^4[0,1]$. This in turn implies $y \in C^6[0,1]$. By induction, we find $y \in C^\infty[0,1]$. So even though our admissible class $\mathscr{Y}$ includes functions that may have only one continuous derivative, a minimizer for J must actually have continuous derivatives of all orders.

3.6.3 Minimizers Might Not be in $\mathbf{C^2}$

We remark that it may not always be possible to conclude that a C^1-minimizer for a functional has a continuous second derivative. It is true that a minimizer for a functional J with integrand f must satisfy

$$f_{,2}(x, y(x), y'(x)) = \frac{d}{dx}\left[f_{,3}(x, y(x), y'(x))\right] \quad \text{for all } x \in [a,b], \tag{E-L$_1$}$$

and by Lemma 3.4, we know that the derivative on the right hand side of (E-L)$_1$ exists and is continuous. It is the composition $x \mapsto f_{,3}(x, y(x), y'(x))$, however, that is continuously differentiable. Without more information regarding $f_{,3}$, Lemma 3.4 tells us nothing about the differentiability of y'.

Indeed, there are many examples of functions $f, g : \mathbb{R} \to \mathbb{R}$ such that the composition $x \mapsto f(g(x))$ is continuously differentiable, but g itself is not differentiable.

The following example has a C^1-minimizer that is not of class C^2. Let

$$\mathscr{Y} = \left\{ y \in C^1[0,1] \,:\, y(0) = 0,\, y(0) = 1 \right\}$$

and define $J : \mathscr{Y} \to \mathbb{R}$ by

$$J(y) := \int_0^1 (4y'(x)^2 - 9x)^2 \, dx \quad \text{for all } y \in \mathscr{Y},$$

and notice that $J(y) \geq 0$ for all $y \in \mathscr{Y}$. Is is straightforward to check that $J(y_*) = 0$ for the function $y_* \in \mathscr{Y}$ given by $y_*(x) = x^{3/2}$., and consequently y_* minimizes J on $\mathscr{Y}$. Notice that $y_* \notin C^2[0,1]$.

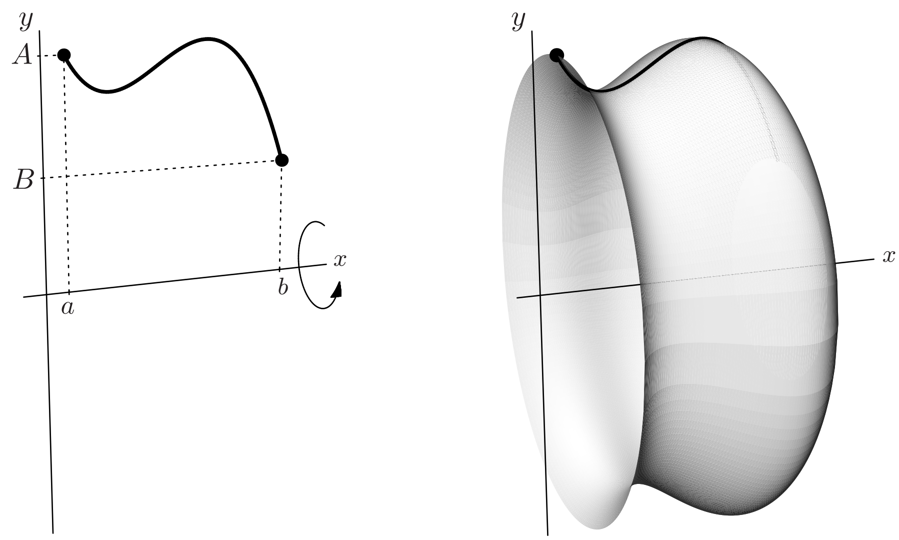

Figure 3.5: A surface of revolution

3.7 Minimal Surface of Revolution Problem

Let $a, b, A, B \in \mathbb{R}$ with $a < b$ and $A, B > 0$ be given and put

$$\mathscr{Y} := \left\{ y \in C^1[a,b] \,:\, y(a) = A \,, y(b) = B \,, y(x) > 0 \quad \text{for all } x \in [a,b] \right\}.$$

Given a function $y \in \mathscr{Y}$, we consider the surface generated by revolving the graph of y about the x-axis. The area of such a surface is

$$S = 2\pi \int_a^b y(x)\sqrt{1 + y'(x)^2}\, dx.$$

Our problem is to minimize, over $\mathscr{Y}$, the functional $J : \mathscr{Y} \to \mathbb{R}$ defined by

$$J(y) := \int_a^b y(x)\sqrt{1 + y'(x)^2}\, dx \quad \text{for all } y \in \mathscr{Y}.$$

We begin by observing that the theory that we have discussed so far does not apply directly to this problem because of the pointwise constraint $y(x) > 0$ in the definition of $\mathscr{Y}$. However, since $A, B > 0$, it is not difficult to show that for each $y \in \mathscr{Y}$, the class of admissible variations at y is

$$\mathscr{V}_y = \left\{ v \in C^1[a,b] \,:\, v(a) = v(b) = 0 \right\}.$$

(See Exercise) Therefore, the derivation $(\text{E-L})_1$ given in Section still applies. The integrand $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is

$$f(x, y, z) := y(1 + z^2)^{\frac{1}{2}} \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R} \times \mathbb{R},$$

so that

$$f_{,2}(x, y, z) = (1 + z^2)^{\frac{1}{2}} \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R} \times \mathbb{R}$$

and

$$f_{,3}(x, y, z) = \frac{yz}{(1 + z^2)^{\frac{1}{2}}} \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R} \times \mathbb{R}.$$

Thus the first Euler-Lagrange equation is

$$(1 + y'(x)^2)^{\frac{1}{2}} = \frac{d}{dx}\left[\frac{y(x)y'(x)}{(1 + y'(x)^2)^{\frac{1}{2}}}\right] \quad \text{for all } x \in [a, b]. \tag{E-L$_1$}$$

Although we do not know *a priori* that a minimizer will belong to $C^2[a, b]$, we do know that if y is a minimizer then the mapping

$$x \mapsto \frac{y(x)y'(x)}{(1 + y'(x)^2)^{\frac{1}{2}}}$$

is continuously differentiable on [a,b]. This fact can be used to show that if y is a minimizer, then $y \in C^2[a, b]$. (See Exercise .) Once it is known that $y \in C^2[a, b]$, $(\text{E-L})_1$ can be simplified to

$$y''(x)y(x) = 1 + y'(x)^2 \quad \text{for all } x \in [a, b]. \tag{3.40}$$

(See Exercise .)

It is also useful to look at the second Euler-Lagrange equation. Since the integrand for J has no explicit dependence on x, the second Euler-Lagrange equation reduces to

$$y(x)(1 + y'(x)^2)^{\frac{1}{2}} - \frac{y(x)y'(x)^2}{(1 + y'(x)^2)^{\frac{1}{2}}} = c_1 \quad \text{for all } x \in [a, b] \tag{E-L$_2$}$$

for some constant $c_1 \in \mathbb{R}$. Upon putting the left hand side of $(\text{E-L})_2$ over a common denominator, we find that y satisfies $(\text{E-L})_2$ if

$$\frac{y(x)}{(1 + y'(x)^2)^{\frac{1}{2}}} = c_1 \quad \text{for all } x \in [a, b]. \tag{3.41}$$

At this point we observe that the denominator in (3.41) is always strictly positive and therefore the following hold:

(1) if $c_1 = 0$, then $y(x) = 0$ at each $x \in [a, b]$;

(2) if $c_1 < 0$, then $y(x) < 0$ at each $x \in [a, b]$;

(3) if $c_1 > 0$, then $y(x) > 0$ at each $x \in [a, b]$.

Consequently, if y satisfies $(\text{E-L})_2$, then y must have the same sign throughout the interval $[a, b]$. Since we seek a solution with $y(a), y(b) > 0$, we assume that $c_1 > 0$. This tells us that $y(x) > 0$ for all $x \in [a, b]$. Using (3.40), we conclude that $y''(x) > 0$ for all $x \in [a, b]$. Using (3.41) and the fact that $c_1 > 0$ we find that $y(x) \geq c_1$ for all $x \in [a, b]$.

It follows from (3.41) that

$$y'(x)^2 = \left(\frac{y(x)}{c_1}\right)^2 - 1 \quad \text{for all } x \in [a, b], \tag{3.42}$$

which suggests a hyperbolic-cosine substitution for y. Using the properties that we have established about a minimizer y, it is possible to show that there is a function $u \in C^1[a, b]$ such that

$$y(x) = c_1 \cosh u(x) := \frac{c_1}{2}\left(e^{u(x)} + e^{-u(x)}\right) \quad \text{for all } x \in [a, b]. \tag{3.43}$$

(See Problem .) It follows that

$$y'(x) = c_1(\sinh u(x))u'(x) = \frac{c_1}{2}\left(e^{u(x)} - e^{-u(x)}\right)u'(x) \quad \text{for all } x \in [a, b]. \tag{3.44}$$

Substituting (3.43) and (3.44) into (3.41) and using the identity $\cosh^2 z - 1 = \sinh^2 z$ we find that

$$(c_1 u'(x))^2 = 1 \quad \text{for all } x \in [a, b],$$

which gives

$$u'(x) = \pm\frac{1}{c_1} \quad \text{for all } x \in [a, b].$$

Since u' is continuous on $[a, b]$, we must have either

$$u(x) = \frac{x + c_2}{c_1} \quad \text{for all } x \in [a, b],$$

or

$$u(x) = \frac{-x + c_2}{c_1} \quad \text{for all } x \in [a, b],$$

for some constant c_2. Thus we have

$$y(x) = c_1 \cosh\left(\frac{x + c_2}{c_1}\right) \quad \text{for all } x \in [a, b],$$

or

$$y(x) = c_1 \cosh\left(\frac{-x + c_2}{c_1}\right) \quad \text{for all } x \in [a, b].$$

Since the hyperbolic cosine is an even function (and c_2 is unspecified), we can ignore the second formula for y. (Indeed, if we replace c_2 by $-c_2$ in the second formula for y, we get the first formula.)

It is straightforward to check that the functions we have just found do indeed satisfy $(\text{E-L})_1$ for our functional. Moreover, we have found all possible solutions to $(\text{E-L})_1$ with $y(x) > 0$ for all $x \in [a, b]$, since every C^2-solution of $(\text{E-L})_1$ is also a solution of $(\text{E-L})_2$. In other words, our argument above proves that a function $y \in C^1[a, b]$ with $y(x) > 0$ for all $x \in [a, b]$ is a solution of $(\text{E-L})_1$ if and only if y is of the form

$$y(x) = c_1 \cosh\left(\frac{x + c_2}{c_1}\right) \quad \text{for all } x \in [a, b] \tag{3.45}$$

for some $c_1, c_2 \in \mathbb{R}$ with $c_1 > 0$.

Thus to find a solution of $(\text{E-L})_1$ that is in $\mathscr{Y}$, it only remains to choose $c_1, c_2 \in \mathbb{R}$ with $c_1 > 0$ such that

$$A = c_1 \cosh\left(\frac{a + c_2}{c_1}\right) \text{ and } B = c_1 \cosh\left(\frac{b + c_2}{c_1}\right). \tag{3.46}$$

This is a system of two (nonlinear) equations with two unknowns. Whether or not it is possible to choose c_1, c_2 satisfying (3.46) depends upon the values of a, A, b, B. It turns out that in some cases there is a unique choice, but in other cases there is either no way or two ways in which to choose c_1 and c_2.

Let us now make an observation. Consider the parametric curve $(x, y) : [0, 1] \to \mathbb{R}^2$ given by

$$(x(t), y(t)) := \begin{cases} \left(a, 3A\left(\frac{1}{3} - t\right)\right), & 0 \le t < \frac{1}{3}; \\ \left(a + 3(b - a)\left(t - \frac{1}{3}\right), 0\right), & \frac{1}{3} \le t < \frac{2}{3}; \\ \left(b, 3B\left(t - \frac{2}{3}\right)\right), & \frac{2}{3} \le t \le 1 \end{cases}$$

(which is not the graph of a function). The surface of revolution generated by this curve consists of a disk with a radius A, a disk with radius B and a line joining the centers of the two disks. The area of this surface is simply $\left(A^2 + B^2\right)\pi$. Although, the parametric curve given above is not in $\mathscr{Y}$, we can approximate it using functions from $\mathscr{Y}$; moreover, we can choose our approximations so that the associated surface areas approach $\left(A^2 + B^2\right)\pi$. For example, for each $n \in \mathbb{N}$ define $y_n \in \mathscr{Y}$ by

$$y_n(x) := A\frac{(b - x)^n}{(b - a)^n} + B\frac{(x - a)^n}{(b - a)^n} \quad \text{for all } x \in [a, b].$$

As n gets large, the functions y_n closely approximate the parametric curve, and the area of the surface generated by y_n becomes very close to the sum of the areas for the two disks with radii A and B. What this tells us is that the minimal surface area must be at least as small as $\left(A^2 + B^2\right)\pi$. That is, if there is a minimal value for J, then it cannot be larger than $\frac{1}{2}\left(A^2 + B^2\right)$.

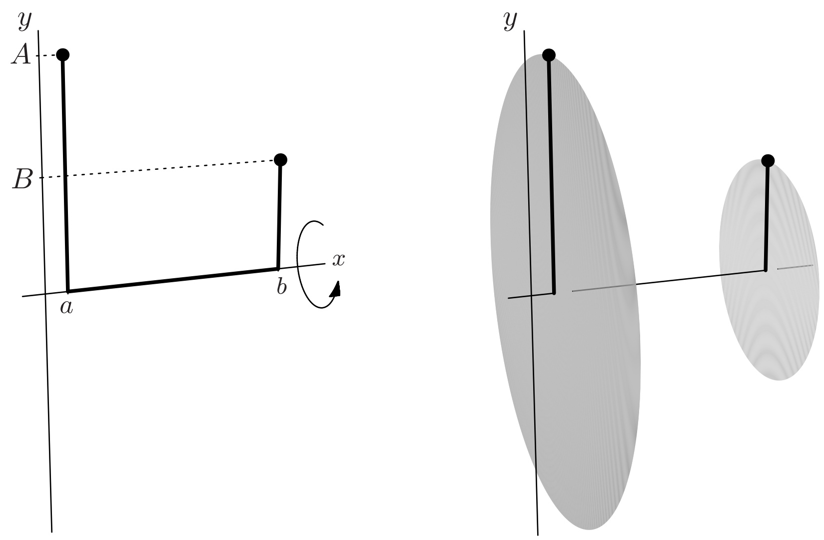

Figure 3.6: Surface of revolution generated by $(x(t), y(t))$

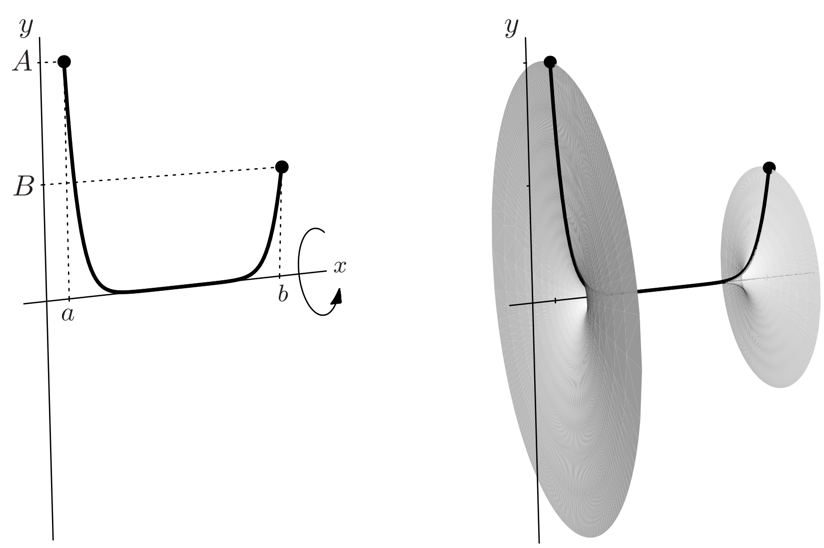

Figure 3.7: Surface of revolution generated by y_n