Chapter 9

Relaxing the Smoothness Requirements for the Admissible Functions

In this chapter, we consider the possibility of minimizing a functional of the form

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

over classes of admissible functions that need not have continuous derivatives. Recall that when we first minimized such functionals, we assumed that the admissible functions were of class C^2; we then developed a theory in which the admissible functions were only required to be of class C^1. It is natural to see if we can further relax the smoothness of the admissible functions. An important reason for this is that if we are trying to find the "true minimum" of J, then we need to consider the largest possible appropriate class of admissible functions. However, we shall always require our admissible functions y to satisfy the fundamental theorem of calculus in the sense that y' must be (Riemann) integrable (possibly in the improper sense) and

$$y(x) = y(a) + \int_a^x y'(t)\, dt \quad \text{for all } x \in [a, b]. \tag{9.1}$$

Indeed, if the admissible functions are not required to satisfy this important property, then the entire character of the minimization problem changes dramatically. There are important problems in which it is not appropriate to require the admissible functions to satisfy (9.1), but in such problems the functional J often includes additional terms to penalize jumps in y, and the treatment of such problems is beyond the scope of these notes. Notice that (9.1) implies that

y is continuous on $[a, b]$. For simplicity, we consider only problems in which the values of the admissible functions are prescribed at both endpoints.

9.1 Example 9.1

Suppose that we wish to minimize

$$J(y) := \int_0^1 \left(y'(x)^2 - 1\right)^2 \, dx$$

subject to $y(0) = y(1) = 0$. If we allow we allow admissible functions to have jump discontinuities in their derivatives, then it is clear that the minimizers for J will those functions y such that $y'(x) = \pm 1$ at each $x \in [0, 1]$ and the minimum value of J will be zero. If we require y to be in $C^1[0, 1]$, however, then we always find that $J(y) > 0$.

We can construct "approximate minimizers" for J using C^1 functions. More precisely, for each $n \in \mathbb{N}$, there exists $y_n \in C^1[0, 1]$ with $y_n(0) = y_n(1) = 0$ such that

$$0 < J(y_n) \leq \frac{1}{n}.$$

9.2 Example 9.2 (Heinricher & Mizel, 1986)

Suppose that we want to minimize

$$J(y) := \int_0^1 \left(y(x)^2 - x\right)^2 y'(x)^6 \, dx$$

subject to $y(0) = 0$ and $y(1) = 1$. Notice that $J(y)$ is always nonnegative. We see that if $y_*(x) = \sqrt{x}$ for all $x \in [0, 1]$, then $J(y_*) = 0$, but $y_* \notin C^1[0, 1]$, since its derivative $y_*'(x)$ blows up as x approaches zero.

One might guess that, as in the previous example, we can approximate the "true minimum for" J using some $y_n \in C^1[0, 1]$ such that $y_n(0) = 0$ and $y_n(1) = 1$ and letting $n \to \infty$. For this example, such an approximation is impossible. Heinricher and Mizel showed that if $y \in C^1[0, 1]$ and satisfies $y(0) = 0$ and $y(1) = 1$, then

$$J(y) \geq \frac{1}{6} \left(\frac{3}{5}\right)^6.$$

Consequently, if one restricts the class of admissible functions to those with continuous derivatives over $[0, 1]$, then the "true minimum" for J cannot even be approached. This pathology, originally discovered in 1926 by Lavrentiev, is known as Lavrentiev's phenomenon.

9.3 Continuous Piecewise $\mathbf{C^1}$ Functions

In this section, we introduce a class of functions that is appropriate for minimizing functionals such as the one in Example 9.1 Let $a, b \in \mathbb{R}$ with $a < b$ be given.

Definition 9.1 *Let $y : [a, b] \to \mathbb{R}$ be given. We define the* singular set for y *by*

$$S(y) := \{x \in (a, b) \mid y \textit{ is not differentiable at } x\}.$$

Definition 9.2 *Let $y : [a, b] \to \mathbb{R}$ be given. We say that y* is a continuous piecewise C^1 function *provided the following conditions are satisfied:*

(1) $y \in C[a, b]$;

(2) $S(y)$ is a finite set;

(3) y' is continuous on the set $(a, b)\backslash S(y) := \{x \in (a, b) \mid x \notin S(y)\}$;

(4) for each $c \in S(y)$, we have $\lim_{x\to c^-} y'(x)$ and $\lim_{x\to c^+} y'(x)$ exist in $\mathbb{R}$;

(5) $\lim_{x\to a^+} y'(x)$ and $\lim_{x\to b^-} y'(x)$ exist in $\mathbb{R}$.

The class of all functions satisfying items (1)–(5) will be denoted by $\widehat{C}^1[a, b]$.

If a function y belongs to $\widehat{C}^1[a, b]$, then y is continuous, has no vertical tangents and has at most a finite number of "corners"; the corners in the graph of y are at the points $(c, y(c))$ with $c \in S(y)$. For notational convenience, if $y \in \widehat{C}^1[a, b]$ and $c \in (a, b)$, we write $y'(c^-)$ for $\lim_{x\to c^-} y'(x)$ and $y'(c^+)$ for $\lim_{x\to c^+} y'(x)$.

9.4 Minimizers with "Corners"

We first establish a version of the fundamental theorem of calculus and integration by parts for functions in $\widehat{C}[a, b]$. Before ststing the theorem we make a remark concerning integrability of derivatives of $\widehat{C}^1$-functions.

Remark 9.1 *Let $y \in \widehat{C}^1[a, b]$ be given. Then, for every $\gamma \in (a, b]$, the restriction of y to $[a, \gamma]$ belongs $\widehat{C}^1[a, \gamma]$ and y' is Riemann integrable on $[a, \gamma]$. The fact that there is a (possibly nonempty) finite subset of $[a, \gamma]$ on which y' is not well-defined does not cause any problems. If you feel more comfortable having an integrand that is everywhere defined on $[a, \gamma]$, you can assign values to $y'(c)$ for $c \in S(y)$ in any manner you wish. Changing the values of y' on a finite set does not alter the value of its integral.*

Theorem 9.1 (Fundamental Theorem of Calculus) *Let $y \in \widehat{C}^1[a, b]$ be given. Then, for every $\gamma \in (a, b]$ we have*

$$\int_a^\gamma y'(x)\, dx = y(\gamma) - y(a).$$

Proof. Let $\gamma \in (a,b]$ be given and put $P := \{a,\gamma\} \cup (S(y) \cap (a,\gamma))$. Notice that P is a non-empty and finite set with $a, \gamma \in P$ and $P \subset [a,\gamma]$. Therefore, we may write $P = \{x_0, x_1, \cdots, x_n\}$ with $a = x_0 < x_1 < \cdots < x_n = \gamma$. Using the standard fundamental theorem of calculus, we find that

$$\begin{aligned}
\int_a^b y'(x)\,dx &= \int_{x_0}^{x_n} y'(x)\,dx \\
&= \int_{x_0}^{x_1} y'(x)\,dx + \int_{x_1}^{x_2} y'(x)\,dx + \cdot + \int_{x_{n-1}}^{x_n} y'(x)\,dx \\
&= y(x_1) - y(x_0) + y(x_2) - y(x_1) + \cdots + y(x_n) - y(x_{n-1}) \\
&= y(x_n) - y(x_0) = y(b) - y(a).
\end{aligned} \tag{9.2}$$

$\square$

Theorem 9.2 (Integration by Parts) *Let $u, v \in \widehat{C}^1[a,b]$ be given. Then*

$$\int_a^b u(x)v'(x)\,dx = u(b)v(b) - u(a)v(a) - \int_a^b u'(x)v(x)\,dx.$$

Proof. Set $P := \{a,b\} \cup S(u) \cup S(v)$. We may write $P = \{x_0, x_1, \cdots, x_n\}$ with $a = x_0 < x_1 < \cdots < x_n = b$. Using the standard integration by parts formula on each subinterval $[x_{i-1}, x_i]$, we find that

$$\begin{aligned}
\int_a^b u(x)v'(x)\,dx &= \int_{x_0}^{x_n} u(x)v'(x)\,dx \\
&= \int_{x_0}^{x_1} u(x)v'(x)\,dx + \int_{x_1}^{x_2} u(x)v'(x)\,dx + \cdots + \int_{x_{n-1}}^{x_n} u(x)v'(x)\,dx \\
&= u(x)v(x)\Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} u'(x)v(x)\,dx + u(x)v(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} u'(x)v(x)\,dx \\
&\qquad + \cdots + u(x)v(x)\Big|_{x_{n-1}}^{x_n} - \int_{x_{n-1}}^{x_n} u'(x)v(x)\,dx \\
&= u(x_1)v(x_1) - u(x_0)v(x_0) + u(x_2)v(x_2) - u(x_1)v(x_1) \\
&\qquad + \cdots + u(x_n)v(x_n) - u(x_{n-1})v(x_{n-1}) - \int_{x_0}^{x_n} u'(x)v(x)\,dx \\
&= u(x_n)v(x_n) - u(x_0)v(x_0) - \int_{x_0}^{x_n} u'(x)v(x)\,dx \\
&= u(b)v(b) - u(a)v(a) - \int_a^b u'(x)v(x)\,dx.
\end{aligned}$$

□

We now pose our minimization problem. Let $A, B \in \mathbb{R}$ be given and put

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

Let $f : [a,b] \times \mathbb{R} \times \mathbb{R}$ be a given function with continuous first-order partial derivatives and define $J : \mathscr{Y} \to \mathbb{R}$ by

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

As usual, we wish to minimize J over $\mathscr{Y}$. (The fact that there might be some finite set on which y' is not defined does not cause any difficulties with defining the integral.)

For each $y \in \mathscr{Y}$, the class of admissible variations is easily seen to be

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

Given $y \in \mathscr{Y}$ and $v \in \mathscr{V}$, it is straightforward to verify that the Gâteaux variation at y in the direction v exists and is given by the usual formula

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

(Indeed, we can take a partition for [a,b] consisting of a, b together with all corner points for y and all corner points for v, break the integral for $J(y + \varepsilon v$ into a sum of integrals, differentiate each integral separately, and re-assemble the pieces.)

Suppose that J attains a minimum over $\mathscr{Y}$ at $y_* \in \mathscr{Y}$. Then $\delta J(y_*; v) = 0$ for every $v \in \mathscr{V}$. Define $G, H : [a,b] \to \mathbb{R}$ by

$$G(x) := f_{,3}(x, y_*(x), y_*'(x)) \text{ and } H(x) := \int_a^x f_{,2}(t, y_*(t), y_*'(t))\, dt \quad \text{for all } x \in [a,b].$$

(For $x \in S(y_*)$ we can take $G(x)$ to be any convenient real number.) Notice that $H \in \widehat{C}^1[a,b]$. Using our integration by parts theorem and our definition of $\mathscr{V}$, we may write

$$\begin{aligned} \delta J(y_*; v) &= \int_a^b \Big\{ H'(x)v(x) + G(x)v'(x) \Big\}\, dx \\ &= H(b)v(b) - H(a)v(a) + \int_a^b \Big\{ G(x) - H(x) \Big\} v'(x)\, dx \\ &= \int_a^b \Big\{ G(x) - H(x) \Big\} v'(x)\, dx \quad \text{for all } v \in \mathscr{V}. \end{aligned}$$

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

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

so that we have that

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

We now proceed along the lines of the proof of the du Bois-Reymond Lemma (Lemma 3.3). Let k be a constant to be specified later. Since $\int_a^b g(x)v'(x)\, dx = 0$ for every $v \in \mathscr{V}$, we have

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

We want to find a $v_* \in \mathscr{V}$ such that $v_*'(x) = g(x) - k$ at each $x \in (a,b)\backslash S(v_*)$. Since we want $v_*(a) = 0$ and $v_*'(x) = g(x) - k$ for all $x \in (a,b)\backslash S(v_*)$, we put

$$v_*(x) := \int_a^x [g(t) - k]\, dt \quad \text{for all } x \in [a,b].$$

Observe that with this definition we have $v_* \in \widehat{C}^1[a,b]$, $v_*(a) = 0$ and $v_*'(x) = g(x) - k$ for all $x \in (a,b)\backslash S(v_*)$. We need to choose $k \in \mathbb{R}$ so that $v_*(b) = 0$. From our definition for v_*, we see that

$$v_*(b) = \int_a^b [g(t) - k]\, dt = \int_a^b g(t)\, dt - (b-a)k,$$

so clearly, we should choose

$$k = \frac{1}{b-a} \int_a^b g(t)\, dt.$$

With this choice of k, we find that $v_* \in \mathscr{V}$ and

$$\int_a^b [g(x) - k] v_*'(x)\, dx = 0 \Rightarrow \int_a^b [g(x) - k]^2\, dx = 0.$$

We conclude that $g(x) = k$ for all $x \in [a,b]$ at which g is continuous. It follows that there exists a $k \in \mathbb{R}$ such that

$$f_{,3}(x, y_*(x), y_*'(x)) - \int_a^x f_{,2}(t, y_*(t), y_*'(t))\, dt = k \quad \text{for all } x \in (a,b)\backslash S(y_*). \tag{IE-L$_1$}$$

The above equation is often referred to as the integrated first Euler-Lagrange equation.

We have shown that if $y_* \in \mathscr{Y}$ is a minimizer for J over $\mathscr{Y}$, then y_* must satisfy $(\text{IE-L})_1$ for all $x \in (a,b)\backslash S(y_*)$. The points $c \in S(y_*)$ are called corner points. For $x \notin S(y_*)$, we can differentiate both sides of $(\text{IE-L})_1$ to obtain

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

Set $P := \{a,b\} \cup S(y_*) = \{x_0, x_1, \cdots, x_n\}$, with $x_0, x_1, \cdots < x_n$ and suppose that $(\text{E-L})_1$ holds at each $x \in (a,b)\backslash S(y_*) = (x_0,x_1) \cup (x_1,x_2) \cup \cdots \cup (x_{n-1},x_n)$. Then, there exist $k_1, k_2, \cdots, k_n \in \mathbb{R}$ such that

$$f_{,3}(x, y(x), y'(x)) - \int_a^x f_{,2}(t, y(t), y'(t))\, dt = k_i \quad \text{for all } x \in (x_{i-1}, x_i).$$

If the integrated first Euler-Lagrange equation holds, then all the k_i's must be the same.

We now state our first corner condition.

Theorem 9.3 (1^{st} Weierstrass-Erdmann Corner Condition) *Let $y \in \widehat{C}^1[a,b]$ be given, and suppose that y satisfies (IE-L)$_1$. Then, for each $c \in S(y)$ we find that*

$$f_{,3}(c, y(c), y'(c^+)) = f_{,3}(c, y(c), y'(c^-)).$$

Proof. Notice that $x \mapsto \int_a^x f_{,2}(t, y(t), y'(t))\,dt$ is continuous on $[a,b]$. In particular

$$\lim_{x\to c^-} \int_a^x f_{,2}(t, y(t), y'(t))\,dt = \lim_{x\to c^+} \int_a^x f_{,2}(t, y(t), y'(t)) \quad \text{for all } c \in S(y).$$

Thus (IE-L)$_1$ implies

$$\lim_{x\to c^-} f_{,3}(x, y(x), y'(x)) = \lim_{x\to c^+} f_{,3}(x, y(x), y'(x)) \quad \text{for all } c \in S(y),$$

and therefore

$$f_{,3}(c, y(c), y'(c^-)) = f_{,3}(c, y(c), y'(c^+)) \quad \text{for all } c \in S(y),$$

since $f_{,3}$ is continuous and $y \in \widehat{C}^1[a,b]$. □

Corollary 9.1 *Suppose that $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has continuous second-order partial derivatives. Suppose further that $y_* \in \mathscr{Y}$ minimizes J over $\mathscr{Y}$. Let $c \in S(y_*)$ be given, and put $\alpha := y_*'(c^-)$ and $\beta := y_*'(c^+)$. Then, there exists a λ between α and β such that*

$$F'(\lambda) = f_{,3,3}(c, y_*(c), \lambda) = 0.$$

Proof. Define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(z) := f_{,3}(c, y_*(c), z) \quad \text{for all } z \in \mathbb{R}.$$

Since y_* minimizes J over $\mathscr{Y}$, it must satisfy (IE-L)$_1$. Therefore it satisfies the 1^{st} Weierstrass-Erdmann corner condition. That is

$$F(\alpha) = f_{,3}(c, y_*(c), \alpha) = f_{,3}(c, y_*(c), \beta) = F(\beta).$$

By Rolle's theorem, there exists a λ between α and β such that $F'(\lambda) = 0$. From the definition of F, we have

$$f_{,3,3}(c, y_*(c), \lambda) = 0.$$

□

Notice, in particular, that the corollary above tells us that when $f_{,3,3}(x,y,z) > 0$ at each $(x,y,z) \in [a,b] \times \mathbb{R} \times \mathbb{R}$, then no minimizer for J can have corners.

Before formulating our second corner condition, we state the following fact: if $y_* \in \mathscr{Y}$ minimizes J over $\mathscr{Y}$, then there exists a $K \in \mathbb{R}$ such that y_* satisfies

$$f(x, y_*(x), y_*'(x)) - y_*'(x) f_{,3}(x, y_*(x), y_*'(x)) = K + \int_a^x f_{,1}(t, y_*(t), y_*'(t))\,dt \quad \text{for all } x \in (a,b)\backslash S(y_*). \tag{E-L$_2$}$$

In other words, the second Euler-Lagrange Equation holds. Before proving this fact, we shall record an important consequence of (E-L)$_2$, namely another corner condition, and look at a couple of examples. The proof that (E-L)$_2$ holds will be given in Section 9.5

Theorem 9.4 (2$^{\text{nd}}$ Weierstrass-Erdmann Corner Condition) *Let $y \in \widehat{C}^1[a,b]$ be given, and assume that y satisfies (E-L)$_2$. Then, for each $c \in S(y)$, we have*

$$\begin{aligned} f(c, y(c), y'(c^-)) - y'(c^-) f_{,3}(c, y(c), y'(c^-)) \\ = f(c, y(c), y'(c^+)) - y'(c^+) f_{,3}(c, y(c), y'(c^+)). \end{aligned}$$

Proof. The result follows from the continuity of $x \mapsto \int_a^x f_{,1}(t, y(t), y'(t))\,dt$. $\square$

9.4.1 Example 9.4.1

Let

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

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

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

The integrand $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for J is given by

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

Thus for each $(x,y,z) \in [a,b] \times \mathbb{R} \times \mathbb{R}$, we find that

$$f_{,1}(x,y,z) = 0; \;\; f_{,2}(x,y,z) = 2y(z-1)^2;$$

$$f_{,3}(x,y,z) = 2y^2(z-1); \text{ and } f_{,3,3}(x,y,z) = 2y^2.$$

Notice that

$$f_{,3,3}(c, y(c), \lambda) = 0 \Rightarrow y(c) = 0,$$

so if c is a corner point for a minimizer y, then $y(c)$ must be zero.

Suppose that y minimizes J over $\mathscr{Y}$ and that $c \in S(y)$ is a corner point. Put $\alpha := y'(c^-)$ and $\beta := y'(c^+)$. The 1^{st} corner condition becomes

$$2y(c)^2(\alpha - 1) = 2y(c)^2(\beta - 1).$$

This, however, provides no new information, since at a corner point $y(c) = 0$. The 2^{nd} corner condition becomes

$$y(c)^2(\alpha - 1)^2 - 2\alpha y(c)^2(\alpha - 1) = y(c)^2(\beta - 1)^2 - 2\beta y(c)^2(\beta - 1).$$

Again since we know that $y(c) = 0$, we get no new information from the above equation.

Let us now look at $(\text{IE-L})_1$ and $(\text{E-L})_2$ for J. We have

$$2y(x)^2(y'(x) - 1) = k + \int_a^x 2y(t)(y'(t) - 1)^2\,dt \quad \text{for all } x \in (a,b)\backslash S(y) \quad (\text{IE-L})_1$$

and

$$y(x)^2(y'(x) - 1)^2 - 2y(x)^2 y'(x)(y'(x) - 1) = K \quad \text{for all } x \in (a,b)\backslash S(y), \quad (\text{E-L})_2$$

for some $k, K \in \mathbb{R}$. If there is a corner point c for y, then $y(c) = 0$, and this tells us that $K = 0$. So if y has a corner point, then $(\text{E-L})_2$ reduces to

$$y(x)^2(y'(x) - 1)^2 - 2y(x)^2 y'(x)(y'(x) - 1) = 0 \quad \text{for all } x \in (a,b)\backslash S(y).$$

For those $x \in (a,b)\backslash S(y)$ with $y(x) \neq 0$, the above equation implies that $y'(x) = \pm 1$. We conclude that if a minimizer y has at least one corner point then

(1) for those $x \in (a,b)\S(y)$ where $y(x) \neq 0$, we have $y'(x) = \pm 1$;

(2) if c is a corner point for y, then $y(c) = 0$.

Thus either $y'(x) = \pm 1$ or $y(x) = 0$. We note, for example, that if $a = 0, b = 1, A = 0, B = 2$, there can be no corner points for a minimizer, since minimizers with corners can only have graphs with slopes of 0 or ± 1. More generally, a necessary condition for there to be a minimizer with at least one corner point is

$$\frac{|B - A|}{|b - a|} < 1.$$

9.4.2 Example 9.4.2

Put

$$\mathscr{Y} := \left\{ y \in \widehat{C}^1[a,b] \mid y(a) = A \,:\, y(b) = B \right\},$$

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

$$J(y) := \int_a^b \left[y'(x)^4 - y'(x)^2 \right] dx.$$

We wish to find out what happens at a corner point of a minimizer for J over $\mathscr{Y}$.

Remark 9.2 *We could use the fact that*

$$y'(x)^4 - y'(x)^2 = \left(y'(x)^2 - \frac{1}{2}\right)^2 - \frac{1}{4}$$

to analyze J directly and see that $\pm\frac{1}{\sqrt{2}}$ are possible for values of $y'(c^-), y'(c^+)$ at a corner point c. For purposes of illustartion we shall use the Weierstrass-Erdmann corner conditions and see how the conclusions compare with what we just observed by above completing the square in the integrand.

The integrand $f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for J is given by

$$f(x, y, z) := z^4 - z^2 \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R} \times \mathbb{R},$$

so that

$$f_{,1}(x, y, z) = 0;\ f_{,2}(x, y, z) = 0;\ f_{,3}(x, y, z) = 4z^3 - 2z$$

and

$$f_{,3,3}(x, y, z) = 12z^2 - 2 \quad \text{for all } (x, y, z) \in [a, b] \times \mathbb{R} \times \mathbb{R}.$$

We notice that if c is a corner point for a minimizer $y \in \mathscr{Y}$, then

$$f_{,3,3}(c, y(c), \lambda) = 12\lambda^2 - 2 = 0 \Rightarrow \lambda = \pm\sqrt{\frac{1}{6}}.$$

By Corollary 9.1, if c is a corner point for y, then at least one of $\pm\sqrt{\frac{1}{6}}$ must lie between $y'(c^-)$ and $y'(c^+)$.

Now let us look at the 1$^{\text{st}}$ and 2$^{\text{nd}}$ corner conditions. Suppose that c is a corner point for a minimizer y, and put $\alpha := y'(c^-)$ and $\beta := y'(c^+)$. For the 1$^{\text{st}}$ corner condition, we have

$$4\alpha^3 - 2\alpha = 4\beta^3 - 2\beta. \tag{9.3}$$

The 2$^{\text{nd}}$ corner condition gives

$$\alpha^4 - \alpha^2 - \alpha(4\alpha^3 - 2\alpha) = \beta^4 - \beta^2 - \beta(4\beta^3 - 2\beta). \tag{9.4}$$

Since we are supposing that c is a corner point for y, we seek solutions to (9.3) and (9.4) such that $\alpha \neq \beta$. Using this, we may rewrite (9.3) as

$$\begin{aligned} 4\alpha^3 - 4\beta^3 = 2\alpha - 2\beta &\Rightarrow 4(\alpha - \beta)(\alpha^2 + \alpha\beta + \beta^2) = 2(\alpha - \beta) \\ &\Rightarrow \alpha^2 + \alpha\beta + \beta^2 = \frac{1}{2}. \end{aligned} \tag{9.5}$$

We also rewrite (9.4) as

$$\begin{aligned} 3\alpha^4 - \alpha^2 = 3\beta^4 - \beta^2 &\Rightarrow 3\alpha^4 - 3\beta^4 = \alpha^2 - \beta^2 \\ &\Rightarrow 3(\alpha - \beta)(\alpha + \beta)(\alpha^2 + \beta^2) = (\alpha - \beta)(\alpha + \beta) \\ &\Rightarrow 3(\alpha + \beta)(\alpha^2 + \beta^2) = \alpha + \beta. \end{aligned} \tag{9.6}$$

From this last expression, we deduce that either $\alpha = -\beta$ or $\alpha^2 + \beta^2 = \frac{1}{3}$.

First, we consider the case where $\alpha^2 + \beta^2 = \frac{1}{3}$. Substituting this back into (9.5) yields

$$\alpha\beta = \frac{1}{6} \Rightarrow \beta = \frac{1}{6\alpha}. \tag{9.7}$$

Now, using $\alpha^2 + \beta^2 = \frac{1}{3}$ gives us

$$\begin{aligned} \alpha^2 + \frac{1}{36\alpha^2} = \frac{1}{3} &\Rightarrow \alpha^4 + \frac{1}{36} = \frac{1}{3}\alpha^2 \\ &\Rightarrow \left(\alpha^2 - \frac{1}{6}\right)^2 = 0 \\ &\Rightarrow \alpha = \pm\sqrt{\frac{1}{6}}. \end{aligned}$$

If $\alpha = \sqrt{\frac{1}{6}}$, then (9.7) implies that $\beta = \sqrt{\frac{1}{6}} = \alpha$; also, if $\alpha = -\sqrt{\frac{1}{6}}$, we find that $\beta = -\sqrt{\frac{1}{6}} = \alpha$. In either case, our solutions violate the condition that $\alpha \neq \beta$, so these are not valid solutions.

We now turn to the cases where $\alpha = -\beta$. Substituting this back into (9.5) yields

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

If $\alpha = \sqrt{\frac{1}{2}}$, then $\beta = -\sqrt{\frac{1}{2}}$; and if $\alpha = -\sqrt{\frac{1}{2}}$, then $\beta = \sqrt{\frac{1}{2}}$.

In conclusion, if c is a corner point for a minimizer $y \in \mathscr{Y}$ for J, then

$$y'(c^-) = \pm\sqrt{\frac{1}{2}} \text{ and } y'(c^+) = -y'(c^-).$$

These conclusions are completely consistent with Remark 9.2.

9.5 The Second Euler Lagrange Equation

(TO BE FILLED IN)

9.6 The Cauchy-Schwarz Inequality

We now wish to prove our statement regarding the Heinricher & Mizel Example. Namely, we will show that the "true minimum" (namely 0) for the functional J defined in Section 9.2 cannot be approached using functions in $C^1[0,1]$ satisfying $y(0) = 0$ and $y(1) = 1$. To do this, we need to establish an inequality for integrals. The proof of this inequality makes use of a preliminary result (Hölder's inequality) which we develop in this and the next section.

Hölder's inequality is a generalization of the Cauchy-Schwarz inequality. The Cauchy-Schwarz inequality can be obtained easily from the following elementary algebraic inequality.

Proposition 9.1 *Let $\alpha, \beta \in \mathbb{R}$ be given. We have*

$$|\alpha\beta| \leq \frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2.$$

Proof. Observe that

$$\begin{aligned}(\alpha+\beta)^2 \geq 0 \Rightarrow & \alpha^2 + 2\alpha\beta + \beta^2 \geq 0 \\ \Rightarrow & -\alpha\beta \leq \frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2.\end{aligned}$$

Similarly

$$\begin{aligned}(\alpha-\beta)^2 \geq 0 \Rightarrow & \alpha^2 - 2\alpha\beta + \beta^2 \geq 0 \\ \Rightarrow & \alpha\beta \leq \frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2.\end{aligned}$$

It follows that

$$|\alpha\beta| \leq \frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2.$$

□

Using this simple inequality, we can prove a very useful inequality for integrals.

Definition 9.3 *For functions $f \in C[a,b]$, put*

$$\|f\|_2 := \left(\int_a^b f(x)^2 \, dx \right)^{\frac{1}{2}}$$

Notice that for every $c \in \mathbb{R}$, we have $\|cf\|_2 = |c| \cdot \|f\|_2$. Furthermore for $f \in C[a,b]$ we have $\|f\|_2 = 0$ if and only if $f(x) = 0$ for all $x \in [a,b]$.

We now state the Cauchy-Schwarz inequality for integrals.

Theorem 9.5 *Cauchy-Schwarz Inequality Let $f, g \in C[a,b]$ be given. Then*

$$\int_a^b |f(x)g(x)| \, dx \leq \|f\|_2 \|g\|_2.$$

Proof. If either $\|f\|_2 = 0$ or $\|g\|_2 = 0$, then the inequality is trivially satisfied because both sides vanish. We therefore may assume that $\|f\|_2 \neq 0$ and $\|g\|_2 \neq 0$. Let us define $F, G \in C[a,b]$ by

$$F(x) := \frac{f(x)}{\|f\|_2} \text{ and } G(x) := \frac{g(x)}{\|g\|_2} \quad \text{for all } x \in [a,b].$$

Using Proposition 9.1 and our definitions of F and G, we have

$$\begin{aligned}
\int_a^b |F(x)G(x)|\,dx \leq & \frac{1}{2}\int_a^b |F(x)|^2\,dx + \frac{1}{2}\int_a^b |G(x)|^2\,dx \\
= & \frac{1}{2\|f\|_2^2}\int_a^b |f(x)|^2\,dx + \frac{1}{2\|g\|_2^2}\int_a^b |g(x)|^2\,dx \\
= & \frac{1}{2} + \frac{1}{2} = 1.
\end{aligned}$$

Thus

$$\begin{aligned}
\int_a^b |F(x)G(x)|\,dx \leq 1 \Rightarrow & \frac{1}{\|f\|_2\|g\|_2}\int_a^b |f(x)g(x)\,dx \leq 1 \\
\Rightarrow & \int_a^b |f(x)g(x)|\,dx \leq \|f\|_2\|g\|_2
\end{aligned}$$

□

Remark 9.3 (Triangle Inequality) *Let $f, g \in C[a,b]$ be given. Then we have*

$$\begin{aligned}
\|f+g\|_2^2 &= \int_a^b \left\{ f(x)^2 + 2f(x)g(x) + g(x)^2 \right\} dx \\
&= \|f\|_2^2 + \|g\|_2^2 + 2\int_a^b f(x)g(x)\,dx \leq \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2.
\end{aligned}$$

Since $\|f\|_2, \|g\|_2 \geq 0$ we conclude that

$$\|f+g\|_2 \leq \|f\|_2 + \|g\|_2 \quad \text{for all } f, g \in C[a,b].$$

It follows that $\|\cdot\|_2$ is a norm on $C[a,b]$.

9.7 Hölder's Inequality

In this section, we will prove Hölder's inequality, which is a generalization of Theorem 9.5. In the next section, we use Hölder's inequality to prove a special case of Jensen's inequality. Once we have Jensen's inequality in hand, we will prove our statement regarding the Heinricher & Mizel example.

First, we give a geometrical explanation of Proposition 9.1. Suppose that $\alpha, \beta > 0$. The proposition states that

$$\alpha\beta \leq \frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2. \tag{9.8}$$

The left-hand side of (9.8) is the area of a rectangle with sides of length α and β. The right hand side of (9.8) is sum of the area of two triangles: one with base and height α, the other with base and height β. From Figure 9.1, we see

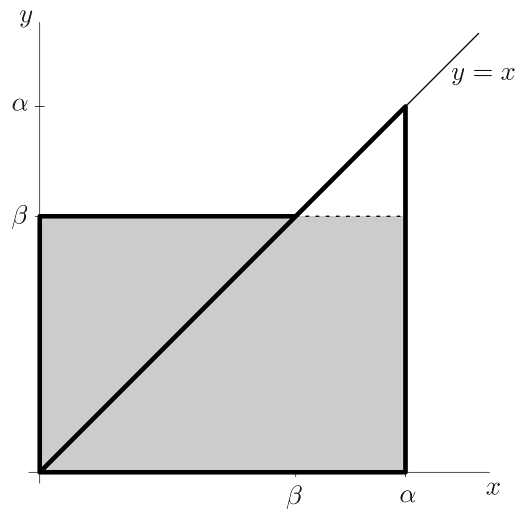

Figure 9.1: A geometrical interpretation of Proposition 9.1

that the area of the rectangle is indeed at least as small as the sum of the areas for the triangles. (The figure was drawn for the case $\alpha > \beta$. If $\beta > \alpha$, we can simply interchange α and β.) Observe also from the figure that equality holds in (9.8) if and only if $\alpha = \beta$.

For the Cauchy-Schwarz inequality, the areas of the two triangles are separated by the line $y = x$. We will generalize (9.8) by using regions separated by the curve $y = x^\gamma$, with $\gamma > 0$, rather than by a straight line (see Figure 9.2). From the figure, it is apparent that the area of the rectangle with sides of length α and β is not larger than the sum of the areas A_1 and A_2. That is

$$\alpha\beta \leq A_1 + A_2. \tag{9.9}$$

(Figure 9.2 was drawn for the case $\alpha > \beta$. You should convince yourself by drawing an appropriate figure that (9.9) remains valid when $\beta \geq \alpha$.) We find that

$$A_1 = \int_0^\alpha x^\gamma \, dx = \frac{1}{\gamma + 1} \alpha^{\gamma+1}$$

and

$$A_2 = \int_0^\beta y^{\frac{1}{\gamma}} \, dy = \frac{\gamma}{\gamma + 1} \beta^{\frac{\gamma+1}{\gamma}}.$$

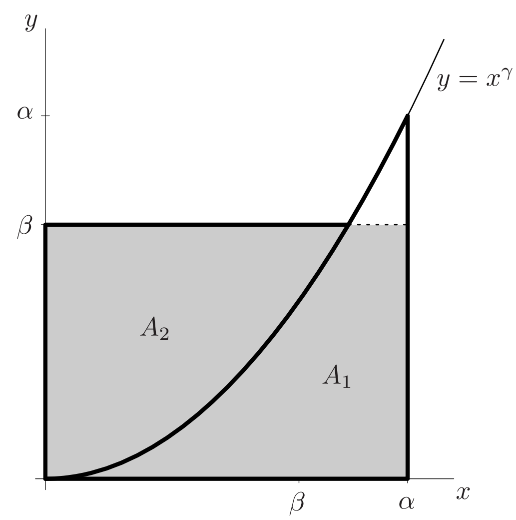

Figure 9.2: A geometrical interpretation of Young's inequality

Thus

$$\alpha\beta \leq \frac{1}{\gamma+1}\alpha^{\gamma+1} + \frac{\gamma}{\gamma+1}\beta^{\frac{\gamma+1}{\gamma}}.$$

Put $p = \gamma + 1$ and $q = \frac{\gamma+1}{\gamma}$. Then we have

$$\alpha\beta \leq \frac{1}{p}\alpha^p + \frac{1}{q}\beta^q. \tag{9.10}$$

Observe that $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. The exponents p and q are often called conjugate exponents. Now, we have shown (9.10) for $\alpha, \beta > 0$, but it is straightforward to extend the inequality to arbitrary $\alpha, \beta \in \mathbb{R}$, with appropriate absolute values inserted. Observe also that starting with $p > 1$ we can define $\gamma = p - 1$ and obtain (9.10).

Theorem 9.6 (Young's Inequality) *Let $\alpha, \beta, p \in \mathbb{R}$ with $p > 1$ be given. Choose $q \in \mathbb{R}$ so that $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$|\alpha\beta| \leq \frac{1}{p}|\alpha|^p + \frac{1}{q}|\beta|^q.$$

Notice that when $p = 2$, we have $q = 2$ and Young's inequality reduces to (9.8). Observe also that given $p > 1$ there is exactly one $q \in \mathbb{R}$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, namely

$$q = \frac{p}{p-1}$$

and this value of q is strictly graeter than 1.

Before generalizing Theorem 9.5, we need a definition.

Definition 9.4 *For each $p > 1$ and function $f \in C[a,b]$, put*

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}}.$$

The quantity $\|f\|_p$ is usually referred to as the L^p-norm of f. We point out that for every $c \in \mathbb{R}$, we have $\|cf\|_p = |c| \cdot \|f\|_p$. Furthermore, for $f \in C[a,b]$ we have $\|f\|_p = 0$ if and only if $f(x) = 0$ for all $x \in [a,b]$.

We will now state and prove

Theorem 9.7 (Hölder's Inequality) *Let $p, q \in \mathbb{R}$ be such that $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $f, g \in C[a,b]$ be given. Then we have*

$$\int_a^b |f(x)g(x)| \le \|f\|_p \|g\|_q.$$

Proof. The proof proceeds along the lines of the proof for Theorem 9.5.

If either $\|f\|_p = 0$ or $\|g\|_q = 0$, then there is nothing to prove. So we assume that $\|f\|_p \neq 0$ and $\|g\|_q \neq 0$.

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

$$F(x) := \frac{f(x)}{\|f\|_p} \text{ and } G(x) := \frac{g(x)}{\|g\|_q} \quad \text{for all } x \in [a,b].$$

As before, we find that

$$\int_a^b |F(x)G(x)| \, dx \le 1,$$

since $\|F\|_p = \|G\|_q = 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. It follows that

$$\int_a^b |f(x)g(x)| \, dx \le \|f\|_p \|g\|_q.$$

□

Remark 9.4 (Triangle Inequality) *It can be shown that for each $p > 1$ we have*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p \quad \text{for all } f, g \in C[a,b].$$

Consequently $\|\cdot\|_p$ is a norm on $C[a,b]$.

9.8 Jensen's Inequality (A Special Case)

Using Hölder's inequality, we can easily prove

Theorem 9.8 (Jensen's Inequality (a special case)) *Let $p \in \mathbb{R}$ with $p > 1$ and $f \in C[a,b]$ be given. Then we have*

$$\frac{1}{(b-a)^{p-1}} \left(\int_a^b |f(x)|\, dx \right)^p \leq \int_a^b |f(x)|^p\, dx.$$

Proof. Put $q := \frac{p}{p-1}$, so $\frac{1}{p} + \frac{1}{q} = 1$. Using Hölder's inequality, we may write

$$\begin{aligned} \int_a^b |f(x)|\, dx &= \int_a^b |f(x)|^p \cdot 1\, dx \\ &\leq \left(\int_a^b |f(x)|^p\, dx \right)^{\frac{1}{p}} \left(\int_a^b 1^q\, dx \right)^{\frac{1}{q}} \\ &= \left(\int_a^b |f(x)|^p\, dx \right)^{\frac{1}{p}} (b-a)^{\frac{1}{q}}. \end{aligned}$$

Thus

$$\begin{aligned} \left(\int_a^b |f(x)|\, dx \right)^p &\leq (b-a)^{\frac{p}{q}} \int_a^b |f(x)|^p\, dx \\ \Rightarrow \frac{1}{(b-a)^{\frac{p}{q}}} \left(\int_a^b |f(x)|\, dx \right)^p &\leq \int_a^b |f(x)|^p\, dx \\ \Rightarrow \frac{1}{(b-a)^{p-1}} \left(\int_a^b |f(x)|\, dx \right)^p &\leq \int_a^b |f(x)|^p\, dx. \end{aligned}$$

□

9.9 Example 9.9 (cf. Example 9.2; Heinricher & Mizel, 1986)

Let us recall the example given in Section 9.2. We set

$$\mathscr{Y} \in \{y \in C1[0,1] \,:\, y(0) = 0 \text{ and } y(1) = 1\}$$

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

$$J(y) := \int_0^1 \left(y(x)^2 - x\right)^2 y'(x)^6 \, dx \quad \text{for all } y \in \mathscr{Y}.$$

As pointed out in Section 9.2, if $y_*(x) := \sqrt{x}$ for all $x \in [0,1]$, then $J(y_*) = 0$ but $y_* \notin \mathscr{Y}$ since $y_* \notin C^1[0,1]$. In this section, we will prove

$$J(y) \geq \frac{1}{32}\left(\frac{3}{5}\right)^6 \quad \text{for all } y \in \mathscr{Y}.$$

In other words, we will show that no matter which y from $\mathscr{Y}$ is used, the value for J can never be smaller than $\frac{1}{32}\left(\frac{3}{5}\right)^6$. This phenomenon is called Lavrentiev's phenomenon and is quite remarkable since we know $y_*(x) = \sqrt{x}$ makes $J(y_*)$ zero.

To prove our statement, let $y \in \mathscr{Y}$ be given. Since $y \in C[0,1]$, we may choose $M > 0$ such that

$$|y(x_2) - y(x_1)| \leq M|x_2 - x_1| \quad \text{for all } x_1, x_2 \in [0,1].$$

In particular, we have

$$|y(x) - y(0)| = |y(x)| \leq M|x - 0| = Mx \quad \text{for all } x \in [0,1].$$

Thus

$$y(x)^2 \leq M^2x^2 \leq \frac{1}{2}x \quad \text{for all } x \in [0, \tfrac{1}{2M^2}].$$

Therefore

$$-\sqrt{\frac{x}{2}} \leq y(x) \leq \sqrt{\frac{x}{2}} \quad \text{for all } x \in [0, \tfrac{1}{2M^2}].$$

This inequality tells us that near $x = 0$, the graph of y must be between the graphs of the functions $x \mapsto -\sqrt{\frac{x}{2}}$ and $x \mapsto \sqrt{\frac{x}{2}}$. Since $y(1) = 1$, it follows that there must be at least one point where the graph of y crosses the graph of $x \mapsto \sqrt{\frac{x}{2}}$. We may therefore choose $\beta \in (0,1)$ to be the smallest strictly positive x-value at which the graph of y crosses the graph of one of the functions $x \mapsto -\sqrt{\frac{x}{2}}$ or $x \mapsto \sqrt{\frac{x}{2}}$. That is, we choose $\beta \in (0,1)$ such that for every $x \in (0,\beta)$

$$-\sqrt{\frac{x}{2}} \leq y(x) \leq \sqrt{\frac{x}{2}}$$

and

$$|y(\beta)| = \sqrt{\frac{\beta}{2}}.$$

With β so chosen, we deduce that

$$\begin{aligned} x \geq 2y(x)^2 \Rightarrow & x - y(x)^2 \geq y(x)^2 \\ & \Rightarrow \left(y(x)^2 - x\right)^2 \geq y(x)^4 \quad \text{for all } x \in [0,\beta]. \end{aligned}$$

We now prove a positive lower bound for the value of $J(y)$. Using the above inequality and the fact that $0 < \beta < 1$, we have

$$
\begin{aligned}
J(y) &= \int_0^1 \left(y(x)^2 - x\right)^2 y'(x)^6 \, dx \\
&\geq \int_0^\beta \left(y(x)^2 - x\right)^2 y'(x)^6 \, dx \\
&\geq \int_0^\beta y(x)^4 y'(x)^6 \, dx \\
&\geq \int_0^\beta \left| y(x)^{\frac{2}{3}} y'(x) \right|^6 \, dx
\end{aligned}
$$

Now we use Jensen's inequality, with $p = 6$, we write

$$
\begin{aligned}
J(y) &\geq \frac{1}{(\beta - 0)^5} \left[\int_0^\beta |y(x)^{\frac{2}{3}} y'(x)| \, dx \right]^6 \\
&\geq \frac{1}{\beta^5} \left[\frac{3}{5} y(x)^{\frac{5}{3}} \Big|_0^\beta \right]^6 \\
&\geq \frac{1}{\beta^5} \left(\frac{3}{5}\right)^6 \left[y(\beta)^{\frac{5}{3}} - y(0)^{\frac{5}{3}} \right]^6 \\
&\geq \frac{1}{\beta^5} \left(\frac{3}{5}\right)^6 \left(\frac{1}{2}\right)^5 \beta^5 \\
&\geq \frac{1}{32} \left(\frac{3}{5}\right)^6 .
\end{aligned}
$$

We have shown that for any $y \in \mathscr{Y}$, the value of J is no smaller than $\frac{1}{32}\left(\frac{3}{5}\right)^5$. With more advanced techniques, one can actually show that the value of J is no smaller than $\frac{1}{6}\left(\frac{3}{5}\right)^6$. In any case, there is no way that the value for J can be driven to zero using functions from $\mathscr{Y}$ despite the fact that $J(y_*)$ is zero.