Chapter One

Introduction and Preview

In adolescence, I hated life and was continually on the verge of suicide, from which, however, I was restrained by the desire to know more mathematics.

BERTRAND RUSSELL, *The Conquest of Happiness*

Just as "nonlinear" is understood in mathematics to mean "not necessarily linear," we intend the term "nonsmooth" to refer to certain situations in which smoothness (differentiability) of the data is not necessarily postulated. One of the purposes of this book is to demonstrate that much of optimization and analysis which have evolved under traditional smoothness assumptions can be developed in a general nonsmooth setting; another purpose is to point out the benefits of doing so. We shall make the following points:

1. Nonsmooth phenomena in mathematics and optimization occur naturally and frequently, and there is a need to be able to deal with them. We are thus led to study differential properties of nondifferentiable functions.
2. There is a recent body of theory (*nonsmooth analysis*) and associated techniques which are well suited to this purpose.
3. The interest and the utility of the tools and methods of nonsmooth analysis and optimization are not confined to situations in which nonsmoothness is present.

Our complete argument in support of these contentions is the entirety of this book. In this chapter we get under way with a nontechnical overview of the theory and some of its applications. The final sections are devoted to placing in context the book's contributions to dynamic optimization (i.e., the

calculus of variations and optimal control), the single largest topic with which it deals.

1.1 EXAMPLES IN NONSMOOTH ANALYSIS AND OPTIMIZATION

For purposes of exposition, it is convenient to define five categories of examples.

1.1.1 Existing Mathematical Constructs

The first example is familiar to anyone who has had to prepare a laboratory report for a physics or chemistry class. Suppose that a set of observed data points $(x_0, y_0), (x_1, y_1), \ldots, (x_N, y_N)$ in the x-y plane is given, and consider the problem of determining the straight line in the x-y plane that best fits the data. Assuming that the given data points do not all lie on a certain line (any lab instructor would be suspicious if they did), the notion of "best" must be defined, and any choice is arbitrary. For a given line $y = mx + b$, the error e_i at the ith data point (x_i, y_i) is defined to be $|mx_i + b - y_i|$. A common definition of best approximating line requires that the slope m and the intercept b minimize $\{\sum_{i=0}^{N} e_i^2\}^{1/2}$ over all m and b (or, equivalently, $\sum_{i=0}^{N} e_i^2$). On the face of it, it seems at least as natural to ask instead that the total error $\sum_{i=0}^{N} e_i$ be minimized. The characteristics of the resulting solution certainly differ. In Figure 1.1, for example, the dashed line represents the "least total error" solution (see Example 2.3.17), and the solid line represents the "least total square error" solution. Note that the former ignores the anomalous data point which presumably corresponds to a gross measurement error. The least squares solution, in contrast, is greatly affected by that point. One or the other of these solutions may be preferable; the point we wish to make is that the function $\sum_{i=0}^{N} e_i$ is nondifferentiable as a function of m and b. Thus the usual methods for minimizing differentiable functions would be inapplicable to this function, and different methods would have to be used. Of course, the reason that the least square definition is the common one is that it leads to the minimization of a smooth function of m and b.

The two functions being minimized above are actually special cases of the L^2 and L^1 norms. The differentiability (or otherwise) of norms and of other classes of functions has been and remains a central problem in functional analysis. One of the first results in this area is due to Banach, who characterized the continuous functions x on $[0, 1]$ at which the supremum norm

$$\|x\| := \max_{0 \leq t \leq 1} |x(t)|$$

is differentiable. (His result is rederived in Section 2.8.)

An interesting example of a nondifferentiable function is the *distance function* d_C of a nonempty closed subset C of R^n. This is the function defined by

$$d_C(x) := \min\{|x - c| : c \in C\},$$

where $|\cdot|$ refers to the Euclidean norm. (It is a consequence of the results of Section 2.5 that when C is convex, for example, d_C must fail to be differentiable at any point on the boundary of C.) The distance function has been a useful tool in the geometrical theory of Banach spaces; it will serve us as well, acting as a bridge between the analytic and the geometric concepts developed later. As an illustration, consider the natural attempt to define directions of tangency to C, at a point x lying in C, as the vectors v satisfying $d_C'(x; v) = 0$, where the notation refers to the customary one-sided directional derivative. Since such directional derivatives do not necessarily exist (unless extra smoothness or convexity hypotheses are made), this approach is only feasible (see Section 2.4) when an appropriate nonsmooth calculus exists.

As a final example in this category, consider the initial-value problem

$$\frac{d}{dt}x(t) = f(t, x(t)), \qquad x(0) = u.$$

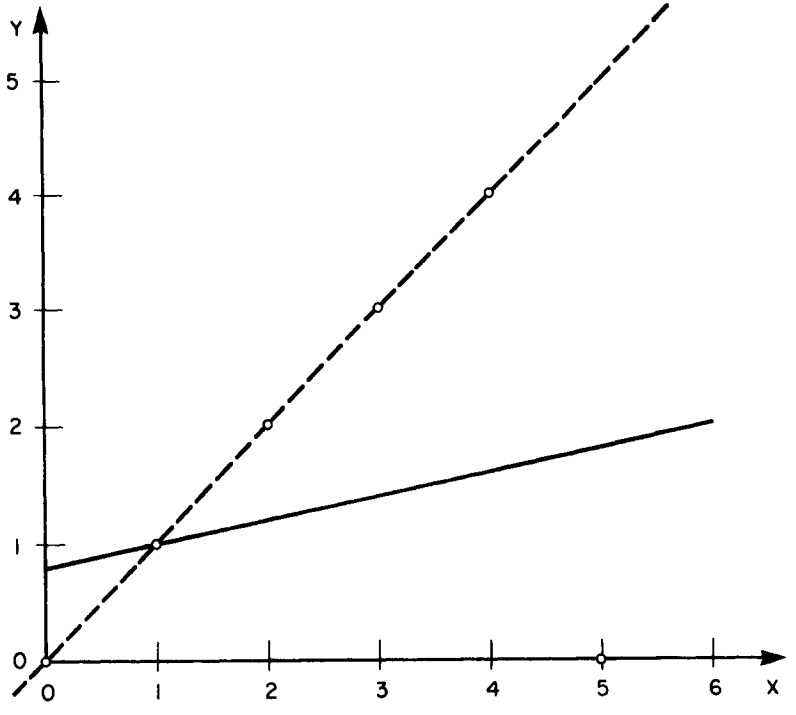

Figure 1.1 Two lines fitted to six data points (indicated by circles).

It is well known that the natural framework for studying existence and uniqueness of solutions is that of functions f which satisfy a Lipschitz condition in the x variable. It is desirable, then, to be able to study in this same framework the closely related issue of how solutions depend on the initial value u. The classical theory, however, hinges upon the resolvent, which is defined in terms of derivatives of f. This confines the analysis to smooth functions f. In Section 7.4 we extend the theory to the Lipschitz setting.

1.1.2 Direction Phenomena

Consider an elastic band whose upper end is fixed, and whose lower end is tied to a unit point mass. When the band is stretched a positive amount x, it exerts an upward (restoring) force proportional to x (Hooke's Law). When unstretched, no force is exerted. (This contrasts to a spring, which also exerts a restoring force when compressed.) If the mass is oscillating vertically, and if $x(t)$ measures the (positive or negative) amount by which the distance from the mass to the upper end of the band exceeds the natural (unstretched) length of the band, Newton's Law yields $\ddot{x}(t) = f(x(t))$, where f is given by

$$f(x) = \begin{cases} g - kx & \text{if } x \geqslant 0 \\ g & \text{if } x \leqslant 0. \end{cases}$$

(g is the acceleration of gravity, and k is the proportionality constant for Hooke's Law; friction and the weight of the band have been neglected.) Note that the function f is continuous but not differentiable at 0.

As another example, consider a flat solar panel in space. When the sun's rays meet its surface, the energy produced is proportional to $\cos \alpha$, where α is the (positive) angle of incidence (see Figure 1.2). When the panel's back is turned to the sun (i.e., when α exceeds $\pi/2$), no energy is produced. It follows then that the energy produced is proportional to the quantity $f(\alpha)$, where f is given by

$$f(\alpha) = \begin{cases} \cos \alpha & \text{if } \alpha \leqslant \pi/2 \\ 0 & \text{if } \alpha \geqslant \pi/2. \end{cases}$$

This again is a function that fails to be differentiable.

As a last illustration in this category, consider the electrical circuit of Figure 1.3 consisting of a diode, a capacitor, and an impressed voltage. A diode is a resistor whose resistance depends upon the direction of the current. If I is the current and V is the voltage drop across the diode, one has the following nonsmooth version of Ohm's Law:

$$I = \begin{cases} V/R_+ & \text{if } V \geqslant 0 \\ V/R_- & \text{if } V \leqslant 0, \end{cases}$$

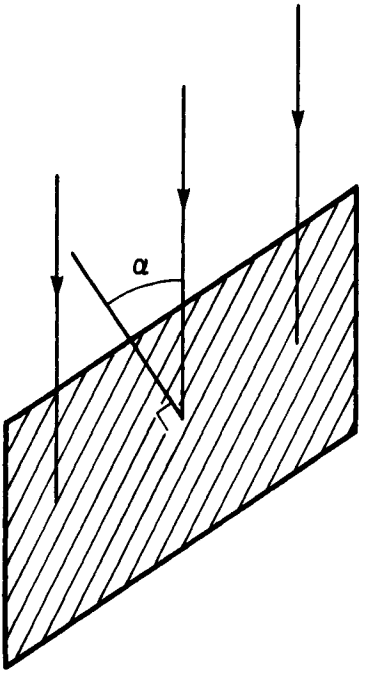

Figure 1.2 The angle of incidence between the sun's rays and a solar panel.

where R_+ and R_- are different positive constants. If x is the voltage across the capacitor, application of Kirchhoff's Laws yields

$$\frac{dx}{dt} = \begin{cases} \alpha(u - x) & \text{if } x \leqslant u \\ -\beta(x - u) & \text{if } x \geqslant u, \end{cases}$$

where u is the impressed voltage. In Section 5.3 we shall formulate and solve an optimal control problem with these nonsmooth dynamics.

1.1.3 Objective Functions

The first example is drawn from engineering design in which, in the study of physical systems, an appealing criterion is to minimize the greatest eigenvalue

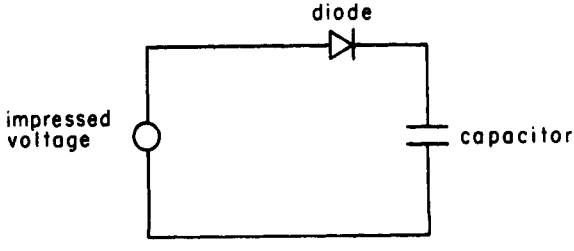

Figure 1.3 An electrical network.

of the system; that is, the greatest eigenvalue of a matrix A associated with it. Suppose that a vector x of parameters is open to choice, so that the objective function f to be minimized becomes

$$f(x) := \text{greatest eigenvalue of } A(x).$$

Even if the matrix A depends smoothly on x, the function f is not differentiable in general. We shall pursue this topic in Section 2.8 (Example 2.8.7).

As a second example, suppose we have a process which constructs electronic parts within a certain tolerance, and consider the problem of minimizing the error in manufacture. Specifically, suppose that when a state x is specified, the process actually produces the state $x + \tau$ for some τ in the tolerance set T, and let $\theta(x + \tau)$ measure the resulting distortion. Since τ is not known ahead of time, the worst-case distortion is

$$f(x) := \max_{\tau \in T} \theta(x + \tau).$$

We seek to minimize f over some feasible set of x values. The objective function f, which is not differentiable in general, is of a type which is studied in Section 2.8.

These two examples are minimization problems in which the possible constraints on x are not specified. In most cases such constraints exist and are themselves dependent upon parameters. Consider, for example, the problem of minimizing $f(x)$ subject to the constraints

$$g(x) + p \leqslant 0, \qquad h(x) + q = 0, \quad x \in C, \tag{1}$$

where C is a given closed subset of some Euclidean space X, and where $g\colon X \to R^n$, $h\colon X \to R^m$ are given functions. Here p and q are fixed vectors in R^n and R^m, respectively. The value (minimum) of this prototypical mathematical programming problem clearly depends upon p and q; let us designate it $V(p, q)$. The function V is of fundamental interest in the analysis of the problem, from both the theoretical and the computational points of view. Not only does V fail to be differentiable in general, it even fails to be finite everywhere. For example, our main interest could be in the case $p = 0$, $q = 0$, and $V(0, 0)$ could be finite. Yet there may easily be values of p and q arbitrarily near 0 for which there exist no values of x satisfying Eq. (1). In that case, $V(p, q)$ is assigned the value $+\infty$, in accord with the usual convention that the infimum over the empty set is $+\infty$. We shall analyze the differential properties of V in Chapter 6.

1.1.4 Threshold Phenomena

There are many economic and engineering situations which undergo a change when certain thresholds are breached. The canonical example is that of a dam holding a reserve of water whose level fluctuates with time (due, for example,

to tidal or other action). If $h(t)$ is the reservoir's natural water level at time t and if h_0 is the height of the dam, the amount of water that will have to be coped with downstream from the dam is proportional to $f(t)$, where

$$f(t) := \max\{h(t) - h_0, 0\}.$$

Of course, f is nondifferentiable in general, even if $h(\cdot)$ is smooth.

Let us turn to an economic example (of which there are a great number). Suppose that a firm is faced with a *demand function* $q(p)$ for its product. This means that the quantity demanded (in some units) is $q(p)$ when the asking price is set at p. Let Q be the largest quantity of the product that the firm can produce. Then, corresponding to any price p, the quantity that the firm can actually sell is

$$\min\{q(p), Q\},$$

so that effectively any firm with bounded output has a nonsmooth demand function.

One of the growth laws of the classical theory of capital growth is known as *fixed proportions*. It refers to situations in which inputs to production can only be used in certain proportions (e.g., one worker, one shovel). If, for example, L and K measure in some units two inputs that must be used in equal proportions, then the effective quantity of either input that is available to the production process is $\min\{L, K\}$. Most dynamic problems in capital growth or resource theory are treated by continuous models (i.e., neither time nor input levels are discrete) which incorporate some production law. If the fixed-proportions law is used, the resulting optimal control problem has nonsmooth dynamics. In Section 3.3 we study and solve such a problem in the theory of nonrenewable resources.

1.1.5 Mathematical Techniques

There exist mathematical methods which involve nonsmoothness in a fundamental way, and which can be very useful even in situations which seem smooth at first glance. We shall illustrate this with three such techniques, all of which will be used subsequently. The first is called *exact penalization*. To begin, let us consider the admittedly trivial problem of minimizing the function $f(x) = x$ over the reals, subject to the constraint $x = 0$. The minimum is 0, attained at the only point x satisfying the constraint (i.e., $x = 0$). The technique of exact penalization transforms constrained problems to unconstrained ones by adding to the original objective function a term that penalizes violation of the constraint. For example, as the reader may verify, the function $x + 2|x|$ attains its unique unconstrained minimum at $x = 0$ (which is the solution to the original constrained problem). Note that the penalty term $2|x|$ is nonsmooth. This is not due simply to the choice that was made; no smooth penalty function (i.e., a nonnegative function vanishing at 0) will ever produce

a problem whose minimum occurs at 0. For example, the function $x + kx^2$ does not attain a minimum at 0, no matter how large k is.

This simple example illustrates that exact penalization hinges upon nonsmoothness. We shall use the technique in Chapter 3 to replace a constraint $\dot{x} \in F(x)$, where x is a function on $[a, b]$, by a penalty term $\int_a^b \rho(x, \dot{x})\, dt$, where ρ is defined by

$$\rho(x, v) := \text{Euclidean distance from } v \text{ to } F(x).$$

This is reminiscent of the distance function cited earlier. In fact, distance functions frequently arise in connection with exact penalization (see Proposition 2.4.3).

Our second technique is the method which classically is known as the variational principle. It consists of obtaining some desired conclusion by constructing a functional F and finding a point u which minimizes it. Applying the stationarity conditions (e.g., $F'(u) = 0$) yields the required conclusion. Many of the initial applications of this method took place within the calculus of variations (hence the name). (A famous example is Dirichlet's Principle, which seeks to find solutions of Laplace's equation $u_{xx} + u_{yy} = 0$ by minimizing the functional $\int\int (u_x^2 + u_y^2)\, dx\, dy$.) Now it is clear that the method can only become more powerful as the functionals that can be considered become more general. In particular, the admissibility of nonsmooth functionals is useful, as we shall see.

Suppose that in applying the method of variational principles, or through some other means, we are led to consider a functional F which does not attain its infimum. (This is more likely to happen in infinite than in finite dimensions.) Further suppose that u is a point which "almost minimizes" F. Ekeland's Theorem, which is presented in Chapter 7, is a useful tool which says something about this situation. It asserts, roughly speaking, that there is a functional $\tilde{F}$ which is a slight perturbation of F and which attains a minimum at a point $\tilde{u}$ near u. The perturbed functional $\tilde{F}$ is of the form $\tilde{F}(v) = F(v) + k\|v - \tilde{u}\|$ for a positive constant k (where $\|\cdot\|$ is the norm on the space in question), so that the theorem leads to nonsmooth functionals.

The final example of how nonsmoothness can intervene in new mathematical techniques is provided by a *dual principle of least action* due to the author, and which we shall describe here only briefly. (The principle is used in Sections 7.7 and 7.8.) Hamilton's Principle of least action in classical mechanics amounts to the assertion that a physical system evolves in such a way as to minimize (or, more precisely, render stationary) the *action*, which is the variational functional

$$\int \{\langle p, \dot{x} \rangle - H(x, p)\}\, dt. \tag{2}$$

This leads to the Hamiltonian equations $-\dot{p} = \nabla_x H$, $\dot{x} = \nabla_p H$, which are

basic to classical mechanics. The functional (2) suffers from the serious mathematical defect of being indefinite; that is, it admits no local maxima or minima. This has limited its role, for example, in the qualitative theory of existence of solutions to Hamilton's equations.

Consider now the function G which is conjugate to H in the sense of convex analysis:

$$G(u, v) := \sup_{x, p} \{\langle u, x\rangle + \langle v, p\rangle - H(x, p)\}, \tag{3}$$

and let us define the *dual action* to be the functional

$$\int \{\langle \dot{p}, x\rangle + G(-\dot{p}, \dot{x})\}\, dt. \tag{4}$$

The function G, and hence the functional (4), are nonsmooth in general. In Chapter 4 we shall define and derive stationarity conditions for such functionals. It turns out that when H is convex, these stationarity conditions for the functional (4) almost coincide with the Hamiltonian equations. (To be precise, if (x, p) satisfies the generalized stationarity conditions for Eq. (4), then there exist translates of x and p by constants which satisfy the Hamiltonian equations.)

The dual action, Eq. (4), in contrast to the classical action, Eq. (2), can be shown to attain a minimum for certain classes of Hamiltonians H. In consequence, the dual action, which is a nonsmooth and nonconvex functional, has proven to be a valuable tool in the study of classical Hamiltonian systems, notably in the theory of periodic solutions.

1.2 GENERALIZED GRADIENTS

The generalized gradient is a replacement for the derivative. It can be defined for a very general class of functions (and will be in Chapter 2). Our purpose here is to give a nontechnical summary of the main definitions for those whose primary interest lies in the results of later chapters. We begin the discussion with the simplest setting, that of a locally Lipschitz real-valued function defined on R^n (n-dimensional Euclidean space).

Let $f\colon R^n \to R$ be a given function, and let x be a given point in R^n. The function f is said to be *Lipschitz near* x if there exist a scalar K and a positive number ε such that the following holds:

$$|f(x'') - f(x')| \leqslant K|x'' - x'| \quad \text{for all } x'', x' \text{ in } x + \varepsilon B.$$

(Here B signifies the open unit ball in R^n, so that $x + \varepsilon B$ is the open ball of radius ε about x.) This is also referred to as a Lipschitz condition of rank K.

We wish to study differential properties of f. Note to begin with that f need not be differentiable at x. Indeed, it is not hard to produce a function f which is Lipschitz near x and which fails to admit even one-sided directional derivatives at x.

We define instead the *generalized directional derivative* of f which, when evaluated at x in the direction v, is given by

$$f^\circ(x; v) := \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{f(y + \lambda v) - f(y)}{\lambda}. \tag{1}$$

The difference quotient whose upper limit is being taken is bounded above by $K|v|$ (for y sufficiently near x and λ sufficiently near 0) in light of the Lipschitz condition, so that $f^\circ(x; v)$ is a well-defined finite quantity. One might conceive of using other expressions in defining a generalized directional derivative. What makes f° above so useful is that, as a function of v, $f^\circ(x; v)$ is positively homogeneous and subadditive. This fact allows us to define a nonempty set $\partial f(x)$, the *generalized gradient* of f at x, as follows:

$$\partial f(x) := \{\zeta \in R^n : f^\circ(x; v) \geq \langle v, \zeta \rangle \text{ for all } v \text{ in } R^n\}. \tag{2}$$

By considering the properties of f°, it is possible to show that $\partial f(x)$ is a nonempty convex compact subset of R^n. One has, for any v,

$$f^\circ(x; v) = \max\{\langle \zeta, v \rangle : \zeta \in \partial f(x)\}, \tag{3}$$

so that knowing f° is equivalent to knowing $\partial f(x)$.

When f is smooth (continuously differentiable), $\partial f(x)$ reduces to the singleton set $\{\nabla f(x)\}$, and when f is convex, then $\partial f(x)$ coincides with what is called the subdifferential of convex analysis; that is, the set of vectors ζ in R^n satisfying

$$f(x + u) - f(x) \geq \langle u, \zeta \rangle \quad \text{for all } u \text{ in } R^n.$$

The computation of $\partial f(x)$ from this definition may appear to be a formidable task. In fact, it is something that one seeks to avoid, just as in differential calculus, where one rarely computes derivatives from the definition. Instead, one appeals to a body of theory that characterizes generalized gradients of certain kinds of functions, and to rules which relate the generalized gradient of some compound function (e.g., sum, composition) to those of its simpler components. This calculus of generalized gradients is developed in detail in Chapter 2.

We shall see that there are several equivalent ways of defining the generalized gradient. One alternate to the route taken above hinges upon Rademacher's Theorem, which asserts that a locally Lipschitz function is

differentiable almost everywhere (in the sense of Lebesgue measure). Let Ω_f be the set of points in $x + \varepsilon B$ at which f fails to be differentiable, and let S be any other set of measure zero. We shall obtain in Section 2.5 the following characterization of the generalized gradient:

$$\partial f(x) = \operatorname{co}\Big\{ \lim_{i \to \infty} \nabla f(x_i) : x_i \to x,\ x_i \notin S,\ x_i \notin \Omega_f \Big\}. \tag{4}$$

In words; $\partial f(x)$ is the convex hull of all points of the form $\lim \nabla f(x_i)$, where $\{x_i\}$ is any sequence which converges to x while avoiding $S \cup \Omega_f$.

Normals and Tangents

Let C be a nonempty closed subset of R^n. An interesting, nonsmooth, Lipschitz function related to C is its *distance function* d_C, defined by

$$d_C(x) = \min\{|x - c| : c \in C\}.$$

The generalized directional derivative defined earlier can be used to develop a notion of tangency that does not require C to be smooth or convex. The *tangent cone* $T_C(x)$ to C at a point x in C is defined as follows:

$$T_C(x) := \{v \in R^n : d_C^\circ(x; v) = 0\}. \tag{5}$$

It can be shown that this condition does in fact specify a closed convex cone.

Having defined a tangent cone, the likely candidate for the normal cone is the one obtained from $T_C(x)$ by polarity. Accordingly, we define $N_C(x)$, the *normal cone* to C at x, as follows:

$$N_C(x) := \{\zeta : \langle \zeta, v \rangle \leqslant 0 \text{ for all } v \text{ in } T_C(x)\}. \tag{6}$$

It follows then that $N_C(x)$ is the closed convex cone generated by $\partial d_C(x)$.

It is natural to ask whether T_C and N_C can be defined directly without recourse to the distance function. We shall prove in Section 2.4 that a vector v belongs to $T_C(x)$ if and only if (iff) it satisfies the following property:

(7) For every sequence x_i in C converging to x and every sequence t_i in $(0, \infty)$ converging to 0, there is a sequence v_i converging to v such that $x_i + t_i v_i$ belongs to C for all i.

The direct definition of $N_C(x)$ requires the notion of *perpendicular*. A nonzero vector v is *perpendicular* to C at a point x in C (this is denoted $v \perp C$ at x) if $v = x' - x$, where the point x' has unique closest point x in C. (Equivalently, $v = x' - x$, where there is a closed ball centered at x' which

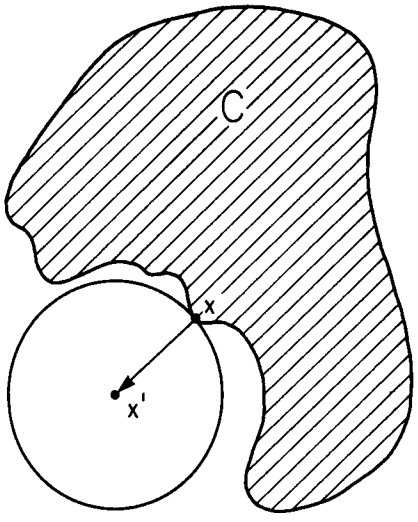

Figure 1.4 The vector $x' - x$ is perpendicular to C at x.

meets C only at x; see Figure 1.4.) It will be shown that one has

$$(8) \qquad N_C(x) = \overline{\text{co}}\left\{ \lambda \lim \frac{v_i}{|v_i|} : \lambda \geqslant 0,\ v_i \perp C \text{ at } x_i,\ x_i \to x,\ v_i \to 0 \right\}.$$

(An overbar denotes closure.) This is probably the easiest way to see how the normal cones of Figure 1.5 are obtained; tangents follow by polarity.

Everything Is One

We now have two analytic notions for functions (f° and ∂f) and two geometric notions for sets (T_C and N_C). Each of these pairs is dual: f° and ∂f are obtainable one from the other via Eq. (3), while T_C and N_C are each related to the other by polarity. It turns out moreover that these analytic and geometric concepts can be linked to one another. This is an important factor in their utility, just as chess pieces gain in strength when used cooperatively. The key to the relationship is the *epigraph* of a function f, denoted epi f, which is

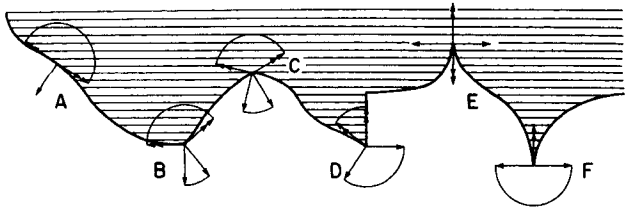

Figure 1.5 Five normal (single-arrow) and tangent (double-arrow) cones to a set.

the set

$$\{(x, r) \in R^n \times R : f(x) \leqslant r\}$$

(i.e., the set of points on or above the graph of f). In Chapter 2 we shall prove the following formulas:

$$T_{\text{epi}\, f}(x, f(x)) = \text{epi}\, f^\circ(x; \cdot) \tag{9}$$

$$\partial f(x) = \{\zeta : (\zeta, -1) \in N_{\text{epi}\, f}(x, f(x))\}. \tag{10}$$

(This last relationship reflects the familiar calculus result that the vector $[f'(x), -1]$ is normal to the graph of f at the point $(x, f(x))$.) It follows that T_C, N_C, ∂f, and f° form a circle of concepts so that any of them can be defined (by Eq. (7), (8), (4), or (1), respectively) as the nucleus from which the others evolve.

Thus far the functions we have been considering have been locally Lipschitz. This turns out to be a useful class of functions, and not only because a very satisfactory theory can be developed for it. As we shall see, the local Lipschitz property is a reasonable and verifiable hypothesis in a wide range of applications, and one that is satisfied in particular when smoothness or convexity is present. Locally Lipschitz functions possess the important property of being closed under the main functional operations (sums, compositions, etc.).

Nonetheless, there are good reasons for wanting to develop the theory for non-Lipschitz functions as well, and we shall do so in Chapter 2. For now, let us mention only that the key is to use Eq. (10) as a *definition* of ∂f in the extended setting, rather than as a derived result. When f is not necessarily Lipschitz near x, $\partial f(x)$ may not be compact, and may be empty.

If the set in Figure 1.5 is interpreted as the epigraph of a function f, then (using Eq. (10)) we find that at the x value corresponding to point A we have $\partial f = \{-1\}$, and f is smooth there. The function is Lipschitz near the points corresponding to B and C, and one has $\partial f = [0, 1]$ and $[-\frac{1}{2}, 1]$, respectively. At D there is a discontinuity, and $\partial f = [-1, \infty)$. At E, ∂f is empty, while at F it coincides with the real line.

Chapter 2 goes beyond the theory of this section in two further directions, in considering vector-valued functions, and in allowing the domain of f to be a (possibly infinite-dimensional) Banach space.

1.3 THREE PARADIGMS FOR DYNAMIC OPTIMIZATION

In later chapters we shall be dealing with three distinct though overlapping problems in optimal control and the calculus of variations. We propose now to describe them and discuss their origins and relative merits.

The Problem of Bolza: P_B

The calculus of variations will soon be celebrating its three-hundredth birthday (some would say posthumously). It is a beautiful body of mathematics in which the origins of several of today's mathematical subdisciplines can be found, and to which most of mathematics' historic names have contributed. The basic problem in the calculus of variations is the following: to minimize, over all functions x in some class mapping $[a, b]$ to R^n, the integral functional

$$\int_a^b L(x(t), \dot{x}(t))\, dt. \tag{1}$$

In the main developments of the theory (which, along with other matters, will be discussed in the following section), the given function L is smooth (four times continuously differentiable will do). A common version of the problem is that in which the functions x admitted above must satisfy $x(a) = x_0$, $x(b) = x_1$, where x_0 and x_1 are given (the *fixed endpoint* problem). Another classical version of the problem is the so-called *problem of Bolza*, in which the functional to be minimized is given by

$$l(x(a), x(b)) + \int_a^b L(x(t), \dot{x}(t))\, dt. \tag{2}$$

An *arc* is an absolutely continuous function mapping $[a, b]$ to R^n. We shall adopt the problem of minimizing the functional (2) over all arcs x as our first paradigm; we label it P_B. The problem as we shall treat it, however, is much more general than the classical problem of Bolza, due to the fact we shall allow l and L to be extended-valued. For example, the following choice for l is permissible:

$$l(u, v) = \begin{cases} 0 & \text{if } u = x_0 \text{ and } v = x_1 \\ +\infty & \text{otherwise.} \end{cases}$$

With this choice of l, minimizing (2) is equivalent to minimizing (1) under the constraint $x(a) = x_0$, $x(b) = x_1$. Thus the fixed endpoint problem is a special case of P_B. In similar fashion, we may define L to take account of such constraints as, say,

$$g(x(t), \dot{x}(t)) \leqslant 0, \qquad a \leqslant t \leqslant b.$$

While the extension of the problem of Bolza just described takes us beyond the pale of the classical calculus of variations, the formal resemblance that we take with us is a valuable guide in building an extended theory, as we shall see presently. Let us first define the second paradigm.

The Optimal Control Problem: P_C

The optimal control problem, which is generally thought to have originated in the engineering literature of the 1940s and 1950s, can be formulated as follows. Let the word *control* signify any member of some given class of functions $u(\cdot)$ mapping $[a, b]$ to a given subset U of R^m. We associate to each control u an arc x (called the *state*) by means of the differential equation and initial condition

$$\dot{x}(t) = \phi(x(t), u(t)), \qquad x(a) = x_0, \tag{3}$$

where ϕ is given, and where the point x_0 is free to be chosen within a given set C_0. The problem P_C consists of choosing u and x_0 (and hence x) so as to minimize a given functional of the type

$$f(x(b)) + \int_a^b F(x(t), u(t))\, dt, \tag{4}$$

subject to the condition that $x(b)$ lie in a given set C_1. (In Chapter 5 we allow U, F, and ϕ to depend on t, and we admit a further *state constraint* $g(t, x(t)) \leqslant 0$, but in this chapter we suppress these possibilities in order to unburden the exposition.)

A large number of physical systems can be modeled by differential equations in which some parameter values can be chosen within a certain range. This is precisely the framework of Eq. (3), which explains in part why the optimal control problem has found such wide application. From the strictly mathematical point of view, the optimal control problem subsumes the classical calculus of variations. For example, the fixed-endpoint problem of minimizing (1) subject to $x(a) = x_0$, $x(b) = x_1$ is seen to be the special case of P_C in which

$$\phi(x, u) \equiv u, \qquad U = R^n, \qquad F = L,$$

$$C_0 = \{x_0\}, \qquad C_1 = \{x_1\}, \qquad f \equiv 0.$$

Conversely, it is also the case that many (but not all) optimal control problems can be rephrased as standard variational problems with inequality and equality constraints (which are admitted in the classical theory).

Accordingly, it seems that optimal control theory can simply be viewed as the modern face of the calculus of variations. Yet we would argue that this fails to reflect the actual state of affairs, and that in fact a "paradigm shift" has taken place which outweighs in importance the mathematical links between the two. There are significant distinctions in outlook; many of the basic questions differ, and different sorts of answers are required. We have found that by and large the methodology and the areas of application of the two disciplines have

little in common. We believe that this is due in large part to the fact that [as Young (1969) has pointed out] the smoothness requirements of the classical variational calculus are inappropriate in optimal control theory. Indeed, we would argue that in establishing a bridge between the two, nondifferentiability intervenes fundamentally; more anon.

We now complete our triad of central problems.

The Differential Inclusion Problem: P_D

A *multifunction* F from R^n to itself is a mapping such that, for each x in R^n, $F(x)$ is a subset of R^n. Given such a multifunction F, along with two subsets C_0 and C_1 of R^n and a function f on R^n, we define the problem P_D to consist of minimizing $f(x(b))$ over all arcs x satisfying

$$\dot{x}(t) \in F(x(t)), \qquad a \leqslant t \leqslant b \tag{5}$$

$$x(a) \in C_0, \qquad x(b) \in C_1.$$

(We refer to (5) as a *differential inclusion*.) In contrast to P_B and P_C, this problem has very little history or tradition associated with it. It is in a sense the antithesis of P_B, which suppresses all explicit constraints by incorporating them (via extended values) in the objective functional. The only explicit minimization in P_D, on the other hand, involves one endpoint; all else is constraint. The optimal control problem P_C is a hybrid from this point of view.

We now turn to the relationships among these three problems and their roles in what follows.

Equivalence: Why Three?

Some relationships between the problems P_B, P_C, and P_D are well known. For example, let us suppose that a problem P_C is given, and let us attempt to recast it in the mold of P_D. To begin with, we may suppose that F in P_C is identically zero, by using the following device: let $\bar{x}$ signify the point (x°, x) in $R \times R^n$, and define

$$\bar{\phi}(\bar{x}, u) = [F(x, u), \phi(x, u)], \qquad \bar{f}(\bar{x}) = f(x) + x^\circ$$

$$\bar{C}_0 = \{0\} \times C_0, \qquad \bar{C}_1 = R \times C_1.$$

Then the original P_C is equivalent to the problem of minimizing $\bar{f}(\bar{x}(b))$ over the arcs $\bar{x}$ and controls u, satisfying

$$\dot{\bar{x}}(t) = \bar{\phi}(\bar{x}(t), u(t))$$

$$\bar{x}(a) \in \bar{C}_0, \qquad \bar{x}(b) \in \bar{C}_1.$$

This continues to have the form of P_C, but with F identically zero. Assuming then that F in P_C is zero, we define the data of a problem P_D as follows: f, C_0, C_1 remain as they are, and the multifunction F is defined by $F(x) := \phi(x, U)$. Under very mild assumptions on ϕ, it follows that the arc x satisfies $\dot{x} \in F(x)$ iff there is a control u such that $\dot{x} = \phi(x, u)$. (This fact is known as Filippov's Lemma.) We deduce therefore that the original P_C is equivalent to the resulting special case of P_D, in the sense that x solves the latter iff there is a control u, which, along with x, solves the former.

In turn, it is easy to see that any problem of the form P_D can be viewed as a special case of P_B. We have merely to define l and L as follows:

$$l(u, v) = \begin{cases} f(v) & \text{if } u \in C_0 \text{ and } v \in C_1 \\ +\infty & \text{otherwise} \end{cases}$$

$$L(s, v) = \begin{cases} 0 & \text{if } v \in F(s) \\ +\infty & \text{otherwise.} \end{cases}$$

Together with the foregoing, then, we can summarize notationally as follows:

$$P_C \subset P_D \subset P_B. \tag{6}$$

The relationship $P_C \subset P_B$ can also be obtained directly without first reducing to $F = 0$ in the following way: define l as above, and define L by

$$L(s, v) = \inf\{F(s, u) : u \in U, v = \phi(s, u)\}. \tag{7}$$

(Note that in keeping with custom, the infimum over the empty set is taken to be $+\infty$.)

Thus far in the discussion we have been neglecting an important factor: the hypotheses under which each of the three problems is analyzed. For example, the function L produced by Eq. (7) will in general be extended-valued and not even continuous where it is finite, so that the transformation of P_C to P_B which uses Eq. (7) is of little use if we are not prepared to deal with such functions (rest assured, we are). Thus the practical import of such reductions depends upon hypotheses, and these in turn depend upon the issue under discussion (e.g., existence, necessary conditions, etc.). By and large, the relations embodied in (6) accurately reflect the hierarchy that exists. An example of a formal reduction that is incompatible with our hypotheses in general is the one in which P_B is phrased in the form of P_C by taking

$$\phi(x, u) \equiv u, \qquad U = R^n, \qquad F(x, u) = L(x, u).$$

(We ignore boundary terms.)

Why treat three problems if one of them subsumes the other two? One reason is that the differing structure of the problems (together with some

prodding from established custom) leads to certain results (e.g., necessary conditions) of differing nature according to the problem, and this can be advantageous. Many problems encountered in applications can be phrased as any one of the three central types we have defined. As we shall illustrate, the methods which correspond to one problem may yield more precise information than those of another.

Each of the three paradigms has advantages specific to it. P_B, by its very form, facilitates our drawing inspiration from the calculus of variations. It is possible within the framework of P_B to achieve a unification of the classical calculus of variations and optimal control theory. Indeed, the extension to P_B of classical variational methods, which requires consideration of nonsmoothness, turns out to be a powerful tool in optimal control theory. A disadvantage of P_B is that a rather high level of technical detail is required to deal with the full generality of the integrands L that may arise.

The standard optimal control problem P_C is a paradigm that has proven itself natural and useful in a wide variety of modeling situations. It is the only one of the three in which consideration of nondifferentiability can be circumvented to a great extent by certain smoothness hypotheses. As we shall see in the next section and in Section 5.3, however, the other formulations can be better in some cases.

The main advantage of P_D derives from the fact that its structure is the simplest of the three. For this reason, most of the results that we shall obtain for the problems P_B and P_C are derived as consequences of the theory developed for P_D in Chapter 3. P_D, being of a novel form, has of course not been used very much in modeling applications. It is our belief that some problems can most naturally be modeled this way; an example is analyzed in Section 3.3.

We turn now to a discussion of the main issues that arise in connection with variational and control problems, and the existing theory of such problems. We shall then be able to place in context the results of the three chapters on dynamic optimization.

1.4 THE THEORY OF OPTIMAL CONTROL AND THE CALCULUS OF VARIATIONS

The three central themes common to any optimization problem are (1) existence—is there a solution to the problem?; (2) necessary conditions—what clues are there to help identify solutions?; (3) sufficient conditions—how can a candidate be confirmed to be a solution? There are of course other issues that depend on the particular structure of the problem (such as controllability and sensitivity, which will also be studied), but we shall organize the present discussion around the three themes given above. We shall lead off with necessary conditions; we begin at the beginning.

Necessary Conditions in the Calculus of Variations

Consider the basic problem in the calculus of variations: that of minimizing the functional

$$\int_a^b L(x(t), \dot{x}(t))\, dt \tag{1}$$

over a class of arcs x with prescribed values at a and b, where $(s, v) \to L(s, v)$ is a given function. When L is sufficiently smooth, any (local) solution to this problem satisfies the *Euler–Lagrange equation*

$$\frac{d}{dt}\{\nabla_v L(x(t), \dot{x}(t))\} = \nabla_s L(x(t), \dot{x}(t)). \tag{2}$$

Another condition that a solution x must satisfy is known as the *Weierstrass condition*. This is the assertion that for each t, the following holds:

(3)

$$L(x(t), \dot{x}(t) + w) - L(x(t), \dot{x}(t)) \geqslant \langle w, \nabla_v L(x(t), \dot{x}(t))\rangle \quad \text{for all } w.$$

The other main necessary condition is known as the *Jacobi condition*; it is described in Section 4.3.

Classical mechanics makes much use of a function called the *Hamiltonian*. It is the function H derived from L (the *Lagrangian*) via the Legendre transform, as follows. Suppose that (at least locally) the equation $p = \nabla_v L(x, v)$ defines v as a (smooth) function of (x, p). Then we define

$$H(x, p) = \langle p, v(x, p)\rangle - L(x, v(x, p)). \tag{4}$$

It follows that if x satisfies the Euler–Lagrange equation (2), then the function $p(t) := \nabla_v L(x(t), \dot{x}(t))$ satisfies

$$-\dot{p}(t) = \nabla_x H(x(t), p(t)), \qquad \dot{x}(t) = \nabla_p H(x(t), p(t)), \tag{5}$$

a celebrated system of differential equations called *Hamilton's equations*.

In order to facilitate future comparison, let us depart from classical variational theory by defining the *pseudo-Hamiltonian* H_P:

$$H_P(x, p, v) := \langle p, v\rangle - L(x, v). \tag{6}$$

The reader may verify that the Euler–Lagrange equation (2) and the Weierstrass

condition (3) can be summarized in terms of H_P as follows:

$$(7)\qquad -\dot{p} = \nabla_x H_P(x, p, \dot{x}), \qquad \dot{x} = \nabla_p H_P(x, p, \dot{x})$$
$$H_P(x, p, \dot{x}) = \max_v H_P(x, p, v).$$

Now suppose that side constraints of the type, say, $g(x, \dot{x}) \leqslant 0$ are added to the problem. Inspired by the technique of Lagrange multipliers, we may seek to prove that for an appropriate multiplier function λ, a solution x satisfies the necessary conditions above for L replaced by $L + \langle \lambda, g \rangle$ (see Section 4.5). The proof of this *multiplier rule* (in its fullest generality) was the last great quest of the classical school of the calculus of variations; it was completed in the work of Bliss (1930) and McShane (1939). (See Hestenes (1966) for a thorough treatment of the approach.) It has been eclipsed in prominence, however, by the work of Pontryagin et al. (1962), who proved the *maximum principle*, which we shall now describe.

Necessary Conditions for P_C, P_D, and P_B

Suppose that we now constrain the arcs x under consideration to satisfy

$$(8)\qquad \dot{x}(t) = \phi(x(t), u(t)),$$

where $u(t)$ belongs to U. Then $L(x, \dot{x})$ can be expressed in the form $F(x, u)$, and the basic problem of minimizing (1), but subject now to (8), is seen to be the optimal control problem P_C described in the preceding section. If in the definition (6) of the pseudo-Hamiltonian H_P we take account of Eq. (8) by replacing v (which stands in for $\dot{x}$) with the term $\phi(x, u)$, we obtain H_P as a function of (x, p, u):

$$(9)\qquad H_P(x, p, u) = \langle p, \phi(x, u) \rangle - F(x, u).$$

The maximum principle is simply the assertion that the classical necessary conditions (7) continue to hold in this situation.

We have seen two approaches to giving necessary conditions for constrained problems: the multiplier rule (which is a Lagrangian approach) and the maximum principle (which employs the pseudo-Hamiltonian, and which is therefore essentially Lagrangian in nature also). Clearly, neither of these is a direct descendant of the Hamiltonian theory. What in fact is the true Hamiltonian H of a constrained problem?

To answer this question, we first note the extension of the Legendre transform defined by Fenchel. Applied to L, it expresses the Hamiltonian H as follows:

$$(10)\qquad H(x, p) = \sup_v \{\langle p, v \rangle - L(x, v)\}.$$

The advantage of this version of the transform as opposed to the classical procedure (see Eq. (4)) is that H can be defined without referring to derivatives of L. Now we have seen in the preceding section that the control problem P_C is equivalent to minimizing $\int L(x, \dot{x})\, dt$, where L is defined by

$$L(s, v) = \inf\{F(s, u) : u \in U, v = \phi(s, u)\}. \tag{11}$$

If this Lagrangian is substituted in Eq. (10), a simple calculation gives

$$H(x, p) = \sup_{u \in U} \{\langle p, \phi(x, u)\rangle - F(x, u)\}. \tag{12}$$

This is the true Hamiltonian for P_C; observe that (in light of Eq. (9))

$$H(x, p) = \sup_{u \in U} H_P(x, p, u).$$

Why has the true Hamiltonian been neglected? Because it is nondifferentiable in general, in contrast to H_P, which is as smooth as the data of the problem.

This causes no difficulty, however, if the nonsmooth calculus described in Section 1.2 can be brought to bear. We shall prove that if x solves the optimal control problem, then an arc p exists such that one has

$$\begin{bmatrix} -\dot{p}(t) \\ \dot{x}(t) \end{bmatrix} \in \partial H(x(t), p(t)), \tag{13}$$

which reduces to the classical Hamiltonian system (5) when H is smooth. We shall confirm in discussing existence and sufficiency that H is heir to the rich Hamiltonian theory of the calculus of variations.

The Hamiltonian H for the more general problem P_B is simply the function defined by Eq. (10); it shall figure prominently in Chapter 4. As noted earlier, the differential inclusion problem P_D is the case of P_B in which

$$L(x, v) = \begin{cases} 0 & \text{if } v \in F(x) \\ +\infty & \text{otherwise.} \end{cases}$$

When this Lagrangian is substituted in Eq. (10), we obtain

$$H(x, p) = \sup\{\langle p, v\rangle : v \in F(x)\},$$

a function which lies at the heart of things in Chapter 3. Thus, the three main problems P_C, P_D, and P_B all admit necessary conditions featuring (13); one need only interpret the Hamiltonian appropriately. (We are omitting in this section other components of the necessary conditions that deal with boundary terms.) The approach just outlined above is different than, for example, that of the maximum principle; the resulting necessary conditions are not equivalent.

The problem treated in Section 5.3 illustrates that (13) can in some cases provide more precise information than the maximum principle; the opposite is also possible.

Existence and Sufficient Conditions

The classical existence theorem of Tonelli in the calculus of variations postulates two main conditions under which the basic problem of minimizing (1) admits a solution. These are: first, the convexity of the function $v \to L(s, v)$ for each s, and second, certain growth conditions on L, a simple example being

$$L(s, v) \geqslant \alpha + \beta|v|^2 \quad \text{for all } s \text{ and } v \qquad (\beta > 0). \tag{14}$$

When L is convex in v, L and its Hamiltonian H (see Eq. (10)) are mutually conjugate functions, each obtainable from the other by the Fenchel transform. It follows that growth conditions on L can be expressed in terms of H. For example, (14) is equivalent to

$$H(s, p) \leqslant \frac{p^2}{4\beta} - \alpha \quad \text{for all } s \text{ and } p.$$

In Section 4.1 we state an extension in Hamiltonian terms due to R. T. Rockafellar (1975) of the classical existence theory to the generalized problem of Bolza P_B. Since P_C is a special case of P_B, this result applies to the optimal control problem as well (see Section 5.4).

Let us now turn to the topic of sufficient conditions. There are four main techniques in the calculus of variations to prove that a given arc actually does solve the basic problem. We examine these briefly with an eye to how they extend to constrained problems.

1. *Elimination.* This refers to the situation in which we know that a solution exists, and in which we apply the necessary conditions to eliminate all other candidates. This method is common to all optimization problems. Note the absolute necessity of an appropriate existence theorem, and of knowing that the necessary conditions apply.

2. *Convexity.* Every optimization problem has its version of the general rule that for "convex problems," the necessary conditions are also sufficient. It is easy to prove that if $L(s, v)$ is convex (jointly in (s, v)), then any (admissible) solution x of the Euler–Lagrange equation (2) is a solution of the basic problem. (As simple and useful as it is, this fact is not known in the classical theory.) As for P_B, we may use the fact that convexity of L is equivalent to the Hamiltonian $H(s, p)$ being concave in s (it is always convex in p) in order to state sufficient conditions in Hamiltonian terms (see Section 4.3). This approach can be extended to derive sufficient conditions for the optimal control problem (see Section 5.4).

3. *Conjugacy and Fields.* There is a beautiful chapter in the calculus of variations that ties together certain families of solutions to the Euler–Lagrange equation (*fields*), the zeros of a certain second-order differential equation (*conjugate points*), and solutions of a certain partial differential equation (the *Hamilton–Jacobi equation*), and uses these ingredients to derive sufficient conditions. There has been only partial success in extending this theory to constrained problems (e.g., P_C; see the method of geodesic coverings (Young, 1969), which is based on fields). Recently Zeidan (1982) has used the Hamiltonian approach and the classical technique of canonical transformations to develop a conjugacy approach to sufficient conditions for P_B (see Section 4.3).

4. *Hamilton–Jacobi Methods.* The classical Hamilton–Jacobi equation is the following partial differential equation for a function $W(t, x)$:

$$W_t(t, x) + H(x, W_x(t, x)) = 0.$$

It has long been known in classical settings that this equation is closely linked to optimality. A discretization of the equation leads to the useful numerical technique known as dynamic programming. Our main interest in it stems from its role as a verification technique. Working in the context of the differential inclusion problem P_D, we shall use generalized gradients and the true Hamiltonian to define a generalized Hamilton–Jacobi equation (see Section 3.7). To a surprising extent, it turns out that the existence of a certain solution to this extended Hamilton–Jacobi equation is both a necessary and a sufficient condition for optimality.