

Chapter 16

Absolutely continuous solutions

The theory of the calculus of variations at the turn of the twentieth century lacked a critical component: it had no existence theorems. These constitute an essential ingredient of the *deductive method* for solving optimization problems, the approach whereby one combines existence, rigorous necessary conditions, and examination of candidates to arrive at a solution. (The reader may recall that in Chapter 9, we discussed at some length the relative merits of the inductive and deductive approaches to optimization.)

The deductive method, when it applies, often leads to the conclusion that a *global* minimum exists. Contrast this, for example, to Jacobi's theorem 14.12, which asserts only the existence of a local minimum. In mechanics, a local minimum is a meaningful goal, since it generally corresponds to a stable configuration of the system. In many modern applications however (such as in engineering or economics), only global minima are of real interest.

Along with the quest for the multiplier rule (which we discuss in the next chapter), it was the longstanding question of existence that dominated the scene in the calculus of variations in the first half of the twentieth century.¹

16.1 Example. Suppose that we are asked to solve the following instance of the basic problem:

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

We proceed to apply the necessary conditions for a solution x_* . Since the problem is autonomous, the Erdmann condition (Prop. 14.4) implies

$$(1+x_*(t))x_*'(t)^2 = c, \quad t \in [0,1]$$

¹ One of Hilbert's famous problems, in the list that he composed in 1900, concerned this issue; another concerned the regularity of solutions.

for some constant c . If $c = 0$, then x_* cannot leave its initial value 0, and thus $x_* \equiv 0$, a contradiction. We deduce therefore that $c > 0$. It follows that $x_*' \neq 0$, and therefore $x_*'(t) > 0$ for all t , and x_* satisfies

$$\sqrt{1+x_*(t)} x_*'(t) = \sqrt{c}, \quad t \in [0,1].$$

We easily solve this separable differential equation and invoke the boundary conditions to find the unique extremal

$$x_*(t) = (7t+1)^{2/3} - 1,$$

with associated cost $J(x_*) = (14/3)^2$. Having achieved this, we may very well have the impression of having solved the problem.

This is far from being the case, however, since the infimum in the problem is $-\infty$. This can be seen as follows. Consider a function y that is affine between $(0,0)$ and $(1/3,-3)$, and also between $(2/3,-3)$ and $(1,3)$. Between $(1/3,-3)$ and $(2/3,-3)$, we take y to be of “sawtooth” form, with values of y between -4 and -2 , and derivative y' satisfying $|y'| = M$ a.e. By taking M increasingly large, we can arrange for $J(y)$ to approach $-\infty$. (The sawtooth function can be approximated in order to achieve the same conclusion using functions in $C^2[a,b]$.)

The serious mistake in the analysis consists of assuming that a solution exists. The purpose of the example is to demonstrate the fallacy of using deductive reasoning when we don't know this to be true.

The sufficient conditions of Theorem 14.12 can be applied to show that x_* is a weak local minimizer; this is an inductive approach. It is also possible (deductively) to show that the function x_* found above is the unique global minimizer when the state constraint $x(t) \geq 0$ is added to the problem (see Exer. 21.16). \square

The key step in developing existence theory is to extend the context of the basic problem to functions that belong to the larger class $AC[a,b]$ of absolutely continuous functions, rather than $C^2[a,b]$ or $Lip[a,b]$ as in the preceding sections. Of course, this step could not be taken until Lebesgue had done his great work.

As we did in Chapter 12, we refer to an absolutely continuous function x mapping an interval $[a,b]$ to \mathbb{R}^n as an **arc**; recall that the notation $AC[a,b]$ is used for arcs on $[a,b]$, even in the vector-valued case. The fact that an arc x has a derivative x' that may be unbounded means that we have to pay some attention to whether $J(x)$ is well defined. (Under our previous hypotheses, this was automatic.)

The phrase “basic problem” now refers to

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

An arc x is **admissible** if it satisfies the constraints of the problem, and if $J(x)$ is well defined and finite. An arc x_* admissible for the basic problem (P) is said to be a solution (or a minimizer) if $J(x_*) \leq J(x)$ for all other admissible arcs x .

16.1 Tonelli's theorem and the direct method

The celebrated theorem of Tonelli identifies certain hypotheses under which a solution to (P) exists in the class of arcs. It features a Lagrangian Λ that is continuous and bounded below. Note that, in this case, for any arc x , the function $t \mapsto \Lambda(t, x(t), x'(t))$ is measurable and bounded below. In this setting, then, the integral $J(x)$ is well defined for any arc x , possibly as $+\infty$.

The following result, a turning point in the theory, was the concrete predecessor of the abstract direct method depicted in Theorem 5.51. Note, however, that the functional J is *not* necessarily convex here, which adds to the complexity of the problem.

16.2 Theorem. (Tonelli 1915) *Let the Lagrangian $\Lambda(t, x, v)$ be continuous, convex in v , and coercive of degree $r > 1$: for certain constants $\alpha > 0$ and β we have*

$$\Lambda(t, x, v) \geq \alpha |v|^r + \beta \quad \forall (t, x, v) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^n.$$

Then the basic problem (P) admits a solution in the class $AC[a, b]$.

Proof. It is clear that there exist admissible arcs x for which $J(x)$ is finite (for example, take x to be the unique affine admissible arc). Accordingly, there exists a minimizing sequence x_i of admissible functions for (P):

$$\lim_{i \rightarrow \infty} J(x_i) = \inf(\text{P}) \quad (\text{finite}).$$

For all i sufficiently large, in view of the coercivity, we have

$$\int_a^b \{ \alpha |x'_i|^r + \beta \} dt \leq \int_a^b \Lambda(t, x_i, x'_i) dt \leq \inf(\text{P}) + 1.$$

This implies that the sequence x'_i is bounded in $L^r[a, b]^n$. By reflexivity and weak sequential compactness, we may assume without loss of generality that each component converges weakly in $L^r[a, b]$; we label the vector limit v_* .

We proceed to define an element x_* of $AC^r[a, b]^n$ via

$$x_*(t) = A + \int_a^t v_*(s) ds, \quad t \in [a, b].$$

For each $t \in [a, b]$, the weak convergence implies that

$$\int_a^t x_i'(s) ds = \int_a^b x_i'(s) \chi_{(a,t)}(s) ds \rightarrow \int_a^b v_*(s) \chi_{(a,t)}(s) ds = \int_a^t v_*(s) ds$$

(where $\chi_{(a,t)}$ is the characteristic function of the interval (a,t)), from which we deduce that $x_i(t)$ converges pointwise to $x_*(t)$. (The convergence can be shown to be uniform.) It follows that $x_*(b) = B$, so that x_* is admissible for (P).

We now proceed to invoke the integral semicontinuity theorem 6.38, taking Q to be $[a,b] \times \mathbb{R}^n$, $z_i = x_i'$ and $u_i = x_i$; the convexity of Λ in v is essential here. (Of course, Tonelli does not refer to convex functions in his proof.) We conclude that

$$J(x_*) \leq \lim_{i \rightarrow \infty} J(x_i) = \inf (P).$$

Since x_* is admissible for (P), it follows that x_* is a global minimizer. \square

16.3 Exercise. In the following, we outline a more elementary proof of Tonelli's theorem in the case of a "separated" Lagrangian having the form

$$\Lambda(t, x, v) = f(t, x) + g(t, v).$$

The technicalities are sharply reduced in this case (since no appeal to Theorem 6.38 is required), but the main ideas are the same. We assume (consistently with Theorem 16.2) that f and g are continuous, g is convex in v , f is bounded below, and that we have (for some $\alpha > 0$ and $r > 1$)

$$g(t, v) \geq \alpha |v|^r \quad \forall (t, v) \in [a, b] \times \mathbb{R}^n.$$

- (a) Prove that a minimizing sequence x_i exists, and that x_i' is bounded in $L^r(a, b)$.
- (b) Prove that the sequence x_i is bounded and equicontinuous, and that, for some subsequence (not relabeled), there is an arc x_* admissible for (P) such that

$$x_i \rightarrow x_* \text{ uniformly, } x_i' \rightarrow x_*' \text{ weakly in } L^r(a, b).$$

- (c) Prove that

$$\int_a^b f(t, x_*(t)) dt = \lim_{i \rightarrow \infty} \int_a^b f(t, x_i(t)) dt.$$

- (d) Prove that the mapping

$$v \mapsto \int_a^b g(t, v(t)) dt$$

is lower semicontinuous on $L^r(a, b)$.

- (e) Prove that

$$\int_a^b g(t, x_*'(t)) dt \leq \liminf_{i \rightarrow \infty} \int_a^b g(t, x_i'(t)) dt$$

and then conclude that x_* is a solution of (P). \square

We proceed now to illustrate how Tonelli's theorem fails if either of the coercivity or the convexity hypotheses is absent. We begin with the convexity.

16.4 Example. Consider the basic problem with $n = 1$,

$$J(x) = \int_0^1 (x(t)^2 + [x'(t)^2 - 1]^2) dt,$$

and constraints $x(0) = 0$, $x(1) = 0$. Then the Lagrangian $\Lambda(t, x, v)$, which here is the function $x^2 + (v^2 - 1)^2$, is continuous, and is also coercive of degree 4. We claim that nonetheless, the problem has no solution.

To see this, note that $J(x) > 0$ for any arc x , since we cannot have *both* $x \equiv 0$ and $|x'(t)| = 1$ a.e. But the infimum in the problem is 0, since, for any positive ε , there is a sawtooth function x whose derivative is ± 1 a.e. and which satisfies $\|x\| < \varepsilon$ (whence $J(x) < \varepsilon^2$).

It is clearly the convexity hypothesis (in v) that is missing here, which is why Tonelli's theorem is inapplicable. Informally, we may say that the sawtooth function "chatters" between the derivative values ± 1 (the locus where $(v^2 - 1)^2$ attains its minimum) and "almost succeeds" in giving $x = 0$ as well (the locus where x^2 attains its minimum). But complete success is not possible: in the limit, x goes to 0, but x' converges weakly to 0 (see Exer. 6.19). \square

The following shows that the degree of coercivity in Tonelli's theorem cannot be lowered to $r = 1$.

16.5 Exercise. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be given by $g(v) = v[1 + \min\{v, 0\}]$.

(a) Show that g is convex, continuously differentiable, and satisfies

$$g(v) \geq \max\{v, |v| - 1\} \quad \forall v.$$

(b) Deduce that any arc x admissible for the problem

$$\min J(x) = \int_0^1 (x^2 + g(x')) dt : \text{subject to } x \in \text{AC}[0, 1], x(0) = 0, x(1) = 1$$

satisfies $J(x) > 1$.

(c) Show that the functions

$$x_i(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq 1 - 1/i \\ i[t - 1 + 1/i] & \text{if } 1 - 1/i < t \leq 1 \end{cases}$$

satisfy $\lim_{i \rightarrow \infty} J(x_i) \rightarrow 1$.

(d) Conclude that the problem defined in (b) admits no solution. What hypothesis of Tonelli's theorem is missing? \square

16.6 Exercise. Let $T > 0$ be fixed. Prove that the following problem has a solution:

$$\begin{aligned} \text{minimize } \int_0^T \left\{ \frac{1}{2} m (\ell \theta'(t))^2 - mg\ell(1 - \cos \theta(t)) \right\} dt : \\ \theta \in \text{AC}[0, T], \theta(0) = 0 = \theta(T). \end{aligned}$$

Note that the integral is the action of Example 14.6. Show that if T is large, the solution² spends a lot of time near the unstable equilibrium $\theta = \pi$. \square

Simple variants of the theorem. The proof of Tonelli's theorem adapts to numerous variants of the basic problem. Some typical ones appear below; they maintain the central elements (coercivity of the Lagrangian, and its convexity in v).

16.7 Exercise. Show that Tonelli's theorem holds when Λ is measurable in t and continuous in (x, v) (rather than continuous in all its variables). An extra hypothesis is now required, however, to ensure the existence of an admissible arc x for which $J(x) < \infty$ (and hence, of a minimizing sequence). A simple one that suffices: for every x and v in \mathbb{R}^n , the function $t \mapsto \Lambda(t, x + tv, v)$ is summable. \square

16.8 Exercise. Under the same hypotheses as the previous exercise, verify that the proof of Tonelli's theorem is unaffected by the presence in the underlying problem of a *unilateral state constraint*

$$x(t) \in S, \quad t \in [a, b],$$

where S is a given *closed* subset of \mathbb{R}^n . Note that in this setting, the coercivity of Λ need only hold when x lies in S . \square

16.9 Exercise. Extend Tonelli's theorem to the problem

$$\text{minimize } \ell(x(a), x(b)) + \int_a^b \Lambda(t, x(t), x'(t)) dt, \quad (x(a), x(b)) \in E,$$

where Λ satisfies the hypotheses of the previous exercise, ℓ is continuous, E is closed, and one of the following projections is bounded:

$$\{y \in \mathbb{R}^n : \exists z \in \mathbb{R}^n : (y, z) \in E\}, \quad \{z \in \mathbb{R}^n : \exists y \in \mathbb{R}^n : (y, z) \in E\}. \quad \square$$

16.10 Exercise. One may extend Tonelli's theorem to certain cases in which the Lagrangian is not necessarily bounded below. A simple instance of this is obtained by weakening the coercivity condition as follows:

$$\Lambda(t, x, v) \geq \alpha |v|^r - \gamma |x|^s + \beta \quad \forall (t, x, v) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^n,$$

² This example shows that the principle of least action does not describe physical reality in the long term, although, as we have seen, it does do so in the short term (and locally).

where $\gamma \geq 0$, $0 < s < r$, and where, as before, $r > 1$, $\alpha > 0$. Note that the (positive) growth in v is of higher order than the (possibly negative) growth in x . With this reduced growth condition, the other hypotheses on Λ being unchanged, show that Tonelli's theorem continues to hold. \square

The exercise above allows one to assert existence when the Lagrangian is given by $|v|^2 - |x|$, for example.

16.11 Exercise. Let Λ satisfy the hypotheses of Tonelli's theorem, except that the coercivity is weakened to

$$\Lambda(t, x, v) \geq \frac{\alpha |v|^r}{(1 + |x|)^s} + \beta \quad \forall (t, x, v) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^n,$$

where $0 < s < r$. Establish the existence of a solution to the basic problem. \square

The Direct Method. In view of the variants evoked above, the reader may well feel that some general (albeit ungainly) theorem might be fabricated to cover a host of special cases. Indeed, we could take a stab at this, but experience indicates that circumstances will inevitably arise which will not be covered. It is better to master the method, now known as the *direct method*. (Tonelli, upon introducing it, had called it *a direct method*.)

The underlying approach, then, combines three ingredients: the analysis of a minimizing sequence x_i in order to establish the existence of a subsequence converging in an appropriate sense; the lower semicontinuity of the cost with respect to the convergence; the persistence of the constraints after taking limits. The convergence is generally in the sense that x'_i converges weakly and x_i pointwise or uniformly; the lower semicontinuity typically results from the integral semicontinuity theorem 6.38. To deduce the persistence in the limit of the constraints, the weak closure theorem 6.39 is often helpful. This approach applies to a broad range of problems in dynamic optimization.

In applying the direct method, it is essential to grasp the distinction between constraints that persist under uniform convergence (of x_i) and those that survive weak convergence (of x'_i): convexity is the key to the latter. Consider, for example, the basic problem in the presence of two additional unilateral state and velocity constraints:

$$(a): x(t) \in S \quad \forall t \in [a, b], \quad \text{and} \quad (b): x'(t) \in V, \quad t \in [a, b] \text{ a.e.}$$

Uniform (or pointwise) convergence of the sequence x_i will preserve the constraint (a) in the limit, provided only that S is closed. In order for weak convergence to preserve the constraint (b) in the limit, however, we require that V be *convex* as well as closed (as in the weak closure theorem 6.39); compactness does not suffice (see Exer. 8.45). Thus, the appropriate hypotheses for an existence theorem in this context would include: S closed, V closed *and* convex.

As regards the hypotheses giving the weak sequential compactness of the sequence x'_i , we would only require the coercivity of Λ to hold for points $x \in S$. The coercivity could also be replaced by compactness of V . (Note that the integral semicontinuity theorem 6.38 does not require a coercive Lagrangian.) Hybrid possibilities can also be envisaged; for example, coercivity of Λ with respect to certain coordinates of v , and compactness of V with respect to the others.

An example of these considerations occurs in the second problem of Exer. 14.23, in which we saw that the problem

$$\text{minimize } \int_0^\pi x(t)^2 dt \quad \text{subject to } \int_0^\pi x'(t)^2 dt = \pi/2$$

does not admit a solution. In order that the isoperimetric constraint

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

be preserved in the limit, ψ needs to be linear with respect to v . On the other hand, the robustness under convergence of a constraint of the form

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

requires only convexity of ψ with respect to v (as in the integral semicontinuity theorem). In this setting, the required weak compactness could result from various hypotheses. The Lagrangian Λ itself being coercive (as before) would do; but we could require, instead, that the Lagrangian ψ of the isoperimetric constraint be coercive. In general, then, a combination of circumstances will come into play.

16.12 Exercise. Prove that the problem of Exer. 14.23 admits a solution if $C^2[0, \pi]$ is replaced by $AC[0, \pi]$. (Note, however, that Theorem 14.21 cannot be applied in order to identify it; existence has been achieved, but the necessary conditions have been lost. The analysis is continued in Exer. 17.11.) \square

16.2 Regularity via growth conditions

Now that we are armed with an existence theory, we would like to use it in the deductive method, the next step in which is the application of necessary conditions. In examining Theorem 15.2, which asserts the integral Euler equation for the basic problem (P), however, we spot a potential difficulty when absolutely continuous functions are involved. The proof invoked the dominated convergence theorem, which no longer seems to be available when x'_* is unbounded; it appears, in fact, that the differentiability of g cannot be asserted.

Is this problem in proving the existence of $g'(0)$ simply a technical difficulty in adapting the argument, or is it possible that the basic necessary condition for (P) can actually fail? It turns out to be the latter: even for an analytic Lagrangian satisfying the hypotheses of Tonelli's theorem, the integral Euler equation may not hold³ at the unique minimizing arc x_* . The reader may detect a certain irony here: in order to be able to apply the deductive approach, the basic problem has been extended to $AC[a, b]$. However, with solutions in this class, the necessary conditions can no longer be asserted.

Another disturbing fact about the extension to arcs is the possibility of the *Lavrentiev phenomenon*. This is said to occur when the infimum in the basic problem over $AC[a, b]$ is strictly less than the infimum over $Lip[a, b]$, and it can happen even for smooth Lagrangians satisfying the hypotheses of Tonelli's theorem. From the computational point of view, this is disastrous, since most numerical methods hinge upon minimizing the cost over a class of smooth (hence Lipschitz) functions (for example, polynomials). In the presence of the Lavrentiev phenomenon, such methods cannot approach the minimum over $AC[a, b]$. The extension from $Lip[a, b]$ to $AC[a, b]$, then, is not necessarily a faithful one (a completion), as was the extension from $C^2[a, b]$ to $Lip[a, b]$.

However, all is not lost. There is a way to recover, in many cases, the happy situation in which we can both invoke existence and assert the necessary conditions, while excluding the Lavrentiev phenomenon. This hinges upon identifying additional structural hypotheses on Λ which serve to rule out the pathological situations cited above. We shall see two important examples of how to do this. The first one below recovers the integral Euler equation under an "exponential growth" hypothesis on the Lagrangian.

Remark. Local minima in the class of arcs are defined essentially as before. For example, an admissible arc x_* is a weak local minimizer if, for some $\varepsilon > 0$, we have $J(x_*) \leq J(x)$ for all admissible arcs x satisfying $\|x - x_*\| \leq \varepsilon$ and $\|x' - x_*'\| \leq \varepsilon$. Recall that the meaning of the word "admissible" in the preceding sentence includes the integral being well defined.

16.13 Theorem. (Tonelli-Morrey) *Let Λ admit gradients Λ_x, Λ_v which, along with Λ , are continuous in (t, x, v) . Suppose further that for every bounded set S in \mathbb{R}^n , there exist a constant c and a summable function d such that, for all $(t, x, v) \in [a, b] \times S \times \mathbb{R}^n$, we have*

$$|\Lambda_x(t, x, v)| + |\Lambda_v(t, x, v)| \leq c(|v| + |\Lambda(t, x, v)|) + d(t). \quad (*)$$

Then any weak local minimizer x_ satisfies the Euler equation in integral form.*

Proof. It follows from the hypotheses on Λ that the function

³ The first examples of this phenomenon are quite recent, and exhibit the feature that the function $\Lambda_x(t, x_*(t), x_*'(t))$, which is the derivative of the costate, is not summable.

$$f(t) = \Lambda_v(t, x_*(t), x'_*(t)) - \int_a^t \Lambda_x(s, x_*(s), x'_*(s)) ds$$

lies in $L^1(a, b)^n$. Let $y : [a, b] \rightarrow \mathbb{R}^n$ be a function which belongs to $\text{Lip}[a, b]$, vanishes at a and b , and satisfies $\|y\| + \|y'\| \leq 1$. We prove that

$$\int_a^b f(t) \cdot y'(t) dt = 0$$

for any such y , which, by Example 9.5, implies the integral Euler equation.

For any $t \in [a, b]$ such that $x'_*(t)$ and $y'(t)$ exist, and such that $|y(t)| + |y'(t)| \leq 1$ (thus, for almost every t), we define

$$g(t, s) = \Lambda(t, x_*(t) + sy(t), x'_*(t) + sy'(t)) - \Lambda(t, x_*(t), x'_*(t)), \quad s \in [0, 1].$$

Note that $g(t, 0) = 0$ a.e., and, since x_* is a weak local minimizer,

$$\int_a^b g(t, s) dt \geq 0 \quad \text{for } s \text{ sufficiently small.} \quad (1)$$

The structural hypothesis (*) yields, for almost every t , for $s \in [0, 1]$ a.e.,

$$\begin{aligned} \left| \frac{d}{ds} g(t, s) \right| &\leq c \{ 1 + |x'_*(t)| + |\Lambda(t, x_*(t), x'_*(t))| + |g(t, s)| \} + d(t) \\ &= c|g(t, s)| + k(t), \end{aligned}$$

for a certain summable function k . This estimate, together with Gronwall's lemma, leads to $|g(t, s)| \leq sMk(t)$ for a certain constant M , for all s sufficiently small. In view of this, we may invoke Lebesgue's dominated convergence theorem to deduce (with the help of (1))

$$\begin{aligned} 0 &\leq \lim_{s \downarrow 0} \int_a^b \frac{g(t, s) - g(t, 0)}{s} dt = \int_a^b \frac{d}{ds} g(t, s) dt \\ &= \int_a^b \{ \Lambda_x(t, x_*(t), x'_*(t)) \cdot y(t) + \Lambda_v(t, x_*(t), x'_*(t)) \cdot y'(t) \} dt \\ &= \int_a^b f(t) \cdot y'(t) dt, \end{aligned}$$

after an integration by parts. Since y may be replaced by $-y$, equality must hold, and the proof is complete. \square

16.14 Exercise. Let $\Lambda(t, x, v)$ have the form $f(t, x) + g(v)$, where f and g are continuously differentiable and, for some constant c , the function g satisfies

$$|\nabla g(v)| \leq c(1 + |v| + |g(v)|) \quad \forall v \in \mathbb{R}^n.$$

Prove that any weak local minimizer for (P) satisfies the integral Euler equation. \square

Nagumo growth. The following localized and weakened form of coercivity is useful in regularity theory. It asserts that Λ has superlinear growth in v along x_* .

16.15 Definition. We say that Λ has Nagumo growth along x_* if there exists a function $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}$ satisfying $\lim_{t \rightarrow \infty} \theta(t)/t = +\infty$, such that

$$t \in [a, b], v \in \mathbb{R}^n \implies \Lambda(t, x_*(t), v) \geq \theta(|v|).$$

As an illustrative example, we observe that when the Lagrangian satisfies the hypothesis of Exer. 16.10, then Nagumo growth holds along any arc.

16.16 Corollary. Under the hypotheses of Theorem 16.13, if $\Lambda(t, x, v)$ is convex in v and has Nagumo growth along x_* , then x_* is Lipschitz.

Proof. Under the additional hypotheses, the costate p is the subgradient of Λ in v along x_* , whence

$$\Lambda(t, x_*(t), 0) - \Lambda(t, x_*(t), x_*'(t)) \geq -p(t) \bullet x_*'(t) \text{ a.e.}$$

Nagumo growth, along with this inequality, reveals:

$$\theta(|x_*'(t)|) \leq \Lambda(t, x_*(t), x_*'(t)) \leq \Lambda(t, x_*(t), 0) + p(t) \bullet x_*'(t) \text{ a.e.,}$$

which implies that x_*' is essentially bounded, since θ has superlinear growth and both $\Lambda(t, x_*(t), 0)$ and $p(t)$ are bounded. □

The desirability of Lipschitz regularity. Note that when the basic problem (P) admits a global solution x_* which is Lipschitz, then the Lavrentiev phenomenon does not occur, and the necessary conditions can be asserted. This is why the *Lipschitz regularity* of the solution is a desirable property. It offers the further advantage of giving us access to the higher regularity results of §15.2, which would allow us to deduce the smoothness of the solution

16.17 Exercise. Use the results above to prove that any solution θ_* to the problem of Exer. 16.6 is Lipschitz. Proceed to show by the results of §15.2 that θ_* is C^∞ . □

Remark. Future developments will make it possible to assert the Euler equation and Lipschitz regularity under a weaker growth condition than (*) of Theorem 16.13. Specifically, we shall obtain Theorem 16.13 and Cor. 16.16 in §17.3 under the following structural assumption: There exist constants $\varepsilon > 0$ and c , and a summable function d , such that, for almost every t ,

$$|\Lambda_x(t, x, x_*'(t))| \leq c |\Lambda(t, x, x_*'(t))| + d(t) \quad \forall x \in B(x_*(t), \varepsilon).$$

16.3 Autonomous Lagrangians

We now prove that under hypotheses of Tonelli type, solutions to the basic problem in the calculus of variations are Lipschitz when the Lagrangian is autonomous. The reader will recall that the problem (P), or its Lagrangian Λ , are said to be *autonomous* when Λ has no dependence on the t variable.

16.18 Theorem. (Clarke-Vinter) *Let $x_* \in AC[a, b]$ be a strong local minimizer for the problem (P), where the Lagrangian is continuous, autonomous, convex in v , and has Nagumo growth along x_* . Then x_* is Lipschitz.*

Proof. Let x_* be a solution of (P) relative to $\|x - x_*\| \leq \varepsilon$. By uniform continuity, there exists $\delta \in (0, 1/2)$ with the property that

$$t, \tau \in [a, b], |t - \tau| \leq (b - a)\delta/(1 - \delta) \implies |x_*(t) - x_*(\tau)| < \varepsilon.$$

A. Let us consider any measurable function $\alpha : [a, b] \rightarrow [1 - \delta, 1 + \delta]$ satisfying the equality $\int_a^b \alpha(t) dt = b - a$. For any such α , the relation

$$\tau(t) = a + \int_a^t \alpha(s) ds$$

defines a bi-Lipschitz one-to-one mapping from $[a, b]$ to itself; it follows readily that the inverse mapping $t(\tau)$ satisfies

$$\frac{d}{d\tau} t(\tau) = \frac{1}{\alpha(t(\tau))}, \quad |t(\tau) - \tau| \leq (b - a)\delta/(1 - \delta) \text{ a.e.}$$

Proceed now to define an arc y by $y(\tau) = x_*(t(\tau))$. Then y is admissible for the problem (P), and satisfies $\|y - x_*\| < \varepsilon$ (by choice of δ), whence

$$\int_a^b \Lambda(y(\tau), y'(\tau)) d\tau \geq J(x_*).$$

Applying the change of variables $\tau = \tau(t)$ to the integral on the left, and noting that $y'(\tau) = x'_*(t(\tau))/\alpha(t(\tau))$ a.e., we obtain

$$\begin{aligned} \int_a^b \Lambda(y(\tau), y'(\tau)) d\tau &= \int_a^b \Lambda(y(\tau(t)), y'(\tau(t))) \tau'(t) dt \\ &= \int_a^b \Lambda(x_*(t), x'_*(t)/\alpha(t)) \alpha(t) dt \geq J(x_*). \end{aligned}$$

Note that equality holds when α is the function $\alpha_* \equiv 1$, so we see that α_* solves a certain minimization problem. Let us formulate this problem more explicitly by introducing

$$\Phi(t, \alpha) = \Lambda(x_*(t), x'_*(t)/\alpha) \alpha$$

It is straightforward to verify that for each t , the function $\Phi(t, \cdot)$ is convex on the interval $(0, \infty)$. Consider the functional f given by

$$f(\alpha) = \int_a^b \Phi(t, \alpha(t)) dt.$$

Then $f(\alpha)$ is well defined when α is measurable and has values in the interval $[1 - \delta, 1 + \delta]$, possibly as $+\infty$, and it follows that f is convex.

For almost every t , by continuity, there exists $\delta(t) \in (0, \delta]$ such that

$$\Phi(t, 1) - 1 \leq \Phi(t, \alpha) \leq \Phi(t, 1) + 1 \quad \forall \alpha \in [1 - \delta(t), 1 + \delta(t)].$$

It follows from measurable selection theory (§6.2) that we may take $\delta(\cdot)$ measurable.⁴ We define S to be the convex subset of $X := L^\infty[a, b]$ whose elements α satisfy $\alpha(t) \in [1 - \delta(t), 1 + \delta(t)]$ a.e.

Consider now an optimization problem (Q) defined on the vector space X . It consists of minimizing f over S subject to the equality constraint

$$h(\alpha) = \int_a^b \alpha(t) dt - (b - a) = 0.$$

The argument given above shows that the function $\alpha_* \equiv 1$ solves (Q).

B. We now apply the multiplier rule, more precisely, the version given by Theorem 9.4. We obtain a nonzero vector $\zeta = (\eta, \lambda)$ in \mathbb{R}^2 (with $\eta = 0$ or 1) such that

$$\eta f(\alpha) + \lambda h(\alpha) \geq \eta f(\alpha_*) \quad \forall \alpha \in S.$$

It follows easily from Theorem 6.31 that $\eta = 1$. Rewriting the conclusion, we have, for any α in S , the inequality

$$\int_a^b \{ \Lambda(x_*(t), x'_*(t)/\alpha(t)) \alpha(t) + \lambda \alpha(t) \} dt \geq \int_a^b \{ \Lambda(x_*(t), x'_*(t)) + \lambda \} dt.$$

Invoking Theorem 6.31, we deduce that, for almost every t , the function

$$\alpha \mapsto \theta_t(\alpha) := \Lambda(x_*(t), x'_*(t)/\alpha) \alpha + \lambda \alpha$$

attains a minimum over the interval $[1 - \delta(t), 1 + \delta(t)]$ at the interior point $\alpha = 1$. Fix such a value of t . Then the generalized gradient of θ_t at 1 must contain zero. It follows from nonsmooth calculus (see Exer. 13.23, strangely relevant) that

$$\Lambda(x_*(t), x'_*(t)) - \langle x'_*(t), \zeta(t) \rangle = -\lambda \quad \text{a.e.}, \quad (1)$$

⁴ Because $(0, \delta]$ is not closed, there is an implicit (but not difficult) exercise involved here.

where $\zeta(t)$ lies in the subdifferential at $x_*'(t)$ of the convex function $v \mapsto \Lambda(x_*(t), v)$.

C. The last step in the proof is to show (using (1)) that $x_*'(t)$ is essentially bounded. Let t be such that $x_*'(t)$ exists, and such that (1) holds. We have

$$\begin{aligned} & \Lambda(x_*(t), x_*'(t) \{1 + |x_*'(t)|\}^{-1}) - \Lambda(x_*(t), x_*'(t)) \\ & \geq [\{1 + |x_*'(t)|\}^{-1} - 1] \langle x_*'(t), \zeta(t) \rangle \text{ (by the subgradient inequality)} \\ & = [\{1 + |x_*'(t)|\}^{-1} - 1] \{ \Lambda(x_*(t), x_*'(t)) + \lambda \}, \end{aligned}$$

in light of (1). Letting M be a bound for all values of Λ at points of the form $(x_*(t), w)$ with $t \in [a, b]$ and $w \in B$, this leads to (in view of the Nagumo growth)

$$\theta(|x_*'(t)|) \leq \Lambda(x_*(t), x_*'(t)) \leq M + (M + |\lambda|) |x_*'(t)|.$$

The superlinearity of θ implies that $|x_*'(t)|$ is essentially bounded, as required. \square

Remark. When Λ is taken to be differentiable in v , then the $\zeta(t)$ that appears in the proof is none other than the costate $p(t) = \Lambda_v(x_*(t), x_*'(t))$, and we see that (1) extends the Erdmann condition (see Prop. 14.4) to the current setting (with $h = \lambda$). It has now been obtained for x_* merely Lipschitz rather than C^2 ; the simple proof used before no longer pertains.

The reader may verify that the coercivity was not used to obtain the Erdmann condition, and that its proof goes through unchanged in the presence of an explicit state constraint $x(t) \in S$ in the problem (P), and also when a constraint $x'(t) \in C$ is imposed, provided that C is a cone. We summarize these observations:

16.19 Corollary. *Let Λ be continuous and autonomous and, with respect to v , be convex and differentiable. Let x_* be a strong local minimizer for the problem*

$$\text{minimize } J(x) : x \in AC[a, b], x(t) \in S, x'(t) \in C, x(a) = A, x(b) = B,$$

where S is a subset of \mathbb{R}^n and C is a cone in \mathbb{R}^n . Then, for some constant h , the arc x_* satisfies the **Erdmann condition**

$$\langle x_*'(t), \Lambda_v(x_*(t), x_*'(t)) \rangle - \Lambda(x_*(t), x_*'(t)) = h \text{ a.e.}$$

If in addition Λ has Nagumo growth along x_* , then x_* is Lipschitz.

The reader will notice that as a result of the above, and in contrast to Chapter 14, the Erdmann condition is now available as a separate necessary condition for optimality in certain situations in which the Euler equation cannot be asserted (because of the additional constraints, or simply because x_* is not known to be Lipschitz). Exers. 21.15 and 21.16 illustrate its use in such situations.

16.20 Exercise. Consider the problem of Exer. 16.6. Letting its solution be θ_* , show that the Erdmann condition asserts

$$\frac{1}{2}m(\ell\theta'_*(t))^2 + mg\ell(1 - \cos\theta_*(t)) = h.$$

Note that this corresponds to *conservation of energy*, which is often the interpretation of the Erdmann condition in classical mechanics. \square

16.21 Example. We illustrate now the use of the existence and regularity theorems, and also their role in studying boundary-value problems in ordinary differential equations. Consider the following version of the basic problem (P), with $n = 1$:

$$\min \int_0^T \left\{ \frac{x(t)^4}{4} - \frac{x(t)^2}{2} + \frac{x'(t)^2}{2} \right\} dt : x \in \text{AC}[0, T], x(0) = 0, x(T) = 0.$$

The Lagrangian

$$\Lambda(x, v) = v^2/2 + x^4/4 - x^2/2$$

is continuous and convex in v , and (it can easily be shown) coercive of degree 2. According to Tonelli's theorem, there exists a solution x_* of (P).

It follows now from Theorem 16.18 that x_* is Lipschitz, since Λ is autonomous. An alternative to calling upon Theorem 16.18 is to argue as follows. We have

$$\frac{|\Lambda_x| + |\Lambda_v|}{1 + |v| + |\Lambda(x, v)|} \leq \frac{|v| + |x|^3 + |x|}{1 + |v|} \leq 1 + |x|^3 + |x|,$$

which shows that the structural hypothesis (*) of Theorem 16.13 holds. This allows us to invoke Cor. 16.16 in order to conclude that x_* is Lipschitz.

In either case, it follows that x_* satisfies the integral Euler equation. We then appeal to Theorem 15.7 in order to deduce that $x_* \in C^\infty[0, T]$. This allows us to write the Euler equation in fully differentiated form:

$$x''(t) = x(t)(x(t)^2 - 1).$$

In summary, there is a solution x_* of the boundary-value problem

$$x''(t) = x(t)(x(t)^2 - 1), \quad x \in C^\infty[0, T], \quad x(0) = 0, \quad x(T) = 0, \quad (2)$$

one that also solves the problem (P).

However, it is clear that the zero function is a solution of (2), and we would wish to know when there is a *nontrivial* solution. This will certainly be the case if the zero function, which is evidently an extremal for Λ , admits a conjugate point in the interval $(0, T)$. For in that case, it cannot be a solution of (P), by the necessary condition of Jacobi (Theorem 14.12), whence $x_* \neq 0$.

The Jacobi equation (for the zero function) is $u''(t) + u(t) = 0$, which yields the conjugate point $\tau = \pi$. We arrive therefore at the following conclusion: there exists a nontrivial solution of (2) when $T > \pi$. \square

16.22 Exercise. We consider the following problem (P):

$$\text{minimize } \int_0^1 \exp \{ x(t) + x'(t)^2 \} dt : x \in AC[0,1], x(0) = 0, x(1) = 1.$$

- (a) Use the direct method to prove that (P) admits a solution.
- (b) Prove that (P) admits a unique solution.
- (c) Observe that the Lagrangian does *not* satisfy hypothesis (*) of Theorem 16.13.
- (d) Prove that the solution of (P) is Lipschitz.
- (e) Deduce the existence of a *unique* solution x to the following boundary-value problem:

$$x''(t) = \frac{1 - x'(t)^2}{1 + x'(t)^2}, \quad x \in C^\infty[0,1], x(0) = 0, x(1) = 1. \quad \square$$

Remark. We are now able to reflect with hindsight on the role of each of the three different function spaces that have figured in the theory. The choice of $C^2[a, b]$ is agreeable for evident reasons of simplicity and smoothness. We venture on to $\text{Lip}[a, b]$ because this space still leads to a good theory, including the basic necessary conditions, while allowing nonsmooth solutions; further, there are regularity results that establish a bridge back to $C^2[a, b]$. Finally, we advance to $AC[a, b]$ because it makes existence theorems possible; and again, there exist bridges from $AC[a, b]$ that lead back to $\text{Lip}[a, b]$ in many cases.⁵

⁵ Note that the class PWS of piecewise-smooth functions would not fit into this scheme: there are no bridges from $AC[a, b]$ to PWS.