Chapter 8

Sufficiency via Field Theory

Let $a, b \in \mathbb{R}$ with $a < b$, $n \in \mathbb{N}$, $A, B \in \mathbb{R}^n$, and $f \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ be given. Put

$$\mathscr{Y} = \{y \in C^1([a,b];\mathbb{R}^n) \,:\, y(a) = A \text{ and } y(b) = B\},$$

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

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

Recall that if $y_* \in \mathscr{Y}$ minimizes J over $\mathscr{Y}$ then y_* satisfies the first Euler-Lagrange system

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

We also know that if $f(x,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is convex for every $x \in [a,b]$, then any function $y_* \in \mathscr{Y}$ that satisies $(\text{E-L})_1$ does actually minimize J over $\mathscr{Y}$. In this chapter, we develop conditions that are weaker than convexity of $f(x,\cdot,\cdot)$ and still ensure that a solution y_* of $(\text{E-L})_1$ does, in fact, minimize J over $\mathscr{Y}$. Roughly speaking these conditions will involve convexity of the mapping $z \mapsto f(x,y,z)$ for fixed values of $x \in [a,b]$, $y \in \mathbb{R}^n$, together with existence of an suitable collection of stationary functions.

8.1 The Method of Weierstrass

In his lectures of 1879, Weierstrass presented an approach for comparing the value of $J(\widetilde{y})$ for a given (but arbitrary) $\widetilde{y} \in \mathscr{Y}$ to the value of $J(y_*)$, where $y_* \in \mathscr{Y}$ satisfies $(\text{E-L})_1$. This approach requires that for each $\xi \in (a,b]$ there is exactly one solution $\psi(\cdot,\xi)$ of $(\text{E-L})_1$ on $[a,\xi]$ satisfying

$$\psi(a,\xi) = A \text{ and } \psi(\xi,\xi) = \widetilde{y}(\xi).$$

Assuming the existence of such a family of solutions of $(\text{E-L})_1$ (and some additional "regularity" of ψ), there is an extremely elegant formula which ensures that if f is convex in its third argument, then $J(y) \geq J(y_*)$. The "fly in the ointment" here is that it is generally quite difficult to verify that a family of solutions of $(\text{E-L})_1$ having the requisite properties exists (even in the scalar case $n = 1$).

8.1.1 Example 8.1.1

Before discussing the method of Weierstrass in generality, it seems instructive to look at an example where $f(x, \cdot, \cdot)$ fails to be convex, but Weierstrass' approach works and there are closed-form expressions for the important ingredients needed for this approach.

Here, we assume that $n = 1$. Let $L \in (0, \pi)$ be given, put

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

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

$$J(y) = \int_0^L [y'(x)^2 - y(x)^2]\, dx \quad \text{for all } y \in \mathscr{Y}.$$

Solutions of the first Euler-Lagrange equation for this integrand are automatically twice continuously differentiable and obey the differential equation

$$y'' + y = 0. \tag{8.1}$$

Solutions of (8.1) satisfying $y(0) = 0$ are given by

$$y(x) = k \sin x,$$

where k is a constant.

Let $\widetilde{y} \in \mathscr{Y}$ be given. For each $\xi \in (a, b]$, there is exactly one solution $\psi(\cdot, \xi)$ of (8.1) satisfying $\psi(0, \xi) = 0$ and $\psi(\xi, \xi) = \widetilde{y}(\xi)$. This solution is given by

$$\psi(x, \xi) = \frac{\widetilde{y}(\xi) \sin x}{\sin \xi} \quad \text{for all } x \in [0, \xi]. \tag{8.2}$$

Notice that a difficulty occurs if $L \geq \pi$ because the denominator in (8.2) vanishes when $\xi = \pi$. (For $L = \pi$, this difficulty can be overcome by a limiting arguemnt. However, if $L > \pi$ then J is unbounded below and consequently the difficulty cannot possibly be overcome in this case.)

The unique solution y_* of (8.2) on $[0, L]$ satisfying $y_*(0) = y_*(L) = 0$ is given by $y_*(x) = 0$ for all $x \in [0, L]$. Observe that $y_*(x) = \psi(L, x)$ for all $x \in [0, l]$.

Using L'Hôpital's rule, we find that

$$\lim_{\xi \to 0^+} \frac{\widetilde{y}(\xi)}{\sin \xi} = \widetilde{y}'(0).$$

It is therefore natural to put $\psi(0,0)=0$.

With ψ given by (8.2) we have

$$\psi_{,1}(x,\xi)=\frac{\widetilde{y}(\xi)\cos x}{\sin\xi} \tag{8.3}$$

and

$$\psi_{,2}(x,\xi)=\left(\frac{\widetilde{y}'(\xi)\sin\xi-\widetilde{y}(\xi)\cos\xi}{\sin^2\xi}\right)\sin x. \tag{8.4}$$

Remark 8.1 *For future reference, we note that*

(i) For fixed ξ the function $x\mapsto\psi_{,1}(x,\xi)$ is differentiable and we have

$$\psi_{,1,1}(x,\xi)=-\frac{\widetilde{y}(\xi)\sin x}{\sin\xi}.$$

(ii) For fixed x the function $\xi\mapsto\psi_{,1}(x,\xi)$ is differentiable and we have

$$\psi_{,1,2}(x,\xi)=\left(\frac{\widetilde{y}'(\xi)\sin-\widetilde{y}(\xi)\cos\xi}{\sin^2\xi}\right)\cos x.$$

(iii) For fixed ξ the function $x\mapsto\psi_{,2}(x,\xi)$ is differentiable and we have

$$\psi_{,2,1}(x,\xi)=\left(\frac{\widetilde{y}'(\xi)\sin\xi-\widetilde{y}'(\xi)\cos\xi}{\sin^2\xi}\right)\cos x.$$

(iv) The partial derivative $\psi_{,2,2}(x,\xi)$ does not exist unless $\widetilde{y}$ is twice differentiable at ξ. Nevertheless we have equality of mixed partials, namely $\psi_{,1,2}=\psi_{,2,1}$.

(v) There is a type of uniform convergence of $\psi(\cdot,\xi)$ to y_* and $\psi_{,1}(\cdot,\xi)$ to y_*' as $\xi\to L^-$. More precisely, for every $\epsilon>0$ there exists $\delta>0$ such that

$$|\psi(x,\xi)-y_*(x)|<\epsilon \text{ and } |\psi_{,1}(x,\xi)-y_*'(x)|<\epsilon \text{ whenever } \xi>L-\delta \text{ and } 0<x<\xi.$$

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

$$\sigma(\xi)=-\int_0^\xi[\psi_{,1}(x,\xi)^2-\psi(x,\xi)^2]\,dx-\int_\xi^L[\widetilde{y}'(x)^2-\widetilde{y}(x)^2]\,dx. \tag{8.5}$$

Ignorning the minus signs in (8.5), what we are doing in this formula is evaluating J along a path that rides the graph of a solution of the Euler-Lagrange equation from $(0,0)$ to $(\xi,\widetilde{y}(\xi))$ and then follows the graph of $\widetilde{y}$ from $(\xi,\widetilde{y}(\xi))$ to $(L,0)$.

Observe that

$$\sigma(L)=-\int_0^L[\psi_{,1}(x,L)^2-\psi(x,L)^2]\,dx=-\int_0^L[y_*(x)^2-y_*(x)^2]\,dx=-J(y_*)$$

and

$$\sigma(0) = -\int_0^L [\widetilde{y}'(x)^2 - \widetilde{y}(x)^2]\, dx = -J(\widetilde{y}).$$

It follows that

$$J(\widetilde{y}) - J(y_*) = \sigma(L) - \sigma(0).$$

If we can show that σ is continuous on $[0, L]$, differentiable on $(0, L)$, and $\sigma'(\xi) \geq 0$ for all $\xi \in (0, L)$ then we will know that $\sigma(L) \geq \sigma(0)$, which will ensure that $J(\widetilde{y}) \geq J(y_*)$.

Substitution of (8.2) and (8.3) into (8.5) gives

$$\sigma(\xi) = -\frac{\widetilde{y}(\xi)^2}{\sin^2 \xi} \int_0^\xi [\cos^2 x - \sin^2 x]\, dx - \int_\xi^L [\widetilde{y}'(x)^2 - \widetilde{y}(\xi)^2]\, dx. \tag{8.6}$$

The first integral can be computed explicitly using some trig identities. (We shall do so shortly.) The second integral in (8.6) cannot be computed explicitly; however, we can easily find the derivative (with respect to ξ) of this integral simply by applying the fundamental theorem of calculus.

To differentiate the term involving the first integral on the right-hand side of (8.6), let us put

$$F(\xi) = \frac{\widetilde{y}(\xi)^2}{\sin^2 \xi} \int_0^\xi [\cos^2 x - \sin^2 x]\, dx.$$

Using the double angle formulas

$$\cos^2 x - \sin^2 x = \cos 2x \text{ and } 2 \sin x \cos x = \sin 2x,$$

we see that

$$\int_0^\xi [\cos^2 x - \sin^2 x]\, dx = \sin \xi \cos \xi,$$

and we conclude that

$$F(\xi) = \widetilde{y}(\xi)^2 \cot \xi \quad \text{for all } \xi \in (0, L].$$

It follows that

$$\sigma(\xi) = -\widetilde{y}(\xi)^2 \cot \xi - \int_\xi^L [\widetilde{y}'(x)^2 - \widetilde{y}(x)^2]\, dx \quad \text{for all } \xi \in (0, L]. \tag{8.7}$$

Using L'Hôpital's rule, we find that

$$\lim_{\xi \to 0^+} \widetilde{y}(\xi)^2 \cot \xi = 0,$$

so that σ is continuous at 0. Continuity of σ at L is evident.

Employing the chain rule, product rule, fundamental theorem of calculus, and the identity $\csc^2 x - 1 = \cot^2 x$, we can differentiate the expression in (8.6) to obtain

$$\begin{aligned}
\sigma'(\xi) &= -2\widetilde{y}(\xi)\widetilde{y}'(\xi)\cot\xi + \widetilde{y}(\xi)^2\csc^2\xi + \widetilde{y}'(\xi)^2 - \widetilde{y}(\xi)^2 \\
&= \widetilde{y}(\xi)^2\cot^2\xi - 2\widetilde{y}(\xi)\widetilde{y}'(\xi)\cot\xi + \widetilde{y}'(x)^2 \\
&= (\widetilde{y}(\xi)\cot\xi - \widetilde{y}'(\xi))^2 \\
&\geq 0 \quad \text{for all } \xi \in (0, L).
\end{aligned}$$

We conclude that

$$J(\widetilde{y}) \geq J(y_*) = 0 \quad \text{for all } \widetilde{y} \in \mathscr{Y}.$$

In other words, the zero function minimizes J over $\mathscr{Y}$.

8.1.2 General Formula and the Weierstrass Excess Function

We now return to the general setting described at the beginning of the chapter. (The admissible functions take values in $\mathbb{R}^n$, where $n \in \mathbb{N}$ is a given natural number, not necessarily 1, and $a, b \in \mathbb{R}$, $A, B \in \mathbb{R}^n$, and $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are unspecified.)

Let $\widetilde{y}, y_* \in \mathscr{Y}$ be given and assume that y_* satisfies $(\text{E-L})_1$. We require that Assumption (WM) below holds. [(WM) stands for Weierstrass Method here.]

ASSUMPTION (WM): For every $\xi \in (a, b]$ there is exactly one solution $\psi(\cdot, \xi)$ of $(\text{E-L})_1$ satisfying

$$\psi(a, \xi) = A \text{ and } \psi(\xi, \xi) = \widetilde{y}(\xi). \tag{8.8}$$

We put

$$\psi(a, a) = A.$$

As in Section 8.1.1, we define $\sigma : [a, b] \to \mathbb{R}$ by

$$\sigma(\xi) = -\int_a^\xi f(x, \psi(x, \xi), \psi_{,1}(x, \xi))\, dx - \int_\xi^b f(x, \widetilde{y}(x), \widetilde{y}'(x))\, dx, \tag{8.9}$$

so that

$$\sigma(a) = -J(\widetilde{y}) \text{ and } \sigma(b) = -J(y_*).$$

We want to know that σ is continuous on $[a, b]$, differentiable on (a, b) and satisfies $\sigma'(\xi) \geq 0$ for all $\xi \in (a, b)$.

Continuity of σ at 0 will be ensured if we know that

$$\begin{aligned}\exists M \in \mathbb{R}, \delta > 0 \quad & \text{such that } |f(x,\psi(x,\xi),\psi_{,1}(x,\xi))| \leq M \\ & \text{for all } x,\xi \text{ with } a < x < \xi.\end{aligned} \tag{8.10}$$

(The condition in (8.10) guarantees that the first integral on the right-hand side of (8.9) remains tends to 0 as $\xi \to a^+$.) Continuity of σ at b will be ensured if we have

$$\begin{aligned}&\forall \epsilon > 0, \quad \exists \delta > 0 \text{ such that} \\ &\qquad |\psi(x,\xi) - y_*(x)| < \epsilon \text{ and} \\ &\qquad |\psi_{,1}(x,\xi) - y_*'(x)| < \epsilon \text{ whenever } \xi > L - \delta \text{ and } 0 < x < \xi.\end{aligned} \tag{8.11}$$

The second integral in (8.9) is automatically differentiable and its derivative is given by the fundamental theorem of calculus. In order to differentiate the first integral we shall make use of a result known as Leibniz' rule (for differentiating integrals). We record below a special case of this result that is sufficient for the present purpose.

Remark 8.2 (Special Case of Leibniz' Rule) *Let*

$$S := \{(x,\xi) \in \mathbb{R}^2 : a \leq x \leq \xi,\ a \leq \xi \leq b\} \tag{8.12}$$

and let $\rho : S \to \mathbb{R}$ be given.Assume that ρ and $\rho_{,2}$ are continuous on S and put

$$R(\xi) = \int_a^\xi \rho(x,\xi)\,dx \quad \text{for all } \xi \in [a,b].$$

Then R is differentiable and

$$R'(\xi) = \rho(\xi,\xi) + \int_a^\xi \rho_{,2}(x,\xi)\,dx \quad \text{for all } \xi \in [a,b].$$

In other words

$$\frac{d}{d\xi}\int_a^\xi \rho(x,\xi)\,dx = \rho(\xi,\xi) + \int_a^\xi \frac{\partial}{\partial\xi}\rho(x,\xi)\,dx.$$

Let us put

$$F(\xi) = \int_a^\xi f(x,\psi(x,\xi),\psi_{,1}(x,\xi))\,dx \tag{8.13}$$

and

$$G(\xi) = \int\limits_{\xi}^{b} f(x, \widetilde{y}(x), \widetilde{y}'(x))\,dx. \tag{8.14}$$

We want to apply Remark 8.2 with

$$\rho(x,\xi) = f(x, \psi(x,\xi), \psi_{,1}(x,\xi)).$$

In order to do so, we need to make some assumptions concerning $\psi_{,2}$ and $\psi_{,1,2}$. (Looking ahead, in order to make use of the fact that $\psi(\cdot,\xi)$ satisfies the first Euler Lagrange equation, it will useful to make some assumptions concerning $\psi_{,1,1}$ and $\psi_{,2,1}$.) We assume that

$$\psi_{,2}, \psi_{1,1}, \psi_{1,2}, \psi_{2,1} \text{ exist and are continuous on } S, \tag{8.15}$$

and

$$\psi_{1,2}(x,\xi) = \psi_{2,1}(x,\xi) \quad \text{for all } (x,\xi) \in S, \tag{8.16}$$

where S is given by (8.12). [These are assumptions are at least plausible in veiew of Remark 8.1.]

Using Remark 8.2 and the fundamental theorem of calculus we find that

$$F'(\xi) = f(\xi, \psi(\xi,\xi), \psi_{,1}(\xi,\xi)) + \int\limits_{a}^{\xi} \frac{\partial}{\partial\xi} f(x, \psi(x,\xi), \psi_{,1}(x,\xi))\,dx, \tag{8.17}$$

$$G'(\xi) = -f(\xi, \widetilde{y}(\xi), \widetilde{y}'(\xi)). \tag{8.18}$$

Since (for fixed ξ) $\psi(\cdot,\xi)$ satisfies $(\text{E-L})_1$ we know that

$$f_{,2}(x, \psi(x,\xi), \psi_{,1}(x,\xi) = \frac{d}{dx} f_{,3}(x, \psi(x,\xi), \psi_{,1}(x,\xi)) \quad \text{for all } x \in [a,\xi]. \tag{8.19}$$

Using the chain rule, (8.16), and (8.19) we see that

$$\begin{aligned}
\tfrac{\partial}{\partial\xi} f(x, \psi(x,\xi), \psi_{,1}(x,\xi)) &= f_{,2}(x, \psi(x,\xi), \psi_{,1}(x,\xi)) \cdot \psi_{,2}(x,\xi) \\
&\quad + f_{,3}(x, \psi(x,\xi), \psi_{,1}(x,\xi)) \cdot \psi_{,1,2}(x,\xi) \\
&= \left(\tfrac{d}{dx} f_{,3}(x, \psi(x,\xi), \psi_{,1}(x,\xi))\right) \cdot \psi_{,2}(x,\xi) \\
&\quad + f_{,3}(x, \psi(x,\xi), \psi_{,1}(x,\xi)) \cdot \psi_{.2.1}(x,\xi) \\
&= \tfrac{\partial}{\partial x}[f_{,3}(x, \psi(x,\xi), \psi_{,1}(x,\xi)) \cdot \psi_{,2}(x, \psi(x,\xi))].
\end{aligned}$$

It follows that

$$\begin{aligned}\int_a^\xi \tfrac{\partial}{\partial\xi} f(x,\psi(x,\xi),\psi_{,1}(x,\xi))\,dx &= \int_a^\xi \tfrac{\partial}{\partial x}[f_{,3}(x,\psi(x,\xi),\psi_{,1}(x,\xi))\cdot\psi_{,2}(x,\xi)]\,dx \\ &= f_{,3}(\xi,\psi(\xi,\xi),\psi_{,1}(\xi,\xi))\cdot\psi_{,2}(\xi,\xi) \\ &\quad -f_{,3}(a,\psi(a,\xi),\psi_{,1}(a,\xi))\cdot\psi_{,2}(a,\xi).\end{aligned} \tag{8.20}$$

Since $\psi(a,\xi)=A$ for all $\xi\in[a,b]$ we conclude that

$$\psi_{,2}(a,\xi)=0 \quad \text{for all } \xi\in[a,b]. \tag{8.21}$$

Furthermore, since $\psi(\xi,\xi)=\widetilde{y}(\xi)$ for all $\xi\in[a,b]$ we conclude that

$$\psi_{,1}(\xi,\xi)+\psi_{,2}(\xi,\xi)=\widetilde{y}'(\xi) \quad \text{for all } \xi\in[a,b]. \tag{8.22}$$

Substituting (8.21) and (8.22) into (8.20) (and recalling that $\psi(\xi,\xi)=\widetilde{y}(\xi)$) we find that

$$\int_a^\xi \frac{\partial}{\partial\xi} f(x,\psi(x,\xi),\psi_{,1}(x,\xi))\,dx = f_{,3}(\xi,\widetilde{y}(\xi),\psi_{,1}(\xi,\xi))\cdot(\widetilde{y}'(\xi)-\psi_{,1}(\xi,\xi)). \tag{8.23}$$

Combining (8.17) and (8.23) we find that

$$F'(\xi)=f(\xi,\widetilde{y}(\xi),\psi_{,1}(\xi,\xi)+f_{,3}(\xi,\widetilde{y}(\xi),\psi_{,1}(\xi,\xi))\cdot(\widetilde{y}'(\xi)-\psi_{,1}(\xi,\xi)). \tag{8.24}$$

Recalling that $\sigma(\xi)=-F(\xi)-G(\xi)$, and using (8.18), (8.24) we obtain

$$\begin{aligned}\sigma'(\xi)=f(\xi,\widetilde{y}(\xi),\widetilde{y}'(\xi)) \quad &-f(\xi,\widetilde{y}(\xi),\psi_{,1}(\xi,\xi)) \\ &-f_{,3}(\xi,\widetilde{y}(\xi),\psi_{,1}(\xi,\xi))\cdot(\widetilde{y}'(\xi)-\psi_{,1}(\xi,\xi)).\end{aligned} \tag{8.25}$$

The following defininition is quite natural in view of (8.25).

Definition 8.1 (Weierstrass Excess Function) *Given $f\in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n)$, the* Weierstrass excess function *associated with f is the function $\mathcal{E}:[a,b]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ defined by*

$$\begin{aligned}\mathcal{E}(x,y,z,w)= \quad & f(x,y,w)-f(x,y,z)-f_{,3}(x,y,z)\cdot(w-z) \\ & \text{for all } x\in[a,b],\ y,z,w\in\mathbb{R}^n.\end{aligned} \tag{8.26}$$

The next is result an immediate consequence of Theorem 5.1 and Definition 8.1.

Proposition 8.1 *Let $f\in C^1([a,b]\times\mathbb{R}^n\times\mathbb{R}^n)$ be given. The following two statements are equivalent.*

(i) $\mathcal{E}(x,y,z.w)\geq 0$ for all $x\in[a,b]$, $y,z,w\in\mathbb{R}^n$.

(ii) *For every $x \in [a,b]$, $y \in \mathbb{R}^n$ the mapping $z \to f(x,y,z)$ is convex on $\mathbb{R}^n$.*

Putting all of the pieces together we see that

$$J(\widetilde{y}) - J(y_*) = \int_a^b \mathcal{E}(x, \widetilde{y}(x), \psi_{,1}(x,x), \widetilde{y}'(x))\, dx. \tag{8.27}$$

Consequently, if $y_* \in \mathscr{Y}$ satisfies (E-L$_1$ and there is a family ψ having all of the properties above and for every $x \in [a,b]$, $y \in \mathbb{R}^n$ the function $z \mapsto f(x,y,z)$ is convex on $\mathbb{R}^n$, then we can be sure that $J(\widetilde{y}) \geq J(y_*)$. Unfortunately, it seems to be extremely difficult to impose reasonable coonditions directly on f that will the existence of such a family ψ.

8.2 Fields

A powerful and beautiful alternative approach to obtaining a conclusion similar in spirit to Weierstrass' formula (8.27) is due to Hilbert and is based on the notion of *fields* of stationary functions.

8.2.1 Basic Definitions

Let $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$ be given and put

$$D = (\alpha, \beta) \times \mathbb{R}^n.$$

It is convenient to have the integrand f defined on an open set that "matches up" nicely with D. We therefore consider integrands $f : (\alpha,\beta) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.

Definition 8.2 *By a* field[1] *on D we mean a function $\Phi \in C^1(D;\mathbb{R}^n)$. The differential equation*

$$y'(x) = \Phi(x, y(x)) \tag{8.28}$$

is called the field equation *associated with Φ.*

Before proceeding, we pause to make a few comments about differential equations of the form (8.28). By a solution of (8.28) we mean a function $y_* \in C^1(I;\mathbb{R}^n)$, where $I \subset (\alpha,\beta)$ is an interval (with nonempty interior) such that $y_*'(x) = \Phi(x, y_*(x))$ for all $x \in I$. It is a standard result from the theory of differential equations that every solution $y_* : I \to \mathbb{R}^n$ of (8.28) can be extended to be defined on a maximal interval of existence $I_{max} \supset I$ in such a way that the extended function is a solution of (8.28); the maximal interval of existence is necessarily open. (To say that I_{max} is the maximal interval of existence of a solution means that the solution cannot be extended to an interval $I' \supsetneq I_{max}$ in such a way that it remains a solution.) Unless stated otherwise, when we speak of a solution of (8.28) it is understood that this solution has been extended to be defined on its maximal interval of existence.

[1] The term "field" is used with different meanings in other branches of mathematics.

Definition 8.3 *Solutions of the field equation* (8.28) *are called* field trajectories*.*

Remark 8.3 *Let Φ be a field on D. It is a consequence of the basic existence-uniqueness theory for first-order differential equations that through each point $(\xi, \eta) \in D$ passes exactly one field trajectory. In other words, given $(\xi, \eta) \in D$, the field equation (8.28) has exactly one (maximally extended) solution y_* satisfying $y_*(\xi) = \eta$.*

Fields having the property that all field trajectories satisfy the first Euler-Lagrange equation for a given $f : (\alpha, \beta) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are of special importance in the calculus of variations.

Definition 8.4 *Let $f \in C^1((\alpha, \beta) \times \mathbb{R}^n \times \mathbb{R}^n)$ be given and let Φ be a field on $D = (\alpha, \beta) \times \mathbb{R}^n$. We say that Φ is a* stationary field *for f on D provided that every field trajectory satisfies the first Euler-Lagrange equation for f on its maximal interval of existence; in other words, if y_* is a field trajectory with maximal interval of existence I_{max} then the mapping $x \mapsto f_{,3}(x, y_*(x), y_*'(x))$ is continuously differentiable on I_{max} and*

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

Around 1900, Hilbert realized the relevance to variational problems of certain line integrals of the form

$$\int_\Gamma P(x, y)\, dx + Q(x, y) \cdot dy$$

where $P : D \to \mathbb{R}$ and $Q : D \to \mathbb{R}^n$ are constructed in a special way from a variational integrand f and a field Φ.

Definition 8.5 *Given $f \in C^1((\alpha, \beta) \times \mathbb{R}^n \times \mathbb{R}^n)$ and a field Φ on $D = (\alpha, \beta) \times \mathbb{R}^n$, define $P : D \to \mathbb{R}$ and $Q :\to \mathbb{R}^n$ by*

$$Q(x, y) = f_{,3}(x, y, \Phi(x, y)) \quad \text{for all } (x, y) \in D, \tag{8.29}$$

$$P(x, y) = f(x, y, \Phi(x, y)) - \Phi(x, y) \cdot Q(x, y) \quad \text{for all } (x, y) \in D. \tag{8.30}$$

We say that Φ is an exact field *for f provided that there is a function $U \in C^1(D)$ such that*

$$P(x, y) = U_{,1}(x, y), \quad Q(x, y) = U_{,2}(x, y) \quad \text{for all } (x, y) \in D. \tag{8.31}$$

U is called a potential *for the pair (P, Q).*

We recall an important result from calculus concerning exact differentials. (Note that $(\alpha, \beta) \times \mathbb{R}^n$ is simply connected.)

Remark 8.4 *Assume that $f \in C^1((\alpha, \beta)) \times \mathbb{R}^n \times \mathbb{R}^n)$ and that the functions P and Q defined by (8.30) and (8.29) are continuously differentiable on $D = (\alpha, \beta) \times \mathbb{R}^n$. Then Φ is exact if and only if*

$$(Q_{,2}(x, y))^T = Q_{,2}(x, y) \text{ and } Q_{,1}(x, y) = P_{,2}(x, y) \quad \text{for all } (x, y) \in D.$$

8.2.2 Some Formulas Involving Gradients

Before discussing the significance of exact fields, we make a few conventions and observations concerning gradients of vector-valued functions.

Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. We identify the gradient of F evaluated at $y \in \mathbb{R}^n$, denoted by $\nabla F(y)$, with an $n \times n$ matrix. The i^{th} row of $\nabla F(y)$ is simply the gradient of the the i^{th} component of F evaluated at y; in other words

$$(\nabla F(y))_{ij} = F_{i,j}(y).$$

Remark 8.5 *[A Chain Rule]: If $g : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable and we define $h : \mathbb{R}^n \to \mathbb{R}$ by*

$$h(y) = g(F(y)) \quad \text{for all } y \in \mathbb{R}^n,$$

then h is continuously differentiable and

$$\nabla h(y) = \nabla g(F(y))\nabla F(y). \tag{8.32}$$

(Notice that the right-hand side of (8.32) is a row vector (with n components) times an $n \times n$ matrix and is therefore a row vector.)

Remark 8.6 *[Another Chain Rule] Let $I \subset \mathbb{R}$ be an interval. If $G : \mathbb{R}^n \to \mathbb{R}^n$ and $y_* : I \to \mathbb{R}^n$ are continuously differentiable and*

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

then H is continuously differentiable and

$$H'(x) = y_*'(x)(\nabla G(y_*(x)))^{\mathsf{T}} \quad \text{for all } x \in [a, b]. \tag{8.33}$$

(Notice that the right-hand side of (8.33) is a row vector (with n components) times an $n \times n$ matrix and is therefore a row vector.)

Remark 8.7 *[A Product Rule] If $F, G : \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable then*

$$\nabla(F \cdot G)(y) = F(y)\nabla G(y) + G(y)\nabla F(y) \quad \text{for all } y \in \mathbb{R}^n. \tag{8.34}$$

(Notice that each summand on the right-hand side of (8.34) is a row vector (with n components) times an $n \times n$ matrix and is therefore a row vector.)

8.2.3 Exact Fields are Stationary

We now show that every exact field is stationary.

Theorem 8.1 *Let $f \in C^1((\alpha, \beta) \times \mathbb{R}^n \times \mathbb{R}^n)$ be given and assume that $f_{,3}$ is continuously differentiable. Put $D = (\alpha, \beta) \times \mathbb{R}^n$ and define $P : D \to \mathbb{R}$, $Q : D \to \mathbb{R}^n$ by (8.29) and (8.30), and let Φ be a field on D. If Φ is exact then Φ is stationary.*

Proof. Assume that Φ is exact and let y_* be a field trajectory with maximal interval of existence I_{max}. We need to show that y_* satisfies $(\text{E-L})_1$ on I_{max}.

Recall that

$$\begin{aligned} Q(x,y) &= f_{,3}(x,y,\Phi(x,y)), \\ P(x,y) &= f(x,y,\Phi(x,y)) - \Phi(x,y)\cdot Q(x,y) \quad \text{for all } (x,y)\in D. \end{aligned}$$

Since $f_{,3}$ and Φ are assumed to be continuously differentiable, we know that P and Q are continuously differentiable. Using Remarks 8.5 and 8.7 we find that

$$\begin{aligned} P_{,2}(x,y) &= f_{,2}(x,y,\Phi(x,y)) + f_{,3}(x,y,\Phi(x,y))\Phi_{,2}(x,y) \\ &\quad -\Phi(x,y)Q_{,2}(x,y) - Q(x,y)\Phi_{,2}(x,y) \quad \text{for all } (x,y)\in D. \end{aligned} \tag{8.35}$$

Using (8.29), (8.35), and the fact that $P_{,2} = Q_{,1}$, we see that

$$Q_{,1}(x,y) = f_{,2}(x,y,\Phi(x,y)) - \Phi(x,y)Q_{,2}(x,y) \quad \text{for all } (x,y)\in D,$$

so that

$$f_{,2}(x,y,\Phi(x,y)) = Q_{,1}(x,y) + \Phi(x,y)Q_{,2}(x,y) \quad \text{for all } (x,y)\in D. \tag{8.36}$$

Using (8.36) and the fact that y_* is a field trajectory, we obtain

$$f_{,2}(x,y_*(x),y_*'(x)) = Q_{,1}(x,y_*(x)) + y_*'(x)Q_{,2}(x,y_*(x)) \quad \text{for all } x\in I_{max}. \tag{8.37}$$

On the other hand, since

$$f_{,3}(x,y_*(x),y_*'(x)) = Q(x,y_*(x)) \quad \text{for all } x\in I_{max},$$

we have

$$\begin{aligned} \tfrac{d}{dx} f_{,3}(x,y_*(x)y_*'(x)) &= \tfrac{d}{dx} Q(x,y_*(x)) \\ &= Q_{,1}(x,y_*(x)) + y_*'(x)(Q_{,2}(x,y_*(x)))^{\mathsf{T}} \\ &\quad \text{for all } x\in I_{max}. \end{aligned} \tag{8.38}$$

Observe that Remark 8.6 was used in obtaining (8.38). Using the fact that $Q_{,2}(x,y) = (Q_{,2}(x,y))^{\mathsf{T}}$ (see Remark 8.4), it follows from (8.37) and (8.38) that

$$f_{,2}(x,y_*(x),y_*'(x)) = \frac{d}{dx} f_{,3}(x,y_*(x),y_*'(x)) \quad \text{for all } x\in I_{max},$$

and the proof is complete. □

The converse of Theorem 8.1 is not true in general, but *is* true when $n = 1$.

Theorem 8.2 *Assume that $n = 1$. Let $f \in C^1((\alpha,\beta)\times\mathbb{R}\times\mathbb{R})$ be given and assume that $f_{,3}$ is continuously differentiable. Put $D = (\alpha,\beta)\times\mathbb{R}$ and define $P : D \to \mathbb{R}$, $Q : D \to \mathbb{R}^n$ by (8.29) and (8.30), and let Φ be a field on D. If Φ is stationary then Φ is exact.*

The proof of Theorem 8.2 is an exercise. (See Problem) It is important that the set D in Theorem 8.2 is simply connected.

8.3 Exact Fields and Hilbert's Invariant Integral

Assume that Φ is an exact field for f and that P and Q are given by (8.30) and (8.29). Then (using "traditional" notation) the line integral

$$\int_{\Gamma} P(x,y)\,dx + Q(x,y)\cdot dy \tag{8.39}$$

is path independent, meaning that given $(\xi_1,\eta_1),(\xi_2,\eta_2)\in D$ the value of the line integral in (8.39) is the same for all piecewise smooth directed paths Γ in D having initial point (ξ_1,η_1) and terminal point (ξ_2,η_2). (In fact, the value of the integral is $U(\xi_2,\eta_2)-U(\xi_1,\eta_1)$, where U is any potential for the pair (P,Q).) When P and Q are given by (8.30) and (8.29), the integral in (8.39) is called *Hilbert's Invariant Integral* for f associated with the field Φ. The idea will be to consider directed paths generated by functions in $\widetilde{y}\in C^1([a,b];\mathbb{R}^n)$ and exploit the fact that the value of the integral depends only on $(a,\widetilde{y}(a))$ and $(b,\widetilde{y}(b))$.

Formally, if $\widetilde{\Gamma}$ is the graph of a function $\widetilde{y}\in C^1([a,b];\mathbb{R}^n)$, oriented so that $(a,\widetilde{y}(a))$ is the initial point and $(b,\widetilde{y}(b))$ is the terminal point then

$$\int_{\widetilde{\Gamma}} P(x,y)\,dx + Q(x,y)\cdot dy = \int_a^b \left\{P(x,\widetilde{y}(x)) + Q(x,\widetilde{y}(x))\cdot\widetilde{y}'(x)\right\}dx. \tag{8.40}$$

Rather than attemp to give a precise meaning to line integrals of the type (8.39), we shall work directly with integrals of the type appearing on the right-hand side of (8.40) and express the values of these integrals in terms of a potential U for the pair (P,Q) and $(a,\widetilde{y}(a))$, $(b,\widetilde{y}(b))$.

Choose a potential $U\in C^1(D)$ for the pair (P,Q) and let $a,b\in\mathbb{R}$ be given with

$$\alpha < a < b < \beta.$$

Then, for each $\widetilde{y}\in C^1([a,b];\mathbb{R}^n)$, we have

$$\int_a^b \left\{P(x,\widetilde{y}(x)) + Q(x,\widetilde{y}(x))\cdot\widetilde{y}'(x)\right\}dx = U(b,\widetilde{y}(b)) - U(a,\widetilde{y}(a)). \tag{8.41}$$

If y_* is a field trajectory with maximal interval of existence $I_{max}\supset[a,b]$ then, since $y_*'(x)=\Phi(x,y_*(x))$ for all $x\in[a,b]$, we have

$$P(x,y_*(x)) + Q(x,y_*(x))\cdot y_*'(x) = f(x,y_*(x),y_*'(x)) \quad \text{for all } x\in[a,b]$$

and consequently

$$\int_a^b \left\{P(x,y_*(x)) + Q(x,y_*(x))\cdot y_*'(x)\right\}dx = \int_a^b f(x,y_*(x),y_*'(x))\,dx; \tag{8.42}$$

in other words, we have

$$\int_a^b f(x, y_*(x), y_*'(x))\, dx = U(b, y_*(b)) - U(a, y_*(a)). \tag{8.43}$$

Let $\widetilde{y} \in C^1([a,b];\mathbb{R}^n)$ be given. By adding and subtracting terms, and using the definitions of P and Q, we see that

$$\begin{aligned} f(x,\widetilde{y}(x),\widetilde{y}'(x)) &= f(x,\widetilde{y}(x),\widetilde{y}'(x)) - f(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)) \\ &\quad + f(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)) \\ &= f(x,\widetilde{y}(x),\widetilde{y}'(x)) - f(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)) \\ &\quad + P(x,\widetilde{y}(x)) + Q(x,\widetilde{y}(x)) \cdot \Phi(x,\widetilde{y}(x)) \\ &= f(x,\widetilde{y}(x),\widetilde{y}'(x)) - f(x,\widetilde{y}(x),\Phi(x.\widetilde{y}(x))) \\ &\quad + Q(x,\widetilde{y}'(x)) \cdot (\Phi(x,\widetilde{y}(x)) - \widetilde{y}'(x)) \\ &\quad + \Big\{ P(x,\widetilde{y}(x)) + Q(x,\widetilde{y}(x)) \cdot \widetilde{y}'(x) \Big\} \\ &\quad \text{for all } x \in [a,b]. \end{aligned} \tag{8.44}$$

Using the fact that $Q(x,\widetilde{y}(x)) = f_{,3}(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)))$, and recalling the definition of the Weierstrass excess function $\mathcal{E}$, we see that

$$\begin{aligned} f(x,\widetilde{y}(x)\widetilde{y}'(x)) &= \Big\{ P(x,\widetilde{y}(x)) + Q(x,\widetilde{y}(x)) \Big\} + \mathcal{E}(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)),\widetilde{y}'(x)) \\ &\quad \text{for all } x \in [a,b]. \end{aligned} \tag{8.45}$$

Integrating (8.45) from a to b and using (8.41) we obtain

$$\begin{aligned} \int_a^b f(x,\widetilde{y}(x),\widetilde{y}'(x))\, dx &= U(b,\widetilde{y}(b)) - U(a,\widetilde{y}(a)) \\ &\quad + \int_a^b \mathcal{E}(x,\widetilde{y}(x),\Phi(x,\widetilde{y}(x)),\widetilde{y}'(x))\, dx \end{aligned} \tag{8.46}$$

Remark 8.8 *Notice that since $\mathcal{E}(x, y_*(x), y_*'(x), y_*'(x)) = 0$ for all $x \in [a,b]$, it follows from (8.46) that if y_* is a field trajectory then (8.43) holds.*

We summarize the result of the computations above in a theorem.

Theorem 8.3 *Let $f \in C^1((\alpha,\beta)\times\mathbb{R}^n\times\mathbb{R}^n)$ be given, define P and Q by (8.30) and (8.29) and assume that Φ is an exact field for f on $D = (\alpha,\beta)\times\mathbb{R}^n$. Let U be a potential for the pair (P,Q) and let $a, b \in \mathbb{R}$ be given with $\alpha < a < b < \beta$. Then (8.46) holds for every $\widetilde{y} \in C^1([a,b];\mathbb{R}^n)$.*

8.3.1 Fixed Endpoints

We now apply Theorem 8.3 to problems with fixed endpoints.

Theorem 8.4 *Let $a, b, \alpha, \beta \in \mathbb{R}$ with $\alpha < a < b < \beta$, $A, B \in \mathbb{R}^n$, and $f \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$ be given. Put*

$$\mathscr{Y} = \{y \in C^1([a,b];\mathbb{R}^n) \,:\, y(a) = A \text{ and } y(b) = B\},$$

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

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

Assume that Φ is an exact field for f on $D = (\alpha, \beta) \times \mathbb{R}^n$. Let $y_*, \widetilde{y} \in \mathscr{Y}$ be given and assume that y_* is (the restriction to $[a,b]$ of) a field trajectory. Then we have

$$J(\widetilde{y}) - J(y_*) = \int_a^b \mathcal{E}(x, \widetilde{y}(x), \Phi(x, \widetilde{y}(x)), \widetilde{y}'(x))\,dx. \tag{8.47}$$

Proof. Choose a potential U for the pair (P, Q), where P and Q are given by (8.30) and (8.29). By (8.43) we have

$$J(y_*) = U(b, B) - U(a, A), \tag{8.48}$$

and by Theorem 8.3 we have

$$J(\widetilde{y}) = U(b, B) - U(a, A) + \int_a^b \mathcal{E}(x, \widetilde{y}(x), \Phi(x, \widetilde{y}(x)), \widetilde{y}'(x))\,dx. \tag{8.49}$$

Combining (8.48) and (8.49) we arrive at (8.47). $\square$

Remark 8.9 *There is, of course, a connection between (8.47) and the formula (8.14) from Section 8.1.2. Suppose that for each fixed ξ, $\psi(\cdot, \xi)$ is a stationary function and also that $\psi(x, x) = \widetilde{y}(x)$ for all x. If there were a field Φ such that $\psi_{,1}(x, \xi) = \Phi(x, \psi(x, \xi))$ then we would have $\psi_{,1}(x, x) = \Phi(x, \widetilde{y}(x))$ and, at least formally, (**??**) would look identical to (8.47). However, for the family ψ whose existence is postulated in Assumption (WM), we have $\psi(a, \xi) = A$ for all ξ (multiple stationary functions passing through the same point) and consequently (a, A) cannot be a point of smoothness of Φ because this would violate the condition that through each point in D passes exactly one field trajectory.*

Remark 8.10 *If the assumptions of Theorem 8.4 hold and the mapping $z \mapsto f(x, y, z)$ is convex for all $x \in [a,b]$, $y \in \mathbb{R}^n$ then y_* minimizes J on $\mathscr{Y}$.*

8.3.2 Example 8.3.2

Let $n = 1$, $\alpha = 0$, and $\beta = \pi$, so that $D = (0, \pi)\times\mathbb{R}$. Define $f : (0, \pi)\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ by

$$f(x, y, z) = z^2 - y^2 \quad \text{for all } x \in (0, \pi), y, z \in \mathbb{R}.$$

Solutions of the first Euler-Lagrange equation for this integrand are of class C^2 and satisfy

$$y'' + y = 0.$$

In order to construct a suitable field Φ for f on D we want a one-parameter family of solutions of the Euler-Lagrange equation having the property that exactly one solution from this family passes through each point in D. Such a family of solutions is given by

$$y(x) = k \sin x,$$

where k is a constant. To find a formula for Φ, we observe that since this family of solutions is characterized by

$$\frac{y(x)}{\sin x} \text{ is constant,}$$

we know that

$$\frac{d}{dx}\left(\frac{y(x)}{\sin x}\right) = 0,$$

which gives

$$\frac{y'(x) \sin x - y(x) \cos x}{\sin^2 x} = 0,$$

so that

$$y'(x) = y(x) \cot x \quad \text{for all } x \in (0, \pi). \tag{8.50}$$

We put

$$\Phi(x, y) = y \cot x \quad \text{for all } (x, y) \in D. \tag{8.51}$$

This field was constructed in such a way that we know that it is stationary for f. Therefore, since $n = 1$, we know it is also exact by Theorem 8.2. For purposes of illustration, we shall verify directly that this field is exact by explicitly exhibiting a potential U. Using (8.30) and (8.29), the formula for f, and the fact that $f_{,3}(x, y, z) = 2z$, we find that

$$Q(x, y) = 2y \cot x,$$

$$P(x, y) = y^2 \cot^2 x - y^2 - 2y^2 \cot^2 x = -y^2(1 + \cot^2 x) = -y^2 \csc^2 x.$$

We seek $U : D \to \mathbb{R}$ such that

$$U_{,1}(x, y) = -y^2 \csc^2 x, \quad U_{,2}(x, y) = 2y \cot x \quad \text{for all } (x, y) \in D. \tag{8.52}$$

It is easy to see that the function $U : D \to \mathbb{R}$ defined by

$$U(x, y) = y^2 \cot x \quad \text{for all } (x, y) \in D$$

satisfies 8.52.

Let $a, b \in (0, \pi)$ with $a < b$ be given and put

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

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

Put $y_*(x) = 0$ for all $x \in (0, \pi)$, and observe that y_* is a field trajectory. The Weierstrass excess function for this integrand is given by

$$\begin{aligned} \mathcal{E}(x,y,z,w) &= w^2 - z^2 - (z^2 - y^2) - 2z(w-z) \\ &= w^2 - z^2 - 2zw + 2z^2 = w^2 - 2zw + z^2 = (w-z)^2. \end{aligned}$$

It follows that y_* minimizes J on $\mathscr{Y}$.

Remark 8.11 *(i) Using a limiting argument, it can be shown that*

$$luint0\pi[y'(x)^2 - y(x)^2]\, dx \geq 0$$

For all $y \in C^1[0, \pi]$ with $y(0) = y(\pi) = 0$.

(ii) Let $\mu \in \mathbb{R}$ be given and define $\Phi_\mu : (\mu, \pi + \mu) \times \mathbb{R}\mathbb{R}$ by

$$\Phi_\mu(x,y) = y \cot(x - \mu) \quad \text{for all } (x,y) \in (\mu, \pi + \mu).$$

Then Φ_μ is an exact field for the integrand $f(x,y,z) = z^2 - y^2$ on $(\mu, \pi + \mu) \times \mathbb{R}$.

8.3.3 Free Endpoints

8.3.4 Example

8.3.5 The Hamilton Jacobi Equation

There is an extremely important connection between exact fields and a partial differential equation know as the *Hamilton Jacobi equation.* We shall merely touch the surface here.

For fixed $(x, y) \in D$ and $q \in \mathbb{R}^n$, we consider the equation

$$f_{,3}(x,y,z) = q \tag{8.53}$$

and assume that there is exactly one solution z that can be expressed as a smooth function of (x, y, q). More specifically, we assume that there is a smooth function $g : (\alpha, \beta) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ such that

$$f_{,3}(x,y,g(x,y,q)) = q \quad \text{for all } x \in (\alpha, \beta), y, q \in \mathbb{R}^n. \tag{8.54}$$

We define the *Hamiltonian* $H : (\alpha, \beta) \times \mathbb{R}^r \times \mathbb{R}^n \to \mathbb{R}$ by

$$-H(x, y, q) = f(x, y, g(x, y, q)) - q \cdot g(x, y, q). \tag{8.55}$$

Given a field $Phi \in C^1(D; \mathbb{R}^n)$, we put

$$??\begin{cases} Q(x, y) = f_{,3}(x, y, \Phi(x, y)) \\ P(x, y) = f(x, y, \Phi(x, y)) - Q(x, y) \cdot \Phi(x, y) \quad \text{for all } (x, y) \in D. \end{cases} \tag{8.56}$$

Notice that

$$\Phi(x, y) = g(x, y, Q(x, y)) \quad \text{for all } (x, y) \in D. \tag{8.57}$$

Using the definitions of H and Q and (8.57), we find that

$$\begin{aligned} -H(x, y, Q(x, y)) &= f(x, y, g(x, y, Q(x, y))) - Q(x, y) \cdot \Phi(x, y) \\ &= P(x, y) \quad \text{for all } (x, y) \in D. \end{aligned} \tag{8.58}$$

Suppose now that Φ is exact. Then we may choose a potential $U \in C^1(D)$ such that

$$P(x, y) = U_{,1}(x, y), \quad Q(x, y) = U_{,2}(x, y) \quad \text{for all } (x, y) \in D. \tag{8.59}$$

Combining (8.58) and (8.59) we arrive at

$$W_{,1}(x, y) + H(x, y, W_{,2}(x, y)) = 0 \tag{8.60}$$

Equation (8.60) is called the *Hamilton-Jacobi equation.* It is a first-order non-linear partial differential equation.

If one has a solution U of (8.60) on an open set $\widetilde{D} \subset D$, it can be used to construct an exact field Φ on $\widetilde{D}$.

8.3.6 Example

Let $n = 1$ and put $D = (0, \pi) \times \mathbb{R}$ and define $f : (0, \pi) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x, y, z) = z^2 - y^2 \quad \text{for all } x \in (0, \pi) \times \mathbb{R} \times \mathbb{R}.$$

For this example, equation (8.53) becomes

$$2z = q,$$

which is equivalent to $z = \frac{1}{2}q$, so we can put

$$g(x, y, q) = \frac{1}{2}q.$$

We find that

$$\begin{aligned} -H(x, y, q) &= -y^2 + \left(\tfrac{1}{2}q\right)^2 - q\left(\tfrac{1}{2}q\right) \\ &= -y^2 - \tfrac{1}{4}q^2. \end{aligned}$$

The Hamilton-Jacobi equation becomes

$$U_{,1}(x,y) + y^2 + \frac{1}{4}U_{,2}(x,y)^2 = 0. \tag{8.61}$$

In Example we found that the function $U : (0,\pi) \times \mathbb{R}$ defined by

$$U(x,y) = y^2 \cot x \quad \text{for all } (x,y) \in (0,\pi) \times \mathbb{R} \tag{8.62}$$

is the potential for an exact field. Let us verify that U satisfies the Hamilton Jacobi equation. We have

$$U_{,1}(x,y) = -y^2 \csc^2 x, \quad U_2(x,y) = 2y \cot x,$$

and consequently

$$\begin{aligned} U_{,1}(x,y) + y^2 + \tfrac{1}{4}U_{,2}(x,y)^2 &= -y^2 \csc^2 x + y^2 + \tfrac{1}{4}(4y^2)\cot^2 x \\ &= y^2(-\csc^2 x + 1 + \cot^2 x) \\ &= 0, \end{aligned}$$

which verifies that U is indeed a solution of the Hamilton Jacobi equation.