Chapter 4

Lagrange Multipliers

We now turn our attention to problems where the admissible functions are required to lie on a level set of a constraining functional. Unless the constraining functional is linear, the set over which we wish to optimize will not lead to classes of admissible variations that are large enough to be helpful. The method of Lagrange multipliers provides a flexible and straightforward way to deal with these types of problems. We will develop the method within a sufficiently abstract framework so that it can be applied to many different kinds of problems.

4.1 Too Few Admissible Variations

In this section we investigate a simple (and natural) situation where there are simply not enough admissible variations to provide *any* information about maximizers or minimizers. Let $\mathfrak{X} = C^1[0,\pi]$ and put

$$\mathscr{S} = \{y \in C^1[0,\pi] \,:\, y(0) = y(\pi) = 0 \;,\; \int_0^\pi y(x)^2\,dx = 1\}.$$

For definiteness, suppose that we wish to minimize the functional $J : \mathscr{S} \to \mathbb{R}$ defined by

$$J(y) = \int_0^\pi y'(x)^2\,dx \quad \text{for all } y \in \mathscr{S}.$$

In order to apply the approach of Chapter 2, we need to identify the class of $\mathscr{S}$-admissible variations for each $y \in \mathscr{S}$. To this end, let $y \in \mathscr{S}, v \in \mathfrak{X}$ be given be given. In order to have $v \in \mathscr{V}_y$, it is necessary (and sufficient) to have an open interval I with $0 \in I$ such that $[y + \varepsilon v] \in \mathscr{S}$ for all $\varepsilon \in I$. Suppose that such an interval I exists. Then we must have

$$\begin{aligned}\int_0^\pi [y(x)+\varepsilon v(x)]^2\,dx &= \int_0^\pi y(x)^2\,dx + 2\varepsilon\int_0^\pi y(x)v(x)\,dx + \varepsilon^2\int_0^\pi v(x)^2\,dx \\ &= 1 + 2\varepsilon\int_0^\pi y(x)v(x)\,dx + \varepsilon^2\int_0^\pi v(x)^2\,dx \\ &= 1.\end{aligned}$$

It follows easily that

$$2\int_0^\pi y(x)v(x)\,dx + \varepsilon\int_0^\pi v(x)^2 = 0 \quad \text{for all } \varepsilon\in I\backslash\{0\}, \tag{4.1}$$

which implies that

$$\int_0^\pi v(x)^2\,dx = 0. \tag{4.2}$$

(Indeed, if the equality in (4.1) holds for two distinct values of ε, we can immediately conclude that (4.2) holds.) It follows from (4.2) that $v(x)=0$ for all $x\in[0,\pi]$. Therefore we have $\mathscr{V}_y=\{0\}$ for all $y\in\mathscr{S}$. Since $\delta J(y;0)=0$ for every $y\in\mathscr{S}$, we can conclude nothing from the fact that the Gâteaux variations must vanish for all admissible variations.

The idea will be use the larger domain $\mathscr{Y}=\{y\in C^1[0,\pi] : y(0)=y(\pi)=0\}$, for which there is a rich class of admissible variations, and consider two functionals $J,G:\mathscr{Y}\to\mathbb{R}$ defined by

$$J(y)=\int_0^\pi y'(x)^2\,dx \quad\text{and}\quad G(y)=\int_0^\pi y(x)^2\,dx \quad \text{for all } y\in\mathscr{Y}.$$

Notice that

$$\mathscr{S}=\{y\in\mathscr{Y}\mid G(y)=1\}.$$

By studying the behavior of the pair (J,G) of functionals on $\mathscr{Y}$, we can draw some very useful conclusions about maxima and minima of J on $\mathscr{S}$.

4.2 The Lagrange Multiplier Method in Real Linear Spaces

Let $\mathfrak{X}$ be a real linear space and $\mathscr{Y}\subset\mathfrak{X}$, $J,G:\mathscr{Y}\to\mathbb{R}$, $c\in\mathbb{R}$ be given. Define

$$\mathscr{S}:=\{y\in\mathscr{Y}\mid G(y)=c\}.$$

If y is a member of $\mathscr{S}$, then in addition to being in $\mathscr{Y}$, it must satisfy the constraint $G(y)=c$. Our problem is to minimize (or maximize) J on $\mathscr{S}$. If the functional G is nonlinear, then the class of $\mathscr{S}$-admissible variations may be too small to be useful, as the example in Section 4.1 shows.

Rather than focusing on the class $\mathscr{S}$ as being the domain of the functional to be minimized, we think about the pair (J, G) of functionals with domain $\mathscr{Y}$. We define $F : \mathscr{Y} \to \mathbb{R}^2$ by

$$F(y) := (J(y), G(y)) \quad \text{for all } y \in \mathscr{Y}.$$

Suppose that $y_* \in \mathscr{S}$ is a minimizer for J over $\mathscr{S}$, i.e. $J(y) \geq J(y_*)$ for every $y \in \mathscr{S}$. Let us investigate what this tells about F. For one thing, since $y_* \in \mathscr{S}$, we have

$$F(y_*) = (J(y_*), c).$$

Now if $J(y) \geq J(y_*)$ for every $y \in \mathscr{S}$, then there cannot be any $y \in \mathscr{Y}$ such that

$$F(y) = (\alpha, c)$$

with $\alpha < J(y_*)$. Another way of saying this is that for every $\alpha < J(y_*)$, the point (α, c) is not in the range of F.

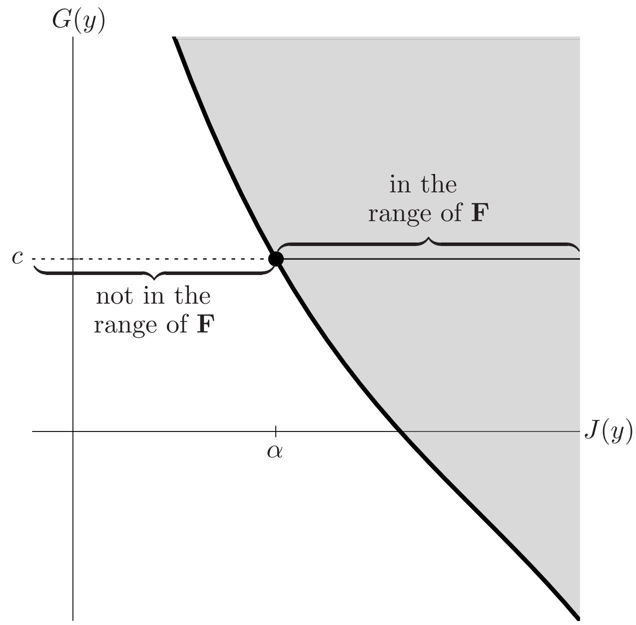

Figure 4.1: (α, c) is in the boundary of the range of F

This means that there is a point $z_* = F(y_*)$ in the range of F for which there are points arbitrarily close to z_* that are not in the range of F. (In other words, $F(y_*)$ is not an interior point of the range of F.) This property has important implications regarding F. In order to take advantage of this observation, we will make use of the following

Lemma 4.1 *Let $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuously differentiable mapping, and let $x_* \in \mathbb{R}^2$ be given. Suppose that* $\det \nabla\Phi(x_*) \neq 0$. *Then there exists $\delta > 0$ such that*

$$\left\{z \in \mathbb{R}^2 \,:\, \|z - \Phi(x_*)\| < \delta\right\} \subset \mathrm{Range}(\Phi).$$

In order to use Lemma 4.1 it will be convenient to employ a pair of admissible directions, so that we can use F to create a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$. Given $y \in \mathscr{Y}$, we want $u, w \in \mathfrak{X}$ such that

$$y + \eta u + \mu w \in \mathscr{Y}$$

for all (η, μ) in some neighborhood of $(0, 0)$ in $\mathbb{R}^2$. In order to highlight the important ideas, we shall treat in detail a vety important case in which the situation regarding admissible variations is rather straightforward. For simplicity we assume that $\mathscr{V}$ is a subspace of $\mathfrak{X}$ such that

$$y + v \in \mathscr{Y} \quad \text{for all } y \in \mathscr{Y} \text{ and } v \in \mathscr{V}. \tag{4.3}$$

This implies that for every $y \in \mathscr{Y}$ we have

$$y + \eta u + \mu w \in \mathscr{Y} \quad \text{for all } u, w \in \mathscr{V} \text{ and } \eta, \mu \in \mathbb{R}. \tag{4.4}$$

Observe that if $\mathscr{V}$ is the 0 subspace then (4.3) is automatically, but this is of no use. However, in situations when $\mathscr{Y}$ is characterized by linear constraints, then the set $\mathscr{V}_y$ of admissible variations is a subspace that is independent of $\mathscr{Y}$ and this common subspace $\mathscr{V}$ satisfies (4.3) (and hence also (4.4)).

To use Lemma 4.1, let $u, w \in \mathscr{V}$ be given and define $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\Phi(\eta, \mu) := (J(y_* + \eta u + \mu w), G(y_* + \eta u + \mu w)) \quad \text{for all } (\eta, \mu) \in \mathbb{R}^2. \tag{4.5}$$

Notice that $\Phi(0, 0) = (J(y_*), G(y_*)) = (J(y_*), c)$.

Let us assume that Φ is continuously differentiable. Then from our discussion above, there is no neighborhood of $\Phi(0, 0)$ that is contained in the Range(Φ), since y_* minimizes J on $\mathscr{S}$. In other words, for every $\delta > 0$

$$\left\{(\eta, \mu) \in \mathbb{R}^2 \,:\, \Phi(\eta, \mu) - \Phi(0, 0)\| < \delta\right\} \not\subset \mathrm{Range}(\Phi).$$

Hence, Lemma 4.1 implies that $\det \nabla\Phi(0, 0) = 0$. So if y_* is a minimizer for J over $\mathscr{S}$, then $\det \nabla\Phi(0, 0) = 0$.

We thus have an important condition on the gradient of Φ at the point $(0, 0)$. Let us now determine $\nabla\Phi(0, 0)$. By definition

$$\Phi(\eta, \mu) = (\Phi_1(\eta, \mu), \Phi_2(\eta, \mu)) = (J(y_*+\eta u+\mu w), G(y_*+\eta u+\mu w)) \quad \text{for all } (\eta, \mu) \in \mathbb{R}^2.$$

Computing the partial derivatives of the components of Φ with respect to η at $(0, 0)$, we find that

$$\Phi_{1,1}(0, 0) = \frac{\partial}{\partial \eta}\left[J(y_* + \eta u + \mu w)\right]\Big|_{(\eta,\mu)=(0,0)} = \frac{\partial}{\partial \eta}\left[J(y_* + \eta u)\right]\Big|_{\eta=0} = \delta J(y_*; u)$$

and

$$\Phi_{2,1}(0,0) = \frac{\partial}{\partial\eta}\left[G(y_* + \eta u + \mu w)\right]\Big|_{(\eta,\mu)=(0,0)} = \frac{\partial}{\partial\eta}\left[G(y_* + \eta u)\right]\Big|_{\eta=0} = \delta G(y_*;u).$$

Similarly, the partial derivatives of the components of Φ with respect to μ at $(0,0)$ are given by

$$\Phi_{1,2}(0,0) = \delta J(y_*;w)$$

and

$$\Phi_{2,2}(0,0) = \delta G(y_*;w).$$

Consequently, we have

$$\nabla\Phi(0,0) = \begin{pmatrix} \delta J(y_*;u) & \delta J(y_*;w) \\ \delta G(y_*;u) & \delta G(y_*;w) \end{pmatrix}.$$

It follows that

$$\det \nabla\Phi(0,0) = \delta J(y_*;u)\delta G(y_*;w) - \delta J(y_*;w)\delta G(y_*;u).$$

Therefore, if y_* minimizes J over $\mathscr{S}$, we must have

$$\delta J(y_*;u)\delta G(y_*;w) - \delta J(y_*;w)\delta G(y_*;u) = 0 \quad \text{for all } u, w \in \mathscr{V}. \tag{4.6}$$

There are only two ways that (4.6) can hold:

(1) either $\delta G(y_*;v) = 0$ for all $v \in \mathscr{V}$, or

(2) there exists a $\lambda \in \mathbb{R}$ such that $\delta J(y_*;v) = \lambda\delta G(y_*;v)$ for all $v \in \mathscr{V}$.

To see why this is the case, suppose that (1) does not hold. Then we may choose $w \in \mathscr{V}$ such that $\delta G(y_*;w) \neq 0$. For (4.6) to hold, we must have

$$\delta J(y_*;u) = \left[\frac{\delta J(y_*;w)}{\delta G(y_*;w)}\right]\delta G(y_*;u) \quad \text{for all } u \in \mathscr{V}.$$

So the λ in (2) may be set to the value of $\frac{\delta J(y_*;w)}{\delta G(y_*;w)}$.

4.2.1 Lagrange Multipliers in $\mathbb{R}^n$

We will apply our results from the previous section to the case when $\mathfrak{X} = \mathscr{Y} = \mathbb{R}^n$. Assume that $f, g : \mathbb{R}^n \to \mathbb{R}$ have continuous first-order partial derivatives, and let $c \in \mathbb{R}$ be given. Put

$$\mathscr{S} := \{x \in \mathbb{R}^n \mid g(x) = c\}.$$

We may take $\mathscr{V} = \mathbb{R}^n$.

Suppose that f attains a minimum over $\mathscr{S}$ at the point x_*. The Gateaux variations of f and g are given by

$$\delta f(x_*;v) = \nabla f(x_*) \cdot v,\ \delta g(x_*;v) = \nabla g(x_*) \cdot v \quad \text{for all } v \in \mathbb{R}^n.$$

It follows from the chain rule that mapping Φ will be continuously differentiable. Thus by our argument in the previous section, we conclude that

(1) either $\nabla g(x_*) = 0$, or

(2) there exists a $\lambda \in \mathbb{R}$ such that $\nabla f(x_*) = \lambda \nabla g(x_*)$.

This is the standard method of Lagrange multipliers in $\mathbb{R}^n$. The quantity λ is usually referred to as a Lagrange multiplier.

Remark 4.1 *Assume that $c \in \mathrm{Range}(g)$ and that $\mathscr{S}$ is bounded. Since g is continuous it follows that $\mathscr{S}$ is closed, and therefore compact. (It is also nonempty.) Consequently, f attains a minimum and a maximum on $\mathscr{S}$, i.e., there exist $x_*, y_* \in \mathscr{S}$ such*

$$f(x_*) \leq f(x) \leq f(y_*) \quad \text{for all } x \in \mathbb{R}^n.$$

4.2.2 Example of the Lagrange Multiplier Method in $\mathbb{R}^2$

Define $f, g : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x_1, x_2) = x_1 x_2 \text{ and } g(x_1, x_2) = x_1^2 + x_2^2 \quad \text{for all } x = (x_1, x_2) \in \mathbb{R}^2.$$

Set

$$\mathscr{S} := \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) = 1 \right\}.$$

We consider the problem of minimizing f over the class $\mathscr{S}$.

First, we determine those $x \in \mathscr{S}$ such that $\nabla g(x) = 0$. For the gradient of g, we have

$$\nabla g(x_1, x_2) = (2x_1, 2x_2) \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2.$$

So

$$\nabla g(x) = 0 \Leftrightarrow x = 0,$$

but $0 \notin \mathscr{S}$. Therefore 0 is not a possible minimizer for f over $\mathscr{S}$.

Let us now look for those $x \in \mathscr{S}$ for which there exists a $\lambda \in \mathbb{R}$ such that $\nabla f(x) = \lambda \nabla g(x)$. The gradient of f is given by

$$\nabla f(x_1, x_2) = (x_2, x_1) \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2,$$

so we want those points in $\mathscr{S}$ where

$$(x_2, x_1) = \lambda(2x_1, 2x_2).$$

We therefore seek solutions to

$$\begin{cases} x_2 = 2\lambda x_1; \\ x_1 = 2\lambda x_2; \\ x_1^2 + x_2^2 = 1. \end{cases} \tag{4.7}$$

Notice that this is a system of three equations with three unknowns. Further note that if λ, x_1 or x_2 is zero, then $x_1 = x_2 = 0$. Since $0 \notin \mathscr{S}$, we must have

$\lambda \neq 0$, $x_1 \neq 0$ and $x_2 \neq 0$. So we may substitute the second relation in (4.7) into the first and divide by x_2. Doing so, we see that

$$x_2 = 2\lambda(2\lambda x_2) \Rightarrow \lambda = \pm\frac{1}{2}.$$

Now if $\lambda = \frac{1}{2}$, then $x_1 = x_2$; while $\lambda = -\frac{1}{2}$ implies that $x_1 = -x_2$. Using the third condition in (4.7), we find $x_1 = \pm\frac{1}{\sqrt{2}}$. Thus the possible minimizers for f over $\mathscr{S}$ are

$$\begin{array}{rl} \text{with } \lambda = \frac{1}{2} & \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \text{ and } \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right); \\ \text{with } \lambda = -\frac{1}{2} & \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \text{ and } \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right). \end{array}$$

The last two points are the minimizers for f over $\mathscr{S}$, while the first two points are actually maximizers.

4.2.3 Eigenvalues of a Real Symmetric Matrix

There are very important connections between Lagrange multipliers and eigenvalues. In this section we use the method of Lagrange multipliers to show that every real symmetric matrix has at least one (real) eigenvalue.

Let $n \in \mathbb{N}$ be given and let A be an $n \times n$ matrix with real entries. We assume that A is symmetric, i.e. that $A = A^T$, where A^T is the transpose of A. Recall that a real number λ is said to be an *eigenvalue* of A provided that there is a nonzero vector $x \in \mathbb{R}^n$ such that

$$Ax^T = \lambda x^T.$$

Since we regard the elements of $\mathbb{R}^n$ as row vectors, we have taken the transpose of x to turn it into a column vector (i.e., an $n \times 1$ matrix before multiplying it on the left by an $n \times n$ matrix.

Let us define $f, g : \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) = xAx^T \quad \text{for all } x \in \mathbb{R}^n,$$

$$g(x) = x \cdot x \quad \text{for all } x \in \mathbb{R}^n,$$

and put

$$\mathscr{S} = \{x \in \mathbb{R}^n : g(x) = 1\}.$$

It is straightforward to verify that f and g are continuously differentiable and

$$(\nabla f(x))^T = 2Ax^T, \ \ (\nabla g(x))^T = 2x^T \quad \text{for all } x \in \mathbb{R}^n. \tag{4.8}$$

We note that the assumption $A = A^T$ has been used in an essential way to derive the formula for ∇f. In fact this is the only part of the argument where symmetry of A is used.

It is clear that $1 \in \text{Range}(g)$ and that $\mathscr{S}$ is bounded, so by Remark 4.1, we may choose $x_* \in \mathscr{S}$ that minimizes f over $\mathscr{S}$. Since $x_* \in \mathscr{S}$, we know that

$\nabla g(x_*) \neq 0$ by virtue of our formula for ∇g. Therefore, we may choose $\lambda \in \mathbb{R}$ such that $\nabla f(x_*) = \lambda \nabla g(x_*)$. In view of (4.8), we conclude that

$$Ax_*^T = \lambda x_*^T,$$

and λ is an eigenvalue of A since $0 \notin \mathscr{S}$.

The number λ produced by the minimization procedure is the *smallest* eigenvalue of A. We know that f also attains a maximum on $\mathscr{S}$. If we maximize f on $\mathscr{S}$, the corresponding number λ will be the *largest* eigenvalue of A. (Of course, there is no guarantee that the largest and smallest eigenvalues of A will be different, as we can see by taking A to be the identity matrix.)

4.3 An Example of the Lagrange Multiplier Method in the Calculus of Variations

We now illustrate the Lagrange multiplier method for a calculus of variations problem with an integral constraint, namely the problem from Section 4.1. We put $\mathfrak{X} = \mathscr{C}^1[0,\pi]$,

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

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

$$J(y) := \int_0^\pi y'(x)^2 \, dx \text{ and } G(y) := \int_0^\pi y(x)^2 \, dx \quad \text{for all } y \in \mathscr{Y}.$$

Our problem is to minimize J over

$$\mathscr{S} := \{ y \in \mathscr{Y} \, : \, G(y) = 1 \}.$$

The subspace $\mathscr{V}$ will be

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

Notice that this is simply the class of $\mathscr{Y}$-admissible variations. For now, we take it for granted that the function Φ will be continuously differentiable; this technical issue will be addressed later.

We first determine the functions $y \in \mathscr{S}$ such that $\delta G(y; v) = 0$ for every $v \in \mathscr{V}$. These functions must satisfy the first Euler-Lagrange equations for G:

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

This, however, implies $y(x)$ is zero at each $x \in [0,\pi]$, but the zero function is not a member of $\mathscr{S}$, so 0 is not a possible minimizer for J over $\mathscr{S}$.

We now seek those $y \in \mathscr{S}$ where there is a $\lambda \in \mathbb{R}$ such that $\delta J(y; v) - \lambda \delta G(y; v) = 0$ for every $v \in \mathscr{V}$. It is straightforward to verify that such y must satisfy the first Euler-Lagrange equation for the functional $J - \lambda G$. Therefore, we seek those y satisfying

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

and the conditions

$$y(0) = y(\pi) = 0 \text{ and } \int_0^\pi y(x)^2\, dx = 1.$$

We now solve

$$\begin{cases} y''(x) + \lambda y(x) = 0; \\ y(0) = y(\pi) = 0; \\ \int_0^\pi y(x)^2\, dx = 1. \end{cases} \qquad (4.9)$$

There are three cases to consider: the first case is when $\lambda < 0$, the second is $\lambda = 0$ and the last is $\lambda > 0$.

For the first case, we may choose $\omega \in \mathbb{R}$ so that $\omega > 0$ and $\lambda = -\omega^2$. Let us first look for the general solution to

$$y''(x) - \omega^2 y(x) = 0 \quad \text{for all } x \in [0,\pi]. \qquad (4.10)$$

The characteristic equation for (4.10) is $r^2 - \omega^2 = 0$. The roots of the characteristic equation are therefore $r = \pm\omega$ and the general solution to (4.10) is

$$y(x) = c_1 e^{\omega x} + c_2 e^{-\omega x}. \qquad (4.11)$$

It is easily checked that $c_1 = c_2 = 0$ in order for the function y given in (4.11) to satisfy $y(0) = y(\pi) = 0$. So for this case, the first two conditions in (4.9) imply that $y(x) = 0$ at each $x \in [0,\pi]$. Clearly, the last condition in (4.9) cannot be satisfied. So there is no solutions to (4.9) when $\lambda < 0$.

We now look at the second case, when $\lambda = 0$. The first condition in (4.9) reduces to

$$y''(x) = 0 \quad \text{for all } x \in [0,\pi].$$

For this equation, the general solution is

$$y(x) = c_1 x + c_2.$$

As in the first case, the condition $y(0) = y(\pi) = 0$ is satisfied only when $c_1 = c_2 = 0$. Thus, we again have no solutions for this case.

In the last case, we have $\lambda > 0$. So we may choose $\omega \in \mathbb{R}$ so that $\omega > 0$ and $\lambda = \omega^2$. We want the general solution for

$$y''(x) + \omega^2 y(x) = 0 \quad \text{for all } x \in [0,\pi]. \qquad (4.12)$$

The roots to the characteristic equation $r + \omega^2 = 0$ for (4.12) are $r = \pm i\,\omega$. The general solution to (4.12) is

$$y(x) = c_1 \sin\omega x + c_2 \cos\omega x. \qquad (4.13)$$

Let us now find c_1, c_2 such that $y(0) = y(\pi) = 0$. We have

$$y(0) = 0 \Rightarrow c_2 = 0 \Rightarrow y(x) = c_1 \sin \omega x,$$

and

$$y(\pi) = 0 \Rightarrow c_1 \sin \omega\pi = 0 \Rightarrow c_1 = 0 \text{ or } \sin \omega\pi = 0.$$

Now if $c_1 = 0$, then $y(x) = 0$ at each $x \in [0, \pi]$, and we have already seen that this does not satisfy all the conditions in (4.9). So we want $\sin \omega\pi$ to be zero, or in other words, we want ω to be an integer. Since $\omega > 0$, we take ω to be a natural number $n \in \mathbb{N}$. Therefore, we have

$$y(x) = c_1 \sin nx \quad \text{for all } x \in [0, \pi],$$

for some $c_1 \in \mathbb{R}\backslash\{0\}$ and $n \in \mathbb{N}$. Now we will try to satisfy the third condition in (4.9). We need

$$\int_0^\pi c_1^2 \sin^2 nx\, dx = c_1^2 \int_0^\pi \sin^2 nx\, dx = 1.$$

Thus

$$c_1^2 \frac{\pi}{2} = 1 \Rightarrow c_1 = \pm\sqrt{\frac{2}{\pi}}.$$

So for every $n \in \mathbb{N}$, the functions

$$y(x) = \pm\sqrt{\frac{2}{\pi}} \sin nx \tag{4.14}$$

are all the possible minimizers for J over the class $\mathscr{S}$.

Let us compute the value of J for the functions given in (4.14). For y given by (4.14), we find that

$$y'(x) = \pm n\sqrt{\frac{2}{\pi}} \cos nx.$$

Thus

$$J(y) = \int_0^\pi y'(x)^2\, dx = \frac{2n^2}{\pi} \int_0^\pi \cos^2 nx\, dx = \frac{2n^2}{\pi}\frac{\pi}{2} = n^2, \tag{4.15}$$

where $n \in \mathbb{N}$. The smallest value of J for the functions given in (4.14) is 1, and this value is provided by the functions

$$y(x) = \pm\sqrt{\frac{2}{\pi}} \sin x.$$

So these are the only two possible minimizers for J over $\mathscr{S}$. We note that the computation in (4.15) shows that there is no maximizer for J over $\mathscr{S}$, since we may take $n \in \mathbb{N}$ as large as desired.

4.4 Summary of the Lagrange Multiplier Method for Basic Problems

Let $a, b, A, B, c \in \mathbb{R}$ with $a < b$ be given, and let $f, g : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be functions with continuous first partial derivatives. In order to handle problems with fixed endpoints and free endpoints simultaneously, let $\alpha, \beta \in \{0, 1\}$ be given and put

$$\mathscr{Y} := \left\{ y \in C^1[a, b] \, : \, \alpha y(a) = \alpha A \text{ and } \beta y(b) = \beta B \right\}.$$

Notice that if $\alpha = 1$ then the value of y at a is prescribed to be A, while if $\alpha = 0$, then we have a free end at a. (Similar comments apply to β and the end at b.) Define $J, G : \mathscr{Y} \to \mathbb{R}$ by

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

and let

$$\mathscr{S} := \{ y \in \mathscr{Y} \, : \, G(y) = c \}.$$

We take $\mathscr{V}$ to be

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

In order to apply the Lagrange multiplier procedure, we need to know that the mapping Φ is continuously differentiable. The required smoothness of Φ is assured by the following lemma.

Lemma 4.2 *Assume that $M : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has continuous first partial derivatives and let $y, u, w \in C^1[a, b]$ be given. Define $\Psi : \mathbb{R}^2 \to \mathbb{R}$ by*

$$\Psi(\eta, \mu) = \int_a^b M(x, y(x) + \eta u(x) + \mu w(x), y'(x) + \eta u'(x) + \mu w'(x))\, dx \quad \text{for all } \eta, \mu \in \mathbb{R}.$$

Then Ψ is continuously differentiable.

We define the augmented integrand $L : [a, b] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$L(x, y, z, \lambda) := f(x, y, z) - \lambda g(x, y, z) \quad \text{for all } (x, y, z, \lambda) \in [a, b] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

Consider the functional $H : \mathscr{Y} \times \mathbb{R} \to \mathbb{R}$ given by

$$H(y, \lambda) := \int_a^b L(x, y(x), y'(x), \lambda) dx \quad \text{for all } y \in \mathscr{Y} \text{ and } \lambda \in \mathbb{R}.$$

There are two basic steps to finding all the possible minimizers (and maximizers) for J over the class $\mathscr{S}$:

(Step 1) find all solutions (if any exist) to $(\text{E-L})_1$ for the functional G that belong to $\mathscr{S}$ and satisfy any relevant natural boundary conditions.

(Step 2) holding λ fixed, find all solutions (if any exist) to $(\text{E-L})_1$ for the functional H that belong to $\mathscr{S}$ any relevant natural boundary conditions (NBC).

For (Step 2), one treats λ as a constant in order to find $(\text{E-L})_1$ for H. So the first Euler-Lagrange equation for H is given by

$$L_{,2}(x, y(x), y'(x), \lambda) = \frac{d}{dx}\left[L_{,3}(x, y(x), y'(x), \lambda)\right]. \qquad (\text{E-L})_1$$

The set of possible minimizers (and maximizers) for J over $\mathscr{S}$ consists of all the solutions found from (Step 1) and (Step 2). We emphasize that the solutions, if any, from (Step 1) must be considered as *potential* minimizers (and maximizers).

4.5 An Isoperimetric Problem

For this problem, let

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

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

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

We want to maximize the functional J over the class

$$\mathscr{S} := \{y \in \mathscr{Y} \mid G(y) = l\},$$

where $l > 2$ is some fixed constant. The integrands for J and G are given by

$$f(x, y, z) = y \text{ and } g(x, y, z) = \left(1 + z^2\right)^{\frac{1}{2}} \quad \text{for all } (x, y, z) \in [-1,1] \times \mathbb{R} \times \mathbb{R},$$

respectively, and the augmented integrand for $J - \lambda G$ is given by

$$L(x, y, z, \lambda) = y - \lambda\left(1 + z^2\right)^{\frac{1}{2}} \quad \text{for all } (x, y, z, \lambda) \in [-1,1] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}.$$

We will now try to solve our problem using the method of Lagrange multipliers.

For (Step 1), we need to find all $y \in \mathscr{S}$ that satisfy the first Euler-Lagrange equation for G, which is

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

The general solution to this equation is easily found to be

$$y(x) = c_1 x + c_2 \quad \text{for all } x \in [-1,1].$$

For y to be a member of $\mathscr{Y}$, we need $y(-1) = y(1) = 0$ and this implies $c_1 = c_2 = 0$. Thus $y(x) = 0$ at each $x \in [-1, 1]$. For y to be in $\mathscr{S}$, we also need $G(y)$ to be equal to l. With $y(x) = 0$ at each $x \in [-1, 1]$, however, we find

$$G(y) = \int_{-1}^{1} \sqrt{1 + y'(x)^2}\, dx = \int_{-1}^{1} dx = 2 < l.$$

So there are no functions $y \in \mathscr{S}$ that satisfy the first Euler-Lagrange equation for G. This completes (Step 1).

We now proceed to (Step 2). The first Euler-Lagrange equation for $J - \lambda G$ is

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

Upon computing the necessary partial derivatives of L, we find that $y \in \mathscr{S}$ must satisfy

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

Clearly $\lambda \neq 0$. Letting $\mu = -\lambda$, we see that y satisfies (E-L)_1 if and only if

$$\frac{y'(x)}{\left(1 + y'(x)^2\right)^{\frac{1}{2}}} = \frac{x + c_1}{\mu} \quad \text{for all } x \in [-1, 1]. \tag{4.16}$$

(It is important to keep track of the sign of the Lagrange multiplier because this information plays a role in determining whether the candidate we find will actually provide a maximum.) At this point we observe that the denominator on the left hand side of (4.16) is always positive. So (4.16) implies that $y'(x)$ always has the same sign as $\frac{x+c_1}{\mu}$. Upon solving (4.16) for y', we have

$$y'(x) = \frac{\frac{x+c_1}{\mu}}{\sqrt{1 - \left(\frac{x+c_1}{\mu}\right)^2}} \quad \text{for all } x \in [-1, 1]. \tag{4.17}$$

Integrating (4.17) we find that

$$y(x) = -\mu\sqrt{1 - \left(\frac{x + c_1}{\mu}\right)^2} + c_2 = \lambda\sqrt{1 - \left(\frac{x + c_1}{\lambda}\right)^2} + c_2 \quad \text{for all } x \in [-1, 1].$$

This is the equation for a circle of radius $|\lambda|$ centered at the point $(-c_1, c_2)$. We need y to satisfy $y(-1) = y(1) = 0$ for it to be a member of $\mathscr{Y}$, which leads to

$$\lambda\sqrt{1 - \left(\frac{-1 + c_1}{\lambda}\right)^2} + c_2 = \lambda\sqrt{1 - \left(\frac{1 + c_1}{\lambda}\right)^2} + c_2 \Rightarrow (c_1 - 1)^2 = (c_1 + 1)^2$$
$$\Rightarrow c_1 = 0.$$

Thus

$$y(x) = \lambda\sqrt{1 - \left(\frac{x}{\lambda}\right)^2} + c_2 \quad \text{for all } x \in [-1, 1]. \tag{4.18}$$

For y to be in $\mathscr{S}$, we now need to satisfy the constraint $G(y) = l$. Computing the derivative for y given in (4.18), we find that

$$y'(x) = -\frac{\frac{x}{\lambda}}{\sqrt{1 - \left(\frac{x}{\lambda}\right)^2}} \Rightarrow 1 + y'(x)^2 = \frac{1}{1 - \left(\frac{x}{\lambda}\right)^2} \quad \text{for all } x \in [-1, 1].$$

Therefore

$$G(y) = \int_{-1}^{1} \sqrt{1 + y'(x)^2}\, dx = \int_{-1}^{1} \frac{1}{\sqrt{1 - \left(\frac{x}{\lambda}\right)^2}}\, dx = 2|\lambda| \arcsin \frac{1}{|\lambda|}.$$

Thus for y to be in $\mathscr{S}$, we must have

$$|\lambda| \arcsin \frac{1}{|\lambda|} = \frac{l}{2}. \tag{4.19}$$

It is not difficult to prove that there is a value for $|\lambda|$ satisfying (4.19) if and only if $2 < l \leq \pi$. Provided that $2 < l \leq \pi$, the condition $y(1) = 0$ implies that

$$c_2 = -\lambda\sqrt{1 - \frac{1}{\lambda^2}}.$$

Whence

$$y(x) = \lambda\left[\sqrt{1 - \frac{x^2}{\lambda^2}} - \sqrt{1 - \frac{1}{\lambda^2}}\right],$$

with $|\lambda|$ satisfying (4.19). When $l = \pi$, the function y above does not belong to $C^1[-1, 1]$ because there are vertical tangents at the endpoints. We shall show in the next chapter that when λ is positive, our solution is a maximizer, and when λ is negative, we have a minimizer.