Order Number 9117998

Convergence and approximation for primal-dual methods in large-scale optimization

Wright, Stephen Edward, Ph.D.

University of Washington, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Convergence and Approximation for Primal-Dual
Methods in Large-Scale Optimization

by Stephen E. Wright

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

1990

Approved by R T Rockafellar
(Chairperson of Supervisory Committee)

Program Authorized
to Offer Degree Department of Mathematics

Date Dec 19, 1990

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Stefan E. Wing

Date 21 December 1990

University of Washington

Abstract

Convergence and Approximation for Primal-Dual Methods in Large-Scale Optimization

by Stephen E. Wright

Chairperson of the Supervisory Committee: Professor R. Tyrrell Rockafellar
Department of Mathematics

Large-scale problems in convex optimization often can be reformulated in primal-dual (minimax) representations having special decomposition properties. Approximation of the resulting high-dimensional problems by restriction to low-dimensional subspaces leads to a family of minimax problems dependent on a parameter. The continuity and convergence properties of this dependence are explored in this dissertation. Examples in optimal control and stochastic programming are considered in which discretizations give rise to large-scale optimization problems. A possible approach to the numerical solution of the discretized problems is described, as well as details of its computer implementation.

Table of Contents

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor R. T. Rockafellar for his inspirational teaching and for providing many opportunities for research and discovery; to Dr. Mark J. Nielsen for our endless conversations on mathematics, politics and life in general; and to my wife Jennie for her love and tolerance. I owe a tremendous debt to each of these three people for their continual encouragement during my years as a graduate student.

Special thanks go to Professor R. J.-B. Wets for suggesting the line of research that lead to this dissertation. I also benefited from numerous stimulating discussions with Professor Domokos Vermes, Dr. James V. Burke and Dr. Alan J. King.

Finally, I gratefully acknowledge the financial support I have received as a graduate student. The research and preparation of this dissertation were supported in part by the National Science Foundation grant number DMS-8819586 and the Air Force Office of Scientific Research grant number AFOSR-89-0081, under the supervision of Professor Rockafellar.

Chapter 1. Introduction

This dissertation deals with approximations of primal-dual approaches for solving large-scale problems in convex optimization. The term "large-scale" refers to problems which involve a huge (perhaps infinite) number of variables or constraints, but which also have a highly specialized structure. Large-scale problems typically appear as the conjunction of many smaller problems which are interrelated in a very special way. For example, in a dynamic optimization problem there is a time framework which may consist of either discrete stages $\{1, 2, \ldots, T\}$ or an interval $[t_0, t_1]$. The decisions allowed at a given point in time depend on the "state" of the system at that time, where the state is governed by a difference or differential equation with coefficients dependent on the decisions. In a stochastic problem, there are uncertain elements which need to be taken into account: the weather, stock prices, etc. The decision-maker must weigh the cost associated with each possible contingency against its likelihood of occurring. The interrelationships between the subproblems in this case may be given by a probability distribution. When both dynamic *and* stochastic elements are present, there may be additional structure representing the exchange of information between time periods.

First let's review the basic concepts of the primal-dual approach. Abstractly, a constrained optimization problem has the form

$$(\mathcal{P}_0) \qquad \text{minimize } f_0(x) \text{ over all } x \in X \text{ satisfying } h(x) \in C,$$

where X is a set whose properties are relatively simple in comparison with the constraint set $\{x | h(x) \in C\}$. Primal-dual approaches to solving $(\mathcal{P}_0)$ consist of introducing a *dual space* Y and a *bivariate* functional $L : X \times Y \mapsto \overline{\mathbf{R}}$. (We use the symbol $\overline{\mathbf{R}}$ to denote the set of *extended-real numbers*: $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, \infty\}$.) Instead of $(\mathcal{P}_0)$ one then solves the associated *saddle point problem* (or "minimax" problem):

$$(\mathcal{S}) \qquad \begin{array}{l} \text{find } (\bar{x}, \bar{y}) \in X \times Y \text{ so that } L(\bar{x}, y) \leq L(\bar{x}, \bar{y}) \leq L(x, \bar{y}) \\ \text{for all } (x, y) \in X \times Y. \end{array}$$

A solution pair $(\bar{x}, \bar{y})$ for the problem $(\mathcal{S})$ is called a *saddle point* for L relative to $X \times Y$. For such a pair, the point $\bar{x}$ is a solution to the "primal" problem

$$(\mathcal{P}) \qquad \text{minimize the function } f(x) = \sup_{y \in Y} L(x, y) \text{ over all } x \in X.$$

For appropriate choices of Y and L, the problem $(\mathcal{P})$ can be made equivalent to the original problem $(\mathcal{P}_0)$; for other choices, $(\mathcal{P})$ may be considered as a perturbation of $(\mathcal{P}_0)$. As an example, suppose in the problem $(\mathcal{P}_0)$ that h is a real-valued function and C is the singleton $\{0\}$. Taking $Y = \mathbf{R}$ and $L(x,y) = f_0(x) + y \cdot h(x)$ yields the classical approach of *Lagrange multipliers* for equality-constrained minimization problems.

This dissertation deals with *internal* approximations of the minimax problem $(\mathcal{S})$, i.e. approximations of the sort

$$(\mathcal{S}_n) \qquad \begin{array}{l} \text{find } (\bar{x},\bar{y}) \in X_n \times Y_n \text{ so that } L(\bar{x},y) \le L(\bar{x},\bar{y}) \le L(x,\bar{y}) \\ \text{for all } (x,y) \in X_n \times Y_n, \end{array}$$

for subspaces $X_n \subset X$ and $Y_n \subset Y$. The focus is on functionals L having the form

$$L(x,y) = \varphi(x) - \psi(y) - \gamma(x,y), \tag{1.1}$$

where the functions φ and ψ are extended-real-valued and lower semicontinuous, and the functional γ is jointly sequentially continuous. We demonstrate that, for increasing sequences X_n and Y_n, the problems $(\mathcal{S}_n)$ *epi/hypo-converge* to the problem $(\mathcal{S})$: among other things, this guarantees that the solutions of $(\mathcal{S}_n)$ are in some sense approximations of the solution(s) of $(\mathcal{S})$. Our strongest general results are for the case where φ and ψ are lower semicontinuous, convex functions on Banach spaces X and Y, and γ is a continuous bilinear functional on $X \times Y$. The notion of "epi/hypo-convergence" was developed recently by Attouch and Wets ([2], [3]) as a tool for studying perturbations of minimax problems: it generalizes the widely applied concept of *epi-convergence* of minimization problems. Both of these concepts are defined and discussed in Chapter 2. For the moment, the following result illustrates one of the desirable consequences of epi/hypo-convergence . In Chapter 2, this proposition is derived as a corollary of a more general result.

Proposition. *Let L be a bivariate functional on a product $X \times Y$ of topological spaces, and consider a sequence $\{L_n\}$ of bivariate functionals epi/hypo-converging to L. Suppose that $\{L_{n_k}\}$ is a subsequence and, for each $k \in \mathbf{N}$, the pair $(\bar{x}_k, \bar{y}_k)$ is a saddle point for L_{n_k} relative to $X \times Y$. If $\bar{x}_k \to \bar{x}$ and $\bar{y}_k \to \bar{y}$, then $(\bar{x},\bar{y})$ is a*

saddle point for L relative to $X_n \times Y_n$. Furthermore, the values $L(\bar{x}_k, \bar{y}_k)$ converge to $L(\bar{x}, \bar{y})$.

Another important property is that epi/hypo-convergence is preserved under a wide class of perturbations: if the sequence $\{L_n\}$ epi/hypo-converges to L, then $\{L_n + G\}$ epi/hypo-converges to $L + G$ whenever G is a continuous real-valued function on $X \times Y$. (See [2].)

Primal-dual representations of the sort given in (1.1) arise naturally in many problems in large-scale optimization. Since either of the spaces X or Y (or both) may have high dimensionality, computational approaches may be based on low-dimensional internal approximations of minimax representations. The main results in this dissertation are concerned with approximations for dynamic and stochastic problems involving integral cost functionals: specifically, we consider convex problems in the optimal control of linear systems (Chapter 3) and in multistage stochastic programming (Chapter 4). These problems are briefly described below. They are posed in a manner which facilitates dualization via (integral) bivariate functionals of the type in (1.1). The approximations used are discretizations by increasingly fine partitions of the underlying measure spaces. (An important aspect of this class of approximations is that the resulting problems have a structure which is similar to the undiscretized problems.) The primary theorems state that under reasonable hypotheses, the approximate minimax problems *epi/hypo-converge* to the original minimax problem.

We now describe the optimal control problem that is the subject of Chapter 3. The particular problem formulation used here was recently introduced by Rockafellar [27], [31]. The primal problem takes the form

$$(\mathcal{P}) \qquad \begin{aligned} &\text{minimize the functional} \\ &F(u, u_e) = \int_{t_0}^{t_1} [p_t \cdot u_t + \varphi_t(u_t) - c_t \cdot x_t + \psi_t(q_t - D_t u_t - C_t x_t)]dt \\ &\qquad + p_e \cdot u_e + \varphi_e(u_e) - c_e \cdot x_{t_1} + \psi_e(q_e - D_e u_e - C_e x_{t_1}) \\ &\text{over the control space } \mathcal{U} = \mathcal{L}_k^1 \times \mathbf{R}^{ke}, \text{ with the dynamics given by} \\ &\dot{x}_t = A_t x_t + B_t u_t + b_t \text{ a.e.}, \quad x_{t_0} = B_e u_e + b_e. \end{aligned}$$

Here φ_t and ψ_t are assumed to be proper, lower semicontinuous, convex functions for each $t \in [t_0, t_1]$, and φ_e and ψ_e are similarly (proper) lower semicontinuous and

convex. (We also assume that φ_t and ψ_t vary *epi-continuously* with t.) Many of the usual problems in the optimal control of linear systems with convex costs can be put in this form. Since the functions φ_t, ψ_t, φ_e and ψ_e are allowed to be extended-real-valued, constraints on both controls and states (as well as endpoints) may be included. This formulation has the advantage that it is easily dualized, at least in the case where φ_t, φ_e are "coercive" and ψ_t, ψ_e are finite-valued. We introduce the dual space $\mathcal{V} = \mathcal{L}_l^1 \times \mathbf{R}^{l_e}$ and the bivariate functional

$$\mathcal{J}(u, u_e; v, v_e) = \int_{t_0}^{t_1} J_t(u_t, v_t)dt + J_e(u_e, v_e) - \gamma(u, u_e; v, v_e)$$

where

$$J_t(u_t, v_t) = \begin{cases} \infty & \text{if } \varphi_t(u_t) = \infty, \\ p_t \cdot u_t + \varphi_t(u_t) - v_t \cdot D_t u_t + q_t \cdot v_t - \psi_t^*(v_t) & \text{otherwise,} \end{cases}$$

$$J_e(u_e, v_e) = \begin{cases} \infty & \text{if } \varphi_e(u_e) = \infty, \\ p_e \cdot u_e + \varphi_e(u_e) - v_e \cdot D_e u_e + q_e \cdot v_e - \psi_e^*(v_e) & \text{otherwise,} \end{cases}$$

and

$$\gamma(u, u_e; v, v_e) = \int_{t_0}^{t_1} x_t \cdot (C_t^* v_t + c_t)dt + x_{t_1} \cdot (C_e^* v_e + c_e).$$

Here the superscript '$*$' denotes the transpose for a matrix, and the *convex conjugate* for a function: $h^*(r) = \sup_s\{r \cdot s - h(s)\}$. It turns out that the duality is also possible, *without* the hypothesis of coercivity, in the important case where the functions $\varphi_t, \varphi_e, \psi_t^*, \psi_e^*$ take the form of quadratic programs: for example

$$\varphi_t(u_t) = \begin{cases} \frac{1}{2} u_t \cdot P_t u_t & \text{if } u_t \in U_t, \\ \infty & \text{if } u_t \notin U_t, \end{cases}$$

where P_t is a symmetric, positive semi-definite matrix and U_t is a convex polyhedron. A formal statement (without proof) of the duality for both the quadratic and coercive cases is given in Chapter 3; for the full arguments we refer the reader to the papers of Rockafellar [27], [31].

Of course, this duality is useful only if the convex conjugates of the functions ψ_t and ψ_e are easily calculated. Nonetheless, this still allows for a large range of problems. As an example, consider the case where $\psi_t(s) = |s|^2$: we can think of the expression $\psi_t(q_t - D_t u_t - C_t x_t)$ as representing a "penalty" that is applied whenever

the equality $q_t = D_t u_t - C_t x_t$ is violated. The conjugate of ψ_t is $\psi_t^*(r) = (1/4)|r|^2$. Various modelling possibilities are presented in the paper of Rockafellar [27].

A number of authors have studied methods of solution for constrained optimal control problems. Some have used direct approximations of the primal problem [7], [8], [9], while others have introduced penalties into the primal formulation [5], [6], [36]. Others yet have worked with various dual formulations [11], [14], [20]. Most attempts involve Ritz-type approximations; Cullum's papers [6], [7], [8] have treated finite difference schemes. The main distinction between our approach and that of earlier investigators is the direct use of minimax formulations and their analysis through epi/hypo-convergence concepts. There are several advantages gained from this. First, many nonsmooth problems and most exact penalty representations can be written equivalently as minimax problems with smooth data and simple constraints. Second, epi/hypo-convergence involves convergence of both solutions *and* multipliers, affording better sensitivity analysis at approximate solutions. Finally, the perturbational properties of epi/hypo-convergence make it easier to extend convergence results to a broader class of approximations.

An important aspect of approximation that is not discussed in this dissertation is the rate of convergence. At the present there is very little known about rates of convergence for *any* type of approximation scheme for optimal control problems, and virtually nothing known about rates for primal-dual methods. Sharp rates have been established only for some primal *or* dual approaches to problems of very simple structure. The papers of Bosarge and Johnson [4] and Mathis and Reddien [16] treat unconstrained problems with quadratic objective functions; Hager [11] used a dual approach to include linear inequality constraints on states and controls; Chen and Mills [5] analyzed the effect of adding a penalty for a simple terminal state constraint. These four papers are concerned with Ritz-type approximation schemes. Hager [12] has also considered finite-difference schemes for unconstrained problems. Results on convergence rates typically require the data to have Lipschitz continuous derivatives of several orders; some also require the system to be controllable. We have opted to obtain convergence under only general assumptions on the data, such as continuity. For this reason, we consider only approximate controls which are piecewise constant. The arguments can be extended easily to other classes of approximate controls. We also assume that the trajectory corresponding

to an approximate control can be calculated precisely at the grid points. This assumption is not as restrictive as it may appear: the perturbational properties of epi/hypo-convergence allow the extension of our results to include approximation of trajectories as well. An example is worked out in Chapter 3 concerning a finite difference scheme. Alternatively, the problems can be reformulated (by augmenting the control space) so that the time-derivative of the state is also approximated.

Our second application is a piecewise linear-quadratic model in multistage stochastic programming, which was introduced by Rockafellar and Wets [35]. In a sense, this model generalizes the more familiar two-stage "recourse" problem, but the differences run deeper than simply tacking on more stages. A fundamental aspect of the newer model concerns the exchange of information between stages. The model is formulated using the notions of "controls" and "states". The controls represent the decision variables, as usual; the state represents a means of transferring information from one time stage to another. Introducing the state also aids in clarifying duality results. It has the added benefit of simplifying the notation both for theoretical and computational purposes. The model differs from the usual models in dynamic programming and discrete-time stochastic control in the way measurability requirements are used to model the information structure. For example, feedback considerations are not imposed here. We refer the reader to the above paper for further commentary on the distinctions. The problem we shall be working with is

$(\mathcal{P})$ minimize the functional $F(u)$ over the control space $\mathcal{U}$,

where

$$F(u) = \mathbf{E}\left\{ \begin{array}{l} \displaystyle\sum_{\tau=1}^{T} \left[\begin{array}{l} p_\tau \cdot u_\tau + \frac{1}{2} u_\tau \cdot P_\tau u_\tau - c_\tau \cdot x_{\tau-1} \\ \quad + \rho_{V_\tau, Q_\tau}(q_\tau - \mathbf{E}^{\mathcal{G}_\tau}[D_\tau u_\tau + C_\tau x_{\tau-1}]) \end{array} \right] \\ \quad + p_0 \cdot u_0 + \frac{1}{2} u_0 \cdot P_0 u_0 - c_{T+1} \cdot x_T \\ \quad\quad + \rho_{V_{T+1}, Q_{T+1}}(q_{T+1} - \mathbf{E}^{\mathcal{G}_{T+1}}[C_{T+1} x_T]) \end{array} \right\},$$

$$\mathcal{U} = \mathcal{U}_0 \times \cdots \times \mathcal{U}_T = \prod_{\tau=0}^{T} \{u_\tau(\cdot) \in \mathcal{L}^2_{k_\tau} | u_\tau \text{ is } \mathcal{F}_\tau\text{-measurable}, u_\tau \in U_\tau \text{ a.s.}\},$$

and the dynamics are given by

$$x_\tau = A_\tau x_{\tau-1} + B_\tau u_\tau + b_\tau \text{ a.s. for } \tau = 1,\dots,T, \quad x_0 = B_0 u_0 + b_0 \text{ a.s.}$$

The symbol $\mathbf{E}^{\mathcal{G}}$ indicates conditional expectation with respect to the field $\mathcal{G}$. The functions $\rho_{V,Q}$ are piecewise linear-quadratic, parametrized by the convex polyhedron V and the positive semidefinite matrix Q as:

$$\rho_{V_\tau,Q_\tau}(s) = \sup_{v\in V_\tau}\{s\cdot v - \frac{1}{2}v\cdot Q_\tau(\omega)v\}.$$

For details on the modelling possibilities of such ρ-functions, we refer the reader to [27], [35]. The minimax representation for $(\mathcal{P})$ is given by taking the dual space to be $\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_{T+1} = \prod_{\tau=1}^{T+1}\{v_\tau(\cdot) \in \mathcal{L}^2_{l_\tau} | v_\tau \text{ is } \mathcal{G}_\tau\text{-measurable}, v_\tau \in V_\tau \text{ a.s.}\}$ and defining the bivariate functional by

$$\mathcal{J}(u,v) = \begin{cases} \mathbf{E}\{J(u,v) - \gamma(u,v)\} & \text{if } u \in \mathcal{U}, v \in \mathcal{V}, \\ -\infty & \text{if } u \in \mathcal{U}, v \notin \mathcal{V}, \\ \infty & \text{if } u \notin \mathcal{U}, \end{cases}$$

where

$$\begin{aligned} J(u,v) = &\sum_{\tau=1}^{T}[p_\tau\cdot u_\tau + \tfrac{1}{2}u_\tau\cdot P_\tau u_\tau - v_\tau\cdot D_\tau u_\tau + q_\tau\cdot v_\tau - \tfrac{1}{2}v_\tau\cdot Q_\tau v_\tau] \\ &+ p_0\cdot u_0 + \tfrac{1}{2}u_0\cdot P_0 u_0 + q_{T+1}\cdot v_{T+1} - \tfrac{1}{2}v_{T+1}\cdot Q_{T+1}v_{T+1} \end{aligned}$$

and

$$\gamma(u,v) = \sum_{\tau=0}^{T} y_{\tau+1}\cdot(B_\tau u_\tau + b_\tau) = \sum_{\tau=1}^{T+1} x_{\tau-1}\cdot(C^*_\tau v_\tau + c_\tau).$$

In their paper [35], Rockafellar and Wets treat only (finite) discrete probability spaces. They suggest that such spaces may tend to be very large, and often result from the discretization of a continuous distribution. In Chapter 4 we derive a statement of strong duality in this model for arbitrary probability spaces. Next we discretize the probability space (again by partitioning) and show that the resulting problem has the same form as $(\mathcal{P})$. Finally, we demonstrate that for successively finer partitions, the approximate problems epi/hypo-converge to the original.

The reader may wonder why we use a partitioning scheme instead of sampling. The main reason is that the dynamic information structure is considered to be of primary importance in the model, and we wish to preserve this as much as possible. However, the perturbational properties of epi/hypo-convergence allow one to partition first, and afterward use sampling techniques to approximate the measure on the partition. This can be done in a way that is analogous to the extension to finite differences of the approximations for optimal control problems in Chapter 3. However, we won't elaborate on this further in this dissertation.

The final issue addressed in this dissertation is the means of solving the potentially large approximate problems that arise from discretizing the above models. In Chapter 5 we describe a general algorithm for minimax problems which allows great flexibility in adapting to the special structures associated with large-scale optimization problems: the *finite envelope method.* This method was first proposed by Rockafellar and Wets [32] as an approach to solving a piecewise linear-quadratic problem in stochastic programming; a computer implementation for that problem was undertaken by King [15]. We demonstrate that the algorithm can be applied to non-quadratic problems, and that it converges in this more general setting. We also indicate some details of its implementation for solving (piecewise linear-quadratic) optimal control problems and give some computational results.

We now summarize the organization of the rest of the dissertation. In Chapter 2 we review the concepts of convergence for sets, epi-convergence for (univariate) functions, and epi/hypo-convergence for bivariate functions. We then prove some general facts about convergence of saddle points and internal approximations of minimax problems. Chapter 3 is dedicated to the specialization of the concepts in Chapter 2 to the approximation of problems in the optimal control of linear systems with convex costs. In Chapter 4, we extend the duality theory for piecewise linear-quadratic problems in multistage stochastic programming to the case of general probability measures, and then demonstrate the convergence of approximations given by partitioning the probability space. Finally, Chapter 5 is devoted to a discussion of the finite envelope method for solving saddle point problems, and includes the description of an implementation of this algorithm to the solution of discretized optimal control problems.

Chapter 2. Variational Approximation

1. Epi-convergence

Let (X, τ) be a linear topological space. For a sequence $\{C_n\}$ of subsets of X we define the *Painlevé-Kuratowski limits* of $\{C_n\}$ to be

$$\liminf C_n = \bigcap_{M \in \mathcal{N}^\#} \mathrm{cl}[\bigcup_{m \in M} C_m],$$
$$\limsup C_n = \bigcap_{n \in \mathbf{N}} \mathrm{cl}[\bigcup_{m \geq n} C_m],$$

where $\mathbf{N}$ is the set of natural numbers and $\mathcal{N}^\#$ denotes the collection of all infinite subsets of $\mathbf{N}$. These limit sets are both closed and they satisfy

$$\liminf C_n \subset \limsup C_n.$$

The sequence $\{C_n\}$ is said to converge if $\liminf C_n = \limsup C_n$, in which case the common value is written as $\lim C_n$. Often, we shall consider more than one topology on X, in which case we write τ-liminf instead of lim inf, and similarly τ-limsup and τ-lim. If τ is first countable (e.g. metrizable), then we also have the representations

$$\begin{aligned} \liminf C_n &= \{x \in X | x = \lim x_n \text{ for some } x_n \in C_n\} \\ \limsup C_n &= \{x \in X | x = \lim x_k \text{ for some } x_k \in C_{n_k}\}. \end{aligned} \tag{2.1.1}$$

It turns out that the sets defined by the right-hand sides in (2.1.1) are also useful for topologies that are not necessarily first countable. We define the following "sequential" limits for a sequence of sets as

$$\begin{aligned} \text{seq-}\liminf C_n &= \{x \in X | x = \lim x_n \text{ for some } x_n \in C_n\} \\ \text{seq-}\limsup C_n &= \{x \in X | x = \lim x_k \text{ for some } x_k \in C_{n_k}\}. \end{aligned}$$

Clearly, seq-lim inf $C_n \subset$ seq-lim sup C_n; if they are equal we write seq-lim C_n for the common value. If τ is not first countable, then neither of these sequential limits need be closed. On the other hand, if the sets C_n are eventually convex, then both lim inf C_n and seq-lim inf C_n are convex. Of course, neither lim sup C_n nor seq-lim sup C_n need be convex even if all the sets C_n are.

Applying the above definitions to the *epigraphs* $\operatorname{epi} f_n = \{(x,\alpha)|f_n(x) \leq \alpha\} \subset X \times \mathbf{R}$ of a sequence of functions $f_n : X \to \overline{\mathbf{R}}$ leads to a notion of convergence for functions. The sequence $\{f_n\}$ is said to *epi-converge* to f if

$$\operatorname{epi} f = \lim(\operatorname{epi} f_n),$$

and we write $f = \mathrm{e}_\tau\text{-}\lim f_n$. The subscript τ is omitted when no confusion will arise. If instead we have

$$\operatorname{epi} f = \text{seq-}\lim(\operatorname{epi} f_n)$$

then we say that f is the *sequential* epi-limit of $\{f_n\}$, denoted $f = \text{seq-}\mathrm{e}_\tau\text{-}\lim f_n$. By carrying over the facts about limits of sets we see that $\text{e-}\lim f_n$ is always lower semi-continuous and, if the f_n are eventually convex, then both $\text{e-}\lim f_n$ and $\text{seq-e-}\lim f_n$ are convex.

The upper and lower limits of the epigraphs can also be shown to be epigraphs. Consider the function $x \mapsto (\text{e-ls}\, f_n)(x)$ defined by $\operatorname{epi}(\text{e-ls}\, f_n) = \liminf(\operatorname{epi} f_n)$. It is easily verified that

$$(\mathrm{e}_\tau\text{-ls}\, f_n)(x) = \sup_{U \in \mathcal{N}_\tau(x)} \limsup_{n\to\infty} \inf_{u \in U} f_n(u),$$

whence the notation (here $\mathcal{N}_\tau(x)$ denotes the τ-neighborhood system of x). We call this function the *epi-limit superior* of f_n and denote it by $\text{e-ls}\, f_n$. Similarly we define the epi-limit inferior by $\operatorname{epi}(\text{e-li}\, f_n) = \limsup(\operatorname{epi} f_n)$, and observe that

$$(\text{e-li}\, f_n)(x) = \sup_{U \in \mathcal{N}_\tau(x)} \liminf_{n\to\infty} \inf_{u \in U} f_n(u).$$

Note that $\text{e-ls}\, f_n \geq \text{e-li}\, f_n$. We say that f_n *epi-converges to* f *at* x if

$$(\text{e-ls}\, f_n)(x) \leq f(x) \leq (\text{e-li}\, f_n)(x). \tag{2.1.2}$$

Thus the sequence f_n epi-converges if and only if it epi-converges at every point.

Analogously, we may define the "sequential" epi-limits inferior and superior through the upper and lower sequential limits of epigraphs:

$$\operatorname{epi}(\text{seq-e-ls}\, f_n) = \text{seq-}\liminf(\operatorname{epi} f_n)$$
$$\operatorname{epi}(\text{seq-e-li}\, f_n) = \text{seq-}\limsup(\operatorname{epi} f_n).$$

The use of the notation "li" and "ls" in the sequential case is justified by the following representations:

$$(\text{seq-e}_\tau\text{-ls}\, f_n)(x) = \inf_{x_n \overset{\tau}{\to} x} \limsup_{n\to\infty} f_n(x_n),$$
$$(\text{seq-e}_\tau\text{-li}\, f_n)(x) = \inf_{x_n \overset{\tau}{\to} x} \liminf_{n\to\infty} f_n(x_n).$$

As expected, sequential epi-convergence is characterized by

$$\text{seq-e-ls}\, f_n \le f \le \text{seq-e-li}\, f_n. \tag{2.1.3}$$

Thus f_n epi-converges sequentially to f if and only if the following two statements are true:

(a) $\forall x, \exists x_n \to x$ such that $\limsup f_n(x_n) \le f(x)$,

(b) $\liminf f_{n_k}(x_k) \ge f(x)$, $\forall x_k \to x$ and $\forall$ subsequences $\{f_{n_k}\}$.

One of the primary motivations for the study of epi-convergence is the following:

Proposition 2.1.1. *Suppose* $x_n \in \operatorname{argmin} f_n$ *and* $x_n \to x$. *If either* $f = \text{e-lim}\, f_n$ *or* $f = \text{seq-e-lim}\, f_n$ *then*

$$x \in \operatorname{argmin} f \quad \text{and} \quad \inf f = \lim f_n(x_n) = \lim(\inf f_n).$$

This result is an immediate consequence of the representations in (2.1.2) and (2.1.3). (It is also a special case of Corollary 2.3.5.)

Often, a given space X is equipped with more than one topology. Besides using the set-limits and epi-limits defined above for each topology, it is sometimes necessary to combine them. The best-known example of such a combination is the concept of Mosco convergence, introduced by U. Mosco in 1969 [18].

Definition. *Let* X *be a reflexive Banach space, and let* s *and* w *denote the strong (norm) and weak topologies on* X. *The sequence* $C_n \subset X$ *is said to* Mosco converge *to* C, *written* $C = \text{M-lim}\, C_n$, *if*

$$\text{seq-}w\text{-limsup}\, C_n = C = \text{seq-}s\text{-liminf}\, C_n.$$

A sequence $\{f_n\}$ *of functions on* X *is said to* Mosco epi-converge *to* f *if the epigraphs* $\text{epi}\, f_n$ *Mosco converge to* $\text{epi}\, f$ *or, equivalently, if*

$$\text{seq-e}_\text{w}\text{-li}\, f_n = f = \text{seq-e}_\text{s}\text{-ls}\, f_n.$$

We write $f = \text{M-e-lim}\, f_n$.

Using the facts that s is first countable and $w \subset s$, we observe that

$$\begin{array}{ccccc} \text{seq-}w\text{-limsup}\, C_n & \supset & \text{seq-}s\text{-limsup}\, C_n & = & s\text{-limsup}\, C_n \\ \cup & & & & \cup \\ \text{seq-}w\text{-liminf}\, C_n & \supset & \text{seq-}s\text{-liminf}\, C_n & = & s\text{-liminf}\, C_n. \end{array}$$

Hence C_n Mosco converges if and only if

$$\text{seq-}w\text{-limsup}\, C_n \subset s\text{-liminf}\, C_n.$$

Moreover, M-lim C_n is always strongly closed. If the C_n are eventually convex, then M-lim C_n will be convex and thus weakly closed. Translating this to functions, we find that the Mosco epi-limit of a sequence of convex functions is a (strongly and weakly) lower semicontinuous convex function.

The *Legendre-Fenchel transform* of a convex function f on X is defined to be the function f^* on X^* given by

$$x^* \longmapsto \sup_{x \in X} [\langle x^*, x \rangle - f(x)].$$

This transformation plays a central role in both convex analysis and optimization theory. The following result indicates the importance of Mosco epi-convergence in this context. It was proved by Mosco in 1971 [19], and generalizes the finite-dimensional version obtained by Wijsman in 1966 [38].

Theorem 2.1.2. *Suppose that* $\{f_n\}$ *and* f *are proper*[1], *lower semicontinuous convex functions on a reflexive Banach space. Then*

$$f = \text{M-lim}\, f_n \text{ on } X \iff f^* = \text{M-lim}\, f_n^* \text{ on } X^*.$$

In complete analogy to epi-convergence, there is a notion of hypo-convergence for a sequence $\{g_n\}$ of $\overline{\mathbf{R}}$-valued functions on X. This is introduced by applying the above definitions of the limits of sets to the *hypographs* $\text{hypo}\, g_n = \{(x, \alpha) | g_n(x) \geq \alpha\}$

[1] A convex function is said to be *proper* if it is not identically $+\infty$ and nowhere takes the value $-\infty$.

in $X \times \mathbf{R}$ or, equivalently, by using the definitions of epi-limits of functions applied to $-g_n$. Accordingly, we define the functions

$$(\mathrm{h}_\tau\text{-ls}\, g_n)(y) = \inf_{V \in \mathcal{N}_\tau(y)} \limsup_{n\to\infty} \sup_{v \in V} g_n(v)$$
$$(\mathrm{h}_\tau\text{-li}\, g_n)(y) = \inf_{V \in \mathcal{N}_\tau(y)} \liminf_{n\to\infty} \sup_{v \in V} g_n(v),$$

as well as their sequential counterparts

$$(\text{seq-h}_\tau\text{-ls}\, g_n)(y) = \sup_{y_n \overset{\tau}{\to} y} \limsup_{n\to\infty} g_n(y_n)$$
$$(\mathrm{h}_\tau\text{-li}\, g_n)(x) = \sup_{y_n \overset{\tau}{\to} y} \liminf_{n\to\infty} g_n(y_n).$$

Clearly we have the inequalities

$$\text{h-ls}\, g_n \geq \text{h-li}\, g_n$$
$$\text{seq-h-ls}\, g_n \geq \text{seq-h-li}\, g_n.$$

The sequence $\{g_n\}$ is said to *hypo-converge* if

$$\text{h-ls}\, g_n = \text{h-li}\, g_n$$

and we write $\text{h-lim}\, g_n$ to denote the common value. Similarly, we can define *sequential hypo-convergence*. Parallel to the situation for epi-limits, the functions $\text{h-ls}\, g_n$ and $\text{h-li}\, g_n$ are both upper semicontinuous. If the g_n are eventually all concave, then both $\text{h-ls}\, g_n$ and $\text{seq-h-ls}\, g_n$ are concave. Also, if X is a reflexive Banach space, we say that g_n *Mosco hypo-converges* if

$$\text{seq-h}_\text{w}\text{-ls}\, g_n \leq \text{h}_\text{s}\text{-li}\, g_n.$$

Finally, there are the obvious parallel statements of Proposition 2.1.1 and Theorem 2.1.2 for hypo-convergence, where we substitute "sup" for "inf," "concave" for "convex," etc.

2. Closed Saddle Functions

In this section we review the concept of *closed saddle functions,* which play a role in minimax theory similar to that played by closed convex functions in minimization

theory. For a more detailed presentation, we refer the reader to the paper by Rockafellar [21]; see also [22].

Let (X, τ) and (Y, σ) be locally convex Hausdorff spaces. Consider a bivariate function $K : X \times Y \to \overline{\mathbf{R}}$. We shall think of such functions as a representing a "minimax" problem where we minimize with respect to $x \in X$ and maximize with respect to $y \in Y$. Note that the inequality

$$\inf_{x \in X} \sup_{y \in Y} K(x, y) \geq \sup_{y \in Y} \inf_{x \in X} K(x, y)$$

is always valid. In the case that

$$\inf_{x \in X} \sup_{y \in Y} K(x, y) = \sup_{y \in Y} \inf_{x \in X} K(x, y)$$

the common value is called the *saddle value* for K. A pair $(\bar{x}, \bar{y})$ is said to be a *saddle point* for K if, for all $(x, y) \in X \times Y$, it is true that

$$K(\bar{x}, y) \leq K(\bar{x}, \bar{y}) \leq K(x, \bar{y}).$$

The *minimax* problem associated with K is that of finding the saddle value and saddle points for K, whenever either of these exist.

We define the *effective domain* of K as

$$\operatorname{dom} K := \operatorname{dom}_1 K \times \operatorname{dom}_2 K := \{x | K(x, \cdot) < \infty\} \times \{y | K(\cdot, y) > -\infty\}.$$

K is said to be *proper* if $\operatorname{dom} K \neq \emptyset$. As with minimization problems, where we interpret the effective domain of an extended-real-valued objective function as specifying the "constraint" set, the effective domain of a bivariate function is the set to which we are necessarily restricted in the search for saddle points.

To make full use of the usual theory of minimization (or maximization), we need to impose some sort of regularity hypotheses on K. Ideally, we would like to assume that $K(x, y)$ is lower semicontinuous in x and upper semicontinuous in y. It turns out that this requirement is too restrictive: it strictly prohibits the use of bivariate functions which take both of the values ∞ and $-\infty$, thus excluding the possibility of modelling constraints on both x and y through the

use of infinite penalties. To deal with this difficulty, Rockafellar [21] introduced an equivalence relation for saddle functions, which we extend to general bivariate functions. Within this framework, the natural regularity condition to impose is that a bivariate function be equivalent to certain upper and lower regularizations of itself. Functions satisfying this condition will be called *closed.*

We define the *epi-closure* $\mathrm{cl}_1 K$ of K to be the lower semicontinuous regularization of K in x. Equivalently, $\mathrm{cl}_1 K$ is the function which, for each $y \in Y$, satisfies

$$\mathrm{epi}(\mathrm{cl}_1 K(\cdot, y)) = \mathrm{cl}(\mathrm{epi}\, K(\cdot, y)).$$

The *lower closure* $\underline{\mathrm{cl}}_1 K$ is then defined as the function which satisfies, for each $y \in Y$,

$$\underline{\mathrm{cl}}_1 K(\cdot, y) = \begin{cases} \mathrm{cl}_1 K(\cdot, y) & \text{if } \mathrm{cl}_1 K(\cdot, y) > -\infty, \\ -\infty & \text{otherwise.} \end{cases}$$

Similarly, we define the *hypo-closure* of K by

$$\mathrm{hypo}(\mathrm{cl}_2 K(x, \cdot)) = \mathrm{cl}(\mathrm{hypo}\, K(x, \cdot)) \quad \text{for all } x \in X,$$

and the *upper closure* of K by

$$\overline{\mathrm{cl}}_2 K(x, \cdot) = \begin{cases} \mathrm{cl}_2 K(x, \cdot) & \text{if } \mathrm{cl}_2 K(x, \cdot) < +\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

Note that if L is another bivariate function on $X \times Y$ for which $K \leq L$, then $\underline{\mathrm{cl}}_1 K \leq \underline{\mathrm{cl}}_1 L$ and $\overline{\mathrm{cl}}_2 K \leq \overline{\mathrm{cl}}_2 L$. We say that two bivariate functions K and L are *equivalent* if

$$\underline{\mathrm{cl}}_1 K = \underline{\mathrm{cl}}_1 L \text{ and } \overline{\mathrm{cl}}_2 K = \overline{\mathrm{cl}}_2 L.$$

The function K is said to be *closed* if it is equivalent to both $\underline{\mathrm{cl}}_1 K$ and $\overline{\mathrm{cl}}_2 K$, in which case we write $\overline{K} = \overline{\mathrm{cl}}_2 K$ and $\underline{K} = \underline{\mathrm{cl}}_1 K$. If K is closed and $\underline{K} \leq L \leq \overline{K}$, then we write $L \in [\underline{K}, \overline{K}]$. The next two propositions are immediate consequences of the definitions.

Proposition 2.2.1. *Suppose K is closed.*

(i) *$L \in [\underline{K}, \overline{K}]$ if and only if K and L are equivalent.*

(ii) *If L is also closed, then L is equivalent to K if and only if $\underline{K} \leq \overline{L}$ and $\underline{L} \leq \overline{K}$.*

Proposition 2.2.2. *The following are true for any bivariate function K:*

(i) *If $\mathrm{dom}_1 K = \emptyset$ then $\overline{\mathrm{cl}}_2 K \equiv \infty$.*

(ii) *If* $\operatorname{dom}_2 K = \emptyset$ *then* $\underline{\operatorname{cl}}_1 K \equiv -\infty$.

In particular, the only closed improper bivariate functions are $K \equiv \infty$ *and* $K \equiv -\infty$.

Our primary interest with bivariate functions, as mentioned above, is in their interpretation as minimax problems. The following shows why the term *equivalent* is justified as applied to bivariate functions.

Proposition 2.2.3. *Suppose that* K *is closed and is equivalent to* L.

(i) $\operatorname{cl}(\operatorname{dom} L) = \operatorname{cl}(\operatorname{dom} K)$.

(ii) $L(x,y) = \begin{cases} -\infty & \text{if } x \in \operatorname{dom}_1 K,\ y \in Y \setminus \operatorname{cl}(\operatorname{dom}_2 K), \\ \infty & \text{if } x \in X \setminus \operatorname{cl}(\operatorname{dom}_1 K),\ y \in \operatorname{dom}_2 K. \end{cases}$

(iii) *If a saddle value exists for* K *then it is also the saddle value for* L.

(iv) *If* $(\bar{x}, \bar{y})$ *is a saddle point for* K, *then it is also a saddle point for* L.

(v) L *is finite on* $\operatorname{dom}_1 \overline{K} \times \operatorname{dom}_2 \underline{K}$.

A *saddle function* on $X \times Y$ is a bivariate function K for which $x \mapsto K(x,y)$ is a convex function for each $y \in Y$, and $y \mapsto K(x,y)$ is a concave function for each $x \in X$. For such a function we define the *convex* and *concave parents* F and G of K by the following partial conjugation formulas:

$$F(x,y^*) = \sup_y \{K(x,y) - \langle y^*, y\rangle\} \quad \text{for } x \in X, y^* \in Y^*,$$
$$G(x^*,y) = \inf_x \{K(x,y) - \langle x^*, x\rangle\} \quad \text{for } x^* \in X^*, y \in Y.$$

Here (X^*, τ^*) and (Y^*, σ^*) denote locally convex spaces paired with X and Y for which the pairings are compatible with the respective topologies. The parents depend only on the equivalence class of K and, in fact, two saddle functions are equivalent if and only if they have the same parents. The equivalence class for K may be recovered by the formulas

$$\overline{\operatorname{cl}}_2 K(x,y) = \inf_{y^*} \{F(x,y^*) + \langle y^*, y\rangle\},$$
$$\underline{\operatorname{cl}}_1 K(x,y) = \sup_{x^*} \{G(x^*,y) + \langle x^*, x\rangle\}.$$

In particular, a proper saddle function K is closed if and only if its parents are "reverse conjugates" to each other, i.e if and if only they satisfy

$$\begin{aligned} F(x,y^*) &= \sup_{x^*,y} \{G(x^*,y) - \langle y^*, y\rangle + \langle x^*, x\rangle\}, \\ G(x^*,y) &= \inf_{x,y^*} \{F(x,y^*) - \langle x^*, x\rangle + \langle y^*, y\rangle\}. \end{aligned} \tag{2.2.1}$$

Also, for any $L \in [\underline{K}, \overline{K}]$,

$$\operatorname{dom}_1 L = \operatorname{proj}_1[\operatorname{dom} F] \text{ and } \operatorname{dom}_2 L = \operatorname{proj}_2[\operatorname{dom} G],$$

where $\operatorname{proj}_1, \operatorname{proj}_2$ are the projections onto the first and second arguments respectively. From the facts above, we see that there are one-to-one correspondences between the class of closed proper saddle functions on $X \times Y$, the class of closed proper convex functions on $X \times Y^*$ and the class of closed proper concave functions on $X^* \times Y$.

Finally, it should be noted that a bivariate function L which is equivalent to a saddle function K is not necessarily a saddle function. However, we may define the convex and concave parents for any bivariate function. We see then that a proper bivariate function K is equivalent to some closed (proper) saddle function if and only if its convex and concave parents satisfy the conjugacy conditions (2.2.1), in which case both of $\overline{\operatorname{cl}}_2 K = \overline{K}$ and $\underline{\operatorname{cl}}_1 K = \underline{K}$ are saddle functions.

3. Epi/hypo-convergence

In this section we review the definitions of epi/hypo-convergence, state some of the basic facts from the theory, and also prove some general results concerning internal approximations. The definitions given here generalize those in §2.1. Let (X, τ) and (Y, σ) be linear topological spaces and let $\{K_n\}$ be a sequence of bivariate functions on $X \times Y$. We define the *epi/hypo-limits superior* and *inferior* of $\{K_n\}$ to be

$$(\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, K_n)(x,y) = \sup_{U \in \mathcal{N}_\tau(x)} \inf_{V \in \mathcal{N}_\sigma(y)} \limsup_{n\to\infty} \sup_{v \in V} \inf_{u \in U} K_n(u,v),$$

$$(\mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, K_n)(x,y) = \inf_{V \in \mathcal{N}_\sigma(y)} \sup_{U \in \mathcal{N}_\tau(x)} \liminf_{n\to\infty} \inf_{u \in U} \sup_{v \in V} K_n(u,v).$$

In general, these two functions are **not** comparable. We say that K_n *τ-epi/σ-hypo-converges* to K if

$$\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, K_n \le K \le \mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, K_n.$$

In this case, we write $K = \mathrm{e}_\tau/\mathrm{h}_\sigma\text{-lim}\, K_n$ (or simply e/h-lim K_n if the topologies are understood) even though this "limit" may not be unique. This definition is due to Attouch and Wets. In their paper [3], they note that the function $(\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, K_n)(\cdot, y)$ is τ-lower semicontinuous for each $y \in Y$. (This can be seen from the easily verified

fact that $x \mapsto \sup_{U \in \mathcal{N}_\tau(x)} \gamma(U)$ is τ-lower semicontinuous for any extended-real-valued function γ defined on the subsets of X.) Similarly, $(\mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, K_n)(x,\cdot)$ is σ-upper semicontinuous for each $x \in X$.

As with epi-convergence, it is also useful to have a "sequential" version of epi/hypo-convergence. For this purpose we define the following:

$$(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(x,y) = \sup_{y_n \overset{\sigma}{\to} y} \inf_{x_n \overset{\tau}{\to} x} \limsup_{n\to\infty} K_n(x_n,y_n),$$
$$(\text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n)(x,y) = \inf_{x_n \overset{\tau}{\to} x} \sup_{y_n \overset{\sigma}{\to} y} \liminf_{n\to\infty} K_n(x_n,y_n).$$

Again, these functions are not, in general, comparable. In the case that

$$\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n \le K \le \text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n,$$

we say that K_n *epi/hypo-converges sequentially* to K and indicate this by writing $K = \text{seq-e}_\tau/\text{h}_\sigma\text{-lim}\, K_n$, although this "limit" need not be unique. Also note that these sequential limits may be rewritten as

$$(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(x,y) = \sup_{y_n \overset{\sigma}{\to} y} [(\text{seq-e}_\tau\text{-ls}\, K_n(\cdot,y_n))(x)],$$
$$(\text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n)(x,y) = \inf_{x_n \overset{\tau}{\to} x} [(\text{seq-e}_\sigma\text{-li}\, K_n(x_n,\cdot))(y)].$$

The following proposition was proved in [2]. We include a simpler proof based on the well-known fact that the Kuratowski and sequential limits of sets coincide for first countable topologies.

Proposition 2.3.1. *If σ and τ are first countable, then*

$$\mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, K_n = \text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n,$$
$$\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, K_n = \text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n.$$

Proof. For $y \in Y$ and $U \subset X$, define $g_n(y|U) = \inf_{u\in U} K_n(u,y)$. Then we have

$$\begin{aligned}(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(\bar{x},\bar{y}) &= \sup_{y_n \overset{\sigma}{\to} y} [(\text{seq-e}_\tau\text{-ls}\, K_n(\cdot,y_n))(x)] \\ &= \sup_{y_n \overset{\sigma}{\to} y} [(\mathrm{e}_\tau\text{-ls}\, K_n(\cdot,y_n))(x)]\end{aligned}$$

$$
\begin{aligned}
&= \sup_{y_n \overrightarrow{\sigma} y} \left[\sup_{U \in \mathcal{N}_\tau(\bar{x})} \limsup_{n \to \infty} \inf_{u \in U} K_n(u, y_n) \right] \\
&= \sup_{U \in \mathcal{N}_\tau(\bar{x})} \left[\sup_{y_n \overrightarrow{\sigma} y} \limsup_{n \to \infty} g_n(y_n | U) \right] \\
&= \sup_{U \in \mathcal{N}_\tau(\bar{x})} [(\text{seq-h}_\sigma\text{-ls}\, g_n(\cdot | U))(\bar{y})] \\
&= \sup_{U \in \mathcal{N}_\tau(\bar{x})} [(\text{h}_\sigma\text{-ls}\, g_n(\cdot | U))(\bar{y})] \\
&= \sup_{U \in \mathcal{N}_\tau(\bar{x})} \left[\inf_{V \in \mathcal{N}_\sigma(\bar{y})} \limsup_{n \to \infty} \sup_{v \in V} g_n(v | U) \right] \\
&= (\text{e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(\bar{x}, \bar{y}).
\end{aligned}
$$

The proof that $\text{h}_\sigma/\text{e}_\tau\text{-li}\, K_n = \text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n$ is similar. □

Proposition 2.3.2. *If τ is first countable then the function* $(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(\cdot, y)$ *is τ-lower semicontinuous for each $y \in Y$. Similarly, if σ is first countable then, for each $x \in X$, the function* $(\text{seq-h}_\sigma/\text{e}_\tau\text{-li}\, K_n)(x, \cdot)$ *is σ-upper semicontinuous.*

Proof. We have

$$
\begin{aligned}
(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, K_n)(\bar{x}, \bar{y}) &= \sup_{y_n \overrightarrow{\sigma} y} [(\text{seq-e}_\tau\text{-ls}\, K_n(\cdot, y_n))(x)] \\
&= \sup_{y_n \overrightarrow{\sigma} y} [(\text{e}_\tau\text{-ls}\, K_n(\cdot, y_n))(x)].
\end{aligned}
$$

The desired lower semicontinuity follows from that of $\text{e}_\tau\text{-ls}\, K_n(\cdot, y_n)$ and the fact that the supremum of lower semicontinuous functions is also lower semicontinuous. The other part of the proposition follows by symmetry. □

It is clear from the definitions that if $\{K_{n_k}\}$ is any subsequence of $\{K_n\}$, then one has

$$
\begin{aligned}
\text{h/e-li}\, K_n &\geq \text{h/e-li}\, K_{n_k}, \\
\text{seq-h/e-li}\, K_n &\geq \text{seq-h/e-li}\, K_{n_k},
\end{aligned}
$$

and that similar inequalities hold for the epi/hypo-limits superior. Thus we see that if $K = \text{e/h-lim}\, K_n$ then $K = \text{e/h-lim}\, K_{n_k}$; similarly, if $K = \text{seq-e/h-lim}\, K_n$ then $K = \text{seq-e/h-lim}\, K_{n_k}$.

As with the concept of epi-convergence, there are various results relating the convergence of bivariate functions to the convergence of their saddle values and

saddle points. We shall now prove several general results along this line. First we give a small technical result.

Lemma 2.3.3. *Suppose that* $\bar{x} = \tau\text{-lim}\,\bar{x}_n$ *and* $\bar{y}_n \in \operatorname{argmax}_{y \in Y} K_n(\bar{x}_n, y)$. *If either*

(i) $\mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\,K_n \geq K$, *or*

(ii) $\text{seq-h}_\sigma/\mathrm{e}_\tau\text{-li}\,K_n \geq K$,

then $K(\bar{x}, y) \leq \liminf_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$ *for all* $y \in Y$.

Proof. Let $y \in Y$ be given. First assume that (i) holds. Assume also that $K(\bar{x}, y) > -\infty$ (if $K(\bar{x}, y) = -\infty$, we're done) and consider $\alpha < K(\bar{x}, y)$. By (i), since $Y \in \mathcal{N}_\sigma(y)$, we can find $U \in \mathcal{N}_\tau(\bar{x})$ and $N \in \mathbf{N}$ so that $\sup_{v \in Y} K_n(u, v) \geq \alpha$ for all $u \in U$ and all $n \geq N$. Choose $N_1 \geq N$ so that $\bar{x}_n \in U$ whenever $n \geq N_1$. Then, for $n \geq N_1$, we have

$$\alpha \leq \sup_v K_n(\bar{x}_n, v) \leq K_n(\bar{x}_n, \bar{y}_n)$$

since $\bar{y}_n \in \operatorname{argmax} K_n(\bar{x}_n, \cdot)$. Thus $\liminf_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n) \geq \alpha$. Since α can be arbitrarily close to $K(\bar{x}, y)$, the proof is complete in the case of (i). If (ii) holds, then there exists $y_n \overset{\sigma}{\to} y$ such that $\liminf_{n\to\infty} K_n(\bar{x}_n, y_n) \geq K(\bar{x}, y)$. Since the inequality $K_n(\bar{x}_n, \bar{y}_n) \geq K_n(\bar{x}_n, y_n)$ holds for each n, the desired result follows. □

The following theorem indicates an important aspect of epi/hypo-convergence. It is a slight generalization of a similar result of Attouch and Wets [2],[3]. Note that the condition (i) is actually more general than epi/hypo-convergence since it mentions two different topologies on each of X and Y.

Theorem 2.3.4. *Let* τ_1 *and* τ_2 *be topologies on* X *and* σ_1 *and* σ_2 *be topologies on* Y. *Suppose that* $\bar{x} = \tau_1\text{-lim}\,\bar{x}_n$, $\bar{y} = \sigma_1\text{-lim}\,\bar{y}_n$ *and that* $(\bar{x}_n, \bar{y}_n)$ *is a saddle point for* K_n. *If either*

(i) $\mathrm{e}_{\tau_2}/\mathrm{h}_{\sigma_1}\text{-ls}\,K_n \leq K \leq \mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\,K_n$, *or*

(ii) $\text{seq-e}_{\tau_2}/\mathrm{h}_{\sigma_1}\text{-ls}\,K_n \leq K \leq \text{seq-h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\,K_n$,

then $(\bar{x}, \bar{y})$ *is a saddle point for* K *and* $K(\bar{x}, \bar{y}) = \lim_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$.

Proof. By the above lemma, $K(\bar{x}, y) \leq \liminf_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$ for all y. Similarly, we have $K(x, \bar{y}) \geq \limsup_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$ for all x. Since $\liminf \leq \limsup$,

we may combine these inequalities to obtain $K(\bar{x}, y) \leq K(x, \bar{y})$. On the other hand, we may use these inequalities with $x = \bar{x}$ and $y = \bar{y}$ to obtain

$$\limsup_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n) \leq K(\bar{x}, \bar{y}) \leq \liminf_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n),$$

so that $K(\bar{x}, \bar{y}) = \lim_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$. □

Corollary 2.3.5. *Let $\{K_n\}$ be sequence of bivariate functions, and consider a subsequence $\{K_{n_k}\}$. Suppose that $\bar{x} = \tau\text{-lim}\, \bar{x}_k$ and $\bar{y} = \sigma\text{-lim}\, \bar{y}_k$, where $(\bar{x}_k, \bar{y}_k)$ is a saddle point for K_{n_k}. If either $K = \mathrm{e}_\tau/\mathrm{h}_\sigma\text{-lim}\, K_n$ or $K = \text{seq-}\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-lim}\, K_n$ then $(\bar{x}, \bar{y})$ is a saddle point for K and $K(\bar{x}, \bar{y}) = \lim_{k\to\infty} K_{n_k}(\bar{x}_k, \bar{y}_k)$.*

A similar result is given by the following theorem, which is new. The conditions (i) and (ii) represent a kind of convergence for equivalence classes. A related concept is "Mosco epi/hypo-convergence", which is described in the next section.

Theorem 2.3.6. *Let K be a (τ_2, σ_2)-closed bivariate function on $X \times Y$. Suppose that $(\bar{x}_n, \bar{y}_n)$ is a saddle point for K_n and that $\bar{x}_n \overrightarrow{\tau_1} \bar{x}$ and $\bar{y}_n \overrightarrow{\sigma_1} \bar{y}$. Assume that one of the following holds:*

(i) $\underline{K} \leq \mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\, K_n$ *and* $\overline{K} \geq \mathrm{e}_{\tau_2}/\mathrm{h}_{\sigma_1}\text{-ls}\, K_n$.

(ii) $\underline{K} \leq \text{seq-}\mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\, K_n$, $\overline{K} \geq \text{seq-}\mathrm{e}_{\tau_2}/\mathrm{h}_{\sigma_1}\text{-ls}\, K_n$, *and both τ_2 and σ_2 are first countable.*

(The "closures" $\underline{K}$ and $\overline{K}$ of K are with respect to τ_2 and σ_2.)
Then $(\bar{x}, \bar{y})$ is a saddle point for $[\underline{K}, \overline{K}]$ and $K(\bar{x}, \bar{y}) = \lim_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n)$.

Proof. Assume that (i) holds: the proof for (ii) is similar. Recall that

$$(\mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\, K_n)(x, \cdot)$$

is σ_2-upper semicontinuous for each x. Hence $\underline{K} \leq \mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\, K_n$ implies that $\mathrm{cl}_2\, \underline{K} \leq \mathrm{h}_{\sigma_2}/\mathrm{e}_{\tau_1}\text{-li}\, K_n$. If $\mathrm{cl}_2\, \underline{K}(\bar{x}, \cdot) = \overline{K}(\bar{x}, \cdot)$, then Lemma 2.3.3 gives us

$$\overline{K}(\bar{x}, y) \leq \liminf K_n(\bar{x}_n, \bar{y}_n) \text{ for all } y \in Y \text{ and}$$
$$\overline{K}(\bar{x}, y) \geq \liminf K_n(\bar{x}_n, \bar{y}_n) \text{ for all } x \in X,$$

and we proceed exactly as in the proof of Theorem 2.3.4. If instead we have $\mathrm{cl}_2\, \underline{K}(\bar{x}, \cdot) \neq \overline{K}(\bar{x}, \cdot)$, then there exists y' for which $\mathrm{cl}_2\, \underline{K}(\bar{x}, y') = \infty$. By Lemma 2.3.3.,

$$\mathrm{cl}_2\, \underline{K}(\bar{x}, y) \leq \liminf K_n(\bar{x}_n, \bar{y}_n), \text{ for all } y. \tag{2.3.1}$$

Similarly, we may obtain,

$$\limsup K_n(\bar{x}_n, \bar{y}_n) \leq \overline{\mathrm{cl}}_2\, \underline{K}(x, \bar{y}), \text{ for all } x.$$

Hence $\mathrm{cl}_2\, \underline{K}(\bar{x}, y) \leq \overline{\mathrm{cl}}_2\, \underline{K}(x, \bar{y})$ for all x and y, so

$$\infty = \mathrm{cl}_2\, \underline{K}(\bar{x}, y') \leq \overline{\mathrm{cl}}_2\, \underline{K}(x, \bar{y})$$

for all x. Thus $\overline{K}(\cdot, \bar{y}) \equiv \infty$, so $\overline{K}(\bar{x}, y) \leq \overline{K}(x, \bar{y})$ for all x and y, i.e. $(\bar{x}, \bar{y})$ is a saddle point for $[\underline{K}, \overline{K}]$. By (2.3.1), with $y = y'$, we see that

$$\lim_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n) = \infty,$$

as desired. □

Numerical procedures typically can find only approximate solutions even for approximate problems. The pair $(\bar{x}, \bar{y})$ is said to be an ε-saddle point ($\varepsilon \geq 0$) for K if

$$\sup_y K(\bar{x}, y) - \varepsilon \leq K(\bar{x}, \bar{y}) \leq \inf_x K(x, \bar{y}) + \varepsilon.$$

Our next result illustrates the relationship between ε-saddle points for a sequence $\{K_n\}$ and ε-saddle points for the epi/hypo-limit of $\{K_n\}$. A partial converse is given by Attouch and Wets [3].

Proposition 2.3.7. *Suppose $(\bar{x}_n, \bar{y}_n)$ is an ε_n-saddle point for K_n, with $\bar{x}_n \overset{\tau}{\to} \bar{x}$, $\bar{y}_n \overset{\sigma}{\to} \bar{y}$, and $\varepsilon_n \to \varepsilon \geq 0$. If either*

(i) $\mathrm{e}_{\tau_1}/\mathrm{h}_{\sigma}\text{-ls}\, K_n \leq K \leq \mathrm{h}_{\sigma_1}/\mathrm{e}_{\tau}\text{-li}\, K_n$, *or*

(ii) $\text{seq-e}_{\tau_1}/\mathrm{h}_{\sigma}\text{-ls}\, K_n \leq K \leq \text{seq-h}_{\sigma_1}/\mathrm{e}_{\tau}\text{-li}\, K_n$,

then $(\bar{x}, \bar{y})$ is a 2ε-saddle point for K and

$$K(\bar{x}, \bar{y}) \in [\limsup_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n) - \varepsilon, \liminf_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n) + \varepsilon].$$

Sketch of Proof. This requires only a modification of the proof of Theorem 2.3.4. For example, suppose (ii) holds. First we generalize Lemma 2.3.3 by replacing all occurences of "argmax" by "ε-argmax", all occurences of "α" by "$\alpha + \varepsilon$", etc. Then, for all y, there exist $y_n \overset{\tau}{\to}_1 y$ so that $\liminf K_n(\bar{x}_n, y_n) \geq K(\bar{x}, y)$. Since

$K_n(\bar{x}_n, \bar{y}_n) + \varepsilon_n \geq K_n(\bar{x}_n, y_n)$ we see that $K(\bar{x}, y) - \varepsilon \leq \liminf K_n(\bar{x}_n, \bar{y}_n)$. Similarly, for any x, $K(x, \bar{y}) + \varepsilon \geq \limsup K_n(\bar{x}_n, \bar{y}_n)$. Now, since $\limsup \geq \liminf$, we have $K(x, \bar{y}) + \varepsilon \geq K(\bar{x}, y) - \varepsilon$ for all x and y. Therefore $(\bar{x}, \bar{y})$ is a 2ε-saddle point for K. By choosing $x = \bar{x}, y = \bar{y}$ we obtain

$$\limsup K_n(\bar{x}_n, \bar{y}_n) - \varepsilon \leq K(\bar{x}, \bar{y}) \leq \liminf K_n(\bar{x}_n, \bar{y}_n) + \varepsilon,$$

as desired. □

In the remainder of this section we prove our general results concerning "internal" approximations of minimax problems. By this we mean replacing the problem of finding a saddle point of K relative to $X \times Y$ by the problem of finding a saddle point relative to $X_n \times Y_n$, where $\{X_n\}$ and $\{Y_n\}$ are increasing sequences of subsets of X and Y respectively. Equivalently, we approximate K by bivariate functionals of the form

$$K_n(x, y) = \begin{cases} K(x, y) & \text{if } x \in X_n, y \in Y_n \\ -\infty & \text{if } x \in X_n, y \notin Y_n \\ +\infty & \text{if } x \notin X_n, y \in Y_n. \end{cases}$$

It doesn't matter what values we choose for $K_n(x, y)$ when $x \notin X_n$ and $y \notin Y_n$. We start with a somewhat surprising fact: internal approximations always epi/hypo-converge to *something*. This is related to the fact (see [3]) that monotone sequences of bivariate functionals are epi/hypo-convergent. Note however that if neither X_n nor Y_n are constant with respect to n, then the sequence $\{K_n(x, y)\}$ will not be monotone for any (x, y).

Proposition 2.3.8. *Consider an increasing sequence $\{X_n\}$ of subsets of X. Likewise, let $\{Y_n\}$ be an increasing sequence in Y. Suppose $K : X \times Y \to \overline{\mathbf{R}}$ and define*

$$K_n(x, y) = \begin{cases} K(x, y) & \text{if } x \in X_n, y \in Y_n \\ -\infty & \text{if } x \in X_n, y \notin Y_n \\ +\infty & \text{if } x \notin X_n. \end{cases}$$

Then K_n τ-epi/σ-hypo-converges.

Proof. Let $(\bar{x}, \bar{y})$ be a fixed pair in $X \times Y$. For any choice of $U \in \mathcal{N}_\tau(\bar{x})$, we see that $\inf_{x \in U} K_n(x, y')$ is eventually a decreasing sequence for fixed y': if $y' \in Y_k$ then $\inf_{x \in U} K_n(x, y') = \infty$ while $U \cap X_n = \emptyset$ and decreases as X_n increases after

$U \in X_n \neq \emptyset$; if $y' \notin \cup Y_k$ then

$$\inf_{x \in U} K_n(x, y') = \begin{cases} \infty & \text{if } U \cap X_n = \emptyset \\ -\infty & \text{if } U \cap X_n \neq \emptyset. \end{cases}$$

Likewise, for any choice of $V \in \mathcal{N}_\sigma(\bar{y})$, $\sup_{y \in V} K_n(x', y)$ is eventually increasing for fixed x': if $x' \in X_k$ then

$$\sup_{y \in V} K_n(x', y) = \begin{cases} \infty & \text{if } n < k \text{ and } x' \notin X_n, \\ -\infty & \text{if } n \geq k \text{ and } V \cap Y_n = \emptyset, \end{cases}$$

and $\sup_{y \in V} K_n(x', y)$ is increasing after $V \cap Y_n \neq \emptyset$. Thus $\sup_{y \in V} \inf_{x \in U} K_n(x, y)$ is eventually decreasing, whereas $\inf_{x \in U} \sup_{y \in V} K_n(x, y)$ is eventually increasing, so we may write

$$\limsup_{n \to \infty} \sup_{y \in V} \inf_{x \in U} K_n(x, y) = \lim_{n \to \infty} \sup_{y \in V} \inf_{x \in U} K_n(x, y),$$

$$\liminf_{n \to \infty} \inf_{x \in U} \sup_{y \in V} K_n(x, y) = \lim_{n \to \infty} \inf_{x \in U} \sup_{y \in V} K_n(x, y).$$

Applying $\sup \inf \leq \inf \sup$ twice we obtain first

$$\limsup_{n \to \infty} \sup_{y \in V} \inf_{x \in U} K_n(x, y) \leq \liminf_{n \to \infty} \inf_{x \in U} \sup_{y \in V} K_n(x, y)$$

and then $(\mathrm{e}_\tau / \mathrm{h}_\sigma\text{-ls}\, K_n)(\bar{x}, \bar{y}) \leq (\mathrm{h}_\sigma / \mathrm{e}_\tau\text{-li}\, K_n)(\bar{x}, \bar{y})$, as desired. □

We now consider a very special class of bivariate functions. Suppose that Φ and Ψ are extended-real-valued functions on X and Y respectively, and that $\Gamma : X \times Y \to \mathbf{R}$. Assume that Φ and Ψ are (inf-) proper, i.e. neither takes the value $-\infty$ and neither is identically $+\infty$. Suppose $\{X_n\}$ is a sequence of subsets of X and $\{Y_n\}$ is a sequence of subsets of Y. Define

$$J(x, y) = \begin{cases} \infty & \text{if } \Phi(x) = \infty, \\ \Phi(x) - \Psi(y) - \Gamma(x, y) & \text{otherwise;} \end{cases}$$

$$J_n(x, y) = \begin{cases} \infty & \text{if } \Phi_n(x) = \infty, \\ \Phi_n(x) - \Psi_n(y) - \Gamma(x, y) & \text{otherwise,} \end{cases} \tag{2.3.2}$$

where $\Phi_n = \Phi + \delta_{X_n}$ and $\Psi_n = \Psi + \delta_{Y_n}$. (The function δ_C is that which takes the value zero on the set C and the value $+\infty$ elsewhere.)

Theorem 2.3.9. *Suppose that the pair $(\bar{x}, \bar{y})$ satisfies the following hypotheses:*

(i) *Ψ is σ-lower semicontinuous at $\bar{y}$;*

(ii) *Γ is $\tau \times \sigma$-lower semicontinuous at $(\bar{x}, \bar{y})$;*

(iii) *X_n is increasing and $\Phi(\bar{x}) = (\mathrm{e}_\tau\text{-lim}\, \Phi_n)(\bar{x})$.*

Then $(\mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, J_n)(\bar{x}, \bar{y}) \leq J(\bar{x}, \bar{y})$.

Proof. If $J(\bar{x}, \bar{y}) = \infty$ we're done. Assume then that $J(\bar{x}, \bar{y}) < \infty$, so that $\Phi(\bar{x}) < \infty$. We need to show that, for any $U \in \mathcal{N}_\tau(\bar{x})$ and any $\alpha > J(\bar{x}, \bar{y})$, there exist $V \in \mathcal{N}_\sigma(\bar{y})$ and $n \in \mathbf{N}$ so that, whenever $v' \in V$ and $k \geq n$, one has $J_k(u', v') < \alpha$ for some $u' \in U$. Let $U \in \mathcal{N}_\tau(\bar{x})$ and $\alpha > J(\bar{x}, \bar{y})$ be given. Choose α_1 with $J(\bar{x}, \bar{y}) < \alpha_1 < \alpha$. Since $\Phi(\bar{x}) - \Psi(\bar{y}) - \Gamma(\bar{x}, \bar{y}) < \alpha_1$ and Ψ is σ-lower semicontinuous at $\bar{y}$, there exists $V_1 \in \mathcal{N}_\sigma(\bar{y})$ such that $-\Psi(v) < \alpha_1 - \Phi(\bar{x}) - \Gamma(\bar{x}, \bar{y})$ whenever $v \in V_1$. Also, there exist $V_2 \in \mathcal{N}_\sigma(\bar{y})$ and $U_1 \in \mathcal{N}_\tau(\bar{x})$ so that

$$-\Gamma(u, v) < -\Gamma(\bar{x}, \bar{y}) + \frac{\alpha - \alpha_1}{2}$$

whenever $u \in U_1$ and $v \in V_2$. Let $V = V_1 \cap V_2$ and $\bar{U} = U \cap U_1$. Next, since X_n is increasing, Φ_n is decreasing. Hence Φ_n τ-epi-converges to Φ at $\bar{x}$ if and only if, for each $U' \in \mathcal{N}_\tau(\bar{x})$ and $\varepsilon > 0$, there exists $u' \in U'$ with $u' \in \cup X_k$ and $\Phi(u') \leq \Phi(\bar{x}) + \varepsilon$. Take $U' = \bar{U}$ and $\varepsilon = (\alpha - \alpha_1)/2$. Finally, choose n so that $u' \in X_n$. Now, whenever $v' \in V$ and $k \geq n$, we have

$$\begin{aligned} J_k(u', v') &= \Phi_k(u') - \Psi_k(v') - \Gamma(u', v') \\ &\leq \Phi(u') - \Psi(v') - \Gamma(u', v') \\ &< [\frac{\alpha - \alpha_1}{2} + \Phi(\bar{x})] + [\alpha_1 - \Phi(\bar{x}) + \Gamma(\bar{x}, \bar{y})] + [\frac{\alpha - \alpha_1}{2} - \Gamma(\bar{x}, \bar{y})] \\ &= \alpha. \end{aligned}$$

□

Corollary 2.3.10. *Let J and J_n be defined as in (2.3.2), with Φ, Ψ, Γ, X_n and Y_n satisfying the following hypotheses:*

(i) *Φ and Ψ are (inf-) proper functions;*

(ii) *Φ is τ-lower semicontinuous and Ψ is σ-lower semicontinuous;*

(iii) *Γ is $\tau \times \sigma$-continuous;*

(iv) *$\{X_n\}$ and $\{Y_n\}$ are increasing with $\Phi = \mathrm{e}_\tau\text{-lim}\, \Phi_n$ and $\Psi = \mathrm{e}_\sigma\text{-lim}\, \Psi_n$.*

Then J is a closed proper bivariate function with

$$J = \overline{J} \geq \mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, J_n \text{ and } \underline{J} \leq \mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, J_n.$$

Moreover, in the case that $\operatorname{dom}\Phi$ *is closed, one actually has* $J \leq \mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, J_n$, *so that* J_n *epi/hypo-converges to* J.

Proof. Clearly, J is proper. We see that

$$\operatorname{hypo} J(x,\cdot) = \begin{cases} Y \times \mathbf{R} & \text{if } \Phi(x) = \infty, \\ (0,\Phi(x)) - \operatorname{epi}(\Psi + \Gamma(x,\cdot)) & \text{otherwise.} \end{cases}$$

Hence $\operatorname{hypo} J(x,\cdot)$ is closed for all x, so $J = \mathrm{cl}_2\, J$. Since $J(x,y) = \infty$ if and only if $\Phi(x) = \infty$, we must have $J = \overline{\mathrm{cl}}_2\, J$. Next, we have

$$\operatorname{epi} J(\cdot,y) = \begin{cases} \operatorname{epi}(\Phi - \Gamma(\cdot,y)) - (0,\Psi(y)) & \text{if } \Psi(y) < \infty, \\ (\operatorname{dom}\Phi) \times \mathbf{R} & \text{if } \Psi(y) = \infty, \end{cases}$$

so that

$$\mathrm{cl}(\operatorname{epi} J(\cdot,y)) = \begin{cases} \operatorname{epi}(\Phi - \Gamma(\cdot,y)) - (0,\Psi(y)) & \text{if } \Psi(y) < \infty, \\ \mathrm{cl}(\operatorname{dom}\Phi) \times \mathbf{R} & \text{if } \Psi(y) = \infty. \end{cases}$$

Hence

$$\mathrm{cl}_1\, J(x,y) = \begin{cases} \Phi(x) - \Gamma(x,y) - \Psi(y) & \text{if } \Psi(y) < \infty, \\ -\infty & \text{if } \Psi(y) = \infty \text{ and } x \in \mathrm{cl}(\operatorname{dom}\Phi), \\ \infty & \text{if } x \notin \mathrm{cl}(\operatorname{dom}\Phi), \end{cases}$$

so that

$$\underline{\mathrm{cl}}_1\, J(x,y) = \begin{cases} -\infty & \text{if } \Psi(y) = \infty, \\ \Phi(x) - \Psi(y) - \Gamma(x,y) & \text{otherwise.} \end{cases}$$

By symmetry, we see that $\overline{\mathrm{cl}}_2(\underline{\mathrm{cl}}_1\, J) = \overline{\mathrm{cl}}_2\, J = J$ and thus J is closed.

By Theorem 2.3.9, we have $J \geq \mathrm{e}_\tau/\mathrm{h}_\sigma\text{-ls}\, J_n$ and $\underline{J} \leq \mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, J_n$.

Finally, assume that $\operatorname{dom}\Phi$ is closed. Fix $\bar{x} \in X$ and $\bar{y} \in Y$. If $\Phi(\bar{x}) < \infty$ then $J(\bar{x},\bar{y}) = \underline{J}(\bar{x},\bar{y})$. Suppose then that $\Phi(\bar{x}) = \infty$. Then $U = X \setminus \operatorname{dom}\Phi$ is a τ-neighborhood of $\bar{x}$, so $J_n(u,v) = \infty$ for all $n \in \mathbf{N}$, $u \in U$, and $v \in Y$. Thus $(\mathrm{h}_\sigma/\mathrm{e}_\tau\text{-li}\, J_n)(\bar{x},\bar{y}) = \infty = J(\bar{x},\bar{y})$, which completes the proof. □

Consider the above theorem and corollary in the situation where X and Y are normed linear spaces and Γ is a biaffine functional which is separately norm-continuous. Then the hypothesis that Γ be $\tau \times \sigma$-continuous is satisfied if τ and σ are the respective norm topologies. In many circumstances, however, it is necessary to take σ to be the *weak* topology on Y. In this case, Γ will necessarily fail to be jointly continuous, unless Y is finite-dimensional. On the other hand, it may well be that Γ is *sequentially* $\tau \times \sigma$-continuous. It is therefore important to have a sequential version of Theorem 2.3.9.

Theorem 2.3.11. *Suppose that the pair $(\bar{x}, \bar{y})$ satisfies the following hypotheses:*

(i) *Ψ is sequentially σ-lower semicontinuous at $\bar{y}$;*

(ii) *Γ is sequentially $\tau \times \sigma$-lower semicontinuous at $(\bar{x}, \bar{y})$;*

(iii) *X_n is increasing and $\Phi(\bar{x}) = (\text{seq-e}_\tau\text{-lim}\, \Phi_n)(\bar{x})$.*

Then $(\text{seq-e}_\tau/\text{h}_\sigma\text{-ls}\, J_n)(\bar{x}, \bar{y}) \leq J(\bar{x}, \bar{y})$.

Proof. If $J(\bar{x}, \bar{y}) = \infty$, we're done. Assume then that $J(\bar{x}, \bar{y}) < \infty$, so that $\Phi(\bar{x}) < \infty$. Suppose $\{y_n\}$ σ-converges to $\bar{y}$, and fix $\alpha > J(\bar{x}, \bar{y})$. We need to find a sequence $\{x_n\}$, τ-converging to $\bar{x}$, such that $\limsup J_n(x_n, y_n) \leq \alpha$. Choose $\alpha_1 \in (J(\bar{x}, \bar{y}), \alpha)$. By hypothesis (iii), there is a sequence $\{x_n\}$ converging to $\bar{x}$, with $x_n \in X_n$, such that

$$\limsup \Phi(x_n) \leq \Phi(\bar{x}) + \frac{\alpha - \alpha_1}{2}.$$

This, along with the sequential lower semicontinuity hypotheses on Ψ and Γ, allows us to find a positive integer N so that the following three inequalities hold whenever $n \geq N$:

$$\begin{aligned} -\Psi(y_n) &< \alpha_1 - \Phi(\bar{x}) + \Gamma(\bar{x}, \bar{y}), \\ -\Gamma(x_n, y_n) &< -\Gamma(\bar{x}, \bar{y}) + \frac{\alpha - \alpha_1}{2}, \\ \Phi(x_n) &\leq \Phi(\bar{x}) + \frac{\alpha - \alpha_1}{2}. \end{aligned}$$

Thus, for all $n \geq N$, we have

$$\begin{aligned} J(x_n, y_n) &= \Phi_n(x_n) - \Psi_n(y_n) - \Gamma(x_n, y_n) \\ &\leq \Phi(x_n) - \Psi(y_n) - \Gamma(x_n, y_n) \\ &< [\Phi(\bar{x}) + \frac{\alpha - \alpha_1}{2}] + [\alpha_1 - \Phi(\bar{x}) + \Gamma(\bar{x}, \bar{y})] + [\frac{\alpha - \alpha_1}{2} - \Gamma(\bar{x}, \bar{y})] \\ &= \alpha, \end{aligned}$$

as desired. □

In the next section, we strengthen this result for the case of saddle functions on reflexive Banach spaces.

4. Mosco Convergence of Saddle Functions

In this section we discuss the notion of Mosco epi/hypo-convergence, which is a generalization of Mosco epi- and hypo-convergence. Throughout the section, X

and Y will be reflexive Banach spaces with respective duals X^* and Y^*. Note that a saddle function K on $X \times Y$ is closed (in the sense of §2.2) with respect to the strong (norm) topologies on X and Y if and only if it is closed with respect to the weak topologies. In what follows, s denotes the strong topology and w denotes the weak topology; no notational distinction is made to indicate which space these refer to: the context will make it unambiguous.

Let K_n and K be closed proper saddle functions on $X \times Y$. The sequence K_n is said to *Mosco epi/hypo-converge* to K if the following three conditions are satisfied:

(a) The sequence K_n is upper modulated, i.e. there exists a weakly convergent sequence $\{x_n\}$ in X and a real number $r \geq 0$ such that $K_n(x_n, y) \leq r(\|y\|+1)$ for all $y \in Y$ and all $n \in \mathbf{N}$.

(b) $\underline{K} \leq \text{seq-h}_s/\text{e}_w\text{-li}\, \underline{K}_n$.

(c) $\overline{K} \geq \text{seq-e}_s/\text{h}_w\text{-ls}\, \overline{K}_n$.

We write $K = \text{M-e/h-lim}\, K_n$. Note that this is actually a convergence of classes. (In the paper by Attouch, Aze and Wets [1] this form of convergence is called "Mosco epi/hypo-convergence in the *extended sense*".) The following theorem explains the use of the name "Mosco" in this situation. Its proof is given by Do [10].

Theorem 2.4.1[10]. *Suppose K_n and K are closed, proper saddle functions and let F_n, F and G_n, G be their respective convex and concave parents. The following are equivalent:*

(i) *K_n Mosco epi/hypo-converges to K.*

(ii) *F_n Mosco epi-converges to F.*

(iii) *G_n Mosco hypo-converges to G.*

Corollary 2.4.2[10]. *Let K_n and K be closed, proper saddle functions, and suppose that $J_n, L_n \in [\underline{K}_n, \overline{K}_n]$. Then the following are equivalent:*

(i) *K_n Mosco epi/hypo-converges to K.*

(ii) *The sequence K_n is upper modulated and*

$$\underline{K} \leq \text{seq-h}_s/\text{e}_w\text{-li}\, J_n, \tag{2.4.1}$$

$$\overline{K} \geq \text{seq-e}_s/\text{h}_w\text{-ls}\, L_n. \tag{2.4.2}$$

(iii) *The sequence K_n is lower modulated, i.e. there exists a weakly convergent sequence $\{y_n\}$ in X and a real number $r \geq 0$ such that $K_n(x, y_n) \geq -r(\|x\|+1)$ for all $x \in X, n \in \mathbf{N}$, and the inequalities (2.4.1) and (2.4.2) hold.*

The results of section 2.3 concerning convergence of saddle points and saddle values extend to Mosco epi/hypo-convergence, as illustrated by the following proposition.

Proposition 2.4.3. *Suppose that* $K = \text{M-e/h-lim}\, K_n$ *and that* $(\bar{x}_n, \bar{y}_n)$ *is a saddle point for* K_n. *If* $\bar{x}_n \overset{w}{\to} \bar{x}$ *and* $\bar{y}_n \overset{w}{\to} \bar{y}$, *then* $(\bar{x}, \bar{y})$ *is a saddle point for* K *and the saddle value for* K *is*

$$K(\bar{x}, \bar{y}) = \lim_{n\to\infty} K_n(\bar{x}_n, \bar{y}_n).$$

Proof. This is just Theorem 2.3.6(b) using the weak topologies on X and Y for τ_1 and σ_1, and the strong topologies for τ_2 and σ_2. □

Before turning to internal approximations in the Mosco framework, we first prove a general convergence result for saddle functions of the form

$$(x, y) \mapsto \Phi(x) - \Psi(y) - \Gamma(x, y).$$

This is an extension of a similar result proved in Do's thesis [10].

Proposition 2.4.4. *Let* $\Gamma : X \times Y \to \mathbf{R}$ *be a continuous biaffine map. Suppose* $\Phi_n, \Phi : X \to \overline{\mathbf{R}}$ *and* $\Psi_n, \Psi : Y \to \overline{\mathbf{R}}$ *are proper, lower semicontinuous convex functions such that* $\Phi = \text{M-e-lim}\, \Phi_n$ *and* $\Psi = \text{M-e-lim}\, \Psi_n$. *Let* J_n *and* J *be any bivariate functions on* $X \times Y$ *for which*

$J_n(x, y) = \Phi_n(x) - \Psi_n(y) - \Gamma(x, y)$ whenever $\infty - \infty$ does not occur, and

$J(x, y) = \Phi(x) - \Psi(y) - \Gamma(x, y)$ whenever $\infty - \infty$ does not occur.

Then J_n *and* J *are equivalent to closed proper saddle functions and* J_n *Mosco epi/hypo-converges to* J.

Proof. It easy to see that $\overline{\text{cl}}_2\, J$ and $\underline{\text{cl}}_1\, J$ are given by

$$\overline{\text{cl}}_2\, J(x, y) = \begin{cases} \infty & \text{if } \Phi(x) = \infty, \\ \Phi(x) - \Psi(y) - \Gamma(x, y) & \text{otherwise,} \end{cases}$$

$$\underline{\text{cl}}_1\, J(x, y) = \begin{cases} -\infty & \text{if } \Psi(y) = \infty, \\ \Phi(x) - \Psi(y) - \Gamma(x, y) & \text{otherwise.} \end{cases}$$

It is clear from these descriptions that $\overline{\text{cl}}_2\, J$ and $\underline{\text{cl}}_1\, J$ are both saddle functions, and that $\underline{\text{cl}}_1(\overline{\text{cl}}_2\, J) = \underline{\text{cl}}_1\, J$ and $\overline{\text{cl}}_2(\underline{\text{cl}}_1\, J) = \overline{\text{cl}}_2\, J$. By the same argument, each J_n

is equivalent to a closed saddle function. Since Γ is continuous and biaffine, there exist a continuous linear map $D : X \to Y^*$, elements $b \in X^*$ and $a \in Y^*$, and a real number c such that

$$\Gamma(x,y) = \langle Dx, y\rangle + \langle a, y\rangle + \langle b, x\rangle + c.$$

Thus the convex parent of J_n can be written as

$$\begin{aligned} F_n(x,y^*) &= \Phi_n(x) - \langle b,x\rangle - c + \sup_y\{-\Psi_n(y) - \langle Dx+a, y\rangle - \langle y^*, y\rangle\} \\ &= \Phi_n(x) - \langle b,x\rangle - c + \Psi_n^*(-Dx - a - y^*). \end{aligned}$$

The convex parent F of J has the same form, but without the n subscripts. We shall show that $F = \text{M-e-lim}\, F_n$. Let (x,y^*) be given. There exists a sequence $\{x_n\}$ strongly converging to x with $\limsup \Phi_n(x_n) \le \Phi(x)$, since $\Phi = \text{M-e-lim}\,\Phi_n$. Similarly, we can find a sequence $\{z_n\}$ in Y^* converging strongly to $z = -Dx - a - y^*$ with $\limsup \Psi_n^*(z_n) \le \Psi^*(z)$, since $\Psi = \text{M-e-lim}\,\Psi_n$ implies $\Psi^* = \text{M-e-lim}\,\Psi_n^*$. Taking $y_n^* = -z_n - Dx_n - a$, we see that $\limsup F_n(x_n, y_n^*) \le F(x,y^*)$. Hence $e_S\text{-ls}\, F_n \le F$. Now suppose that $\{(x_k, y_k^*)\}$ converges weakly to (x,y^*). Then, for any subsequence $\{\Phi_{n_k}\}$, we have $\limsup \Phi_{n_k}(x_k) \ge \Phi(x)$, by Mosco epi-convergence of the Φ_n. Similarly, since $\Psi = \text{M-e-lim}\,\Psi_n$ implies $\Psi^* = \text{M-e-lim}\,\Psi_n^*$, we have $\limsup \Psi_{n_k}(-y_k^* - a - Dx_k) \ge \Psi(-y^* - a - Dx)$ so $\liminf F_{n_k}(x_k, y_k^*) \ge F(x,y^*)$, i.e. $\text{seq-e}_W\text{-li}\, F_n \ge F$. □

In order to apply the above proposition to internal approximations for saddle functions, we need some general criteria for the epi-convergence of internal approximations for univariate functions. This is furnished by the next result.

Proposition 2.4.5. *Suppose that Φ is a proper, lower semicontinuous convex function on X, and that $C = \text{M-lim}\, C_n$ where each C_n is a nonemtpy closed convex set in X. Then $\Phi + \delta_{C_n}$ Mosco epi-converges to $\Phi + \delta_C$ if and only if for every $x \in C$ there exist $x_n \overset{s}{\to} x$ such that $x_n \in C_n$ and $\limsup \Phi(x_n) \le \Phi(x)$.*

Proof. Define $\Phi_n = \Phi + \delta_{C_n}$. Given x, we shall consider the following three statements:

(*) There exist $x_n \overset{s}{\to} x$ with $x_n \in C_n$ and $\limsup \Phi(x_n) \le \Phi(x)$.

(i) There exist $x_n \overset{s}{\to} x$ with $\limsup \Phi_n(x_n) \le \Phi(x)$.

(ii) If $x_k \overset{w}{\to} x$, then $\liminf \Phi_{n_k}(x_k) \geq \Phi(x)$ for any subsequence $\{\Phi_{n_k}\}$ of $\{\Phi_n\}$. Clearly, $\Phi = \text{M-e-lim}\,\Phi_n$ if and only if (i) and (ii) hold for every x. First we claim that the hypotheses that Φ is closed and that $C = \text{M-lim}\,C_n$ imply condition (ii). Suppose $x_n \overset{w}{\to} x$ and let $\{\Phi_{n_k}\}$ be a subsequence of $\{\Phi_n\}$. If $x \notin C$ then there exists k' such that $x_k \notin C_{n_k}$ whenever $k \geq k'$ (otherwise, there would be $x_{km} \overset{w}{\to} x$ with $x_{km} \in C_{n_{km}}$, so $x \in C = \text{M-lim}\,C_n$). Thus $\Phi_{n_k}(x_k) = \infty$ for all $k \geq k'$, so (ii) holds. If $x \in C$ then we need only look at the subsequence $\{x_{km}\}$ where $x_{km} \in C_{n_{km}}$ (if no such subsequence exists then $\liminf \Phi_{n_k}(x_k) = \infty$, and (ii) holds trivially). Thus we have

$$\liminf \Phi_{n_k}(x_k) \geq \liminf \Phi_{n_{km}}(x_{km}) = \liminf \Phi(x_{km}) \geq \Phi(x),$$

by the lower semicontinuity of Φ.

Notice that (i) is trivially satisfied if $x \notin C$. Since the hypotheses imply (ii), it remains to show that (i) and (*) are equivalent when $x \in C$. Clearly, (*) implies (i). Suppose (i) is satisfied at $x \in C$. Then there exist $\bar{x}_n \overset{s}{\to} x$ with $x_n \in C_n$. If $\Phi(x) < \infty$, then only finitely many of the $\bar{x}_n \notin C_n$, so we may replace these by elements in the C_n. If we denote the new sequence by $\{x_n\}$, then we see that (*) holds. □

The basic result on internal approximations for saddle functions of the form $(x, y) \mapsto \Phi(x) - \Psi(y) - \Gamma(x, y)$ may now be stated as a corollary of the above two propositions. It can be seen as an extension of Theorem 2.3.11 to Mosco epi/hypo-convergence, and represents the underlying theme of the proofs of the main approximation theorems in Chapters 3 and 4.

Corollary 2.4.6. *Suppose that $\Phi : X \to \overline{\mathbf{R}}$ and $\Psi : Y \to \overline{\mathbf{R}}$ are proper, lower semicontinuous convex functions and that $\Gamma : X \times Y \to \mathbf{R}$ is a continuous biaffine functional. Suppose $\{X_n\}$ and $\{Y_n\}$ are increasing sequences of closed convex sets in X and Y respectively, which satisfy*

(i) *For every $x \in X$ there exists a sequence $\{x_n\}$ converging strongly to x for which $x_n \in X_n$ and $\limsup \Phi(x_n) \leq \Phi(x)$;*

(ii) *For every $y \in Y$ there exists a sequence $\{y_n\}$ converging strongly to y for which $y_n \in Y_n$ and $\limsup \Psi(y_n) \leq \Psi(y)$.*

If we define

$$J(x,y) = \begin{cases} \infty & \text{if } \Phi(x) = \infty, \\ \Phi(x) - \Psi(y) - \Gamma(x,y) & \text{otherwise;} \end{cases}$$

$$J_n(x,y) = \begin{cases} J(x,y) & \text{if } x \in X_n, y \in Y_n, \\ -\infty & \text{if } x \in X_n, y \notin Y_n, \\ \infty & \text{if } x \notin X_n, \end{cases}$$

then J and J_n are closed proper saddle functions and $J = \text{M-e/h-lim}\, J_n$.

Proof. Define $\Phi_n = \Phi + \delta_{X_n}$ and $\Psi_n = \Psi + \delta_{Y_n}$. It is easy to show that an increasing sequence of convex sets Mosco converges to the closure of its union. Since (i) implies that $X \subset s\text{-liminf}\, X_n$, we see that $X = \text{M-lim}\, X_n$. Similarly, (ii) implies that $Y = \text{M-lim}\, Y_n$. In view of Proposition 2.4.5, $\Phi = \text{M-e-lim}\, \Phi_n$ and $\Psi = \text{M-e-lim}\, \Psi_n$, so Proposition 2.4.4 tells us that J and J_n are closed proper saddle functions and $J = \text{M-e/h-lim}\, J_n$. □

Chapter 3. Applications in Optimal Control

1. Linear-quadratic Models

In this section we present a model in (continuous-time) optimal control, list some duality results, and describe an approximation scheme which gives rise to a discrete-time optimal control problem. The development of the duality theory itself is due to Rockafellar [27], [28]. The problem we consider has a *piecewise linear-quadratic* formulation. The extension to models with nonquadratic objective functions is treated in the next section. The problem we shall be working with is the following:
$(\mathcal{P}^1)$

minimize the functional

$$F(u,u_e) = \int_{t_0}^{t_1}[p_t \cdot u_t + \tfrac{1}{2}u_t \cdot P_t u_t - c_t \cdot x_t + \rho_{V_t,Q_t}(q_t - D_t u_t - C_t x_t)]dt$$
$$+ p_e \cdot u_e + \tfrac{1}{2}u_e \cdot P_e u_e - c_e \cdot x_{t_1} + \rho_{V_e,Q_e}(q_e - D_e u_e - C_e x_{t_1})$$

over the control space

$$\mathcal{U}^1 = \{(u,u_e) \in \mathcal{L}_k^1 \times \mathbf{R}^{ke} | u_t \in U_t \text{ a.e., } u_e \in U_e\}$$

with the dynamics given by

$$\dot{x}_t = A_t x_t + B_t u_t + b_t \text{ a.e.}, \quad x_{t_0} = B_e u_e + b_e.$$

The dual problem to $(\mathcal{P}^1)$ will be
$(\mathcal{Q}^1)$

maximize the functional

$$G(v,v_e) = \int_{t_0}^{t_1}[q_t \cdot v_t - \tfrac{1}{2}v_t \cdot Q_t v_t - b_t \cdot y_t - \rho_{U_t,P_t}(D_t^* v_t + B_t^* y_t - p_t)]dt$$
$$+ q_e \cdot v_e - \tfrac{1}{2}v_e \cdot Q_e v_e - b_e \cdot y_{t_0} - \rho_{U_e,P_e}(D_e^* v_e + B_e y_{t_0} - p_e)$$

over the control space

$$\mathcal{V}^1 = \{(v,v_e) \in \mathcal{L}_l^1 \times \mathbf{R}^{le} | v_t \in V_t \text{ a.e., } v_e \in V_e\}$$

with the dynamics given by

$$-\dot{y}_t = A_t^* y_t + C_t^* v_t + c_t \text{ a.e.}, \quad y_{t_1} = C_e^* v_e + c_e.$$

We assume that the data elements $p_t, q_t, P_t, Q_t, U_t, V_t, A_t, B_t, b_t, C_t, c_t, D_t$ are all continuous with respect to $t \in [t_0, t_1]$. In addition, the matrices P_t, P_e, Q_t and Q_e are symmetric, positive semidefinite, and the sets V_t, V_e, U_t and U_e are assumed to be polyhedral convex.

In addition to the problems $(\mathcal{P}^1)$ and $(\mathcal{Q}^1)$, we shall consider the problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^r)$ (for $r \in [1,\infty]$) given by replacing the spaces $\mathcal{U}^1$ and $\mathcal{V}^1$ by

$$\mathcal{U}^r = \{(u,u_e) \in \mathcal{L}_k^r \times \mathbf{R}^{ke} | u_t \in U_t \text{ a.e., } u_e \in U_e\}$$

and

$$\mathcal{V}^r = \{(v,v_e) \in \mathcal{L}_l^r \times \mathbf{R}^{le} | v_t \in V_t \text{ a.e., } v_e \in V_e\}.$$

The expression $\rho_{V_t,Q_t}(s)$ represents a piecewise linear-quadratic function on $\mathbf{R}^l$, given explicitly by

$$\rho_{V_t,Q_t}(s) = \sup_{v_t \in V_t} \{s \cdot v_t - \frac{1}{2} v_t \cdot Q_t v_t\}.$$

Its effective domain

$$L_t = \{s \in \mathbf{R}^l | \rho_{V_t,Q_t}(s) < \infty\}$$

is a nonempty convex polyhedron that can be decomposed into finitely many polyhedral convex sets, on each of which ρ_{V_t,Q_t} is quadratic or linear.

Proposition 3.1.1 [27]. *The map $(t,s) \mapsto \rho_{V_t,Q_t}(s)$ is lower semicontinuous jointly in t and s. The effective domain*

$$L_t = \{s \in \mathbf{R}^l | \rho_{V_t,Q_t}(s) < \infty\}$$

depends lower semicontinuously on t.

The same is true for $(t,r) \mapsto \rho_{U_t,P_t}(r)$ and its effective domain

$$K_t = \{r \in \mathbf{R}^k | \rho_{U_t,P_t}(r) < \infty\}.$$

The duality theory for these problems has been developed by Rockafellar [27]. Theorems 3.1.2 through 3.1.6 give the main results concerning this duality. We state these without proof: analogous results will be obtained for the stochastic problems in Chapter 4, and the proofs given there are similar to those needed here.

Theorem 3.1.2 [27]. *For each $r \in [1,\infty]$, $\mathcal{U}^r$ is a (nonempty) closed convex subset of $\mathcal{L}_k^r \times \mathbf{R}^{ke}$ and F is well defined, lower semicontinuous and convex on $\mathcal{L}_k^r \times \mathbf{R}^{ke}$, with values that are finite or ∞.*

Similarly, $\mathcal{V}^r$ is a (nonempty) closed convex subset of $\mathcal{L}_l^r \times \mathbf{R}^{le}$ and G is well defined, upper semicontinuous and concave, with values that are finite or $-\infty$.

The duality between $(\mathcal{P})$ and $(\mathcal{Q})$ may be demonstrated through the introduction of an appropriate bivariate function. We define this by

$$\mathcal{J}(u, u_e; v, v_e) = \int_{t_0}^{t_1} J_t(u_t, v_t)dt + J_e(u_e, v_e) - \gamma(u, u_e; v, v_e)$$

where

$$J_t(u_t, v_t) = p_t \cdot u_t + \tfrac{1}{2} u_t \cdot P_t u_t - v_t \cdot D_t u_t + q_t \cdot v_t - \tfrac{1}{2} v_t \cdot Q_t v_t,$$
$$J_e(u_e, v_e) = p_e \cdot u_e + \tfrac{1}{2} u_e \cdot P_e u_e - v_e \cdot D_e u_e + q_e \cdot v_e - \tfrac{1}{2} v_e \cdot Q_e v_e,$$

and

$$\begin{aligned}\gamma(u, u_e; v, v_e) &= \int_{t_0}^{t_1} y_t \cdot (B_t u_t + b_t)dt + y_{t_0} \cdot (B_e u_e + b_e) \\ &= \int_{t_0}^{t_1} x_t \cdot (C_t^* v_t + c_t)dt + x_{t_1} \cdot (C_e^* v_e + c_e).\end{aligned}$$

In evaluating the integrals defining $\mathcal{J}$ we use the convention that $\infty - \infty = \infty$: this is the same as requiring $\mathcal{J}(u, v) = \infty$ if and only if $J_t(u_t, v_t)$ is not majorized by any integrable function.

Theorem 3.1.3 [27]. *Consider $r, r' \in [1, \infty]$. The problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^{r'})$ are the primal and dual problems associated with the problem of finding a saddle point of $\mathcal{J}$ on $\mathcal{U}^r \times \mathcal{V}^{r'}$. In fact, we have*

$$F(u, u_e) = \sup_{(v, v_e) \in \mathcal{V}^{r'}} \mathcal{J}(u, u_e; v, v_e) = \sup_{(v, v_e) \in \mathcal{V}^{\infty}} \mathcal{J}(u, u_e; v, v_e)$$

and

$$G(v, v_e) = \inf_{(u, u_e) \in \mathcal{U}^r} \mathcal{J}(u, u_e; v, v_e) = \inf_{(u, u_e) \in \mathcal{U}^{\infty}} \mathcal{J}(u, u_e; v, v_e).$$

The following is a direct consequence of Theorem 3.1.3. (In this dissertation, we write "min" to indicate an infimum which is actually attained; similarly, "max" is used to denote a supremum that is attained.)

Proposition 3.1.4 (Weak Duality). *For $r, r' \in [1, \infty]$ with $r \leq r'$, it is always true that* $\inf(\mathcal{P}^{r'}) \geq \inf(\mathcal{P}^r) \geq \sup(\mathcal{Q}^r) \geq \sup(\mathcal{Q}^{r'})$. *Moreover, for any $s \in [1, \infty]$, $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ is a saddle point of $\mathcal{J}$ over $\mathcal{U}^r \times \mathcal{V}^s$ if and only if $(\bar{u}, \bar{u}_e)$ solves $(\mathcal{P}^r)$, $(\bar{v}, \bar{v}_e)$ solves $(\mathcal{Q}^s)$ and*

$$\min(\mathcal{P}^r) = \max(\mathcal{Q}^s) \quad (\textit{finite}).$$

In order to obtain strong duality results, i.e. the existence of saddle points, we need to impose some sort of "finiteness" conditions. We say that the *primal finiteness condition* is satisfied if $\rho_{V_t,Q_t}(\cdot)$ and $\rho_{V_e,Q_e}(\cdot)$ are finite everywhere. Similarly, the *dual finiteness condition* is satisfied if $\rho_{U_t,P_t}(\cdot)$ and $\rho_{U_e,P_e}(\cdot)$ are finite everywhere. These conditions are satisfied in the case where U_t, U_e, V_t, V_e are all bounded. They would also be satisfied in the case where the matrices P_t, P_e, Q_t, Q_e are all positive definite.

Theorem 3.1.5 (Strong Duality) [27],[28]. *If the primal finiteness condition is satisfied, then*

$$\inf(\mathcal{P}^1) = \max(\mathcal{Q}^1) < \infty;$$

likewise, if the dual finiteness condition is satisfied, then

$$\min(\mathcal{P}^1) = \sup(\mathcal{Q}^1) > -\infty.$$

If both conditions are satisfied, then, for all $r, r' \in [1, \infty]$, solutions exist to both $(\mathcal{P}^r)$ and $(\mathcal{Q}^{r'})$, and

$$\min(\mathcal{P}^r) = \max(\mathcal{Q}^{r'}) \quad \textit{finite}.$$

Furthermore, when both finiteness conditions hold, every optimal solution $(\bar{u}, \bar{u}_e)$ of $(\mathcal{P}^r)$ in fact has $\bar{u} \in \mathcal{L}^\infty$, whereas each optimal solution of $(\bar{v}, \bar{v}_e)$ of $(\mathcal{Q}^{r'})$ has $\bar{v} \in \mathcal{L}^\infty$.

The following result says that the saddle point condition actually "decomposes" with respect to time.

Theorem 3.1.6 (Minimaximum Principle) [27]. *Consider $r, s \in [1, \infty]$. For the control pair $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ to be a saddle point for $\mathcal{J}$ on $\mathcal{U}^r \times \mathcal{V}^s$, it is necessary and sufficient for the following three conditions to hold:*

(i) $\bar{u}_t$ *and* $\bar{v}_t$ *are in* $\mathcal{L}_k^r$ *and* $\mathcal{L}_l^s$ *respectively (with trajectories* $\bar{x}_t$ *and* $\bar{y}_t$*).*

(ii) $(\bar{u}_t, \bar{v}_t)$ *is a saddle point over* $U_t \times V_t$ *for*

$$J_t(u_t, v_t) - u_t \cdot B_t^* \bar{y}_t - v_t \cdot C_t \bar{x}_t.$$

(iii) $(\bar{u}_e, \bar{v}_e)$ *is a saddle point over* $U_e \times V_e$ *for*

$$J_e(u_e, v_e) - u_e \cdot B_e^* \bar{y}_{t_0} - v_e \cdot C_e \bar{x}_{t_1}.$$

We now introduce an approximation scheme for finding a saddle point of $\mathcal{J}$ on $\mathcal{U}^r \times \mathcal{V}^{r'}$. The idea is to approximate the controls by feasible *step functions* on $[t_0, t_1]$. Clearly, it is possible that either $\mathcal{U}^r$ or $\mathcal{V}^{r'}$ may not contain any step functions at all. For this reason we shall make the assumption that the multifunctions U_t and V_t are *constant*. In what follows, this assumption could actually be relaxed somewhat. For example, we could just as well assume that $U_t = W_t \bar{U} + w_t$ and $V_t = Z_t \bar{V} + z_t$, where W_t, w_t, Z_t, z_t are continuous with respect to t, and the sets $\bar{U}$ and $\bar{V}$ are fixed convex polyhedrons. It can also be shown that a similar procedure can be applied if the graphs of U_t and V_t are convex in $\mathbf{R}^k \times \mathbf{R}$ and $\mathbf{R}^l \times \mathbf{R}$.

Let $\pi = (t_0 = a_0 < a_1 < \ldots < a_T = t_1)$ be a partition of the interval $[t_0, t_1]$. Consider the following subsets of $\mathcal{U}^\infty$ and $\mathcal{V}^\infty$:

$$\mathcal{U}_\pi = \{(u, u_e) \in \mathcal{U}^\infty | u_t \text{ is constant a.e. on } [a_{\tau-1}, a_\tau] \text{ for } \tau = 0, \ldots, T\},$$
$$\mathcal{V}_\pi = \{(v, v_e) \in \mathcal{V}^\infty | v_t \text{ is constant a.e. on } [a_{\tau-1}, a_\tau] \text{ for } \tau = 0, \ldots, T\}.$$

The approximate problem is

$(\mathcal{S}_\pi)$ find a saddle point of $\mathcal{J}$ relative to $\mathcal{U}_\pi \times \mathcal{V}_\pi$.

This is clearly a finite-dimensional problem. As we shall see below, this problem can be reformulated as a problem of discrete-time optimal control. But first we show that this is in some sense an approximation for the original problem. Our main theorem for this section is the following.

Theorem 3.1.7. *Let* $\{\pi_\nu : \nu \in \mathrm{N}\}$ *be an increasing sequence of partitions of* $[t_0, t_1]$ *with* $|a_i - a_{i+1}| \to 0$ *uniformly in* i *as* $\nu \to 0$. *Define*

$$\mathcal{J}_\nu(u, u_e; v, v_e) = \begin{cases} \mathcal{J}(u, u_e; v, v_e) & \text{if } (u, u_e) \in \mathcal{U}_{\pi_\nu}, (v, v_e) \in \mathcal{V}_{\pi_\nu}, \\ -\infty & \text{if } (u, u_e) \in \mathcal{U}_{\pi_\nu}, (v, v_e) \notin \mathcal{V}_{\pi_\nu}, \\ \infty & \text{if } (u, u_e) \notin \mathcal{U}_{\pi_\nu}. \end{cases}$$

For any $r, r' \in [2, \infty)$, the approximating sequence $\mathcal{J}_\nu$ Mosco-epi/hypo-converges to $\mathcal{J}$ over $(\mathcal{L}_k^r \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^{r'} \times \mathbf{R}^{le})$.

In the statement of this theorem, $\mathcal{J}$ refers to the same function as before, except that we now extend it to $(\mathcal{L}_k^1 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^1 \times \mathbf{R}^{le})$ by defining it to be $-\infty$ if $(u, u_e) \in \mathcal{U}^1$ but $(v, v_e) \notin \mathcal{V}^1$, and to be ∞ if $(u, u_e) \notin \mathcal{U}^1$. Clearly this gives rise to the same minimax problem, but allows us to view it in terms of the concepts in Chapter 2. Note that under the primal and dual finiteness conditions, all saddle points are essentially bounded (see Theorem 3.1.5) and so the restriction to $\mathcal{L}^2$ in the above theorem is not quite so restrictive as it may appear. To prove Theorem 3.1.7, we will use a couple of basic facts about constrained approximation of measurable functions by simple functions and step functions. Since these facts will be used in several different contexts within this paper, we state them explicitly as lemmas. The proofs are typical of those in elementary measure theory, so we omit them.

Lemma 3.1.8. *Suppose $1 \leq p < \infty$. Consider $f \in \mathcal{L}_d^p(\Omega, \mathcal{A}, \mu)$ and a closed convex set $M \subset \mathbf{R}^d$. If $f(\omega) \in M$ a.e. $[\mu]$, then there is a sequence $\{f_\nu\}$ of simple integrable functions satisfying:*

(i) *$f_\nu(\omega) \in M$ for all $\omega \in \Omega$.*
(ii) *$|f_\nu(\omega)| \leq |f(\omega)|$ a.e. $[\mu]$.*
(iii) *$f_\nu \to f$ a.e. $[\mu]$.*
(iv) *$f_\nu \to f$ in $\mathcal{L}_d^p(\Omega, \mathcal{A}, \mu)$.*

Lemma 3.1.9. *Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space with $\mathcal{A} = \sigma(\mathcal{F})$, where $\mathcal{F}$ is a field on Ω. Let $p \in [1, \infty)$. Suppose that $\varphi : \mathbf{R}^d \to \overline{\mathbf{R}}$. If $f : \Omega \to \mathbf{R}^d$ is an $\mathcal{A}$-simple function, then, for any $\varepsilon > 0$, there exists an $\mathcal{F}$-simple function $\tilde{f}$ with the same range as f for which $\|\tilde{f} - f\|_p < \varepsilon$ and*

$$\mu\{\omega \in \Omega | \varphi(\tilde{f}(\omega)) > \varphi(f(\omega))\} < \varepsilon.$$

Proof of Theorem 3.1.7. We shall prove the theorem for the special case where $k_e = l_e = 0$ and $b_e = c_e = 0$: the proof of the general case is similar. In what follows, we use $\varphi_t(u_t)$ to denote $p_t \cdot u_t + \frac{1}{2} u_t \cdot P_t u_t$ and $\psi_t(v_t)$ to denote $-q_t \cdot v_t + \frac{1}{2} v_t \cdot Q_t v_t$. For $u \in \mathcal{L}_k^r$ and $v \in \mathcal{L}_l^{r'}$, define

$$\Phi(u) = \int_{t_0}^{t_1} \varphi_t(u_t) dt + \delta_{\mathcal{U}^r}(u), \quad \Psi(v) = \int_{t_0}^{t_1} \psi_t(v_t) dt + \delta_{\mathcal{V}^{r'}}(v),$$

and

$$\Gamma(u,v) = \int_{t_0}^{t_1} v_t \cdot D_t u_t dt + \gamma(u,v).$$

Then, for all $(u,v) \in \mathcal{L}_k^r \times \mathcal{L}_l^{r'}$, we have

$$\mathcal{J}(u,v) = \begin{cases} \infty & \text{if } \Phi(u) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u,v) & \text{otherwise.} \end{cases}$$

Also, we introduce the (finite-dimensional) sets

$$X_\nu = \{u \in \mathcal{L}_k^r | u_t \text{ is measurable relative to } \pi_\nu\}$$

and

$$Y_\nu = \{v \in \mathcal{L}_l^{r'} | v_t \text{ is measurable relative to } \pi_\nu\},$$

so that $\mathcal{U}_{\pi_\nu} = X_\nu \cap \mathcal{U}^r$ and $\mathcal{V}_{\pi_\nu} = Y_\nu \cap \mathcal{V}^{r'}$.

We shall show that $\Phi, \Psi, \Gamma, X_\nu$ and Y_ν satisfy the hypotheses for Corollary 2.4.6, thereby obtaining the stated Mosco convergence. It is clear that Φ and Ψ are proper, lower semicontinuous, convex functions, and that Γ is a (norm) continuous biaffine functional. Fix $\bar{u} \in \mathcal{L}_k^r$. Suppose $\Phi(\bar{u}) = \infty$. Then $\limsup \Phi(u^\nu) \leq \Phi(\bar{u})$ for any sequence converging (in norm) to $\bar{u}$; there exists such a sequence with $u^\nu \in X_\nu$ since $\cup X_\nu$ is dense in $\mathcal{L}_k^r$. Assume then that $\Phi(\bar{u}) < \infty$. Then $\bar{u} \in \mathcal{U}^r$. By Lemma 3.1.8, there exists a sequence $\tilde{u}^m$ of simple functions, with $\tilde{u}_t^m \in U$ for all $t \in [t_0, t_1]$, such that $\|\tilde{u}^m - \bar{u}\|_r \to 0$ as $m \to \infty$. For each $m \in \mathbf{N}$, there exists, by Lemma 3.1.9, u^m which is measurable with respect to one of the partitions π_ν, has the same range as $\tilde{u}^m$, and has $\|\tilde{u}^m - u^m\|_r < 1/m$. Thus, u^m converges (in norm) to $\bar{u}$. Since $u^m \in \mathcal{U}^r$, we have $\Phi(u^m) = \int_{t_0}^{t_1} \varphi_t(u_t^m) dt \to \Phi(\bar{u})$, because the map $u \mapsto \int_{t_0}^{t_1} \varphi_t(u_t) dt$ is continuous on $\mathcal{L}_k^r$ (for $r \geq 2$). Choose ν_m so that $u^m \in X_{\nu_m}$ and so that $\nu_m < \nu_{m+1}$. We now define a sequence $\{\bar{u}^\nu\}$ such that $\bar{u}^\nu \in X_\nu$, $\|\bar{u}^\nu - \bar{u}\|_r \to 0$ and $\limsup \Phi(\bar{u}^\nu) \leq \Phi(\bar{u})$. If $\nu_1 = 1$, set $\bar{u}^1 = u^1$; otherwise, choose $\bar{u}^1$ to be an arbitrary element of X_1 and set $\bar{u}^\nu = \bar{u}^1$ for $\nu = 2, \ldots, \nu_1 - 1$. For $\nu = \nu_m, \ldots, \nu_{m-1}$ set $\bar{u}^\nu = u^m$. Thus condition (i) of Corollary 2.4.6 holds. A similar argument shows that condition (ii) is also satisfied. We have therefore demonstrated the desired Mosco epi/hypo-convergence. □

Theorem 3.1.7 allows us to say something about the approximation of saddle points and saddle values. Suppose $(\bar{u}^\nu, \bar{u}_e^\nu; \bar{v}^\nu, \bar{v}_e^\nu)$ is a saddle point for $\mathcal{J}^\nu$. If, for

a some $r \in [2,\infty)$, we can guarantee that $(\bar{u}^\nu, \bar{u}_e^\nu; \bar{v}^\nu, \bar{v}_e^\nu)$ converges in the weak topology on $\mathcal{L}^r$, then Proposition 2.4.3 tells us that the limit is a saddle point for $\mathcal{J}$ relative to $(\mathcal{L}_k^r \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^r \times \mathbf{R}^{le})$. (It is also a saddle point relative to $(\mathcal{L}_k^1 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^1 \times \mathbf{R}^{le})$, by Proposition 3.1.4.) Note that for $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ to be a saddle point, it is sufficient to show that the sequence of approximates merely clusters (weakly) at $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$. This clustering is guaranteed, for example, when the sets U_t, V_t are bounded for all t (and therefore uniformly bounded, by continuity), or more generally if $U_t \in \alpha_t \mathbf{B}_k$ and $V_t \in \beta_t \mathbf{B}_l$ for some $\mathcal{L}^2$ functions α and β.

In view of these remarks, it would be desirable to have epi/hypo-convergence with respect to the weakest possible topology in which convergence of controls implies uniform convergence of trajectories. The weakest topology that we have available here is the ordinary weak topology $w(\mathcal{L}^1, \mathcal{L}^\infty)$ on $\mathcal{L}^1$. The proof of the theorem above doesn't extend directly to this case: the map $(u,v) \mapsto \int_{t_0}^{t_1} v_t \cdot D_t u_t dt$ is not necessarily a continuous bilinear functional on $\mathcal{L}_k^1 \times \mathcal{L}_l^1$. In addition, the functionals $u \mapsto \int_{t_0}^{t_1} \frac{1}{2} u_t \cdot P_t u_t dt$ and $v \mapsto \int_{t_0}^{t_1} \frac{1}{2} v_t \cdot Q_t v_t dt$ are unlikely to be continuous. Of course, $\mathcal{L}^1$ is also nonreflexive, so Mosco convergence doesn't make sense. However, there is an important case where at least the continuity difficulties disappear, namely, the case where P_t, Q_t and D_t are the zero matrices for all t. This is the case treated by our next result.

Theorem 3.1.10. *Let $\{\pi_\nu : \nu \in \mathbf{N}\}$ be an increasing sequence of partitions of $[t_0, t_1]$ with $|a_i - a_{i+1}| \to 0$ uniformly in i as $\nu \to 0$. Define $\mathcal{J}_\nu$ as in the statement of Theorem 3.1.7. Suppose that $P_t = 0$, $Q_t = 0$ and $D_t = 0$ for all $t \in [t_0, t_1]$. Then for any $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \in (\mathcal{L}_k^1 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^1 \times \mathbf{R}^{le})$, one has*

$$\begin{aligned} \text{seq-e}_\text{s}/\text{h}_\text{w}\text{-ls}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) &\le \mathcal{J}(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \\ &\le \text{seq-h}_\text{s}/\text{e}_\text{w}\text{-li}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e), \end{aligned} \tag{3.1.1}$$

where s denotes the norm topology on $\mathcal{L}^1$ and w denotes the weak topology on $\mathcal{L}^1$.

Proof. We shall prove the theorem for the special case where $k_e = l_e = 0$ and $b_e = c_e = 0$: the proof of the general case is similar. For $u \in \mathcal{L}_k^1$ and $v \in \mathcal{L}_l^1$, define

$$\Phi(u) = \int_{t_0}^{t_1} p_t \cdot u_t dt + \delta_{\mathcal{U}1}(u) \text{ and } \Psi(v) = \int_{t_0}^{t_1} q_t \cdot v_t dt + \delta_{\mathcal{V}1}(v)$$

Then, for all $(u,v) \in \mathcal{L}_k^1 \times \mathcal{L}_l^1$, we have

$$\mathcal{J}(u,v) = \begin{cases} \infty & \text{if } \Phi(u) = \infty, \\ \Phi(u) - \Psi(v) - \gamma(u,v) & \text{otherwise.} \end{cases}$$

Also, we introduce the (finite-dimensional) sets

$$X_\nu = \{u \in \mathcal{L}_k^1 | u_t \text{ is measurable relative to } \pi_\nu\}$$

and

$$Y_\nu = \{v \in \mathcal{L}_l^1 | v_t \text{ is measurable relative to } \pi_\nu\},$$

so that $\mathcal{U}_{\pi_\nu} = X_\nu \cap \mathcal{U}^1$ and $\mathcal{V}_{\pi_\nu} = Y_\nu \cap \mathcal{V}^1$.

Fix $(\bar{u},\bar{v}) \in \mathcal{L}_k^1 \times \mathcal{L}_l^1$. Let $\Phi_\nu(u) = \Phi(u) + \delta_{X_\nu} = \int_{t_0}^{t_1} \varphi_t(u_t)dt + \delta_{\mathcal{U}_{\pi_\nu}}$. In the proof of Theorem 3.1.7, we showed Mosco epi-convergence of the $\{\Phi_\nu\}$ to Φ (in the reflexive setting). A similar argument here gives $\Phi(\bar{u}) = (\text{seq-e}_\text{s}\text{-lim}\,\Phi_\nu)(\bar{u})$. Also, Ψ is lower semicontinuous on $\mathcal{L}_l^1$ with respect to the norm topology, and thus with respect to the weak topology. It is also clear that γ is sequentially $s \times w$-continuous. Therefore the hypotheses of Theorem 2.3.11 are satisfied at $(\bar{u},\bar{v})$, so the left-hand inequality in (3.1.1) is valid. By similar reasoning, we may apply Theorem 2.3.11 to obtain $\underset{\sim}{\mathcal{J}}(\bar{u},\bar{v}) \leq \text{seq-h}_\text{s}/\text{e}_{\text{w}\infty}\text{-li}\,\mathcal{J}_\nu(\bar{u},\bar{v})$, where

$$\underset{\sim}{\mathcal{J}}(u,v) = \begin{cases} -\infty & \text{if } \Psi(v) = \infty, \\ \Phi(u) - \Psi(v) - \gamma(u,v) & \text{otherwise.} \end{cases}$$

(Actually, we apply Theorem 2.3.11 to $\underset{\sim}{\mathcal{J}}$ and the function

$$\underset{\sim}{\mathcal{J}}_\nu(u,v) = \begin{cases} -\infty & \text{if } \Psi_\nu(v) = \infty, \\ \Phi_\nu(u) - \Psi_\nu(v) - \gamma(u,v) & \text{otherwise,} \end{cases}$$

and then use the inequality $\underset{\sim}{\mathcal{J}}_\nu \leq \mathcal{J}_\nu$.) Thus, the right-hand inequality in (3.1.1) holds if either $\bar{u} \in \mathcal{U}^1$ or $\bar{v} \in \mathcal{V}^1$, since then we have $\underset{\sim}{\mathcal{J}}(\bar{u},\bar{v}) = \mathcal{J}(\bar{u},\bar{v})$. It remains to show that the right-hand inequality in (3.1.1) is valid when we have both $\bar{u} \notin \mathcal{U}^1$ and $\bar{v} \notin \mathcal{V}^1$. Note that $\mathcal{U}^1$ is weakly closed. Thus the complement $\mathcal{W}$ of $\mathcal{U}^1$ in $\mathcal{L}_k^1$ is a weak neighborhood of $\bar{u}$ such that $\Phi_\nu(u) = \Phi(u) = \infty$ for all $u \in \mathcal{W}$ and for all ν. Hence, $\mathcal{J}_\nu(u^\nu, v^\nu) = \infty$ eventually whenever $\{u^\nu\}$ converges weakly to $\bar{u}$ and $\{v^\nu\}$ norm-converges to $\bar{v}$. Therefore, we have $(\text{seq-h}_\text{s}/\text{e}_\text{w}\text{-li}\,\mathcal{J}_\nu)(\bar{u},\bar{v}) = \infty = \mathcal{J}(\bar{u},\bar{v})$, which completes the proof. □

We now describe how the approximate problem of finding a saddle point of $\mathcal{J}$ over $\mathcal{U}_\pi \times \mathcal{V}_\pi$ leads to a pair of optimal control problems in *discrete time* which are dual to each other. To simplify the discussion, we shall assume that $t_0 = 0$ and $t_1 = 1$, and also that $D_e = 0$. Furthermore, we will work only with the special case where the given partition is $\pi = (0, \frac{1}{T}, \ldots, 1 - \frac{1}{T}, 1)$. It will be clear how to extend the process to the general case.

Let $\mathcal{A}$ be the fundamental matrix for the homogeneous differential equation $\dot{x}_t = A_t x_t$, i.e. the $\mathbf{R}^{n\times n}$-valued function on $[0,1]$ which satisfies

$$\dot{\mathcal{A}}_t = A_t \mathcal{A}_t, \quad \mathcal{A}_0 = I.$$

Then the unique solution to the initial-value problem

$$\dot{x}_t = A_t x_t + B_t u_t + b_t \text{ a.e.}, \quad x_0 = B_e u_e + b_e$$

is given by

$$x_t = \mathcal{A}_t \left[B_e u_e + b_e + \int_0^t \mathcal{A}_s^{-1}(B_s u_s + b_s) ds \right]. \tag{3.1.2}$$

Now suppose that u_t is constant on $((\tau-1)/T, \tau/T]$ for each $\tau = 0, \ldots, T$. Using (3.1.2) we can write

$$x_{(\tau-1)/T} = \mathcal{A}_{(\tau-1)/T} \left[x_0 + \int_0^{(\tau-1)/T} \mathcal{A}_s^{-1}(B_s u_s + b_s) ds \right],$$

$$x_{\tau/T} = \mathcal{A}_{\tau/T} \left[x_0 + \int_0^{\tau/T} \mathcal{A}_s^{-1}(B_s u_s + b_s) ds \right].$$

We can combine these two equations to get the following representation for $x_{\tau/T}$:

$$x_{\tau/T} = \left[\mathcal{A}_{\tau/T} \mathcal{A}_{(\tau-1)/T}^{-1} \right] x_{(\tau-1)/T} + \left[\int_{(\tau-1)/T}^{\tau/T} \mathcal{A}_{\tau/T} \mathcal{A}_s^{-1} B_s ds \right] u_{\tau/T}$$
$$+ \left[\int_{(\tau-1)/T}^{\tau/T} \mathcal{A}_{\tau/T} \mathcal{A}_s^{-1} b_s \right].$$

This formula leads us to introduce a discrete-time control system with time periods $\tau = 0, \ldots, T$, controls $\tilde{u}_\tau = u_{\tau/T}$ and states $\tilde{x}_\tau = x_{\tau/T}$. The evolution of this system is described by

$$\tilde{x}_\tau = \tilde{A}_\tau \tilde{x}_{\tau-1} + \tilde{B}_\tau \tilde{u}_\tau + \tilde{b}_\tau, \ \tau = 1, \ldots, T$$
$$\tilde{x}_0 = \tilde{B}_0 \tilde{u}_0 + \tilde{b}_0,$$

where we define

$$\begin{aligned}
\tilde{A}_\tau &= \mathcal{A}_{\tau/T}\mathcal{A}^{-1}_{(\tau-1)/T}, \\
\tilde{B}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \mathcal{A}_{\tau/T}\mathcal{A}_s^{-1}B_s ds, \text{ for } \tau = 1,\ldots,T \\
\tilde{B}_0 &= B_e, \\
\tilde{b}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \mathcal{A}_{\tau/T}\mathcal{A}_s^{-1}b_s ds, \text{ for } \tau = 1,\ldots,T \\
\tilde{b}_0 &= b_e.
\end{aligned}$$

In a similar fashion, for a v_t which is constant on each $[(\tau-1)/T, \tau/T)$, we can introduce a dual control system:

$$\begin{aligned}
\tilde{y}_\tau &= \tilde{A}^*_\tau \tilde{y}_{\tau+1} + \tilde{C}^*_\tau \tilde{v}_\tau + \tilde{c}_\tau, \ \tau = T,\ldots,1 \\
\tilde{y}_{T+1} &= \tilde{C}^*_{T+1}\tilde{v}_{T+1} + \tilde{c}_{T+1},
\end{aligned}$$

where we define

$$\begin{aligned}
\tilde{C}_\tau &= \int_{(\tau-1)/T}^{\tau/T} C_s\mathcal{A}_s\mathcal{A}^{-1}_{(\tau-1)/T} ds, \text{ for } \tau = 1,\ldots,T \\
\tilde{C}_{T+1} &= C_e, \\
\tilde{c}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \left(\mathcal{A}^{-1}_{(\tau-1)/T}\right)^* \mathcal{A}_s^* c_s ds, \text{ for } \tau = 1,\ldots,T \\
\tilde{c}_{T+1} &= c_e.
\end{aligned}$$

For $(u, u_e; v, v_e) \in \mathcal{U}_\pi \times \mathcal{V}_\pi$ we can reexpress the value of $\mathcal{J}(u, u_e; v, v_e)$ in terms of the corresponding $(\tilde{u}, \tilde{v})$ as follows:

$$\begin{aligned}
\mathcal{J}(u, u_e; v, v_e) &= \tilde{\mathcal{J}}(\tilde{u}, \tilde{v}) \\
&:= \sum_{\tau=1}^{T} [\tilde{p}_\tau \cdot \tilde{u}_\tau + \tfrac{1}{2}\tilde{u}_\tau \cdot \tilde{P}_\tau \tilde{u}_\tau - \tilde{v}_\tau \cdot \tilde{D}_\tau \tilde{u}_\tau - \tfrac{1}{2}\tilde{v}_\tau \cdot \tilde{Q}_\tau \tilde{v}_\tau + \tilde{q}_\tau \cdot \tilde{v}_\tau - \tilde{d}_\tau] \\
&\quad + \tilde{p}_0 \cdot \tilde{u}_0 + \tfrac{1}{2}\tilde{u}_0 \cdot \tilde{P}_0 \tilde{u}_0 + \tilde{q}_{T+1} \cdot \tilde{v}_{T+1} - \tfrac{1}{2}\tilde{v}_{T+1} \cdot \tilde{Q}_{T+1}\tilde{v}_{T+1} - \tilde{\gamma}(\tilde{u}, \tilde{v}),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\gamma}(\tilde{u}, \tilde{v}) &= \sum_{\tau=1}^{T+1} x_{\tau-1}(\tilde{C}^*_\tau \tilde{v}_\tau + \tilde{c}_\tau) \\
&= \sum_{\tau=0}^{T} y_{\tau+1}(\tilde{B}_\tau \tilde{u}_\tau + \tilde{b}_\tau).
\end{aligned}$$

The new coefficients for this functional are given by

$$\begin{aligned}
\tilde{P}_\tau &= \int_{(\tau-1)/T}^{\tau/T} P_t dt, \text{ for } \tau = 1,\ldots,T \\
\tilde{P}_0 &= P_e, \\
\tilde{Q}_\tau &= \int_{(\tau-1)/T}^{\tau/T} Q_t dt, \text{ for } \tau = 1,\ldots,T \\
\tilde{Q}_{T+1} &= Q_e, \\
\tilde{p}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \left[p_t - \left(\int_{(\tau-1)/T}^{t} \mathcal{A}_s^{-1} B_s ds \right)^* \mathcal{A}_t^* c_t \right] dt, \text{ for } \tau = 1,\ldots,T \\
\tilde{p}_0 &= p_e, \\
\tilde{q}_{T+1} &= q_e, \\
\tilde{q}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \left[q_t - C_t \mathcal{A}_t \left(\int_{(\tau-1)/T}^{t} \mathcal{A}_s^{-1} b_s ds \right) \right] dt, \text{ for } \tau = 1,\ldots,T \\
\tilde{D}_\tau &= \int_{(\tau-1)/T}^{\tau/T} \left[D_t + C_t \mathcal{A}_t \left(\int_{(\tau-1)/T}^{t} \mathcal{A}_s^{-1} B_s ds \right) \right] dt \\
&= \int_{(\tau-1)/T}^{\tau/T} \left[D_t + \left(\int_{t}^{\tau/T} C_s \mathcal{A}_s ds \right) \mathcal{A}_t^{-1} B_t \right] dt, \\
\tilde{d}_\tau &= \int_{(\tau-1)/T}^{\tau/T} (\mathcal{A}_t^* c_t) \cdot \left(\int_{(\tau-1)/T}^{t} \mathcal{A}_s^{-1} b_s ds \right) dt \\
&= \int_{(\tau-1)/T}^{\tau/T} \left(\int_{t}^{\tau/T} \mathcal{A}_s^* c_s ds \right) \cdot \mathcal{A}_t^{-1} b_t dt.
\end{aligned}$$

Clearly $\tilde{\mathcal{J}}$ is a finite-valued saddle function on $(\mathbf{R}^{ke} \times \mathbf{R}^{k \cdot T}) \times (\mathbf{R}^{l \cdot T} \times \mathbf{R}^{le})$. If restricted to the constraints for $(\mathcal{P})$ and $(\mathcal{Q})$ it gives rise to to the following primal and dual problems

$$(\tilde{\mathcal{P}}) \qquad \text{minimize} \quad \tilde{F}(\tilde{u}) = \sup_{\tilde{v} \in V} \tilde{\mathcal{J}}(\tilde{u}, \tilde{v}) \quad \text{over all } \tilde{u} \in U_e \times \left(\prod_{\tau=1}^{T} U \right)$$

and

$$(\tilde{\mathcal{Q}}) \qquad \text{maximize} \quad \tilde{G}(\tilde{v}) = \inf_{\tilde{u} \in U} \tilde{\mathcal{J}}(\tilde{u}, \tilde{v}) \quad \text{over all } \tilde{v} \in \left(\prod_{\tau=1}^{T} V \right) \times V_e.$$

We see that $\tilde{F}$ is convex, while $\tilde{G}$ is concave. Both are piecewise linear-quadratic functions as can be seen from the following representations:

$$\tilde{F}(\tilde{u}) = \sum_{\tau=1}^{T}[\tilde{p}_\tau \cdot \tilde{u}_\tau + \tfrac{1}{2}\tilde{u}_\tau \cdot \tilde{P}_\tau \tilde{u}_\tau - \tilde{c}_\tau \cdot \tilde{x}_{\tau-1} + \rho_{V_\tau,Q_\tau}(\tilde{q}_\tau - \tilde{D}_\tau \tilde{u}_\tau - \tilde{C}_\tau \tilde{x}_{\tau-1}) - \tilde{d}_\tau]$$
$$+ \tilde{p}_0 \cdot \tilde{u}_0 + \tfrac{1}{2}\tilde{u}_0 \cdot \tilde{P}_0 \tilde{u}_0 - \tilde{c}_{T+1} \cdot \tilde{x}_T + \rho_{V_{T+1},Q_{T+1}}(\tilde{q}_{T+1} - \tilde{C}_{T+1}\tilde{x}_T),$$

$$\tilde{G}(\tilde{v}) = \sum_{\tau=1}^{T}[\tilde{q}_\tau \cdot \tilde{v}_\tau - \tfrac{1}{2}\tilde{v}_\tau \cdot \tilde{Q}_\tau \tilde{v}_\tau - \tilde{b}_\tau \cdot \tilde{y}_{\tau+1} - \rho_{U_\tau,P_\tau}(\tilde{D}_\tau^* \tilde{v}_\tau + \tilde{B}_\tau^* \tilde{y}_{\tau+1} - \tilde{p}_\tau) - \tilde{d}_\tau]$$
$$+ \tilde{q}_{T+1} \cdot \tilde{v}_{T+1} - \tfrac{1}{2}\tilde{v}_{T+1} \cdot \tilde{Q}_{T+1}\tilde{v}_{T+1} - \tilde{b}_0 \cdot \tilde{y}_1 + \rho_{U_0,P_0}(\tilde{B}_0 \tilde{y}_1 - \tilde{p}_0).$$

The duality theory for such discrete-time control problems has been developed by Rockafellar and Wets [35]. One of the interesting facts about these problems is that they can be reformulated as *quadratic programs*, and therefore could be solved by existing numerical routines if the dimension T is not too large. However, when the problems arise from the discretization of a continuous-time problem (as they do here), it is likely that the dimension T will be quite large, and the number of variables could easily exceed the capabilities of ordinary quadratic programming routines. On the other hand, these problems have a very special structure which allows the application of various "decomposition" techniques. One example of such a technique is the *finite envelope method*, which also has been studied recently by Rockafellar and Wets. A computer implementation for solving problems $(\tilde{\mathcal{P}})$ and $(\tilde{\mathcal{Q}})$ via this method has been developed by the author [39]; the code has been used by Zhu and Rockafellar [40] as the basis for their numerical experiments in solving large-scale optimization problems. An overview of the finite envelope method is given in Chapter 5, along with some convergence results and details of its implementation for solving problems like $(\tilde{\mathcal{P}})$.

A drawback to using the discretization described in this section is the computational (and numerical) burden of calculating the integrals defining the coefficients for the discrete-time problem, in addition to the calculation of the fundamental matrix $\mathcal{A}$. A computationally simpler scheme using finite differences is examined in section 3 of this chapter, where we explore (from a variational point of view) its relationship to the discretization used here.

2. Nonquadratic Models

In this section we will extend the approximation results of the previous section to a more general class of nonquadratic models in optimal control. As before, the approximation scheme gives rise to a discrete-time optimal control problem. The problem we shall consider is the following:

minimize the functional

$$(\mathcal{P}^1) \qquad \begin{aligned} F(u,u_e) = \int_{t_0}^{t_1} [p_t \cdot u_t + \varphi_t(u_t) - c_t \cdot x_t + \psi_t^*(q_t - D_t u_t - C_t x_t)]dt \\ + p_e \cdot u_e + \varphi_e(u_e) - c_e \cdot x_{t_1} + \psi_e^*(q_e - D_e u_e - C_e x_{t_1}) \end{aligned}$$

over the control space $\mathcal{U}^1 = \mathcal{L}_k^1 \times \mathbf{R}^{ke}$, with the dynamics given by

$$\dot{x}_t = A_t x_t + B_t u_t + b_t \text{ a.e.}, \quad x_{t_0} = B_e u_e + b_e.$$

The dual problem to $(\mathcal{P}^1)$ will be

maximize the functional

$$(\mathcal{Q}^1) \qquad \begin{aligned} G(v,v_e) = \int_{t_0}^{t_1} [q_t \cdot v_t - \psi_t(v_t) - b_t \cdot y_t - \varphi_t^*(D_t^* v_t + B_t^* y_t - p_t)]dt \\ + q_e \cdot v_e - \psi_e(v_t) - b_e \cdot y_{t_0} - \varphi_e^*(D_e^* v_e + B_e y_{t_0} - p_e) \end{aligned}$$

over the control space $\mathcal{V}^1 = \mathcal{L}_l^1 \times \mathbf{R}^{le}$ with the dynamics given by

$$-\dot{y}_t = A_t^* y_t + C_t^* v_t + c_t \text{ a.e.}, \quad y_{t_1} = C_e^* v_e + c_e.$$

As before, we shall also work with the problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^r)$ (for $r \in [1,\infty]$) given by replacing the spaces $\mathcal{U}^1$ and $\mathcal{V}^1$ by $\mathcal{U}^r = \mathcal{L}_k^r \times \mathbf{R}^{ke}$ and $\mathcal{V}^r = \mathcal{L}_l^r \times \mathbf{R}^{lc}$.

The functions φ_t, ψ_t (for each $t \in [t_0,t_1]$) and φ_e and ψ_e are assumed to be proper, lower semicontinuous convex functions, with φ_t and ψ_t varying epi-continuously with t. We also assume that the data elements p_t, q_t, A_t, B_t, b_t, C_t, c_t, D_t are all continuous with respect to $t \in [t_0,t_1]$. In addition, we shall impose the following finiteness assumption:

The functions $\varphi_t^*, \psi_t^*, \varphi_e^*, \psi_e^*$ are assumed to be finite everywhere.

Equivalently, the functions $\varphi_t, \psi_t, \varphi_e, \psi_e$ are assumed to be coercive. (An extended-real-valued function h on $\mathbf{R}^d$ is said to be *coercive* if $\lim_{|w|\to\infty} h(w)/|w| = \infty$. For example, h is coercive if the effective domain of h is bounded. Also, strong convexity implies coercivity.)

It is readily seen that these problems are more general than those treated in section 3.1. Specifically, the previous situation is the case where we have

$$\varphi_t(u_t) = \tfrac{1}{2}u_t \cdot P_t u_t + \delta_{U_t}(u_t), \quad \varphi_e(u_e) = \tfrac{1}{2}u_e \cdot P_e u_e + \delta_{U_e}(u_e),$$
$$\psi_t(v_t) = \tfrac{1}{2}v_t \cdot Q_t v_t + \delta_{V_t}(v_t), \quad \psi_e(v_e) = \tfrac{1}{2}v_e \cdot Q_e v_e + \delta_{V_e}(v_e).$$

In this case, the finiteness assumption given above corresponds to the conjunction of the primal and dual finiteness conditions of the previous section.

Instead of approximating the problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^r)$ directly, we will work with a corresponding minimax representation for optimality. To this end, we introduce the functional

$$\mathcal{J}(u, u_e; v, v_e) = \int_{t_0}^{t_1} J_t(u_t, v_t)dt + J_e(u_e, v_e) - \gamma(u, u_e; v, v_e),$$

where

$$J_t(u_t, v_t) = \begin{cases} \infty & \text{if } \varphi_t(u_t) = \infty, \\ p_t \cdot u_t + \varphi_t(u_t) - v_t \cdot D_t u_t + q_t \cdot v_t - \psi_t(v_t) & \text{otherwise,} \end{cases}$$

$$J_e(u_e, v_e) = \begin{cases} \infty & \text{if } \varphi_e(u_e) = \infty, \\ p_e \cdot u_e + \varphi_c(u_e) - v_e \cdot D_e u_e + q_e \cdot v_e - \psi_e(v_e) & \text{otherwise,} \end{cases}$$

and

$$\gamma(u, u_e; v, v_e) = \int_{t_0}^{t_1} y_t \cdot (B_t u_t + b_t)dt + y_{t_0} \cdot (B_e u_e + b_e)dt$$
$$= \int_{t_0}^{t_1} x_t \cdot (C_t^* v_t + c_t)dt + x_{t_1} \cdot (C_e^* v_e + c_e).$$

To avoid ambiguity, we use the convention here that $\infty - \infty = \infty$. This is used twice: first in evaluating the integrals, and then in adding the integrals to $J_e(u_e, v_e)$. The expression $\int_{t_0}^{t_1} J_t(u_t, v_t)dt$ is taken to be ∞ if and only if $J_t(u_t, v_t)$ is not majorized by any integrable function; it is taken to be $-\infty$ if and only if $J_t(u_t, v_t)$ is majorized by an integrable function but not minorized by an integrable function. Of course, the term $\gamma(u, u_e; v, v_e)$ is always finite.

A theory of duality for problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^{r'})$ (via the functional $\mathcal{J}$) was developed recently by Rockafellar [31]. The next three theorems give a summary of the facts that are most relevant to our purpose here. Basically, they say that the strong duality of the previous section extends to these more general problems. However, the arguments given by Rockafellar in deriving these results (including the properness of the functionals F and G) depend heavily on the finiteness assumption made earlier in this section.

Theorem 3.2.1 [31]. *Consider $r, r' \in [1, \infty]$. The functional F is (inf-)proper, lower semicontinuous and convex on $\mathcal{U}^r$, whereas G is (sup-)proper, upper semicontinuous and concave on $\mathcal{U}^{r'}$. The problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^{r'})$ are the primal and dual problems associated with the problem of finding a saddle point of $\mathcal{J}$ on $\mathcal{U}^r \times \mathcal{V}^{r'}$, i.e.*

$$F(u, u_e) = \sup_{(v, v_e) \in \mathcal{V}^{r'}} \mathcal{J}(u, u_e; v, v_e) \text{ and } G(v, v_e) = \inf_{(u, u_e) \in \mathcal{U}^r} \mathcal{J}(u, u_e; v, v_e).$$

Theorem 3.2.2 [31]. *Consider $r, r' \in [1, \infty]$. The problems $\mathcal{P}^r$ and $\mathcal{Q}^{r'}$ both admit optimal solutions, and*

$$\min(\mathcal{P}^r) = \max(\mathcal{Q}^{r'}) \quad (\textit{finite}).$$

A pair $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ is a saddle point of $\mathcal{J}$ over $\mathcal{U}^r \times \mathcal{V}^{r'}$ if and only if $(\bar{u}, \bar{u}_e)$ solves $(\mathcal{P}^r)$, $(\bar{v}, \bar{v}_e)$ solves $(\mathcal{Q}^{r'})$. Furthermore, any optimal solution of $(\mathcal{P}^r)$ is actually in $\mathcal{U}^\infty$, and any optimal solution of $(\mathcal{Q}^{r'})$ is actually in $\mathcal{V}^\infty$.

Theorem 3.2.3 (Minimaximum Principle) [31]. *Consider $r, r' \in [1, \infty]$. For the control pair $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e)$ to be a saddle point for $\mathcal{J}$ on $\mathcal{U}^r \times \mathcal{V}^{r'}$, it is necessary and sufficient for the following three conditions to hold:*

(i) *$\bar{u}_t$ and $\bar{v}_t$ are in $\mathcal{L}_k^r$ and $\mathcal{L}_l^{r'}$ respectively (with trajectories $\bar{x}_t$ and $\bar{y}_t$).*

(ii) *$(\bar{u}_t, \bar{v}_t)$ is a saddle point over $U_t \times V_t$ for*

$$J_t(u_t, v_t) - u_t \cdot B_t^* \bar{y}_t - v_t \cdot C_t \bar{x}_t.$$

(iii) *$(\bar{u}_e, \bar{v}_e)$ is a saddle point over $U_e \times V_e$ for*

$$J_e(u_e, v_e) - u_e \cdot B_e^* \bar{y}_{t_0} - v_e \cdot C_e \bar{x}_{t_1}.$$

Theorems 3.2.1 and 3.2.2 tell us that, in the search for solutions to problems $(\mathcal{P}^r)$ and $(\mathcal{Q}^{r'})$, we need only consider the problem of finding a saddle point of $\mathcal{J}$ over $\mathcal{U}^\infty \times \mathcal{V}^\infty$. Indeed, in applications it is often more natural to restrict attention to essentially bounded controls anyway.

Even so, we still need to work with the larger spaces $\mathcal{U}^r$ and $\mathcal{V}^{r'}$ in our discussion of approximations. The results we give are of the epi/hypo-convergence variety. Thus, any cluster point of a sequence of solutions to the approximate problems will solve the original problem. The difficulty in applying such a result is

that of establishing whether the approximate solutions actually cluster at all. Of course, clustering is more likely in a weaker topology than in a stronger one. Thus the sharpest result is that which guarantees epi/hypo-convergence relative to the weakest topology available. On the other hand, many applications require that the trajectories corresponding to approximate controls converge uniformly (or cluster in the uniform norm) to an optimal trajectory. This requires working with the weak topologies for $\mathcal{L}^r$ on the controls.

We will use the same approximation scheme as in section 1, namely approximation of controls by (feasible) step functions. Of course, there might not be any feasible step functions, so our hypotheses will typically require the domains $\operatorname{dom}\varphi_t$ and $\operatorname{dom}\psi_t$ to be constant with respect to t. Note, however, that this does not necessarily restrict φ_t and ψ_t from varying with t.

Let $\pi = (t_0 = a_0 < a_1 < \ldots < a_T = t_1)$ be a partition of the interval $[t_0, t_1]$. Consider the following subsets of $\mathcal{U}^\infty$ and $\mathcal{V}^\infty$:

$$\mathcal{U}_\pi = \{(u, u_e) \in \mathcal{U}^\infty | u_t \text{ is constant a.e. on } [a_{\tau-1}, a_\tau] \text{ for } \tau = 0, \ldots, T\},$$
$$\mathcal{V}_\pi = \{(v, v_e) \in \mathcal{V}^\infty | v_t \text{ is constant a.e. on } [a_{\tau-1}, a_\tau] \text{ for } \tau = 0, \ldots, T\}.$$

The approximate problem is

$$(\mathcal{S}_\pi) \qquad \text{find a saddle point of } \mathcal{J} \text{ relative to } \mathcal{U}_\pi \times \mathcal{V}_\pi.$$

This is a finite-dimensional problem, and we shall show that it can be restated as a problem of discrete-time optimal control. First we give our epi/hypo-convergence results.

Let $\{\pi_\nu : \nu \in \mathbf{N}\}$ be an increasing sequence of partitions of $[t_0, t_1]$ such that $|a_i - a_{i+1}| \to 0$ uniformly in i as $\nu \to 0$. Define

$$\mathcal{J}_\nu(u, u_e; v, v_e) = \begin{cases} \mathcal{J}(u, u_e; v, v_e) & \text{if } (u, u_e) \in \mathcal{U}_{\pi_\nu}, (v, v_e) \in \mathcal{V}_{\pi_\nu}, \\ -\infty & \text{if } (u, u_e) \in \mathcal{U}_{\pi_\nu}, (v, v_e) \notin \mathcal{V}_{\pi_\nu}, \\ \infty & \text{if } (u, u_e) \notin \mathcal{U}_{\pi_\nu}. \end{cases}$$

The following theorems will be concerned with this sequence of problems. In addition, the theorems will make use of the following condition on $r, r' \in [1, \infty)$:

$$\text{The map } (u, v) \mapsto \int_{t_0}^{t_1} v_t \cdot D_t u_t dt \text{ defines a (norm-) continuous bilinear functional on } \mathcal{L}_k^r \times \mathcal{L}_l^{r'}. \tag{3.2.1}$$

This condition is satisfied if $r \geq r'/(r'-1)$ (or equivalently, if $r' \geq r/(r-1)$), in which case $\mathcal{L}_k^r \subset (\mathcal{L}_l^{r'})^*$ and $\mathcal{L}_l^{r'} \subset (\mathcal{L}_k^r)^*$. If $D \equiv 0$, then the condition is satisfied for any choice of $r, r' \in [1, \infty)$.

Our first result applies to problems where φ_t has the form $\bar{\varphi}_t + \delta_U$ and ψ_t has the form $\bar{\psi}_t + \delta_V$, where we require that $\bar{\varphi}_t$ and $\bar{\psi}_t$ are finite-valued.

Theorem 3.2.4. *Consider $r, r' \in [1, \infty)$ satisfying condition (3.2.1). Let $\bar{\varphi}_t$ and $\bar{\psi}_t$ be continuous convex functions (on $\mathbf{R}^k$ and $\mathbf{R}^l$ respectively) which vary epi-continuously with t, and let $U \subset \mathbf{R}^k$ and $V \subset \mathbf{R}^l$ be nonempty, closed convex sets. Assume that $\varphi_t = \bar{\varphi}_t + \delta_U$ and $\psi_t = \bar{\psi}_t + \delta_V$, and that the functionals $u \mapsto \int_{t_0}^{t_1} \bar{\varphi}_t(u_t)dt$ and $v \mapsto \int_{t_0}^{t_1} \bar{\psi}_t(v_t)dt$ are continuous on $\mathcal{U}^r$ and $\mathcal{V}^{r'}$ respectively. If $r, r' \in (1, \infty)$ then $\mathcal{J}_\nu$ Mosco-epi/hypo-converges to $\mathcal{J}$ over $\mathcal{U}^r \times \mathcal{V}^{r'}$. If either $r = 1$ or $r' = 1$, then for any $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \in (\mathcal{U}^r \times \mathcal{V}^{r'})$ one has*

$$\begin{aligned} \text{seq-e}_{s_r}/\text{h}_{w_{r'}}\text{-ls}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) &\leq \mathcal{J}(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \\ &\leq \text{seq-h}_{s_{r'}}/\text{e}_{w_r}\text{-li}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e), \end{aligned} \tag{3.2.2}$$

where s_r and $s_{r'}$ denote the norm topologies on $\mathcal{U}^r$ and $\mathcal{V}^{r'}$, and w_r and $w_{r'}$ are the respective weak topologies.

Proof. We shall prove the theorem for the special case where $k_e = l_e = 0$ and $b_e = c_e = 0$: the proof of the general case is similar. For $u \in \mathcal{L}_k^r$ and $v \in \mathcal{L}_l^{r'}$, define

$$\Phi(u) = \int_{t_0}^{t_1} \varphi_t(u_t)dt, \quad \Psi(v) = \int_{t_0}^{t_1} \psi_t(v_t)dt$$

and

$$\Gamma(u, v) = \int_{t_0}^{t_1} [v_t \cdot D_t u_t - p_t \cdot u_t - q_t \cdot v_t]dt + \gamma(u, v).$$

Then, for all $(u, v) \in \mathcal{L}_k^r \times \mathcal{L}_l^{r'}$, we have

$$\mathcal{J}(u, v) = \begin{cases} \infty & \text{if } \Phi(u) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u, v) & \text{otherwise.} \end{cases}$$

We first treat the case where $r, r' \in (1, \infty)$. Here we need to show that $\Phi, \Psi, \Gamma, \mathcal{U}_{\pi_\nu}$ and $\mathcal{V}_{\pi_\nu}$ satisfy the hypotheses for Corollary 2.4.6, thereby obtaining the stated Mosco convergence. It is clear that Φ and Ψ are proper, lower semicontinuous, convex functions, and that Γ is a (norm) continuous biaffine functional.

Fix $\bar{u} \in \mathcal{L}_k^r$. Suppose $\Phi(\bar{u}) = \infty$. Then $\limsup \Phi(u^\nu) \leq \Phi(\bar{u})$ for any sequence converging (in norm) to $\bar{u}$; there exists such a sequence with $u^\nu \in \mathcal{U}_{\pi_\nu}$ since $\cup \mathcal{U}_{\pi_\nu}$ is dense in $\mathcal{L}_k^r$. Assume then that $\Phi(\bar{u}) < \infty$. Then $\bar{u}_t \in U$ for almost every t. By Lemma 3.1.8, there exists a sequence $\tilde{u}^m$ of simple functions, with $\tilde{u}_t^m \in U$ for all $t \in [t_0, t_1]$, such that $\|\tilde{u}^m - u\|_r \to 0$ as $m \to 0$. For each $m \in \mathbf{N}$, there exists, by Lemma 3.1.9, u^m which is measurable with respect to one of the partitions π_ν, has the same range as $\tilde{u}^m$, and has $\|\tilde{u}^m - u^m\|_r < 1/m$. Thus, u^m converges (in norm) to $\bar{u}$. Since $u_t^m \in U$, we have $\Phi(u^m) = \int_{t_0}^{t_1} \bar{\varphi}_t(u_t^m)dt \to \Phi(\bar{u})$, because the map $u \mapsto \int_{t_0}^{t_1} \bar{\varphi}_t(u_t)dt$ is continuous on $\mathcal{L}_k^r$. Choose $\{\nu_m\}$ with $\nu_m < \nu_{m+1}$ so that $u^m \in \mathcal{U}_{\pi_\nu}$ for $\nu = \nu_m$, We now define a sequence $\{\bar{u}^\nu\}$ such that $\bar{u}^\nu \in \mathcal{U}_{\pi_\nu}$, $\|\bar{u}^\nu - \bar{u}\|_r \to 0$ and $\limsup \Phi(\bar{u}^\nu) \leq \Phi(\bar{u})$. If $\nu_1 = 1$, set $\bar{u}^1 = u^1$; otherwise, choose $\bar{u}^1$ to be an arbitrary element of $\mathcal{U}_{\pi_1}$ and set $\bar{u}^\nu = \bar{u}^1$ for $\nu = 2, \ldots, \nu_1 - 1$. For $\nu = \nu_m, \ldots, \nu_{m-1}$ set $\bar{u}^\nu = u^m$. Thus condition (i) of Corollary 2.4.6 holds. A similar argument shows that condition (ii) is also satisfied. We have therefore demonstrated the desired Mosco epi/hypo-convergence.

We now turn to the case where either $r = 1$ or $r' = 1$, and prove the sequential epi/hypo-limit given by (3.2.2). Fix $(\bar{u}, \bar{v}) \in \mathcal{L}_k^r \times \mathcal{L}_l^{r'}$. The same argument as above can be used to show that $\Phi(\bar{u}) = (\text{seq-e}_{s_r}\text{-lim}\, \Phi_\nu)(\bar{u})$. Also, Ψ is lower semicontinuous on $\mathcal{L}_l^{r'}$ with respect to the norm topology, and thus with respect to the weak topology. It is also clear that Γ is sequentially $s_r \times w_{r'}$-continuous. Therefore the hypotheses of Theorem 2.3.11 are satisfied at $(\bar{u}, \bar{v})$, so the left-hand inequality in (3.2.2) is valid. By similar reasoning, we may apply Theorem 2.3.11 to obtain $\underset{\sim}{\mathcal{J}}(\bar{u}, \bar{v}) \leq \text{seq-h}_{s_{r'}}/\text{e}_{w_r}\text{-li}\, \mathcal{J}_\nu(\bar{u}, \bar{v})$, where

$$\underset{\sim}{\mathcal{J}}(u, v) = \begin{cases} -\infty & \text{if } \Psi(v) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u, v) & \text{otherwise.} \end{cases}$$

(Actually, we apply Theorem 2.3.11 to $\underset{\sim}{\mathcal{J}}$ and the functions

$$\underset{\sim}{\mathcal{J}}_\nu(u, v) = \begin{cases} -\infty & \text{if } \Psi_\nu(v) = \infty, \\ \Phi_\nu(u) - \Psi_\nu(v) - \Gamma(u, v) & \text{otherwise,} \end{cases}$$

and then use the inequality $\underset{\sim}{\mathcal{J}}_\nu \leq \mathcal{J}_\nu$.) Thus, the right-hand inequality in (3.2.2) holds if either $\bar{u}_t \in U$ a.e. or $\bar{v}_t \in V$ a.e., since then we have $\underset{\sim}{\mathcal{J}}(\bar{u}, \bar{v}) = \mathcal{J}(\bar{u}, \bar{v})$. Define

$$\overline{\mathcal{U}} = \{u \in \mathcal{L}_k^r | u_t \in U a.e.\} \text{ and } \overline{\mathcal{V}} = \{v \in \mathcal{L}_l^{r'} | v_t \in V a.e.\},$$

so that we can write $\Phi(u) = \int_{t_0}^{t_1} \bar{\varphi}_t(u_t)dt + \delta_{\overline{U}}(u)$ and $\Psi(v) = \int_{t_0}^{t_1} \bar{\psi}_t(v_t)dt + \delta_{\overline{V}}(v)$. It remains to show that the right-hand inequality in (3.2.2) is valid when we have both $\bar{u} \notin \overline{U}$ and $\bar{v} \notin \overline{V}$. Note that $\overline{U}$ is w_r-closed. Thus the complement $\mathcal{W}$ of $\overline{U}$ in $\mathcal{L}_k^r$ is a w_r-neighborhood of $\bar{u}$ such that $\Phi_\nu(u) = \Phi(u) = \infty$ for all $u \in \mathcal{W}$ and for all ν. Hence, $\mathcal{J}_\nu(u^\nu, v^\nu) = \infty$ eventually whenever $\{u^\nu\}$ w_r-converges to $\bar{u}$ and $\{v^\nu\}$ norm-converges to $\bar{v}$. Therefore, we have $(\text{seq-h}_{\text{s}_{r'}}/\text{e}_{\text{w}_r}\text{-li}\, \mathcal{J}_\nu)(\bar{u}, \bar{v}) = \infty = \mathcal{J}(\bar{u}, \bar{v})$, which completes the proof. □

Our next approximation theorem requires φ_t and ψ_t to be constant with respect to t, but without any finiteness conditions. The proof uses the following modified version of Lemma 3.1.8, concerning approximation by simple functions.

Lemma 3.2.5. *Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure and suppose $1 \le p \le \infty$. Consider $f \in \mathcal{L}_d^p(\Omega, \mathcal{A}, \mu)$ and an inf-compact convex function h on $\mathbf{R}^d$. If $f(\omega) \in \operatorname{dom} h$ a.e. $[\mu]$, then there is a sequence $\{f_\nu\}$ of simple functions satisfying:*

(i) $h(f_\nu(\omega)) \le h(f_{\nu+1}(\omega)) \le h(f(\omega))$ a.e. $[\mu]$.

(ii) $f_\nu \to f$ a.e. $[\mu]$.

(iii) $f_\nu \to f$ *in* $\mathcal{L}_d^p(\Omega, \mathcal{A}, \mu)$.

Theorem 3.2.6. *Consider $r, r' \in [1, \infty)$ satisfying condition (3.2.1). Let $\bar{\varphi}$ and $\bar{\psi}$ be proper, lower semicontinuous convex functions on $\mathbf{R}^k$ and $\mathbf{R}^l$, respectively. Assume that $\varphi_t = \bar{\varphi}$ and $\psi_t = \bar{\psi}$ for all t. If $r, r' \in (1, \infty)$ then $\mathcal{J}_\nu$ Mosco-epi/hypo-converges to $\mathcal{J}$ over $\mathcal{U}^r \times \mathcal{V}^{r'}$. If either $r = 1$ or $r' = 1$, then*

$$\begin{aligned} \text{seq-e}_{\text{s}_r}/\text{h}_{\text{w}_{r'}}\text{-ls}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) &\le \mathcal{J}(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \\ &\le \text{seq-h}_{\text{s}_{r'}}/\text{e}_{\text{w}_r}\text{-li}\, \mathcal{J}_\nu(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e), \end{aligned} \tag{3.2.3}$$

is satisfied at any $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \in (\mathcal{U}^r \times \mathcal{V}^{r'})$ for which either $\int_{t_0}^{t_1} \bar{\varphi}(\bar{u}_t)dt + \varphi_e(\bar{u}_e)$ or $\int_{t_0}^{t_1} \bar{\psi}(\bar{v}_t)dt + \psi_e(\bar{v}_e)$ is finite. If the effective domain of $u \mapsto \int_{t_0}^{t_1} \bar{\varphi}(u_t)dt$ in $\mathcal{L}_k^r$ is closed, then (3.2.3) is satisfied for all $(\bar{u}, \bar{u}_e; \bar{v}, \bar{v}_e) \in (\mathcal{U}^r \times \mathcal{V}^{r'})$.

Proof. We shall prove the theorem for the special case where $k_e = l_e = 0$ and $b_e = c_e = 0$: the proof of the general case is similar. For $u \in \mathcal{L}_k^r$ and $v \in \mathcal{L}_l^{r'}$, define

$$\Phi(u) = \int_{t_0}^{t_1} \varphi_t(u_t)dt, \quad \Psi(v) = \int_{t_0}^{t_1} \psi_t(v_t)dt$$

and

$$\Gamma(u,v) = \int_{t_0}^{t_1} [v_t \cdot D_t u_t - p_t \cdot u_t - q_t \cdot v_t]dt + \gamma(u,v).$$

Then, for all $(u,v) \in \mathcal{L}_k^r \times \mathcal{L}_l^{r'}$, we have

$$\mathcal{J}(u,v) = \begin{cases} \infty & \text{if } \Phi(u) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u,v) & \text{otherwise.} \end{cases}$$

We first treat the case where $r, r' \in (1,\infty)$. As in the proof of Theorem 3.2.4, the crux of the matter is to find, for each $\bar{u} \in \mathcal{L}_k^r$, a sequence $\{u^\nu\}$ converging to $\bar{u}$, such that $u^\nu \in \mathcal{U}_{\pi_\nu}$ for all ν and $\limsup \Phi(u^\nu) \leq \Phi(\bar{u})$. In the case where $\Phi(\bar{u}) = \infty$, this is trivial. Assume then that $\Phi(\bar{u}) < \infty$. Then $\bar{u}_t \in \operatorname{dom} \bar{\varphi}$ for almost every t. Since $\bar{\varphi}$ is coercive, it is inf-compact. By Lemma 3.2.5, there exists a sequence $\{\tilde{u}^m\}$ of simple functions such that $\|\tilde{u}^m - \bar{u}\|_r \to 0$ as $m \to \infty$ and $\bar{\varphi}(\tilde{u}_t^m) \leq \bar{\varphi}(\bar{u}_t)$ almost everwhere. In particular we have $\Phi(\tilde{u}^m) \leq \Phi(\bar{u})$ for all m. For each $m \in \mathbf{N}$ there exists, by Lemma 3.1.9, u^m which is measurable with respect to one of the partitions π_ν and has $\|\tilde{u}^m - u^m\|_r < 1/m$. Moreover, u^m may be chosen so as to have the same range as $\tilde{u}^m$ and to satisfy

$$\mu\{t \in [t_0,t_1] | \bar{\varphi}(u_t^m) > \bar{\varphi}(\tilde{u}_t^m)\} < \frac{1}{m \cdot \max\{1, \bar{\alpha}\}},$$

where $\bar{\alpha} = \max\{\bar{\varphi}(\tilde{u}_t^m) - \bar{\varphi}(\tilde{u}_{t'}^m) | t, t' \in [t_0,t_1]\}$. Hence u^m may be chosen so that $\Phi(u^m) < \Phi(\tilde{u}^m) + (1/m)$. Thus, the sequence $\{u^m\}$ converges (in norm) to $\bar{u}$ and $\limsup_{m\to\infty} \Phi(u^m) \leq \Phi(\bar{u})$. The rest of the argument giving Mosco epi/hypo-convergence is the same as in the proof of Theorem 3.2.4.

Assuming now that either $r = 1$ or $r' = 1$, we turn to the sequential epi/hypo-limit given by (3.2.3). Fix $(\bar{u},\bar{v}) \in \mathcal{L}_k^r \times \mathcal{L}_l^{r'}$. The same argument as above can be used to show that $\Phi(\bar{u}) = (\text{seq-e}_{s_r}\text{-lim}\, \Phi_\nu)(\bar{u})$. Also, Ψ is lower semicontinuous on $\mathcal{L}_l^{r'}$ with respect to the norm topology, and thus with respect to the weak topology. It is also clear that Γ is sequentially $s_r \times w_{r'}$-continuous. Therefore the hypotheses of Theorem 2.3.11 are satisfied at $(\bar{u},\bar{v})$, so the left-hand inequality in (3.2.3) is valid. By similar reasoning, we may apply Theorem 2.3.11 to obtain

$$\mathcal{J}(\bar{u},\bar{v}) \leq \text{seq-h}_{s_{r'}}/\text{e}_{w_r}\text{-li}\, \mathcal{J}_\nu(\bar{u},\bar{v}),$$

where

$$\mathcal{J}(u,v) = \begin{cases} -\infty & \text{if } \Psi(v) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u,v) & \text{otherwise.} \end{cases}$$

(Actually, we apply Theorem 2.3.11 to $\mathcal{J}$ and the functions

$$\underset{\sim}{\mathcal{J}}_\nu(u,v) = \begin{cases} -\infty & \text{if } \Psi_\nu(v) = \infty, \\ \Phi_\nu(u) - \Psi_\nu(v) - \Gamma(u,v) & \text{otherwise,} \end{cases}$$

and then use the inequality $\underset{\sim}{\mathcal{J}}_\nu \le \mathcal{J}_\nu$.) Thus, the right-hand inequality in (3.2.3) holds if either $\bar{u} \in \operatorname{dom}\Phi$ or $\bar{v} \in \operatorname{dom}\Psi$, since then we have $\underset{\sim}{\mathcal{J}}(\bar{u},\bar{v}) = \mathcal{J}(\bar{u},\bar{v})$.

It remains to show that the right-hand inequality in (3.2.2) is valid when we have both $\bar{u} \notin \operatorname{dom}\Phi$ and $\bar{v} \notin \operatorname{dom}\Psi$, in the case where $\operatorname{dom}\Phi$ is closed. The complement $\mathcal{W}$ of $\operatorname{dom}\Phi$ in $\mathcal{L}_k^r$ is a w_r-neighborhood of $\bar{u}$ where $\Phi_\nu = \Phi \equiv \infty$ for all ν. Hence, $\mathcal{J}_\nu(u^\nu, v^\nu) = \infty$ eventually whenever $\{u^\nu\}$ w_r-converges to $\bar{u}$ and $\{v^\nu\}$ norm-converges to $\bar{v}$. Therefore, we have $(\text{seq-h}_{s_{r'}}/e_{w_r}\text{-li}\,\mathcal{J}_\nu)(\bar{u},\bar{v}) = \infty = \mathcal{J}(\bar{u},\bar{v})$, which completes the proof. □

The approximate problem of finding a saddle point of $\mathcal{J}$ over $\mathcal{U}_\pi \times \mathcal{V}_\pi$ leads to a pair of optimal control problems in *discrete time* which are dual to each other, in much the same way as done in section 1. The new primal and dual problems are

$$(\tilde{\mathcal{P}}) \qquad \text{minimize } \tilde{F}(\tilde{u}) = \sup_{\tilde{v}} \tilde{\mathcal{J}}(\tilde{u},\tilde{v}) \text{ over all } \tilde{u} \in \mathbf{R}^{ke} \times \Big(\prod_{\tau=1}^{T} \mathbf{R}^k\Big)$$

and

$$(\tilde{\mathcal{Q}}) \qquad \text{maximize } \tilde{G}(\tilde{v}) = \inf_{\tilde{u}} \tilde{\mathcal{J}}(\tilde{u},\tilde{v}) \text{ over all } \tilde{v} \in \Big(\prod_{\tau=1}^{T} \mathbf{R}^l\Big) \times \mathbf{R}^{le}.$$

The functions $\tilde{F}$ and $\tilde{G}$ are defined by

$$\begin{aligned} \tilde{F}(\tilde{u}) = {} & \sum_{\tau=1}^{T} [\tilde{p}_\tau \cdot \tilde{u}_\tau + \tilde{\varphi}_\tau(\tilde{u}_\tau) - \tilde{c}_\tau \cdot \tilde{x}_{\tau-1} + (\tilde{\psi}_\tau)^*(\tilde{q}_\tau - \tilde{D}_\tau \tilde{u}_\tau - \tilde{C}_\tau \tilde{x}_{\tau-1}) - \tilde{d}_\tau] \\ & + \tilde{p}_0 \cdot \tilde{u}_0 + \tilde{\varphi}_0(\tilde{u}_0) - \tilde{c}_{T+1} \cdot \tilde{x}_T + (\tilde{\psi}_{T+1})^*(\tilde{q}_{T+1} - \tilde{C}_{T+1}\tilde{x}_T), \\ \tilde{G}(\tilde{v}) = {} & \sum_{\tau=1}^{T} [\tilde{q}_\tau \cdot \tilde{v}_\tau - \tilde{\psi}_\tau(\tilde{v}_\tau) - \tilde{b}_\tau \cdot \tilde{y}_{\tau+1} - (\tilde{\varphi}_\tau)^*(\tilde{D}_\tau^* \tilde{v}_\tau + \tilde{B}_\tau^* \tilde{y}_{\tau+1} - \tilde{p}_\tau) - \tilde{d}_\tau] \\ & + \tilde{q}_{T+1} \cdot \tilde{v}_{T+1} - \tilde{\psi}_{T+1}(\tilde{v}_{T+1}) - \tilde{b}_0 \cdot \tilde{y}_1 + (\tilde{\varphi}_0)^*(\tilde{B}_0 \tilde{y}_1 - \tilde{p}_0), \end{aligned}$$

where the "trajectories" $\tilde{x}$ and $\tilde{y}$ are calculated as follows:

$$\begin{aligned} \tilde{x}_\tau &= \tilde{A}_\tau \tilde{x}_{\tau-1} + \tilde{B}_\tau \tilde{u}_\tau + \tilde{b}_\tau, \ \tau = 1,\dots,T \\ \tilde{x}_0 &= \tilde{B}_0 \tilde{u}_0 + \tilde{b}_0, \end{aligned}$$

and

$$\tilde{y}_\tau = \tilde{A}_\tau^* \tilde{y}_{\tau+1} + \tilde{C}_\tau^* \tilde{v}_\tau + \tilde{c}_\tau, \ \tau = T, \ldots, 1$$
$$\tilde{y}_{T+1} = \tilde{C}_{T+1}^* \tilde{v}_{T+1} + \tilde{c}_{T+1}.$$

The coefficients $\tilde{A}_\tau, \tilde{B}_\tau, \tilde{b}_\tau, \tilde{C}_\tau, \tilde{c}_\tau, \tilde{p}_\tau, \tilde{q}_\tau, \tilde{D}_\tau, \tilde{d}_\tau$ are all calculated by the same formulas as in section 1. The principle difference here is the introduction of the functions $\tilde{\varphi}_\tau$ and $\tilde{\psi}_\tau$, which are given by

$$\tilde{\varphi}_0(\tilde{u}_0) = \varphi_e(u_0),$$
$$\tilde{\varphi}_\tau(\tilde{u}_\tau) = \int_{(\tau-1)/T}^{\tau/T} \varphi_t(u_\tau)dt, \ \text{for } \tau = 1, \ldots, T$$
$$\tilde{\psi}_\tau(\tilde{v}_\tau) = \int_{(\tau-1)/T}^{\tau/T} \psi_t(v_\tau)dt, \ \text{for } \tau = 1, \ldots, T$$
$$\tilde{\psi}_{T+1}(\tilde{v}_{T+1}) = \psi_e(v_{T+1}).$$

The problems $(\tilde{\mathcal{P}})$ and $(\tilde{\mathcal{Q}})$ belong to the realm of convex programming. Some duality theory concerning a special subclass of such problems is given in section 5.1; in section 5.2 an algorithm is proposed that might be used to solve these problems.

3. Approximation by Finite Differences

In section 1 of this chapter, we introduced a model in linear-quadratic control and demonstrated that, under certain assumptions, discretization by partitioning of the time interval can be considered as a form of variational approximation (Theorems 3.1.7 and 3.1.10). Such a discretization leads to a problem in *discrete-time* optimal control, which may be reformulated as an ordinary quadratic program. The coefficients for the discretized problem are given by integral formulas involving the fundamental matrix associated with the linear dynamics. These integrals could be calculated numerically by some quadrature scheme. In this section we will go one step further and discretize the original problem directly by a finite difference scheme. It will be shown that this also leads to a variational approximation. To simplify the presentation we will assume the original optimal control problem is *autonomous*: all the coefficients are constant with respect to time. In addition, only the Euler forward difference scheme is considered. Most of the usual multistep methods will also work.

As before, we shall actually approximate the following saddle point problem which is associated with the original optimal control problem:

($\mathcal{S}$) find a saddle point $(u, u_e; v, v_e)$ of $\mathcal{J}$ relative to $(\mathcal{L}_k^2 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^2 \times \mathbf{R}^{le})$.

Here we have

$$\mathcal{J}(u, u_e; v, v_e) = \int_0^1 [\varphi(u_t) - \psi(v_t) - v_t \cdot D_t u_t]dt + \varphi_e(u_e) - \psi_e(v_e) - \gamma(u, u_e; v, v_e),$$

where

$$\begin{aligned}\gamma(u, u_e; v, v_e) &= \int_0^1 x_t \cdot (C^* v_t + c)dt + x_1 \cdot (C_e^* v_e + c_e) \\ &= \int_0^1 y_t \cdot (Bu_t + b)dt + y_0 \cdot (B_e u_e + b_e),\end{aligned}$$

and the dynamics are given by

$$\begin{aligned}\dot{x}_t &= Ax_t + Bu_t + b \text{ a.e. }, \quad x_0 = B_e u_e + b_e, \\ -\dot{y}_t &= A^* y_t + C^* v_t + c \text{ a.e. }, \quad y_1 = C_e^* v_e + c_e.\end{aligned}$$

The functions φ, ψ, φ_e and ψ_e are proper, lower semicontinuous and convex. We assume they satisfy suitable conditions for $\mathcal{J}$ to be a closed saddle function (e.g. see §1 and §2).

We partition the unit interval with a uniform stepsize $h = 1/T$, where T is a positive integer. Associated with this stepsize will be two approximate saddle point problems. The first is the discrete-time problem given in sections 1 and 2, which consists in using the original saddle function and dynamics, but restricting the controls to be constant on each interval $[\tau h, (\tau + 1)h]$:

($\widetilde{\mathcal{S}}_h$) find a saddle point $(u_0^h, \ldots, u_T^h; v_1^h, \ldots, v_{T+1}^h)$ of $\widetilde{\mathcal{J}}_h$ relative to $(\mathbf{R}^{ke} \times \mathbf{R}^{kT}) \times (\mathbf{R}^{lT} \times \mathbf{R}^{le})$.

Here we define

$$\begin{aligned}&\widetilde{\mathcal{J}}_h(u_0^h, \ldots, u_T^h; v_1^h, \ldots, v_{T+1}^h) \\ &\qquad = \sum_{\tau=1}^{T} [\widetilde{\varphi}_h(u_\tau^h) - \widetilde{\psi}_h(v_\tau^h) - v_\tau^h \widetilde{D}_h u_\tau^h - d_h] \\ &\qquad\quad + \varphi_e(u_0^h) - \psi_e(v_{T+1}^h) - \widetilde{\gamma}_h(u_0^h, \ldots, u_T^h; v_1^h, \ldots, v_{T+1}^h),\end{aligned}$$

where

$$\widetilde{\gamma}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h) = \sum_{\tau=1}^{T} \widetilde{x}_{\tau-1}^h \cdot (\widetilde{C}_h^* v_\tau^h + \widetilde{c}_h) + \widetilde{x}_T^h \cdot (C_e^* v_{T+1}^h + c_e)$$
$$= \sum_{\tau=1}^{T} \widetilde{y}_{\tau+1}^h \cdot (\widetilde{B}_h u_\tau^h + \widetilde{b}_h) + \widetilde{y}_1^h \cdot (B_e u_0^h + b_e)$$

and

$$\begin{aligned}
\widetilde{x}_\tau^h &= \widetilde{A}_h \widetilde{x}_{\tau-1}^h + \widetilde{B}_h u_\tau^h + \widetilde{b}_h, \text{ for } \tau = 1,\ldots,T \\
\widetilde{x}_0^h &= B_e u_0^h + b_e, \\
\widetilde{y}_\tau^h &= \widetilde{A}_h^* \widetilde{y}_{\tau+1}^h + \widetilde{C}_h^* v_\tau^h + \widetilde{c}_h, \text{ for } \tau = T,\ldots,1 \\
\widetilde{y}_{T+1}^h &= C_e^* v_{T+1}^h + c_e.
\end{aligned}$$

With the notation $M_h = \int_0^h e^{sA} ds$ and $S_h = \int_0^h \int_0^t e^{sA} ds\, dt$, we may write the above coefficients as:

$$\begin{aligned}
&\widetilde{A}_h = e^{hA}, \\
&\widetilde{B}_h = M_h B, \quad \widetilde{b}_h = M_h b, \\
&\widetilde{C}_h = C M_h, \quad \widetilde{c}_h = M_h^* c, \\
&\widetilde{D}_h = hD + C S_h B, \quad \widetilde{d}_h = c \cdot S_h b.
\end{aligned}$$

The functions $\widetilde{\varphi}_h$ and $\widetilde{\psi}_h$ are defined by

$$\begin{aligned}
\widetilde{\varphi}_h(u_\tau^h) &= h\varphi(u_\tau^h) - c \cdot S_h B u_\tau^h, \\
\widetilde{\psi}_h(v_\tau^h) &= h\psi(v_\tau^h) + b \cdot S_h^* C^* v_\tau^h.
\end{aligned}$$

Note that the matrices M_h and S_h commute with each other and with A and e^{hA}. It can also be shown that $e^{hA} = I + M_h A$ and $M_h = hI + S_h A$.

The following approximate problem is given using forward finite differences:

$(\widehat{\mathcal{S}}_h)$ find a saddle point $(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)$ of $\widehat{\mathcal{J}}_h$ relative to $(\mathbf{R}^{ke} \times \mathbf{R}^{kT}) \times (\mathbf{R}^{lT} \times \mathbf{R}^{le})$.

This problem has the same form as $(\widetilde{\mathcal{S}}_h)$, but we replace $\widetilde{\varphi}_h$, $\widetilde{\psi}_h$, $\widetilde{A}_h$, $\widetilde{B}_h$, $\widetilde{b}_h$, $\widetilde{C}_h$, $\widetilde{c}_h$, $\widetilde{D}_h$, and $\widetilde{d}_h$ with functions $\widehat{\varphi}_h$, $\widehat{\psi}_h$, and coefficients $\widehat{A}_h$, $\widehat{B}_h$, $\widehat{b}_h$, $\widehat{C}_h$, $\widehat{c}_h$, $\widehat{D}_h$, $\widehat{d}_h$ defined by

$$\begin{aligned}
\widehat{\varphi}_h(u_\tau^h) &= h\varphi(u_\tau^h), \quad \widehat{\psi}_h(v_\tau^h) = h\psi(v_\tau^h), \\
\widehat{A}_h &= I + hA, \\
\widehat{B}_h &= hB, \quad \widehat{b}_h = hb, \\
\widehat{C}_h &= hC, \quad \widehat{c}_h = hc, \\
\widehat{D}_h &= hD, \quad \widehat{d}_h = 0.
\end{aligned}$$

The dynamics in this problem are given by

$$\begin{aligned}
\widehat{x}_\tau^h &= \widehat{A}_h \widehat{x}_{\tau-1}^h + \widehat{B}_h u_\tau^h + \widehat{b}_h, \text{ for } \tau = 1, \ldots, T \\
\widehat{x}_0^h &= B_e u_0^h + b_e, \\
\widehat{y}_\tau^h &= \widehat{A}_h^* \widehat{y}_{\tau+1}^h + \widehat{C}_h^* v_\tau^h + \widehat{c}_h, \text{ for } \tau = T, \ldots, 1 \\
\widehat{y}_{T+1}^h &= C_e^* v_{T+1}^h + c_e.
\end{aligned}$$

Note that the trajectories associated with problem $(\widehat{\mathcal{S}}_h)$ are denoted by $(\widehat{x}_0^h, \ldots, \widehat{x}_T^h)$ and $(\widehat{y}_1^h, \ldots, \widehat{y}_{T+1}^h)$, whereas the trajectories for $(\widetilde{\mathcal{S}}_h)$ are indicated by tildes.

We shall also think of $(\widetilde{\mathcal{S}}_h)$ and $(\widehat{\mathcal{S}}_h)$ as problems on $(\mathcal{L}_k^2 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^2 \times \mathbf{R}^{le})$ by identifying $\mathbf{R}^{ke} \times \mathbf{R}^{kT}$ and $(\mathbf{R}^{lT} \times \mathbf{R}^{le})$ with the subspaces $\mathcal{U}_h$ and $\mathcal{V}_h$ given by

$$\mathcal{U}_h = \{(u, u_e) \in \mathcal{L}_k^2 \times \mathbf{R}^{ke} \mid u_t \text{ is constant a.e. on } [(\tau-1)h, \tau h) \text{ for } \tau = 1, \ldots, T\},$$

$$\mathcal{V}_h = \{(v, v_e) \in \mathcal{L}_l^2 \times \mathbf{R}^{le} \mid v_t \text{ is constant a.e. on } [(\tau-1)h, \tau h) \text{ for } \tau = 1, \ldots, T\}.$$

Thus the point $(u_0^h, u_1^h, \ldots, u_T^h)$ is identified with the pair (u, u_e), where $u_e = u_0^h$ and $u_t = u_\tau^h$ a.e. on $[(\tau-1)h, \tau h)$ for $\tau = 1, \ldots, T$. The norm on $\mathbf{R}^{ke} \times \mathbf{R}^{kT}$ is also given through the identification with $\mathcal{U}_h$:

$$\|(u_0^h, \ldots, u_T^h)\|_h = (h \sum_{\tau=1}^{T} |u_\tau^h|^2 + |u_0^h|^2)^{1/2}.$$

Here $|\cdot|$ denotes the Euclidean norm on $\mathbf{R}^k$ (or $\mathbf{R}^{ke}$). A similar formula gives the norm on $\mathbf{R}^{lT} \times \mathbf{R}^{le}$. The functionals $\widetilde{\mathcal{J}}_h$ are extended to $(\mathcal{L}_k^2 \times \mathbf{R}^{ke}) \times (\mathcal{L}_l^2 \times \mathbf{R}^{le})$ by taking $\widetilde{\mathcal{J}}_h(u, u_e; v, v_e)$ to be $-\infty$ if $(u, u_e) \in \mathcal{U}_h$ but $(v, v_e) \notin \mathcal{V}_h$, and to be ∞ if $(u, u_e) \notin \mathcal{U}_h$. We extend $\widehat{\mathcal{J}}_h$ in the same manner.

As one may expect, the saddle functions $\widetilde{\mathcal{J}}_h$ and $\widehat{\mathcal{J}}_h$ are closely related. An important aspect of this relationship is given by the following proposition.

Proposition 3.3.1. *There exists a real number $r > 0$ such that*

$$\begin{aligned}\widetilde{\mathcal{J}}_h(u,u_e;v,v_e) - r(\|(u,u_e)\|+1)(\|(v,v_e)\|+1)|S_h|/h \\ \leq \widehat{\mathcal{J}}_h(u,u_e;v,v_e) \\ \leq \widetilde{\mathcal{J}}_h(u,u_e;v,v_e) + r(\|(u,u_e)\|+1)(\|(v,v_e)\|+1)|S_h|/h\end{aligned} \tag{3.3.1}$$

for all $(u,u_e;v,v_e) \in (\mathcal{L}^2_k \times \mathbf{R}^{ke}) \times (\in \mathcal{L}^2_l \times \mathbf{R}^{le})$ and all $h = 1/T$ with $T \in \mathbf{N}$.

Remarks.

(i) It is easily verified that $|S_h|/h$ is increasing as a function of $h \in (1,\infty)$, and that $|S_h|/h \to 0$ as $h \downarrow 0$.

(ii) We are tacitly assuming here that the values of $\widetilde{\mathcal{J}}_h$ are defined using the same conventions regarding $\infty - \infty$ as are used in defining $\widehat{\mathcal{J}}_h$. If this were not the case, then we could simply replace (3.3.1) by similar inequalities involving $\overline{\mathrm{cl}}_2\,\widetilde{\mathcal{J}}_h$ and $\overline{\mathrm{cl}}_2\,\widehat{\mathcal{J}}_h$ (or, equivalently, $\underline{\mathrm{cl}}_1\,\widetilde{\mathcal{J}}_h$ and $\underline{\mathrm{cl}}_1\,\widehat{\mathcal{J}}_h$).

(iii) Note that (3.3.1) is equivalent to the inequalities given by reversing the rôles of $\widetilde{\mathcal{J}}_h$ and $\widehat{\mathcal{J}}_h$.

Proof. Clearly, $\widetilde{\mathcal{J}}_h(u,u_e;v,v_e)$ and $\widehat{\mathcal{J}}_h(u,u_e;v,v_e)$ agree whenever $(u,u_c) \notin \mathcal{U}_h$ or $(v,v_e) \notin \mathcal{V}_h$. Similarly, if the control pair $(u,u_e;v,v_e) \in \mathcal{U}_h \times \mathcal{V}_h$ is identified with $(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)$, then $\widetilde{\mathcal{J}}_h(u,u_e;v,v_e) = \widehat{\mathcal{J}}_h(u,u_e;v,v_e)$ whenever $\max_\tau\{\varphi(u_\tau^h),\psi(v_\tau^h)\} = \infty$ or whenever $\max\{\varphi_e(u_0^h),\psi_e(v_{T+1}^h)\} = \infty$. In these cases, the inequalities in (3.3.1) are therefore satisfied trivially, regardless of h or r.

Now suppose that $\varphi(u_\tau^h)$ and $\varphi(v_\tau^h)$ are finite for each $\tau = 1,\ldots,T$, and that $\varphi_e(u_0^h)$ and $\psi(v_{T+1}^h)$ are finite. In this case we may rewrite (3.3.1) as

$$\begin{aligned}|\widehat{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h) - \widetilde{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)| \\ \leq r(\|(u,u_e)\|+1)(\|(v,v_e)\|+1)|S_h|/h\end{aligned} \tag{3.3.2}$$

since both $\widehat{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)$ and $\widetilde{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)$ are finite. By the definitions of $\widehat{\varphi}_h,\widehat{\psi}_h,\widetilde{\varphi}_h,\widetilde{\psi}_h$, we have

$$\begin{aligned}&|\widehat{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h) - \widetilde{\mathcal{J}}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)| \\ &\leq \left|(B^*S_h^*c)\cdot(\sum_{\tau=1}^T u_\tau^h) + (CS_hb)\cdot(\sum_{\tau=1}^T v_\tau^h) + \sum_{\tau=1}^T v_\tau^h\cdot(CS_hB)u_\tau^h + \sum_{\tau=1}^T(c\cdot S_hb)\right|\end{aligned}$$

$$
\begin{aligned}
&+ |\widetilde{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h)| \\
&\le [(|B||c|)\|(u,u_e)\| + (|C||b|)\|(v,v_e)\| + (|C||B|)\|(u,u_e)\|\|(v,v_e)\| + |b||c|]\,(|S_h|/h) \\
&\quad + |\widetilde{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h)| \\
&\le r_1(\|(u,u_e)\|+1)(\|(v,v_e)\|+1)|S_h|/h \\
&\quad + |\widetilde{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h)|,
\end{aligned}
$$

for $r_1 = \max\{|B|,|b|\}\cdot\max\{|C|,|c|\}$. Thus, to prove (3.3.2) it suffices to find $r_2 > 0$ so that

$$
\begin{aligned}
|\widetilde{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h)| \qquad (3.3.3)\\
\le r_2(\|(u,u_e)\|+1)(\|(v,v_e)\|+1)|S_h|/h.
\end{aligned}
$$

Observe that $\widetilde{\gamma}_h = \widehat{\gamma}_h$ if $A = 0$, so we need only consider the case where $A \neq 0$.

First we note that for each $\tau = 1,\dots,T$ we have

$$
\widetilde{x}_\tau^h = \widetilde{A}_h^\tau \widetilde{x}_0^h + \sum_{\tau'=1}^{\tau} \widetilde{A}_h^{\tau-\tau'}(\widetilde{B}_h u_{\tau'}^h + \widetilde{b}_h)
$$

and

$$
\widehat{x}_\tau^h = \widehat{A}_h^\tau \widehat{x}_0^h + \sum_{\tau'=1}^{\tau} \widehat{A}_h^{\tau-\tau'}(\widehat{B}_h u_{\tau'}^h + \widehat{b}_h).
$$

Substituting these into the formulas for $\widetilde{\gamma}_h$ and $\widehat{\gamma}_h$, and rearranging terms, yields

$$
\begin{aligned}
&\widetilde{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\dots,u_T^h;v_1^h,\dots,v_{T+1}^h) \\
&\quad = (B_e u_0^h + b_e)\cdot\left\{\sum_{\tau=1}^{T}\left[(\widetilde{A}_h^{\tau-1})^*(\widetilde{C}_h^* v_\tau^h + \widetilde{c}_h) - (\widehat{A}_h^{\tau-1})^*(\widehat{C}_h^* v_\tau^h + \widehat{c}_h)\right]\right\} \\
&\qquad + (C_e^* v_{T+1}^h + c_e)\cdot\left\{\sum_{\tau=1}^{T}\left[(\widetilde{A}_h^{\tau-1})(\widetilde{B}_h u_\tau^h + \widetilde{b}_h) - (\widehat{A}_h^{\tau-1})(\widehat{B}_h u_\tau^h + \widehat{b}_h)\right]\right\} \\
&\qquad + (B_e u_0^h + b_e)\cdot(\widetilde{A}_h^T - \widehat{A}_h^T)^*(C_e^* v_{T+1}^h + c_e) \\
&\qquad + \sum_{\tau=1}^{T}\sum_{\tau'=1}^{\tau-1}\left[\begin{array}{l}(\widetilde{C}_h^* v_\tau^h + \widetilde{c}_h)\cdot(\widetilde{A}_h^{\tau-\tau'-1})(\widetilde{B}_h u_{\tau'}^h + \widetilde{b}_h) \\ \quad -(\widehat{C}_h^* v_\tau^h + \widehat{c}_h)\cdot(\widehat{A}_h^{\tau-\tau'-1})(\widehat{B}_h u_{\tau'}^h + \widehat{b}_h)\end{array}\right] \\
&\quad = (B_e u_0^h + b_e)\cdot\left[\sum_{\tau=1}^{T}(M_h\widetilde{A}_h^{\tau-1} - h\widehat{A}_h^{\tau-1})^*(C_h^* v_\tau^h + c_h)\right]
\end{aligned}
$$

$$+ (C_e^* v_{T+1}^h + c_e) \cdot \left[\sum_{\tau=1}^{T} (M_h \widetilde{A}_h^{\tau-1} - h\widehat{A}_h^{\tau-1})(B_h u_\tau^h + b_h) \right]$$
$$+ (B_e u_0^h + b_e) \cdot (\widetilde{A}_h^T - \widehat{A}_h^T)^* (C_e^* v_{T+1}^h + c_e)$$
$$+ \sum_{\tau=1}^{T} \sum_{\tau'=1}^{\tau-1} \left[(C_h^* v_\tau^h + c_h) \cdot (M_h^2 \widetilde{A}_h^{\tau-\tau'-1} - h^2 \widehat{A}_h^{\tau-\tau'-1})(B_h u_{\tau'}^h + b_h) \right].$$

Thus we obtain the bound

$$\begin{aligned}
&|\widetilde{\gamma}_h(u_0^h, \ldots, u_T^h; v_1^h, \ldots, v_{T+1}^h) - \widehat{\gamma}_h(u_0^h, \ldots, u_T^h; v_1^h, \ldots, v_{T+1}^h)| \qquad (3.3.4)\\
&\le [(|B_e| + |B|)\|(u, u_e)\| + |b| + |b_e|][(|C_e| + |C|)\|(v, v_e)\| + |c| + |c_e|]\\
&\qquad \cdot \frac{1}{h} \max_{\tau=1,\ldots,T} |M_h \widetilde{A}_h^{\tau-1} - h\widehat{A}_h^{\tau-1}|\\
&\quad + (|B_e|\|(u, u_e)\| + |b_e|)(|C_e|\|(v, v_e)\| + |c_e|)|\widetilde{A}_h^T - \widehat{A}_h^T|\\
&\quad + (|C|\|(v, v_e)\| + |c|)(|B|\|(u, u_e)\| + |b|)\\
&\qquad \cdot \frac{1}{h^2} \max_{\substack{\tau=1,\ldots,T\\ \tau'=1,\ldots,\tau-1}} |M_h^2 \widetilde{A}_h^{\tau-\tau'-1} - h^2 \widehat{A}_h^{\tau-\tau'-1}|\\
&\le \bar{r}_2(\|(u, u_e)\| + 1)(\|(v, v_e)\| + 1)\\
&\quad \cdot \left[\begin{array}{l} |\widetilde{A}_h^T - \widehat{A}_h^T| + (1/h) \max\limits_{\tau=1,\ldots,T} |M_h \widetilde{A}_h^{\tau-1} - h\widehat{A}_h^{\tau-1}| \\ \qquad + (1/h^2) \max\limits_{\substack{\tau=1,\ldots,T\\ \tau'=1,\ldots,\tau-1}} |M_h^2 \widetilde{A}_h^{\tau-\tau'-1} - h^2 \widehat{A}_h^{\tau-\tau'-1}| \end{array} \right],
\end{aligned}$$

for $\bar{r}_2 = \max\{|B_e| + |B|, |b_e| + |b|\} \cdot \max\{|C_e| + |C|, |c_e| + |c|\}$. Now for $m = 1, \ldots, T$ we have

$$\begin{aligned}
|\widetilde{A}_h^m - \widehat{A}_h^m| &\le |\widetilde{A}_h - \widehat{A}_h| \left[\sum_{j=0}^{m-1} |\widetilde{A}|^j |\widehat{A}|^{m-j-1} \right] \qquad (3.3.5)\\
&= |S_h| \left[\sum_{j=0}^{m-1} |e^{hA}|^j |I + hA|^{m-j-1} \right]\\
&\le |S_h| \left[\sum_{j=0}^{m-1} e^{jh|A|} e^{(m-j-1)h|A|} \right] \le \frac{|S_h|}{h} e^{|A|}.
\end{aligned}$$

This in turn gives

$$\begin{aligned}|M_h\widetilde{A}_h^m - h\widehat{A}_h^m| &\le \max\{|M_h|,h\}|\widetilde{A}_h^m - \widehat{A}_h^m| + \max\{|\widetilde{A}_h|^m,|\widehat{A}_h|^m\}|M_h - hI| \\ &\le \max\{|M_h|,h\}\frac{|S_h|}{h}e^{|A|} + \frac{|S_h|}{h}e^{mh|A|} \\ &\le \frac{|S_h|}{h}e^{|A|}\left(\frac{e^{|A|}-1}{|A|}+1\right), \end{aligned} \tag{3.3.6}$$

where we have also used the bound

$$(1/h)\max\{|M_h|,h\} \le (1/h)\frac{1}{|A|}(e^{h|A|}-1) \le \frac{e^{|A|}-1}{|A|}. \tag{3.3.7}$$

Next we have

$$\begin{aligned}|M_h^2\widetilde{A}_h^m - h^2\widehat{A}_h^m| &\le \max\{|M_h|^2,h^2\}|\widetilde{A}_h^m - \widehat{A}_h^m| + \max\{|\widetilde{A}_h|^m,|\widehat{A}_h|^m\}|M_h^2 - h^2I| \\ &\le \left[\left(\frac{e^{h|A|}-1}{|A|}\right)^2\frac{|S_h|}{h}e^{|A|}\right] + \left[e^{|A|}\max\{|M_h|,h\}|S_h|\right] \\ &\le |S_h|e^{|A|}\left(\frac{e^{h|A|}-1}{|A|}\right)\left(\frac{e^{|A|}-1}{|A|}+1\right). \end{aligned} \tag{3.3.8}$$

Finally, by applying inequalities (3.3.5), (3.3.6) and (3.3.8) to the estimate in (3.3.4), we obtain

$$\begin{aligned}&|\widetilde{\gamma}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h) - \widehat{\gamma}_h(u_0^h,\ldots,u_T^h;v_1^h,\ldots,v_{T+1}^h)| \\ &\quad\le \bar{r}_2(\|(u,u_e)\|+1)(\|(v,v_e)\|+1) \\ &\qquad\cdot\frac{|S_h|}{h}e^{|A|}\left[1+\left(\frac{e^{|A|}-1}{|A|}+1\right)\left(1+\frac{e^{h|A|}-1}{|hA|}\right)\right] \\ &\quad\le \bar{r}_2e^{|A|}\left[1+\left(\frac{e^{|A|}-1}{|A|}\right)^2\right]\frac{|S_h|}{h}(\|(u,u_e)\|+1)(\|(v,v_e)\|+1), \end{aligned}$$

where we have again used (3.3.7). Thus (3.3.3) is satisfied for

$$r_2 = \bar{r}_2e^{|A|}\left[1+\left(\frac{e^{|A|}-1}{|A|}\right)^2\right],$$

which completes the proof. □

The proposition just proved tells us, in particular, that the nets $\{\widetilde{\mathcal{J}}_h\}$ and $\{\widehat{\mathcal{J}}_h\}$ are "uniformly cofinal" on bounded sets: for a fixed $\rho > 0$ and $r_\rho = r(\rho+1)^2$, we have

$$\widetilde{\mathcal{J}}_h(u,u_e;v,v_e) - r_\rho|S_h|/h \le \widehat{\mathcal{J}}_h(u,u_e;v,v_e) \le \widetilde{\mathcal{J}}_h(u,u_e;v,v_e) + r_\rho|S_h|/h$$

whenever $\|(u,u_e)\| \le \rho$ and $\|(v,v_e)\| \le \rho$. The uniformity in (3.3.1) actually guarantees that the problems $(\widetilde{\mathcal{S}}_h)$ and $(\widehat{\mathcal{S}}_h)$ are close from a variational standpoint. This is illustrated by the next theorem. Combined with the results of sections 1 and 2, it gives us sufficient conditions for the problems $\{(\widehat{\mathcal{S}}_h)\}$ to be variational approximations to $(\mathcal{S})$.

Theorem 3.3.2. *Suppose $\{T_m\}$ is an increasing sequence of positive integers and let $h_m = 1/T_m$ and consider a proper closed saddle function $\mathcal{K}$ defined on $(\mathcal{L}^2_k \times \mathbf{R}^{ke}) \times (\mathcal{L}^2_l \times \mathbf{R}^{le})$. The sequence $\{\widehat{\mathcal{J}}_{hm}\}$ Mosco epi/hypo-converges to $\mathcal{K}$ if and only if $\{\widetilde{\mathcal{J}}_{hm}\}$ Mosco epi/hypo-converges to $\mathcal{K}$.*

Proof. By Corollary 2.4.2 and the definitions of the sequential epi/hypo-upper and lower limits, a sequence $\{\mathcal{K}_m\}$ Mosco epi/hypo-converges to $\mathcal{K}$ if and only if the following conditions are satisfied:

(a) There exists a weakly convergent sequence $\{(u^m,u_e^m)\}$ and a real number $\bar{r} \ge 0$ such that $\mathcal{K}_m(u^m,u_e^m;v,v_e) \le \bar{r}(\|(v,v_e)\|+1)$ for all (v,v_e) and all m.

(b) For all (u,u_e) and $\{(u^m,u_e^m)\}$ converging weakly to (u,u_e), one has

$$\text{seq-h}_\text{s}\text{-li}\,\mathcal{K}_m(u^m,u_e^m;\cdot,\cdot) \ge \mathcal{K}(u,u_e;\cdot,\cdot).$$

(c) For all (v,v_e) and $\{(v^m,v_e^m)\}$ converging weakly to (v,v_e), one has

$$\text{seq-e}_\text{s}\text{-ls}\,\mathcal{K}_m(\cdot,\cdot;v^m,v_e^m) \ge \mathcal{K}(\cdot,\cdot;v,v_e).$$

The desired result then follows by simply combining Proposition 3.3.1 with the fact that weakly convergent sequences are norm-bounded. For example, by (3.3.1) the sequence $\{\widehat{\mathcal{J}}_{hm}\}$ satisfies condition (a) for some choice of $\{(u^m,u_e^m)\}$ and some $\bar{r} \ge 0$ if and only if $\{\widetilde{\mathcal{J}}_{hm}\}$ satisfies (a) with the same choice of $\{(u^m,u_e^m)\}$ and for some $\bar{r}' \ge 0$ (possibly different from $\bar{r}$). □

An analogous result to Theorem 3.3.2 is possible in the case where the data elements A, B, b, C, c, D are allowed to vary Lipschitz continuously with t. In this case, the term $|S_h|/h$ in the estimate (3.3.1) is replaced by a more complicated expression in h; the argument needed is essentially the same as that given for Proposition 3.3.1, but requires far more space to write out.

Chapter 4. Applications in Multistage Stochastic Programming

1. Description of the Model

In this section we describe a model in multistage stochastic programming. Our development of the duality theory mimics that of Rockafellar [27] in his treatment of the continuous-time optimal control models that were discussed in the previous chapter. In order to simplify the presentation, we restrict our attention here to a *piecewise linear-quadratic* formulation. Rockafellar and Wets have developed a duality theory for this problem in the very special case of (finite) discrete probability spaces [35]. Their result on strong duality is based on the fact that the problem is equivalent to a (very large) quadratic programming problem in that case. In this section, we show how to extend the duality theory to arbitrary probability spaces. For simplicity, we set the problem in $\mathcal{L}^2$-space. The problem we shall be working with is

$$(\mathcal{P}) \qquad \text{minimize the functional} \quad F(u) \quad \text{over the control space} \quad \mathcal{U},$$

where

$$F(u) = \mathbf{E}\Bigg\{\sum_{\tau=1}^{T}\Big[p_\tau \cdot u_\tau + \tfrac{1}{2}u_\tau \cdot P_\tau u_\tau - c_\tau \cdot x_{\tau-1} + \rho_{V_\tau,Q_\tau}(q_\tau - \mathbf{E}^{\mathcal{G}_\tau}[D_\tau u_\tau + C_\tau x_{\tau-1}])\Big] + p_0 \cdot u_0 + \tfrac{1}{2}u_0 \cdot P_0 u_0 - c_{T+1} \cdot x_T + \rho_{V_{T+1},Q_{T+1}}(q_{T+1} - \mathbf{E}^{\mathcal{G}_{T+1}}[C_{T+1}x_T])\Bigg\},$$

$$\mathcal{U} = \mathcal{U}_0 \times \cdots \times \mathcal{U}_T = \prod_{\tau=0}^{T}\{u_\tau(\cdot) \in \mathcal{L}^2_{k_\tau} | u_\tau \text{ is } \mathcal{F}_\tau\text{-measurable}, u_\tau \in U_\tau \text{ a.s.}\},$$

and the dynamics are given by

$$x_\tau = A_\tau x_{\tau-1} + B_\tau u_\tau + b_\tau \text{ a.s. for } \tau = 1, \ldots, T, \quad x_0 = B_0 u_0 + b_0 \text{ a.s.}$$

The symbol $\mathbf{E}^{\mathcal{G}}$ indicates conditional expectation with respect to the field $\mathcal{G}$.

For each $\tau = 0, \ldots, T$, we assume that p_τ, P_τ and U_τ are $\mathcal{F}_\tau$-measurable, where $\mathcal{F}_\tau$ is a sub-σ-algebra of $\mathcal{E}$, and $(\Omega, \mathcal{E}, \mu)$ is the underlying probability space.

Similarly, for $\tau = 1, \ldots, T+1$, we assume that q_τ, Q_τ and V_τ are $\mathcal{G}_\tau$-measurable, where $\mathcal{G}_\tau$ is a sub-σ-algebra of $\mathcal{E}$. The matrix-valued functions A_τ, B_τ, b_τ, C_τ, c_τ, D_τ are all assumed to be $\mathcal{E}$-measurable. We require the random coefficients P_τ, Q_τ, A_τ, B_τ, C_τ, D_τ to be essentially bounded, but allow p_τ, q_τ, b_τ and c_τ to be *square summable*. In addition, for each $\omega \in \Omega$, $P_\tau(\omega)$ and $Q_\tau(\omega)$ are assumed to be symmetric and positive semidefinite, while $V_\tau(\omega)$ and $U_\tau(\omega)$ are assumed to be polyhedral convex sets. Further, we shall posit $\mathcal{L}^2(\mathcal{G}_\tau)$- and $\mathcal{L}^2(\mathcal{F}_\tau)$-selections for V_τ and U_τ respectively.

The expression $\rho_{V_\tau,Q_\tau}(s)$ represents a piecewise linear-quadratic function on $\mathbf{R}^{l_\tau}$, given explicitly by

$$\rho_{V_\tau(\omega),Q_\tau(\omega)}(s) = \sup_{v \in V_\tau(\omega)} \{s \cdot v - \frac{1}{2} v \cdot Q_\tau(\omega) v\}.$$

For each $\omega \in \Omega$ and each τ, this function is lower semicontinuous and convex. Its effective domain

$$L_\tau(\omega) = \{s \in \mathbf{R}^{l_\tau} | \rho_{V_\tau(\omega),Q_\tau(\omega)}(s) < \infty\}$$

is a nonempty convex polyhedron that can be decomposed into finitely many polyhedral convex sets, on each of which $\rho_{V_\tau(\omega),Q_\tau(\omega)}$ is quadratic or linear.

Proposition 4.1.1. *The map $(\omega, s) \mapsto \rho_{V_\tau(\omega),Q_\tau(\omega)}(s)$ is a proper $\mathcal{G}_\tau$-normal integrand. If σ_ω is $\mathcal{G}_\tau$-measurable then $\omega \mapsto \partial\rho_{V_\tau(\omega),Q_\tau(\omega)}(\sigma_\omega)$ is $\mathcal{G}_\tau$-measurable.*

Proof. Define $E_\tau(\omega) = \{(s, \alpha) \in \mathbf{R}^{l_\tau} \times \mathbf{R} | \alpha \geq \rho_{V_\tau(\omega),Q_\tau(\omega)}(s)\}$. We need to show that E_τ is closed-valued and $\mathcal{G}_\tau$-measurable. Since $\rho_{V,Q}$ is the conjugate of $v \mapsto \frac{1}{2} v \cdot Qv + \delta_V(v)$, it is lower semicontinuous and hence has closed epigraph. Thus E_τ is closed-valued. Define $\varphi_\tau : \Omega \times \mathbf{R}^{l_\tau} \to \overline{\mathbf{R}}$ by

$$\varphi_\tau(\omega, v) = \frac{1}{2} v \cdot Q_\tau(\omega) v + \delta_{V_\tau(\omega)}(v).$$

Then φ_τ is a proper normal integrand by Propositions 2H and 2M of [26]. Proposition 2S in [26] gives the $\mathcal{G}_\tau$-normality of $(\omega, u) \mapsto \rho_{V_\tau(\omega),Q_\tau(\omega)}(u)$ (as well as properness). The statement concerning $\rho_{V_\tau(\omega),Q_\tau(\omega)}(\sigma_\omega)$ follows from Corollary 2X of [26]. □

Proposition 4.1.2. *The set $\mathcal{U}$ is a closed convex subset of $\mathcal{L}^2_{k_0} \times \cdots \times \mathcal{L}^2_{k_T}$ and F is well defined, lower semicontinuous and convex, with values that are finite or ∞.*

Proof. Note that $\mathcal{U}$ is nonempty by the assumed existence of an $\mathcal{L}^2(\mathcal{F}_\tau)$-selection for U_τ. The convexity and closedness follows from that of each $U_\tau(\omega)$. The mapping $u \mapsto x$ from $\prod_{\tau=0}^T \mathcal{L}^2_{k_\tau}$ into $\prod_{\tau=0}^T \mathcal{L}^2_{n_\tau}$ is affine and continuous, so

$$u \mapsto \mathbf{E}\{\sum_{\tau=0}^{T}(p_\tau \cdot u_\tau - c_\tau \cdot x_{\tau-1})\}$$

gives a continuous affine functional of u. Likewise, the mapping of $\prod_{\tau=0}^T \mathcal{L}^2_{k_\tau}$ into $\prod_{\tau=0}^T \mathcal{L}^2_{l_\tau}$ given by

$$\begin{aligned} s_\tau &= q - \mathbf{E}^{\mathcal{G}_\tau}\{C_\tau x_{\tau-1} + D_\tau u_\tau\},\ \tau = 1,\ldots,T \\ s_{T+1} &= q_{T+1} - \mathbf{E}^{\mathcal{G}_{T+1}}\{C_{T+1}x_T\} \end{aligned}$$

is affine and continuous. Also, it is clear that $\mathbf{E}\{\frac{1}{2}u_\tau \cdot P_\tau u_\tau\}$ defines a continuous convex functional on $\mathcal{L}^2_{k_\tau}$. It remains to show that $I_\tau(s_\tau) = \mathbf{E}\{\rho_{V_\tau,Q_\tau}(s_\tau)\}$ is a well-defined, lower semicontinuous convex functional on $\mathcal{L}^2_{l_\tau}$. Define

$$\varphi_\tau(v_\tau) = \tfrac{1}{2}v_\tau \cdot Q_\tau v_\tau + \delta_{V_\tau}(v_\tau)$$

on $\mathbf{R}^{l_\tau}$, and take $\bar{v}_\tau$ to be an $\mathcal{L}^2(\mathcal{G}_\tau)$-selection of V_τ. We see that $\varphi_\tau = (\rho_{V_\tau,Q_\tau})^*$ and $\mathbf{E}\{\varphi_\tau(\bar{v}_\tau)\} < +\infty$, so I_τ is lower semicontinuous and nowhere $-\infty$ by Corollary 3D of [26], since ρ_{V_τ,Q_τ} is $\mathcal{G}_\tau$-normal. □

The dual problem to $(\mathcal{P})$ is:

$(\mathcal{Q})$ maximize the functional $G(v)$ over the control space $\mathcal{V}$.

Here we have

$$\begin{aligned} G(v) = \mathbf{E}\Bigg\{\sum_{\tau=1}^{T}\Big[& q_\tau \cdot v_\tau - \tfrac{1}{2}v_\tau \cdot Q_\tau v_\tau - b_\tau \cdot y_{\tau+1} \\ & - \rho_{U_\tau,P_\tau}(\mathbf{E}^{\mathcal{F}_\tau}[D_\tau^* v_\tau + B_\tau^* y_{\tau+1}]p_\tau)\Big] \\ & + q_{T+1} \cdot v_{T+1} + \tfrac{1}{2}v_{T+1} \cdot Q_{T+1}v_{T+1} - b_0 \cdot y_1 \\ & - \rho_{U_0,P_0}(\mathbf{E}^{\mathcal{F}_0}[B_0^* y_1] - p_0)\Bigg\}, \end{aligned}$$

$$\mathcal{V} = \mathcal{V}^1 \times \cdots \times \mathcal{V}^{T+1} = \prod_{\tau=1}^{T+1}\{v_\tau(\cdot) \in \mathcal{L}^2_{l_\tau} | v_\tau \text{ is } \mathcal{G}_\tau\text{-measurable}, v_\tau \in V_\tau \text{ a.s.}\},$$

and the dual dynamics are given by

$$y_\tau = A_\tau^* y_{\tau+1} + C_\tau^* v_\tau + c_\tau \text{ a.s. for } \tau = T, \ldots, 1, \quad y_{T+1} = C_{T+1}^* v_{T+1} + c_{T+1} \text{ a.s.}$$

By the symmetry between $(\mathcal{P})$ and $(\mathcal{Q})$, Proposition 4.1.2 gives us

Corollary 4.1.3. *The set $\mathcal{V}$ is a closed convex subset of $\mathcal{L}^2_{l_1} \times \cdots \times \mathcal{L}^2_{l_{T+1}}$ and G is well defined, upper semicontinuous and concave, with values that are finite or $-\infty$.*

The duality between $(\mathcal{P})$ and $(\mathcal{Q})$ may be demonstrated through the introduction of an appropriate saddle function. We define this by

$$\mathcal{J}(u,v) = \begin{cases} \mathbf{E}\{J(u,v) - \gamma(u,v)\} & \text{if } u \in \mathcal{U}, v \in \mathcal{V}, \\ -\infty & \text{if } u \in \mathcal{U}, v \notin \mathcal{V}, \\ \infty & \text{if } u \notin \mathcal{U}, \end{cases}$$

where

$$\begin{aligned} J(u,v) = \sum_{\tau=1}^{T} [p_\tau \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau - v_\tau \cdot D_\tau u_\tau + q_\tau \cdot v_\tau - \tfrac{1}{2} v_\tau \cdot Q_\tau v_\tau] \\ + p_0 \cdot u_0 + \tfrac{1}{2} u_0 \cdot P_0 u_0 + q_{T+1} \cdot v_{T+1} - \tfrac{1}{2} v_{T+1} \cdot Q_{T+1} v_{T+1} \end{aligned}$$

and

$$\gamma(u,v) = \sum_{\tau=0}^{T} y_{\tau+1} \cdot (B_\tau u_\tau + b_\tau) = \sum_{\tau=1}^{T+1} x_{\tau-1} \cdot (C_\tau^* v_\tau + c_\tau).$$

It is easy to see that $\mathcal{J}$ is a closed proper saddle function with $\mathcal{J} = \overline{\mathcal{J}}$ and effective domain given by $(u,v) \in \mathcal{U} \times \mathcal{V}$.

Proposition 4.1.4. *The problems $(\mathcal{P})$ and $(\mathcal{Q})$ are the primal and dual problems associated with the problem of finding a saddle point of $\mathcal{J}$. Thus we have*

$$F(u) = \sup_{v \in \mathcal{V}} \mathcal{J}(u,v) \quad \text{and} \quad G(v) = \inf_{u \in \mathcal{U}} \mathcal{J}(u,v).$$

Proof. Fix $u \in \mathcal{U}$ and use the second expression in the definition of γ, so that

$$\begin{aligned} \mathcal{J}(u,v) = \mathbf{E}\{\sum_{\tau=0}^{T} [p_\tau \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau - c_\tau \cdot x_{\tau-1}]\} \\ + \mathbf{E}\{\sum_{\tau=1}^{T+1} [(q_\tau - D_\tau u_\tau - C_\tau x_{\tau-1}) \cdot v_\tau - \tfrac{1}{2} v_\tau \cdot Q_\tau v_\tau]\}. \end{aligned}$$

By the definition of ρ_{V_τ,Q_τ}, it is clear that $F(u) \geq \mathcal{J}(u,v)$ for all $u \in \mathcal{U}, v \in \mathcal{V}$. It suffices then to show that for arbitrary $s_\tau \in \mathcal{L}^2_{l_\tau}(\mathcal{G}_\tau)$ one has

$$\sup_{\substack{v_\tau(\cdot) \in \mathcal{L}^2(\mathcal{G}_\tau) \\ v_\tau \in V_\tau \text{ a.s.}}} \mathbf{E}\{v_\tau \cdot s_\tau - \tfrac{1}{2} v_\tau \cdot Q_\tau v_\tau\} = \mathbf{E}\{\rho_{V_\tau,Q_\tau}(s_\tau)\}.$$

This equation may be rewritten as

$$\sup_{v_\tau(\cdot) \in \mathcal{L}^2(\mathcal{G}_\tau)} \mathbf{E}\{v_\tau \cdot s_\tau - \varphi_\tau(v_\tau)\} = \mathbf{E}\{\varphi^*_\tau(s_\tau)\}$$

where $\varphi_\tau(v_\tau) = \frac{1}{2} v_\tau \cdot Q_\tau v_\tau + \delta_{V_\tau}(v_\tau)$. But this new equation holds by Theorem 3C of [26] since φ_τ is a normal integrand and we have assumed that there is an $\mathcal{L}^2(\mathcal{G}_\tau)$-selection $\bar{v}_\tau$ for V_τ. That $G(v) = \inf_{u \in \mathcal{U}} \mathcal{J}(u,v)$ follows by symmetry. □

The following is a direct consequence of Proposition 4.1.4.

Proposition 4.1.5 (Weak Duality). *It is always true that* $\inf(\mathcal{P}) \geq \sup(\mathcal{Q})$. *Moreover,* $(\bar{u}, \bar{v})$ *is a saddle point of* $\mathcal{J}$ *if and only if* $\bar{u}$ *solves* $(\mathcal{P})$, $\bar{v}$ *solves* $(\mathcal{Q})$ *and*

$$\min(\mathcal{P}) = \max(\mathcal{Q}) \quad (\textit{finite}).$$

In order to obtain strong duality results, i.e. the existence of saddle points, we need to introduce some sort of "finiteness" conditions on the integrands

$$\rho_{U_\tau(\cdot),P_\tau(\cdot)}(\cdot) \quad \text{and} \quad \rho_{V_\tau(\cdot),Q_\tau(\cdot)}(\cdot).$$

Primal Finiteness Condition. *There exist* $\alpha \geq 0$ *and* $a \in \mathcal{L}^1_1$ *such that, for almost every* ω,

$$\rho_{V_\tau(\omega),Q_\tau(\omega)}(s) \leq \alpha |s|^2 + a(\omega) \; \forall s \in \mathbf{R}^{l_\tau}, \; \forall \tau = 1, \ldots, T+1.$$

Equivalent to the primal finiteness condition is the statement that there exist $\alpha > 0$ and $a \in \mathcal{L}^1_1$ such that, for almost every ω and for each $\tau = 1, \ldots, T+1$,

$$s \cdot Q_\tau(\omega) s \geq \alpha |s|^2 - a(\omega) \; \forall s \in V_\tau(\omega).$$

The condition is satisfied in the case where $V_\tau(\omega) \subset \beta(\omega) \mathbf{B}_{l_\tau}$ for almost every ω and for each $\tau = 1, \ldots, T+1$, where β is some function in $\mathcal{L}^2_1$ and $\mathbf{B}_m$ is the unit ball in $\mathbf{R}^m$. It would also be satisfied in the case where the eigenvalues of $Q_\tau(\omega)$ are all bounded away from 0 uniformly in ω, for all τ. We also state the

Dual Finiteness Condition. *There exist* $\alpha \geq 0$ *and* $a \in \mathcal{L}_1^1$ *such that, for almost every* ω *and for each* $\tau = 0, \ldots, T$,

$$\rho_{U_\tau(\omega), P_\tau(\omega)}(r) \leq \alpha |r|^2 + a(\omega) \; \forall r \in \mathbf{R}^{k_\tau}.$$

Proposition 4.1.6 (Strong Duality). *If the primal finiteness condition is satisfied, then*

$$\inf(\mathcal{P}) = \max(\mathcal{Q}) < \infty;$$

likewise, if the dual finiteness condition is satisfied, then

$$\min(\mathcal{P}) = \sup(\mathcal{Q}) > -\infty.$$

Thus, if both conditions are satisfied, then solutions exist to both $(\mathcal{P})$ *and* $(\mathcal{Q})$, *and*

$$\min(\mathcal{P}) = \max(\mathcal{Q}) \quad \textit{finite.}$$

Proof. Suppose the dual finiteness condition is satisfied. We shall first work with

$$\underline{\mathcal{J}}(u, v) = \begin{cases} \mathbf{E}\{J(u,v) - \gamma(u,v)\} & \text{if } u \in \mathcal{U}, v \in \mathcal{V}, \\ \infty & \text{if } u \notin \mathcal{U}, v \in \mathcal{V}, \\ -\infty & \text{if } v \notin \mathcal{V} \end{cases}$$

Fix $\bar{v} \in \mathcal{V}$. To obtain $\min(\mathcal{P}) = \sup(\mathcal{Q}) > -\infty$, it suffices, by a minimax theorem of Moreau [17], to show that $\underline{\mathcal{J}}(\cdot, \bar{v})$ is weakly inf-compact. We have

$$\begin{aligned} \underline{\mathcal{J}}(u, \bar{v}) &= \mathbf{E}\{\sum_{\tau=1}^{T+1} [q_\tau \cdot \bar{v}_\tau - \tfrac{1}{2}\bar{v}_\tau \cdot Q_\tau \bar{v}_\tau - b_\tau \cdot \bar{y}_{\tau+1}]\} \\ &\quad + \mathbf{E}\{\sum_{\tau=0}^{T} [(p_\tau - D_\tau^* \bar{v}_\tau - B_\tau^* \bar{y}_{\tau+1}) \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau]\} \\ &= \text{ constant } + \mathbf{E}\{\sum_{\tau=0}^{T} [\tfrac{1}{2} u_\tau \cdot P_\tau u_\tau - \bar{r}_\tau \cdot u_\tau]\}, \end{aligned}$$

where $\bar{r}_\tau = D_\tau^* \bar{v}_\tau + B_\tau^* \bar{y}_{\tau+1} - p_\tau \in \mathcal{L}_{k_\tau}^2$ and $\bar{y}$ is the trajectory associated with $\bar{v}$. Thus, it suffices to show that, for all $\beta \in \mathbf{R}$, the set

$$\left\{ u \in \mathcal{U} \;\middle|\; \mathbf{E}[\sum_{\tau=0}^{T} (\tfrac{1}{2} u_\tau \cdot P_\tau u_\tau - \bar{r}_\tau \cdot u_\tau)] \leq \beta \right\}$$

is weakly compact. By assumption, $\mathcal{U}_\tau$ is nonempty so $u \mapsto \mathbf{E}[\frac{1}{2}u_\tau \cdot P_\tau u_\tau + \delta_{U_\tau}(u_\tau)]$ is finite for some $u_\tau \in \mathcal{L}^2_{k_\tau}(\mathcal{F}_\tau)$. This fact, combined with the dual finiteness condition, allows us to apply Theorem 3K of [26] to obtain the desired weak compactness. The rest of the theorem follows by symmetry (where we use $\overline{\mathcal{J}}$). □

In the next section, we shall show how the problem of finding a saddle point of $\mathcal{J}$ relative to $\mathcal{U} \times \mathcal{V}$ can be approximated by partitioning the probability space.

2. Approximation via Partitioning

In §1 we introduced a piecewise linear-quadratic model in multistage stochastic programming and developed some duality results based on a saddle function representation. In this section, we make further use of this representation in analyzing a particular method of approximation. This approximation leads to a new problem of similar structure, but of lower dimensionality. The problem, as we will work with it here, is that of finding a saddle point for $\mathcal{J}(u,v)$ relative to $\mathcal{U} \times \mathcal{V}$. To facilitate the discussion we introduce some new notation. Let $\mathbf{U}$ denote the product space $\mathcal{L}^2_{k_0}(\mathcal{F}_0) \times \ldots \times \mathcal{L}^2_{k_T}(\mathcal{F}_T)$. Similarly, we use $\mathbf{V}$ to represent the space $\mathcal{L}^2_{l_1}(\mathcal{G}_1) \times \ldots \times \mathcal{L}^2_{l_{T+1}}(\mathcal{G}_{T+1})$. Consider the following functions:

$$\Phi(u) = \mathbf{E}\{\sum_{\tau=0}^{T}[p_\tau \cdot u_\tau + \tfrac{1}{2}u_\tau \cdot P_\tau u_\tau + \delta_{U_\tau}(u_\tau)]\},$$

$$\Psi(v) = \mathbf{E}\{\sum_{\tau=1}^{T+1}[-q_\tau \cdot v_\tau + \tfrac{1}{2}v_\tau \cdot Q_\tau v_\tau + \delta_{V_\tau}(v_\tau)]\},$$

$$\begin{aligned}\Gamma(u,v)) &= \mathbf{E}\{\sum_{\tau=1}^{T} v_\tau \cdot D_\tau u_\tau + \sum_{\tau=0}^{T} y_{\tau+1} \cdot (B_\tau u_\tau + b_\tau)\} \\ &= \mathbf{E}\{\sum_{\tau=1}^{T} v_\tau \cdot D_\tau u_\tau + \sum_{\tau=1}^{T+1} x_{\tau-1} \cdot (C_\tau^* v_\tau + c_\tau)\}.\end{aligned}$$

It is clear that Φ is a proper, lower semicontinuous convex function on $\mathbf{U}$ with $\operatorname{dom}\Phi = \mathcal{U}$. Likewise, Ψ is a proper, lower semicontinuous convex function on $\mathbf{V}$ with $\operatorname{dom}\Psi = \mathcal{V}$. The functional Γ on $\mathbf{U} \times \mathbf{V}$ is continuous and biaffine. By the definition of $\mathcal{J}$, we see that

$$\mathcal{J}(u,v) = \begin{cases} \infty & \text{if } \Phi(u) = \infty, \\ \Phi(u) - \Psi(v) - \Gamma(u,v) & \text{otherwise,} \end{cases}$$

and thus $\mathcal{J}$ is a proper closed saddle function on $\mathbf{U} \times \mathbf{V}$, with $\mathcal{J} = \overline{\mathcal{J}}$ and $\operatorname{dom} \mathcal{J} = \mathcal{U} \times \mathcal{V}$.

We now define

$$\mathbf{U}^\nu = \{u \in \mathbf{U} | u_\tau \in \mathcal{L}^2_{k_\tau}(\mathcal{F}^\nu_\tau) \text{ for } \tau = 0, \ldots, T\},$$
$$\mathbf{V}^\nu = \{v \in \mathbf{V} | v_\tau \in \mathcal{L}^2_{l_\tau}(\mathcal{G}^\nu_\tau) \text{ for } \tau = 1, \ldots, T+1\},$$

for increasing subfields $\mathcal{F}^\nu_\tau$ of $\mathcal{F}_\tau$ and increasing subfields $\mathcal{G}^\nu_\tau$ of $\mathcal{G}_\tau$. For each ν, the approximate problem is that of finding a saddle point for the functional

$$\mathcal{J}_\nu(u,v) = \begin{cases} \mathcal{J}(u,v) & \text{if } u \in \mathbf{U}^\nu, v \in \mathbf{V}^\nu, \\ -\infty & \text{if } u \in \mathbf{U}^\nu, v \notin \mathbf{V}^\nu, \\ +\infty & \text{if } u \notin \mathbf{U}^\nu. \end{cases}$$

For the remainder of this section we will assume that $U_\tau(\omega)$ and $V_\tau(\omega)$ are constant (a.s.) with respect to ω. The next result shows that, in some sense, $\mathcal{J}_\nu$ may indeed be an approximation for $\mathcal{J}$.

Theorem 4.2.1. *Let $\mathcal{J}$ and $\mathcal{J}_\nu$ be defined as above. Suppose that the σ-fields $\mathcal{F}_\tau$ are generated as $\mathcal{F}_\tau = \sigma(\cup_{\nu=1}^\infty \mathcal{F}^\nu_\tau)$ for each $\tau = 0, \ldots, T$. Likewise, assume $\mathcal{G}_\tau = \sigma(\cup_{\nu=1}^\infty \mathcal{G}^\nu_\tau)$ for each $\tau = 1, \ldots, T+1$. Then $\mathcal{J}_\nu$ Mosco epi/hypo-converges to $\mathcal{J}$ on $\mathbf{U} \times \mathbf{V}$.*

Proof. By Corollary 2.4.6, we need only show

(a) for every $u \in \mathbf{U}$ there exist $\{u^\nu\}$ converging strongly to $\mathbf{V}$ such that $u^\nu \in \mathbf{U}^\nu$ and $\limsup \Phi(u^\nu) \leq \Phi(u)$,

(b) for every $v \in \mathbf{V}$ there exist $\{v^\nu\}$ converging strongly to $\mathbf{V}$ such that $v^\nu \in \mathbf{V}^\nu$ and $\limsup \Psi(v^\nu) \leq \Psi(v)$.

Let $u \in \mathbf{U}$. Define $\overline{\mathcal{U}}(u) = \begin{cases} \mathbf{U} & \text{if } u \notin \mathcal{U}, \\ \mathcal{U} & \text{if } u \in \mathcal{U}. \end{cases}$ By Lemma 3.1.8, there is a sequence $\bar{u}^m \in \overline{\mathcal{U}}(u)$ converging strongly to u for which $\bar{u}^m_\tau$ is $\mathcal{F}_\tau$-simple for each $\tau = 0, \ldots, T+1$. Let $\varepsilon_m = \|u - \bar{u}^m\|_2$. By Lemma 3.1.9, there exists $\tilde{u}^m \in \overline{\mathcal{U}}(u)$ with $\varepsilon_m > \|\tilde{u}^m - \bar{u}^m\|_2$ and for which $\tilde{u}^m_\tau$ is $(\cap_{\nu=1}^\infty \mathcal{F}^\nu_\tau)$-simple. In particular, $\tilde{u}^m_\tau$ must be $\mathcal{F}^{\nu_{m,\tau}}_\tau$-simple for some $\nu_{m,\tau} \geq m$. Take $\nu_m = \max\{\nu_{m,0}, \ldots, \nu_{m,T}\}$, let u^1 be any element of $\mathbf{U}^1$, and set $u^\nu = \tilde{u}^{\nu_m}$ for $\nu_m \leq \nu < \nu_{m+1}$. Then $u_\nu \in \mathbf{U}^\nu$ and $u^\nu \to u$ strongly. If $u \notin \mathcal{U}$, then $\Phi(u) = \infty \geq \limsup \Phi(u^\nu)$. If $u \in \mathcal{U}$, then all of the u^ν are in $\mathcal{U}$, so $\Phi(u) = \lim \Phi(u^\nu) = \limsup \Phi(u^\nu)$. Thus (a) holds. A similar argument gives (b). $\square$

Note that the above result does not require the existence of a saddle point for $\mathcal{J}$. Under the primal and dual finiteness conditions of the previous section (which guarantee the existence of saddle points), we can also guarantee that any sequence of saddle points for the $\{\mathcal{J}_\nu\}$ cluster weakly and therefore cluster about saddle points of $\mathcal{J}$. This is shown in the next result.

Proposition 4.2.2. *Assume the primal and dual finiteness conditions of section 1 hold. Suppose that the pair $(\bar{u}^\nu, \bar{v}^\nu)$ is a saddle point of $\mathcal{J}_\nu$. Then the sequence $\{(\bar{u}^\nu, \bar{v}^\nu)\}$ is bounded, and hence clusters weakly about saddle points of $\mathcal{J}$.*

Proof. Fix $(u, v) \in \mathcal{U} \times \mathcal{V}$. There exist $u^\nu \in \mathbf{U}^\nu$ and $v^\nu \in \mathbf{V}^\nu$ such that

$$\limsup \Phi(u^\nu) \leq \Phi(u) \text{ and } \limsup \Psi(v^\nu) \leq \Psi(v).$$

Thus we have

$$\mathcal{J}(\bar{u}^\nu, v^\nu) \leq \mathcal{J}(\bar{u}^\nu, \bar{v}^\nu) \leq \mathcal{J}(u^\nu, \bar{v}^\nu),$$

so that

$$\Psi(\bar{v}^\nu) + \Gamma(\bar{u}^\nu, \bar{v}^\nu) \leq \Psi(v^\nu) + \Gamma(\bar{u}^\nu, v^\nu)$$

and

$$\Phi(\bar{u}^\nu) - \Gamma(\bar{u}^\nu, \bar{v}^\nu) \leq \Phi(u^\nu) - \Gamma(u^\nu, \bar{v}^\nu).$$

Adding these last two inequalities gives

$$\Phi(\bar{u}^\nu) + \Psi(\bar{v}^\nu) \leq \Phi(u^\nu) + \Psi(v^\nu) + \Gamma(\bar{u}^\nu, v^\nu) - \Gamma(u^\nu, \bar{v}^\nu). \tag{4.2.1}$$

Suppose that $\{\bar{u}^\nu, \bar{v}^\nu\}$ is unbounded, and let $M_\nu = \max\{1, \|\bar{u}^\nu\|, \|\bar{v}^\nu\|\}$. As mentioned in the previous section, the primal and dual finiteness conditions are equivalent to the existence of real numbers $\alpha_1, \alpha_2 > 0$ and functions $a_1, a_2 \in \mathcal{L}_1^1$ such that, for almost every ω and for each $\tau = 1, \ldots, T+1$ and $\tau' = 0, \ldots, T$,

$$s \cdot Q_\tau(\omega) s \geq \alpha_1 |s|^2 - a_1(\omega) \; \forall s \in V_\tau(\omega),$$
$$r \cdot P_{\tau'}(\omega) r \geq \alpha_2 |r|^2 - a_2(\omega) \; \forall r \in U_{\tau'}(\omega).$$

Thus one has

$$\Psi(v) \geq \frac{1}{2}\alpha_1 \|v\| - \mathbf{E}\{\frac{1}{2}a_1 + \sum_{\tau=1}^{T+1} q_\tau \cdot v_\tau\} \text{ and } \Phi(u) \geq \frac{1}{2}\alpha_2 \|u\| - \mathbf{E}\{\frac{1}{2}a_2 - \sum_{\tau=0}^{T} p_\tau \cdot u_\tau\},$$

so that, by (4.2.1), we obtain

$$\begin{aligned}
\infty &= \lim_{\nu\to\infty} \frac{1}{M_\nu}[\Phi(\bar{u}^\nu) + \Psi(\bar{v}^\nu)] \\
&\leq \limsup_{\nu\to\infty} \frac{1}{M_\nu}[\Phi(u^\nu) + \Psi(v^\nu) + \Gamma(\bar{u}^\nu, v^\nu) - \Gamma(u^\nu, \bar{v}^\nu)] \\
&\leq \limsup_{\nu\to\infty} \left[\frac{\Phi(u) + \Psi(v)}{M_\nu} + c_1\frac{(\|v^\nu\| + 1)(\|\bar{u}^\nu + 1)\|}{M_\nu} + c_2\frac{(\|u^\nu\| + 1)(\|\bar{v}^\nu + 1)\|}{M_\nu}\right] \\
&\leq c_1 + c_2,
\end{aligned}$$

a contradiction. Therefore $\{(\bar{u}^\nu, \bar{v}^\nu)\}$ is bounded, and so has weak cluster points. By Proposition 2.4.3, such cluster points are saddle points for $\mathcal{J}$. □

Especially interesting is the *fully quadratic* case, i.e. when the primal and dual conditions are replaced by the condition that there exists a real number $\alpha > 0$ such that, for almost every ω and for each $\tau = 1, \ldots, T+1$ and $\tau' = 0, \ldots, T$, one has

$$\begin{aligned}
s \cdot Q_\tau(\omega)s &\geq \alpha|s|^2 \quad \forall s \in \mathbf{R}^{l_\tau}, \\
r \cdot P_{\tau'}(\omega)r &\geq \alpha|r|^2 \quad \forall r \in \mathbf{R}^{k_\tau}.
\end{aligned}$$

Since this condition is strictly stronger than the primal and dual finiteness conditions, Theorem 4.1.6 guarantees the existence of a saddle point for $\mathcal{J}$. It is easy to show that this saddle point is in fact unique in the fully quadratic case. But much more can be said: solutions to the approximate problems actually converge *strongly* to this solution.

Corollary 4.2.3. *Suppose that the pair $(\bar{u}^\nu, \bar{v}^\nu)$ is a saddle point of $\mathcal{J}_\nu$. In the fully quadratic case, the sequence $\{(\bar{u}^\nu, \bar{v}^\nu)\}$ converges in norm to the (unique) saddle point of $\mathcal{J}$.*

Proof. By the proof of the previous proposition we see that the any subsequence of the sequence $\{(\bar{u}^\nu, \bar{v}^\nu)\}$ must cluster weakly at the saddle point $(\bar{u}, \bar{v})$ of $\mathcal{J}$. Hence the entire sequence converges weakly to $(\bar{u}, \bar{v})$. By the Mosco convergence of $\mathbf{U}^\nu$ to $\mathbf{U}$ and of $\mathbf{V}^\nu$ to $\mathbf{V}$, there exist u^ν converging strongly to $\bar{u}$ and v^ν converging strongly to $\bar{v}$, with $u^\nu \in \mathbf{U}^\nu \cap \mathcal{U}$ and $v^\nu \in \mathbf{V}^\nu \cap \mathcal{V}$. This gives

$$\frac{1}{2}\alpha\|\bar{u}^\nu - \bar{u}\|^2 + \frac{1}{2}\alpha\|\bar{u}^\nu - \bar{u}\|^2 \leq$$

$$\leq \mathbf{E}\left\{\sum_{\tau=0}^{T}\frac{1}{2}(\bar{u}_\tau^\nu-\bar{u}_\tau)\cdot P_\tau(\bar{u}_\tau^\nu-\bar{u}_\tau)+\sum_{\tau=1}^{T+1}\frac{1}{2}(\bar{v}_\tau^\nu-\bar{v}_\tau)\cdot Q_\tau(\bar{v}_\tau^\nu-\bar{v}_\tau)\right\}$$

$$=\mathbf{E}\left\{\sum_{\tau=0}^{T}\frac{1}{2}(\bar{u}_\tau^\nu)\cdot P_\tau(\bar{u}_\tau^\nu)+\sum_{\tau=0}^{T}[-(\bar{u}_\tau^\nu)\cdot P_\tau(\bar{u}_\tau^\nu)+\frac{1}{2}(\bar{u}_\tau)\cdot P_\tau(\bar{u}_\tau)]\right.$$
$$\left.+\sum_{\tau=1}^{T+1}\frac{1}{2}(\bar{v}_\tau^\nu)\cdot Q_\tau(\bar{v}_\tau^\nu)+\sum_{\tau=1}^{T+1}[-(\bar{v}_\tau^\nu)\cdot Q_\tau(\bar{v}_\tau^\nu)+\frac{1}{2}(\bar{v}_\tau)\cdot Q_\tau(\bar{v}_\tau)]\right\}$$

$$=\Phi(\bar{u}^\nu)+\Psi(\bar{v}^\nu)+\mathbf{E}\left\{\sum_{\tau=0}^{T}[-(\bar{u}_\tau^\nu)\cdot P_\tau(\bar{u}_\tau^\nu)+\frac{1}{2}(\bar{u}_\tau)\cdot P_\tau(\bar{u}_\tau)-p_\tau\cdot\bar{u}_\tau]+\right.$$
$$\left.+\sum_{\tau=1}^{T+1}[-(\bar{v}_\tau^\nu)\cdot Q_\tau(\bar{v}_\tau^\nu)+\frac{1}{2}(\bar{v}_\tau)\cdot Q_\tau(\bar{v}_\tau)+q_\tau\cdot\bar{v}_\tau]\right\}$$

$$\leq \Phi(u^\nu)+\Psi(v^\nu)+\Gamma(\bar{u}^\nu,v^\nu)-\Gamma(u^\nu,\bar{v}^\nu)$$
$$+\mathbf{E}\left\{\sum_{\tau=0}^{T}[-(\bar{u}_\tau^\nu)\cdot P_\tau(\bar{u}_\tau^\nu)+\frac{1}{2}(\bar{u}_\tau)\cdot P_\tau(\bar{u}_\tau)-p_\tau\cdot\bar{u}_\tau]\right.$$
$$\left.+\sum_{\tau=1}^{T+1}[-(\bar{v}_\tau^\nu)\cdot Q_\tau(\bar{v}_\tau^\nu)+\frac{1}{2}(\bar{v}_\tau)\cdot Q_\tau(\bar{v}_\tau)+q_\tau\cdot\bar{v}_\tau]\right\},$$

where the last inequality follows from inequality 4.2.1. Rearranging terms, we obtain

$$\frac{1}{2}\alpha\|\bar{u}^\nu-\bar{u}\|^2+\frac{1}{2}\alpha\|\bar{u}^\nu-\bar{u}\|^2$$
$$\leq \mathbf{E}\left\{\sum_{\tau=0}^{T}[\tfrac{1}{2}\bar{u}_\tau\cdot P_\tau\bar{u}_\tau+\frac{1}{2}u_\tau^\nu\cdot P_\tau u_\tau^\nu-\bar{u}_\tau^\nu\cdot P_\tau\bar{u}_\tau+p_\tau\cdot(u_\tau^\nu-\bar{u}_\tau^\nu)]\right\}$$
$$+\mathbf{E}\left\{\sum_{\tau=1}^{T+1}[\tfrac{1}{2}\bar{v}_\tau\cdot Q_\tau\bar{v}_\tau+\frac{1}{2}v_\tau^\nu\cdot Q_\tau v_\tau^\nu-\bar{v}_\tau^\nu\cdot Q_\tau\bar{v}_\tau+q_\tau\cdot(\bar{v}_\tau^\nu-v_\tau^\nu)]\right\}$$
$$+\Gamma(\bar{u}^\nu,v^\nu)-\Gamma(u^\nu,\bar{v}^\nu).$$

Since (u^ν,v^ν) converges in norm to $(\bar{u},\bar{v})$ and $(\bar{u}^\nu,\bar{v}^\nu)$ converges weakly to $(\bar{u},\bar{v})$, the right-hand side of this inequality tends to zero as $\nu\to\infty$. Thus the left-hand side also tends to zero, so that $(\bar{u}^\nu,\bar{v}^\nu)$ converges in norm to $(\bar{u},\bar{v})$, as desired. □

Observe that $\mathcal{J}_\nu$ leads to primal and dual problems of multistage stochastic programming which are of the same type as $(\mathcal{P})$ and $(\mathcal{Q})$. Let $\mathcal{E}^\nu$ be the sub-σ-field

of $\mathcal{E}$ generated by the approximate fields $\mathcal{F}_0^\nu, \ldots, \mathcal{F}_t^\nu, \mathcal{G}_1^\nu, \ldots, \mathcal{G}_{T+1}^\nu$. We replace the data in $(\mathcal{P})$ and $(\mathcal{Q})$ by the following:

$$\begin{aligned}
&\mathcal{U}^\nu = \mathcal{U} \cap \mathbf{U}^\nu, \quad \mathcal{V}^\nu = \mathcal{V} \cap \mathbf{V}^\nu, \\
&p_\tau^\nu = \mathbf{E}^{\mathcal{F}_\tau^\nu}[p_\tau],\ P_\tau^\nu = \mathbf{E}^{\mathcal{F}_\tau^\nu}[P_\tau],\ q_\tau^\nu = \mathbf{E}^{\mathcal{G}_\tau^\nu}[q_\tau],\ Q_\tau^\nu = \mathbf{E}^{\mathcal{G}_\tau^\nu}[Q_\tau], \\
&D_\tau^\nu = \mathbf{E}^{\mathcal{E}^\nu}[D_\tau],\ b_\tau^\nu = \mathbf{E}^{\mathcal{E}^\nu}[b_\tau],\ B_\tau^\nu = \mathbf{E}^{\mathcal{E}^\nu}[B_\tau],\ c_\tau^\nu = \mathbf{E}^{\mathcal{E}^\nu}[c_\tau],\ C_\tau^\nu = \mathbf{E}^{\mathcal{E}^\nu}[C_\tau].
\end{aligned}$$

The primal problem for $\mathcal{J}_\nu$ is then given by

$$(\mathcal{P}^\nu) \qquad \text{minimize} \quad \mathcal{F}^\nu(u) \quad \text{over all} \quad u \in \mathcal{U}^\nu,$$

where

$$\begin{aligned}
F^\nu(u) = \mathbf{E}\Bigg\{ \sum_{\tau=1}^{T} \Bigg[& p_\tau^\nu \cdot u_\tau + \frac{1}{2} u_\tau \cdot P_\tau^\nu u_\tau - c_\tau^\nu \cdot x_{\tau-1} \\
& + \rho_{V_\tau, Q_\tau^\nu}(q_\tau^\nu - \mathbf{E}^{\mathcal{G}_\tau^\nu}[D_\tau^\nu u_\tau + C_\tau^\nu x_{\tau-1}]) \Bigg] \\
& + p_0^\nu \cdot u_0 + \frac{1}{2} u_0 \cdot P_0^\nu u_0 - c_{T+1}^\nu \cdot x_T \\
& + \rho_{V_{T+1}, Q_{T+1}^\nu}(q_{T+1}^\nu - \mathbf{E}^{\mathcal{G}_{T+1}^\nu}[C_{T+1}^\nu x_T]) \Bigg\}.
\end{aligned}$$

The dual problem $(\mathcal{Q}^\nu)$ is given by making the same substitutions in $(\mathcal{Q})$.

An important special case to consider is where each of the approximate fields is finite. Such approximations arise naturally when the measurable space $(\Omega, \mathcal{E})$ consists of the Borel sets on $\mathbf{R}^d$. Any probability measure on $(\Omega, \mathcal{E})$ would then be a Borel measure. In this case any sequence of finite partitions of $\mathbf{R}^d$ for which the diameter of the cells within any bounded set tends to zero uniformly will generate a sequence of fields satisfying the hypotheses of Theorem 4.2.1.

The remainder of this section is devoted to the discussion of the model when the field $\mathcal{E}$ is finite. Obviously the problem is now finite-dimensional and the sets $\mathcal{U}$ and $\mathcal{V}$ are both polyhedral. It is easy to see that the objective function F is piecewise linear-quadratic. It turns out that the problem $(\mathcal{P})$ is equivalent to a quadratic program [35]. Typically the number of variables would far exceed the capabilities of conventional quadratic programming routines. However the problem has a very special structure which can be exploited by various "decomposition" techniques like the *finite envelope method,* which is described in Chapter 5.

We shall assume, with no real loss of generality, that $\mathcal{E}$ is actually the field generated by the union of the fields $\mathcal{F}_0, \mathcal{G}_1, \ldots, \mathcal{F}_T, \mathcal{G}_{T+1}$. Note that any finite field on Ω is generated by a *partition* of Ω. Let $\mathcal{A}_\tau$ denote the partition generating the field $\mathcal{F}_\tau$. For convenience, we include the empty set in $\mathcal{A}_\tau$. We define $\mathcal{A}$ to be the cartesian product $\mathcal{A}_0 \times \cdots \times \mathcal{A}_T$. Similarly, we define the collections $\mathcal{B}_\tau$ for the partitions generating the $\mathcal{G}_\tau$. For $\alpha \in \mathcal{A}$ we denote the intersection $\cap_{\tau=0}^{T} \alpha_\tau$ by $\cap\alpha$, and likewise write $\cap\beta$ for $\cap_{\tau=1}^{T+1} \beta_\tau$ and $\cap(\alpha, \beta)$ for $(\cap\alpha) \cap (\cap\beta)$. Note that $\cap\alpha$, $\cap\beta$ and $\cap(\alpha, \beta)$ are $\mathcal{E}$-measurable sets, and that any set in $\mathcal{E}$ can be represented as $\cap(\alpha, \beta)$ for some (α, β). Furthermore, when $\alpha \in \mathcal{A}$ is nonempty one has

$$\alpha = \bar{\alpha} \Longleftrightarrow \cap\alpha_\tau = \cap\bar{\alpha}_\tau,$$

with similar statements for $\mathcal{B}$ and $\mathcal{A} \times \mathcal{B}$. Thus we may define the probabilities

$$\begin{aligned}
\pi(\alpha, \beta) &= \mu[\cap(\alpha, \beta)], \\
\pi(\alpha) &= \mu[\cap\alpha], \\
\pi(\beta) &= \mu[\cap\beta], \\
\pi(\alpha_\tau) &= \mu[\alpha_\tau], \\
\pi(\beta_\tau) &= \mu[\beta_\tau].
\end{aligned}$$

Observe that

$$1 = \sum_{\alpha_\tau \in \mathcal{A}_\tau} \pi(\alpha_\tau) = \sum_{\alpha \in \mathcal{A}} \pi(\alpha) = \sum_{\beta_\tau \in \mathcal{B}_\tau} \pi(\beta_\tau) = \sum_{\beta \in \mathcal{B}} \pi(\beta) = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} \pi(\alpha, \beta).$$

and

$$\begin{aligned}
\pi(\bar{\alpha}_\tau) &= \sum_{\substack{\alpha \in \mathcal{A} \\ \text{with } \alpha_\tau = \bar{\alpha}_\tau}} \pi(\alpha) = \sum_{\substack{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B} \\ \text{with } \alpha_\tau = \bar{\alpha}_\tau}} \pi(\alpha, \beta), \\
\pi(\bar{\beta}_\tau) &= \sum_{\substack{\beta \in \mathcal{B} \\ \text{with } \beta_\tau = \bar{\beta}_\tau}} \pi(\beta) = \sum_{\substack{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B} \\ \text{with } \beta_\tau = \bar{\beta}_\tau}} \pi(\alpha, \beta), \\
\pi(\bar{\alpha}) &= \sum_{\beta \in \mathcal{B}} \pi(\bar{\alpha}, \beta), \\
\pi(\bar{\beta}) &= \sum_{\alpha \in \mathcal{A}} \pi(\alpha, \bar{\beta}).
\end{aligned}$$

The mapping p_τ is $\mathcal{F}_\tau$-measurable if and only if p_τ is constant on each (nonvoid) set in the partition $\mathcal{A}_\tau$. Thus we can write $p_\tau(\alpha_\tau)$ for the common value of $p_\tau(\omega)$ for $\omega \in \alpha_\tau$. Similarly, we may write $P_\tau(\alpha_\tau)$, $U_\tau(\alpha_\tau)$, $Q_\tau(\beta_\tau)$, and so on. Since p_τ is also constant on $\cap\alpha$ for each $\alpha \in \mathcal{A}$, we may also write $p_\tau(\alpha)$. In addition, all the data elements are $\mathcal{E}$-measurable, so we may write $A_\tau(\alpha,\beta)$, $D_\tau(\alpha,\beta)$, $p_\tau(\alpha,\beta)$. Of course, none of these make sense if the argument is the empty set. In this case, we simply define the value to be the origin in the appropriate space (or the empty set, if the mapping is set-valued). With this notation, the expectation of an $\mathcal{F}_\tau$-measurable function z on Ω can be reexpressed as

$$\mathbf{E}z = \sum_{\alpha_\tau \in \mathcal{A}_\tau} \pi(\alpha_\tau) z(\alpha_\tau) = \sum_{\alpha \in \mathcal{A}} \pi(\alpha) z(\alpha) = \sum_{(\alpha,\beta) \in \mathcal{A}\times\mathcal{B}} \pi(\alpha,\beta) z(\alpha,\beta).$$

The conditional expectation with respect to $\mathcal{F}_\tau$ of any $\mathcal{E}$-measurable function z is given by

$$(\mathbf{E}^{\mathcal{F}_\tau}[z])(\bar{\alpha}_\tau) = \begin{cases} \sum\left\{ \begin{smallmatrix} (\alpha,\beta)\in\mathcal{A}\times\mathcal{B} \\ \text{with } \alpha_\tau = \bar{\alpha}_\tau \end{smallmatrix} \right\} \pi(\alpha,\beta) z(\alpha,\beta)/\pi(\bar{\alpha}_\tau) & \text{if } \pi(\bar{\alpha}_\tau) \neq 0, \\ 0 & \text{if } \pi(\bar{\alpha}_\tau) = 0. \end{cases}$$

With all of this notation, we are able to state some fairly useful results concerning the decomposition of the saddle point condition in the finite-dimensional case. The next theorem is an example of such a result.

Proposition 4.1.7. *Consider $\bar{u} \in \mathcal{U}$ and $\bar{v} \in \mathcal{V}$ with corresponding trajectories $\bar{x}$ and $\bar{y}$. Define*

$$\begin{aligned} \bar{r}_\tau &= q_\tau - \mathbf{E}^{\mathcal{G}_\tau}[C_\tau \bar{x}_{\tau-1} - D_\tau \bar{u}_\tau] \quad \text{for} \quad \tau = 1,\dots,T, \\ \bar{r}_{T+1} &= q_{T+1} - \mathbf{E}^{\mathcal{G}_{T+1}}[C_{T+1}\bar{x}_T], \\ \bar{s}_\tau &= \mathbf{E}^{\mathcal{F}_\tau}[B_\tau^* \bar{y}_{\tau+1} + D_\tau \bar{v}_\tau] - p_\tau \quad \text{for} \quad \tau = 1,\dots,T, \\ \bar{s}_0 &= \mathbf{E}^{\mathcal{F}_0}[B_0^* \bar{y}_1] - p_0. \end{aligned}$$

Then $\tilde{u} \in \operatorname{argmin}_{u\in\mathcal{U}} \mathcal{J}(u,\bar{v})$ if and only, for each $\tau = 0,\dots,T$ and each $\alpha_\tau \in \mathcal{A}_\tau$, one has

$$\tilde{u}_\tau(\alpha_\tau) \in \operatorname*{argmax}_{u_\tau(\alpha_\tau)\in U_\tau(\alpha_\tau)} \{\bar{s}_\tau(\alpha_\tau)\cdot u_\tau(\alpha_\tau) - \frac{1}{2} u_\tau(\alpha_\tau)\cdot P_\tau(\alpha_\tau) u_\tau(\alpha_\tau)\}.$$

Similarly, $\tilde{v} \in \operatorname{argmin}_{v \in \mathcal{V}} \mathcal{J}(\bar{u}, v)$ *if and only, for each* $\tau = 1, \ldots, T+1$ *and each* $\beta_\tau \in \mathcal{B}_\tau$, *one has*

$$\tilde{v}_\tau(\beta_\tau) \in \operatorname*{argmax}_{v_\tau(\beta_\tau) \in V_\tau(\beta_\tau)} \{\bar{r}_\tau(\beta_\tau) \cdot v_\tau(\beta_\tau) - \frac{1}{2} v_\tau(\beta_\tau) \cdot Q_\tau(\beta_\tau) v_\tau(\beta_\tau)\}.$$

Proof. We simply expand the expression for $\mathcal{J}(u, \bar{v})$:

$$\begin{aligned}
\mathcal{J}(u, \bar{v}) &= \sum_{\tau=0}^{T} \mathbf{E}\{p_\tau \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau\} + \sum_{\tau=1}^{T+1} \mathbf{E}\{q_\tau \cdot \bar{v}_\tau - \tfrac{1}{2} \bar{v}_\tau \cdot Q_\tau \bar{v}_\tau\} \\
&\quad - \sum_{\tau=1}^{T} \mathbf{E}\{\bar{v}_\tau \cdot D_\tau u_\tau\} - \sum_{\tau=0}^{T} \mathbf{E}\{\bar{y}_{\tau+1} \cdot (B_\tau u_\tau + b_\tau)\} \\
&= \sum_{\tau=1}^{T} \mathbf{E}\{[p_\tau - D_\tau^* \bar{v}_\tau - B_\tau^* \bar{y}_{\tau+1}] \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau\} \\
&\quad + \mathbf{E}\{[p_0 - B_0^* \bar{y}_1] \cdot u_0 + \tfrac{1}{2} u_0 \cdot P_0 u_0\} \\
&= \sum_{\tau=1}^{T} \mathbf{E}\{(p_\tau - \mathbf{E}^{\mathcal{G}_\tau}[D_\tau^* \bar{v}_\tau + B_\tau^* \bar{y}_{\tau+1}]) \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau\} \\
&\quad + \mathbf{E}\{[p_0 - B_0^* \bar{y}_1] \cdot u_0 + \tfrac{1}{2} u_0 \cdot P_0 u_0\} \\
&= \sum_{\tau=0}^{T} \mathbf{E}\{-\bar{s} \cdot u_\tau + \tfrac{1}{2} u_\tau \cdot P_\tau u_\tau\}.
\end{aligned}$$

The proof of the second statement is the same. □

Chapter 5. Numerical Solution of Discretized Problems

In Chapters 3 and 4, we introduced some infinite-dimensional optimization problems, and showed how certain discretizations of these lead to finite-dimensional problems of similar structure. More importantly, we demonstrated that the discretized versions can actually be considered as variational approximations to the original problems.

The discretized problems can typically be represented as convex programs of extremely high dimension. Theoretically, the usual methods of solution for convex programs could be applied to these, but the dimensionality of the problems is an obstacle to most approaches. On the other hand, such a problem has a very special structure: it consists essentially of a very large number of relatively small problems which are related to each other in a simple manner. We are thus lead to seek out methods of solution which exploit this structure, requiring only simple computations and perhaps allowing the use of parallel processors. In this chapter, we examine one such approach to solving these problems: the *finite envelope method* of Rockafellar and Wets [33], [34] (also known as the *finite generation method*).

In the first section, we introduce a general finite-dimensional model to which the algorithm can be applied. In the section 2, we describe the algorithm and prove some convergence results. Section 3 is devoted to the specialization of the algorithm to the solution of the discrete-time optimal control problem of Chapter 3. In particular, we consider the so-called "box-diagonal" case, which is a nearly separable problem of piecewise linear-quadratic programming.

1. Description of the Model

The finite envelope method was originally developed by Rockafellar and Wets as a dual approach to solving a very special linear-quadratic problem in stochastic programming [32]. In recent years, they have built up a complete duality theory of *piecewise* (or *extended*) *linear-quadratic programming*; at the same time, Rockafellar [29] has extended the finite envelope method to this larger class of problems. In this section and the next, we show that this method may be extended to an even broader class of problems.

We shall be concerned with the following problems:

$$(\mathcal{P}) \qquad \begin{aligned} &\text{minimize} \quad f(u) \text{ over all } u \in U, \\ &\text{where } f(u) = p \cdot u + \varphi(u) + (\psi + \delta_V)^*(q - Du) \end{aligned}$$

and

$$(\mathcal{Q}) \qquad \begin{aligned} &\text{maximize} \quad g(v) \text{ over all } v \in V, \\ &\text{where } g(v) = q \cdot v - \psi(v) - (\varphi + \delta_U)^*(D^*v - p). \end{aligned}$$

Here U and V are (nonempty) closed convex sets in $\mathbf{R}^k$ and $\mathbf{R}^l$ (respectively) and φ and ψ are (finite) differentiable convex functions on $\mathbf{R}^k$ and $\mathbf{R}^l$. By the *quadratic case*, we shall mean the case where the sets U and V are polyhedral and the functions φ and ψ have the form $\varphi(u) = \frac{1}{2}u \cdot Pu$ and $\psi(v) = \frac{1}{2}v \cdot Qv$.

The problems $\mathcal{P}$ and $\mathcal{Q}$ are dual to each other, as can be seen by introducing the functional

$$J(u, v) = p \cdot u + \varphi(u) + q \cdot v - \psi(v) - v \cdot Du$$

and noting that

$$f(u) = \sup_{v \in V} J(u, v) \text{ and } g(v) = \inf_{u \in U} J(u, v).$$

In particular, we have the following well-known result on weak duality. (See [23] for example.)

Proposition 5.1.1. *It is always true that* $\inf(\mathcal{P}) \geq \sup(\mathcal{Q})$. *A given pair* $(\bar{u}, \bar{v})$ *is a saddle point for* J *relative to* $U \times V$ *if and only if* $\bar{u}$ *solves* $\mathcal{P}$, $\bar{v}$ *solves* $\mathcal{Q}$, *and* $\min(\mathcal{P}) = \max(\mathcal{Q})$.

Throughout this chapter we make the following assumption concerning the functions φ and ψ and the sets U and V.

Finiteness Assumption. The functions $(\varphi + \delta_U)^*$ and $(\psi + \delta_V)^*$ are finite everywhere. Equivalently, the functions $\varphi + \delta_U$ and $\psi + \delta_V$ are *coercive*, i.e.

$$\lim_{\lambda \to \infty} (\varphi + \delta_U)(\lambda u)/\lambda = \infty \text{ for all } u \neq 0,$$
$$\lim_{\lambda \to \infty} (\psi + \delta_V)(\lambda v)/\lambda = \infty \text{ for all } v \neq 0.$$

Under this assumption, f and g are finite everywhere, so that there are no (strictly enforced) constraints in $\mathcal{P}$ and $\mathcal{Q}$ other than the requirement that $u \in U$ and $v \in V$. We remark that $\varphi + \delta_U$ is coercive if either U is bounded or φ is strongly convex. Similarly, $\psi + \delta_V$ is coercive if either V is bounded or ψ is strongly convex.

The finiteness assumption represents a normalization. For example, suppose $C \subset \mathbf{R}^k$ and $D \subset \mathbf{R}^l$ are compact convex sets, and that $(\bar{u}, \bar{v})$ is a saddle point of J relative to $(C \cap U) \times (D \cap V)$. If $\bar{u} \in \operatorname{int} C$ and $\bar{v} \in \operatorname{int} D$, then $(\bar{u}, \bar{v})$ is also a saddle point of J relative to $U \times V$. Alternatively, we could add quadratic "proximal" terms to φ and ψ to make them strongly convex. More will be said about this latter approach in the next section.

One of the main reasons for imposing the finiteness assumption is that it enables us to close the duality gap between problems $(\mathcal{P})$ and $(\mathcal{Q})$, as shown in the following theorem. This will prove valuable in the analysis of the algorithm given in the next section.

Proposition 5.1.2 Strong Duality. *Optimal solutions exist to both $(\mathcal{P})$ and $(\mathcal{Q})$, and $\min(\mathcal{P}) = \max(\mathcal{Q})$ (finite). Moreover, $(\bar{u}, \bar{v})$ is a saddle point of J relative to $U \times V$ if and only if $\bar{u}$ solves $(\mathcal{P})$ and $\bar{v}$ solves $(\mathcal{Q})$.*

Proof. The coercivity of the function $\varphi + \delta_U$ implies that it has no "directions of recession" (see [22]), and similarly for $\psi + \delta_V$. We may therefore apply Theorem 37.6 of [22] to obtain the existence of saddle points for J relative to $U \times V$. The rest of the proposition then follows from Proposition 5.2.1 above. □

We remark that in the quadratic case, strong duality may be achieved without resorting to the Finiteness Assumption: in this case, one has $\inf(\mathcal{P}) = \sup(\mathcal{Q})$ if either $(\mathcal{P})$ or $(\mathcal{Q})$ admits a feasible solution. For further details on this, see [33].

2. The Finite Envelope Method

In this section we describe the basic finite envelope method for solving the saddle point problem associated with the primal and dual problems $(\mathcal{P})$ and $(\mathcal{Q})$ of the previous section. First we introduce some notation. For each $u \in \mathbf{R}^k$, we define

$$F(u) = \operatorname*{argmax}_{v \in V} J(u, v) = \{v \in V | J(u, v) = f(u)\}.$$

Similarly, for each $v \in \mathbf{R}^l$, we define

$$G(v) = \operatorname*{argmin}_{u \in U} J(u,v) = \{u \in U | J(u,v) = g(v)\}.$$

Under the finiteness assumption made in section 1, the sets $F(u)$ and $G(v)$ are nonempty for all u and v. In particular, we may write

$$f(u) = \max_{v \in V} J(u,v) \text{ and } g(v) = \min_{u \in U} J(u,v).$$

We may now state the algorithm.

Finite Envelope Method. *Given initial guesses $\bar{u}_0 \in U$ and $\bar{v}_0 \in V$, we generate sequences $\{\bar{u}_\nu\} \subset U$ and $\{\bar{v}_\nu\} \subset V$ as follows.*
Step 1 (Optimality Test). *Set $\varepsilon^\nu = f(\bar{u}_\nu) - g(\bar{v}_\nu)$. If ε_ν is sufficiently small, then terminate. Otherwise, proceed with Step 2.*
Step 2 (Envelope Generation). *Calculate $\tilde{u}_\nu \in G(\bar{v}_\nu)$ and $\tilde{v}_\nu \in F(\bar{u}_\nu)$. Choose closed convex sets $U^\nu \subset U$ and $V^\nu \subset V$ such that*

$$\{\bar{u}_\nu, \tilde{u}_\nu\} \subset U^\nu \text{ and } \{\bar{v}_\nu, \tilde{v}_\nu\} \subset V^\nu.$$

Step 3 (Envelope Subproblem). *Determine a saddle point $(\hat{u}_\nu, \hat{v}_\nu)$ for J relative to $U^\nu \times V^\nu$.*
Step 4 (Line Search). *Find a minimum point $\bar{u}_{\nu+1}$ of f over the line segment from $\bar{u}_\nu$ to $\hat{u}_\nu$. Likewise, find a maximum point $\bar{v}_{\nu+1}$ of g over the line segment from $\bar{v}_\nu$ to $\hat{v}_\nu$. Return to Step 1.*

The algorithm given here presupposes that, for any $u \in U$, it is computationally easy to calculate both the value $f(u)$ and at least one element of $F(u)$. Similarly, we assume that for each $v \in V$, the calculation of $g(v)$ and an element of $G(v)$ is straightforward. In addition, it will be advantageous to have methods available to efficiently minimize f over any line segment in U, and to maximize g over any line segment in V.

The primary obstacle to the effective application of this algorithm is the saddle point computation in Step 3. In the quadratic case (i.e. where $\varphi(u) = \frac{1}{2}u \cdot Pu$ and $\psi(v) = \frac{1}{2}v \cdot Qv$), this subproblem may be reformulated equivalently as a quadratic

program of low dimension, which can be easily solved by existing routines (see section 3). But even in this case, experience to date indicates that Step 3 requires far more computational effort than the other calculations in the algorithm.

For the remainder of this section, we shall discuss some general aspects of the algorithm, including convergence. First, it is natural to ask why the size of the duality gap ε^ν in Step 1 should be considered as a stopping criterion. The next result answers this question.

Proposition 5.2.1. *Let $\bar{u}$ be optimal for $(\mathcal{P})$ and $\bar{v}$ be optimal for $(\mathcal{Q})$. If u^* is feasible for $(\mathcal{P})$ and v^* is feasible for $(\mathcal{Q})$ with $f(u^*) - g(v^*) \leq \varepsilon$, then*

$$|f(u^*) - f(\bar{u})| \leq \varepsilon \text{ and } |g(v^*) - g(\bar{v})| \leq \varepsilon.$$

Furthermore, if φ has the form $\varphi(u) = \varphi_0(u) + \frac{1}{2}u \cdot Pu$ and likewise ψ has the form $\psi(v) = \psi_0(v) + \frac{1}{2}v \cdot Qv$ (where P and Q are symmetric, positive semi-definite matrices), then also

$$\|u^* - \bar{u}\|_P \leq (2\varepsilon)^{1/2} \text{ and } \|v^* - \bar{v}\|_Q \leq (2\varepsilon)^{1/2},$$

where we define $\|u\|_P = (u \cdot Pu)^{1/2}$ and $\|v\|_Q = (v \cdot Qv)^{1/2}$.

Proof. The estimates for $|f(u^*) - f(\bar{u})|$ and $|g(v^*) - g(\bar{v})|$ follow immediately from the saddle point representation of optimality given in Theorem 5.1.2. Now write $J(u,v) = J_0(u,v) + \varphi_0(u) - \psi(v)$, where

$$J_0(u,v) = p \cdot u + \tfrac{1}{2}u \cdot Pu + q \cdot v - \tfrac{1}{2}v \cdot Qv - v \cdot Du.$$

Then we have

$$\begin{aligned}
\varepsilon \geq f(u^*) - f(\bar{u}) &\geq f(u^*) - J(\bar{u},\bar{v}) \geq J(u^*,\bar{v}) - J(\bar{u},\bar{v}) \\
&= \varphi_0(u^*) - \varphi_0(\bar{u}) + \nabla_u J_0(\bar{u},\bar{v}) \cdot (u^* - \bar{u}) + \frac{1}{2}\|u^* - \bar{u}\|_P^2 \\
&\geq \nabla_u J(\bar{u},\bar{v}) \cdot (u^* - \bar{u}) + \frac{1}{2}\|u^* - \bar{u}\|_P^2 \\
&\geq \frac{1}{2}\|u^* - \bar{u}\|_P^2,
\end{aligned}$$

where the last inequality holds because $\bar{u}$ minimizes $J(u,\bar{v})$ over all $u \in U$. Similarly, $-\varepsilon \leq g(v^*) - g(\bar{v}) \leq -\frac{1}{2}\|v^* - \bar{v}\|_Q^2$. □

The above proof is a slight modification of that given for the quadratic case by Rockafellar [29]. In the same paper he proves the following lemma, which we will use (sometimes without mention) in the proof of convergence.

Lemma 5.2.2. *Suppose that* $u^{**} \in G(v^*)$ *and* $v^{**} \in F(u^*)$. *Then* $f(u) \geq J(u, v^{**})$ *for all* u *(with equality if* $u = u^{**}$*), and* $\nabla_u J(u^*, v^{**}) \in \partial f(u^*)$. *Similarly,* $g(v) \leq J(u^{**}, v)$ *for all* v *(with equality if* $v = v^{**}$*), and* $\nabla_v J(u^{**}, v^*) \in \partial g(v^*)$.

We can use Lemma 5.2.2 to shed some light on the nature of the direction-finding subproblem given by Steps 2 and 3 of the algorithm. We introduce the primal and dual subproblems

$$(\mathcal{P}^\nu) \qquad \begin{aligned} &\text{minimize } \ f^\nu(u) \text{ over all } u \in U^\nu, \\ &\text{where } f^\nu(u) = p \cdot u + \varphi(u) + (\psi + \delta_{V^\nu})^*(q - Du) \end{aligned}$$

and

$$(\mathcal{Q}^\nu) \qquad \begin{aligned} &\text{maximize } \ g^\nu(v) \text{ over all } v \in V^\nu, \\ &\text{where } g^\nu(v) = q \cdot v - \psi(v) - (\varphi + \delta_{U^\nu})^*(D^* v - p). \end{aligned}$$

It is clear that $f(u) \geq f^\nu(u)$ for all $u \in \mathbf{R}^k$, with equality if $F(u) \cap V^\nu \neq \emptyset$. Similarly, $g(v) \geq g^\nu(v)$ for all $v \in \mathbf{R}^l$, with equality if $G(v) \cap U^\nu \neq \emptyset$. Thus f^ν is a lower envelope approximation for f which agrees with f at $\bar{u}_\nu$. Moreover, if both f and f^ν are differentiable at $\bar{u}_\nu$, then $\nabla f^\nu(\bar{u}_\nu) = \nabla f(\bar{u}_\nu)$, by Lemma 5.2.2. In fact, we can prove somewhat more, as the next proposition (and its corollary) show. The proof is given in [29].

Proposition 5.2.3. *In the algorithm, suppose that* $F(\bar{u}_\nu)$ *is a singleton, as would be the case if* $\psi + \delta_V$ *were strictly convex. Then* f *and* f^ν *are differentiable at* $\bar{u}_\nu$ *with* $\nabla f(\bar{u}_\nu) = \nabla f^\nu(\bar{u}_\nu)$. *If* $\bar{u}_\nu$ *is not optimal for* $(\mathcal{P})$, *then* $\hat{u}_\nu - \bar{u}_\nu$ *is a direction of descent for* f *at* $\bar{u}_\nu$ *relative to* U.

Corollary 5.2.4. *If* $F(\bar{u}_\nu)$ *and* $G(\bar{v}_\nu)$ *are singletons, then the* ν*-th iteration leads to a reduction in the duality gap (i.e.* $\varepsilon^{\nu+1} < \varepsilon^\nu$*), unless* $\bar{u}_\nu$ *and* $\bar{v}_\nu$ *are already optimal for* $(\mathcal{P})$ *and* $(\mathcal{Q})$.

In general, we cannot guarantee that the duality gap ε^ν will converge to zero, unless we have more information on the specific nature of the sets U^ν and V^ν chosen in Step 2 of the algorithm. Fortunately, in the case where φ and ψ are strongly convex, it can be shown that ε^ν decreases to zero at a geometric rate. This is the subject of the next theorem, which was proved for the quadratic case in [29] by Rockafellar. (See also [33].)

Theorem 5.2.5. *Suppose that* $\varphi = \varphi_0 + \frac{1}{2}\|\cdot\|_P^2$ *for some symmetric, positive definite matrix* P *and some (differentiable) convex function* φ_0. *(Here, as in Proposition 5.2.1, we use the notation* $\|u\|_P = (u \cdot Pu)^{1/2}$.*) Similarly, assume that* $\psi = \psi_0 + \|\cdot\|_Q^2$, *with* Q *symmetric positive definite and* ψ_0 *convex. Then, in this case, one has*

$$\varepsilon^{\nu+1} \leq \theta\varepsilon^{\nu} \text{ for all } \nu \in \mathbf{N}, \tag{5.2.1}$$

where $\theta = 1 - \frac{1}{4(1+\gamma^2)} < 1$ *with* γ *given by*

$$\gamma^2 = \max\left\{\sup_{x\neq 0}\frac{x \cdot D^*Q^{-1}Dx}{x \cdot Px}, \sup_{x\neq 0}\frac{x \cdot DP^{-1}D^*x}{x \cdot Qx}\right\}.$$

If $\bar{u}$ *is the solution to* $(\mathcal{P})$ *and* $\bar{v}$ *is the solution to* $(\mathcal{Q})$, *then we also have*

$$\|\bar{u}_\nu - \bar{u}\|_P^2 + \|\bar{v}_\nu - \bar{v}\|_Q^2 \leq 2\theta^\nu\varepsilon^0 \text{ for all } \nu. \tag{5.2.2}$$

Proof. First we define

$$J_0(u,v) = p \cdot u + \tfrac{1}{2}u \cdot Pu + q \cdot v - \tfrac{1}{2}v \cdot Qv - v \cdot Du$$

so that $J(u,v) = J_0(u,v) + \varphi_0(u) - \psi_0(v)$. Then we can write

$$\begin{aligned} J(u,v) = {} & \varphi_0(u) - \psi_0(v) + J_0(\bar{u}_\nu, \tilde{v}_\nu) \\ & + \nabla_u J_0(\bar{u}_\nu, \tilde{v}_\nu) \cdot (u - \bar{u}_\nu) + \nabla_v J_0(\bar{u}_\nu, \tilde{v}_\nu) \cdot (v - \tilde{v}_\nu) \\ & + \frac{1}{2}(u - \bar{u}_\nu) \cdot P(u - \bar{u}_\nu) - \frac{1}{2}(v - \bar{v}_\nu) \cdot Q(v - \bar{v}_\nu) \\ & - (v - \bar{v}_\nu) \cdot D(u - \bar{u}_\nu). \end{aligned} \tag{5.2.3}$$

By definition of $\tilde{v}_\nu$, we have

$$f(\bar{u}_\nu) = \varphi_0(\bar{u}_\nu) - \psi_0(\tilde{v}_\nu) + J_0(\bar{u}_\nu, \tilde{v}_\nu) = \sup_{v\in V} J(\bar{u}_\nu, v).$$

Thus $J_0(\bar{u}_\nu, \tilde{v}_\nu) = f(\bar{u}_\nu) - \varphi_0(\bar{u}_\nu) + \psi_0(\tilde{v}_\nu)$, and

$$(-\nabla\psi_0(\tilde{v}_\nu) + \nabla_v J_0(\bar{u}_\nu, \tilde{v}_\nu)) \cdot (v - \tilde{v}_\nu) \leq 0 \text{ for all } v \in V$$

so that

$$\nabla_v J_0(\bar{u}_\nu, \tilde{v}_\nu) \cdot (v - \tilde{v}_\nu) \leq \psi_0(v) - \psi_0(\tilde{v}_\nu) \text{ for all } v \in V.$$

Also,

$$\begin{aligned}\nabla_u J_0(\bar{u}_\nu, \tilde{v}_\nu) \cdot (u - \bar{u}_\nu) &= \nabla_u J(\bar{u}_\nu, \tilde{v}_\nu) \cdot (u - \bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu) \cdot (u - \bar{u}_\nu) \\ &= [\nabla f(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu)] \cdot (u - \bar{u}_\nu).\end{aligned}$$

Applying these facts to (5.2.3), we get

$$\begin{aligned}J(u,v) &\le \varphi_0(u) - \psi_0(v) + [f(\bar{u}_\nu) - \varphi_0(\bar{u}_\nu) + \psi_0(\tilde{v}_\nu)] \\ &\quad + [\nabla f(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu)] \cdot (u - \bar{u}_\nu) + [\psi_0(v) - \psi_0(\tilde{v}_\nu)] \\ &\quad + \frac{1}{2}(u - \bar{u}_\nu) \cdot P(u - \bar{u}_\nu) - \frac{1}{2}(v - \bar{v}_\nu) \cdot Q(v - \bar{v}_\nu) - (v - \bar{v}_\nu) \cdot D(u - \bar{u}_\nu) \\ &= [\varphi_0(u) - \varphi_0(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu) \cdot (u - \bar{u}_\nu)] + [f(\bar{u}_\nu) + \nabla f(\bar{u}_\nu) \cdot (u - \bar{u}_\nu)] \\ &\quad + \frac{1}{2}(u - \bar{u}_\nu) \cdot P(u - \bar{u}_\nu) - \frac{1}{2}(v - \bar{v}_\nu) \cdot Q(v - \bar{v}_\nu) - (v - \bar{v}_\nu) \cdot D(u - \bar{u}_\nu)\end{aligned}$$

for all $v \in V$. Since $f(u) = \sup_{v \in V} J(u,v)$, this inequality gives us

$$\begin{aligned}f(u) &- [\varphi_0(u) - \varphi_0(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu) \cdot (u - \bar{u}_\nu)] - [f(\bar{u}_\nu) + \nabla f(\bar{u}_\nu) \cdot (u - \bar{u}_\nu)] \\ &- \frac{1}{2}(u - \bar{u}_\nu) \cdot P(u - \bar{u}_\nu) \\ &\le \max_{v \in V}\{-(v - \bar{v}_\nu) \cdot D(u - \bar{u}_\nu) - \frac{1}{2}(v - \bar{v}_\nu) \cdot Q(v - \bar{v}_\nu)\} \\ &\le \sup_{w \in \mathbf{R}^l}\{-w \cdot D(u - \bar{u}_\nu) - \frac{1}{2} w \cdot Qw\} \\ &= \frac{1}{2}(u - \bar{u}_\nu) \cdot D^* Q^{-1} D(u - \bar{u}_\nu).\end{aligned}$$

Suppose u has the form $u = \bar{u}_\nu + \lambda(\hat{u}_\nu - \bar{u}_\nu)$ for $\lambda \in [0,1]$, so that we have

$$\begin{aligned}f(\bar{u}_\nu &+ \lambda(\hat{u}_\nu - \bar{u}_\nu)) - f(\bar{u}_\nu) \\ &\le [\varphi_0(\bar{u}_\nu + \lambda(\hat{u}_\nu - \bar{u}_\nu)) - \varphi_0(\bar{u}_\nu) - \lambda\nabla\varphi_0(\bar{u}_\nu) \cdot (\hat{u}_\nu - \bar{u}_\nu)] \\ &\quad + \lambda[\nabla f(\bar{u}_\nu) \cdot (\hat{u}_\nu - \bar{u}_\nu)] + \frac{1}{2}\lambda^2 \|\hat{u}_\nu - \bar{u}_\nu\|^2_{P_0},\end{aligned}$$

where $P_0 = P + D^* Q^{-1} D$. The requirement that $\bar{u}_{\nu+1}$ give the minimum of f over the line segment between $\bar{u}_\nu$ and $\hat{u}_\nu$ yields

$$\begin{aligned}f(\bar{u}_{\nu+1}) &- f(\bar{u}_\nu) \\ &\le \min_{\lambda \in [0,1]} \left\{ \lambda \left\{ \begin{array}{c} [\varphi_0(\hat{u}_\nu) - \varphi_0(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu) \cdot (\hat{u}_\nu - \bar{u}_\nu)] \\ + [\nabla f(\bar{u}_\nu) \cdot (\hat{u}_\nu - \bar{u}_\nu)] \end{array} \right\} \right. \\ &\qquad \left. + \frac{1}{2}\lambda^2 \|\hat{u}_\nu - \bar{u}_\nu\|^2_{P_0} \right\}\end{aligned} \tag{5.2.4}$$

We now estimate the coefficient of λ^2 in (5.2.4). First note that

$$\begin{aligned}\|\hat{u}_\nu - \bar{u}_\nu\|_{P_0}^2 &= \|\hat{u}_\nu - \bar{u}_\nu\|_P^2 + \|\hat{u}_\nu - \bar{u}_\nu\|_{D^*Q^{-1}D}^2 \\ &\leq \|\hat{u}_\nu - \bar{u}_\nu\|_P^2(1+\gamma^2).\end{aligned} \tag{5.2.5}$$

Next, since $(\hat{u}_\nu, \hat{v}_\nu)$ is a saddle point of J relative to $U^\nu \times V^\nu$, we see that

$$\begin{aligned}f^\nu(\bar{u}_\nu) - f^\nu(\hat{u}_\nu) &\geq J(\bar{u}_\nu, \hat{v}_\nu) - J(\hat{u}_\nu, \hat{v}_\nu) \\ &= \varphi_0(\bar{u}_\nu) - \varphi_0(\hat{u}_\nu) + \nabla_u J(\hat{u}_\nu, \hat{v}_\nu)\cdot(\bar{u}_\nu - \hat{u}_\nu) + \frac{1}{2}\|\hat{u}_\nu - \bar{u}_\nu\|_P^2 \\ &\geq \nabla\varphi_0(\bar{u}_\nu)\cdot(\bar{u}_\nu - \hat{u}_\nu) + \nabla_u J(\hat{u}_\nu, \hat{v}_\nu)\cdot(\bar{u}_\nu - \hat{u}_\nu) + \frac{1}{2}\|\hat{u}_\nu - \bar{u}_\nu\|_P^2 \\ &= \nabla f^\nu(\hat{u}_\nu)\cdot(\bar{u}_\nu - \hat{u}_\nu) + \frac{1}{2}\|\hat{u}_\nu - \bar{u}_\nu\|_P^2 \geq \frac{1}{2}\|\hat{u}_\nu - \bar{u}_\nu\|_P^2,\end{aligned}$$

where we have used Lemma 5.2.2 to get $\nabla f^\nu(\hat{u}_\nu) = \nabla_u J(\hat{u}_\nu, \hat{v}_\nu)$. Combining this with (5.2.5) yields

$$\|\hat{u}_\nu - \bar{u}_\nu\|_{P_0}^2 \leq [f^\nu(\bar{u}_\nu) - f^\nu(\hat{u}_\nu)](1+\gamma^2). \tag{5.2.6}$$

We turn then to the coefficient of λ in (5.2.4). By (5.2.3), we have

$$\begin{aligned}&f^\nu(\hat{u}_\nu) - f^\nu(\bar{u}_\nu) \\ &\quad\geq J(\hat{u}_\nu, \tilde{v}_\nu) - J(\bar{u}_\nu, \tilde{v}_\nu) \\ &\quad= [\varphi_0(\hat{u}_\nu) - \psi_0(\tilde{v}_\nu) + J_0(\bar{u}_\nu, \tilde{v}_\nu) + \nabla_u J_0(\bar{u}_\nu, \tilde{v}_\nu)\cdot(\hat{u}_\nu - \bar{u}_\nu) \\ &\qquad\quad + \frac{1}{2}(\hat{u}_\nu - \bar{u}_\nu)\cdot P(\hat{u}_\nu - \bar{u}_\nu)] \\ &\qquad - [\varphi_0(\bar{u}_\nu) - \psi_0(\tilde{v}_\nu) + J_0(\bar{u}_\nu, \tilde{v}_\nu)] \\ &\quad= [\varphi_0(\hat{u}_\nu) - \varphi_0(\bar{u}_\nu)] + [\nabla_u J(\bar{u}_\nu, \tilde{v}_\nu) - \nabla\varphi_0(\bar{u}_\nu)]\cdot(\hat{u}_\nu - \bar{u}_\nu) \\ &\qquad + \frac{1}{2}\|\hat{u}_\nu - \bar{u}_\nu\|_P^2 \\ &\quad\geq [\varphi_0(\hat{u}_\nu) - \varphi_0(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu)\cdot(\hat{u}_\nu - \bar{u}_\nu)] + \nabla_u J(\bar{u}_\nu, \tilde{v}_\nu)\cdot(\hat{u}_\nu - \bar{u}_\nu) \\ &\quad\geq [\varphi_0(\hat{u}_\nu) - \varphi_0(\bar{u}_\nu) - \nabla\varphi_0(\bar{u}_\nu)\cdot(\hat{u}_\nu - \bar{u}_\nu)] + \nabla f(\bar{u}_\nu)\cdot(\hat{u}_\nu - \bar{u}_\nu).\end{aligned}$$

We may now use this inequality and (5.2.6) to obtain an upper bound for the right-hand side in (5.2.4), giving

$$\begin{aligned}f(\bar{u}_{\nu+1}) - f(\bar{u}_\nu) &\leq [f^\nu(\bar{u}_\nu) - f^\nu(\hat{u}_\nu)] \min_{\lambda\in[0,1]}\{-\lambda + (1+\gamma^2)\lambda^2\} \\ &= [f^\nu(\bar{u}_\nu) - f^\nu(\hat{u}_\nu)][-4(1+\gamma^2)]^{-1} \\ &= [f(\bar{u}_\nu) - J(\hat{u}_\nu, \hat{v}_\nu)][-4(1+\gamma^2)]^{-1}\end{aligned} \tag{5.2.7}$$

By a similar argument, we can show

$$g(\bar{v}_{\nu+1}) - g(\bar{v}_\nu) \geq [f(\bar{u}_\nu) - J(\hat{u}_\nu, \hat{v}_\nu)][-4(1+\gamma^2)]^{-1}.$$

Subtracting (5.2.7) from this yields $\varepsilon^\nu - \varepsilon^{\nu+1} \geq \varepsilon^\nu[-4(1+\gamma^2)]^{-1}$, so that

$$\varepsilon^\nu \left[1 - \frac{1}{4(1+\gamma^2)}\right] \geq \varepsilon^{\nu+1}$$

which is the same as (5.2.1).

The estimate in (5.2.2) now follows from Proposition 5.2.1. □

Corollary 5.2.6. *Suppose that φ is strongly convex with modulus γ_φ and that ψ is strongly convex with modulus γ_ψ. Then $\varepsilon^{\nu+1} \leq \theta\varepsilon^\nu$ for all $\nu \in \mathbf{N}$, where*

$$\theta = 1 - \frac{1}{4(1+\gamma^2)} < 1 \text{ with } \gamma = \|D\|_2/\sqrt{\gamma_\varphi\gamma_\psi}.$$

If $\bar{u}$ is optimal for $(\mathcal{P})$ and $\bar{v}$ is optimal for $(\mathcal{Q})$, then

$$\gamma_\varphi\|\bar{u}_\nu - \bar{u}\|^2 + \gamma_\psi\|\bar{v}_\nu - \bar{v}\|^2 \leq 2\theta^\nu\varepsilon^0 \text{ for all } \nu.$$

Proof. It is easily verified that $\varphi_0 = \varphi - \frac{1}{2}\gamma_\varphi\|\cdot\|^2$ and $\psi_0 = \psi - \frac{1}{2}\gamma_\psi\|\cdot\|^2$ are convex functions. We may then apply Theorem 5.2.5 with $P = \gamma_\varphi I_k$ and $Q = \gamma_\psi I_l$ (where I_k and I_l denote the $k \times k$ and $l \times l$ identity matrices). From linear algebra, we have the following chain of identities:

$$\sup_{x\neq 0} \frac{x \cdot DD^*x}{x \cdot x} = \|D^*\|_2^2 = \|D\|_2^2 = \sup_{x\neq 0} \frac{x \cdot D^*Dx}{x \cdot x}.$$

The value of γ may be then calculated using these. □

In the case where either φ or ψ fails to be strongly convex, we may still be able to approximate solutions to $(\mathcal{P})$ and $(\mathcal{Q})$ by adding *proximal* terms. For example, suppose that $(u^\dagger, v^\dagger) \in U \times V$ is the current "best guess" for a saddle point of J relative to $U \times V$. We define

$$J^\dagger(u,v) = J(u,v) + \frac{1}{2r^\dagger}\|u - u^\dagger\|^2 - \frac{1}{2r^\dagger}\|v - v^\dagger\|^2,$$

where $r^\dagger$ is some positive constant. Theorem 5.2.5 says that we can then use the finite envelope method to find an approximate solution $(\bar{u}^\dagger, \bar{v}^\dagger)$ of $J^\dagger$ relative to $U \times V$. This constitutes one iteration of the *proximal point algorithm*, which can be shown to converge under various mild assumptions. We refer the reader to the papers [24], [25] of Rockafellar for the theoretical details. In [33], Rockafellar and Wets present an algorithm based on using the proximal point algorithm in conjunction with the finite envelope method to solve a special quadratic version of $(\mathcal{P})$ and $(\mathcal{Q})$. Their analysis includes rates of convergence and a discussion of stopping criteria.

3. Implementation of Algorithm: Discrete-Time Optimal Control

A highly desirable feature of the finite envelope method is that its implementation can make extensive use of the specialized structure present in many large-scale optimization problems. In this section, we briefly indicate how this can be done for (deterministic) discrete-time optimal control problems, such as those arising from the discretizations discussed in Chapter 3. The author [39] has developed an experimental code (the DYNFGM program) for the solution of such problems in the case of extended linear-quadratic objectives with constant coefficients. This is also the case described here.

We consider the following extended linear-quadratic problem of discrete-time optimal control, given in its saddle point representation:

$$(\mathcal{S}) \qquad \begin{array}{c} \text{find a saddle point of } \mathcal{J}(u_0, \ldots, u_N;\, v_1, \ldots, v_{N+1}) \\ \text{relative to } (U_e \times (U)^N) \times ((V)^N \times V_e), \end{array}$$

where

$$\mathcal{J}(u_0, \ldots, u_N;\, v_1, \ldots, v_{N+1}) = \sum_{\tau=1}^{N} J(u_\tau, v_\tau) + J_e(u_0, v_{N+1}) - \langle (u_0, \ldots, u_N);\, (v_1, \ldots v_{N+1}) \rangle.$$

Here we have

$$J(u_\tau, v_\tau) = p \cdot u_\tau + q \cdot v_\tau + \frac{1}{2} u_\tau \cdot P u_\tau - \frac{1}{2} v_\tau \cdot Q v_\tau - v_\tau \cdot D u_\tau$$

$$J_e(u_0, v_{N+1}) = p_e \cdot u_0 + q_e \cdot v_{N+1} + \frac{1}{2} u_0 \cdot P_e u_0 - \frac{1}{2} v_{N+1} \cdot Q_e v_{N+1}$$

and

$$\langle (u_0,\dots,u_N);(v_1,\dots,v_{N+1})\rangle = \sum_{\tau=1}^{N} y_{\tau+1}(Bu_\tau + b) + y_1\cdot(B_e u_0 + b_e)$$
$$= \sum_{\tau=1}^{N} x_{\tau-1}(C^* v_\tau + c) + x_N\cdot(C_e^* v_{N+1} + c_e).$$

The trajectories are given by

$$x_\tau = Ax_{\tau-1} + Bu_\tau + b \text{ for } \tau = 1,\dots,N, \qquad (x_\tau \in \mathbf{R}^m)$$
$$x_0 = B_e u_0 + b_e$$
$$y_\tau = A^* y_{\tau+1} + C^* v_\tau + c \text{ for } \tau = 1,\dots,N, \qquad (y_\tau \in \mathbf{R}^m).$$
$$y_{N+1} = C_e^* v_{N+1} + c_e$$

The sets U, U_e, V and V_e are convex polyhedrons, and the matrices P, P_e, Q, Q_e are symmetric, positive semidefinite. The primal and dual problems are

minimize
$$f(u_0,\dots,u_N) = \sup_{(V)^N\times V_e} \mathcal{J}(u_0,\dots,u_N;\,\cdot\,)$$
over all $(u_0,\dots,u_N) \in U_e \times (U)^N$

and

maximize
$$g(v_1,\dots,v_{N+1}) = \inf_{U_e\times(U)^N} \mathcal{J}(\,\cdot\,;v_1,\dots,v_{N+1})$$
over all $(v_1,\dots,v_{N+1}) \in (V)^N \times V_e$.

The problem $(\mathcal{S})$ is a finite-dimensional saddle point problem with a quadratic, convex-concave saddle function of the type in sections 1 and 2. Recall that the general linear-quadratic saddle point problem can be stated as

$(\widetilde{\mathcal{S}})$ find a saddle point of
$$\mathcal{J}(u,v) = \tilde{p}\cdot u + \frac{1}{2}u\cdot\tilde{P}u + \tilde{q}\cdot v - \frac{1}{2}v\cdot\tilde{Q}v - v\cdot\tilde{D}u$$
relative to $\mathcal{U}\times\mathcal{V}$.

That $(\mathcal{S})$ fits into this framework can be seen by considering the case where the data in $(\widetilde{\mathcal{S}})$ is defined as follows:

$$\mathcal{U} := U_e \times (U)^N \subseteq \mathbf{R}^{ke} \times (\mathbf{R}^k)^N \text{ and } \mathcal{V} = (V)^N \times V_e \subseteq (\mathbf{R}^l) \times \mathbf{R}^{le},$$

and

$$\tilde{P} = \operatorname{diag}[P_e, P, \ldots, P], \quad \tilde{p} = \begin{bmatrix} p_e \\ p \\ \vdots \\ p \end{bmatrix} - \tilde{B}^* \tilde{A}^* \tilde{c},$$

$$\tilde{Q} = \operatorname{diag}[Q, \ldots, Q, Q_e], \quad \tilde{q} = \begin{bmatrix} q \\ \vdots \\ q \\ q_e \end{bmatrix} - \tilde{C} \tilde{A} \tilde{b},$$

$$\tilde{D} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & D & 0 & 0 & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & 0 & 0 & D & 0 \\ D_e & 0 & \ldots & 0 & 0 \end{bmatrix} + \tilde{C} \tilde{A} \tilde{B}.$$

Here we have used the matrices

$$\tilde{A} = \begin{bmatrix} \begin{pmatrix} I & & & & \\ A & I & & & \\ \vdots & \vdots & \ddots & & \\ A^{N-1} & A^{N-2} & \ldots & I & \\ A^N & A^{N-1} & \ldots & A & I \end{pmatrix} \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} b_e \\ b \\ \vdots \\ b \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} c \\ \vdots \\ c \\ c_e \end{bmatrix},$$

$$\tilde{B} = \operatorname{diag}[B_e, B, \ldots, B], \quad \tilde{C} = \operatorname{diag}[C, \ldots, C, C_e].$$

The DYNFGM program is designed to efficiently solve the problem in the very important *box-diagonal* case: we assume that P, Q, P_e, Q_e are diagonal matrices with positive diagonal elements, and that U, U_e, V, V_e are "boxes", i.e. Cartesian products of closed, bounded intervals. There are many advantages in having a code specially written for this case, rather than solving it with a routine for the more general quadratic problem $(\widetilde{\mathcal{S}})$. The most obvious is the savings in data storage: if k and l are small but N is large, the corresponding matrices would be quite sparse. For example, the matrix $\tilde{P}$ would require $(Nk + k_e)^2$ storage elements. But since P

and P_e are diagonal, and because the problem is autonomous (i.e., the data elements are time-independent) we require only $k + k_e$ storage elements to accommodate P and P_e.

A somewhat less apparent source of savings is in the necessary work: it turns out that $\mathcal{J}$ is *separable* in the variables

$$u_{i0}\ (i = 1, \ldots, k_e), \quad u_{i\tau}\ (i = 1, \ldots, k;\ \tau = 1, \ldots, N)$$
$$v_{i,N+1}\ (i = 1, \ldots, l_e), \quad v_{i\tau}\ (i = 1, \ldots, l;\ \tau = 1, \ldots, N).$$

In fact, for any $(u_0, \ldots, u_N) = u$ we can *explicitly* calculate the function value $f(u_0, \ldots, u_N)$ as well as its derivative, and similarly for the dual variables. For example, suppose that we know $\bar{u}$ and its corresponding trajectory $\bar{x}$, and assume that $V_e = \prod_{i=1}^{l_e}[v_{i0}^-, v_{i0}^+]$ and $V = \prod_{i=1}^{l}[v_i^-, v_i^+]$. Then we can write

$$\begin{aligned}
f(\bar{u}) &:= \sup_{\mathcal{V}} \mathcal{J}(u, v) \\
&= \sup_{\mathcal{V}} \left\{ \sum_{\tau=1}^{N} J(\bar{u}_\tau, v_\tau) + J_e(\bar{u}_0, v_{N+1}) - \langle \bar{u}, v \rangle \right\} \\
&= \sup_{\mathcal{V}} \left\{ \sum_{\tau=1}^{N} \left[\begin{array}{c} p \cdot \bar{u}_\tau + \frac{1}{2}\bar{u}_\tau \cdot P\bar{u}_\tau + q \cdot v_\tau - \frac{1}{2} v_\tau \cdot Q v_\tau \\ -v_\tau \cdot D\bar{u}_\tau - \bar{x}_{\tau-1} \cdot (C^* v_\tau + c) \end{array} \right] \right. \\
&\qquad \left. + \left[\begin{array}{c} p_e \cdot \bar{u}_0 + \frac{1}{2}\bar{u}_0 \cdot P_e \bar{u}_0 + q_e \cdot v_{N+1} - \frac{1}{2} v_{N+1} \cdot Q_e v_{N+1} \\ -\bar{x}_N \cdot (C_e^* v_{N+1} + c_e) \end{array} \right] \right\} \\
&= \sum_{\tau=1}^{N} \left\{ \begin{array}{l} p \cdot \bar{u}_\tau + \frac{1}{2}\bar{u}_\tau \cdot P\bar{u}_\tau - \bar{x}_{\tau-1} \cdot c \\ \qquad + \sup_{v_\tau \in V} [(q - D\bar{u}_\tau - C\bar{x}_{\tau-1}) \cdot v_\tau - \frac{1}{2} v_\tau \cdot Q v_\tau] \end{array} \right\} \\
&\quad + \left[\begin{array}{l} p_e \cdot \bar{u}_0 + \frac{1}{2}\bar{u}_0 \cdot P_e \bar{u}_0 - \bar{x}_N \cdot c_e \\ \qquad + \sup_{v_{N+1} \in V_e} [(q_e - C_e \bar{x}_N) \cdot v_{N+1} - \frac{1}{2} v_{N+1} \cdot Q_e v_{N+1}] \end{array} \right] \\
&= \left[\sum_{\tau=1}^{N} (p \cdot \bar{u}_\tau + \frac{1}{2}\bar{u}_\tau \cdot P\bar{u}_\tau - \bar{x}_{\tau-1} \cdot c) + p_e \cdot \bar{u}_0 + \frac{1}{2}\bar{u}_0 \cdot P_e \bar{u}_0 - \bar{x}_N \cdot c_e \right] +
\end{aligned}$$

$$+\sum_{\tau=1}^{N}\sum_{i=1}^{l}\sup_{v'\in[v_i^-,v_i^+]}\left[\left(q_i-\sum_{j=1}^{k}D_{ij}\bar{u}_{j\tau}-\sum_{j=1}^{m}C_{ij}\bar{x}_{j,\tau-1}\right)v'-\frac{1}{2}Q_{ii}(v')^2\right]$$

$$+\sum_{i=1}^{l_e}\sup_{v'\in[v_{i,N+1}^-,v_{i,N+1}^+]}\left[\left(q_{ei}-\sum_{j=1}^{m}(C_e)_{ij}\bar{x}_{jN}\right)v'-\frac{1}{2}(Q_e)_{ii}(v')^2\right].$$

So the maximization problem to determine $f(u)$ reduces to $Nl+l_e$ one-dimensional maximization problems, each of the form:

$$\text{maximize } \alpha v'-\frac{1}{2}\beta(v')^2 \text{ over } v'\in[v^-,v^+].$$

These can be solved *explicitly*. This special structure greatly simplifies many tasks, e.g., the line searches in Step 4 of the algorithm. Problems without the box-diagonal structure are still separable with respect to the time periods, allowing parallelization of much of the algorithm.

In Step 2 of the ν^{th} iteration, the DYNFGM code generates finite sets $\mathcal{U}^\nu\subset\mathcal{U}$ and $\mathcal{V}^\nu\subset\mathcal{V}$ satisfying $\{\bar{u}^\nu,\tilde{u}^\nu\}\subset\text{co}\,\mathcal{U}^\nu$ and $\{\bar{v}^\nu,\tilde{v}^\nu\}\subset\text{co}\,\mathcal{V}^\nu$, where $\bar{u}^\nu$ is the "best guess" for the primal problem and $\bar{v}^\nu$ is "best" for the dual problem (the "co" denotes the convex hull of the set). The vectors $\tilde{u}^\nu$, $\tilde{v}^\nu$ are given by

$$\tilde{u}^\nu\in G(\bar{v}^\nu),\quad \tilde{v}^\nu\in F(\bar{u}^\nu)$$

where

$$G(v):=\operatorname*{argmin}_{u\in\mathcal{U}}\mathcal{J}(u,v),\qquad F(u):=\operatorname*{argmax}_{v\in\mathcal{V}}\mathcal{J}(u,v).$$

In this implementation, for given controls $\bar{u}^\nu$, $\tilde{u}^\nu$ and $\hat{v}^\nu$, $\tilde{v}^\nu$ we generate the finite sets

$$\mathcal{U}^\nu=\{\bar{u}^\nu,\tilde{u}^\nu,u^{(1)},\ldots,u^{(L)}\},\qquad \mathcal{V}^\nu=\{\bar{v}^\nu,\tilde{v}^\nu,v^{(1)},\ldots,v^{(L)}\}$$

by the rule

$$u^{(j)}:=G(v^{(j-1)}),\qquad v^{(j)}:=F(u^{(j-1)}),$$

for $j=1,\ldots,L$ (for some fixed positive integer L), where we take $u^{(0)}=\tilde{u}^\nu$ and $v^{(0)}=\tilde{v}^\nu$. The generation of these elements is made especially easy by the separability discussed in the previous paragraph.

The envelope subproblem (Step 3) concerns the solution of the saddle point problem

(S^ν) find a saddle point $(\hat{u}, \hat{v})$ of $J(u,v)$ relative to $\operatorname{co}\mathcal{U}^\nu \times \operatorname{co}\mathcal{V}^\nu$.

We can rewrite this by using the substitution

$$u = \sum_{i=1}^{m} \xi^i u^i, \qquad v = \sum_{j=1}^{n} \eta^j v^j$$

to get

$$\begin{aligned}
\hat{J}(\xi,\eta) &:= J\left(\sum_{i=1}^{m} \xi^i u^i, \sum_{j=1}^{n} \eta^j v^j\right) \\
&= \sum_{i=1}^{m} \xi^i \left[p_e \cdot u_0^i + \sum_{\tau=1}^{N} p \cdot u_\tau^i\right] + \frac{1}{2} \sum_{i,k=1}^{m,m} \xi^i \xi^k \left[u_0^i \cdot P_e u_0^k + \sum_{\tau=1}^{N} u_\tau^i \cdot P u_\tau^k\right] \\
&\quad - \sum_{i,j=1}^{m,n} \xi^i \eta^j \left[\sum_{\tau=1}^{N} v_\tau^j \cdot D u_\tau^i + \langle u^i, v^j \rangle\right] + \sum_{j=1}^{n} \eta^j \left[q_e \cdot v_{N+1}^j + \sum_{\tau=1}^{N} q \cdot v_\tau^j\right] \\
&\quad - \frac{1}{2} \sum_{j,l=1}^{n,n} \eta^j \eta^l \left[v_{N+1}^j \cdot Q_e v_{N+1}^l + \sum_{\tau=1}^{N} v_\tau^j \cdot Q v_\tau^l\right] \\
&=: \hat{p} \cdot \xi + \frac{1}{2} \xi \cdot \hat{P} \xi - \eta \cdot \hat{D} \xi + \hat{q} \cdot \eta - \frac{1}{2} \eta \cdot \hat{Q} \eta.
\end{aligned}$$

By exploiting the Kuhn-Tucker conditions and duality, it turns out that the reformulated problem

($\widehat{S}$) find a saddle point $(\hat{\xi}, \hat{\eta})$ of $\hat{J}(\xi,\eta)$ relative to $(\xi,\eta) \in \{\xi \geq 0 | \mathbf{1} \cdot \xi = 1\} \times \{\eta \geq 0 | \mathbf{1} \cdot \eta = 1\}$

is equivalent to

(C) minimize $\hat{p} \cdot \xi + \frac{1}{2} \xi \cdot \hat{P} \xi + \frac{1}{2} \zeta \cdot \hat{Q} \zeta + \tau$
over all ξ, ζ, τ satisfying
$\xi \geq 0, \mathbf{1} \cdot \xi = 1, \mathbf{1} \cdot \zeta = 1, \tau$ free
$\hat{D}\xi + \hat{Q}\zeta + \tau \mathbf{1} \geq \hat{q}$

where η is the Lagrange multiplier vector for the constraint $\bar{D}\xi+\bar{Q}\zeta+\tau\mathbf{1} \geq q$. Notice that (C) is a low-dimensional convex quadratic program, which may be solved using standard quadratic programming codes.

The line search part of the algorithm (Step 4) may be implemented in several different ways. Typically, one would parametrize u by $u = \lambda\hat{u}^\nu + (1-\lambda)\bar{u}^\nu$ for the problem of finding $\bar{u}^{\nu+1} \in [\hat{u}^\nu, \bar{u}^\nu]$ to minimize f. This leads to the problem of minimizing $\alpha(\lambda)$ subject to $\lambda \in [0,1]$, where

$$\alpha(\lambda) := \max_{v \in \mathcal{V}} J(\lambda\hat{u}^\nu + (1-\lambda)\bar{u}^\nu, v).$$

For each λ we may explicitly find the v giving this maximum, as well as the derivative of α at λ. (At the optimal λ this v will be the new element $\tilde{v}^{\nu+1}$.) It turns out that α is a convex, piecewise linear-quadratic function of λ, and that the "breakpoints" of α may also be explicitly computed. Hence, one could trace λ from breakpoint to breakpoint to find where the derivative changes sign. Clearly, standard line search procedures could also be used, or even various combinations of these with the above method.

Similar decompositions to those described in this section are possible also for stochastic problems [33], [34]. For details of such an implementation for quadratic problems in stochastic programming with recourse, we refer the reader to the paper of King [15].

BIBLIOGRAPHY

[1] H. Attouch, D. Azé, R. J.-B. Wets, "Convergence of convex-concave saddle functions: applications to convex programming and mechanics," *Ann. Inst. Henri Poincaré, Analyse non Linèaire* **5** No. 6 (1988), 537–572.

[2] H. Attouch and R. J.-B. Wets, "A convergence for bivariate functions aimed at the convergence of saddle values," in *Mathematical Theory of Optimization*, J. P. Cecconi and T. Zolezzi, eds., Springer-Verlag Lecture Notes in Mathematics, No. 979, 1981, 1–42.

[3] H. Attouch and R. J.-B. Wets, "A convergence theory for saddle functions," *Trans. Amer. Math. Soc.* **280** No. 1 (1983), 1–41.

[4] W. E. Bosarge,Jr. and O. G. Johnson, "Error bounds of high order accuracy for the state regulator problem via piecewise polynomial approximations," *SIAM J. Control* **9** (1971), 15–28.

[5] G. Chen and W. H. Mills, "Finite elements and terminal penalization for quadratic cost optimal control problems governed by ordinary differential equations," *SIAM J. Control Optim.* **19** No.6 (1981), 744–764.

[6] J. Cullum, "Penalty functions and nonconvex continuous optimal control problems," in *Computing Methods in Optimization Problems–2*, L. A. Zadeh, L. W. Neustadt, and A. V. Balakrishnan, eds., Academic Press, New York, 1969, 55-67.

[7] J. Cullum, "Discrete approximations to continuous optimal control problems," *SIAM J. Control* **7** No.1 (1969), 32–49.

[8] J. Cullum, "An explicit procedure for discretizing continuous optimal control problems," *J. Optim. Th. Appl.* **8** No.1 (1971), 15–34.

[9] J. W. Daniel, "The Ritz-Galerkin Method for abstract optimal control problems," *SIAM J. Control* **11** No.1 (1973), 53–63.

[10] C. N. Do, *Second-Order Nonsmooth Analysis and Sensitivity in Optimization Problems Involving Convex Integral Functionals*, Ph.D. dissertation, Univ. of Washington, 1989.

[11] W. W. Hager, "The Ritz-Trefftz method for state and control constrained optimal control problems," *SIAM J. Numer. Anal.* **12** (1975), 854–867.

[12] W. W. Hager, "Rates of convergence for discrete approximations to unconstrained control problems in a finite dimensional space," *SIAM J. Numer. Anal.* **13** (1976), 449–472.

[13] W. W. Hager, "Approximations to the multiplier method," *SIAM J. Numer. Anal.* **22** (1985), 16–46.

[14] W. W. Hager and G. D. Ianculescu, "Dual approximations in optimal control," *SIAM J. Control Optim.* **22** No.3 (1984), 423–465.

[15] A. J. King, "An implementation of the Lagrangian finite generation method," in *Numerical Techniques for Stochastic Programming Problems,* Y. Ermoliev and R. J.-B. Wets, eds., Springer-Verlag, Berlin, New York, 1988.

[16] F. H. Mathis, and G. W. Reddien, "Ritz-Trefftz approximations in optimal control," *SIAM J. Control Optim.* **17** (1979), 307–310.

[17] J. J. Moreau, "Théorèmes 'inf-sup,'" *C. R. Acad. Sci. Paris* **258** (1964), 2720-2722.

[18] U. Mosco, "Convergence of convex sets and of solutions of variational inequalities," *Advances Math.* **3** (1969), 510–585.

[19] U. Mosco, "On the continuity of the Young-Fenchel transform," *J. Math. Anal. Appl.* **35** (1971), 518–535.

[20] O. Pironneau and E. Polak, "A dual method for optimal control problems with initial and final boundary constraints," *SIAM J. Control* **11** (1973), 534–549.

[21] R. T. Rockafellar, "Minimax theorems and conjugate saddle functions," *Math. Scand.* **14** (1964), 151–173.

[22] R. T. Rockafellar, *Convex Analysis,* Princeton Univ. Press, Princeton, N.J., 1970.

[23] R. T. Rockafellar, *Conjugate Duality and Optimization,* Regional Conference Series in Applied Mathematics, 61, Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[24] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Math. Op. Res.* **1** (1976), 97–116.

[25] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Optim.* **14** No. 5 (1976), 877–898.

[26] R. T. Rockafellar, "Integral functionals, normal integrands and measurable selections," in *Nonlinear Operators and the Calculus of Variations,* L. Waelbrock, ed., Lecture Notes in Mathematics, No. 543, Springer-Verlag, 1976, 157-207.

[27] R. T. Rockafellar, "Linear-quadratic programming and optimal control," *SIAM J. Control Optim.* **25** No. 3, (1987) 781-814.

[28] R. T. Rockafellar, "On the essential boundedness of solutions to problems in piecewise linear-quadratic optimal control," in *Analyse Mathématique et Applications,* F. Murat and O. Pironneau, eds., Gauthier-Villars, Paris, 1988, 437–443.

[29] R. T. Rockafellar, "Computational schemes for large-scale problems in extended linear-quadratic programming," *Math. Programming* **48** (1990), 447–474.

[30] R. T. Rockafellar, "Generalized second derivatives of convex functions and saddle functions," preprint.

[31] R. T. Rockafellar, "Hamiltonian trajectories and duality in the optimal control of linear systems with convex costs," preprint.

[32] R. T. Rockafellar and R. J.-B. Wets, "A dual solution procedure for quadratic stochastic programs with simple recourse," in *Numerical Methods,* V. Pereyra and A. Reinoza, eds., Lecture Notes in Mathematics, No. 1005, Springer-Verlag, Berlin, 1983, 252–265.

[33] R. T. Rockafellar and R. J.-B. Wets, "A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming," *Math. Programming Study* **28** (1986), 63–93.

[34] R. T. Rockafellar and R. J.-B. Wets, "Linear-quadratic programming problems with stochastic penalties: the finite generation algorithm," in *Numerical Techniques for Stochastic Optimization Problems,* Y. Ermoliev and R.J.-B. Wets, eds., Springer-Verlag, Berlin, New York, 1986.

[35] R. T. Rockafellar and R. J.-B. Wets, "Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete-time," *SIAM J. Control Optim.* **28** No. 4, (1990) 810-822.

[36] D. L. Russell, "Penalty functions and bounded phase coordinate control," *SIAM J. Control* **2** (1965), 409–422.

[37] R. Wijsman, "Convergence of sequences of convex sets, cones and functions," *Bull. Amer. Math. Soc.* **70** (1964), 186–188.

[38] R. Wijsman, "Convergence of sequences of convex sets, cones and functions II," *Trans. Amer. Math. Soc.* **123** (1966), 32–45.

[39] S. E. Wright, "DYNFGM: A computer program for solving extended linear-quadratic problems in optimal control by saddle point generation," Technical report, Dept. of Mathematics, University of Washington, March 1989.

[40] C. Zhu and R. T. Rockafellar, "Finite-envelope gradient projection methods for extended linear-quadratic programming," preprint.

Biographical Note

Stephen E. Wright was born on October 6, 1962 in Indianapolis, Indiana. He attended secondary school in Miles City, Montana and received a Bachelor of Arts in mathematics from the University of Montana in 1985. He was awarded the Doctor of Philosophy in December 1990 by the University of Washington.