

## Chapter 14

### The classical theory

*We will now discuss in a little more detail the Struggle for Existence.*

Charles Darwin (The Origin of Species)

*Life is grey, but the golden tree of theory is always green.*

Goethe (Journey by Moonlight)

*It's the question that drives us, Neo.*

Morpheus (The Matrix)

The basic problem in the subject that is referred to as the *calculus of variations* consists in minimizing an integral functional of the type

$$J(x) = \int_a^b \Lambda(t, x(t), x'(t)) dt$$

over a class of functions  $x$  defined on the interval  $[a, b]$ , and which take prescribed values at  $a$  and  $b$ .

The study of this problem (and its numerous variants) is over three centuries old, yet its interest has not waned. Its applications are numerous in geometry and differential equations, in mechanics and physics, and in areas as diverse as engineering, medicine, economics, and renewable resources. It is not surprising, then, that modeling and numerical analysis play a large role in the subject today. In the following chapters, however, we present a course in the calculus of variations which focuses on the core mathematical issues: necessary conditions, sufficient conditions, existence theory, regularity of solutions.

For those like the reader who have a sensitive mathematical conscience, the statement of the basic problem, as rendered above, may well create an uneasiness, a craving for precision. What, exactly, is the class of functions in which  $x$  lies? What hypotheses are imposed on the function  $\Lambda$ ? Is the integral well defined? Does a solution exist?

In the early days of the subject, these questions went unaddressed, at least explicitly. (Implicitly: everything was taking place in a very smooth universe in which problems evidently had solutions.) Our era, more attuned to the limits of smoothness, requires a more deliberate approach, and a well-defined setting. And this is just as well, for, as the reader will come to understand, the history, the development, and the most useful insights into the subject are inextricably wrapped up with the very questions just posed.

**Hypotheses.** The focus of our attention is the integral functional  $J(x)$  defined above, where  $\Lambda$  is a function of three variables, and where  $[a, b]$  is a given interval in  $\mathbb{R}$ .  $\Lambda(t, x, v)$  is referred to as the **Lagrangian**, and the generic notation for its three variables is  $(t, x, v)$ : time, state, velocity.

This chapter deals with the case in which these variables are one-dimensional and all the functions involved are smooth. We take  $\Lambda : \mathbb{R}^3 \rightarrow \mathbb{R}$  to be a twice continuously differentiable function, and we limit attention to functions  $x : [a, b] \rightarrow \mathbb{R}$  that belong to  $C^2[a, b]$ . This means that  $x$  lies in  $C[a, b]$ , the derivatives  $x'$  and  $x''$  exist and are continuous in  $(a, b)$ , and both  $x'$  and  $x''$  admit continuous extensions to  $[a, b]$ . It is clear that this is more than adequate to guarantee that the integral defining  $J(x)$  is well defined for each competing  $x$ .

Given in addition two points  $A$  and  $B$  in  $\mathbb{R}$ , we now consider the **basic problem** in the calculus of variations:

$$\text{minimize } J(x) : x \in C^2[a, b], x(a) = A, x(b) = B. \quad (\text{P})$$

$J(x)$  is referred to as the **cost** corresponding to  $x$ . A function  $x : [a, b] \rightarrow \mathbb{R}$  is termed **admissible** if it satisfies the boundary constraints and lies in the appropriate class, in this case  $C^2[a, b]$ . A **solution**  $x_*$  of (P) refers to an admissible function  $x_*$  such that  $J(x_*) \leq J(x)$  for all other admissible functions  $x$ . We also refer to  $x_*$  as a **minimizer** for the problem.

**14.1 Example. (A minimal surface problem)** The well-known problems to which the calculus of variations was first applied arise in geometry and mechanics. A famous example of the problem (P) that goes back to Euler's seminal monograph of 1744 is to find the shape of the curve  $x(t)$  joining  $(a, A)$  to  $(b, B)$  whose associated surface of rotation (about the  $t$ -axis) has minimal area.

This can be given a physical interpretation: when a soap surface is spanned by two concentric rings of radius  $A$  and  $B$ , the resulting surface will be a surface of rotation of a curve  $x(t)$ , and we expect the area of the surface to be a minimum. This expectation (confirmed by experiment) is based upon *d'Alembert's principle*, which affirms that in static equilibrium, the observed configuration minimizes total potential energy (which, for a soapy membrane desperately seeking to contract, is proportional to its area).

In concrete terms, the soap bubble problem consists of minimizing

$$\int_a^b x(t) \sqrt{1 + x'(t)^2} dt \quad \text{subject to } x(a) = A, x(b) = B.$$

This is the case of the basic problem (P) in which  $\Lambda(t, x, v) = x \sqrt{1 + v^2}$ . (The surface area is in fact given by  $2\pi$  times the integral, but we can omit this multiplicative factor, which does not effect the minimization.) We shall be seeing this problem again later.  $\square$

## 14.1 Necessary conditions

The following result identifies the first *necessary condition* that a minimizing  $x$  must satisfy; it is in effect an analogue of Fermat's rule that  $f'(x) = 0$  at a minimum.

**Notation:** The partial derivatives of the function  $\Lambda(t, x, v)$  with respect to  $x$  and to  $v$  are denoted by  $\Lambda_x$  and  $\Lambda_v$ .

**14.2 Theorem. (Euler 1744)** *If  $x_*$  is a solution of (P), then  $x_*$  satisfies the Euler equation:*

$$\frac{d}{dt} \{ \Lambda_v(t, x_*(t), x_*'(t)) \} = \Lambda_x(t, x_*(t), x_*'(t)) \quad \forall t \in [a, b]. \quad (1)$$

**Proof.** Euler's proof of this result used discretization, but the now standard proof given here uses Lagrange's idea: a *variation*, from which the subject derives its name. In the present context, a variation means a function  $y \in C^2[a, b]$  such that  $y(a) = y(b) = 0$ . We fix such a  $y$ , and proceed to consider the following function  $g$  of a single variable:

$$g(\lambda) = J(x_* + \lambda y) = \int_a^b \Lambda(t, x_* + \lambda y, x_*' + \lambda y') dt. \quad (2)$$

(The reader will notice that we have yielded to the irresistible temptation to leave out certain arguments in the expression for the integral; thus  $x_*$  should be  $x_*(t)$  and  $y$  is really  $y(t)$ , and so on. Having already succumbed the first time, we shall do so routinely hereafter.) It follows from standard results in calculus that  $g$  is differentiable, and that we can "differentiate through the integral" to obtain

$$g'(\lambda) = \int_a^b [ \Lambda_x(t, x_* + \lambda y, x_*' + \lambda y') y + \Lambda_v(t, x_* + \lambda y, x_*' + \lambda y') y' ] dt. \quad (3)$$

Observe now that for each  $\lambda$ , the function  $x_* + \lambda y$  is admissible for (P), whence

$$g(\lambda) = J(x_* + \lambda y) \geq J(x_*) = g(0).$$

It follows that  $g$  attains a minimum at  $\lambda = 0$ , and hence that  $g'(0) = 0$ ; thus:

$$\int_a^b [ \alpha(t)y(t) + \beta(t)y'(t) ] dt = 0,$$

where we have set

$$\alpha(t) = \Lambda_x(t, x_*(t), x_*'(t)), \quad \beta(t) = \Lambda_v(t, x_*(t), x_*'(t)).$$

Using integration by parts, we deduce

$$\int_a^b [\alpha(t) - \beta'(t)]y(t) dt = 0.$$

Since this is true for any variation  $y$ , it follows that the continuous function which is the coefficient of  $y$  under the integral sign must vanish identically on  $[a, b]$  (left as an exercise). But this conclusion is precisely Euler's equation.  $\square$

A function  $x \in C^2[a, b]$  satisfying Euler's equation is referred to as an **extremal**. The Euler equation (1) is (implicitly) a differential equation of order two for  $x_*$ , and one may expect that, in principle, the two boundary conditions will single out a unique extremal. We shall see, however, that it's more complicated than that.

### 14.3 Exercise.

- Show that all extremals for the Lagrangian  $\Lambda(t, x, v) = \sqrt{1 + v^2}$  are affine. Why is this to be expected?
- Show that the Euler equation for the Lagrangian  $\Lambda(t, x, v) = x^2 + v^2$  is given by  $x'' - x = 0$ .
- Find the unique admissible extremal for the problem

$$\min \int_0^1 (x'(t)^2 + x(t)^2) dt \quad : \quad x \in C^2[0, 1], \quad x(0) = 0, \quad x(1) = 1. \quad \square$$

**Local minima are extremals.** The Euler equation is the first-order necessary condition for the calculus of variations problem (P), and we would expect it to hold for merely *local* minima (suitably defined). We develop this thought now.

A function  $x_*$  admissible for (P) is said to provide a **weak local minimum** if, for some  $\varepsilon > 0$ , for all admissible  $x$  satisfying  $\|x - x_*\| \leq \varepsilon$  and  $\|x' - x_*'\| \leq \varepsilon$ , we have  $J(x) \geq J(x_*)$ . The anonymous norm referred to in a context such as this one will always be that of  $L^\infty[a, b]$  (or  $C[a, b]$ ); thus, for example, the notation above refers to

$$\|x - x_*\| = \max \{ |x(t) - x_*(t)| : t \in [a, b] \}.$$

The proof of the necessity of Euler's equation goes through for a local minimizer just as it did for a global one: the function  $g$  defined in (2) attains a local minimum at 0 rather than a global one; but we still have  $g'(0) = 0$ , which is what leads to the Euler equation. Thus any weak local minimizer for (P) must be an extremal.

**The Erdmann condition.** The Lagrangian  $\Lambda$  is said to be **autonomous** if it has no explicit dependence on the  $t$  variable. The following consequence of the Euler equation can be a useful starting point in the identification of extremals.

**14.4 Proposition.** *Let  $x_*$  be a weak local minimizer for (P), where  $\Lambda$  is autonomous. Then  $x_*$  satisfies the **Erdmann condition**: for some constant  $h$ , we have*

$$x_*'(t) \Lambda_v(x_*(t), x_*'(t)) - \Lambda(x_*(t), x_*'(t)) = h \quad \forall t \in [a, b].$$

**Proof.** It suffices to show that the derivative of the function on the left side is zero, which follows from the Euler equation: we leave this as an exercise.  $\square$

**14.5 Example. (continued)** We return to the soap bubble problem (Example 14.1), armed now with some theory. Suppose that  $x_*$  is a weak local minimizer for the problem, with  $x_*(t) > 0 \quad \forall t$ . The reader may verify that the Euler equation is given by

$$x''(t) = (1 + x'(t)^2)/x(t).$$

We deduce from this that  $x_*'$  is strictly increasing (thus  $x_*$  is strictly convex). Since  $\Lambda$  is autonomous, we may invoke the Erdmann condition (Prop. 14.4). This yields the existence of a positive constant  $k$  such that

$$(x_*(t)')^2 = x_*(t)^2/k^2 - 1, \quad t \in [a, b].$$

If  $x_*'$  is positive throughout  $[a, b]$ , we may solve this to reveal the separated differential equation

$$\frac{k dx}{\sqrt{x^2 - k^2}} = dt.$$

Mathematics students used to know by heart a primitive for the left side: the function  $k \cosh^{-1}(x/k)$ . It follows that  $x_*$  is of the form

$$x_*(t) = k \cosh\left(\frac{t+c}{k}\right).$$

A curve of this type is called a *catenary*.<sup>1</sup>

If, instead of being positive,  $x_*'$  is negative, a similar analysis shows that, once again,  $x_*$  is a catenary:

$$x_*(t) = \kappa \cosh\left(\frac{t+\sigma}{\kappa}\right),$$

for certain constants  $\sigma, \kappa$  (different from  $c$  and  $k$ , on the face of it). Since  $x_*'$  is strictly increasing, the general case will have  $x_*'$  negative, up to a point  $\tau$  (say), and then positive thereafter. Thus,  $x_*$  is a catenary (with constants  $\sigma, \kappa$ ) followed by another catenary (with constants  $c, k$ ). The smoothness of  $x_*$ , it can be shown, forces the constants to coincide, so we can simply assert that  $x_*$  is a catenary.

Notice that the detailed analysis of this problem is not a trivial matter. Furthermore, the only rigorous conclusion reached at the moment is the following: if there exists a weak local minimizer  $x_* \in C^2[a, b]$  which is positive on  $[a, b]$ , then  $x_*$  is a catenary.

---

<sup>1</sup> Not every choice of interval  $[a, b]$  and prescribed endpoints  $A, B$  will define a catenary; we do not pursue this issue, which is carefully analyzed in Bliss [5].

Our impulse may be to accept that conclusion, especially since soap bubbles do demonstrably exist, and have the good grace to cooperate (frequently) by being catenaries. But the example illustrates a general fact in optimization: “real” problems are rarely so impressed by the theory that they immediately reveal their secrets.

An *ad hoc* analysis, sometimes difficult, is often required. In the soap bubble case, for example, note the use of the (presupposed) regularity of  $x_*$ , which allowed us to match up the catenaries. (We return to this regularity issue later.) In this text, we stress the theory, rather than the details of particular problems. But we prefer to have warned the reader that no amount of general theory reduces a difficult problem to a simple exercise.  $\square$

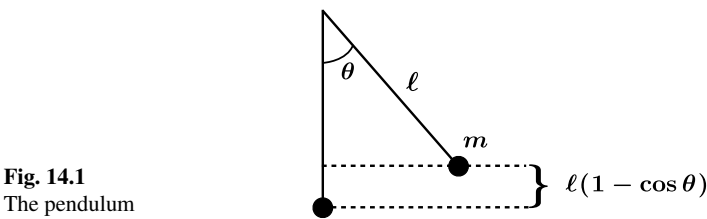
Our next example involves an important topic in classical mechanics.

**14.6 Example. (Least action principle)** In 1744, Euler (in that same monograph we mentioned before) extended d’Alembert’s principle to mechanical systems which are in motion, rather than in static equilibrium. His celebrated Principle of Least Action<sup>2</sup> postulates that the movement between two time instants  $t_1$  and  $t_2$  minimizes the *action*

$$\int_{t_1}^{t_2} (K - V) dt,$$

where  $K$  refers to kinetic energy and  $V$  to potential energy.

We proceed to illustrate the principle of least action in a simple case: the (unforced) oscillation in the plane of a pendulum of length  $\ell$  whose mass  $m$  is entirely in the bob. The angle  $\theta$  (see Fig. 14.1) is a convenient choice of *generalized coordinate*; the motion of the pendulum is described by the corresponding function  $\theta(t)$ . In terms of  $\theta$ , the kinetic energy  $K = mv^2/2$  is given by  $m(\ell\theta')^2/2$ .



**Fig. 14.1**  
The pendulum

If one uses  $\theta = 0$  as the reference level for calculating potential energy  $mgh$ , then it is given in terms of  $\theta$  by  $mg\ell(1 - \cos\theta)$ , as a little trigonometry shows. Thus the action between two instants  $t_1$  and  $t_2$  is given by

$$\int_{t_1}^{t_2} \left\{ \frac{1}{2} m (\ell \theta'(t))^2 - mg\ell (1 - \cos \theta(t)) \right\} dt.$$

<sup>2</sup> Sometimes mistakenly attributed to Maupertuis; see the discussion in [27].

We apply the least action principle: it follows that the resulting motion  $\theta(t)$  satisfies Euler's equation for the action functional. The reader may check that this yields the following differential equation governing the pendulum's movement:

$$\theta''(t) + (g/\ell) \sin \theta(t) = 0.$$

This equation, which can also be deduced from Newton's law, is in fact the one which describes the movement of the pendulum. But does it really minimize the action? Perhaps in a local sense?

Consider the analogy with the minimization of a function  $f(x)$  (on  $\mathbb{R}$ , say). The Euler equation corresponds to the necessary condition  $f'(x_*) = 0$ , the stationarity of  $f$  at a given point  $x_*$ . Further evidence of a local minimum would be the second-order condition  $f''(x_*) \geq 0$ . And if we knew in addition that  $f''(x_*) > 0$ , then we could say with certainty that  $x_*$  provides at least a local minimum. In this light, it seems reasonable to pursue second-order conditions in the calculus of variations. The honor of first having done so belongs to Legendre, although, to some extent, he was scorned for his efforts, for reasons that we shall see.  $\square$

In studying second-order conditions, and for the rest of this chapter, we strengthen the regularity hypothesis on the Lagrangian by assuming that  $\Lambda$  is  $C^3$ .

**14.7 Theorem. (Legendre's necessary condition, 1786)** *Let  $x_*$  be a weak local minimizer for (P). Then we have*

$$\Lambda_{vv}(t, x_*(t), x_*'(t)) \geq 0 \quad \forall t \in [a, b].$$

**Proof.** We consider again the function  $g$  defined by (2). We observe that the formula (3) for  $g'(\lambda)$  implies that  $g'$  is itself differentiable. We proceed to develop an expression for  $g''(0)$ . Differentiating under the integral in  $g'(\lambda)$ , and then setting  $\lambda = 0$ , we obtain

$$g''(0) = \int_a^b \left[ \Lambda_{xx}(t) y^2 + 2\Lambda_{xv}(t) y y' + \Lambda_{vv}(t) y'^2 \right] dt,$$

where  $\Lambda_{xx}(t)$  (for example) is an abbreviation for  $\Lambda_{xx}(t, x_*(t), x_*'(t))$ , and where we have invoked the fact that  $\Lambda_{xv}$  and  $\Lambda_{vx}$  coincide. We proceed to define

$$P(t) = \Lambda_{vv}(t, x_*(t), x_*'(t)) \tag{4}$$

$$Q(t) = \Lambda_{xx}(t, x_*(t), x_*'(t)) - \frac{d}{dt} \Lambda_{xv}(t, x_*(t), x_*'(t)). \tag{5}$$

(Note that  $Q$  is well defined, in part because  $\Lambda$  is  $C^3$ .) Using this notation, integration by parts shows that the last expression for  $g''(0)$  may be written

$$g''(0) = \int_a^b \left[ P(t) y'^2(t) + Q(t) y^2(t) \right] dt. \tag{6}$$

Since  $g$  attains a local minimum at 0, we have  $g''(0) \geq 0$ . We now seek to exploit the fact that this holds for every variation  $y$ . To begin with, a routine approximation argument (see Ex. 21.13) shows that for any Lipschitz (rather than  $C^2$ ) variation  $y$  in  $\text{Lip}_0[a, b]$  (the class of Lipschitz functions on  $[a, b]$  that vanish at  $a$  and  $b$ ), we still have the inequality (6).

Now let  $[c, d]$  be any subinterval of  $[a, b]$ , and let  $\varepsilon$  be any positive number. We define a function  $\varphi \in \text{Lip}_0[a, b]$  as follows:  $\varphi$  vanishes on  $[a, c]$  and  $[d, b]$ , and, in between (that is, on  $[c, d]$ ),  $\varphi$  is a sawtooth function whose derivative is alternately  $+1$  and  $-1$ , with the effect that for all  $t \in [c, d]$  we have  $|\varphi(t)| < \varepsilon$ . Then, by taking  $y = \varphi$  in (6), we deduce

$$\int_c^d [P(t) + |Q(t)|\varepsilon^2] dt \geq 0.$$

Since  $\varepsilon > 0$  is arbitrary, we conclude that the integral of the continuous function  $P(t)$  over  $[c, d]$  is nonnegative. Since in turn the subinterval  $[c, d]$  is arbitrary, we have proved, as required, that  $P$  is nonnegative on  $[a, b]$ .  $\square$

## 14.2 Conjugate points

In contrast to the Euler equation, the Legendre necessary condition has the potential to distinguish between a maximum and a minimum: at a local maximizer  $x_*$  of  $J$  (that is, a local minimizer of  $-J$ ), we have  $\Lambda_{v,v}(t, x_*(t), x'_*(t)) \leq 0$ .

To illustrate the distinction, consider the functional

$$J(x) = \int_a^b \sqrt{1 + x'(t)^2} dt.$$

Legendre's condition tells us that it is useless to seek local *maxima* of  $J$ , since here we have  $\Lambda_{v,v} = \{1 + v^2\}^{-3/2} > 0$ ; only local (or global) minima may exist.

Legendre proceeded to prove (quite erroneously) that his necessary condition, when strengthened to strict inequality, was also a *sufficient* condition for a given extremal to provide a weak local minimum. He was scathingly criticized by Lagrange for his sins.<sup>3</sup>

Legendre's proof went as follows. Using the same function  $g$  as above, together with the (Lagrange!) expansion

$$g(1) - g(0) = g'(0) + (1/2)g''(\lambda) = (1/2)g''(\lambda)$$

---

<sup>3</sup> The reader will be relieved to know that Legendre's reputation recovered; his name is one of 72 inscribed on the Eiffel tower. . .



(the Euler equation corresponds to  $g'(0) = 0$ ), routine calculations (given in the proof of Theorem 14.8 below) lead to the inequality

$$J(x_* + y) - J(x_*) \geq \frac{1}{2} \int_a^b [(P - \delta)y'^2 + Qy^2] dt \quad (1)$$

for all variations  $y$  in a suitable weak neighborhood of 0 (that is, such that  $\|y\|$  and  $\|y'\|$  are sufficiently small). Here,  $P$  and  $Q$  have the same meaning as before, and  $\delta > 0$  is chosen so that  $P(t) - \delta > 0$  on  $[a, b]$  (this is where the supposed strict positivity of  $P$  is used). There remains only to show, therefore, that the integral term in (1), which we label  $I$ , is nonnegative.

To this end, let  $w$  be any continuously differentiable function, and note that

$$\begin{aligned} I &= \int_a^b [(P - \delta)y'^2 + Qy^2] dt = \int_a^b [(P - \delta)y'^2 + Qy^2 + (wy^2)'] dt \\ &= \int_a^b (P - \delta) \left[ y'^2 + \frac{Q + w'}{P - \delta} y^2 + 2 \frac{w}{P - \delta} y y' \right] dt \\ &= \int_a^b (P - \delta) \left[ y' + \frac{w}{P - \delta} y \right]^2 dt \geq 0, \end{aligned}$$

where the factorization in the last integral expression depends upon having chosen the function  $w$  (heretofore arbitrary) to satisfy

$$\frac{Q + w'}{P - \delta} = \left( \frac{w}{P - \delta} \right)^2 \iff w' = \frac{w^2}{P - \delta} - Q.$$

The proof appears to be complete. It has a serious defect, however.

Even in the present sophisticated era, students are sometimes surprised that such an innocent-looking differential equation as the one above can *fail* to have a solution  $w$  defined on the entire interval  $[a, b]$ . In fact, the equation is nonlinear, and we know only that a solution exists if  $b$  is taken sufficiently close to  $a$ , from well-known local existence theorems. In light of this, perhaps we can forgive Legendre his error (taking for granted the existence of  $w$ ), especially since his approach, suitably adapted, did in fact turn out to be highly fruitful.

We summarize what can be asserted on the basis of the discussion so far.

**14.8 Theorem.** *Let  $x_* \in C^2[a, b]$  be an admissible extremal satisfying the strengthened Legendre condition  $\Lambda_{yy}(t, x_*(t), x_*'(t)) > 0 \forall t \in [a, b]$ . Suppose there exists a function  $w \in C^1[a, b]$  satisfying the differential equation*

$$w'(t) = \frac{w(t)^2}{P(t)} - Q(t), \quad t \in [a, b].$$

*Then  $x_*$  is a weak local minimizer for (P).*

**Proof.** We begin with the estimate mentioned above.

**Lemma.** *There is a constant  $M$  such that the following inequality holds for every variation  $y$  having  $\|y\| + \|y'\| \leq 1$ :*

$$J(x_* + y) - J(x_*) \geq \frac{1}{2} \int_a^b [Py'^2 + Qy^2] dt - M \{ \|y\| + \|y'\| \} \int_a^b y'^2 dt.$$

To prove this, we first observe that, for some  $\lambda \in (0, 1)$ ,

$$J(x_* + y) - J(x_*) = g(1) - g(0) = \frac{1}{2} g''(\lambda),$$

by the second-order mean value theorem of Lagrange (also known as the Taylor expansion), since  $g'(0) = 0$  ( $x_*$  being an extremal). Calculating as we did in the proof of Theorem 14.7, we find

$$g''(\lambda) = \int_a^b [\Lambda_{vv}^\lambda(t) y'^2 + 2\Lambda_{xv}^\lambda(t) y y' + \Lambda_{xx}^\lambda(t) y^2] dt, \quad (2)$$

where, for example,  $\Lambda_{vv}^\lambda(t)$  is shorthand for

$$\Lambda_{vv}(t, x_* + \lambda y, x_*' + \lambda y').$$

The partial derivatives  $\Lambda_{vv}$ ,  $\Lambda_{xv}$  and  $\Lambda_{xx}$ , being continuously differentiable, admit a common Lipschitz constant  $K$  on the ball around  $(0, 0, 0)$  of radius

$$|a| + |b| + \|x_*\| + \|x_*'\| + 1.$$

This allows us to write, for any variation  $y$  as described in the statement of the lemma,

$$|\Lambda_{vv}^\lambda(t) - \Lambda_{vv}^0(t)| \leq K |\lambda| |(y(t), y'(t))| \leq K |(y(t), y'(t))|,$$

and similarly for the other two terms in (2). This leads to

$$\begin{aligned} J(x_* + y) - J(x_*) &\geq \frac{1}{2} \int_a^b [\Lambda_{vv}^0(t) y'^2 + 2\Lambda_{xv}^0(t) y y' + \Lambda_{xx}^0(t) y^2] dt \\ &\quad - 2[\|y\| + \|y'\|] \int_a^b [y'^2 + 2|y y'| + y^2] dt. \end{aligned}$$

The proof of Theorem 14.7 showed that the first term on the right is precisely

$$\frac{1}{2} \int_a^b [Py'^2 + Qy^2] dt.$$

To complete the proof of the lemma, therefore, it now suffices to know that for some constant  $c$ , we have

$$\int_a^b [2|yy'| + y^2] dt \leq c \int_a^b y'^2 dt.$$

This consequence of the Cauchy-Schwartz inequality is entrusted to the reader as an exercise.

We now complete the proof of the theorem. We proceed to pick  $\delta > 0$  sufficiently small so that  $P(t) > \delta$  on  $[a, b]$ , and (calling upon a known fact in differential equations), also so that the following differential equation admits a solution  $w_\delta$ :

$$w_\delta'(t) = \frac{w_\delta(t)^2}{P(t) - \delta} - Q(t), \quad t \in [a, b].$$

(The hypothesized existence of the solution  $w = w_0$  for  $\delta = 0$  is crucial here for being able to assert that the solution  $w_\delta$  of the perturbed equation exists for suitably small  $\delta$ ; see [28].) We then pick any variation  $y$  satisfying

$$\|y\| + \|y'\| \leq \min(\delta/4M, 1).$$

We then find (exactly as in Legendre's argument)

$$\int_a^b [(P - \delta)y'^2 + Qy^2] dt = \int_a^b (P - \delta) \{y' + (w_\delta y)/(P - \delta)\}^2 dt \geq 0.$$

Applying the lemma, we deduce

$$\begin{aligned} J(x_* + y) - J(x_*) &\geq \frac{1}{2} \int_a^b [Py'^2 + Qy^2] dt - \frac{\delta}{4} \int_a^b y'^2 dt \\ &= \frac{1}{2} \int_a^b [(P - \delta)y'^2 + Qy^2] dt + \frac{\delta}{4} \int_a^b y'^2 dt \geq 0. \quad \square \end{aligned}$$

**14.9 Example.** The function  $x_* \equiv 0$  is an admissible extremal for the problem

$$\min \int_0^1 \left\{ \frac{1}{2} x'(t)^2 + t x'(t) \sin x(t) \right\} dt : x \in C^2[0, 1], \quad x(0) = x(1) = 0,$$

as the reader may verify. We wish to show that it provides a weak local minimum. We calculate

$$P(t) = 1, \quad Q(t) = -1.$$

Thus, the strengthened Legendre condition holds, and it suffices (by Theorem 14.8) to exhibit a solution  $w$  on  $[0, 1]$  of the differential equation  $w' = w^2 + 1$ . The function  $w(t) = \tan t$  serves the purpose. Note that this  $w$  would fail to be defined, however, if the underlying interval were  $[0, 2]$  (say).  $\square$

The following exercise demonstrates that merely *pointwise* conditions such as the Euler equation and the strengthened Legendre condition cannot always imply optimality; some extra element must be introduced which refers to the interval  $[a, b]$

or, more precisely, its length. We arrive in this way at a new idea: extremals can be optimal *in the short term*, without being optimal globally on their entire interval of definition.

**14.10 Exercise. (Wirtinger's inequality)** We study the putative inequality

$$\int_0^T x^2(t) dt \leq \int_0^T x'(t)^2 dt$$

for smooth functions  $x : [0, T] \rightarrow \mathbb{R}$  which vanish at 0 and  $T$  (here,  $T > 0$  is fixed). We rephrase the issue as follows: we study whether the function  $x_* \equiv 0$  solves the following special case of the problem (P):

$$\min J(x) = \int_0^T [x'^2 - x^2] dt : \text{subject to } x \in C^2[a, b], x(0) = x(T) = 0. \quad (*)$$

- (a) Show that, whatever the value of  $T$ , the function  $x_*$  is an extremal and satisfies the strengthened Legendre condition on  $[0, T]$ .
- (b) For any  $x$  admissible for  $(*)$ , we have  $J(\lambda x) = \lambda^2 J(x) \quad \forall \lambda > 0$ . Deduce from this homogeneity property that  $x_*$  is a weak local minimizer for  $(*)$  if and only if it is a global minimizer.
- (c) Let  $T \leq 1$ , and let  $x$  be admissible for  $(*)$ . Use the Cauchy-Schwarz inequality to prove

$$|x(t)|^2 \leq \int_0^T x'(s)^2 ds, \quad t \in [0, T].$$

Deduce that  $J(x) \geq 0 = J(x_*)$ , so that  $x_*$  does solve problem  $(*)$ .

- (d) Now let  $T \geq 4$ . Show that the function  $x$  defined by  $x(0) = 0$  and

$$x'(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1 \\ 0 & \text{if } 1 < t < T-1 \\ -1 & \text{if } T-1 \leq t \leq T \end{cases}$$

satisfies  $J(x) < 0$ . Note that  $x$  is Lipschitz, but not  $C^2$ ; however, approximation (see Exer. 21.13) leads to the conclusion that for  $T \geq 4$ , the extremal  $x_*$  *fails* to solve  $(*)$ .

It follows that the optimality of  $x_*$  for  $(*)$  *ceases* to hold at some value of  $T$  between 1 and 4. We surmise that it must be a notable number; we shall identify it in due course.  $\square$

**Conjugate points.** We know from local existence theorems that the differential equation that appears in the statement of Theorem 14.8 admits a solution  $w$  on an interval  $[a, a + \varepsilon]$ , for some  $\varepsilon > 0$  (in the presence of the other hypotheses). It follows that the extremal  $x_*$ , restricted to  $[a, a + \varepsilon]$ , is a weak local minimizer for

the truncated version of the basic problem for which it is admissible (that is, the problem whose boundary values at  $a$  and  $a + \varepsilon$  are those of  $x_*$ ). The difficulty is that  $a + \varepsilon$  may have to be strictly less than  $b$ .

A half-century after Legendre, Jacobi found a way to calibrate the extent of an extremal's optimality. We now examine this theory, which is based upon a certain second-order differential equation. Let  $x_*$  be an extremal on  $[a, b]$ , and let  $P$  and  $Q$  be defined as before:

$$P(t) = \Lambda_{vv}(t, x_*(t), x_*'(t))$$

$$Q(t) = \Lambda_{xx}(t, x_*(t), x_*'(t)) - \frac{d}{dt} \Lambda_{xv}(t, x_*(t), x_*'(t)).$$

The **Jacobi equation** corresponding to  $x_*$  is the following second-order differential equation:

$$-\frac{d}{dt} \{P(t)u'(t)\} + Q(t)u(t) = 0, \quad u \in C^2[a, b].$$

The somewhat unusual way in which this differential equation is expressed (as in a classical Sturm-Liouville problem) is traditional.

**14.11 Definition.** *The point  $\tau$  in  $(a, b]$  is said to be conjugate to  $a$  (relative to the given extremal  $x_*$ ) if there is a nontrivial solution  $u$  of the associated Jacobi equation which satisfies  $u(a) = u(\tau) = 0$ .*

In seeking conjugate points, it turns out that any nontrivial solution  $u$  of Jacobi's equation vanishing at  $a$  can be used: any other such  $u$  will generate the same conjugate points (if any). Let us see why this is so. Consider two such functions  $u_1$  and  $u_2$ . We claim that  $u_1'(a) \neq 0$ . The reason for this is that the only solution  $u$  of Jacobi's equation (a linear homogeneous differential equation of order two) satisfying  $u(a) = u'(a) = 0$  is the zero function (by the well-known uniqueness theorem for the initial-value problem). Similarly, we have  $u_2'(a) \neq 0$ . It follows that for certain nonzero constants  $c, d$ , we have

$$cu_1'(a) + du_2'(a) = 0.$$

But then the function  $u := cu_1 + du_2$  is a solution of Jacobi's equation which vanishes, together with its derivative, at  $a$ . Thus  $u \equiv 0$ ; that is,  $u_2$  is a nonzero multiple of  $u_1$ . It follows, then, that  $u_1$  and  $u_2$  have the same zeros, and hence determine the same conjugate points.

A nontrivial solution  $u$  of Jacobi's equation that vanishes at  $a$  has a *first* zero  $\tau > a$ , if it has one at all (for otherwise we find  $u'(a) = 0$ , a contradiction). Thus, it makes sense to speak of the *nearest* conjugate point  $\tau$  (if any), which is located at a strictly positive distance to the right of  $a$ .

In the study of conjugate points, it is always assumed that the underlying extremal  $x_*$  satisfies the strengthened Legendre condition:  $P(t) > 0 \quad \forall t \in [a, b]$ .

**14.12 Theorem. (Jacobi 1838)** *Let  $x_* \in C^2[a, b]$  be an extremal of the basic problem (P) which satisfies the boundary conditions of (P) as well as the strengthened Legendre condition. Then*

- (a) (Necessary condition) *If  $x_*$  is a weak local minimizer for (P), there is no conjugate point to  $a$  in the interval  $(a, b)$ .*
- (b) (Sufficient condition) *Conversely, if there is no point conjugate to  $a$  in the interval  $(a, b]$ , then  $x_*$  is a weak local minimizer for (P).*

**Proof.** The proof of necessity is postponed (see Ex. 15.6). We prove here the sufficiency. Accordingly, let  $x_*$  be an admissible extremal satisfying the strengthened Legendre condition, and admitting no conjugate point in  $(a, b]$ .

**Lemma.** *Jacobi's equation*

$$-\frac{d}{dt}\{P(t)u'(t)\} + Q(t)u(t) = 0$$

*admits a solution  $\bar{u}$  on  $[a, b]$  which is nonvanishing.*

To see this, let us consider first the solution  $u_0$  on  $[a, b]$  of Jacobi's equation with initial condition  $u(a) = 0$ ,  $u'(a) = 1$  (such a solution exists because of the linearity of Jacobi's equation).

Since there is no conjugate point in the interval  $(a, b]$  (by hypothesis), it follows that  $u_0$  is nonvanishing on the interval  $(a, b]$ . Because  $u'_0$  is continuous, we can therefore find  $\varepsilon > 0$  and  $d \in (a, b)$  such that

$$u'_0(t) > \varepsilon \quad (t \in [a, d]), \quad |u_0(t)| > \varepsilon, \quad (t \in [d, b]).$$

Now consider the solution  $u_\eta$  of the Jacobi equation satisfying the initial conditions  $u(a) = \eta$ ,  $u'(a) = 1$ , where  $\eta$  is a small positive parameter. According to the "imbedding theorem" (see [28]) whereby solutions of differential equations depend continuously upon initial conditions, for  $\eta$  sufficiently small we have

$$|u'_\eta(t) - u'_0(t)| < \frac{\varepsilon}{2}, \quad |u_\eta(t) - u_0(t)| < \frac{\varepsilon}{2}, \quad t \in [a, b].$$

(In order to apply the imbedding theorem, we rewrite the second-order differential equation as a system of two first-order equations, in the usual way; the positivity of  $P$  is used in this.) Note that  $u_\eta$  clearly cannot vanish on  $[d, b]$ ; it also cannot vanish on  $[a, d]$ , since we have  $u_\eta(a) > 0$  and  $u'_\eta > 0$  on  $[a, d]$ . Thus  $u_\eta$  is nonvanishing on  $[a, b]$ . This proves the lemma: take  $\bar{u} = u_\eta$ .

We now complete the proof of the theorem. We set

$$w(t) = -\bar{u}'(t)P(t)/\bar{u}(t),$$

which is possible because  $\bar{u}$  is nonvanishing. This clever change of variables was discovered by Jacobi; it linearizes the differential equation that appears in the statement of Theorem 14.8, in the sense that, as follows by routine calculation, the resulting function  $w$  satisfies

$$w'(t) = \frac{w(t)^2}{P(t)} - Q(t), \quad t \in [a, b].$$

Thus  $x_*$  is revealed to be a weak local minimizer, by Theorem 14.8.  $\square$

**14.13 Corollary.** *Let  $x_* \in C^2[a, b]$  be an extremal of the basic problem (P) which satisfies the boundary conditions of (P) as well as the strengthened Legendre condition. Suppose there exists a solution  $u$  of Jacobi's equation which does not vanish on  $[a, b]$ . Then  $x_*$  is a weak local minimizer.*

**Proof.** As shown in the proof of Theorem 14.12,  $u$  induces a solution of the differential equation that appears in the statement of Theorem 14.8; thus, the conclusion follows from that result.  $\square$

**14.14 Example.** Consider the basic problem

$$\min \int_0^1 x'(t)^3 dt : x \in C^2[0, 1], x(0) = 0, x(1) = 1.$$

The Euler equation is

$$\frac{d}{dt} \{ 3x'^2 \} = 0,$$

which implies that  $x'$  is constant. Thus, the unique admissible extremal is  $x_*(t) = t$ . The corresponding functions  $P$  and  $Q$  are given by  $P(t) = 6$  and  $Q = 0$ . It follows that the strengthened Legendre condition holds along  $x_*$ , and the Jacobi equation is

$$-\frac{d}{dt} \{ 6u' \} = 0.$$

A suitable solution of this equation (for finding conjugate points) is  $u(t) = t$ . Since this has no zeros in  $(0, 1]$ , and hence generates no conjugate points, we deduce from Theorem 14.12 that  $x_*$  provides a weak local minimum.  $\square$

**14.15 Exercise.** Prove that the problem considered in Example 14.14 admits no weak local maximizer.  $\square$

In the classical situations of mechanics, extremals of the action functional satisfy the strengthened Legendre condition, because the kinetic energy term is positive definite and quadratic. Jacobi's theorem therefore confirms the principle of least action as a true minimization assertion: the action is in fact minimized *locally* and *in the short term* by any extremal (relative to the points that it joins).

**14.16 Exercise.** Consider the soap bubble problem (Example 14.5), with

$$[a, b] = [0, T], \quad (T > 0), \quad A = 1, \quad B = \cosh T.$$

Let  $x_*$  be the catenary  $\cosh t$ . Prove that, if  $T$  is sufficiently small, then  $x_*$  provides a weak local minimum for the problem.  $\square$

**14.17 Exercise. (Wirtinger's inequality, continuation)** Find Jacobi's equation for the extremal and Lagrangian of the problem arising from the Wirtinger inequality (Exer. 14.10). Show that the point  $\tau = \pi$  (a notable number) is the first point conjugate to  $a = 0$ . Deduce the following result:<sup>4</sup>

**Proposition.** Let  $T \in [0, \pi]$ . Then, for any  $x \in C^2[0, T]$  which vanishes at 0 and  $T$ , we have

$$\int_0^T x(t)^2 dt \leq \int_0^T x'(t)^2 dt.$$

The general inequality fails if  $T > \pi$ .  $\square$

**14.18 Exercise.** We study the following problem in the calculus of variations:

$$\text{minimize } \int_0^T e^{-\delta t} (x'(t)^2 - x(t)^2) dt : x \in C^2[0, T], \quad x(0) = 0, \quad x(T) = 0,$$

where  $T > 0$  and  $\delta \geq 0$  are given.

- Show that  $x_* \equiv 0$  is a weak local minimizer when  $\delta \geq 2$ , for any value of  $T > 0$ .
- If  $\delta \geq 2$ , prove that  $x_*$  is in fact a global minimizer. [Hint:  $J(\lambda x) = \lambda^2 J(x)$ .]
- When  $\delta < 2$ , show that, for a certain  $\tau > 0$ ,  $x_*$  is a local minimizer if  $T < \tau$ , but fails to provide a local minimum if  $T > \tau$ .  $\square$

### 14.3 Two variants of the basic problem

We conclude this chapter's survey of the classical theory with two well-known variants of the underlying problem.

**The transversality condition.** In certain variational problems, the endpoint values of the competing functions  $x$  are not fully prescribed. In such cases, the extra flexibility at the boundary gives rise to additional conclusions in the necessary conditions, conclusions that say something about the initial and/or final values of  $x$ . These are known as *transversality conditions*.

<sup>4</sup> See Exer. 21.14 for an equivalent version of Wirtinger's inequality, one that is formulated on intervals of arbitrary length.



The following provides a simple example. We consider the problem of minimizing

$$\ell(x(b)) + \int_a^b \Lambda(t, x(t), x'(t)) dt$$

over the functions  $x \in C^2[a, b]$  satisfying the initial condition  $x(a) = A$ . The given function  $\ell$  (which we take to be continuously differentiable) corresponds to an extra cost term that depends on the (now unprescribed) value of  $x(b)$ .

**14.19 Theorem.** *Let  $x_*$  be a weak local minimizer of the problem. Then  $x_*$  is an extremal for  $\Lambda$ , and  $x_*$  also satisfies the following transversality condition:*

$$-\Lambda_v(b, x_*(b), x_*'(b)) = \ell'(x_*(b)).$$

The main point to be retained is that the extra information provided by the transversality condition exactly compensates for the fact that  $x_*(b)$  is now unknown. Thus the overall balance between known and unknown quantities is preserved: there is *conservation of information*. We shall see other instances later of this general principle for necessary conditions.

**Proof.** It is clear that  $x_*$  is a weak local minimizer for the version of the original problem (P) in which we impose the final constraint corresponding to  $B := x_*(b)$ . Thus,  $x_*$  is an extremal by Theorem 14.2.

Let us now choose any function  $y \in C^2[a, b]$  for which  $y(a) = 0$  (but leaving  $y(b)$  unspecified). We define  $g$  as follows:

$$g(\lambda) = \ell(x_*(b) + \lambda y(b)) + J(x_* + \lambda y).$$

It follows that  $g$  has a local minimum at  $\lambda = 0$ ; thus  $g'(0) = 0$ . As in the proof of Theorem 14.2, this leads to

$$\ell'(x_*(b))y(b) + \int_a^b [\alpha(t)y(t) + \beta(t)y'(t)] dt = 0.$$

Since  $\alpha = \beta'$  (as we know from the Euler equation), integration by parts shows that the integral is equal to  $\beta(b)y(b)$ . We derive therefore

$$[\ell'(x_*(b)) + \beta(b)]y(b) = 0.$$

Since  $y(b)$  is arbitrary, we deduce  $\ell'(x_*(b)) + \beta(b) = 0$ , which is the desired conclusion.  $\square$

**14.20 Exercise.** Given that the following problem has a unique solution  $x_*$ :

$$\text{minimize } \int_0^3 \left(\frac{1}{2}x'(t)^2 + x(t)\right) dt : x \in C^2[0, 3], x(0) = 0,$$

show that  $x_*$  is of the form  $t^2/2 + ct$ . Use the transversality condition to show that  $c = -3$ . Find the solution when the problem is modified as follows:

$$\text{minimize } x(3) + \int_0^3 \left( \frac{1}{2} x'(t)^2 + x(t) \right) dt : x \in C^2[0,3], x(0) = 0. \quad \square$$

**The isoperimetric problem.** This phrase refers to the classical problem of minimizing the same functional  $J(x)$  as in the basic problem (P), under the same boundary conditions, but under an additional equality constraint defined by a functional of the same type as  $J$ :

$$\int_a^b \psi(t, x(t), x'(t)) dt = 0.$$

The method of multipliers was introduced in this context by Euler (yes, in that very same monograph), but, as we know, is most often named after Lagrange, who made systematic use of it in his famous treatise on mechanics. Because of our experience with the multiplier rule in optimization, the reader will find that the next result has a familiar look.

**14.21 Theorem.** *Let  $x_* \in C^2[a, b]$  be a weak local minimizer for the isoperimetric problem, where  $\Lambda$  and  $\psi$  are  $C^2$ . Then there exists  $(\eta, \lambda) \neq 0$ , with  $\eta = 0$  or 1, such that  $x_*$  is an extremal for the Lagrangian  $\eta\Lambda + \lambda\psi$ .*

**Proof.** We merely sketch the proof; a much more general multiplier rule will be established later. The idea is to derive the conclusion from Theorem 9.1, for the purposes of which we define

$$f(x) = J_\Lambda(x_* + x), \quad h(x) = J_\psi(x_* + x),$$

where  $J_\Lambda$  and  $J_\psi$  are the integral functionals corresponding to  $\Lambda$  and  $\psi$ , and where  $x$  lies in the vector space  $X = C_0^2[a, b]$  consisting of those elements in  $C^2[a, b]$  which vanish at  $a$  and  $b$ . We may conveniently norm  $X$  by  $\|x\|_X = \|x''\|_\infty$ , turning it into a Banach space.

Then  $f(x)$  attains a local minimum in  $X$  relative to  $h(x) = 0$  at  $x = 0$ . It is elementary to show that  $f$  and  $h$  are continuously differentiable. It follows from Theorem 9.1 that for  $(\eta, \lambda)$  as described, we have  $(\eta f + \lambda h)'(0; y) = 0$  for every  $y \in X$ . As in the proof of Theorem 14.2, this implies that  $x_*$  is an extremal of  $\eta\Lambda + \lambda\psi$ .  $\square$

**14.22 Exercise.** A homogeneous chain of length  $L$  is attached to two points  $(a, A)$  and  $(b, B)$ , where  $a < b$ . Hanging in equilibrium, it describes a curve  $x(t)$  which minimizes the potential energy, which can be shown to be of the form

$$\sigma \int_a^b x(t) \sqrt{1 + x'(t)^2} dt,$$

where  $\sigma > 0$  is the (constant) mass density. Thus, we seek to minimize this functional relative to all curves in  $C^2[a, b]$  having the same endpoints and the same length:

$$x(a) = A, \quad x(b) = B, \quad \int_a^b \sqrt{1 + x'(t)^2} dt = L.$$

We assume that the chain is long enough to make the problem meaningful:  $L$  is greater than the distance between  $(a, A)$  and  $(b, B)$ .

Show that if  $x_*$  is a solution to the problem, then, for some constant  $\lambda$ , the function  $x_* + \lambda$  is an extremal of the Lagrangian  $x\sqrt{1 + v^2}$ . Invoke Example 14.5 to reveal that  $x_*$  is a translate of a catenary.  $\square$

We remark that the problem treated in the exercise above is a continuous version of the discrete one considered in Exer. 13.5. The conclusion explains the etymology: “catena” means “chain” in Latin.

**14.23 Exercise.** Assuming that a solution exists (this will be confirmed later), solve the following isoperimetric problem:

$$\min \int_0^\pi x'(t)^2 dt : x \in C^2[0, \pi], \quad \int_0^\pi x(t)^2 dt = \pi/2, \quad x(0) = x(\pi) = 0.$$

(The analysis is continued in Exer. 16.12.)  $\square$

**14.24 Exercise.** By examining the necessary conditions for the problem

$$\min \int_0^\pi x(t)^2 dt : x \in C^2[0, \pi], \quad \int_0^\pi x'(t)^2 dt = \pi/2, \quad x(0) = x(\pi) = 0,$$

show that it does *not* admit a solution. What is the infimum in the problem?  $\square$