2

Extremal Problems and Optimal Control

Summary. *This chapter is devoted to the study of various kinds of variational problems and of the mathematical tools needed to deal with them. We start with the so-called "direct method", that leads to a systematic study of lower semicontinuous functionals. Then we consider optimization problems with constraints using the Lagrange multipliers methods. This method is closely connected to nonlinear eigenvalue problems. We also develop a general duality theory for convex optimization problems and study saddle points, KKM-multimaps, coincidence theorems, variational inequalities and the Fenchel duality theory. Subsequently, we present the Ekeland variational principle and some of its most remarkable consequences. We also show that it is equivalent to a whole family of other important results of nonlinear analysis. With these mathematical tools, we then pass to the study of calculus of variations problems (Euler equation and canonical Hamiltonian equations) and of optimal control problems (existence theory, relaxation and Pontryagin's maximum principle).

Introduction

This chapter is devoted to the study of variational problems and of the mathematical tools that are necessary to do this.

So, in Section 2.1 we conduct a detailed investigation of the topological notion of lower semicontinuity. Since the work of Tonelli, the concept of lower semicontinuity has played a central role in the study of variational problems. Our study outlines the so-called direct method of the calculus of variations and also introduces the notion of relaxed functional which we encounter also in Section 2.6.

In Section 2.2 we deal with infinite-dimensional optimization problems with constraints and we analyze them by developing the so-called method of Lagrange multipliers. This method is closely connected to nonlinear eigenvalue problems.

Section 2.3 studies saddle points and develops a general duality theory for convex optimization problems. Saddle points appear in game theory and in general in control problems where the controllers exhibit conflicting interests. In connection with saddle points we also introduce and study KKM-maps which are useful in producing coincidence theorems for families of multifunctions and also lead to existence

N.S. Papageorgiou, S.Th. Kyritsi-Yiallourou, *Handbook of Applied Analysis*,
Advances in Mechanics and Mathematics 19, DOI 10.1007/b120946_2,

theorems for variational inequalities. Duality is at the core of convex analysis. In introductory functional analysis we encounter the first instances of duality with the separation theorems for convex sets, which was extended in Definition 1.2.15 with the introduction of the conjugate of a convex function. Here using a “perturbation” approach to a convex minimization problem, we associate a concave maximization one and investigate the precise relation between the two. As a special case, we present the so-called Fenchel duality theory.

In Section 2.4 first we present the Ekeland variational principle and discuss some remarkable consequences of it. Then we introduce some other fundamental results of nonlinear analysis, namely the Caristi fixed point theorem, the Takahashi variational principle, and the drop theorem, and we show that they are all equivalent to the Ekeland variational principle. Finally we prove a principle concerning ordered spaces, which we show generates all the aforementioned results.

In Section 2.5 we present some basic aspects of the calculus of variations. We consider problems with integral cost functional and fixed endpoints. We develop the Euler equation for an optimal state as well as the canonical Hamiltonian equations.

In Section 2.6 we deal with optimal control problems. We limit ourselves to systems monitored by ordinary differential equations (lumped parameter systems). First we develop the existence theory for such optimal control problems. This theory reveals the importance of convex structure in the problem. When this convexity is missing, we need to augment the system in order to assure existence of optimal pairs. This is the process of relaxation. We present four relaxation methods that we show are equivalent under reasonable conditions on the data. Finally we develop necessary conditions for optimality in the form of a maximum principle. Our approach is based on the Ekeland variational principle.

2.1 Lower Semicontinuity

In Section 1.2 we examined in some detail proper, lower semicontinuous, convex functions defined on a locally convex space. In this section we drop the convexity condition and concentrate on lower semicontinuous functions. The reason for this is that such functions are involved in the so-called direct method of the calculus of variations. Already in Section 1.5 we had some first results involving the notion of lower semicontinuity. Here we conduct a systematic study of this concept.

Our setting is purely topological. So let (X, τ) be a Hausdorff topological space, with τ denoting its topology. Additional conditions on X are introduced as needed. For $x \in X$, by $\mathcal{N}(x)$ we denote the filter of neighborhoods of x. Also recall that $\mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$. We start by recalling the definition of τ-lower semicontinuity (both local and global) of a function $\varphi\colon X \longrightarrow \mathbb{R}^*$.

DEFINITION 2.1.1 A function $\varphi\colon X \longrightarrow \mathbb{R}^*$ is said to be *τ-lower semicontinuous at x*, if

$$\varphi(x) = \liminf_{u \to x} \varphi(u) = \sup_{U \in \mathcal{N}(x)} \inf_{u \in U} \varphi(u).$$

If φ is τ-lower semicontinuous at every $x \in X$, then we say that φ is τ-lower semicontinuous. If $-\varphi$ is τ-lower semicontinuous (at x), then we say that φ is *τ-upper semicontinuous* (at x).

REMARK 2.1.2 If the topology τ is clearly understood, then we drop τ from the definition and say that φ is lower semicontinuous (at $x \in X$). Also if $x \in X$ is a point where $\varphi(x) > -\infty$, then φ is τ-lower semicontinuous at x if and only if for every $\lambda \in \mathbb{R}$ with $\lambda < \varphi(x)$, we can find $U \in \mathcal{N}(x)$ such that $\lambda < \varphi(u)$ for all $u \in U$.

PROPOSITION 2.1.3 *For a function $\varphi\colon X \longrightarrow \mathbb{R}^*$ the following statements are equivalent.*

(a) *φ is τ-lower semicontinuous.*
(b) *For every $\lambda \in \mathbb{R}$, the set $L_\lambda = \{x \in X : \varphi(x) \leq \lambda\}$ is τ-closed.*
(c) epi $\varphi = \{(x,\lambda) \in X \times \mathbb{R} : \varphi \leq \lambda\}$ *is closed in $X \times \mathbb{R}$.*

PROOF: (a)$\Rightarrow$(b) We show that L_λ^c is open. Let $x \in L_\lambda$. Then $\lambda < \varphi(x)$ and by virtue of Remark 2.1.2 we can find $U \in \mathcal{N}(x)$ such that $\lambda < \varphi(u)$ for all $u \in U$, which shows that L_λ^c is open; hence L_λ is closed.

(b)$\Rightarrow$(a) If $\varphi(x) = -\infty$, then it is clear from Definition 2.1.1 that φ is τ-lower semicontinuous at $x \in X$. So suppose that $-\infty < \varphi(x)$ and let $\lambda < \varphi(x)$. Note that $x \in L_\lambda^c$ and because by hypothesis L_λ is closed, we can find $U \in \mathcal{N}(x)$ such that $U \subseteq L_\lambda^c$, hence $\lambda < \varphi(u)$ for all $u \in U$, which proves the τ-lower semicontinuity of φ at $x \in X$.

So we have established the equivalence of statements (a) and (b).

Next let $h\colon X \times \mathbb{R} \longrightarrow \mathbb{R}^*$ be defined by $h(x,\lambda) = \varphi(x) - \lambda$. Clearly φ is τ-lower semicontinuous on X if and only if h is lower semicontinuous on the product space $X \times \mathbb{R}$. Because epi $\varphi = \{(x,\lambda) \in X \times \mathbb{R} : h(x,\lambda) \leq 0\}$, the equivalence of (a) and (b) established above implies the equivalence of (b) and (c). $\square$

COROLLARY 2.1.4 *If $\varphi_i\colon X \longrightarrow \mathbb{R}^*$, $i \in I$, is a family of τ-lower semicontinuous functions, then*

(a) *$\sup\limits_{i \in I} \varphi_i$ is τ-lower semicontinuous.*
(b) *If I is finite, then $\inf\limits_{i \in I} \varphi_i$ is τ-lower semicontinuous.*

COROLLARY 2.1.5 *A set $C \subseteq X$ is τ-closed if and only if i_C is τ-lower semicontinuous.*

From Definition 2.1.1 we see that $\varphi\colon X \longrightarrow \mathbb{R}^*$ is τ-lower semicontinuous at $x \in X$; then for every $x_n \longrightarrow x$ in X, we have $\varphi(x) \leq \liminf\limits_{n \to \infty} \varphi(x_n)$. The converse is not in general true.

EXAMPLE 2.1.6 Let X be an infinite-dimensional Banach space furnished with the weak topology denoted by w. Also let $C \subseteq X$ be a nonempty set which is sequentially w-closed but not w-closed. Then the indicator function i_C satisfies $i_C(x) \leq \liminf\limits_{n \to \infty} i_C(x_n)$ for every sequence $x_n \xrightarrow{w} x$ in X, but it is not w-lower semicontinuous. For example in $X = l^1$, due to the Schur property the set $C = \partial B_1 = \{x \in l^1 : \|x\|_{l^1} = 1\}$ is sequentially w-closed but it is not w-closed (in fact, $\overline{C}^{\,w} = \overline{B}_1 = \{x \in l^1 : \|x\|_{l^1} \leq 1\}$).

DEFINITION 2.1.7 We say that $\varphi\colon X \longrightarrow \mathbb{R}^*$ is *sequentially τ-lower semicontinuous at $x \in X$*, if $\varphi(x) \leq \liminf_{n\to\infty} \varphi(x_n)$ for every sequence $x_n \longrightarrow x$ in X. We say that φ is *sequentially τ-lower semicontinuous*, if it is sequentially τ-lower semicontinuous at every $x \in X$.

REMARK 2.1.8 Let τ_{seq} denote the topology of X whose sets are sequentially τ-closed; then $\varphi\colon X \longrightarrow \mathbb{R}^*$ is sequentially τ-lower semicontinuous if and only if it is τ_{seq}-lower semicontinuous. In general the τ_{seq} topology is stronger than the τ-topology and so the notion of sequential τ-lower semicontinuous is more general than the τ-lower semicontinuity. Clearly $\tau = \tau_{\text{seq}}$ if and only if X is first countable. So we can state the following proposition.

PROPOSITION 2.1.9 *If (X, τ) is first countable, $\varphi\colon X \longrightarrow \mathbb{R}^*$ and $x \in X$, then φ is τ-lower semicontinuous at $x \in X$ if and only if φ is sequentially τ-lower semicontinuous at $x \in X$.*

The next theorem summarizes the direct method of the calculus of variations and underlines the importance of the notion of lower semicontinuity in variational analysis. The result is known as the *Weierstrass theorem*.

THEOREM 2.1.10 *If $\varphi\colon X \longrightarrow \mathbb{R}^*$ is τ-coercive and τ-lower semicontinuous (resp., sequentially τ-coercive and sequentially τ-lower semicontinuous), then*

(a) *There exists $x \in X$ such that $\varphi(x) = \inf\limits_X \varphi$.*

(b) *If $\{x_n\}_{n\geq 1}$ is a minimizing sequence and x is a cluster point of $\{x_n\}_{n\geq 1}$ (resp., x is a subsequential limit of $\{x_n\}_{n\geq 1}$), then $\varphi(x) = \inf\limits_X \varphi$.*

(c) *If φ is not identically $+\infty$, then every minimizing sequence has a cluster point (resp., a convergent subsequence).*

PROOF: If $\varphi = +\infty$, then all $x \in X$ minimize φ and so (a) and (b) hold.

So assume that φ is not identically $+\infty$. Let $\{x_n\}_{n\geq 1}$ be a minimizing sequence for φ. We have $\varphi(x_n) \downarrow \inf\limits_X \varphi < +\infty$. Let $\lambda \in \mathbb{R}$ be such that $\inf\limits_X \varphi < \lambda$. Then we can find $n_0 \geq 1$ such that $\varphi(x_n) \leq \lambda$ for all $n \geq n_0$. Due to the τ-coercivity (resp., sequential τ-coercivity) of φ, we have that $\{x_n\}_{n\geq 1}$ has a cluster point $x \in X$ (resp., a subsequential limit point $x \in X$); see Definition 1.5.26. Then due to the τ-lower semicontinuity (resp. sequential τ-lower semicontinuity) of φ, we have

$$
\begin{aligned}
& \inf_X \varphi \leq \varphi(x) \leq \lim_{n\to\infty} \varphi(x_n) = \inf_X \varphi, \\
\Rightarrow \quad & \varphi(x) = \inf_X \varphi.
\end{aligned}
$$

□

COROLLARY 2.1.11 *If X is a reflexive Banach space, $\varphi \in \Gamma_0(X)$, and $\varphi(x) \longrightarrow +\infty$ as $\|x\| \longrightarrow \infty$, then we can find $x \in X$ such that $-\infty < \varphi(x) = \inf\limits_X \varphi$.*

PROOF: Note that due to the convexity of φ, the function is lower semicontinuous if and only if it is w-lower semicontinuous. Moreover, due to the reflexivity of X bounded sets are relatively w-compact (in fact relatively sequentially w-compact by the Eberlein–Smulian theorem). □

REMARK 2.1.12 If $\varphi\colon X \longrightarrow \mathbb{R}^*$ is lower semicontinuous convex and there exists $x_0 \in X$ such that $\varphi(x_0) = -\infty$, then φ is nowhere finite on X. For this reason, when convexity is present, we consider functions with values in $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

DEFINITION 2.1.13 Let $\varphi\colon X \longrightarrow \mathbb{R}^*$. The *$\tau$-lower semicontinuous regularization* (or the *τ-lower semicontinuous envelope*) of φ, is the function $\overline{\varphi}^{\,\tau}\colon X \longrightarrow \mathbb{R}^*$ defined by

$$\overline{\varphi}^{\,\tau}(x) = \liminf_{u \to x} \varphi(u) = \sup_{U \in \mathcal{N}(x)} \inf_{u \in U} \varphi(u).$$

REMARK 2.1.14 It is clear from this definition that $\overline{\varphi}^{\,\tau}$ is τ-lower semicontinuous.

In the next proposition we show that $\overline{\varphi}^{\,\tau}$ is the biggest τ-lower semicontinuous function majorized by φ.

PROPOSITION 2.1.15 *If $\varphi\colon X \longrightarrow \mathbb{R}^*$ and $\mathcal{S}(\varphi) = \big\{h : X \longrightarrow \mathbb{R}^* :$ h is τ-lower semicontinuous, $h \le \varphi\big\}$, then* $\operatorname{epi} \overline{\varphi}^{\,\tau} = \overline{\operatorname{epi} \varphi}$ *and* $\overline{\varphi}^{\,\tau}(x) = \sup\big[h(x) : h \in \mathcal{S}(\varphi)\big]$ *for all $x \in X$.*

PROOF: Evidently $\overline{\operatorname{epi} \varphi} \subseteq \operatorname{epi} \overline{\varphi}^{\,\tau}$. Let $(x,\lambda) \in \operatorname{epi} \overline{\varphi}^{\,\tau}$. Then $\overline{\varphi}^{\,\tau}(x) \le \lambda$. So for every $U \in \mathcal{N}(x)$ and for every $\varepsilon > 0$ we have $\inf_{u \in U} \varphi(u) \le \overline{\varphi}^{\,\tau}(x) \le \lambda < \lambda + \varepsilon$. So we can find $u \in U$ and $\mu \in (\lambda, \lambda+\varepsilon)$ such that $\varphi(u) < \mu$. Hence $\big(U \times (\lambda-\varepsilon, \lambda+\varepsilon)\big) \cap \operatorname{epi} \varphi \ne \emptyset$, from which it follows that $(x,\lambda) \in \overline{\operatorname{epi} \varphi}$. Therefore $\operatorname{epi} \overline{\varphi}^{\,\tau} = \overline{\operatorname{epi} \varphi}$.

Now note that $\overline{\varphi}^{\,\tau} \in \mathcal{S}(\varphi)$. Also if $h \in \mathcal{S}(\varphi)$, then $\operatorname{epi} \varphi \subseteq \operatorname{epi} h$ and so $\operatorname{epi} \overline{\varphi}^{\,\tau} = \overline{\operatorname{epi} \varphi} \subseteq \overline{\operatorname{epi} h} = \operatorname{epi} h$ (see Proposition 2.1.3). Therefore $h \le \overline{\varphi}^{\,\tau}$ and so we conclude that

$$\overline{\varphi}^{\,\tau}(x) = \sup\big[h(x) : h \in \mathcal{S}(\varphi)\big] \qquad \text{for all } x \in X.$$

$\square$

COROLLARY 2.1.16 *If $\varphi\colon X \longrightarrow \mathbb{R}^*$ and $\lambda \in \mathbb{R}$, then $\big\{\overline{\varphi}^{\,\tau} \le \lambda\big\} = \bigcap_{\mu > \lambda} \overline{\{\varphi \le \mu\}}^{\,\tau}$.*

COROLLARY 2.1.17 *If $C \subseteq X$, then $\overline{i_C}^{\,\tau} = i_{\overline{C}^{\,\tau}}$.*

COROLLARY 2.1.18 *If $\varphi, \psi\colon X \longrightarrow \mathbb{R}^*$, then*

(a) $\overline{(\varphi + \psi)}^{\,\tau} \ge \overline{\varphi}^{\,\tau} + \overline{\psi}^{\,\tau}$.

(b) *If ψ is continuous and finite everywhere, then $\overline{(\varphi + \psi)}^{\,\tau} = \overline{\varphi}^{\,\tau} + \psi$.*

The following proposition characterizes $\overline{\varphi}^{\,\tau}$ in terms of sequences and follows from Proposition 1.5.20 (see also Remark 1.5.2).

PROPOSITION 2.1.19 *If (X,τ) is first countable, $\varphi\colon X \longrightarrow \mathbb{R}^*$ and $x \in X$, then $\overline{\varphi}^{\,\tau}$.*

(a) *For every sequence $x_n \longrightarrow x$ in X we have $\overline{\varphi}^{\,\tau}(x) \le \liminf_{n \to \infty} \varphi(x_n)$.*

(b) *There exists a sequence* $x_n \longrightarrow x$ *in* X *such that* $\overline{\varphi}^{\,\tau}(x) = \lim_{n\to\infty} \varphi(x_n)$.

The next theorem summarizes the relaxation method in the study of variational problems.

THEOREM 2.1.20 *If* $\varphi\colon X \longrightarrow \mathbb{R}^*$ *is* τ*-coercive, then*

(a) $\overline{\varphi}^{\,\tau}$ *is* τ*-coercive and* τ*-lower semicontinuous.*
(b) *There exists* $x \in X$ *such that* $\overline{\varphi}^{\,\tau}(x) = \inf_X \overline{\varphi}^{\,\tau}$.
(c) $\min_X \overline{\varphi}^{\,\tau} = \inf_X \varphi$.
(d) *Every cluster point of a minimizing sequence for* φ *is a minimizer of* $\overline{\varphi}^{\,\tau}$.
(e) *If* (X,τ) *is first countable,*
then the minimizers of $\overline{\varphi}^{\,\tau}$ *are the limits of minimizing sequences for* φ.

PROOF: (a) We already know that $\overline{\varphi}^{\,\tau}$ is τ-lower semicontinuous. The τ-coercivity of $\overline{\varphi}^{\,\tau}$ follows from Theorem 1.5.29 and Remark 1.5.2 (see also Proposition 2.1.19).

(b) Follows from part (a) and Theorem 2.1.10.

(c) Let $h(x) = \inf_X \varphi$ for all $x \in X$. Then $h \in \mathcal{S}(\varphi)$ and so $h \leq \overline{\varphi}^{\,\tau}$ (see Proposition 2.1.15). It follows that

$$\inf_X \varphi \leq \min_X \overline{\varphi}^{\,\tau}.$$

Because the opposite inequality is always true (recall that $\overline{\varphi}^{\,\tau} \leq \varphi$), we conclude that

$$\min_X \overline{\varphi}^{\,\tau} = \inf_X \varphi.$$

(d) Let $\{x_n\}_{n\geq 1}$ be a minimizing sequence for φ and x a cluster point of $\{x_n\}_{n\geq 1}$. Then

$$\overline{\varphi}^{\,\tau}(x) \leq \limsup_{n\to\infty} \overline{\varphi}^{\,\tau}(x_n) \leq \lim_{n\to\infty} \varphi(x_n) = \inf_X \varphi = \min_X \overline{\varphi}^{\,\tau}$$

(see part (c)),

$$\Rightarrow \overline{\varphi}^{\,\tau}(x) = \min_X \overline{\varphi}^{\,\tau}.$$

(e) Let $x \in X$ be such that $\overline{\varphi}^{\,\tau}(x) = \min_X \overline{\varphi}^{\,\tau}$. Then from Proposition 2.1.19 and (c), we can find $x_n \longrightarrow x$ such that

$$\inf_X \varphi = \overline{\varphi}^{\,\tau}(x) = \lim_{n\to\infty} \varphi(x_n).$$

□

REMARK 2.1.21 This theorem shows that the relaxed functional captures the asymptotic behavior of the minimizing sequences for φ.

For convex functions in a Banach space, the relaxation process is simpler because by Mazur's lemma for convex sets, the closures in the norm and weak topologies coincide. So we have the following.

PROPOSITION 2.1.22 *If X is a Banach space and $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is convex, then $\overline{\varphi}^{\,s} = \overline{\varphi}^{\,w}$ (here by s we denote the strong topology on X and by w the weak topology on X).*

Next we prove a lower semicontinuity result for integral functionals defined on products of Lebesgue spaces. The mathematical setting is the following. Let (Ω, Σ, μ) be a complete, finite nonatomic measure space and $\varphi\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^m \longrightarrow \overline{\mathbb{R}}$ a function that satisfies the following hypotheses.

$\mathbf{H_0}$:

(i) φ is $\Sigma \times \mathbf{B}(\mathbb{R}^N) \times \mathbf{B}(\mathbb{R}^m)$-measurable with $\mathbf{B}(\mathbb{R}^N)$ (resp., $\mathbf{B}(\mathbb{R}^m)$) being the Borel σ-field of $\mathbb{R}^N$ (resp., of $\mathbb{R}^m$).

(ii) For μ-almost all $\omega \in \Omega$, $(x,u) \longrightarrow \varphi(\omega, x, u)$ is lower semicontinuous.

(iii) For μ-almost all $\omega \in \Omega$ and all $x \in \mathbb{R}^N$, $u \longrightarrow \varphi(\omega, x, u)$ is convex.

Let $1 \leq p, r \leq +\infty$ and consider the integral function $I_\varphi(x,u) : L^p(\Omega, \mathbb{R}^N) \times L^r(\Omega, \mathbb{R}^m) \longrightarrow \mathbb{R}^*$ defined by

$$I_\varphi(x,u) = \int_\Omega \varphi\big(\omega, x(\omega), u(\omega)\big)\, d\mu \qquad \text{for all } (x,u) \in L^p(\Omega, \mathbb{R}^N) \times L^r(\Omega, \mathbb{R}^m).$$

We want to establish a sequential lower semicontinuity result when we consider the strong topology on $L^p(\Omega, \mathbb{R}^N)$ and the weak topology on $L^r(\Omega, \mathbb{R}^m)$ (the weak* topology on $L^\infty(\Omega, \mathbb{R}^m)$ when $r = +\infty$). To do this we need the following approximation result, whose proof can be found in Buttazzo [126, p. 42].

PROPOSITION 2.1.23 *If $\varphi\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^m \longrightarrow \overline{\mathbb{R}}$ satisfies H_0 above and one of the following conditions.*

(iv) *For μ-almost all $\omega \in \Omega$, we can find a continuous function $u_0\colon \mathbb{R}^N \longrightarrow \mathbb{R}^m$ such that the function $x \longrightarrow \varphi\big(\omega, x, u_0(x)\big)$ is finite and continuous; or*

(v) *There exists a function $\vartheta\colon \mathbb{R} \longrightarrow \mathbb{R}$ such that*

$$\lim_{s \to +\infty} \frac{\vartheta(s)}{s} = +\infty$$

and $\vartheta(\|u\|_{\mathbb{R}^m}) \leq \varphi(\omega, x, u)$ for μ-a.a. $\omega \in \Omega$, all $x \in \mathbb{R}^N$ and all $u \in \mathbb{R}^m$,

then we can find two sequences of Carathéodory functions $a_n\colon \Omega \times \mathbb{R}^N \longrightarrow \mathbb{R}^m$ and $c_n\colon \Omega \times \mathbb{R}^N \longrightarrow \mathbb{R}$ such that

$$\varphi(\omega, x, u) = \sup_{n \geq 1} \big[\big(a_n(\omega, x), u\big)_{\mathbb{R}^m}) + c_n(\omega, x)\big]$$

for μ-almost all $\omega \in \Omega$ and all $x \in \mathbb{R}^N$, all $u \in \mathbb{R}^m$.

REMARK 2.1.24 Recall that $h\colon \Omega \times \mathbb{R}^k \longrightarrow \mathbb{R}^s$ $(1 \leq s, k)$ is a Carathéodory function, if for all $x \in \mathbb{R}^k$, $\omega \longrightarrow h(\omega, x)$ is Σ-measurable and for μ-almost all $\omega \in \Omega$, $x \longrightarrow h(\omega, x)$ is continuous. Such a function is automatically $\Sigma \times \mathbf{B}(\mathbb{R}^k)$-measurable. If in the above proposition $\varphi \geq 0$, then it is easy to see the following approximation for φ holds

$$\varphi(\omega, x, u) = \sup_{n \geq 1} \big[\big(a_n(\omega, x), u\big)_{\mathbb{R}^m} + c_n(\omega, x)\big]^+$$

with $a_n\colon \Omega \times \mathbb{R}^N \longrightarrow \mathbb{R}^m$ and $c_n\colon \Omega \times \mathbb{R}^N \longrightarrow \mathbb{R}$ bounded Carathéodory functions.

We can prove the following sequential lower semicontinuity result for the integral functional $I_\varphi(x,u), (x,u) \in L^p(\Omega,\mathbb{R}^N) \times L^r(\Omega,\mathbb{R}^m)$.

THEOREM 2.1.25 *If $\varphi\colon \Omega\times\mathbb{R}^N\times\mathbb{R}^m \longrightarrow \overline{\mathbb{R}}$ satisfies H_0 and the following holds,*

"for every sequence $\{x_n\}_{n\geq 1}$ converging in $L^p(\Omega,\mathbb{R}^N)$ and every sequence $\{u_n\}_{n\geq 1}$ weakly converging in $L^r(\Omega,\mathbb{R}^m)$ (weakly* if $r=+\infty$) such that

$$\int_\Omega \varphi^+\big(\omega,x_n(\omega),u_n(\omega)\big)d\mu \leq c + \int_\Omega \varphi^-\big(\omega,x_n(\omega),u_n(\omega)\big)d\mu,$$

for some $c>0$ and all $n\geq 1$, the sequence $\Big\{\varphi^-\big(\cdot,x_n(\cdot),u_n(\cdot)\big)\Big\}_{n\geq 1} \subseteq L^1(\Omega)$ is relatively weakly compact,"

then the integral function I_φ is well-defined on $L^p(\Omega,\mathbb{R}^N) \times L^r(\Omega,\mathbb{R}^m)$ with values in $\overline{\mathbb{R}}$ and it is sequentially lower semicontinuous when we consider $L^p(\Omega,\mathbb{R}^N)$ with the norm topology and $L^r(\Omega,\mathbb{R}^m)$ with the weak topology (weak* if $r=+\infty$).

PROOF: Let $x \in L^p(\Omega,\mathbb{R}^N),\ u \in L^r(\Omega,\mathbb{R}^m)$ and suppose that for some $c>0$ we have

$$\int_\Omega \varphi^+\big(\omega,x(\omega),u(\omega)\big)d\mu \leq c + \int_\Omega \varphi^-\big(\omega,x(\omega),u(\omega)\big)d\mu. \tag{2.1}$$

Then by hypothesis $\varphi^-\big(\cdot,x(\cdot),u(\cdot)\big) \subseteq L^1(\Omega)$ and so using (2.1) we conclude that $\varphi\big(\cdot,x(\cdot),u(\cdot)\big) \subseteq L^1(\Omega)$. If (2.1) is not satisfied, then

$$\int_\Omega \varphi^-\big(\omega,x(\omega),u(\omega)\big)d\mu < +\infty \qquad \text{and} \qquad \int_\Omega \varphi^+\big(\omega,x(\omega),u(\omega)\big)d\mu = +\infty,$$
$$\Rightarrow \int_\Omega \varphi\big(\omega,x(\omega),u(\omega)\big)d\mu = +\infty.$$

So indeed I_φ is well-defined with values in $\overline{\mathbb{R}} = \mathbb{R}\cup\{+\infty\}$.

We prove the desired sequential lower semicontinuity of I_φ, in steps.

Step 1: We assume that there exists a function $\vartheta\colon \mathbb{R}\longrightarrow\mathbb{R}$ such that

$$\lim_{s\to+\infty} \frac{\vartheta(s)}{s} = +\infty \qquad \text{and} \qquad \vartheta(\|u\|_{\mathbb{R}^m}) \leq \varphi(\omega,x,u) \tag{2.2}$$

for μ-a.a. $\omega \in \Omega$, all $x \in \mathbb{R}^N$, and all $u \in \mathbb{R}^m$.

Because of (2.2) we can apply Proposition 2.1.23 (see also Remark 2.1.24) and obtain that

$$\varphi(\omega,x,u) = \sup_{n\geq 1} \Big[\big(a_n(\omega,x),u\big)_{\mathbb{R}^N} + c_n(\omega,x)\Big]^+ \tag{2.3}$$

with $a_n\colon \Omega\times\mathbb{R}^N \longrightarrow \mathbb{R}^m$ and $c_n\colon \Omega\times\mathbb{R}^N \longrightarrow \mathbb{R},\ n\geq 1$, bounded Carathéodory functions. Recall that

$$\begin{aligned}&\int_\Omega \big[\big(a_n(\omega,x),u\big)_{\mathbb{R}^m} + c_n(\omega,x)\big]^+ d\mu \\ = \quad & \sup\Big[\int_C \big[\big(a_n(\omega,x),u\big)_{\mathbb{R}^m} + c_n(\omega,x)\big]d\mu : C \in \Sigma\Big].\end{aligned} \tag{2.4}$$

So in view of (2.3) and (2.4) it suffices to show that for all $n\geq 1$ and all $C \in \Sigma$

$$(x,u) \longrightarrow \int_C \left[\big(a_n(\omega,x(\omega)),u(\omega)\big)_{\mathbb{R}^N} + c_n\big(\omega,x(\omega)\big)\right]d\mu$$

is sequentially lower semicontinuous on $L^p(\Omega,\mathbb{R}^N)\times L^r(\Omega,\mathbb{R}^m)$ the first space in the product furnished with the norm topology and the second with the weak topology (weak* topology if $r=+\infty$). But this is immediate because a_n and c_n are bounded Carathéodory functions.

Step 2: Now we assume that φ is bounded below.
Without any loss of generality, we may assume that $\varphi \geq 0$. Suppose that $x_n \longrightarrow x$ in $L^p(\Omega,\mathbb{R}^N)$ and $u_n \xrightarrow{w} u$ in $L^r(\Omega,\mathbb{R}^m)$ (weakly* if $r=+\infty$). Because $u_n \xrightarrow{w} u$ in $L^r(\Omega,\mathbb{R}^m)$ (weakly* if $r=+\infty$) and μ is finite, we have that $u_n \xrightarrow{w} u$ in $L^1(\Omega,\mathbb{R}^m)$ and so by the Dunford–Pettis theorem the sequence $\{u_n\}_{n\geq 1}\subseteq L^1(\Omega,\mathbb{R}^m)$ is uniformly integrable. So according to the De La Vallée–Poussin theorem we can find $\vartheta\colon \mathbb{R}\longrightarrow\mathbb{R}_+$ such that

$$\lim_{s\to+\infty}\frac{\vartheta(s)}{s}=+\infty \qquad \text{and} \qquad \int_\Omega \vartheta(\|u_n(\omega)\|_{\mathbb{R}^m})d\mu \leq 1 \qquad \text{for all } n\geq 1.$$

Let $\varepsilon>0$ and set

$$\varphi_\varepsilon(\omega,x,u)=\varphi(\omega,x,u)+\varepsilon\vartheta(\|u\|_{\mathbb{R}^m}).$$

Because of Step 1, we have

$$\begin{aligned}
&\int_\Omega \varphi\big(\omega,x(\omega),u(\omega)\big)d\mu \leq \int_\Omega \varphi_\varepsilon\big(\omega,x(\omega),u(\omega)\big)d\mu\\
&\leq \liminf_{n\to\infty}\int_\Omega \varphi_\varepsilon\big(\omega,x_n(\omega),u_n(\omega)\big)d\mu\\
&\leq \liminf_{n\to\infty}\int_\Omega \varphi\big(\omega,x_n(\omega),u_n(\omega)\big)d\mu+\varepsilon.
\end{aligned}$$

Because $\varepsilon>0$ was arbitrary, we let $\varepsilon\downarrow 0$ to conclude that

$$\int_\Omega \varphi\big(\omega,x(\omega),u(\omega)\big)d\mu \leq \liminf_{n\to\infty}\int_\Omega \varphi\big(\omega,x_n(\omega),u_n(\omega)\big)d\mu.$$

Step 3: General case.
Again suppose that $x_n\longrightarrow x$ in $L^p(\Omega,\mathbb{R}^N)$ and $u_n\xrightarrow{w} u$ in $L^r(\Omega,\mathbb{R}^m)$ (weakly* if $r=+\infty$). For every $k\geq 1$ we define

$$\varphi_k(\omega,x,u)=\max\{\varphi(\omega,x,u),-k\},\quad g_n(\omega)=\varphi^-\big(\omega,x_n(\omega),u_n(\omega)\big)$$

and $C_{n,k}=\{\omega\in\Omega : g_n(\omega)>k\}$.

Without any loss of generality we may assume that $I_\varphi(x_n,u_n)$ tends to a finite limit as $n\to\infty$. So by virtue of the hypothesis of the theorem $\{g_n\}_{n\geq 1}\subseteq L^1(\Omega)$ is relatively weakly compact. Then for every $k\geq 1$, because of Step 2, we have

$$\begin{aligned}
I_\varphi(x,u)\leq I_{\varphi_k}(x,u) &\leq \liminf_{n\to\infty} I_{\varphi_k}(x_n,u_n)\\
&=\liminf_{n\to\infty}\left[I_\varphi(x_n,u_n)+\int_{C_{n,k}}\big(g_n(\omega)-k\big)\mathrm{d}\mu\right]\\
&\leq \liminf_{n\to\infty} I_\varphi(x_n,u_n)+\limsup_{n\to\infty}\int_{C_{n,k}}\big(g_n(\omega)-k\big)\mathrm{d}\mu.
\end{aligned}$$

The Dunford–Pettis theorem implies that $\{g_n\}_{\{x_n\}_{n\geq 1}\geq 1} \subseteq L^1(\Omega)$ is uniformly integrable. So if we pass to the limit as $k \longrightarrow \infty$, we conclude that

$$I_\varphi(x, u) \leq \liminf_{n\to\infty} I_\varphi(x_n, u_n).$$

$\square$

COROLLARY 2.1.26 *If $\Omega \subseteq \mathbb{R}^k$ is a bounded domain with Lipschitz boundary and $\varphi\colon \Omega\times\mathbb{R}^N \times\mathbb{R}^{Nk} \longrightarrow \mathbb{R}_+$ satisfies*

(i) *φ is Borel measurable.*
(ii) *For almost all $z \in \Omega$, $(x, u) \longrightarrow \varphi(z, x, u)$ is lower semicontinuous.*
(iii) *For almost all $z \in \Omega$ and all $x \in \mathbb{R}^N$, $u \longrightarrow \varphi(z, x, u)$ is convex,*

then the integral functional $x \longrightarrow J_\varphi(x) = \int_\Omega \varphi\big(z, x(z), Dx(z)\big)\mathrm{d}z$ is sequentially w-lower semicontinuous on $W^{1,1}_{loc}(\Omega, \mathbb{R}^N)$.

REMARK 2.1.27 This result is optimal when $N = 1$ (the so-called scalar case), in the sense that the convexity of $\varphi(z, x, \cdot)$ is a necessary condition for the sequential weak lower semicontinuity of $J_\varphi(\cdot)$ on $W^{1,p}(\Omega)$, $1 \leq p \leq \infty$. In the case $N > 1$ (the so-called vector case), convex integrands produce only a small subclass of the sequentially weak lower semicontinuous functionals on $W^{1,p}(\Omega, \mathbb{R}^N)$ (see Section 2.5).

We can have a version of Theorem 2.1.25 for integrands defined on infinite-dimensional Banach spaces. The result is due to Balder [53] and a different proof based on Young measures can be found in Hu–Papageorgiou [316, p. 31].

THEOREM 2.1.28 *If X is a separable Banach space, Y is a separable reflexive Banach space, and $\varphi\colon \Omega\times X\times Y \longrightarrow \overline{\mathbb{R}}$ satisfies*

(i) *φ is $\Sigma \times \mathbf{B}(X) \times \mathbf{B}(Y)$ measurable $\big($with $\mathbf{B}(X)$ (resp., $\mathbf{B}(Y)$) the Borel σ-field of X (resp., Y)$\big)$;*
(ii) *For μ-almost all $\omega \in \Omega$, $(x, y) \longrightarrow \varphi(\omega, x, y)$ is lower semicontinuous;*
(iii) *For μ-almost all $\omega \in \Omega$ and all $x\in X$, $y \longrightarrow \varphi(\omega, x, y)$ is convex;*
(iv) *There exist $M > 0$ and $\vartheta \in L^1(\Omega)$ such that for $\mu-$almost all $\omega \in \Omega$ and all $(x, y) \in X \times Y$ we have*

$$\varphi(\omega, x, y) \geq \vartheta(\omega) - M\big(\|x\|_X + \|y\|_Y\big),$$

then $(x, u) \longrightarrow I_\varphi(x, u) = \int_\Omega \varphi\big(\omega, x(\omega), u(\omega)\big)\mathrm{d}\mu$ is sequentially lower semicontinuous on $L^1(\Omega, X) \times L^1(\Omega, Y)_w$, where by $L^1(\Omega, Y)_w$ we denote the Lebesgue–Bochner space $L^1(\Omega, Y)$ furnished with the weak topology.

2.2 Constrained Minimization Problems

In many applied situations we are not just looking for the minimum of an objective functional on an open set U, but we want to determine the minimum of φ subject to certain restrictions on the points $x \in U$. In this section we examine such minimization problems with side conditions and develop the method of Lagrange multipliers.

DEFINITION 2.2.1 Let X, Y be Banach spaces, $U \subseteq X$ and $V \subseteq Y$ nonempty open sets and $\varphi\colon U \longrightarrow V$ a Fréchet differentiable map.

(a) We say that $x_0 \in U$ is a *critical point* of φ if and only if $\varphi'(x_0) \in \mathcal{L}(X, Y)$ is not surjective.

(b) We say that $x_0 \in U$ is a *regular point* of φ if and only if $\varphi'(x_0) \in \mathcal{L}(X, Y)$ is surjective.

REMARK 2.2.2 If $Y = \mathbb{R}$, then $x_0 \in U$ is a critical (resp., regular) point of the function $\varphi\colon U \longrightarrow \mathbb{R}$ only if $\varphi'(x_0) = 0$ (resp., $\varphi'(x_0) \neq 0$). If $X = \mathbb{R}$, then $x_0 \in (a, b)$ is a critical point of φ if and only if $\varphi'(x_0) = 0$.

EXAMPLE 2.2.3 Let $X = H =$ a Hilbert space, $A \in \mathcal{L}(H, H)$ is a self-adjoint isomorphism (i.e., $A^{-1} \in \mathcal{L}(H, H)$) and $\psi\colon H \longrightarrow \mathbb{R}$ is Fréchet differentiable. Let $\varphi\colon H \longrightarrow \mathbb{R}$ be defined by

$$\varphi(x) = \frac{1}{2} \langle Ax, x \rangle_H - \psi(x),$$

where by $\langle \cdot, \cdot \rangle_H$ we denote the inner product of H. Evidently φ is Fréchet differentiable and

$$\varphi'(x) = Ax - \psi'(x).$$

Therefore $x_0 \in H$ is a critical point of φ if and only if $x_0 = A^{-1}\psi'(x_0)$. If $\psi(x) = \lambda/2\|x\|_H^2$, $\lambda \in \mathbb{R}$, then $\psi'(x_0) = \lambda x_0$ and so $x_0 \in H$ is a critical point of the function $\varphi(x) = 1/2 \langle Ax, x \rangle_H - \lambda/2\|x\|_H^2$ if and only if $Ax_0 = \lambda x_0$; that is, x_0 is an eigenvector of the operator A with eigenvalue λ.

Next we extend Definition 1.4.1 to nonconvex sets C.

DEFINITION 2.2.4 Let X be a Banach space and $C \subseteq X$ nonempty.

(a) A vector $h \in X$ is said to be *tangent* to the set C at $x_0 \in \overline{C}$, if there exist $\varepsilon > 0$ and a map $r : [0, \varepsilon] \longrightarrow X$ such that

$$x_0 + \lambda h + r(\lambda) \in C \quad \text{for all } \lambda \in [0, \varepsilon] \quad \text{and} \quad \frac{\|r(\lambda)\|_X}{\lambda} \longrightarrow 0 \text{ as } \lambda \longrightarrow 0^+.$$

(b) The set of all vectors $h \in X$ that are tangent to the set C at $x_0 \in \overline{C}$ form a closed cone (which is nonempty because it contains the origin) denoted by $T_C(x_0)$. If this cone is a subspace, then it is called the *tangent space* to the set C at the point $x_0 \in \overline{C}$.

We also need the following notion from Banach space theory.

DEFINITION 2.2.5 Let X be a Banach space and Y a closed subspace of it. We say that a subspace V is a *topological complement* of Y if
(a) V is closed and
(b) $Y \cap V = \{0\}$ and $Y + V = X$. In this case we write $X = Y \oplus V$.

REMARK 2.2.6 A closed subspace Y of X admits a topological complement if and only if there exists a projection operator onto Y; that is, there exists $p_Y \in \mathcal{L}(X)$ such that $p_Y\big|_Y = \mathrm{id}_Y$. It is easy to see that if Y is finite-dimensional, then it admits a topological complement. Similarly, if Y has finite codimension (i.e., $\dim\big(X/Y\big) < +\infty$), then Y has a topological complement. In a Hilbert space every closed subspace admits a topological complement (consider the orthogonal complement of Y). In a Banach space that is not isomorphic to a Hilbert space, we can always find a closed subspace which has no topological complement.

Now let X, Y be two Banach spaces and $\psi\colon X \longrightarrow Y$ a map. We consider the following constraint set

$$C = \{x \in X : \psi(x) = 0\}.$$

The next theorem produces a very convenient characterization of the tangent space $T_C(x_0)$, $x_0 \in C$. The result is known as *Ljusternik's theorem.*

THEOREM 2.2.7 *If X, Y are Banach spaces, $U \subseteq X$ is a nonempty open set, $\psi\colon U \longrightarrow Y$ is continuously Fréchet differentiable, $C = \{x \in U : \psi(x) = 0\}$, and $x_0 \in C$ is a regular point of φ at which $N\big(\psi'(x_0)\big) = \ker \psi'(x_0)$ has a topological complement, then $T_C(x_0) = N\big(\psi'(x_0)\big)$.*

PROOF: Let $V = N\big(\psi'(x_0)\big)$. This is a closed subspace of X that by hypothesis has a topological complement $W \subseteq X$. So we can find $p_V,\ p_W \in \mathcal{L}(X)$ such that

$$\begin{aligned} &V = p_V(X) = N(p_W),\ W = p_W(X) = N(p_V),\\ &p_V^2 = p_V,\ p_W^2 = p_W \end{aligned}$$

(i.e., they are linear projection operators)

$$\text{and} \quad p_V + p_W = \mathrm{id}_X$$

(i.e., $X = V \oplus W$).

Because $U \subseteq X$ is an open set, we can find $r > 0$ such that

$$x_0 + B_r(0) + B_r(0) \subseteq U$$

(recall that $B_r(0) = \big\{x \in X : \|x\|_X < r\big\}$). Let $f\colon \big(V \cap B_r(0)\big) \times \big(W \cap B_r(0)\big) \longrightarrow Y$ be defined by

$$f(v, w) = \psi(x_0 + v + w) \qquad \text{for all } v \in V \cap B_r(0) \text{ and all } w \in W \cap B_r(0).$$

The function f has the following properties.

$$\begin{aligned} & f(0,0) = \psi(x_0) = 0 && && (2.5)\\ \text{and} \quad & f \in C^1(V \times W, Y) && \text{with} && f_1'(0,0) = \psi'(x_0)\big|_V = 0\\ \text{and} \quad & f_2'(0,0) = \psi'(x_0)\big|_W \in \mathcal{L}(W, Y). && && (2.6) \end{aligned}$$

Because $f_2'(0,0) \in \mathcal{L}(W, Y)$ is bijective, from Banach's theorem we know that $f_2'(0,0)^{-1} \in \mathcal{L}(Y, W)$. These properties of f permit the use of the implicit function theorem (see Theorem 1.1.23), which gives a $0 < \delta < r$ and a uniquely determined continuously Fréchet differentiable $g\colon V \cap B_\delta(0) \longrightarrow W$ such that

$$f\big(v, g(v)\big) = 0 \qquad \text{for all } v \in V \cap B_\delta(0) \tag{2.7}$$

$$\text{and} \qquad g(0) = 0, \; g'(0) = f_2'(0,0)^{-1} f_1'(0,0). \tag{2.8}$$

From (2.6) and (2.8), we have that $g'(0) = 0$ and so

$$\lim_{v \to 0} \frac{\|g(v)\|_Y}{\|v\|_X} = 0. \tag{2.9}$$

If we define $\xi \colon x_0 + V \cap B_\delta(0) \longrightarrow X$ by

$$\xi(x_0 + v) = x_0 + v + g(v),$$

then clearly ξ is continuous. Moreover, from (2.7) we have

$$0 = f\big(v, g(v)\big) = \psi\big(x_0 + v + g(v)\big),$$

hence ξ has values in C. Note that $v, g(v)$ are in complementary subspaces of X. Hence ξ is injective and so we can consider its inverse on $D = \big\{x_0 + v + g(v) : v \in V \cap B_\delta(0)\big\} \subseteq C$. So

$$\xi^{-1}\big(x_0 + v + g(v)\big) = x_0 + v = x_0 + p_V\big(v + g(v)\big),$$
$$\Rightarrow \; \xi^{-1} \text{ is continuous too.}$$

Therefore ξ is a homeomorphism of $U' = x_0 + V \cap B_\delta(0)$ onto $D \subseteq C$. This together with (2.9) imply that $T_C(x_0) = N\big(\psi'(x_0)\big)$. $\square$

REMARK 2.2.8 If every $x_0 \in C$ is a regular point and $N\big(\psi'(x_0)\big)$ admits a topological complement, then $C = \{x \in X : \psi(x) = 0\}$ is a C^1-manifold. If in addition $\psi \in C^r(U, Y)$, then C is a C^r-manifold. Note that the rather technical hypothesis that $N\big(\psi'(x_0)\big)$ admits a topological complement is automatically satisfied in the following three cases.

(a) Y is finite-dimensional.
(b) $X = H =$ a Hilbert space.
(c) For all $x \in U$ $\varphi'(x) \in \mathcal{L}(X, Y)$ is a Fredholm operator.

Cases (a) and (b) follow from Remark 2.2.6, and Case (c) is considered in Section 3.1.

COROLLARY 2.2.9 *If X, Y, Z are Banach spaces, $U \subseteq X$ is nonempty open, $\psi \in C^1(U, Y)$, $C = \{x \in U : \psi(x) = 0\}$, $x_0 \in C$ is a regular point of φ at which $N\big(\psi'(x_0)\big)$ admits a topological complement, and $\varphi \colon U \longrightarrow Z$ is Fréchet differentiable at x_0, then there exists a neighborhood U_0 of the origin in $T_C(x_0)$, a neighborhood D_0 of x_0 in C, and a continuously Fréchet differentiable homeomorphism g of U_0 onto D_0 such that $\varphi\big(g(u)\big) = \varphi(x_0) + \varphi'(x_0)u + o(\|u\|_X)$ with $\big(o(\|u\|_X)\big)/\|u\|_X \longrightarrow 0$ as $\|x\|_X \longrightarrow 0$.*

Using Theorem 2.2.7, we can produce necessary conditions for the existence of solutions for a constrained minimization problem. This theorem is the first result establishing the well-known *method of Lagrange multipliers*.

THEOREM 2.2.10 *If X, Y are Banach spaces, $U \subseteq X$ is a nonempty open set, $\psi \in C^1(U, Y)$, $\varphi\colon U \longrightarrow \mathbb{R}$ is Fréchet differentiable, $C = \{x \in U : \psi(x) = 0\}$, $x_0 \in C$ is a regular point of ψ at which $N\big(\psi'(x_0)\big)$ admits a topological complement, and x_0 is a local minimizer of φ on the constraint set C, then there exists $y_0^* \in Y^*$ such that*

$$\varphi'(x_0) = y_0^* \circ \psi'(x_0).$$

PROOF: Let $V = N\big(\psi'(x_0)\big)$ and let W be its topological complement (it exists by hypothesis). Because $x_0 \in C$ is a regular point of ψ,

$$K = \psi'(x_0)\Big|_W \in \mathcal{L}(W, Y)$$

is an isomorphism (i.e., $K^{-1} \in \mathcal{L}(Y, W)$ by Banach's theorem).

According to Theorem 2.2.7 we can find D_0 a neighborhood of x_0 in C such that for every $x_0 \in D_0$, we have

$$x = x_0 + u + g(u) \qquad \text{for all } u \in V \cap B_\delta(0),\ V = N\big(\psi'(x_0)\big)$$

with $\big(\|g(u)\|_X\big)/\|u\|_X \longrightarrow 0$ as $u \longrightarrow 0$ in X. Because x_0 is a local minimizer of φ on C, we can find $0 < \delta_1 \leq \delta$ such that

$$\begin{aligned}
&\varphi(x_0) \leq \varphi\big(x_0 + u + g(u)\big) \qquad \text{for all } u \in V \cap B_\delta(0),\\
\Rightarrow\ & 0 \leq \big\langle \varphi'(x_0), u + g(u)\big\rangle_X + o_{x_0}(\|u + g(u)\|_X) \qquad \text{for all } u \in V \cap B_{\delta_1}(0),\\
\Rightarrow\ & \big\langle \varphi'(x_0), u\big\rangle_X = 0 \quad \text{for all } u \in V \cap B_{\delta_1}(0) \text{ (recall } g \text{ is a homeomorphism)},\\
\Rightarrow\ & V \subseteq N\big(\varphi'(x_0)\big) = \ker\ \varphi'(x_0).
\end{aligned} \tag{2.10}$$

Let $p_W \in \mathcal{L}(X)$ be the canonical projection on W, the topological complement of V. Because of (2.10) we can define $\widehat{\varphi}\,'(x_0) \in W^*$ by

$$\big\langle \widehat{\varphi}\,'(x_0), p_W x\big\rangle_W = \big\langle \varphi'(x_0), x\big\rangle_X \qquad \text{for all } x \in X.$$

Then $y_0^* = \widehat{\varphi}\,'(x_0) \circ K^{-1} \in Y^*$. If $p_V \in \mathcal{L}(X)$ is the canonical projection of X onto V, then for every $x \in X$ we have $x = p_V x + p_W x$ and so

$$\begin{aligned}
\big\langle y_0^* \circ \psi'(x_0), x\big\rangle_X &= \big\langle y_0^* \circ \psi'(x_0), p_W x\big\rangle_X\\
&= \langle y_0^* \circ K, p_W x\rangle_X\\
&= \big\langle \widehat{\varphi}\,'(x_0), p_W x\big\rangle_X = \big\langle \varphi'(x_0), x\big\rangle_X\\
\Rightarrow y_0^* \circ \psi'(x_0) &= \varphi'(x_0) \quad \text{in } X^*.
\end{aligned}$$

$\square$

REMARK 2.2.11 According to Theorem 2.2.10, $x_0 \in C$ is a critical point of the function $L(x) = \varphi(x) - (y_0^* \circ \psi)(x)$, known as the *Lagrangian* function of the constrained minimization problem $\inf\limits_C \varphi$. Also the element $y^* \in Y^*$ is a *Lagrange multiplier*.

COROLLARY 2.2.12 *If $U \subseteq \mathbb{R}^N$ is a nonempty open set, $\varphi\colon U \longrightarrow \mathbb{R}$ is Fréchet differentiable, $\psi \in C^1(U, \mathbb{R}^m)$, $C = \{x \in U : \psi(x) = 0\}$, $x_0 \in C$ is a regular point* *of ψ, and it is also a local minimizer of φ on C, then we can find $\widehat{\lambda} = (\lambda_k)_{k=1}^m \in \mathbb{R}^m$ such that*

$$\frac{\partial \varphi}{\partial x_i}(x_0) = \sum_{k=1}^{m} \lambda_k \frac{\partial \psi}{\partial x_i}(x_0) \qquad \textit{for all} \quad i = 1, \ldots, N.$$

EXAMPLE 2.2.13 Let $X = H$ a Hilbert space and $A \in \mathcal{L}(H)$ a self-adjoint operator. Let $\varphi\colon H \longrightarrow \mathbb{R}$ be defined by $\varphi(x) = \langle Ax, x \rangle_H$. Also let $\psi\colon H \longrightarrow \mathbb{R}_+$ be defined by $\psi(x) = \|x\|_H^2 - 1$. Suppose that $x_0 \in H$, $\|x_0\|_H = 1$ is a minimizer of φ on $\partial B(0) = \{x \in H : \psi(x) = 0\}$. According to Theorem 2.2.10 (with $Y = \mathbb{R}$), we can find $\lambda \in \mathbb{R}$ such that $\varphi'(x_0) = \lambda \psi'(x_0)$, hence $2A(x_0) = 2\lambda x_0$ and so $A(x_0) = \lambda x_0$. In other words $x_0 \in H$ is an eigenvector of A for the eigenvalue λ. This example shows how we can establish the existence of eigenvalues for bounded self-adjoint linear operators on a Hilbert space.

Now we want to improve Theorem 2.2.10 by weakening the hypotheses on the constraint function ψ. To do this we need the following generalization of Ljusternik's theorem (see Theorem 2.2.7) due to Ioffe–Tichomirov [327, p. 34], where the interested reader can find its proof.

THEOREM 2.2.14 *If X, Y are Banach spaces, $U \subseteq X$ is a nonempty open set, $\psi\colon U \longrightarrow \mathbb{R}$ is Fréchet differentiable, $C = \{x \in U : \psi(x) = 0\}$ and $x_0 \in C$ is a regular point of ψ (i.e., $\psi'(x_0)(X) = Y$), then $T_C(x_0) = N\big(\psi'(x_0)\big)$.*

REMARK 2.2.15 Note that compared to Theorem 2.2.7, we have dropped the splitting property for $N\big(\psi'(x_0)\big)$.

We also need the following general result from operator theory, which underlines the idea of the Lagrange multiplier method.

PROPOSITION 2.2.16 *If X, Y are Banach spaces, $A \in \mathcal{L}(X, Y)$, $x_0^* \in X^*$, $\mathcal{R}(A)$ is closed and $N(A) \subseteq N(x_0^*)$, then there exists $y_0^* \in Y^*$ such that $x_0^* + y_0^* \circ A = 0$. Moreover, if $A(X) = Y$, then y_0^* is unique.*

PROOF: We know that $\mathcal{R}(A^*) = N(A)^\perp = \big\{x^* \in X^* : \langle x^*, x \rangle_X = 0 \text{ for all } x \in N(A)\big\}$. By hypothesis $x_0^* \in N(A)^\perp$. So it follows that $x_0^* = -A^* y_0^*$ for some $y_0^* \in Y^*$. Therefore

$$x_0^* + y_0^* \circ A = 0. \tag{2.11}$$

If $\mathcal{R}(A) = Y$, then $N(A^*) = \mathcal{R}(A)^\perp = \big\{y^* \in Y^* : \langle y^*, y \rangle_Y = 0 \text{ for all } y \in \mathcal{R}(A)\big\} = \{0\}$. So A^* is injective and by (2.11) this means that y_0^* is uniquely determined by x_0^*. $\square$

COROLLARY 2.2.17 *If X, Y are Banach spaces, $A \in \mathcal{L}(X, Y)$, $x_0^* \in X^*$, $\mathcal{R}(A)$ is closed, and $\mathcal{R}(A) \neq Y$, then there exists $y_0^* \in Y^*$, $y_0^* \neq 0$ such that $y_0^* \circ A = 0$.*

PROOF: Let $y \in Y$ such that $y \notin \mathcal{R}(A)$. Then we can find $y_0^* \in Y^*$ with $\|y_0^*\|_{Y^*} = 1$ such that

$$\langle y_0^*, y \rangle_Y = 0 \qquad \text{and} \qquad \langle y_0^*, u \rangle_Y = 0 \qquad \text{for all } u \in \mathcal{R}(A).$$

Therefore it follows that $y_0^* \neq 0$ and $y_0^* \circ A = 0$. □

We are led to the following generalization of Theorem 2.2.10.

THEOREM 2.2.18 *If X, Y are Banach spaces, $U \subseteq X$ is a nonempty open set, $\varphi \colon U \longrightarrow \mathbb{R}$ is Fréchet differentiable, $\psi \in C^1(U, Y)$, $C = \{x \in U : \psi(x) = 0\}$, $x_0 \in C$, $\mathcal{R}\big(\psi'(x_0)\big)$ is closed, and $x_0 \in C$ is a local minimizer of φ on C, then there exist $\lambda_0 \in \mathbb{R}$ and $y_0^* \in Y^*$ not both equal to zero such that*

$$\lambda_0 \varphi'(x_0) = y_0^* \circ \psi'(x_0). \tag{2.12}$$

PROOF: In the degenerate case $\mathcal{R}\big(\psi'(x_0)\big) \neq Y$, (2.12) is true with $\lambda_0 = 0$ by virtue of Corollary 2.2.17.

If $\mathcal{R}\big(\psi'(x_0)\big) = Y$, then provided that $N\big(\psi'(x_0)\big)$ admits a topological complement, the validity of (2.12) with $\lambda_0 = 1$ is a consequence of Theorem 2.2.10.

Finally, if we do not have the splitting of $N\big(\psi'(x_0)\big)$, then the result follows if we argue as in the proof of Theorem 2.2.10 using this time Theorem 2.2.14. □

Theorems 2.2.10 and 2.2.18 indicate that there is a close relation between Lagrange multipliers and constrained critical points of a function φ.

DEFINITION 2.2.19 Let X be a Banach space, $U \subseteq X$ nonempty open, $\varphi \colon U \longrightarrow \mathbb{R}$ a Fréchet differentiable map, $C \subseteq X$ a nonempty closed set and $x_0 \in C$. We say that x_0 is *a critical point of φ subject to the constraint C*, if for all curves $u : (-\varepsilon, \varepsilon) \longrightarrow X$ such that $u(t) \in C$ for all $t \in (-\varepsilon, \varepsilon)$, $u(0) = x_0$, $u'(0)$ exists, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\, \varphi\big(u(t)\big)\Big|_{t=0} = 0.$$

REMARK 2.2.20 If $x_0 \in \operatorname{int} C$ (the interior taken in X), then x_0 is a usual critical point of φ (also known as the *free critical point* of φ). A critical point that is not a local extremum of φ (i.e., it is not a local minimizer or a local maximizer of φ), is called a *saddle point*. So $x_0 \in C$ is a saddle point of φ subject to the constraint C, if for every $U \in \mathcal{N}(x_0) =$ filter of neighborhoods of x_0, we can find $y, v \in C \cap U$ such that $\varphi(y) < \varphi(x_0) < \varphi(v)$.

The next theorem establishes the connection between constrained critical points and Lagrange multipliers.

THEOREM 2.2.21 *If X, Y are Banach spaces, $U \subseteq X$ is nonempty open, $\varphi \colon U \longrightarrow \mathbb{R}$ is Fréchet differentiable, $\psi \in C^1(U, Y)$, $C = \{x \in X : \psi(x) = 0\}$, and $x_0 \in C$ is a regular point of φ at which $N\big(\varphi'(x_0)\big)$ admits a topological complement, then x_0 is a critical point of φ subject to the constraint C if and only if we can find $y_0^* \in Y^*$ such that $\varphi'(x_0) = y_0^* \circ \psi'(x_0)$.*

PROOF: $\Rightarrow$: From Definition 2.2.19 it follows that $\langle \varphi'(x_0), h \rangle_X = 0$ for all $h \in N\big(\psi'(x_0)\big)$. But from Theorem 2.2.10 we have that $N\big(\psi'(x_0)\big) = T_C(x_0)$. Invoking Proposition 2.2.16, we can find $y_0^* \in Y^*$ such that $\varphi'(x_0) = y_0^* \circ \psi'(x_0)$.

$\Leftarrow$: Let $u : (-\varepsilon, \varepsilon) \longrightarrow X$ be a curve as in Definition 2.2.19. Then $\psi\big(u(t)\big) = 0$ for all $t \in (-\varepsilon, \varepsilon)$ and $\psi'(x_0)u'(0) = 0$. It follows that $\langle \varphi'(x_0), u'(0)\rangle_X = 0$, hence $(\mathrm{d}/\mathrm{d}t)\ \varphi\big(u(t)\big)\Big|_{t=0} = 0$ (by the chain rule). So $x_0 \in C$ is a critical point of φ subject to the constraint C. $\square$

We conclude this section with the so-called *Dubovickii–Milyutin theorem*, which is crucial for the development of the Dubovickii–Milyutin formalism for the analysis of constrained optimization problems. For the proof we refer to Girsanov [266, p. 37].

THEOREM 2.2.22 *If X is a Banach space, $\{K_m\}_{m=1}^{n+1}$ are nonempty convex cones in X, and $\{K_m\}_{m=1}^{n}$ are open, then $\bigcap\limits_{m=1}^{n+1} K_m = \emptyset$ if and only if we can find $x_m^* \in K_m^* = \{x^* \in X^* : \langle x^*, x\rangle_X \geq 0$ for all $x \in K_m\}$, $m = 1, \ldots, n+1$, not all zero, such that $\sum\limits_{m=1}^{n+1} x_m^* = 0$.*

2.3 Saddle Points and Duality

Recall (see Remark 2.2.20), that a saddle point is a critical point of a function that is neither a local minimum nor local maximum. In this section we define a saddle point of a function on a product space $C \times D$. This new concept that we introduce is independent of the notion introduced in the previous section and it is global in nature.

DEFINITION 2.3.1 Let C, D be two nonempty sets and $\varphi\colon C \times D \longrightarrow \mathbb{R}$ a function. We say that (x_0, y_0) is a *saddle point* of φ, if we have

$$\varphi(x_0, y) \leq \varphi(x_0, y_0) \leq \varphi(x, y_0) \qquad \text{for all } (x, y) \in C \times D.$$

REMARK 2.3.2 The terminology of Definition 2.3.1 is justified by picturing the function $\varphi(x, y) = x^2 - y^2$ in $\mathbb{R}^3$ (a hyperbolic paraboloid). For this function $(0, 0)$ is a saddle point.

Note that we always have

$$\sup_{y \in D} \inf_{x \in C} \varphi(x, y) \leq \inf_{x \in C} \sup_{y \in D} \varphi(x, y). \tag{2.13}$$

If equality holds in (2.13), then the common value is called the *saddle value* of φ on $C \times D$.

DEFINITION 2.3.3 Let C, D be two nonempty sets and $\varphi\colon C \times D \longrightarrow \mathbb{R}$ a function. We say that φ satisfies a *minimax equality* on $C \times D$, if the following three conditions hold.

(a) φ has a saddle value (i.e., equality holds in (2.13)).

(b) There is $x_0 \in C$ such that $\sup\limits_{y \in D} \varphi(x_0, y) = \inf\limits_{x \in C} \sup\limits_{y \in D} \varphi(x, y)$.

(c) There is $y_0 \in D$ such that $\inf\limits_{x \in C} \varphi(x, y_0) = \sup\limits_{y \in D} \inf\limits_{x \in C} \varphi(x, y)$.

REMARK 2.3.4 In this case the inf and sup operations can be replaced by min and max, respectively, and we can write that $\max\limits_{y \in D} \min\limits_{x \in C} \varphi(x, y) = \min\limits_{x \in C} \max\limits_{y \in D} \varphi(x, y)$. This is the reason for the terminology minimax equality.

The following proposition is an immediate consequence of Definitions 2.3.1 and 2.3.3.

PROPOSITION 2.3.5 *If C, D are nonempty sets and $\varphi\colon C \times D \longrightarrow \mathbb{R}$, then φ has a saddle point in $C \times D$ if and only if φ satisfies a minimax equality on $C \times D$.*

DEFINITION 2.3.6 Let X be a vector space and $B \subseteq X$ a nonempty set. A multifunction (set-valued function) $F\colon B \longrightarrow 2^X \setminus \{\emptyset\}$ is said to be a *Knaster–Kuratowski–Mazurkiewicz map* (a KKM-map for short), if for every finite set $\{x_k\}_{k=1}^m$

$$\operatorname{conv}\{x_k\}_{k=1}^m \subseteq \bigcup_{k=1}^m F(x_k).$$

The basic property of KKM-maps is the following.

THEOREM 2.3.7 *If X is a Hausdorff topological vector space, $B \subseteq X$ is a nonempty set, and $F\colon B \longrightarrow 2^X \setminus \{\emptyset\}$ is a KKM-map with closed values, then the family $\{F(x)\}_{x \in B}$ of sets has the finite intersection property; that is, the intersection of each finite subfamily is nonempty.*

PROOF: We argue indirectly. So suppose that $\bigcap\limits_{k=1}^m F(x_k) = \emptyset$. Let $C = \operatorname{conv}\{x_k\}_{k=1}^m$ and consider the function $\vartheta\colon C \longrightarrow \mathbb{R}_+$ defined by $\vartheta(u) = \sum\limits_{k=1}^m d\big(u, F(x_k)\big)$. Evidently $\vartheta(u) > 0$ for all $u \in C$ and so we can define the continuous map $h\colon C \longrightarrow C$ by

$$h(u) = \frac{1}{\vartheta(u)} \sum_{k=1}^m d\big(u, F(x_k)\big) x_k. \tag{2.14}$$

By Brouwer's theorem (see Theorem 3.5.3), we can find $u_0 \in C$ such that

$$h(u_0) = u_0.$$

Let $K = \big\{k \in \{1, \dots, m\} : d\big(u_0, F(x_k)\big) \neq 0\big\}$. Then $u_0 \notin \bigcup\limits_{k \in K} F(x_k)$. On the other hand

$$u_0 = h(u_0) \in \operatorname{conv}\{x_k\}_{k \in K} \subseteq \bigcup_{k \in K} F(x_k),$$

(see (2.14) and recall that F is a KKM-map). So we have a contradiction, which finishes the proof. $\square$

REMARK 2.3.8 The hypotheses of this theorem can be weakened in the following way. We can drop the requirement that F has closed values and only assume that each $F(x)$ is finitely closed. Namely, if V is a finite-dimensional flat in X, then $V \cap F(x)$ is closed in the Euclidean topology of V. Evidently in this case we need only assume that X is a vector space. The finitely closed sets define a topology on X, known as the finite topology. It is stronger than any Hausdorff linear topology on X.

COROLLARY 2.3.9 *If X is a Hausdorff topological vector space, $B \subseteq X$ is a nonempty set, and $F\colon B \longrightarrow 2^X \setminus \{\emptyset\}$ is a KKM-map with closed values one of which is compact, then $\bigcap\limits_{x \in B} F(x) \neq \emptyset$.*

Using this corollary we can prove the following useful coincidence theorem for set-valued maps.

PROPOSITION 2.3.10 *If X, Y are Hausdorff topological vector spaces, $C \subseteq X$ and $D \subseteq Y$ are nonempty compact and convex sets, and $F, G\colon C \longrightarrow 2^D$ are two multifunctions such that*

(i) *$F(x)$ is open in D and $G(x)$ is nonempty convex for every $x \in C$.*

(ii) *$G^{-1}(y) = \{x \in C : y \in G(x)\}$ is open in C and $F^{-1}(y) = \{x \in C : y \in F(x)\}$ is convex for every $y \in D$,*

then we can find $x_0 \in C$ such that $F(x_0) \cap G(x_0) \neq \emptyset$.

PROOF: Let $E = C \times D$ and let $H\colon C \times D \longrightarrow 2^{X \times Y}$ be defined by

$$H(x, y) = E \cap \big(G^{-1}(y) \times F(x)\big)^c.$$

Because of hypotheses (i) and (ii), for every $(x, y) \in E$ the set $H(x, y)$ is closed in E, hence compact. Also given any $(x_0, y_0) \in E$, choose $(x, y) \in F^{-1}(y_0) \times G(x_0)$ (by hypotheses (i) and (ii), $F^{-1}(y_0) \times G(x_0) \neq \emptyset$). Hence $(x_0, y_0) \in G^{-1}(y) \times F(x)$ and so we infer that

$$E = \bigcup_{(x,y) \in E} \big(G^{-1}(y) \times F(x)\big). \tag{2.15}$$

From (2.15) and the definition of the multifunction H, it follows that

$$\bigcap_{(x,y) \in E} H(x, y) = \emptyset. \tag{2.16}$$

Because of (2.16) H cannot be a KKM-map. So according to Definition 2.3.6, we can find elements $\Big\{(x_k, y_k)\Big\}_{k=1}^m \subseteq E$ such that

$$\begin{aligned} &\operatorname{conv}\Big\{(x_k, y_k)\Big\}_{k=1}^m \not\subseteq \bigcup_{k=1}^m H(x_k, y_k) \\ \Rightarrow\ & v = \sum_{k=1}^m \lambda_k (x_k, y_k) \notin \bigcup_{k=1}^m H(x_k, y_k) \qquad \text{with } \lambda_k \geq 0,\ \sum_{k=1}^m \lambda_k = 1. \end{aligned}$$

Note that due to the convexity of $E = C \times D$ we have $v \in E$ and so $v \in E \setminus \bigcup_{k=1}^{m} H(x_k, y_k) = \bigcap_{k=1}^{m} G^{-1}(y_k) \cap F(x_k)$. Hence

$$\sum_{k=1}^{m} \lambda_k x_k \in G^{-1}(y_k) \qquad \text{and} \qquad \sum_{k=1}^{m} \lambda_k y_k \in F(x_k) \qquad \text{for all} \quad k = 1, \dots, m,$$

$$\Rightarrow\ y_k \in G\Big(\sum_{k=1}^{m} \lambda_k x_k\Big) \qquad \text{and} \qquad x_k \in F^{-1}\Big(\sum_{k=1}^{m} \lambda_k y_k\Big) \qquad \text{for all} \quad k = 1, \dots, m,$$

$$\Rightarrow\ \sum_{k=1}^{m} \lambda_k y_k \in G\Big(\sum_{k=1}^{m} \lambda_k x_k\Big) \qquad \text{and} \qquad \sum_{k=1}^{m} \lambda_k x_k \in F^{-1}\Big(\sum_{k=1}^{m} \lambda_k y_k\Big),$$

$$\Rightarrow\ \sum_{k=1}^{m} \lambda_k y_k \in G\Big(\sum_{k=1}^{m} \lambda_k x_k\Big) \qquad \text{and} \qquad \sum_{k=1}^{m} \lambda_k y_k \in F\Big(\sum_{k=1}^{m} \lambda_k x_k\Big).$$

So, if we set $x_0 = \sum_{k=1}^{m} \lambda_k x_k$ and $y_0 = \sum_{k=1}^{m} \lambda_k y_k$, we see that $y_0 \in F(x_0) \cap G(x_0) \neq \emptyset$. ∎

Using this proposition we can prove a basic minimax theorem. First we have a definition.

DEFINITION 2.3.11 Let V be a vector space, $B \subseteq V$ a nonempty convex set, and $\psi\colon B \longrightarrow \mathbb{R}$. We say that ψ is *quasiconvex* (resp., *quasiconcave*), if for all $\lambda \in \mathbb{R}$ the set $\{x \in B : \psi(x) \leq \lambda\}$ (resp., the set $\{x \in B : \psi(x) \geq \lambda\}$) is convex in B.

REMARK 2.3.12 Evidently a convex (resp., concave) function is quasiconvex (resp., quasiconcave). The converse is not in general true. Consider, for example, the function $x \longrightarrow \ln(|x| + 1)$ which is quasiconvex, but not convex.

THEOREM 2.3.13 *If X, Y are Hausdorff topological vector spaces, $C \subseteq X$ and $D \subseteq Y$ are nonempty, compact, and convex sets, and $\varphi\colon C \times D \longrightarrow \mathbb{R}$ satisfies*

(i) *For every $y \in D$, $x \longrightarrow \varphi(x, y)$ is lower semicontinuous and quasiconvex.*

(ii) *For every $x \in C$, $y \longrightarrow \varphi(x, y)$ is upper semicontinuous and quasiconcave,*

then φ has a saddle point on $C \times D$.

PROOF: By virtue of Proposition 2.3.5 it suffices to show that φ satisfies a minimax equality on $C \times D$, namely that

$$\min_{x \in C} \max_{y \in D} \varphi(x, y) = \max_{y \in D} \min_{x \in C} \varphi(x, y). \tag{2.17}$$

Because of the upper semicontinuity of $\varphi(x, \cdot)$ and the compactness of D, $\max_{y \in D} \varphi(x, y)$ exists for every $x \in C$ (see Theorem 2.1.10). Also if $h_1(x) = \max_{y \in D} \varphi(x, y)$, then we claim that $x \longrightarrow h_1(x)$ is lower semicontinuous. Indeed let $x_\alpha \longrightarrow x$ in C and suppose $h_1(x_\alpha) \leq \lambda$ for all $\alpha \in J$. We have $\varphi(x_\alpha, y) \leq \lambda$ for all $\alpha \in J$ and all $y \in D$. Because $\varphi(\cdot, y)$ is lower semicontinuous, we have

$\varphi(x,y) \leq \liminf_{\alpha\in J} \varphi(x_\alpha, y) \leq \lambda$ for all $y \in D$, hence $h_1(x) \leq \lambda$ and this by virtue of Proposition 2.1.3(b) implies that $h_1(\cdot)$ is lower semicontinuous. Therefore $\min_{x\in C} h_1(x) = \min_{x\in C}\max_{y\in D} \varphi(x,y)$ exists. Similarly we show that $\max_{y\in D}\min_{x\in C}\varphi(x,y)$ exists. We know that

$$\max_{y\in D}\min_{x\in C}\varphi(x,y) \leq \min_{x\in C}\max_{y\in D}\varphi(x,y) \quad \text{(see (2.13))}. \tag{2.18}$$

We show that in (2.18) we cannot have strict inequality. We proceed by contradiction. Suppose we can find $\lambda \in \mathbb{R}$ such that

$$\max_{y\in D}\min_{x\in C}\varphi(x,y) < \lambda < \min_{x\in C}\max_{y\in D}\varphi(x,y). \tag{2.19}$$

We introduce the multifunctions $F, G \colon C \longrightarrow 2^D$ defined by

$$F(x) = \{y \in D : \varphi(x,y) < \lambda\} \qquad \text{and} \qquad G(x) = \{y \in D : \varphi(x,y) > \lambda\}.$$

Note that because of hypothesis (ii) for each $x \in C$, $F(x)$ is open in D and $G(x)$ is convex and nonempty (see (2.19)). Also $G^{-1}(y) = \{x \in C : \varphi(x,y) > \lambda\}$ is open in C due to hypothesis (i), whereas $F^{-1}(y) = \{x \in C : \varphi(x,y) < \lambda\}$ is convex again due to hypothesis (i). So we can apply Proposition 2.3.10 and produce a point $(\widehat{x}_0, \widehat{y}_0) \in C \times D$ such that $\widehat{y}_0 \in F(\widehat{x}_0) \cap G(\widehat{x}_0)$. But then

$$\lambda < \varphi(\widehat{x}_0, \widehat{y}_0) < \lambda,$$

a contradiction. This proves that (2.17) holds and so φ admits a saddle point. $\square$

Another saddle point theorem useful in applications is the following.

THEOREM 2.3.14 *If X, Y are reflexive Banach spaces, $C \subseteq X$ and $D \subseteq Y$ are nonempty, closed, and convex sets, and $\varphi \colon C \times D \longrightarrow \mathbb{R}$ is a function such that*

(i) *For every $y \in D$, $x \longrightarrow \varphi(x,y)$ is convex and lower semicontinuous.*

(ii) *For every $x \in C$, $y \longrightarrow \varphi(x,y)$ is concave and upper semicontinuous.*

(iii) *C is bounded or there exists $\widehat{y}_0 \in D$ such that $\varphi(x, \widehat{y}_0) \longrightarrow +\infty$ as $\|x\|_X \longrightarrow +\infty$.*

(iv) *D is bounded or there exists $\widehat{x}_0 \in C$ such that $\varphi(\widehat{x}_0, y) \longrightarrow -\infty$ as $\|y\|_Y \longrightarrow +\infty$,*

then φ has a saddle point on $C \times D$.

PROOF: Due to the reflexivity of X and Y, if $C \subseteq X$ and $D \subseteq Y$ are bounded, they are w-compact. Also due to convexity the functions $x \longrightarrow \varphi(x,y)$ and $y \longrightarrow -\varphi(x,y)$ are weakly lower semicontinuous. So in this case the result follows from Theorem 2.3.13.

So suppose that at least one of the sets $C \subseteq X$ and $D \subseteq Y$ is unbounded. For every $n \geq 1$ set

$$C_n = \{x \in C : \|x\|_X \leq n\} \qquad \text{and} \qquad D_n = \{y \in D : \|y\|_Y \leq n\}.$$

For $n \geq 1$ large $C_n \neq \emptyset$ and $D_n \neq \emptyset$. Then from the bounded case we know that we can find $(x_n, y_n) \in C_n \times D_n$ such that

$$\varphi(x_n, y) \le \varphi(x_n, y_n) \le \varphi(x, y_n) \qquad \text{for all } (x,y) \in C_n \times D_n,\ n \ge 1. \tag{2.20}$$

Let $x = \widehat{x}_0$ and $y = \widehat{y}_0$ (see hypotheses (iii) and (iv)). Using those hypotheses, from (2.20) we infer that $\{x_n\}_{n\ge 1} \subseteq X$, $\{y_n\}_{n\ge 1} \subseteq Y$, and $\{\varphi(x_n, y_n)\}_{n\ge 1} \subseteq \mathbb{R}$ are all bounded. So passing to a subsequence if necessary, we may assume that

$$x_n \xrightarrow{w} x_0 \text{ in } X, \qquad y_n \xrightarrow{w} y_0 \text{ in } Y \qquad \text{and} \qquad \varphi(x_n, y_n) \longrightarrow \xi \in \mathbb{R}.$$

Evidently $(x_0, y_0) \in C \times D$. Also because of hypotheses (i) and (ii) we have

$$\varphi(x_0, y) \le \liminf_{n\to\infty} \varphi(x_n, y) \le \xi \le \limsup_{n\to\infty} \varphi(x, y_n) \le \varphi(x, y_0)$$

for all $(x,y) \in C \times D$,

$$\Rightarrow \ \varphi(x_0, y) \le \varphi(x_0, y_0) \le \varphi(x, y_0),$$

that is, $(x_0, y_0) \in C \times D$ is a saddle point of φ.

□

REMARK 2.3.15 Functions $\varphi(x,y)$ such as those in Theorems 2.3.13 and 2.3.14 that are convex in $x \in C$ and concave in $y \in D$, are usually called *saddle functions* (or *convex–concave functions*).

Next we derive a duality theory for convex optimization. Saddle functions are helpful in this respect. The mathematical setting is the following. We are given two Banach spaces X, Y and a convex function $\varphi \colon X \times Y \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. We know that $(X \times Y)^* = X^* \times Y^*$ and for the duality brackets we have

$$\langle (x^*, y^*), (x,y) \rangle_{X\times Y} = \langle x^*, x \rangle_X + \langle y^*, y \rangle_Y\,,$$

for all $x \in X$, $x^* \in X^*$, $y \in Y$, $y^* \in Y^*$.

We can think of $y \in Y$ as a perturbation and $\varphi(x,y)$ as a perturbed functional with $\varphi(x,0)$ as the original unperturbed functional. Then the primal problem is:

$$(\mathrm{P}) : \ \inf_{x\in X} \varphi(x, 0). \tag{2.21}$$

We also consider the value function $m \colon Y \longrightarrow \mathbb{R}^*$ of the perturbed problem

$$m(y) = \inf_{x\in X} \varphi(x, y). \tag{2.22}$$

Clearly $m(\cdot)$ is convex. We have

$$\begin{aligned} m^*(y^*) &= \sup_{y\in Y} \big[\langle y^*, y \rangle_Y - \inf_{x\in X} \varphi(x,y) \big] \quad \text{(see (2.22))} \\ &= \sup_{(x,y)\in X\times Y} \big[\langle 0, x \rangle_X + \langle y^*, y \rangle_Y - \varphi(x,y) \big] = \varphi^*(0, y^*). \end{aligned} \tag{2.23}$$

Then the dual problem associated to (P) (see (2.21)), is defined by

$$(\mathrm{D}) : \ \sup_{y^*\in Y^*} \big(-\varphi^*(0, y^*) \big). \tag{2.24}$$

PROPOSITION 2.3.16 *We always have* $\sup(\mathrm{D}) \le \inf(\mathrm{P})$ (*weak duality*).

PROOF: For all $x \in X$ and all $y^* \in Y^*$, by the Young–Fenchel inequality (see Proposition 1.2.18), we have

$$\begin{aligned} &\varphi(x,0) + \varphi^*(0,y^*) \geq \langle 0,x\rangle_X + \langle y^*,0\rangle_Y = 0 \\ \Rightarrow\ &\inf(\mathrm{P}) \geq \sup(\mathrm{D}). \end{aligned}$$

$\square$

In the next proposition we provide conditions for strong duality to hold (i.e., to have equality of primal and dual problems).

PROPOSITION 2.3.17 *If $\partial m(0) \neq \emptyset$ and $y^* \in \partial m(0)$, then y^* is a solution of* (D) *and* $\inf(\mathrm{P}) = \max(\mathrm{D})$.

PROOF: Because $\partial m(0) \neq \emptyset$, we have $m(0) = \inf(\mathrm{P}) \in \mathbb{R}$ and for all $y^* \in Y^*$, $m^*(y^*) \geq -m(0) > -\infty$. From Proposition 1.2.31, we have

$$\begin{aligned} &\inf(\mathrm{P}) + m^*(y^*) = m(0) + m^*(y^*) = \langle y^*,0\rangle_Y = 0, \\ \Rightarrow\ &\sup(\mathrm{D}) = \inf(\mathrm{P}) \quad \text{(see (2.23), (2.24), and Proposition 2.3.16)}. \end{aligned}$$

$\square$

In order to analyze further the strong duality situation, we need the following lemma.

LEMMA 2.3.18 *If V is a Banach space, $f\colon V \longrightarrow \mathbb{R}^*$ is a convex function, and there exist $(v_0,\lambda_0) \in V \times \mathbb{R}$ such that $f(v_0) \in \mathbb{R}$ and $(v_0,\lambda_0) \in$ int epi f, then $f(v) > -\infty$ for all $v \in V$.*

PROOF: Let $\mu_0 < f(v_0)$. Then $(v_0,\mu_0) \notin \mathrm{epi}\, f$ and so by the separation theorem for convex sets we can find $(v^*,t) \in V^* \times \mathbb{R}$, $(v^*,t) \neq (0,0)$ such that

$$\langle v^*,v_0\rangle_V + t\mu_0 \leq \langle v^*,v\rangle_V + t\lambda \quad \text{for all} \quad (v,\lambda) \in \mathrm{epi}\, f. \tag{2.25}$$

Because λ can increase up to $+\infty$, from (2.25) it follows that $t \geq 0$. If $t = 0$, then

$$\langle v^*,v_0\rangle_V \leq \langle v^*,v\rangle_V \quad \text{for all} \quad v \in \mathrm{dom}\, f. \tag{2.26}$$

But int dom $f \neq \emptyset$ (see Theorem 1.2.3(d)). So from (2.26) we infer that $v^* = 0$, a contradiction to the fact that $(v^*,t) \neq (0,0)$. Hence $t > 0$ and we may assume that $t = 1$. We have

$$\begin{aligned} &\langle v^*,v_0 - v\rangle_V + \mu_0 \leq \lambda \qquad \text{for all } (v,\lambda) \in \mathrm{epi}\, f, \\ \Rightarrow\ &\langle v^*,v_0 - v\rangle_V + \mu_0 \leq f(v) \qquad \text{for all } v \in V \text{ (i.e., } f(v) > -\infty \quad \text{for all } v \in V). \end{aligned}$$

$\square$

Using this auxiliary result we can have a theorem on the duality of problems (P) and (D).

THEOREM 2.3.19 *If there exists $x_0 \in X$ such that $\varphi(x_0,\cdot)$ is finite and continuous at $y = 0$ and* $\inf(\mathrm{P}) \in \mathbb{R}$, *then*

(a) *m is continuous at $y = 0$.*

(b) $y^* \in \partial m(0)$ *if and only if* $y^* \in Y^*$ *is a solution of the dual problem* (D).
(c) $\inf(\mathrm{P}) = \max(\mathrm{D}) \in \mathbb{R}$ (*strong duality*).

PROOF: We can find $r>0$ and $M>0$ such that

$$\begin{aligned} &\varphi(x_0, y) \le M \qquad \text{for all } \ y \in \overline{B}_r(0), \\ \Rightarrow \ & m(y) \le M \qquad \text{for all } \ y \in \overline{B}_r(0). \end{aligned}$$

Then from Theorem 1.2.3 we have that m is continuous at $y=0$ and $(0,\lambda) \in \operatorname{int}\operatorname{epi} m$ for all $\lambda > M$. Because $m(0) = \inf(\mathrm{P}) \in \mathbb{R}$, from Lemma 2.3.18 we infer that $-\infty < m(y)$ for all $y \in Y$, hence m is a proper convex function continuous at $y = 0$. By virtue of Theorem 1.2.34 we have that $\partial m(0) \neq \emptyset$. Invoking Proposition 2.3.17, we know that every $y^* \in \partial m(0)$ is a solution of (D) and $\inf(\mathrm{P}) = \max(\mathrm{P})$. Finally, if $y^* \in Y^*$ is a solution of (D), we have

$$m^*(y^*) \le m^*(u^*) \quad \text{for all } \ u^* \in Y^*. \tag{2.27}$$

Because m is continuous at $y = 0$, using Proposition 1.2.24 (with $\varphi = m$ and $\psi = i_{\{0\}}$) we obtain

$$\begin{aligned} & m(0) = \sup_{u^* \in Y^*} \left[\langle u^*, 0 \rangle_Y - m^*(u^*) \right] \le -m^*(y^*) \quad \text{(see (2.27))}, \\ \Rightarrow \ & m(0) + m^*(y^*) \le 0 \quad \text{(i.e., } y^* \in \partial m(0); \ \text{see Proposition 1.2.31).} \end{aligned}$$

□

REMARK 2.3.20 In the above theorem, the hypothesis that $\inf(\mathrm{P}) \in \mathbb{R}$ is important and cannot be dropped. To see this consider the case where $X = Y = \mathbb{R}$; let $C(y) = \{x \in \mathbb{R} : x \le y\}$ and

$$\varphi(x, y) = x + i_{C(y)}(x).$$

Then for every $x<0$, $\varphi(x,\cdot)$ is continuous at 0, but $m \equiv -\infty$ and so $m^* \equiv +\infty$.

DEFINITION 2.3.21 The *Lagrangian function* corresponding to the dual pair of problems $\{(\mathrm{P}),(\mathrm{D})\}$, is the function $L\colon X \times Y^* \longrightarrow \mathbb{R}^*$ defined by

$$L(x, y^*) = \inf \left[\varphi(x,y) - \langle y^*, y \rangle_Y : y \in Y \right] \qquad \text{for all } \ (x, y^*) \in X \times Y^*.$$

REMARK 2.3.22 Evidently $\inf\limits_{x \in X} L(x,y) = -\varphi^*(0, y^*)$ and L is a saddle function.

EXAMPLE 2.3.23 Consider the mathematical programming situation. Namely let

$$(\mathrm{P}) : \inf_{x \in X} [f(x) : f_k(x) \le 0 \qquad \text{for all } \ k = 1, \ldots, m],$$

with $f, f_k\colon X \longrightarrow \overline{\mathbb{R}}$, $k = 1, \ldots, m$, convex functions. Then $\varphi\colon X \times \mathbb{R}^m \longrightarrow \overline{\mathbb{R}}$ is defined by

$$\varphi(x,y) = f(x) + i_{C(y)}(x) \qquad \text{for all } \ x \in X, \ y = (y_k)_{k=1}^m \in \mathbb{R}^m,$$

with $C(y) = \{x \in X : f_k(x) \le y_k \quad \text{for all } \ k = 1, \ldots, m\}$. Therefore

$$L(x, y^*) = \begin{cases} f(x) - \sum_{k=1}^m f_k(x) y_k^* & \text{if } y^* \in -\mathbb{R}_+^m \\ -\infty & \text{if } y^* \notin -\mathbb{R}_+^m \end{cases},$$

for all $x \in X$, $y^* = (y_k^*)_{k=1}^m \in \mathbb{R}^m$.

PROPOSITION 2.3.24 *If $\varphi \in \Gamma_0(X \times Y)$, then $\varphi(x,0) = \sup_{y^* \in Y^*} L(x,y^*)$.*

PROOF: Because of Theorem 1.2.21, we have

$$\varphi(x,0) = \sup\left[\langle y^*, 0\rangle_Y - \big(\varphi(x,\cdot)\big)^*(y^*) : y^* \in Y^*\right] = \sup_{y^* \in Y^*} L(x,y^*).$$

□

We can characterize the situation of strong duality via the saddle points of the Lagrangian.

THEOREM 2.3.25 *If $\varphi \in \Gamma_0(X \times Y)$ and L is the corresponding Lagrangian function, then the following statements are equivalent.*

(a) *$(x_0, y_0^*) \in X \times Y^*$ is a saddle point of L.*
(b) *x_0 is a solution of* (P)*, y_0^* is a solution of* (D) *and* min(P)= max(D).
(c) *$\varphi(x_0, 0) = -\varphi^*(0, y_0^*)$.*

PROOF: (a)⇒**(b),(c):** We have

$$L(x_0,y_0^*) = \inf_{x \in X} L(x,y_0^*) = -\varphi^*(0,y_0^*) \quad \text{(see Remark 2.3.22)} \tag{2.28}$$

$$\text{and} \quad L(x_0,y_0^*) = \sup_{y^* \in Y^*} L(x_0,y^*) = \varphi(x_0,0) \quad \text{(see Proposition 2.3.24).} \tag{2.29}$$

So from (2.28) and (2.29), we have

$$\inf(\mathrm{P}) \le \varphi(x_0,0) = -\varphi^*(0,y_0^*) \le \inf(\mathrm{D}). \tag{2.30}$$

Combining (2.30) with Proposition 2.3.16, we conclude that

$$\min(\mathrm{P}) = \max(\mathrm{D}) = \varphi(x_0,0) = -\varphi^*(0,y_0^*).$$

(b)⇒(a), (c): We have

$$\varphi(x_0,0) = \min(\mathrm{P}), \quad -\varphi^*(0,y_0^*) = \max(\mathrm{D}) \quad \text{and} \quad \varphi(x_0,0) = -\varphi^*(0,y_0^*). \tag{2.31}$$

Note that

$$\varphi(x_0,0) = \sup_{y^* \in Y^*} L(x_0,y^*) \ge L(x_0,y^*) \quad \text{(see Proposition 2.3.24)} \tag{2.32}$$

$$\text{and} \quad -\varphi^*(0,y_0^*) = \inf_{x \in X} L(x,y_0^*) \le L(x,y_0^*) \quad \text{(see Remark 2.3.22).} \tag{2.33}$$

From (2.31) through (2.33) it follows that $(x_0,y_0^*) \in X \times Y^*$ is a saddle point of the Lagrangian.

(c)⇒(a), (b): We have

$$\inf (\mathrm{P}) \leq \varphi(x_0, 0) = -\varphi^*(0, y_0^*) \leq \sup (\mathrm{D}),$$
$$\Rightarrow \inf (\mathrm{P}) = \sup (\mathrm{D}) = \varphi(x_0, 0) = -\varphi^*(0, y_0^*) \quad \text{(see Proposition 2.3.16).}$$

□

Next we consider an important special case of the above general duality theory. So let $f \in \Gamma_0(X),\ g \in \Gamma_0(Y)$, and $A \in \mathcal{L}(X, Y)$. The primal problem is:

$$(\mathrm{P})' : \inf[f(x) + g\big(A(x)\big) : x \in X]. \tag{2.34}$$

The convex perturbation functional $\varphi\colon X \times Y \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is given by

$$\varphi(x, y) = f(x) + g\big(A(x) + y\big) \qquad \text{for all } (x, y) \in X \times Y.$$

Then the dual problem is:

$$(\mathrm{D})' : \sup \big[- f^*\big(A^*(y^*)\big) - g^*(-y^*) : y^* \in Y^*\big]. \tag{2.35}$$

Also the Lagrangian function for the pair of problems $\big\{(\mathrm{P})', (\mathrm{D})'\big\}$ (see (2.34) and (2.35)) is given by

$$L(x, y^*) = \begin{cases} f(x) + \langle y^*, A(x)\rangle_Y - g^*(y^*) & \text{if } f(x) < +\infty \\ +\infty & \text{if } f(x) = +\infty \end{cases} \tag{2.36}$$

Using Theorem 2.3.25, we obtain the following.

PROPOSITION 2.3.26 *A pair $(x_0, y_0^*) \in X \times Y^*$ is a saddle point of the Lagrangian L defined by (2.36) if and only if*

$$y_0^* \in \partial g\big(A(x_0)\big) \qquad \text{and} \qquad - A^*(y_0^*) \in \partial f(x_0).$$

Let us close this section with a simple application of duality theory on a calculus of variations problem.

EXAMPLE 2.3.27 Let $X = W_0^{1,2}(0,1),\ Y = L^2(0,1),\ A \in \mathcal{L}(X, Y)$ defined by $A(x) = x'$, and $f \in \Gamma_0(X),\ g \in \Gamma_0(Y)$ defined by

$$f(x) = \langle x_0^*, x\rangle_X = -\int_0^1 h(t)x(t)\mathrm{d}t = -(h, x)_{L^2(0,1)} \qquad \text{for some } h \in L^2(0,1)$$
$$\text{and} \quad g(y) = \frac{1}{2}\|y\|^2_{L^2(0,1)}.$$

Then identifying $L^2(0,1)$ with its dual (hence $X \subseteq Y \subseteq X^* = W^{-1,2}(0,1)$), we have

$$f^*(x^*) = i_{\{x_0^*\}}(x^*) \qquad \text{for all } x^* \in X^* \qquad \text{and} \qquad g^*(y^*) = \frac{1}{2}\|y^*\|^2_{L^2(0,1)}.$$

Note that $-A^*(y^*) = x_0^*$ if and only if

$$-\int_0^1 y^*(t)x'(t)\mathrm{d}t = -\int_0^1 h(t)x(t)\mathrm{d}t \qquad \text{for all } x \in W_0^{1,2}(0,1),$$
$$\Rightarrow\ y^* \in W^{1,2}(0,1) \qquad \text{and} \qquad (y^*)' = -h.$$

Then the primal problem is

$$(\mathrm{P})'' : \inf \Big[\frac{1}{2} \|x'\|^2_{L^2(0,1)} - (h,x)_{L^2(0,1)} : x \in W_0^{1,2}(0,1) \Big]$$

and the dual problem is

$$(\mathrm{D})'' : \sup \Big[- \frac{1}{2} \|y^*\|^2_{L^2(0,1)} : y^* \in W^{1,2}(0,1),\ (y^*)' = -h \Big].$$

Then problems $(\mathrm{P})''$ and $(\mathrm{D})''$ have solutions $x_0 \in W_0^{1,2}(0,1) \cap W^{2,2}(0,1)$ and $y_0^* \in W^{1,2}(0,1)$ (see Corollary 2.1.11) and they are unique due to the strict convexity of the functionals. Moreover, we have

$$-x_0''(t) = h(t) \text{ a.e. on } (0,1), \quad x_0(0) = x_0(1) = 0 \qquad \text{and} \qquad y_0^* = x_0'.$$

2.4 Variational Principles

In Theorem 2.1.10 and Corollary 2.1.11 we saw that if the objective functional $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is lower semicontinuous in a topology and the domain exhibits some kind of (local) compactness in the same topology, then a minimizing point exists. This is the cornerstone of the so-called direct method of the calculus of variations. However, in many situations the above convenient setting is not present. Think, for example, of infinite-dimensional Banach space functionals that are strongly lower semicontinuous, but not weakly. An effective tool to approach such problems is provided by the so-called Ekeland variational principle. Since its appearance in 1974, this result found many significant applications in different parts of analysis. Also it turned out to be equivalent to some other important results of nonlinear analysis and also served to provide new and elegant proofs to known results. In addition it initiated the production of some other related variational principles.

In this section we survey this area exhibiting all those features that are significant for problems in nonlinear analysis. Of course our presentation is centered on the Ekeland variational principle.

We start by giving the general form of the Ekeland variational principle.

THEOREM 2.4.1 *If (X,d) is a complete metric space, $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is a proper lower semicontinuous function that is bounded below, and $\varepsilon > 0$, $\bar{x} \in X$ satisfy*

$$\varphi(\bar{x}) \leq \inf_X \varphi + \varepsilon,$$

then for any $\lambda > 0$, we can find $x_\lambda \in X$ such that

$$\varphi(x_\lambda) \leq \varphi(\bar{x}), \quad d(x_\lambda, \bar{x}) \leq \lambda \quad \textit{and} \quad \varphi(x_\lambda) \leq \varphi(x) + \frac{\varepsilon}{\lambda} d(x, x_\lambda) \qquad \textit{for all } x \in X.$$

PROOF: In what follows $d_\lambda = (1/\lambda) d$. We introduce the following relation on $X \times X$.

$$x \preceq u \qquad \text{if and only if } \varphi(x) \leq \varphi(u) - \varepsilon d_\lambda(x,u). \tag{2.37}$$

It is easy to see that $\preceq$ is reflexive, antisymmetric, and transitive, hence a partial order on X. We define a sequence of nonempty sets $C_n \subseteq X$, $n \geq 1$ as follows. Let $x_1 = \bar{x}$ and set

$$C_1 = \{x \in X : x \leq x_1\}.$$

Choose $x_2 \in C_1$ such that $\varphi(x_2) \leq \inf_{C_1} \varphi + (\varepsilon/2^2)$ and define $C_2 = \{x \in X : x \leq x_2\}$. We continue this way and inductively define

$$\begin{aligned} C_n = \{x \in X : x \leq x_n\} \qquad &\text{and} \qquad x_{n+1} \in C_n \qquad \text{such that} \\ \varphi(x_{n+1}) \leq \inf_{C_n} \varphi + \frac{\varepsilon}{2^{n+1}},\ n \geq 1. \end{aligned} \tag{2.38}$$

Evidently $\{C_n\}_{n\geq 1}$ is a decreasing sequence and each set C_n is closed. Indeed, suppose $\{u_m\}_{m\geq 1} \subseteq C_n$ such that $u_m \longrightarrow u$. We have

$$\begin{aligned} &\varphi(u_m) \leq \varphi(x_n) - \varepsilon d_\lambda(u_m, x_n), \\ \Rightarrow\ &\varphi(u) \leq \liminf_{m\longrightarrow\infty} \varphi(u_m) \leq \varphi(x_n) - \varepsilon d_\lambda(u, x_n) \quad (\text{i.e., } u \in C_n). \end{aligned}$$

We claim that $\operatorname{diam} C_n \longrightarrow 0$ as $n \to \infty$. To this end let $u \in C_n$. We have $u \leq x_n$ and so

$$\varphi(u) \leq \varphi(x_n) - \varepsilon d_\lambda(u, x_n) \quad (\text{see } (2.37)). \tag{2.39}$$

Also note that $u \in C_{n-1}$, because $\{C_n\}_{n\geq 1}$ is a decreasing sequence. Therefore

$$\varphi(x_n) \leq \varphi(u) + \frac{\varepsilon}{2^n} \quad (\text{see } (2.38)). \tag{2.40}$$

Combining (2.39) and (2.40), we obtain

$$d_\lambda(u, x_n) \leq \frac{1}{2^n}. \tag{2.41}$$

Because $u \in C_n$ was arbitrary, from (2.41) and the triangle inequality it follows that

$$\begin{aligned} &\operatorname{diam} C_n \leq \frac{1}{2^{n-1}}, \\ \Rightarrow\ &\operatorname{diam} C_n \longrightarrow 0 \qquad \text{as } n \to \infty. \end{aligned}$$

Then by Cantor's theorem, we have that $\bigcap_{n\geq 1} C_n = \{x_\lambda\}$. Because $x_\lambda \in C_1$, we have $x_\lambda \leq \bar{x}$ and so $\varphi(x_\lambda) \leq \varphi(\bar{x})$ (see (2.37)). Also from the triangle inequality and (2.41), we have

$$\begin{aligned} &d_\lambda(\bar{x}, x_n) \leq \sum_{k=1}^{n-1} d_\lambda(x_k, x_{k+1}) \leq \sum_{k=1}^{n-1} \frac{1}{2^k} \\ \Rightarrow\ &d_\lambda(\bar{x}, x_\lambda) = \lim_{n\to\infty} d_\lambda(\bar{x}, x_n) \leq 1 \quad (\text{i.e., } d_\lambda(\bar{x}, x_\lambda) \leq \lambda). \end{aligned}$$

Finally let $x \neq x_\lambda$. Evidently we cannot have $x \leq x_\lambda$ or otherwise $x \in \bigcap_{n\geq 1} C_n$, a contradiction. Then by virtue of (2.37)

$$\varphi(x_\lambda) < \varphi(x) + \frac{\varepsilon}{\lambda} d(x, x_\lambda).$$

$\square$

REMARK 2.4.2 In the conclusions of Theorem 2.4.1, the relations $d(\bar{x}, x_\lambda) \leq \lambda$ and $\varphi(x_\lambda) \leq \varphi(x) + (\varepsilon/\lambda)d(x, x_\lambda)$ for all $x \in X$, are in a sense complementary. Indeed, the choice of $\lambda > 0$ allows us to strike a balance between them. So if $\lambda > 0$ is large the inequality $d(x, x_\lambda) \leq \lambda$ gives little information on the whereabouts of x_λ, and the inequality $\varphi(x_\lambda) \leq \varphi(x) + (\varepsilon/\lambda)d(x, x_\lambda)$, $x \in X$ becomes sharper and says that x_λ is close to being a global minimizer of φ. The situation is reversed if $\lambda > 0$ is small. Two important cases are when $\lambda = 1$ (which means we do not care about the whereabouts of x_λ) and when $\lambda = \sqrt{\varepsilon}$ (which means that we need to have information from both inequalities). We state both cases as corollaries of Theorem 2.4.1.

COROLLARY 2.4.3 *If (X, d) is a complete metric space and $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is a proper lower semicontinuous function that is bounded below, then for any $\varepsilon > 0$, we can find $x_\varepsilon \in X$ such that*

$$\varphi(x_\varepsilon) \leq \inf_X \varphi + \varepsilon \qquad \text{and} \qquad \varphi(x_\varepsilon) \leq \varphi(x) + \varepsilon d(x, x_\varepsilon) \qquad \text{for all } x \in X.$$

COROLLARY 2.4.4 *If (X, d) is a complete metric space, $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is a proper lower semicontinuous function that is bounded below, and $\varepsilon > 0$, $x_\varepsilon \in X$ satisfy*

$$\varphi(x_\varepsilon) \leq \inf_X \varphi + \varepsilon,$$

then we can find $u_\varepsilon \in X$ such that

$$\varphi(u_\varepsilon) \leq \varphi(x_\varepsilon), \quad d(x_\varepsilon, u_\varepsilon) \leq \sqrt{\varepsilon} \quad \text{and} \quad \varphi(u_\varepsilon) \leq \varphi(x) + \sqrt{\varepsilon} d(x, u_\varepsilon) \quad \text{for all } x \in X.$$

Theorem 2.4.1 has some important consequences if we assume more structure on the space X and the function φ.

THEOREM 2.4.5 *If X is a Banach space, $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is a function that is bounded below and Gâteaux differentiable, and $\varepsilon > 0$, $x_\varepsilon \in X$ satisfy*

$$\varphi(x_\varepsilon) \leq \inf_X \varphi + \varepsilon,$$

then there exists $u_\varepsilon \in X$ such that

$$\varphi(u_\varepsilon) \leq \varphi(x_\varepsilon), \quad \|u_\varepsilon - x_\varepsilon\|_X \leq \sqrt{\varepsilon} \qquad \text{and} \qquad \|\varphi'(u_\varepsilon)\|_{X^*} \leq \sqrt{\varepsilon}.$$

PROOF: Apply Corollary 2.4.4 to obtain $u_\varepsilon \in X$ such that

$$\begin{aligned} &\varphi(u_\varepsilon) \leq \varphi(x_\varepsilon), \quad \|u_\varepsilon - x_\varepsilon\|_X \leq \sqrt{\varepsilon} \qquad \text{and} \\ &-\sqrt{\varepsilon}\|x - u_\varepsilon\|_X \leq \varphi(x) - \varphi(u_\varepsilon) \qquad \text{for all } x \in X. \end{aligned} \tag{2.42}$$

From the third inequality in (2.42), taking $x = u_\varepsilon + \lambda h$ with $\lambda > 0$ and $h \in X$ with $\|h\|_X \leq 1$, we have

$$\begin{aligned} & -\sqrt{\varepsilon} \leq \frac{\varphi(u_\varepsilon + \lambda h) - \varphi(u_\varepsilon)}{\lambda}, \\ \Rightarrow \quad & -\sqrt{\varepsilon} \leq \left\langle \varphi'(u_\varepsilon), h \right\rangle_X \qquad \text{for all } h \in \overline{B}_1(0), \\ \Rightarrow \quad & \|\varphi'(u_\varepsilon)\|_{X^*} \leq \sqrt{\varepsilon}. \end{aligned}$$

$\square$

COROLLARY 2.4.6 *If X is a Banach space and $\varphi\colon X \longrightarrow \mathbb{R}$ is a function that is bounded below and Gâteaux differentiable, then for every minimizing sequence $\{x_n\}_{n\geq 1}$ of φ, we can find a minimizing sequence $\{u_n\}_{n\geq 1}$ of φ such that*

$$\varphi(u_n) \leq \varphi(x_n), \quad \|u_n - x_n\|_X \longrightarrow 0 \quad \text{and} \quad \|\varphi'(u_n)\|_{X^*} \longrightarrow 0 \quad \text{as } n \to \infty.$$

The following proposition is useful in critical point theory (see Section 4.1).

PROPOSITION 2.4.7 *If X is a Banach space and $\varphi\colon X \longrightarrow \mathbb{R}$ is a lower semicontinuous function that is bounded below, Gâteaux differentiable, and satisfies the following compactness-type condition*

> *every sequence $\{x_n\}_{n\geq 1} \subseteq X$ that satisfies $|\varphi(x_n)| \leq M$ for some $M > 0$, all $n \geq 1$ and $\|\varphi'(x_n)\|_{X^*} \longrightarrow 0$, it admits a strongly convergent subsequence,*

then we can find $\widehat{x}_0 \in X$ such that $\varphi(\widehat{x}_0) = \inf\limits_X \varphi$ and $\varphi'(\widehat{x}_0) = 0$.

PROOF: Because of Theorem 2.4.5 for every $n \geq 1$, we can find $x_n \in X$ such that

$$\varphi(x_n) \leq \inf_X \varphi + \frac{1}{n} \qquad \text{and} \qquad \|\varphi'(x_n)\|_{X^*} \leq \frac{1}{n}.$$

Then by virtue of the compactness-type condition that φ satisfies, by passing to a subsequence if necessary, we may assume that $x_n \longrightarrow \widehat{x}_0$ in X. Clearly $\varphi(\widehat{x}_0) = \inf\limits_X \varphi$ and from this it follows that $\varphi'(\widehat{x}_0) = 0$. ☐

REMARK 2.4.8 In critical point theory (see Section 4.1), we encounter the compactness-type condition of Proposition 2.4.7 in the context of functions that are C^1 on X with values in $\mathbb{R}$. In such a setting it is called the *Palais–Smale condition* (*PS-condition* for short) and it is a basic tool in the derivation of minimax characterizations of the critical values of functions $\varphi \in C^1(X)$.

PROPOSITION 2.4.9 *If X is a Banach space and $\varphi\colon X \longrightarrow \mathbb{R}$ is a lower semicontinuous function that is bounded below, Gâteaux differentiable, and*

$$\varphi(x) \geq c_1\|x\| - c_2 \qquad \text{for some } c_1, c_2 > 0 \text{ and all } x \in X, \tag{2.43}$$

then $\varphi'(X)$ is dense in $c_1\overline{B}_1^*$ where $\overline{B}_1^* = \{x^* \in X^* : \|x^*\|_{X^*} \leq 1\}$.

PROOF: Let $x^* \in c_1\overline{B}_1^*$ and consider the function $\psi(x) = \varphi(x) - \langle x^*, x\rangle_X$ for all $x \in X$. Clearly ψ is lower semicontinuous, bounded below (because of (2.43)) and Gâteaux differentiable. So by virtue of Theorem 2.4.5 we can find $x_\varepsilon \in X$ such that $\|\psi'(x_\varepsilon)\|_{X^*} \leq \varepsilon$, hence $\|\varphi'(x_\varepsilon) - x^*\|_{X^*} < \varepsilon$, which proves the density of $\varphi'(X)$ in $c_1\overline{B}_1^*$. ☐

COROLLARY 2.4.10 *If X is a Banach space and $\varphi\colon X \longrightarrow \mathbb{R}$ is a lower semicontinuous function that is Gâteaux differentiable and satisfies $\varphi(x) \geq r(\|x\|)$ for all $x \in X$, with $r\colon \mathbb{R}_+ \longrightarrow \mathbb{R}$ a continuous function such that $\big(r(\lambda)/\lambda\big) \longrightarrow +\infty$ as $\lambda \longrightarrow +\infty$, then $\varphi'(X)$ is dense in X^*.*

PROOF: Let $c_1 > 0$ be given. Then we can find $\lambda_0 > 0$ such that for all $\lambda \geq \lambda_0$ we have $r(\lambda) \geq c_1 \lambda$. Hence $\varphi(x) \geq c_1 \|x\|$ for all $x \in X$ with $\|x\|_X > \lambda_0$. On the other hand if $\|x\|_X < \lambda_0$, then

$$\varphi(x) \geq m = \inf[r(\lambda) : 0 \leq \lambda \leq \lambda_0].$$

Let $c_2 = |m|$. Then we have

$$\varphi(x) \geq c_1 \|x\| - c_2$$

and we can apply Proposition 2.4.9 and obtain that $\varphi'(X)$ is dense in $c_1 \overline{B}_1^*$. Because $c_1 > 0$ was arbitrary we conclude that $\varphi'(X)$ is dense in X^*. $\square$

Next we show that the Ekeland variational principle is actually equivalent to some other well-known results of nonlinear analysis. We start with *Caristi's fixed point theorem.*

THEOREM 2.4.11 *If (X, d) is a complete metric space, $\varphi\colon X \longrightarrow \overline{\mathbb{R}}$ is a proper, lower semicontinuous function that is bounded below, and $F : X \longrightarrow 2^X \setminus \{\emptyset\}$ is a multifunction such that*

$$\varphi(y) \leq \varphi(x) - d(x, y) \qquad \textit{for all } x \in X \textit{ and all } y \in F(x), \tag{2.44}$$

then there exists $x_0 \in X$ such that $x_0 \in F(x_0)$.

PROOF: Apply Corollary 2.4.3 with $\varepsilon = 1$, to obtain $x_0 \in X$ such that

$$\varphi(x_0) < \varphi(x) + d(x, x_0) \qquad \text{for all } x \neq x_0. \tag{2.45}$$

We claim that $x_0 \in F(x_0)$. If this is not the case, then for every $y \in F(x_0)$ we have $y \neq x_0$. Then from (2.44) and (2.45) we have $\varphi(y) \leq \varphi(x_0) - d(x_0, y) < \varphi(y)$, a contradiction. $\square$

So we have seen that the Ekeland variational principle in the form of Corollary 2.4.3 implies Theorem 2.4.11 (Caristi's fixed point theorem). Next we show that the opposite is also true.

THEOREM 2.4.12 *Caristi's fixed point theorem (Theorem 2.4.11), implies the Ekeland variational principle in the form of Corollary 2.4.3.*

PROOF: We argue indirectly. Suppose that we cannot find $x_\varepsilon \in X$ satisfying $\varphi(x_\varepsilon) < \varphi(x) + \varepsilon d(x, x_\varepsilon)$ for all $x \neq x_\varepsilon$. Let $F(x) = \big\{y \in X : \varphi(x) \geq \varphi(y) + \varepsilon d(y, x),\ y \neq x\big\}$. Then $F(x) \neq \varnothing$ for all $x \in X$. Invoking Theorem 2.4.11 we can find $x_0 \in X$ such that $x_0 \in F(x_0)$, which is impossible. So the Ekeland variational principle holds. $\square$

REMARK 2.4.13 Banach's fixed point theorem (see Section 3.4), in its existence part, can be deduced from Theorem 2.4.11. Indeed, if $F\colon X \longrightarrow X$ is a contraction, that is, $d\big(F(x), F(y)\big) \leq k\, d(x, y)$ for some $k \in [0, 1)$, then it satisfies (2.44) with $\varphi(x) = \big(1/1 - k\big) d\big(x, F(x)\big)$. However, Banach's fixed point theorem includes more information than the mere existence of a fixed point (a computational scheme to find the fixed point, rate of convergence, error estimates, etc.).

Another result equivalent to the Ekeland variational principle and to the Caristi fixed point theorem, is the so-called *Takahashi variational principle*.

THEOREM 2.4.14 *If (X,d) is a complete metric space, $\varphi\colon X \longrightarrow \overline{\mathbb{R}} = \mathbb{R}\cup\{+\infty\}$ is a proper lower semicontinuous function that is bounded below, and for each $u \in X$ with $\varphi(u) > \inf\limits_X \varphi$ we can find $v \in X$ such that $v \neq u$ and $\varphi(v) + d(u,v) \leq \varphi(u)$, then there exists $x_0 \in X$ such that $\varphi(x_0) = \inf\limits_X \varphi$.*

PROOF: We proceed by contradiction. So suppose that $\inf\limits_X \varphi$ is not attained. We introduce the multifunction $F : X \longrightarrow 2^X$ defined by $F(x) = \{y \in X : \varphi(y) + d(x,y) \leq \varphi(x),\ y \neq x\}$. By hypothesis $F(x) \neq \varnothing$ for all $x \in X$. Invoking Theorem 2.4.11, we can find $x_0 \in F(x_0)$, a contradiction. So $\inf\limits_X \varphi$ is attained. $\square$

Thus we have seen that Caristi's fixed point theorem (Theorem 2.4.11) implies Takahashi's variational principle (Theorem 2.4.14). Next we show that the converse is also true.

THEOREM 2.4.15 *Takahashi's variational principle (Theorem 2.4.14) implies Caristi's fixed point theorem (Theorem 2.4.11).*

PROOF: We proceed by contradiction. So suppose that in Caristi's fixed point theorem (Theorem 2.4.11), the multifunction F has no fixed point, (i.e., $x \notin F(x)$ for all $x \in X$). Thus from the property of the multifunction F, we see that for every $x \in X$ we can find $y \neq x$ such that

$$\varphi(y) + d(x,y) \leq \varphi(x).$$

This is the hypothesis in Theorem 2.4.14. By virtue of that result we can find $x_0 \in X$ such that

$$\varphi(x_0) = \inf_X \varphi.$$

Let $y_0 \in F(x_0)$ be such that $y_0 \neq x_0$ and

$$\varphi(y_0) + d(x_0,y_0) \leq \varphi(x_0).$$

Then we have

$$0 < d(x_0,y_0) \leq \varphi(x_0) - \varphi(y_0) \leq \varphi(y_0) - \varphi(y_0) = 0,$$

a contradiction. So F has a fixed point and Caristi's fixed point theorem (Theorem 2.4.11) holds. $\square$

Another result of nonlinear analysis related to the Ekeland variational principle (in fact equivalent to it), is the so-called *drop theorem*.

THEOREM 2.4.16 *If X is a Banach space, $C \subseteq X$ is a nonempty closed set, $y \in X \setminus C$, $R = d(y,C)$, and $0 < r < R < \varrho$, then there exists $x_0 \in C$ such that $\|y - x_0\| \leq \varrho$ and $D(x_0;y,r) \cap C = \{x_0\}$, where $D(x_0;y,r) = \operatorname{conv}\big(\overline{B}_r(y) \cup \{x_0\}\big)$.*

PROOF: By translating things if necessary, without any loss of generality, we may assume that $y = 0$. Let $S = \overline{B}_r(0) \cap C$. Then S is a closed set in X, hence a complete metric space. Consider the function $\varphi\colon S \longrightarrow \mathbb{R}$ defined by

$$\varphi(x) = \frac{\varrho + r}{R - r}\|x\|_X.$$

By virtue of Corollary 2.4.3 with $\varepsilon = 1$, we can find $x_0 \in S$ such that

$$\varphi(x_0) < \varphi(x) + \|x - x_0\|_X \qquad \text{for all } \ x \in S,\ x \neq x_0. \tag{2.46}$$

Because we have assumed that $y = 0$, we have $\|y - x_0\|_X = \|x_0\|_X \le \varrho$.

Next we show that $D(x_0; 0, r) \cap C = \{x_0\}$. Suppose that this is not the case and we can find $x \neq x_0$ such that $x \in D(x_0; 0, r) \cap C$. Then

$$x \in C \qquad \text{and} \qquad x = (1 - \lambda)x_0 + \lambda u \quad \text{with } \ u \in \overline{B}_r(0) \ \text{ and } \ 0 \le \lambda \le 1. \tag{2.47}$$

Evidently $0 < \lambda < 1$ (recall that $\varrho < R$). From (2.47) we have

$$\begin{aligned} & \|x\|_X \le (1 - \lambda)\|x_0\|_X + \lambda\|u\|_X, \\ \Rightarrow\ & \lambda(R - r) \le \lambda\big(\|x_0\|_X - \|x\|_X\big) \le \|x_0\|_X - \|x\|_X. \end{aligned} \tag{2.48}$$

From (2.46) and (2.47) it follows that

$$\frac{\varrho + r}{R - r}\|x_0\|_X < \frac{\varrho + r}{R - r}\|x\|_X + \|x - x_0\|_X = \frac{\varrho + r}{R - r}\|x\|_X + \lambda\|x_0 - u\|_X. \tag{2.49}$$

Also from (2.48) we have

$$\lambda \le \frac{\|x_0\|_X - \|x\|_X}{R - r}. \tag{2.50}$$

Using (2.50) in (2.49), we obtain

$$\begin{aligned} (\varrho + r)\|x_0\|_X &< (\varrho + r)\|x\|_X + \|x_0 - u\|_X(\|x_0\|_X - \|x\|_X) \\ &\le (\varrho + r)\|x\|_X + (\varrho + r)(\|x_0\|_X - \|x\|_X) \end{aligned}$$

(because $\|x_0\|_X \le \varrho$, $u \in \overline{B}_r(0)$),

$$\Rightarrow\ \|x_0\|_X < \|x\|_X + \|x_0\|_X - \|x\|_X = \|x_0\|_X, \qquad \text{a contradiction.}$$

Therefore we conclude that $D(x_0; 0, r) \cap C = \{x_0\}$ and the proof of the theorem is finished. $\square$

REMARK 2.4.17 The set $D(x_0; y, r)$ is called a *drop*, in view of its evocative geometry.

In the above proof, we saw that the Ekeland variational principle (in the form of Corollary 2.4.3) implies the drop theorem (Theorem 2.4.16). In fact the converse is also true. For a proof of this we refer to Penot [494]. Thus we have the following.

THEOREM 2.4.18 *The drop theorem* (*Theorem 2.4.16*) *implies the Ekeland variational principle* (*Corollary 2.4.3*).

Summarizing we have the following.

THEOREM 2.4.19 *Corollary 2.4.3, Theorem 2.4.11, Theorem 2.4.14 and Theorem 2.4.16 are all equivalent.*

Now we present a general principle on partially ordered sets, which implies all the above equivalent theorems. Recall that on a set X, $\leq$ is a partial order if $x \leq x$ (reflexive) $x \leq y$ and $y \leq x$ imply $x = y$ (antisymmetric) and $x \leq y,\ y \leq z$ imply $x \leq z$ (transitive).

THEOREM 2.4.20 *If $(X, \leq)$ is a partially ordered set, every increasing sequence $\{x_n\}_{n\geq 1} \subseteq X$ has an upper bound in X (i.e. if $x_n \leq x_{n+1}$ for all $n \geq 1$, there exists $\overline{y} \in X$ such that $x_n \leq \overline{y}$ for all $n \geq 1$) and $\varphi\colon X \longrightarrow \mathbb{R}$ is an increasing function that is bounded above, then there exists $\widehat{x}_0 \in X$ such that $\widehat{x}_0 \leq y$ implies $\varphi(\widehat{x}_0) = \varphi(y)$.*

PROOF: Let $x_1 \in X$ be any element. Inductively we produce an increasing sequence $\{x_n\}_{n\geq 1} \subseteq X$. Suppose we have generated the element x_n. We define

$$C_n = \{x \in X : x_n \leq x\} \qquad \text{and} \qquad M_n = \sup_{C_n} \varphi.$$

If for x_n, we have that $x_n \leq y$ implies that $\varphi(x_n) = \varphi(y)$, then clearly we are done. Otherwise $\varphi(x_n) < M_n$ and so we can find $x_{n+1} \in C_n$ such that

$$M_n \leq \varphi(x_{n+1}) + \frac{1}{2}\big(M_n - \varphi(x_n)\big). \tag{2.51}$$

So by induction, we have produced an increasing sequence $\{x_n\}_{n\geq 1} \subseteq X$. By hypothesis we can find $\widehat{x}_0 \in X$ such that

$$x_n \leq \widehat{x}_0 \qquad \text{for all } n \geq 1. \tag{2.52}$$

We claim that $\widehat{x}_0 \in X$ is the desired solution. Suppose that this is not the case. Then we can find $y \in X$ such that $\widehat{x}_0 \leq y$ and $\varphi(\widehat{x}_0) < \varphi(y)$. The sequence $\big\{\varphi(x_n)\big\}_{n\geq 1} \subseteq \mathbb{R}$ is increasing and bounded above by $\varphi(y)$. Therefore it converges. We have

$$\lim_{n\to\infty} \varphi(x_n) \leq \varphi(\widehat{x}_0) \quad \text{(see (2.52))}. \tag{2.53}$$

Because of (2.52), $y \in C_n$ for all $n \geq 1$. So from (2.51) it follows that

$$\begin{aligned} &\varphi(y) \leq M_n \leq 2\varphi(x_{n+1}) - \varphi(x_n) \qquad \text{for all } n \geq 1 \\ \Rightarrow\ &\varphi(y) \leq \varphi(\widehat{x}_0) \qquad \text{(see (2.53))}, \end{aligned}$$

a contradiction to the choice of y. This proves the theorem. ☐

REMARK 2.4.21 Theorem 2.4.20 can have a physical interpretation. We can think of φ as a function measuring the entropy of a system. The theorem guarantees the existence of a state of maximal entropy. To these states correspond stable equilibrium states of the system.

COROLLARY 2.4.22 *If X is a Hausdorff topological space equipped with a partial order $\leq$ and $\psi\colon X \longrightarrow \mathbb{R}$ a function that is bounded below such that*

(i) *For every $x \in X$ the set $\{y \in X : x \leq y\}$ is closed;*
(ii) *$x \leq y$ and $x \neq y$ imply $\psi(y) < \psi(x)$ (strictly decreasing ψ);*

(iii) *Any increasing sequence in X is relatively compact,*

then for each $x \in X$ we can find $\widehat{x}_0 \in X$ such that $x \leq \widehat{x}_0$ and $\widehat{x}_0$ is maximal.

PROOF: Let $\{x_n\}_{n\geq 1} \subseteq X$ be an increasing sequence in X. By virtue of hypothesis (iii) $\{x_n\}_{n\geq 1} \subseteq X$ is relatively compact and so we can find a subsequence $\{x_{n_k}\}_{k\geq 1}$ such that $x_{n_k} \longrightarrow y$ in X. We claim that $x_n \leq y$ for all $n \geq 1$. Indeed given $n \geq 1$, we have $n \leq n_k$ for $k \geq k_n$ and so $x_n \leq x_{n_k}$ for $k \geq k_n$. From this and hypothesis (i) it follows that $y \in C_n$ for all $n \geq 1$; that is, $x_n \leq y$ for all $n \geq 1$. Taking $\varphi = -\psi$, because of hypothesis (ii) we can apply Theorem 2.4.20 starting from $x_1 = x$ and obtain $\widehat{x}_0 \in X$ such that $x \leq \widehat{x}_0$ and $\widehat{x}_0$ is maximal. $\square$

COROLLARY 2.4.23 *Theorem 2.4.20 implies the Ekeland variational principle in the form of Corollary 2.4.3.*

PROOF: Without any loss of generality we take $\varepsilon = 1$ and define a partial order $\leq$ by

$$x \leq u \qquad \text{if and only if } \ \varphi(u) - \varphi(x) \leq -d(x,u). \tag{2.54}$$

For any increasing sequence $\{x_n\}_{n\geq 1} \subseteq X$, $\{\varphi(x_n)\}_{n\geq 1}$ is decreasing and bounded below, so it converges. Therefore once more from (2.54), we infer that $\{x_n\}_{n\geq 1} \subseteq X$ is Cauchy, hence due to the completeness of X it converges to an element in X. Therefore we can apply Corollary 2.4.22 and finish the proof. $\square$

REMARK 2.4.24 Thus Theorem 2.4.20 also implies Caristi's fixed point theorem, Takahashi's variational principle, and the drop theorem.

We conclude this section with a generalization of Theorem 2.4.1, which is used in Section 4.1. For a proof of it we refer to Zhong [626].

THEOREM 2.4.25 *If $h\colon \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a continuous nondecreasing function such that $\int\limits_0^{+\infty} \frac{1}{1+h(r)} \mathrm{d}r = +\infty$, (X,d) is a complete metric space, $x_0 \in X$ is fixed, $\varphi\colon X \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a proper, lower semicontinuous, bounded below function, $\varepsilon > 0$, $\varphi(y) \leq \inf\limits_X \varphi + \varepsilon$ and $\lambda > 0$, then there exists $x_\lambda \in X$ such that*

$$\begin{aligned} &\varphi(x_\lambda) \leq \varphi(y),\ d(x_\lambda, x_0) \leq r_0 + \bar{r} \\ \textit{and} \quad &\varphi(x_\lambda) \leq \varphi(x) + \frac{\varepsilon}{\lambda\big(1 + h\big(d(x_0, x_\lambda)\big)\big)} d(x_\lambda, x) \qquad \textit{for all } x \in X, \end{aligned}$$

where $r_0 = d(x_0, y)$ and $\bar{r} > 0$ such that $\int\limits_0^{r_0+\bar{r}} \frac{1}{1+h(r)}\, \mathrm{d}r \geq \lambda$.

REMARK 2.4.26 If $h \equiv 0$ and $x_0 = y$, then Theorem 2.4.25 reduces to Theorem 2.4.1.

2.5 Calculus of Variations

The calculus of variations deals with the minimization or maximization of functions defined on function spaces. It is as old as calculus itself and has important applications in mechanics and physics.

In this section we present some basic aspects of the theory concerning scalar problems. This means that the unknown functions are curves defined on closed intervals in $\mathbb{R}$ with values in a Banach space. For this reason we make the following definition.

DEFINITION 2.5.1 Let $T = [0, b]$ and X a Banach space. By $C^1(T, X)$ we denote the vector space of all functions $u \in C(T, X)$ which are differentiable on $(0, b)$ and the derivative $u' : (0, b) \longrightarrow X$ is bounded and uniformly continuous.

REMARK 2.5.2 This means that u' admits a unique continuous extension on $T = [0, b]$ and

$$u'(0) = \lim_{t \longrightarrow 0^+} u'(t), \quad u'(b) = \lim_{t \longrightarrow b^-} u'(t).$$

In fact $u'(0)$ is the right derivative of u at $t = 0$ and $u'(b)$ is the left derivative of u at $t = b$.

We furnish $C^1(T, X)$ with the norm

$$\|u\|_{C^1(T,X)} = \max_{t \in T} \|u(t)\|_X + \max_{t \in T} \|u'(t)\|_X. \tag{2.55}$$

It is easy to see that (2.55) indeed defines a norm on $C^1(T, X)$. Also we have the following.

PROPOSITION 2.5.3 *The space $C^1(T, X)$ equipped with the norm $\|\cdot\|_{C^1(T,X)}$ becomes a Banach space.*

Now we introduce the setting of the calculus of variations problem that we study. So let $T = [0, b]$, X be a Banach space, U be a nonempty open set in $\mathbb{R} \times X \times X$, and $L \colon U \longrightarrow \mathbb{R}$ a continuous function, usually called the Lagrangian. A map $u \in C^1(T, X)$ is said to be *admissible*, if for $t \in T$, we have $\big(t, u(t), u'(t)\big) \in U$. We introduce the integral functional $I_L \colon D \subseteq C^1(T, X) \longrightarrow \mathbb{R}$ defined by

$$I_L(u) = \int_0^b L\big(t, u(t), u'(t)\big) \mathrm{d}t$$

for every admissible map $u \in C^1(T, X)$.

PROPOSITION 2.5.4 *The set $D \subseteq C^1(T, X)$ of all admissible maps is an open subset of $C^1(T, X)$.*

PROOF: We show that $D^c \subseteq C^1(T, X)$ is closed. To this end let $\{x_n\}_{n \geq 1} \subseteq D^c$ and suppose that $x_n \longrightarrow x$ in $C^1(T, X)$. Because for all $n \geq 1$ $x_n' \notin D$, we can find $t_n \in T$ such that $\big(t_n, x_n(t_n), x_n'(t_n)\big) \in U^c$. By passing to a subsequence if necessary, we may assume that $t_n \longrightarrow t \in T$. Because $x_n \longrightarrow x$ and $x_n' \longrightarrow x'$ in $C^1(T, X)$, we have that $x_n(t_n) \longrightarrow x(t)$ and $x_n'(t_n) \longrightarrow x'(t)$ in X as $n \to \infty$. So in the limit as $n \to \infty$, we have $\big(t, x(t), x'(t)\big) \in U^c$, hence $x \notin D^c$, which finishes the proof. $\square$

PROPOSITION 2.5.5 *If Y, Z are Banach spaces, $W \subseteq Y$ is a nonempty open set, $\vartheta \colon T \times W \longrightarrow Z$ is a continuous map, and we set $\xi(y) = \int_0^b \vartheta(t,y)\mathrm{d}t$ for all $y \in W$, then*

(a) *$\xi \colon W \longrightarrow Z$ is a continuous map.*
(b) *If in addition for all $t \in T$, $y \longrightarrow \vartheta(t,y)$ is Fréchet differentiable and $(t,y) \longrightarrow \vartheta'_y(t,y)$ is continuous from $T \times W$ into $\mathcal{L}(Y,Z)$, then $\xi \in C^1(W,Z)$ and $\xi'(y) = \int_0^b \vartheta'_y(t,y)\mathrm{d}t$.*

PROOF: (a) Let $\{y_n\}_{n\geq 1} \subseteq W$ and suppose that $y_n \longrightarrow y \in W$. Then $\{y_n, y\}_{n\geq 1} = C$ is compact in Y and $C \subseteq W$. So $h(T \times C)$ is compact in Z. Therefore we can find $M > 0$ such that

$$\|h(t,y_n)\|_Z \leq M \qquad \text{for all } t \in T \text{ and all } n \geq 1.$$

Because $h(t,y_n) \longrightarrow h(t,y)$ in Z for all $t \in T$, from the dominated convergence theorem we have that

$$\xi(y_n) = \int_0^b h(t,y_n)\mathrm{d}t \longrightarrow \int_0^b h(t,y)\mathrm{d}t = \xi(y) \ \text{ in } Z \ \text{ as } n \to \infty,$$
$$\Rightarrow \ \xi \colon W \longrightarrow Z \ \text{ is continuous.}$$

(b) We show that ξ is continuously Gâteaux differentiable. Then it follows that $\xi \in C^1(W,Z)$ (see Proposition 1.1.10). So let $h \in Y$ and $\lambda \neq 0$. We have

$$\frac{\xi(y+\lambda h) - \xi(y)}{\lambda} = \int_0^b \frac{\vartheta(t,y+\lambda h) - \vartheta(t,y)}{\lambda}\mathrm{d}t. \tag{2.56}$$

From Proposition 1.1.6 (the mean value theorem), we know that

$$\|\vartheta(t,y+\lambda h) - \vartheta(t,y)\| \leq \sup_{\mu \in [0,1]} \|\vartheta'_y(t,y+\mu\lambda h)\|_{\mathcal{L}(Y,Z)} \lambda \|h\|_Y. \tag{2.57}$$

Because by hypothesis $(t,y) \longrightarrow \vartheta'_y(t,y)$ is continuous from $T \times W$ into $\mathcal{L}(Y,Z)$, we can find $\overline{M} > 0$ such that

$$\|\vartheta'_y(t,y+\mu\lambda h)\|_{\mathcal{L}(Y,Z)} \leq \overline{M} > 0 \qquad \text{for all } t \in T, \text{ all } \mu \in [0,1] \text{ and all } \lambda \in [-1,1]. \tag{2.58}$$

Therefore

$$\Big\| \frac{\vartheta(t,y+\lambda h) - \vartheta(t,y)}{\lambda} \Big\|_Z \leq \overline{M}\|h\|_Y \qquad \text{for all } t \in T \text{ and all } \lambda \in [-1,1] \quad (\text{see (2.57) and (2.58)}). \tag{2.59}$$

Recall that

$$\frac{\vartheta(t,y+\lambda h) - \vartheta(t,y)}{\lambda} \longrightarrow \vartheta'_y(t,y)h \qquad \text{for all } t \in T \text{ as } \lambda \longrightarrow 0. \tag{2.60}$$

Because of (2.59), (2.60), and the dominated convergence theorem, if we pass to the limit as $\lambda \longrightarrow 0$, we obtain

$$\xi'(y)h = \int_0^b \vartheta_y'(t,y)h \,\mathrm{d}t \qquad \text{for all } \ h \in Y.$$

Therefore by virtue of part (a), $\xi \in C^1(W,Z)$ and

$$\xi'(y) = \int_0^b \vartheta_y'(t,y) \,\mathrm{d}t \qquad \text{for all } \ y \in Y.$$

$\square$

We use this proposition to establish the continuous Fréchet differentiability of the integral functional I_L and determine its derivative in terms of the partial derivatives of the Lagrangian L.

THEOREM 2.5.6 *If the Lagrangian $L\colon U \longrightarrow \mathbb{R}$ is continuously Fréchet differentiable, then $I_L : D \subseteq C^1(T,X) \longrightarrow \mathbb{R}$ is continuously Fréchet differentiable too and*

$$\begin{aligned}\big\langle I_L'(u), v\big\rangle_{C^1(T,X)} = &\int_0^b \big\langle L_2'\big(t,u(t),u'(t)\big), v(t)\big\rangle_X \,\mathrm{d}t \\ &+ \int_0^b \big\langle L_3'\big(t,u(t),u'(t)\big), v'(t)\big\rangle_X \,\mathrm{d}t\end{aligned}$$

for all $v \in C^1(T,X)$. Here by $\langle \cdot,\cdot\rangle_{C^1(T,X)}$ we denote the duality brackets for the pair $\big(C^1(T,X), C^1(T,X)^*\big)$, by $\langle\cdot,\cdot\rangle_X$ the duality brackets for the pair (X,X^*), and L_k' $k=2,3$ the partial derivative of $L(t,x,y)$ with respect to the second and third variables, respectively.

PROOF: From Proposition 2.5.4 we know that the set $D \subseteq C^1(T,X)$ of admissible maps is open. We consider the map $\vartheta\colon T \times D \longrightarrow \mathbb{R}$ defined by

$$\vartheta(t,u) = L\big(t,u(t),u'(t)\big) \qquad \text{for all } \ (t,u) \in T \times D.$$

Let $\vartheta_1\colon T \times D \longrightarrow U$ and $\vartheta_2\colon U \longrightarrow \mathbb{R}$ be defined by

$$\vartheta_1(t,u) = \big(t,u(t),u'(t)\big) \qquad \text{and} \qquad \vartheta_2(t,x,y) = L(t,x,y)$$

for all $(t,u) \in T \times D$ and all $(t,x,y) \in U$. Clearly ϑ_1, ϑ_2 are continuous and $\vartheta = \vartheta_2 \circ \vartheta_1$. So ϑ is continuous too.

Moreover, because for $t \in T$ fixed the maps $u \longrightarrow u(t)$ and $u \longrightarrow u'(t)$ are continuous linear, we see that for every $t \in T$, $\vartheta(t,\cdot)$ is continuously Fréchet differentiable. Then by virtue of Propositions 1.1.14 and 1.1.17, we have

$$\big\langle \vartheta_2'(t,u), v\big\rangle_{C^1(T,X)} = \big\langle L_2'\big(t,u(t),u'(t)\big), v(t)\big\rangle_X + \big\langle L_3'\big(t,u(t),u'(t)\big), v'(t)\big\rangle_X$$

for all $v \in C^1(T,X)$. It follows that $(t,u) \longrightarrow \vartheta_2'(t,u)$ is continuous from $T \times D$ into $C^1(T,X)^*$ and so we can apply Proposition 2.5.5 and conclude that

$$\begin{aligned}\big\langle I_L'(u), v\big\rangle_{C^1(T,X)} = &\int_0^b \big\langle L_2'\big(t,u(t),u'(t)\big), v(t)\big\rangle_X \,\mathrm{d}t \\ &+ \int_0^b \big\langle L_3'\big(t,u(t),u'(t)\big), v'(t)\big\rangle_X \,\mathrm{d}t\end{aligned}$$

for all $v \in C^1(T, X)$. $\square$

We consider the problem with fixed endpoints. Namely we require that all admissible maps satisfy $u(0)=x_0$ and $u(b) = x_b$, where $x_0, x_b \in X$ are given. So let

$$C^1_{x_0,x_b}(T, X) = \big\{u \in C^1(T, X) : u(0) = x_0,\ u(b) = x_b\big\}.$$

This is a closed affine subspace of $C^1(T, X)$ produced by translation of the closed subspace $C^1_0(T, X) = C^1_{0,0}(T, X)$.

Now let us formally introduce the concept of the local (relative) extremal point of a functional.

DEFINITION 2.5.7 Let Y be a Banach space, $C \subseteq Y$ a nonempty set, and $\varphi\colon C \longrightarrow \mathbb{R}$. We say that $y_0 \in C$ is a local (relative) extremum of φ, if we can find V a neighborhood of y_0 in Y such that one of the following two holds.

(a) $\varphi(y_0) \leq \varphi(y)$ for all $y \in V \cap C$ (*local minimum*).
(b) $\varphi(y_0) \geq \varphi(y)$ for all $y \in V \cap C$ (*local maximum*).

If the inequalities in (a) and (b) are strict for $y \neq y_0$, then we speak of a strict local extremum (strict local minimum and strict local maximum).

If the inequalities in (a) and (b) are true for all $y \in C$, then we speak of a (global) extremum on C $\big($(global) minimum on C and (global) maximum on $C\big)$.

In the next proposition we produce a necessary condition for I_L to admit a local extremum on $C^1_{x_0,x_b}(T, X)$. In what follows we assume that $D \cap C^1_{x_0,x_b}(T, X)$ is nonempty.

PROPOSITION 2.5.8 *If $L\colon U \longrightarrow \mathbb{R}$ is a continuously Fréchet differentiable Lagrangian and I_L admits a local relative extremum on $C^1_{x_0,x_b}(T, X)$ at the map $u \in C^1_{x_0,x_b}(T, X)$, then for every $v \in D \cap C^1_0(T, X)$ we have*

$$\begin{aligned}\big\langle I_L'(u), v\big\rangle_{C^1(T,X)} &= \int_0^b \big\langle L_2'\big(t, u(t), u'(t)\big), v(t)\big\rangle_X \,\mathrm{d}t \\ &\quad + \int_0^b \big\langle L_3'\big(t, u(t), u'(t)\big), v'(t)\big\rangle_X \,\mathrm{d}t. \end{aligned} \tag{2.61}$$

PROOF: From Proposition 2.5.4 we know that D is an open subset of $C^1(T, X)$. Hence $D \cap C^1_{x_0,x_b}(T, X)$ is an open subset of the closed affine subspace $C^1_{x_0,x_b}(T, X)$. By Theorem 2.5.6 $I_L : D \longrightarrow \mathbb{R}$ is continuously Fréchet differentiable, thus we have that its restriction on $C^1_{x_0,x_b}(T, X)$ is also continuously Fréchet differentiable. Because I_L attains its local relative minimum or local relative maximum at a map $u \in C^1_{x_0,x_b}(T, X)$, we have

$$\Big(I_L\Big|_{C^1_{x_0,x_b}(T,X)}\Big)'(u) = 0.$$

But

$$\Big(I_L\Big|_{C^1_{x_0,x_b}(T,X)}\Big)'(u) = I_L'(u)\Big|_{C^1_0(T,X)}$$

and so (2.61) follows. $\square$

Equation (2.61) is not yet in a convenient form. To achieve this we need some auxiliary results that are basic in the theory of calculus of variations. The first auxiliary result is usually known in the literature as the *Lagrange lemma*.

LEMMA 2.5.9 *If $\xi \in C(T, X^*)$, then $\xi \equiv 0$ if and only if $\int_0^b \langle \xi(t), u(t)\rangle_X \, dt = 0$ for all $u \in C_0^1(T, X)$.*

PROOF: $\Rightarrow$: Obvious.

$\Leftarrow$: Suppose that ξ is not identically zero. Then we can find $t_0 \in (0, b)$ such that $\xi(t_0) \neq 0$. Hence we can find $x \in X$ such that $\langle \xi(t_0), x\rangle_X \neq 0$ and without any loss of generality we may assume that $\langle \xi(t_0), x\rangle_X > 0$. The function $t \longrightarrow \langle \xi(t), x\rangle_X$ is continuous on T and so we can find $\delta > 0$ that $[t_0 - \delta, t_0 + \delta] \subseteq (0, b)$ and $\langle \xi(t), x\rangle_X > 0$ for all $t \in [t_0 - \delta, t_0 + \delta]$. Consider a function $\vartheta \in C^1(T)$ such that

$$\vartheta \geq 0 \qquad \text{and} \qquad \vartheta(t) = 0 \qquad \text{for all } \ t \notin [t_0 - \delta, t_0 + \delta].$$

A possible such function (known as a cut-off function) is

$$\vartheta(t) = \begin{cases} 0 & \text{if } \ 0 \leq t \leq t_0 - \delta \\ \exp\left(-\frac{1}{\delta^2 - (t - t_0)^2}\right) & \text{if } \ t_0 - \delta \leq t \leq t_0 + \delta. \\ 0 & \text{if } \ t_0 + \delta \leq t \leq b \end{cases}$$

Then let $u_0(t) = \vartheta(t)x$. Evidently $u_0 \in C_0^1(T, X)$ and the function $t \longrightarrow \langle \xi(t_0), u_0(t)\rangle_X$ is continuous on T, nonnegative, and $\langle \xi(t_0), u_0(t)\rangle_X > 0$. Therefore it follows that

$$\int_0^b \langle \xi(t), u_0(t)\rangle_X \, dt > 0,$$

a contradiction to the hypothesis of the lemma. $\square$

The second auxiliary result is known as *Du Bois–Reymond's lemma*.

LEMMA 2.5.10 *If $\xi \in C(T, X^*)$, then ξ is constant if and only if for all $u \in C(T, X)$ with mean value zero (i.e., $\int_0^b u(t) dt = 0$) we have $\int_0^b \langle \xi(t), u(t)\rangle_X \, dt = 0$.*

PROOF: $\Rightarrow$: Suppose that $\xi(t) = x^* \in X^*$ for all $t \in T$. Also assume that $u \in C(T, X)$ has mean value zero. Then from the properties of the Bochner integral, we have

$$\int_0^b \langle \xi(t), u(t)\rangle_X \, dt = \int_0^b \langle x^*, u(t)\rangle_X \, dt = \left\langle x^*, \int_0^b u(t) dt \right\rangle_X = 0.$$

$\Leftarrow$: Suppose that ξ is not constant. Because $\xi \in C(T, X)$ we can find $0 < t_1 < t_2 < b$ such that $\xi(t_1) \neq \xi(t_2)$. So there exists $x \in X$ such that $\langle \xi(t_1), x\rangle_X \neq \langle \xi(t_2), x\rangle_X$. Without any loss of generality we may assume that

$$\langle \xi(t_1), x\rangle_X < \langle \xi(t_2), x\rangle_X \,.$$

By virtue of the continuity of the function $t \longrightarrow \langle \xi(t), x\rangle_X$ on T, we can find $\delta > 0$ such that $0 < t_1 - \delta < t_1 + \delta < t_2 - \delta < t_2 + \delta < b$ and

$$\max\left\{ \langle \xi(t), x\rangle_X : t \in [t_1-\delta, t_1+\delta]\right\} < \min\left\{ \langle \xi(t), x\rangle_X : t \in [t_2-\delta, t_2+\delta]\right\}.$$

From this inequality it follows that

$$\langle \xi(t_2+s) - \xi(t_1+s), x\rangle_X > 0 \qquad \text{for all } \ s \in [-\delta, \delta].$$

We consider a continuous cut-off function ϑ for the interval $[-\delta, \delta]$; that is, $\vartheta \geq 0$, $\vartheta(s) > 0$ for $s \in (-\delta, \delta)$ and $\vartheta(s) = 0$ for $s \notin (-\delta, \delta)$. A possible such function is given by

$$\vartheta(s) = \begin{cases} \frac{\delta - |s|}{\delta} & \text{if } \ |s| \leq \delta. \\ 0 & \text{if } \ |s| > \delta \end{cases}$$

Let $\eta(t) = \int_0^b \big(\vartheta(s-t_2) - \vartheta(s-t_1)\big) \mathrm{d}s$ for all $t \in T$ and set $u(t) = \eta(t)x$. Evidently $u \in C^1(T, X)$ and $u(0) = u(b) = 0$. Then

$$\begin{aligned} \int_0^b \langle \xi(t), u'(t)\rangle_X \, \mathrm{d}t &= \int_0^b \big(\vartheta(s-t_2) - \vartheta(s-t_1)\big) \langle \xi(t), x\rangle_X \, \mathrm{d}t \\ &= \int_\delta^\delta \vartheta(s) \langle \xi(t_2+s) - \xi(t_1+s), x\rangle_X \, \mathrm{d}s, \end{aligned}$$

a contradiction to the hypothesis. $\square$

The previous two lemmata lead to the third auxiliary result which provides the tools to simplify the extremality equation (2.61).

LEMMA 2.5.11 *If $\xi_1, \xi_2 \in C(T, X^*)$, then $\int_0^b \left[\langle \xi_1(t), u(t)\rangle_X + \langle \xi_2(t), u'(t)\rangle_X \right] \mathrm{d}t = 0$ for all $u \in C_0^1(T, X)$ if and only if ξ_2 is differentiable on T and $\xi_2'(t) = \xi_1(t)$ for all $t \in T$.*

PROOF: Let $\beta(t) = \int_0^t \xi_1(s)\mathrm{d}s$ for all $t \in T$. Then $\beta \in C^1(T, X^*)$ and $\beta'(t) = \xi_1(t)$ for all $t \in T$. Let $u \in C_0^1(T, X)$. We have

$$\langle \xi_1(t), u(t)\rangle_X = \frac{\mathrm{d}}{\mathrm{d}t} \langle \beta(t), u(t)\rangle_X - \left\langle \beta(t), u'(t)\right\rangle_X \qquad \text{for all } \ t \in T,$$

$$\begin{aligned} \Rightarrow \int_0^b \langle \xi_1(t), u(t)\rangle_X \, \mathrm{d}t &= \langle \beta(b), u(b)\rangle_X - \langle \beta(0), u(0)\rangle_X - \int_0^b \left\langle \beta(t), u'(t)\right\rangle_X \mathrm{d}t \\ &= -\int_0^b \left\langle \beta(t), u'(t)\right\rangle_X \mathrm{d}t \quad (\text{because } \ u(0) = u(b) = 0) \end{aligned}$$

$$\Rightarrow \int_0^b \left[\langle \xi_1(t), u(t)\rangle_X + \left\langle \xi_2(t), u'(t)\right\rangle_X \right] \mathrm{d}t = \int_0^b \left\langle \xi_2(t) - \beta(t), u'(t)\right\rangle_X \mathrm{d}t.$$

Then by virtue of Lemma 2.5.10, we have

$$\int_0^b \left\langle \xi_2(t) - \beta(t), u'(t)\right\rangle_X \mathrm{d}t = 0 \quad \text{for all } \ u \in C_0^1(T, X)$$

if and only if

$$\xi_2(t) - \beta(t) = x_0^* \in X^* \qquad \text{for all } \ t \in T.$$

Therefore $\xi_2 \in C^1(T, X)$ and $\xi_2'(t) = \xi_1(t)$ for all $t \in T$. $\square$

Using this lemma we can simplify equation (2.61).

THEOREM 2.5.12 *If $L\colon U \longrightarrow \mathbb{R}$ is a continuously differentiable Lagrangian and I_L admits a local relative extremum on $C^1_{x_0,x_b}(T,X)$ at $u \in C^1_{x_0,x_b}(T,X)$, then*

$$\frac{\mathrm{d}}{\mathrm{d}t} L_3'\big(t,u(t),u'(t)\big) = L_2'\big(t,u(t),u'(t)\big) \qquad \textit{for all } t \in T. \tag{2.62}$$

REMARK 2.5.13 Equation (2.62) is known as *Euler's equation* (in Lagrange form). The functions $u \in C^1_{x_0,x_b}(T,X)$ along which (2.62) is valid are called *extremals*.

Next we present some special cases where the Euler equation has easily determined integrals.

COROLLARY 2.5.14 *If the Lagrangian function $L(t,x,y)$ is actually independent of the third variable $y \in X$, then a necessary condition for the extremality of the map $u \in C^1_{x_0,x_b}(T,X)$ is given by*

$$L_2'\big(t,u(t)\big) = 0 \qquad \textit{for all } t \in T.$$

COROLLARY 2.5.15 *If the Lagrangian function $L(t,x,y)$ is actually independent of the second variable $x \in X$, then the Euler equation (2.62) admits the following solution*

$$p(t) = L_3'\big(t,u'(t)\big) = p_0^* \in X^* \qquad \textit{for all } t \in T.$$

REMARK 2.5.16 In mechanics $p(t)$ is the momentum function and Corollary 2.5.15 says that along an extremal the momentum is constant.

COROLLARY 2.5.17 *If the Lagrangian function $L(t,x,y)$ actually does not depend on the time variable $t \in T$ and the extremal $u \in C^2(T,X) \cap C^1_{x_0,x_b}(T,X)$, then the Euler equation (2.62) admits an energy integral*

$$\begin{aligned} H(t) &= \big\langle p(t),u'(t)\big\rangle_X - L\big(u(t),u'(t)\big) \\ &= \big\langle L_2'\big(u(t),u'(t)\big),u'(t)\big\rangle_X - L\big(u(t),u'(t)\big) = H_0 \in \mathbb{R} \qquad \textit{for all } t \in T. \end{aligned}$$

PROOF: From the chain rule we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} H(t) = &\left\langle \frac{\mathrm{d}}{\mathrm{d}t} L_2'\big(u(t),u'(t)\big),u'(t)\right\rangle + \big\langle L_2'\big(u(t),u'(t)\big),u''(t)\big\rangle_X \\ &- \big\langle L_1'\big(u(t),u'(t)\big),u'(t)\big\rangle_X - \big\langle L_2'\big(u(t),u'(t)\big),u''(t)\big\rangle_X = 0 \end{aligned}$$

(see (2.62)). Hence $H(t) = H_0 \in \mathbb{R}$ for all $t \in T$. ☐

REMARK 2.5.18 The above corollary says that along an extremal the energy is constant (conservation of energy).

To be able to analyze further the Euler equation (2.62), we need to introduce second-order derivatives.

DEFINITION 2.5.19 Let X, Y be Banach spaces, $U \subseteq X$ nonempty open, $x_0 \in U$, and $\varphi\colon U \longrightarrow Y$ a Fréchet differentiable map. If the Fréchet derivative $x \longrightarrow \varphi'(x)$ from U into $\mathcal{L}(X,Y)$ with the operator norm topology is differentiable at x_0, then φ is said to be *twice differentiable* at x_0 and this second-order derivative is denoted by $\varphi''(x_0)$.

REMARK 2.5.20 According to this definition $\varphi''(x_0) \in \mathcal{L}\big(X, \mathcal{L}(X,Y)\big)$.

DEFINITION 2.5.21 Let X_1, X_2 and Y be Banach spaces. A map $L\colon X = X_1 \times X_2 \longrightarrow Y$ is said to be *bilinear* if $x_1 \longrightarrow L(x_1,x_2)$ from X_1 into Y and $x_2 \longrightarrow L(x_1,x_2)$ from X_2 into Y are both linear. We say that L is *continuous*, if there exists $M \geq 0$ such that

$$\|L(x)\|_Y \leq M\|x_1\|_X\|x_2\|_X \qquad \text{for all } x = (x_1,x_2) \in X = X_1 \times X_2. \tag{2.63}$$

The infimum of all $M \geq 0$ for which (2.63) is true is the norm of L. Equipped with this norm the space of continuous bilinear maps $L\colon X_1 \times X_2 \longrightarrow Y$ is a Banach space denoted by $\mathcal{L}_2(X_1 \times X_2; Y)$.

REMARK 2.5.22 Note that on $X = X_1 \times X_2$, we consider the norm $\|x\|_X = \max\{\|x_1\|_{X_1}, \|x_2\|_{X_2}\}$ for all $x = (x_1,x_2) \in X = X_1 \times X_2$; then $\|L\|_{\mathcal{L}_2} = \sup\big\{\|L(x)\|_Y : x \in X,\ \|x\|_X \leq 1\big\}$.

PROPOSITION 2.5.23 *If X_1, X_2, and Y are Banach spaces, then $\mathcal{L}_2(X_1 \times X_2; Y)$, $\mathcal{L}\big(X_1, \mathcal{L}(X_2,Y)\big)$, and $\mathcal{L}\big(X_2, \mathcal{L}(X_1,Y)\big)$ are isometrically isomorphic Banach spaces.*

PROOF: Let $L \in \mathcal{L}_2(X_1 \times X_2; Y)$ and $x_1 \in X_1$. We consider the map $x_2 \longrightarrow K(x_1)(x_2) = L(x_1,x_2)$ from X_2 into Y. We have

$$\begin{aligned} &\|K(x_1)(x_2)\|_Y \leq \|L\|_{\mathcal{L}_2}\|x_1\|_{X_1}\|x_2\|_{X_2}, \\ \Rightarrow\ & K(x_1) \in \mathcal{L}(X_2,Y) \quad \text{and} \quad \|K(x_1)\|_{\mathcal{L}} \leq \|L\|_{\mathcal{L}_2}\|x_1\|_X. \end{aligned} \tag{2.64}$$

Also the map $x_1 \longrightarrow K(x_1)$ is linear and (2.64) implies that it is also continuous and satisfies

$$\|K\|_{\mathcal{L}\left(X_1,\mathcal{L}(X_2,Y)\right)} \leq \|L\|_{\mathcal{L}_2}. \tag{2.65}$$

On the other hand, if $K \in \mathcal{L}\big(X_1, \mathcal{L}(X_2,Y)\big)$, for all $x_1 \in X_1$ we set

$$L(x_1,x_2) = K(x_1)x_2.$$

Evidently $L\colon X_1 \times X_2 \longrightarrow Y$ is a bilinear map and

$$\begin{aligned} &\|L(x_1,x_2)\|_Y \leq \|K(x_1)\|_{\mathcal{L}}\|x_2\|_{X_2} \leq \|K\|_{\mathcal{L}\left(X_1,\mathcal{L}(X_2,Y)\right)}\|x_1\|_{X_1}\|x_2\|_{X_2}, \\ \Rightarrow\ & L \in \mathcal{L}_2(X_1 \times X_2; Y) \quad \text{and} \quad \|L\|_{\mathcal{L}_2} \leq \|K\|_{\mathcal{L}\left(X_1,\mathcal{L}(X_2,Y)\right)}. \end{aligned} \tag{2.66}$$

From (2.65) and (2.66) we conclude that $\mathcal{L}_2(X_1 \times X_2; Y)$ and $\mathcal{L}\big(X_1, \mathcal{L}(X_2,Y)\big)$ are isometrically isomorphic and similarly for $\mathcal{L}\big(X_2, \mathcal{L}(X_1,Y)\big)$. $\square$

By virtue of Proposition 2.5.23, if $\varphi\colon U \longrightarrow Y$ is a twice differentiable map and $x_0 \in U$, then

$$\varphi''(x_0) \in \mathcal{L}_2(X \times X; Y).$$

In fact we can say more.

PROPOSITION 2.5.24 *If X, Y are Banach spaces, $U \subseteq X$ is nonempty open, $x_0 \in X$, and $\varphi\colon U \longrightarrow X$ is a map which is twice differentiable at x_0 (see Definition 2.5.19), then $\varphi''(x_0) \in \mathcal{L}_2(X \times X; Y)$ is symmetric; that is, for all $x, u \in X$ we have*

$$\varphi''(x_0)(x,u) = \varphi''(x_0)(u,x).$$

PROOF: Let $r > 0$ be such that $B_r(x_0) \subseteq U$ and consider the map $\psi\colon [0,1] \longrightarrow Y$ defined by

$$\psi(t) = \varphi(x_0 + tx + u) - \varphi(x_0 + tx) \qquad \text{for } \|x\|_X, \|u\|_Y \leq r.$$

From the mean value theorem (see Proposition 1.1.6), we have

$$\|\psi(1) - \psi(0) - \psi'(0)\|_X \leq \sup_{t \in [0,1]} \|\psi'(t) - \psi'(0)\|_X. \tag{2.67}$$

Also from the chain rule (see Proposition 1.1.14), we have

$$\begin{aligned}\psi'(t) &= \big(\varphi'(x_0 + tx + u) - \varphi'(x_0 + tx)\big)x \\ &= \big(\varphi'(x_0 + tx + u) - \varphi(x_0) - \varphi''(x_0)tx\big)x \\ &\quad - \big(\varphi'(x_0 + tx) - \varphi(x_0) - \varphi''(x_0)tx\big)x.\end{aligned}$$

Given $\varepsilon > 0$, we can find $\delta > 0$ such that if $\|x\|_X, \|u\|_X \leq \delta$ and $0 \leq t \leq 1$, then

$$\begin{aligned} &\|\varphi'(x_0 + tx + u) - \varphi'(x_0) - \varphi''(x_0)(tx + u)\|_{\mathcal{L}} \leq \varepsilon(\|x\|_X + \|u\|_X) \\ \text{and} \quad &\|\varphi'(x_0 + tx) - \varphi'(x_0) - \varphi''(x_0)(tx)\|_{\mathcal{L}} \leq \varepsilon\|x\|_X.\end{aligned}$$

These imply

$$\|\psi'(t) - \big(\varphi''(x_0)u\big)x\|_Y \leq 2\varepsilon(\|x\|_X + \|u\|_X) \tag{2.68}$$

$$\begin{aligned}\text{and} \quad \|\psi(1) - \psi(0) - \big(\varphi''(x_0)u\big)x\|_Y &\leq \|\psi(1) - \psi(0) - \varphi'(0)\|_Y \\ &\quad + \|\varphi'(0) - \big(\varphi''(x_0)u\big)x\|_Y \\ &\leq 6\varepsilon(\|x\|_X + \|u\|_X)\|x\|_X \end{aligned} \tag{2.69}$$

(see (2.67) and (2.68)).

But $\psi(1) - \psi(0) = \varphi(x_0 + tx + u) - \varphi(x_0 + x) - \varphi(x_0 + u) + \varphi(x_0)$ is symmetric in $x, u \in X$ and so in (2.68) we may interchange x and u and obtain

$$\|\big(\varphi''(x_0)u\big)x - \big(\varphi''(x_0)x\big)u\|_Y \leq 6\varepsilon(\|x\|_X + \|u\|_X)^2 \qquad \text{for all } \|x\|_X, \|u\|_X \leq \delta. \tag{2.70}$$

If we replace x, u by $\lambda x, \lambda u$ with $\lambda > 0$, then both sides of (2.70) are multiplied with λ^2. So (2.70) holds also for $\|x\|_X = \|u\|_X = 1$. Hence

$$\|\varphi''(x_0)(u,x) - \varphi''(x_0)(x,u)\|_Y \leq 24\varepsilon \qquad \text{for all } \|x\|_X = \|u\|_X = 1. \tag{2.71}$$

Because $\varepsilon > 0$ was arbitrary, from (2.71) we conclude that

$$\varphi''(x_0)(x,u) = \varphi''(x_0)(u,x) \qquad \text{for all } x, u \in X.$$

$\square$

Now that we have the second-order derivative at our disposal, we can further analyze Euler's equation (2.62). Assume that the extremal u belongs in $C^2(T, X)$.

In order to keep track with respect to which variable we differentiate at each step, we consider the variables (t, x, y) of the Lagrangian and by L'_t (resp., L'_x, L'_y) we denote the partial derivative of L with respect to t (resp., with respect to x, y). Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} L'_3\big(t, u(t), u'(t)\big) &= \frac{\mathrm{d}}{\mathrm{d}t} L'_y\big(t, u(t), u'(t)\big)\\ &= L'_{yt}\big(t, u(t), u'(t)\big) + L'_{yx}\big(t, u(t), u'(t)\big)u'(t)\\ &\quad + L'_{yy}\big(t, u(t), u'(t)\big)u''(t) \text{ in } X^*.\end{aligned}$$

Therefore the Euler equation can be written as follows.

$$\begin{aligned}&L'_{yy}\big(t, u(t), u'(t)\big)u''(t) + L'_{yx}\big(t, u(t), u'(t)\big)u'(t)\\ &\quad + L'_{yt}\big(t, u(t), u'(t)\big) - L'_x\big(t, u(t), u'(t)\big) = 0 \qquad \text{in } X^*.\end{aligned} \tag{2.72}$$

This is a second-order differential equation with the function u unknown. Of course (2.72) is not in canonical form, because it is not solved for u''. In the sequel by imposing a suitable condition on $L'_{yy}\big(t, u(t), u'(t)\big) \in \mathcal{L}(X, X^*)$, we are able to put (2.72) in canonical form.

As this point we focus on a difficulty that we encounter. Equation (2.72) was derived under the assumption that $u \in C^2(T, X)$. On the other hand the Euler equation (2.62) in Theorem 2.5.12 was obtained under the assumption that the extremal u belongs in $C^1(T, \mathbb{R}^N)$. In what follows we show how we can overcome this difficulty.

PROPOSITION 2.5.25 *If $L\colon U \longrightarrow \mathbb{R}$ is a twice differentiable Lagrangian function (i.e., $L \in C^2(U)$), $(t_0, x_0, y_0) \in U$ with $0 < t_0 < b$, and $L''_{yy}(t_0, x_0, y_0)$ is an isomorphism from X into X^*, then*

(a) *There exists a solution u of the Euler equation (2.62) defined on an open interval I_0 containing t_0 such that $u(t_0) = x_0, u'(t_0) = y_0$ and u is C^2.*
(b) *Every other solution $\widehat{u}$ of the Euler equation (2.62) defined on an open interval $\widehat{I}$ containing t_0 and satisfying $\widehat{u}(t_0) = x_0$, $\widehat{u}'(t_0) = y_0$ coincides with u on an open interval containing t_0 and $\widehat{u}$ is C^2 on that interval.*

PROOF: Consider the map $\psi\colon U \longrightarrow \mathbb{R} \times X \times X$ defined by

$$\psi(t, x, y) = \big(t, x, L'_y(t, x, y)\big). \tag{2.73}$$

Because by hypothesis $L'_{yy}(t_0, x_0, y_0)$ is an isomorphism from X onto X^*, we can apply the inverse theorem on the function L'_y(see Theorem 1.1.25). So according to that theorem, we can find a neighborhood V of $(t_0, x_0, y_0) \in \mathbb{R} \times X \times X$ and a neighborhood W of $\big(t_0, x_0, p_0 = L'_y(t_0, x_0, y_0)\big) \in \mathbb{R} \times X \times X^*$, such that the function ψ defined by (2.73) is a C^1-diffeomorphism from V onto W. The inverse function ψ^{-1} (also a C^1-function) is of the form $(t, x, p) \longrightarrow \big(t, x, y = g(t, x, p)\big)$. We consider the following system on $X \times X^*$

$$\left\{\begin{aligned} x'(t) &= g\big(t, x(t), p(t)\big)\\ p'(t) &= L'_x\big(t, x(t), g\big(t, x(t), p(t)\big)\big)\end{aligned}\right\}. \tag{2.74}$$

Note that (2.74) is in canonical form (it is solved with respect to the higher-order derivatives x', p' of the unknown functions x, p). The vector field of (2.74),

$$(t,x,p) \longrightarrow \big(g(t,x,p), L'_x\big(t,x,g(t,x,p)\big)\big),$$

is a C^1-function from W into $X \times X^*$ and the triple (t_0, x_0, p_0) corresponds to the Cauchy data of (2.74). So by the basic existence and uniqueness theorem for initial value problems in Banach spaces, we can find a unique maximal solution $t \longrightarrow \big(x(t), p(t)\big)$ of (2.74) such that $x(t_0) = x_0$ and $p(t_0) = p_0$. Let I be the open interval containing t_0 on which the solution is defined. We know that on this interval the solution is C^2. Let $u(t) = x(t)$. Then $u \in C^2(I, X)$ satisfies the Euler equation (2.62) and the initial conditions $u(t_0) = x_0,\ u'(t_0) = y_0$.

Next let $\widehat{u}\colon \widehat{I} \longrightarrow X$ be another solution defined on an interval $\widehat{I}$ containing t_0 satisfying $\widehat{u}(t_0) = x_0,\ \widehat{u}'(t_0) = y_0$. For t close to t_0, $\big(t, \widehat{u}(t), \widehat{u}'(t)\big) \in V$. So we can consider $\big(\psi(t), \widehat{u}(t), \widehat{u}'(t)\big)$ (see (2.73)). Then the function $t \longrightarrow \big(\widehat{u}(t), \widehat{u}'(t)\big)$ solves (2.74) and satisfies the Cauchy data (t_0, x_0, y_0). By virtue of the uniqueness of the solution, we conclude that for all $t \in \widehat{I} \cap I$ we have $\widehat{u}(t) = u(t)$. $\square$

REMARK 2.5.26 This proposition shows that in order to hope to be able to transform (2.72) to a canonical first-order system we need to assume that the Lagrangian function L is C^2.

Before proceeding to the Hamiltonian formulation of the necessary conditions for extremality, let us illustrate the use of Euler's equation through some characteristic examples.

EXAMPLE 2.5.27 (a) Let $L(y) = \sqrt{1+y^2}$ and consider the problem $\inf\big[I_L(u)\colon u \in C^1([0,1]),\ u(0) = x_0,\ u(1) = x_1\big]$. Then the Euler equation is $(\mathrm{d}/\mathrm{d}t)(\partial/\partial y)\sqrt{1+u'(t)^2} = 0,\ u(0) = x_0,\ u(1) = x_1$, which gives $u(t) = (x_1 - x_0)t + x_0$ for all $t \in T$. This is a global minimum of $I_L(\cdot)$ over $C^1_{x_0,x_1}([0,1])$ and simply says that the curve with the minimum length joining x_0 and x_1 is the straight line connecting the two points.

(b) Let $L(y)=y^3$ and consider the problem $\inf[I_L(u) : u \in C^1([0,1]), u(0)=0,\ u(1)=1]$. Then the Euler equation is $\big(u'(t)^2\big)' = 0,\ u(0) = 0,\ u(1) = 1$. The unique solution of this boundary value problem is $u(t) = t$ for $t \in [0,1]$. This produces a local minimum for I_L on $C^1_{0,1}([0,1])$. Let $v \in C^1_0([0,1])$. Then $u + v$ is admissible and we have

$$\begin{aligned} I_L(u+v) &= \int_0^1 \big[\big(t+v(t)\big)'\big]^3 \mathrm{d}t \\ &= I_L(u) + 3\int_0^1 v'(t)\mathrm{d}t + \int_0^1 \big(3v'(t)^2 + v'(t)^3\big)\mathrm{d}t \\ &= I_L(u) + \int_0^1 \big(3v'(t)^2 + v'(t)^3\big)\mathrm{d}t \qquad (\text{because } v(0) = v(1) = 0). \end{aligned}$$

Evidently if $3v'(t)^2 + v'(t)^3 \geq 0$ for all $t \in [0,1]$, then $I_L(u+v) \geq I_L(u)$. In particular if $\|v\|_{C^1([0,1])} \leq 3$, then $3v'(t)^2 + v'(t)^3 \geq 0$ for all $t \in [0,1]$ and so $I_L(u+v) \geq I_L(u)$ which proves that the extremal u produces a local minimum of I_L on $C^1_{0,1}([0,1])$.

(c) If L is not C^1, then the extremal function need not be C^1. Indeed let $L(t,y) = t^{2/3}y^2$ and consider the problem $\inf[I_L(u) : u \in C^1([0,1]),\ u(0) = 0,\ u(1) = 1]$. In this case the Euler equation has the form

$$\big(2t^{2/3}u'(t)\big)' = 0, \quad u(0) = 0, \quad u(1) = 1.$$

Solving this boundary value problem, we obtain $u(t) = t^{2/3}$ which does not belong in $C^1([0,1])$, but it is a global minimum of I_L over $C^1_{0,1}([0,1])$. This example is due to Hilbert and shows that a variational problem does not always have a solution in the given class of curves under consideration.

(d) In this example the Euler equation has no solutions and moreover, there is no solution even in the larger space of absolutely continuous functions. So let $L(t,y) = (ty)^2$ and consider the problem $\inf[I_L(u) : u \in C^1([0,1]),\ u(0) = 0,\ u(1) = 1]$. In this case the Euler equation has the form

$$\big(2t^2u'(t)\big)' = 0, \quad u(0) = 0, \quad u(1) = 1.$$

The general solution of the differential equation is $u(t) = (c/t) + d$. First we observe that u does not belong in $C^1([0,1])$ and none of these functions satisfies the boundary conditions $u(0) = 0,\ u(1) = 1$. In fact the variational problem has no solution in the space of absolutely continuous functions with the given boundary values. Indeed, if $v(t)$ is such a function, then $I_L(v) > 0$ and the value of the problem is zero. To see this consider

$$v_n(t) = \begin{cases} nt & \text{if } t \in [0, \frac{1}{n}] \\ 1 & \text{if } t \in [\frac{1}{n}, 1] \end{cases}.$$

Then we have $I_L(x_n) \longrightarrow 0$. This example is due to Weierstrass and was produced as an argument against Riemann's justification of the Dirichlet principle.

Now we introduce the Legendre transform, which is the "forefather" of the Legendre–Fenchel transform, which was introduced in Definition 1.2.15 and which is the basic tool in the duality theory of convex functions.

DEFINITION 2.5.28 Let X be a Banach space, $U \subseteq \mathbb{R} \times X \times X$ a nonempty open set, and $L\colon U \longrightarrow \mathbb{R}$ a C^2 Lagrangian function.

(a) The *Legendre transform* of L is the function $\mathcal{L}\colon U \longrightarrow \mathbb{R} \times X \times X^*$ defined by

$$\mathcal{L}(t,x,y) = \big(t, x, p = L'_y(t,x,y)\big).$$

(b) We say that the Lagrangian is *regular* if the corresponding Legendre transform $\mathcal{L}$ is a local diffeomorphism. We say that the Lagrangian is *hyperregular* if the corresponding Legendre transform $\mathcal{L}$ is a diffeomorphism of U on an open set $\mathcal{L}(U) \subseteq \mathbb{R} \times X \times X^*$.

REMARK 2.5.29 From the inverse theorem (see Theorem 1.1.25), the Lagrangian L is regular if and only if the Fréchet derivative of $\mathcal{L}$ at all $(t,x,y) \in U$ is an isomorphism from $T \times X \times X$ onto $T \times X \times X^*$. In what follows let D denote the Fréchet derivative with respect to all three variables $(t,x,y) \in U$ of the function involved. We have

$$D\mathcal{L}(t,x,y)(\lambda,v,w) = \big(\lambda, v, DL'_y(t,x,y)(\lambda,v,w)\big).$$

Note that

$$DL'_y(t,x,y)(\lambda,v,w) = \lambda L''_{yt}(t,x,y) + L''_{yx}(t,x,y)v + L''_{yy}(t,x,y)w \quad \text{in } X^*.$$

Then $D\mathcal{L}(t,x,y) \in \mathcal{L}(\mathbb{R}\times X\times X, \mathbb{R}\times X\times X^*)$ is an isomorphism if and only if $L''_{yy}(t,x,y) \in \mathcal{L}(X,X^*)$ is an isomorphism. In Proposition 2.5.25 we have assumed that $L''_{yy}(t_0,x_0,y_0) \in \mathcal{L}(X,X^*)$ is an isomorphism. Recall that in $\mathcal{L}(X,X^*)$ with the operator norm topology, the set of isomorphisms is open. So for all (t,x,y) near (t_0,x_0,y_0) in $\mathbb{R}\times X\times X$, we have that $L''_{yy}(t,x,y) \in \mathcal{L}(X,X^*)$ is an isomorphism. Therefore we see that the hypothesis in Proposition 2.5.25 implies that the Lagrangian L is regular in a neighborhood of the point (t_0,x_0,y_0).

THEOREM 2.5.30 *If $L\colon U \longrightarrow \mathbb{R}$ is a C^2 Lagrangian function that is hyperregular, $\mathcal{L}$ is the Legendre transform of L with $\mathcal{L}^{-1}$ its inverse $\big(\mathcal{L}^{-1}(t,x,y) = \big(t,x,y = g(t,x,p)\big)\big)$, and $H\colon \mathcal{L}(U) \longrightarrow \mathbb{R}\times X\times X^*$ is defined by*

$$H(t,x,p) = \langle p, g(t,x,p)\rangle_X - (L\circ\mathcal{L}^{-1})(t,x,p),$$

then the Euler equation

$$\frac{\mathrm{d}}{\mathrm{d}t} L'_y\big(t,u(t),u'(t)\big) = L'_x\big(t,u(t),u'(t)\big) \qquad \text{for all } t\in T=[0,b] \tag{2.75}$$

is equivalent to the following first-order canonical system

$$\left.\begin{cases} x'(t) = H'_p\big(t,x(t),p(t)\big) & \\ p'(t) = -H'_x\big(t,x(t),p(t)\big) & \text{for all } t\in T=[0,b] \end{cases}\right\}. \tag{2.76}$$

The function H is the *Hamiltonian* associated with the Lagrangian L and system (2.76) is called a *Hamiltonian system*. The equivalence of (2.75) and (2.76) is in the following sense: if $t\longrightarrow u(t)$ is a solution of the Euler equation (2.75), then $t\longrightarrow \big(u(t),\ L'_y\big(t,u(t),u'(t)\big)\big)$ is a solution of the Hamiltonian system (2.76). Conversely, if $t\longrightarrow \big(x(t),p(t)\big)$ is a solution of the Hamiltonian system (2.76), then $t\longrightarrow u(t) = x(t)$ is a solution of the Euler equation (2.75).

PROOF: In the proof of Proposition 2.5.25 we proved the local equivalence of (2.75) with the canonical system (2.74). In fact due to the hyperregularity of the Legendre transform $\mathcal{L}$, this equivalence is actually global. So it suffices to show the equivalence of (2.74) and (2.76).

By definition we have

$$H(t,x,p) = \langle p, g(t,x,p)\rangle_X - L\big(t,x,g(t,x,p)\big).$$

We have $H'_p(t,x,p) \in X^{**}$. So for all $x^*\in X^*$

$$\begin{aligned}\big\langle H'_p(t,x,p), x^*\big\rangle_{X^*} &= \langle x^*, g(t,x,p)\rangle_X + \big\langle p, g'_p(t,x,p)x^*\big\rangle_X \\ &\quad - \big\langle L'_y\big(t,x,g(t,x,p)\big), g'_p(t,x,p)x^*\big\rangle_X.\end{aligned} \tag{2.77}$$

But $p = L'_y\big(t,x,g(t,x,p)\big)$. Hence from (2.77) we infer that

$$\begin{aligned}&\big\langle H'_p(t,x,p), x^*\big\rangle_{X^*} = \langle x^*, g(t,x,p)\rangle_X \qquad \text{for all } x^*\in X^* \\ \Rightarrow\ & H'_p(t,x,p) = g(t,x,p) \in X.\end{aligned} \tag{2.78}$$

Then we see that the first equations in (2.74) and (2.76) are the same. Also if $v \in X$, then

$$\begin{aligned}\left\langle H_x'(t,x,p), v\right\rangle_X &= \left\langle p, h_x'(t,x,p)v\right\rangle_X - \left\langle L_x'\big(t,x,g(t,x,p)\big), v\right\rangle_X \\ &\quad - \left\langle L_y'\big(t,x,g(t,x,p)\big), h_x'(t,x,p)v\right\rangle_X . \end{aligned} \tag{2.79}$$

Again because $p = L_x'\big(t,x,g(t,x,p)\big)$, from (2.79) we obtain

$$\begin{aligned}&\left\langle H_x'(t,x,p), v\right\rangle_X = -\left\langle L_x'\big(t,x,g(t,x,p)\big), v\right\rangle_X \qquad \text{for all } v \in X, \\ \Rightarrow\ & H_x'(t,x,p) = L_x'\big(t,x,g(t,x,p)\big).\end{aligned}$$

$\square$

EXAMPLE 2.5.31 *The Poincaré half-plane*: Let $P = \{(x,y) \in \mathbb{R}^2 : y > 0\}$ and $L\colon P \times \mathbb{R}^2 \longrightarrow \mathbb{R}_+$ is defined by

$$L(x,y,u,v) = \frac{1}{2}\frac{u^2+v^2}{y^2}.$$

The integral functional I_L is defined by

$$I_L(u) = \frac{1}{2}\int_0^b \frac{x'(t)^2 + y'(t)^2}{y(t)^2}\mathrm{d}t,$$

where $u : T = [0,b] \longrightarrow P$ and $u(t) = \big(x(t), y(t)\big)$ for all $t \in T$. Then

$$\mathcal{L}(x,y,u,v) = (x,y,p,q)$$

with $p = u/y^2$, $q = v/y^2$. This is a diffeomorphism (i.e., $\mathcal{L}$ is hyperregular). Also

$$H(x,y,p,q) = \frac{1}{2}y^2(p^2+q^2).$$

Then the Hamiltonian system (2.76) takes the following form:

$$\left\{\begin{array}{l} x'(t) = y(t)^2\, p(t) \\ y'(t) = y(t)^2\, q(t) \end{array}\right. \qquad \left\{\begin{array}{l} p'(t) = 0 \\ q'(t) = -y(t)\big(p(t)^2 + q(t)^2\big) \end{array}\right. .$$

This system is a model for the hyperbolic geometry. For this reason its study is important in the understanding of non-Euclidean geometries.

2.6 Optimal Control

In this section we study some basic aspects of optimal control theory. Specifically we focus our attention on the existence theory, the relaxation methods, and the derivation of the maximum principle (a necessary condition for optimality of an admissible state-control pair). To avoid technical complications that require the development of a substantial mathematical background, we limit ourselves to finite dimensional systems (i.e., systems driven by ordinary differential equations, also known as lumped parameter systems).

So the mathematical setting is the following. The state space is $\mathbb{R}^N$ and the control space is a Polish space Y. Recall that a Polish space is a Hausdorff topological space which is separable and can be metrized by means of a complete metric. The time horizon is $T = [0, b]$. Let $\Gamma(T, Y) = \{u \colon T \longrightarrow Y : u \text{ is measurable}\}$. Any element of $\Gamma(T, Y)$ is called a control. Also $f \colon T \times \mathbb{R}^N \times Y \longrightarrow \mathbb{R}^N$ is a vector field and the controlled dynamical system we are considering is the following

$$x'(t) = f\big(t, x(t), u(t)\big) \qquad \text{a.e. on } T, \ x(0) = x_0 \in \mathbb{R}^N. \tag{2.80}$$

Any solution $x(\cdot)$ of (2.80) is referred to as a state trajectory of the control system corresponding to the initial state $x_0 \in \mathbb{R}^N$ and the control $u(\cdot)$. Note that we do not assume the uniqueness and/or the existence of solutions for (2.80). So for any control $u(\cdot)$ and any initial state x_0, we may have more than one or no response for the system. Moreover, in general we have some constraints on the control described by a multifunction (set-valued function) $U \colon T \longrightarrow 2^Y \setminus \{\emptyset\}$. For these reasons, the following definition is necessary.

DEFINITION 2.6.1 A pair $\big(x(\cdot), u(\cdot)\big) \in W^{1,1}\big((0, b), \mathbb{R}^N\big) \times \Gamma(T, Y)$ is said to be *admissible* (or *feasible*) state-control pair, if (2.80) is satisfied and $u(t) \in U(t)$ a.e. on T. In what follows by $\alpha_{\mathbf{ad}}$ we denote the set of all admissible state-control pairs.

We are also given an integral cost functional

$$J(x, u) = \int_0^b L\big(t, x(t), u(t)\big) \mathrm{d}t \qquad \text{for all } \ (x, u) \in \alpha_{\mathbf{ad}}.$$

Then our optimal control problem can be stated as follows:

(P) Find $(\widehat{x}, \widehat{u}) \in \alpha_{\mathbf{ad}}$ such that

$$J(\widehat{x}, \widehat{u}) = \inf \big[J(x, u) : (x, u) \in \alpha_{\mathbf{ad}}\big] = m. \tag{2.81}$$

DEFINITION 2.6.2 If we can find $(\widehat{x}, \widehat{u}) \in \alpha_{\mathbf{ad}}$ such that (2.81) is satisfied, then we say that $(\widehat{x}, \widehat{u})$ is an *optimal admissible pair*. The function $\widehat{x} \in W^{1,1}\big((0, b), \mathbb{R}^N\big)$ is said to be an *optimal trajectory* and the function $u \in \Gamma(T, Y)$ is said to be an *optimal control*.

The hypotheses on the data of the optimal control problem (P), are the following.

H(f): $f \colon T \times \mathbb{R}^N \times Y \longrightarrow \mathbb{R}^N$ is a function, such that

(i) For all $(x, u) \in \mathbb{R}^N \times Y$, $t \longrightarrow f(t, x, u)$ is measurable.
(ii) For almost all $t \in T$, $(x, u) \longrightarrow f(t, x, u)$ is continuous.
(iii) For almost all $t \in T$, all $x \in \mathbb{R}^N$ and all $u \in U(t)$, we have

$$\|f(t, x, u)\| \leq \alpha(t) + c(t)\|x\| \qquad \text{with } \alpha, c \in L^1(T)_+.$$

H(U): $U \colon T \longrightarrow 2^Y \setminus \{\emptyset\}$ is a multifunction with compact values such that

$$\mathrm{Gr} U = \{(t, u) \in T \times Y : u \in U(t)\} \in \mathbf{B}(T) \times \mathbf{B}(Y),$$

with $\mathbf{B}(T)$ (resp., $\mathbf{B}(Y)$) being the Borel σ-field of T (resp., of Y).

REMARK 2.6.3 From measure theory we know that $\mathbf{B}(T) \times \mathbf{B}(Y) = \mathbf{B}(T \times Y)(=$ the Borel σ-field of $T \times Y)$.

H(L): $L\colon T\times\mathbb{R}^N\times Y\longrightarrow\overline{\mathbb{R}}=\mathbb{R}\cup\{+\infty\}$ is an integrand such that

(i) $(t,x,u)\longrightarrow L(t,x,u)$ is measurable.
(ii) For almost all $t\in T$, $(x,u)\longrightarrow L(t,x,u)$ is lower semicontinuous proper.
(iii) For almost all $t\in T$, all $x\in\mathbb{R}^N$, and all $u\in U(t)$

$$\alpha_0(t)-c_0\|x\|\leq L(t,x,u)\qquad\text{with } \alpha_0\in L^1(T),\ c_0>0.$$

In the existence theory the following convexity-type hypothesis plays a central role.
$\mathbf{H_c}$**:** For all $(t,x)\in T\times\mathbb{R}^N$ the set $Q(t,x)=\{(v,\lambda)\in\mathbb{R}^N\times\mathbb{R}: v=f(t,x,u),\ u\in U(t),\ L(t,x,u)\leq\lambda\}$ is convex.

REMARK 2.6.4 Hypotheses H(f), H(u), and H(L) imply that for all $(t,x)\in T\times\mathbb{R}^N$, $Q(t,x)$ is also closed. Hypothesis H_c means that the optimal control problem has enough "convex structure". Note that if the control variable enters linearly in the dynamics of the system (i.e., $f(t,x,u)=f_1(t,x)+f_2(t,x)u$), $Y=\mathbb{R}^m$, the control constraint multifunction has also convex values and the cost integrand $L(t,x,u)$ is in addition convex in $u\in\mathbb{R}^m$, then the hypothesis H_c is satisfied. Note that this hypothesis implies that for almost all $t\in T$ and all $x\in\mathbb{R}^N$, $F(t,x)=f\big(t,x,U(t)\big)$ is convex. In Chapter 6, we produce equivalent conditions for H_c to hold.

In what follows let $\alpha^{\mathbf{1}}_{\mathbf{ad}}\subseteq W^{1,1}\big((0,b),\mathbb{R}^N\big)$ be the set of admissible states (trajectories). We start by establishing the nonemptiness and compactness in $C(T,\mathbb{R}^N)$ of $\alpha^{\mathbf{1}}_{\mathbf{ad}}$. Recall that the Sobolev space $W^{1,1}\big((0,b),\mathbb{R}^N\big)$ is embedded continuously (but not compactly) in $C(T,\mathbb{R}^N)$. To determine the properties of $\alpha^{\mathbf{1}}_{\mathbf{ad}}$, we study the following differential inclusion.

$$\left\{\begin{array}{l} x'(t)\in F\big(t,x(t)\big)\ \text{ a.e. on } T=[0,b],\\ x(0)=x_0\end{array}\right\}. \tag{2.82}$$

We impose the following conditions on the multifunction $F(t,x)$:

H(F): $F\colon T\times\mathbb{R}^N\longrightarrow 2^{\mathbb{R}^N}\setminus\{\emptyset\}$ is a multifunction with compact and convex values such that

(i) For all $x\in\mathbb{R}^N$, $\mathrm{Gr}F(\cdot,x)=\{(t,v)\in T\times\mathbb{R}^N: v\in F(t,x)\}\in\mathbf{B}(T)\times\mathbf{B}(\mathbb{R}^N)$.
(ii) For almost all $t\in T$, $\mathrm{Gr}F(t,\cdot)=\{(x,v)\in\mathbb{R}^N\times\mathbb{R}^N: v\in F(t,x)\}$ is closed.
(iii) For almost all $t\in T$, all $x\in\mathbb{R}^N$, and all $v\in F(t,x)$, we have

$$\|v\|\leq\alpha(t)+c(t)\|x\|\qquad\text{with } \alpha,c\in L^1(T)_+.$$

REMARK 2.6.5 If $F(t,x)=f\big(t,x,U(t)\big)$, then by virtue of hypotheses H(f), H(U), H(L), and H_c, the multifunction F satisfies hypotheses H(F).

DEFINITION 2.6.6 By a solution of (2.82), we mean a function $x\in W^{1,1}\big((0,b),\mathbb{R}^N\big)$ such that $x(0)=x_0$ and $x'(t)\in F\big(t,x(t)\big)$ for almost all $t\in T$. We denote the set of solutions of (2.82) by $S_F(x_0)$. Then $S_F(x_0)\subseteq W^{1,1}\big((0,b),\mathbb{R}^N\big)\subseteq C(T,\mathbb{R}^N)$.

To be able to solve (2.82) and eventually study the optimal control problem, we need some results from multivalued analysis, which we state here for easy reference and postpone their proof until Chapter 6, where we conduct a more systematic and detailed study of multifunctions.

The first result is known as the *Kakutani–Ky Fan fixed point theorem*.

THEOREM 2.6.7 *If X is a locally convex space, $C \subseteq X$ is nonempty, compact and convex, and $G\colon C \longrightarrow 2^C$ is a multifunction with nonempty, closed and convex values that has a closed graph (i.e., $Gr\, G = \{(x,y) \in C \times C : y \in G(x)\}$ is closed in $X \times X$), then there exists $x \in X$ such that $x \in G(x)$.*

The second result is known as the *Yankov–von Neumann–Aumann selection theorem*. It is stated in a less general form, which however suffices for our needs here. The general form of the result can be found in Theorem 6.3.20.

THEOREM 2.6.8 *If (Ω, Σ, μ) is a σ-finite measure space, X is a Polish space, and $G\colon \Omega \longrightarrow 2^X$ is a multifunction such that $Gr\, G = \{(\omega, x) \in \Omega \times X : y \in G(\omega)\} \in \Sigma \times \mathbf{B}(X)$, then there exists a Σ-measurable function $g\colon \Omega \longrightarrow X$ such that $g(\omega) \in G(\omega)$ μ-a.e. on Ω.*

It is well known that given a Borel set in $\mathbb{R}^2$, its projection on a coordinate axis need not be Borel. The next theorem provides conditions which guarantee that the projection is indeed Borel.

THEOREM 2.6.9 *If S, X are Polish spaces, $G \in \mathbf{B}(S \times X) = \mathbf{B}(S) \times \mathbf{B}(X)$, and for every $s \in S$ the section $G(s) = \{x \in X : (s,x) \in G\}$ is σ-compact, then $\mathrm{proj}_S G \in \mathbf{B}(S)$.*

Now we are ready to determine the properties of the solution set $S(x_0)$ of problem (2.82).

THEOREM 2.6.10 *If hypotheses H(F) hold, then $S_F(x_0) \neq \emptyset$ and $S_F(x_0)$ is compact in $C(T, \mathbb{R}^N)$.*

PROOF: We start by establishing an a priori bound for the elements of $S_F(x_0)$. So suppose that $x \in S_F(x_0)$. Then for some $v \in L^1(T, \mathbb{R}^N)$ that satisfies $v(t) \in F\big(t, x(t)\big)$ a.e. on T, we have

$$\|x(t)\| \leq \|x_0\| + \int_0^t \|v(s)\| \mathrm{d}s \qquad \text{for all } t \in T,$$

$$\Rightarrow \|x(t)\| \leq \|x_0\| + \int_0^t \big(\alpha(s) + c(s)\|x(s)\|\big) \qquad \text{for all } t \in T.$$

Invoking Gronwall's inequality, we can find $M_1 > 0$ such that

$$\|x(t)\| \leq M_1 \qquad \text{for all } t \in T \text{ and all } x \in S(x_0).$$

Let $p_{M_1} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the M_1-radial contraction; that is,

$$p_{M_1}(x) = \begin{cases} \frac{M_1 x}{\|x\|} & \text{if } \|x\| > M_1 \\ x & \text{if } \|x\| \le M_1 \end{cases}.$$

Clearly $p_{M_1}(\cdot)$ is Lipchitz continuous. Let $F_1(t,x) = F\big(t, p_{M_1}(x)\big)$. Then F_1 still satisfies H(F)(i) and (ii) (with F replaced by F_1) and in addition for almost all $t \in T$, all $x \in \mathbb{R}^N$, and all $v \in F_1(t,x)$, we have

$$\|v\| \le \eta(t) \qquad \text{a.e. on } T, \text{ with } \eta(\cdot) = \alpha(\cdot) + c(\cdot)M \in L^1(T)_+.$$

Set $C = \big\{g \in L^1(T,\mathbb{R}^N) : \|g(t)\| \le \eta(t) \text{ a.e. on } T\big\}$. By virtue of the Dunford–Pettis theorem, C furnished with the relative weak topology is a compact convex set. Let $\xi \colon C \longrightarrow C(T,\mathbb{R}^N)$ be defined by

$$\xi(g)(t) = x_0 + \int_0^t g(s)\mathrm{d}s.$$

Via the Arzela–Ascoli theorem, we check that ξ is sequentially continuous from C with the relative weak topology into $C(T,\mathbb{R}^N)$ with the norm topology. Then let $G_1 \colon C \longrightarrow 2^C$ be defined by

$$G_1(g) = \{h \in L^1(T,\mathbb{R}^N) : h(t) \in F_1\big(t, \xi(g)(t)\big) \qquad \text{a.e. on } T\}.$$

Let $\xi(g) \in C(T,\mathbb{R}^N)$ and let $\{s_n\}_{n\ge 1}$ be step functions from T into $\mathbb{R}^N$ such that $s_n(t) \longrightarrow x(t)$ a.e. on T as $n \to \infty$. Then because for all $y \in \mathbb{R}^N$, $t \longrightarrow F_1(t,y)$ has a measurable graph, applying Theorem 2.6.8 for every $n \ge 1$ we can find $h_n : T \longrightarrow \mathbb{R}^N$ a measurable map such that $h_n(t) \in F_1\big(t, s_n(t)\big)$ a.e on T. Evidently $\|h_n(t)\| \le \eta(t)$ a.e on T and so by the Dunford–Pettis theorem and by passing to a suitable subsequence if necessary, we may assume that $h_n \xrightarrow{w} h$ in $L^1(T,\mathbb{R}^N)$. Then using Mazur's lemma and the fact that for almost all $t \in T$, $y \longrightarrow F_1(t,y)$ has closed graph, we have $h(t) \in F_1\big(t, x(t)\big)$ a.e. on T (i.e., $h \in G_1(g)$). So G_1 has nonempty values, which clearly are closed and convex in C. Moreover, the above argument reveals that Gr $G \subseteq C \times C$ is sequentially closed in $C \times C$, when C is equipped with the relative weak topology. Recall that the relative weak topology on C is metrizable. So we can apply Theorem 2.6.7 and obtain $g \in C$ such that $g \in G_1(g)$. Then if $x = \xi(g)$, we have

$$\begin{aligned} & x'(t) \in F_1\big(t, x(t)\big) \qquad \text{a.e. on } T,\ x(0) = x_0, \\ \Rightarrow\ & \|x(t)\| \le \|x_0\| + \int_0^t \|v(s)\|\mathrm{d}s \qquad \text{for all } t \in T, \text{ with } v \in L^1(T,\mathbb{R}^N), \\ & v(s) \in F_1\big(s, x(s)\big) \qquad \text{a.e. on } T. \end{aligned} \tag{2.83}$$

But from the definition of F_1, we see that $\|v(s)\| \le \alpha(s) + c(s)\|x\|$ a.e. on T and so from (2.83) and Gronwall's inequality as before it follows that

$$\begin{aligned} & \|x(t)\| \le M_1 \qquad \text{for all } t \in T, \\ \Rightarrow\ & F_1\big(t, x(t)\big) = F\big(t, x(t)\big) \qquad \text{for all } t \in T \quad (\text{i.e., } x \in S_F(x_0)). \end{aligned}$$

Therefore we have proved the nonemptiness of $S_F(x_0)$.

Clearly $S(x_0) \subseteq C(T,\mathbb{R}^N)$ is closed, whereas from the Arzela–Ascoli theorem we have that $S(x_0)$ is relatively compact in $C(T,\mathbb{R}^N)$. Then $S_F(x_0)$ is compact in $C(T,\mathbb{R}^N)$. ∎

Now if $F(t,x) = f\big(t,x,U(t)\big)$ and $x \in S(x_0)$, then we set

$$V(t) = \{u \in U(t) : x'(t) = f\big(t, x(t), u\big)\},$$
$$\Rightarrow \text{Gr } V = \{(t,u) : x'(t) = f\big(t,x(t),u\big)\} \cap \text{Gr } U.$$

Because of hypotheses H(f)(i),(ii), $(t,u) \longrightarrow x'(t) - f\big(t,x(t),u\big)$ is a Carathéodory function (i.e., for all $u \in Y$, $t \longrightarrow x'(t) - f\big(t,x(t),u\big)$ is measurable, whereas for almost all $t \in T$, $u \longrightarrow x'(t) - f\big(t,x(t),u\big)$ is continuous). Hence $(t,u) \longrightarrow x'(t) - f\big(t,x(t),u\big)$ is measurable and combining this with hypothesis H(U) we see that

$$\text{Gr } V \in \mathcal{L}(T) \times \mathbf{B}(Y)$$

with $\mathcal{L}(T)$ being the Lebesgue σ-field of T and $\mathbf{B}(Y)$ the Borel σ-field of Y. So we can apply Theorem 2.6.8 and obtain $u : T \longrightarrow Y$ a measurable map such that $u(t) \in V(t)$ a.e. on T. Then $(x,u) \in \alpha_{\mathbf{ad}}$ and so Theorem 2.6.10 implies that $\alpha^{\mathbf{1}}_{\mathbf{ad}} = S_F(x_0) \subseteq C(T, \mathbb{R}^N)$ is compact.

Next we solve the optimal control problem (P). The approach is the following. We use hypothesis H$_c$ to transform (P) into a control-free variational problem with convex structure (calculus of variations problem) with the same values as (P). We solve the calculus of variations problem using the *direct method*. This method is known as the *reduction method*.

THEOREM 2.6.11 *If hypotheses H(f), H(U), H(L), and H$_c$ hold and $m < +\infty$, then problem (P) admits an optimal state-control pair $(\widehat{x}, \widehat{u}) \in \alpha^{\mathbf{1}}_{\mathbf{ad}}$.*

PROOF: We start implementing the reduction method described above. To this end let

$$A(t,x,v) = \{u \in U(t) : v = f(t,x,u)\}.$$

This is the set of all admissible controls which at time $t \in T$ and when the system is at state $x \in \mathbb{R}^N$, produce the "velocity" $v \in \mathbb{R}^N$. Evidently, by modifying f, U, and L on a Lebesgue-null set, we have

$$\begin{aligned} \text{Gr } A = &\big\{(t,x,v,u) \in T \times \mathbb{R}^N \times \mathbb{R}^N \times Y : (t,u) \in \text{Gr } U,\ v - f(t,x,u) = 0\big\} \\ &\in \mathbf{B}(T) \times \mathbf{B}(\mathbb{R}^N) \times \mathbf{B}(\mathbb{R}^N) \times \mathbf{B}(Y). \end{aligned}$$

Let $p \colon T \times \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be defined by

$$p(t,x,v) = \inf[L(t,x,u) : u \in A(t,x,v)] \tag{2.84}$$

which represents the minimum cost to generate "velocity" v, at time $t \in T$, when the state of the system is $x \in \mathbb{R}^N$ and we use admissible controls. As always $\inf\, \emptyset = +\infty$.

Claim 1: $(t,x,v) \longrightarrow p(t,x,v)$ is Borel-measurable.

For every $\lambda \in \mathbb{R}$ we have

$$\begin{aligned} &\{(t,x,v) \in T \times \mathbb{R}^N \times \mathbb{R}^N : p(t,x,v) \le \lambda\} \\ &= \text{proj}_{T \times \mathbb{R}^N \times \mathbb{R}^N} \big\{(t,x,v,u) \in T \times \mathbb{R}^N \times \mathbb{R}^N \times Y : L(t,x,u) \le \lambda, \\ &\quad u \in A(t,x,v)\big\} \in \mathbf{B}(T) \times \mathbf{B}(\mathbb{R}^N) \times \mathbf{B}(\mathbb{R}^N) \ \text{(see Theorem 2.6.9).} \end{aligned}$$

Claim 2: For every $t \in T$, the function $(x,v) \longrightarrow p(t,x,v)$ is lower semicontinuous.

We need to show that for every $\lambda \in \mathbb{R}$, the set $S(\lambda) = \{(x,v) \in \mathbb{R}^N \times \mathbb{R}^N : p(t,x,v) \leq \lambda\}$ is closed. So let $\{(x_n, v_n)\}_{n\geq 1} \subseteq S(\lambda)$ and assume that $x_n \longrightarrow x,\ v_n \longrightarrow v$ in $\mathbb{R}^N$ as $n \to \infty$. From (2.84) we know that we can find $u_n \in A(t, x_n, v_n)$ such that

$$p(t, x_n, v_n) = L(t, x_n, u_n).$$

We have $u_n \in U(t)$ and $v_n = f(t, x_n, u_n)$. By passing to a suitable subsequence if necessary, we may assume that $u_n \longrightarrow u \in U(t)$. Then $v = f(t,x,u)$ (see hypothesis H(f)(ii)) and so $u \in A(t,x,v)$. Therefore

$$p(t,x,v) \leq L(t,x,u) \leq \liminf_{n\to\infty} L(t,x_n,u_n) = \liminf_{n\to\infty} p(t,x_n,v_n) \leq \lambda$$

(see hypothesis H(L)(ii)),

$$\Rightarrow\ (x,v) \in S(\lambda), \quad \text{(i.e., } p(t,\cdot,\cdot) \text{ is lower semicontinuous).}$$

Claim 3: For all $(t,x) \in T \times \mathbb{R}^N,\ v \longrightarrow p(t,x,v)$ is convex.

Note that

$$\begin{aligned} \operatorname{epi}(t,x,\cdot) &= \{(v,\lambda) \in \mathbb{R}^N \times \mathbb{R} : p(t,x,v) \leq \lambda\} \\ &= \bigcap_{\varepsilon>0} \{(v,\lambda) \in \mathbb{R}^N \times \mathbb{R} : L(t,x,u) \leq \lambda + \varepsilon, \\ &\qquad\qquad u \in U(t),\ v = f(t,x,u)\}. \end{aligned}$$

So by virtue of hypothesis H_c, $\operatorname{epi}(t,x,\cdot)$ is convex; that is, $v \longrightarrow p(t,x,v)$ is convex as claimed.

Now let $\{(x_n,u_n)\}_{n\geq 1} \subseteq \alpha_{\mathbf{ad}}$ be a minimizing sequence for problem (P); that is, $J(x_n,u_n) \downarrow m$ as $n \to \infty$. Because of Theorem 2.6.10, we may assume that $x_n \longrightarrow \widehat{x}$ in $C(T,\mathbb{R}^N)$. Also from hypothesis H(f)(iii) we see that $\{x_n'(\cdot)\}_{n\geq 1} \subseteq L^1(T,\mathbb{R}^N)$ is uniformly integrable and so by the Dunford–Pettis theorem, we may assume that $x_n' \xrightarrow{w} v$ in $L^1(T,\mathbb{R}^N)$. Evidently $v = \widehat{x}'$. Using Claims 1, 2, and 3 and Theorem 2.1.28, we obtain

$$\begin{aligned} \int_0^b p\big(t,\widehat{x}(t),\widehat{x}'(t)\big)\mathrm{d}t &\leq \liminf_{n\to\infty} \int_0^b p\big(t,x_n(t),x_n'(t)\big)\mathrm{d}t \\ &\leq \lim_{n\to\infty} J(x_n,u_n) = m < +\infty. \end{aligned}$$

So due to hypothesis H(L)(iii) and by redefining $t \longrightarrow p\big(t,\widehat{x}(t),\widehat{x}'(t)\big)$ on a Lebesgue-null set, we can say that for all $t \in T$, $p\big(t,\widehat{x}(t),\widehat{x}'(t)\big)$ is finite. A straightforward application of Theorem 2.6.8 produces a measurable function $\widehat{u}\colon T \longrightarrow Y$ such that $\widehat{u}(t) \in A\big(t,\widehat{x}(t),\widehat{x}'(t)\big)$ for almost all $t \in T$ and

$$\begin{aligned} & p\big(t,\widehat{x}(t),\widehat{x}'(t)\big) = L\big(t,\widehat{x}(t),\widehat{u}'(t)\big) \text{ a.e. on } T, \\ \Rightarrow\ & J(\widehat{x},\widehat{u}) \leq m \quad \text{and } (\widehat{x},\widehat{u}) \in \alpha_{\mathbf{ad}}, \\ \Rightarrow\ & J(\widehat{x},\widehat{u}) = m \quad \text{(i.e., } (\widehat{x},\widehat{u}) \text{ is an optimal state–control pair).} \end{aligned}$$

□

Reviewing the proof of the above theorem, we see that hypothesis H_c was crucial in the argument because it provided the needed convex structure for the direct method of the calculus of variations to work. If H_c fails, we need not have a solution. Then, in order to capture the asymptotic behavior of the minimizing sequences, we embed the original problem in a "larger" system exhibiting the necessary convex structure that guarantees existence of optimal pairs. This method of augmenting the system is known as *relaxation*. There is no unique approach to relaxation. Nevertheless, we can agree that a reasonable relaxation method should meet the following three criteria.

(i) Every original state is also a relaxed state.

(ii) The set of original states is dense in the set of relaxed states.

(iii) The relaxed problem admits a solution and the values of the two problems (relaxed and original) are equal.

The first two requirements concern the dynamics of the system, and the third also involves the cost functional. The requirement that the values of the two problems are equal is known as *relaxability*. If relaxability fails, it can be said that the original system has been enlarged too much. A relaxation method that meets the three criteria is said to be *admissible*.

Next we present some different ways to relax (convexify) the original problem (P). The first relaxation method has its roots in the following observation. We recall that a sequence $\{u_n\}_{n\geq 1} \subseteq L^1(T, \mathbb{R}^m)$ that converges weakly but not strongly to a limit u oscillates violently around its limit. However, in the limit u a great deal of information about the faster and faster oscillations is forgotten and only an average value is registered in the limit. Clearly this is not satisfactory if the control enters in a nonlinear fashion in the dynamics. The idea then is to assign as a limit not a usual $\mathbb{R}^m$-valued function, but a probability-valued function, known as a transition measure (or parametrized measure or Young measure).

DEFINITION 2.6.12 Let (Ω, Σ, μ) be a finite measure space and Y a Polish space. Let $M_+^1(Y)$ be the space of all probability measures on Y. A *relaxed control* (or *transition probability* or *Young measure*) is a function $\lambda \colon \Omega \longrightarrow M_+^1(Y)$, such that for all $C \in \mathbf{B}(Y) =$ the Borel σ-field of Y, $\omega \longrightarrow \lambda(\omega)(C)$ is Σ-measurable. We denote the set of all relaxed controls from Ω into Y by $R(\Omega, Y)$.

REMARK 2.6.13 The "narrow topology" on $M_+^1(Y)$ is the weakest topology on $M_+^1(Y)$ that makes continuous all the maps $\psi_u \colon M_+^1(Y) \longrightarrow \mathbb{R}$ with $u \in C_b(Y) = \{u : Y \longrightarrow \mathbb{R} : u \text{ is continuous and bounded}\}$ defined by $\psi_u(\lambda) = \int_Y u(y)\lambda(\mathrm{d}y)$. If we topologize $M_+^1(Y)$ this way, then $\lambda \in R(\Omega, Y)$ if and only if $\lambda \colon \Omega \longrightarrow M_+^1(Y)$ is Σ-measurable. So the narrow topology is the topology $w\big(M_+^1(Y), C_b(Y)\big)$. We can define an analogue of it on the space $R(\Omega, Y)$ of relaxed controls.

DEFINITION 2.6.14 Let $\mathrm{Car}(\Omega \times Y)$ be the space of L^1-Carathéodory integrands. Namely all functions $\varphi \colon \Omega \times Y \longrightarrow \mathbb{R}$ such that

(i) For all $y \in Y$, $\omega \longrightarrow \varphi(\omega, y)$ is Σ-measurable.
(ii) For μ-almost all $\omega \in \Omega$, $y \longrightarrow \varphi(\omega, y)$ is continuous.
(iii) There exists $h \in L^1(\Omega)_+$ such that for μ-almost all $\omega \in \Omega$ and all $y \in \mathbb{R}$, $|\varphi(\omega, y)| \leq h(\omega)$. The *narrow topology* on $R(\Omega, Y)$ is the weakest topology on $R(\Omega, Y)$ with respect to which the functionals

$$\lambda \longrightarrow I_\varphi(\lambda) = \int_\Omega \int_Y \varphi(\omega, y)\lambda(\omega)(\mathrm{d}y)\mathrm{d}\mu, \quad \varphi \in \mathrm{Car}(\Omega \times Y)$$

are continuous.

So the narrow topology on $R(\Omega, Y)$ is the topology $w\big(R(\Omega, Y), \mathrm{Car}(\Omega \times Y)\big)$.

REMARK 2.6.15 We say that an integrand $\varphi\colon \Omega\times Y \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is *normal* if

(i) φ is $\Sigma \times \mathbf{B}(Y)$-measurable.
(ii) For μ-almost all $\omega \in \Omega$, $y \longrightarrow \varphi(\omega, y)$ is proper and lower semicontinuous.

A positive normal integrand can be approximated pointwise from below by integrands in $\mathrm{Car}(\Omega \times Y)$. So equivalently the narrow topology on $R(\Omega, Y)$ can be defined as the weakest topology that makes all functionals $\lambda \longrightarrow I_\varphi(\lambda) = \int_\Omega \int_Y \varphi(\omega, y)\lambda(\omega)(\mathrm{d}y)\mathrm{d}\mu$ lower semicontinuous as φ ranges over all positive normal integrands. If Y is compact, then $R(\Omega, Y) \subseteq L^\infty\big(\Omega, M(Y)_{w^*}\big) = L^1\big(\Omega, C(Y)\big)^*$ (here $M(Y)$ is the space of all Radon measures on Y; recall that $C(Y)^* = M(Y)$) and the narrow topology on $R(\Omega, Y)$ coincides with the relative w^*-topology in $L^\infty\big(\Omega, M(Y)_{w^*}\big)$. If Σ is countably generated, then $L^1\big(\Omega, C(Y)\big)$ is separable and so on bounded subsets of $L^\infty\big(\Omega, M(Y)_{w^*}\big)$ (such as $R(\Omega, Y)$), the relative w^*-topology is metrizable. In the sequel this is the context in which we use the narrow topology on $R(\Omega, Y)$.

In what follows for simplicity we set $Y = \mathbb{R}^m$.

$\mathbf{H(U)_1}$**:** $U\colon T \longrightarrow 2^{\mathbb{R}^m} \setminus \{\emptyset\}$ is a multifunction with compact values such that Gr $U \in \mathbf{B}(T) \times \mathbf{B}(\mathbb{R}^m)$ and there exists $r > 0$ such that for almost all $t \in T$ and all $u \in U(t)$, $\|u\| \le r$.

We set $\overline{B}_r = \{u \in \mathbb{R}^m : \|u\| \le r\}$ and we introduce the constraint set for the relaxed controls, namely

$$\Sigma(t) = \big\{\mu \in M_+^1(\overline{B}_r) : \mu\big(U(t)\big) = 1\big\}.$$

Then given a state function $x \in W^{1,1}\big((0,b), \mathbb{R}^N\big) \subseteq C(T, \mathbb{R}^N)$, the set of admissible relaxed controls is given by

$$S_\Sigma = \{\lambda \in R(T, \overline{B}_r) : \lambda(t) \in \Sigma(t) \text{ a.e. on } T\}.$$

Now we have all the necessary tools to introduce and study the first relaxation of problem (P) which is based on Young measures (see Definition 2.6.12):

$$\left\{\begin{array}{ll} & J_r^1(x,\lambda) = \int_0^b \int_{\overline{B}_r} L\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \longrightarrow \inf = m_r^1 \\ \text{s.t.} & x'(t) = \int_{\overline{B}_r} f\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \text{ a.e. on } T \\ & x(0) = x_0, \ \lambda \in S_\Sigma \end{array}\right\}. \tag{2.85}$$

Note that in (2.85) in both the dynamics and the cost functional, the relaxed control enters linearly. This provides problem (2.85) with the necessary convex structure to obtain an optimal pair. For problem (2.85) every original control $u(\cdot)$ can be viewed as a relaxed control by considering the Dirac transition probability $\delta_{u(\cdot)}$. So the first requirement for admissibility is satisfied.

PROPOSITION 2.6.16 *If hypotheses $H(f)$, $H(U)_1$, and $H(L)$ hold and $m<+\infty$, then problem (2.85) has an optimal pair $(\widehat{x},\widehat{\lambda}) \in W^{1,1}\big((0,b),\mathbb{R}^N\big) \times S_\Sigma$; that is, $J_r^1(\widehat{x},\widehat{\lambda}) = m_r^1$.*

PROOF: Evidently $-\infty < m_r^1 \leq m < \infty$. Also let $\{(x_n,\lambda_n)\}_{n\geq 1}$ be a minimizing sequence for problem (2.85). Note that if $F(t,x)=\big\{ \int_{\overline{B}_r} f(t,x,u)\lambda(t)(\mathrm{d}u) : \lambda\in S_\Sigma\big\}$, then F satisfies hypotheses H(F). On the other hand by virtue of Remark 2.6.15, $S_\Sigma \subseteq L^\infty\big(T, M(\overline{B}_r)_{w^*}\big)$ is w^*-compact. So we may assume that

$$x_n \longrightarrow x \text{ in } C(T,\mathbb{R}^N) \quad \text{(see Theorem 2.6.10) and}$$
$$\lambda_n \xrightarrow{w^*} \lambda \in S_\Sigma \quad \text{in } L^\infty\big(T, M(\overline{B}_r)_{w^*}\big).$$

Also from the dynamics of (2.85) and hypothesis H(f)(iii), we see that

$$x_n' \xrightarrow{w} x' \quad \text{in } L^1(T,\mathbb{R}^N).$$

We have

$$\begin{aligned}\|f\big(t,x_n(t),\cdot\big) - f\big(t,x(t),\cdot\big)\|_{C(T,\mathbb{R}^m)} &= \max_{u\in\overline{B}_r} \|f\big(t,x_n(t),u\big) - f\big(t,x(t),u\big)\| \\ &= \|f\big(t,x_n(t),u_n\big) - f\big(t,x(t),u_n\big)\|,\end{aligned}$$

with $u_n \in \overline{B}_r$.

We may assume that $u_n \longrightarrow u$ in $\mathbb{R}^m$. Then by virtue of hypothesis H(f)(ii)

$$\begin{aligned} &\|f\big(t,x_n(t),u_n\big) - f\big(t,x(t),u_n\big)\| \longrightarrow 0 \quad \text{as } n\to\infty, \\ \Rightarrow\ & \widehat{f}_n(t) = f\big(t,x_n(t),\cdot\big) \longrightarrow f\big(t,x(t),\cdot\big) = \widehat{f}(t) \text{ in } C(\overline{B}_r,\mathbb{R}^m) \quad \text{as } n\to\infty, \\ \Rightarrow\ & \widehat{f}_n \longrightarrow \widehat{f} \text{ in } L^1\big(T, C(\overline{B}_r,\mathbb{R}^m)\big) \text{ as } n\to\infty. \end{aligned}$$

Because $\lambda_n \xrightarrow{w^*} \lambda$ in $L^\infty\big(T, M(\overline{B}_r)_{w^*}\big)$, for every $C \in \mathbf{B}(T)$ we have

$$\int_C\int_{\overline{B}_r} f\big(t,x_n(t),u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t \longrightarrow \int_C\int_{\overline{B}_r} f\big(t,x(t),u\big)\lambda(t)(\mathrm{d}u)$$

and $\int_C x_n'(t)\mathrm{d}t \longrightarrow \int_C x'(t)\mathrm{d}t$. So in the limit we have

$$\begin{aligned} &\int_C x'(t)\mathrm{d}t = \int_C\int_{\overline{B}_r} f\big(t,x(t),u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \qquad \text{for all } C \in \mathbf{B}, \\ \Rightarrow\ & x'(t) = \int_{\overline{B}_r} f\big(t,x(t),u\big)\lambda(t)(\mathrm{d}u) \qquad \text{for almost all } t\in T,\ x(0) = x_0. \end{aligned}$$

Therefore the pair (x,λ) is admissible for the relaxed problem (2.85).

Let $\widehat{L}_n(t,u) = L\big(t,x_n(t),u\big)$. Because of hypotheses H(L), this is a normal integrand (see Remark 2.6.15) and so we can find $\widehat{L}_n^k(t,u)$, $k \geq 1$, Carathéodory integrands such that

$$\begin{aligned} &\widehat{L}_n^k(t,u) \uparrow \widehat{L}_n(t,u) \qquad \text{for almost all } t\in T, \text{ all } u \in \overline{B}_r \text{ as } k \longrightarrow\infty \\ \text{and} \quad & \alpha_0(t) - c_0\|x_n(t)\| \leq \widehat{L}_n^k(t,u) \qquad \text{for almost all } t\in T, \text{ all } u \in \overline{B}_r. \end{aligned}$$

Then because $x_n \longrightarrow x$ in $C(T,\mathbb{R}^N)$ and $\lambda_n \xrightarrow{w^*} \lambda$ in $L^\infty\big(T, M(\overline{B}_r)_{w^*}\big)$, we have

$$\int_0^b \int_{\overline{B}_r} L_n^k\big(t, x_n(t), u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t \longrightarrow \int_0^b \int_{\overline{B}_r} L^k\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t,$$

for all $k \geq 1$, as $n \to \infty$. Also from the monotone convergence theorem, we have

$$\int_0^b \int_{\overline{B}_r} L^k\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \uparrow \int_0^b \int_{\overline{B}_r} L\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t$$
$$\text{as } k \longrightarrow \infty.$$

Therefore according to Proposition 1.5.25 we can find a sequence $\{k(n)\}_{n \geq 1}$ such that $k(n) \longrightarrow +\infty$ as $n \to \infty$ and

$$\int_0^b \int_{\overline{B}_r} L^{k(n)}\big(t, x_n(t), u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t \longrightarrow \int_0^b \int_{\overline{B}_r} L\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t.$$

Because $\int_0^b \int_{\overline{B}_r} L^{k(n)}\big(t, x_n(t), u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t \leq \int_0^b L\big(t, x_n(t), u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t$ for all $n \geq 1$, we have

$$\int_0^b \int_{\overline{B}_r} L\big(t, x(t), u\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \leq \liminf_{n \to \infty} \int_0^b \int_{\overline{B}_r} L\big(t, x_n(t), u\big)\lambda_n(t)(\mathrm{d}u)\mathrm{d}t$$
$$\Rightarrow \ J_r^1(x, \lambda) \leq \liminf_{n \to \infty} J_r^1(x_n, \lambda_n) = m_r^1.$$

Because (x, λ) is admissible for the relaxed problem (2.85), we conclude that

$$J_r^1(x, \lambda) = m_r^1.$$

$\square$

To establish the admissibility of the relaxed problem (2.85), we need to have that the original states are dense in the relaxed ones for the $C(T, \mathbb{R}^N)$-norm. This gives equality of the values of the original and relaxed problems. To do this we need one more result from multivalued analysis which we state here for convenience and postpone its proof until Section 6.4.

PROPOSITION 2.6.17 *If Ω, Σ, μ is a σ-finite nonatomic measure space, X is a separable Banach space, $F\colon \Omega \longrightarrow 2^X \setminus \{\emptyset\}$ a multifunction such that* $\mathrm{Gr}\, F = \{(\omega, x) \in \Omega \times X : x \in F(\omega)\} \in \Sigma \times \mathbf{B}(X)$, *and for some $1 \leq p \leq \infty$, $S_F^p = \{u \in L^p(\Omega, X) : u(\omega) \in F(\omega)$ μ-a.e.$\} \neq \emptyset$, then denoting by w (resp., w^*) the weak (resp., weak*) topology on $L^p(\Omega, X), 1 \leq p < \infty$ (resp., on $L^\infty(\Omega, X)$) we have*

$$\overline{S_F^p}^{\,w} = S^p_{\overline{\mathrm{conv}} F} \quad 1 \leq p < \infty \qquad \textit{(resp., } \overline{S_F^p}^{\,w^*} = S^p_{\overline{\mathrm{conv}} F} \textit{ for } p = +\infty).$$

Here by $\overline{\mathrm{conv}} F$, we denote the multifunction $\omega \longrightarrow \overline{\mathrm{conv}}\, F(\omega)$ and

$$S^p_{\overline{\mathrm{conv}} F} = \big\{u \in L^p(\Omega, X) : u(\omega) \in \overline{\mathrm{conv}}\, F(\omega) \ \mu\text{-a.e. on } \Omega\big\}.$$

With the help of this abstract result for multifunctions, we are able to show the density for the $C(T, \mathbb{R}^N)$-norm topology of the original states in the relaxed ones (second requirement for the admissibility of a relaxation method). For this we need to strengthen our hypotheses on f as follows.

$\mathbf{H(f)_1}$: $f\colon T\times\mathbb{R}^N\times\mathbb{R}^m\longrightarrow\mathbb{R}^N$ is a function, such that

(i) For all $(x,u)\in\mathbb{R}^N\times\mathbb{R}^m$, $t\longrightarrow f(t,x,u)$ is measurable.
(ii) For almost all $t\in T$, all $x,y\in\mathbb{R}^N$, and all $u\in U(t)$, we have

$$\|f(t,x,u)-f(t,y,u)\|\le k(t)\|x-y\| \qquad \text{with } k\in L^1(T)_+$$

and for almost all $t\in T$ and all $x\in\mathbb{R}^N$, $u\longrightarrow f(t,x,u)$ is continuous.
(iii) For almost all $t\in T$, all $x\in\mathbb{R}^N$ and all $u\in U(t)$

$$\|f(t,x,u)\|\le\alpha(t)+c(t)\|x\| \qquad \text{with } \alpha,c\in L^1(T)_+.$$

In what follows by $\alpha^{\mathbf{1}}_{\mathbf{ad}}$ we denote the set of admissible states for the original system (see problem (P)) and by $\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_1}}$ we denote the set of admissible states for the relaxed system (2.85).

PROPOSITION 2.6.18 *If hypotheses $H(f)_1$, $H(U)_1$ and $H(L)$ hold, then $\overline{\alpha^{\mathbf{1}}_{\mathbf{ad}}}=\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_1}}$ the closure taken in the $C(T,\mathbb{R}^N)$-norm.*

PROOF: For every $(t,x)\in T\times\mathbb{R}^N$, let

$$F(t,x)=f\big(t,x,U(t)\big)$$
$$\text{and}\quad G(t,x)=\Big\{\int_{\overline{B}_r}f(t,x,u)\mu(\mathrm{d}u):\mu\in\Sigma(t)\Big\}.$$

We claim that conv $F(t,x)=G(t,x)$. Note that $f\big(t,x,U(t)\big)$ is compact and then so is conv $F(t,x)$. Also it is clear that $F(t,x)\subseteq G(t,x)$ and the latter is a convex set. Moreover, $M^1_+(\overline{B}_r)$ furnished with the narrow topology (see Remark 2.6.13), which due to the compactness of $\overline{B}_r$ coincides with the relative w^*-topology on $M(\overline{B}_r)$ = space of Radon measures on $\overline{B}_r$ (recall that $C(\overline{B}_r)^*=M(\overline{B}_r)$), is compact metrizable. Let $\{v_n\}_{n\ge1}\subseteq G(t,x)$ and suppose $v_n\longrightarrow v$ in $\mathbb{R}^N$. We have

$$v_n=\int_{\overline{B}_r}f(t,x,u)\mu_n(\mathrm{d}u) \qquad \text{with } \mu_n\in\Sigma(t).$$

We may assume that $\mu_n\longrightarrow\mu$ narrowly (equivalently in the relative w^*-topology) in $M^1_+(\overline{B}_r)$. So from the portmanteau theorem we have

$$\limsup_{n\to\infty}\mu_n\big(U(t)\big)\le\mu\big(U(t)\big),$$
$$\Rightarrow\ 1\le\mu\big(U(t)\big).$$

Because $\mu\in M^1_+(\overline{B}_r)$, we have $\mu\big(U(t)\big)=1$ and so $\mu\in\Sigma(t)$. Also

$$v_n=\int_{\overline{B}_r}f(t,x,u)\mu_n(\mathrm{d}u)\longrightarrow\int_{\overline{B}_r}f(t,x,u)\mu(\mathrm{d}u)=v\in G(t,x).$$

Therefore $G(t,x)$ is closed and so

$$\text{conv } F(t,x)\subseteq G(t,x). \tag{2.86}$$

On the other hand if $v \in G(t,x)$, we have $v = \int_{\overline{B}_r} f(t,x,u)\mu(\mathrm{d}u), \mu \in \Sigma(t)$. We can find $s_n = \sum\limits_{k=1}^{n} r_k \delta_{u_k}$ with $u_k \in U(t)$ and $\{r_k\}_{n \geq 1} \subseteq [0,1]$, $\sum\limits_{k=1}^{n} r_k = 1$ such that $s_n \xrightarrow{w^*} \mu$ in $M_+^1(\overline{B}_r)$. Then

$$v_n = \int_{\overline{B}_r} f(t,x,u) s_n(\mathrm{d}u) = \sum_{k=1}^{n} r_k f(t,x,u_k) \in \operatorname{conv} F(t,x)$$

and $v_n \longrightarrow v = \int_{\overline{B}_r} f(t,x,u)\mu(\mathrm{d}u)$. So it follows

$$\operatorname{conv} F(t,x) = G(t,x) \quad (\text{see } (2.86)). \tag{2.87}$$

Let $S_{\operatorname{conv}F}(x_0)$ be the solution set of (2.82) when the multivalued right hand side is conv F and by $S_G(x_0)$ the solution set of (2.82) when the multivalued right hand is G. Then from (2.87) it follows that

$$S_{\operatorname{conv}F}(x_0) = S_G(x_0). \tag{2.88}$$

Because of hypothesis H(f)(ii) for almost all $t \in T$, $x \longrightarrow \operatorname{conv}F(t,x)$ is Lipchitz continuous for the Hausdorff metric. So invoking the relaxation theorem for differential inclusions (see Denkowski–Migórski–Papageorgiou [195, p. 262], we know that

$$S_{\operatorname{conv}F}(x_0) = \overline{S_F(x_0)} \qquad \text{the closure taken in } C(T,\mathbb{R}^N). \tag{2.89}$$

Because $\alpha^{\mathbf{1}}_{\mathbf{ad}} = S_F(x_0)$ and $\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_1}} = S_G(x_0)$, from (2.88) and (2.89) we conclude that

$$\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_1}} = \overline{\alpha^{\mathbf{1}}_{\mathbf{ad}}} \qquad \text{the closure taken in } C(T,\mathbb{R}^N).$$

$\square$

REMARK 2.6.19 It is well-known that the relaxation theorem for differential inclusions fails if the multivalued nonlinearity $F(t,x)$ is only continuous in the Hausdorff metric in the $x \in \mathbb{R}^N$-variable. For this reason we had to strengthen the hypothesis on the vector field $x \longrightarrow f(t,x,u)$ (see hypothesis H$(f)_1$(iii)).

The final step in the analysis of the relaxed problem (2.85) is to show that $m_r^1 = m$ (relaxability). This implies that the relaxation method of problem (2.85) is in fact admissible. To achieve this, we strengthen our hypotheses on the cost integrand L as follows.

H(L)$_1$: $L\colon T \times \mathbb{R}^N \times \mathbb{R}^m \longrightarrow \mathbb{R}$ is an integrand such that

(i) For all $(x,u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \longrightarrow L(t,x,u)$ is measurable.
(ii) For almost all $t \in T$, $(x,u) \longrightarrow L(t,x,u)$ is continuous.
(iii) For almost all $t \in T$, all $x \in \mathbb{R}^N$ with $\|x\| \leq n, n \geq 1$, and all $u \in U(t)$,

$$|L(t,x,u)| \leq \eta_n(t) \qquad \text{with } \eta_n \in L^1(T)_+.$$

THEOREM 2.6.20 *If hypotheses $H(f)_1$, $H(U)_1$, and $H(L)_1$ hold, then $m_r^1 = m$.*

PROOF: From Proposition 2.6.16 we know that we can find $(\widehat{x}, \widehat{\lambda})$ an admissible state-control pair for the relaxed problem (2.85) such that

$$J_r^1(\widehat{x}, \widehat{\lambda}) = m_r^1.$$

Also by virtue of Propositions 2.6.17 and 2.6.18 we can find $(x_n, u_n) \in \alpha_{\mathbf{ad}}$ such that

$$x_n \longrightarrow \widehat{x} \text{ in } C(T, \mathbb{R}^N) \qquad \text{and} \qquad \delta_{u_n} \xrightarrow{w^*} \widehat{\lambda} \text{ in } L^\infty\big(T, M(\overline{B}_r)_{w^*}\big) \text{ as } n \to \infty.$$

Set $\widehat{L}_n(t)(u) = L\big(t, x_n(t), u\big)$ and $\widehat{L}(t)(u) = L\big(t, \widehat{x}(t), u\big)$. Evidently $\widehat{L}_n \longrightarrow \widehat{L}$ in $L^1\big(T, C(\overline{B}_r)\big)$. So if by $\langle \cdot, \cdot \rangle$ we denote the duality brackets for the pair $\Big(L^1\big(T, C(\overline{B}_r)\big), L^\infty\big(T, M(\overline{B}_r)_{w^*}\big)\Big)$, we have

$$\begin{aligned} \int_0^b L\big(t, x_n(t), u_n(t)\big) \mathrm{d}t &= \Big\langle \widehat{L}_n, \delta_{u_n} \Big\rangle \longrightarrow \Big\langle \widehat{L}, \widehat{\lambda} \Big\rangle \\ &= \int_0^b \int_{\overline{B}_r} L\big(t, \widehat{x}(t), u\big) \widehat{\lambda}(t)(\mathrm{d}u) \mathrm{d}t = m_r^1, \\ &\Rightarrow m \leq m_r^1. \end{aligned}$$

Because the opposite inequality is clearly true, we conclude that $m = m_r^1$. $\square$

So we have established the admissibility of the relaxation method based on Young measures.

The next relaxation method is suggested by the reduction method used in the proof of Theorem 2.6.11. We introduce the following relaxed problem:

$$\left\{ \begin{array}{l} \qquad J_r^2(x, x') = \int_0^b p^{**}\big(t, x(t), x'(t)\big) \mathrm{d}t \longrightarrow \inf = m_r^2 \\ \text{s.t.} \quad x'(t) \in \operatorname{conv} F\big(t, x(t)\big) \text{ a.e. on } T, \quad x(0) = x_0 \end{array} \right\}. \tag{2.90}$$

Here p^{**} stands for the second conjugate (in the sense of Definition 1.2.15) of the function $v \longrightarrow p(t, x, v)$. Also as before $F(t, x) = f\big(t, x, U(t)\big)$. Note that problem (2.90) is control-free (a calculus of variations problem with multivalued dynamic constraints). In what follows by $\alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_2}}$ we denote the set of admissible states for this new problem (2.90).

PROPOSITION 2.6.21 *If hypotheses $H(f)_1$, $H(U)_1$, and $H(L)$ hold, then $\alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_1}} = \alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_2}}$, $m_r^1 = m_r^2$, and there exists $\widehat{x} \in \alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_2}}$ such that*

$$J_r^2(\widehat{x}, \widehat{x}') = m_r^2.$$

PROOF: From the proof of Proposition 2.6.18, we know that

$$\operatorname{conv} F(t, x) = \Big\{ \int_{\overline{B}_r} f(t, x, u) \mu(\mathrm{d}u) : \mu \in \Sigma(t) \Big\}. \tag{2.91}$$

From (2.91) it follows that $\alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_1}} = \alpha^{\mathbf{1}}_{\mathbf{ad}, \mathbf{r_2}}$.

By definition we have

$$p^{**}(t,x,v) = \inf\Big\{\eta \in \mathbb{R} : (v,\eta) \in \overline{\text{conv}}\ \text{epi}\ p(t,x,\cdot)\Big\}. \tag{2.92}$$

As always $\inf \emptyset = +\infty$. Set

$$H(t,x) = \text{epi}\ p(t,x,\cdot) = \big\{\big(f(t,x,u),\eta\big) : u \in U(t),\ L(t,x,u) \le \eta\big\}.$$

As in the proof of Proposition 2.6.18, we can show that

$$\begin{aligned}\overline{\text{conv}}\ H(t,x) &= \overline{\text{conv}}\ \text{epi}\ p(t,x,\cdot)\\ &= \Big\{\Big(\int_{\overline{B}_r} f(t,x,u)\mu(\mathrm{d}u),\eta\Big) : \mu \in \Sigma(t),\ \int_{\overline{B}_r} L(t,x,u)\mu(\mathrm{d}u) \le \eta\Big\}.\end{aligned}$$

Hence we have

$$p^{**}(t,x,v) = \inf\Big\{\int_{\overline{B}_r} L(t,x,u)\mu(\mathrm{d}u) : \mu \in \Sigma(t),\ v = \int_{\overline{B}_r} f(t,x,u)\mu(\mathrm{d}u)\Big\}. \tag{2.93}$$

Because $\mu \longrightarrow \int_{\overline{B}_r} L(t,x,u)\mu(\mathrm{d}u)$ is narrowly lower semicontinuous and $M_+^1(\overline{B}_r)$ is narrowly compact, from Theorem 2.1.10, we see that the infimum in (2.93) is attained and an application of Theorem 2.6.8 implies that the minimizing measure depends measurable on $t \in T$. Combining this fact with Proposition 2.6.16, we infer that $m_r^1 = m_r^2$ and for $\widehat{x}$ the optimal state for problem (2.85), we also have

$$J_r^2(\widehat{x},\widehat{x}') = m_r^2.$$

□

Combining Propositions 2.6.18 and 2.6.21, we see the following.

PROPOSITION 2.6.22 *If hypotheses H(f), H(U)$_1$, and H(L) hold, then* $\overline{\alpha^{\mathbf{1}}_{\mathbf{ad}}} = \alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_2}}$, *the closure taken in the* $C(T,\mathbb{R}^N)$ *norm.*

Now we can establish the admissibility of the second relaxed problem (2.90).

THEOREM 2.6.23 *If hypotheses H(f), H(U)$_1$, and H(L)$_1$ hold, then* $m_r^2 = m$.

PROOF: From Theorem 2.6.20 we know that $m_r^1 = m$. On the other hand Proposition 2.6.21 says that $m_r^1 = m_r^2$. So finally $m_r^2 = m$. □

So the second relaxation method, which was motivated from the reduction method, is also admissible.

The third relaxed problem has its roots in the well-known theorem of Carathéodory for the representation of convex sets in $\mathbb{R}^N$. Let us recall that theorem.

THEOREM 2.6.24 *If* $C \subseteq \mathbb{R}^N$ *is nonempty, then every point of the set* conv C *is a convex combination of no more than* $n+1$ *distinct points of the set* C.

Motivated by this result, we introduce the following relaxed problem. In what follows $\widehat{u} = (u_k)_{k=1}^{N+1}$ and $\widehat{r} = (r_k)_{k=1}^{N+1}$.

$$\left\{\begin{array}{l} J_r^3(x,\widehat{u},\widehat{r}) = \int_0^b \sum\limits_{k=1}^{N+1} r_k(t) L\big(t, x(t), u_k\big) \mathrm{d}t \longrightarrow \inf = m_r^3 \\ \text{s.t.} \quad x'(t) = \sum\limits_{k=1}^{N+1} r_k(t) f\big(t, x(t), u_k(t)\big) \text{ a.e. on } T,\ x(0) = x_0 \\ u_k \in L^1(T, \mathbb{R}^m),\ u(t) \in U(t) \text{ a.e. on } T \\ r_k : T \longrightarrow [0,1] \text{ measurable } \sum_{k=1}^{N+1} r_k(t) = 1 \qquad \text{for all } t \in T \end{array}\right\}. \tag{2.94}$$

By $\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_3}}$ we denote the admissible states for problem (2.94). From the proof of Proposition 2.6.21 it is clear that we have the following.

PROPOSITION 2.6.25 *If hypotheses $H(f)$, $H(U)_1$, and $H(L)$ hold, then $\alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_1}} = \alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_2}} = \alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_3}}$, $m_r^1 = m_r^2 = m_r^3$, $\overline{\alpha^{\mathbf{1}}_{\mathbf{ad}}} = \alpha^{\mathbf{1}}_{\mathbf{ad},\mathbf{r_3}}$ the closure taken in the $C(T,\mathbb{R}^N)$-norm, and there exists an optimal triple $(\widehat{x},\widehat{u}^*,\widehat{r}^*) = m_r^3$ from problem (2.94); that is,*

$$J_r^3(\widehat{x},\widehat{u}^*,\widehat{r}^*) = m_r^3.$$

Also we have the admissibility of the new relaxed problem (2.94).

THEOREM 2.6.26 *If hypotheses $H(f)$, $H(U)_1$, and $H(L)_1$ hold, then $m_r^3 = m$.*

Next we derive a fourth relaxation method based on semicontinuity techniques, which are different from the ones employed so far. The new approach uses the so-called multiple Γ-operators which are related to the Γ-convergence that we studied in Section 1.5.

DEFINITION 2.6.27 Let X_1, X_2 be two Hausdorff topological spaces, and let $\varphi\colon X_1 \times X_2 \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be a proper function and $(x_1, x_2) \in X_1 \times X_2$. In what follows by $\mathcal{Z}(+)$ we denote the *sup* operator and by $\mathcal{Z}(-)$ the *inf* operator. Also for $k = 1, 2$, let S_k be the set of sequences in X_k that converge to x_k and let β_k be one of the signs $+$ and $-$. We define

$$\Gamma_{\mathrm{seq}}(X_1^{\beta_1}, X_2^{\beta_2})\varphi(x_1,x_2) = \underset{(x_n^1)\in S_1}{\mathcal{Z}(\beta_1)}\ \underset{(x_n^2)\in S_2}{\mathcal{Z}(\beta_2)}\ \underset{m\geq 1}{\mathcal{Z}(-\beta_1)}\ \underset{n\geq m}{\mathcal{Z}(\beta_1)}\ \varphi(x_n^1, x_n^2).$$

REMARK 2.6.28 Here for simplicity we have restricted ourselves to the sequential definition of multiple Γ-operators. Evidently we can have a more general topological version of the above notion, as we did for the Γ-convergence in Section 1.5.

EXAMPLE 2.6.29 By virtue of Definition 2.6.27, we have

$$\Gamma_{\mathrm{seq}}(X_1^-, X_2^-)\varphi(x_1,x_2) = \inf_{x_n^1 \to x_1} \inf_{x_n^2 \to x_2} \liminf_{n\to\infty} \varphi(x_n^1, x_n^2)$$

$$\text{and} \qquad \Gamma_{\mathrm{seq}}(X_1^-, X_2^+)\varphi(x_1,x_2) = \inf_{x_n' \to x_1} \sup_{x_n^2 \to x_2} \liminf_{n\to\infty} \varphi(x_n^1, x_n^2)$$

If the Γ_{seq}-limit is independent of the sign $+$ or $-$ associated with one of the spaces, then this sign is omitted. So if

$$\Gamma_{\text{seq}}(X_1^-, X_2^+)\varphi(x_1, x_2) = \Gamma_{\text{seq}}(X_1^+, X_2^+)\varphi(x_1, x_2)$$

then we simply write $\Gamma_{\text{seq}}(X_1, X_2^+)\varphi(x_1, x_2)$. Note that

$$\Gamma_{\text{seq}}(X_1^{\beta_1}, X_2^{\beta_2})\varphi(x_1, x_2) = -\Gamma_{\text{seq}}(X_1^{-\beta_1}, X_2^{-\beta_2})(-\varphi)(x_1, x_2)$$

and if both X_1, X_2 are metrizable, then

$$\Gamma_{\text{seq}}(X_1^-, X_2^-)\varphi(x_1, x_2) = \Gamma_{\text{seq}}\big((X_1 \times X_2)^-\big)\varphi(x_1, x_2).$$

Although in general multiple $\Gamma-$operators are not distributive with respect to addition, nevertheless we have some useful inequalities.

PROPOSITION 2.6.30 *If* $\varphi, \psi\colon X_1 \times X_2 \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ *are proper functions, then*

$$\begin{aligned}\Gamma_{\text{seq}}(X_1^-, X_2^-)\varphi + \Gamma_{\text{seq}}(X_1^-, X_2^-)\psi &\le \Gamma_{\text{seq}}(X_1^-, X_2^-)(\varphi + \psi)\\ &\le \Gamma_{\text{seq}}(X_1^-, X_2^+)\varphi + \Gamma_{\text{seq}}(X_1^+, X_2^-)\psi.\end{aligned}$$

PROOF: It is a consequence of

$$\inf_{i\in I} \lambda_i + \inf_{i\in I} \mu_i \le \inf_{i\in I}(\lambda_i + \mu_i) \le \inf_{i\in I} \lambda_i + \sup_{i\in I} \mu_i$$

for all $\{\lambda_i\}_{i\in I},\ \{\mu_i\}_{i\in I} \subseteq \overline{\mathbb{R}}$. □

COROLLARY 2.6.31 *If for* $(x_1, x_2) \in X_1 \times X_2$, $\Gamma_{\text{seq}}(X_1^-, X_2)\varphi(x_1, x_2)$, *and* $\Gamma_{\text{seq}}(X_1, X_2^-)\psi(x_1, x_2)$ *exist, then*

$$\Gamma_{\text{seq}}(X_1, X_2^-)(\varphi + \psi)(x_1, x_2) = \Gamma_{\text{seq}}(X_1^-, X_2)\varphi(x_1, x_2) + \Gamma_{\text{seq}}(X_1, X_2^-)\psi(x_1, x_2).$$

In abstract terms the idea of this new relaxation method is the following. Suppose X is the state space, Y the control space (both Hausdorff topological spaces), $\varphi\colon X \times Y \longrightarrow \overline{\mathbb{R}}$ is the cost functional, and C is the set of all admissible state-control pairs (in C we have incorporated all constraints of the problem dynamic and nondynamic). Then the optimal control problem is

$$\inf\big[(\varphi + i_C)(x, u) : (x, u) \in X \times Y\big], \tag{2.95}$$

where i_C is the indicator function of the set C; that is,

$$i_C(x, u) = \begin{cases} 0 & \text{if } (x, u) \in C \\ +\infty & \text{otherwise} \end{cases}.$$

Then the relaxed problem corresponding to (2.95) is

$$\inf\big[\Gamma_{\text{seq}}(X^-, U^-)(\varphi + i_C)(x, u) : (x, u) \in X \times Y\big];$$

in other words in the relaxed problem we minimize the Γ_{seq}-relaxation of the extended cost functional $\varphi + i_C$. So our goal is to determine $\Gamma_{\text{seq}}(X^-, U^-)(\varphi + i_C)$. We do this by employing the *auxiliary variable method*, which is outlined in the next proposition.

PROPOSITION 2.6.32 *If Z is a third Hausdorff topological space, $\xi\colon X\times Y\longrightarrow Z$ is a map that satisfies*

> $(\mathbf{C_0})$ *Every sequence $\{(x_n,u_n)\}_{n\geq 1}$ that is convergent in $X\times Y$ and $\{(\varphi+i_C)(x_n,u_n)\}_{n\geq 1}$ is bounded, the sequence $\{\xi(x_n,y_n)\}_{n\geq 1}$ has a convergent subsequence in Z,*

$$\text{and}\quad \psi(x,u,v)=\begin{cases}(\varphi+i_C)(x,u) & \text{if } v=\xi(x,u)\\ +\infty & \text{otherwise}\end{cases}, \tag{2.96}$$

then $\Gamma_{\text{seq}}(X^-,Y^-)(\varphi+i_C)(x,u)=\inf\big[\Gamma_{\text{seq}}\big(X^-,(Y\times Z)^-\big)\psi(x,u,v):v\in Z\big]$.

PROOF: Let $(x,u,v)\in X\times Y\times Z$ and suppose that $x_n\to x$ in X, $u_n\longrightarrow u$ in Y, and $v_n\longrightarrow v$ in Z. We have

$$\begin{aligned}
&(\varphi+i_C)(x_n,u_n)\leq\psi(x_n,u_n,v_n),\ n\geq 1 \qquad \text{(see (2.96))},\\
\Rightarrow\ &\Gamma_{\text{seq}}(X^-,Y^-)(\varphi+i_C)(x,u)\leq\Gamma_{\text{seq}}\big(X^-,(Y\times Z)^-\big)\psi(x,u,v)\\
&\quad\text{for all } v\in Z,\\
\Rightarrow\ &\Gamma_{\text{seq}}(X,^-,Y^-)(\varphi+i_C)(x,u)\\
&\leq\inf\big[\Gamma_{\text{seq}}\big(X^-,(X\times Y)^-\big)\psi(x,u,v):v\in Z\big].
\end{aligned}\tag{2.97}$$

On the other hand let $(x,u)\in X\times Y$ and consider $x_n\longrightarrow x$ in X and $u_n\longrightarrow u$ in Y. Without any loss of generality we assume that $\lim\limits_{n\to\infty}(\varphi+i_C)(x_n,u_n)$ exists and it is finite. Then by virtue of condition (C_0), $v_n=\xi(x_n,u_n)\longrightarrow v\in Z$. Hence using Definition 2.6.27 we have

$$\begin{aligned}
&\Gamma_{\text{seq}}\big(X^-,(Y\times Z)^-\big)\psi(x,u,v)\leq\liminf_{n\to\infty}\psi(x_n,u_n,v_n)=\lim_{n\to\infty}\varphi(x_n,u_n)\\
\Rightarrow\ &\inf\big[\Gamma_{\text{seq}}\big(X^-,(Y\times Z)^-\big)\psi(x,u,v):v\in Z\big]\leq\Gamma_{\text{seq}}(X^-,Y^-)(\varphi+i_C)(x,u).
\end{aligned}\tag{2.98}$$

From (2.97) and (2.98), we conclude that equality holds. $\square$

The next proposition is helpful in the evaluation of $\Gamma_{\text{seq}}\big(X^-,(Y\times Z)^-\big)\psi$.

PROPOSITION 2.6.33 *If the cost functional φ satisfies*

$$\varphi(x,u)\leq\varphi(x',u)+w(x,x')h(x',u)\qquad \text{for all } x,x'\in X,\ u\in Y \tag{2.99}$$

with $w\colon X\times X\longrightarrow\overline{\mathbb{R}}_+$ and $h\colon X\times Y\longrightarrow\overline{\mathbb{R}}_+$ satisfying

$$x_n\longrightarrow x\quad \text{in } X \text{ implies that } w(x,x_n),\ \ w(x_n,x)\longrightarrow 0 \tag{2.100}$$

and $\{(x_n,u_n)\}_{n\geq 1}$ is convergent in $X\times Y$ and $\{\varphi(x_n,u_n)\}_{n\geq 1}$ is bounded imply

$$\{h(x_n,u_n)\}_{n\geq 1}\ \text{ is bounded,} \tag{2.101}$$

then for all $(x,u)\in X\times Y$, $\Gamma_{\text{seq}}(X,Y^-)\,\varphi(x,u)$ exists and

$$\Gamma_{\text{seq}}(X,Y^-)\,\varphi(x,u)=\Gamma_{\text{seq}}(\delta_X,Y^-)\,\varphi(x,u)$$

with δ_X being the discrete topology on X.

PROOF: Let $(x_n, u_n) \longrightarrow (x, u)$ in $X \times Y$. Without any loss of generality we assume that $\{\varphi(x_n, u_n)\}_{n\geq 1}$ converges to a finite limit. From (2.99) we have

$$\varphi(x, u_n) \leq \varphi(x_n, u_n) + w(x, x_n)h(x_n, u_n), \quad n \geq 1.$$

Then by virtue of (2.100) and (2.101) we have

$$\begin{aligned} &\liminf_{n\to\infty} \varphi(x, u_n) \leq \liminf_{n\to\infty} \varphi(x_n, u_n), \\ \Rightarrow\ &\Gamma_{\text{seq}}(X, Y^-)\, \varphi(x, u) \leq \Gamma_{\text{seq}}(X, Y^-)\, \varphi(x, u). \end{aligned} \tag{2.102}$$

On the other hand let $(x_n, u_n) \longrightarrow (x, u)$ in $X \times Y$, and without any loss of generality assume that $\{\varphi(x, u_n)\}_{n\geq 1}$ converges to a finite limit. Then because of (2.99), we have

$$\varphi(x_n, u_n) \leq \varphi(x, u_n) + w(x_n, x)h(x, u_n), \quad n \geq 1.$$

Then by virtue of (2.100) and (2.101), we have

$$\begin{aligned} &\liminf_{n\to\infty} \varphi(x_n, u_n) \leq \lim \varphi(x, u_n), \\ \Rightarrow\ &\Gamma_{\text{seq}}(\delta_X, Y^-)\, \varphi(x, u) \leq \Gamma_{\text{seq}}(\delta_X, Y^-)\, \varphi(x, u). \end{aligned} \tag{2.103}$$

From (2.102) and (2.103) we conclude that equality must hold. $\square$

Now we introduce the specific optimal problem that we study. The state space X is $W^{1,1}\big((0,b), \mathbb{R}^N\big)$ endowed with the $C(T, \mathbb{R}^N)$-norm topology (recall that $W^{1,1}\big((0,b), \mathbb{R}^N\big)$ is embedded continuously in $C(T, \mathbb{R}^N)$) and the control space Y is $L^1(T, \mathbb{R}^m)$ endowed with the weak topology. The optimal control problem is the following.

$$\left\{ \begin{array}{lll} & J(x,u) = \int_0^b L\big(t, x(t), u(t)\big)\mathrm{d}t \longrightarrow \inf = \widehat{m}, & \\ \text{s.t.} & x'(t) = A\big(t, x(t)\big) + C\big(t, x(t)\big) g\big(t, u(t)\big) & \text{a.e. on } T \\ & x(0) = x_0, \ \ u \in L^1(T, \mathbb{R}^m), \ \ u(t) \in U(t) & \text{a.e. on } T \end{array} \right\}. \tag{2.104}$$

Here $A\colon T \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$, $C\colon T \times \mathbb{R}^N \longrightarrow \mathbb{R}^{N\times k}$, and $g\colon T \times \mathbb{R}^m \longrightarrow \mathbb{R}^k$ are measurable functions, $U\colon T \longrightarrow 2^{\mathbb{R}^m} \setminus \{\emptyset\}$ is the control constraint multifunction, and $L\colon T \times \mathbb{R}^N \times \mathbb{R}^m \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is the cost integrand. We introduce the following hypotheses on the above data of problem (2.104).

H(A): $A\colon T \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is a function such that

(i) For all $x \in \mathbb{R}^N$, $t \longrightarrow A(t, x)$ is measurable.
(ii) For every $r > 0$, there exists $\beta_r \in L^1(T)_+$ such that for almost all $t \in T$ and all $\|x\|, \|y\| \leq r$, we have

$$\|A(t, x) - A(t, y)\| \leq \beta_r(t)\|x - y\|.$$

(iii) For almost all $t \in T$ and all $x \in \mathbb{R}^N$, we have

$$\|A(t, x)\| \leq \alpha(t) + c(t)\|x\| \quad \text{with } \alpha,\ c \in L^1(T)_+.$$

H(C): $C\colon T\times\mathbb{R}^N\longrightarrow\mathbb{R}^{N\times k}$ is a function such that

(i) For all $x\in\mathbb{R}^N$, $t\longrightarrow C(t,x)$ is measurable.
(ii) For every $r>0$, there exists $\gamma_r\in L^\infty(T)_+$ such that for almost all $t\in T$ and all $\|x\|,\|y\|\leq r$, we have

$$\|C(t,x)-C(t,y)\|\leq\gamma_r(t)\|x-y\|.$$

(iii) For almost all $t\in T$ and all $x\in\mathbb{R}^N$, we have

$$\|C(t,x)\|\leq\alpha_1(t)+c_1(t)\|x\|\quad\text{ with } \alpha_1,\ c_1\in L^\infty(T)_+.$$

H(g): $g\colon T\times\mathbb{R}^m\longrightarrow\mathbb{R}^k$ is a Borel function.

H(U)$_2$: $U\colon T\longrightarrow 2^{\mathbb{R}^m}\setminus\{\emptyset\}$ is a multifunction with compact values, such that Gr $U=\{(t,u)\in T\times\mathbb{R}^m : u\in U(t)\}\in\mathbf{B}(T)\times\mathbf{B}(\mathbb{R}^m)$ and there exists $u\in L^1(T,\mathbb{R}^m)$ such that $u(t)\in U(t)$ a.e. on T and $g\big(\cdot,u(\cdot)\big)$ is integrable on T.

H(L)$_2$: $L\colon T\times\mathbb{R}^N\times\mathbb{R}^m\longrightarrow\overline{\mathbb{R}}_+=\mathbb{R}_+\cup\{+\infty\}$ is an integrand such that

(i) $(t,x,y)\longrightarrow L(t,x,y)$ is measurable.
(ii) For every $r>0$, almost all $t\in T$, all $\|x\|,\|y\|\leq r$ and all $u\in U(t)$, we have

$$L(t,x,u)\leq L(t,y,u)+\vartheta_r(t,\|x-y\|)+\sigma_r(\|x-y\|)L(t,y,u),$$

where $\vartheta_r\colon T\times\mathbb{R}_+\longrightarrow\mathbb{R}_+$ and $\sigma_r\colon\mathbb{R}_+\longrightarrow\mathbb{R}_+$ are such that $\vartheta_r(\cdot,\lambda)$ is integrable and $\vartheta_r(t,\cdot)$, $\sigma_r(\cdot)$ are continuous increasing with $\vartheta_r(t,0)=\sigma_r(0)=0$.

(iii) For almost all $t\in T$ and all $u\in U(t)$

$$\nu(\|g(t,x)\|+\|u\|)-\zeta(t)\leq L(t,0,u)$$

with $\nu\colon\mathbb{R}_+\longrightarrow\mathbb{R}_+$ increasing convex function satisfying $\lim\limits_{r\to+\infty}\big(\nu(\tau)/\tau\big)=+\infty$ and $\zeta\in L^1(T)_+$.
(iv) There exist $u_0\in L^1(T,\mathbb{R}^m)$ such that $t\longrightarrow L\big(t,0,u_0(t)\big)$ is integrable on T.

We implement the method of the auxiliary variable, outlined earlier. Let $S=\{(x,u)\in W^{1,1}\big((0,b),\mathbb{R}^N\big)\times L^1(T,\mathbb{R}^m) : (x,u)$ is an admissible pair for problem (2.104)$\}$ and set $\varphi=J+i_S$. The next lemma shows that the method of auxiliary variable indeed can be used within the context of problem (2.104).

LEMMA 2.6.34 *If $Z=L^1(T,\mathbb{R}^k)$ furnished with the weak topology and $\xi\colon X\times Y\longrightarrow Z$ is defined by*

$$\xi(x,u)(t)=\begin{cases}g\big(t,u(t)\big) & \textit{if } g\big(\cdot,u(\cdot)\big) \textit{ is integrable,}\\ 0 & \textit{otherwise}\end{cases},$$

then φ satisfies the compactness condition (C_0) of Proposition 2.6.32.

PROOF: Suppose $(x_n, u_n) \longrightarrow (x,u)$ in $X \times Y$ and assume that $|\varphi(x_n,u_n)| \leq M$ for some $M>0$, all $n \geq 1$. Let $v_n(t) = g\big(t, u_n(t)\big)$. If $\varepsilon_n = \|x_n - x\|_{C(T,\mathbb{R}^N)}$ and $r = \sup\limits_{n\geq 1} \|x_n\|_{C(T,\mathbb{R}^N)}$ (recall that on the state space $X = W^{1,1}\big((0,b),\mathbb{R}^N\big)$ we consider the $C(T,\mathbb{R}^N)$-norm topology), using hypothesis H(L), we have

$$\begin{aligned}\int_0^b \Big(\psi(\|v_n(t)\|) - \zeta(t)\Big)\mathrm{d}t &\leq \int_0^b L\big(t,0,u_n(t)\big)\mathrm{d}t \\ &\leq \int_0^b \Big(L\big(t,x_n(t),u_n(t)\big)\mathrm{d}t + \vartheta_r(t,\varepsilon_n)\Big)\mathrm{d}t \\ &\quad + \int_0^b \sigma_r(\varepsilon_n)\; L\big(t,x_n(t),u_n(t)\big)\mathrm{d}t \\ &\leq M_1\big(1+\sigma_r(\varepsilon_n)\big) + \int_0^b \vartheta_r(t,\varepsilon_n)\mathrm{d}t \end{aligned} \tag{2.105}$$

for some $M_1 > 0$, all $n \geq 1$.

Then from (2.105), hypothesis H$(L)_2$(iii), and the De La Vallée–Poussin theorem, we conclude that $\{v_n\}_{n\geq 1} \subseteq L^1(T,\mathbb{R}^k)$ has a weakly convergent subsequence. $\square$

This lemma permits the use of Proposition 2.6.32, which says the relaxation of the extended cost functional $\varphi = J + i_S$ is reduced to the relaxation in $X \times (Y \times Z)$ of

$$\psi(x,u,v) = \widehat{J}(x,u,v) + i_S(x,u,v),$$

where $\widehat{J}(x,u,v) = \int_0^b \widehat{L}\big(t,x(t),u(t),v(t)\big)\mathrm{d}t$ with $\widehat{L}(t,x,u,v) = L(t,x,u) + i_{\{v=g(t,u)\}}$ and $\widehat{S} = \Big\{(x,u,v) \in X \times Y \times Z : \;\; x'(t) = A\big(t,x(t)\big) + C\big(t,x(t)\big)v(t)$ a.e. on $T,\ x(0) = x_0\Big\}$.

We compute separately $\Gamma_{\mathrm{seq}}\big(X,(Y\times Z)^-\big)\widehat{J}(x,u,v)$ and $\Gamma_{\mathrm{seq}}(X^-, Y \times Z) i_{\widehat{C}}(x,u,v)$ and then use Corollary 2.6.31 to have $\Gamma_{\mathrm{seq}}\big(X^-,(Y\times Z)^-\big)\psi(x,u,v)$, which is what we want.

PROPOSITION 2.6.35 *For all $(x,u,v) \in X \times Y \times Z$ we have*

$$\Gamma_{\mathrm{seq}}\big(X,(Y\times Z)^-\big)\widehat{J}(x,u,v) = \int_0^b \widehat{L}^{**}\big(t,x(t),u(t),v(t)\big)\mathrm{d}t,$$

where $\widehat{L}^{**}(t,x,u,v)$ denotes the second conjugate of $\widehat{L}(t,x,\cdot,\cdot)$.

PROOF: From hypothesis H$(L)_2$(ii) we have

$$\begin{aligned}\widehat{J}(x,u,v) \leq \widehat{J}(y,u,v) &+ \int_0^b \vartheta_r(t, \|x-y\|_{C(T,\mathbb{R}^N)})\mathrm{d}t \\ &+ \;\sigma_r(\|x-y\|_{C(T,\mathbb{R}^N)})\widehat{J}(y,u,v)\end{aligned}$$

for all $x,y \in X$, with $\|x\|_{C(T,\mathbb{R}^N)},\ \|y\|_{C(T,\mathbb{R}^N)} \leq r, r>0$ and all $u \in Y,\ v \in Z$. Set

$$w(x,y) = \int_0^b \vartheta_r(t,\|x-y\|_{C(T,\mathbb{R}^N)})\mathrm{d}t + \sigma_r(\|x-y\|_{C(T,\mathbb{R}^N)})$$

and $\quad h(x,u,v) = 1 + \widehat{J}(x,u,v).$

We can apply Proposition 2.6.33 and have

$$\Gamma_{\text{seq}}\big(X, (Y \times Z)^-\big)\widehat{J}(x,u,v) = \Gamma_{\text{seq}}\big(\delta_X, (Y \times Z)^-\big)\widehat{J}(x,u,v).$$

But $\Gamma_{\text{seq}}\big(\delta_X, (Y \times Z)^-\big)\widehat{J}(x,u,v) = \int_0^b \widehat{L}^{**}\big(t, x(t), u(t), v(t)\big)\mathrm{d}t$ (see Buttazzo [126, p. 74], and Denkowski–Migórski–Papageorgiou [195, p. 588]). $\square$

PROPOSITION 2.6.36 *For all* $(x,u,v) \in X \times Y \times Z$ *we have*

$$\Gamma_{\text{seq}}(X^-, Y \times Z)\chi_{\widehat{S}}(x,u,v) = \chi_{\widehat{S}}(x,u,v).$$

PROOF: We need to show that

$$\begin{aligned} &\text{if } (x_n, u_n, v_n) \longrightarrow (x,u,v) \text{ in } X \times Y \times Z \text{ and } (x_n, u_n, v_n) \in \widehat{S} \\ &\quad \text{for all } n \geq 1, \text{ then } (x,u,v) \in \widehat{S} \end{aligned} \tag{2.106}$$

and

$$\begin{aligned} &\text{if } (x,u,v) \in \widehat{S} \text{ and } (u_n, v_n) \longrightarrow (u,v) \text{ in } Y \times Z, \\ &\text{then there exists } x_n \longrightarrow x \text{ in } X \text{ with } (x_n, u_n, v_n) \in \widehat{S},\ n \geq 1. \end{aligned} \tag{2.107}$$

First we prove (2.106). We have

$$\left\{ \begin{array}{l} x_n'(t) = A\big(t, x_n(t)\big) + C\big(t, x_n(t)\big)v_n(t) \text{ a.e. on } T \\ x_n(0) = x_0 \end{array} \right\}.$$

By virtue of hypotheses H(A) and H(C) and because $v_n \xrightarrow{w} v$ in $L^1(T, \mathbb{R}^k)$ (recall that $Z = L^1(T, \mathbb{R}^k)$ is furnished with the weak topology), we have

$$A\big(\cdot, x_n(\cdot)\big) + C\big(\cdot, x_n(\cdot)\big)v_n(\cdot) \xrightarrow{w} A\big(\cdot, x(\cdot)\big) + C\big(\cdot, x(\cdot)\big)v(\cdot) \text{ in } L^1(T, \mathbb{R}^N).$$

Because $x_n \longrightarrow x$ in $C(T, \mathbb{R}^N)$ (recall that $X = W^{1,1}\big((0,b), \mathbb{R}^N\big)$ is furnished with the $C(T, \mathbb{R}^N)$-norm topology), in the limit as $n \to \infty$, we have

$$\left\{ \begin{array}{l} x'(t) = A\big(t, x(t)\big) + C\big(t, x(t)\big)v(t) \text{ a.e. on } T \\ x(0) = x_0 \end{array} \right\}. \tag{2.108}$$

So we have proved (2.106).

Next let $(x,u,v) \in X \times Y \times Z$ be such that it satisfies (2.108) and suppose $(u_n, v_n) \longrightarrow (u,v)$ in $Y \times Z$. For every $n \geq 1$, let $x_n \in W^{1,1}\big((0,b), \mathbb{R}^N\big)$ be the unique solution of the Cauchy problem

$$\left\{ \begin{array}{l} x_n'(t) = A\big(t, x_n(t)\big) + C\big(t, x_n(t)\big)v_n(t) \text{ a.e. on } T \\ x_n(0) = x_0 \end{array} \right\}. \tag{2.109}$$

Because $\sup_{n \geq 1} \|v_n\|_{L^1(T, \mathbb{R}^k)} \leq M_2 < +\infty$, then from (2.109), after integration over $[0,t]$ and using hypothesis H(A)(iii), H(C)(iii), and Gronwall's inequality, we obtain $r > 0$ such that $\sup_{n \geq 1} \|v_n\|_{L^1(T, \mathbb{R}^k)} \leq r$. From the Arzela–Ascoli theorem it follows that $\{x_n\}_{n \geq 1} \subseteq C(T, \mathbb{R}^N)$ is relatively compact and so we may assume that $x_n \longrightarrow \widehat{x}$ in $C(T, \mathbb{R}^N)$. In the limit we have

$$\left\{\begin{array}{l} \widehat{x}'(t) = A\big(t, \widehat{x}(t)\big) + C\big(t, \widehat{x}(t)\big)v(t) \text{ a.e. on } T \\ x(0) = x_0 \end{array}\right\}.$$

Given the uniqueness of the solution of (2.108), it follows that $\widehat{x} = x$ and so we have proved (2.107) and also the proposition. $\square$

Combining Propositions 2.6.35 and 2.6.36 with Corollary 2.6.31, we have

$$\Gamma_{\text{seq}}(X^-, Y^-)\varphi(x,u) = \inf\Big\{ \int_0^b \widehat{L}^{**}\big(t, x(t), u(t), v(t)\big)\mathrm{d}t : v \in Z,$$
$$x'(t) = A\big(t, x(t)\big) + C\big(t, x(t)\big)v(t) \text{ a.e. on T}, \ x(0) = x_0\Big\}.$$

The final step in the derivation of the relaxed problem, is to eliminate the variable $v \in Z$ from the right-hand side expression. To this end let

$$G(t,u) = \big\{ v \in \mathbb{R}^k : (u,v) \in \overline{\text{conv}}\{(u', v') \in \mathbb{R}^m \times \mathbb{R}^k : v' = g(t, u')\}\big\}$$
$$\overline{L}(t, x, u, w) = \inf\big\{\widehat{L}^{**}(t, x, u, v) : w = A(t,x) + C(t,x)v\big\}$$
$$\overline{S} = \big\{(x,u) \in X \times Y : \ x'(t) \in A\big(t, x(t)\big) + C\big(t, x(t)\big)G\big(t, u(t)\big)$$
$$\text{a.e. on } T, \ x(0) = x_0\big\}.$$

THEOREM 2.6.37 *If hypotheses $H(A)$, $H(C)$, $H(g)$, $H(U)_2$, and $H(L)_2$ hold, then* $\Gamma_{\text{seq}}(X^-, Y^-)\ \varphi(x,u) = \int_0^b \overline{L}\big(t, x(t), u(t), x'(t)\big)\mathrm{d}t + i_{\overline{S}}\ (x,u)$.

PROOF: Note that $v \in G(t,u)$, when $\widehat{L}^{**}(t, x, u, v) < +\infty$. Therefore

$$\Gamma_{\text{seq}}(X^-, Y^-)\ \varphi(x,u) = \inf\Big\{ \int_0^b \widehat{L}^{**}\big(t, x(t), u(t), x'(t)\big)\mathrm{d}t : v \in Z, \ x'(t)$$
$$= A\big(t, x(t)\big) + C\big(t, x(t)\big)v(t) \text{ a.e. on } T\Big\} + i_{\overline{S}}\ (x,u).$$

Note that if $x'(t) = A\big(t, x(t)\big) + C\big(t, x(t)\big)v(t)$ a.e. on T, we have

$$\overline{L}\big(t, x(t), u(t), x'(t)\big) \le \widehat{L}^{**}\big(t, x(t), u(t), v(t)\big) \quad \text{a.e. on } T,$$
$$\Rightarrow \int_0^b \overline{L}\big(t, x(t), u(t), x'(t)\big)\mathrm{d}t + i_{\overline{S}}\ (x,u) \le \Gamma_{\text{seq}}(X^-, Y^-)\ \varphi(x,u). \tag{2.110}$$

On the other hand if $(x,u) \in \overline{S}$ and $\int_0^b \overline{L}\big(t, x(t), u(t), x'(t)\big)\mathrm{d}t < +\infty$, then $\overline{L}\big(t, x(t), u(t), x'(t)\big) < +\infty$ a.e. on T and so we must have

$$\overline{L}\big(t, x(t), u(t), x'(t)\big) = \widehat{L}^{**}\big(t, x(t), u(t), v(t)\big) \quad \text{a.e. on } T,$$

with $v \in L^1(T, \mathbb{R}^k)$ such that $x'(t) = A\big(t, x(t)\big) + C\big(t, x(t)\big)v(t)$ a.e. on T (see Theorem 2.6.8 and hypothesis $\mathrm{H}(L)_2$). It follows that

$$\Gamma_{\text{seq}}(X^-, Y^-)\ \varphi(x,u) \le \int_0^b \widehat{L}^{**}\big(t, x(t), u(t), v(t)\big)\mathrm{d}t + i_{\overline{S}}\ (x,u)$$
$$= \int_0^b \overline{L}\big(t, x(t), u(t), x'(t)\big)\mathrm{d}t + i_{\overline{S}}\ (x,u). \tag{2.111}$$

Comparing (2.110) and (2.111), we obtain the desired equality. □

So the new relaxed problem corresponding to (2.104) is:

$$\inf\Big[\int_0^b \overline{L}\big(t,x(t),u(t),x'(t)\big)\mathrm{d}t + i_{\overline{S}}\,(x,u) : (x,u)\in X\times Y\Big] = \widehat{m}_r^4. \tag{2.112}$$

Then by virtue of Theorem 2.1.20, we have the following.

THEOREM 2.6.38 *If hypotheses $H(A)$, $H(L)$, $H(g)$, $H(U)_2$, and $H(L)_2$ hold, then problem (2.112) has a solution $(\widehat{x},\widehat{u})\in X\times Y$ and $\widehat{m}=\widehat{m}_r^4$.*

To relate problem (2.112) with the relaxed problems introduced earlier we need to strengthen the hypotheses on g and U.

$\mathbf{H(g)_1}$: $g\colon T\times\mathbb{R}^m \longrightarrow \mathbb{R}^k$ is a Carathéodory function (i.e., measurable in $t\in T$ and continuous in $x\in\mathbb{R}^m$) and for almost all $t\in T$ and all $\|u\|\le r$, we have $\|g(t,u)\|\le p(t)$ with $p\in L^1(T)_+$ (here $r>0$ as in $\mathrm{H}(U)_1$).

For the control constraint set U we assume hypothesis $\mathrm{H}(U)_1$. Then from (2.85) with $f(t,x,u)=A(t,x)+C(t,x)g(t,u)$, we have the following alternative relaxation of (2.104).

$$\left\{\begin{array}{ll} & J_r^1(x,\lambda)=\int_0^b\int_{\overline{B}_r} L\big(t,x(t),u(t)\big)\lambda(t)(\mathrm{d}u)\mathrm{d}t \longrightarrow \inf = \widehat{m}_r^1 \\ \text{s.t.} & x'(t)=A\big(t,x(t)\big)+C\big(t,x(t)\big)\int_{\overline{B}_r} g(t,u)\lambda(t)(\mathrm{d}u) \quad \text{a.e. on } T \\ & x(0)=x_0,\ \lambda\in S_\Sigma \end{array}\right\}. \tag{2.113}$$

DEFINITION 2.6.39 Given an admissible original control $u(\cdot)$, the *barycenter* of $u(\cdot)$ is defined by

$$\mathrm{Bar}(u)=\Big\{\lambda\in S_\Sigma : u(t)=\int_{\overline{B}_r} u\lambda(t)(\mathrm{d}u) \text{ a.e. on } T\Big\}.$$

In what follows, for convenience we set

$$\overline{\varphi}(x,u)=\Gamma_{\mathrm{seq}}(X^-,Y^-)\ \varphi(x,u). \tag{2.114}$$

PROPOSITION 2.6.40 *If hypotheses $H(A)$, $H(C)$, $H(g)$, $H(U)_1$, and $H(L)_2$ hold, then $\overline{\varphi}(x,u)=\min\big\{J_r^1(x,\lambda) : (x,\lambda)$ is admissible for (2.113) and $\lambda\in\mathrm{Bar}(u)\big\}$.*

PROOF: Let (x,λ) be an admissible state-control pair for problem (2.113) with $\lambda\in\mathrm{Bar}(u)$. Then by virtue of Proposition 2.6.17 we can find original controls $\{u_n\}_{n\ge1}$ such that $\delta_{u_n}\longrightarrow\lambda$ narrowly in $R(T,\overline{B}_r)$. Let $x_n\in W^{1,1}\big((0,b),\mathbb{R}^N\big)$ be the unique state produced by control u_n. From Theorem 2.6.10 we know that $\{x_n\}_{n\ge1}\subseteq C(T,\mathbb{R}^N)$ is relatively compact and so we may assume that $x_n\longrightarrow x$ in $C(T,\mathbb{R}^N)$. Also because of hypothesis $\mathrm{H}(U)_1$, we may assume that $u_n\xrightarrow{w}\widehat{u}$ in $L^1(T,\mathbb{R}^m)$. For all $D\in\mathbf{B}(T)$ we have

$$\int_D u_n(t)\mathrm{d}t\longrightarrow\int_D\widehat{u}(t)\mathrm{d}t. \tag{2.115}$$

Moreover, because $\delta_{u_n} \longrightarrow \lambda$ narrowly and $\lambda \in \mathrm{Bar}\ u$, we have

$$\int_D u_n(t)\mathrm{d}t = \int_D \int_{\overline{B}_r} u\delta_{u_n(t)}(\mathrm{d}u)\mathrm{d}t \longrightarrow \int_D \int_{\overline{B}_r} u\lambda(t)(\mathrm{d}u)\mathrm{d}t = \int_D u(t)\mathrm{d}t. \quad (2.116)$$

From (2.115), (2.116), and because $D \in \mathbf{B}(T)$ was arbitrary, we infer that $\widehat{u} = u$ and so $u_n \xrightarrow{w} u$ in $L^1(T, \mathbb{R}^m)$. Because (x_n, u_n) is admissible for the original optimal control problem, we have

$$\varphi(x_n, u_n) = J(x_n, u_n),$$

$$\Rightarrow \liminf_{n\to\infty} \varphi(x_n, u_n) = \liminf_{n\to\infty} J(x_n, u_n) = \liminf_{n\to\infty} J_r^1(x_n, \delta_{u_n}) = J_r^1(x, \lambda)$$

(see $\mathrm{H}(L)_2$),

$$\Rightarrow \overline{\varphi}(x, u) \leq \inf \left[J_r^1(x, \lambda) : (x, \lambda) \text{ is admissible for (2.113)}, \lambda \in \mathrm{Bar}\ u\right]. \quad (2.117)$$

On the other hand, if $\varphi(x, u) < +\infty$, from the definition of $\overline{\varphi}$ (see (2.114)), given $\varepsilon > 0$, we can find $x_n \longrightarrow x$ in $C(T, \mathbb{R}^N)$ and $u_n \xrightarrow{w} u$ in $L^1(T, \mathbb{R}^m)$ such that

$$\liminf_{n\to\infty} J(x_n, u_n) \leq \overline{\varphi}(x, u) + \varepsilon. \quad (2.118)$$

By Alaoglu's theorem we may also assume that $\delta_{u_n} \xrightarrow{w^*} \lambda$ in $L^\infty\big(T, M(\overline{B}_r)_{w^*}\big)$ (hence narrowly in $R(T, \overline{B}_r)$). Then for all $D \in \mathbf{B}(T)$ we have

$$\int_D u_n(t)\mathrm{d}t = \int_D \int_{\overline{B}_r} u\delta_{u_n(t)}(\mathrm{d}u)\mathrm{d}t \longrightarrow \int_D \int_{\overline{B}_r} u\lambda(t)(\mathrm{d}u)\mathrm{d}t \quad (2.119)$$

$$\text{and} \quad \int_D u_n(t)\mathrm{d}t \longrightarrow \int_D u(t)\mathrm{d}t. \quad (2.120)$$

From (2.119), (2.120), and because $D \in \mathbf{B}(T)$ was arbitrary it follows that $u(t) = \int_{\overline{B}_r} u\lambda(t)(\mathrm{d}u)$ a.e. on T (i.e., $\lambda \in \mathrm{Bar}\ u$). Also because of hypotheses $\mathrm{H}(L)_2$ we have $J(x_n, u_n) \longrightarrow J_r^1(x, \lambda)$. Therefore

$$J_r^1(x, \lambda) \leq \overline{\varphi}(x, u) + \varepsilon \quad \text{(see (2.118))}.$$

Let $\varepsilon \downarrow 0$, to conclude that

$$J_r^1(x, \lambda) \leq \overline{\varphi}(x, u). \quad (2.121)$$

Comparing (2.117) and (2.121), we obtain the desired equation. □

This proposition leads at once to the equivalence of relaxed problems (2.112) and (2.113).

THEOREM 2.6.41 *If hypotheses $H(A)$, $H(C)$, $H(g)$, $H(U)_1$, and $H(L)_2$ hold, then $\widehat{m} = \widehat{m}_r^1 = \widehat{m}_r^4$ and problems (2.112) and (2.113) are equivalent; that is, if $(\widehat{x}, \widehat{u})$ solves (2.112), then we can find $\widehat{\lambda} \in \mathrm{Bar}(u)$ such that $(\widehat{x}, \widehat{\lambda})$ solves (2.113) and conversely if $(\widehat{x}, \widehat{\lambda})$ solves (2.113) we can find an admissible original control $\widehat{u}$ such that $(\widehat{x}, \widehat{u})$ solves (2.112). Both problems have optimal pairs.*

REMARK 2.6.42 Evidently (2.112) is also equivalent to the other two relaxed problems defined in (2.90) and (2.94) with $f(t,x,u) = A(t,x) + C(t,x)g(t,u)$.

In the last part of this section we derive the *Pontryagin maximum principle* for the following optimal control problem.

$$\left\{ \begin{array}{ll} & J(x,u) = \int_0^b L\big(t,x(t),u(t)\big)\mathrm{d}t \longrightarrow \inf = m \\ \text{s.t.} & x'(t) = f\big(t,x(t),u(t)\big) \quad \text{a.e. on } T=[0,b] \\ & \big(x(0),x(b)\big) \in C,\ u \in U_{\mathrm{ad}} \end{array} \right\}. \tag{2.122}$$

Here $C \subseteq \mathbb{R}^N \times \mathbb{R}^N$ is nonempty and $U_{\mathrm{ad}} = \{u\colon T \longrightarrow Y \text{ measurable}\}$, where Y is a separable metric space (the control space). Note the general endpoint constraints that include the periodic problem as a special case. Our derivation of the Pontryagin maximum principle is based on the Ekeland variational principle (see Theorem 2.4.1).

The precise mathematical setting for problem (2.122) is the following.

The state space is $X = \mathbb{R}^N$ and the control space Y is a separable metric space. We impose the following conditions on the data of (2.122).

$\mathbf{H(f)_2}$: $f\colon T \times \mathbb{R}^N \times Y$ is a function such that

(i) For all $(x,u) \in \mathbb{R}^N \times Y$, $t \longrightarrow f(t,x,u)$ is measurable.
(ii) For all $t \in T$ and all $u \in Y$, $x \longrightarrow f(t,x,u)$ is differentiable and $(x,u) \longrightarrow f'_x(t,x,u)$ is continuous;.
(iii) There exists $M > 0$ such that $\|f(t,0,u)\|$, $\|f'_x(t,x,u)\| \leq M$ for all $(t,x,u) \in T \times \mathbb{R}^N \times Y$.

$\mathbf{H(C)}$: $C \subseteq \mathbb{R}^N \times \mathbb{R}^N$ is nonempty, closed, and convex.

$\mathbf{H(L)_3}$: $L\colon T \times \mathbb{R}^N \times Y \longrightarrow \mathbb{R}$ is a function such that

(i) For all $(x,u) \in \mathbb{R}^N \times Y$, $t \longrightarrow L(t,x,u)$ is measurable.
(ii) For all $t \in T$ and all $u \in Y, x \longrightarrow L(t,x,u)$ is differentiable and $(x,u) \longrightarrow L'_x(t,x,u)$ is continuous.
(iii) There exists $M_1 > 0$ such that $\|L(t,0,u)\|$, $\|L'_x(t,x,u)\| \leq M_1$ for all $(t,x,u) \in T \times \mathbb{R}^N \times Y$.

The control space Y is only a separable metric space and does not have a linear structure, therefore we cannot talk about convexity. For this reason the variations of the control function $u(\cdot)$ are of special type and are called *spike perturbations* of u. For this reason we need to use measure-theoretic arguments when dealing with such perturbations. In particular we need to know how we can approximate the constant function 1 by an oscillatory function of the form $(1/\lambda)\,\chi_C$, $\lambda \in (0,1)$ (here χ_C denotes the characteristic function of C).

LEMMA 2.6.43 *If (Ω,Σ,μ) is a finite nonatomic measure space, X is a separable Banach space, and for any $\lambda \in (0,1)$, $\mathcal{S}_\lambda = \big\{C \in \Sigma : \mu(C) = \lambda\mu(\Omega)\big\}$, then for any $g \in L^1(\Omega,X)$ we have*

$$\inf_{C \in \mathcal{S}_\lambda} \Big\| \int_\Omega \Big(\frac{1}{\lambda}\,\chi_C(\omega) - 1\Big) g(\omega)\mathrm{d}\mu \Big\| = 0. \tag{2.123}$$

PROOF: Given $\varepsilon > 0$, we can find a simple function $s_\varepsilon(\omega) = \sum\limits_{k=1}^{n} \chi_{A_k}(\omega) x_k$ with $A_k \in \Sigma,\ x_k \in X$ such that

$$\|g - s_\varepsilon\|_{L^1(\Omega,X)} < \varepsilon. \tag{2.124}$$

Due to the nonatomicity of μ, for each $k \in \{1,\ldots,n\}$ we can find $A_k^\lambda \subseteq A_k$ such that

$$\mu(A_k^\lambda) = \lambda\mu(A_k).$$

Let $A_\lambda = \bigcup\limits_{k=1}^{n} A_k^\lambda \in \mathcal{S}_\lambda$. Then

$$\begin{aligned}
&\int_\Omega \Big(\frac{1}{\lambda}\,\chi_{A_\lambda}(\omega) - 1\Big) s_\varepsilon(\omega)\mathrm{d}\mu = \sum_{k=1}^{n} \Big(\frac{1}{\lambda}\mu(A_k^\lambda) - \mu(A_k)\Big) x_k = 0,\\
\Rightarrow\ &\|\int_\Omega \Big(\frac{1}{\lambda}\,\chi_{A_\lambda}(\omega) - 1\Big) g(\omega)\mathrm{d}\mu\|_X \le \|\int_\Omega \Big(\frac{1}{\lambda}\,\chi_{A_\lambda}(\omega) - 1\Big) s_\varepsilon(\omega)\mathrm{d}\mu\|_X + \\
&\qquad\qquad + \Big(1 + \frac{1}{\lambda}\Big)\|s_\varepsilon - g\|_{L^1(\Omega,X)}\\
&\qquad\qquad \le \Big(1 + \frac{1}{\lambda}\Big)\,\varepsilon \quad \text{(see (2.124))}.
\end{aligned}$$

Because $\varepsilon > 0$, was arbitrary, we conclude that (2.123) holds. $\square$

In dealing with the endpoint constraints, we use the distance function from the set C. Recall that if Z is a Banach space and D is a nonempty, closed, and convex subset of Z, the distance function from D defined by $d_D(z) = \inf\{\|z-u\|_Z : u \in D\}$ is both convex and Lipschitz continuous with Lipschitz constant 1. So the convex subdifferential (see Definition 1.2.28) and the generalized subdifferential (see Definition 1.3.5) coincide and we have

$$\begin{aligned}
\partial d_D(z) &= \{z^* \in Z^* : \big\langle z^*, z' - z\big\rangle_Z \le d_D(z') - d_D(z) \quad \text{for all } \ z' \in Z\}\\
&= \{z^* \in Z^* : \langle z^*, h\rangle_Z \le d_D^0(z;h) \quad \text{for all } \ h \in Z\}.
\end{aligned} \tag{2.125}$$

LEMMA 2.6.44 *If Z is a Banach space and $D \subseteq Z$ is nonempty, closed, and convex, then for any $z \notin D$ and any $z^* \in \partial d_D(z)$ we have $\|z^*\|_{Z^*} = 1$.*

PROOF: Because $d_D(\cdot)$ is globally Lipschitz with constant 1, we have $\|z^*\|_{Z^*} \le 1$. Because $z \notin D$, given any $0 < \delta < 1$, we can find $u_\delta \in D$ such that

$$0 < (1-\delta)\|z - u_\delta\|_Z \le d_D(z).$$

So by virtue of (2.125) we have

$$\begin{aligned}
&\langle z^*, u_\delta - z\rangle_Z \le -d_D(z)\\
\Rightarrow\ &(1-\delta)\|z - u_\delta\|_Z \le d_D(z) \le -\langle z^*, u_\delta - z\rangle_Z \le \|z^*\|_{Z^*}\|u_\delta - z\|_Z\\
\Rightarrow\ &(1-\delta) \le \|z^*\|_{Z^*}.
\end{aligned}$$

Because $\delta \in (0,1)$ was arbitrary, we let $\delta \downarrow 0$, to conclude that $\|z^*\|_{Z^*} = 1$. $\square$

COROLLARY 2.6.45 *If Z is a Banach space with a strictly convex dual Z^* and $D \subseteq Z$ is a nonempty, closed, and convex set, then for any $z \notin D$ and the set $\partial d_D(z)$ is a singleton.*

As we already mentioned, the derivation of the Pontryagin maximum principle is based on the Ekeland variational principle. We need to endow the set of admissible controls U_{ad} with a metric structure. So for $u,\ w \in U_{\mathrm{ad}}$ we define

$$\widehat{d}(u,w) = \big|\{t \in T : u(t) \neq w(t)\}\big|_1.$$

Recall that by $|\cdot|_1$ we denote the one-dimensional Lebesgue measure. It is easy to check that $\widehat{d}$ is a metric and so $(U_{\mathrm{ad}}, \widehat{d})$ is a metric space.

LEMMA 2.6.46 *$(U_{\mathrm{ad}}, \widehat{d})$ is a complete metric space.*

PROOF: Let $\{u_n\}_{n\geq 1} \subseteq U_{\mathrm{ad}}$ be a $\widehat{d}$-Cauchy sequence. By passing to a suitable subsequence, we may assume that

$$\widehat{d}(u_{n+1}, u_n) \leq \frac{1}{2^n}, \quad n \geq 1.$$

Set $D_{nm} = \{t \in T : u_n(t) \neq u_m(t)\},\ n,m \geq 1$, and $C_k = \bigcup\limits_{n\geq k} D_{n(n+1)},\ k \geq 1$. Clearly $\big\{C_k\big\}_{k\geq 1}$ is decreasing and we have

$$|C_k|_1 \leq \sum_{n\geq k} \frac{1}{2^n} = \frac{1}{2^{k-1}}, \quad k \geq 1.$$

It follows that $\Big|\bigcup\limits_{k\geq 1} C_k^c\Big|_1 = b$. We set $\widehat{u}(t) = u_k(t)$ if $t \in C_k^c,\ k \geq 1$. From the definition of C_k it is clear that $\widehat{u}$ is well-defined and $\widehat{u} \in U_{\mathrm{ad}}$. Moreover,

$$\widehat{d}(u_k, \widehat{u}) \leq |C_k|_1 \longrightarrow 0 \quad \text{as } k \longrightarrow \infty.$$

□

REMARK 2.6.47 A careful reading of the above proof reveals that the lemma remains valid if Y is only a measurable space without any metric structure. In addition T can be replaced by any nonatomic measure space.

Now we are ready to state and prove the Pontryagin maximum principle for problem (2.122).

THEOREM 2.6.48 *If hypotheses $H(f)_2$, $H(C)$ and $H(L)_3$ hold and $(\widehat{x}, \widehat{u})$ is an optimal pair for problem (2.122), then there exists $(\psi_0, \psi) \in \mathbb{R} \times W^{1,1}\big((0,b), \mathbb{R}^N\big)$, $(\psi_0, \psi) \neq (0,0)$ such that*

(a) $\psi_0 \leq 0$.

(b) $\psi'(t) = f'_x\big(t, \widehat{x}(t), \widehat{u}(t)\big)^* \psi(t) + \psi_0 L'_x\big(t, \widehat{x}(t), \widehat{u}(t)\big)$ *a.e. on T.*

(c) $\big(\psi(0), x_0 - \widehat{x}(0)\big)_{\mathbb{R}^N} \leq \big(\psi(b), x_1 - \widehat{x}(b)\big)_{\mathbb{R}^N}$ *for all* $(x_0, x_1) \in C$.

(d) *If* $H(t,x,u,v_0,v) = \big(v, f(t,x,u)\big)_{\mathbb{R}^N} + v_0 L(t,x,u)$ *for all* $(t,x,u,v_0,v) \in T \times \mathbb{R}^N \times Y \times \mathbb{R} \times \mathbb{R}^N$, *then* $H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big) = \max_{u \in Y} H\big(t, \widehat{x}(t), u, \psi_0, \psi(t)\big)$ *for almost all* $t \in T$.

PROOF: First let us briefly outline the strategy of the proof. It is divided into a series of steps.

(i) We introduce a penalty function that is used to define an approximation of problem (2.122) (called the approximate problem) which has no end point constraint.
(ii) We use the Ekeland variational principle to produce an optimal pair for the approximate problem. This pair is actually close to the original optimal pair.
(iii) We derive necessary conditions for the approximate optimal pair.
(iv) We pass to the limit to obtain necessary conditions for the original optimal pair.

On $\mathbb{R}^N \times U_{\mathrm{ad}}$ we introduce the metric

$$\widehat{d}_0\big((x,u),(z,w)\big) = \|x - z\|_{\mathbb{R}^N} + \widehat{d}(u,w).$$

Clearly $(\mathbb{R}^N \times U_{\mathrm{ad}}, \widehat{d}_0)$ is a complete metric space (see Lemma 2.6.46). In what follows by $x(x_0,u)(\cdot) \in W^{1,1}\big((0,b),\mathbb{R}^N\big)$ we denote the unique state generated by $u \in U_{\mathrm{ad}}$ and the initial condition $x_0 \in \mathbb{R}^N$. Then we can write the cost functional $J(x_0,u)$. If $\widehat{x}(0) = \widehat{x}_0$, then by translating the cost functional if necessary (i.e., considering the cost functional $J(x_0,u) - J(\widehat{x}_0,\widehat{u})$), we may assume without any loss of generality that $J(\widehat{x}_0,\widehat{u}) = 0$.

We start executing the four steps of the method of proof outlined above.

Step i: We introduce the penalty function.

For $\varepsilon > 0$ and $(x_0,u) \in \mathbb{R}^N \times U_{\mathrm{ad}}$, we set

$$J_\varepsilon(x_0,u) = \Big[d_C^2\big(x_0, x(x_0,u)(b)\big) + [\big(J(x_0,u) + \varepsilon\big)^+]^2\Big]^{1/2}. \tag{2.126}$$

Because of hypotheses $\mathrm{H}(f)_2$ and $\mathrm{H}(L)_3$ and Gronwall's inequality for all (x_0,u), $(\overline{x}_0,\overline{u}_0) \in \mathbb{R}^N \times U_{\mathrm{ad}}$ we have

$$\|x(x_0,u)(\cdot) - x(\overline{x}_0,\overline{u}_0)(\cdot)\|_{\mathbb{R}^N} \leq c(1 + \max\{\|x_0\|, \|\overline{x}_0\|\}) \times \widehat{d}_0\big((x_0,u),(\overline{x}_0,\overline{u}_0)\big) \tag{2.127}$$

and
$$|J(x_0,u) - J(\overline{x}_0,\overline{u}_0)| \leq c\big(1 + \max\{\|x_0\|, \|\overline{x}_0\|\}\widehat{d}_0\big((x_0,u),(\overline{x}_0,\overline{u}_0)\big),$$
$$\text{for some } c > 0. \tag{2.128}$$

From (2.126) through (2.128) it follows that $J_\varepsilon(\cdot,\cdot)$ is continuous on $(\mathbb{R}^N \times U_{\mathrm{ad}}, \widehat{d}_0)$.

Step ii: Solution (via the Ekeland variational principle) of the problem with the same dynamics and cost functional a new expression involving J_ε.

From (2.126) we see that $J_\varepsilon \geq 0$ and $J_\varepsilon(\widehat{x}_0,\widehat{u}) = \varepsilon \leq \inf_{\mathbb{R}^N \times U_{\mathrm{ad}}} J_\varepsilon(x_0,u) + \varepsilon$ (recall that $\widehat{x}(0) = \widehat{x}_0$ and that we have assumed that $J(\widehat{x}_0,\widehat{u}) = 0$). So we can apply Corollary 2.4.3 and obtain $(x_0^\varepsilon, u^\varepsilon) \in \mathbb{R}^N \times U_{\mathrm{ad}}$ such that

$$\widehat{d}_0\big((x_0^\varepsilon, u^\varepsilon), (\widehat{x}_0, \widehat{u})\big) \le \sqrt{\varepsilon} \tag{2.129}$$

$$\text{and} \qquad J_\varepsilon(x_0^\varepsilon, u^\varepsilon) \le J_\varepsilon(x_0, u) + \sqrt{\varepsilon}\, \widehat{d}_0\big((x_0, u), (x_0^\varepsilon, u^\varepsilon)\big)$$
$$\text{for all } (x_0, u) \in \mathbb{R}^N \times U_{\mathrm{ad}}. \tag{2.130}$$

So if $x^\varepsilon = x(x_0^\varepsilon, u^\varepsilon)$, then $(x^\varepsilon, u^\varepsilon)$ is a solution of the approximate optimal control problem with the same dynamics and cost functional as the right-hand side of (2.130). Note that because of (2.129) the optimal pair of the approximate problem is located close to the optimal pair of the original problem.

Step iii: We derive necessary conditions for the optimal pair $(x^\varepsilon, u^\varepsilon)$ of the approximate problem.

Fix $(e_0, w) \in \overline{B}_1 \times U_{\mathrm{ad}}$ ($\overline{B}_1 = \{e_0 \in \mathbb{R}^N : \|e_0\|_{\mathbb{R}^N} \le 1\}$) and $\lambda \in (0, 1]$ and set

$$g(s) = \begin{pmatrix} L\big(s, x^\varepsilon(s), w(s)\big) - L\big(s, x^\varepsilon(s), u^\varepsilon(s)\big) \\ f\big(s, x^\varepsilon(s), w(s)\big) - f\big(s, x^\varepsilon(s), u^\varepsilon(s)\big) \end{pmatrix}.$$

From Lemma 2.6.43 we know that given any $\delta > 0$. we can find $C_\lambda \subseteq T$ with $|C_\lambda|_1 = \lambda b$ such that

$$\Big\| \int_0^b g(s)\mathrm{d}s - \frac{1}{\lambda} \int_0^b \chi_{C_\lambda}(s) g(s)\mathrm{d}s \Big\|_{\mathbb{R}^N} \le \delta. \tag{2.131}$$

We introduce the following spike perturbation of the optimal control $u^\varepsilon(\cdot)$,

$$u_\lambda^\varepsilon(t) = \begin{cases} u^\varepsilon(t) & \text{if } t \in T \setminus C_\lambda \\ w(t) & \text{if } t \in C_\lambda \end{cases}. \tag{2.132}$$

Clearly $u_\lambda^\varepsilon \in U_{\mathrm{ad}}$ and $\widehat{d}(u_\lambda^\varepsilon, u^\varepsilon) \le |C_\lambda|_1 = \lambda b$. Set $x_\lambda^\varepsilon = x(x_0^\varepsilon + \lambda e_0, u_\lambda^\varepsilon)$. We define $(z^\varepsilon, z_0^\varepsilon) \in W^{1,1}\big((0, b), \mathbb{R}^N\big) \times \mathbb{R}$ as follows

$$(z^\varepsilon)'(t) = f_x'\big(t, x^\varepsilon(t), u^\varepsilon(t)\big) z^\varepsilon(t) + f\big(t, x^\varepsilon(t), w(t)\big) - f\big(t, x^\varepsilon(t), u^\varepsilon(t)\big)$$
$$\text{a.e. on } T, \; z^\varepsilon(0) = e_0 \quad \text{and}$$
$$\begin{aligned} z_0^\varepsilon &= \int_0^b L_x'\big(t, x^\varepsilon(t), u^\varepsilon(t)\big) z^\varepsilon(t)\mathrm{d}t \\ &\quad + \int_0^b \Big(L\big(t, x^\varepsilon(t), w(t)\big) - L\big(t, x^\varepsilon(t), u^\varepsilon(t)\big) \Big)\mathrm{d}t. \end{aligned}$$

Exploiting hypotheses $\mathrm{H}(f)_2$ and $\mathrm{H}(L)_3$, via Gronwall's inequality, we can easily check that

$$\|x_\lambda^\varepsilon - x^\varepsilon - \lambda z^\varepsilon\|_{C(T, \mathbb{R}^N)} = o(\lambda) \tag{2.133}$$

$$\text{and} \qquad \big| J(x_0^\varepsilon + \lambda e_0, u_\lambda^\varepsilon) - J(x_0^\varepsilon, u^\varepsilon) - \lambda z_0^\varepsilon \big| = o(\lambda), \tag{2.134}$$

where $\big(o(\lambda)/\lambda\big) \longrightarrow 0$ as $\lambda \downarrow 0$. Recalling that J_ε is continuous on $(\mathbb{R}^N \times U_{\mathrm{ad}}, \widehat{d}_0)$, from (2.133) and (2.133) and the definition of J_ε, we see that

$$J_\varepsilon(x_0^\varepsilon + \lambda e_0, u_\lambda^\varepsilon) = J_\varepsilon(x_0^\varepsilon, u^\varepsilon) + \varepsilon(\lambda) \quad \text{with } \varepsilon(\lambda) \downarrow 0. \tag{2.135}$$

From (2.130) we have

$$
\begin{aligned}
&-\sqrt{\varepsilon}(\|e_0\|_{\mathbb{R}^N}+b) \le \frac{1}{\lambda}\big(J_\varepsilon(x_0^\varepsilon+\lambda e_0,u_\lambda^\varepsilon)-J_\varepsilon(x_0^\varepsilon,u^\varepsilon)\big)\\
&= \frac{1}{J_\varepsilon(x_0^\varepsilon+\lambda e_0,u_\lambda^\varepsilon)+J_\varepsilon(x_0^\varepsilon,u^\varepsilon)}\Big[\frac{1}{\lambda}\big(J_\varepsilon(x_0^\varepsilon+\lambda e_0,u_\lambda^\varepsilon)^2-J_\varepsilon(x_0^\varepsilon,u^\varepsilon)^2\big)\Big]\\
&= \frac{1}{2J_\varepsilon(x_0^\varepsilon,u^\varepsilon)+\varepsilon(\lambda)}\Big[\frac{1}{\lambda}\big(d_C^2(x_0^\varepsilon+\lambda e_0,x_\lambda^\varepsilon(b))-d_C^2(x_0^\varepsilon,x^\varepsilon(b))\big)\Big]\\
&\quad+\frac{1}{\lambda}\Big(\big[\big(J\big(x_0^\varepsilon+\lambda e_0,u^\varepsilon\big)+\varepsilon\big)^+\big]^2-\big[\big(J(x_0^\varepsilon,u^\varepsilon)+\varepsilon\big)^+\big]^2\Big) \text{ (see (2.133)).}
\end{aligned}
\tag{2.136}
$$

We know that $d_C^2\in C^1(\mathbb{R}^N\times\mathbb{R}^N)$ and

$$
\nabla d_C^2(x_0,x_1)=\begin{cases}2\, d_C(x_0,x_1)\,\partial d_C(x_0,x_1) & \text{if } (x_0,x_1)\notin C\\ 0 & \text{if } (x_0,x_1)\in C\end{cases}.
$$

Recall that by virtue of Corollary 2.6.45 $\partial d_C(x_0,x_1)=(\alpha_0,\alpha_1)\in\mathbb{R}^N\times\mathbb{R}^N$ and $\|\alpha_0\|_{\mathbb{R}^N}^2+\|\alpha_1\|_{\mathbb{R}^N}^2=1$ (see Lemma 2.6.44), when $(x_0,x_1)\notin C$. On the other hand because d_C is Lipschitz continuous with Lipschitz constant 1, we always have $\partial d_C(x_0,x_1)\subseteq\overline{B}_1=\{(x,z)\in\mathbb{R}^N\times\mathbb{R}^N:\|x\|_{\mathbb{R}^N}^2+\|z\|_{\mathbb{R}^N}^2\le 1\}$. So we can write with no ambiguity that

$$
\begin{aligned}
\nabla d_C^2(x_0,x_1)=2\, d_C(x_0,x_1)(\alpha_0,\alpha_1) \quad &\text{with } (\alpha_0,\alpha_1)\in\partial d_C(x_0,x_1),\\
\|\alpha_0\|_{\mathbb{R}^N}^2+\|\alpha_1\|_{\mathbb{R}^N}^2&=1.
\end{aligned}
$$

Because of (2.133), we have

$$
\begin{aligned}
&\lim_{\lambda\downarrow 0}\frac{d_C^2\big(x_0^\varepsilon+\lambda e_0,x_\lambda^\varepsilon(b)\big)-d_C^2\big(x_0^\varepsilon,x^\varepsilon(b)\big)}{\lambda}\\
&= 2\, d_C\big(x_0^\varepsilon,x^\varepsilon(b)\big)\Big((\alpha_0^\varepsilon,e_0)_{\mathbb{R}^N}+\big(\alpha_1^\varepsilon,z^\varepsilon(b)\big)_{\mathbb{R}^N}\Big)
\end{aligned}
$$

where

$$
(\alpha_0^\varepsilon,\alpha_1^\varepsilon)\in\partial d_C\big(x_0^\varepsilon,x^\varepsilon(b)\big) \text{ and } \|\alpha_0^\varepsilon\|_{\mathbb{R}^N}^2+\|\alpha_1^\varepsilon\|_{\mathbb{R}^N}^2=1. \tag{2.137}
$$

Also because of (2.133) we have

$$
\begin{aligned}
&\lim_{\lambda\downarrow 0}\frac{1}{\lambda}\Big[\big[\big(J\big(x_0^\varepsilon+\lambda e_0,u^\varepsilon\big)+\varepsilon\big)^+\big]^2-\big[\big(J(x_0^\varepsilon,u^\varepsilon)+\varepsilon\big)^+\big]^2\Big]\\
&=2\,\big(J(x_0^\varepsilon,u^\varepsilon)+\varepsilon\big)^+z_0^\varepsilon.
\end{aligned}
\tag{2.138}
$$

Returning to (2.136), passing to the limit as $\lambda\downarrow 0$ and using (2.137) and (2.138) we obtain

$$
-\sqrt{\varepsilon}(\|e_0\|_{\mathbb{R}^N}+b)\le(\widehat{\varphi}_\varepsilon,e_0)_{\mathbb{R}^N}+\big(\widehat{\psi}_\varepsilon,z_\varepsilon(b)\big)_{\mathbb{R}^N}+\widehat{\psi}_0^\varepsilon z_0^\varepsilon \tag{2.139}
$$

where $\widehat{\varphi}_\varepsilon=\frac{d_C\left(x_0^\varepsilon,x^\varepsilon(b)\right)}{J_\varepsilon(x_0^\varepsilon,u^\varepsilon)}\,\alpha_0^\varepsilon$, $\widehat{\psi}_\varepsilon=\frac{d_C\left(x_0^\varepsilon,x^\varepsilon(b)\right)}{J_\varepsilon(x_0^\varepsilon,u^\varepsilon)}\,\alpha_1^\varepsilon$, and

$$
\widehat{\psi}_0^\varepsilon=\frac{\big(J(x_0^\varepsilon,u^\varepsilon)+\varepsilon\big)^+}{J_\varepsilon(x_0^\varepsilon,u^\varepsilon)}. \tag{2.140}
$$

Because of (2.137) we have

$$\|\widehat{\varphi}_\varepsilon\|^2_{\mathbb{R}^N} + \|\widehat{\psi}_\varepsilon\|^2_{\mathbb{R}^N} + (\widehat{\psi}_0^\varepsilon)^2 = 1, \tag{2.141}$$

and from the definition of the convex subdifferential and because $\partial d_C\big(x_0^\varepsilon, x^\varepsilon(b)\big)$ is a cone we have

$$(\widehat{\varphi}_\varepsilon, x_0 - x_0^\varepsilon)_{\mathbb{R}^N} + \big(\widehat{\psi}_\varepsilon, x_1 - x^\varepsilon(b)\big)_{\mathbb{R}^N} \le 0 \qquad \text{for all } (x_0, x_1) \in C. \tag{2.142}$$

Conditions (2.138)⟶(2.142) are necessary conditions for optimality of $(x^\varepsilon, u^\varepsilon)$.

Step iv: Passage to the limit as $\varepsilon \downarrow 0$.

Let $(z, z^0) \in W^{1,1}\big((0,b), \mathbb{R}^N\big) \times \mathbb{R}$ be defined by

$$\begin{aligned} z'(t) &= f'_x\big(t, \widehat{x}(t), \widehat{u}(t)\big) z(t) + f\big(t, \widehat{x}(t), w(t)\big) - f\big(t, \widehat{x}(t), \widehat{u}(t)\big) \\ &\qquad \text{a.e. on } T,\ z(0) = e_0 \\ z_0 &= \int_0^b L'_x\big(t, \widehat{x}(t), \widehat{u}(t)\big) z(t) \mathrm{d}t + \int_0^b \Big(L\big(t, \widehat{x}(t), w(t)\big) - L\big(t, \widehat{x}(t), \widehat{u}(t)\big)\Big) \mathrm{d}t. \end{aligned}$$

Then because of (2.129) and as before via Gronwall's inequality, we have

$$\lim_{\varepsilon \downarrow 0} \left[\|x_0^\varepsilon - \widehat{x}_0\|_{\mathbb{R}^N} + \|z^\varepsilon - z\|_{C(T,\mathbb{R}^N)} + |z_0^\varepsilon - z_0|\right] = 0.$$

So from (2.141) and (2.142), we have

$$\begin{aligned} (\widehat{\varphi}_\varepsilon, x_0 - \widehat{x}_0)_{\mathbb{R}^N} + \big(\widehat{\psi}_\varepsilon, x_1 - \widehat{x}(b)\big)_{\mathbb{R}^N} &\le \big(\|x_0^\varepsilon - \widehat{x}_0\|^2_{\mathbb{R}^N} + \|x^\varepsilon(b) - \widehat{x}(b)\|^2_{\mathbb{R}^N} \\ &= \beta_\varepsilon \longrightarrow 0 \ \text{ as } \ \varepsilon \downarrow 0. \end{aligned}$$

Combining this inequality with (2.139), we obtain

$$\begin{aligned} &\big(\widehat{\varphi}_\varepsilon, e_0 - (x_0 - \widehat{x}_0)\big)_{\mathbb{R}^N} + \Big(\widehat{\psi}_\varepsilon, z(b) - (x_1 - \widehat{x}(b))\Big)_{\mathbb{R}^N} + \widehat{\psi}_0^\varepsilon z_0 \\ \ge\ & -\sqrt{\varepsilon}(\|e_0\|_{\mathbb{R}^N} + b) - \beta_\varepsilon - \|z^\varepsilon(b) - z(b)\|_{\mathbb{R}^N} - |z_0^\varepsilon - z_0| \ge -\gamma_\varepsilon \\ &\text{for all } (x_0, x_1) \in C, \end{aligned} \tag{2.143}$$

where $\gamma_\varepsilon > 0$ independent of $u \in U_{\mathrm{ad}}$, $e_0 \in \overline{B}_1$, and $\gamma_\varepsilon \downarrow 0$ as $\varepsilon \downarrow 0$. Because of (2.141) we may assume that

$$(\widehat{\varphi}_\varepsilon, \widehat{\psi}_\varepsilon, \widehat{\psi}_0^\varepsilon) \longrightarrow (\widehat{\varphi}, \widehat{\psi}, \widehat{\psi}_0) \quad \text{in } \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \text{ as } \varepsilon \downarrow 0.$$

Clearly $(\widehat{\varphi}, \widehat{\psi}, \widehat{\psi}_0) \ne (0,0,0)$ (see (2.141)). Let $\varepsilon \downarrow 0$ in (2.143). We obtain

$$\big(\widehat{\varphi}, e_0 - (x_0 - \widehat{x}_0)\big)_{\mathbb{R}^N} + \Big(\widehat{\psi}, z(b) - \big(x_1 - \widehat{x}(b)\big)\Big)_{\mathbb{R}^N} + \widehat{\psi}_0 z_0 \ge 0 \tag{2.144}$$

for all $(x_0, x_1) \in C$, all $w \in U_{\mathrm{ad}}$, and all $e_0 \in \overline{B}_1$. Let $\psi \in W^{1,1}\big((0,b), \mathbb{R}^N\big)$ be the unique solution of the equation in statement (b) of the theorem $\psi(b) = -\widehat{\psi}$ and $\psi_0 = -\widehat{\psi}_0 \le 0$ (see (2.140)). We rewrite (2.144) as follows,

$$(\widehat{\varphi}, x_0 - \widehat{x}_0 - e_0)_{\mathbb{R}^N} - \big(\psi(b), x_1 - \widehat{x}(b) - z(b)\big)_{\mathbb{R}^N} + \psi_0 z_0 \le 0 \tag{2.145}$$

for all $(x_0, x_1) \in C$, all $w \in U_{\mathrm{ad}}$, and all $e_0 \in \overline{B}_1$.

A straightforward computation gives

$$\begin{aligned}&\big(\psi(b), z(b)\big)_{\mathbb{R}^N} - \big(\psi(0), e_0\big)_{\mathbb{R}^N} + \psi_0 z_0 \\ &= \int_0^b \Big(H\big(t, \widehat{x}(t), w(t), \psi_0, \psi(t)\big) - H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big)\Big)\mathrm{d}t,\end{aligned} \tag{2.146}$$

for all $e_0 \in \overline{B}_1$ and all $w \in U_{\mathrm{ad}}$. In (2.145) we set $e_0 = 0,\ x_0 = \widehat{x}_b = \widehat{x}(0)$ and use the resulting inequality in (2.146) to obtain

$$\int_0^b \Big(H\big(t, \widehat{x}(t), w(t), \psi_0, \psi(t)\big) - H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big)\Big)\mathrm{d}t \le 0 \tag{2.147}$$

for all $w \in U_{\mathrm{ad}}$.

Recall that Y is separable. So we can find a countable dense set $\{u_m\}_{m\ge 1} \subseteq Y$. Then if

$$h_m(t) = H\big(t, \widehat{x}(t), u_m, \psi_0, \psi(t)\big) - H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big), \quad m \ge 1,$$

we have $h_m \in L^1(T)$ and we can find $E_m \subseteq T$ with $|E_m|_1 = b$ such that

$$\lim_{\delta \downarrow 0} \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} h_m(s)\mathrm{d}s = h_m(t) \qquad \text{for all } \ t \in E_m.$$

If $t \in E_m$, we define

$$w_\delta(s) = \begin{cases} \widehat{u}(s) & \text{if } \ |s-t| > \delta \\ u_m & \text{if } \ |s-t| \le \delta \end{cases}.$$

From (2.147) we have

$$\int_{t-\delta}^{t+\delta} h_m(s)\mathrm{d}s \le 0 \qquad \text{for all } \ \delta > 0.$$

Divide with 2δ and let $\delta \downarrow 0$, to obtain $h_m(t) \le 0$. Thus,

$$H\big(t, \widehat{x}(t), u_m, \psi_0, \psi(t)\big) - H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big) \le 0$$

for all $t \in E = \bigcap_{m\ge 1} E_m,\ |E|_1 = b$.

Exploiting the continuity of the Hamiltonian H, we have

$$H\big(t, \widehat{x}(t), \widehat{u}(t), \psi_0, \psi(t)\big) = \max_{u \in Y} H\big(t, \widehat{x}(t), u, \psi_0, \psi(t)\big) \ \text{ a.e. on } T.$$

Also, if in (2.145) we let $w = \widehat{u},\ x_0 = \widehat{x}_0,\ x_1 = \widehat{x}(b)$ and we use (2.146) and (2.147), we obtain

$$\begin{aligned}&(\widehat{\varphi}, e_0)_{\mathbb{R}^N} \ge \big(\psi(b), z(b)\big)_{\mathbb{R}^N} + \psi_0 z_0 \ge \big(\psi(0), e_0\big)_{\mathbb{R}^N} \qquad \text{for all } \ e_0 \in \overline{B}_1, \\ \Rightarrow\ &\widehat{\varphi} = \psi(0).\end{aligned}$$

Therefore, if in (2.145) $e_0 = 0$ and $w = \widehat{u}$, we obtain

$$\big(\psi(0), x_0 - \widehat{x}(0)\big)_{\mathbb{R}^N} \le \big(\psi(b), x_1 - \widehat{x}(b)\big)_{\mathbb{R}^N} \qquad \text{for all } \ (x_0, x_1) \in C.$$

Finally note that $(\psi_0, \psi) \neq (0,0)$, or otherwise we contradict the fact that $(\widehat{\varphi}, \widehat{\psi}, \widehat{\psi}_0) \neq (0,0,0)$. $\square$

REMARK 2.6.49 The function $\psi(\cdot)$ is known as the *costate* or *adjoint state*. The system in (b) is known as the *adjoint system*. The inequality in (c) is known as the *transversality condition* and finally statement (d) is the *maximum condition*. In analogy with mechanics H is the *Hamiltonian*.

2.7 Remarks

2.1: The notion of lower semicontinuity for real functions was introduced by Borel [89], but systematic use of it in variational problems is due to Tonelli [585]. Lower semicontinuity is studied in the books of Barbu–Precupanu [59], Cesari [140], Dal Maso [171], Denkowski–Migórski–Papageorgiou [194], Ekeland–Temam [222], and Ioffe–Tichomirov [327]. Theorem 2.1.25 was first proved by De Giorgi [189] for the case φ is a nonnegative Carathéodory integrand. The version presented here is due to Ioffe [326]. Earlier results on the lower semicontinuity properties of integral functionals $J(x,u)=\int_\Omega \varphi\big(\omega, x(\omega), u(\omega)\big)\mathrm{d}\mu$ on the Lebesgue spaces $L^p(\Omega,\mathbb{R}^N)\times L^r(\Omega,\mathbb{R}^m)$, were provided by Berkovitz [71], Cesari [138, 139], and Olech [467]. Corresponding results for the weak lower semicontinuity of integral functionals $J(x)=\int_\Omega \varphi\big(z, x(z), Dx(z)\big)\mathrm{d}z$ ($\Omega\subseteq\mathbb{R}^N$ a bounded domain) on the Sobolev spaces $W^{1,p}(\Omega)$ were proved by Buttazzo [126], Dacorogna [170], Ekeland–Temam [222], Giaquinta [263], and Morrey [443]. Lower semicontinuity theorems in the weak topologies of Sobolev spaces when the function $x(\cdot)$ is vector-valued can be found in Acerbi–Fusco [1], Ball–Zhang [56], Dacorogna [170], and Morrey [443].

2.2: The method of Lagrange multipliers is discussed in Alexeev–Tichomirov–Fomin [7], Tichomirov [582] and Zeidler [620]. Theorem 2.2.7 is due to Ljusternik [392]. It can also be found in the book of Ljusternik–Sobolev [395]. Nonsmooth extensions of the Lagrange multipliers method can be found in Aubin [37], Clarke [149, 153], Rockafellar [522], and Rockafellar–Wets [530]. The Dubovitskii–Milyutin formalism for infinite dimensional optimization problems, was first formulated by Dubovitskii–Milyutin [208]. A comprehensive and readable presentation of the theory is contained in the monograph of Girsanov [266].

2.3: The first minimax theorem was formulated for bilinear functionals on finite-dimensional simplices by von Neumann [454]. It was a basic tool in the study of his model for the growth of an economy. Since then there have been several generalizations of the result of von Neumann. We mention the works of Fan [232], Sion [557], Browder [119, 121], Brezis–Nirenberg–Stampacchia [101], Terkelsen [580], and Simons [556]. Theorem 2.3.7 was first proved by Knaster–Kuratowski–Mazurkiewicz [358] in the special case in which B is the set of vertices of a simplex in $\mathbb{R}^N$. The form presented here is due to Dugundji–Granas [211, 212]. An earlier version was proved by Fan [233] who also gave numerous applications of it (see, e.g., Fan [234, 235]; in fact Fan [235] obtained the coincidence result proved in Proposition 2.3.10). Theorem 2.3.13 is due to Sion [557]. The duality theory for convex optimization, which uses the idea of perturbed problems and conjugate convex functionals, was developed for the finite-dimensional case by Rockafellar [522]. The infinite-dimensional case presented here is due to Ekeland–Temam [222]. The special case considered in (2.34) (problem $(\mathrm{P})'$) is due to Fenchel [241] when $A=I$ and to Rockafellar [518], the general situation with $A\in\mathcal{L}(X,Y)$.

2.4: Theorem 2.4.1 was proved by Ekeland [221]. A detailed discussion of the various applications that this theorem has can be found in Ekeland [223, 226]. Theorem 2.4.11 is due to Caristi [129] who proved it using transfinite induction. Takahashi [571] proved Theorem 2.4.14. Theorem 2.4.16 is due to Danes [175], who proved it using a result of Krasnoselskii–Zabreiko. Theorem 2.4.20 is due to Brezis–Browder [103]. A generalization of the Ekeland variational principle, in which the Lipschitz continuous perturbations are replaced by quadratic ones, was proved by Borwein–Preiss [92]. Additional results in this direction can be found in Deville–Godefroy–Zizler [197].

2.5: The calculus of variations is of course a much broader subject. Detailed presentation of the subject can be found in the books Bliss [80], Bolza [83], Cesari [140], Gelfand–Fomin [260], Hestenes [292], Ioffe–Tichomirov [327], Morrey [443], Morse [444], Troutman [586], and Young [617]. The more recent developments on vectorial problems (with significant applications in elasticity theory) can be found in Dacorogna [170].

2.6: Hypothesis H_c in the existence theory is related to the property (Q) of Cesari [138]. For a detailed discussion of the existence theory for finite dimensional optimal control problems and of the relevant property (Q), we refer to Cesari [140]. Similar results can also be found in Berkovitz [70], Fleming–Richel [249] (finite-dimensional problems) and Ahmed [4], Ahmed–Teo [3], Hu–Papageorgiou [316], Lions [388] and Tiba [581] (infinite-dimensional problems). Scrutinizing the proof of Theorem 2.6.11, we see that what makes it work is hypothesis H_c. So in order to be able to guarantee existence of an optimal pair in an optimal control problem, we need a convexity-type hypothesis. If such a condition is not present the problem need not have a solution. This difficulty was surmounted independently in the late fifties and early sixties by Filippov [247], Warga [601] and Gamkrelidze [255]. They realized that we need to suitably augment the system with the introduction of new admissible controls in order to be able to capture the asymptotic behavior of the minimizing sequences of the original problem. Gamkrelidze and Warga worked with parametrized measures (see problem (2.42)), whereas Filippov used a control-free problem (see problem (2.47)). Here we also present a third relaxation method based on Carathéodory's theorem for convex sets. The fourth relaxation method based on the Γ-regularization of the extended cost functional, is due to Buttazzo–Dal Maso [125] (see also Buttazzo [126]), who for that purpose developed the formalism of multiple Γ-limits. Further results on the relaxation of the calculus of variations and optimal control problems can be found in Buttazzo [126]), Dal Maso [171], Fattorini [239], and Roubicek [532]. Necessary conditions for optimality were first obtained by Pontryagin and his co-workers (see Pontryagin [499] and Pontryagin–Boltyanski–Gamkrelidze–Mischenko [500]. It has become a common terminology to call these necessary conditions the Pontryagin maximum principle. Proofs of the maximum principle for different types of optimal control problems can be found in Berkovitz [70], Cesari [140], Fleming–Richel [249], Hermes–La Salle [289], Ioffe–Tichomirov [327], Seierstadt–Sydsaeter [547] (finite-dimensional systems), and Ahmed [4], Ahmed–Teo [3], Fattorini [239], Li–Yong [381], Lions [388], and Tiba [581] (infinite-dimensional systems).

http://www.springer.com/978-0-387-78906-4

Handbook of Applied Analysis
Papageorgiou, N.S.; Kyritsi-Yiallourou, S.T.
2009, XXIII, 793 p., Hardcover
ISBN: 978-0-387-78906-4