c.1

PROXIMAL NORMAL ANALYSIS

IN DYNAMIC OPTIMIZATION

By

PHILIP DANIEL LOEWEN

B.Sc., The University of Alberta, 1981

M.Sc., The University of British Columbia, 1983

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Mathematics

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1985

Philip Loewen

Department of Mathematics

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date 5 December 1985

Abstract

Proximal normal analysis is a relatively new technique whose power and breadth of applicability are only now being realized. Given an optimization problem, there are many ways to define a "value function" which describes the changes in the problem's minimum value as certain parameters are varied. The epigraph of this function, namely the set of points lying on or above its graph, is a set whose geometry is intimately connected both with necessary conditions for optimality in the original problem and with the problem's sensitivity to perturbations. Proximal normal analysis is the geometrical technique which allows such information to be derived from a study of this fundamental set. In the first chapter we illustrate the technique in the simple model framework of a finite-dimensional mathematical programming problem, and describe its consequences for parameter sensitivity in optimal control.

Chapter II presents a detailed proof of the fundamental geometric result, called the "proximal normal formula", in Hilbert space. The proof is distilled from the more general work of Borwein and Strojwas (1985), who were the first to make this basic ingredient of the method available in infinite dimensions. This extension is of considerable practical interest: in Chapter III it makes possible a proximal normal analysis of state constraints in optimal control, which gives rise to a new form of the maximum principle for state constrained problems.

Limiting techniques and existence theorems are key ingredients in proximal normal analysis. Chapter IV gives a new existence theorem for open-loop stochastic optimal control problems in which compactness of the control set is not required, but instead a growth condition is imposed on the problem's running cost. In addition to their independent interest, the methods and results of Chapter IV enable us to use proximal normal analysis to investigate parameter sensitivity in stochastic optimal control in Chapter V. A byproduct of this analysis is a new proof of the Stochastic Maximum Principle which is more direct (if slightly more technical) than the proofs current in the literature, and which provides a rigorous interpretation of the multipliers.

Table of Contents

List of Figures

Acknowledgement

It gives me great pleasure to thank Frank Clarke for serving as my research supervisor. Professor Clarke's optimism and love for mathematics are infectious; his breadth of knowledge and seemingly infallible intuition nothing less than awe-inspiring. Besides encouraging me implicitly by his example, Frank also spent countless hours discussing mathematics with me, and made enough suggestions to keep me busy for the foreseeable future. Among the many people to whom this project owes its completion and I owe my thanks, he deserves the first position.

I must also thank Professor Perkins for teaching me probability and for many helpful suggestions about Chapter IV. Professor Haussmann introduced me to the field of dynamic optimization in my first year of graduate studies, for which I will always be grateful. He also inspired Chapters IV and V during a 1984–85 course on the Stochastic Maximum Principle and took care of many administrative matters relating to my graduation.

David Hare also helped me overcome the logistic difficulties inherent in being a student three thousand miles removed from his University. The McGill School of Computer Science generously allowed me to use TeX on its UNIX system to produce the final version of the thesis. The support of NSERC through a 1967 Science Scholarship has also been indispensable. Thank you all.

Finally, I thank my wife Kimberley. Her periodic exhortations to take heart and work had great effect, being spoken in love. This work is respectfully dedicated to Kimberley Loewen.

Chapter I. An Overview of Proximal Normal Analysis

Proximal normal analysis is a young technique which is quickly establishing itself as an indispensable tool in the study of optimization. This thesis is an attempt to explain and add momentum to this process. The body of the work stands on two legs: exposition and research. First, Chapter I introduces proximal normal analysis in the model framework of finite-dimensional mathematical programming. This simple context clarifies the general form of the method and suggests the sort of results it can be expected to yield in more general settings. Chapter I continues with a review of parameter sensitivity in optimal control, one of the earliest triumphs of proximal normal analysis in dynamic optimization. After gathering the necessary technical equipment in Chapter II, most notably an infinite-dimensional version of the crucial "proximal normal formula", we go on to demonstrate two new applications of proximal normal analysis in the field of dynamic optimization. In the first of these (Chapter III) we show how it allows the derivation of a new form of the maximum principle for deterministic problems with state constraints. In the second (Chapter V) we apply it to stochastic control problems to obtain a new proof of the Stochastic Maximum Principle which affords a precise interpretation of the multipliers. Chapter IV contains an existence theorem for a stochastic optimal control problem in Bolza form which is used in Chapter V, but is also of independent interest. In each case, proximal normal analysis reveals an interplay between value functions, geometrical objects, and necessary conditions which is deep enough to yield impressive new results and wide enough to suggest many other areas for fruitful investigation.

Section 1. On Lagrange Multipliers

Consider the optimization problem

$$\min_{x \in X} \{\ell(x) \ : \ g(x) = 0\}, \tag{1.1}$$

where ℓ and g are smooth real-valued functions defined on a Banach space X. No matter how large the space X may be, this problem has a natural "image" in $\mathbf{R}^2$, namely the set

$$C = \{(g(x), \ell(x)) \ : \ x \in X\}. \tag{1.2}$$

In the example illustrated in Figure 0 below, C is the lightly drawn curve.

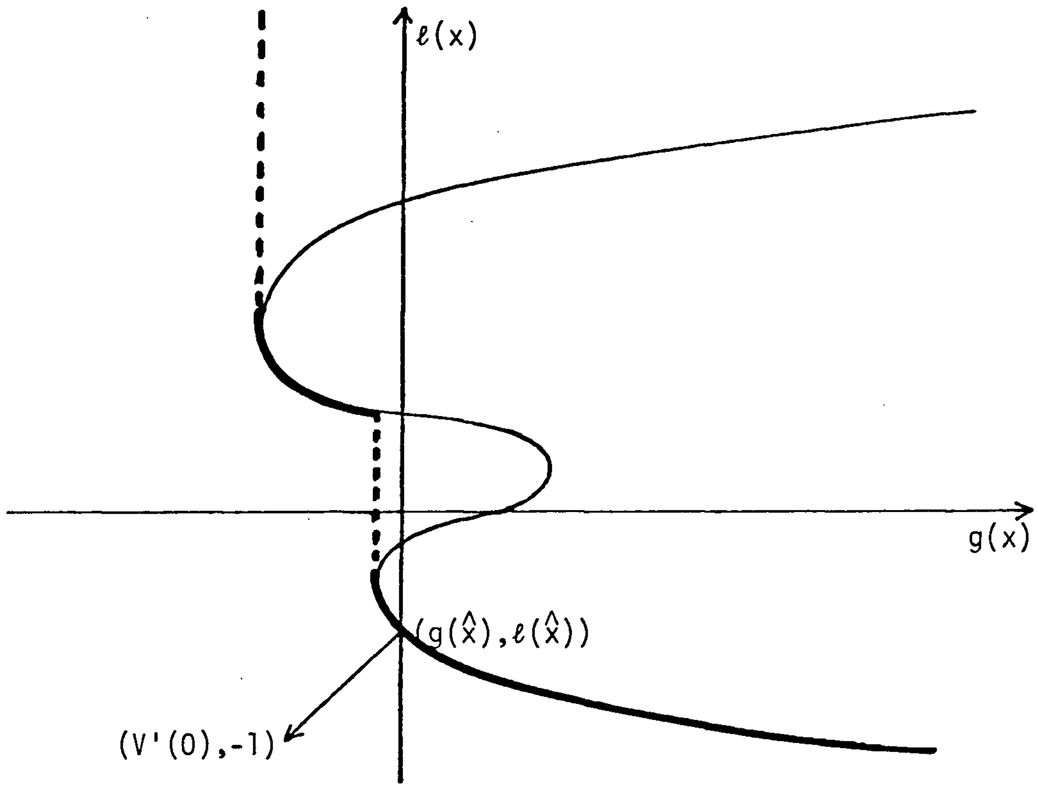

Fig. 0. The "image" of problem (1.1) in $\mathbf{R}^2$.

The minimization problem above can now be viewed as a two-step process: first find the lowest point on the y-axis lying in C, and then find an element of X realizing this image. If we were to seek the lowest points on other vertical lines, the following *value function* $V: \mathbf{R} \to \mathbf{R} \cup \{+\infty\}$ would emerge:

$$\begin{aligned} V(\alpha) &= \min \{v \ : \ (\alpha, v) \in C\} \\ &= \min_{x \in X} \{\ell(x) \ : \ g(x) = \alpha\}. \end{aligned}$$

In Fig. 0, the heavy line traces the graph of V. Note that in spite of the smoothness of g and ℓ, the function V may easily be discontinuous or even take the value $+\infty$. Without more hypotheses, we can only reasonably expect V to be lower semicontinuous. But suppose for the moment that V is finite-valued and differentiable at $\alpha = 0$, with $V(0) = \ell(\widehat{x})$ for some $\widehat{x} \in X$. Then the point $\big(V'(0), -1\big)$ is the downward normal to the graph of V at the point $\big(0, V(0)\big)$. If V is convex near 0, as illustrated in Fig. 0, it follows that one has

$$\big\langle \big(V'(0), -1\big), \big(\alpha, V(\alpha)\big) - \big(0, V(0)\big) \big\rangle \leq 0 \tag{1.3}$$

for all α near 0. In particular, for all x near $\widehat{x}$, one has $g(x)$ near $g(\widehat{x}) = 0$, so

$$\begin{aligned} & \big\langle \big(V'(0), -1\big), \big(g(x), \ell(x)\big) - \big(g(\widehat{x}), \ell(\widehat{x})\big) \big\rangle \leq 0 \\ \iff \quad & \ell(\widehat{x}) - V'(0)g(\widehat{x}) \leq \ell(x) - V'(0)g(x). \end{aligned} \tag{1.4}$$

Hence the right side has a local minimum over X at $\widehat{x}$, which forces its derivative to vanish:

$$D\ell(\widehat{x}) - V'(0)Dg(\widehat{x}) = 0. \tag{1.5}$$

(Here $D\ell$ is the Fréchet derivative of ℓ.) This is the familiar Lagrange multiplier rule for constrained minimization problems, with the "multiplier" being $-V'(0)$. Thus in the usual Lagrange multiplier rule one may interpret the multiplier corresponding to a given solution as an index of the marginal effects of perturbations of the corresponding constraint.

This simple introduction demonstrates three things:

1. Differential analysis of the value function can be used to prove multiplier rules. The analysis proceeds by applying known necessary conditions for an *unconstrained* problem to a "perpendicular inequality" like (1.3).
2. Conversely, known multiplier rules involve constituents which may have an interpretation as the marginal value associated with some appropriate perturbation.
3. Even for problems with smooth data, the value function can behave very badly.

The promising tone of articles 1 and 2 is in sharp contrast with the sad facts of life set down in article 3: differential analysis of V may indeed be profitable, but is often impossible. Besides the cowardly option of simply giving up in despair, there are two ways to deal with this conflict. The first is to impose a sufficient number of additional conditions (linearity, convexity, etc.) to guarantee that V is smooth. Such conditions often disallow the study of many problems of substantial practical interest. The second possibility is to broaden the scope of the phrase "differential analysis" in some way which allows its application to ill-behaved functions. This has been attempted by many authors for many reasons, but the "generalized gradient" discovered by Frank H. Clarke in 1973 has proven especially useful in the study of dynamic optimization. We now review those aspects of Clarke's theory most essential to the study of proximal normal analysis.

Section 2. Generalized Gradients

Let X be a Banach space, and let $f: X \to \mathbf{R} \cup \{+\infty\}$ be an extended-real-valued function on X. Suppose f is lower semicontinuous, and that $\widehat{x} \in X$ is a point where $f(\widehat{x})$ is finite. It is quite possible that f may fail to be differentiable at $\widehat{x}$: one way to explain this situation is that there is no single point $Df(\widehat{x}) \in X^*$ which adequately summarizes the local behaviour of f near $\widehat{x}$. Often there are several points in X^* which each give partial information about f's local behaviour. Clarke's *generalized gradient of f at $\widehat{x}$*, denoted $\partial f(\widehat{x})$, is a weak*-closed convex subset of X^* which contains all such points. In this section we discuss the precise definition and properties of the set-valued operation $\partial f(\cdot)$. All the definitions and theorems are taken from Clarke (1983).

Lipschitz Functions. Let us begin by considering the case when f is *Lipschitz near $\widehat{x}$*, that is, when there is a constant K and a neighbourhood U of $\widehat{x}$ such that

$$|f(y) - f(x)| \leq K \|y - x\| \qquad \forall x, y \in U.$$

In this case, the following quantity is finite for any $v \in X$:

$$f^0(\widehat{x}; v) = \limsup_{\substack{x \to \widehat{x} \\ h \downarrow 0}} \frac{f(x + hv) - f(x)}{h}. \tag{2.1}$$

This quantity is called the *generalized directional derivative of f at $\widehat{x}$ in direction v*. As a function of v, it is positively homogeneous and subadditive, so we may define

$$\partial f(\widehat{x}) = \{\varsigma \in X^* \,:\, \langle \varsigma, v \rangle \leq f^0(\widehat{x}; v) \quad \forall v \in X\}. \tag{2.2}$$

This is the *generalized gradient of f at $\hat{x}$*: it is a nonempty, weak*-compact, convex subset of X^*. Notice that the set $\partial f(\hat{x})$ is truly a generalization of the usual gradient. For if f is continuously differentiable ("smooth") near $\hat{x}$, then $f^0(\hat{x}; v) = \langle Df(\hat{x}), v\rangle \;\; \forall v \in X$, and thus $\partial f(\hat{x}) = \{Df(\hat{x})\}$. The following is a useful converse.

2.1 Proposition. *If $\partial f(\hat{x}) = \{\varsigma\}$ for some $\varsigma \in X^*$, then*

$$\lim_{\substack{x \to \hat{x} \\ h \downarrow 0}} \frac{f(x + hv) - f(x)}{h} = \langle \varsigma, v\rangle \qquad \forall v \in X$$

and convergence in uniform on every compact set of v-values. In Clarke's terminology, f is strictly differentiable at $\hat{x}$, with $D_s f(\hat{x}) = \varsigma$.

The utility of generalized gradients is based not only upon the evident similarity between their definitions and those of the classical concepts, but also because this similarity is preserved in many of the rules of calculus. For example, the classical rules concerning the derivative of a sum of two smooth functions, or of a scalar multiple of a smooth function, have natural counterparts in terms of generalized gradients. Moreover, differentiation rules for certain nonclassical means of combining functions can be easily expressed in terms of generalized gradients.

2.2 Proposition. *Let $f_1, \ldots, f_k : X \to \mathbf{R}$ be functions Lipschitz near $\hat{x}$.*

(a) For any scalars c_1 and c_2,

$$\partial\Big(c_1 f_1 + c_2 f_2\Big)(\hat{x}) \subseteq c_1 \partial f_1(\hat{x}) + c_2 \partial f_2(\hat{x}).$$

Equality holds if $c_2 = 0$, or if f_2 is continuously differentiable near $\hat{x}$.

(b) $\partial\Big(f_1 f_2\Big)(\hat{x}) \subseteq f_1(\hat{x}) \partial f_2(\hat{x}) + f_2(\hat{x}) \partial f_1(\hat{x})$.

(c) The minimum function $V(x) := \min\{f_i(x) \;:\; i = 1, \ldots, k\}$ is Lipschitz near $\hat{x}$ and

$$\partial V(\hat{x}) \subseteq \operatorname{co} \bigcup \{\partial f_i(\hat{x}) \;:\; i \in I(\hat{x})\},$$

where $I(\hat{x}) := \arg\min\{f_i(\hat{x}) \;:\; i = 1, \ldots, k\}$ is the set of indices at which the minimum defining V is attained.

(d) If f_1 takes on a local minimum value at $\hat{x}$, then $0 \in \partial f_1(\hat{x})$.

Part (c) of Prop. 2.2 is especially suggestive in the present context because it can be viewed as a very simple instance of our general problem, namely the differential analysis of a certain minimum-value function V as perturbations x displace the data of the associated minimization problem. The problem in (c) is admittedly elementary—one is simply required to choose the smallest of k scalars—but the issues noted in Section 1 are already perceptible. For instance, the function V may well fail to be smooth even when each f_i is linear, so classical differential analysis of V is impossible. However, an elegant and illuminating description of the local behaviour of V is available in terms of generalized gradients. At the risk of some oversimplification, the goal of this thesis may be summarized as an attempt to find calculus results analogous to (c) in which the minimum defining V is taken over more general sets than $\{1, 2, \ldots, k\}$.

Of course, Clarke's conception of the generalized gradient as a set-valued mapping instead of a point-valued one raises issues with no obvious precursors in classical analysis. For instance, one must consider the closure properties of the mapping $\partial f(\cdot)$.

2.3 Proposition. *Let x_i and ς_i be sequences in X and X^* such that $\varsigma_i \in \partial f(x_i)$ $\forall i$. Suppose that x_i converges to $\widehat{x}$ and that $\widehat{\varsigma}$ is a weak*-cluster point of $\{\varsigma_i\}$. Then $\widehat{\varsigma} \in \partial f(\widehat{x})$.*

In our discussion of Lagrange multipliers in Section 1, the geometrical significance of the classical derivative was prominent, as it led to line (1.3). Generalized gradients also have their geometrical side, which is best understood in terms of the "epigraph" of the function f. The *epigraph of* f is the set of points (x, r) in $X \times \mathbf{R}$ lying on or above the graph of f:

$$\operatorname{epi} f := \{(x, r) \in X \times \mathbf{R} \,:\, r \geq f(x)\}.$$

Let us first consider, for fixed $\widehat{x}$, the set

$$\operatorname{epi} f^0(\widehat{x}; \cdot) = \left\{(v, r) \in X \times \mathbf{R} \,:\, r \geq f^0(\widehat{x}; v)\right\}. \tag{2.3}$$

Since $f^0(\widehat{x}; \cdot)$ is sublinear, $\operatorname{epi} f^0(\widehat{x}; \cdot)$ is a closed convex cone with vertex at 0. This cone is the *(Clarke) tangent cone to* $\operatorname{epi} f$ *at* $(\widehat{x}, f(\widehat{x}))$. In the case of a smooth function $f : \mathbf{R}^n \to \mathbf{R}$, $f^0(\widehat{x}; v) = \langle Df(\widehat{x}), v\rangle$ is a linear function of v and its epigraph, which we are calling $T_{\operatorname{epi} f}(\widehat{x}, f(\widehat{x}))$, is the half-space lying above the hyperplane through $(\widehat{x}, f(\widehat{x}))$ and supporting the graph of f. Thus the Clarke tangent cone corresponds to our usual notion of tangency in the classical case. Moreover the *(Clarke) normal cone to* epi f *at* $(\widehat{x}, f(\widehat{x}))$, defined by

$$(2.4) \qquad N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x})) = \{(\varsigma, \beta) \in X^* \times \mathbf{R} \,:\, \langle(\varsigma, \beta), (v, r)\rangle \leq 0 \quad \forall (v, r) \in T_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))\}$$

is a closed convex cone with vertex at 0 obeying the rule

$$(2.5) \qquad \varsigma \in \partial f(\widehat{x}) \quad \Longleftrightarrow \quad (\varsigma, -1) \in N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x})).$$

This corresponds to the well-known classical result that if f is C^1 near $\widehat{x}$, then the vector $(Df(\widehat{x}), -1)$ is normal to the graph of f at $(\widehat{x}, f(\widehat{x}))$.

Non-Lipschitz Functions. The Clarke tangent cone can be defined in considerably more generality than the previous paragraph suggests. Indeed, generalized notions of tangency and normality are the foundations of a whole theory of "nonsmooth analysis", a structure whose keystone is the generalized gradient. These foundations are also basic features of this thesis, and we have devoted all of Chapter II to their study. Let us therefore be content with a sketch of their significance.

For any closed set $C \subseteq X$ and any point c in C, it is possible to define the *tangent cone to* C *at* c, denoted $T_C(c)$, in a way consistent with that cited above when $C = \mathrm{epi}\, f$ for f Lipschitz. (This is done in Definition II.3.1 below.) The *normal cone to* C *at* c is then given by (compare line (2.4))

$$N_C(c) = \{\varsigma \in X^* \,:\, \langle \varsigma, v\rangle \leq 0 \quad \forall v \in T_C(c)\}.$$

By analogy with line (2.5), it is now possible to define a set $\partial f(\widehat{x})$ via

$$(2.6) \qquad \partial f(\widehat{x}) = \{\varsigma \in X^* \,:\, (\varsigma, -1) \in N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))\}.$$

This set may also be called the generalized gradient of f at $\widehat{x}$ since it agrees with the previous definition when f is Lipschitz near $\widehat{x}$. However, the definition in line (2.6) makes sense for any function f for which epi f is locally closed near $(\widehat{x}, f(\widehat{x}))$. In particular, the generalized gradient is thus defined for any lower semicontinuous function $f\colon X \to \mathbf{R} \cup \{+\infty\}$ and point $\widehat{x}$ where $f(\widehat{x}) < +\infty$. However, in this broad class of functions the generalized gradient is known only to be a weak*-closed set (perhaps not weak*-compact); furthermore, it may be empty. To obtain $\partial f(\widehat{x}) = \emptyset$, it is necessary that each vector in $N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))$ have zero in its second component. The local information about f near $\widehat{x}$ is then concentrated in the *asymptotic generalized gradient of f at $\widehat{x}$*, namely

$$\partial^\infty f(\widehat{x}) = \{\varsigma : (\varsigma, 0) \in N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))\}. \tag{2.7}$$

The asymptotic generalized gradient is always a closed convex cone containing 0, which can be intuitively understood as the set of directions in which f is particularly badly-behaved (i.e., nonLipschitz). Together, $\partial f(\widehat{x})$ and $\partial^\infty f(\widehat{x})$ contain all the local information inherent in $N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))$:

$$N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x})) = \{\lambda(\varsigma, -1) : \lambda > 0,\ \varsigma \in \partial f(\widehat{x})\} \cup \{(\varsigma, 0) : \varsigma \in \partial^\infty f(\widehat{x})\}. \tag{2.8}$$

The Finite-Dimensional Case. When $X = \mathbf{R}^n$, the following desirable results become available.

2.4 Proposition. *Let $f: \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and a point $\widehat{x}$ where $f(\widehat{x}) < +\infty$ be given. If* epi f *is locally closed near $(\widehat{x}, f(\widehat{x}))$, then one has $N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x})) \neq \{0\}$. In particular,*

$$\partial f(\widehat{x}) \cup \left(\partial^\infty f(\widehat{x}) \setminus \{0\}\right) \neq \emptyset.$$

2.5 Proposition. *Let $f: \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $\widehat{x} \in \mathbf{R}^n$ be given. If $f(\widehat{x}) < +\infty$ and* epi f *is locally closed near $(\widehat{x}, f(\widehat{x}))$, then the following are equivalent:*

(a) $\partial f(\widehat{x})$ is a bounded nonempty set;

(b) $\partial^\infty f(\widehat{x}) = \{0\}$;

(c) f is Lipschitz near $\widehat{x}$.

An invaluable tool in actually computing $\partial f(\widehat{x})$ for a function f satisfying the conditions of Prop. 2.5 is the "proximal normal formula", which indicates how to calculate $N_{\mathrm{epi}\, f}(\widehat{x}, f(\widehat{x}))$. For any closed set $C \subseteq \mathbf{R}^n$ and point $\widehat{c} \in C$, we say that a vector $v \in \mathbf{R}^n$ is *perpendicular to C at $\widehat{c}$* and write $v \perp C$ at $\widehat{c}$, if the closed ball $(\widehat{c} + v) + |v|\overline{B}$ meets C in the single point $\widehat{c}$. (Here B denotes the open unit ball of $\mathbf{R}^n$.) In terms of an inequality, $v \perp C$ at $\widehat{c}$ means

$$\langle v, c - \widehat{c} \rangle < \frac{1}{2}|c - \widehat{c}|^2 \qquad \forall c \in C \setminus \{\widehat{c}\}. \tag{2.9}$$

The proximal normal formula states that

$$N_C(\hat{c}) = \overline{\mathrm{co}} \bigcup_{\lambda \geq 0} \lambda \left\{ \lim_{i\to\infty} \frac{v_i}{|v_i|} \ : \ v_i \perp C \text{ at } c_i \to \hat{c},\ v_i \to 0 \right\}. \tag{2.10}$$

Chapter II proves a version of this formula valid in any Hilbert space, and provides more details concerning the general theory of normal and tangent cones. Before beginning that study, however, let us reconsider the problem treated informally in Section 1.

Section 3. Lagrange Multipliers Revisited

In Section 1 we suggested that differential analysis of a certain value function was the key to both the proof of a Lagrange multiplier rule and the interpretation of the multipliers, and moreover that Clarke's nonsmooth analysis allowed such progress despite the inapplicability of classical calculus. This section justifies that claim, relying to a certain extent on Clarke (1983), Section 6.5, and on Rockafellar (1982). We consider locally Lipschitz functions $\ell\colon \mathbf{R}^n \to \mathbf{R}$, $g\colon \mathbf{R}^n \to \mathbf{R}^a$, and define $V\colon \mathbf{R}^a \to \mathbf{R} \cup \{+\infty\}$ by

$$V(\alpha) := \min\{\ell(x) \ : \ x \in \mathbf{R}^n,\ g(x) + \alpha = 0\}.$$

The minimization problem defining $V(\alpha)$ is called $P(\alpha)$: we say that x *solves* $P(\alpha)$ if $g(x) + \alpha = 0$ and $\ell(x) = V(\alpha)$. Let us make the following hypothesis throughout this section:

(H) $V(0) < +\infty$, and for some $\varepsilon > 0$ the level set $K = \{x \in \mathbf{R}^n \ : \ \ell(x) \leq V(0) + \varepsilon\}$ is compact.

Hypothesis (H) guarantees that a satisfactory existence theory can be developed for this family of problems, and is crucial to the convergence arguments of Prop. 3.3 below.

3.1 Lemma. *(a) For any α such that $V(\alpha) < V(0) + \varepsilon$, problem $P(\alpha)$ has a solution lying in* $\operatorname{int} K$.

(b) The set $\operatorname{epi} V$ *is locally closed near* $(0, V(0))$.

Proof. (a) If $V(\alpha) < V(0) + \varepsilon$, then no generality is lost in restricting the set of admissible values of x to a compact subset K' of $\operatorname{int} K = \{x \in \mathbf{R}^n \ : \ \ell(x) < V(0) + \varepsilon\}$. Let x_i be a minimizing sequence in K'. Then by passing to a subsequence if necessary, one has $x_i \to x$ for some $x \in K'$, and also

$$g(x_i) + \alpha = 0, \qquad \ell(x_i) \leq V(\alpha) + \frac{1}{i} \qquad \forall i.$$

In the limit as $i \to \infty$ we find that $g(x) + \alpha = 0$ and $\ell(x) \leq V(\alpha)$. Hence $x \in K'$ solves $P(\alpha)$.

(b) Let (α_i, v_i) be a sequence in $\operatorname{epi} V$ converging to a point (α, v) with $v < V(0) + \varepsilon$. Then without loss of generality we may assume $v_i < V(0) + \varepsilon \ \forall i$. By part (a), problem $P(\alpha_i)$ has a solution x_i in K' for which

$$g(x_i) + \alpha_i = 0, \qquad V(\alpha_i) = \ell(x_i) \leq v_i \qquad \forall i.$$

Along a subsequence, $x_i \to x$ for some $x \in K'$. Taking limits above then gives $g(x) + \alpha = 0$ and $\ell(x) \leq v$. In particular, $V(\alpha) \leq v$ implies $(\alpha, v) \in \operatorname{epi} V$, as required. ////

Lemma 3.1 assures that $N_{\operatorname{epi} V}(0, V(0))$ is well-defined. To evaluate this set using the proximal normal formula (2.10), let us first analyse the properties of a single perpendicular.

3.2 Proposition. *If $(\beta, -\lambda) \perp \operatorname{epi} V$ at (α, v) for some $(\alpha, v) \in (0, V(0)) + \varepsilon B$, then $P(\alpha)$ has a solution x lying in K for which*

$$0 \in \partial\Big[\widetilde{\lambda}\ell(\cdot) + \Big\langle \widetilde{\beta}, g(\cdot)\Big\rangle\Big](x), \quad \textit{where } (\widetilde{\beta}, -\widetilde{\lambda}) = \frac{(\beta, -\lambda)}{|(\beta, -\lambda)|}.$$

Proof. If $(\beta, -\lambda) \perp \operatorname{epi} V$ at (α, v), then inequality (2.9) asserts that

$$(*) \qquad \langle (\beta, -\lambda), (\alpha', v') - (\alpha, v)\rangle \leq \frac{1}{2}\left|(\alpha', v') - (\alpha, v)\right|^2 \qquad \forall (\alpha', v') \in \operatorname{epi} V.$$

Now since $\alpha \in \varepsilon B$ and $V(\alpha) \leq v < V(0) + \varepsilon$, problem $P(\alpha)$ has a solution x lying in K. Moreover, for any $x' \in \mathbf{R}^n$, we have

$$\begin{aligned} & V(-g(x')) \leq \ell(x') \leq \ell(x') + v - \ell(x) \\ \Rightarrow \qquad & (-g(x'), \ell(x') + v - \ell(x)) \in \operatorname{epi} V. \end{aligned}$$

Hence $(*)$ implies that for all $x' \in \mathbf{R}^n$, one has

$$\langle \beta, g(x)\rangle + \lambda\ell(x) \leq \langle \beta, g(x')\rangle + \lambda\ell(x') + \frac{1}{2}\left|g(x') - g(x)\right|^2 + \frac{1}{2}\left|\ell(x') - \ell(x)\right|^2.$$

That is, the locally Lipschitz function of x' on the right side has a global minimum at $x' = x$. By Prop. 2.2(a)(d), it follows that

$$0 \in \partial\Big[\lambda\ell(\cdot) + \langle \beta, g(\cdot)\rangle\Big](x).$$

Dividing both sides by $|(\beta, -\lambda)| \neq 0$ gives the result. ////

The next step is to consider a limit of normalized perpendiculars.

3.3 Proposition. *If* $(\widetilde{\beta}, -\widetilde{\lambda}) = \lim_{i\to\infty} (\widetilde{\beta}_i, -\widetilde{\lambda}_i)$ *for a sequence of perpendiculars* $(\beta_i, -\lambda_i) \perp \operatorname{epi} V$ *at* $(\alpha_i, v_i) \to (0, V(0))$, *then* $P(0)$ *has a solution* x *for which*

$$0 \in \partial\Big[\widetilde{\lambda}\ell(\cdot) + \Big\langle \widetilde{\beta}, g(\cdot)\Big\rangle\Big](x).$$

Proof. Write

$$(\dagger) \qquad \widetilde{\lambda}_i\ell(\cdot) + \Big\langle \widetilde{\beta}_i, g(\cdot)\Big\rangle = \widetilde{\lambda}\ell(\cdot) + \Big\langle \widetilde{\beta}, g(\cdot)\Big\rangle + (\widetilde{\lambda}_i - \widetilde{\lambda})\ell(\cdot) + \Big\langle \widetilde{\beta}_i - \widetilde{\beta}, g(\cdot)\Big\rangle.$$

For each i one has $x_i \in K$ such that

$$\begin{aligned} 0 &\in \partial\Big[\widetilde{\lambda}_i\ell(\cdot) + \Big\langle \widetilde{\beta}_i, g(\cdot)\Big\rangle\Big](x_i) \\ &\subseteq \partial\Big[\widetilde{\lambda}(\cdot) + \Big\langle \widetilde{\beta}, g(\cdot)\Big\rangle\Big](x_i) + \varepsilon_i\overline{B}, \end{aligned}$$

where ε_i is the Lipschitz rank of the last two terms in $(\dagger)$ on K. Note that $\varepsilon_i \to 0$ as $i \to \infty$. Along a subsequence, $x_i \to x$ in K for some x solving $P(0)$. (This follows from Lemma 3.1.) Then the closure property of the generalized gradient (Prop. 2.3) gives

$$0 \in \partial\Big[\widetilde{\lambda}\ell(\cdot) + \Big\langle \widetilde{\beta}, g(\cdot)\Big\rangle\Big](x),$$

as required. ////

Now let Y denote the set of solutions to $P(0)$. The proximal normal analysis above leads to the following Lagrange multiplier rule.

3.4 Theorem (Lagrange Multipliers). *If* $x \in Y$, *then there exists* $(\beta, -\lambda) \in \mathbf{R}^a \times \mathbf{R}$ *such that*

$$\lambda \in \{0, 1\}, \qquad 0 \in \partial\Big(\lambda\ell(\cdot) + \langle \beta, g(\cdot)\rangle\Big)(x), \qquad \lambda + |\beta| > 0.$$

Proof. Assume first that $Y = \{x\}$. By Prop. 2.4, $N_{\operatorname{epi} V}(0, V(0)) \neq \{0\}$. Hence the proximal normal formula (2.10) asserts that some sequence of normalized perpendiculars has a nonzero limit. This limit must obey the conclusions of Prop. 3.3. If $\widetilde{\lambda} = 0$ in those conclusions, the statement of our theorem holds with $(\beta, -\lambda) = (\widetilde{\beta}, 0)$. If $\widetilde{\lambda} > 0$, then our statement holds with $(\beta, -\lambda) = (\widetilde{\beta}/\widetilde{\lambda}, -1)$.

If the set Y contains more than one point, fix any $\widehat{x} \in Y$ and repeat the development above with ℓ replaced by $\widehat{\ell}(x) = \ell(x) + |x - \widehat{x}|^2$. Hypothesis (H) remains valid for this new problem, which has the unique solution $\hat{x}$. It follows from the previous paragraph that there exists $(\beta, -\lambda) \in \mathbf{R}^a \times \mathbf{R}$ such that

$$\lambda \in \{0,1\}, \qquad 0 \in \partial\Big(\lambda\ell(\cdot) + \lambda\,|(\cdot) - \hat{x}|^2 + \langle\beta, g(\cdot)\rangle\Big)(\hat{x}), \qquad \lambda + |\beta| > 0.$$

Since the second term in the sum whose generalized gradient is computed above is smooth, with derivative 0 at $\hat{x}$, the desired conclusion follows from Prop. 2.2(a). ////

The development leading to Thm. 3.4 shows how proximal normal analysis permits the derivation of a Lagrange multiplier rule. More specifically, it allows necessary conditions for a constrained problem to be derived from known necessary conditions for an unconstrained problem. (See the proof of Prop. 3.2.) This proof of the Lagrange multiplier rule is more elementary than that given by Clarke (1983), Thm. 6.1.1, p. 228, which relies on Ekeland's theorem. But in addition to its debatable aesthetic advantages, Thm. 3.4 has two unquestionable merits. First, if the functions ℓ and g are smooth, the auxiliary problem leading to Prop. 3.2 is also smooth. In fact, the whole proof of the Lagrange multiplier rule then requires no nonsmooth analysis at all, except for the definition of $N_{\mathrm{epi}\,V}(0, V(0))$ which suggests the appropriate auxiliary problem. This is in contrast with the proof based on Ekeland's theorem, in which nonsmooth terms enter the auxiliary problems even for smooth data. (Another general technique for proving necessary conditions, called "exact penalization", also introduces nonsmooth auxiliary elements into problems whose data are smooth. Chapter III below shows that by avoiding this, proximal normal analysis allows a significant improvement in the necessary conditions for optimal control problems with smooth state constraints.) The second main advantage of the current proof is that it adds rigor to the traditional interpretation of Lagrange multipliers as marginal values. This is the content of Thm. 3.6 below, which can be viewed as a precise version of the results suggested in Section 1. (Theorem 3.6 is due to Rockafellar (1982); the development in this section follows Clarke (1983), Section 6.5.)

3.5 Definition. Let $x \in \mathbf{R}^n$ and $\lambda \geq 0$. A vector $\beta \in \mathbf{R}^a$ is an *index* λ *multiplier corresponding to* x if

$$0 \in \partial\Big[\lambda\ell(\cdot) + \langle\beta, g(\cdot)\rangle\Big](x).$$

The set of all such vectors is denoted $M^\lambda(x)$; we also write

$$M^\lambda(Y) = \bigcup_{x \in Y} M^\lambda(x).$$

3.6 Theorem. *Assume (H). Then $Y \neq \emptyset$, and one has*

$$\partial V(0) = \overline{\mathrm{co}}\Big[M^1(Y) \cap \partial V(0) + M^0(Y) \cap \partial^\infty V(0)\Big].$$

If $M^0(Y) = \{0\}$, then $\partial^\infty V(0) = \{0\}$ and the previous equation becomes

$$\partial V(0) = \mathrm{co}\Big[M^1(Y) \cap \partial V(0)\Big].$$

Proof. To prove the first statement, we will apply Prop. II.6.2 with

$$D = M^1(Y) \cap \partial V(0) \quad \text{and} \quad D^\infty = M^0(Y) \cap \partial^\infty V(0).$$

It therefore suffices to show that $N_{\mathrm{epi}\,V}(0, V(0)) = \overline{\mathrm{co}}[N \cup N^\infty]$, where

$$\begin{aligned} N &= \{\lambda(\varsigma, -1) \,:\, \varsigma \in M^1(Y) \cap \partial V(0),\ \lambda \geq 0\} \\ N^\infty &= \{\ (\varsigma, 0) \,:\, \varsigma \in M^0(Y) \cap \partial^\infty V(0)\}. \end{aligned}$$

The inclusion $N_{\mathrm{epi}\,V}(0, V(0)) \supseteq \overline{\mathrm{co}}[N \cup N^\infty]$ is an obvious consequence of the definitions of $\partial V(0)$ and $\partial^\infty V(0)$ (lines (2.6) and (2.7) above). To see the reverse inclusion, note that by the proximal normal formula (2.10), $N_{\mathrm{epi}\,V}(0, V(0))$ is the closed convex cone generated by certain limits of perpendiculars. If $(\widetilde{\beta}, -\widetilde{\lambda})$ is such a limit with $\widetilde{\lambda} = 0$, then $\widetilde{\beta} \in M^0(Y)$ by Prop. 3.3, while $(\widetilde{\beta}, 0) \in N_{\mathrm{epi}\,V}(0, V(0))$ by construction implies $\widetilde{\beta} \in \partial^\infty V(0)$. Hence $(\widetilde{\beta}, 0) \in N^\infty$. And if $(\widetilde{\beta}, -\widetilde{\lambda})$ is such a limit with $\widetilde{\lambda} > 0$, then $\beta = \widetilde{\beta}/\widetilde{\lambda} \in M^1(Y)$ by Prop. 3.3, while $(\beta, -1) \in N_{\mathrm{epi}\,V}(0, V(0))$ by construction implies $\beta \in \partial^\infty V(0)$. Hence $\widetilde{\lambda}(\beta, -1) \in N$. Combining these two possibilities gives

$$N_{\mathrm{epi}\,V}(0, V(0)) \subseteq \overline{\mathrm{co}}[N \cup N^\infty],$$

as required.

To prove the second statement we note that if $M^0(Y) = \{0\}$ then the cone $D^\infty = \{0\}$ is certainly pointed. The result would then follow from Prop. II.6.5 if we could prove $D^\infty \supseteq 0^+D$. (This notation is introduced in Lemma II.6.3.) It is quite easy to show that $M^0(Y) \supseteq 0^+M^1(Y)$, while $\partial^\infty V(0) \supseteq 0^+\partial V(0)$ for any function V is a well-known result of nonsmooth analysis. Combining these two statements justifies the application of Prop. II.6.5. ////

Corollary 1. *If $M^0(Y) = \{0\}$ and $M^1(Y)$ is the singleton $\{\beta\}$, then V is strictly differentiable at 0 and $D_s V(0) = \beta$.*

Proof. By Thm. 3.6, $M^0(Y) = \{0\}$ implies $\partial^\infty V(0) = \{0\}$. Consequently V is Lipschitz near 0 by Prop. 2.5. In fact, Thm. 3.6 implies that $\partial V(0) = \{\beta\}$, so the result follows from Prop. 2.1. ////

Among the interesting byproducts of the first-order information provided by Thm. 3.6 is the following sufficient condition for local surjectivity of the mapping $g: \mathbf{R}^n \to \mathbf{R}^a$. In the smooth case, it reduces to the well-known Surjective Mapping Theorem, which states that if the $a \times n$ matrix $Dg(x)$ has rank a, then g is locally surjective near x.

Corollary 2. *Suppose $M^0(x) = \{0\}$ for some $x \in \mathbf{R}^n$. Then there exist positive constants η and M such that for all $\alpha \in \eta B$, there exists some $y \in \mathbf{R}^n$ obeying $g(y) + \alpha = 0$ and $|y - x| \leq M\,|\alpha|$.*

Proof. Define $\ell(y) = |y - x|$ and consider the value function

$$V(\alpha) := \min\{\ell(y) \;:\; g(y) + \alpha = 0\}.$$

The unique solution to $P(0)$ is x, and $M^0(x) = \{0\}$ implies $\partial^\infty V(0) = \{0\}$ by Thm. 3.6. According to Prop. 2.5, V is Lipschitz near 0. That is, there exists $M > 0$ and $\eta \in \big(0, \varepsilon/M\big)$ such that

$$V(\alpha) \leq V(0) + M\,|\alpha| \qquad \forall \alpha \in \eta B.$$

Noting that $V(0) = 0$ and applying Lemma 3.1(a) gives the desired result. ////

A result like Corollary 2 is of interest in optimization theory because it says something about the stability of the set of feasible points under perturbations of the data. Clarke (1983), Section 6.6 proves a more general result involving inequality constraints and an abstract constraint as well as equality constraints.

Section 4. Parameter Sensitivity in Optimal Control

The previous section showed how proximal normal analysis allows both the derivation of necessary conditions and a sensitivity analysis for a simple problem. In this section we review some results of Loewen (1983) which show that even when necessary conditions are known in advance, proximal normal analysis can be used to interpret them and perform an independent analysis of parameter sensitivity. The first-order dependence of a problem's minimum value on various parameters is of considerable interest in its own right. It finds theoretical application in dynamic programming and the Hamilton-Jacobi equation, for example, and also has consequences for such practical issues as controllability. There is also its obvious utility in identifying which parameters have the greatest effect on the problem's value. In linear programming, the latter application is sufficient to explain "why the dual vector is cherished by oil company vice-presidents and not just by mathematicians." (Franklin (1980).)

The Problem. In this section we study perturbations of the following *differential inclusion problem*:

$$(P) \qquad \min\left\{\ell\big(x(0),x(T)\big) \;:\; \dot{x}(t)\in F(t,x(t)) \text{ a.e. } [0,T],\ \big(x(0),x(T)\big)\in S\right\}.$$

The objective of problem (P) is to choose an arc (i.e., an absolutely continuous function) $x\colon[0,T]\to\mathbf{R}^n$ satisfying the dynamic constraint

$$(4.1) \qquad \dot{x}(t)\in F(t,x(t)) \text{ a.e. } [0,T]$$

and the endpoint constraint $\big(x(0),x(T)\big)\in S$ while minimizing ℓ over all such arcs. Line (4.1) is a *differential inclusion* phrased in terms of a set-valued mapping, or *multifunction*, $F\colon[0,T]\times\mathbf{R}^n\to\mathbf{R}^n$. An arc x obeying (4.1) is called an F-trajectory; if it also obeys $\big(x(0),x(T)\big)\in S$, then x is called an *admissible* F-trajectory. A special case of problem (P) is the Mayer problem arising when one is given a control set $U\subseteq\mathbf{R}^m$ and a function $f\colon[0,T]\times\mathbf{R}^n\times\mathbf{R}^m\to\mathbf{R}^n$ and instructed to solve (P) with the modified dynamics

$$(4.2) \qquad \dot{x}(t)=f(t,x(t),u(t)) \text{ a.e. } [0,T] \text{ for some measurable } u\colon[0,T]\to U.$$

Filippov's lemma asserts that if we take $F(t,x) := f(t,x,U)$, then any F-trajectory actually obeys (4.2). The converse is clear, so the admissible arcs for the Mayer problem built around (4.2) are the same as those for (P).

Under the following hypotheses, necessary conditions for problem (P) are known. Throughout this section, B denotes the open unit ball in $\mathbf{R}^n$.

(h1) The multifunction $F:[0,T] \times \mathbf{R}^n \to \mathbf{R}^n$ has nonempty compact convex values. For each fixed $x \in \mathbf{R}^n$, $F(\cdot, x)$ is measurable on $[0,T]$. That is, for any closed set $C \subseteq \mathbf{R}^n$, the following "inverse image" is a Lebesgue measurable set: $\{t \in [0,T] \,:\, F(t,x) \cap C \neq \emptyset\}$.

(h2) There is a function $k(t) \in L^1[0,T]$ such that

(a) $F(t,x) \subseteq k(t)\overline{B} \quad \forall t \in [0,T],\ x \in \mathbf{R}^n$,

(b) for each fixed $t \in [0,T]$ and $x \in \mathbf{R}^n$, one has

$$F(t,y) \subseteq F(t,x) + k(t)\,|y-x|\,\overline{B} \qquad \forall y \in \mathbf{R}^n.$$

We define $K_F = \exp\left(\int_0^T k(t)\,dt\right)$.

(h3) The constraint set $S \subseteq \mathbf{R}^n \times \mathbf{R}^n$ is closed, and $\{x \,:\, (x,y) \in S\}$ is compact.

(h4) The objective function $\ell: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is Lipschitz of rank K_ℓ on $\mathbf{R}^n \times \mathbf{R}^n$.

Note that the convexity hypothesis in (h1) means that we are working with the "relaxed problem" in the sense of Warga (1972).

The necessary conditions are phrased in terms of the *Hamiltonian* $H:[0,T] \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ and the *distance function* $d_S: \mathbf{R}^{2n} \to \mathbf{R}$ defined by

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

$$d_S(v) := \inf\{|v-s| \,:\, s \in S\}$$

4.1 Theorem. *Assume (h1)–(h4), and fix any $r > (2K_\ell + 2)(1 + K_F \ln K_F)$. If the arc $x(\cdot)$ solves (P), then there exist a scalar μ and an arc $p:[0,T] \to \mathbf{R}^n$, not both zero, such that*

(a) $\mu \in \{0,1\}$,

(b) $(-\dot{p}(t), \dot{x}(t)) \in \partial H(t, x(t), p(t)) \quad$ a.e. $[0,T]$,

(c) $\big(p(0), -p(T)\big) \in \mu\varsigma + r\,|(\mu, E)|\,\partial d_S(x(0), x(T))$ *for some* $\varsigma \in \partial\ell(x(0), x(T))$.

Here ∂H signifies the generalized gradient of H in the (x,p) variable only, and $E := \mu\varsigma + (-p(0), p(T))$.

Proof. This follows easily from Clarke (1983), Thm. 3.5.2, p. 147. ////

A full discussion of the relationship between Thm. 4.1 and Pontryagin's maximum principle is given by Clarke (1983).

Note that since $N_S(s) = \overline{\bigcup_{r \geq 0} r\partial d_S(s)}$ (Clarke (1983), Prop. 2.4.2, p. 51), condition (c) implies the simpler form

(c′) $\left(p(0), -p(T)\right) \in \mu\varsigma + N_S(x(0), x(T))$ for some $\varsigma \in \partial\ell(x(0), x(T))$.

Conclusion (c′) is only slightly weaker than (c) for large values of r, and has the advantage of containing no explicit r-dependence.

Hypotheses (h1)–(h4) specifying (P) restrict the application of Thm. 4.1 to global solutions $x(\cdot)$. This is not an essential restriction. In fact, for any $\delta > 0$, the result remains valid whenever the arc $x(\cdot)$ only solves (P) relative to all feasible arcs $y(\cdot)$ obeying the relationship $\|y - x\|_\infty < \delta$. This is equivalent to the requirement that the graph of y be contained in the set

$$T_\delta(x) := \{(t, y) \in [0, T] \times \mathbf{R}^n \ : \ |y - x(t)| < \delta\},$$

called the *tube of radius δ about x*. The key point here is that for the purposes of Thm. 4.1, hypotheses (h1)–(h4) need only hold within $T_\delta(x)$. This observation is important in Chapter III, where we will use it to reduce the size of the constants K_F and K_ℓ once a solution is known.

Perturbations. Let us now consider the effect of a finite-dimensional parameter α in $\mathbf{R}^a$ on problem (P). We define a family of problems $P(\alpha)$ by

$$P(\alpha) \qquad \min\left\{\ell(x(0), x(T), \alpha) \ : \ \dot{x}(t) \in F(t, x(t), \alpha) \ \text{ a.e. } [0, T], \ \left(x(0), x(T), \alpha\right) \in S\right\}.$$

The *value function* $V : \mathbf{R}^a \to \mathbf{R} \cup \{+\infty\}$ is then given by $V(\alpha) := \inf P(\alpha)$. Here we intend to explain (omitting most proofs, which are given in Loewen (1983)) how proximal normal analysis leads to a valuable formula for $\partial V(0)$. First, however, we must place some mild restrictions on the α-dependence of the data defining $P(\alpha)$. These hypotheses also involve the set Y, which consists of all arcs solving the nominal problem $P(0)$.

(H1) The multifunction $F : [0, T] \times \mathbf{R}^n \times \mathbf{R}^a \to \mathbf{R}^n$ has nonempty compact convex values. For each fixed $(x, \alpha) \in \mathbf{R}^n \times \mathbf{R}^a$, $F(\cdot, x, \alpha)$ is measurable on $[0, T]$.

(H2) There is a function $k(t) \in L^1[0, T]$ such that

(a) $F(t,x,\alpha) \subseteq k(t)\overline{B} \quad \forall (t,x,\alpha) \in [0,T] \times \mathbf{R}^n \times \mathbf{R}^a$,

(b) for each fixed $t \in [0,T]$ and $(x,\alpha) \in \mathbf{R}^n \times \mathbf{R}^a$, one has

$$F(t,y,\beta) \subseteq F(t,x,\alpha) + k(t)\,|(y,\beta) - (x,\alpha)|\,\overline{B} \qquad \forall (y,\beta) \in \mathbf{R}^n \times \mathbf{R}^a.$$

(H3) The set $S \subseteq \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^a$ is closed and $\{x : (x,y,\alpha) \in S\}$ is compact. Moreover, for any $x(\cdot) \in Y$, the multifunction $N_S(\cdot)$ is closed at $(x(0), x(T), 0)$.

(H4) The objective function $\ell\colon \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^a \to \mathbf{R}$ is globally Lipschitz of rank K_ℓ.

The closure hypothesis of (H3) requires that if (s_i, v_i) is a sequence converging to (s,v) and obeying $v_i \in N_S(s_i)\ \forall i$, then one has $v \in N_S(s)$ at least whenever the limit point s has the form $(x(0), x(T), 0)$ for some $x \in Y$. This mild assumption holds automatically if the cone $N_S(x(0), x(T), 0)$ is pointed. (A set is *pointed* if zero cannot be obtained as a positive linear combination of its nonzero elements.) A detailed discussion of other conditions ensuring the validity of (H3) is given by Loewen (1983).

4.2 Lemma. *(a) If $V(\alpha) < +\infty$ then problem $P(\alpha)$ has a solution.*

(b) The function V is lower semicontinuous near 0.

Proof. See Loewen (1983). ////

Just as in Section 3, consideration of a single perpendicular vector leads to a certain auxiliary problem which can be solved by known methods. Since the auxiliary problem depends on α, we must introduce a more general Hamiltonian via

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

4.3 Proposition. *Let $(\beta, -\lambda) \perp \operatorname{epi} V$ at $(\hat\alpha, \hat v)$. Then $P(\hat\alpha)$ has a solution $\hat x(\cdot)$ to which there corresponds an arc $(p,q)\colon [0,T] \to \mathbf{R}^n \times \mathbf{R}^a$ and a constant μ such that*

(a) $\lambda \geq 0, \quad \mu \in \{0,1\}, \quad \mu + \|(p,q)\| > 0$,

(b) $(-\dot p(t), -\dot q(t), \dot{\hat x}(t)) \in \partial\mathcal{H}(t, \hat x(t), \hat\alpha, p(t))$ a.e. $[0,T]$,

(c) $(p(0), -p(T), -q(T)) \in \lambda\mu\varsigma + N_S(\hat x(0), \hat x(T), \hat\alpha)$ *for some* $\varsigma \in \partial\ell(\hat x(0), \hat x(T), \hat\alpha)$.

(d) $-q(0) = \mu\beta$.

Proof. Since $V(\widehat{\alpha}) \leq \widehat{v} < +\infty$, Lemma 4.2(a) ensures that $P(\widehat{\alpha})$ has a solution $\widehat{x}$. For any $\alpha \in \mathbf{R}^a$ and any F-trajectory $x(\cdot)$ admissible for $P(\alpha)$, one has

$$V(\alpha) \leq \ell(x(0), x(T), \alpha) \leq \ell(x(0), x(T), \alpha) + \widehat{v} - \ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha}),$$

so $\big(\alpha, \ell(x(0), x(T), \alpha) - \ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha}) + \widehat{v}\big)$ lies in epi V. From inequality (2.9), it follows that

$$\lambda\ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha}) - \langle\beta, \widehat{\alpha}\rangle \leq \lambda\ell(x(0), x(T), \alpha) - \langle\beta, \alpha\rangle + \frac{1}{2}\left|\big(\alpha - \widehat{\alpha},\ \ell(x(0), x(T), \alpha) - \ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha})\big)\right|^2$$

Since equality holds when $x = \widehat{x}$ and $\alpha = \widehat{\alpha}$, we find that the arc $\big(\widehat{x}(\cdot), \widehat{\alpha}\big)$ provides a global solution for an *unperturbed* differential inclusion problem whose data are

$$\widetilde{\ell}(x, x_1, y, y_1) = \lambda\ell(x, y, y_1) - \langle\beta, x_1\rangle + \frac{1}{2}|y_1 - \widehat{\alpha}|^2 + \frac{1}{2}|\ell(x, y, y_1) - \ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha})|^2,$$

$$\widetilde{F}(t, x, x_1) = F(t, x, x_1) \times \{0\},$$

$$\widetilde{S} = \{(x, x_1, y, y_1) \ :\ (x, y, y_1) \in S,\ x_1 \in \mathbf{R}^a\}.$$

The Hamiltonian for this unperturbed problem is

$$\begin{aligned}\widetilde{H}(t, x, x_1, p, q) &= \sup\{\langle(p, q), (v, 0)\rangle \ :\ (v, 0) \in \widetilde{F}(t, x, x_1)\} \\ &= \mathcal{H}(t, x, x_1, p).\end{aligned}$$

Applying Thm. 4.1, we find that there is an arc $(p, q)\colon [0, T] \to \mathbf{R}^n \times \mathbf{R}^a$ and a scalar μ such that

(a) $\mu \in \{0, 1\}$, $\mu + \|(p, q)\| > 0$.

(b) $\big(-\dot{p}(t), -\dot{q}(t), \dot{\widehat{x}}(t), 0\big) \in \partial\widetilde{H}(t, \widehat{x}(t), \widehat{\alpha}, p(t), q(t))$ a.e. $[0, T]$. This implies

$\big(-\dot{p}(t), -\dot{q}(t), \dot{\widehat{x}}(t)\big) \in \partial\mathcal{H}(t, \widehat{x}(t), \widehat{\alpha}, p(t))$ a.e. $[0, T]$.

(c′) $\big(p(0), q(0), -p(T), -q(T)\big) \in \mu\left[\lambda(\xi, 0, \eta, \eta_1) - (0, \beta, 0, 0)\right] + N_{\widetilde{S}}(\widehat{x}(0), \widehat{\alpha}, \widehat{x}(T), \widehat{\alpha})$ for some $(\xi, \eta, \eta_1) \in \partial\ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha})$.

Condition (c′) reduces to the pair of conditions

$$\big(p(0), -p(T), -q(T)\big) \in \mu\lambda\varsigma + N_S(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha}) \text{ for some } \varsigma \in \partial\ell(\widehat{x}(0), \widehat{x}(T), \widehat{\alpha}),$$

$$q(0) = -\mu\beta.$$

This proves the lemma. ////

The next step motivated by the proximal normal formula (2.10) is to take limits of normalized perpendicular vectors. To obtain meaningful results, we require the following hypothesis.

(H5) $P(0)$ has no solution x for which some nonzero arc $(p,q)\colon [0,T] \to \mathbf{R}^n \times \mathbf{R}^a$ obeys

$$\big(-\dot{p}(t), -\dot{q}(t), \dot{x}(t)\big) \in \partial \mathcal{H}(t, x(t), 0, p(t)) \quad \text{a.e. } [0,T],$$

$$\big(p(0), -p(T), -q(T)\big) \in N_S(x(0), x(T), 0),$$

$$q(0) = 0.$$

If (H5) fails, we say that the perturbation structure of problem $P(\alpha)$ is *degenerate*. In the trivial case when none of the data defining $P(\alpha)$ have explicit α-dependence, so that $V(\alpha) \equiv V(0)$, then (H5) is equivalent to the assumption that $P(0)$ is normal. (See Clarke (1983) for a definition of normality.) But in general, the nondegeneracy condition is a weak hypothesis which does not require normality: Loewen (1983) presents several significant examples.

Suppose now that $V(0) < +\infty$ and that a sequence of perpendiculars $(\beta_i, -\lambda_i)$ to $\operatorname{epi} V$ at points $(\alpha_i, v_i) \to (0, V(0))$ is given, such that $\dfrac{(\beta_i, -\lambda_i)}{|(\beta_i, -\lambda_i)|}$ tends to the limit $(\beta, -\lambda)$. Then there exist solutions x_i to $P(\alpha_i)$ and corresponding arcs (p_i, q_i) and scalars μ_i as in Prop. 4.3. Hypothesis (H5) implies that $\mu_i = 1$ for all i sufficiently large. Since the conclusions of Prop. 4.3 are stable under limiting operations, we obtain the following result.

4.4 Proposition. *Let $(\beta, -\lambda)$ be the limit of normalized perpendiculars described above. Then there is a solution $x(\cdot)$ of $P(0)$ and an arc $(p,q)\colon [0,T] \to \mathbf{R}^n \times \mathbf{R}^a$ such that*

(a) $\lambda \geq 0, \quad |(\beta, -\lambda)| = 1,$

(b) $\big(-\dot{p}(t), -\dot{q}(t), \dot{x}(t)\big) \in \partial \mathcal{H}(t, x(t), 0, p(t))$ a.e. $[0,T]$,

(c) $\big(p(0), -p(T), -q(T)\big) \in \lambda\varsigma + N_S(x(0), x(T), 0)$ *for some* $\varsigma \in \partial \ell(x(0), x(T), 0)$,

(d) $-q(0) = \beta$.

Proof. See Loewen (1983), pp. 34–36. ////

The appropriate definition of the multiplier sets is now clear.

4.5 Definition. Let x solve $P(0)$. An arc $(p,q)\colon[0,T] \to \mathbf{R}^n \times \mathbf{R}^a$ is an *index λ multiplier corresponding to x* if it satisfies conditions (b) and (c) of Prop. 4.4. The set of all such arcs is denoted $M^\lambda(x)$, and we define

$$M^\lambda(Y) = \bigcup_{x\in Y} M^\lambda(x).$$

We also define a mapping Δ from the space of multipliers to $\mathbf{R}^a$ as follows:

$$\Delta(p,q) = -q(0).$$

The desired result describing the first-order dependence of V on α near 0 is the following analogue of Thm. 3.6.

4.6 Theorem. *Assume (H1)–(H5) and suppose $V(0) < +\infty$. Then $Y \neq \emptyset$, and one has*

$$\partial V(0) = \overline{\mathrm{co}}\Big(\Delta\big[M^1(Y)\big] \cap \partial V(0) + \Delta\big[M^0(Y)\big] \cap \partial^\infty V(0)\Big).$$

If $\Delta\big[M^0(Y)\big]$ is pointed, then the closure operation is superfluous and one also has

$$\partial^\infty V(0) = \mathrm{co}\,\Big(\Delta\big[M^0(Y)\big] \cap \partial^\infty V(0)\Big).$$

Proof. Similar to Thm. 3.6. ////

Although sensitivity analysis was our primary objective, a new multiplier rule for perturbed problems is a byproduct of this investigation.

Corollary 1. *If x solves $P(0)$ then it has an index λ multiplier (p,q) for some $\lambda \geq 0$, with $\lambda + |q(0)| > 0$.*

Proof. See Loewen (1983). ////

Other desirable results also emerge as corollaries. The fundamental issue of controllability, for example, can be addressed as follows. (Note that since the hypothesis below involves only index 0 multipliers, it is independent of the cost function ℓ.)

Corollary 2. *If* $\Delta[M^0(x)] = \{0\}$ *for some arc* x *admissible for* $P(0)$, *then there exists* $\varepsilon > 0$ *such that for every* $\alpha \in \varepsilon B$, *one has an arc* $y(\cdot)$ *obeying*

$$\dot{y}(t) \in F(t, y(t), \alpha) \quad \text{a.e. } [0, T], \qquad \big(y(0), y(T), \alpha\big) \in S.$$

See Loewen (1983), Chapter IV, for a more detailed controllability result which generalizes Clarke (1983), Thm. 3.5.3.

Extensions. Loewen (1983) also presents a sensitivity analysis for differential inclusion problems with free terminal time. These results, together with a detailed investigation of their consequences for controllability and the nonlinear time-optimal control problem, may be found in Clarke and Loewen (1984) and Clarke and Loewen (to appear).

The following chapters extend the sensitivity analysis of this section in two more significant ways. In Chapter III, infinite-dimensional perturbations are used, and a new form of the maximum principle for state-constrained problems ensues. In Chapter V we revert to finite-dimensional perturbations, but apply them to a stochastic optimal control problem.

Chapter II. A Proximal Normal Formula in Hilbert Space

Clarke's calculus of generalized gradients is distinguished by its mutually complementary geometric and analytic aspects, which extend those of ordinary calculus. In calculus, one considers a smooth function $f: X \to \mathbf{R}$ for which the gradient $\nabla f(x)$ is defined analytically in terms of limits. The resulting object has a geometric interpretation: the ray in direction $(\nabla f(x), -1)$ is the outward normal to the epigraph of f at the point $(x, f(x))$. (The epigraph of f is the set $\operatorname{epi} f := \{(x, r) \in X \times \mathbf{R} \, : \, r \geq f(x)\}$.) The pedagogical progression from analysis to geometry in the smooth case accurately reflects the sequence of steps used to solve many applied problems: one first computes a gradient, and then uses its geometric properties. In Clarke's calculus, a *generalized gradient* can be defined for any lower semicontinuous $f: X \to \mathbf{R} \cup \{+\infty\}$: it is a weak*-closed convex subset of X^*. The resulting object has a geometrical side which is perfectly analogous to that for smooth functions—if $N_{\operatorname{epi} f}(x, f(x))$ denotes the (Clarke) normal cone to the closed set $\operatorname{epi} f$ at the point $(x, f(x))$, then

$$\partial f(x) = \{\varsigma \in X^* \, : \, (\varsigma, -1) \in N_{\operatorname{epi} f}(x, f(x))\}. \tag{1}$$

The fundamental difference between Clarke's theory and the ordinary calculus becomes clear when one turns to applications. In the problems to be considered in subsequent chapters, the desired results are analytic statements which could be derived from a formula for $\partial f(x)$; the appropriate solution is based on *doing the geometry first,* i.e. computing $N_{\operatorname{epi} f}(x, f(x))$, *and then drawing the analytic conclusions* from (1). The geometric approach is valuable because, whereas the analytic definition of $\partial f(x)$ is difficult to apply, an elegant sequential characterization of $N_{\operatorname{epi} f}(x, f(x))$ is available. This result, called a *proximal normal formula,* is presented in Clarke (1983), Section 2.5.

Theorem. *Let $C \subseteq \mathbf{R}^n$ be a closed set containing a point c. Then*

$$N_C(c) = \overline{\text{co}}\{\lim_{i\to\infty} v_i : v_i \text{ is a bounded sequence of proximal}$$

$$\text{normals to } C \text{ at base points } c_i \to c \text{ in } C\}.$$

Until very recently, this formula was known only for finite-dimensional sets C, so the geometrical approach to $\partial f(x)$ outlined above could be applied only to functions whose domain was $\mathbf{R}^n$. Even in the finite-dimensional context, however, the proximal normal formula has made significant progress possible in the study of sensitivity, controllability, and time-optimality in optimal control. (See Section I.4.) This useful formula has recently been extended to reflexive Banach spaces X by Borwein and Strojwas (1985), under the hypotheses that the norm of X is Fréchet differentiable away from 0 and Kadec. This chapter shows that any Hilbert space satisfies these requirements and presents a simple proof of the Borwein-Strojwas theorem in this context. Chapter III presents a significant application of the theorem.

Section 1. Elementary Geometry of Hilbert Space

Throughout this section we consider a real Hilbert space H equipped with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$. The open unit ball of H is denoted by B, and its boundary by S.

1.1 Proposition. *Let $x, y \in H$. Then $\|x\| = 1$ and $\|x - y\| < \delta$ imply $\left\|\frac{x+y}{2}\right\|^2 > 1 - \delta$. (This shows that H is uniformly convex.)*

Proof. By the parallelogram identity, any $x, y \in H$ obey

$$\begin{aligned}
\|x+y\|^2 &= 2\|x\|^2 + 2\|y\|^2 - \|x-y\|^2 \\
&\geq 2\|x\|^2 + 2(\|x-y\| - \|x\|)^2 - \|x-y\|^2 \\
&= 4\|x\|^2 - 4\|x\|\,\|x-y\| + \|x-y\|^2 \\
&\geq 4\|x\|^2 - 4\|x\|\,\|x-y\|.
\end{aligned}$$

Hence if $\|x\| = 1$ and $\|x - y\| < \delta$ we obtain $\left\|\frac{x+y}{2}\right\|^2 > 1 - \delta$, as required. ////

1.2 Proposition. *Let x, y be any unit vectors in H. Then for any $\delta \geq 0$, $\langle x, y \rangle \geq 1 - \delta$ implies $\|x - y\| \leq \sqrt{2\delta}$. In particular, $\langle x, y \rangle = 1$ implies $x = y$.*

Proof. By definition of the Hilbert space norm, any $x, y \in H$ obey

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle .$$

Hence $\|x\| = \|y\| = 1$ and $\langle x, y \rangle \geq 1 - \delta$ imply the desired inequality:

$$\|x - y\|^2 \leq 2 - 2(1 - \delta) = 2\delta. \qquad ////$$

1.3 Proposition. *The norm topology and the weak topology of H coincide on the set S. (In other words, a Hilbert space norm is a Kadec norm.)*

Proof. Let N be a strongly open subset of S. We must show that N is also weakly open. To do so, we will show that any point $x \in N$ has a weak neighbourhood contained in N.

Fix $x \in N$. Since N is strongly open, there exists $\varepsilon > 0$ such that $(x + \varepsilon B) \cap S \subseteq N$. Now consider the weak-open set $U = \{y \in H \,:\, \langle x, y \rangle > 1 - \frac{1}{2}\varepsilon^2\}$: certainly $x \in U$. Moreover, if $y \in U \cap S$, then

$$\|x - y\|^2 = 1 - 2\langle x, y \rangle + 1 = 2(1 - \langle x, y \rangle) < \varepsilon^2.$$

Thus $U \cap S \subseteq (x + \varepsilon B) \cap S \subseteq N$. ////

1.4 Corollary. *Let $\{x_n\}$ be a sequence in H. If $x_n \xrightarrow{\text{w}} x \neq 0$ and $\|x_n\| \to \|x\|$, then $x_n \to x$.*

Proof. Consider the sequence $y_n = \dfrac{x_n}{\|x_n\|}$ in S. For any $z \in H$, we have

$$\langle z, y_n \rangle = \frac{1}{\|x_n\|} \langle z, x_n \rangle \to \frac{1}{\|x\|} \langle z, x \rangle .$$

Hence $y_n \xrightarrow{\text{w}} y = \dfrac{x}{\|x\|}$. By Proposition 1.3, $y_n \to y$, i.e.

$$x = \lim_{n \to \infty} \left(\frac{\|x\|}{\|x_n\|} \right) x_n.$$

Since $\dfrac{\|x\|}{\|x_n\|} \to 1$ as $n \to \infty$, it follows that $x_n \to x$. ////

1.5 Proposition. *For any $x \in H \setminus \{0\}$, the norm of H is Fréchet differentiable at x, with derivative* $\dfrac{x}{\|x\|}$.

Proof. Observe that

$$\|v\|^2 = \|x+v\|^2 - \|x\|^2 - 2\langle x, v\rangle$$

$$= (\|x+v\| - \|x\|)(\|x+v\| + \|x\|) - 2\langle x, v\rangle .$$

Rearranging terms in this equation yields the identity

$$\frac{\|x+v\| - \|x\| - \left\langle \frac{x}{\|x\|}, v\right\rangle}{\|v\|} = \frac{\|v\|}{\|x+v\| + \|x\|} + \left\langle \frac{x}{\|x\|}, \frac{v}{\|v\|}\right\rangle \left[\frac{2\|x\|}{\|x+v\| + \|x\|} - 1\right].$$

Both terms on the right side evidently tend to 0 as $v \to 0$. ////

1.6 Corollary. *Let $x \neq 0$ be given in H. Then the following assertions about a vector $v \in H$ are equivalent.*

(a) $\|v\| = 1$ *and* $\langle v, x\rangle = \|x\|$.

(b) $v = \dfrac{x}{\|x\|}$.

(c) v is the Fréchet derivative of the norm at x.

Proof. (b)⇔(c) was the content of Proposition 1.5; (b)⇒(a) is obvious; and (a)⇒(b) is an instance of Proposition 1.2 with $\delta = 0$. ////

Section 2. Best Approximation by Closed Sets

We continue to study the geometry of the Hilbert space H. If $C \subseteq H$ is a nonempty closed subset of H, we may define the *distance function* from C as follows:

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

(The alternate notation $d(x; C) \equiv d_C(x)$ will also be used below.) A point $c \in C$ at which the infimum defining $d_C(x)$ is attained is called a *best approximation to x in C,* or else a *nearest point to x in C.* When H is simply finite-dimensional Euclidean space, the Heine-Borel theorem implies that every $x \in H$ admits a best approximation from C. For a general Hilbert space H, however, compactness is a more elusive property and best approximations become harder to find. If the set C is *convex,* then every $x \in H$ admits a unique best approximation (Rudin (1973), Thm. 12.3), but if C is not convex then there may be points x with no best approximation. For example, consider the Hilbert space ℓ^2 of square-summable real sequences with basis elements $e_1 = (1, 0, 0, \ldots)$, $e_2 = (0, 1, 0, \ldots)$, etc. In this space, the origin has no best approximation from the closed set $C = \{(1 + \frac{1}{n})e_n : n \in \mathbf{N}\}$. It can be shown, however, that "most" of the points in ℓ^2 do have best approximations from C. The objective of this section is to prove a general theorem of this type, due to Ka-Sing Lau.

The following technical lemma about the shape of the unit ball in H will be used in the proof of Lau's theorem. For any unit vector $x \in H$, $r \in (0,1)$, and $\delta \in (0,1)$, we define

$$M_r(x, \delta) = \left[rx + (1 - r + \delta)\overline{B}\right] \setminus B.$$

2.1 Lemma. *Let $r \in (0,1)$ be fixed. Then for any $\varepsilon > 0$ there exists $\delta > 0$ so small that for any $x \in S$, one has $\|y_1 - y_2\| < \varepsilon \quad \forall y_1, y_2 \in M_r(x, \delta)$.*

Proof. For any $y \in M_r(x, \delta)$,

$$\|y\| \leq \|y - rx\| + \|rx\| \leq (1 - r + \delta) + r = 1 + \delta.$$

Likewise,

$$\begin{aligned} 1 \leq \|y\|^2 &= \|rx + (y - rx)\|^2 \\ &= \|rx\|^2 + 2\langle rx, y - rx\rangle + \|y - rx\|^2 \\ &\leq r^2 + 2r\langle x, y\rangle - 2r^2 + (1 - r + \delta)^2. \end{aligned}$$

Rearrangement of terms gives

$$\begin{aligned} & 2r\langle x, y\rangle \geq 1 + r^2 - (1 - r + \delta)^2 \\ \iff \quad & \langle x, y\rangle \geq 1 - \left(\frac{1}{r} - 1\right)\delta - \frac{1}{2r}\delta^2. \end{aligned}$$

By Prop. 1.2, it follows that for $\delta \in (0,1)$,

$$\|x - y\| \leq \sqrt{2\delta\left(\frac{1}{r} - 1 + \frac{\delta}{2r}\right)} < K\sqrt{\delta},$$

where $K = \sqrt{\dfrac{3 - 2r}{r}}$ is independent of x.

Hence for any $x \in S$ and $\delta \in (0,1)$, $y_1, y_2 \in M_r(x, \delta)$ imply

$$\|y_1 - y_2\| \leq \|y_1 - x\| + \|x - y_2\| < 2K\sqrt{\delta}.$$

Given any $\varepsilon \in (0, 2K)$, one needs only to choose $\delta = \left(\dfrac{\varepsilon}{2K}\right)^2$ to verify the Lemma. ////

2.2 Remark. Let us define a nonnegative real-valued function d on H by

$$d(x) := \lim_{\delta \to 0+} \sup \left\{ \left\langle \frac{z-x}{\|z-x\|}, c_2 - c_1 \right\rangle \ : \ d(z; C_\delta(x)) < \delta,\ c_1, c_2 \in C_\delta(x) \right\},$$

where $C_\delta(x) = [x + (d_C(x) + \delta)\overline{B}] \cap C$. Observe that $d(c) = 0$ for all c in C. For if $c \in C$, then $C_\delta(c) = (c + \delta \overline{B}) \cap C$. Hence $c_1, c_2 \in C_\delta(c)$ imply $\|c_2 - c_1\| \leq 2\delta$: thus $d(c) = 0$. Contraposition shows that if $d(x) > 0$ then $x \notin C$, so $d_C(x) > 0$.

2.3 Theorem. *Let $C \subseteq H$ be closed and nonempty. Then every point x in the set $G = \{x \in H \ : \ d(x) = 0\}$ has a best approximation from C.*

Proof. Fix any $x \in G$. We will construct a nearest point to x in C.

Consider any sequence $\{c_n\}$ for which $c_n \in C_{\frac{1}{n}}(x)$. Then $\{\|c_n\|\}$ is bounded, so along some subsequence (we do not relabel), c_n converges weakly to some point $c_x \in H$. This is our candidate for the nearest point to x in C. Note that since $\|c_n - x\| \leq d_C(x) + \frac{1}{n}$ for each n, we have $\|c_x - x\| \leq d_C(x)$. Therefore it suffices to prove $c_x \in C$.

Fix any $\varepsilon > 0$. The definition of d yields a corresponding δ_0 such that the indicated supremum is less than ε whenever $0 < \delta < \delta_0$. Choose $N \geq 1/\delta_0$. Then whenever $m > n > N$, we have

$$\varepsilon > \left\langle \frac{c_n - x}{\|c_n - x\|}, c_n - c_m \right\rangle = \left\langle \frac{c_n - x}{\|c_n - x\|}, c_n - x \right\rangle + \left\langle \frac{c_n - x}{\|c_n - x\|}, x - c_m \right\rangle$$

$$\Longleftrightarrow \qquad \left\langle \frac{c_n - x}{\|c_n - x\|}, c_m - x \right\rangle > \|c_n - x\| - \varepsilon.$$

Letting $m \to \infty$ here, we find

$$d_C(x) \geq \|c_x - x\| \geq \left\langle \frac{c_n - x}{\|c_n - x\|}, c_x - x \right\rangle > \|c_n - x\| - \varepsilon \geq d_C(x) - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\|c_x - x\| = d_C(x)$ and that $\|c_n - x\| \to \|c_x - x\|$. Since $(c_n - x)$ converges weakly to $(c_x - x)$ also, we deduce that $c_n \to c_x$ (Corollary 1.4). Since C is closed, we have $c_x \in C$ as required. ////

The following theorem justifies the statement that most points of H have best approximations from C by demonstrating that in fact, most points of H lie in G.

2.4 Theorem. *The set $G = \{x \in H : d(x) = 0\}$ is a dense G_δ subset of H.*

Proof. Consider the increasing sequence of sets $F_n = \{x \in H : d(x) \geq 1/n\}$. It suffices to show that each F_n is closed and nowhere dense, since then the G_δ set $G = \bigcap_{n=1}^{\infty} (H \setminus F_n)$ is dense by Baire's theorem.

To show that F_n is closed, fix any $x_0 \notin F_n$. Then there exists $\Delta > 0$ such that $0 < \delta \leq \Delta$ implies

$$\left\langle \frac{z - x_0}{\|z - x_0\|}, c_2 - c_1 \right\rangle < \frac{1}{n} - \Delta \qquad \forall z \in C_\delta(x_0) + \delta B, \quad c_1, c_2 \in C_\delta(x_0).$$

Pick $\delta = \frac{1}{3}\Delta$. Then $x \in x_0 + \delta \overline{B}$ implies $d_C(x) \leq d_C(x_0) + \delta$: hence $C_\delta(x) \subseteq C_\Delta(x_0)$. In particular, $d(z; C_\delta(x)) < \delta$ implies $d(z - (x - x_0); C_\delta(x)) < 2\delta$, and thus $d(z - (x - x_0); C_\Delta(x_0)) < 2\delta < \Delta$. Consequently $x \in x_0 + \delta \overline{B}$ implies that for all z such that $d(z; C_\delta(x)) < \delta$ and all $c_1, c_2 \in C_\delta(x_0) \subseteq C_\Delta(x_0)$,

$$\left\langle \frac{z - (x - x_0) - x_0}{\|z - (x - x_0) - x_0\|}, c_2 - c_1 \right\rangle < \frac{1}{n} - \Delta.$$

By definition, $d(x) < 1/n$, i.e. $x_0 + \delta \overline{B} \subseteq H \setminus F_n$. So F_n is indeed closed.

To see that $\operatorname{int} F_n = \emptyset$, suppose the contrary. By translation, assume that $0 \in \operatorname{int} F_n$, so that there is some $r \in (0, 1)$ such that $rB \subseteq F_n$. We may also take $r < d_C(0)$, since $d(0) > 0$ implies $d_C(0) > 0$ by Remark 2.2. Finally, we lose no generality in scaling the inner product in such a way as to make $d_C(0) = 1$. By Lemma 2.1, there exists $\Delta > 0$ such that $\|y_2 - y_1\| < \dfrac{1}{n+1} \quad \forall y_1, y_2 \in M_r(x, \Delta), \quad x \in S$. Now for $\delta = \frac{1}{2}\Delta$, pick any $x_0 \in C_\delta(0)$: Define $x_1 = \dfrac{x_0}{\|x_0\|}$ and $x_r = r x_1$. Then $C_\delta(x_r) \subseteq M_r(x_1, 2\delta)$, so we have

$$\left\langle \frac{z - x_r}{\|z - x_r\|}, c_2 - c_1 \right\rangle \leq \|c_2 - c_1\| < \frac{1}{n+1} \qquad \forall c_1, c_2 \in C_\delta(x_r), \quad z \in H.$$

In particular, $d(x_r) < 1/n$, a contradiction. Thus each F_n is nowhere dense. ////

Section 3. Proximal Normals and Clarke's Normal Cone

As we saw in the introduction, the normal cone to a closed nonempty subset C of H is a key ingredient in the theory of generalized gradients. However, it is not a basic ingredient: rather, it is defined indirectly in terms of the *tangent cone,* which we now introduce.

3.1 Definition. Let C be a nonempty closed subset of H containing some point c. The *tangent cone to C at c,* denoted $T_C(c)$, is the set of vectors $y \in H$ obeying any one of the following three equivalent conditions.

(a) $\limsup_{\substack{x \to c \\ t \searrow 0}} \frac{d_C(x+ty) - d_C(x)}{t} = 0.$

(b) For every pair of sequences $c_i \to c$ in C and $t_i \searrow 0$, there is a sequence $y_i \to y$ in H such that $c_i + t_i y_i \in C \quad \forall i$.

(c) For every $\varepsilon > 0$ there exist $\delta > 0$, $\lambda > 0$ such that

$$C \cap \left[x + t(y + \varepsilon B)\right] \neq \emptyset \qquad \forall x \in C \cap (c + \delta B), \quad t \in (0, \lambda].$$

The equivalence of definitions (a) and (b) is proven in Clarke (1983), Theorem 2.4.5; the same method establishes (a$\Rightarrow$c), while (c$\Rightarrow$b) is obvious.

Here is a particularly useful characterization of $T_C(c)$, quoted from Treiman (1983), Thm. 2.1.

3.2 Proposition. *Let c and C be as in Definition 3.1. Then $y \notin T_C(c)$ if and only if there exist $\varepsilon > 0$ and sequences $c_i \to c$ in C and $\lambda_i > 0$ such that*

$$C \cap \left[c_i + (0, \lambda_i](y + \varepsilon B)\right] = \emptyset \qquad \forall i.$$

Now the normal cone is defined by polarity with the tangent cone.

3.3 Definition. Let C be a nonempty closed subset of H containing some point c. The *normal cone to C at c,* denoted $N_C(c)$, is defined by

$$N_C(c) = T_C^\circ(c) = \{v \in H \, : \, \langle v, y \rangle \leq 0 \quad \forall y \in T_C(c)\}.$$

Since the geometric features of the generalized gradient are more closely linked to the normal cone than to the tangent cone, a direct characterization of $N_C(c)$ would certainly be a useful companion to Definition 3.3. Such a characterization is the goal of this chapter: it is given in Section 4. The fundamental idea behind it is that there ought to be a relationship between normals and perpendiculars.

3.4 Definition. A vector $v \in H$ is *perpendicular to C at a point $c \in C$* if $c + v$ is a point whose best approximation from C is c. This situation is written as $v \perp C$ at c, and each of the vectors tv, $t \geq 0$, is called a *proximal normal to C at c*. The set of all proximal normals to C at c is denoted $PN_C(c)$.

3.5 Proposition. *The following assertions about $v \in H$ and $\widehat{c} \in C$ are equivalent.*

(a) v is a proximal normal to C at $\widehat{c}$.

(b) For some $t > 0$, $\|tv\| = d_C(\widehat{c} + tv)$.

(c) For some $t > 0$, $\langle v, c - \widehat{c}\rangle \leq \dfrac{1}{2t}\|c - \widehat{c}\|^2 \quad \forall c \in C$.

Moreover, the same t will serve in both (b) and (c).

Proof. (a$\Leftrightarrow$b) restates the definition of a proximal normal. To prove (b$\Leftrightarrow$c), observe that the following three statements about $c \in C$ are equivalent for any $t > 0$:

$$\|tv\|^2 = \|(\widehat{c} + tv) - \widehat{c}\|^2 \leq \|(\widehat{c} + tv) - c\|^2,$$

$$t^2\|v\|^2 \leq \|\widehat{c} - c\|^2 + 2\langle \widehat{c} - c, tv\rangle + t^2\|v\|^2,$$

$$\langle v, c - \widehat{c}\rangle \leq \frac{1}{2t}\|\widehat{c} - c\|^2. \qquad ////$$

3.6 Remarks. 1. Note that 0 is perpendicular to C at every point. Thus $PN_C(c)$ is never empty. Moreover, the results of Section 2 indicate that most points of H have a best approximation from C. Each of these points lying outside C gives rise to a nonzero element of $PN_C(c)$ at some point $c \in C$.

2. In Clarke (1983), the definition of perpendicular requires that c provide the *unique* best approximation in C for the vector $c + v$. Since the ball of H is strictly convex, the proximal normal vectors corresponding to Clarke's more restrictive definition are identical to those discussed here.

3. Another widely used concept of normality to a given closed set involves "Fréchet normals." A vector $v \in H$ is *Fréchet normal to C at a point $\widehat{c} \in C$* if for every $\varepsilon > 0$ there is a neighbourhood N_ε of $\widehat{c}$ such that

$$\langle v, c - \widehat{c}\rangle \leq \varepsilon\|c - \widehat{c}\| \qquad \forall c \in C \cap N_\varepsilon.$$

From Proposition 3.5(c), we see that every proximal normal to C at $\widehat{c}$ is automatically a Fréchet normal to C at $\widehat{c}$. The converse is false even in $\mathbf{R}^2$, where one may take $C = \{(x, y) : y \leq |x|^{3/2}\}$, $\widehat{c} = (0, 0)$, and $v = (0, 1)$. In this example v is a Fréchet normal to C at $\widehat{c}$, but not a proximal normal.

Section 4. The Proximal Normal Formula

The main result of this chapter is a formula which allows the computation of Clarke's normal cone in terms of the proximal normal vectors discussed in Section 3. It asserts that $N_C(c) = R_C(c)$, where $R_C(c)$ is the closed convex cone

$$R_C(c) = \overline{\text{co}}\{\text{w-lim}\, v_i \,:\, v_i \text{ is a bounded sequence of proximal normals}$$
$$\text{to } C \text{ at corresponding base points } c_i \to c\}.$$

As one might expect, this is proven by establishing two set inclusions, one of which is relatively straightforward.

4.1 Proposition. $N_C(c) \supseteq R_C(c)$.

Proof. Let v lie in the set whose closed convex hull is computed to yield $R_C(c)$. By definition, there is a bounded sequence $\{v_i\}$ tending weakly to v and a corresponding sequence of base points $c_i \to c$ such that $v_i \in PN_C(c_i)$ for each i. By Proposition 3.5(c), there exists a sequence $\varepsilon_i > 0$ such that $\langle v_i, x - c_i \rangle \le \varepsilon_i \|x - c_i\|^2 \quad \forall x \in C$. Choose a sequence t_i decreasing to 0 such that $t_i\varepsilon_i \to 0$. Then for any $y \in T_C(c)$, Definition 3.1(b) asserts that there is a sequence y_i in H converging to y such that $c_i + t_i y_i \in C \quad \forall i$. The proximal normal inequality quoted above gives

$$\langle v_i, (c_i + t_i y_i) - c_i \rangle \le \varepsilon_i \|(c_i + t_i y_i) - c_i\|^2 \quad \forall i,$$

$$\iff \qquad \langle v_i, y_i \rangle \le t_i \varepsilon_i \|y_i\|^2 \qquad \forall i.$$

Since $t_i\varepsilon_i \to 0$ and $\{\|y_i\|\}$ is bounded, we find that $\limsup_{i\to\infty} \langle v_i, y_i \rangle \le 0$.

Now since v_i converges weakly to v, we have

$$\begin{aligned} \langle v, y \rangle &= \lim_{i\to\infty} \langle v_i, y \rangle \\ &\le \limsup_{i\to\infty} \langle v_i, y_i \rangle + \limsup_{i\to\infty} \langle v_i, y - y_i \rangle . \end{aligned}$$

The first term on the right side is known to be non-positive; the second is zero because $\{v_i\}$ is bounded and $y_i \to y$. Hence $\langle v, y \rangle \le 0$.

Since $y \in T_C(c)$ was arbitrary, we find that $\langle v, y \rangle \le 0 \quad \forall y \in T_C(c)$, i.e. $v \in N_C(c)$. Since $N_C(c)$ is known to be closed and convex, it must contain the closed convex hull of all such points v, namely $R_C(c)$. ////

The assertion that $N_C(c) \subseteq R_C(c)$ is considerably more difficult to prove. It is equivalent to the statement that $T_C(c) \supseteq R_C^\circ(c)$, or $H \setminus T_C(c) \subseteq H \setminus R_C^\circ(c)$. We prove this last fact by considering any unit vector $y \in H \setminus T_C(c)$ and exhibiting a $v \in R_C(c)$ such that $\langle v, y\rangle > 0$.

The construction is based on Proposition 3.2. Since $y \notin T_C(c)$, there is an $\varepsilon > 0$, a sequence of points $c_i \to c$ in C, and constants $\lambda_i > 0$ such that

$$C \cap [c_i + (0, \lambda_i](y + \varepsilon B)] = \emptyset \quad \forall i.$$

For each i, we construct a unit proximal normal vector v_i to C at a point c_i' near c_i such that each v_i has a rather large inner product with y. The weak limit of a subsequence of $\{v_i\}$ then provides the desired v.

The direction y will clearly play a significant part in this effort. We let $y^\perp$ denote the orthogonal complement of the subspace $\mathbf{R}y$ in H. Then every vector $v \in H$ can be written as $v = \lambda x + \mu y$ for some $x \in y^\perp \cap S$ and $\lambda, \mu \in \mathbf{R}$. In this decomposition, $\|\lambda x + \mu y\|^2 = \lambda^2 + \mu^2$.

The next few lemmas deal with the situation for a fixed i. No generality is lost in translating the problem so that $c_i = 0$. For the $\varepsilon > 0$ given above, we consider the cone

$$K = \bigcup_{t \geq 0} t(y + \varepsilon B).$$

We also fix $\alpha \in (0, 1)$ (a specific choice will be made later) and define the set

$$E = \{\lambda x + \mu y : \frac{\lambda^2}{\alpha^2} + \mu^2 < 1, \quad x \in y^\perp \cap S\}.$$

(In the simplest of all special cases, when $H = \mathbf{R}^2$, E is an ellipse whose major axis lies along the y-axis, and each of the following five lemmas has a simple geometrical interpretation. Figures 1–6 at the end of this section are included to justify this statement and to make the proof easy to follow.)

4.2 Lemma. *Let $x \in y^\perp \cap S$. For any $\lambda, \mu \in \mathbf{R}$ one has*

$$\lambda x + \mu y \in K \quad \Leftrightarrow \quad \mu \geq \frac{\sqrt{1 - \varepsilon^2}}{\varepsilon} |\lambda|.$$

Proof. The point $\lambda x + \mu y$ lies in K if and only if there is a scalar $\rho > 0$ such that

$$\|\rho(\lambda x + \mu y) - y\|^2 \leq \varepsilon^2.$$

This inequality reduces to $\lambda^2 \rho^2 + (\mu\rho - 1)^2 \leq \varepsilon^2$, the left side of which has a minimum value of $\dfrac{\lambda^2}{\lambda^2 + \mu^2}$ when $\rho = \dfrac{\mu}{\lambda^2 + \mu^2}$. Thus $\lambda x + \mu y \in K$ if and only if $\mu \geq 0$ and $\dfrac{1}{1 + (\mu/\lambda)^2} \leq \varepsilon^2$. The result follows. ////

4.3 Lemma. *Let $\rho > 0$. If $\lambda x + \mu y \in \big(y + (1+\rho)E\big) \setminus K$ for some $x \in y^\perp \cap S$ and constants $\lambda, \mu \in \mathbf{R}$, then*

$$\mu < \frac{1 + \sqrt{1 + \rho(2+\rho)(1 + \alpha^{-2}m^{-2})}}{1 + \alpha^{-2}m^{-2}}, \qquad \text{where } m = \frac{\sqrt{1-\varepsilon^2}}{\varepsilon}.$$

(Note that the same estimate for μ holds if we assume $t(\lambda x + \mu y) \in t\big(y + (1+\rho)E\big) \setminus K$ for any $t > 0$.)

Proof. If $\mu \leq 0$ the conclusion is evident, so assume $\mu > 0$. Observe that $\lambda x + \mu y \in y + (1+\rho)E$ is equivalent to

$$(*) \qquad \frac{1}{\alpha^2}\lambda^2 + (\mu - 1)^2 \leq (1+\rho)^2,$$

while $\lambda x + \mu y \notin K$ is equivalent to $\mu < m\,|\lambda|$. Substituting this into $(*)$ yields

$$\left(1 + \alpha^{-2}m^{-2}\right)\mu^2 - 2\mu - \rho(2+\rho) < 0.$$

The positive root of the quadratic function of μ on the left here gives the maximum permissible value of μ, which is precisely the estimate written above. ////

4.4 Lemma. *Suppose points $v_0 = \lambda_0 x_0 + \mu_0 y$ and $v_1 = \lambda_1 x_1 + \mu_1 y$ in H are given $(x_0, x_1 \in y^\perp \cap S)$ such that v_1 lies on the boundary of $v_0 + \rho E$ for some $\rho > 0$. If $\langle y, v_0 - v_1 \rangle \geq \rho\eta$ for some $\eta > 0$, then for the unit vector $\widehat{n}$ in direction $-(\lambda_1 x_1 - \lambda_0 x_0) - \alpha^2(\mu_1 - \mu_0)y$ one has*

$$\langle \widehat{n}, y \rangle \geq \frac{\alpha}{\left[\eta^{-2} - (1 - \alpha^2)\right]^{1/2}}.$$

Proof. Without loss of generality take $v_0 = 0$, $\rho = 1$. Then observe that

$$\left\|-\lambda_1 x_1 - \alpha^2 \mu_1 y\right\|^2 = \lambda_1^2 + \alpha^4 \mu_1^2 = \alpha^2(1 - \mu_1^2) + \alpha^4 \mu_1^2 = \alpha^2[1 - (1 - \alpha^2)\mu_1^2].$$

This function of μ_1 takes its maximum value on $(-\infty, -\eta]$ at $\mu_1 = -\eta$. Therefore

$$\langle \widehat{n}, y \rangle = \frac{\alpha^2(-\mu_1)}{\|-\lambda_1 x_1 - \alpha^2 \mu_1 y\|} \geq \frac{\alpha^2 \eta}{\sqrt{\alpha^2[1 - (1 - \alpha^2)\eta^2]}}.$$

The indicated inequality is a rearrangement of this one. ////

Now the reason for studying E so closely is that it can be considered as the unit ball of a new norm on H. Indeed, if we define a bijective linear mapping $L: H \to H$ by

$$L(\lambda x + \mu y) = \frac{\lambda}{\alpha} x + \mu y \qquad \forall x \in y^\perp \cap S, \quad \lambda, \mu \in \mathbf{R},$$

then a new inner product on the vector space H is obtained by defining

$$\begin{aligned} \langle\!\langle \lambda_1 x_1 + \mu_1 y, \lambda_2 x_2 + \mu_2 y \rangle\!\rangle &= \langle L(\lambda_1 x_1 + \mu_1 y), L(\lambda_2 x_2 + \mu_2 y) \rangle \\ &= \frac{\lambda_1 \lambda_2}{\alpha^2} \langle x_1, x_2 \rangle + \mu_1 \mu_2. \end{aligned}$$

The new inner product induces the norm $|||\lambda x + \mu y|||^2 = \dfrac{\lambda^2}{\alpha^2} + \mu^2$ whose unit ball is E. Clearly the new norm is equivalent to the original one, so $\big(H, \langle\!\langle \cdot, \cdot \rangle\!\rangle\big)$ is complete—i.e. is a Hilbert space in its own right. The key fact to be used below is that $|||\cdot|||$-proximal normals correspond to $\|\cdot\|$-proximal normals.

4.5 Lemma. *Suppose that $\lambda_1 x_1 + \mu_1 y$ has a closest point $\bar{c} = (\lambda_1 x_1 + \mu_1 y) + (\lambda_0 x_0 + \mu_0 y) \in C$ $(x_0, x_1 \in y^\perp \cap S)$ with respect to the norm $|||\cdot|||$. Then a proximal normal vector to C at $\bar{c}$ is*

$$\hat{n} = -\left(\frac{\lambda_0 x_0 + \alpha^2 \mu_0 y}{\|\lambda_0 x_0 + \alpha^2 \mu_0 y\|} \right).$$

Proof. Without loss of generality, assume $\lambda_1 = \mu_1 = 0$. Let $r = |||\bar{c}|||$, and note that $\|\lambda_0 x_0 + \alpha^2 \mu_0 y\| = |(\lambda_0, \alpha^2 \mu_0)|$ where $|\cdot|$ is the Euclidean norm in $\mathbf{R}^2$. Then it suffices to show that $\bar{c} + \alpha^2 r \hat{n}$ has $\|\cdot\|$-closest point $\bar{c}$ in C. For this, we need only prove that $\bar{c} + \alpha^2 r \hat{n} + \alpha^2 r B \subseteq rE$, i.e. that for any $x \in y^\perp \cap S$ and any $\lambda, \mu \in \mathbf{R}$ with $\lambda^2 + \mu^2 \le 1$, one has

$$\begin{aligned} & |||\bar{c} + \alpha^2 r \hat{n} + \alpha^2 r(\lambda x + \mu y)|||^2 \le r^2 \\ \iff \quad & \frac{1}{\alpha^2} \left\| \lambda_0 x_0 - \alpha^2 r \frac{\lambda_0 x_0}{|(\lambda_0, \alpha^2 \mu_0)|} + \alpha^2 r \lambda x \right\|^2 + \left[\mu_0 - \frac{\alpha^4 r \mu_0}{|(\lambda_0, \alpha^2 \mu_0)|} + \alpha^2 r \mu \right]^2 \le r^2 \\ \iff \quad & \frac{1}{\alpha^2} \left[\lambda_0^2 \left(1 - \frac{\alpha^2 r}{|(\lambda_0, \alpha^2 \mu_0)|}\right)^2 + 2\alpha^2 r \lambda \lambda_0 \left(1 - \frac{\alpha^2 r}{|(\lambda_0, \alpha^2 \mu_0)|}\right) \langle x_0, x \rangle + \alpha^4 r^2 \lambda^2 \right] \\ & \qquad + \left[\mu_0 - \frac{\alpha^4 r \mu_0}{|(\lambda_0, \alpha^2 \mu_0)|} + \alpha^2 r \mu \right]^2 \le r^2. \end{aligned}$$

This would be assured if we could prove that

$$\begin{aligned} & \frac{1}{\alpha^2} \left[\lambda_0^2 \left(1 - \frac{\alpha^2 r}{|(\lambda_0, \alpha^2 \mu_0)|}\right)^2 + 2\alpha^2 r \left(1 - \frac{\alpha^2 r}{|(\lambda_0, \alpha^2 \mu_0)|}\right) |\lambda_0 \lambda| + \alpha^4 r^2 \lambda^2 \right] \\ & \qquad + \left[\mu_0 - \frac{\alpha^4 r \mu_0}{|(\lambda_0, \alpha^2 \mu_0)|} + \alpha^2 r \mu \right]^2 \le r^2. \end{aligned}$$

This, however, is precisely the statement of Corollary 7.2, in which one must simply replace (x, y) by (λ, μ) throughout. ////

4.6 Lemma. *Let $c \in C$ be a point for which there exists $\Lambda > 0$ such that $C \cap [c + (0, \Lambda](y + \varepsilon B)] = \emptyset$. Then for any $\eta > 0$ there exists a point $\bar{c} \in C$ and a unit vector $v \in H$ such that*

(a) $\|\bar{c} - c\| \leq \eta$,

(b) v *is a* $\|\cdot\|$*-proximal normal to* C *at* $\bar{c}$,

(c) $\langle v, y\rangle \geq \dfrac{(4m)^{-1}}{[15 + (4m)^{-2}]^{1/2}}$, *where* $m = \dfrac{\sqrt{1-\varepsilon^2}}{\varepsilon}$.

Proof. Without loss of generality, we may take $c = 0$.

Let $\eta > 0$ be given. Choose $\alpha = (4m)^{-1}$, where $m = \dfrac{\sqrt{1-\varepsilon^2}}{\varepsilon}$. Then for any $\rho \in (0, \frac{1}{6})$ and $0 < t < \eta/4$, we argue as follows.

By Theorem 2.4, the open set $t(y + \rho E)$ in $(H, \langle\!\langle \cdot, \cdot \rangle\!\rangle)$ contains a point $\bar{p} = t(\lambda x + \mu y)$ which has a best approximation $\bar{c} = \bar{\lambda}\bar{x} + \bar{\mu} y$ in C. Note that since $0 \in C$,

$$(\dagger) \qquad |||\bar{p} - \bar{c}||| \leq |||\bar{p}||| \leq |||\bar{p} - ty||| + |||ty||| \leq t\rho + t.$$

Hence $\|\bar{c}\| \leq |||\bar{c}||| \leq |||\bar{c} - \bar{p}||| + |||\bar{p}||| \leq 2t(1+\rho)$ implies (a), by our choices of ρ, t. If we choose t so small that $2t(1+\rho) \leq (1-\varepsilon)\Lambda$, then this also gives $\|\bar{c}\| \leq (1-\varepsilon)\Lambda$. Since $\bar{c} \in C$, it follows that $\bar{c}$ must lie outside the cone K (or else $\bar{c} = 0$, for which the following arguments also remain valid). Inequality ($\dagger$) also asserts that

$$|||\bar{c} - ty||| \leq |||\bar{c} - \bar{p}||| + |||\bar{p} - ty||| \leq t(1+\rho) + t\rho = t(1 + 2\rho).$$

It follows from Lemma 4.3 that

$$\bar{\mu} \leq t \cdot \frac{1 + \sqrt{1 + 2\rho(2+2\rho)(1 + \alpha^{-2}m^{-2})}}{1 + \alpha^{-2}m^{-2}} = t \cdot \frac{1 + \sqrt{1 + 68\rho(1+\rho)}}{17} \leq \frac{1}{3}t$$

by our choice of α. Consequently

$$\langle y, \bar{p} - \bar{c}\rangle = t\mu - \bar{\mu} \geq t(\mu - \frac{1}{3}) \geq t(1 - \rho - \frac{1}{3}) \geq \frac{1}{2}t$$

by our choice of ρ. By Lemma 4.4, with $\eta = \dfrac{t/2}{|||\bar{p} - \bar{c}|||} \geq \dfrac{1/2}{1+\rho} \geq 1/4$, we obtain

$$\langle v, y\rangle \geq \frac{\alpha}{[16 - 1 + \alpha^2]^{1/2}} = \frac{(4m)^{-1}}{[15 + (4m)^{-2}]^{1/2}},$$

where v is the $\|\cdot\|$-unit vector in direction

$$-(\bar{\lambda}\bar{x} - t\lambda x) - \alpha^2(\bar{\mu} - t\mu)y.$$

By Lemma 4.5, this v is a proximal normal to C at $\bar{c}$. ////

We may now complete the proof of the proximal normal formula.

4.7 Theorem. $N_C(c) \subseteq R_C(c)$.

Proof. We follow the program set forth at the beginning of this section. Pick any unit vector $y \notin T_C(c)$. Then by Proposition 3.2, there is some $\varepsilon > 0$ and a sequence of points $c_i \to c$ in C and constants $\lambda_i \searrow 0$ such that

$$C \cap \left[c_i + (0, \lambda_i](y + \varepsilon B)\right] = \emptyset \qquad \forall i.$$

Let $m = \dfrac{\sqrt{1-\varepsilon^2}}{\varepsilon}$ and $\delta = \dfrac{(4m)^{-1}}{[15 + (4m)^{-2}]^{1/2}} > 0$. By Lemma 4.6, there exists for each i a point $\overline{c_i} \in (c_i + \frac{1}{i}B) \cap C$ at which there is a unit vector v_i proximal normal to C and obeying $\langle v_i, y\rangle \geq \delta$. Now the bounded sequence $\{v_i\}$ must have a subsequence converging weakly to some $v \in H$. By definition, $v \in R_C(x)$. Along this subsequence (which we do not relabel), $\langle v_i, y\rangle \geq \delta \quad \forall i$ implies $\langle v, y\rangle \geq \delta > 0$. Hence $y \notin R_C^\circ(c)$. This completes the proof. ////

In the case where $H = \mathbf{R}^n$, weak convergence and strong convergence are indistinguishable. In particular, if v_i is a bounded sequence converging (weakly) to v, then either $v = 0$ or $\dfrac{v_i}{\|v_i\|} \to \dfrac{v}{\|v\|}$ (weakly). (This is false in ℓ^2–take $v_i = e_1 + e_i$.) Hence Clarke's (1983) Prop. 2.5.7, p. 68, is embedded in the proof given here.

4.8 Corollary. *Let $C \subseteq \mathbf{R}^n$ be a closed set containing a point c. Then $N_C(c)$ is the closed convex cone generated by the set*

$$\left\{ \lim_{i\to\infty} \frac{v_i}{|v_i|} \; : \; v_i \perp C \text{ at } c_i \to c, \; v_i \to 0 \right\} \cup \{0\}.$$

Diagrams. The elementary nature of our proof of Thm. 4.7 becomes apparent when one coinsiders the simple geometric significance of each step in the argument. Figures 1–6 below facilitate this by presenting the situation in $\mathbf{R}^2$. We have chosen $y = (0,1)$ and $x = (1,0)$ in these diagrams, and labelled the axes λ and μ to reflect the general case. Figure 1 is introductory, and the enumeration of Figs. 2–6 corresponds to that of Lemmas 4.2–4.6.

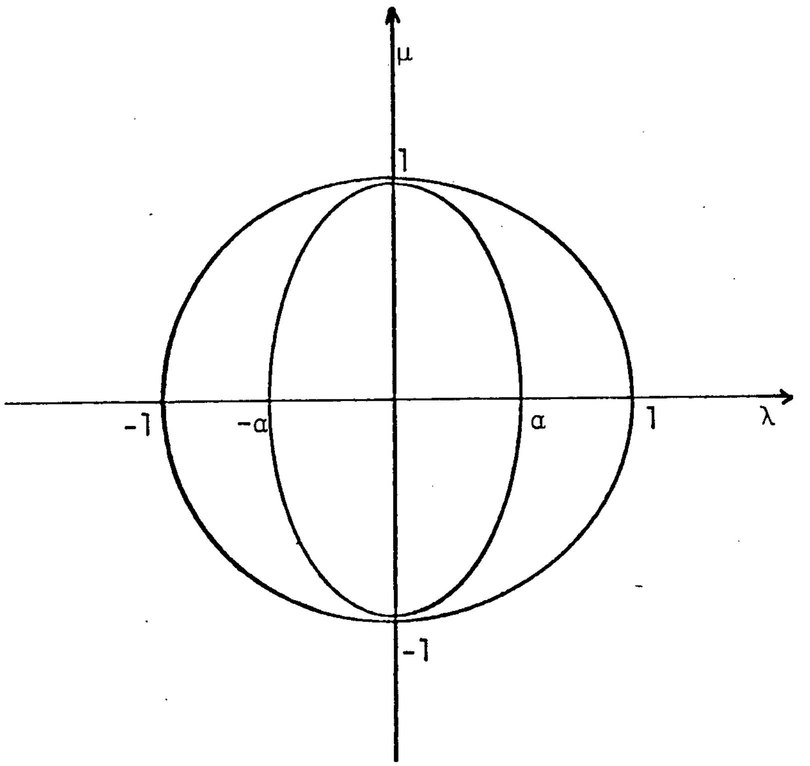

Fig. 1. The sets E and B.

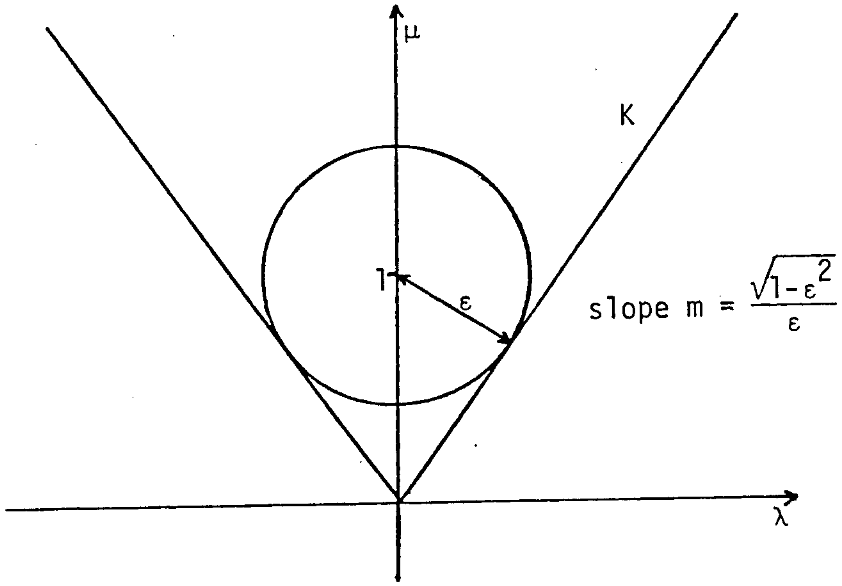

Fig. 2. Lemma 4.2 gives the slope of the sides of cone K.

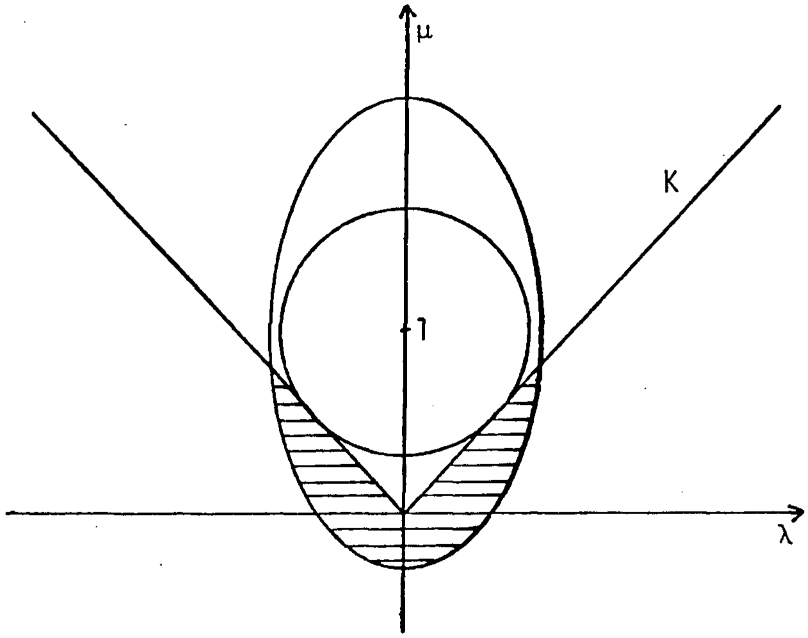

Fig. 3. Lemma 4.3 gives an explicit upper bound on μ for all points (λ, μ) in $[y + (1+\rho)E] \setminus K$, the shaded region.

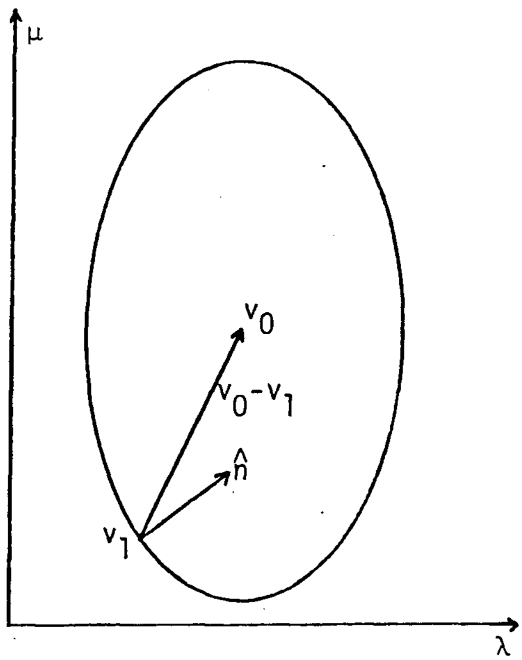

Fig. 4. Lemma 4.4 asserts that if $v_0 - v_1$ has a large y-component, then so does the related vector $\hat{n}$.

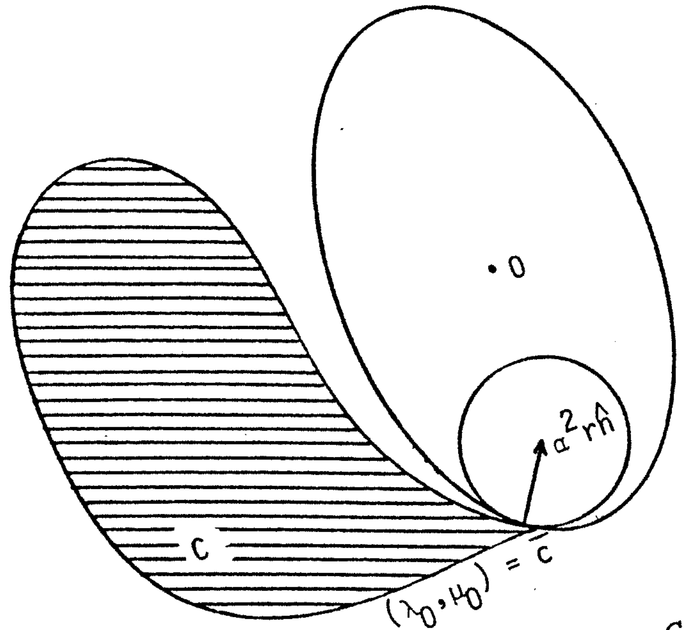

Fig. 5. Lemma 4.5 gives a vector $\hat{n}$ such that if $\bar{c} \in C$ is the $|||\cdot|||$-closest point in C to $(0,0) \notin C$, then $\bar{c}$ is the $\|\cdot\|$-closest point in C to $\bar{c} + \alpha^2 r\hat{n}$.

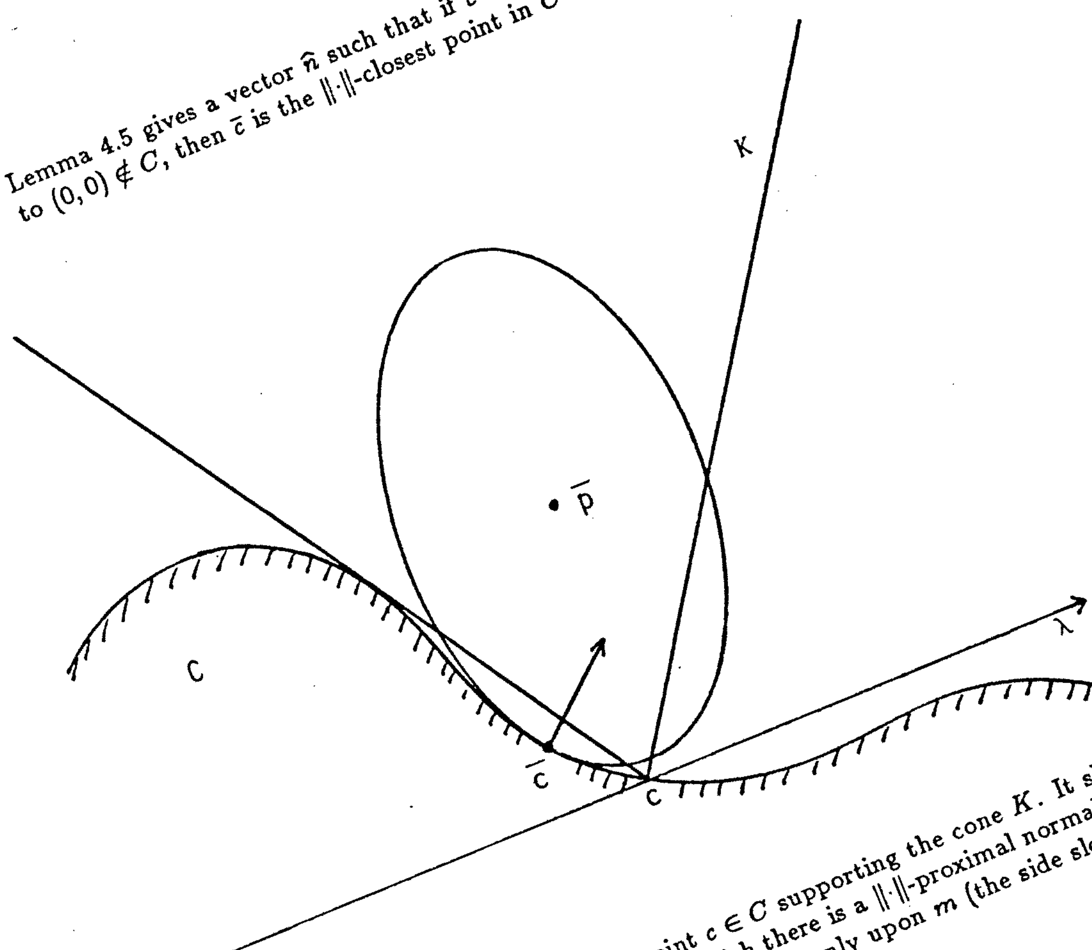

Fig. 6. Lemma 4.6 applies to a point $c \in C$ supporting the cone K. It shows how to find a point $\bar{c} \in C$ near c at which there is a $\|\cdot\|$-proximal normal vector whose y-component is positive and depends only upon m (the side slope of cone K).

Section 5. Proper Points

Given a nonempty closed set C in the Hilbert space H, a point $c \in C$ is called *proper* if $N_C(c)$ contains nonzero elements. In the finite-dimensional setting, the proximal normal formula implies that the proper points of C are precisely the boundary points of C.

5.1 Proposition. *Let $C \subseteq \mathbf{R}^n$ be a closed set containing a point c. Then*

$$N_C(c) = \{0\} \iff c \in \operatorname{int} C.$$

Proof. ($\Leftarrow$) If $c \in \operatorname{int} C$, then $T_C(c) = \mathbf{R}^n$ by Def. 3.1, and $N_C(c) = \{0\}$ follows by polarity.

($\Rightarrow$) Conversely, if $c \notin \operatorname{int} C$, then there is a sequence x_i converging to c such that each x_i lies outside C and has a closest point c_i in C. Let $v_i = x_i - c_i$. Then $c_i \to c$ because

$$|c_i - c| \leq |c_i - x_i| + |x_i - c| \leq 2\,|x_i - c|,$$

and $v_i \to 0$ by the same estimate. The proximal normal formula asserts that $N_C(c)$ contains all (weak) limit points of the sequence $\dfrac{v_i}{|v_i|}$. Since we are in a finite-dimensional setting, some unit vector is such a limit point. Hence $N_C(c) \neq \{0\}$. ////

The proof of Prop. 5.1 would carry over to arbitrary spaces H if only every sequence of unit vectors had a subsequence converging weakly to a nonzero limit. However, this need not be the case. A detailed account of what can go wrong in the infinite-dimensional case is provided by the following proposition, which shows that a sequence of unit vectors converges weakly to 0 if and only if it is "almost orthogonal".

5.2 Proposition. *For a weakly convergent sequence of unit vectors $\{v_i\} \subseteq H$, the following assertions are equivalent.*

(a) $v_i \xrightarrow{w} 0$.

(b) $\forall K \in \mathbf{N}, \qquad \lim\limits_{i \to \infty} \sup\limits_{k=1,\ldots,K} |\langle v_i, v_k \rangle| = 0.$

Proof. ($a \Rightarrow b$) Obvious.

($b \Rightarrow a$) We will prove "not (a) implies not (b)". Thus, let v_∞ denote the weak limit of v_i, and assume $\|v_\infty\|^2 = \varepsilon > 0$. Then there is a sequence of convex combinations of the form

$$w_k = \sum_{i=1}^{k} \lambda_i^{(k)} v_i, \qquad \lambda_i^{(k)} \geq 0, \qquad \sum_{i=1}^{k} \lambda_i^{(k)} = 1$$

for which $w_k \to v_\infty$ strongly. Consequently there exists some $K \in \mathbf{N}$ so big that $\|w_K - v_\infty\| < \varepsilon/2$. Write

$$\begin{aligned}\langle v_i, w_K \rangle &= \langle v_i, w_K - v_\infty \rangle + \langle v_i, v_\infty \rangle \geq \langle v_i, v_\infty \rangle - \|w_K - v_\infty\| \\ &\geq \langle v_i, v_\infty \rangle - \varepsilon/2.\end{aligned}$$

The RHS tends to the limit $\varepsilon - \varepsilon/2 > 0$, so we find

$$\liminf_{i \to \infty} \langle v_i, w_K \rangle \geq \varepsilon/2 > 0.$$

The left side of this inequality is itself majorized by

$$\liminf_{i \to \infty} \sup_{k=1,\ldots,K} \langle v_i, v_k \rangle,$$

so (b) fails as required. ////

Now suppose c is a boundary point of C. The only way to obtain $N_C(c) = \{0\}$ is if every sequence of proximal normal vectors (and there are a great many such sequences) is "almost orthogonal". This requires that the boundary of C be extraordinarily ill-behaved—so rough, in fact, that one might expect a very mild regularity condition on C (weak closure?) to eliminate the possibility completely. No such global result is known yet, but the proximal normal formula may be the key to solving this problem. In the remainder of this section we outline three additional approaches.

One generic result on the propriety of C's boundary is given by Borwein and Strojwas (1985b), Thm. 5.1. Its proof involves Ekeland's theorem; in our setting, the statement reduces to the following.

5.3 Theorem. *If $C \subseteq H$ is a closed set, then the set of points c for which $N_C(c) \neq \{0\}$ is dense in the boundary of C.*

A generic fact like Thm. 5.3 is tantalizing, but until proper points are completely characterized, one often wishes to know whether a specific point is proper for C. Our next proposition shows that if $c \in C$ is a point supporting (locally) a cone with nonempty interior, then c is proper.

5.4 Proposition. *Let $C \subseteq H$ be a closed set containing a cluster point c. If K is a cone with vertex at 0 for which there exists $\eta > 0$ such that $C \cap \big(c + (K \cap \eta B)\big) = \{c\}$, then*

$$\operatorname{int} K \cap T_C(c) = \emptyset.$$

In particular, if $\operatorname{int} K \neq \emptyset$, then $T_C(c) \neq H$ and $N_C(c) \neq \{0\}$.

Proof. Observe that $0 \notin \operatorname{int} K$, since this would force $K = H$ and then $c + \big(K \cap \eta B\big) = c + \eta B$ would meet C in a set strictly larger than $\{c\}$. So for any $y \in \operatorname{int} K$, there exists some $\varepsilon \in (0,1)$ such that $0 \notin y + \varepsilon B \subseteq K$. By assumption,

$$C \cap \left[c + \left(0, \frac{\eta}{1 + \|y\|}\right] (y + \varepsilon B)\right] \subseteq C \cap [c + (K \cap \eta B) \setminus \{0\}] = \emptyset.$$

Hence $y \notin T_C(c)$ by Prop. 3.2; this completes the proof. ////

Another approach to proving that a specific point $c \in C$ is proper relies on the following notion, due to Borwein and Strojwas (1984).

5.5 Definition. The closed set $C \subseteq H$ is called *compactly epi-Lipschitzian at $c \in C$* if there exist $\delta > 0$, $\varepsilon > 0$, $\lambda > 0$, and a compact set $K \subseteq H$ such that

$$C \cap (c + \delta B) + t\varepsilon B \subseteq C + tK \qquad \forall t \in (0, \lambda).$$

Note that any finite-dimensional set is compactly epi-Lipschitzian at all its points.

5.6 Theorem (Borwein-Strojwas (1984)). *Let $C \subseteq H$ be a closed set which is compactly epi-Lipschitzian at c. Then*

$$N_C(c) = \{0\} \iff c \in \operatorname{int} C.$$

Section 6. Rockafellar's Theorem

Previous sensitivity results based on proximal normals have relied on a geometrical result due to Rockafellar (1982), Prop. 15. His statement and proof of this result are intrinsically finite dimensional. In this section we extend Rockafellar's proposition to arbitrary normed spaces, beginning with a topology-free version of the result.

6.1 Lemma. *Let X be a real vector space containing a nonempty subset D and a cone D^∞ with vertex at 0. Define cones N, $N^\infty \subseteq X \times \mathbf{R}$ via*

$$N = \{\lambda(d, -1) \,:\, \lambda \geq 0, d \in D\},$$

$$N^\infty = \{(d^\infty, 0) \,:\, d^\infty \in D^\infty\}.$$

Then (a) $\{v \,:\, (v, -1) \in \mathrm{co}\,[N \cup N^\infty]\} = \mathrm{co}\,[D + D^\infty]$,

(b) $\{v \,:\, (v, 0) \in \mathrm{co}\,[N \cup N^\infty]\} = \mathrm{co}\, D^\infty$.

Proof. Note that both (a) and (b) are automatic if $D = \emptyset$. We therefore assume $D \neq \emptyset$.

(a) Let $L = \{v \,:\, (v, -1) \in \mathrm{co}\,[N \cup N^\infty]\}$ denote the left-hand side of (a). For any $d \in D$ and $d^\infty \in D^\infty$, we have $2(d, -1) \in N$ and $(2d^\infty, 0) \in N^\infty$, so

$$(d + d^\infty, -1) = \frac{1}{2}[2(d, -1)] + \frac{1}{2}(2d^\infty, 0) \in \mathrm{co}\,[N \cup N^\infty].$$

Thus $d + d^\infty \in L$. This shows $D + D^\infty \subseteq L$: since L is convex, we obtain $\mathrm{co}\,[D + D^\infty] \subseteq L$.

To prove the reverse inclusion, pick any $v \in L$. We will show that $v \in \mathrm{co}\,[D + D^\infty]$. By definition, $(v, -1) \in \mathrm{co}\,[N \cup N^\infty]$: thus there exist $n \in \mathbf{N}$ and $\mu_i > 0$ ($i = 1, \ldots, n$) with $\sum \mu_i = 1$ such that

$$(v, -1) = \sum_{i=1}^{k} \mu_i (d_i^\infty, 0) + \sum_{i=k+1}^{n} \mu_i \lambda_i (d_i, -1)$$

for some $d_1^\infty, \ldots, d_k^\infty \in D^\infty$, $\lambda_{k+1}, \ldots, \lambda_n > 0$, and $d_{k+1}, \ldots, d_n \in D$. Now if $k = 0$ this shows that $v \in \mathrm{co}\,[D + \{0\}] \subseteq \mathrm{co}\,[D + D^\infty]$, so we assume $k \geq 1$. Also, the second component of this equation forces $k < n$, and

$$\sum_{i=k+1}^{n} \mu_i \lambda_i = 1.$$

Now simply rewrite v as

$$\begin{aligned} v &= \sum_{i=1}^{k} \mu_i \left(\tfrac{1}{k}\mu_n \lambda_n d_n + d_i^\infty\right) + \sum_{i=k+1}^{n-1} \mu_i \lambda_i d_i + \mu_n \lambda_n \left[1 - \tfrac{1}{k} \sum_{i=1}^{k} \mu_i\right] d_n \\ &= \sum_{i=1}^{k} \mu_i \cdot \tfrac{1}{k}\mu_n \lambda_n \left(d_n + \frac{k}{\mu_n \lambda_n} d_i^\infty\right) + \sum_{i=k+1}^{n-1} \mu_i \lambda_i (d_i + 0) + \mu_n \lambda_n \left[1 - \tfrac{1}{k} \sum_{i=1}^{k} \mu_i\right] (d_n + 0). \end{aligned}$$

This expresses v as a convex combination of n points in $D + D^\infty$. (The coefficients are nonnegative since $\sum_{i=1}^{k} \mu_i \leq 1$, and their sum is

$$\tfrac{1}{k}\mu_n \lambda_n \sum_{i=1}^{k} \mu_i + \sum_{i=k+1}^{n-1} \mu_i \lambda_i + \mu_n \lambda_n - \tfrac{1}{k}\mu_n \lambda_n \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{n} \mu_i \lambda_i = 1.)$$

Thus indeed $L \subseteq \mathrm{co}\,[D + D^\infty]$.

(b) The left-hand side of (b) is convex and contains D^∞, so

$$\{v \, : \, (v, 0) \in \mathrm{co}\,[N \cup N^\infty]\} \supseteq \mathrm{co}\, D^\infty.$$

For the reverse inclusion, note that any v in the left-hand side must be the convex combination of finitely many vectors of the form $(d_i^\infty, 0)$ for $d_i^\infty \in D^\infty$. (Clearly, no nonzero vectors from N may be included in this convex combination.) Hence $v \in \mathrm{co}\, D^\infty$, as desired. ////

6.2 Proposition (Rockafellar). *Let X be a normed vector space containing nonempty subsets D and D^∞; assume that D^∞ is a cone with vertex at 0. Define cones N, N^∞ as in Lemma 6.1. Then one has*

$$\{v \, : \, (v, -1) \in \overline{\mathrm{co}}[N \cup N^\infty]\} = \overline{\mathrm{co}}[D + D^\infty].$$

Moreover, if $\mathrm{co}\,[N \cup N^\infty]$ *is closed, then one obtains*

$$\{v \, : \, (v, -1) \in \overline{\mathrm{co}}[N \cup N^\infty]\} = \mathrm{co}\,[D + D^\infty]$$

$$\{v \, : \, (v, 0) \in \overline{\mathrm{co}}[N \cup N^\infty]\} = \mathrm{co}\, D^\infty.$$

Proof. The stronger conclusions follow immediately from Lemma 6.1 when one writes in $\overline{\mathrm{co}}[N \cup N^\infty]$ for $\mathrm{co}\,[N \cup N^\infty]$. The first statement is the only one requiring elaboration.

According to Lemma 6.1, it suffices to show that

$$\{v \, : \, (v, -1) \in \overline{\mathrm{co}}[N \cup N^\infty]\} = \overline{\{v \, : \, (v, -1) \in \mathrm{co}\,[N \cup N^\infty]\}}.$$

Upon defining the convex cone $C = \mathrm{co}\,[N \cup N^\infty]$ and the closed affine subspace $S = X \times \{-1\}$ of $X \times \mathbf{R}$, we find that this reduces to showing

$$S \cap \mathrm{cl}\, C = \mathrm{cl}[S \cap C],$$

where cl denotes closure. Since $S \cap \mathrm{cl}\, C$ is a closed set containing $S \cap C$, we only need to prove $S \cap \mathrm{cl}\, C \subseteq \mathrm{cl}[S \cap C]$. Suppose, therefore, that $(v_\infty, -1) \in S \cap \mathrm{cl}\, C$. Then there must be some sequence $(v_i, -r_i) \in C$ tending to $(v_\infty, -1)$. The norm on $X \times \mathbf{R}$ forces $v_i \to v_\infty$, $r_i \to 1$. Consequently the sequence $(v_i/r_i, -1) \in S \cap C$ tends to $(v_\infty, -1)$, i.e. $(v_\infty, -1) \in \mathrm{cl}[S \cap C]$ as required. ////

Sufficient conditions for $\mathrm{co}\,[N \cup N^\infty]$ to be closed may be based on the following simple lemma.

6.3 Lemma. *Let $D \subseteq X$ be a closed set, and let $D^\infty \subseteq X$ be a closed cone with vertex at 0 such that*

$$D^\infty \supseteq 0^+D = \left\{ \lim_{i\to\infty} \lambda_i d_i \,:\, \lambda_i \to 0,\ d_i \in D \right\}.$$

Then $N \cup N^\infty$ is closed.

Proof. Clearly $N^\infty = D^\infty \times \{0\}$ is closed. It therefore suffices to show that if a sequence $\lambda_i(d_i, -1)$ in N converges to some point $(v, -\lambda)$, then this limit point lies in $N \cup N^\infty$. When $\lambda > 0$, the sequence d_i converges to v/λ, which we may denote by $d \in D$ since the set D is closed. Hence $(v, -\lambda) = \lambda(d, -1) \in N$. Alternatively, when $\lambda = 0$ we find $v \in 0^+D \subseteq D^\infty$ by assumption, so $(v, 0) \in N^\infty$. ////

The following lemma, proven in Rockafellar (1982), leads to a useful simplification of Prop. 6.2 when $X = \mathbf{R}^n$.

6.4 Lemma. *Let $C \subseteq \mathbf{R}^n$ be a closed cone with vertex at 0. If C is pointed, then* co C *is closed and pointed.*

6.5 Proposition (Rockafellar). *Take $X = \mathbf{R}^n$. Suppose that the sets D and D^∞ of Prop. 6.2 are closed, and that $D^\infty \supseteq 0^+D$. If D^∞ is pointed, then* co $[N \cup N^\infty]$ *is closed and hence the stronger conclusions of Prop. 6.2 hold.*

Proof. The set $N \cup N^\infty$ is certainly a cone with vertex at 0. It is closed by Lemma 6.3, so according to Lemma 6.4 it suffices to show that $N \cup N^\infty$ is pointed. If this were not the case, then there would be positive scalars $\lambda_1^\infty, \lambda_2^\infty, \ldots, \lambda_\ell^\infty$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\sum_{k=1}^{\ell} \lambda_k^\infty (d_k^\infty, 0) + \sum_{k=1}^{m} \lambda_k (d_k, -1) = (0, 0)$$

for some points $d_k \in D$, $d_k^\infty \in D^\infty \setminus \{0\}$. This clearly forces $m = 0$, and thus implies that D^∞ is not pointed. This contradicts our assumptions. ////

The only intrinsically finite-dimensional link in the chain of arguments supporting Prop. 6.5 is Lemma 6.4. It can be false even in the Hilbert space ℓ^2 of square-summable real sequences. To see this, let e_i denote the sequence whose i^{th} entry is 1 while all others are 0, and define

$$C = \bigcup_{\lambda \geq 0} \lambda \{e_i \, : \, k \in \mathbf{N}\}.$$

The set C is evidently a closed pointed cone with vertex 0 in ℓ^2. However, $\operatorname{co} C$ is not closed. For the point $x = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots)$ lies outside $\operatorname{co} C$, all of whose members are sequences with only finitely many nonzero terms, but inside $\overline{\operatorname{co}}\, C$.

Section 7. Appendix: Some Euclidean Geometry

The Hilbert-space computations of Section 4 involve a distinguished direction y, and treat all vectors in $y^\perp$ alike. Thus a good model for these computations can be displayed in the Cartesian plane: we take the y-axis to correspond to the distinguished direction y in the general theory, and visualize all of $y^\perp$ as the x-axis. For a fixed $\alpha \in (0,1)$, we now consider the ellipse

$$E = \{(x,y) \, : \, \frac{x^2}{\alpha^2} + y^2 < 1\}.$$

7.1 Theorem. *Let (x_0, y_0) lie on the boundary of E. Then the outward unit normal vector to E at (x_0, y_0) is given by $\widehat{n}(x_0, y_0) = \dfrac{(x_0, \alpha^2 y_0)}{|(x_0, \alpha^2 y_0)|}$. The ball of radius α^2 centred at $(x_0, y_0) - \alpha^2 \widehat{n}(x_0, y_0)$ lies entirely inside the ellipse E and touches the boundary at (x_0, y_0). These two facts are unchanged by scaling and translation.*

Proof. The normal vector to E is given by the gradient of $\dfrac{x^2}{\alpha^2} + y^2$. The radius of curvature of E at a point (x_0, y_0) is given by

$$\rho(x_0, y_0) = \frac{1}{\alpha} \left[\alpha^2 + \left(\frac{1}{\alpha^2} - 1 \right) x_0^2 \right]^{3/2}.$$

This has a minimum value of α^2 when $x_0 = 0$. Since E is smooth and convex, the circle of curvature lies inside E, as claimed. Scale and translation invariance are obvious (but true). ////

7.2 Corollary. *For any* (x_0, y_0) *obeying* $\frac{x_0^2}{\alpha^2} + y_0^2 = r^2$ *and any* (x, y) *with* $x^2 + y^2 \le 1$, *one has*

$$\frac{1}{\alpha^2}\left[\left(1 - \frac{\alpha^2 r}{|(x_0, \alpha^2 y_0)|}\right)^2 x_0^2 + 2\alpha^2 r\left(1 - \frac{\alpha^2 r}{|(x_0, \alpha^2 y_0)|}\right)|x_0 x| + \alpha^4 r^2 x^2\right] + \left[y_0 - \frac{\alpha^4 r y_0}{|(x_0, \alpha^2 y_0)|} + \alpha^2 r y\right]^2 \le r^2.$$

Proof. The analytic inequality corresponding to the second statement of Theorem 7.1 is

$$\frac{1}{\alpha^2}\left[x_0 - \frac{\alpha^2 r x_0}{|(x_0, \alpha^2 y_0)|} + \alpha^2 r x\right]^2 + \left[y_0 - \frac{\alpha^4 r y_0}{|(x_0, \alpha^2 y_0)|} + \alpha^2 r y\right]^2 \le r^2.$$

Expanding the first term leads to the desired inequality when one observes that it is valid for $-x$ as well as for x. ////

Chapter III. State Constraints in Optimal Control

In this chapter we study a differential inclusion problem in which the admissible arcs must satisfy not only the usual differential and endpoint constraints, but also a *state constraint* of the form

$$g(t, x(t)) \leq 0 \quad \forall t \in [0, T] \text{ a.e.}$$

Here g is a function with values in some finite-dimensional space $\mathbf{R}^a$, and the relation $g \leq 0$ holds iff each component of g is nonpositive. Infinite-dimensional perturbations of the state constraint give rise to a "value function" $V\colon L^2([0,T], \mathbf{R}^a) \to \mathbf{R} \cup \{+\infty\}$; proximal normal analysis then leads to a new form of the maximum principle for state-constrained problems. The necessary conditions presented here say more about an optimal arc than do the appropriate special cases of Vinter and Pappas (1982) or Clarke (1983). The price to be paid for this improvement is twofold: first, a constraint qualification called "calmness" (valid in many cases) is required; second, the function g must be smooth.

Section 1. The Value Function

For each $\alpha \in L^2([0,T], \mathbf{R}^a)$, we consider the following differential inclusion problem $P(\alpha)$:

$$P(\alpha) \qquad \begin{aligned} \min\{\ell(x(T)) \,:\, & x(0) \in C_0, \\ & \dot{x}(t) \in F(t, x(t)) \text{ a.e. } [0,T], \\ & g(t, x(t)) + \alpha(t) \leq 0 \text{ a.e. } [0,T]\}. \end{aligned}$$

The *value function* $V\colon L^2([0,T], \mathbf{R}^a) \to \mathbf{R} \cup \{+\infty\}$ is defined by $V(\alpha) := \inf P(\alpha)$. No matter how smooth ℓ, F, g, and C_0 may be, we must expect V to be very badly behaved. For example, suppose $V(0) < +\infty$ (so that $P(0)$ has an admissible arc) and that $g(t,x) = x$. Under very mild assumptions on C_0 and F, it follows that there is a constant $M > 0$, independent of α, such that every admissible arc x for $P(\alpha)$ obeys $\|x\|_\infty \le M$. Now it is easy to find a function $\alpha \in L^2$ with arbitrarily small L^2-norm such that $\alpha \ge M$ on a set of positive measure. For such an α, problem $P(\alpha)$ has no admissible arcs and $V(\alpha) = +\infty$. So we must be prepared to deal with the possibility, even if $V(0) < +\infty$, that every neighbourhood of 0 in L^2 contains points α where $V(\alpha) = +\infty$. Even though the data of the problem are smooth, the resulting value function may be astonishingly discontinuous.

Let Y denote the collection of all arcs $x(\cdot)$ solving $P(0)$. (We will shortly show that $Y \neq \emptyset$.) The following hypotheses will be used in this chapter:

(H1) The multifunction $F : [0,T] \times \mathbf{R}^n \to \mathbf{R}^n$ has nonempty compact convex values. For each fixed $x \in \mathbf{R}^n$, $F(\cdot, x)$ is measurable.

(H2) There is a function $k(t) \in L^2[0,T]$ such that

(a) $F(t,x) \subseteq k(t)\overline{B}$ $\forall t \in [0,T]$, $x \in \mathbf{R}^n$,

(b) for each fixed $t \in [0,T]$ and $x \in \mathbf{R}^n$, one has

$$F(t,y) \subseteq F(t,x) + k(t)\,|y-x|\,\overline{B} \qquad \forall y \in \mathbf{R}^n.$$

We define $K_F = \exp\left(\int_0^T k(t)\,dt\right)$.

(H3) The set C_0 constraining the initial point is compact.

(H4) The terminal cost $\ell: \mathbf{R}^n \to \mathbf{R}$ is Lipschitz of rank K_ℓ on $\mathbf{R}^n$.

(H5) The state constraint function $g: [0,T] \times \mathbf{R}^n \to \mathbf{R}^a$ is Lebesgue measurable in t and continuously differentiable in x, with $|g_x(t,x)| \le K_g$ for all $(t,x) \in [0,T] \times \mathbf{R}^n$. Moreover, $\int_0^T |g(t,x)|^2\,dt < +\infty$ $\forall x \in \mathbf{R}^n$.

Hypotheses (H1)–(H5) allow the existence of solutions to $P(\alpha)$ to be proven by the direct method.

1.1 Theorem. *(a) Let a sequence x_i of F-trajectories be given, with $x_i(0) \in C_0$ $\forall i$. Then $\{x_i\}$ has a subsequence converging uniformly to an F-trajectory x such that $x(0) \in C_0$.*

(b) If there is a sequence $\alpha_i \in L^2([0,T], \mathbf{R}^a)$ for which the sequence x_i in (a) also obeys $g(t, x_i(t)) + \alpha_i(t) \le 0$ a.e. $\forall i$, and if $\alpha_i \xrightarrow{w} \alpha$ for some α, then along the subsequence in (a), one also obtains

$$g(t, x(t)) + \alpha(t) \le 0 \quad \text{a.e.}$$

(c) If $V(\alpha) < +\infty$ for some α, then $P(\alpha)$ has a solution.

(d) The function V is weakly sequentially lower semicontinuous. (In particular, V is norm-lower semicontinuous.)

Proof. (a) This is a Corollary to Clarke (1983), Thm. 3.1.7, p. 118.

(b) According to (H5), the difference

$$\sup_t |g(t, x_i(t)) - g(t, x(t))| \leq K_g \|x_i - x\|_\infty$$

tends to 0 along the subsequence described in (a). Hence the functions $g(t, x_i(t)) + \alpha_i(t)$ converge weakly in L^2 to the function $g(t, x(t)) + \alpha(t)$. Now each element of this sequence lies in $-P$, where P is the "positive cone" defined by

$$P = \{r(\cdot) \in L^2 \, : \, r(t) \geq 0 \;\; \text{a.e.}\}.$$

The cone P is convex and strongly closed, hence weakly closed. Therefore the limit function $g(t, x(t)) + \alpha(t)$ lies in $-P$ also, as required.

(c) Let x_i be a minimizing sequence of F-trajectories. According to (a) and (b), in which we take $\alpha_i \equiv \alpha$, there is a subsequence along which x_i tends uniformly to an arc x satisfying all the constraints of problem $P(\alpha)$. And by (H4), $\ell(x(T)) = \lim_{i \to \infty} \ell(x_i(T)) = V(\alpha)$. Thus x solves $P(\alpha)$.

(d) Let $\alpha_i \xrightarrow{w} \alpha$ in L^2, with $V(\alpha_i) \to v$. We must show $v \geq V(\alpha)$. If $v = +\infty$ this is trivial, so assume $v < +\infty$. Then (c) gives solutions x_i to $P(\alpha_i)$ for which

$$V(\alpha_i) = \ell(x_i(T)) \quad \text{and} \quad g(t, x_i(t)) + \alpha_i(t) \leq 0 \;\; \text{a.e.}$$

By parts (a) and (b), we have $v = \ell(x(T))$ for some limiting F-trajectory x obeying $g(t, x(t)) + \alpha(t) \leq 0$ a.e. Hence $V(\alpha) \leq v$ as required. ////

Section 2. Proximal Normals

Let $(\beta, -\lambda)$ be proximal normal to the (weakly sequentially-) closed set $\operatorname{epi} V$ at $(\widehat{\alpha}, \widehat{v})$. Then $V(\widehat{\alpha}) \leq \widehat{v} < +\infty$ implies $P(\widehat{\alpha})$ has a solution $\widehat{x}$ with $\widehat{v} \geq \ell(\widehat{x}(T))$. Now for any F-trajectory $x(\cdot)$ starting in C_0, one has

$$V(-g(\cdot, x(\cdot))) \leq \ell(x(T)).$$

Moreover, the inequality is preserved if a nonnegative s is added to the right side and a nonnegative-valued function $r \in L^2$ is subtracted from the argument of V. Thus

$$\Big(-g(\cdot, x(\cdot)) - r(\cdot), \ell(x(T)) + s\Big) \in \operatorname{epi} V.$$

By definition of a proximal normal, there exists $M > 0$ such that

$$\big\langle (\beta, -\lambda), \big(-g(t,x) - r, \ell(x) + s\big) - (\widehat{\alpha}, \widehat{v})\big\rangle \le M \left\| \big(g(t,x) + r, \ell(x) + s\big) - (\widehat{\alpha}, \widehat{v}) \right\|^2. \tag{2.1}$$

First put $x = \widehat{x}$: then one finds

$$0 \le \langle \beta, r + g(t, \widehat{x}) + \widehat{\alpha} \rangle + \lambda\big[s + \ell(\widehat{x}) - \widehat{v}\big] + M \left\| r + g(t, \widehat{x}) + \widehat{\alpha} \right\|^2 + M \left| s + \ell(\widehat{x}) - \widehat{v} \right|^2. \tag{2.2}$$

Next put $r = -\widehat{\alpha} - g(t, \widehat{x})$ to obtain

$$0 \le \lambda\big[s + \ell(\widehat{x}) - \widehat{v}\big] + M \left| s + \ell(\widehat{x}) - \widehat{v} \right|^2 \qquad \forall s \ge 0. \tag{2.3}$$

This implies that $\widehat{s} = \widehat{v} - \ell(\widehat{x}(T))$ gives a global minimum to the (smooth) right-hand side over the set $\big[0, +\infty\big)$. Hence the right derivative of this expression must vanish if $\widehat{s} > 0$, or at least be nonnegative if $\widehat{s} = 0$. We write this as follows.

$$\lambda \ge 0, \qquad \lambda\big[\widehat{v} - \ell(\widehat{x}(T))\big] = 0. \tag{2.4}$$

If we now take $s = \widehat{s}$ in (2.2), we obtain

$$0 \le \langle \beta, r - \widehat{r} \rangle + M \left\| r - \widehat{r} \right\|^2 \qquad \forall r \in P, \tag{2.5}$$

where $P = \big\{ r \in L^2 \,:\, r(t) \ge 0 \text{ a.e.} \big\}$ and $\widehat{r}(t) = -\widehat{\alpha}(t) - g\big(t, \widehat{x}(t)\big)$. This statement is the definition of $-\beta \in PN_P(\widehat{r})$. By the proximal normal formula, it follows that $-\beta \in N_P(\widehat{r})$. Finally, since P is closed and convex, N_P is the normal cone in the sense of convex analysis. Therefore (2.5) gives

$$\langle \beta, r - \widehat{r} \rangle \ge 0 \qquad \forall r \in P. \tag{2.6}$$

If we now put $r = \widehat{r}$ and $s = \widehat{s}$ in (2.1), the result is

$$\langle \beta, g(t, \widehat{x}) \rangle + \lambda \ell(\widehat{x}) \le \langle \beta, g(t,x) \rangle + \lambda \ell(x) + M \left\| g(t,x) - g(t, \widehat{x}) \right\|^2 + M \left| \ell(x) - \ell(\widehat{x}) \right|^2. \tag{2.7}$$

This shows that $\widehat{x}$ solves the optimal control problem of minimizing the right-hand side over all F-trajectories x with $x(0) \in C_0$.

We are about to use the inner product in L^2 for the first time. Since it has not been used before, lines (2.4), (2.6), (2.7) remain valid in any Hilbert space of functions G satisfying the mild conditions

(i) $g(\cdot, x(\cdot)) \in G$ for all F-trajectories x;

(ii) P is a closed convex subset of G.

Different choices for G yield different necessary conditions and raise different technical problems in the proofs below. Such difficulties arise mainly in the weak convergence arguments of Section 3 below: the salubrious properties of the weak topology on L^2 explain why we have chosen this space for G.

Here are our conclusions about proximal normality, phrased in terms of the *Hamiltonian*

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

2.1 Theorem. *Let $(\beta, -\lambda)$ be proximal normal to* epi V *at $(\widehat{\alpha}, \widehat{v})$. Define the constant $\widehat{R}$ depending on $(\beta, -\lambda)$ by equation (2.9) below. Then for any $R \geq \widehat{R}$, problem $P(\widehat{\alpha})$ has a solution $\widehat{x}$ to which there corresponds an arc $p\colon [0, T] \to \mathbf{R}^n$ obeying the following conditions.*

(a) $\lambda \geq 0, \ \ \langle \beta, r - \widehat{r}\rangle \geq 0 \quad \forall r \in P,$

(b) $\begin{pmatrix} -\dot{p}(t) + \beta(t)' g_x(t, \widehat{x}(t)) \\ \dot{\widehat{x}}(t) \end{pmatrix} \in \partial H(t, x(t), p(t)) \quad$ a.e. $[0, T],$

(c) $p(0) \in R\sqrt{1 + |p(0)|^2}\,\partial d_{C_0}(x(0)), \quad -p(T) \in \lambda \partial \ell(\widehat{x}(T)).$

Here $\widehat{r}(t) = -\widehat{\alpha}(t) - g(t, \widehat{x}(t))$, $P = \{r \in L^2 \ : \ r(t) \geq 0$ a.e.$\}$, and prime denotes transpose.

Proof. Conclusion (a) is a transcription of lines (2.4) and (2.6). The other two conclusions follow from line (2.7), as we now show. The objective functional minimized by $\widehat{x}$ is

$$\lambda\ell(x(T)) + M\,|\ell(x(T)) - \ell(\widehat{x}(T))|^2 + \int_0^T \beta(t)' g(t, x(t))\, dt + M \int_0^T |g(t, x(t)) - g(t, \widehat{x}(t))|^2 \, dt.$$

This functional can easily be transformed into Mayer form by introducing a new state variable $y(t) \in \mathbf{R}$ obeying

$$\dot{y}(t) = \beta(t)' g(t, x(t)) + M\,|g(t, x(t)) - g(t, \widehat{x}(t))|^2, \qquad y(0) = 0.$$

We deduce that the arc $(\hat{x}(t), \hat{y}(t))$, where $\hat{y}(t) = \int_0^t \beta(r)' g(r, \hat{x}(r))\, dr$, minimizes

$$\hat{\ell}(x(T), y(T)) = \lambda \ell(x(T)) + y(T) + M\,|\ell(x(T)) - \ell(\hat{x}(T))|^2$$

over all $(n+1)$-dimensional trajectories for the multifunction

$$\hat{F}(t, x, y) = F(t, x) \times \left\{\beta(t)' g(t, x) + M\,|g(t, x) - g(t, \hat{x}(t))|^2\right\}$$

obeying $(x(0), y(0)) \in C_0 \times \{0\}$. This is a problem for which Thm. I.4.1 provides necessary conditions. These conditions involve the constants $K_{\hat{\ell}}$ and $K_{\hat{F}}$, which we now estimate. On a sufficiently small tube about the solution $\hat{x}$ (see the remarks following Thm. I.4.1) the Lipschitz rank of the quadratic terms in $\hat{\ell}$ and $\hat{F}$ can be made arbitrarily small, so any choices of these constants obeying the following inequalities will suffice:

$$\begin{aligned} &K_{\hat{\ell}} > \lambda K_\ell + 1, \\ &K_{\hat{F}} > \exp\left(\int_0^T \left[k(t) + |\beta(t)|\, K_g\right] dt\right). \end{aligned}$$

In particular, let us choose

$$\begin{aligned} &K_{\hat{\ell}} = \lambda K_\ell + 2, \\ &K_{\hat{F}} > \exp\left(\int_0^T k(t)\, dt + K_g \sqrt{T}\, \|\beta\|_2\right) = K_F \exp\left(K_g \sqrt{T}\, \|\beta\|_2\right). \end{aligned}$$

Then we obtain the explicit expression

$$\hat{R} = (2\lambda K_\ell + 6)\left(2 + K_F \exp(K_g \sqrt{T}\, \|\beta\|_2)\left[\ln K_F + K_g \sqrt{T}\, \|\beta\|_2\right]\right). \tag{2.9}$$

Now in the terminology of Thm. I.4.1, the endpoint constraint set for the problem we are studying is $\hat{S} = C_0 \times \{0\} \times \mathbf{R}^n \times \mathbf{R}$. This implies $\mu = 1$. The transversality condition (c) of Thm. I.4.1 implies that for some $\varsigma \in \partial \ell(\hat{x}(T))$, one has

$$\begin{aligned} &(p(0), q(0), -p(T), -q(T)) \in \lambda(0, 0, \varsigma, 0) + (0, 0, 0, 1) + \hat{R}\,|(1, E)|\left[\partial d_{C_0}(\hat{x}(0)) \times \overline{B} \times \{(0, 0)\}\right], \\ &E = (0, 0, \lambda\varsigma, 1) - (p(0), q(0), -p(T), -q(T)) = (-p(0), -q(0), 0, 0). \end{aligned}$$

We deduce that

$$\begin{aligned} &p(0) \in \hat{R}\,|(1, E)|\, \partial d_{C_0}(\hat{x}(0)), \\ &-p(T) = \lambda\varsigma, \\ &-q(T) = 1. \end{aligned} \tag{2.10}$$

The Hamiltonian for our auxiliary problem is

$$\begin{aligned}\widehat{H}(t,x,y,p,q) &= \sup\left\{\langle (p,q),(v,w)\rangle \,:\, (v,w)\in \widehat{F}(t,x,y)\right\}\\ &= H(t,x,p) + q\left(\beta(t)'g(t,x) + M\left|g(t,x)-g(t,\widehat{x}(t))\right|^2\right).\end{aligned}$$

Since $\widehat{H}$ is independent of y, the costate q is constant; we find $q \equiv -1$. Thus $E = (-p(0),1,0,0)$ and $|(1,E)| = \sqrt{2+|p(0)|^2}$. Conclusion (c) now follows from (2.10).

The Hamiltonian inclusion for the auxiliary problem leads to

$$\begin{aligned}\begin{pmatrix}-\dot{p}(t)\\ \dot{\widehat{x}}(t)\end{pmatrix} &\in \partial_{(x,p)}\left[H(t,x,p) - \beta(t)'g(t,x) - M\left|g(t,x)-g(t,\widehat{x}(t))\right|^2\right]\Big|_{(t,\widehat{x}(t),p(t))}\\ &= \partial H(t,\widehat{x}(t),p(t)) - \left(\beta(t)'g_x(t,\widehat{x}(t)),0\right).\end{aligned}$$

Conclusion (b) follows from this. ////

The second statement of conclusion (a) can be regarded as a complementary slackness condition on the function β. The following result makes this precise.

2.2 Theorem. *Let $\beta \in L^2[0,T]$ and $\widehat{r} \in P$ be given. The following are equivalent:*

(a) $\langle \beta, r-\widehat{r}\rangle \geq 0 \quad \forall r \in P$,

(b) $\beta(t) \geq 0$ a.e., $\quad \beta(t)'\widehat{r}(t) = 0$ a.e.

Proof. $(a \Rightarrow b)$ Let $b(t) = \min\{\beta_i(t) \,:\, i = 1,2,\ldots,a\}$. Define $E = \{t \,:\, b(t) < 0\}$ and

$$r(t) = \begin{cases}\widehat{r}(t) & \text{if } t \notin E,\\ \widehat{r}(t) + e_i & \text{if } t \in E \text{ and } i = \min\arg\min_i\{\beta_i(t)\}.\end{cases}$$

Then (a) implies

$$0 \leq \int_0^T \beta(t)'\left(r(t)-\widehat{r}(t)\right)dt = \int_E b(t)\,dt \leq 0,$$

so indeed $m(E) = 0$.

Next, consider $F = \{t \,:\, \beta(t)'\widehat{r}(t) > 0\}$. Since $\widehat{r}(t) \geq 0$ a.e. by assumption and $\beta(t) \geq 0$ a.e. is now known, we have $\beta(t)'\widehat{r}(t) \geq 0$ a.e. Consider

$$r(t) = \tfrac{1}{2}\widehat{r}(t)I_F(t) + \widehat{r}(t)I_{[0,T]\setminus F}(t).$$

Certainly $r(t) \geq 0$ a.e., so (a) implies

$$0 \leq \int_0^T \beta(t)'\left(r(t)-\widehat{r}(t)\right)dt = -\tfrac{1}{2}\int_F \beta(t)'\widehat{r}(t)\,dt \leq 0.$$

Thus $m(F) = 0$ and (b) is established.

($b \Rightarrow a$) If (b) holds, then certainly $\int_0^T \beta(t)'\hat{r}(t)\,dt = 0$, so

$$\langle \beta, r - \hat{r} \rangle = \int_0^T \beta(t)' r(t)\,dt.$$

For any $r \in P$, the integrand is nonnegative almost everywhere, whence the integral cannot be negative. ////

Section 3. Convergence

Suppose that $V(0) < +\infty$, and that $(\beta, -\lambda)$ is obtained as the weak limit of a bounded sequence of proximal normals $(\beta_i, -\lambda_i)$ to $\operatorname{epi} V$ at base points (α_i, v_i) converging (strongly) to $(0, V(0))$. Let x_i be the corresponding solutions to $P(\alpha_i)$, with multipliers p_i as in Thm. 2.1.

By Thm. 1.1 we have $x_i \to x$ uniformly along some subsequence, where x solves $P(0)$. Also, of course, $\lambda = \lim_{i\to\infty} \lambda_i$ is nonnegative. Just as in the proof of Thm. 1.1(b), $g(t, x_i(t))$ converges to $g(t, x(t))$ strongly in L^2. In particular, the inner products $\langle \beta_i(\cdot), -\alpha_i(\cdot) - g(\cdot, x_i(\cdot)) \rangle$ converge to $\langle \beta(\cdot), -g(\cdot, x(\cdot)) \rangle$. Since we also have $\langle \beta_i, r \rangle \to \langle \beta, r \rangle$ $\forall r \in P$, it follows from Thm. 2.2 that

$$\langle \beta, r(\cdot) + g(\cdot, x(\cdot)) \rangle \geq 0 \quad \forall r \in P. \tag{3.1}$$

Next, observe that the constants R_i defined in terms of $(\beta_i, -\lambda_i)$ by (2.9) form a bounded sequence. Hence we may use the same number $R = \sup\{R_i \,:\, i \in \mathbf{N}\}$ for each i when applying transversality condition (c) of Thm. 2.1. This condition implies

$$\frac{p_i(0)}{\sqrt{2 + |p_i(0)|^2}} \in R\partial d_{C_0}(x_i(0)) \qquad \forall i.$$

Since the sequence on the left side is bounded, it has a convergent subsequence, whose limit we denote by $p(0)/\sqrt{2 + |p(0)|^2}$. By Prop. I.2.3, it follows that

$$p(0) \in R\sqrt{2 + |p(0)|^2}\,\partial d_{C_0}(x(0)) \subseteq N_{C_0}(x(0)). \tag{3.2}$$

Similarly, we have $-p_i(T) \in \lambda_i \partial \ell(x_i(T))$ with $\lambda_i \to \lambda$, $x_i \to x$, and $\partial \ell(x_i(T)) \subseteq K_\ell \overline{B}$ $\forall i$. Hence along a further subsequence, $p_i(T)$ converges to a point we denote by $p(T)$, for which Prop. I.2.3 gives

$$-p(T) \in \lambda \partial \ell(x(T)). \tag{3.3}$$

Let us show that $\sup_i \|p_i\|_\infty < \infty$. Integrating the first component of the Hamiltonian inclusion (Thm. 2.1(b)) and using the Lipschitz condition on H computed by Clarke (1983), Prop. 3.2.4, we find

$$p_i(t) - p_i(T) + \int_t^T \beta_i(s)' g_x(s, x_i(s))\, ds \in \int_t^T k(s)\, |p_i(s)|\, \overline{B}\, ds$$

$$\implies \qquad |p_i(t)| \le \lambda_i K_\ell + T^{\frac{1}{2}} K_g \left[\int_t^T |\beta_i(s)|^2\, ds \right] + \int_t^T k(s)\, |p_i(s)|\, ds,$$

The first RHS term is uniformly bounded in i because λ_i converges; the second, because β_i is a bounded sequence in L^2. Thus there is a constant C for which

$$|p_i(t)| \le C + \int_t^T k(s)\, |p_i(s)|\, ds.$$

Application of Gronwall's lemma gives

$$|p_i(t)| \le C \exp\left(\int_t^T k(s)\, ds \right) \le C K_F.$$

The right side is independent of i and t as required.

To see that the sequence $\{p_i\}$ is equicontinuous, note that much as above,

$$\begin{aligned} |p_i(t) - p_i(s)| &\le \int_s^t |\beta_i(r)' g_x(r, x_i(r))|\, dr + \int_s^t k(r)\, |p_i(r)|\, dr \\ &\le (t-s)^{\frac{1}{2}} K_g \left[\int_s^t |\beta_i(r)|^2\, dr \right]^{\frac{1}{2}} + M \int_s^t k(r)\, dr. \end{aligned}$$

Here we have written M for the finite number $\sup_i \|p_i\|_\infty$. Again, since β_i converges weakly in L^2, it is a bounded sequence in L^2. Hence the first RHS term is majorized by $K(t-s)^{\frac{1}{2}}$ for some K independent of i. Thus uniform equicontinuity of the family $\{p_i\}$ follows.

Since the family $\{p_i\}$ is uniformly bounded and equicontinuous, it has a subsequence converging uniformly to some continuous function p obeying (3.2) and (3.3).

Consider next the functions $u_i(t) = \beta_i(t)' g_x(t, x_i(t)) - \dot{p}_i(t)$. Hypothesis (H2) and Clarke (1983), Prop. 3.2.4, p. 121, imply that

$$|u_i(t)| \le k(t)\, |p_i(t)|$$

$$\implies \qquad \|u_i\|_2 \le \|k(\cdot)\|_2 \sup_i \|p_i\|_\infty$$

so u_i is a bounded sequence in L^2 which must therefore admit a subsequence converging weakly to some function u. Now the bounded convergence theorem implies that $g_x(t, x_i(t))$ converges strongly in $L^2[0,T]$ to $g_x(t, x(t))$. Hence $\beta_i(t)'g_x(t, x_i(t))$ converges weakly in L^2 to $\beta(t)'g_x(t, x(t))$ and the relationship

$$-\dot{p}_i(t) = u_i(t) - \beta_i(t)'g_x(t, x_i(t))$$

implies that $-\dot{p}_i$ converges weakly in L^2 to $u - \beta' g_x(t, x)$. This allows us to pass to the limit in the relationship

$$p_i(t) = p_i(0) - \int_0^t \left[u_i(s) - \beta_i(s)'g_x(s, x_i(s))\right] ds \qquad \forall t$$

and obtain

$$p(t) = p(0) - \int_0^t \left[u(s) - \beta(s)'g_x(s, x(s))\right] ds \qquad \forall t.$$

Hence $p(t)$ is an arc.

Now for each i, conclusion (c) of Thm. 2.1 implies that

$$\begin{pmatrix} -\dot{p}_i(t) + \beta_i(t)'g_x(t, x_i(t)) \\ \dot{x}_i(t) \end{pmatrix} \in \partial H(t, x_i(t), p_i(t)) \text{ a.e.}$$

The proof of Clarke (1983), Thm. 3.1.7, p. 118 shows that this implies

$$\begin{pmatrix} -\dot{p}(t) + \beta(t)'g_x(t, x(t)) \\ \dot{x}(t) \end{pmatrix} \in \partial H(t, x(t), p(t)) \text{ a.e.} \tag{3.4}$$

The following result summarizes conclusions (3.1)–(3.4).

3.1 Theorem. *Assume (H1)–(H5). Let $(\beta, -\lambda)$ be a weak limit of a bounded sequence of proximal normals as described above. Then $P(0)$ has a solution x for which some absolutely continuous function $p(\cdot)$ obeys*

(a) $\lambda \geq 0, \qquad \beta(t) \geq 0$ a.e., $\qquad \beta(t)'g(t, x(t)) = 0$ a.e.

(b) $\begin{pmatrix} -\dot{p}(t) + \beta(t)'g_x(t, x(t)) \\ \dot{x}(t) \end{pmatrix} \in \partial H(t, x(t), p(t))$ a.e.

(c) $p(0) \in N_{C_0}(x(0)), \qquad -p(T) = \lambda \partial \ell(x(T))$.

Section 4. Constraint Qualifications and Necessary Conditions

Conclusions (a)–(c) of Thm. 3.1 are those we wish to propose as a new set of necessary conditions for problem $P(0)$. However, Thm. 3.1 does not immediately justify this because it contains no nontriviality condition. Lau's nearest point theorem (see Section II.2) guarantees that many sequences of proximal normal unit vectors exist, but cannot rule out the possibility that all of them converge weakly to zero. If we take $(\beta, -\lambda) = (0, 0)$ in Thm. 3.1 then conclusions (a)–(c) hold trivially for the arc $p \equiv 0$. So the final step in proving necessary conditions for problem $P(0)$ is to demonstrate that *some* sequence of proximal normal unit vectors has a nonzero limit. The proximal normal formula of Chap. II comes in here. All the hard work in that chapter was devoted to proving the inclusion

$$N_{\mathrm{epi}\,V}(0, V(0)) \subseteq R_{\mathrm{epi}\,V}(0, V(0)), \tag{4.1}$$

where R denotes the closed convex cone generated by weak limits of proximal normals. (See Section II.4.) Our labours now bear fruit.

4.1 Theorem. *If $N_{\mathrm{epi}\,V}(0, V(0))$ contains nonzero points, then there is a solution x to problem $P(0)$ for which one can find a scalar $\lambda \geq 0$, a function $\beta \in L^2([0,T], \mathbf{R}^a)$, and an arc $p(\cdot)$ such that $\lambda + \|\beta\|_2 > 0$ and conclusions (a)–(c) of Thm. 3.1 hold.*

Before comparing these conditions to those current in the literature, let us investigate the condition $N_{\mathrm{epi}\,V}(0, V(0)) \neq \{0\}$.

Calmness. Problem P is called *calm at* $\widehat{\alpha}$ if $V(\widehat{\alpha}) < \infty$ and

$$\liminf_{\alpha \to \widehat{\alpha}} \frac{V(\alpha) - V(\widehat{\alpha})}{\|\alpha - \widehat{\alpha}\|} = m > -\infty.$$

Calmness is a well-respected constraint qualification in other settings—the calculus of variations and mathematical programming, for example. See Clarke (1983) for a discussion of these matters. Our present concern is to show that calmness guarantees the nontriviality of the necessary conditions introduced in Thm. 4.1.

4.2 Proposition. *If problem P is calm at 0, then $N_{\operatorname{epi} V}(0, V(0)) \neq \{0\}$. In fact, $\partial V(0) \neq \emptyset$.*

Proof. We will apply Prop. II.5.4, with $H = L^2([0,T], \mathbf{R}^a) \times \mathbf{R}$, $C = \operatorname{epi} V$, and $c = (0, V(0))$. Now $\operatorname{epi} V$ is locally closed near $(0, V(0))$ and this point is a cluster point of $\operatorname{epi} V$. So it suffices to exhibit a cone K with nonempty interior such that for some $\eta > 0$,

$$(\dagger) \qquad \operatorname{epi} V \cap \Big[(0, V(0)) + (K \cap \eta B)\Big] = \{(0, V(0))\}.$$

Our candidate for K is

$$K = \bigcup_{t \geq 0} t\Big[(0, -1) + \varepsilon B\Big],$$

where $\varepsilon > 0$ is chosen so small that $\dfrac{-\sqrt{1-\varepsilon^2}}{\varepsilon} \leq m - 1$. Evidently, $(0, -1) \in \operatorname{int} K$. Moreover, Lemma II.4.2 shows that $(\alpha, v) \in K$ if and only if $v \leq \dfrac{-\sqrt{1-\varepsilon^2}}{\varepsilon} \|\alpha\|$. Thus for any $\eta > 0$, any $(\alpha, v) \in (0, V(0)) + K \cap \eta B$ with $\alpha \neq 0$ obeys

$$(*) \qquad 0 < \|\alpha\| \leq \eta \quad \text{and} \quad \frac{v - V(0)}{\|\alpha\|} \leq \frac{-\sqrt{1-\varepsilon^2}}{\varepsilon} \leq m - 1.$$

Now by the calmness condition, there exists $\eta > 0$ so small that whenever $(\alpha, v) \in \operatorname{epi} V$, one has

$$(**) \qquad 0 < \|\alpha\| \leq \eta \quad \Longrightarrow \quad \frac{v - V(0)}{\|\alpha\|} \geq \frac{V(\alpha) - V(0)}{\|\alpha\|} \geq m - \frac{1}{2}.$$

Clearly no point (α, v) can obey both $(*)$ and $(**)$, so $(\dagger)$ follows. According to Prop. II.5.4, $N_{\operatorname{epi} V}(0, V(0)) \neq \{0\}$.

In fact, Prop. II.5.4 says more than this. It affirms that for any $\delta \in (0, \varepsilon)$ and any sufficiently small value of t, the open convex set

$$t\big[(0, -1) + \delta B\big] \subseteq \operatorname{int} K \subseteq H$$

is disjoint from the closed convex set $T_{\operatorname{epi} V}(0, V(0))$. Hence (Rudin (1973), Thm. III.3.4(a), p. 58) there is a unit vector $(\beta, -\lambda)$ in H for which

$$\sup \langle (\beta, -\lambda), T_{\operatorname{epi} V}(0, V(0)) \rangle \leq \inf \big\langle (\beta, -\lambda), t\big[(0, -1) + \delta B\big] \big\rangle.$$

Noting that $0 \in T_{\operatorname{epi} V}(0, V(0))$ always holds and evaluating the right side gives

$$0 \leq \sup \langle (\beta, -\lambda), T_{\operatorname{epi} V}(0, V(0)) \rangle \leq t(\lambda - \delta).$$

This clearly implies $\lambda \geq \delta > 0$. But since $T_{\mathrm{epi}\,V}(0, V(0))$ is a cone, the middle expression must continue to satisfy the indicated inequality when multiplied by any positive number ρ. Thus it cannot be positive, and the equation

$$\sup \langle (\beta, -\lambda), T_{\mathrm{epi}\,V}(0, V(0)) \rangle = 0$$

implies $(\beta, -\lambda) \in N_{\mathrm{epi}\,V}(0, V(0))$. It follows that $\beta/\lambda \in \partial V(0)$. ////

Although the calmness condition is difficult to verify without further information about the specific problem under investigation, it is possible to say confidently that "many problems are calm." Indeed, the following result implies that whenever ℓ, F, and g are given satisfying (H1)–(H5), the set of α for which P is calm at α is dense in

$$\mathrm{Dom}\, V = \left\{ \alpha \in L^2 \; : \; V(\alpha) < +\infty \right\}.$$

It is quoted from Aubin and Ekeland (1984), Thm. VI.6.5, p. 281.

4.3 Proposition. *Let X be a smooth Banach space and $V : X \to \mathbf{R} \cup \{+\infty\}$ a lower semicontinuous function on X. For any $\varepsilon > 0$, there is a dense subset of* $\mathrm{Dom}\, V$ *in which each $\widehat{\alpha}$ obeys the following: there is a continuous linear functional $\beta \in X^*$ and a scalar $\eta > 0$ such that*

$$V(\alpha) - V(\widehat{\alpha}) \geq \langle \beta, \alpha - \widehat{\alpha} \rangle - \varepsilon \left\| \alpha - \widehat{\alpha} \right\| \qquad \forall \alpha \in \widehat{\alpha} + \eta B.$$

In our setting $X = L^2$ is certainly a smooth space, and a continuous linear functional $\beta \in X^*$ can be identified with an element of L^2. The key inequality of Prop. 4.3 becomes

$$\frac{V(\alpha) - V(\widehat{\alpha})}{\|\alpha - \widehat{\alpha}\|} \geq \left\langle \beta, \frac{\alpha - \widehat{\alpha}}{\|\alpha - \widehat{\alpha}\|} \right\rangle - \varepsilon \geq - \|\beta\| - \varepsilon \qquad \forall \alpha \in \widehat{\alpha} + \eta (B \setminus \{0\}).$$

Thus P is calm at $\widehat{\alpha}$.

Multiple Solutions. If problem $P(0)$ has a unique solution x then the conclusions of Thm. 4.1 provide new necessary conditions which x must obey. The only prerequisite on the data is the requirement that $N_{\mathrm{epi}\,V}(0, V(0)) \neq \{0\}$. However, if the set Y of solutions to $P(0)$ contains more than one element, this condition only guarantees that *some* solution of $P(0)$ satisfies the new necessary conditions. To assert that *every* solution of $P(0)$ has a nontrivial multiplier, we modify the problem somewhat. Let $\hat{x}$ be any fixed element of Y, and define a problem $\hat{P}(\alpha)$ and a value function $\hat{V}(\alpha)$ by

$$\begin{aligned}\hat{V}(\alpha) := \min\{\ell(x(T)) + |y(T)|^2 \ : x(0) \in C_0, \quad & \dot{x}(t) \in F(t, x(t)) \ \text{ a.e.},\\ y(0) = 0, \quad & \dot{y}(t) = |x(t) - \hat{x}(t)|^2 \ \text{ a.e.},\\ & g(t, x(t)) + \alpha(t) \le 0 \ \text{ a.e.}\}.\end{aligned}$$

Note that the admissible arcs for $\hat{P}(\alpha)$ are the same as those for $P(\alpha)$, and that the modified objective function majorizes the original one. Hence $\hat{V}(\alpha) \ge V(\alpha) \ \ \forall \alpha$. Clearly, $\hat{x}$ is the unique solution of $\hat{P}(0)$, and in particular $\hat{V}(0) = V(0)$. We wish to use Thm. 4.1 to find necessary conditions for $\hat{x}$. The prerequisite for this is that $N_{\operatorname{epi}\hat{V}}(0, V(0)) \ne \{0\}$. Unfortunately, this is not an obvious consequence of $N_{\operatorname{epi}V}(0, V(0)) \ne \{0\}$, even though $\hat{V}(\alpha) \ge V(\alpha) \ \ \forall\alpha$. (For instance, it is not necessary that $N_{\operatorname{epi}\hat{V}}(0, V(0)) \supseteq N_{\operatorname{epi}V}(0, V(0))$—the reader is invited to find continuous functions $V, \hat{V}: \mathbf{R} \to \mathbf{R}$ such that $\hat{V}(\alpha) \ge V(\alpha) \ \ \forall\alpha$, $\hat{V}(0) = V(0)$, and $N_{\operatorname{epi}\hat{V}}(0, V(0)) \cap N_{\operatorname{epi}V}(0, V(0)) = \{0\}$.) So this program will only succeed if we strengthen the constraint qualification. Calmness will do.

4.4 Theorem. *Suppose P is calm at 0. Then for any $x \in Y$ there is a nonzero $(\beta, -\lambda) \in L^2 \times \mathbf{R}$ and an arc $p(\cdot)$ such that*

(a) $\lambda \ge 0, \qquad \beta(t) \ge 0$ a.e., $\qquad \beta(t)' g(t, x(t)) = 0$ a.e.

(b) $\begin{pmatrix} -\dot{p}(t) + \beta(t)' g_x(t, x(t)) \\ \dot{x}(t) \end{pmatrix} \in \partial H(t, x(t), p(t))$ a.e.

(c) $p(0) \in N_{C_0}(x(0)), \qquad -p(T) = \lambda \partial \ell(x(T))$.

Proof. Fix any $\hat{x} \in Y$, and consider problem $\hat{P}(\alpha)$ defined above. Since $\hat{V}(\alpha) \ge V(\alpha)$ for all α and $\hat{V}(0) = V(0)$, the calmness of problem P at 0 implies the calmness of problem $\hat{P}$ at 0. Hence $N_{\operatorname{epi}\hat{V}}(0, V(0)) \ne \{0\}$, and Thm. 4.1 gives a nonzero $(\beta, -\lambda) \in L^2 \times \mathbf{R}$ and an arc $(p(\cdot), q(\cdot))$ satisfying certain conditions. These conditions imply that $q \equiv 0$, and that $p(\cdot)$ obeys the desired conclusions (a)–(c). ////

Sensitivity Analysis. In addition to allowing the derivation of necessary conditions, Thm. 3.1 lends itself to an analysis of the marginal effects of perturbations to the state constraint.

4.5 Definition. Let x be an admissible F-trajectory and λ a nonnegative scalar. Then the pair $(p, \beta) \in AC([0,T], \mathbf{R}^n) \times L^2([0,T], \mathbf{R}^a)$ is an *index λ multiplier corresponding to x* if conclusions (a)–(c) of Thm. 4.4 hold. The collection of all such pairs is denoted $M^\lambda(x)$, and $M^\lambda(Y)$ is the union of the sets $M^\lambda(x)$ over $x \in Y$. We define a mapping Δ from the space of multipliers to L^2 via $\Delta(p, \beta) = \beta$.

4.6 Theorem. *If $V(0) < +\infty$ and (H1)–(H5) hold, then $Y \neq \emptyset$ and*

$$\partial V(0) = \overline{\mathrm{co}}\Big(\Delta\big[M^1(Y)\big] \cap \partial V(0) + \Delta\big[M^0(Y)\big] \cap \partial^\infty V(0)\Big).$$

Proof. Similar to the first part of Thm. I.3.6. ////

Note that Thm. 4.6 holds trivially if $\partial V(0) = \emptyset$, so the only way to get useful information from it is to introduce conditions excluding this possibility. As Prop. 4.2 shows, the calmness condition is sufficient to do this.

Section 5. Comparison to Known Conditions

The necessary conditions of Vinter and Pappas (1982) or of Clarke (1983), Thm. 3.2.6 are as follows. For simplicity, we take the state constraint dimension $a = 1$.

5.1 Theorem. *Assume (H1)–(H5), and assume moreover that $g(t,x)$ is lower semicontinuous in t. If x solves $P(0)$, there exist a scalar $\lambda \geq 0$, a nonnegative Radon measure μ, and an arc q such that*

(a) μ *is supported on the set* $S = \{t \in [0,T] \,:\, g(t, x(t)) = 0\}$.

(b) $\begin{pmatrix} -\dot q(t) \\ \dot x(t) \end{pmatrix} \in \partial H\left(t, x(t), q(t) + \int_{[0,t)} g_x(s, x(s))\, \mu(ds)\right)$.

(c) $q(0) \in N_{C_0}(x(0)), \qquad -q(T) - \int_{[0,T]} g_x(s, x(s))\, \mu(ds) \in \lambda \partial \ell(x(T))$.

We can obtain these conclusions from Thm. 4.4 by defining a new arc $q(\cdot)$ as follows:

$$q(t) = p(t) - \int_0^t g_x(s, x(s))\beta(s)\, ds.$$

The conclusions of Thm. 4.4 then become

(a) $\lambda \geq 0, \qquad \beta(t) \geq 0$ a.e., $\qquad \beta(t)' g(t, x(t)) = 0$ a.e.

(b) $\begin{pmatrix} -\dot{q}(t) \\ \dot{x}(t) \end{pmatrix} \in \partial H\left(t, x(t), q(t) + \int_0^t g_x(s, x(s))\beta(s)\, ds\right).$

(c) $q(0) \in N_{C_0}(x(0)), \qquad -q(T) - \int_0^T g_x(s, x(s))\beta(s)\, ds \in \lambda\partial\ell(x(T)).$

The advantages of Thm. 4.4 now become clear: it shows that the measure μ appearing in the known conditions can be assumed to be absolutely continuous with respect to Lebesgue measure, with a square-integrable density β. We obtain this desirable conclusion by an equally attractive method of proof, based on the geometrical structure of a certain epigraph intimately related to the state constraint itself. Moreover, g is allowed to be merely measurable in t. The cost of these advances is a constraint qualification (calmness) which is not assumed in the standard results cited above, but which is known to hold arbitrarily near any problem of interest. Smoothness of g in x, a condition many problems obey, is a requirement of our theory but is merely a special case of the results cited above.

Chapter IV: Existence Theory for a Stochastic Bolza Problem

The previous chapters have confirmed the intimate connection between sensitivity to perturbations and necessary conditions in optimization problems. The analysis itself, however, makes frequent and essential use of limiting arguments—both in proving vital existence theorems and in the proximal normal formula itself. For this reason, any attempt to study the sensitivity of a stochastic control problem with proximal normals must begin with an investigation of appropriate limiting techniques. In this chapter we introduce the techniques of convergence in distribution, tightness, and martingale representation by showing them at work in a new existence theorem for a constrained Bolza problem of stochastic optimal control.

These methods have been used before to study the existence of optimal stochastic control laws, but never in such generality as that of Thm. 5.1 below. Our work owes its basic approach to Kushner (1975), but it improves on his result by allowing the control set U to be unbounded, by treating objective functionals of Bolza form, and by allowing the incorporation of soft constraints. We also invoke stronger martingale representation theorems, and thereby weaken some of Kushner's technical hypotheses. Our method of proof also has the pedagogical advantage of replacing the use of Skorokhod's theorem with a well-known closure theorem from deterministic optimal control which clarifies the connection between deterministic and stochastic existence theories. Finally, our hypotheses are stated more explicitly than Kushner's, making them easier to verify in practice.

This chapter reviews the required probability theory in Sections 1–3 and formulates the stochastic control problem precisely in Section 4. The main existence theorem is proven in Section 5, after which Section 6 adds several significant extensions. Section 7 is devoted to a comparison between the main existence theorem and its counterpart involving a compact control set.

Section 1. The Probability Background

Standard textbooks on the general theory of stochastic processes like Jacod (1979), Dellacherie-Meyer (1975), and Ikeda-Watanabe (1981) present the extensive theoretical basis for our Chapters IV and V in detail. Our aim in this section is simply to collect those aspects of the theory which are critical to the development below. We begin with a summary of commonly used notation.

Notation.

$[0,T]$ denotes a fixed time interval used throughout this chapter; $\forall t$ means $\forall t \in [0,T]$.

S denotes a metric space; S is *Polish* if it is complete and separable.

$\mathcal{B}(S)$ denotes the *Borel σ-field* of S. ($\mathcal{B}(S) = \sigma\{U \subseteq S : U \text{ is open.}\}$.)

$\mathcal{B}^m$ denotes $\mathcal{B}(\mathbf{R}^m)$.

C_t^m denotes $C([0,t], \mathbf{R}^m)$.

$\mathcal{C}_t^m$ denotes $\mathcal{B}(C_t^m)$, completed with respect to Wiener measure. See text below.

$C^m, \mathcal{C}^m$ denote C_T^m, $\mathcal{C}_T^m$, respectively.

$s \wedge t$ denotes $\min\{s,t\}$.

$x^t(s)$ denotes $x(s \wedge t)$, when $x(\cdot)$ is a function.

$\mathcal{F} \times \mathcal{S}$ denotes the product σ-field, when two σ-fields $\mathcal{F}$ and $\mathcal{S}$ are given.

We write $x: (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ if the mapping $x: \Omega \to S$ is measurable with respect to the σ-fields $\mathcal{F}$ and $\mathcal{S}$. When $\mathcal{S} = \mathcal{B}(S)$, we sometimes express this by writing $x \in \mathcal{F}$.

Stochastic Processes. A *filtered probability space* $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ consists of a set Ω equipped with a σ-field $\mathcal{F}$ on which a probability measure P is defined. The *filtration* $\{\mathcal{F}_t : t \in [0,T]\}$ is a family of sub-σ-fields of $\mathcal{F}$ obeying

$$\mathcal{F}_s \subseteq \mathcal{F}_t \quad \text{whenever } 0 \leq s \leq t \leq T.$$

The usual hypotheses regarding such a filtered space are the following three conditions:

(i) The measure space $(\Omega, \mathcal{F}, P)$ is complete.

(ii) The filtration $\mathcal{F}_t$ is right-continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$ $\forall t$, where $\mathcal{F}_{t+} = \bigcap_{s>0} \mathcal{F}_{t+s}$.

(iii) $\mathcal{F}_0 \supseteq \{A \in \mathcal{F} : P(A) = 0\}$.

An S-valued *stochastic process* is a measurable mapping $x: [0,T] \times \Omega \to S$; the process x is called $\mathcal{F}_t$-adapted if $x(t,\cdot)$ is $\mathcal{F}_t$-measurable for each $t \in [0,T]$. It is conventional to simplify the notation involving the process x by leaving the ω-dependence implicit whenever possible. Thus x_t or $x(t)$ are often used to denote either the mapping $x(t,\cdot)$ or one of its values $x(t,\omega)$—context distinguishes the two possibilities. A given process x_t defines a filtration $\{\mathcal{F}_t^x : t \in [0,T]\}$ as follows:

$$\mathcal{F}_t^x = \sigma\Big(\{\{x_s \in A\} : s \leq t,\ A \in \mathcal{B}(S)\} \cup \{A \in \mathcal{F} : P(A) = 0\}\Big).$$

We say that a second process y is x_t-adapted if it is $\mathcal{F}_t^x$-adapted.

An important σ-field on the product space $[0,T] \times \Omega$ is the *predictable σ-field,* denoted by $\mathcal{P}$, which is the σ-field generated by all continuous $\mathcal{F}_t$-adapted processes defined on $[0,T] \times \Omega$. An S-valued process x is called *predictable* if it is $\mathcal{P}$-measurable when considered as a mapping of $[0,T] \times \Omega$ into $(S, \mathcal{B}(S))$. Most of the processes to be discussed below are continuous and hence predictable; however, we need this terminology for the statement of certain key results in Section 3.

Brownian Motion. Let a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual hypotheses be given. An $\mathcal{F}_t$-adapted stochastic process w with values in $\mathbf{R}^d$ is an *$\mathcal{F}_t$-Brownian motion* if it obeys conditions (i)–(iii) below.

(i) $w_0(\omega) = 0$ for all $\omega \in \Omega$;

(ii) the $\mathbf{R}^d$-valued random vector $w_t - w_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s < t$;

(iii) the $\mathbf{R}^d$-valued random vector $w_t - w_s$ has a Gaussian distribution with mean 0 and covariance matrix $(t-s)I$ for any $0 \leq s < t \leq T$;

(iv) for P-almost all ω, the function $w(\cdot,\omega)$ lies in C^d.

The continuity hypothesis in condition (iv) implies that a Brownian motion can equally well be considered as a measurable mapping w from Ω into C^d. Indeed, a particularly important example of Brownian motion is obtained when $(\Omega, \mathcal{F}) = (C^d, \mathcal{C}^d)$ and the probability measure P is *Wiener measure*, denoted by W. In this case the identity map is a Brownian motion, and $(\Omega, \mathcal{F}) = (C^d, \mathcal{C}^d)$ is called *canonical path space*. The (completed) filtration generated by the identity map is simply $\mathcal{C}_t^d = \overline{\mathcal{B}\left(C_t^d\right)}$.

Itô's Integral. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered space satisfying the usual hypotheses. Stochastic integration in the sense of K. Itô is an operation which exchanges one stochastic process for another. If the input is an $\mathbf{R}^{n\times d}$-valued process σ_t adapted to $\mathcal{F}_t$ and obeying

$$\mathbf{E}\int_0^T |\sigma_r|^2\,dr < +\infty, \tag{1.1}$$

then for any $\mathcal{F}_t$-Brownian motion w in $\mathbf{R}^d$, the output is an $\mathbf{R}^n$-valued process denoted by $\int_0^t \sigma_r\,dw_r$. Integral notation for the new process is justified by the many similarities between the operational properties of the stochastic integral and those of conventional measure-theoretic integrals. However, it is not meant to suggest any particular computational strategy. For instance, it is usually incorrect to try to compute $\int_0^t \sigma_r(\omega)\,dw_r(\omega)$ as a Riemann-Stieltjes integral for each fixed ω. (See Fleming and Rishel (1975), p. 112, for a standard counterexample.) This tactic only works on rather simple integrands—for example, those which are simple predictable processes in the technical sense.

We will use the two inequalities below throughout this chapter.

1.1 Proposition (An inequality of Burkholder). *Let w_t be a Brownian motion on $\mathbf{R}^d$, and let σ_t be an $n \times d$-matrix valued process obeying (1.1). Then for any $\delta, \varepsilon > 0$ one has*

$$P\left\{\left\|\int_0^{(\cdot)} \sigma_r\,dw_r\right\|_t > \varepsilon\right\} \le \delta/\varepsilon^2 + P\left\{\int_0^t |\sigma_r|^2\,dr > \delta\right\} \qquad \forall t \in [0,T].$$

1.2 Proposition (Burkholder-Davis-Gundy). *For any exponent $p \in [0,\infty)$ there is a constant C_p, depending only upon p and n, such that for any process σ as in Proposition 1.1,*

$$\mathbf{E}\left\|\int_0^{(\cdot)} \sigma_r\,dw_r\right\|_t^p \le C_p\mathbf{E}\left(\int_0^t |\sigma_r|^2\,dr\right)^{p/2} \qquad \forall t \in [0,T]. \tag{BDG}$$

The Metric Space $Z[0,T]$.

Let $Z = Z[0,T]$ denote the set of all functions $z\colon [0,T] \to \mathbf{R}$ with the properties

(i) $z(0) = 0$,

(ii) $\lim_{t\to s+} z(t) = z(s)$,

(iii) $0 \le s \le t \le T$ implies $z(s) \le z(t)$.

There is a one-to-one correspondence between functions $z \in Z[0,T]$ and finite nonnegative Borel measures defined on the compact set $[0,T]$. Indeed, for a given $z \in Z$ the corresponding finite measure μ_z is the unique Borel measure for which

$$\mu_z(a,b] = z(b) - z(a) \qquad \text{whenever } 0 \leq a < b \leq T.$$

(See Royden (1968), Sect. 12.3.) Thus $Z[0,T]$ may just as well be viewed as the set of all such measures on $[0,T]$. (We will therefore feel free to write $z(E)$ for the μ_z-measure of a Borel set E in the following discussion.) As such, Z may be given the topology of weak convergence of measures. Since $[0,T]$ is a compact metric space, this topology is metrizable: the *Prokhorov metric* is defined by

$$d(y,z) = \inf\{\varepsilon > 0 \,:\, y(A) \leq z(A^\varepsilon) + \varepsilon,\ z(A) \leq y(A^\varepsilon) + \varepsilon,\ \forall A \in \mathcal{B}\}.$$

(Here A^ε denotes the set $\{t \in [0,T] \,:\, \mathrm{dist}(t,A) < \varepsilon\}$.) In fact d makes $Z[0,T]$ into a complete separable metric space. These facts are special cases of the results in Billingsley (1968) Appendix III; they appear more explicitly in Prokhorov (1956). In terms of the increasing functions used to define the space $Z[0,T]$ in the first place, the topology of weak convergence corresponds to pointwise convergence at continuity points (Billingsley (1968), Sect. 3, p. 17). The following proposition summarizes these observations and incorporates the Helly-Bray selection theorem.

1.3 Proposition. *$Z[0,T]$ is a complete separable metric space in which a sequence of functions $\{z^k\}$ converges to z if and only if $z^k(t) \to z(t)$ for each point t where z is continuous, and for $t = T$. A closed subset S of $Z[0,T]$ is compact if and only if there is a constant $M > 0$ such that $z(T) \leq M$ for all $z \in S$.*

Section 2. Convergence in Distribution

When proving existence theorems for optimization problems by the "direct method," one seeks to isolate a solution as the limit of a well-chosen minimizing sequence. Proximal normal analysis also relies on limiting arguments. These two considerations motivate a search for the appropriate notion of limits in stochastic optimal control problems. Weak convergence of probability measures appears to be the correct answer: in this section we review this mode of convergence.

2.1 Definition. Let any metric space S be given, and suppose that P and P_k, $k = 1, 2, \ldots$, are probability measures on $(S, \mathcal{B}(S))$. Then the sequence of measures P_k *converges weakly to* P, denoted $P_k \xrightarrow{w} P$, if and only if one of the following three equivalent conditions is satisfied.

(a) $\int_S f(s)\, dP_k(s) \to \int_S f(s)\, dP(s)$ for all bounded, uniformly continuous $f: S \to \mathbf{R}$.

(b) $\limsup_{k \to \infty} P_k(F) \leq P(F)$ for all closed sets $F \subseteq S$.

(c) $\lim_{k \to \infty} P_k(A) = P(A)$ for all sets $A \in \mathcal{B}(S)$ such that $P(\operatorname{bdy} A) = 0$.

(The equivalence of conditions (a)–(c) is proven in Billingsley (1968), p. 11.)

It is clear that the weak limit of a sequence of probability measures is unique.

Weak convergence of measures is intimately related to the notion of "tightness." A family Π of probability measures on $(S, \mathcal{B}(S))$ is *tight* if for every $\varepsilon > 0$ there is a compact set $K \subseteq S$ for which every P in Π obeys $P(K) \geq 1 - \varepsilon$. The following theorem, proven in Billingsley (1968), Section 6, p. 35, shows why tightness is so important.

2.2 Theorem (Prokhorov). *Let S be a Polish space on which a family Π of probability measures is given. Then Π is tight if and only if every sequence chosen from Π has a weakly convergent subsequence.*

Whenever a random element $x: (\Omega, \mathcal{F}, P) \to (S, \mathcal{B}(S))$ is given, a measure P_x is induced on $(S, \mathcal{B}(S))$ as follows:

$$P_x(A) = \left(P \circ x^{-1}\right)(A) = P\{\omega \,:\, x(\omega) \in A\} \qquad \forall A \in \mathcal{B}(S).$$

This observation sets up a correspondence between random elements and probability measures which allows the preceding notions to be reworded as follows. Suppose $x_\iota: (\Omega_\iota, \mathcal{F}_\iota, P_\iota) \to (S, \mathcal{B}(S))$, $\iota \in I$, is a family of random elements of S. This family is *tight* if for every $\varepsilon > 0$ there is a compact subset K of S such that

$$\left(P_\iota \circ x_\iota^{-1}\right)(K) \geq 1 - \varepsilon \qquad \forall \iota \in I.$$

If $I = \mathbf{N}$, then the sequence x_ι *converges in distribution* to the limit $x: (\Omega, \mathcal{F}, P) \to (S, \mathcal{B}(S))$ if the induced measures $P_\iota \circ x_\iota^{-1}$ converge weakly to $P \circ x^{-1}$. We denote this by $x_\iota \xrightarrow{\mathcal{D}} x$. Prokhorov's theorem takes the following form in this alternate terminology.

2.3 Theorem (Prokhorov). *Let S be a Polish space equipped with a family of random elements*

$$\Xi = \{x_\iota : (\Omega_\iota, \mathcal{F}_\iota, P_\iota) \to (S, \mathcal{B}(S))\}_{\iota \in I}.$$

The family Ξ is tight if and only if every sequence x_i chosen from Ξ has a subsequence converging in distribution to some random element $x: (\Omega, \mathcal{F}, P) \to (S, \mathcal{B}(S))$.

An important consequence of convergence in distribution arises from Def. 2.1(b). If $x_i \xrightarrow{\mathcal{D}} x$ and F is a closed subset of S, then

$$P\{x \in F\} \geq \limsup_{i\to\infty} P_i\{x_i \in F\}.$$

This fact will be used in Section 5.

The fundamental role of tightness makes it important to be able to detect this property. Billingsley (1968), Thm. 8.2, p. 55, gives the following criterion in the space $S = C^n$.

2.4 Proposition. *A sequence $x_i: (\Omega_i, \mathcal{F}_i, P_i) \to (C^n, \mathcal{C}^n)$ of random elements of C^n is tight if and only if the following two conditions hold:*

(i) $\lim\limits_{R\to\infty} \sup\limits_{i} P_i\{|x_i(0)| > R\} = 0,$

(ii) $\lim\limits_{\substack{\delta\to 0^+\\ N\to\infty}} \sup\limits_{i\geq N} P_i\left\{\sup\limits_{\substack{0<t-s<\delta\\ 0\leq s<t\leq T}} |x_i(t) - x_i(s)| > \varepsilon\right\} = 0 \qquad \forall \varepsilon > 0.$

Section 3. Martingales and Their Representations

Let us fix a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual hypotheses throughout this section. An $\mathcal{F}_t$-adapted process m_t taking values in $\mathbf{R}^n$ and obeying $m_0 = 0$ is an *$\mathcal{F}_t$-martingale* if $\mathbf{E}\,|m_t| < +\infty$ $\forall t \in [0, T]$ and if

$$0 \leq s \leq t \leq T \quad \Longrightarrow \quad \mathbf{E}\big[m_t \bigm| \mathcal{F}_s\big] = m_s \ \text{ a.s.} \tag{3.1}$$

Clearly any $\mathcal{F}_t$-martingale is also an $\mathcal{F}_t^m$-martingale.

A *square-integrable martingale* m_t is a martingale which obeys

$$\mathbf{E}\,|m_t|^2 < +\infty \qquad \forall t \in [0, T]; \tag{3.2}$$

if m_t has continuous sample paths, then there is a unique continuous increasing nonnegative definite $n \times n$ matrix-valued process q_t such that $q_0 = 0$ and $(m_t)(m_t)' - q_t$ is a matrix-valued $\mathcal{F}_t$-martingale. The process q_t is called the *quadratic variation of* m_t, and denoted by $\langle m \rangle_t$. For each $i, j = 1, 2, \ldots, n$, let $\langle m^i, m^j \rangle_t$ denote the scalar process defined by the (i,j)-component of $\langle m \rangle_t$. Then $m_t^i m_t^j - \langle m^i, m^j \rangle_t$ is a scalar-valued $\mathcal{F}_t$-martingale.

The following proposition shows that for continuous processes, convergence in distribution preserves the martingale property. It uses the notion of "uniform integrability," for which an excellent reference is Billingsley (1968), pp. 32–33.

3.1 Proposition. *Let* $(x^k, m^k): (\Omega^k, \mathcal{F}^k, P^k) \to (C^{n+\ell}, \mathcal{C}^{n+\ell})$ *be a sequence of random elements of* $C^{n+\ell}$. *Defining* $\mathcal{F}_t^k$ *to be the filtration generated by* (x_t^k, m_t^k), *suppose that* m_t^k *is an* $\mathcal{F}_t^k$*-martingale for each* k. *If* $(x^k, m^k) \xrightarrow{D} (x, m)$ *for some random element* $(x, m): (\Omega, \mathcal{F}, P) \to (C^{n+\ell}, \mathcal{C}^{n+\ell})$ *and if the sequence* $\{\|m^k\|\}$ *is uniformly integrable, then the limit process* m_t *is an* $\mathcal{F}_t$*-martingale. Here* $\mathcal{F}_t$ *is the filtration generated by* (x, m).

Proof. Fix any $N \in \mathbf{N}$ and choose any $0 \le s < t \le T$ and $0 \le t_1 \le t_2 \le \cdots \le t_N \le s$. Let $g: \mathbf{R}^{(Nn+N\ell)} \to \mathbf{R}$ be an arbitrary bounded, uniformly continuous function. Then for each k, the martingale character of m^k implies

$$\mathbf{E}^k\Big[g\left(x^k(t_1), \ldots, x^k(t_N), m^k(t_1), \ldots, m^k(t_N)\right)\left(m^k(t) - m^k(s)\right)\Big] = 0.$$

Now since $\{\|m^k\|\}$ is a uniformly integrable sequence, so is the sequence of real-valued random variables whose expectations are computed above. Moreover, this sequence can be viewed as the image of the sequence (x^k, m^k) under a continuous map $G: C^{n+\ell} \to \mathbf{R}$. Hence the sequence of integrands converges in distribution to $G(x, m)$: by uniform integrability it follows that we can let $k \to \infty$ above to obtain

$$\mathbf{E}\Big[g\left(x(t_1), \ldots, x(t_N), m(t_1), \ldots, m(t_N)\right)\left(m(t) - m(s)\right)\Big] = 0.$$

Since N, g, and $t_1, \ldots, t_N$ are arbitrary, this shows that $\mathbf{E}\big[m_t - m_s \mid \mathcal{F}_s\big] = 0$, as required. ////

Suppose that $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ carries an $\mathcal{F}_t$-Brownian motion w_t with values in $\mathbf{R}^d$. If an $\mathcal{F}_t$-adapted process σ_t with values in $\mathbf{R}^{n\times d}$ is given satisfying (1.1), we can define $\int_0^t \sigma_r \, dw_r$. *This process is a martingale.* In other words,

$$0 \le s \le t \le T \quad \Longrightarrow \quad \mathbf{E}\left[\int_0^t \sigma_r \, dw_r \,\middle|\, \mathcal{F}_s\right] = \int_0^s \sigma_r \, dw_r \quad \text{a.s.}$$

The quadratic variation of this martingale is the $n \times n$ matrix-valued process

$$\left\langle \int_0^{(\cdot)} \sigma_r \, dw_r \right\rangle_t = \int_0^t \sigma_r \sigma_r' \, dr; \tag{3.3}$$

in particular,

$$\mathbf{E}\left|\int_0^t \sigma_r \, dw_r\right|^2 = \mathbf{E}\int_0^t |\sigma_r|^2 \, dr, \tag{3.4}$$

where $|\sigma|^2 = \operatorname{tr}(\sigma\sigma')$. Propositions 1.1 and 1.2 above are actually special cases of inequalities valid for arbitrary martingales.

It is more than a happy coincidence that the process $\int_0^t \sigma_r \, dw_r$ turns out to be a martingale. The relationship between martingales and stochastic integrals is strong enough that a sort of converse to (3.3) is available. In the case $\sigma \equiv I$, it takes the following form:

3.2 Lemma (Doob). *If m_t is an $\mathcal{F}_t$-martingale with continuous $\mathbf{R}^d$-valued sample paths for which $\mathbf{E}\,|m_t|^2 < +\infty \ \forall t \ge 0$ and $\langle m \rangle_t = tI$, then m is an $\mathcal{F}_t$-Brownian motion. The converse is also true.*

Doob's lemma can be generalized substantially. In fact, virtually any continuous martingale whose quadratic variation is absolutely continuous with respect to Lebesgue measure can be represented as a stochastic integral. Let us make this precise. Suppose a continuous $\mathbf{R}^n$-valued martingale m is given on our fixed space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, and that $\langle m \rangle_t = \int_0^t \sigma_r \sigma_r' \, dr$ for some predictable $n \times d$ matrix valued process σ with $d \le n$. In a certain rigorously definable sense, it follows that there is a Brownian motion w in $\mathbf{R}^d$ with respect to which $m_t = \int_0^t \sigma_r \, dw_r$. This is the content of Prop. 3.3, below.

Before stating the representation theorem precisely, let us note that it may be impossible to construct w on the given space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. For example, if $m_t \equiv 0$ and the σ-fields $\mathcal{F}$ and $\mathcal{F}_t$ are all trivial, then there is simply no room for an $\mathcal{F}_t$-Brownian motion of any dimension on this space. To eliminate this possibility, we augment the given space with a copy of canonical d-dimensional Wiener space (C^d, C^d, C_t^d, W) as follows:

$$
\begin{aligned}
(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) &= (\Omega \times C^d, \mathcal{F} \times C^d, P \times W) \\
\tilde{\mathcal{F}}_t &= \bigcap_{h>0} \left(\mathcal{F}_{t+h} \times C_{t+h}^d\right).
\end{aligned} \tag{3.5}
$$

Now $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{P})$ is a filtered space satisfying the usual hypotheses, and any random variable x originally defined on Ω can be readily replaced by a random variable $\tilde{x} = x \circ \pi$ on $\tilde{\Omega}$, where $\pi(\omega, \omega') = \omega$ is the natural projection. Clearly $\tilde{x} \overset{\mathcal{D}}{\equiv} x$.

The fine points of the extension (3.5) are explored in detail by Jacod (1979), Section X.2(b), p. 332. He shows that for the natural embedding $\pi^{-1}(\mathcal{F})$ of the original σ-field $\mathcal{F}$ in $\tilde{\mathcal{F}}$, a random variable $\tilde{x}$ defined on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ is $\pi^{-1}(\mathcal{F})$-measurable if and only if $\tilde{x} = x \circ \pi$ for some random variable x on $(\Omega, \mathcal{F}, P)$ with $x \overset{\mathcal{D}}{\equiv} \tilde{x}$. Also, $\pi^{-1}(\mathcal{F}_t) = \pi^{-1}(\mathcal{F}) \cap \tilde{\mathcal{F}}_t$, and any $\pi^{-1}(\mathcal{F})$-measurable random variable $\tilde{x} = x \circ \pi$ obeys

$$
\mathbf{E}[\tilde{x} \mid \tilde{\mathcal{F}}_t] = \mathbf{E}[\tilde{x} \mid \pi^{-1}(\mathcal{F}_t)] = \mathbf{E}[x \mid \mathcal{F}_t] \circ \pi \qquad \tilde{P}-\text{ a.s.}
$$

Thus all relevant properties of x and $\mathcal{F}_t$ are retained in the passage to $\tilde{x}$, $\tilde{\mathcal{F}}_t$, while the enlarged space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{P})$ most assuredly has room to contain a d-dimensional Brownian motion.

We may now state the promised representation theorem, essentially due to Doob. The version here, which allows $d \leq n$ and imposes no nonnegativity condition on σ, is taken from Jacod (1979), Thm. (14.45), p. 466.

3.3 Proposition. *Let m_t be a continuous $\mathbf{R}^n$-valued $\mathcal{F}_t$-martingale with $\langle m \rangle_t = \int_0^t \sigma_r \sigma_r' \, dr$ for some $\mathcal{F}_t$-predictable $n \times d$ matrix valued process σ with $d \leq n$. Then there is a d-dimensional $\tilde{\mathcal{F}}_t$-Brownian motion $\tilde{w}_t$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{P})$ such that, with $\tilde{m}_t = m_t \circ \pi$ and $\tilde{\sigma}_r = \sigma_r \circ \pi$, one has*

$$
\tilde{m}_t = \int_0^t \tilde{\sigma}_r \, d\tilde{w}_r \qquad \forall t \in [0, T], \ \tilde{P}-\text{ a.s.}
$$

Moreover, if the rank of the matrix σ is identically equal to d, this conclusion remains valid with $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{P}) = (\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and $\pi =$ identity. That is, no extension of the given space is required.

A significant application of Prop. 3.3, first noted by Wong (1971), deals with "quasimartingales." Let x be an $\mathcal{F}_t$-adapted process with cadlag sample paths in $\mathbf{R}^n$. For any partition $0 \leq t_1 < t_2 <$

$\cdots < t_k \leq T$ of $[0,T]$, define

$$\text{var}\,(x;t_1,\ldots,t_k) = \sum_{i=1}^{k} \left|\mathbf{E}\left[x(t_{i+1}) - x(t_i) \mid \mathcal{F}_{t_i}\right]\right| + |x(T)|, \tag{3.6}$$
$$\text{Var}\,(x) = \sup\left\{\mathbf{E}\,\text{var}\,(x;t_1,\ldots,t_k) \,:\, k \in \mathbf{N},\ 0 \leq t_1 < \cdots < t_k \leq T\right\}.$$

The process x is a *quasimartingale* if $\text{Var}(x) < +\infty$. Note that any $\mathcal{F}_t$-martingale m_t is automatically a quasimartingale, since for any partition

$$\text{var}\,(m;t_1,\ldots,t_k) = 0 + |m_T|$$

and $\mathbf{E}\,|m_T| < +\infty$. (In particular, a Brownian motion is a quasimartingale even though almost all its sample functions have unbounded variation on every interval.) The linearity of conditional expectation implies that the sum of a quasimartingale with a martingale remains a quasimartingale. Conversely, we might expect that a typical quasimartingale should be decomposable into the sum of a martingale and some other process. The properties of the other summand are described in the following theorem, originally proven by Fisk, but quoted here from Jacod (1979), Thm (5.36), p. 174.

3.4 Proposition. *A given process x on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a quasimartingale if and only if*

$$x_t - x_0 = m_t + a_t \qquad \forall t \in [0,T],\ P-\ \text{a.s.}, \tag{3.7}$$

where m_t is an $\mathcal{F}_t$-martingale and a_t is a predictable process whose sample paths have finite variation on $[0,T]$ $P-$ a.s. This decomposition is unique.

Jacod (1979), Prop. (9.14), p. 285, also shows that if a given quasimartingale x on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ happens to be $\mathcal{G}_t$-adapted, where $\mathcal{G}_t$ is a filtration satisfying the usual hypotheses and $\mathcal{G}_t \subseteq \mathcal{F}_t\ \forall t$, then x is also a quasimartingale with respect to $(\Omega, \mathcal{F}, \mathcal{G}_t, P)$. Of course, the canonical decomposition of x given in (3.7) may change when one passes to this smaller filtration. Eugene Wong (1971) has given an explicit construction of the $\mathcal{G}_t$-representation of a quasimartingale whose canonical $\mathcal{F}_t$-decomposition takes the form

$$x_t = x_0 + \int_0^t f_r\,dr + m_t. \tag{3.8}$$

We quote his result as Thm. 3.6 below, assuming that f is an $\mathcal{F}_t$-adapted process obeying

$$\mathbf{E}\int_0^T |f_r|\,dr < +\infty \tag{3.9}$$

and that m_t is a continuous second-order $\mathcal{F}_t$-martingale with

$$\mathbf{E}\langle m\rangle_T < +\infty. \tag{3.10}$$

Wong's original proof is phrased somewhat differently, and is valid in the more general case when m is a "locally square integrable martingale." It relies on the following lemma.

3.5 Lemma (Wong). *Let $\mathcal{F}_t$ and $\mathcal{G}_t$ be two filtrations of $(\Omega, \mathcal{F}, P)$ satisfying the usual hypotheses, and suppose that m_t and n_t are continuous second-order $\mathcal{F}_t$- and $\mathcal{G}_t$-martingales. If $m_t - n_t$ is a process whose sample paths have bounded variation, then $\langle m\rangle_t = \langle n\rangle_t$ $\forall t \in [0,T]$* a.s.

3.6 Theorem (Wong). *Let x be an $\mathbf{R}^n$-valued $\mathcal{F}_t$-martingale of the form (3.8) satisfying (3.9) and (3.10). Suppose that there is an $\mathcal{F}_t$-predictable $n \times d$ matrix valued process σ with $d \leq n$ such that $\langle m\rangle_t = \int_0^t \sigma_r \sigma_r'\, dr$. Then for any filtration $\mathcal{G}_t$ obeying the usual hypotheses and $\mathcal{F}_t^x \subseteq \mathcal{G}_t \subseteq \mathcal{F}_t$, there is a d-dimensional $\widetilde{\mathcal{G}}_t$-Brownian motion $\widetilde{w}_t$ on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{G}}_t, \widetilde{P})$ such that*

$$\widetilde{x}_t = \widetilde{x}_0 + \int_0^t \mathbf{E}[\widetilde{f}_r \mid \widetilde{\mathcal{G}}_r]\, dr + \int_0^t \widetilde{\sigma}_r\, d\widetilde{w}_r \qquad \forall t \in [0,T],\ \widetilde{P}-\ \text{a.s.}$$

(Here $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ is the product space defined in (3.5).)

Proof. We follow Wong (1971), Thm. 4.2, p. 629.

It is known that x_t is a quasimartingale with respect to the filtration $\mathcal{G}_t$. Therefore there is a $\mathcal{G}_t$-martingale n_t and a $\mathcal{G}_t$-predictable process a_t of bounded variation such that

$$x_t = x_0 + a_t + n_t \qquad \forall t \in [0,T],\ P-\ \text{a.s.} \tag{3.11}$$

Wong's proof (his p. 630) shows that

$$a_t = \int_0^t \mathbf{E}[f_r \mid \mathcal{G}_r]\, dr \qquad \forall t \in [0,T],\ P-\ \text{a.s.} \tag{3.12}$$

Subtracting representation (3.11) from the original representation (3.8) leads to

$$m_t - n_t = a_t - \int_0^t f_r\, dr \qquad \forall t \in [0,T],\ P-\ \text{a.s.}$$

Now the right side here is a process whose sample paths have bounded variation, $P-$ a.s. So by Lemma 3.5, we have

$$\langle n\rangle_t = \langle m\rangle_t = \int_0^t \sigma_r \sigma_r'\, dr \qquad \forall t \in [0,T],\ P- \text{ a.s.}$$

Hence Prop. 3.3 gives an $\mathbf{R}^d$-valued Brownian motion $\widetilde{w}$ on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{G}}_t, \widetilde{P})$ such that

$$\widetilde{n}_t = \int_0^t \widetilde{\sigma}_r\, d\widetilde{w}_r \qquad \forall t \in [0,T],\ \widetilde{P}- \text{ a.s.} \tag{3.13}$$

As we have seen above, (3.12) implies that on the extended space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$,

$$\widetilde{a}_t = \int_0^t \mathbf{E}[\widetilde{f}_r \mid \widetilde{\mathcal{G}}_r]\, dr \qquad \forall t \in [0,T],\ \widetilde{P}- \text{ a.s.} \tag{3.14}$$

Combining (3.11), (3.13), and (3.14) gives the desired result. ////

Because of the simple structure of the product space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ and the fact that all processes x on the original space $(\Omega, \mathcal{F}, P)$ retain their essential properites in the passage to $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$, the superscript tilde is often suppressed in applications of Wong's theorem. We will use this convention in the sections to follow.

Section 4. Problem Formulation

Stochastic Dynamics. We study a random dynamical system whose state x evolves in $\mathbf{R}^n$ under the influence of a Brownian motion w in $\mathbf{R}^d$ $(d \leq n)$ and a control signal u chosen from a preassigned closed set $U \subseteq \mathbf{R}^m$. The motion takes place on the (given) finite time interval $[0,T]$, starting from a fixed initial value x_0. Thus the dynamics are described by the Itô equation

$$x_t = x_0 + \int_0^t f(r,x,u_r)\, dr + \int_0^t \sigma(r,x)\, dw_r. \tag{4.1}$$

It is conventional to write (4.1) in differential form as follows

$$dx_r = f(r,x,u_r)\, dr + \sigma(r,x)\, dw_r, \qquad x(0) = x_0. \tag{4.2}$$

Here the coefficients $f\colon [0,T] \times C^n \times \mathbf{R}^m \to \mathbf{R}^n$ and $\sigma\colon [0,T] \times C^n \to \mathbf{R}^{n\times d}$ must satisfy hypotheses (H1)–(H3) below.

(H1) $f(t,x,u)$ and $\sigma(t,x)$ are Lebesgue measurable in t, continuous and $\mathcal{C}_t^n$-measurable in x; the continuity of $\sigma(t,\cdot)$ is uniform in t; and $f(t,x,u)$ is continuous in u.

(H2) There are constants $\kappa_1 > 0$ and $\beta \in (0,1)$ such that

$$|\sigma(t,x)| \leq \kappa_1 \left(1 + \|x\|_t^\beta\right) \qquad \forall t \in [0,T],\ x \in C^n.$$

This assumption evidently implies that for some k_1, one has

$$|\sigma(t,x)| \leq k_1 \left(1 + \|x\|_t\right) \qquad \forall t \in [0,T],\ x \in C^n;$$

we assume that k_1 also satisfies

$$|f(t,x,u)| \leq k_1 \left(1 + \|x\|_t + |u|\right) \qquad \forall t \in [0,T],\ x \in C^n,\ u \in U.$$

(H3) The initial value $x_0 \in \mathbf{R}^n$ is fixed and non-random.

Under these three hypotheses, a solution to equation (4.1) is to be understood as follows. The solution consists of a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ obeying the usual hypotheses and carrying an $\mathcal{F}_t$-Brownian motion w in $\mathbf{R}^d$, an $\mathcal{F}_t$-adapted stochastic process x_t with values in $\mathbf{R}^n$, and an x_t-adapted stochastic process u_t with values in U, all related by equation (4.1). Note that conditions (H1)–(H3) are too weak to imply any known existence or uniqueness theorems, so we must specify all these elements when discussing a solution. However, for the sake of brevity, we will often speak simply of "control-state pairs" (u,x), leaving implicit the associated space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and Brownian motion w.

Regardless of the underlying space, bounds on the moments of u and on the initial conditions lead to bounds on the moments of any process x satisfying (4.1).

4.1 Lemma. *Let $p \geq 2$, and assume (H1)–(H2). If $\mathbf{E}\,|x_0|^p \leq K_0$ and $\mathbf{E}\int_0^T |u_r|^p\,dr \leq K$, then for any control-state pair (u,x) one has*

$$\mathbf{E}\,\|x\|^p \leq M,$$

for a constant M which depends only on T, k_1, n, p, K_0, and K. In particular, M is independent of the specific choices of x_0 and u obeying the indicated conditions.

Proof. Hypothesis (H2) and (BDG) sponsor the following calculation.

$$
\begin{aligned}
&|x_t| \le |x_0| + \int_0^t |f(r,x,u_r)|\,dr + \left|\int_0^t \sigma(r,x)\,dw_r\right| \\
\Longrightarrow\quad &|x_t|^p \le K_p\left(|x_0|^p + T^{p-1}\int_0^t k_1^p(1+\|x\|_r+|u_r|)^p\,dr + \left|\int_0^t \sigma(r,x)\,dw_r\right|^p\right) \\
\Longrightarrow\quad &\mathbf{E}\,\|x\|_t^p \le K_p\left(K_0 + T^{p-1}k_1^pK_p\left(T+\int_0^t \mathbf{E}\,\|x\|_r^p\,dr + K\right) + K_pC_p\int_0^t (1+\mathbf{E}\,\|x\|_r^p)\,dr\right) \\
\Longrightarrow\quad &\mathbf{E}\,\|x\|_t^p \le C_0 + C_1\int_0^t \mathbf{E}\,\|x\|_r^p\,dr
\end{aligned}
$$

for some constants K_p, C_0, C_1. By Gronwall's inequality, $M = C_0e^{C_1T}$ suffices. ////

The Objective Functional. The cost of a given control-state pair (u,x) is measured by the functional

$$\Lambda[u,x] := \mathbf{E}\left[\ell(x_T,\mathbf{E}x_T) + \int_0^T L(r,x,u_r)\,dr\right]. \tag{4.3}$$

An *admissible pair* (u,x) is a control-state pair for which $\Lambda[u,x] < +\infty$. The pointwise cost $\ell\colon \mathbf{R}^{2n} \to \mathbf{R}$ and the running cost $L\colon [0,T]\times C^n\times U \to \mathbf{R}$ in (4.3) must satisfy (H4)–(H5) below.

(H4) $L(t,x,u)$ is measurable in t, continuous and $\mathcal{C}_t^n$-measurable in x, and continuous in u; ℓ is continuous.

(H5) The function ℓ takes on nonnegative values, and there exists $\alpha > 0$ such that $L(t,x,u) \ge \alpha\,|u|^2$ for all $(t,x,u) \in [0,T]\times C^n\times U$. (Note that if U is a compact set, this requirement can be replaced by the assumption that $L \ge 0$.)

Cesari's Condition. Even in the deterministic special case of our existence Thm. 5.1 below, a certain upper semicontinuity property must be assumed. This hypothesis needs no fortification in the stochastic case. It is stated in terms of the multifunction $Q\colon [0,T]\times C^n \hookrightarrow \mathbf{R}^n\times\mathbf{R}$ defined by

$$Q(t,x) = \left\{\left(f(t,x,u), L(t,x,u)+r\right) \;:\; u\in U,\ r\ge 0\right\}.$$

Cesari's condition is the following hypothesis.

(H6) For all $(t,x)\in[0,T]\times C^n$, with the possible exception of a set whose projection onto the t-axis has Lebesgue measure zero, one has

$$Q(t,x) = \bigcap_{\epsilon>0} \overline{\text{co}} \bigcup_{\substack{|s-t|<\epsilon \\ \|y-x\|<\epsilon}} Q(s,y). \tag{Q}$$

Note that under (H6), $Q(t,x)$ must be closed and convex because it is the intersection of closed convex sets. It is easy to verify that under our standing assumption that U is closed, the growth condition of (H5) and our continuity conditions on f, L automatically make $Q(t,x)$ closed. Fleming and Rishel (1975), Sect. III.4, p. 68 show this. Moreover, they prove that if f and L are continuous in t at some point t_0 and $Q(t_0,x_0)$ is known to be convex for some x_0, then property (Q) holds at (t_0,x_0). (Their Lemma III.5.4, p. 72.) More general conditions implying (H6) are given by Cesari (1983).

Problem (P). The stochastic control problem (P) is to choose an admissible pair $(\widehat{u},\widehat{x})$ such that $\Lambda[\widehat{u},\widehat{x}]$ equals the infimum of $\Lambda[u,x]$ over all admissible pairs (u,x) as defined above. This latter number is denoted $\inf(P)$: it is defined even if no optimal pair $(\widehat{u},\widehat{x})$ exists. (If there are no admissible pairs, then $\inf(P)=+\infty$.)

Discussion of Hypotheses. The measurability and continuity conditions of (H1) are standard. In Kushner (1975) they are supplemented by an assumption, denoted (A3), that $d=n$ and that σ is the positive square root of $\Sigma=\sigma\sigma'$. These conditions may both be traced to Kushner's use of Wong's theorem, whose original form required these conditions. However, the version of Wong's theorem presented above as Thm. 3.6 avoids these hypotheses—an improvement resulting directly from Jacod's careful formulation of our Prop. 3.3.

The growth conditions of (H2) are used not only in Lemma 4.1, but also to establish tightness and uniform integrability in Section 5. Similar conditions are required even in the deterministic case. The presence of two conditions on σ, the second implied by the first, reflects a desire to make do with the second one throughout. Our distinction between κ_1 and k_1 clarifies the fact that the strict inequality $\beta<1$ is needed only in the proof of Prop. 5.5: everywhere else we use the weaker condition involving k_1. To be more specific, the proofs of Prop. 5.5 and Lemma 4.1 show that the size of β is limited only by the modulus of integrability of the admissible controls. If the growth condition of (H5) is replaced by $L(t,x,u)\geq\alpha\left|u\right|^{2p}$ for some $p\geq 1$, then any $\beta<p$ will serve in (H2).

Hypothesis (H3) can be relaxed considerably: see paragraph 6.5.

The measurability and continuity hypotheses of (H4) are standard.

Hypothesis (H5) is substantially weaker than the assumption that U is a compact set used by Kushner (1975) and throughout the literature. This is the main contribution of the current chapter.

The requirement that the control process u be x_t-adapted in the definition of a control-state pair can be defended on the practical grounds that at any time t, the controller should use only the information obtainable from observing the state process itself at times before t. This information is defined by the σ-field $\mathcal{F}_t^x$. For theoretical completeness, however, we should mention the apparently larger class of controls u which are allowed to depend on anything in the known universe which takes place before time t: that is, controls which are adapted to $\mathcal{F}_t$ rather than to $\mathcal{F}_t^x$. We will now show that this family of controls offers no advantage (as measured by Λ) over the more practical family chosen above.

Suppose a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and an $\mathcal{F}_t$-Brownian motion w_t are given such that u_t is an $\mathcal{F}_t$-adapted process satisfying (4.1) for some $\mathcal{F}_t$-adapted process x_t. Assume $\mathbf{E}\int_0^T |u_r|^2\, dr < +\infty$. Then by Wong's Theorem (Thm. 3.6), there is an extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ carrying a Brownian motion $\widetilde{w}_t$ such that

$$x_t = x_0 + \int_0^t \mathbf{E}\big[f(r,x,u_r) \mid \mathcal{F}_r^x\big]\, dr + \int_0^t \sigma(r,x)\, d\widetilde{w}_r. \tag{4.4}$$

(We continue to denote this extension by $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$.) The objective value can be written as follows, using conditioning and Fubini's theorem:

$$\Lambda[u,x] = \mathbf{E}\left[\ell(x_T, \mathbf{E}x_T) + \int_0^T \mathbf{E}\big[L(t,x,u_t) \mid \mathcal{F}_t^x\big]\, dt\right]. \tag{4.5}$$

Now the process $\left(\widetilde{f}(t,\omega), \widetilde{L}(t,\omega)\right) = \mathbf{E}\left[\left(f(t,x,u_t), L(t,x,u_t)\right) \,\middle|\, \mathcal{F}_t^x\right]$ is $\mathcal{F}_t^x$-adapted and obeys

$$\left(\widetilde{f}(t,\omega), \widetilde{L}(t,\omega)\right) \in Q\big(t, x(\cdot,\omega)\big) \qquad \text{a.e. } t \in [0,T], \quad \text{a.s.} \tag{4.6}$$

Equation (4.6) is justified by the following Lemma.

4.2 Lemma. *(a) Let $(\Omega, \mathcal{F}, P)$ be a given probability space carrying a set-valued mapping $\Gamma\colon \Omega \to \mathbf{R}^n$ and an $\mathcal{F}$-measurable mapping $g\colon \Omega \to \mathbf{R}^n$ with $\mathbf{E}|g| < +\infty$. Suppose that Γ has nonempty closed convex values, and that for some σ-field $\mathcal{G} \subseteq \mathcal{F}$, the mapping $\omega \to \sup\{p\cdot\gamma \,:\, \gamma \in \Gamma(\omega)\}$ is*

$\mathcal{G}$-measurable for each $p \in \mathbf{R}^n$. Suppose further that the set $\{p \in \mathbf{R}^n : \sup p \cdot \Gamma(\omega) < +\infty\}$ has nonempty interior almost surely. Then under these conditions,

$$g(\omega) \in \Gamma(\omega) \ \text{ a.s.} \quad \Longrightarrow \quad \mathbf{E}\big[g \mid \mathcal{G}\big](\omega) \in \Gamma(\omega) \ \text{ a.s.}$$

(b) For each fixed $t \in [0,T]$, $\Gamma(\omega) = Q\big(t, x(\cdot,\omega)\big)$ and $\mathcal{G} = \mathcal{F}_t^x$ obey the hypotheses in (a).

Proof. (a) For each $p \in \mathbf{Q}^n$, $p \cdot \mathbf{E}\big[g \mid \mathcal{G}\big] = \mathbf{E}\big[p \cdot g \mid \mathcal{G}\big] \le \mathbf{E}\big[\sup(p \cdot \Gamma) \mid \mathcal{G}\big] = \sup(p \cdot \Gamma)$ a.s. For fixed ω, the sublinear function of p on the RHS is finite on a convex set with nonempty interior: the inequality therefore holds for all $p \in \mathbf{R}^n$ and (a) follows.

(b) The linear growth of f and the superquadratic growth of L imply

$$\operatorname{int}\big\{p \in \mathbf{R}^{n+1} : \sup p \cdot Q\big(t, x(\cdot,\omega)\big) < +\infty\big\} = \mathbf{R}^n \times (-\infty, 0) \qquad \text{a.s.} \qquad\qquad ////$$

Let us now consider the measure space $M = [0,T] \times \Omega$ equipped with the σ-field $\mathcal{M}$ generated by all sets $F \in \mathcal{B} \times \mathcal{F}$ such that for each fixed t, the projection $\{\omega : (t,\omega) \in F\}$ lies in $\mathcal{F}_t^x$ and for each fixed ω, the projection $\{t : (t,\omega) \in F\}$ lies in $\mathcal{B}$. Beneš (1971) shows that a mapping of $[0,T] \times \Omega$ into $\mathbf{R}^{n+1}$ is x_t-adapted if and only if it is $\mathcal{M}$-measurable. Hence in particular $\big(\widetilde{f}, \widetilde{L}\big)$ is $\mathcal{M}$-measurable.

Beneš (1971), Lemma 5, p. 460, gives an implicit function lemma almost perfectly suited to this situation. Using his notation, we take $(M, \mathcal{M})$ as defined above, $A = \mathbf{R}^{n+1}$, and $\widetilde{U} = U \times [0, +\infty)$. A typical element of $\widetilde{U}$ will be denoted by (u, v). Let us define $k \colon M \times \widetilde{U} \to A$ and $y \colon M \to \widetilde{U}$ by

$$\begin{aligned} k(t,\omega,u,v) &= \Big(f\big(t, x(\cdot,\omega), u\big),\ L\big(t, x(\cdot,\omega), u\big) + v\Big) \\ y(t,\omega) &= \Big(\widetilde{f}(t,\omega),\ \widetilde{L}(t,\omega)\Big). \end{aligned}$$

Then $k(t,\omega,u,v)$ is $\mathcal{M}$-measurable in (t,ω) for each fixed (u,v), and continuous in (u,v) for each fixed (t,ω). Also y is $\mathcal{M}$-measurable and obeys $y(t,\omega) \in k(t,\omega,\widetilde{U}) = Q\big(t, x(\cdot,\omega)\big)$ by (4.6). Now the set $\widetilde{U}$ is admittedly non-compact, but it is closed and σ-compact, a case in which Beneš's lemma is easily seen to remain valid. The conclusion of this lemma is that there is an $\mathcal{M}$-measurable mapping $(\widetilde{u}, \widetilde{v}) \colon [0,T] \times \Omega \to \widetilde{U}$ such that

$$\Big(\widetilde{f}(t,\omega), \widetilde{L}(t,\omega)\Big) = \Big(f\big(t, x(\cdot,\omega), \widetilde{u}(t,\omega)\big),\ L\big(t, x(\cdot,\omega), \widetilde{u}(t,\omega)\big) + \widetilde{v}(t,\omega)\Big).$$

Now $(\tilde{u}, \tilde{v})$ is $\mathcal{F}_t^x$-adapted since it is $\mathcal{M}$-measurable, and (4.4) (4.5) become

$$x_t = x_0 + \int_0^t f(r, x, \tilde{u}_r)\, dr + \int_0^t \sigma(r, x)\, d\tilde{w}_r$$
$$\Lambda[u, x] = \mathbf{E}\left[\ell(x_T, \mathbf{E}x_T) + \int_0^T L(t, x, \tilde{u}_t)\, dt + \int_0^T \tilde{v}_t\, dt\right] \geq \Lambda[\tilde{u}, x].$$

This demonstrates that any admissible pair (u, x) for which u is merely $\mathcal{F}_t$-adapted can be replaced by an admissible pair $(\tilde{u}, x)$ for which $\tilde{u}$ is $\mathcal{F}_t^x$-adapted, without increasing the cost. In fact, we can say more. Beneš (1971), pp. 450–451, shows that the $\mathcal{M}$-measurability of $\tilde{u}$ implies that there is a measurable map $v: [0, T] \times C^n \to U$ such that v is C_t^n-adapted and $\tilde{u}(t, \omega) = v(t, x(\cdot, \omega))$. Thus the use of $\mathcal{F}_t^x$-adapted controls is equivalent to the use of feedback controls.

In summary, there are two ways to define problem (P). The first uses *nonanticipative controls*, i.e. U-valued stochastic processes u defined on a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with $\mathcal{F}_t$-Brownian motion such that u_t is merely $\mathcal{F}_t$-adapted and solves (4.1) together with some $\mathcal{F}_t$-adapted process x_t. The second seems more restrictive: it uses *feedback controls*. These are C_t^n-adapted functions $v: [0, T] \times C^n \to U$ with the property that for some space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ carrying a Brownian motion w_t, there is an $\mathcal{F}_t$-adapted process x_t obeying

$$x_t = x_0 + \int_0^t f(r, x, v(r, x))\, dr + \int_0^t \sigma(r, x)\, dw_r. \tag{4.7}$$

The arguments above show that the use of feedback controls is not restrictive. Indeed, for any admissible nonanticipative control there is an admissible feedback control giving rise to an objective value at least as small. Of course feedback controls are not more general either, since an admissible feedback pair (v, x) solving (4.7) gives rise to a nonanticipative solution pair (u, x) for (4.1) by simply defining $u(t, \omega) = v(t, x(\cdot, \omega))$. Thus the number $\inf(P)$ is the same in the feedback and nonanticipative formulations, and the feedback problem has a solution if and only if the nonanticipative problem does.

Section 5. Existence Theory

This whole section is devoted to the proof of the following fact.

5.1 Theorem. *Assume (H1)–(H6).*

(a) *If problem (P) has an admissible pair then it has a solution.*

(b) *Indeed, let any sequence of admissible pairs $\{(u^k, x^k)\}$ be given, such that the objective values $\lambda^k = \Lambda[u^k, x^k]$ converge to some real number λ. Then there is an admissible pair (u, x) for (P) such that, along a subsequence, $x^k \xrightarrow{D} x$ in C^n and $\Lambda[u, x] \leq \lambda$.*

Clearly conclusion (a) follows from conclusion (b). For (H5) ensures that $\inf(P) \geq 0$, while $\inf(P) < +\infty$ because an admissible pair exists. Hence one can construct a sequence of admissible pairs as described in (b) for which the objective values tend to $\lambda = \inf(P)$. Then the admissible pair (u, x) given by (b) solves the problem. However, statement (b) is somewhat more general than a simple existence theorem. It can be viewed as a lower-semicontinuity conclusion about Λ which holds globally and not just near optimality. The significance of this will become clear later. We now turn our attention to the proof of (b).

Let any sequence $\{(u^k, x^k)\}$ as described in (b) be given. Each entry (u^k, x^k) in this sequence carries with it a probability space $(\Omega^k, \mathcal{F}^k, P^k)$ and a Brownian motion w^k. We will try to simplify the notation in the following arguments by suppressing the superscript k on P^k when evaluating the probability of an event which is clearly taken from $\mathcal{F}^k$, and agreeing that for a real-valued function g defined on $(\Omega^k, \mathcal{F}^k)$, we will write $\mathbf{E}g$ for expectation with respect to P^k.

Now the convergent sequence $\Lambda[u^k, x^k]$ is certainly bounded. Since $\Lambda[u^k, x^k] \geq \alpha \mathbf{E} \int_0^T |u_r^k|^2 \, dr$ by (H5), there is some constant $k_3 > 0$ such that

$$\mathbf{E} \int_0^T |u_r^k|^2 \, dr \leq k_3 \qquad \forall k.$$

The first step in proving (b) is to study the convergence properties of the state processes x^k. This commences with Prop. 5.2, using the notation

$$x^{1,k}(t) = \int_0^t f(r, x^k, u_r^k) \, dr,$$
$$x^{2,k}(t) = \int_0^t \sigma(r, x^k) \, dw_r,$$
$$z^k(t) = \int_0^t L(r, x^k, u_r^k) \, dr.$$

5.2 Proposition. *The sequence of quadruples $(x^k, x^{1,k}, x^{2,k}, z^k)$, considered as a collection of random vectors in $C^{3n} \times Z$, is tight.*

Proof. To prove tightness, we may consider each component sequence individually.

We begin with the sequence $\{x^{1,k}\}$. For each $\omega \in \Omega$ and $k \in \mathbf{N}$, one has

$$\begin{aligned}\sup_{0<t-s<\delta} \left|x^{1,k}(t) - x^{1,k}(s)\right| &\le \sup_{0<t-s<\delta} \int_s^t \left|f(r, x^k, u_r^k)\right| dr \\ &\le \sup_{0<t-s<\delta} \int_s^t k_1(1 + \left\|x^k\right\|_r + \left|u_r^k\right|) \, dr \\ &\le k_1 \left[\delta + \delta \left\|x^k\right\| + \delta^{\frac{1}{2}} \left(\int_0^T \left|u_r^k\right|^2 dr\right)^{\frac{1}{2}}\right].\end{aligned}$$

Taking expectations and using $\mathbf{E}\left\|x^k\right\| \le M$ (from Lemma 4.1) and $\mathbf{E}\int_0^T \left|u_r^k\right|^2 dr \le k_3$ gives

$$\mathbf{E} \sup_{0<t-s<\delta} \left|x^{1,k}(t) - x^{1,k}(s)\right| \le k_1 \left[\delta + \delta M + \delta^{\frac{1}{2}} k_3^{\frac{1}{2}}\right].$$

It follows from Chebyshev's inequality that there is some constant $R > 0$ such that

$$P\left\{\sup_{0<t-s<\delta} \left|x^{1,k}(t) - x^{1,k}(s)\right| > \varepsilon\right\} \le R\delta^{\frac{1}{2}}/\varepsilon.$$

This verifies condition (ii) of Prop. 2.4. Condition (i) is immediate since $x^{1,k}(0) = 0$ for all k. Hence the sequence $\{x^{1,k}\}$ is tight.

Let us now consider the sequence $\{x^{2,k}\}$. Denoting the standard basis vectors of $\mathbf{R}^n$ by $e_1, \ldots, e_n$, we first fix k and write $\sigma^k(t)$ for the matrix $\sigma(t, x^k)$. The i-component of the n-vector $x^{2,k}$ is

$$e_i' x^{2,k}(t) = \sum_{j=1}^d \int_0^t \sigma_{ij}^k(r) \, dw_j^k(r).$$

According to Ikeda-Watanabe (1981), Thm. II-7.2′, p. 91, there are Brownian motions $\tilde{w}_j^k$, $j = 1, \ldots, d$ (based perhaps on an extension of $(\Omega^k, \mathcal{F}^k, P^k)$) such that

$$e_i' x^{2,k}(t) = \sum_{j=1}^d \tilde{w}_j^k \left(\int_0^t \left|\sigma_{ij}^k(r)\right|^2 dr\right).$$

Therefore for any given $\delta, \varepsilon > 0$ we have

$$P\left\{\sup_{0<t-s<\delta} \left|e_i'(x^{2,k}(t) - x^{2,k}(s))\right| > \varepsilon\right\} \le$$
$$P\left\{\sup_{0<t-s<\delta} \sum_{j=1}^d \left|\tilde{w}_j^k \left(\int_0^t \left|\sigma_{ij}^k(r)\right|^2 dr\right) - \tilde{w}_j^k \left(\int_0^s \left|\sigma_{ij}^k(r)\right|^2 dr\right)\right| > \varepsilon\right\}.$$

For each ω, σ^k is a bounded measurable function on $[0,T]$ so $\|\sigma^k\|$ is defined in $L^\infty[0,T]$. We partition the RHS into the sets $\left\{\|\sigma^k\|^2 > R\right\}$ and $\left\{\|\sigma^k\|^2 \le R\right\}$, whereupon it is evident that the LHS is not larger than

$$P\left\{\|\sigma^k\|^2 > R\right\} + P\left\{\sup_{0<t'-s'<R\delta} \sum_{j=1}^{d} \left|\tilde{w}_j^k(t') - \tilde{w}_j^k(s')\right| > \varepsilon\right\}$$
$$\le \mathbf{E}\left\|\sigma^k\right\|^2 / R + \sum_{j=1}^{d} P\left\{\sup_{0<t'-s'<R\delta} \left|\tilde{w}_j^k(t') - \tilde{w}_j^k(s')\right| > \varepsilon/d\right\}.$$

Now for all k at once, Lemma 4.1 gives a constant upper bound for

$$\mathbf{E}\left\|\sigma^k\right\|^2 \le 2k_1\left(1 + \mathbf{E}\left\|x^k\right\|^2\right).$$

Hence for any given $\varepsilon > 0$ and $\eta > 0$, we first choose R so that $\mathbf{E}\left\|\sigma^k\right\|^2/R < \eta/2$ for all k. Then we note that the second term of the RHS above tends to 0 as $\delta \to 0$ because almost-sure convergence implies convergence in probability; hence the second term is also smaller than $\eta/2$ for all $\delta > 0$ sufficiently small. Also, the numerical value of the second term is independent of k and i. It is therefore possible to combine the componentwise estimates for $i = 1, 2, \ldots, n$ into the following:

$$\lim_{\delta\to 0+} \sup_k P\left\{\sup_{0<t-s<\delta} \left|x^{2,k}(t) - x^{2,k}(s)\right| > \varepsilon\right\} = 0.$$

Prop. 2.4 now shows that $\left\{x^{2,k}\right\}$ is tight.

Since $x^k = x_0 + x^{1,k} + x^{2,k}$, the tightness of $\left\{x^k\right\}$ follows from the fact that $\left\{x^{1,k}\right\}$ and $\left\{x^{2,k}\right\}$ are tight.

Finally, the tightness of $\left\{z^k\right\}$ can be verified directly by Prop. 1.3. Simply consider the compact sets $K_N = \{z \in Z \,:\, z(T) \le N\}$ for $N \in \mathbf{N}$. Then

$$P\left\{z^k \notin K_N\right\} = P\left\{\int_0^T L(r, x^k, u_r^k)\, dr > N\right\}$$
$$\le \frac{1}{N}\mathbf{E}\left(\int_0^T L(r, x^k, u_r^k)\, dr\right).$$

Since the supremum over k of the RHS is bounded, it follows that

$$\lim_{N\to\infty} \sup_k P\left\{z^k \notin K_N\right\} = 0.$$

That is, the sequence $\left\{z^k\right\}$ is tight. ////

Let us now apply Prokhorov's Theorem (Prop. 2.3). It states that there is a subsequence (which we do not relabel) along which $(x^k, x^{1,k}, x^{2,k}, z^k)$ converges in distribution to a random element

$$(x, x^1, x^2, z) : (\Omega, \mathcal{F}, P) \to C^{3n} \times Z.$$

We wish to use this limiting quadruple to construct the admissible pair whose existence is asserted by Thm. 5.1(b). Let us start by considering the following subsets of $C^{3n} \times Z$:

$$\begin{aligned} S_N = \Big\{ (x, x^1, x^2, z) \ : (x, x^1, x^2) \in C^{3n}, \ & z \in Z, \\ & x(t) = x_0 + x^1(t) + x^2(t), \quad x^1(0) = x^2(0) = 0, \\ & x^1 \in AC^n, \quad \int_0^T \left|\dot{x}^1(r)\right|^2 dr \le N^2, \\ & z(T) \le N, \\ & (\dot{x}^1(t), \dot{z}(t)) \in Q(t, x(\cdot)) \text{ a.e.} \Big\}. \end{aligned}$$

Each set S_N is closed in $C^{3n} \times Z$. To justify this statement, let $(x^k, x^{1,k}, x^{2,k}, z^k) \in S_N$ be a sequence converging to some point (x, x^1, x^2, z). It follows immediately that $x(t) = x_0 + x^1(t) + x^2(t) \quad \forall t$, that $x^1(0) = x^2(0) = 0$, and that $z(T) \le N$. The condition $\left\|\dot{x}^{1,k}\right\|_2 \le N \quad \forall k$ implies that $x^1 \in AC^n$ and $\left\|\dot{x}^1\right\|_2 \le N$. (This is a standard fact in the calculus of variations, central to the proof of Tonelli's classical existence theorem.) It remains only to check that $(\dot{x}^1(t), \dot{z}(t)) \in Q(t, x(\cdot))$ a.e. This is a corollary of the following well-known closure theorem used in deterministic control theory.

5.3 Proposition. *Assume that the sets $Q(t, x(\cdot))$ satisfy hypothesis (H6). Let $(x^k, z^k) \in AC^n \times Z$ be a sequence of functions obeying*

$$(\dot{x}^k(t), \dot{z}^k(t)) \in Q(t, x^k(\cdot)) \qquad \text{a.e. on } [0,T], \ \forall k.$$

If $x^k \to x$ uniformly for some $x \in AC$ and $z^k \to z$ pointwise a.e. for some $z \in Z$, then one has

$$(\dot{x}(t), \dot{z}(t)) \in Q(t, x(\cdot)) \text{ a.e. on } [0,T].$$

Proof. The statement is very similar to that of closure theorem 15.2.i given by Cesari (1983), p. 444. Two differences are worth mentioning, however.

First, Cesari's theorem involves a multifunction Q depending on $(t, x_t) \in [0,T] \times \mathbf{R}^n$ instead of on $(t, x(\cdot)) \in [0,T] \times C^n$. His proof requires only minor changes to treat the more general case.

Second, Cesari uses a sequence of absolutely continuous functions for z^k instead of using elements of Z. His exact conclusion is that if the pointwise limit $z(t)$ has a representation as the sum of an absolutely continuous function $c(t)$ and a singular function $s(t)$, then $(x, c) \in AC^{n+1}$ solves the indicated differential inclusion. Now in our context, the limit $z \in Z$ certainly has such a decomposition, so our conclusion stated above follows from Cesari's. However, the requirement that each z^k be absolutely continuous is important to Cesari's proof. We must therefore explain why his argument remains valid despite the possibility of a strictly positive r in the statement

$$\int_a^b \dot{z}^k(t)\,dr = z^k(b) - z^k(a) + r \quad \text{for some } r \geq 0.$$

Here the special shape of the sets Q intervenes. The identity

$$Q(t,x) = Q(t,x) + \{0\} \times [0, +\infty) \qquad \forall (t,x) \in [0,T] \times C^n$$

is precisely the observation needed to extend Cesari's proof to include the current proposition. ////

Now for each fixed k, we investigate $P\left\{(x^k, x^{1,k}, x^{2,k}, z^k) \in S_N\right\}$. The follo wing conditions hold with probability one:

$$x^k(t) = x_0 + x^{1,k}(t) + x^{2,k}(t) \quad \forall t, \qquad x^{1,k}(0) = x^{2,k}(0) = 0,$$

$$x^{1,k} \in AC^n, \qquad (\dot{x}^{1,k}(t), \dot{z}^k(t)) \in Q(t, x^k(\cdot)) \text{ a.e.}$$

Only the upper bounds on $\left\|\dot{x}^{1,k}\right\|_2$ and $z^k(T)$ remain to consider. By (H2), we have

$$\begin{aligned} P\left\{\int_0^T \left|\dot{x}^{1,k}(r)\right|^2 dr > N^2\right\} &\leq \frac{1}{N^2}\mathbf{E}\int_0^T \left|f(r, x^k, u_r^k)\right|^2 dr \\ &\leq \frac{1}{N^2} 3k_1 \mathbf{E}\int_0^T \left(1 + \left\|x^k\right\|_r^2 + \left|u_r^k\right|^2\right) dr, \end{aligned}$$

and the expectation on the right-hand side is bounded uniformly in k by Lemma 4.1 and the observation following Theorem 5.1. Also,

$$P\{z^k(T) > N\} \leq \frac{1}{N}\mathbf{E}\int_0^T L(r, x^k, u_r^k)\,dr.$$

Here the RHS is uniformly bounded in k because we started with a convergent sequence of objective values. We therefore have

$$\lim_{N\to\infty} \inf_k P\left\{(x^k, x^{1,k}, x^{2,k}, z^k) \in S_N\right\} = 1.$$

By Def. 2.1(b) (see text following Thm. 2.3), it follows that the limiting 4-tuple (x, x^1, x^2, z) lies in the set $\bigcup_{N=1}^{\infty} S_N$ a.s.This proves the following assertion.

5.4 Proposition. *The limiting quadruple (x, x^1, x^2, z) of Prop. 5.2 obeys the following conditions with probability one:*

$$x(t) = x_0 + x^1(t) + x^2(t) \quad \forall t, \qquad x^1(0) = x^2(0) = 0,$$
$$x^1 \in AC^n, \qquad \int_0^T \left|\dot{x}^1(r)\right|^2 dr < \infty, \qquad z(T) < \infty,$$
$$\left(\dot{x}^1(t), \dot{z}(t)\right) \in Q(t, x(\cdot)) \text{ a.e.}$$

Let us now use Wong's theorem to show that x_t^2 can be represented as a stochastic integral of the required form. We first define a filtration $\mathcal{F}_t$ on $(\Omega, \mathcal{F}, P)$ as the filtration generated by (x_t, x_t^2).

5.5 Proposition. *There exists an extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ on which the conclusions of Prop. 5.4 remain valid, and which supports a d-dimensional Bownian motion $\tilde{w}_t$ such that*

$$x_t = x_0 + \int_0^t \mathbf{E}\left[\dot{x}_r^1 \mid \mathcal{F}_r^x\right] dr + \int_0^t \sigma(r, x)\, d\tilde{w}_r \qquad \forall t \in [0, T], \text{ a.s.}$$

Proof. From Prop. 5.4 we have

$$x_t = x_0 + \int_0^t \dot{x}_r^1\, dr + x_t^2 \qquad \forall t \in [0, T], \text{ a.s.}$$

Condition (3.9) of Wong's theorem is evident; it remains to verify the martingale properties of x_t^2.

We first show that $\left\{\left\|x^{2,k}\right\|^2\right\}$ is a uniformly integrable sequence. Indeed, (BDG) gives

$$\begin{aligned}\mathbf{E}\left\|x^{2,k}\right\|^{2/\beta} &\le C_p \int_0^T \mathbf{E}\left|\sigma(r, x^k)\right|^{2/\beta} dr \\ &\le C_p \kappa_1 \int_0^T \left(1 + \mathbf{E}\left\|x^k\right\|^2\right) dr.\end{aligned}$$

The RHS is bounded uniformly in k; uniform integrability follows because $2/\beta > 2$.

Now for each k, $x_t^{2,k}$ is a martingale with respect to the filtration generated by $(x_t^k, x_t^{2,k})$. So by Prop. 3.1, it follows that x_t^2 is an $\mathcal{F}_t$-martingale.

The quadratic variation of $x_t^{2,k}$ is $q_t^k = \int_0^t \sigma(r, x^k)\sigma(r, x^k)'\, dr$. By definition, this means that $\left(x^{2,k}\right)\left(x^{2,k}\right)' - q_t^k$ is a martingale for each k. The calculation above shows that this martingale has a uniformly integrable sequence of supremum norms. Moreover, q_t^k is a continuous image of the convergent (in distribution) sequence x^k, so on the space of continuous matrix-valued functions we have

$$\left(x^{2,k}\right)\left(x^{2,k}\right)' - q^k \xrightarrow{\mathcal{D}} \left(x^2\right)\left(x^2\right)' - q,$$

where $q_t = \int_0^t \sigma(r,x)\sigma(r,x)'\,dr$. According to Prop. 3.1, it follows that $(x_t^2)(x_t^2)' - q_t$ is an $\mathcal{F}_t$-martingale. By definition, the quadratic variation of x_t^2 is q_t, i.e.

$$\langle x^2 \rangle_t = \int_0^t \sigma(r,x)\sigma(r,x)'\,dr.$$

The process $\sigma(r,x)$ is clearly $\mathcal{F}_t^x$-predictable, so Wong's theorem (Thm. 3.6) applies. It concludes that on a certain extension of $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ there is a d-dimensional Brownian motion $\tilde{w}_t$ such that

$$x_t = x_0 + \int_0^t \mathbf{E}[\dot{x}_r^1 \mid \mathcal{F}_r^x]\,dr + \int_0^t \sigma(r,x)\,d\tilde{w}_r \qquad \forall t \in [0,T], \quad \text{a.s.}$$

The explicit construction of this extension in Section 3 makes it clear that the properties of the limiting quadruple described in Prop. 5.4 remain intact. ////

Let us now examine the following conclusion of Prop. 5.4:

$$\left(\dot{x}^1(t,\omega), \dot{z}(t,\omega)\right) \in Q(t, x(\cdot,\omega)).$$

Lemma 4.2 implies that if we define $\left(\tilde{f}(t,\omega),\ \tilde{L}(t,\omega)\right) = \mathbf{E}\left[(\dot{x}_t^1, \dot{z}_t) \;\middle|\; \mathcal{F}_t^x\right]$, then

$$\left(\tilde{f}(t,\omega), \tilde{L}(t,\omega)\right) \in Q(t, x(\cdot,\omega)) \qquad \forall t \in [0,T], \quad \text{a.s.}$$

In this notation, the argument given in the text following Lemma 4.2 establishes the following result.

5.6 Proposition. *There is an $\mathcal{F}_t^x$-adapted process $(u,v)\colon [0,T] \times \Omega \to U \times [0,+\infty)$ for which*

$$\tilde{f}(t,\omega) = f(t,x,u_t) \qquad \text{a.e. } [0,T], \quad \text{a.s.,}$$

$$\tilde{L}(t,\omega) = L(t,x,u_t) + v_t \qquad \text{a.e. } [0,T], \quad \text{a.s.}$$

The following Proposition completes the proof of Thm. 5.1.

5.7 Proposition. *The pair (u,x) is admissible, and $\Lambda[u,x] \le \lambda$.*

Proof. Propositions 5.4–5.6 show that $u\colon [0,T] \times \Omega \to U$ is an $\mathcal{F}_t^x$-adapted process for which

$$dx_t = f(t,x,u_t)\,dt + \sigma(t,x)\,dw_t, \qquad x(0) = x_0.$$

Thus (u,x) is a control-state pair for the dynamics (4.1).

Since $z^k(T) = \int_0^T L(r, x^k, u_r^k)\, dr$ converges in distribution to $z(T)$, a version of Fatou's lemma (Billingsley (1968), Thm. 5.3, p. 32) gives

$$\mathbf{E}z(T) \leq \liminf_{k\to\infty} \mathbf{E}z^k(T).$$

Now $z(T) \geq \int_0^T \dot{z}(t)\, dt$ a.s., so

$$\liminf_{k\to\infty} \mathbf{E}z^k(T) \geq \mathbf{E}\int_0^T \widetilde{L}(t)\, dt = \mathbf{E}\left[\int_0^T L(r, x, u_r)\, dr + \int_0^T v_r\, dr\right],$$

$$\Longrightarrow \qquad \mathbf{E}\int_0^T L(r, x, u_r)\, dr \leq \liminf_{k\to\infty} \mathbf{E}z^k(T) - \mathbf{E}\int_0^T v_r\, dr. \qquad (*)$$

Since $x^k(T) \xrightarrow{\mathcal{D}} x(T)$ and $\mathbf{E}\left|x^k(T)\right|^2$ is bounded uniformly in k by Lemma 4.1, uniform integrability implies $\mathbf{E}x^k(T) \to \mathbf{E}x(T)$. The technical Lemma 5.8 below shows that this implies

$$\ell(x^k(T), \mathbf{E}x^k(T)) \xrightarrow{\mathcal{D}} \ell(x(T), \mathbf{E}x(T)).$$

A second application of Fatou's lemma now gives

$$\mathbf{E}\ell(x(T), \mathbf{E}x(T)) \leq \liminf_{k\to\infty} \mathbf{E}\ell(x^k(T), \mathbf{E}x^k(T)). \qquad (**)$$

Combining $(*)$ and $(**)$, we obtain the inequality proclaimed by Thm. 5.1(b):

$$\Lambda[u, x] \leq \liminf_{k\to\infty} \Lambda[u^k, x^k] - \mathbf{E}\int_0^T v(r)\, dr \leq \lambda. \qquad ////$$

5.8 Lemma. *Suppose a sequence x^k of $\mathbf{R}^n$-valued random variables converges in distribution to a random variable x, and that moreover $\mathbf{E}x^k \to \mathbf{E}x$. Then for any continuous function $\ell\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, one has $\ell(x^k, \mathbf{E}x^k) \xrightarrow{\mathcal{D}} \ell(x, \mathbf{E}x)$.*

Proof. Let any $\varepsilon > 0$ be given. Since $\{x^k\}$ converges in distribution, there is a compact set K_0 such that $P\{x^k \in K_0\} > 1 - \varepsilon \quad \forall k$. Choose a compact set K_1 containing the sequence $\{\mathbf{E}x^k\}$. Then on $K = K_0 \times K_1$, the function ℓ is uniformly continuous, so there exists $\delta > 0$ so small that $|(x, y) - (x', y')| < \delta$ for $(x, y), (x', y') \in K$ implies $|\ell(x, y) - \ell(x', y')| < \varepsilon$. Then choose $N \in \mathbf{N}$ such that $k \geq N$ forces $\left|\mathbf{E}x^k - \mathbf{E}x\right| < \delta$. It follows that

$$P\{\left|\ell(x^k, \mathbf{E}x^k) - \ell(x^k, \mathbf{E}x)\right| > \varepsilon\} \leq P\{x^k \notin K_0\} < \varepsilon.$$

Hence $\left|\ell(x^k, \mathbf{E}x^k) - \ell(x^k, \mathbf{E}x)\right| \xrightarrow{\mathrm{P}} 0$. Now since $\ell(\cdot, \mathbf{E}x)$ is continuous, $\ell(x^k, \mathbf{E}x) \xrightarrow{\mathcal{D}} \ell(x, \mathbf{E}x)$. By Billingsley (1968), Thm. 4.1, p. 25, it follows that $\ell(x^k, \mathbf{E}x^k) \xrightarrow{\mathcal{D}} \ell(x, \mathbf{E}x)$. $\qquad ////$

Section 6. Extensions of Theorem 5.1

The standing assumption that the running cost L and the terminal cost ℓ are nonnegative can clearly be relaxed. We need only assume that L and ℓ are bounded below, and that L satisfies a growth condition of the form

$$L(t,x,u) \geq \alpha\,|u|^2 - \gamma \text{ for some } \alpha > 0,\ \gamma \geq 0.$$

This section is devoted to a collection of similar observations—most not quite so obvious—regarding the hypotheses under which existence is assured.

6.1 The Deterministic Case. The case $\sigma \equiv 0$, in which the stochastic dynamics of Section 4 reduce to deterministic functional differential equations, is permitted by the hypotheses governing problem (P). When (H1) is further specialized so that f depends only upon the current position of the state process x and not on its past, Thm. 5.1 reduces to a standard existence theorem in deterministic optimal control. See, for example, Fleming and Rishel (1975), Thm. III.4.1, p. 68. However, several issues raised by this comparison deserve comment. First, our result requires no Lipschitz condition on f, whereas Fleming and Rishel make this assumption in their condition (2.4), p. 62. The Lipschitz condition is used to ensure the existence and uniqueness of solutions to the governing equation for every possible choice of u: our theory avoids this by ignoring uniqueness completely and making the existence of a corresponding solution one of the prerequisites for the admissibility of u. When existence and uniqueness are required in the stochastic context for other reasons (as they will be in Chap. V), the Itô conditions are used. These are more stringent than Fleming and Rishel's hypothesis (2.4) on p. 62.

Second, the deterministic theory can be developed under the growth condition

$$L(t,x,u) \geq \alpha\,|u|^p - \gamma \text{ for some } p > 1,$$

whereas we require $p \geq 2$. Here the stochastic theory cannot be strengthened. The exponent 2 is essential because it is the reciprocal of Brownian motion's index of Hölder continuity. This accounts for the special role of 2 in Propositions 1.1 and 1.2, which are crucial to our proof.

Finally, note that setting $\sigma \equiv 0$ really does render the problem completely deterministic, since we have shown that to any admissible control $u(t,\omega)$ there corresponds a feedback control $v(t,x)$ for which the dynamics become

$$\dot{x}(t) = f\big(t, x, v(t,x)\big), \quad x(0) = x_0.$$

No external randomization is present here. However, Thm. 5.1 must not be interpreted as implying the existence of an optimal feedback control law for deterministic problems. This is because the phrase "feedback control" unfortunately has two different meanings. In the stochastic theory, it refers only to the functional dependence of the control law on the state process—something which is always trivially present in the deterministic case. In the deterministic theory, "optimal feedback control" designates a single function $v(t,x)$ which solves all the versions of problem (P) regardless of the initial point in (t,x)-space. We have fixed the initial point $(0,x_0)$ throughout the arguments above, so the optimal control generated by Thm. 5.1 is a feedback control only in the sense of functional dependence, and not necessarily in the sense used by dynamic programmers.

6.2 Stochastic Calculus of Variations. A very important special case of problem (P) arises when $f(t,x,u) = u$, $U = \mathbf{R}^n$. Then the dynamics are simply

$$dx_t = u_t\,dt + \sigma(t,x)\,dw_t, \quad x(0) = x_0, \quad u \in \mathbf{R}^n,$$

and problem (P) becomes a *stochastic calculus of variations* problem. See Fleming (1983). If we assume that L is continuous in t, then hypothesis (H6) holds whenever $L(t,x,\cdot)$ is a convex function for each pair (t,x).

6.3 Incorporating Constraints. Suppose the criteria for admissibility in problem (P) are tightened by adding a system of soft constraints such as

$$\begin{aligned}
\mathbf{E}\left[\ell_i(x_T, \mathbf{E}x_T) + \int_0^T L_i(r,x,u_r)\,dr\right] &\le \lambda_i, \quad i = -1,-2,\ldots,-I,\\
\mathbf{E}\left[\ell_j(x_T, \mathbf{E}x_T) + \int_0^T L_j(r,x,u_r)\,dr\right] &= \lambda_j, \quad j = 1,2,\ldots,J.
\end{aligned}$$

Under certain hypotheses, Thm. 5.1 can be extended to include this case. Let us write ℓ_0, L_0 for the objective functional's constituents ℓ, L.

The inequality constraints can be added by imagining a multidimensional objective functional in Thm. 5.1. Thus we replace the scalar point cost ℓ and running cost L with the vectors $\ell = (\ell_0, \ell_{-1}, \ldots, \ell_{-I})$ and $L = (L_0, L_{-1}, \ldots, L_{-I})$ and assume that these vector-valued functions obey (H4). Instead of (H5), we require only that each function ℓ_i, L_i, $i = 0, -1, \ldots, -I$, be nonnegative, and that *any one of the functions* L_i, $i = 0, \ldots, -I$, obeys the superquadratic growth condition $L_i(t, x, u) \geq \alpha |u|^2$. The growth of any single component of L suffices to imply that $\mathbf{E} \int_0^T |u_r^k|^2 \, dr$ is bounded uniformly in k for any sequence of admissible arcs with converging vector objective values, and this leads to the boundedness of $\mathbf{E} \|x^k\|^2$ by Lemma 4.1. The remainder of the proof carries through as before, except that we now consider the $(I+1)$-fold product Z^{I+1} as the metric space in which $z(t) = \int_0^t L(r, x, u_r) \, dr$ takes its values. Hypothesis (H6) also involves the vector function L in a natural way. (We will discuss this further below.)

To treat the equality constraints, we think of L_j, $j = 1, 2, \ldots, J$, as additional components of the function f and consequently require that L_j obey the linear growth estimates and Lipschitz conditions of (H1)–(H2). Since the proof of Thm. 5.1 shows $\mathbf{E}x^{1,k}(T) \to \mathbf{E}x^1(T)$ along a sequence with convergent objective values, it follows that these hypotheses will preserve the value of each $\mathbf{E} \int_0^T L_j(r, x^k, u_r^k) \, dr$ in the limit as $k \to \infty$. To guarantee that the values $\mathbf{E}\ell_j(x^k(T), \mathbf{E}x^k(T))$ are also preserved in the limit, it suffices to make sure each of these sequences is uniformly integrable. We therefore allow ℓ_j to be any continuous function obeying

$$|\ell_j(x, e)| \leq k_j(1 + |x|^q + |e|^r) \quad \forall (x, e) \in \mathbf{R}^{2n} \tag{6.2}$$

for some constants $k_j > 0$, $q \in (0, 2)$, and $r > 0$. Then the argument replacing the derivation of (**) in the proof of Prop. 5.7 would run as follows. (We suppress the subscript j.) Observe that $\ell(x_T^k, \mathbf{E}x_T^k) \xrightarrow{\mathcal{D}} \ell(x_T, \mathbf{E}x_T)$. Moreover,

$$\mathbf{E} \left| \ell(x_T^k, \mathbf{E}x_T^k) \right|^{2/q} \leq \tilde{k} \left(1 + \mathbf{E} \left| x_T^k \right|^2 + \left| \mathbf{E}x_T^k \right|^{2r/q} \right).$$

The RHS is uniformly bounded in k by Lemma 4.1 and the convergence of $\mathbf{E}x_T^k$, so uniform integrability implies *equality* in

$$\mathbf{E}\ell(x_T, \mathbf{E}x_T) = \lim_{k \to \infty} \mathbf{E}\ell(x_T^k, \mathbf{E}x_T^k).$$

Having briefly described the hypotheses on ℓ_i and L_i corresponding to (H1)–(H5), let us explicitly mention that for the constrained problem the multifunction $Q(t, x(\cdot))\colon [0, T] \times C^n \hookrightarrow \mathbf{R}^{n+J} \times \mathbf{R}^{1+I}$ becomes

$$Q(t, x(\cdot)) = \{(F(t,x,u), L(t,x,u) + r) \; : F = (f, L_1, L_2, \ldots, L_J),$$
$$L = (L_0, L_{-1}, L_{-2}, \ldots, L_{-I}),$$
$$u \in U, \; r \in \mathbf{R}^{1+I} \text{ obeys } r_i \geq 0 \; \forall i\}.$$

The wording of hypothesis (H6) remains the same.

6.4 More General Point Costs. The point cost functional ℓ in problem (P) entered the proof of Thm. 5.1 only in the last part of Prop. 5.7. (Also $\ell \geq 0$ was implicitly used earlier.) A look at that argument shows that ℓ could easily be allowed to depend on any number of points in addition to T. In particular, if $0 < T_1 < T_2 < \cdots < T_N = T$ is any finite partition of $[0, T]$, Thm. 5.1 remains valid when $\ell\colon \mathbf{R}^{2Nn} \to [0, \infty)$ is a continuous function of the form

$$\ell = \ell\big(x(T_1), \mathbf{E}x(T_1), x(T_2), \mathbf{E}x(T_2), \ldots, x(T_N), \mathbf{E}x(T_N)\big).$$

Consideration of such general point-dependence is prompted by Kushner (1972). The remarks of paragraph 6.3 show how such general point costs can also be used in any constraints added to the problem.

6.5 Random Initial Value. Hypothesis (H3) regarding the initial point x_0 can be relaxed considerably. For instance, Thm. 5.1 remains valid under the assumption that x_0 is a random variable with a given distribution on $\mathbf{R}^n$. Under the additional hypothesis that $\mathbf{E}\,|x_0|^2 < +\infty$, Lemma 4.1 remains valid for $p = 2$, and the tightness of any minimizing sequence is established just as before. The proof of Thm. 5.1 then proceeds. The only change is that to deduce Prop. 5.4, one must add a component accounting for x_0 to the sets S_N, which now become subsets of $\mathbf{R}^n \times C^{3n} \times Z$ consisting of all points (x_0, x, x^1, x^2, z) satisfying the same defining conditions as before.

Indeed, the method of the previous paragraph will show that the random variable x_0 can be considered as an additional choice variable, provided some condition is imposed to ensure that along any minimizing sequence, $\mathbf{E}\,|x_0^k|^2$ is uniformly bounded. (This not only preserves the conclusion of Lemma 4.1, but also implies that x_0^k is a tight sequence of random elements of $\mathbf{R}^n$.) Demanding that the random initial value take its values in some predetermined compact set $A \subseteq \mathbf{R}^n$ would certainly accomplish this, as would several weaker conditions. Indeed, by eliminating ω-dependence from these arguments, we can see that existence is also assured if x_0 is assumed to be a *deterministic* choice variable in the compact set A.

Constraints of the form discussed in paragraph 6.3 can be included in any problem where x_0 is free and random, and the general point costs ℓ can then be allowed to depend upon $x_0, \mathbf{E}x_0$ by the explanation of paragraph 6.4.

Section 7. A Compact Control Set

The special case of Thm. 5.1 arising when the control set U is compact is rather widely applicable, but it differs from standard results in several ways that deserve a closer look. In this section we compare this special case to the results obtained when explicit appeals to the compactness of U are allowed throughout the proof. Kushner's work (1975) is representative of the latter. When restricted to the deterministic regime, our discussion may be viewed as a comparison of Theorems III.2.1, p. 63, and III.4.1, p. 68, of Fleming and Rishel (1975).

To make for a meaningful discussion, let us sketch the usual theory for the case when U is compact. It begins with hypotheses allowing random initial conditions (see paragraph 6.5).

(h1) Same as (H1).

(h2) Same as (H2).

(h3) There are an exponent $\overline{q} \geq 2$ and a constant $k_2 > 0$ such that for all admissible initial values x_0, one has $\mathbf{E}\,|x_0|^{\overline{q}} \leq k_2$. (Note: if $\overline{q} > 2$ then we can take $\beta = 1$, $\kappa_1 = k_1$ in (h2).)

(h4) Same as (H4).

(h5) For some constants $q \in [0, \overline{q})$ and $r > 0$, one has a constant $k_3 > 0$ such that

$$|L(t,x,u)| \leq k_3(1 + \|x\|_t^q)$$

$$|\ell(v,e)| \leq k_3(|v|^q + |e|^r)$$

for all $(t,x,u) \in [0,T] \times C^n \times U$ and $(v,e) \in \mathbf{R}^{2n}$.

(h6) The set $Q_0(t,x) = \{(f(t,x,u), L(t,x,u)) \;:\; u \in U\}$ obeys

$$Q_0(t,x) = \bigcap_{\epsilon>0} \overline{\mathrm{co}} \bigcup_{\substack{\|y-x\|<\epsilon \\ |s-t|<\epsilon}} Q_0(s,y)$$

for all $(t,x) \in [0,T] \times C^n$, with the possible exception of a set whose projection onto the t-axis has Lebesgue measure zero.

7.1 Theorem. *Suppose that U is compact, and that (h1)–(h6) hold.*

(a) If problem (P) has an admissible pair then it has a solution.

(b) Indeed, let any sequence of admissible pairs $\{(u^k, x^k)\}$ for (P) be given, such that the objective values $\Lambda[u^k, x^k]$ converge to some real number λ. Then there is an admissible pair (u,x) for (P) such that, along a subsequence, $x^k \xrightarrow{\mathcal{D}} x$ in C^n and $\Lambda[u,x] = \lambda$.

Proof. This is essentially Kushner's (1975) Theorem 3.1, p. 350. We therefore simply sketch the proof, emphasizing its differences from the proof of Theorem 5.1.

The sequence of 5-tuples $(x_0^k, x^k, x^{1,k}, x^{2,k}, z^k)$ is still tight, but in a different metric space. The growth conditions (h5) imply that *$\{z^k\}$ is tight in C.* Indeed, we calculate

$$\sup_{0<t-s<\delta} \left|z^k(t) - z^k(s)\right| \le \sup_{0<t-s<\delta} \int_s^t k_3(1 + \|x^k\|_r^q)\, dr.$$

Taking expectations and using the uniform boundedness of $\mathbf{E}\left\|x^k\right\|^q$ (Lemma 4.1) then gives a constant $R > 0$ independent of k such that

$$\mathbf{E} \sup_{0<t-s<\delta} \left|z^k(t) - z^k(s)\right| \le R\delta$$

$$\implies \qquad \sup_k P\left\{ \sup_{0<t-s<\delta} \left|z^k(t) - z^k(s)\right| > \varepsilon \right\} \le R\delta/\varepsilon.$$

Tightness follows from Prop. 2.4. So instead of Prop. 5.2, we find that along a subsequence, *$(x_0^k, x^k, x^{1,k}, x^{2,k}, z^k) \xrightarrow{\mathcal{D}} (x_0, x, x^1, x^2, z)$ in the space $\mathbf{R}^n \times C^{4n}$.*

Thus we are led to consider the sets

$$\begin{aligned} S_N = \{(x_0, x, x^1, x^2, z) \, &: x_0 \in \mathbf{R}^n, \ (x, x^1, x^2, z) \in C^{4n}, \\ & x(t) = x_0 + x^1(t) + x^2(t), \quad x^1(0) = x^2(0) = 0, \\ & (x^1, z) \in AC^{n+1}, \quad \int_0^T \left|\dot{x}^1(r)\right|^{\bar{q}} dr \le N, \quad \int_0^T |\dot{z}(r)|^{\bar{q}/q} dr \le N, \\ & (\dot{x}^1(t), \dot{z}(t)) \in Q_0(t, x(\cdot)) \text{ a.e.}\} \end{aligned}$$

These sets are closed because of a different closure theorem than Prop. 5.3, namely Cesari (1983), 8.6.i, p. 299. By Def. 2.1(b), we find that the limiting quantity (x_0, x, x^1, x^2, z) lies in $\bigcup_{N=1}^{\infty} S_N$ almost surely. Just as in Prop. 5.5, Wong's theorem now implies that

$$x_t = x_0 + \int_0^t \mathbf{E}\left[\dot{x}_r^1 \mid \mathcal{F}_r^x\right] dr + \int_0^t \sigma(r, x)\, d\widetilde{w}_r.$$

Upon defining $\left(\widetilde{f}(t,\omega),\, \widetilde{L}(t,\omega)\right) = \mathbf{E}\left[\left(\dot{x}_t^1, \dot{z}_t\right) \,\middle|\, \mathcal{F}_t^x\right]$, the measurable selection theorem of Beneš gives an adapted control u_t such that

$$\left(\widetilde{f}(t,\omega), \widetilde{L}(t,\omega)\right) = \Big(f\big(t, x(\cdot,\omega), u(t,\omega)\big), L\big(t, x(\cdot,\omega), u(t,\omega)\big)\Big).$$

Hence (u, x) is a control-state pair solving (4.1).

Moreover, $\widetilde{L}$ is *exactly* L, not something larger as in Prop. 5.7. (This is critical.) Since $z^k(T) \xrightarrow{\mathcal{D}} z(T)$ and $\{z^k(T)\}$ is uniformly integrable by (h5), we have $\mathbf{E}z^k(T) \to \mathbf{E}z(T)$. Also, $\mathbf{E}\ell(x_T^k, \mathbf{E}x_T^k) \to \mathbf{E}\ell(x_T^k, \mathbf{E}x_T^k)$ by (h5) and the uniform integrability argument following equation (6.2) in paragraph 6.3. So indeed $\Lambda[u^k, x^k] \to \Lambda[u, x] = \lambda$. ////

Now let us compare Theorems 5.1 and 7.1. The practitioner who only needs an existence result will undoubtedly prefer Theorem 5.1. For hypotheses (h1)–(h4) are identical to (H1)–(H4) in the case of compact U, whereas (H5) now requires only that ℓ and L be bounded below—a considerably weaker condition than the growth conditions of (h5). Also, (h6) implies that $Q_0(t, x)$ is compact and convex for each (t, x), whence $Q(t, x) = Q_0(t, x) + \{(0, r) \;:\; r \geq 0\}$ is closed and convex: in the very common setting when f and L are continuous in t, this is a situation in which (H6) is known to hold. So in comparing Theorems 5.1(a) and 7.1(a), we see that Thm. 5.1(a) obtains the same conclusion as Thm. 7.1(a) under weaker assumptions.

However, the methods of Thm. 5.1 do not completely eclipse those of Thm. 7.1 because of the differences manifested in their (b) parts. These show that the weaker hypotheses of Thm. 5.1(b) lead only to a "lower-semicontinuity" result on Λ, whereas Thm. 7.1(b) concludes that Λ is "continuous" in some sense. Of course these different statements lose their distinctiveness when applied to a minimizing sequence in an attempt to prove existence, but there are situations in which the continuity conclusions of Thm. 7.1 have other uses. We will discuss such a case in Chapter V.

Appendix. Goor's Existence Theory.

Robert M. Goor (1976, 1979) has recently proposed certain new existence theorems for stochastic optimal control problems which appear to generalize the work of Kushner (1975). This appendix demonstrates that a key lemma used in Goor's "proof" of these results is false, and hence that his assertions about existence must be regarded as conjectures—not as theorems. Goor's statements about existence are distinguished by their strong formulation. In other words, the probability space $(\Omega, \mathcal{F}, P)$ and Brownian motion w_t are fixed throughout his arguments, rather than being considered as additional choice variables. The foundation for the strong approach is Lemma 1.7, p.909 of Goor (1976). It involves a separable Banach space Y and a complete separable metric space X. Both X and Y are equipped with their Borel σ-fields, and the identity map on X is denoted by i_X. A nonatomic probability measure μ is given on X.

Goor's Lemma. *Let $y_k : X \to Y$, $k = 1, 2, \ldots$, be a sequence of measurable maps such that $\int_X \|y_k(x)\| \, d\mu(x) < +\infty$ for each k and such that the sequence of probability measures $P_k = \mu(i_X, y_k)^{-1}$, defined on the Borel subsets of $X \times Y$, converges weakly to some probability measure P. Then there exists a measurable map $y: X \to Y$ such that $P = \mu(i_X, y)^{-1}$, i.e., (i_X, y_k) converges in distribution to (i_X, y).*

Counterexample. Let $X = [0,1]$ and write μ for Lebesgue measure. Taking $Y = \mathbf{R}$, define the sequence of simple measurable functions y_k: $[0,1] \to \mathbf{R}$ via $y_k(t) = t_k$, where $t = 0.t_1t_2t_3\ldots$ is the binary expansion of t. (To ensure that y_k is well-defined, we insist that the binary expansion of every $t \in [0,1)$ contain infinitely many zeros, and set $y_k(1) = 1 \; \forall k$.) It is a simple matter to verify that the probability measures P_k defined on the Borel sets of $[0,1] \times \mathbf{R}$ by

$$P_k(S) = \mu\left\{t \in [0,1] \; : \; \big(t, y_k(t)\big) \in S\right\}$$

converge weakly to the product measure $P = \mu \times \left(\frac{1}{2}\delta_{\{0\}} + \frac{1}{2}\delta_{\{1\}}\right)$. Goor's statement claims that there exists a measurable y: $[0,1] \to \mathbf{R}$ such that $P = \mu(i_X, y)^{-1}$. This is absurd. Indeed, if such a mapping y did exist, then clearly $y(t) = 0$ on a set $Z \subseteq [0,1]$ of measure $\mu(Z) = \frac{1}{2}$. In terms of Z, we may construct the measurable subset $S = Z \times \{1\}$ of $[0,1] \times \mathbf{R}$ for which $\mu\left\{t \; : \; \big(t, y(t)\big) \in S\right\} = 0$, whereas $P(S) = \frac{1}{2}\mu(Z) > 0$. This is a contradiction.

Goor's Proof. Goor's proof of the lemma quoted above is marred by a subtle misuse of multi-index notation. The basic idea is to construct mappings from the interval $[0,1]$ into $X \times Y$ which give rise to the laws P_k and P. Skorokhod (1965), p.10, provides a model for this construction: a sequence of nested partitions of the range space is used to define a corresponding sequence of nested partitions of the interval $[0,1]$. Skorokhod's construction ensures that the correspondence between the partitions of $X \times Y$ and of $[0,1]$ preserves the relationship of set inclusion; Goor's construction does not. So paragraph 5 of his proof does not really "follow Skorokhod's method." The mappings defined in that paragraph are *not* "defined in analogy to the construction of Skorokhod," so convergence may well fail. (In the case of the counterexample above, for instance, the mappings z_0^m do not converge at all.)

Of course, the conclusions of Goor's paragraph 5 can be rescued by taking more care with the correspondence between the partitions of $X \times Y$ and $[0,1]$. However, the incorrect correspondence is essential to the development of paragraph 6. If paragraph 5 is to be salvaged then the assertion that $h = h_k$, $k = 0, 1, 2, \ldots$ in paragraph 6 must be discarded. (This assertion is manifestly false in the counterexample discussed above.) But this statement is the key to the whole proof.

Acknowledgement. The author thanks Ulrich Haussmann and Ed Perkins for insisting on a second look at Goor's Lemma. Professor Perkins suggested the counterexample above.

Chapter V. Parameter Sensitivity in Stochastic Optimal Control

This chapter is devoted to a study of deterministic perturbations of the constrained stochastic control problem introduced in Chapter IV. Our approach is based on a proximal-normal analysis of the problem's value function, and hence makes explicit use of the existence theory of Chapter IV, the limiting techniques and representation theorems of Sections IV.2–3, and the unconstrained Stochastic Maximum Principle. The latter, which necessitates a smooth formulation, is presented in the first three sections below. Section 1 investigates how slight perturbations of the system's initial value and control law affect its evolution; in Section 2, the consequences of this variation for the cost functional are considered. These preliminary results allow the derivation of the unconstrained Stochastic Maximum Principle in Section 3. In that section we discuss the conclusions available in the nonanticipative formulation, and describe a general type of feedback formulation which gives more satisfying results.

Proximal normal analysis is the subject of Section 4, which is the heart of this chapter. There, a family of stochastic control problems indexed by a finite-dimensional parameter is used to define a "value function" whose generalized gradient is captured in Thm. 4.8. Section 5 explores some consequences of this characterization, which include a new proof of the Stochastic Maximum Principle for constrained problems.

Section 1. Perturbed Dynamics

Hypotheses. In this chapter we study stochastic systems of a more specialized form than those of Chap. IV. We still assume that $U \subseteq \mathbf{R}^m$ is a given closed set, but now the system's evolution on the fixed interval $[0,T]$ is determined by the *Markovian* Itô equation

$$x_t = x_0 + \int_0^t f(r, x_r, u_r)\, dr + \int_0^t \sigma(r, x_r)\, dw_r. \tag{1.1}$$

Moreover, at least for the first two sections, we assume that the filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is fixed in advance, and that it carries a d-dimensional $\mathcal{F}_t$-Brownian motion w_t which cannot be changed. The coefficients $f\colon [0,T] \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$ and $\sigma\colon [0,T] \times \mathbf{R}^n \to \mathbf{R}^{n\times d}$ must satisfy (H1)–(H2) below, which are stronger conditions than their counterparts in Chap. IV.

(H1) $f(t,x,\cdot)$ is a continuous function of $u \in U$, uniformly in (t,x); also, for each $(t,u) \in [0,T]\times U$, both $f(t,\cdot,u)$ and $\sigma(t,\cdot)$ are continuously differentiable functions of $x \in \mathbf{R}^n$.

(H2) There is a constant $k_1 > 0$ such that for all $(t,x,u) \in [0,T] \times \mathbf{R}^n \times U$, one has

$$|f(t,x,u)| + |\sigma(t,x)| \le k_1(1+|x|),$$

$$|f_x(t,x,u)| + |\sigma_x(t,x)| \le k_1.$$

Hypotheses (H1) and (H2) imply the Itô conditions. Thus the standard theory of stochastic differential equations implies that (on the fixed space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with w_t) any initial condition x_0 and any $\mathcal{F}_t$-adapted process $u\colon [0,T] \times \Omega \to U$ give rise to a pathwise unique process x_t solving (1.1). We need no longer discuss "control-state pairs" (u,x) as in Chap. IV, since the first element of such a pair uniquely specifies the second. Thus we will concentrate on the controls u. Note also that since the right-hand sides of the inequalities in (H2) are independent of u, Lemma IV.4.1 remains valid without any hypothesis concerning the integrability of u. The moments of the solution process can now be estimated solely in terms of the moments of its initial condition. This form of Lemma IV.4.1 will be used throughout this section.

In this chapter we consider a family of random initial conditions x_0 indexed by a parameter α in $\mathbf{R}^a$. Given a constant $\overline{q} > 2$, we fix our attention on a specific parameter value $\widehat{\alpha}$ and make the following hypothesis.

(H3) There are $\mathcal{F}_0$-measurable random vectors $X_0 \in \mathbf{R}^n$ and $A \in \mathbf{R}^{n\times a}$ such that $\mathbf{E}\,|X_0|^{\overline{q}} < +\infty$ and $|A(\omega)| \le k_2$ for all ω, and $x_0(\alpha) = X_0 + A(\alpha - \widehat{\alpha})$.

Note that when $X_0 = 0$, $a = n$, and $A = I$, (H3) allows the case of a variable deterministic initial condition $x_0(\alpha) = \alpha - \widehat{\alpha}$; and that when $\widehat{\alpha} = 0$, $a = 1$, and A is a bounded random variable in $\mathbf{R}^n$, (H3) allows the scheme $x_0(\alpha) = X_0 + \alpha A$ remeniscent of the calculus of variations.

Variations. Suppose any $\mathcal{F}_t$-adapted control process $\widehat{u}$ is given, and let $\widehat{x}$ denote the corresponding solution of (1.1). We wish to investigate the difference between $\widehat{x}$ and the solution x obtained when a pair (α, u) near $(\widehat{\alpha}, \widehat{u})$ is used. The following Lemma will help.

1.1 Lemma. *Let $\varphi: [0,T] \times \Omega \to \mathbf{R}^q$ be a measurable process for which $\int_0^T |\varphi(t,\omega)|\, dt < +\infty$ a.s. Then there is a null set $\mathcal{N} \subseteq [0,T]$ such that for each $t \notin \mathcal{N}$ one has*

$$\frac{d}{dt} \int_0^t \varphi(r,\omega)\, dr = \varphi(t,\omega) \qquad \text{a.s.}$$

Proof. See Kushner (1972), Lemma 1, p. 556. ////

To apply Lemma 1.1, first take $\varphi(t,\omega) = f(t, \widehat{x}(t,\omega), \widehat{u}(t,\omega))$ to obtain a null set $\mathcal{N}(\widehat{u})$. Then fix any other $\mathcal{F}_t$-adapted control process u_t and take $\varphi(t,\omega) = f(t, \widehat{x}(t,\omega), u(t,\omega))$ to obtain a second null set $\mathcal{N}(u)$. It follows that for all t outside the null set $\mathcal{N}(\widehat{u},u) = \mathcal{N}(\widehat{u}) \cup \mathcal{N}(u) \cup \{0\}$, one has

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{t-\varepsilon}^t \big(f(r,\widehat{x}_r,u_r) - f(r,\widehat{x}_r,\widehat{u}_r)\big)\, dr = f(t,\widehat{x}_t,u_t) - f(t,\widehat{x}_t,\widehat{u}_t) \quad \text{a.s.,} \tag{1.2}$$

To construct a family of perturbed controls, fix any $s \in [0,T] \setminus \mathcal{N}(\widehat{u},u)$, and define u_t^ϵ for each $\varepsilon > 0$ by

$$u_t^\epsilon := \begin{cases} \widehat{u}_t & \text{if } t \in [0, s-\varepsilon] \\ u_t & \text{if } t \in (s-\varepsilon, s] \\ \widehat{u}_t & \text{if } t \in (s,T]. \end{cases}$$

The perturbed parameter values will be $\alpha^\epsilon := \widehat{\alpha} + \varepsilon\alpha$. Let x^ϵ denote the solution of (1.1) corresponding to $(\alpha^\epsilon, u^\epsilon)$. In this section we will investigate the evolution of $\xi_t^\epsilon := x_t^\epsilon - \widehat{x}_t$.

To improve the readability of the results, we will use the following notation:

$$\Delta f^\epsilon(t,x) = f(t,x,u_t^\epsilon) - f(t,x,\widehat{u}_t),$$

$$\widehat{f}(t) = f(t,\widehat{x}(t),\widehat{u}(t)), \qquad \widehat{f}_x(t) = f_x(t,\widehat{x}(t),\widehat{u}(t)),$$

$$\widehat{\sigma}(t) = \sigma(t,\widehat{x}(t)), \qquad \widehat{\sigma}_x(t) = \sigma_x(t,\widehat{x}(t)),$$

$$\delta(\varepsilon) \sim 0 \Leftrightarrow \lim_{\varepsilon \to 0+} \delta(\varepsilon) = 0.$$

The key result is Proposition 1.6: it states that the discrepancy ξ_t^ϵ, which is defined by

$$\begin{aligned} \xi_t^\epsilon = \big(x_0(\alpha^\epsilon) - x_0(\widehat{\alpha})\big) &+ \int_0^t \big(f(r,x_r^\epsilon,\widehat{u}_r) - f(r,\widehat{x}_r,\widehat{u}_r) + \Delta f^\epsilon(r,x_r^\epsilon)\big)\, dr \\ &+ \int_0^t \big(\sigma(r,x_r^\epsilon) - \sigma(r,\widehat{x}_r)\big)\, dw_r, \end{aligned} \tag{1.3}$$

is very well approximated by the variation-of-parameters solution of the inhomogeneous linearized equation

$$y_t^\epsilon = \varepsilon A\alpha + \int_0^t \big(\widehat{f}_x(r) y_r^\epsilon + \Delta f^\epsilon(r,\widehat{x}_r)\big)\, dr + \int_0^t \widehat{\sigma}_x(r) y_r^\epsilon\, dw_r.$$

1.2 Lemma. *For each $p \in [0, \overline{q}]$, there is a constant $M > 0$ for which*

$$\mathbf{E}\, \|\xi^\epsilon\|^p \le M\varepsilon^p \qquad \forall \varepsilon > 0.$$

Proof. Observe first that by Jensen's inequality we have

$$\mathbf{E}\, \|\xi^\epsilon\|^p \le \left(\mathbf{E}\, \|\xi^\epsilon\|^{\overline{q}}\right)^{p/\overline{q}} \qquad \forall p \in [0, \overline{q}].$$

Hence the general result follows from that for the case $p = \overline{q}$, which we now treat.

Note that $\Delta f^\epsilon(r, x) = 0$ unless $r \in (s - \varepsilon, s]$. Hence there is a constant $K_1 > 0$ for which Jensen's inequality gives

$$\begin{aligned} |\xi^\epsilon_\tau|^{\overline{q}} \le K_1 \Big(&|x_0(\alpha^\epsilon) - x_0(\widehat{\alpha})|^{\overline{q}} + \int_0^\tau |f(r, x^\epsilon_r, \widehat{u}_r) - f(r, \widehat{x}_r, \widehat{u}_r)|^{\overline{q}}\, dr \\ &+ \varepsilon^{\overline{q}-1} \int_0^\tau |\Delta f^\epsilon(r, x^\epsilon_r)|^{\overline{q}}\, dr + \left| \int_0^\tau \left(\sigma(r, x^\epsilon_r) - \sigma(r, \widehat{x}_r)\right) dw_r \right|^{\overline{q}} \Big) \end{aligned}$$

for all $\tau \in [0, T]$. Taking the supremum over $\tau \in [0, t]$ and applying (BDG) gives, for some $K_2 > 0$,

$$\begin{aligned} \mathbf{E}\, \|\xi^\epsilon\|^{\overline{q}}_t \le K_2 \mathbf{E} \Big(&|x_0(\alpha^\epsilon) - x_0(\widehat{\alpha})|^{\overline{q}} + \int_0^t |f(r, x^\epsilon_r, \widehat{u}_r) - f(r, \widehat{x}_r, \widehat{u}_r)|^{\overline{q}}\, dr \\ &+ \varepsilon^{\overline{q}-1} \int_0^t |\Delta f^\epsilon(r, x^\epsilon_r)|^{\overline{q}}\, dr + \int_0^t |\sigma(r, x^\epsilon_r) - \sigma(r, \widehat{x}_r)|^{\overline{q}}\, dr \Big). \end{aligned} \tag{1.4}$$

Now $|x_0(\alpha^\epsilon) - x_0(\widehat{\alpha})| \le k_2 \varepsilon\, |\alpha|$ by (H3), and

$$|f(r, x^\epsilon_r, \widehat{u}_r) - f(r, \widehat{x}_r, \widehat{u}_r)| \le k_1\, |\xi_r|,$$

$$|\sigma(r, x^\epsilon_r) - \sigma(r, \widehat{x}_r)| \le k_1\, |\xi_r| \qquad \text{(by (H2))}.$$

Moreover, the linear growth condition of (H2) implies that

$$|\Delta f^\epsilon(r, x^\epsilon_r)| \le |f(r, x^\epsilon_r, u^\epsilon_r)| + |f(r, x^\epsilon_r, \widehat{u}_r)| \le 2k_1\left(1 + |x^\epsilon_r|\right),$$

so there are constants K_3, K_4 such that

$$\mathbf{E}\, |\Delta f^\epsilon(r, x^\epsilon_r)|^{\overline{q}} \le K_3\left(1 + \mathbf{E}\, |x^\epsilon_r|^{\overline{q}}\right) \le K_4.$$

(The second inequality follows from Lemma IV.4.1.) Using these facts in (1.4) implies that, for some constant K_5,

$$\mathbf{E}\, \|\xi^\epsilon\|^{\overline{q}}_t \le K_5 \left(\varepsilon^{\overline{q}} + \int_0^t \mathbf{E}\, \|\xi^\epsilon\|^{\overline{q}}_r\, dr\right) \qquad \forall t \in [0, T].$$

The desired result follows from Gronwall's inequality. ////

1.3 Corollary. *There is a constant $M > 0$ for which one has*

$$\mathbf{E} \sup_{t\in[0,T]} |\Delta f^\epsilon(t, x_t^\epsilon) - \Delta f^\epsilon(t, \widehat{x}_t)|^{\overline{q}} \leq M\varepsilon^{\overline{q}} \qquad \forall \varepsilon > 0.$$

Proof. The Lipschitz condition implied by (H2) gives

$$\begin{aligned} |\Delta f^\epsilon(t, x_t^\epsilon) - \Delta f^\epsilon(t, \widehat{x}_t)| &\leq |f(t, x_t^\epsilon, u_t^\epsilon) - f(t, \widehat{x}_t, u_t^\epsilon)| + |f(t, x_t^\epsilon, \widehat{u}_t) - f(t, \widehat{x}_t, \widehat{u}_t)| \\ &\leq 2k_1 |\xi_t^\epsilon|. \end{aligned}$$

Thus the result follows directly from Lemma 1.2. ////

Now that we know ξ^ϵ is a process relatively small in magnitude, we turn with confidence to the linearization of (1.1) about $(\widehat{\alpha}, \widehat{x}, \widehat{u})$:

$$y_t^\epsilon = \varepsilon A\alpha + \int_0^t \left(\widehat{f}_x(r) y_r^\epsilon + \Delta f^\epsilon(r, \widehat{x}_r)\right) dr + \int_0^t \widehat{\sigma}_x(r) y_r^\epsilon \, dw_r. \tag{1.5}$$

Note that since σ is an $n \times d$ matrix, we must interpret

$$\sigma_x y \, dw = \sum_{k=1}^d \sigma_x^k y \, dw^k,$$

where σ^k is the k-th column of σ and w^k is the k-th component of w. Note also that equation (1.5) satisfies the Itô conditions, so has a pathwise unique solution y_t^ϵ for every $\varepsilon > 0$. Let us compare y_t^ϵ with ξ_t^ϵ.

1.4 Lemma. *For any $p \in [2, \overline{q})$, there is a function $\delta(\varepsilon) \sim 0$ such that*

$$\mathbf{E}\, \|\xi^\epsilon - y^\epsilon\|^p \leq \varepsilon^p \delta(\varepsilon) \qquad \forall \varepsilon > 0.$$

Proof. Let $z_t^\epsilon = \xi_t^\epsilon - y_t^\epsilon$. Then by the mean-value theorem there are a constant θ^ϵ and processes φ_t^ϵ, ψ_t^ϵ, all with values in $[0,1]$, for which

$$z_t^\epsilon = \int_0^t \widehat{f}_x(r) z_r^\epsilon \, dr + \int_0^t \widehat{\sigma}_x(r) z_r^\epsilon \, dw_r + H^\epsilon + I_t^\epsilon + J_t^\epsilon + K_t^\epsilon,$$

where

$$\begin{aligned} H^\epsilon &= \varepsilon\big(Dx_0(\widehat{\alpha} + \varepsilon\theta^\epsilon\alpha) - A\big)\alpha, \\ I_t^\epsilon &= \int_0^t \left[\Delta f^\epsilon(r, x_r^\epsilon) - \Delta f^\epsilon(r, \widehat{x}_r)\right] dr, \\ J_t^\epsilon &= \int_0^t \left[\sigma_x(r, \widehat{x}_r + \psi_t^\epsilon \xi_r^\epsilon) - \widehat{\sigma}_x(r)\right] \xi_r^\epsilon \, dw_r, \\ K_t^\epsilon &= \int_0^t \left[f_x(r, \widehat{x}_r + \varphi_r^\epsilon \xi_r^\epsilon, \widehat{u}_r) - \widehat{f}_x(r)\right] \xi_r^\epsilon \, dr. \end{aligned}$$

Now for any $p \in [2, \overline{q})$,

$$\mathbf{E}\left|H^\epsilon\right|^p \leq \varepsilon^p \left|\alpha\right|^p \mathbf{E}\left|Dx_0(\widehat{\alpha} + \varepsilon\theta^\epsilon\alpha) - A\right|^p =: \varepsilon^p h(\varepsilon). \tag{1.6}$$

The function $h(\varepsilon) \sim 0$ because Dx_0 is continuous for each ω and uniformly bounded by (H3).

Next, Corollary 1.3 gives a constant M such that

$$\begin{aligned}
\mathbf{E}\left\|I^\epsilon\right\|^p &\leq \mathbf{E}\left(\int_{s-\epsilon}^{s} \left|\Delta f^\epsilon(r, x_r^\epsilon) - \Delta f^\epsilon(r, \widehat{x}_r)\right| dr\right)^p \\
&\leq \mathbf{E}\varepsilon^{p-1} \int_{s-\epsilon}^{s} \left|\Delta f^\epsilon(r, x_r^\epsilon) - \Delta f^\epsilon(r, \widehat{x}_r)\right|^p dr \\
&\leq \varepsilon^{p-1} \int_{s-\epsilon}^{s} M\varepsilon^p \, dr \\
&\leq M\varepsilon^{2p}.
\end{aligned} \tag{1.7}$$

To control J_t^ϵ, we use (BDG):

$$\begin{aligned}
\mathbf{E}\left\|J^\epsilon\right\|^p &\leq C_p \mathbf{E} \int_0^T \left|\left(\sigma_x(r, \widehat{x}_r + \psi_r^\epsilon \xi_r^\epsilon) - \sigma_x(r, \widehat{x}_r)\right)\xi_r^\epsilon\right|^p dr \\
&\leq C_p \mathbf{E}\left(\left\|\xi^\epsilon\right\|^p \int_0^T \left|\sigma_x(r, \widehat{x}_r + \psi_r^\epsilon \xi_r^\epsilon) - \widehat{\sigma}_x(r)\right|^p dr\right) \\
&\leq C_p' \left(\mathbf{E}\left\|\xi^\epsilon\right\|^{\overline{q}}\right)^{p/\overline{q}} \left(\int_0^T \mathbf{E}\left|\sigma_x(r, \widehat{x}_r + \psi_r^\epsilon \xi_r^\epsilon) - \widehat{\sigma}_x(r)\right|^{p\overline{q}/(\overline{q}-p)} dr\right)^{(\overline{q}-p)/\overline{q}} \\
&\leq \varepsilon^p j(\varepsilon).
\end{aligned} \tag{1.8}$$

In the last step, the second factor is bounded by a multiple of ε^p by Lemma 1.2, and the function $j(\varepsilon)$ is then defined as the appropriate constant multiple of the third factor. Note that $j(\varepsilon) \sim 0$ because $\sigma_x(r, \cdot)$ is a bounded continuous function and ξ_r^ϵ tends to 0 in probability for each r.

In just the same way, there is a function $k(\varepsilon) \sim 0$ for which

$$\begin{aligned}
\mathbf{E}\left\|K^\epsilon\right\|^p &\leq C_p \left(\mathbf{E}\left\|\xi^\epsilon\right\|^{\overline{q}}\right)^{p/\overline{q}} \left(\int_0^T \mathbf{E}\left|f_x(r, \widehat{x}_r + \varphi_r^\epsilon \xi_r^\epsilon) - \widehat{f}_x(r)\right|^{p\overline{q}/(\overline{q}-p)} dr\right)^{(\overline{q}-p)/\overline{q}} \\
&\leq \varepsilon^p k(\varepsilon).
\end{aligned} \tag{1.9}$$

Let us combine estimates (1.6)–(1.9) into a function $\delta(\varepsilon) \sim 0$ for which $\mathbf{E}\big(\left|H^\epsilon\right|^p + \left\|I^\epsilon\right\|^p + \left\|J^\epsilon\right\|^p + \left\|K^\epsilon\right\|^p\big) \leq \varepsilon^p \delta(\varepsilon)$. Then in the original equation for z_t^ϵ, we find using (BDG) and (H2) that for some $K > 0$,

$$\mathbf{E}\left\|z^\epsilon\right\|_t^{\overline{p}} \leq K\left(\int_0^t \mathbf{E}\left\|z^\epsilon\right\|_r^p dr + \varepsilon^p \delta(\varepsilon)\right).$$

Gronwall's inequality now implies that $\mathbf{E}\left\|z^\epsilon\right\|^p \leq M\varepsilon^p\delta(\varepsilon)$ for some $M > 0$, as required. ////

To complete our study of the evolution of ξ_t^ϵ, we use the variation of parameters formula to write the approximate solution of the inhomogeneous equation (1.5) in terms of its homogeneous counterpart

$$(1.10)\qquad d\varphi_t = \widehat{f}_x(t)\varphi_t\,dt + \widehat{\sigma}_x(t)\varphi_t\,dw_t.$$

Let $\Phi(t,\tau)$ be the fundamental matrix solution of this linear equation: that is, for each $\tau \in [0,T)$, $\Phi(\cdot,\tau)$ obeys

$$\Phi(t,\tau) = I + \int_\tau^t \widehat{f}_x(r)\Phi(r,\tau)\,dr + \int_\tau^t \widehat{\sigma}_x(r)\Phi(r,\tau)\,dw_r.$$

Note that since f_x and σ_x are uniformly bounded by (H2), each of the moments $\mathbf{E}\,|\Phi(t,\tau)|^p$ for $p \in [1,\infty)$ admits a constant upper bound independent of t and τ. (Compare Lemma IV.4.1 and paragraph 3.4, below.) This fact sponsors the final result of this section, which renders precise the approximate statement that for each t,

$$\xi_t^\epsilon \cong \varepsilon\big(\Phi(t,0)A\alpha + I\{t \geq s\}\Phi(t,s)\big[f(s,\widehat{x}_s,u_s) - f(s,\widehat{x}_s,\widehat{u}_s)\big]\big).$$

1.5 Proposition. *For any $p \in [2,\overline{q})$ there is a function $\delta(\varepsilon) \sim 0$ for which*

$$\mathbf{E}\,\big|\xi_t^\epsilon - \varepsilon\big(\Phi(t,0)A\alpha + I\{t \geq s\}\Phi(t,s)\big[f(s,\widehat{x}_s,u_s) - f(s,\widehat{x}_s,\widehat{u}_s)\big]\big)\big|^p \leq \varepsilon^p\delta(\varepsilon)$$

holds for all $t \in [0, s-\varepsilon] \cup [s,T]$ and for all $\varepsilon > 0$. For $t \in (s-\varepsilon, s)$ there is a constant $M > 0$ for which the inequality remains valid if the right-hand side is replaced by $\varepsilon^p M$.

Proof. In view of Lemma 1.4, it suffices to prove this estimate with ξ_t^ϵ replaced by y_t^ϵ. And by definition of Φ, the difference

$$y_t^\epsilon - \varepsilon\big(\Phi(t,0)A\alpha + I\{t \geq s\}\Phi(t,s)\big[f(s,\widehat{x}_s,u_s) - f(s,\widehat{x}_s,\widehat{u}_s)\big]\big)$$

is zero for all $t \in [0, s-\varepsilon]$, a.s. We therefore advance to the case $t \in (s-\varepsilon, s)$, where we have

$$(1.11)\qquad y_t^\epsilon = \varepsilon\Phi(s-\varepsilon,0)A\alpha + \int_{s-\varepsilon}^t \widehat{f}_x(r)y_r^\epsilon\,dr + \int_{s-\varepsilon}^t \widehat{\sigma}_x(r)y_r^\epsilon\,dw_r + \int_{s-\varepsilon}^t \Delta f^\epsilon(r,\widehat{x}_r)\,dr.$$

The first term obeys

$$(1.12a)\qquad \mathbf{E}\,|\varepsilon\Phi(s-\varepsilon,0)A\alpha - \varepsilon\Phi(s,0)A\alpha|^p \leq K\varepsilon^p\mathbf{E}\,|\Phi(s-\varepsilon,0) - \Phi(s,0)|^p =: \varepsilon^p\delta_0\,(\varepsilon).$$

Since Φ is jointly continuous and has moments of all orders, uniform integrability ensures that $\delta_0(\varepsilon) \sim 0$. The first integral obeys

$$(1.12b) \qquad \left|\int_{s-\varepsilon}^{t} \widehat{f}_x(r) y_r^\epsilon \, dr\right|^p \le \varepsilon^{p-1} \|y^\epsilon\|^p \int_{s-\varepsilon}^{s} \left|\widehat{f}_x(r)\right|^p dr \le k_1^p \varepsilon^p \|y^\epsilon\|^p,$$

and $\mathbf{E}\|y^\epsilon\|^p = \varepsilon^{p-1}\delta_0(\varepsilon)$ for some $\delta_0 \sim 0$ by Lemmas 1.2 and 1.4. For the second integral, (BDG) gives

$$(1.12c) \qquad \mathbf{E}\left|\int_{s-\varepsilon}^{t} \widehat{\sigma}_x(r) y_r^\epsilon \, dw_r\right|^p \le C_p \mathbf{E}\|y^\epsilon\|^p \int_{s-\varepsilon}^{s} |\widehat{\sigma}_x(r)|^p \, dr \le C_p k_1^p \varepsilon \mathbf{E}\|y^\epsilon\|^p.$$

So the first two integrals introduce an error bounded by $\varepsilon^p \delta_1(\varepsilon)$ for some $\delta_1 \sim 0$. As for the third, we can only say

$$(1.13) \qquad \mathbf{E}\left|\int_{s-\varepsilon}^{t} \Delta f^\epsilon(r, \widehat{x}_r) \, dr\right|^p \le \varepsilon^{p-1} \int_{s-\varepsilon}^{s} \mathbf{E}\left|\Delta f^\epsilon(r, \widehat{x}_r)\right|^p dr \le K_0 \varepsilon^p$$

for some constant K_0 which provides an upper bound for the integrand. (Such a bound exists by (H2) and Lemma IV.4.1.) So for $t \in (s-\varepsilon, s)$, we have a constant M such that

$$\mathbf{E}\left|y_t^\epsilon - \varepsilon\Phi(s,0)A\alpha\right|^p \le M\varepsilon^p.$$

Now when $t = s$, the three estimates (1.12) remain valid, and (1.13) may be strengthened: we have

$$(1.14) \qquad \begin{aligned} \mathbf{E}\left|\int_{s-\varepsilon}^{s} \Delta f^\epsilon(r, \widehat{x}_r) \, dr - \varepsilon \Delta f^\epsilon(s, \widehat{x}_s)\right|^p &= \varepsilon^p \mathbf{E}\left|\frac{1}{\varepsilon}\int_{s-\varepsilon}^{s} \Delta f^\epsilon(r, \widehat{x}_r) \, dr - \Delta f^\epsilon(s, \widehat{x}_s)\right|^p \\ &=: \varepsilon^p \delta_2(\varepsilon). \end{aligned}$$

Here $\delta_2(\varepsilon) \sim 0$ because the difference tends to 0 a.s. by (1.2), and is uniformly integrable. To prove uniform integrability, simply observe that for some $K > 0$,

$$\begin{aligned} &\mathbf{E}\left|\varepsilon^{-1}\int_{s-\varepsilon}^{s} \Delta f^\epsilon(r, \widehat{x}_r) \, dr - \Delta f^\epsilon(s, \widehat{x}_s)\right|^{\overline{q}} \\ &\qquad \le K\left(\varepsilon^{-1}\int_{s-\varepsilon}^{s} \mathbf{E}\left|\Delta f^\epsilon(r, \widehat{x}_r)\right|^{\overline{q}} dr + \mathbf{E}\left|\Delta f^\epsilon(s, \widehat{x}_s)\right|^{\overline{q}}\right), \end{aligned}$$

and the right side admits a constant upper bound uniformly in ε by Lemma IV.4.1 and the linear growth condition in (H2).

Now for all $t > s$, $\Delta f^\epsilon(t,\hat{x}_t) = 0$, so (1.5) implies that $y_t^\epsilon = \Phi(t,s)y_s^\epsilon$. For any fixed p in $[2,\bar{q})$, we choose a $q \in (p,\bar{q})$. Then the uniform boundedness of the moments of Φ gives a constant $K > 0$ for which

$$\begin{aligned}
\mathbf{E}\left|y_t^\epsilon - \varepsilon\Phi(t,s)\Delta f^\epsilon(s,\hat{x}_s)\right|^p &\le \mathbf{E}\left|\Phi(t,s)\right|^p \left|y_s^\epsilon - \varepsilon\Delta f^\epsilon(s,\hat{x}_s)\right|^p \\
&\le \left(\mathbf{E}\left|\Phi(t,s)\right|^{pq/(q-p)}\right)^{(q-p)/q} \left(\mathbf{E}\left|y_s^\epsilon - \varepsilon\Delta f^\epsilon(s,\hat{x}_s)\right|^q\right)^{p/q} \\
&\le K\left(\varepsilon^q\delta_q(\varepsilon)\right)^{p/q}
\end{aligned}$$

by the result of the previous paragraph. The choice of $\delta(\varepsilon) = K\left(\delta_q(\varepsilon)\right)^{p/q}$ completes the proof. ////

Section 2. The Cost Functional

Hypotheses. We continue to work with the fixed probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and Brownian motion w_t of Section 1 throughout this section. The cost of a given initial condition α and $\mathcal{F}_t$-adapted control process $u\colon [0,T]\times\Omega \to U$ is measured by the functional

$$\Lambda[\alpha,u] := \mathbf{E}\left[\langle\beta,\alpha\rangle + \ell(x_T) + \int_0^T L(t,x_t,u_t)\,dt\right].$$

Here, as in Section 1, x_t denotes the pathwise unique strong solution of the Itô equation (1.1) under Hypotheses (H1)–(H3). The constant vector $\beta \in \mathbf{R}^a$ is included for later theoretical use, even though many applied problems are covered by the case $\beta = 0$. The pointwise cost $\ell\colon \mathbf{R}^n \to \mathbf{R}$ and the running cost $L\colon [0,T]\times\mathbf{R}^n\times U \to \mathbf{R}$ must satisfy (H4) and (H5) below.

(H4) $L(t,x,\cdot)$ is a continuous function of $u \in U$, uniformly in (t,x); also, for each fixed $(t,u) \in [0,T]\times U$, both ℓ and $L(t,\cdot,u)$ are continuously differentiable.

(H5) There is an exponent $q \in [1,\bar{q}-1)$ and a constant $k_3 > 0$ such that for all $(t,x,u) \in [0,T]\times\mathbf{R}^n\times U$, one has

$$|L(t,x,u)| \le k_3(1+|x|^q+|u|^q), \qquad |L_x(t,x,u)| \le k_3(1+|x|^{q-1}+|u|^q),$$
$$|\ell(x)| \le k_3(1+|x|^q), \qquad |\ell_x(x)| \le k_3(1+|x|^{q-1}).$$

The Itô conditions and (H5) imply that $\Lambda[\alpha,u]$ is well-defined and finite for any $\mathcal{F}_t$-adapted stochastic process $u\colon [0,T]\times\Omega \to U$ obeying

$$\mathbf{E}\int_0^T |u_r|^{\bar{q}}\,dr < +\infty. \tag{2.1}$$

Variations. Suppose now that the control strategies $\widehat{u}_t$ and u_t featured in Section 1 actually obey condition (2.1). Then so does u^ϵ for every $\varepsilon > 0$, so $\Lambda[\alpha^\epsilon, u^\epsilon]$ is well-defined for all $\varepsilon > 0$. Moreover, (H4) and (H5) verify the hypotheses of Lemma 1.1 for the functions $\varphi(t,\omega) = L(t, \widehat{x}(t,\omega), u(t,\omega))$ and $\widehat{\varphi}(t,\omega) = L(t, \widehat{x}(t,\omega), \widehat{u}(t,\omega))$. Hence we may enlarge the set $\mathcal{N}(\widehat{u}, u)$ if necessary and assume that the analogue of (1.2) holds with L replacing f. That is, $t \notin \mathcal{N}(\widehat{u}, u)$ implies

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_{t-\varepsilon}^{t} \big(L(r, \widehat{x}_r, u_r) - L(r, \widehat{x}_r, \widehat{u}_r)\big)\, dr = L(t, \widehat{x}_t, u_t) - L(t, \widehat{x}_t, \widehat{u}_t) \quad \text{a.s.} \tag{2.2}$$

We now study the difference $\Lambda[\alpha^\epsilon, u^\epsilon] - \Lambda[\widehat{\alpha}, \widehat{u}]$ as $\varepsilon \to 0^+$.

Corresponding to the notation of Section 1, we set

$$\Delta L^\epsilon(t,x) = L(t,x,u_t^\epsilon) - L(t,x,\widehat{u}_t),$$

$$\widehat{L}(t) = L(t, \widehat{x}_t, \widehat{u}_t), \qquad \widehat{L}_x(t) = L_x(t, \widehat{x}_t, \widehat{u}_t),$$

$$\Lambda^\epsilon = \Lambda[\alpha^\epsilon, u^\epsilon], \qquad \widehat{\Lambda} = \Lambda[\widehat{\alpha}, \widehat{u}].$$

In terms of the difference process $\xi^\epsilon = x^\epsilon - \widehat{x}$, we have

$$\Lambda^\epsilon - \widehat{\Lambda} = \mathbf{E}\left[\varepsilon\,\langle\beta,\alpha\rangle + \ell(\widehat{x}_T + \xi_T^\epsilon) - \ell(\widehat{x}_T) + \int_0^T \big(L(r, x_r^\epsilon, \widehat{u}_r) - \widehat{L}(r) + \Delta L^\epsilon(r, x_r^\epsilon)\big)\, dr\right].$$

The following linearization result is the natural counterpart of Prop. 1.5.

2.1 Proposition. *There is a function $\delta(\varepsilon) \sim 0$ for which $\Lambda^\epsilon - \widehat{\Lambda}$ equals*

$$\begin{aligned} &\varepsilon\mathbf{E}\Bigg[\langle\beta,\alpha\rangle + \left(\ell_x(\widehat{x}_T)\Phi(T,0) + \int_0^T \widehat{L}_x(r)\Phi(r,0)\, dr\right) A\alpha \\ &\qquad + \left(\ell_x(\widehat{x}_T)\Phi(T,s) + \int_s^T \widehat{L}_x(r)\Phi(r,s)\, dr\right)\Delta f^\epsilon(s, \widehat{x}_s) \\ &\qquad + \Delta L^\epsilon(s, \widehat{x}_s)\Bigg] + \varepsilon\delta(\varepsilon). \end{aligned} \tag{2.3}$$

(The $n \times n$ matrix $\Phi(t,\tau)$ was defined in (1.10).)

Proof. First we estimate the variation in the running cost. By the Mean Value Theorem there is a scalar process θ_t with values in $[0,1]$ such that

$$\begin{aligned} \int_0^T \big(L(r, x_r^\epsilon, \widehat{u}_r) - \widehat{L}(r)\big)\, dr &= \int_0^T L_x(r, \widehat{x}_r + \theta_r\xi_r^\epsilon, \widehat{u}_r)\xi_r^\epsilon\, dr \\ &= \varepsilon\int_0^T \widehat{L}_x(r)\big[\Phi(r,0)A\alpha + I\{r \geq s\}\Phi(r,s)\Delta f^\epsilon(s, \widehat{x}_s)\big]\, dr + I^\epsilon + J^\epsilon, \end{aligned} \tag{2.4}$$

where
$$I^\epsilon = \int_0^T \widehat{L}_x(r)\left[\xi_r^\epsilon - \varepsilon\left(\Phi(r,0)A\alpha + I\{r \geq s\}\Phi(r,s)\Delta f^\epsilon(s,\widehat{x}_s)\right)\right] dr,$$
$$J^\epsilon = \int_0^T \left(L_x(r,\widehat{x}_r + \theta_r\xi_r^\epsilon, \widehat{u}_r) - \widehat{L}_x(r)\right)\xi_r^\epsilon\, dr.$$

Now for a constant $p < \overline{q}$ sufficiently near $\overline{q}$ the conjugate exponent obeys $\dfrac{pq}{p-1} < \overline{q}$, so Prop. 1.5 gives a constant $M > 0$ and a function $\delta_0 \sim 0$ such that

$$\begin{aligned}
\mathbf{E}\,|I^\epsilon| &\leq \int_0^T \left(\mathbf{E}\left|\widehat{L}_x(r)\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} \left(\mathbf{E}\left|\xi_r^\epsilon - \varepsilon\left(\Phi(r,0)A\alpha + I\{r\geq s\}\Phi(r,s)\Delta f^\epsilon(s,\widehat{x}_s)\right)\right|^p\right)^{\frac{1}{p}} dr \\
&\leq M\varepsilon \int_{s-\varepsilon}^{s} \left(\mathbf{E}\left|\widehat{L}_x(r)\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} dr + \varepsilon\delta_0(\varepsilon)\int_0^T \left(\mathbf{E}\left|\widehat{L}_x(r)\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} dr \\
&=: \varepsilon\delta_1(\varepsilon).
\end{aligned}$$

Since $\mathbf{E}\left|\widehat{L}_x(r)\right|^{\frac{p}{p-1}} \leq K\left(1 + \mathbf{E}\,|\widehat{x}_r|^{\overline{q}} + \mathbf{E}\,|\widehat{u}_r|^{\overline{q}}\right)$ is integrable, so is its $\dfrac{p-1}{p}$ power: hence $\delta_1(\varepsilon) \sim 0$. Similarly, Lemma 1.2 implies that some $K > 0$ obeys

$$\begin{aligned}
\mathbf{E}\,|J^\epsilon| &\leq K\varepsilon \int_0^T \left(\mathbf{E}\left|L_x(r,\widehat{x}_r + \theta_r\xi_r^\epsilon,\widehat{u}_r) - \widehat{L}_x(r)\right|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}} dr \\
&=: \varepsilon\delta_2(\varepsilon).
\end{aligned}$$

Here $\delta_2(\varepsilon) \sim 0$ because any sequence $\varepsilon_n \to 0^+$ has a subsequence along which $\|\xi^{\epsilon_n}\| \to 0$ a.s. by Lemma 1.2. Uniform integrability gives the result.

Next we study

$$\int_0^T \Delta L^\epsilon(r,x_r^\epsilon)\, dr = \varepsilon\Delta L^\epsilon(s,\widehat{x}_s) + I^\epsilon + J^\epsilon, \tag{2.5}$$

where
$$I^\epsilon = \varepsilon\left(\frac{1}{\varepsilon}\int_{s-\varepsilon}^{s} \Delta L^\epsilon(r,\widehat{x}_r)\, dr - \Delta L^\epsilon(s,\widehat{x}_s)\right),$$
$$J^\epsilon = \int_{s-\varepsilon}^{s} \left(\Delta L^\epsilon(r,x_r^\epsilon) - \Delta L^\epsilon(r,\widehat{x}_r)\right) dr.$$

Now the existence of $i(\varepsilon) \sim 0$ such that $\mathbf{E}\,|I^\epsilon| \leq \varepsilon i(\varepsilon)$ follows from assumption (2.2) just as line (1.14) followed from (1.2). And

$$J^\epsilon = \int_{s-\varepsilon}^{s} \left(L(r,x_r^\epsilon,u_r) - L(r,\widehat{x}_r,u_r)\right) dr + \int_{s-\varepsilon}^{s} \left(L(r,\widehat{x}_r,\widehat{u}_r) - L(r,x_r^\epsilon,\widehat{u}_r)\right) dr.$$

These two integrals can be treated similarly, so we discuss only the first one. By the Mean Value theorem there is a process θ_r with values in $[0,1]$ such that the expectation of the first integral is majorized by

$$\begin{aligned}\mathbf{E}\int_{s-\varepsilon}^{s} &|L_x(r,\widehat{x}_r+\theta_r\xi_r^\varepsilon,u_r)\xi_r^\varepsilon|\,dr\\ &\le \int_{s-\varepsilon}^{s}\left(\mathbf{E}\,|L_x(r,\widehat{x}_r+\theta_r\xi_r^\varepsilon,u_r)|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\left(\mathbf{E}\,\|\xi^\varepsilon\|^p\right)^{\frac{1}{p}}\,dr\\ &\le K\varepsilon\int_{s-\varepsilon}^{s}\left(\mathbf{E}\,|L_x(r,\widehat{x}_r+\theta_r\xi_r^\varepsilon,u_r)|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\,dr\\ &=:\varepsilon j(\varepsilon).\end{aligned}$$

Here $j(\varepsilon)\sim 0$ because the growth conditions of (H5) and (2.1) imply that the quantity in parentheses is integrable.

Finally, we consider the point costs. These obey

$$\mathbf{E}\big(\ell(x_T^\varepsilon)-\ell(\widehat{x}_T)\big)=\mathbf{E}\Big[\varepsilon\ell_x(\widehat{x}_T)\big(\Phi(T,0)A\alpha+\Phi(T,s)\Delta f^\varepsilon(s,\widehat{x}_s)\big)+I^\varepsilon+J^\varepsilon\Big] \tag{2.6}$$

where

$$\begin{aligned}I^\varepsilon&=\ell_x(\widehat{x}_T)\Big[\xi_T^\varepsilon-\varepsilon\big(\Phi(T,0)A\alpha+\Phi(T,s)\Delta f^\varepsilon(s,\widehat{x}_s)\big)\Big],\\ J^\varepsilon&=\Big(\ell_x(\widehat{x}_T+\theta^\varepsilon\xi_T^\varepsilon)-\ell_x(\widehat{x}_T)\Big)\xi_T^\varepsilon\end{aligned}$$

for some $\theta^\varepsilon\in[0,1]$. Hölder's inequality, uniform integrability, and Lemma 1.2 give a function $j(\varepsilon)\sim 0$ for which $\mathbf{E}\,|J^\varepsilon|\le\varepsilon j(\varepsilon)$ much as in the arguments above. Likewise, $\mathbf{E}\,|I^\varepsilon|\le\varepsilon i(\varepsilon)$ for some $i(\varepsilon)\sim 0$ follows from Prop. 1.5.

Combining (2.4), (2.5), and (2.6) gives the desired result. ////

Section 3. Necessary Conditions for an Unconstrained Stochastic Control Problem

The data of the previous sections allow the formulation of several stochastic control problems. The most readily comprehensible of these is *the strong problem with nonanticipative controls:* given $(\Omega,\mathcal{F},\mathcal{F}_t,P)$ and w_t as in Sections 1–2, find $\widehat{\alpha}\in\mathbf{R}^a$ and an $\mathcal{F}_t$-adapted process $\widehat{u}_t$ taking values in U and satisfying (2.1) such that $\Lambda[\widehat{\alpha},\widehat{u}]$ equals the infimum of $\Lambda[\alpha,u]$ over all possible choices of (α,u). Necessary conditions for optimality in this problem can be stated in terms of the *pre-Hamiltonian* $H\colon[0,T]\times\mathbf{R}^n\times\mathbf{R}^n\times U\to\mathbf{R}$ defined by

$$H(t,x,p,u)=p'f(t,x,u)-L(t,x,u).$$

(Prime denotes transpose.) Indeed, upon defining

$$\overline{p}_s' = -\ell_x(\widehat{x}_T)\Phi(T,s) - \int_s^T \widehat{L}_x(r)\Phi(r,s)\,dr, \tag{3.1}$$

the key equation of Prop. 2.1 becomes

$$\Lambda^\epsilon - \widehat{\Lambda} = \varepsilon \mathbf{E}\left[\langle \beta, \alpha\rangle - \overline{p}_0' A\alpha + H(s,\widehat{x}_s,\overline{p}_s,\widehat{u}_s) - H(s,\widehat{x}_s,\overline{p}_s,u_s)\right] + \varepsilon\delta(\varepsilon). \tag{3.2}$$

If $(\widehat{\alpha},\widehat{u})$ solves (P), then the left side must be nonnegative. Dividing equation (3.2) by $\varepsilon > 0$ and letting $\varepsilon \to 0^+$ then gives

$$0 \le \mathbf{E}\left[\langle \beta, \alpha\rangle - \overline{p}_0' A\alpha + H(s,\widehat{x}_s,\overline{p}_s,\widehat{u}_s) - H(s,\widehat{x}_s,\overline{p}_s,u_s)\right]. \tag{3.3}$$

This conclusion is valid for all $s \notin \mathcal{N}(\widehat{u},u)$ and all $\alpha \in \mathbf{R}^a$. We obtain the following version of the stochastic maximum principle.

3.1 Proposition. *Suppose $(\widehat{\alpha},\widehat{u})$ solves the strong problem with nonanticipative controls. Then the process $\overline{p}_t$ defined by (3.1) has the following properties. For any $\mathcal{F}_t$-adapted control $u\colon [0,T]\times\Omega \to U$ obeying the integrability condition (2.1), there is a null set $\mathcal{N}(\widehat{u},u) \subseteq [0,T]$ such that*

$$\mathbf{E}A'\overline{p}_0 = \beta, \tag{3.4}$$

$$\mathbf{E}H(s,\widehat{x}_s,\overline{p}_s,u_s) \le \mathbf{E}H(s,\widehat{x}_s,\overline{p}_s,\widehat{u}_s) \qquad \forall s \notin \mathcal{N}(\widehat{u},u). \tag{3.5}$$

Proof. Line (3.4) holds because (3.3) is valid for all $\alpha \in \mathbf{R}^a$; (3.4) and (3.3) together imply (3.5). ////

The conclusions of Prop. 3.1 unfortunately involve the comparison control u_t in the null set $\mathcal{N}(\widehat{u},u)$ of line (3.5). The results of Kushner (1972) and Haussmann (1985) both offer global versions of this condition. Kushner does this by introducing an explicit assumption regarding the approximability of admissible controls (his Assumption 2.3, p. 552). Haussmann, on the other hand, constructs a rather large family of comparison controls on which a global version of (3.5) holds without further hypotheses. We follow his method here.

Suppose that an $\mathcal{F}_t$-adapted process ψ_t with sample paths in C^ℓ is given. We will show that line (3.5) holds for a single null set $\mathcal{N}(\widehat{u},\psi) \subseteq [0,T]$, provided that the comparison control u_t is of

"ψ-feedback form"—that is, provided u_t is ψ_t-adapted and obeys the integrability condition (2.1). At each instant $t \in [0,T]$, the random variable u_t defined by such a control lies in the set

$$\begin{aligned} \mathcal{U}_t(\psi) &= L^{\overline{q}}(\Lambda, \mathcal{F}_t^{\psi}, P; U) \\ &= \left\{ v\big(\psi(t \wedge \cdot)\big) \;:\; v \in L^{\overline{q}}(C^n, \mathcal{C}^n, P \circ \psi^{-1}; U) \right\}. \end{aligned} \tag{3.6}$$

Consider now the following set of "simple ψ-feedback controls." Let $\mathcal{A}$ be the algebra of subsets of C^ℓ generated by all sets of the form $\{\varphi(\cdot) \in C^\ell \,:\, \varphi(t) \in R\}$, where t is a rational number in $[0,T]$ and $R \subseteq \mathbf{R}^\ell$ is a rectangle with vertices in the countable set $\mathbf{Q}^\ell$. The algebra $\mathcal{A}$ has countably many elements. Next, let $\widetilde{U}$ be a countable dense subset of U. A countable set V of simple ψ-feedback controls $v\colon C^\ell \to U$ is defined by

$$V = \left\{ v(\varphi(\cdot)) = \sum_{i=1}^{k} \widetilde{u}_i I\{\varphi(\cdot) \in A_i\} \;:\; k \in \mathbf{N},\ \widetilde{u}_i \in \widetilde{U},\ A_i \text{ disjoint in } \mathcal{A} \right\}.$$

Once the probability structure and ψ_t are given, the set V gives rise to a countable set $\mathcal{V}$ of ψ_t-adapted control laws as follows:

$$\mathcal{V} = \left\{ u(t,\omega) = v\big(\psi(t \wedge \cdot, \omega)\big) \;:\; v \in V \right\}. \tag{3.7}$$

Note that since every $v \in V$ is a bounded function, every control in $\mathcal{V}$ is bounded, hence admissible. We may define the null set $\mathcal{N}(\widehat{u},\psi)$ as the union of $\mathcal{N}(\widehat{u},u)$ over all $u \in \mathcal{V}$. Outside of this null set, both (1.2) and (2.2) hold simultaneously for all $u \in \mathcal{V}$. Consequently the same is true of line (3.5), which we may now extend by taking limits.

3.2 Theorem (Stochastic Maximum Principle). *Suppose $(\widehat{\alpha},\widehat{u})$ solves the strong problem with nonanticipative controls. Suppose further that a continuous $\mathcal{F}_t$-adapted process ψ_t with values in $\mathbf{R}^\ell$ is given. Then there is a null set $\mathcal{N}(\widehat{u},\psi) \subseteq [0,T]$ such that for any ψ_t-adapted comparison control u_t obeying (2.1), one has*

$$\mathbf{E} A' \overline{p}_0 = \beta, \tag{3.8}$$

$$\mathbf{E} H(s, \widehat{x}_s, \overline{p}_s, u_s) \le \mathbf{E} H(s, \widehat{x}_s, \overline{p}_s, \widehat{u}_s) \qquad \forall s \notin \mathcal{N}(\widehat{u},\psi). \tag{3.9}$$

Proof. The first conclusion follows immediately from (3.4); we need only to show how the second one follows from (3.5). For this, we use the null set $\mathcal{N}(\widehat{u},\psi)$ defined above, outside of which (3.5) holds simultaneously for all $u \in \mathcal{V}$.

The key to the proof is that for each t, the countable set

$$\mathcal{V}_t = \{u(t,\cdot) \, : \, u \in \mathcal{V}\}$$

is a dense subset of $\mathcal{U}_t(\psi)$ in the topology of $L^{\overline{q}}(\Omega, \mathcal{F}_t^{\psi}, P; U)$. This is the content of Halmos (1950), ex. 42(1), p. 177. (See also Halmos, Thm. 40.b, p. 168; a few more details appear in Haussmann (1985), Section 5.) So for any fixed $s \notin \mathcal{N}(\widehat{u}, \psi)$ and random variable $u \in \mathcal{U}_s$, there is a sequence $u^k \to u$ a.s. and in $L^{\overline{q}}$ along which (3.5) holds for each k. We complete the proof by showing that

$$\mathbf{E}H(s, \widehat{x}_s, \overline{p}_s, u^k) \to \mathbf{E}H(s, \widehat{x}_s, \overline{p}_s, u). \tag{3.10}$$

In view of the definition of H, we consider the summands $p'f$ and L separately. Hypothesis (H5) implies that the sequence $\mathbf{E}\left|L(s, \widehat{x}_s, u^k)\right|^{\overline{q}/q}$ is uniformly bounded, so $\mathbf{E}L(s, \widehat{x}_s, u^k) \to \mathbf{E}L(s, \widehat{x}_s, u)$ by uniform integrability. As for $p'f$, first choose any $r \in (q, \overline{q})$ and consider the integral expression (3.1) for $\overline{p}'_s$. Using the fact that $\Phi(t, \tau)$ has moments of all orders, each of which is bounded uniformly in k, repeated application of Hölder's inequality gives a finite k-independent upper bound for $\mathbf{E}\left|\overline{p}_s\right|^{r/q}$. In particular, choose r sufficiently near to $\overline{q}$ that $\dfrac{r}{\overline{q}} \dfrac{\overline{q}-1}{q} > 1$: then there is an exponent ς in the interval $\left(\dfrac{\overline{q}}{\overline{q}-1}, \dfrac{r}{q}\right)$ which will necessarily obey $\dfrac{\varsigma}{\varsigma - 1} < \overline{q}$. Therefore some power $\beta > 1$ obeys $\beta\varsigma < r/q$ and $\beta \dfrac{\varsigma}{\varsigma-1} < \overline{q}$, so we deduce

$$\begin{aligned}
\mathbf{E}\left|\overline{p}'_s f(s, \widehat{x}_s, u^k)\right|^{\beta} &\le \left(\mathbf{E}\left|\overline{p}'_s\right|^{\beta\varsigma}\right)^{1/\varsigma} \left(\mathbf{E}\left|f(s, \widehat{x}_s, u^k)\right|^{\beta\varsigma/(\varsigma-1)}\right)^{(\varsigma-1)/\varsigma} \\
&\le K_1 \left(\mathbf{E}\left|\overline{p}'_s\right|^{r/q}\right)^{1/\varsigma} \left(\mathbf{E}\left[1 + |\widehat{x}|^{\overline{q}}\right]\right)^{(\varsigma-1)/\varsigma}.
\end{aligned}$$

The right side here is bounded uniformly in k, so uniform integrability ensues, and results in

$$\mathbf{E}\overline{p}'_s f(s, \widehat{x}_s, u^k) \to \mathbf{E}\overline{p}'_s f(s, \widehat{x}_s, u) \quad \text{as } k \to \infty.$$

Together with our previous treatment of L, this establishes (3.10). ////

It is important to note that the existence of the continuous process ψ_t places no restrictions on $\widehat{u}$. Indeed, Thm. 3.2 does not require that $\widehat{u}_t$ be ψ_t-adapted, and the choice $\psi_t = \widehat{x}_t$ is always a legitimate possibility. However, if the filtration $\mathcal{F}_t$ happens to equal $\mathcal{F}_t^{\psi}$ for some continuous process ψ_t (for example, this is sometimes the case when $\psi_t = (\widehat{x}_t, w_t)$), then $\mathcal{U}_t(\psi)$ is precisely the set of all $\mathcal{F}_t$-measurable and U-valued random variables with finite $\overline{q}$-th moments. It follows that $\widehat{u}_t \in \mathcal{U}_t(\psi)$ for all t, and that the maximum condition (3.5) is truly global.

Quadratic Penalization. The necessary conditions of Thm. 3.2 do not change if the objective functional incorporates a quadratic penalization term of a certain form. Indeed, let $\widetilde{\ell}$ and $\widetilde{L}$ be functions obeying (H4)–(H5) and defining the functional

$$\widetilde{\Lambda}[\alpha, u] := \mathbf{E}\left[\left\langle\widetilde{\beta}, \alpha\right\rangle + \widetilde{\ell}(x_T) + \int_0^T \widetilde{L}(r, x_r, u_r)\, dr\right].$$

Suppose a pair $(\widehat{\alpha}, \widehat{u})$ is given, and that the unconstrained stochastic control problem of this section is considered with the modified cost functional

$$\Lambda_\rho[\alpha, u] = \Lambda[\alpha, u] + \rho\left|\widetilde{\Lambda}[\alpha, u] - \widetilde{\Lambda}[\widehat{\alpha}, \widehat{u}]\right|^2 + \rho\,|\alpha - \widehat{\alpha}|^2 \tag{3.11}$$

for some fixed $\rho \geq 0$. Now if $(\widehat{\alpha}, \widehat{u})$ minimizes Λ_ρ and we define perturbed pairs $(\alpha^\epsilon, u^\epsilon)$ as above, then Prop. 2.1 applies to both Λ and $\widetilde{\Lambda}$ to give functions $\delta(\varepsilon)$, $\widetilde{\delta}(\varepsilon) \sim 0$ for which

$$\begin{aligned} \Lambda_\rho[\alpha^\epsilon, u^\epsilon] - \Lambda_\rho[\widehat{\alpha}, \widehat{u}] =& \varepsilon\mathbf{E}\left[\langle\beta, \alpha\rangle - \overline{p}_0' A\alpha + \Delta H(s, u_s)\right] + \varepsilon\delta(\varepsilon) \\ &+\varepsilon^2\rho\left|\mathbf{E}\left[\left\langle\widetilde{\beta}, \alpha\right\rangle - \widetilde{\overline{p}}_0' A\alpha + \Delta\widetilde{H}(s, u_s)\right] + \widetilde{\delta}(\varepsilon)\right|^2. \end{aligned} \tag{3.12}$$

Here H and $\overline{p}$ are defined as above, $\widetilde{H}$ and $\widetilde{\overline{p}}$ are their obvious analogues, and

$$\Delta H(s, u) := H(s, \widehat{x}_s, \overline{p}_s, \widehat{u}_s) - H(s, \widehat{x}_s, \overline{p}_s, u).$$

Just as (3.3) follows from (3.2), it also follows from (3.12), and the proof of Thm. 3.2 then proceeds without change. This fact will be useful later.

3.3 Theorem. *The statement of Theorem 3.2 remains true if we consider any objective functional of the form (3.11) in which $\widetilde{\Lambda}$ obeys (H4)–(H5).*

3.4 The Fundamental Matrix. In the convergence analysis to follow in Section 4, the structure of the adjoint process $\overline{p}_t$ must be clearly understood. Most of the properties of $\overline{p}_t$ follow from those of the fundamental matrix $\Phi(t, \tau)$ given by (1.10), which we summarize here.

Let us introduce $\Phi_t (\equiv \Phi(t, 0))$ and Ψ_t as the pathwise unique continuous $n \times n$ matrix processes solving

$$d\Phi_t = \widehat{f}_x(t)\Phi_t\, dt + \widehat{\sigma}_x(t)\Phi_t\, dw_t, \qquad \Phi_0 = I, \tag{3.13a}$$

$$d\Psi_t = -\Psi_t\left[\widehat{f}_x(t) + \widehat{\sigma}_x(t)\widehat{\sigma}_x(t)\right] dt - \Psi_t\widehat{\sigma}_x(t)\, dw_t, \qquad \Psi_0 = I. \tag{3.13b}$$

Here we have used the following notation:

$$\widehat{\sigma}_x(t)\Phi_t\, dw_t = \sum_{k=1}^{d} \widehat{\sigma}_x^k(t)\Phi_t\, dw_t^k,$$

$$\widehat{\sigma}_x(t)\widehat{\sigma}_x(t) = \sum_{k=1}^{d} \widehat{\sigma}_x^k(t)\widehat{\sigma}_x^k(t),$$

$$\Psi_t\widehat{\sigma}_x(t)\, dw_t = \sum_{k=1}^{d} \Psi_t\widehat{\sigma}_x^k(t)\, dw_t^k.$$

With these definitions, equations (3.13) imply that $d(\Psi_t\Phi_t) = 0$, so that $\Psi_t\Phi_t = I$ $\forall t$ a.s. Consequently $\Psi_t = \Phi_t^{-1}$ $\forall t$ a.s.; since the solution Φ of (1.10) is pathwise unique, it must obey

$$\Phi(t,\tau) = \Phi_t\Psi_\tau = \Phi_t\Phi_\tau^{-1} \qquad \forall t,\tau \text{ a.s.} \tag{3.14}$$

Both Φ_t and $\Phi_t^{-1} = \Psi_t$ solve linear SDE's with bounded coefficients. Hence for every $p \geq 2$, the proof of Lemma IV.4.1 yields a constant C_p such that

$$\mathbf{E}\left(\sup_{t\in[0,T]} |\Phi_t|^p + \sup_{t\in[0,T]} \left|\Phi_t^{-1}\right|^p\right) \leq C_p; \tag{3.15}$$

Hölder's inequality can then be used to get

$$\mathbf{E}\sup_{t,\tau} |\Phi(t,\tau)|^p \leq C_{2p}. \tag{3.16}$$

Notice that the bounds on the coefficients of (3.13) are independent of the choice of $\widehat{\alpha}$, $\widehat{u}$, $\widehat{x}$, so the right-hand sides of (3.15) and (3.16) are also independent of this choice.

Section 4. Constraints and the Value Function

Although the precise version of the Stochastic Maximum Principle derived in Section 3 has not been given explicitly before, the linearization techniques used to obtain it are comparatively well known. The main point of this chapter is still to come. In this section we will consider α not as part of the control, but instead as a finite-dimensional perturbation vector which, once specified, dictates the choice of the optimal policy. Our goal is to study the ways in which changes to the parameter α affect the minimum value in a control problem incorporating equality and inequality constraints, and to show how perturbation analysis will allow us to obtain and interpret the results of a Stochastic Maximum Principle for constrained problems.

Proximal normal analysis inevitably involves both existence theory and necessary conditions. These two prerequisites conflict with each other, in that existence theory requires fast growth of the Lagrangian ((H5) in Chap. IV), whereas the SMP requires slow growth ((H5) in this chapter). In fact, the only way to use a coercivity condition to guarantee that (2.1) holds for all admissible controls u is to assume that $L(t,x,u)$ grows at least as fast as $|u|^{\overline{q}}$. But this cannot be reconciled with the requirements of (H5). The two contrary growth conditions on L coalesce only in the case where U is a compact set. *Throughout the remainder of this chapter, we assume that U is compact.* Under this hypothesis, we may use the special existence Thm. IV.7.1.

A somewhat less significant tension between existence theory and necessary conditions also affects our formulation of the constrained problem. In Chap. IV, we proved existence theorems for weak problems in which the probability structure was a choice variable; in Sections 1–3 of this chapter, we developed necessary conditions for strong problems in which the probability structure is fixed in advance. However, the necessary conditions of Thm. 3.2 clearly remain valid for an optimal solution to a weakly formulated problem, since any optimal $\widehat{u}$ for the weak problem must also solve the strong problem obtained when its associated probability structure is regarded as immutable. We therefore study the weak form of the problem.

The set of admissible controls, denoted $\mathcal{U}$, consists of all U-valued stochastic processes u_t associated with some probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and Brownian motion w_t, such that u_t is $\mathcal{F}_t$-adapted. In contrast to the definition of Chap. IV, we do *not* require that u_t be $\mathcal{F}_t^x$-adapted, where x is the corresponding solution of the dynamic equations. Recall, however, that subject to the hypotheses listed below, the value of the infimum and the question of existence are both oblivious to this distinction. Without loss of generality, we may assume that every $u \in \mathcal{U}$ obeys

$$u(t,\omega) \in U \qquad \forall t \in [0,T],\ \forall \omega \in \Omega.$$

Our proximal normal analysis will be based on a rather general perturbation structure, indexed by pairs $(\alpha, \lambda) \in \mathbf{R}^a \times \mathbf{R}^{I+J}$. Typically,

$$\lambda = (\lambda_{-I}, \lambda_{1-I}, \ldots, \lambda_{-1}, \lambda_1, \lambda_2, \ldots, \lambda_J).$$

The family of problems of interest involves the $I+1+J$ functionals

$$\Lambda_k[\alpha,u] := \mathbf{E}\left[\ell_k(x_T) + \int_0^T L_k(r,x_r,u_r)\,dr\right], \quad k=-I,\ 1-I,\ldots,\ J,$$

as follows:

$$\min_{u\in\mathcal{U}}\{\Lambda_0[\alpha,u] : dx_t = f(t,x_t,u_t)\,dt + \sigma(t,x_t)\,dw_t, \quad x(0)=x_0(\alpha),$$

$$P(\alpha,\lambda) \qquad \Lambda_i[\alpha,u] \le -\lambda_i, \quad i=-1,-2,\ldots,-I,$$

$$\Lambda_j[\alpha,u] = -\lambda_j, \quad j=1,2,\ldots,J\}.$$

For ease of notation in the discussion below, we will use subscript + and − signs as follows:

$$\lambda_- = (\lambda_{-I},\lambda_{1-I},\ldots,\lambda_{-1}) \in \mathbf{R}^I; \qquad \lambda_+ = (\lambda_1,\lambda_2,\ldots,\lambda_J) \in \mathbf{R}^J.$$

Thus $\lambda = (\lambda_-,\lambda_+)$. Similarly $\Lambda \in \mathbf{R}^{I+1+J}$ will be defined as $(\Lambda_-,\Lambda_0,\Lambda_+)$, with like notation being inherited by L and ℓ. We will use the symbol $\Lambda_\mp$ for (Λ_-,Λ_+).

The *value function* $V\colon \mathbf{R}^a \times \mathbf{R}^{I+J} \to \mathbf{R}\cup\{+\infty\}$ is defined by $V(\alpha,\lambda) := \inf P(\alpha,\lambda)$. Stated precisely, this section's goal is to compute $\partial V(0)$ and $\partial^\infty V(0)$ in terms of the Lagrange multipliers and adjoint process arising from solutions to $P(0)$.

Hypotheses. For the sake of clarity, we now give the full statements of the standing hypotheses required to apply both existence theorems and necessary conditions in a constrained context. The rapid escalation of technical difficulties in what follows is relieved somewhat by an assumption of continuity in t, whose removal would require a significant fortification of (h6).

(h1) Let $\tilde{f} = (f,L_+) \in \mathbf{R}^n \times \mathbf{R}^J$. Assume that for each $(t,u)\in[0,T]\times U$, both $\tilde{f}(t,\cdot,u)$ and $\sigma(t,\cdot)$ are differentiable; that $\tilde{f}$, $\tilde{f}_x$, σ, σ_x are jointly continuous in all their arguments; and that the continuity of $\tilde{f}(t,x,\cdot)$ and $\tilde{f}_x(t,x,\cdot)$ is uniform in (t,x).

(h2) There is a constant $k_1>0$ such that for all $(t,x,u)\in[0,T]\times\mathbf{R}^n\times U$,

$$\left|\tilde{f}(t,x,u)\right| + |\sigma(t,x)| \le k_1(1+|x|),$$

$$\left|\tilde{f}_x(t,x,u)\right| + |\sigma_x(t,x)| \le k_1.$$

(h3) There are given random vectors $\overline{X}_0\in\mathbf{R}^n$ and $\overline{A}\in\mathbf{R}^{n\times a}$ such that $\mathbf{E}\left|\overline{X}_0\right|^q < +\infty$ and $\left|\overline{A}\right| \le k_2$. For any $u\in\mathcal{U}$ with corresponding probability space $(\Omega,\mathcal{F},\mathcal{F}_t,P)$, the initial distribution of x_t obeys

$$x_0(\alpha) \stackrel{\mathcal{D}}{\equiv} \overline{X}_0 + \overline{A}\alpha.$$

(h4) Let $\widetilde{L} = (L_-, L_0) \in \mathbf{R}^I \times \mathbf{R}$. Assume that for each $(t, \omega) \in [0, T] \times U$, both $\widetilde{L}(t, \cdot, u)$ and $\ell(\cdot)$ are differentiable; that $\widetilde{L}$, $\widetilde{L}_x$, ℓ, ℓ_x are jointly continuous in all their arguments; and that the continuity of $\widetilde{L}(t, x, \cdot)$ and $\widetilde{L}_x(t, x, \cdot)$ is uniform in (t, x).

(h5) There is an exponent $q \in [1, \overline{q} - 1)$ and a constant $k_3 > 0$ such that for all $(t, x, u) \in [0, T] \times \mathbf{R}^n \times U$, one has

$$\left|\widetilde{L}(t, x, u)\right| \leq k_3(1 + |x|^q), \qquad \left|\widetilde{L}_x(t, x, u)\right| \leq k_3(1 + |x|^{q-1}),$$

$$|\ell(x)| \leq k_3(1 + |x|^q), \qquad |\ell_x(x)| \leq k_3(1 + |x|^{q-1}).$$

(h6) For each (t, x), with the possible exception of a set whose projection onto the t-axis has Lebesgue measure zero, the following (compact) set is convex:

$$\left\{ \Big(f(t, x, u),\ L(t, x, u),\ f_x(t, x, u),\ L_x(t, x, u) \Big) \ :\ u \in U \right\}.$$

Hypothesis (h6) is more restrictive than the corresponding convexity condition (h6) in Chap. IV. Technically, this extra convexity assumption on f_x and L_x is required for convergence of the adjoint processes when we take limits as part of the proximal normal analysis to follow. Practically, however, (h6) does not significantly weaken the theory to be developed below because it is automatically satisfied by any problem which is "sufficiently relaxed." See Clarke (1983), Section 5.5, and Warga (1972).

4.1 Lemma. *The value function is lower semicontinuous near* 0.

Proof. Choose any (α, λ) near 0, and let $\{(\alpha^k, \lambda^k)\}$ be any sequence with limit (α, λ). Without loss of generality, we may pass to a subsequence and assume that $V(\alpha^k, \lambda^k) \to v = \liminf_{k\to\infty} V(\alpha^k, \lambda^k)$. We must show $V(\alpha, \lambda) \leq v$. This is evident if $v = +\infty$, so assume $v < +\infty$. In this case, Thm. IV.7.1(a) applies at each k to give a control $u^k \in \mathcal{U}$ and a vector $\rho^k \in \mathbf{R}^I$ with nonnegative components such that $(-\rho^k - \lambda_-^k, V(\alpha^k, \lambda^k), -\lambda_+^k) = \Lambda[\alpha^k, u^k]$. Since ρ^k is bounded by (h5), this sequence has a convergent subsequence, along which Thm. IV.7.1(b) gives a control $u \in \mathcal{U}$ expressing the limit in the form

$$(-\rho - \lambda_-, v, -\lambda_+) = \Lambda[\alpha, u].$$

This shows that $V(\alpha, \lambda) \leq v$, as required. ////

Perpendiculars. Suppose that some vector $(\beta, \varphi_{\mp}, -\varphi_0)$ is perpendicular to epi V at a point $(\widehat{\alpha}, \widehat{\lambda}, \widehat{v})$ near $(0, 0, V(0,0))$. Then since $V(\widehat{\alpha}, \widehat{\lambda}) \leq \widehat{v} < +\infty$, problem $P(\widehat{\alpha}, \widehat{\lambda})$ must have a solution—that is, a control $\widehat{u} \in \mathcal{U}$ for which

$$\Lambda_0[\widehat{\alpha}, \widehat{u}] = V(\widehat{\alpha}, \widehat{\lambda}), \quad \Lambda_-[\widehat{\alpha}, \widehat{u}] \leq -\widehat{\lambda}_-, \quad \Lambda_+[\widehat{\alpha}, \widehat{u}] = -\widehat{\lambda}_+.$$

Now for any other $\alpha \in \mathbf{R}^a$ near 0, any control $u \in \mathcal{U}$ assigns well-defined values to x, Λ. So for any $\rho \geq 0$ in $\mathbf{R}^I$ (this means ρ has no negative component values) and $t \geq 0$, the control u gives rise to the following point:

$$\Big(\alpha, (-\Lambda_-[\alpha, u] - \rho, -\Lambda_+[\alpha, u]), \Lambda_0[\alpha, u] + t\Big).$$

This point lies in epi V by inspection. Thus Prop. II.3.5 implies

$$\begin{aligned} &\Big\langle (\beta, \varphi_{\mp}, -\varphi_0), (\alpha, -\Lambda_- - \rho, -\Lambda_+, \Lambda_0 + t) - (\widehat{\alpha}, \widehat{\lambda}_-, \widehat{\lambda}_+, \widehat{v}) \Big\rangle \\ &\qquad \leq \frac{1}{2} \left| (\alpha, -\Lambda_- - \rho, -\Lambda_+, \Lambda_0 + t) - (\widehat{\alpha}, \widehat{\lambda}_-, \widehat{\lambda}_+, \widehat{v}) \right|^2. \end{aligned} \tag{4.1}$$

If we choose $\alpha = \widehat{\alpha}$ and $u = \widehat{u}$ in (4.1), the constraints of problem $P(\widehat{\alpha}, \widehat{\lambda})$ are satisfied and substantial cancellation occurs. We obtain

$$0 \leq \varphi_0\Big(V(\widehat{\alpha}, \widehat{\lambda}) - \widehat{v} + t\Big) + \Big\langle \varphi_-, \widehat{\Lambda}_- + \widehat{\lambda}_- + \rho \Big\rangle + \frac{1}{2}\left|\Big(0,\ \widehat{\Lambda}_- + \widehat{\lambda}_- + \rho,\ 0,\ \widehat{\Lambda}_0 + t - \widehat{v}\Big)\right|^2, \tag{4.2}$$

an expression valid for all $\rho \geq 0$ and $t \geq 0$. Since equality holds in (4.2) when

$$(t, \rho) = (\widehat{v} - V(\widehat{\alpha}, \widehat{\lambda}), -\widehat{\Lambda}_- - \widehat{\lambda}_-) \geq 0,$$

the RHS of (4.2) must have nonnegative right derivatives in (t, ρ) at this local minimum point. Therefore $\varphi_0 \geq 0$ and $\varphi_- \geq 0$. In fact if any component of the minimizing (t, ρ) is strictly positive then the derivative must actually vanish there. Thus we get the complementary slackness conditions

$$\varphi_0 \geq 0, \qquad \varphi_0\big(V(\widehat{\alpha}, \widehat{\lambda}) - \widehat{v}\big) = 0, \tag{4.3}$$

$$\varphi_- \geq 0, \qquad \Big\langle \varphi_-, \widehat{\Lambda}_- + \widehat{\lambda}_- \Big\rangle = 0. \tag{4.4}$$

Line (4.3) simply restates the geometrically obvious facts that a perpendicular to an epigraph cannot be directed upward, and that if it is based on a vertical side of the epigraph then it cannot be directed downward either. Line (4.4) is the precursor of the usual complementary slackness condition on the multipliers of the constrained Stochastic Maximum Principle.

Let us now fix $t = \widehat{v} - V(\widehat{\alpha}, \widehat{\lambda})$ and $\rho = -(\widehat{\Lambda}_- + \widehat{\lambda}_-)$ in (4.1), and observe that the inequality becomes

$$- \langle \beta, \widehat{\alpha} \rangle + \varphi' \widehat{\Lambda} \le - \langle \beta, \alpha \rangle + \varphi' \Lambda[\alpha, u] + \frac{1}{2} \left| \left(\alpha - \widehat{\alpha}, \Lambda[\alpha, u] - \widehat{\Lambda} \right) \right|^2 . \tag{4.5}$$

This inequality is valid for all $\alpha \in \mathbf{R}^a$, $u \in \mathcal{U}$, and equality holds when $(\alpha, u) = (\widehat{\alpha}, \widehat{u})$. Hence this choice represents the solution to an unconstrained stochastic control problem. This problem has a weak formulation, since every element of $\mathcal{U}$ carries its own probability structure. But if we regard the probabilistic framework specified by $\widehat{u}$ as fixed, then $\widehat{u}$ solves the strong problem with nonanticipative controls in that setting. Therefore Thm. 3.3 provides necessary conditions. Recall from (3.6) the notation

$$\mathcal{U}_t(\psi) = \left\{ u : \Omega \to U \ : \ u \text{ is } \mathcal{F}_t^{\psi}\text{-measurable} \right\};$$

we will choose ψ later.

4.2 Proposition. *Let $(\beta, \varphi_{\mp}, -\varphi_0) \perp \operatorname{epi} V$ at $(\widehat{\alpha}, \widehat{\lambda}, \widehat{v})$, and set $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0) = \dfrac{(\beta, \varphi_{\mp}, -\varphi_0)}{|(\beta, \varphi_{\mp}, -\varphi_0)|}$. Then there is a solution $\widehat{u} \in \mathcal{U}$ to $P(\widehat{\alpha}, \widehat{\lambda})$ for which $\widetilde{\varphi}$ obeys (4.3) and (4.4). Associated with $\widehat{u}$ is a null set $\mathcal{N}(\widehat{u}, \psi) \subseteq [0, T]$ and a probability structure $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, w_t, such that the $\mathbf{R}^n$-valued process*

$$\overline{p}_t' = -\widetilde{\varphi}' \left[\ell_x(\widehat{x}_T) \Phi_T + \int_t^T \widehat{L}_x(r) \Phi_r \, dr \right] \Phi_t^{-1} \tag{4.6}$$

obeys

$$\mathbf{E} A' \overline{p}_0 = -\widetilde{\beta}, \tag{4.7}$$

$$\mathbf{E} H(t, \widehat{x}_t, \overline{p}_t, \widehat{u}_t, \widetilde{\varphi}) \ge \mathbf{E} H(t, \widehat{x}_t, \overline{p}_t, u, \widetilde{\varphi}) \qquad \forall u \in \mathcal{U}_t(\psi), \ \forall t \notin \mathcal{N}(\widehat{u}, \psi). \tag{4.8}$$

Here the pre-Hamiltonian H is defined by

$$H(t, x, p, u, \widetilde{\varphi}) = p' f(t, x, u) - \widetilde{\varphi}' L(t, x, u) \tag{4.9}$$

and the fundamental matrix Φ_t is given in paragraph 3.4.

Convergence. Suppose now that a vector $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0)$ is obtained via

$$(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0) = \lim_{k \to \infty} \frac{(\beta^k, \varphi_{\mp}^k, -\varphi_0^k)}{|(\beta^k, \varphi_{\mp}^k, -\varphi_0^k)|} = \lim_{k \to \infty} (\widetilde{\beta}^k, \widetilde{\varphi}_{\mp}^k, -\widetilde{\varphi}_0^k) \tag{4.10}$$

for some sequence $(\beta^k, \varphi^k_\mp, -\varphi^k_0) \perp \operatorname{epi} V$ at $(\alpha^k, \lambda^k, v^k) \to (0, 0, V(0,0))$. Do the conclusions of Prop. 4.2 hold in the limit? The answer is yes, but the precise version of this statement requires careful treatment of the sequence of underlying probability spaces and the extraction of a subsequence—its proof is rather far from trivial. Indeed, it is highly remeniscent of the existence proofs in Chap. IV.

Let such a sequence of perpendiculars be given. For each k, we have

$$V(\alpha^k, \lambda^k) \le v^k < +\infty,$$

so problem $P(\alpha^k, \lambda^k)$ has an optimal control-state pair (u^k, x^k), i.e. one for which $\Lambda[\alpha^k, u^k] = (-\lambda^k_- - \rho^k, V(\alpha^k, \lambda^k), -\lambda^k_+)$, where $\rho^k \ge 0$. Let the probability structure associated with u^k be labelled $(\Omega^k, \mathcal{F}^k, \mathcal{F}^k_t, P^k)$ and w^k_t. Now since $\Lambda[\alpha^k, u^k]$ is a bounded sequence in $\mathbf{R}^{I+1+J}$ (this follows from Lemma IV.4.1 and the growth conditions (h5)), we may assume that $-\lambda^k_- - \rho^k \to -\rho$ by passing to a suitable subsequence. Also, $V(\alpha^k, \lambda^k) \le v^k$ while $v^k \to V(0,0)$: by Lemma 4.1, it follows that $V(\alpha^k, \lambda^k) \to V(0,0)$. In short, we may assume that $\Lambda[\alpha^k, u^k] \to (-\rho, V(0,0), 0)$ by passing to a subsequence. Now as Thm. IV.7.1 shows, there is a further subsequence (which we do not relabel) along which x^k converges in distribution to a process x which in turn can be realized by a control solving $P(0,0)$. For each k, the conclusions of Prop. 4.2 hold for $(\widetilde{\beta}^k, \widetilde{\varphi}^k_\mp, -\widetilde{\varphi}^k_0)$ and some process $\overline{p}^k$. But before any assertions concerning the convergence of $\overline{p}^k$ can be advanced, we must recall that the methods of Chap. IV used to obtain $x^k \xrightarrow{\mathcal{D}} x$ studiously avoided any assertion that the controls u^k converged in any way at all. In view of this, an elementary approach to convergence in, say, (4.6) is out of the question. We must return to the methods of Chap. IV to show that the control realizing x can actually be chosen to facilitate convergence of the multipliers. Even this is less straightforward than it might appear. The difficulty centres on the issue of $\mathcal{F}_t$-adaptedness, which is critical to the use of Beneš's measurable selection theorem as in Chap. IV. To resolve it, we must make explicit use the definition of $\overline{p}_t$ in terms of the fundamental matrix Φ_t. Each Φ^k_t is $\mathcal{F}^k_t$-adapted, and the estimates of paragraph 3.4 will allow us to show that these matrix processes converge in the appropriate sense and then deduce the convergence of $\overline{p}^k$ from definition (4.6). This programme follows the conceptual lines of Props. IV.5.2–IV.5.7, but is technically more difficult because so many more processes are required to converge simultaneously.

We will use the following notation, based on the constituents of the dynamics, the objective functionals, the adjoint process, and the pre-Hamiltonian.

$$
\begin{aligned}
x^{1,k}(t) &= \int_0^t f(r, x_r^k, u_r^k)\, dr, \\
x^{2,k}(t) &= \int_0^t \sigma(r, x_r^k)\, dw_r^k, \\
z^k(t) &= \int_0^t L(r, x_r^k, u_r^k)\, dr, \\
\overline{p}^{1,k} &= \ell_x(x_T^k)\Phi_T^k, \\
\overline{p}^{2,k}(t) &= \int_0^t L_x(r, x_r^k, u_r^k)\Phi_r^k\, dr, \\
h^{1,k}(t) &= \int_0^t \left(\Phi_r^k\right)^{-1} f(r, x_r^k, u_r^k)\, dr, \\
h^{2,k}(t) &= \int_0^t \overline{p}_r^{2,k} \left(\Phi_r^k\right)^{-1} f(r, x_r^k, u_r^k)\, dr, \\
\Phi^{1,k}(t) &= \int_0^t f_x(r, x_r^k, u_r^k)\Phi_r^k\, dr, \\
\Phi^{2,k}(t) &= \int_0^t \sigma_x(r, x_r^k)\Phi_r^k\, dw_r, \\
\Psi_t^k &= \left(\Phi_t^k\right)^{-1}.
\end{aligned}
$$

4.3 Proposition. *The following sequence of 15-tuples is tight as a sequence of random vectors in the space* $\mathbf{R}^n \times C^{3n} \times C^{I+1+J} \times C^{3n} \times C \times C \times C^{4(n\times n)} \times C^d$*:*

$$
\left(x_0^k(\alpha^k),\ x^k(\cdot),\ x^{1,k}(\cdot),\ x^{2,k}(\cdot),\ z^k(\cdot),\ \overline{p}^k(\cdot),\ \overline{p}^{1,k},\ \overline{p}^{2,k}(\cdot),\ h^{1,k}(\cdot),\ h^{2,k}(\cdot),\ \Phi_\cdot^k,\ \Phi_\cdot^{1,k},\ \Phi_\cdot^{2,k},\ \Psi_\cdot^k,\ w^k(\cdot)\right).
$$

Hence it has a subsequence converging in distribution to a 15-tuple

$$
\left(x_0(0),\ x(\cdot),\ x^1(\cdot),\ x^2(\cdot),\ z(\cdot),\ p(\cdot),\ \overline{p}^1,\ \overline{p}^2(\cdot),\ h^1(\cdot),\ h^2(\cdot),\ \Phi_\cdot,\ \Phi_\cdot^1,\ \Phi_\cdot^2,\ \Psi_\cdot,\ w(\cdot)\right).
$$

Proof. Fortunately, tightness can be proven component by component. Recall also that any sequence which converges in distribution is automatically tight by Prokhorov's theorem. Thus $x_0^k(\alpha^k)$ is tight because it converges in distribution to the random vector $\overline{X}_0$ of (h3), and $w^k(\cdot)$ is tight because this sequence of Brownian motion processes is actually constant in distribution. (Moreover the limiting process $w(\cdot)$ must therefore be a Brownian motion, but we will explore this later.)

Prop. IV.5.2 proves that $\left(x^k(\cdot), x^{1,k}(\cdot), x^{2,k}(\cdot)\right)$ are tight in C^{3n}. In fact, the proof given there that $x^{1,k}$ was tight applies equally well to $\overline{p}^{2,k}$, $h^{1,k}$, $h^{2,k}$, and $\Phi^{1,k}$. (It even simplifies a little bit in the present setting.) Likewise, the treatment of $x^{2,k}$ given in Prop. IV.5.2 applies also to $\Phi^{2,k}$. The relation $\Phi^k = I + \Phi^{1,k} + \Phi^{2,k}$ then implies that Φ^k is tight also. The tightness of $z^k(\cdot)$ is proven in Thm. IV.7.1.

To prove that $\overline{p}^{1,k}$ is tight in $\mathbf{R}^n$, it suffices to show that $\mathbf{E}\left|\overline{p}^{1,k}\right|$ is uniformly bounded. And for this, one simply uses (h5) to write

$$\mathbf{E}\left|\overline{p}^{1,k}\right| \leq \mathbf{E}\left|\ell_x(x_T^k)\right| \left\|\Phi_\cdot^k\right\| \leq \mathbf{E} k_3 \left(1 + \left\|x^k\right\|^{q-1}\right) \left\|\Phi_\cdot^k\right\|.$$

The RHS is uniformly bounded in k by Hölder's inequality, since $\left\|\Phi_\cdot^k\right\|$ is uniformly bounded in any L^p, $p \geq 1$, and $\left\|x^k\right\|$ is uniformly bounded in any L^r, $r \in [1, \overline{q})$.

The proof of tightness for $\Psi^k = \left(\Phi^k\right)^{-1}$ is left to the reader, with the following hint: the function Ψ^k satisfies the linear SDE (3.13b) with bounded coefficients. Writing this equation in integral form and proving the tightness of each term on the resulting RHS separately, just as we have done repeatedly above, will show that Ψ^k is tight also.

Finally, the tightness of $\overline{p}^{1,k}$, $\overline{p}^{2,k}$, Φ^k, and Ψ^k, together with the relation

$$\overline{p}^k(t) = -\left(\tilde{\varphi}^k\right)' \left[\overline{p}^{1,k} + \overline{p}^{2,k}(T) - \overline{p}^{2,k}(t)\right] \Psi_t^k$$

implies that $\overline{p}^k$ is tight also. ////

A word about the initial conditions is appropriate here. To each control u^k there corresponds a space $(\Omega^k, \mathcal{F}^k, P^k)$ on which are defined random variables X_0^k and A^k such that $X_0^k \overset{\mathcal{D}}{\equiv} \overline{X}_0$ and $A^k \overset{\mathcal{D}}{\equiv} \overline{A}$, and $x_0^k(\alpha^k) = X_0^k + A^k \alpha^k$. Now since the pairs $\left\{\left(X_0^k, A^k\right)\right\}$ all have the same distribution, they could easily be adjoined to the 15-tuples considered above. Then the limiting probability space $(\Omega, \mathcal{F}, P)$ would be seen to support random variables $X_0 \overset{\mathcal{D}}{\equiv} \overline{X}_0$ and $A \overset{\mathcal{D}}{\equiv} \overline{A}$ such that $x_0^k(\alpha^k) \xrightarrow{\mathcal{D}} x_0(0) = X_0$, and such that the joint distributions of (X_0^k, A^k) together with the 15-tuples above converge to the joint distribution of (X_0, A) and the limiting 15-tuple.

At this point in the argument, closed sets analogous to the S_N's of Chap. IV come in. (See text preceding Prop. IV.5.3, and the proof of Thm. IV.7.1.) These sets consist of 16-tuples constructed as follows: the first entry is a vector $\tilde{\varphi} \in \mathbf{R}^{I+1+J}$; the next 15 entries are the 15 components of the second displayed vector in Prop. 4.3. With this notation, S_N is the subset of $\mathbf{R}^{I+1+J} \times \mathbf{R}^n \times C^{3n} \times C^{I+1+J} \times C^{3n} \times C \times C \times C^{4(n\times n)} \times C^d$ whose 16-tuples obey the following relationships. Fix $r \in (q, \overline{q})$.

$$
\begin{aligned}
&x(t) = x_0 + x^1(t) + x^2(t) \quad \forall t && x^1(0) = x^2(0) = 0,\\
&\overline{p}(t) = -\widetilde{\varphi}' \left[\overline{p}^1 + \overline{p}^2(T) - \overline{p}^2(t)\right] \Psi_t \quad \forall t, && \overline{p}^2(0) = 0,\\
&\Phi_t = I + \Phi^1_t + \Phi^2_t \quad \forall t, && \Phi^1(0) = \Phi^2(0) = 0,\\
&\Psi_t = (\Phi_t)^{-1} \quad \forall t, &&\\
&|\widetilde{\varphi}| = 1, &&\\
&\overline{p}^1 = \ell_x(x_T)\Phi_T, &&\\
&x^1(\cdot) \in AC^n, && \textstyle\int_0^T \left|\dot{x}^1(r)\right|^2 dr \le N^2,\\
&z(\cdot) \in AC^{I+1+J}, && \textstyle\int_0^T |\dot{z}(r)|^{\overline{q}/q}\, dr \le N^{\overline{q}/q},\\
&\overline{p}^2(\cdot) \in AC^n, && \textstyle\int_0^T \left|\dot{\overline{p}}^2(r)\right|^{\overline{q}/q} dr \le N^{\overline{q}/q},\\
&h^1(\cdot) \in AC, && \textstyle\int_0^T \left|\dot{h}^1(r)\right|^2 dr \le N^2,\\
&h^2(\cdot) \in AC, && \textstyle\int_0^T \left|\dot{h}^2(r)\right|^{r/q} dt \le N^{r/q},\\
&\Phi^1_\cdot \in AC^{n\times n}, && \textstyle\int_0^T \left|\dot{\Phi}^1_r\right|^2 dr \le N^2,\\
&\left(\dot{x}^1(t), \dot{z}(t), \dot{\overline{p}}^2(t), \dot{h}^1(t), \dot{h}^2(t), \dot{\Phi}^1_t\right) \in \Gamma(t, x(t), \overline{p}^2(t), \Phi_t, \Psi_t) \quad \text{a.e. } [0,T]. &&
\end{aligned}
\tag{4.15}
$$

Here the multifunction Γ is defined by

$$
\Gamma(t,x,p,\Phi,\Psi) = \left\{ \left(f(t,x,u), L(t,x,u), L_x(t,x,u)\Phi, \Psi f(t,x,u), p\Psi f(t,x,u), f_x(t,x,u)\Phi \right) \ : \ u \in U \right\}.
$$

For each fixed choice of its arguments, $\Gamma(t,x,p,\Phi,\Psi)$ is a linear image of the (compact) set assumed to be convex in (h6). Hence $\Gamma(t,x,p,\Phi,\Psi)$ is compact-convex-valued. Moreover, the obvious choice of a function $G(t,x,p,\Phi,\Psi,u)$ will realize $\Gamma(t,x,p,\Phi,\Psi) = \{G(t,x,p,\Phi,\Psi,u) \ : \ u \in U\}$. This function G is continuous in all its arguments (including t, by (h4)). Cesari's (1983) argument proving 8.5.vi(a), p. 296, shows that Γ is upper semicontinuous by set inclusion at all points (t,x,p,Φ,Ψ), so by his theorem 8.5.iv, pp. 293–294, the sets Γ have property (Q). This verifies the hypotheses of his closure theorem 8.6.i, p. 299. The assertion of this theorem is that if a convergent sequence of 16-tuples is chosen from the set S_N, then the limiting 16-tuple continues to obey the differential inclusion in the last line of (4.15). The first six lines clearly remain valid in the limit, and the limiting validity of the remaining six lines of (4.15) is a well-known result in the classical calculus of variations. Therefore the sets S_N are closed for each N.

Not only are the sets S_N closed, but one also has

$$\lim_{N\to\infty} \inf_k P\{m^k \in S_N\} = 1,$$

where we have written m^k for the k-th 16-tuple of random vectors described above. This claim has the same form as the statement immediately preceding Prop. IV.5.4, and its justification is very similar. The first six lines of (4.15) and the last line of (4.15) are already known to hold with probability one for each k and N. The essence of the claim deals with the six integral conditions, which are treated by applying Chebyshev's inequality to the expected value of each. Upon doing this, Def. 2.1(b) gives the following analogue of Prop. IV.5.4.

4.4 Proposition. *The limiting 15-tuple of Prop. 4.3 obeys the following conditions with probability one.*

$$\begin{aligned}
&x(t) = x_0(0) + \int_0^t \dot{x}^1(r)\,dr + x^2(t) \quad \forall t, && x^2(0) = 0,\\
&\overline{p}(t) = -\tilde{\varphi}' \left[\ell_x(x_T)\Phi_T + \overline{p}^2(T) - \int_0^t \dot{\overline{p}}^2(r)\,dr\right]\Phi_t^{-1} \quad \forall t, && \overline{p}^2(0) = 0,\\
&\Phi_t = I + \int_0^t \dot{\Phi}_r^1\,dr + \Phi_t^2 \quad \forall t, && \Phi_0^2 = 0,\\
&x^1,\ z,\ \overline{p}^2,\ h^1,\ h^2,\ \Phi^1, \text{ are absolutely continuous,}\\
&(\dot{x}^1(t), \dot{z}(t), \dot{\overline{p}}^2(t), \dot{h}^1(t), \dot{h}^2(t), \dot{\Phi}_t^1) \in \Gamma(t, x(t), \overline{p}^2(t), \Phi_t, \Phi_t^{-1}) \text{ a.e. } [0,T].
\end{aligned}$$

Next we must produce an appropriate control $u(t,\omega)$. This requires that the probability space $(\Omega, \mathcal{F}, P)$ on which the limiting 15-tuple of Prop. 4.3 is defined be equipped with a filtration. A sufficiently large filtration may be defined in terms of the continuous process

$$\psi_t = \left(X_0, A, x_t, x_t^1, x_t^2, z_t, \overline{p}_t, \overline{p}_t^2, h_t^1, h_t^2, \Phi_t, \Phi_t^1, \Phi_t^2, w_t\right).$$

We take for $\mathcal{F}_t$ the filtration generated by ψ_t. (The constant processes in the first two components of ψ ensure that X_0 and A are $\mathcal{F}_0$-adapted.) Evidently ψ_t is the limit in distribution (on the space of continuous functions) of the continuous processes

$$\psi_t^k = \left(X_0^k, A^k, x_t^k, x_t^{1,k}, x_t^{2,k}, z_t^k, \overline{p}_t^k, \overline{p}_t^{2,k}, h_t^{1,k}, h_t^{2,k}, \Phi_t^k, \Phi_t^{1,k}, \Phi_t^{2,k}, w_t^k\right).$$

We will use these processes when applying Prop. 4.2 for each k.

4.5 Proposition. *The limiting process $w(\cdot)$ in Prop. 4.3 is an $\mathcal{F}_t$-Brownian motion in $\mathbf{R}^d$. Moreover, there is an $\mathcal{F}_t$-adapted process $u: [0,T] \times \Omega \to U$ such that*

$$\begin{aligned}
x_t &= x_0(0) + \int_0^t f(r, x_r, u_r)\, dr + \int_0^t \sigma(r, x_r)\, dw_r \qquad \forall t,\\
\overline{p}_t &= -\widetilde{\varphi}' \left[\ell_x(x_T)\Phi_T + \int_t^T L_x(r, x_r, u_r)\Phi_r\, dr \right] \Phi_t^{-1} \qquad \forall t,\\
\Phi_t &= I + \int_0^t f_x(r, x_r, u_r)\Phi_r\, dr + \int_0^t \sigma_x(r, x_r)\Phi_r\, dw_r \qquad \forall t,\\
z_t &= \int_0^t L(r, x_r, u_r)\, dr \qquad \forall t,\\
h_t^1 &= \int_0^t \Phi_r^{-1} f(r, x_r, u_r)\, dr \qquad \forall t,\\
h_t^2 &= \int_0^t \overline{p}_r^2 \Phi_r^{-1} f(r, x_r, u_r)\, dr \qquad \forall t.
\end{aligned}$$

Proof. Being the limit in distribution of Brownian motion processes, $w(\cdot)$ is certainly a Brownian motion. It is adapted to $\mathcal{F}_t$ by construction. We show that it is actually an $\mathcal{F}_t$-Brownian motion in the course of justifying the six integral representations listed above.

First let us treat the stochastic integrals. Since each triple $(w^k, x^{2,k}, \Phi^{2,k})$ is a continuous second-order $\mathcal{F}_t(\psi^k)$-martingale on $(\Omega^k, \mathcal{F}^k, P^k)$ and since $\psi^k \xrightarrow{\mathcal{D}} \psi$, Prop. IV.3.1 implies that the limit process (w, x, Φ^2) is a continuous second-order $\mathcal{F}_t^{\psi}$-martingale. The quadratic variation of the limiting martingale is a little awkward to write down, because Φ denotes an $n \times n$ matrix. Let us temporarily think of Φ as a vector in $\mathbf{R}^{n^2}$ by writing the first column above the second, and so on. Then we are studying continuous martingales v in $\mathbf{R}^{d+n+n^2}$. Along the sequence, the martingale $v^k = (w^k, x^{2,k}, \Phi^{2,k})$ is defined by

$$v^k(t) = \begin{bmatrix} w^k(t) \\ x^{2,k}(t) \\ \Phi^{2,k}(t) \end{bmatrix} = \int_0^t \begin{bmatrix} I \\ \sigma(r, x_r^k) \\ \Sigma(r, x_r^k, \Phi_r^k) \end{bmatrix} dw_r^k,$$

where the $(d + n + n^2) \times d$ integrand matrix is built up of a $d \times d$ identity matrix, an $n \times d$ matrix σ, and an $n^2 \times d$ matrix Σ containing an appropriate arrangement of the elements of $\sigma_x \Phi$. The quadratic variation of the martingale v_t^k is given by

$$\langle v^k \rangle_t = \int_0^t \begin{bmatrix} I & \sigma' & \Sigma' \\ \sigma & \sigma\sigma' & \sigma\Sigma' \\ \Sigma & \Sigma\sigma' & \Sigma\Sigma' \end{bmatrix}^k dr.$$

This means that $v^k (v^k)' - \langle v^k \rangle$ is a continuous $(d+n+n^2) \times (d+n+n^2)$-matrix valued martingale. Since $x^k \xrightarrow{\mathcal{D}} x$ and $\Phi^k \xrightarrow{\mathcal{D}} \Phi$, the continuity properties of σ, Σ, and of integral functionals imply that the processes $\langle v^k \rangle_t$ converge in distribution to the process, for which we use the suggestive notation $\langle v \rangle_t$, given by

$$\langle v \rangle_t = \int_0^t \begin{bmatrix} I & \sigma' & \Sigma' \\ \sigma & \sigma\sigma' & \sigma\Sigma' \\ \Sigma & \Sigma\sigma' & \Sigma\Sigma' \end{bmatrix} dr.$$

Thus $(v^k)(v^k)' - \langle v^k \rangle \xrightarrow{\mathcal{D}} vv' - \langle v \rangle$; since the left side is a uniformly integrable sequence of $\mathcal{F}_t(\psi^k)$-martingales, the right side is an $\mathcal{F}_t^{\psi}$-martingale by Prop. IV.3.1. By definition, it follows that the process we have denoted by $\langle v \rangle_t$ really is the quadratic variation of v_t, and that it is realized here as the integral of the following matrix times its transpose:

$$\begin{bmatrix} I \\ \sigma(r, x_r) \\ \Sigma(r, x_r, \Phi_r) \end{bmatrix}.$$

This $(d + n + n^2) \times d$matrix has (row) rank identically equal to d, so by Prop. IV.3.3, the space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ itself carries an $\mathcal{F}_t$-Brownian motion $\tilde{w}_t$ with values in $\mathbf{R}^d$ such that

$$\begin{bmatrix} w(t) \\ x^2(t) \\ \Phi^2(t) \end{bmatrix} = v(t) = \int_0^t \begin{bmatrix} I \\ \sigma(r, x_r) \\ \Sigma(r, x_r, \Phi_r) \end{bmatrix} d\tilde{w}_r.$$

The first d components of this equation imply that $\tilde{w}_t = w_t$ for all t, except on a negligible set of ω-values, so these processes are indistinguishable. Then, reverting from vectors in $\mathbf{R}^{n^2}$ to matrices in $\mathbf{R}^{n \times n}$, the second two blocks of this identity yield the desired stochastic integral representations:

$$x^2(t) = \int_0^t \sigma(r, x_r)\, dw_r,$$
$$\Phi^2(t) = \int_0^t \sigma_x(r, x_r)\Phi_r\, dw_r.$$

The second step is to produce a suitable control $u(t, \omega)$. As in Chap. IV, this can be done by appealing to the selection lemma given by Beneš (1971). We consider the measure space $M = [0, T] \times \Omega$ together with the σ-field $\mathcal{M}$ with respect to which $\mathcal{M}$-measurability is equivalent to ψ_t-adaptedness. (See text following Lemma IV.4.2.) We take $R = \mathbf{R}^n \times \mathbf{R}^{I+1+J} \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R} \times \mathbf{R}^{n \times n}$, and define $k\colon M \times U \to R$ and $y\colon M \to U$ by

$$k(t, \omega, u) = \Big(f\big(t, x(t, \omega), u\big),\ L\big(t, x(t, \omega), u\big), L_x\big(t, x(t, \omega), u\big)\Phi(t, \omega),$$
$$\Phi(t, \omega)^{-1} f\big(t, x(t, \omega), u\big),\ \bar{p}^2(t, \omega)\Phi(t, \omega)^{-1} f\big(t, x(t, \omega), u\big),\ f_x\big(t, x(t, \omega), u\big)\Phi(t, \omega)\Big),$$

$$y(t, \omega) = \Big(\dot{x}^1(t, \omega), \dot{z}(t, \omega), \dot{\bar{p}}^2(t, \omega), \dot{h}^1(t, \omega), \dot{h}^2(t, \omega), \dot{\Phi}^1(t, \omega)\Big).$$

Then $k(t, \omega, u)$ is $\mathcal{M}$-measurable in (t, ω) for each fixed $u \in U$, and continuous in u for each fixed (t, ω). Also y is $\mathcal{M}$-measurable (because, for example, $\dot{x}_t = \lim_{h \to 0^+} \dfrac{x(t) - x(t - h)}{h}$ is x_t-adapted) and obeys $y(t,\omega) \in k(t,\omega,U) = \Gamma(t, x_t, \overline{p}_t^2, \Phi_t, \Phi_t^{-1})$. The conclusion of Beneš (1971), Lemma 5, p. 460 is that there is an $\mathcal{M}$-measurable mapping $u: [0,T] \times \Omega \to U$ such that $y(t,\omega)$ equals

$$\Big(f(t,x(t,\omega),u(t,\omega)),\ L(t,x(t,\omega),u(t,\omega)),\ L_x(t,x(t,\omega),u(t,\omega))\Phi(t,\omega),$$
$$\Phi(t,\omega)^{-1} f(t,x(t,\omega),u(t,\omega)),\ \overline{p}^2(t,\omega)\Phi(t,\omega)^{-1} f(t,x(t,\omega),u(t,\omega)),\ f_x(t,x(t,\omega),u(t,\omega))\Phi(t,\omega)\Big).$$

Recalling the definition of $y(t,\omega)$ and the representations of x^2 and Φ^2 already given, the integral representations of this proposition now follow from Prop. 4.4. ////

Let us summarize our findings and complete the study of convergence.

4.6 Theorem. *Let $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0)$ be a unit vector obtained from a sequence of perpendiculars as in (4.10). Then solutions x^k to $P(\alpha^k, \lambda^k)$ can be found such that, along a subsequence, $x^k \xrightarrow{\mathcal{D}} x$ for some process x. This limit process is a solution of the dynamic equation (1.1) corresponding to a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, w_t, and a control u which solves $P(0,0)$. Moreover, there exists a Lebesgue null set $\mathcal{N}(u) \subseteq [0,T]$ and a process $\overline{p}_t$ such that for each $t \notin \mathcal{N}(u)$, any $\mathcal{F}_t$-measurable and U-valued random variable v obeys*

$$\mathbf{E}H(t, x_t, \overline{p}_t, u_t, \widetilde{\varphi}) \geq \mathbf{E}H(t, x_t, \overline{p}_t, v, \widetilde{\varphi}), \tag{4.16}$$

$$\mathbf{E}A'\overline{p}_0 = -\widetilde{\beta}. \tag{4.17}$$

Here the pre-Hamiltonian H is defined by $p'f - \widetilde{\varphi}'L$ as in (4.9), and the process $\overline{p}_t$ is given explicitly by

$$\overline{p}_t' = -\widetilde{\varphi}' \left[\ell_x(x_T)\Phi_T + \int_t^T L_x(r, x_r, u_r)\Phi_r \, dr \right] \Phi_t^{-1}, \tag{4.18}$$

where $\Phi.$ is the $n \times n$ matrix process defined in (3.13a) and corresponding to (x,u). The vector $\widetilde{\varphi}$ obeys $\widetilde{\varphi}_0 \geq 0$, together with the complementary slackness condition

$$\widetilde{\varphi}_- \geq 0, \qquad \langle \widetilde{\varphi}_-, \Lambda_-[0,u] \rangle = 0. \tag{4.19}$$

Proof. We have already constucted the limit process x and its corresponding control u, and verified that $\overline{p}_t$ and Φ_t have the correct representations. To see why u solves $P(0,0)$, recall that $\Lambda[\alpha^k, u^k]$ converges to $(-\rho, V(0,0), 0)$ by assumption for some $\rho \geq 0$ in $\mathbf{R}^I$, and that both $z^k(T) \xrightarrow{\mathcal{D}} z(T)$

and $x^k(T) \xrightarrow{\mathcal{D}} x(T)$. Since $\Lambda[\alpha^k, u^k] = \mathbf{E}\left[\ell(x_T^k) + z^k(T)\right]$, a standard uniform integrability argument implies that $\mathbf{E}\left[\ell(x_T) + z(T)\right] = \Lambda[0, u] = \left(-\rho, V(0,0), 0\right)$. This shows that u satisfies all the constraints of $P(0,0)$ and attains the infimum, as claimed.

Since $(\widetilde{\varphi}_-, \widetilde{\varphi}_0)$ is a limit of vectors with nonnegative components, its components must also be nonnegative. To prove the complementary slackness condition (4.19), note that Prop. 4.2 gives

$$\left\langle \widetilde{\varphi}_-^k, \Lambda_-[\alpha^k, u^k] + \lambda_-^k \right\rangle = 0 \quad \forall k.$$

We know that $\widetilde{\varphi}_-^k \to \widetilde{\varphi}_-$ and $\lambda_-^k \to 0$, while $\Lambda_-[\alpha^k, u^k] \to \Lambda_-[0, u]$ has just been shown above. Thus we get (4.19) in the limit as $k \to \infty$.

Only (4.16) and (4.17) remain to check. The easier of these is (4.17), for which it is convenient to observe that for any $r \in (0, \overline{q}/q)$, the sequence $\mathbf{E}\left\|\overline{p}^k\right\|^r$ is bounded. This follows from the representation of each $\overline{p}^k$ analogous to (4.18), the k-independence of the constants in (3.16), and the growth conditions of (h5). (We used it before to prove Thm. 3.2.) It implies that $\mathbf{E}\left|(A^k)'\overline{p}_0^k\right|^r$ is a bounded sequence, and hence uniform integrability gives

$$\mathbf{E}(A^k)'\overline{p}_0^k \to \mathbf{E}A'\overline{p}_0 \quad \text{as } k \to \infty.$$

Since the left side here equals $-\widetilde{\beta}^k$ and we have $\widetilde{\beta}^k \to \widetilde{\beta}$ by assumption, (4.17) holds.

To prove (4.16), let us observe that the pre-Hamiltonian integrals

$$h^k(t) = \int_0^t H(r, x_r^k, \overline{p}_r^k, u_r^k, \widetilde{\varphi}^k)\, dr$$

converge in distribution (on the space C) to the limiting Hamiltonian integral

$$h(t) = \int_0^t H(r, x_r, \overline{p}_r, u_r, \widetilde{\varphi})\, dr.$$

Indeed, we have

$$\begin{aligned} h^k(t) &= \int_0^t \left[(\overline{p}_r^k)' f(r, x_r^k, u_r^k) - (\widetilde{\varphi}^k)' L(r, x_r^k, u_r^k)\right] dr \\ &= -(\widetilde{\varphi}^k)' \left[\int_0^t \left(\ell_x(x_T^k)\Phi_T^k + \int_r^T L_x(\rho, x_\rho^k, u_\rho^k)\Phi_\rho^k\, d\rho\right) \left(\Phi_r^k\right)^{-1} f(r, x_r^k, u_r^k)\, dr + \int_0^t L(r, x_r^k, u_r^k)\, dr\right] \\ &= -(\widetilde{\varphi}^k)' \left[\left(\ell_x(x_T^k)\Phi_T^k + \overline{p}^{2,k}(T)\right) h^{1,k}(t) - h^{2,k}(t) + z^k(t)\right]. \end{aligned}$$

The RHS here is a continuous image of ψ_t^k, which is known to converge in distribution. Since convergence in distribution is preserved by continuous mappings, we conclude that $h^k \xrightarrow{\mathcal{D}} h$.

Now let ℓ denote the dimension of the process ψ, and consider a measurable C_t^ℓ-adapted mapping $v: [0,T] \times C^\ell \to U$ which is continuous in its second argument for every value of the first. The growth conditions of (h2) and (h5) allow one to use the dominated convergence theorem to show that the mapping of $C^n \times C^n \times C^\ell \times \mathbf{R}^{I+1+J}$ into $\mathbf{R}$ defined by

$$(x, p, \psi, \varphi) \longmapsto \int_0^t H(r, x_r, p_r, v(r, \psi), \varphi)\, dr$$

is continuous for each t. Since convergence in distribution is preserved by continuous mappings, it follows that for each fixed t,

$$\int_0^t H(r, x_r^k, \overline{p}_r^k, v(r, \psi^k), \widetilde{\varphi}^k)\, dr \xrightarrow{\mathcal{D}} \int_0^t H(r, x_r, \overline{p}_r, v(r, \psi), \widetilde{\varphi})\, dr.$$

A standard uniform integrability argument shows that the expectations of the left-hand sides converge to that of the right-hand side. Now using ψ^k in Prop. 4.2 for each k, we find that for any $s \in (0,T]$ and $\varepsilon > 0$, one has

$$0 \leq \frac{1}{\varepsilon} \int_{s-\varepsilon}^s \left[\mathbf{E}H(r, x_r^k, \overline{p}_r^k, u_r^k, \widetilde{\varphi}^k) - \mathbf{E}H(r, x_r^k, \overline{p}_r^k, v(r, \psi^k), \widetilde{\varphi}^k)\right] dr.$$

Therefore the arguments above allow us to take the limit as $k \to \infty$ to get

$$0 \leq \frac{1}{\varepsilon} \int_{s-\varepsilon}^s \left[\mathbf{E}H(r, x_r, \overline{p}_r, u_r, \widetilde{\varphi}) - \mathbf{E}H(r, x_r, \overline{p}_r, v(r, \psi), \widetilde{\varphi})\right] dr.$$

Now since the adpated maps $v(r, \psi)$ with continuous ψ-dependence are dense in the measurable adapted maps $v(r, \psi)$ in the sense of almost-sure convergence with respect to $dt \times dP_\psi$, this last relationship must actually hold for all measurable and ψ_t-adapted processes v. The set of all such v's may be called the set of all ψ-feedback controls, of which we have described a countable subset $\mathcal{V}(\psi)$ in the text preceding Thm. 3.2. By Lemma 1.1, there is a null set $\mathcal{N}(u, \psi)$ such that for any $v \in \mathcal{V}(\psi)$ and $s \notin \mathcal{N}(u, \psi)$, one has

$$\lim_{\varepsilon \to 0+} \frac{1}{\varepsilon} \int_{s-\varepsilon}^s \left[H(r, x_r, \overline{p}_r, u_r, \widetilde{\varphi}) - H(r, x_r, \overline{p}_r, v_r, \widetilde{\varphi})\right] dr = H(s, x_s, \overline{p}_s, u_s, \widetilde{\varphi}) - H(s, x_s, \overline{p}_s, v_s, \widetilde{\varphi}).$$

By uniform integrability, we find that

$$\mathbf{E}H(s, x_s, \overline{p}_s, u_s, \widetilde{\varphi}) \geq \mathbf{E}H(s, x_s, \overline{p}_s, v_s, \widetilde{\varphi}) \qquad \forall s \notin \mathcal{N}(u, \psi),\ \forall v \in \mathcal{V}(\psi).$$

This is precisely the situation in which the proof of Thm. 3.2 shows that the density of $\mathcal{V}_s(\psi)$ in $\mathcal{U}_s(\psi)$ implies the global conclusion (4.16). ////

The convergence analysis of this section is the foundation of the applied results to follow. Before turning to these, however, let us present a somewhat stronger Hamiltonian inequality. Consider the adapted adjoint process $p_t = \mathbf{E}[\overline{p}_t \mid \mathcal{F}_t]$. With respect to this process, (4.16) and (4.17) may be replaced by

$$H(t, x_t, p_t, u_t, \widetilde{\varphi}) \geq H(t, x_t, p_t, v, \widetilde{\varphi}) \quad \text{a.s., } \forall t \notin \mathcal{N}(u, \psi),\ \forall v \in \mathcal{U}_t(\psi). \tag{4.21}$$

$$\mathbf{E}A' p_0 = -\widetilde{\beta} \tag{4.22}$$

Equation (4.22) follows from (4.17) by conditioning; to see why (4.21) holds, note that since H is linear in p, and both f and L are $\mathcal{F}_t$-adapted, the inequality in (4.21) is equivalent to

$$\mathbf{E}[H(t, x_t, \overline{p}_t, u_t, \widetilde{\varphi}) \mid \mathcal{F}_t] \geq \mathbf{E}[H(t, x_t, \overline{p}_t, v, \widetilde{\varphi}) \mid \mathcal{F}_t].$$

Thus if (4.21) is false, there must be a time $t \notin \mathcal{N}(u)$, a random variable $u \in \mathcal{U}_t$, and a set $F \in \mathcal{F}_t$ with $P(F) > 0$ for which the reverse inequality prevails. But in this case the new random variable

$$v(\omega) = u_t(\omega) I\{\omega \notin F\} + u(\omega) I\{\omega \in F\}$$

is an element of $\mathcal{U}_t(\psi)$ for which conditioning on $\mathcal{F}_t$ gives

$$\mathbf{E}H(t, x_t, \overline{p}_t, u_t, \widetilde{\varphi}) < \mathbf{E}H(t, x_t, \overline{p}_t, v, \widetilde{\varphi}).$$

Thus the falsity of (4.21) contradicts the proven statement (4.16). Conclusion (4.21) cannot be false.

4.7 Definition (Multiplier Sets). Let a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual hypotheses be given. Suppose it supports an $\mathcal{F}_t$-Brownian motion w_t with values in $\mathbf{R}^d$ and an $\mathcal{F}_t$-adapted process $u\colon [0,T] \times \Omega \to U$ which solves $P(0,0)$. Then a pair (φ, p) consisting of a vector $\varphi \in \mathbf{R}^{I+1+J}$ and an $\mathcal{F}_t$-adapted continuous process $p\colon [0,T] \times \Omega \to \mathbf{R}^n$ is called an *index φ_0 multiplier corresponding to u* if it satisfies the following conditions. There is a null set $\mathcal{N}(u) \subseteq [0,T]$ such that

$$\begin{aligned}
&H(t, x_t, p_t, u_t, \varphi) \geq H(t, x_t, p_t, v, \varphi) \ \text{ a.s. } \ \forall t \notin \mathcal{N}(u)\ \forall v \in \mathcal{U}_t,\\
&p_t' = -\varphi' \mathbf{E}\left[\ell_x(x_T)\Phi_T \Phi_t^{-1} + \int_t^T L_x(r, x_r, u_r)\Phi_r \Phi_t^{-1}\, dr \;\middle|\; \mathcal{F}_t \right],\\
&\Phi_t = I + \int_0^t f_x(r, x_r, u_r)\Phi_r\, dr + \int_0^t \sigma_x(r, x_r)\Phi_r\, dw_r,
\end{aligned}$$

$$\varphi_0 \geq 0, \qquad \varphi_- \geq 0, \qquad \langle \varphi_-, \Lambda_-[0,u] \rangle = 0.$$

Here $\mathcal{U}_t$ denotes the set of all $\mathcal{F}_t$-measurable random variables $v: \Omega \to U$.

The set of all such pairs (φ, p) corresponding to u is denoted $M^{\varphi_0}(u)$. A mapping Δ from the space of multipliers into $\mathbf{R}^{a+I+J}$ is defined as follows:

$$\Delta(\varphi, p) = \Big(-\mathbf{E}A'p_0,\ \varphi_{\mp}\Big).$$

We denote by Y the set of all possible control processes u solving $P(0,0)$: then $M^{\varphi_0}(Y)$ signifies the collection of all multiplier pairs corresponding to some solution of $P(0,0)$, and $\Delta\big[M^{\varphi_0}(Y)\big]$ is the image set of all these pairs under the mapping Δ.

Consider the following two cones.

$$N = \big\{r(\varsigma, -1) \,:\, r > 0,\ \varsigma \in \Delta\big[M^1(Y)\big] \cap \partial V(0)\big\}$$

$$N^\infty = \big\{(\varsigma, 0) \,:\, \varsigma \in \Delta\big[M^0(Y)\big] \cap \partial^\infty V(0)\big\}.$$

For any unit vector $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0)$ obtained as a limit of perpendiculars as in (4.10), Thm. 4.6 asserts that for some process p the pair $(\widetilde{\varphi}, p)$ lies in $M^{\widetilde{\varphi}_0}(Y)$. Indeed, if $\widetilde{\varphi}_0 > 0$ then the positive homogeneity of the conditions defining a multiplier implies that $\Big(\widetilde{\varphi}/\widetilde{\varphi}_0,\ p/\widetilde{\varphi}_0\Big)$ lies in $M^1(Y)$. But Thm. 4.6 also shows that $\Delta\Big(\widetilde{\varphi}/\widetilde{\varphi}_0, p/\widetilde{\varphi}_0\Big) = (\widetilde{\beta}/\widetilde{\varphi}_0, \widetilde{\varphi}_{\mp}/\widetilde{\varphi}_0)$ is a vector for which $\Big(\widetilde{\beta}/\widetilde{\varphi}_0,\ \widetilde{\varphi}_{\mp}/\widetilde{\varphi}_0, -1\Big) \in N_{\operatorname{epi} V}(0, V(0))$ by the proximal normal formula. Upon applying the definition of $\partial V(0)$ in terms of this normal cone, we find that $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, -\widetilde{\varphi}_0) \in N$. On the other hand, if $\widetilde{\varphi}_0 = 0$ then $\Delta(\widetilde{\varphi}, p) = (\widetilde{\beta}, \widetilde{\varphi}_{\mp})$ is a vector for which $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, 0) \in N_{\operatorname{epi} V}(0, V(0))$ by the proximal normal formula. By definition of $\partial^\infty V(0)$, it follows that $(\widetilde{\beta}, \widetilde{\varphi}_{\mp}, 0) \in N^\infty$. These arguments lead to the main result of this chapter.

4.8 Theorem. *The generalized gradient of V obeys*

$$\partial V(0) = \overline{\mathrm{co}}\Big(\Delta\big[M^1(Y)\big] \cap \partial V(0) + \Delta\big[M^0(Y)\big] \cap \partial^\infty V(0)\Big).$$

If the cone $\Delta\big[M^0(Y)\big]$ is pointed then the closure operation is redundant, and one also has

$$\partial^\infty V(0) = \mathrm{co}\left(\Delta\left[M^0(Y)\right] \cap \partial^\infty V(0)\right).$$

Proof. The definitions of $\partial V(0)$ and $\partial^\infty V(0)$ imply that

$$N_{\operatorname{epi} V}(0, V(0)) \supseteq N \cup N^\infty.$$

Indeed, since the normal cone is always closed and convex, one has

$$N_{\mathrm{epi}\,V}(0, V(0)) \supseteq \overline{\mathrm{co}}[N \cup N^{\infty}].$$

But according to the proximal normal formula, the normal cone is contained in the closed convex cone generated by certain limits of perpendiculars. We have just shown that all such limits of perpendiculars are actually elements of $N \cup N^{\infty}$: therefore

$$N_{\mathrm{epi}\,V}(0, V(0)) \subseteq \overline{\mathrm{co}}[N \cup N^{\infty}].$$

This implies that $N_{\mathrm{epi}\,V}(0, V(0)) = \overline{\mathrm{co}}[N \cup N^{\infty}]$, whereupon the assertions of the theorem follow from Prop. II.6.2. ////

Section 5. The Stochastic Maximum Principle for Constrained Problems

Among the many significant consequences of Thm. 4.8 is the stochastic maximum principle (SMP) for problems with soft constraints. The SMP concerns the existence of nontrivial multipliers as defined in 4.7. Suppose first that $M^0(Y) = \{0\}$. Since this cone is pointed, Thm. 4.8 asserts that $\partial^{\infty} V(0) = \{0\}$ and that $\partial V(0) = \mathrm{co}\{\Delta[M^1(Y)] \cap \partial V(0)\}$. According to Prop. I.2.5, we may therefore write

$$M^0(Y) = \{0\} \Longrightarrow \partial^{\infty} V(0) = \{0\} \Longrightarrow \partial V(0) \neq \emptyset \Longrightarrow M^1(Y) \neq \emptyset.$$

This series of implications can be summarized as follows:

$$M^1(Y) \cup [M^0(Y) \setminus \{0\}] \neq \emptyset. \tag{5.1}$$

Line (5.1) is a concise statement of the Stochastic Maximum Principle.

5.1 Theorem (Stochastic Maximum Principle). *Assume (h1)–(h6). If problem $P(0,0)$ has a feasible control process then it has a solution. Moreover, at least one of the optimal controls in the set Y has a multiplier pair as defined in 4.7 for which either $\varphi_0 = 1$ or $(-\mathbf{E}A'p_0, \varphi_{\mp}) \neq (0,0)$.*

Proof. Immediate from Thm. IV.7.1 and line (5.1). ////

We will compare Thm. 5.1 to other forms of this result later. First let us consider some of its consequences.

The interpretation of the multipliers in the SMP Thm. 5.1 is clearest in the "normal" case: problem $P(0,0)$ is said to be *normal* if it has no solution with a nontrivial multiplier (φ, p) for which $\varphi_0 = 0$—in other words, if $M^0(Y) = \{0\}$. In this case we get the following result.

5.2 Proposition. *Suppose $P(0,0)$ is normal. Then V is (finite and) Lipschitz near 0, and $\partial V(0)$ is a compact convex set obeying*

$$\emptyset \neq \operatorname{ext} \partial V(0) \subseteq \Delta\big[M^1(Y)\big].$$

Proof. If $M^0(Y) = \{0\}$ then $\partial^\infty V(0) = \{0\}$, so V is Lipschitz near 0 and $\partial V(0)$ is a nonempty compact convex subset of $\mathbf{R}^{a+I+J}$ by Prop. I.2.5. Such sets have extreme points by the Krein-Milman theorem. And a well-known converse of the Krein-Milman theorem asserts that if $C \subseteq \mathbf{R}^{a+I+J}$ is compact, then $\operatorname{co} C$ is also compact and $\operatorname{ext}(\operatorname{co} C) \subseteq C$. The result follows upon taking $C = \Delta\big[M^1(Y)\big] \cap \partial V(0)$. ////

Corollary. *Suppose $P(0,0)$ is normal and $\Delta\big[M^1(Y)\big]$ is a singleton, say $\big(-\mathbf{E}A'p_0, \varphi_{\mp}\big)$. Then V is strictly differentiable at 0, with strict derivative $D_s V(0) = \big(-\mathbf{E}A'p_0, \varphi_{\mp}\big)$.*

Proof. See Clarke (1983), Prop. 2.2.4, p. 33. (Clarke defines strict differentiability on p. 30.) ////

Prop. 5.2 also shows that the normality of problem $P(0,0)$ implies the stability of the system of soft constraints defining the problem's structure. For the finiteness of V in a neighbourhood of 0 shows that each problem $P(\alpha, \lambda)$ for (α, λ) near $(0,0)$ has a feasible control. Note that the sufficient condition guaranteeing this desirable situation, namely the absence of nontrivial multipliers with $\varphi_0 = 0$, is formulated only in terms of the stochastic dynamics and the soft constraints: the objective functional is irrelevant when $\varphi_0 = 0$. The stability of systems of inequalities in the deterministic case has been studied by S.M. Robinson (1976) and Clarke (1983), Sections 6.3–6.4.

Scholium. The new approach to the SMP afforded by Thm. 4.8 has much to recommend it. First, it relies completely on the simple geometrical notion of perpendicularity rather than on the more theoretical "abstract variational theory of Neustadt" used by Kushner (1972) or the "cone of variations" approach of Haussmann (1985). Second, the relationship between existence theory and necessary conditions is clarified—not only in the broad outlines of the proof of Thm. 4.8, but also in the close parallel between the convergence arguments of Props. 4.3–4.5 and those of Props. IV.5.2–IV.5.5. Finally, Thm. 4.8 gives much more than just the SMP. It provides a rigorous defence of the interpretation of the multipliers corresponding to an optimal trajectory as the marginal costs corresponding to each constraint. These three advantages are unique to our approach, and confirm the value of proximal normal analysis in this context.

Along with its good points, our approach has several shortcomings. First, it makes implicit rather that explicit use of existence theory. By assuming the existence of a solution explicitly, the authors cited above have been able to prove the SMP without assuming that U is compact and without any convexity condition analogous to our (h6). Moreover, their explicit existence assumption allows a strong formulation, in which it is possible to prove that *every* optimal control process has a nontrivial multiplier. Now we have shown in Chap. IV that for the purposes of existence theory, the compactness hypothesis on U can be replaced by a coercivity condition on L. There may be a modification of our arguments in this chapter which will allow that more general existence result to be used to eliminate the compactness condition. As for (h6), it is quite possible that further analysis will show that it can be omitted from the assumptions of Thm. 5.1. Such a result would be based on conditions under which an optimal control process $\widehat{u}$ for a problem violating (h6) retains its optimality when the problem is "convexified" so that (h6) holds. "Relaxation in Stochastic Optimal Control" is a topic well worth considering, but one which lies beyond the scope of this work.

Our inability to show that every optimal control has a nontrivial multiplier can also be traced to the demands of existence theory. For that theory ties us to the weak formulation, in which even the simplest problem has infinitely many representations. Just which of these is selected by the limiting procedure cannot be foreseen with certainty. To compound the difficulty, we complete the limiting procedure by selecting some optimal control which gives a suitable representation to the multipliers:

whether there exist other optimal controls which do not yield such a representation is unknown. In the deterministic theory, difficulties like this can be overcome by assuming that every solution to the nominal problem $P(0,0)$ acts like a unique solution. This assumption is made rigorous by a minor modification of the objective functional which singles out a preassigned solution as the unique solution to a related problem for which the multipliers are the same as those for the original problem. (An example of this approach may be found in our paragraph on "Multiple Solutions" following Prop. III.4.3.) But to effect such a modification in the stochastic case, we would need some numerical measure not just of deviations from a preassigned optimal control on its own base space, but of changes between the base spaces themselves. And the whole philosophy behind the weak formulation insists that all probabilistic structures must be regarded as interchangeable precisely because there is no observable difference between them.

The form of the multipliers in Def. 4.7 raises an interesting question concerning the role of feedback control laws. As we saw in Chap. IV, the value function V may be defined in terms of feedback controls or in terms of adapted controls. The feedback controls are widely considered the more practical choice. It is natural to ask whether the generalized gradient of V is captured by multipliers corresponding to feedback controls, or whether the larger class of adapted controls is necessary. Our result is expressed in terms of the adapted controls. To use the feedback controls instead, we would have to select an $\mathcal{F}_t^x$-adapted control $u(t,\omega)$ in Thm. 4.6 instead of simply a $\mathcal{F}_t^\psi$-adapted one. Such an effort would presumably use Wong's theorem in some way, but the fact that the limiting matrix processes Φ_t may fail to be x_t-adapted makes it difficult to see how to make this approach work. Haussmann (1983) studied convergence in distribution of stochastic extremals in a rather different setting, but apparently came to the same conclusion.

More General Perturbations. It is not difficult to extend the analysis given above to treat problems in which the finite-dimensional parameter α affects more than simply the initial conditions. Suppose, for instance, that we let $P(\alpha,\lambda)$ refer to the problem

$$\min_{u\in\mathcal{U}}\{\Lambda_0[\alpha,u] : dx_t = f(t,x_t,\alpha,u_t)\,dt + \sigma(t,x_t,\alpha)\,dw_t, \quad x(0)=x_0(\alpha),$$

$$\Lambda_i[\alpha,u] \le -\lambda_i, \qquad i=-1,-2,\ldots,-I,$$

$$\Lambda_j[\alpha,u] = -\lambda_j, \qquad j=1,2,\ldots,J\},$$

where $\Lambda_k[\alpha, u] := \mathbf{E}\left[\ell_k(x_T, \alpha) + \int_0^T L_k(t, x_t, \alpha, u_t)\, dt\right]$ $\forall k$. The value function $V(\alpha, \lambda) = \inf P(\alpha, \lambda)$ for this problem is the same as for the problem $\widetilde{P}(\alpha, \lambda)$ defined by

$$\begin{aligned}\min_{u \in \mathcal{U}} \{\widetilde{\Lambda}_0[\alpha, u] \, : \, & dx_t = f(t, x_t, y_t, u_t)\, dt + \sigma(t, x_t, y_t)\, dw_t, \quad x(0) = x_0(\alpha), \\ & dy_t = 0\, dt, \quad y(0) = \alpha, \\ & \widetilde{\Lambda}_i[\alpha, u] \leq -\lambda_i, \qquad i = -1, -2, \ldots, -I, \\ & \widetilde{\Lambda}_j[\alpha, u] = -\lambda_j, \qquad j = 1, 2, \ldots, J\},\end{aligned}$$

where $\widetilde{\Lambda}_k[\alpha, u] := \mathbf{E}\left[\ell_k(x_T, y_T) + \int_0^T L_k(t, x_t, y_t, u_t)\, dt\right]$ $\forall k$. The transformation of α into an additional state variable $y \in \mathbf{R}^a$ makes it clear how to interpret the standing hypotheses (h1)–(h6) in the current setting. Now $\widetilde{P}$ is a problem to which the results above apply. Translating Def. 4.7 into the context of $\widetilde{P}$ gives the corresponding notions appropriate to our new problem P. (Notice the similarity with the necessary conditions of Thm. 3.4 for an unconstrained problem of similar form.)

5.3 Definition. Suppose a feasible control u for problem $P(0,0)$ is given. A triple (φ, p, q) consisting of a vector $\varphi \in \mathbf{R}^{I+1+J}$ and a $\mathcal{F}_t$-adapted continuous process $(p, q) \colon [0, T] \times \Omega \to \mathbf{R}^n \times \mathbf{R}^a$ is an *index φ_0 multiplier for u* if the pair (u, x) solves $P(0,0)$ and obeys the following conditions.

$$\begin{aligned}& H(t, x_t, p_t, u_t, \varphi) \geq H(t, x_t, p_t, v, \varphi) \;\text{ a.s. } \;\forall t \notin \mathcal{N}(u), \;\forall v \in \mathcal{U}_t, \\ & p'_t = -\varphi' \mathbf{E}\left[\ell_x(x_T, 0)\Phi_T\Phi_t^{-1} + \int_t^T L_x(r, x_r, 0, u_r)\Phi_r\Phi_t^{-1}\, dr \;\middle|\; \mathcal{F}_t\right], \\ & q'_t = -\varphi' \mathbf{E}\left[\ell_x(x_T, 0)\Psi(T, t) + \ell_\alpha(x_T, 0) + \int_t^T (L_x(r)\Psi(r, t) + L_\alpha(r))\, dr \;\middle|\; \mathcal{F}_t\right], \\ & \Phi_t = I + \int_\tau^t f_x(r)\Phi_r\, dr + \int_\tau^t \sigma_x(r)\Phi_r\, dw_r, \\ & \Psi(t, \tau) = \int_\tau^t (f_x(r)\Psi(r, \tau) + f_\alpha(r))\, dr + \int_\tau^t (\sigma_x(r)\Psi(r, \tau) + \sigma_\alpha(r))\, dw_r, \\ & \varphi_0 \geq 0, \quad \varphi_- \geq 0, \quad \langle \varphi_-, \Lambda_-[0, u] \rangle = 0.\end{aligned}$$

Here the pre-Hamiltonian $H(t, x, p, u, \varphi) = p' f(t, x, 0, u) - \varphi' L(t, x, 0, u)$, and similar notation prevails throughout these conditions. The set of all such triples is denoted by $M^{\varphi_0}(u)$. Define the mapping Δ on the space of multipliers as follows:

$$\Delta(\varphi, p, q) = \left(-\mathbf{E}p'_0 A - q'_0, \; \varphi_\mp\right).$$

As before, let Y be the set of state trajectories corresponding to solutions to $P(0,0)$.

For the extended problem P, Theorem 4.8 now remains valid exactly as written above, providing we interpret the multiplier sets M and mapping Δ as defined here. The same statements apply to Prop. 5.2 and its corollary, as well as the discussion of stability following. As for the stochastic maximum principle, it may be stated as follows in the general setting.

5.4 Theorem (SMP). *Assume (h1)–(h6). If problem $P(0,0)$ has a feasible control then it has a solution. Moreover, one has*

$$M^1(Y) \cup \big(M^0(Y) \setminus \{0\}\big) \neq \emptyset.$$

A Simple Application. Although the constraints upon this thesis make a detailed investigation of the implications of Thm. 4.8—or indeed of the methods used to prove it—impossible, we can note one of its very elementary consequences. Suppose a *deterministic* optimal control problem with constraints in the form of comparison functionals is given. Assume that the problem obeys conditions (h1)–(h6), appropriately rephrased, and that it has no abnormal multipliers in Pontryagin's sense. The general effects of an additional term $\alpha\, dw_t$ in the deterministic dynamics $dx_t = f(t, x_t, u_t)\, dt$, for some $n \times d$ matrix α and Brownian motion w in $\mathbf{R}^d$, can be studied as a special case of the problem above in which $\sigma(t, x, \alpha) = \alpha$, and x_0, f, ℓ, L are independent of α. In this case it is easy to see that the multipliers of Def. 5.3 are those for which $p_t = \mathbf{E}\big[\overline{p}_t \mid \mathcal{F}_t\big]$, where

$$-\dot{\overline{p}}(t) = H_x(t, x_t, p_t, u_t, \varphi), \quad p(T) = -\varphi' \ell_x(x_T),$$

$$\overline{q}(t) \equiv 0,$$

and the other conditions of Pontryagin's principle hold. For all such multipliers one has $\Delta(\varphi, p, q) = (0, \varphi_{\mp})$. Since $M^0(Y) = \{0\}$ by assumption, Prop. 5.2 implies that the first $n \times d$-component of every element in $\partial V(0,0)$ is zero. In other words, *to first order the introduction of white noise has no effect on the value of a normal deterministic control problem.*

The methods of this chapter seem certain to contain more profound results on the issue of stochastic approximations to deterministic control problems. This is a topic of considerable current interest—see Kushner (1965), Fleming (1971), and Fleming (1983) for example—to which we hope the current study will some day (soon) contribute.

References

Aubin, Jean-Pierre, and Ivar Ekeland (1984). *Applied Nonlinear Analysis.* Wiley-Interscience, New York.

Beneš, V. E. (1971). Existence of optimal stochastic control laws. *SIAM J. Control and Optimization* **9**, pp. 446–472.

Billingsley, Patrick (1968). *Convergence of Probability Measures.* John Wiley and Sons, New York.

Bismut, Jean-Michel (1976). *Théorie probabiliste du contrôle des diffusions.* Memoirs of the AMS, vol. 176.

Borwein, Jonathan B., and H. M. Strojwas (1984). Tangential approximations. *Nonlinear Analysis, Theory, Methods, and Applications,* to appear.

Borwein, Jonathan B., and H. M. Strojwas (1985). Proximal analysis and boundaries of closed sets in Banach spaces, Part I: Theory. Preprint.

Borwein, Jonathan B., and H. M. Strojwas (1985b). Proximal analysis and boundaries of closed sets in Banach spaces, Part II: Applications. Preprint.

Cesari, Lamberto (1983). *Optimization—Theory and Applications: Problems with Ordinary Differential Equations.* Springer-Verlag, New York.

Clarke, Frank H. (1983). *Optimization and Nonsmooth Analysis.* Wiley-Interscience, New York.

Clarke, Frank H., and Philip D. Loewen (1984). The value function in optimal control. *Proceedings of the 23rd IEEE Conference on Decision and Control.* Las Vegas, Nevada.

Clarke, Frank H., and Philip D. Loewen (to appear). The value function in optimal control: sensitivity, controllability, and time-optimality. *SIAM J. Control and Optimization,* to appear.

Dellacherie, Claude, and Paul-André Meyer (1975). *Probabilites et Potentiel.* Hermann, Paris.

Fleming, Wendell (1971). Stochastic control for small noise intensities. *SIAM J. Control and Optimization* **9**, pp. 473–517.

Fleming, Wendell (1983). Stochastic Calculus of Variations and Mechanics. *JOTA* **41**, pp. 55–74.

Fleming, Wendell, and Raymond Rishel (1975). *Deterministic and Stochastic Optimal Control.* Springer-Verlag, Berlin.

Franklin, Joel (1980). *Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed Point Theorems.* Springer-Verlag, New York.

Goor, Robert M. (1976). Existence of an optimal control for systems with jump Markov disturbances. *SIAM J. Control and Optimization* **14**, pp. 899–918. MR **53** #12722.

Goor, Robert M. (1979). Existence of an optimal control for stochastic systems governed by Itô equations. *JOTA* **27**, pp. 377–398. MR 80c:93133.

Halmos, Paul R. (1950). *Measure Theory.* Van-Nostrand Reinhold, New York.

Haussmann, Ulrich (1983). On the approximation of optimal stochastic controls. *JOTA* **40**, pp. 433–450.

Haussmann, Ulrich (1985). *A Stochastic Maximum Principle for Optimal Control of Diffusions.* Preprint, University of British Columbia Mathematics Department.

Haussmann, Ulrich (to appear). Existence of optimal Markovian controls for degenerate diffusions. Preprint, University of British Columbia Mathematics Department.

Ikeda, Nobuyuki, and Shinzo Watanabe (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

Jacod, Jean (1979). *Calcul Stochastique et Problèmes de Martingales.* Lecture Notes in Mathematics, vol. 714. Springer-Verlag, New York.

Kushner, Harold (1965). Near optimal control in the presence of small stochastic perturbations. *Journal of Basic Engineering (Trans. ASME)* **87**, pp. 103–108.

Kushner, Harold (1972). Necessary conditions for continuous parameter stochastic optimization problems. *SIAM J. Control and Optimization* **10**, pp. 550–565.

Kushner, Harold (1975). Existence results for optimal stochastic controls. *JOTA* **15**, pp. 347–359.

Loewen, Philip (1983). *The Sensitivity of Optimal Value Functions in Differential Inclusion Problems.* M.Sc. Thesis, Univ. of British Columbia, Canada.

Prokhorov, Yu. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its Applications* **1**, pp. 157–214.

Robinson, S. M. (1976). Stability theory for systems of inequalities, part II: differentiable nonlinear systems. *SIAM J. Numerical Analysis* **13**, pp. 130–145.

Rockafellar, R. T. (1982). Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming. *Mathematical Programming Study* **17**, pp. 28–66.

Royden, H. L. (1968). *Real Analysis.* MacMillan, New York.

Rudin, Walter (1973). *Functional Analysis.* McGraw-Hill, New York.

Skorokhod, A. V. (1965). *Studies in the Theory of Random Processes.* Addison-Wesley, Reading, Mass.

Treiman, Jay S. (1983). Characterization of Clarke's tangent and normal cones in finite and infinite dimensions. *Nonlinear Analysis* **7**, pp. 771–783.

Vinter, Richard B., and G. Pappas (1982). A maximum principle for non-smooth optimal control problems with state constraints. *J. Math. Anal. Appl.* **89**, pp. 212–232.

Warga, Jack (1972). *Optimal Control of Differential and Functional Equations.* Academic Press, New York.

Wong, Eugene (1971). Representation of martingales, quadratic variation and applications. *SIAM J. Control and Optimization* **9**, pp. 621–633.