LINEAR PROGRAMMING BASED APPROACHES TO OPTIMAL CONTROL PROBLEMS WITH LONG RUN AVERAGE OPTIMALITY CRITERIA

By

Ludmila Manic

Principal Supervisor: Professor Vladimir Gaitsgory
Associate Supervisor: Professor Paul Smith

A thesis submitted to Macquarie University,
Department of Mathematics
For the degree of
Doctor of Philosophy

August 31, 2015

MACQUARIE UNIVERSITY
FACULTY OF SCIENCE

Contents

Glossary

LRAOC	long run average optimal control
LP	linear programming
ID	infinite dimensional
IDLP	infinite dimensional linear programming
SI	semi-infinite
SILP	semi-infinite dimensional linear programming
HJB	Hamilton-Jacobi-Bellman
SP	singularly perturbed
ACG	average control generating
$\square$	indicates the end of a proof
$\stackrel{\text{def}}{=}$	denotes "is defined to be equal"
$\Rightarrow$	denotes "implies"
A^T	the transpose of a matrix A
$\mathbb{R}^n$	n-dimensional Euclidian space
clV	the closure of the set V
$convV$	the convex hull of a set V
$\rho_H(A, B)$	Hausdorff distance between sets A and B

Abstract

The thesis aims at the development of mathematical tools for analysis and construction of near optimal solutions of optimal control problems with long run average optimality criteria (LRAOC). It consists of three parts. In Part I, we establish that near optimal controls of these problems can be constructed on the basis of solutions of semi-infinite dimensional linear programming (SILP) problems and their duals. The latter are shown to be approximations of the Hamilton-Jacobi-Bellman inequality corresponding to the LRAOC problem. In Part II, we extend the consideration of Part I to singularly perturbed LRAOC problems. Our approach to these problems is based on amalgamation of averaging and linear programming based techniques. We show that an asymptotically near optimal solution of the singularly perturbed problem can be constructed on the basis of an optimal solution of the averaged LRAOC problem and we show that the optimal solution of the latter can be found with the help of linear programming based techniques. Some of the results obtained in Parts I and II are stated in the form of algorithms, the convergence of which is discussed and which are illustrated with numerical examples. In Part III, we study families of SILP problems depending on a small parameter. The family of SILP problems is regularly (singularly) perturbed if its optimal value is continuous (discontinuous) at the zero value of the parameter. We introduce a regularity condition such that if it is fulfilled, then the family of SILP problems is regularly perturbed and if it is not fulfilled, then the family is likely to be singularly perturbed. We establish relationships between the regularity condition for SILP problems and regularity conditions used in dealing with perturbed LRAOC problems.

Declaration

I certify that the work in this thesis entitled "Linear programming based approaches to optimal control problems with long run average optimality criteria" has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree to any other university or institution other than Macquarie University.

I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself have been appropriately acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Ludmila Manic (43786306) August 31, 2015

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Vladimir Gaitsgory, for his immeasurable amount of support, guidance, encouragement, great patience and friendship that made this work possible.

I am very grateful to my associate supervisor, Professor Paul Smith, for helpful discussions and for encouraging me during my studies at Macquarie University. Also, I wish to express my gratitude to Professor Jerzy Filar and Associate Professor Vladimir Ejov, for their very helpful discussions, advices and for their continuous support during my studies at University of South Australia and Flinders University.

My special thanks goes to Mr Serguei Rossomakhine for his support, comments and explanations in everything what refers to programming and numerical calculations. I also would like to acknowledge that the numerical results presented in Section 6.3 of the thesis have been obtained with a help of Mr Rossomakhine.

I wish to express my sincere gratitude to all staff members, visitors and postgraduate students in the School of Mathematics at the University of South Australia, Flinders University and Macquarie University, for their support, enthusiasm and friendship.

Last but not least, I would like to give my special gratitude and love to my family for their unfailing support. Without their love, understandings and encouragements this research project simply would not be possible.

Ludmila Manic
August 31, 2015

Introduction

The thesis aims at the development of mathematical tools for analysis and construction of near optimal solutions of long run average optimal control (LRAOC) problems.

Such problems have attracted interest of many leading researchers (see, e.g., [2], [4], [6], [10], [22], [31], [32], [33], [40], [43], [76], [78], [79], [80], [83], [90], [94], [98] and references therein). The interest to this class of optimal control problems is motivated not only by important applications (in chemical and electrical engineering, optimization of manufacturing systems, environmental modelling, etc.) but also by the fact that finding an optimal control in a general nonlinear case still presents a difficult task.

Our approach to the LRAOC problems is based on the idea of "linearizing" the nonlinear optimal control problem by reformulating it as optimization problem on the space of occupational measures generated by the control-state trajectories (see Section 1.1). The main advantage of the occupational measures approach is that it translates nonlinear optimal control problems into associated infinite dimensional linear programming (IDLP) problems. It is based on the fact that the occupational measures generated by admissible controls and the corresponding solutions of a nonlinear system satisfy certain linear equations representing the system's dynamics in a relaxed integral form. Note that fundamental results that justify the use of IDLP formulations in various problems of optimal control of stochastic systems have been obtained in [24], [50], [77], [99]. Important steps in the development of IDLP formulations in deterministic optimal control problems considered on a finite time interval have been made in [66], [67], [73], [95], [104]. A linear programming/occupational measures approach to deterministic LRAOC problems was considered in [58], where it has been established that these problems are asymptotically equivalent to IDLP problems similar to those arising in stochastic control. Also, in [58], it has been shown that these IDLP problems can be approximated by standard finite dimensional linear programming problems (finite dimensional approximations of IDLP problems arising in stochastic control problems and in deterministic problems on finite intervals of time have been studied in [72], [87], and in [95], respectively; finite dimensional approximations of IDLP problems arising in certain problems of calculus of variations have been considered in [43]), the solution of which can be used for construction of the optimal controls.

The thesis consist of three parts.

In Part I (after reviewing some results about relationships between the LRAOC problem and the corresponding IDLP problem), we show that necessary and sufficient optimality conditions for the LRAOC problem can be stated in terms of a solution of the HJB inequality, the latter is shown to be equivalent to the problem dual with respect to the IDLP problem. Being a max-min type variational problem on the space of continuously differentiable functions, this dual problem is approximated by max-min problems on finite dimensional subspaces of the space of continuously differentiable functions, which are dual to the semi-infinite dimensional linear programming (SILP) problems approximating the IDLP problem. We give conditions under which solutions of these duals exist and can be used for construction of near optimal solutions of the LRAOC problem. We establish the convergence of an linear programming (LP) based algorithm for finding optimal solutions of the SILP problems and their duals, and we demonstrate a construction of a near optimal control based on such solutions with a numerical example. The obtained results were published in [60].

In Part II, we develop tools for analysis and construction of near optimal solutions of singularly perturbed LRAOC problems. Our approach to these problems is based on amalgamation of averaging and linear programming based techniques. We show that an asymptotically near optimal solution of the singularly perturbed problem can be constructed on the basis of an optimal solution of the averaged LRAOC problem and we show that the optimal solution of the latter can be found with the help of linear programming based techniques. Key concepts introduced and dealt with, in this part, are those of optimal and near optimal average control generating (ACG) families. Sufficient and necessary conditions for an ACG family to be optimal are established and an algorithm for finding near optimal ACG families is described and justified. The construction of an asymptotically near optimal control is illustrated with a numerical example.

Note that problems of optimal controls of singularly perturbed systems appear in a variety of applications and have received a great deal of attention in the literature (see, e.g., [3], [11], [13], [18], [22], [27], [36], [38], [39], [40], [42], [45], [47], [54], [57], [68], [73], [74], [77], [81], [87], [89], [92], [93], [101], [102], [104] and references therein). In a number of works (see, e.g., [7], [8], [13], [51], [52], [53], [55]) it has been noted that equating of the singular perturbation parameter to zero may not lead to a right approximation of the optimal solution in the general nonlinear case. Various averaging type approaches allowing one to deal with such cases were proposed in [2], [3], [7], [8], [9], [11], [12], [27], [28], [29], [40], [41], [45], [46], [51], [52], [55], [56], [57], [67], [68], [92], [93], [102] (see also references therein). However, despite of the fact that the literature devoted to the topic is very reach, until recently, no algorithms for finding near optimal solutions (in case equating of the singular perturbation parameter to zero does not lead to the right approximation) have been discussed in the literature. In fact, to the best of our knowledge, first such results were obtained in two recent papers, in [64] (for optimal problem with time discounting) and in [61] (for LRAOC problems), with results of [61] constituting the basis for consideration of Part II.

In Part III, we study families of SILP problems depending on a small parameter. The family of SILP problems is regularly (singularly) perturbed if its optimal value is continuous (discontinuous) at the zero value of the parameter. We introduce a regularity condition such that if it is fulfilled, then the family of SILP problems is regularly perturbed and if it is not fulfilled, then the family is likely to be singularly perturbed. We establish relationships between the regularity condition for SILP problems and regularity conditions used in dealing with perturbed LRAOC problems.

Part I

Use of approximations of Hamilton-Jacobi-Bellman inequality for solving long run average problems of optimal control

1

Infinite-dimensional linear programming problem related to long run average optimal control problem

In this chapter, we introduce notations and results that are used further in the text. The chapter consists of three sections. In Section 1.1, we consider various statements of long run average optimal control (LRAOC) problems. In Section 1.2, we establish relationships between the LRAOC problem and a certain infinite-dimensional linear programming (IDLP) problem. In Section 1.3, we describe duality results for the IDLP problem.

1.1 LRAOC problem statements. Occupational measures formulations of the LRAOC problem.

We will be considering the control system written in the form

$$y'(t) = f(u(t), y(t)), \qquad t \geq 0, \tag{1.1.1}$$

where the function $f(u,y) : U \times I\!R^m \to I\!R^m$ is continuous in (u,y) and satisfies Lipschitz condition in y; and $u(\cdot) : [0,S] \to U$ or $u(\cdot) : [0,+\infty) \to U$ (depending on whether the system is considered on the finite time interval $[0,S]$ or on the infinite time interval $[0,+\infty)$) are controls that are assumed to be Lebesgue measurable functions and taking values in a given compact metric space U.

Definition 1.1.1 *A pair $(u(\cdot), y(\cdot))$ will be called admissible on the interval $[0,S]$ if the equation (1.1.1) is satisfied for almost all $t \in [0,S]$ and if the following inclusions are valid:*

$$u(t) \in U, \qquad y(t) \in Y \quad \forall t \geq 0, \tag{1.1.2}$$

where Y is a given compact subset of $I\!R^m$. The pair will be called admissible on $[0,\infty)$ if it is admissible on any interval $[0,S]$, $\;S > 0$.

Note that, the first inclusion in (1.1.2) being valid for almost all t and the second for all $t \in [0,S]$ (the second inclusion is interpreted as the state constraint).

In what follows, it will be assumed that the system (1.1.1) is viable in Y (that is, for any initial condition in Y, there exists a control that keeps the solution of the system in Y; see [16]).

The optimal control problem that we will be dealing with is defined as follows

$$\inf_{(u(\cdot),y(\cdot))} \liminf_{S\to\infty} \frac{1}{S} \int_0^S q(u(t),y(t))dt \stackrel{\text{def}}{=} V^* \;, \tag{1.1.3}$$

where $q(u,y) : U \times I\!R^m \to I\!R^1$ is a given continuous function and *inf* is sought over all admissible pairs of the system (1.1.1). This problem will be referred to as *long run average optimal control* (LRAOC) problem. Note that, the initial condition is not fixed in (1.1.1) and it is, in fact, a part of the optimization problem.

Also, we will be dealing with the optimal control problem considered on the finite time interval

$$\inf_{(u(\cdot),y(\cdot))} \frac{1}{S} \int_0^S q(u(t),y(t))dt \stackrel{\text{def}}{=} V^*(S), \tag{1.1.4}$$

where *inf* is sought over all admissible pairs on the interval $[0,S]$.

Remark 1.1.2 *Note that, the assumption that the controls take values in the compact set U can be replaced by a weaker assumption that the optimal and near optimal controls belong to this set.*

Proposition 1.1.3 *The following inequality is satisfied*

$$\liminf_{S\to\infty} V^*(S) \leq V^*. \tag{1.1.5}$$

Proof. For any $S > 0$ and for any admissible pair $(u(t), y(t))$

$$\frac{1}{S}\int_0^S q(u(t), y(t))dt \geq V^*(S). \tag{1.1.6}$$

Hence,

$$\liminf_{S\to\infty} \frac{1}{S}\int_0^S q(u(t), y(t))dt \geq \liminf_{S\to\infty} V^*(S). \tag{1.1.7}$$

The last inequality implies that

$$V^* \geq \liminf_{S\to\infty} V^*(S), \tag{1.1.8}$$

which proves (1.1.5). □

Along with the problems defined above, we will be referring to the infinite time horizon optimal control problem

$$\inf_{(u(\cdot),y(\cdot))} \lim_{S\to\infty} \frac{1}{S}\int_0^S q(u(t), y(t))dt \stackrel{\text{def}}{=} V^*_\infty \ , \tag{1.1.9}$$

where *inf* is sought over all admissible pairs on the interval $[0, \infty)$ such that the limit in the above expression exists. If this *inf* is sought over the periodic admissible pairs only, that is, over the admissible pairs such that

$$(u(t), y(t)) = (u(t + T), y(t + T)) \qquad \forall t \geq 0 \tag{1.1.10}$$

for some $T > 0$, then (1.1.9) is written as

$$\inf_{(u(\cdot),y(\cdot))_{per}} \lim_{S\to\infty} \frac{1}{S}\int_0^S q(u(t), y(t))dt.$$

Thus, it becomes equivalent to a so-called *periodic optimization problem* (see, e.g., [33])

$$\inf_{T,(u(\cdot),y(\cdot))} \frac{1}{T}\int_0^T q(u(t), y(t))dt \stackrel{\text{def}}{=} V^*_{per}, \tag{1.1.11}$$

where *inf* is over the length T of the time interval and over the admissible pairs defined on $[0, T]$, which satisfies the periodicity condition $y(0) = y(T)$.

A very special family of admissible pairs on $[0, \infty)$ is that consisting of constant valued controls and corresponding steady state solutions of (1.1.1):

$$(u(t), y(t)) = (u, y) \in M \stackrel{\text{def}}{=} \{(u, y) \mid (u, y) \in U \times Y, \quad f(u, y) = 0\}. \tag{1.1.12}$$

If *inf* is sought over admissible pairs from this family then the problem (1.1.9) is reduced to

$$\inf_{(u,y)\in M} q(u,y) \stackrel{\text{def}}{=} V_{ss}^*, \tag{1.1.13}$$

which is called *a steady state optimization problem.*

It is easy to see that the optimal values of the above introduced problems satisfy the inequalities

$$V^* \le V_\infty^* \le V_{per}^* \le V_{ss}^*. \tag{1.1.14}$$

Note that, in the general case,

$$V_{per}^* < V_{ss}^* \tag{1.1.15}$$

(see, e.g., [52], [65], [69] and [71]). Allowing this to be the case, we will be assuming that

$$V^* = V_\infty^* \quad \text{or} \quad V^* = V_{per}^*. \tag{1.1.16}$$

For the sake of our consideration let us reformulate the problems in terms of occupational measures.

The occupational measure generated by an admissible control and the corresponding solution of the system (1.1.1) (that is, it is generated by an admissible pair $(u(t), y(t))$ of the system (1.1.1)) on interval $[0, S]$ is the probability measure defined by the "proportions" of time spent by this admissible pair in different subsets of the control-state space. More precisely, let $\mathcal{P}(U\times Y)$ stands for the space of probability measures defined on the Borel subsets of $U \times Y$.

Definition 1.1.4 *A probability measure $\gamma^{(S,(u(\cdot),y(\cdot))} \in \mathcal{P}(U \times Y)$ is called the occupational measure generated by the admissible pair $(u(t), y(t))$ on the interval $[0, S]$ if, for any Borel set $B \subset U \times Y$,*

$$\gamma^{(S,(u(\cdot),y(\cdot)))}(B) \stackrel{\text{def}}{=} \frac{1}{S}\int_0^S 1_B(u(t), y(t))dt, \tag{1.1.17}$$

where $1_B(\cdot)$ is the indicator function of the set B: $1_B(u,y) = 1 \quad \forall (u,y) \in B$ and $1_B(u,y) = 0 \quad \forall (u,y) \notin B$.

Definition 1.1.5 *The occupational measure generated by the admissible pair on the interval $[0,\infty)$ is the probability measure $\gamma^{(u(\cdot),y(\cdot))} \in \mathcal{P}(U \times Y)$ defined as the limit (assumed to exist)*

$$\gamma^{(u(\cdot),y(\cdot))}(B) \stackrel{\text{def}}{=} \lim_{S\to\infty} \frac{1}{S}\int_0^S 1_B(u(t), y(t))dt. \tag{1.1.18}$$

Note that the occupational measure generated by a steady state admissible pair $(u(t), y(t)) = (u, y) \in M$ (as in (1.1.12)) is just the Dirac measure at (u, y).

Note also, that the Definition 1.1.4 is equivalent to the statement that, for any continuous function $h(\cdot)$ defined on the control-state space, the time average of the integral of the function $h(\cdot)$ along an admissible pair is equal to the integral of $h(\cdot)$ over the occupational measures generated by this admissible pair. Namely, (1.1.17) is equivalent to that

$$\int_{U\times Y} h(u,y)\gamma^{(S,(u(\cdot),y(\cdot)))}(du,dy) = \frac{1}{S}\int_0^S h(u(t),y(t))dt \tag{1.1.19}$$

for any continuous $h(u, y) : U \times I\!R^m \to I\!R^1$.

Similarly, the Definition 1.1.5 is equivalent to the statement that

$$\int_{U\times Y} h(u,y)\gamma^{(u(\cdot),y(\cdot))}(du,dy) = \lim_{S\to\infty}\frac{1}{S}\int_0^S h(u(t),y(t))dt \tag{1.1.20}$$

for any continuous $h(\cdot) \in C(U \times Y)$. Note that, from (1.1.20) it follows, in particular, that there exist the limit

$$\lim_{S\to\infty} \gamma^{(S,(u(\cdot),y(\cdot)))} \stackrel{\text{def}}{=} \gamma^{(u(\cdot),y(\cdot))}.$$

Let us denote by $\Gamma(S) \subset \mathcal{P}(U \times Y)$ the set of all occupational measures generated by the admissible pairs $(u(\cdot), y(\cdot))$ on the interval $[0, S]$. That is,

$$\Gamma(S) \stackrel{\text{def}}{=} \bigcup_{u(\cdot)}\{\gamma^{(S,(u(\cdot),y(\cdot)))}\} \subset \mathcal{P}(U\times Y), \tag{1.1.21}$$

where $\gamma^{(S,(u(\cdot),y(\cdot)))}$ is the occupational measure generated by $(u(\cdot), y(\cdot))$ on the interval $[0, S]$ and the union is over all controls.

Using this notation and the definition of the occupational measures as in (1.1.19), one can rewrite problem (1.1.4) in terms of minimization over measures from the set $\Gamma(S)$ as follows

$$\inf_{\gamma\in\Gamma(S)}\int_{U\times Y} q(u,y)\gamma(du,dy) = V^*(S). \tag{1.1.22}$$

Note that, in what follows, the convergence properties of $V^*(S)$ (as S tends to infinity) are established on the basis of the corresponding convergence properties of $\Gamma(S)$ which are defined at the end of this section.

The periodic optimization problem (1.1.11) also can be rewritten in terms of occupational measures. Namely, based on (1.1.19),

$$\inf_{\gamma\in\Gamma_{per}} \int_{U\times Y} q(u,y)\gamma(du,dy) = V^*_{per}, \tag{1.1.23}$$

where Γ_{per} defines the set of all occupational measures generated by the periodic admissible pairs $(u(\cdot),y(\cdot))_{per}$ (that is, (1.1.10) is satisfied with some positive T).

Note that, due to linearity of the objective function in (1.1.22),

$$\min_{\gamma\in\bar{co}\Gamma(S)} \int_{U\times Y} q(u,y)\gamma(du,dy) = V^*(S), \tag{1.1.24}$$

where $\bar{co}$ stands for the closed convex hull of the set $\Gamma(S)$. Similarly, due to linearity of the objective function in (1.1.23),

$$\min_{\gamma\in\bar{co}\Gamma_{per}} \int_{U\times Y} q(u,y)\gamma(du,dy) = V^*_{per}. \tag{1.1.25}$$

Let us conclude this section with some comments and notation. Given a compact metric space X, $\mathcal{B}(X)$ will stand for the σ-algebra of its Borel subsets and $\mathcal{P}(X)$ will denote the set of probability measures defined on $\mathcal{B}(X)$. The set $\mathcal{P}(X)$ will always be treated as a compact metric space with a metric ρ, which is consistent with its weak* topology (see, e.g., [23] or [91]). That is, a sequence $\gamma^k\in\mathcal{P}(X),\ \ k=1,2,...$, converges to $\gamma\in\mathcal{P}(X)$ in this metric if and only if

$$\lim_{k\to\infty}\int_X c(x)\gamma^k(dx) \ = \ \int_X c(x)\gamma(dx), \tag{1.1.26}$$

for any continuous $c(\cdot): X\to I\!R^1$. There are many ways of how such a metric ρ can be defined. We will use the following definition: $\forall\gamma',\gamma''\in\mathcal{P}(X)$,

$$\rho(\gamma',\gamma'') \stackrel{\text{def}}{=} \sum_{l=1}^{\infty}\frac{1}{2^l} \ \left| \int_X h_l(x)\gamma'(dx) - \int_X h_l(x)\gamma''(dx) \right| , \tag{1.1.27}$$

where $h_l(\cdot),\ \ l=1,2,\ldots$, is a sequence of Lipschitz continuous functions which is dense in the unit ball of $C(X)$ (the space of continuous functions on X).

Using this metric ρ, one can define the "distance" $\rho(\gamma,\Gamma)$ between $\gamma\in\mathcal{P}(X)$ and $\Gamma\subset\mathcal{P}(X)$, and the Hausdorff metric $\rho_H(\Gamma_1,\Gamma_2)$ between $\Gamma_1\subset\mathcal{P}(X)$ and $\Gamma_2\subset\mathcal{P}(X)$, as follows:

$$\rho(\gamma,\Gamma) \stackrel{\text{def}}{=} \inf_{\gamma'\in\Gamma} \rho(\gamma,\gamma'), \qquad \rho_H(\Gamma_1,\Gamma_2) \stackrel{\text{def}}{=} \max\{\sup_{\gamma\in\Gamma_1} \rho(\gamma,\Gamma_2), \sup_{\gamma\in\Gamma_2} \rho(\gamma,\Gamma_1)\}. \tag{1.1.28}$$

Note that, although, by some abuse of terminology, we refer to $\rho_H(\cdot,\cdot)$ as to a metric on the set of subsets of $\mathcal{P}(X)$, it is, in fact, a semi-metric on this set (since $\rho_H(\Gamma_1,\Gamma_2) = 0$ is equivalent to $\Gamma_1 = \Gamma_2$ if and only if Γ_1 and Γ_2 are closed).

The following lemma is implied by the definitions above.

Lemma 1.1.6 *Let Γ be a subset of $\mathcal{P}(X)$ then:*

(i) if $\lim_{S\to\infty} \sup_{\gamma\in\Gamma(S)} \rho(\gamma,\Gamma) = 0$, then, for any continuous $h(x) : X \to \mathbb{R}^1$,

$$\liminf_{S\to\infty} \inf_{\gamma\in\Gamma(S)} \int_X h(x)\gamma(dx) \geq \inf_{\gamma\in\Gamma} \int_X h(x)\gamma(dx);$$

(ii) if $\lim_{S\to\infty} \rho_H(\Gamma(S),\Gamma) = 0$, then

$$\lim_{S\to\infty} \inf_{\gamma\in\Gamma(S)} \int_X h(x)\gamma(dx) = \inf_{\gamma\in\Gamma} \int_X h(x)\gamma(dx).$$

Proof. The proof is obvious. □

It can be verified (see e.g. Lemma П2.4 in [51], p.205) that, with the definition of the metric ρ as in (1.1.27),

$$\rho_H(\bar{co}\Gamma_1, \bar{co}\Gamma_2) \leq \rho_H(\Gamma_1,\Gamma_2)\ , \tag{1.1.29}$$

where $\bar{co}$ stands for the closed convex hull of the corresponding set.

1.2 Infinite dimensional linear programming problem.

Let us define the set $W \subset \mathcal{P}(U \times Y)$ by the equation

$$W \stackrel{\text{def}}{=} \Big\{\gamma \in \mathcal{P}(U\times Y) : \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma(du,dy) = 0 \quad \forall\phi \in C^1\Big\}, \tag{1.2.1}$$

where C^1 is the space of continuously differentiable functions $\phi(y) : I\!R^m \to I\!R^1$, $\nabla\phi(y)$ is a vector column of partial derivatives (the gradient) of $\phi(y)$,

$$\nabla\phi(y) \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\partial\phi}{\partial y_1} \\ \vdots \\ \frac{\partial\phi}{\partial y_m} \end{bmatrix}. \tag{1.2.2}$$

Proposition 1.2.1 *The set W is convex and compact.*

Proof. Note that in order to show that W is compact it is enough to establish that it is closed, since $W \subset \mathcal{P}(U \times Y)$ and $\mathcal{P}(U \times Y)$ is compact. Consider a sequence $\gamma^k \in W$ such that

$$\lim_{k\to\infty} \gamma^k = \bar{\gamma},$$

where $\bar{\gamma}$ is a boundary point. Note that, since $\gamma^k \in W$,

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma^k(du,dy) = 0 \qquad \forall k = 1, 2, ... \quad . \tag{1.2.3}$$

Therefore,

$$\lim_{k\to\infty} \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma^k(du,dy) = \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\bar{\gamma}(du,dy) = 0. \tag{1.2.4}$$

Hence, $\bar{\gamma} \in W$ and thus the set W is closed.

In order to show that the set W is convex one can observe that for any $\gamma', \gamma'' \in W$ and any $\alpha \in [0,1]$,

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y)[\alpha\gamma' - (1-\alpha)\gamma''](du,dy) = \alpha \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma'(du,dy) +$$

$$(1-\alpha) \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma''(du,dy) = 0.$$

□

Note that the set W can be empty. It is easy to see, for example, that W is empty if there exists a continuously differentiable function $\phi(\cdot) \in C^1$ such that

$$\max_{(u,y)\in U\times Y} (\nabla\phi(y))^T \ f(u,y) < 0. \tag{1.2.5}$$

The set W is not empty if the set of steady state or periodic admissible pairs is not empty since the occupational measure generated by each such pair is contained in W.

In fact, let $(u(\cdot), y(\cdot))$ be a periodic admissible pair (that is, (1.1.10) is satisfied with some positive T) and let $\gamma^{(u(\cdot),y(\cdot))}$ be the occupational measure generated by this pair on the interval $[0, T]$. Then, by (1.1.19),

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma^{(u(t),y(t))}(du,dy) = \frac{1}{T}\int_0^T (\nabla\phi(y(t)))^T f(u(t),y(t))dt$$

$$= \frac{\phi(y(T)) - \phi(y(0))}{T} = 0 \qquad \forall\phi(\cdot)\in C^1 \qquad \Rightarrow \qquad \gamma^{(u(\cdot),y(\cdot))} \in W.$$

Proposition 1.2.2 *If the set W is empty, then there exists $S_0 > 0$ such that $\Gamma(S)$ is empty for $S \geq S_0$. If $\Gamma(S)$ is not empty for all $S > 0$ large enough, then W is not empty and*

$$\lim_{S\to\infty} \sup_{\gamma\in\Gamma(S)} \rho(\gamma, W) = 0. \tag{1.2.6}$$

Proof. To prove the validity of (1.2.6), let us define $k(S)$ by the equation

$$k(S) \stackrel{\text{def}}{=} \sup_{\gamma\in\Gamma(S)} \rho(\gamma, W) \tag{1.2.7}$$

and show that $k(S)$ tends to zero as S tends to infinity. Assume it is not the case. Then, there exists a positive number δ and the sequence $S^k \to \infty$, $\gamma^k \in \Gamma(S^k)$, such that

$$\rho(\gamma^k, W) \geq \delta \qquad \forall k = 1, 2, \dots .$$

Without loss of generality, one may assume that there exists

$$\lim_{k\to\infty} \gamma^k \stackrel{\text{def}}{=} \gamma \in \mathcal{P}(U\times Y)$$

(since $\mathcal{P}(U\times Y)$ is compact). From the continuity of the metric it follows that

$$\rho(\gamma, W) \geq \delta. \tag{1.2.8}$$

By the definition of the convergence in $\mathcal{P}(U\times Y)$ (see (1.1.26)),

$$\lim_{k\to\infty} \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma^k(du,dy) = \int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma(du,dy) \tag{1.2.9}$$

for any $\phi \in C^1$. Also, from the fact that $\gamma^k \in \Gamma(S^k)$ it follows that there exists an admissible pair $(u^k(t), y^k(t))$ defined on the interval $[0, S^k]$ such that

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma^k(du,dy) = \frac{1}{S^k}\int_0^{S^k} (\nabla\phi(y^k(t)))^T f(u^k(t), y^k(t))dt. \qquad (1.2.10)$$

The second integral is apparently equal to

$$\frac{\phi(y^k(S^k)) - \phi(y^k(0))}{S^k}$$

and tends to zero as S^k tends to infinity (since $y^k(t) \in Y \quad \forall t \in [0, S^k]$ and Y is a compact set). This and (1.2.9) imply that

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y)\gamma(du,dy) = 0 \qquad \forall \phi \in C^1 \qquad \Rightarrow \qquad \gamma \in W.$$

The latter contradicts (1.2.8) and, hence, $k(S)$ defined in (1.2.7) tends to zero as S tends to infinity. This proves (1.2.6).

From consideration above it follows that, if there exists a sequence of S^k tending to infinity such that $\Gamma(S^k) \neq \emptyset$, then the set W is not empty. Hence, if the latter is empty, then $\Gamma(S) = \emptyset$ for all S large enough. $\square$

In what follows, we will assume that W is not empty.

Let us consider the problem

$$\inf_{\gamma\in W} \int_{U\times Y} q(u,y)\gamma(du,dy) \stackrel{\text{def}}{=} G^*, \qquad (1.2.11)$$

where $q(\cdot)$ is the same as in (1.1.22) (and the same as in (1.1.3)-(1.1.4)).

Note that, since both the objective function in (1.2.11) and the constraints in W are linear in γ, problem (1.2.11) is that of *infinite-dimensional linear programming* (IDLP) (see, e.g., [5]).

Corollary 1.2.3 *The lower limit of the optimal values of (1.1.4) satisfies the inequality*

$$\underline{\lim}_{S\to\infty} V^*(S) = \underline{\lim}_{S\to\infty} \inf_{\gamma\in\Gamma(S)} \int_{U\times Y} q(u,y)\gamma(du,y) \geq G^*. \qquad (1.2.12)$$

Proof. The proof follows from Lemma 1.1.6 (i), Proposition 1.2.2, and the validity of the representation (1.1.22). $\square$

Corollary 1.2.4 *(criteria of optimality)* *(i) If an admissible pair* $(u(\cdot), y(\cdot)) : [0, \infty) \to U \times Y$ *is such that*

$$\lim_{S \to \infty} \frac{1}{S} \int_0^S q(u(t), y(t))dt = G^*,$$

then this pair is a solution of the problem (1.1.9) and $V_\infty^* = G^*$*;*

(ii) If a periodic (with a period T*) admissible pair* $(u(\cdot), y(\cdot))$ *is such that*

$$\frac{1}{T} \int_0^T q(u(t), y(t))dt = G^*,$$

then this pair is a solution of the problems (1.1.9) and (1.1.11), and also $V_\infty^* = V_{per}^* = G^*$*;*

(iii) If a steady state admissible pair $(u(t), y(t)) = (u, y) \in M$ *(as defined in (1.1.12)) is such that*

$$q(u, y) = G^*,$$

then this pair is a solution of the problems (1.1.9), (1.1.11) and (1.1.13), and also $V_\infty^* = V_{per}^* = V_{ss}^* = G^*$*.*

Proof. The proof follows from inequalities (1.1.14) and Corollary 1.2.3. □

It has been shown in [62] (see also [48], [57], [58]) that, if W is not empty and some other mild conditions are satisfied, then the following relationships are valid:

$$\lim_{S \to \infty} \rho_H(co\Gamma(S), W) = 0 \tag{1.2.13}$$

$$\Rightarrow \qquad \lim_{S \to \infty} V^*(S) = G^*. \tag{1.2.14}$$

Note that no assumptions, except of non-emptiness of W, for the validity of these relationships are needed if one allows the use of the relaxed controls (see Theorem 3.1 and 3.3 in [48] or Remark 4.5 in [62]). Note also that the relationships similar to (1.2.13) and (1.2.14) have been established for various optimal control formulations (for problems considered on finite and infinite time horizons and in both deterministic and stochastic settings) in [24], [43], [50], [73], [77], [81], [95], [98] (see also references therein).

From Proposition 1.2.2 and the equality (1.2.14) it follows that if the solution γ^* of the problem (1.2.11) is unique, then, for any $\gamma^S \in \Gamma(S)$ such that

$$\lim_{S\to\infty}\int_{U\times Y} q(u,y)\gamma^S(du,dy) = G^*,$$

$$\lim_{S\to\infty}\rho(\gamma^S,\gamma^*) = 0. \tag{1.2.15}$$

Note also that the solution γ^* of problem (1.2.11) can be unique only if it is an extreme point of W (since (1.2.11) is an LP problem) and that, using (1.2.13), one can show (although not shown here) that, for any extreme point γ of W, there exists $\gamma^S \in \Gamma(S)$ such that $\lim\limits_{S\to\infty}\rho(\gamma^S,\gamma) = 0$.

Let γ^* be a solution of problem (1.2.11) which is an extreme point of W and let $\gamma^S \in \Gamma(S)$ satisfy (1.2.15). Assume that there exists an admissible pair $(u^{\gamma^*}(\cdot), y^{\gamma^*}(\cdot)) : [0,\infty) \to U \times Y$ that generates γ^S on any interval $[0,S]$ (see Definition 1.1.5). Then, for any continuous $h(u,y) : U \times Y \to I\!R^1$,

$$\lim_{S\to\infty}\frac{1}{S}\int_0^S h(u^{\gamma^*}(t), y^{\gamma^*}(t))dt = \int_{U\times Y} h(u,y)\gamma^*(du,dy) \tag{1.2.16}$$

and, in particular, for $h(u,y) = q(u,y)$,

$$\lim_{S\to\infty}\frac{1}{S}\int_0^S q(u^{\gamma^*}(t), y^{\gamma^*}(t))dt = \int_{U\times Y} q(u,y)\gamma^*(du,dy) = G^*. \tag{1.2.17}$$

Thus, by Corollary 1.2.4(i), this pair will be a solution of problem (1.1.9). Also, by Corollary 1.2.4(ii), (iii), this pair will be a solution of the periodic optimization problem (1.1.11) (and the steady state problem (1.1.13)) if it proves to be periodic (and, respectively, steady state).

Note that, in what follows it will be assumed everywhere that $\Gamma(S) \neq \emptyset$ (this implying that $W \neq \emptyset$ and that $V^*(S)$ is well defined; see (1.1.22)) and also that (1.2.13) and (1.2.14) are valid.

1.3 The problem dual to the IDLP problem and duality relationships.

Define the problem dual to IDLP problem (1.2.11) by the equation

$$\sup_{(d,\eta(\cdot))\in\mathcal{D}} d \stackrel{\text{def}}{=} D^*, \tag{1.3.1}$$

with the feasible set $\mathcal{D} \subset I\!R^1 \times C^1$ defined as

$$\mathcal{D} \stackrel{\text{def}}{=} \Big\{(d, \eta(\cdot)) : \; d = \min_{(u,y)\in U\times Y} \{q(u,y) + (\nabla\eta(y))^T f(u,y)\}, \;\; \eta(\cdot) \in C^1\Big\}. \tag{1.3.2}$$

The way how the problem (1.3.1) can be constructed as a "standard" LP dual is discussed in Appendix A.

Note that, if $W \neq \emptyset$, then, for any $\gamma \in W$, from the fact that the pair $(d, \eta(\cdot))$ satisfies the inequality

$$d \leq q(u,y) + (\nabla\eta(y))^T f(u,y) \quad \forall (u,y) \in U \times Y,$$

it follows that

$$d \leq \int_{U\times Y} q(u,y)\gamma(du,dy).$$

Hence, the optimal values of (1.2.11) and its dual (1.3.1) satisfy the inequality

$$D^* \leq G^*. \tag{1.3.3}$$

The following statements establish more elaborate connections between problems (1.2.11) and (1.3.1). Namely, from the theorem stated below follows that the inequality (1.3.3) turns into the equality (that is, there is no duality gap) if and only if W is not empty.

Theorem 1.3.1 *(i) The optimal value of the dual problem (1.3.1) is bounded (that is, $D^* < \infty$) if and only if the set W is not empty;*

(ii) If the optimal value of the dual problem (1.3.1) is bounded, then

$$D^* = G^*; \tag{1.3.4}$$

(iii) The optimal value D^* of the dual problem is unbounded (that is, $D^* = \infty$) if and only if there exists a function $\bar{\eta}(\cdot) \in C^1$ such that

$$\max_{(u,y)\in U\times Y} (\nabla\bar{\eta}(y))^T f(u,y) < 0. \tag{1.3.5}$$

Proof. The proof of the theorem was given in Section 9 of [48] and for completeness it is recalled in Appendix B. □

Note that duality results similar to Theorem 1.3.1 (ii) have been obtained in [24] and [50] in a stochastic setting without state constraints ($Y = I\!R^m$) and in [104] in the deterministic setting with state constraints (for IDLP problems related to optimal control problems considered on a finite time interval).

Remark 1.3.2 *From the statements (i) and (ii) of the theorem stated above it follows that the set W and, hence, the set $\Gamma(S)$ are not empty (see Proposition 1.2.2) if and only if a function $\eta(\cdot) \in C^1$ satisfying (1.3.5) does not exist. Note that, if such a function $\eta(\cdot)$ exists, then the fact that $\Gamma(S)$ are empty for $S \geq S_0$ (for some $S_0 > 0$) follow from the fact that this $\eta(\cdot)$ can be used as a Liapunov function decreasing along the trajectories of the system (1.1.1) and "forcing" them to leave Y in a finite time.*

Note that problem (1.3.1) is equivalent to

$$\sup_{\eta(\cdot)\in C^1} \min_{(u,y)\in U\times Y} \{q(u,y) + (\nabla\eta(y))^T f(u,y)\} = D^*. \tag{1.3.6}$$

Defining Hamiltonian $H(p,y)$ by the equation

$$H(p,y) \stackrel{\text{def}}{=} \min_{u\in U}\{q(u,y) + p^T f(u,y)\} \tag{1.3.7}$$

one can rewrite the equality (1.3.6) (or, equivalently, (1.3.1)) as follows

$$\sup_{\eta(\cdot)\in C^1} \min_{y\in Y} H(\nabla\eta(y), y) = D^*. \tag{1.3.8}$$

Definition 1.3.3 *Assume that $D^* < \infty$. A function $\eta^*(\cdot) \in C^1$ will be called a solution of the dual problem (1.3.1) if*

$$D^* = \min_{(u,y)\in U\times Y}\{q(u,y) + (\nabla\eta^*(y))^T f(u,y)\}. \tag{1.3.9}$$

Note that, by rewriting equation (1.3.9) in the form

$$D^* = \min_{y\in Y} H(\nabla\eta^*(y), y) \qquad \Rightarrow \qquad D^* \leq H(\nabla\eta^*(y), y) \quad \forall y \in Y, \tag{1.3.10}$$

one can come to the conclusion that a solution $\eta^*(\cdot)$ of dual problem (1.3.1) is a smooth viscosity subsolution of the corresponding Hamilton-Jacobi-Bellman equation (see [20] and [49] for relevant definitions and developments).

1.4 Additional comments for Chapter 1

The consideration of Sections 1.1 and 1.2 is based on results obtained in [58] and [57]. The consideration of Section 1.3 is based on results of [48].

2

Necessary and sufficient conditions of optimality. Maxi-min problem and its approximation

In this chapter, we show that sufficient and necessary conditions of optimality in long run average optimal control (LRAOC) problems can be stated in terms of a solution of the corresponding Hamilton-Jacobi-Bellman (HJB) inequality, the latter being equivalent to the problem dual to the infinite dimensional linear programming (IDLP) problem considered in Section 1.2. The latter is a max-min type variational problem considered on the space of continuously differentiable functions. We approximate it with the max-min problems on a finite dimensional subspaces of the space of continuously differentiable functions.

The chapter is organised as follows. In Section 2.1, we define the HJB inequality and show that it can be used to formulate necessary and sufficient conditions of optimality for the LRAOC problem. In Section 2.2, we establish that the HJB inequality is equivalent to the variational max-min problem and consider approximation of this problem by max-min problems on finite dimensional subspaces of the space of continuously differentiable functions. We give conditions under which solutions of these exist and can be used for construction of near optimal solutions of the LRAOC problem. Semi-infinite dimensional linear programming (SILP) problems and their relationships with the approximating max-min problems are discussed in Section 2.3.

2.1 Necessary and sufficient conditions of optimality based on the HJB inequality.

In this section, we give sufficient and necessary conditions for an admissible pair $(u(\cdot), y(\cdot))$ to be optimal and for the equality

$$V^* = G^* \tag{2.1.1}$$

to be valid. Note that the equality (2.1.1) implies that

$$\lim_{S \to \infty} V^*(S) = G^*. \tag{2.1.2}$$

Under the assumption that (2.1.1) is satisfied, the Hamilton-Jacobi-Bellman (HJB) equation for the LRAOC problem is written in the form (see, e.g., Section VII.1.1 in [20])

$$H(\nabla\eta(y), y) = G^*, \tag{2.1.3}$$

where $H(p, y)$ is the Hamiltonian defined in (1.3.7). The equation (2.1.3) is equivalent to the following two inequalities

$$H(\nabla\eta(y), y) \leq G^*, \qquad H(\nabla\eta(y), y) \geq G^*. \tag{2.1.4}$$

As follows from the result below, for a characterization of an optimal control problem (1.1.3), it is sufficient to consider functions that satisfy only the second inequality in (2.1.4), and we will say that a function $\eta(\cdot) \in C^1$ *is a solution of the HJB inequality on Y* if

$$H(\nabla\eta(y), y) \geq G^*, \quad \forall y \in Y. \tag{2.1.5}$$

Note that the concept of a solution of the HJB inequality on Y introduced above is essentially the same as that of a smooth viscosity subsolution of the HJB equation (2.1.3) considered on the interior of Y (see, e.g., [20]).

The following result gives sufficient condition for an admissible pair $(u(\cdot), y(\cdot))$ to be optimal and for the equality (2.1.1) to be valid.

Proposition 2.1.1 *Assume that a solution $\eta(\cdot) \in C^1$ of the HJB inequality (2.1.5) exists. Then an admissible pair $(u(\cdot), y(\cdot))$ is optimal in (1.1.3) and the equality (2.1.1) is valid if*

$$u(t) = argmin_{u \in U}\{q(u, y(t)) + (\nabla\eta(y(t)))^T f(u, y(t))\} \qquad a.e.\ t \in [0, \infty) \tag{2.1.6}$$

and

$$H(\nabla\eta(y(t)), y(t)) = G^*, \qquad \forall t \in [0, \infty). \tag{2.1.7}$$

Proof. Note that from (2.1.6) and (2.1.7) it follows that for almost all $t \in [0, \infty)$

$$q(u(t), y(t)) + (\nabla\eta(y(t)))^T f(u(t), y(t)) = G^*. \tag{2.1.8}$$

Since,

$$\begin{aligned} \lim_{S\to\infty} \frac{1}{S}\int_0^S (\nabla\eta(y(t)))^T f(u(t), y(t))dt &= \lim_{S\to\infty} \frac{1}{S}\int_0^S \frac{d\eta(y(t))}{dt} \\ &= \lim_{S\to\infty} \frac{1}{S}(\eta(y(S)) - \eta(y(0)) = 0. \end{aligned} \tag{2.1.9}$$

From (2.1.8) it follows that

$$\lim_{S\to\infty} \frac{1}{S}\int_0^S q(u(t), y(t))dt = G^*. \tag{2.1.10}$$

The latter implies that $(u(\cdot), y(\cdot))$ is optimal and that the equality (2.1.1) is valid. □

Let us introduce the following assumption.

Assumption 2.1.2 *The following conditions are satisfied:*

(i) the optimal solution of the problem (1.1.3) (that is, an admissible pair $(u^*(\cdot)$, $y^*(\cdot))$ that delivers minimum in (1.1.3)) exists and generates the occupational measure γ^*. That is, for any continuous $h(u, y)$,

$$\lim_{S\to\infty} \frac{1}{S}\int_0^S h(u^*(t), y^*(t))dt = \int_{U\times Y} h(u, y)\gamma^*(du, dy); \tag{2.1.11}$$

(ii) for almost all $t \in [0, \infty)$ and for any $r > 0$, the γ^*- measure of the set

$$B_r(u^*(t), y^*(t)) \stackrel{\text{def}}{=} \{(u, y) : ||u - u^*(t)|| + ||y - y^*(t)|| < r\} \tag{2.1.12}$$

is not zero. That is,

$$\gamma^*(B_r(u^*(t), y^*(t))) > 0. \tag{2.1.13}$$

The following proposition gives sufficient conditions for the validity of Assumption 2.1.2.

Proposition 2.1.3 *The Assumptions 2.1.2 is satisfied if the pair $(u^*(t), y^*(t))$ is T-periodic (T is some positive number) and if $u^*(\cdot)$ is piecewise continuous on $[0, T]$.*

Proof. Let t be a continuity point of $u^*(t)$. Note that, due to the assumed periodicity of the pair $(u^*(\cdot), y^*(\cdot))$, Assumption 2.1.2 (i) is satisfied. That is,

$$\frac{1}{T}\int_0^T h(u^*(t), y^*(t))dt = \int_{U\times Y} h(u, y)\gamma^*(du, dy)$$

and

$$\gamma^*(B_r(u^*(t), y^*(t))) = \frac{1}{T} meas\Big\{s \in [0, T] : \ (u^*(s), y^*(s)) \in B_r(u^*(t), y^*(t))\Big\}. \tag{2.1.14}$$

Since t is a continuity point of $u^*(\cdot)$ and since $y^*(\cdot)$ is continuous, there exists $\alpha > 0$ such that $(u^*(t'), y^*(t')) \in B_r(u^*(t), y^*(t)) \quad \forall t' \in [t - \alpha, t + \alpha]$. Hence, the right-hand-side in (2.1.14) is greater that $\frac{2\alpha}{T}$. This proves the required statement as the number of discontinuity points of $u^*(\cdot)$ is finite (due to the assumed piecewise continuity). $\square$

Proposition 2.1.4 *Let $(u^*(t), y^*(t))$ be an optimal admissible pair such that Assumption 2.1.2 is satisfied and let the equality (2.1.1) be valid then (2.1.6) and (2.1.7) are satisfied.*

Proof. Since $(u^*(\cdot), y^*(\cdot))$ that delivers minimum in (1.1.3) exists and since this pair generates the occupational measure γ^*, (see (2.1.11))

$$\lim_{S\to\infty} \frac{1}{S}\int_0^S q(u^*(t), y^*(t)) = \int_{U\times Y} q(u, y)\gamma^*(du, dy) = G^*. \tag{2.1.15}$$

Note also that from (2.1.13) and from the fact that $(u^*(t), y^*(t)) \in U \times Y$ it follows that for any $r > 0$

$$\gamma^*\Big(B_r(u^*(t), y^*(t)) \bigcap (U \times Y)\Big) > 0. \tag{2.1.16}$$

Since $\gamma^* \in W$,

$$\int_{U\times Y} (\nabla\eta(y))^T f(u, y)\gamma^*(du, dy) = 0. \tag{2.1.17}$$

Hence, by (2.1.15)

$$\int_{U\times Y}[q(u,y)+(\nabla\eta(y))^T f(u,y)-G^*]\gamma^*(du,dy)=0. \tag{2.1.18}$$

Define the set,

$$B\stackrel{\text{def}}{=}\Big\{(u,y)\in U\times Y:\quad q(u,y)+(\nabla\eta(y))^T f(u,y)-G^*>0\Big\}. \tag{2.1.19}$$

Note that from (1.3.7) and (2.1.5) it follows that

$$q(u,y)+(\nabla\eta(y))^T f(u,y)\geq G^*,\qquad \forall(u,y)\in U\times Y. \tag{2.1.20}$$

Hence, from (2.1.18) it follows that

$$\gamma^*(B)=0. \tag{2.1.21}$$

Let us show that

$$meas\{t:(u^*(t),y^*(t))\in B\}=0. \tag{2.1.22}$$

Assume it is not true. That is, $meas\{t:(u^*(t),y^*(t))\in B\}>0$. Then there exists $\bar{t}$ such that

$$(u^*(\bar{t}),y^*(\bar{t}))\in B \tag{2.1.23}$$

and such that

$$\gamma^*\Big(B_r(u^*(\bar{t}),y^*(\bar{t}))\bigcap(U\times Y)\Big)>0\qquad\forall r>0. \tag{2.1.24}$$

From the definition of the set B (see (2.1.19)) it follows that for $r>0$ small enough

$$B_r(u^*(\bar{t}),y^*(\bar{t}))\bigcap(U\times Y)\subset B. \tag{2.1.25}$$

Hence,

$$\gamma^*(B)\geq\gamma^*\Big(B_r(u^*(\bar{t}),y^*(\bar{t}))\bigcap(U\times Y)\Big)>0. \tag{2.1.26}$$

That contradicts (2.1.21) and hence proves (2.1.22). That is, due to (2.1.20)

$$q(u^*(t),y^*(t))+(\nabla\eta(y^*(t)))^T f(u^*(t),y^*(t))=G^*, \tag{2.1.27}$$

for almost all $t>0$. This implies (2.1.6) and (2.1.7). □

Remark 2.1.5 *Note that the difference of Propositions 2.1.1 and 2.1.4 from similar results of optimal control theory is that a solution of the HJB inequality (rather than* *that of the HJB equation) is used in the right-hand-side of (2.1.6), with the relationship (2.1.7) indicating that the HJB inequality takes the form of the equality on the optimal trajectory. Note also that, due to (2.1.5), the equality (2.1.7) is equivalent to the inclusion*

$$y(t) \in argmin_{y\in Y}\{H(\nabla\eta(y), y)\}, \quad \forall t \in [0, S]. \tag{2.1.28}$$

The validity of the following statement is implied by Proposition 2.1.1 and Proposition 2.1.4.

Corollary 2.1.6 *Assume that a solution $\eta(\cdot) \in C^1$ of the HJB inequality (2.1.5) exists. Then a T-periodic admissible pair $(u(t), y(t)) = (u(t+T), y(t+T))$ is optimal in (1.1.11) and the equality*

$$G_{per} = G^* \tag{2.1.29}$$

is valid if and only if the relationships (2.1.6) and (2.1.7) are satisfied.

2.2 Maxi-min problem equivalent to the HJB inequality and its approximation.

Consider the following *max-min* type problem

$$\sup_{\eta(\cdot)\in C^1} \min_{y\in Y} H(\nabla\eta(y), y), \tag{2.2.1}$$

where *sup* is taken over all continuously differentiable functions. Note that, the problem (2.2.1) is dual with respect to the IDLP problem (1.2.11) (see (1.3.8)).

Proposition 2.2.1 *If the optimal value of the problem (2.2.1) is bounded, then it is equal to the optimal value of the IDLP problem (1.2.11). That is,*

$$\sup_{\eta(\cdot)\in C^1} \min_{y\in Y} H(\nabla\eta(y), y) = G^*. \tag{2.2.2}$$

Proof. The equality (2.2.2) follows from the Theorem 1.3.1. Note that from this theorem also follows that *supmin* in (2.2.1) is bounded if and only if $W \neq \emptyset$. $\square$

Definition 2.2.2 *A function $\eta(\cdot) \in C^1$ is called a solution of the problem (2.2.1) if*

$$\min_{y\in Y} H(\nabla\eta(y), y) = G^*. \tag{2.2.3}$$

Proposition 2.2.3 *If $\eta(\cdot) \in C^1$ is a solution of the HJB inequality (2.1.5), then this $\eta(\cdot)$ is also a solution of the problem (2.2.1). Conversely, if $\eta(\cdot) \in C^1$ is a solution of the problem (2.2.1), then it also solves the HJB inequality (2.1.5).*

Proof. Let $\eta(\cdot) \in C^1$ be a solution of the HJB inequality (2.1.5). By (2.2.1) and (2.2.2), the inequality $\min_{y\in Y} H(\nabla\eta(y), y) > G^*$ can not be valid. Hence, $\eta(\cdot)$ solves (2.2.1). The converse statement is obvious too. $\square$

Note that a solution of the max-min problem (2.2.1), defined as a C^1 function satisfying (1.3.10), may not exist, and one can consider the possibility of defining the solution as a nondifferentiable function, which satisfies (1.3.10) in the viscosity sense (see, e.g., [20] p.399). We, however, do not follow this path. Instead, we introduce (following [48]) a way of constructing C^1 functions that solve max-min problem (2.2.1) approximately.

Let $\Big\{\phi_i(\cdot) \in C^1,\ i = 1, 2, ...\Big\}$ be a sequence of functions having continuous partial derivatives of the first and second orders such that any $\eta(\cdot) \in C^1$ and its gradient $\nabla\eta(\cdot)$ can be simultaneously approximated by a linear combination of functions from $\Big\{\phi_i(\cdot),\ i = 1, 2, ...\Big\}$ and their corresponding gradients. That is, for any $\eta(\cdot) \in C^1$ and any $\delta > 0$, there exists real numbers $\beta_1, ..., \beta_N$ such that

$$\max_{y\in Y}\Big\{|\eta(y) - \sum_{i=1}^{N}\beta_i\phi_i(y)| + ||\nabla\eta(y) - \sum_{i=1}^{N}\beta_i\nabla\phi_i(y)||\Big\} \le \delta, \tag{2.2.4}$$

where $||\cdot||$ is a norm in $I\!R^m$. An example of such an approximating sequence is the sequence of monomials $y_1^{i_1}, ..., y_m^{i_m}$, where y_j $(j = 1, 2, ..., m)$ stands for the jth component of y and $i_1, ..., i_m = 0, 1, ...$ (see e.g.[84]). Note that it will always be assumed that $\nabla\phi_i(y),\ \ i = 1, 2, ..., N$ (with $N = 1, 2, ...$), are linearly independent on any open set Q. More specifically, it is assumed that, for any N, the equality

$$\sum_{i=1}^{N} v_i\nabla\phi_i(y) = 0, \quad \forall y \in Q \tag{2.2.5}$$

is valid if and only if $v_i = 0, \ \ i = 1, ..., N$.

Define the finite dimensional space $\Omega_N \subset C^1$ by the equation

$$\Omega_N \stackrel{def}{=} \left\{\eta(\cdot) \in C^1 : \eta(y) = \sum_{i=1}^{N}\lambda_i\phi_i(y), \quad \lambda = (\lambda_i) \in I\!R^N\right\} \tag{2.2.6}$$

and consider the max-min problem

$$\sup_{\eta(\cdot)\in\Omega_N}\min_{y\in Y} H(\nabla\eta(y),y)\stackrel{def}{=} D^N, \tag{2.2.7}$$

which will be referred to as the *N-approximating max-min problem.* Note that, due to the definition of the Hamiltonian (1.3.7), from (2.2.7) it follows that

$$\sup_{\eta(\cdot)\in\Omega_N}\min_{(u,y)\in U\times Y}\{q(u,y)+(\nabla\eta(y))^T f(u,y)\} = D^N. \tag{2.2.8}$$

Proposition 2.2.4 *D^N converges to G^*, that is,*

$$\lim_{N\to\infty} D^N = G^*. \tag{2.2.9}$$

Proof. It is obvious that, for any $N\geq 1$,

$$D^1\leq D^2\leq ...\leq D^N\leq G^*. \tag{2.2.10}$$

Hence, $\lim\limits_{N\to\infty} D^N$ exists, and it is less than or equal to G^*. The fact that it is equal to G^* follows from the fact that, for any function $\eta(\cdot)\in C^1$ and for any $\delta>0$, there exist N large enough and $\eta_\delta(\cdot)\in\Omega_N$ such that

$$\max_{y\in Y}\{|\eta(y)-\eta_\delta(y)|+\parallel\nabla\eta(y)-\nabla\eta_\delta(y)\parallel\}\leq\delta. \tag{2.2.11}$$

□

Definition 2.2.5 *A function $\eta(\cdot)\in C^1$ will be called a solution of the N-approximating max-min problem (2.2.7) if*

$$\min_{y\in Y}\{H(\nabla\eta(y),y)\} = D^N. \tag{2.2.12}$$

The existence of a solution of the N-approximating max-min problem (2.2.7) is guaranteed by the following assumption about controllability properties of the system (1.1.1).

Assumption 2.2.6 *There exists a set $Y^0\subset Y$ such that any two points in Y^0 can be connected by an admissible trajectory of the system (1.1.1)(that is, for any $y',y''\in Y^0$, there exists an admissible pair $(u(\cdot),y(\cdot))$ defined on some interval $[0,S]$ such that $y(0)=y'$ and $y(S)=y''$) and such that the closure of Y^0 has a nonempty interior.*

That is,

$$int(clY^0) \neq \emptyset.$$

Definition 2.2.7 *We shall say that the system (1.1.1) is locally approximately controllable on Y if Assumption 2.2.6 is satisfied.*

Proposition 2.2.8 *Let the system (1.1.1) be locally approximately controllable on Y. Then, for every $N = 1, 2, ...,$ there exists $\lambda^N = (\lambda_i^N)$ such that*

$$\eta^N(y) \stackrel{def}{=} \sum_{i=1}^{N} \lambda_i^N \phi_i(y) \tag{2.2.13}$$

is a solution of the N-approximating max-min problem (2.2.7).

Proof. The proof follows from the following two lemmas. □

Lemma 2.2.9 *Assume that, for*

$$\eta(y) = \sum_{i=1}^{N} \upsilon_i \phi_i(y), \tag{2.2.14}$$

the inequality

$$(\nabla \eta(y))^T f(u, y) \geq 0, \quad \forall (u, y) \in U \times Y \tag{2.2.15}$$

is valid only if $\upsilon_i = 0, \quad \forall i = 1, ..., N$. Then a solution (2.2.13) of the N-approximating max-min problem (2.2.7) exists.

Proof. For any $k = 1, 2, ...,$ let $\upsilon^k = (\upsilon_i^k) \in I\!R^N$ be such that the function

$$\eta^k(y) \stackrel{def}{=} \sum_{i=1}^{N} \upsilon_i^k \phi_i(y), \tag{2.2.16}$$

satisfies the inequality

$$H(\nabla \eta^k(y), y) \geq D^N - \frac{1}{k}, \quad \forall y \in Y. \tag{2.2.17}$$

Hence,

$$q(u, y) + (\nabla \eta^k(y))^T f(u, y) \geq D^N - \frac{1}{k}, \quad \forall (u, y) \in U \times Y. \tag{2.2.18}$$

Let us show that the sequence $\upsilon^k, \quad k = 1, 2, ...,$ is bounded. That is, there exists $\alpha > 0$ such that

$$\| \upsilon^k \| \leq \alpha, \quad k = 1, 2, ... \,. \tag{2.2.19}$$

Assume that the sequence $\upsilon^k, \quad k = 1, 2, ...,$ is not bounded. Then there exists a subsequence $\upsilon^{k_l}, \ l = 1, 2, ...$ such that

$$\lim_{l\to\infty} \| \upsilon^{k_l} \| = \infty, \quad \lim_{l\to\infty} \frac{\upsilon^{k_l}}{\| \upsilon^{k_l} \|} \stackrel{def}{=} \tilde{\upsilon}, \quad \| \tilde{\upsilon} \| = 1. \tag{2.2.20}$$

Dividing (2.2.18) by $\| \upsilon^k \|$ and passing to the limit along the subsequence $\{k_l\}$, one can show that

$$(\nabla\tilde{\eta}(y))^T f(u, y) \geq 0, \quad \forall (u, y) \in U \times Y, \tag{2.2.21}$$

where

$$\tilde{\eta}(y) \stackrel{def}{=} \sum_{i=1}^{N} \tilde{\upsilon}_i \phi_i(y).$$

Hence, by the assumption of the lemma, $\tilde{\upsilon} = (\tilde{\upsilon}_i) = 0$, which is in contradiction with (2.2.20). Thus, the validity of (2.2.19) is established.

Due to (2.2.19), there exists a subsequence $\upsilon^{k_l}, \ l = 1, 2...,$ such that there exists a limit

$$\lim_{l\to\infty} \upsilon^{k_l} \stackrel{def}{=} \upsilon^*. \tag{2.2.22}$$

Passing to the limit in (2.2.18) along this subsequence, one obtains

$$q(u, y) + (\nabla\eta^*(y))^T f(u, y) \geq D^N, \quad \forall (u, y) \in U \times Y, \tag{2.2.23}$$

where

$$\eta^*(y) \stackrel{def}{=} \textstyle\sum_{i=1}^{N} \upsilon_i^* \phi_i(y).$$

From (2.2.23) it follows that

$$H(\nabla\eta^*(y), y) \geq D^N, \quad \forall y \in Y.$$

That is, $\eta^*(y)$ is an optimal solution of the N-approximating max-min problem (2.2.7).
$\square$

Lemma 2.2.10 *If the system (1.1.1) is locally approximately controllable on Y, then the inequality (2.2.15) is valid only if $\upsilon_i = 0$.*

Proof. Assume that

$$\eta(y) = \sum_{i=1}^{N} \upsilon_i \phi_i(y) \tag{2.2.24}$$

and the inequality (2.2.15) is valid. For arbitrary $y', y'' \in Y^0$, there exists an admissible pair $(u(\cdot), y(\cdot))$ such that $y(0) = y'$ and $y(S) = y''$. From (2.2.15) it follows that

$$\phi(y'') - \phi(y') = \int_0^S (\nabla\phi(y(t)))^T f(u(t), y(t))dt \geq 0 \quad \Rightarrow \quad \phi(y'') \geq \phi(y').$$

Since y', y'' are arbitrary points in Y_0, the above inequality allows one to conclude that

$$\phi(y) = const \qquad \forall y \in Y^0 \qquad \Rightarrow \qquad \phi(y) = const \qquad \forall y \in clY^0,$$

the latter implying that $\nabla\eta(y) = 0 \;\; \forall y \in int(clY^0)$ and, consequently leading to the fact that $\upsilon_i = 0,\ i = 1, ..., N$ (due to the linear independence of $\nabla\phi_i(y), \;\; i = 1, 2, ..., N$ (see (2.2.5))). $\square$

Remark 2.2.11 *Note that from Proposition 2.2.4 it follows that solutions of the N-approximating problems (the existence of which is established by Proposition 2.2.8) solve the max-min problem (2.2.1) approximately in the sense that, for any $\delta > 0$, there exists N_δ such that, for any $N \geq N_\delta$,*

$$H(\nabla\eta^N(y), y) \geq G^* - \delta \qquad \forall y \in Y, \tag{2.2.25}$$

where $\eta^N(\cdot)$ is a solution of the N-approximating max-min problem (2.2.7).

2.3 Semi-infinite dimensional LP problem and duality relationships.

Let $\{\phi_i(\cdot)\}, i = 1, 2, ...\}$ be the sequence of functions introduced in Section 2.2. Observe that due to approximating property (see (2.2.4)) of this sequence of functions the set W can be presented in the form of a countable system of equations. That is,

$$W = \Big\{\gamma \in \mathcal{P}(U \times Y) : \int_{U\times Y} (\nabla\phi_i(y))^T f(u, y)\gamma(du, dy) = 0, \quad i = 1, 2, ...\Big\}. \tag{2.3.1}$$

Define the set W_N as follows

$$W_N \stackrel{def}{=} \left\{ \gamma \in \mathcal{P}(U \times Y) : \int_{U\times Y} (\nabla \phi_i(y))^T f(u,y))\gamma(du,dy) = 0. \;\; i = 1, ..., N \right\}. \tag{2.3.2}$$

Note that

$$W_1 \supset ... \supset W_N \supset W. \tag{2.3.3}$$

Consequently, from the fact that W is assumed to be non-empty, it follows that the sets W_N, $N = 1, 2, ...$ are not empty. Also (as can be easily seen), the sets W_N are compact in the weak* topology. Hence, the set of optimal solutions of (2.3.4) is not empty for any $N = 1, 2, ...$.

Let us consider the semi-infinite dimensional linear programming (SILP) problem

$$\min_{\gamma \in W_N} \int_{U\times Y} q(u,y)\gamma(du,dy) \stackrel{def}{=} G^N. \tag{2.3.4}$$

Note that (2.3.3) implies

$$G^* \geq G^N. \tag{2.3.5}$$

Proposition 2.3.1 *If W is not empty, then the following relationships are valid:*

$$\lim_{N\to\infty} \rho_H(W_N, W) = 0, \tag{2.3.6}$$

$$\lim_{N\to\infty} G^N = G^*. \tag{2.3.7}$$

Proof. By Lemma 1.1.6 (ii) the validity of (2.3.7) follows from the validity of (2.3.6). That is, we have to show the validity of (2.3.6).

Since $W \subset W_N$, to prove that (2.3.6) is valid, it is enough to show that

$$\lim_{N\to\infty} \sup_{\gamma \in W_N} \rho(\gamma, W) = 0. \tag{2.3.8}$$

Assume it is not true. Then there exist a positive number δ, a subsequence of positive integers $N' \to \infty$, and a sequence of probability measures $\gamma^{N'} \in W_{N'}$ such that $\rho(\gamma^{N'}, W) \geq \delta$. Due to the compactness of $\mathcal{P}(U \times Y)$, one may assume (without loss of generality) that there exists $\bar{\gamma} \in \mathcal{P}(U \times Y)$ such that

$$\lim_{N'\to\infty} \rho(\gamma^{N'}, \bar{\gamma}) = 0 \qquad \Rightarrow \qquad \rho(\bar{\gamma}, W) \geq \delta. \tag{2.3.9}$$

From the fact that $\gamma^{N'} \in W_{N'}$ it follows that, for any integer i and $N' \geq i$,

$$\int_{U\times Y}(\nabla\phi_i(y))^T f(u,y)\gamma^{N'}(du,dy)=0 \qquad \Rightarrow \qquad \int_{U\times Y}(\nabla\phi_i(y))^T f(u,y)\bar{\gamma}(du,dy)=0.$$

Since the latter is valid for any $i = 1, 2, ...,$ one can conclude that $\bar{\gamma} \in W$, which contradicts (2.3.9). This proves (2.3.6). $\square$

Corollary 2.3.2 *If γ^N is a solution of the problem (2.3.4) and $\lim_{N'\to\infty}\rho(\gamma^{N'},\ \gamma)=0$ for some subsequence of integers N' tending to infinity, then γ is a solution of (1.2.11). If the optimal solution γ^* of the problem (1.2.11) is unique, then, for any optimal solution γ^N of the problem (2.3.4) there exists the limit*

$$\lim_{N\to\infty}\gamma^N=\gamma^*. \tag{2.3.10}$$

Note that every extreme point of the optimal solutions set of (2.3.4) is an extreme point of W_N and that the latter is presented as a convex combination of (no more than $N+1$) Dirac measures (see, e.g., Theorem A.5 in [95]). That is, if γ^N is an extreme point of W_N, which is an optimal solution of (2.3.4), then there exist

$$(u_l^N, y_l^N)\in U\times Y,\quad \gamma_l^N>0,\quad l=1,2,...,K_N\le N+1;\quad \sum_{l=1}^{K_N}\gamma_l^N=1 \tag{2.3.11}$$

such that

$$\gamma^N=\sum_{l=1}^{K_N}\gamma_l^N\delta_{(u_l^N,y_l^N)}, \tag{2.3.12}$$

where $\delta_{(u_l^N,y_l^N)}$ is the Dirac measure concentrated at (u_l^N, y_l^N).

The SILP problem (2.3.4) is related to the N-approximating max-min problem (2.2.7) through the following duality type relationships. Note that these results are similar to the results in Theorem 1.3.1.

Theorem 2.3.3 *(i) The optimal value of the N-approximating max-min problem (2.2.7) is bounded (that is, $D^N<\infty$) if and only if the set W_N is not empty;*

(ii) If the optimal value of the N-approximating max-min problem (2.2.7) is bounded, then

$$G^N=D^N; \tag{2.3.13}$$

(iii) The optimal value of the N-approximating max-min problem (2.2.7) is unbounded (that is, $D^N=\infty$) if and only if there exists $v=(v_1,v_2,...,v_N)$ such that

$$\max_{(u,y)\in U\times Y}(\nabla\eta_v(y))^T f(u,y)<0,\quad \eta_v(y)\stackrel{\text{def}}{=}\sum_{i=1}^{N}v_i\phi_i(y). \tag{2.3.14}$$

Proof. The proof of the theorem follows the same steps as those used in the proof of Theorem 1.3.1 (see Appendix B). □

Proposition 2.3.4 *If γ^N is an optimal solution of (2.3.4) that allows a representation (2.3.12) and if $\eta^N(y) = \sum_{i=1}^N \lambda_i^N \phi_i(y)$ is an optimal solution of (2.2.7), then the concentration points (u_l^N, y_l^N) of the Dirac measures in the expansion (2.3.12) satisfy the following relationships:*

$$y_l^N = \arg\min_{y\in Y}\{H(\nabla\eta^N(y), y)\}, \tag{2.3.15}$$

$$u_l^N = \arg\min_{u\in U}\{q(u, y_l^N) + (\nabla\eta^N(y_l^N))^T f(u, y_l^N)\}, \quad l = 1, ..., K_N. \tag{2.3.16}$$

Proof. Due to (2.3.13) and due to the fact that $\eta^N(y)$ is an optimal solution of (2.2.7) (see (2.2.12)),

$$G^N = \min_{y\in Y}\{H(\nabla\eta^N(y), y)\} = \min_{(u,y)\in U\times Y}\{q(u, y) + (\nabla\eta^N(y))^T f(u, y)\}. \tag{2.3.17}$$

Also, for any $\gamma \in W_N$,

$$\int_{U\times Y} q(u, y)\gamma(du, dy) = \int_{U\times Y}[q(u, y) + (\nabla\eta^N(y))^T f(u, y)]\gamma(du, dy).$$

Consequently, for $\gamma = \gamma^N$,

$$G^N = \int_{U\times Y} q(u, y)\gamma^N(du, dy) = \int_{U\times Y}[q(u, y) + (\nabla\eta^N(y))^T f(u, y)]\gamma^N(du, dy).$$

Hence, by (2.3.12),

$$G^N = \sum_{l=1}^{K_N} \gamma_l^N[q(u_l^N, y_l^N) + (\nabla\eta^N(y_l^N))^T f(u_l^N, y_l^N)]. \tag{2.3.18}$$

Since $(u_l^N, y_l^N) \in U \times Y$, from (2.3.17) and (2.3.18) it follows that, if $\gamma_l^N > 0$, then

$$q(u_l^N, y_l^N) + (\nabla\eta^N(y_l^N))^T f(u_l^N, y_l^N) = \min_{(u,y)\in U\times Y}\{q(u, y) + (\nabla\eta^N(y))^T f(u, y)\}.$$

That is,

$$(u_l^N, y_l^N) = \arg\min_{(u,y)\in U\times Y}\{q(u, y) + (\nabla\eta^N(y))^T f(u, y)\}.$$

The latter is equivalent to (2.3.15) and (2.3.16). □

2.4 Additional comments for Chapter 2

The chapter is mostly based on the results obtained in [60]. In contrast to the aforementioned work, where most of the results were stated for periodic optimization problems, the results of this chapter are stated in the general LRAOC setting, in which the assumption of the periodicity is replaced by a “recurrence” type assumption (Assumption 2.1.2). Note that this assumption is satisfied for the periodical regimes (see Proposition 2.1.3).

3

On construction of a near optimal solution of the long run average optimal control problem

In Section 3.1, we show that a solution of the max-min problems defined on a finite dimensional subspace of the space of continuously differentiable functions can be used for construction of near optimal solution of optimal control problem. In Section 3.2, we present an algorithm that allows one to solve the semi-infinite dimensional linear programming (SILP) problems. The convergence of this algorithm is proved in Section 3.3. In Section 3.4, we demonstrate the construction of a near optimal control with a numerical example.

3.1 Construction of a near optimal control.

In this section, we assume that a solution $\eta^N(\cdot)$ of the N approximating problem (2.2.7) exists for all N large enough (see Proposition 2.2.8) and we show that, under certain additional assumptions, a control $u^N(y)$ defined as a minimizer of the problem

$$\min_{u\in U}\{q(u,y)+(\nabla\eta^N(y))^T f(u,y)\} \tag{3.1.1}$$

(that is, $u^N(y)=\arg\min_{u\in U}\{q(u,y)+(\nabla\eta^N(y))^T f(u,y)\}$) is near optimal in optimization problem (1.1.3). The additional assumptions that we are using to establish this near optimality are introduced below.

Assumption 3.1.1 *The following conditions are satisfied:*

(i) The equality (2.1.1) is valid and the optimal solution γ^* of the IDLP problem (1.2.11) is unique;

(ii) The optimal solution $(u^*(\cdot),y^*(\cdot))$ of the problem (1.2.11) exists and satisfies Assumption 2.1.2.

Remark 3.1.2 *Note that, Assumption 3.1.1 implies that γ^* is the occupational measures generated by $(u^*(\cdot),y^*(\cdot))$.*

Assumption 3.1.3 *For almost all $t\in[0,\infty)$, there exists an open ball $\mathcal{Y}_t\subset\mathbb{R}^m$ centered at $y^*(t)$ such that, the following conditions are satisfied:*

(i) the solution $u^N(y)$ of the problem (3.1.1) is uniquely defined (the problem (3.1.1) has a unique solution) for $y\in\mathcal{Y}_t$;

(ii) the function $u^N(y)$ satisfies Lipschitz conditions on $\mathcal{Y}_t$ (with a Lipschitz constant being independent of N and t). That is,

$$||u^N(y')-u^N(y'')||\leq L(||y'-y''||)\qquad\forall y',y''\in\mathcal{Y}_t, \tag{3.1.2}$$

where L is a constant;

(iii) the solution $y^N(\cdot)$ of the system of differential equations

$$y'(t)=f(u^N(y(t)),y(t)) \tag{3.1.3}$$

exists. Moreover, this solution is unique and is contained in Y for $t\in[0,\infty)$. Also, for any $t>0$, the Lebesgue measure of the set $A_t(N)\stackrel{def}{=}\{t'\in[0,t]:y^N(t')\notin\mathcal{Y}_{t'}\}$ tends to zero as $N\to\infty$. That is,

$$\lim_{N\to\infty} meas\{A_t(N)\}=0. \tag{3.1.4}$$

Assumption 3.1.4 *The pair $(u^N(t), y^N(t))$, where $y^N(t)$ is the solution of (3.1.3) and $u^N(t) = u^N(y^N(t))$, generates the occupational measure $\bar{\gamma}^N$ on the interval $[0,\infty)$, the latter being independent of the initial conditions $y^N(0) = y$ for y in a neighbourhood of $y^*(\cdot)$. Moreover, for any continuous $h(u,y) : U \times Y \to I\!R^1$,*

$$|\frac{1}{S}\int_0^S h(u^N(t), y^N(t))dt - \int_{U\times Y} h(u,y)\bar{\gamma}^N(du,dy)| \leq \delta_h(S), \qquad \lim_{S\to\infty}\delta_h(S) = 0 \quad (3.1.5)$$

(the estimate is uniform in N).

Theorem 3.1.5 *Let U be a compact subset of $I\!R^n$ and let $f(u,y)$ and $q(u,y)$ be Lipschitz continuous in a neighborhood of $U \times Y$. Also, let the system (1.1.1) be locally approximately controllable on Y and let Assumptions 2.1.2, 3.1.1, 3.1.3 and 3.1.4 be satisfied. Then*

$$\lim_{N\to\infty} u^N(y^N(t)) = u^*(t) \quad (3.1.6)$$

for almost all $t \in [0,\infty)$ and

$$\lim_{N\to\infty}\max_{t'\in[0,t]} \| y^N(t') - y^*(t') \| = 0 \qquad \forall t \in [0,\infty). \quad (3.1.7)$$

Also,

$$\lim_{N\to\infty}\bar{\gamma}^N = \gamma^*. \quad (3.1.8)$$

Proof. The proof is given below and it is based on the following Lemma.

Let $d((u,y),Q)$ stands for the distance between a point $(u,y) \in U \times Y$ and a set $Q \subset U \times Y : d((u,y),Q) \stackrel{def}{=} \inf_{(u',y')\in Q}\{\| (u,y) - (u',y') \|\}$.

Lemma 3.1.6 *Let Assumption 2.1.2 (ii) be satisfied. Then for almost all $t \in [0,\infty)$*

$$\lim_{N\to\infty} d((u^*(t), y^*(t)), \Theta^N) = 0, \quad (3.1.9)$$

where $\Theta^N \stackrel{def}{=} \{(u_l^N, y_l^N),\ l = 1,...,K_N\}$ and (u_l^N, y_l^N) are the points as in (2.3.11).

Proof. Let t be such that Assumption 2.1.2 (ii) is true and assume that (3.1.9) is not true for this t. Then there exists a number $r > 0$ and a sequence $N_i, i = 1,2,\dots$, $\lim_{i\to\infty} N_i = \infty$, such that

$$d((u(t), y(t)), \Theta^{N_i}) \geq r \qquad i = 1,2,\dots . \quad (3.1.10)$$

The inequality in (3.1.10) implies that

$$(u_l^{N_i}, y_l^{N_i}) \notin B_r(u(t), y(t)), \quad l = 1, ..., K_{N_i}, \quad i = 1, 2, ...$$

Note that $B_r(\cdot, \cdot)$ is defined as in (2.1.12). By (2.3.12), the latter implies that

$$\gamma^{N_i}(B_r(u(t), y(t))) = 0. \tag{3.1.11}$$

From (2.3.10) it follows that

$$\lim_{i\to\infty} \rho(\gamma^{N_i}, \gamma^*) = 0.$$

Consequently (see, e.g., Theorem 2.1 in [25]),

$$0 = \lim_{i\to\infty} \gamma^{N_i}(B_r(u(t), y(t))) \geq \gamma^*(B_r(u(t), y(t))).$$

The latter contradicts (2.1.13) and thus proves the lemma. $\square$

Proof of Theorem 3.1.5. Let $t \in [0, \infty)$ be such that $\mathcal{Y}_t \neq \emptyset$ (see Assumption 3.1.3) and such that Assumption 2.1.2 (ii) is valid. By (3.1.9), there exists $(u_{l_N}^N, y_{l_N}^N) \in \Theta^N$ such that

$$\lim_{N\to\infty} ||(u_{l_N}^N, y_{l_N}^N) - (u^*(t), y^*(t))|| = 0, \tag{3.1.12}$$

the latter implying, in particular, that $y_{l_N}^N \in \mathcal{Y}_t$ for N large enough. Due to (2.3.16),

$$u_{l_N}^N = u^N(y_{l_N}^N). \tag{3.1.13}$$

Hence,

$$\| u^*(t) - u^N(y^*(t)) \| \leq \| u^*(t) - u_{l_N}^N \| + \| u^N(y_{l_N}^N) - u^N(y^*(t)) \|$$

$$\leq \| u^*(t) - u_{l_N}^N \| + L \| y_{l_N}^N - y^*(t) \|, \tag{3.1.14}$$

where L is a Lipschitz constant of $u^N(\cdot)$. From (3.1.12) it now follows that

$$\lim_{N\to\infty} u^N(y^*(t)) = u^*(t). \tag{3.1.15}$$

By Assumption 3.1.3, the same argument is applicable for almost all $t \in [0, \infty)$. This proves the convergence (3.1.15) for almost all $t \in [0, \infty)$.

Taking an arbitrary $t \in [0, \infty)$ and subtracting the equation

$$y^*(t) = y_0 + \int_0^t f(u^*(t'), y^*(t'))dt' \tag{3.1.16}$$

from the equation

$$y^N(t) = y_0 + \int_0^t f(u^N(y^N(t')), y^N(t'))dt', \tag{3.1.17}$$

one obtains

$$\| y^N(t) - y^*(t) \| \leq \int_0^t \| f(u^N(y^N(t')), y^N(t')) - f(u^*(t'), y^*(t')) \| dt'$$

$$\leq \int_0^t \| f(u^N(y^N(t')), y^N(t')) - f(u^N(y^*(t')), y^*(t')) \| dt'$$

$$+ \int_0^t \| f(u^N(y^*(t')), y^*(t')) - f(u^*(t'), y^*(t')) \| dt'. \tag{3.1.18}$$

It is easy to see that

$$\int_0^t \| f(u^N(y^N(t')), y^N(t')) - f(u^N(y^*(t')), y^*(t')) \| dt'$$

$$\leq \int_{t' \notin A_t(N)} \| f(u^N(y^N(t')), y^N(t')) - f(u^N(y^*(t')), y^*(t')) \| dt'$$

$$+ \int_{t' \in A_t(N)} [\| f(u^N(y^N(t')), y^N(t')) \| + \| f(u^N(y^*(t')), y^*(t')) \|] dt'$$

$$\leq L_1 \int_0^t \| y^N(t') - y^*(t') \| dt' + L_2 \, meas\{A_t(N)\}, \tag{3.1.19}$$

where L_1 is a constant defined (in an obvious way) by Lipschitz constants of $f(\cdot,\cdot)$ and $u^N(\cdot)$, and $L_2 \stackrel{def}{=} 2 \max_{(u,y)\in U\times Y} \{\| f(u,y) \|\}$. Also, due to (3.1.15) and the dominated convergence theorem (see, e.g., p. 49 in [15]),

$$\lim_{N\to\infty} \int_0^t \| f(u^N(y^*(t')), y^*(t')) - f(u^*(t'), y^*(t')) \| dt' = 0. \tag{3.1.20}$$

Let us introduce the notation

$$k_t(N) \stackrel{def}{=} L_2 meas\{A_t(N)\} + \int_0^t \| f(u^N(y^*(t')), y^*(t')) - f(u^*(t'), y^*(t')) \| dt'$$

and rewrite the inequality (3.1.18) in the form

$$\| y^N(t) - y^*(t) \| \leq L_1 \int_0^t \| y^N(t') - y^*(t') \| dt' + k_t(N), \tag{3.1.21}$$

which, by the Gronwall-Bellman lemma (see, e.g., p.218 in [20]), implies that

$$\max_{t'\in[0,t]} \parallel y^N(t) - y^*(t) \parallel \leq k_t(N)e^{L_1 t}. \tag{3.1.22}$$

Since, by (3.1.4) and (3.1.20),

$$\lim_{N\to\infty} k_t(N) = 0, \tag{3.1.23}$$

(3.1.22) implies (3.1.7).

For any $t \in [0,\infty)$ such that $u^N(\cdot)$ is Lipschitz continuous on $\mathcal{Y}_t$, one has

$$\parallel u^N(y^N(t)) - u^*(t) \parallel \leq \parallel u^N(y^N(t)) - u^N(y^*(t)) \parallel + \parallel u^N(y^*(t)) - u^*(t) \parallel$$

$$\leq L \parallel y^N(t) - y^*(t) \parallel + \parallel u^N(y^*(t)) - u^*(t) \parallel .$$

The latter implies (3.1.6) (due to (3.1.22), (3.1.23) and due to (3.1.15)).

Let $t \in [0,\infty)$ be such that $\mathcal{Y}_t \neq \emptyset$ and (2.1.13) is satisfied for an arbitrary $r > 0$. By (2.1.11) and (3.1.5), for any continuous $h(u,y)$ and for any arbitrary small $\alpha > 0$, there exists $S > 0$ such that

$$|\frac{1}{S}\int_0^S h(u^*(t),y^*(t))dt - \int_{U\times Y} h(u,y)\gamma^*(du,dy)| \leq \frac{\alpha}{2} \tag{3.1.24}$$

and

$$|\frac{1}{S}\int_0^S h(u^N(y^N(t)),y^N(t))dt - \int_{U\times Y} h(u,y)\bar{\gamma}^N(du,dy)| \leq \frac{\alpha}{2}. \tag{3.1.25}$$

Using (3.1.24) and (3.1.25), one can obtain

$$|\int_{U\times Y} h(u,y)\bar{\gamma}^N(du,dy) - \int_{U\times Y} h(u,y)\gamma^*(du,dy)| \leq$$

$$|\frac{1}{S}\int_0^S h(u^N(y^N(t)),y^N(t))dt - \frac{1}{S}\int_0^S h(u^*(t),y^*(t))dt| + \alpha. \tag{3.1.26}$$

Due to (3.1.6) and (3.1.7), the latter implies the following inequality

$$\overline{\lim}_{N\to\infty}|\int_{U\times Y} h(u,y)\bar{\gamma}^N(du,dy) - \int_{U\times Y} h(u,y)\gamma^*(du,dy)| \leq \alpha,$$

which in turn, implies

$$\lim_{N\to\infty}|\int_{U\times Y} h(u,y)\bar{\gamma}^N(du,dy) - \int_{U\times Y} h(u,y)\gamma^*(du,dy)| = 0 \tag{3.1.27}$$

(due to the fact that α can be arbitrary small). Since $h(u, y)$ is an arbitrary continuous function from (3.1.27) it follows that

$$\lim_{N\to\infty} \rho(\bar{\gamma}^N, \gamma^*) = 0. \tag{3.1.28}$$

□

Corollary 3.1.7 *Denote by V^N a value of the objective function obtained with the control $u^N(\cdot)$. That is*

$$V^N \stackrel{\text{def}}{=} \lim_{S\to\infty} \frac{1}{S} \int_0^S q(u^N(t), y^N(t))dt = \int_{U\times Y} q(u, y)\bar{\gamma}^N(du, dy). \tag{3.1.29}$$

By (3.1.8)

$$\lim_{N\to\infty} V^N = G^*. \tag{3.1.30}$$

Proposition 3.1.8 *Let the optimal solution of the problem (1.1.3) be periodic with a period T^* and let Assumptions 3.1.1, 3.1.3 and 3.1.4 be satisfied. Assume in addition that, there exists a T^*-periodic solution $\tilde{y}^N(t)$ of the system (1.1.1) obtained with the control $u^N(t) \stackrel{\text{def}}{=} u^N(y^N(t))$ such that*

$$\max_{t\in[0,T^*]} \| \tilde{y}^N(t) - y^N(t) \| \leq \nu_1(N), \qquad \lim_{N\to\infty} \nu_1(N) = 0, \tag{3.1.31}$$

then the pair $(u^N(t), \tilde{y}^N(t))$ is a near optimal solution of the periodic optimization problem (1.1.11) in the sense that

$$\lim_{N\to\infty} \frac{1}{T^*} \int_0^{T^*} q(u^N(t), \tilde{y}^N(t))dt = V^*_{per}. \tag{3.1.32}$$

Proof. Note that from (3.1.31) it follows that

$$\left| \frac{1}{T^*} \int_0^{T^*} q(u^N(t), \tilde{y}^N(t))dt - V^*_{per} \right|$$

$$= \left| \frac{1}{T^*} \int_0^{T^*} q(u^N(t), \tilde{y}^N(t))dt - \frac{1}{T^*} \int_0^{T^*} q(u^*(t), y^*(t))dt \right|$$

$$\leq \frac{1}{T^*} \int_0^{T^*} ||q(u^N(t), \tilde{y}^N(t)) - q(u^N(t), y^N(t))||dt$$

$$+\frac{1}{T^*} \int_0^{T^*} ||q(u^N(t), y^N(t))dt - q(u^*(t), y^*(t))||dt$$

$$\leq \frac{L}{T^*}\int_0^{T^*}[||\tilde{y}^N(t)-y^N(t)||+||y^N(t)-y^*(t)||+||u^N(t)-u^*(t)||]dt,$$

where L is a Lipschitz constant. The latter implies (3.1.32) (due to (3.1.6), (3.1.7) and (3.1.31)). $\square$

In conclusion of this section, let us introduce one more assumption, the validity of which implies the existence of a near optimal periodic admissible pair (see the last part of Proposition 3.1.8).

Assumption 3.1.9 *The solutions of the system (1.1.1) obtained with any initial values* $y_i,\ i=1,2$ *and with any control* $u(\cdot)$ *satisfy the inequality*

$$\| y(t,u(\cdot),y_1)-y(t,u(\cdot),y_2)\| \leq \xi(t)\| y_1-y_2\|, \quad \text{with} \quad \lim_{t\to\infty}\xi(t)=0. \tag{3.1.33}$$

Note that from Lemma 3.1 in [52] it follows that if Assumption 3.1.9 is satisfied and if $\xi(T^*)<1$, then the system

$$y'(t)=f(u^N(t),y(t))$$

(the latter is the system (1.1.1), in which the control $u^N(t)=u^N(y^N(t))$ is used) has a unique T^*- periodic solution. Denote this solution as $\tilde{y}^N(T)$.

Proposition 3.1.10 *Let the assumptions of the Proposition 3.1.8 be satisfied. Also, let Assumption 3.1.9 be valid with*

$$\xi(T^*)<1. \tag{3.1.34}$$

Then, the T^**-periodic solution* $\tilde{y}^N(t)$ *of the system (1.1.1) obtained with the control* $u^N(t)$ *satisfies (3.1.31) and the pair* $(u^N(t),\tilde{y}^N(t))$ *is a near optimal periodic solution of the problem (1.1.3) in the sense that (3.1.32) is true (see Proposition 3.1.8).*

Proof. For any $t\in[0,T^*]$, one has

$$\| \tilde{y}^N(t)-y^N(t)\| \leq \| \tilde{y}^N(0)-y^N(0)\| +\int_0^t \| f(u^N(t'),\tilde{y}^N(t'))-f(u^N(t'),y^N(t'))\|$$

$$\leq \| \tilde{y}^N(0)-y^N(0)\| +L\int_0^t \| \tilde{y}^N(t')-y^N(t')\| dt',$$

which, by the Gronwall-Bellman lemma, implies that

$$\max_{t\in[0,T^*]}\| \tilde{y}^N(t)-y^N(t)\| \leq \| \tilde{y}^N(0)-y^N(0)\| e^{LT^*}. \tag{3.1.35}$$

Due to Assumption 3.1.9 and the periodicity condition $\tilde{y}^N(0) = \tilde{y}^N(T^*)$, the following relationships are valid:

$$\| \tilde{y}^N(0) - y^N(0) \| \leq \| \tilde{y}^N(0) - y^N(T^*) \| + \| y^N(T^*) - y^N(0) \|$$

$$= \| \tilde{y}^N(T^*) - y^N(T^*) \| + \| y^N(T^*) - y^N(0) \|$$
$$\leq \xi(T^*) \| \tilde{y}^N(0) - y^N(0) \| + \| y^N(T^*) - y^N(0) \| .$$

Note that $y^N(0) = y^*(0) = y^*(T^*)$. Hence (see also (3.1.7)),

$$\| y^N(T^*) - y^N(0) \| = \| y^N(T^*) - y^*(T^*) \| \leq \nu(N)$$

$$\Rightarrow \qquad \| \tilde{y}^N(0) - y^N(0) \| \leq \xi(T^*) \| \tilde{y}^N(0) - y^N(0) \| + \nu(N)$$
$$\Rightarrow \qquad \| \tilde{y}^N(0) - y^N(0) \| \leq \frac{\nu(N)}{1 - \xi(T^*)}.$$

Substituting the above inequality into (3.1.35) one obtains

$$\max_{t \in [0,T^*]} \| \tilde{y}^N(t) - y^N(t) \| \leq \frac{\nu(N)}{1 - \xi(T^*)} e^{LT^*}.$$

This proves (3.1.31). The validity of (3.1.32) is established as above. □

3.2 Algorithm for numerical solution of the SILP problem.

In this section, we describe an algorithm for solving the SILP problem (2.3.4). It finds the optimal solution γ^N of the SILP problem via solving a sequence of finite dimensional LP problems, each time augmenting the grid of Dirac measures concentration points with a new point found as an optimal solution of a certain nonlinear optimization problem.

For simplicity, let us denote $X \stackrel{def}{=} U \times Y$, $x \stackrel{def}{=} (u, y)$ and $h_i(x) \stackrel{def}{=} (\nabla \phi_i(y))^T f(u, y)$. Thus, the SILP problem (2.3.4) can be rewritten as follows

$$\min_{\gamma \in W_N} \int_X q(x) \gamma(dx) \stackrel{def}{=} G^N, \tag{3.2.1}$$

where

$$W_N \stackrel{def}{=} \{\gamma \in \mathcal{P}(X) : \int_X h_i(x) \gamma(dx) = 0, \quad i = 1, ..., N\}. \tag{3.2.2}$$

Let points $x_l \in X, \quad l = 1, ..., L \quad$ (note that $L \geq N+1$) be chosen to define an initial grid $\mathcal{X}_0$ on X. That is

$$\mathcal{X}_0 = \{x_l \in X, \quad l = 1, ..., L\}.$$

At every iteration a new point is defined and added to this set. Assume that after J iterations the points $x_{L+1}, ..., x_{L+J}$ have been defined and the set $\mathcal{X}_J$ has been constructed. Namely,

$$\mathcal{X}_J = \{x_l \in X, \quad l = 1, ..., L+J\}.$$

The iteration $J+1 \ \ (J = 0, 1, ...)$ is described as follows:

(i) Find a basic optimal solution $\gamma^J = (\gamma_l^J)$ of the LP problem

$$\min_{\gamma \in W_{\mathcal{X}_J}} \{\sum_{l=1}^{L+J} q(x_l)\gamma_l\} \stackrel{def}{=} g^J, \tag{3.2.3}$$

where

$$W_{\mathcal{X}_J} \stackrel{def}{=} \{\gamma : \gamma = \{\gamma_l\} \geq 0, \ \sum_{l=1}^{L+J} \gamma_l = 1, \ \sum_{l=1}^{L+J} h_i(x_l)\gamma_l = 0, \quad i = 1, ..., N\}. \tag{3.2.4}$$

Note that, no more than $N+1$ components of γ^J are positive, these being called basic components. Also, find an optimal solution $\lambda^J = (\lambda_0^J, \ \lambda_i^J, \ i = 1, ..., N)$ of the problem dual with respect to (3.2.3). The latter is of the form

$$\max\{\lambda_0 : q(x_l) + \sum_{i=1}^{N} h_i(x_l)\lambda_i \geq \lambda_0, \quad \forall l = 1, ..., L+J\}. \tag{3.2.5}$$

(ii) Find an optimal solution x_{L+J+1} of the problem

$$\min_{x \in X}\{q(x) + \sum_{i=1}^{N} h_i(x)\lambda_i^J\} \stackrel{def}{=} a^J, \tag{3.2.6}$$

where $\lambda^J = (\lambda_0^J, \ \lambda_i^J, \ i = 1, ..., N)$ is an optimal solution of the problem (3.2.5).

(iii) Define the set $\mathcal{X}_{J+1}$ by the equation

$$\mathcal{X}_{J+1} = \mathcal{X}_J \bigcup \{x_{L+J+1}\}.$$

Here and in what follows, J stands for the number of an iteration. Note that, by construction

$$g^{J+1} \leq g^J, \quad J = 1, 2, \ldots . \tag{3.2.7}$$

In the next section, we establish that, under certain regularity (non-degeneracy) conditions, the optimal value g^J of the problem (3.2.3) converges (as J tends to infinity) to the optimal value G^N of the problem (2.3.4) (see Theorem 3.3.1 below).

3.3 Convergence of the algorithm.

Let $\mathfrak{X} = \{x_1, x_2, ..., x_{N+1}\} \in X^{N+1}$ and let $\gamma(\mathfrak{X}) = \{\gamma_j(\mathfrak{X}),\ \forall j = 1, ..., N+1\} \geq 0$ satisfy the system of $N+1$ equations

$$\sum_{j=1}^{N+1} \gamma_j(\mathfrak{X}) = 1, \quad \sum_{j=1}^{N+1} h_i(x_j)\gamma_j(\mathfrak{X}) = 0, \quad \forall i = 1, ..., N. \tag{3.3.1}$$

Assume that the solution of the system (3.3.1) is unique (that is, the system is non-singular) and define

$$V(\mathfrak{X}) \stackrel{def}{=} \sum_{j=1}^{N+1} q(x_j)\gamma_j(\mathfrak{X}). \tag{3.3.2}$$

Also, let $\lambda(\mathfrak{X}) = \{\lambda_0(\mathfrak{X}), \lambda_1(\mathfrak{X}), ..., \lambda_N(\mathfrak{X})\}$ be a solution of the system

$$\lambda_0(\mathfrak{X}) - \sum_{i=1}^{N} h_i(x_j)\lambda_i(\mathfrak{X}) = q(x_j), \quad j = 1, ..., N+1. \tag{3.3.3}$$

and let

$$a(\mathfrak{X}) \stackrel{def}{=} \min_{x \in X}\{q(x) + \sum_{i=1}^{N} h_i(x)\lambda_i(\mathfrak{X})\}. \tag{3.3.4}$$

Lemma 3.3.1 *For any $\mathfrak{X} \subset X^{N+1}$,*

$$-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) \leq 0. \tag{3.3.5}$$

Proof. By the definition of D^N (see (2.2.8)), $a(\mathfrak{X}) \leq D^N$. Also, due to the duality theorem (see (2.3.13))

$$D^N = G^N. \tag{3.3.6}$$

Hence,

$$a(\mathfrak{X}) \leq G^N. \tag{3.3.7}$$

Note that, by multiplying the j^{th} equation in (3.3.3) by $\gamma_j(\mathfrak{X})$ and by summing up the resulted equations over $j = 1, ..., N+1$ one can obtain (using (3.3.1) and (3.3.2)) that

$$\lambda_0(\mathfrak{X}) = \sum_{j=1}^{N+1} q(x_j)\gamma_j(\mathfrak{X}) = V(\mathfrak{X}). \tag{3.3.8}$$

Since $V(\mathfrak{X}) \geq G^N$, from (3.3.8) it follows that $\lambda_0(\mathfrak{X}) \geq G^N$. The later and (3.3.7) proves (3.3.5). $\Box$

Corollary 3.3.2 *If*

$$-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) = 0, \tag{3.3.9}$$

then $\gamma(\mathfrak{X}) \stackrel{def}{=} \sum_{j=1}^{N+1} \gamma_j(\mathfrak{X})\delta_{x_j}$ *(where* δ_{x_j} *is the Dirac measure concentrated at* x_j*) is an optimal solution of the SILP problem (2.3.4) and* $\eta(y) \stackrel{def}{=} \sum_{i=1}^{N} \lambda_i(\mathfrak{X})\phi_i(y)$ *is an optimal solution of the N-approximating problem (2.2.8).*

Proof. Due to (3.3.7) and (3.3.8), we have

$$a(\mathfrak{X}) \leq D^N = G^N \leq V(\mathfrak{X}) = \lambda_0(\mathfrak{X}). \tag{3.3.10}$$

From (3.3.9) and (3.3.10) it follows that

$$a(\mathfrak{X}) = D^N \tag{3.3.11}$$

and

$$V(\mathfrak{X}) = G^N. \tag{3.3.12}$$

The equalities (3.3.11) and (3.3.12) prove the statements of the corollary 3.3.2. $\Box$

Definition 3.3.3 *An* $(N+1)$*-tuple* $\mathfrak{X} \subset X^{N+1}$ *is called* regular *if*

(i) the system (3.3.1) has a unique solution (that is, the corresponding $(N+1) \times (N+1)$ *matrix is not singular); and*

(ii) the solution $\gamma(\mathfrak{X})$ *of this system is positive. That is,*

$$\gamma(\mathfrak{X}) \stackrel{def}{=} \{\gamma_j(\mathfrak{X}), \quad \forall j = 1, ..., N+1\} > 0.$$

Assume that $\mathfrak{X}$ is regular and examine the case when

$$-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) < 0. \tag{3.3.13}$$

Denote by $A(\mathfrak{X})$ the $(N+1) \times (N+1)$ matrix of the system (3.3.1). That is

$$A(\mathfrak{X}) = \{H(x_1), H(x_2), ..., H(x_{N+1})\},$$

where columns $H(x_j), j = 1, 2, ..., N+1$ are defined as follows

$$H(x_j) = (1, h_1(x_j), h_2(x_j), ..., h_N(x_j))^T.$$

Using this notations, the solution $\gamma(\mathfrak{X})$ of the system of equations (3.3.1) can be written in the form

$$\gamma(\mathfrak{X}) = A^{-1}(\mathfrak{X})b, \tag{3.3.14}$$

where $b = (1, 0, 0, ..., 0)^T$. Similarly, the solution of the system (3.3.3) can be presented as follows

$$\lambda(\mathfrak{X}) \stackrel{def}{=} (\lambda_0(\mathfrak{X}),\ \lambda_i(\mathfrak{X}),\ \ i = 1, ..., N) = (A^T(\mathfrak{X}))^{-1}c(\mathfrak{X}), \tag{3.3.15}$$

where $c(\mathfrak{X}) \stackrel{def}{=} (\ q(x_1),\ q(x_2),\ \dots\ ,\ q(x_{N+1})\ )^T$.

Let $\mathcal{B}(\mathfrak{X})$ stand for the set of the optimal solutions of the problem (3.3.4). That is,

$$\mathcal{B}(\mathfrak{X}) \stackrel{def}{=} Arg \min_{x \in X}\{q(x) + \sum_{i=1}^{N} h_i(x)\lambda_i(\mathfrak{X})\}. \tag{3.3.16}$$

Choose an arbitrary $\tilde{x} \in \mathcal{B}(\mathfrak{X})$ and consider the system of equations

$$\sum_{j=1}^{N+1} \gamma_j(\theta) + \theta = 1, \quad \sum_{j=1}^{N+1} h_i(x_j)\gamma_j(\theta) + h_i(\tilde{x})\theta = 0, \quad \forall i = 1, ..., N, \tag{3.3.17}$$

where $\theta \geq 0$ is a parameter and $\gamma_j(\theta), j = 1, 2, ..., N+1$ are defined by the value of this parameter. Note that this system also can be presented in the form

$$A(\mathfrak{X})\gamma(\theta) + H(\tilde{x})\theta = b. \tag{3.3.18}$$

Let also $V(\mathfrak{X}, \theta)$ be defined by the equation

$$V(\mathfrak{X}, \theta) \stackrel{def}{=} \sum_{j=1}^{N+1} q(x_j)\gamma_j(\theta) + q(\tilde{x})\theta. \tag{3.3.19}$$

Observe that, by multiplying the j^{th} equation in (3.3.3) by $\gamma_j(\theta)$ and summing up the resulted equations over $j = 1, ..., N+1$ one can obtain (using (3.3.17) and (3.3.19))

$$\lambda_0(\mathfrak{X})(1-\theta) + \sum_{i=1}^{N} h_i(\tilde{x})\lambda_i(\mathfrak{X})\theta = V(\mathfrak{X}, \theta) - q(\tilde{x})\theta. \tag{3.3.20}$$

Due to (3.3.4), (3.3.8) and the fact that $\tilde{x} \in \mathcal{B}(\mathfrak{X})$, from(3.3.20) it follows that

$$V(\mathfrak{X}, \theta) = V(\mathfrak{X}) + (-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}))\theta. \tag{3.3.21}$$

By (3.3.13), the bigger value of θ is, the better is the resulted value of the objective function $V(\mathfrak{X}, \theta)$. The value of θ is, however, bounded from above by $\theta(\mathfrak{X}, \tilde{x}) \stackrel{def}{=} \min_{j,\ d_j>0} \dfrac{\gamma_j(\mathfrak{X})}{d_j(\mathfrak{X}, \tilde{x})}$, where $d(\mathfrak{X}, \tilde{x}) \stackrel{def}{=} A^{-1}(\mathfrak{X})H(\tilde{x})$ (this constraint is implied by the fact that $\gamma(\theta)$ defined by (3.3.18) must remain non-negative). Thus, the improvement, induced by replacing one of the columns of $A(\mathfrak{X})$ with the column $H(\tilde{x})$ in accordance with Simplex Method (see [37]) is determined by the following expression

$$[-\lambda_0(\mathfrak{X}) + a(\mathfrak{X})]\ \theta(\mathfrak{X}, \tilde{x}). \tag{3.3.22}$$

Let us denote $\min_j \gamma_j(\mathfrak{X}) \stackrel{def}{=} \nu(\mathfrak{X}) > 0$ and $\max_j d_j(\mathfrak{X}, \tilde{x}) \stackrel{def}{=} \beta(\mathfrak{X}, \tilde{x}) > 0$. Also, let

$$\max_{\tilde{x} \in \mathcal{B}(\mathfrak{X})} \beta(\mathfrak{X}, \tilde{x}) \stackrel{def}{=} \beta(\mathfrak{X}).$$

Define

$$\mathcal{V}(\mathfrak{X}) \stackrel{def}{=} (-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}))\frac{\nu(\mathfrak{X})}{\beta(\mathfrak{X})}. \tag{3.3.23}$$

Lemma 3.3.4 *If $\mathfrak{X}$ is regular and $a(\mathfrak{X}) < \lambda_0(\mathfrak{X})$, then $\forall \tilde{x} \in \mathcal{B}(\mathfrak{X})$ the replacement of the one of the column in $A(\mathfrak{X})$ by $H(\tilde{x})$ (according to Simplex Method) leads to the improvement of the objective value no less then $\mathcal{V}(\mathfrak{X})$. That is,*

$$(-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}))\theta(\mathfrak{X}, \tilde{x}) \geq \mathcal{V}(\mathfrak{X}). \tag{3.3.24}$$

Proof. The proof follows from the discussion preceding the statement of the lemma. □

Note that if $\mathfrak{X}$ is regular, then any $\mathfrak{X}' \in B_r \stackrel{\text{def}}{=} \{\mathfrak{X}' \ : \| \mathfrak{X}' - \mathfrak{X} \| < r\}$ (where $r > 0$ is small enough) will be regular as well, with $\nu(\mathfrak{X}')$ being continuous function of $\mathfrak{X}'$ and $\beta(\mathfrak{X}', \tilde{x})$ being continuous function of $\mathfrak{X}'$ and $\tilde{x}$.

Lemma 3.3.5 *Let $\mathfrak{X}$ be regular. Then the function $\beta(\cdot)$ is upper semicontinuous at $\mathfrak{X}$. That is,*

$$\overline{\lim_{\mathfrak{X}^l \to \mathfrak{X}}} \beta(\mathfrak{X}^l) \leq \beta(\mathfrak{X}). \tag{3.3.25}$$

Proof. Let us first of all show that

$$\overline{\lim_{\mathfrak{X}^l \to \mathfrak{X}}} \mathcal{B}(\mathfrak{X}^l) \subset \mathcal{B}(\mathfrak{X}). \tag{3.3.26}$$

That is, if $\forall x^l \in \mathcal{B}(\mathfrak{X}^l)$ and $\lim\limits_{l\to\infty} x^l = x \in X$, then

$$x \in \mathcal{B}(\mathfrak{X}). \tag{3.3.27}$$

The fact that $x^l \in \mathcal{B}(\mathfrak{X}^l)$ means that

$$q(x^l) + \sum_{i=1}^{N} \lambda_i(\mathfrak{X}^l) h_i(x^l) = a(\mathfrak{X}^l). \tag{3.3.28}$$

Passing to the limit as $l \to \infty$ in (3.3.28) one can obtain (having in mind that $\lambda_i(\cdot)$ and $a(\cdot)$ are continuous function in a neighbourhood of $\mathfrak{X}$), one obtains the equality

$$q(x) + \sum_{i=1}^{N} \lambda_i(\mathfrak{X}) h_i(x) = a(\mathfrak{X}). \tag{3.3.29}$$

The latter proves (3.3.27) and thus establishes validity of (3.3.26). To prove (3.3.25) let $\mathfrak{X}^l \to \mathfrak{X}$ as $l \to \infty$. Recall that

$$\beta(\mathfrak{X}^l) = \max_{\tilde{x}\in\mathcal{B}(\mathfrak{X}^l)} \beta(\mathfrak{X}^l, \tilde{x}). \tag{3.3.30}$$

Let $\tilde{x}^l \in \mathcal{B}(\mathfrak{X}^l)$ be such that maximum in (3.3.30) is reached. Without loss of generality, one may assume that there exists a limit

$$\lim_{l\to\infty} \tilde{x}^l = \tilde{x} \in \mathcal{B}(\mathfrak{X}),$$

with the inclusion being due to (3.3.26). Thus, passing to the limit as $l \to \infty$ in (3.3.30) one can get

$$\lim_{l\to\infty} \beta(\mathfrak{X}^l) = \lim_{l\to\infty} \beta(\mathfrak{X}^l, \tilde{x}^l) = \beta(\mathfrak{X}, \tilde{x}) \leq \beta(\mathfrak{X}). \tag{3.3.31}$$

□

Lemma 3.3.6 *If* $\mathfrak{X}$ *is regular and* $a(\mathfrak{X}) < \lambda_0(\mathfrak{X})$ *then* $\exists r_0 > 0$ *such that*

$$\mathcal{V}(\mathfrak{X}') \geq \frac{\mathcal{V}(\mathfrak{X})}{2}, \tag{3.3.32}$$

for any $\mathfrak{X}'$ *such that* $||\mathfrak{X}' - \mathfrak{X}|| \leq r_0$, *where* $\mathcal{V}(\cdot)$ *is defined by (3.3.23).*

Proof. To prove (3.3.32) let us show that $\mathcal{V}(\cdot)$ is lower semicontinuous at $\mathfrak{X}$. According to (3.3.23),

$$\mathcal{V}(\mathfrak{X}') \stackrel{def}{=} (-\lambda_0(\mathfrak{X}') + a(\mathfrak{X}')) \frac{\nu(\mathfrak{X}')}{\beta(\mathfrak{X}')}. \tag{3.3.33}$$

By taking the lower limit in (3.3.33) one can obtain (using Lemma 3.3.5)

$$\varliminf_{\mathfrak{X}'\to\mathfrak{X}} \mathcal{V}(\mathfrak{X}') = (-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}))\ \nu(\mathfrak{X})\ \varliminf_{\mathfrak{X}'\to\mathfrak{X}} \frac{1}{\beta(\mathfrak{X}')}$$

$$= (-\lambda(\mathfrak{X}) + a(\mathfrak{X}))\ \nu(\mathfrak{X}) \frac{1}{\varlimsup_{\mathfrak{X}'\to\mathfrak{X}} \beta(\mathfrak{X}')}.$$

Due to (3.3.25), it follows that

$$\varliminf_{\mathfrak{X}'\to\mathfrak{X}} \mathcal{V}(\mathfrak{X}') \geq (-\lambda(\mathfrak{X}) + a(\mathfrak{X}))\frac{\nu(\mathfrak{X})}{\beta(\mathfrak{X})} = \mathcal{V}(\mathfrak{X}). \tag{3.3.34}$$

Thus, $\mathcal{V}(\mathfrak{X})$ is lower semicontinuous at $\mathfrak{X}$. Hence, for any $\epsilon > 0$, there exists $r > 0$ such that for

$$\mathcal{V}(\mathfrak{X}') \geq \mathcal{V}(\mathfrak{X}) - \epsilon \tag{3.3.35}$$

for $\mathfrak{X}'$ such that $||\mathfrak{X}' - \mathfrak{X}|| \leq r$. By taking $\epsilon = \frac{\mathcal{V}(\mathfrak{X})}{2}$ in (3.3.35), one establishes the validity of (3.3.32). $\square$

Let $\mathfrak{X}^J = \{x_1^J, ..., x_{N+1}^J\} \in X^{N+1}$, where $x_j^J,\ j = 1, ...N+1$ are defined by the basic components $\{\gamma_1^J, ..., \gamma_{N+1}^J\}$ of an optimal solution of the problem (3.2.3). That is,

$$\sum_{j=1}^{N+1} \gamma_j^J q(x_j^J) = g^J \tag{3.3.36}$$

and

$$\sum_{j=1}^{N+1} \gamma_j^J = 1, \qquad \sum_{j=1}^{N+1} h_i(x_j^J)\gamma_j^J = 0, \quad i = 1, ..., N. \tag{3.3.37}$$

Let $\Lambda \subset X^{N+1}$ stand for the set of cluster (limit) points of $\{\mathfrak{X}^J\}$ obtained with $J \to \infty$.

Theorem 3.3.1 *Let there exist at least one regular $\mathfrak{X} \in \Lambda$. Then*

$$\lim_{J\to\infty} g^J = G^N. \tag{3.3.38}$$

Also, if $\mathfrak{X} \stackrel{def}{=} \{x_j\} \in \Lambda$ is regular, then $\sum_{j=1}^{N+1} \gamma_j(\mathfrak{X})\delta_{x_j}$ is an optimal solution of the SILP problem (2.3.4) and $\sum_{i=1}^{N} \lambda_i(\mathfrak{X})\phi_i(y)$ is an optimal solution of the N-approximating problem (2.2.8), where $\gamma(\mathfrak{X}) = \{\gamma_j(\mathfrak{X}),\ \forall j = 1, ..., N+1\}$ is the solution of the system (3.3.1) and $\lambda(\mathfrak{X}) = \{\lambda_0(\mathfrak{X}), \lambda_1(\mathfrak{X}), \ \ ... \ , \lambda_{N+1}(\mathfrak{X})\}$ is the solution of the system (3.3.3).

Proof. Let $\mathfrak{X} \in \Lambda$ be regular. By definition, there exists a subsequence $\{\mathfrak{X}^{J_l}\} \in \{\mathfrak{X}^J\}$ such that

$$\lim_{l\to\infty} \mathfrak{X}^{J_l} = \mathfrak{X}. \tag{3.3.39}$$

Note that from the fact that $\mathfrak{X}$ is regular it follows that $\mathfrak{X}^{J_l}$ is regular as well for J_l large enough. Hence,

$$\lim_{l\to\infty} V(\mathfrak{X}^{J_l}) = V(\mathfrak{X}). \tag{3.3.40}$$

Also, due to (3.2.7), there exists a limit

$$\lim_{J\to\infty} g^J \stackrel{def}{=} \widetilde{V}. \tag{3.3.41}$$

Since, by definition, $V(\mathfrak{X}^{J_l}) = G^{J_l}$, it follows that

$$V(\mathfrak{X}) = \widetilde{V}. \tag{3.3.42}$$

By Lemma 3.3.1 there are two possibilities

(i) $-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) = 0$, and

(ii) $-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) < 0$.

If $a(\mathfrak{X}) = \lambda_0(\mathfrak{X})$, then by (3.3.12), $V(\mathfrak{X}) = G^N$ and, hence, the validity of (3.3.38) follows from (3.3.41) and (3.3.42). Let us prove that (ii) leads to a contradiction. Observe that, due to (3.3.41),

$$|g^{J_l+1} - g^{J_l}| < \frac{\mathcal{V}(\mathfrak{X})}{4} \tag{3.3.43}$$

for l large enough. On the other hand, due to (3.3.39)

$$\|\mathfrak{X}^{J_l} - \mathfrak{X}\| < r_0,$$

for l large enough, with r_0 being as in Lemma 3.3.6. Hence, by this lemma and by Lemma 3.3.4,

$$|g^{J_l+1} - g^{J_l}| \geq \frac{\mathcal{V}(\mathfrak{X})}{2}. \tag{3.3.44}$$

The latter contradicts (3.3.43). Thus, the inequality $-\lambda_0(\mathfrak{X})+a(\mathfrak{X}) < 0$ can not be valid and (3.3.38) is proved. Also, by Corollary 3.3.2, from the fact that $-\lambda_0(\mathfrak{X}) + a(\mathfrak{X}) = 0$ it follows that $\sum_{j=1}^{N+1} \gamma_j(\mathfrak{X})\delta_{x_j}$ is an optimal solution of the SILP problem (2.3.4) and $\sum_{i=1}^{N} \lambda_i(\mathfrak{X})\phi_i(y)$ is an optimal solution of the N-approximating problem (2.2.8). $\square$

Remark 3.3.7 *Note that the assumption that there exists at least one regular* $\mathfrak{X} \in \Lambda$ *is much weaker than one used in proving the convergence of a similar algorithm in [63], where it was assumed that the optimal solutions of the LP problems (3.2.3) are "uniformly non-degenerate" (that is, they remain greater than some given positive number for* $J = 1, 2, \ldots$*; see Proposition 6.2 in [63]).*

3.4 Numerical example (a nonlinear pendulum).

Consider the problem of periodic optimization of the nonlinear pendulum

$$x''(t) + 0.3x'(t) + 4\sin(x(t)) = u(t) \tag{3.4.1}$$

with the controls being restricted by the inequality $|u(t)| \leq 1$ and with the objective function being of the form

$$\inf_{u(\cdot),T} \frac{1}{T} \int_0^T (u^2(t) - x^2(t))dt. \tag{3.4.2}$$

By re-denoting $x(t)$ and $x'(t)$ as $y_1(t)$ and $y_2(t)$ respectively, the above problem is reduced to a special case of the periodic optimization problem (1.1.3) with

$$y = (y_1, y_2), \quad f(u, y) = (f_1(u, y), f_2(u, y)) \stackrel{def}{=} (y_2, u - 0.3y_2 - 4\sin(y_1)),$$

$$q(u, y) \stackrel{def}{=} u^2 - y_1^2$$

and with

$$U \stackrel{def}{=} [-1, 1] \in I\!\!R^1, \qquad Y \stackrel{def}{=} \{(y_1, y_2) \mid y_1 \in [-1.7, 1.7], \quad y_2 \in [-4, 4]\} \in I\!\!R^2$$

(note that the set Y is chosen to be large enough to contain all periodic solutions of the system under consideration).

The SILP problem (2.3.4) was formulated for this problem with the use of the monomials $\phi_{i_1,i_2}(y) \stackrel{def}{=} y_1^{i_1} y_2^{i_2}, \quad i_1, i_2 = 0, 1, \ldots, J$, as the functions $\phi_i(\cdot)$ defining W_N in (2.3.2). Note that in this case the number N in (2.3.2) is equal to $(J+1)^2 - 1$. This problem and its dual were solved with the algorithm proposed above (see Section 3.2) for the case $J = 10 (N = 120)$. In particular, the coefficients $\lambda^N_{i_1,i_2}$ defining the optimal

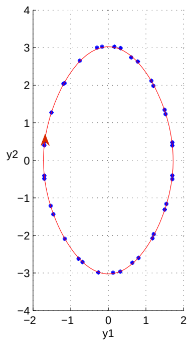


Fig.1: Near optimal state trajectory

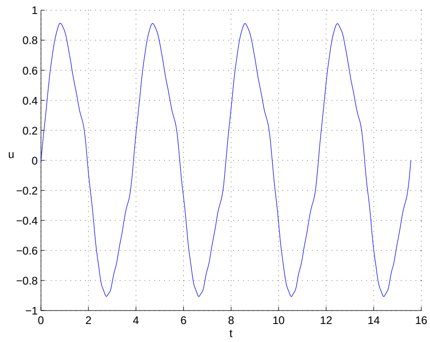


Fig.2: Near optimal control trajectory

solution of the corresponding N-approximating max-min problem

$$\eta^N(y) = \sum_{0<i_1+i_2\leq 10} \lambda^N_{i_1,i_2} y_1^{i_1} y_2^{i_2} \tag{3.4.3}$$

were found (note the change of notations with respect to (2.2.13)), and the optimal value of the SILP was evaluated to be ≈ -1.174.

In this case the problem (3.1.1) takes the form

$$\min_{u\in[-1,1]}\{\frac{\partial\eta^N(y_1,y_2)}{\partial y_1}y_2 + \frac{\partial\eta^N(y_1,y_2)}{\partial y_2}(u-0.3y_2-4sin(y_1)) + (u^2-y_1^2)\}.$$

The solution of the latter leads to the following representation for $u^N(y)$:

$$u^N(y) = \begin{cases} -\frac{1}{2}\frac{\partial\eta^N(y_1,y_2)}{\partial y_2} & \text{if} \quad |\frac{1}{2}\frac{\partial\eta^N(y_1,y_2)}{\partial y_2}| \leq 1, \\ -1 & \text{if} \quad -\frac{1}{2}\frac{\partial\eta^N(y_1,y_2)}{\partial y_2} < -1, \\ 1 & \text{if} \quad -\frac{1}{2}\frac{\partial\eta^N(y_1,y_2)}{\partial y_2} > 1. \end{cases} \tag{3.4.4}$$

Substituting this control into the system (1.1.1) and integrating it with the ode45 solver of MATLAB allows one to obtain the periodic ($T^* \approx 3.89$) state trajectory $\bar{y}^N(t) = (\bar{y}_1^N(t), \bar{y}_2^N(t))$ (see Figure 1) and the control trajectory $u^N(t)$ (see Figure 2). The value of the objective function numerically evaluated on the state control trajectory thus obtained is ≈ -1.174, the latter being the same as in SILP (within the given proximity). Note that the marked dots in Fig. 1 correspond to the concentration points of the measure γ^N (see (2.3.12)) that solves (2.3.4). The fact that the obtained state trajectory passes near these points and, most importantly, the fact that the value of the objective function obtained via integration is the same (within the given proximity) as the optimal value of the SILP problem indicate that the admissible solution found is a good approximation of the optimal one.

3.5 Additional comments for Chapter 3

The chapter is based on results obtained in [60]. Note that the algorithm for solving SILP problems described in Section 3.2 was originally proposed in [63]. Note also that the convergence of the algorithm was established in [63] under significantly more restrictive conditions then those used in Section 3.3 (see Remark 3.3.7).

Part II

On near optimal solution of singularly perturbed long run average optimal control problems

4

Augmented reduced and averaged IDLP problems related to singularly perturbed LRAOC problem

In this chapter, we build a foundation for the developments in the subsequent Chapters 5 and 6. The chapter consists of three sections. In Section 4.1, we introduce the singularly perturbed LRAOC problem and the corresponding IDLP problem. Also, we show that the asymptotic behaviour of the latter can be characterised with the help of a specially constructed "augmented" LP problem. In Section 4.2, we introduce the averaged LRAOC problem and the corresponding "averaged" IDLP problem. In Section 4.3, the augmented and the averaged IDLP problems are shown to be equivalent.

4.1 Singularly perturbed LRAOC problem and related perturbed IDLP problem. The augmented reduced IDLP problem.

Consider the singularly perturbed (SP) control system written in the form

$$\epsilon y'(t) = f(u(t), y(t), z(t)), \tag{4.1.1}$$
$$z'(t) = g(u(t), y(t), z(t)), \tag{4.1.2}$$

where $\epsilon > 0$ is a small parameter; $t \geq 0$; $f(u,y,z) : U \times I\!R^m \times I\!R^n \to I\!R^m$, $g(u,y,z) : U \times I\!R^m \times I\!R^n \to I\!R^n$ are continuous vector functions satisfying Lipschitz conditions in z and y; and where controls $u(\cdot) : [0, \mathcal{T}] \to U$ or $u(\cdot) : [0, +\infty) \to U$ (depending on whether the system is considered on the finite time interval $[0, \mathcal{T}]$ or on the infinite time interval $[0, +\infty)$) are measurable functions of time satisfying the inclusion

$$u(t) \in U, \tag{4.1.3}$$

U being a given compact metric space.

The presence of ϵ in the system (4.1.1)-(4.1.2) implies that the rate with which the y-components of the state variables change their values is of the order $\frac{1}{\epsilon}$ and is, therefore, much higher than the rate of change of the z-components (since ϵ is assumed to be small). Accordingly, the y-components and z-components of the state variables are referred to as *fast* and *slow*, respectively. The parameter ϵ is called *the small singular perturbation parameter.*

Let Y be a given compact subset of $I\!R^m$ and Z be a given compact subset of $I\!R^n$ such that the system (4.1.1)-(4.1.2) is viable in $Y \times Z$ for any $\epsilon > 0$ small enough (see the definition of viability in [16]).

Definition 4.1.1 *Let $u(\cdot)$ be a control and let $(y_\epsilon(\cdot), z_\epsilon(\cdot))$ be the corresponding solution of the system (4.1.1)-(4.1.2). The triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ will be called admissible on the interval $[0, \mathcal{T}]$ if*

$$(y_\epsilon(t), z_\epsilon(t)) \in Y \times Z \qquad \forall t \geq 0. \tag{4.1.4}$$

The triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ will be called admissible on $[0, \infty)$ if it is admissible on any interval $[0, \mathcal{T}]$, $\mathcal{T} > 0$.

We will be dealing with the LRAOC problem formulated as follows

$$\inf_{(u(\cdot),y_\epsilon(\cdot),z_\epsilon(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y_\epsilon(t), z_\epsilon(t))dt \stackrel{\text{def}}{=} V^*(\epsilon), \tag{4.1.5}$$

where $q(\cdot)$ is a continuous function and *inf* is sought over all admissible triplets of the singularly perturbed system. Note that, like in (1.1.3), the initial conditions are not fixed in (4.1.1)-(4.1.2) and finding them is a part of the optimization problem.

Similarly to Part I, along with the problem (4.1.5) we will be considering the following optimal control problem considered on the finite time interval

$$\inf_{(u(\cdot),y_\epsilon(\cdot),z_\epsilon(\cdot))} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y_\epsilon(t), z_\epsilon(t))dt \stackrel{\text{def}}{=} V^*(\epsilon, \mathcal{T}), \tag{4.1.6}$$

where *inf* is sought over all admissible triplets on the interval $[0, \mathcal{T}]$.

The SP optimal control problem (4.1.5) is related to infinite dimensional linear programming problem (see Chapter 1)

$$\min_{\gamma\in W(\epsilon)} \int_{U\times Y\times Z} q(u, y, z)\gamma(du, dy, dz) \stackrel{\text{def}}{=} G^*(\epsilon), \tag{4.1.7}$$

where the set $W(\epsilon) \subset \mathcal{P}(U \times Y \times Z)$ is defined by the equation

$$\begin{aligned} W(\epsilon) \stackrel{\text{def}}{=} \Big\{\gamma \in \mathcal{P}(U \times Y \times Z) \ : \int_{U\times Y\times Z} \Big[\nabla(\phi(y)\psi(z))^T \chi_\epsilon(u, y, z)\Big]\gamma(du, dy, dz) \\ = 0 \quad \forall\phi(\cdot) \in C^1(I\!R^m), \ \ \forall\psi(\cdot) \in C^1(I\!R^n)\Big\}, \end{aligned} \tag{4.1.8}$$

with $\chi_\epsilon(u, y, z)^T \stackrel{\text{def}}{=} (\frac{1}{\epsilon}\ f(u, y, z)^T,\ g(u, y, z)^T)$. Namely, the optimal values of this two problems are related by the inequality

$$V^*(\epsilon) \geq G^*(\epsilon) \qquad \forall\epsilon > 0, \tag{4.1.9}$$

and also, under certain conditions (see [48] and [58]),

$$V^*(\epsilon) = G^*(\epsilon) \qquad \forall\epsilon > 0. \tag{4.1.10}$$

Note that the set $W(\epsilon)$ would not change if the product test functions $\phi(y)\psi(z)$ in (4.1.8) are replaced with the test functions of the general form $\theta(y, z)$. This is due to the fact that any function $\theta(y, z) \in C^1(I\!R^m \times I\!R^n)$ and its gradient $\nabla\theta(y, z)$ can be simultaneously approximated on $Y\times Z$ by linear combinations of the product functions

$\phi(y)\psi(z)$ and their gradients $\nabla(\phi(y)\psi(z))$ (in fact, any function $\theta(y,z) \in C^1(I\!R^m \times I\!R^n)$) and its gradient can be simultaneously approximated on $Y \times Z$ by linear combinations of monomials $y_1^{i_1}...y_m^{i_m} z_1^{j_1}...z_n^{j_n}$, where $i_1,...,i_m = 0,1,...,\quad j_1,...,j_m = 0,1,...$; (see [84]).

In what follows, it will be assumed that (4.1.10) is valid and hence one can use the IDLP problem (4.1.7) to obtain some asymptotic properties of the problem (4.1.5).

Observe that, by multiplying the constraints in (4.1.8) by ϵ and taking into account the fact that

$$\nabla\Big(\phi(y)\psi(z)\Big)^T = \Big(\psi(z)\ (\nabla\phi(y))^T,\ \phi(y)\ (\nabla\psi(z))^T\Big)$$

one can rewrite the set $W(\epsilon)$ as follows

$$W(\epsilon) = \{\gamma \in \mathcal{P}(U \times Y \times Z)\ :\int_{U\times Y\times Z}\Big[\ \psi(z)(\nabla\phi(y))^T f(u,y,z) + \epsilon\Big(\phi(y)(\nabla\psi(z))^T$$

$$g(u,y,z)\Big)\Big]\gamma(du,dy,dz) = 0 \qquad \forall\phi(\cdot) \in C^1(I\!R^m),\ \ \forall\psi(\cdot) \in C^1(I\!R^n)\}. \tag{4.1.11}$$

Taking $\epsilon = 0$ in the expression above, one arrives at the set

$$W = \{\gamma \in \mathcal{P}(U \times Y \times Z)\ :\int_{U\times Y\times Z}[\ \psi(z)(\nabla\phi(y))^T f(u,y,z)\ ]\gamma(du,dy,dz) = 0$$

$$\forall\phi(\cdot) \in C^1(I\!R^m),\ \ \forall\psi(\cdot) \in C^1(I\!R^n)\}. \tag{4.1.12}$$

It is easy to see that $\limsup_{\epsilon\to 0} W(\epsilon) \subset W$. In general case, however, $W(\epsilon) \nrightarrow W$ when $\epsilon \to 0$, this being due to the fact that the equalities defining the set $W(\epsilon)$ contain some "implicit" constraints that are getting lost with equating ϵ to zero. In fact, by taking $\phi(y) = 1$, one can see that, if γ satisfies the equalities in (4.1.11), then it also satisfies the equality

$$\int_{U\times Y\times Z}[\ (\nabla\psi(z))^T g(u,y,z)]\ \gamma(du,dy,dz) = 0\ \ \forall\psi(\cdot) \in C^1(I\!R^n) \tag{4.1.13}$$

for any $\epsilon > 0$ (note that equality (4.1.13) have been reduced by ϵ). That is,

$$W(\epsilon) = W(\epsilon) \cap \mathcal{A} \qquad \forall\epsilon > 0, \tag{4.1.14}$$

where

$$\mathcal{A} \stackrel{\text{def}}{=} \Big\{ \gamma \in \mathcal{P}(U \times Y \times Z) \ : \int_{U\times Y\times Z} [(\nabla\psi(z))^T g(u,y,z)]\gamma(du,dy,dz) = 0 \quad \forall\psi(\cdot) \in C^1(I\!R^n) \Big\}. \tag{4.1.15}$$

Define the set $W^{\mathcal{A}}$ by the equation

$$W^{\mathcal{A}} \stackrel{\text{def}}{=} W \cap \mathcal{A} \tag{4.1.16}$$

and consider the IDLP problem

$$\min_{\gamma\in W^{\mathcal{A}}} \int_{U\times Y\times Z} q(u,y,z)\gamma(du,dy,dz) \stackrel{\text{def}}{=} G^{\mathcal{A}}. \tag{4.1.17}$$

We will be referring to this problem as to *augmented reduced* IDLP problem (the term *reduced* problem is commonly used for the problem obtained from a perturbed family by equating the small parameter to zero).

Proposition 4.1.2 *The following relationships are valid:*

$$\limsup_{\epsilon\to 0} W(\epsilon) \subset W^{\mathcal{A}}, \tag{4.1.18}$$

$$\liminf_{\epsilon\to 0} G^*(\epsilon) \geq G^{\mathcal{A}}. \tag{4.1.19}$$

Proof The validity of (4.1.18) is implied by (4.1.14), and the validity of (4.1.19) follows from (4.1.18). □

In Section 4.3, we will show that the inclusion (4.1.18) and the inequality (4.1.19) are replaced by equalities under the assumption that the averaged system (see Section 4.2) approximates the SP system on the infinite time horizon.

4.2 Averaged optimal control problem. Averaged IDLP problem.

Along with the SP system (4.1.1)-(4.1.2), let us consider a so-called *associated system*

$$y'(\tau) = f(u(\tau), y(\tau), z) \ , \qquad z = \text{const} \ . \tag{4.2.1}$$

Note that the associated system (4.2.1) looks similar to the "fast" subsystem (4.1.1) but, in contrast to (4.1.1), it is evolving in the "stretched" time scale $\tau = \frac{t}{\epsilon}$, with z being a vector of fixed parameters. Everywhere in what follows, it is assumed that the associated system is viable in Y (see the definition of viability in [16]).

Definition 4.2.1 *A pair $(u(\cdot), y(\cdot))$ will be called admissible for the associated system if (4.2.1) is satisfied for almost all τ ($u(\cdot)$ being measurable and $y(\cdot)$ being absolutely continuous functions) and if*

$$u(\tau) \in U, \quad y(\tau) \in Y \qquad \forall \tau \geq 0. \tag{4.2.2}$$

The occupational measure formulation of the problem (4.1.6) defined by associated system (4.2.1) implies definition of the set $\mathcal{M}(z, S, y)$ (see Section 1.1) that is the union of occupational measures generated on the interval $[0, S]$ by the admissible pairs of the associated system that satisfy the initial conditions $y(0) = y$. Namely,

$$\mathcal{M}(z, S, y) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot))} \{\mu^{(u(\cdot),\, y(\cdot))}\} \subset \mathcal{P}(U \times Y),$$

where $\mu^{(u(\cdot), y(\cdot))}$ is the occupational measure generated on the interval $[0, S]$ by an admissible pair of the associated system $(u(\cdot), y(\cdot))$ satisfying the initial condition $y(0) = y$ and the union is over such admissible pairs. Also, denote by $\mathcal{M}(z, S)$ the union of $\mathcal{M}(z, S, y)$ over all $y \in Y$. That is,

$$\mathcal{M}(z, S) \stackrel{\text{def}}{=} \bigcup_{y \in Y} \{\mathcal{M}(z, S, y)\}.$$

Let us define the set $W(z) \subset \mathcal{P}(U \times Y)$ as follows

$$\begin{aligned} W(z) \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) \ : \ & \int_{U \times Y} (\nabla \phi(y))^T f(u, y, z) \mu(du, dy) = 0 \\ & \forall \phi(\cdot) \in C^1(\mathbb{R}^m)\}. \end{aligned} \tag{4.2.3}$$

In [57] it has been established that

$$\limsup_{S \to \infty} \bar{co} \mathcal{M}(z, S) \subset W(z) \tag{4.2.4}$$

and that, under mild conditions (see Theorem 2.1 (i) in [57]),

$$\lim_{S \to \infty} \rho_H(\bar{co} \mathcal{M}(z, S), W(z)) = 0, \tag{4.2.5}$$

where $\bar{co}$ stands for the closed convex hull of the corresponding set. Also, it has been established that, under some additional conditions (see Theorem 2.1(ii),(iii) and Proposition 4.1 in [57]),

$$\lim_{S\to\infty} \rho_H(\mathcal{M}(z,S,y), W(z)) = 0 \quad \forall\, y \in Y, \tag{4.2.6}$$

with the convergence being uniform with respect to $y \in Y$.

Define the function $\tilde{g}(\mu, z) : \mathcal{P}(U \times Y) \times Z \to I\!R^n$ by the equation

$$\tilde{g}(\mu, z) \stackrel{\text{def}}{=} \int_{U\times Y} g(u, y, z)\mu(du, dy) \quad \forall \mu \in \mathcal{P}(U \times Y) \tag{4.2.7}$$

and consider the system

$$z'(t) = \tilde{g}(\mu(t), z(t)), \tag{4.2.8}$$

in which the role of controls is played by measure valued functions $\mu(\cdot)$ that satisfy the inclusion

$$\mu(t) \in W(z(t)). \tag{4.2.9}$$

The system (4.2.8) will be referred to as *the averaged system*. In what follows, it is assumed that the averaged system is viable in Z.

Definition 4.2.2 *A pair $(\mu(\cdot), z(\cdot))$ will be referred to as admissible for the averaged system if (4.2.8) and (4.2.9) are satisfied for almost all t ($\mu(\cdot)$ being measurable and $z(\cdot)$ being absolutely continuous functions) and if*

$$z(t) \in Z \qquad \forall t \geq 0. \tag{4.2.10}$$

From Theorem 2.8 of [56] it follows that, under the assumption that (4.2.6) is satisfied (and under other assumptions including the Lipschitz continuity of the multi-valued map $V(z) \stackrel{\text{def}}{=} \cup_{\mu\in W(z)} \{\tilde{g}(\mu, z)\}$), the averaged system approximates the SP dynamics on the infinite time horizon in the sense that the following two statements are valid:

(i) Given an admissible triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the SP system (4.1.1)-(4.1.2) that satisfies the initial condition

$$(y_\epsilon(0), z_\epsilon(0)) = (y_0, z_0), \tag{4.2.11}$$

there exists an admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ satisfying the initial condition

$$z(0) = z_0 \tag{4.2.12}$$

such that

$$\sup_{t\in[0,\infty)} ||z_\epsilon(t) - z(t)|| \leq \delta(\epsilon), \qquad \text{where} \qquad \lim_{\epsilon\to 0} \delta(\epsilon) = 0 \tag{4.2.13}$$

and, for any Lipschitz continuous functions $h(u, y, z)$,

$$\sup_{\mathcal{T}>0} |\frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(u(t), y_\epsilon(t), z_\epsilon(t))dt - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t))dt| \leq \delta_h(\epsilon), \tag{4.2.14}$$

where $\quad \lim_{\epsilon\to 0} \delta_h(\epsilon) = 0$ and where

$$\tilde{h}(\mu, z) \stackrel{\text{def}}{=} \int_{U\times Y} h(u, y, z)\mu(du, dy) \quad \forall \mu \in \mathcal{P}(U \times Y); \tag{4.2.15}$$

(ii) Let $(\mu(\cdot), z(\cdot))$ be an admissible pair of the averaged system satisfying the initial condition (4.2.12). There exists an admissible triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the SP system satisfying the initial condition (4.2.11) such that the estimates (4.2.13) and (4.2.14) are true.

Note that the statements (i) and (ii) can be interpreted as a justifications of a decomposition of the slow-fast dynamics interaction into two phases. First is the "fast" phase, during which the slow variables almost do not move, and the dynamics of the fast components is approximately described by the associated system (4.2.1), with the set of occupational measures generated by the latter converging to the limit set $W(z)$. Second is a "slow" phase. During this phase, the slow state variables evolve according to the averaged system (4.2.8), with the fast dynamics influencing the slow one only through "limit" occupational distributions which take values in $W(z)$.

Let us introduce the following definition.

Definition 4.2.3 *The averaged system will be said to uniformly approximate the SP system on Y if the statements (i) and (ii) are valid, with the estimates (4.2.13) and (4.2.14) being uniform with respect to the initial conditions $(y_0, z_0) \in Y \times Z$.*

Note that a special case, in which the averaged system uniformly approximates the SP system, is considered below (see Assumption 6.1.2 in Section 6.1).

Consider the optimal control problem

$$\inf_{(\mu(\cdot),z(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(t), z(t))dt \stackrel{\text{def}}{=} \tilde{V}^*, \tag{4.2.16}$$

where

$$\tilde{q}(\mu, z) \stackrel{\text{def}}{=} \int_{U\times Y} q(u, y, z)\mu(du, dy) \tag{4.2.17}$$

and where *inf* is sought over all admissible pairs of the averaged system (4.2.8). This problem will be referred to as *averaged optimal control problem.*

Proposition 4.2.4 *If the averaged system uniformly approximates the SP system, then*

$$\lim_{\epsilon\to 0} V^*(\epsilon) = \tilde{V}^*. \tag{4.2.18}$$

Proof. Let us show that

$$\inf_{(u(\cdot),y(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y(t), z(t))dt \geq \inf_{(\mu(\cdot),z(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(\tau), z(\tau))d\tau - \delta_q(\epsilon) \tag{4.2.19}$$

where $\delta_q(\epsilon)$ is defined as in (4.2.14).

Take an arbitrary admissible triplet $(u(t), y(t), z(t))$ and let $\mathcal{T}_i \to \infty$ be such that

$$\lim_{i\to\infty} \frac{1}{\mathcal{T}_i} \int_0^{\mathcal{T}_i} q(u(\tau), y(\tau), z(\tau))dt = \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y(t), z(t))dt. \tag{4.2.20}$$

Due to our assumption, there exists an admissible pair $(\mu(\tau), z(\tau))$ such that

$$|\frac{1}{\mathcal{T}_i} \int_0^{\mathcal{T}_i} q(u(t), y(t), z(t))dt - \frac{1}{\mathcal{T}_i} \int_0^{\mathcal{T}_i} \tilde{q}(\mu(\tau), z(\tau))d\tau| \leq \delta_q(\epsilon). \tag{4.2.21}$$

From (4.2.20) and (4.2.21) it follows that

$$\liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y(t), z(t))dt \geq \liminf_{i\to\infty} \frac{1}{\mathcal{T}_i} \int_0^{\mathcal{T}_i} \tilde{q}(\mu(\tau)z(\tau))d\tau - \delta_q(\epsilon) \geq$$

$$\liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(\tau), z(\tau))d\tau - \delta_q(\epsilon) \geq \inf_{(\mu,z)} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(\tau), z(\tau))d\tau - \delta_q(\epsilon). \tag{4.2.22}$$

Note that, the obtained inequality is true for any admissible triplet $(u(t), y(t), z(t))$. Hence, (4.2.19) is proved. In a similar way one can show that

$$\inf_{(u(\cdot),y(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t), y(t), z(t))dt \leq \inf_{(\mu(\cdot),z(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(\tau), z(\tau))d\tau - \delta_q(\epsilon). \tag{4.2.23}$$

□

The averaged optimal control problem (4.2.16) is related to the IDLP problem

$$\min_{\xi\in\tilde{W}} \int_F \tilde{q}(\mu, z)\xi(d\mu, dz) \stackrel{\text{def}}{=} \tilde{G}^*, \tag{4.2.24}$$

where F is the graph of the map $W(\cdot)$ (see (4.2.3)),

$$F \stackrel{\text{def}}{=} \{(\mu, z) \;:\; \mu \in W(z), \quad z \in Z\} \;\subset \mathcal{P}(U \times Y) \times Z \;, \tag{4.2.25}$$

and the set $\tilde{W} \subset \mathcal{P}(F)$ is defined by the equation

$$\tilde{W} \stackrel{\text{def}}{=} \{\xi \in \mathcal{P}(F) : \int_F (\nabla\psi(z))^T \tilde{g}(\mu, z)\xi(d\mu, dz) = 0 \quad \forall\psi(\cdot) \in C^1(I\!R^n)\}. \tag{4.2.26}$$

This IDLP problem plays an important role in our consideration and, for convenience, we will be referring to it as to *the averaged IDLP problem*.

The relationships between the problems (4.2.16) and (4.2.24) include, in particular, the inequality between the optimal values

$$\tilde{V}^* \geq \tilde{G}^*, \tag{4.2.27}$$

which, under certain conditions (see Theorem 2.2 of [62]), takes the form of the equality

$$\tilde{V}^* = \tilde{G}^*. \tag{4.2.28}$$

4.3 Equivalence of the augmented reduced and the averaged IDLP problems.

In this section, we are going to establish that the averaged IDLP problem (4.2.24) is equivalent to the augmented reduced IDLP problem (4.1.17).

Let us first of all observe that the set $W^{\mathcal{A}}$ allows another representation which makes use of the fact that an arbitrary $\gamma \in \mathcal{P}(U \times Y \times Z)$ can be "disintegrated" as follows

$$\gamma(du, dy, dz) = \mu(du, dy|z)\nu(dz), \tag{4.3.1}$$

where $\nu(dz) \stackrel{\text{def}}{=} \gamma(U \times Y, dz)$ and $\mu(du, dy|z)$ is a probability measure on $U \times Y$ such that the integral $\int_{U\times Y} h(u, y, z)\mu(du, dy|z)$ is Borel measurable on Z for any continuous

$h(u, y, z)$ and

$$\int_{U\times Y\times Z} h(u, y, z)\gamma(du, dy, dz) = \int_Z \left(\int_{U\times Y} h(u, y, z)\mu(du, dy|z) \right) \nu(dz).$$

The fact that the disintegration (4.3.1) is valid follows from the existence of "regular conditional probabilities" for probability measures defined on the Borel subsets of compact metric spaces (see, e.g., Definition 6.6.3 and Theorems 6.6.5 and 6.6.6 in [14]).

Proposition 4.3.1 *The set $W^{\mathcal{A}}$ can be represented in the form:*

$$\begin{aligned} W^{\mathcal{A}} = \{\gamma = \mu(du, dy|z)\nu(dz) \ : \ \mu(\cdot|z) \in W(z) \ \ for \ \ \nu - almost \ all \ z \in Z, \\ \int_Z [\ (\nabla\psi(z))^T \tilde{g}(\mu(\cdot|z), z)]\nu(dz) = 0 \ \ \forall \ \psi(\cdot) \in C^1(I\!\!R^n)\}, \end{aligned} \tag{4.3.2}$$

where $W(z) \subset \mathcal{P}(U \times Y)$ is defined as in (4.2.3) and $\tilde{g}(\mu(\cdot|z), z) = \int_{U\times Y} g(u, y, z)$ $\mu(du, dy|z)$.

Proof. Let γ belong to the right hand side of (4.3.2). Then, by (4.3.1),

$$\begin{aligned} &\int_{U\times Y\times Z} \psi(z)(\nabla\phi(y))^T f(u, y, z)\gamma(du, dy, dz) \\ &= \int_Z [\psi(z) \int_{U\times Y} (\nabla\phi(y))^T f(u, y, z)\mu(du, dy|z)]\nu(dz) \ = \ 0 \end{aligned}$$

for any $\phi(\cdot) \in C^1(I\!\!R^m)$ and for any $\psi(\cdot) \in C^1(I\!\!R^n)$ (the equality to zero being due to the fact that $\mu(\cdot, \cdot|z) \in W(z)$ for ν- almost all $z \in Z$). Also,

$$\begin{aligned} &\int_{U\times Y\times Z} [\ (\nabla\psi(z))^T g(u, y, z) \ ]\gamma(du, dy, dz) \\ &= \int_Z [\ (\nabla\psi(z))^T \int_{U\times Y} g(u, y, z)\mu(du, dy|z) \ ]\nu(dz) \ = \ 0. \end{aligned}$$

These imply that $\gamma \in W^{\mathcal{A}}$. Assume now that $\gamma \in W^{\mathcal{A}}$. That is, $\gamma \in W$ and $\gamma \in \mathcal{A}$ (see (4.1.12), (4.1.15) and (4.1.16)). Using the fact that $\gamma \in W$ (and taking into account the disintegration (4.3.1)), one can obtain that

$$\int_{z\in Z} [\psi(z) \int_{U\times Y} (\nabla\phi(y))^T f(u, y, z)\mu(du, dy|z) \ ]\nu(dz) = 0, \tag{4.3.3}$$

the latter implying that $\int_{U\times Y} (\nabla\phi(y))^T f(u, y, z)\mu(du, dy|z) = 0$ for ν-almost all $z \in Z$ (due to the fact that $\psi(z)$ is an arbitrary continuously differentiable function). That is, $\mu(\cdot|z) \in W(z)$ for ν-almost all $z \in Z$. This, along with the inclusion $\gamma \in \mathcal{A}$, imply that γ belongs to the right hand side of (4.3.2). This proves (4.3.2). $\Box$

To establish the relationships between the augmented reduced and the averaged IDLP problems, let us introduce the map $\Phi(\cdot) : \mathcal{P}(F) \to \mathcal{P}(U \times Y \times Z)$ defined as follows. For any $\xi \in \mathcal{P}(F)$, let $\Phi(\xi) \in \mathcal{P}(U \times Y \times Z)$ be such that

$$\int_{U\times Y\times Z} h(u,y,z)\Phi(\xi)(du,dy,dz) = \int_F \tilde{h}(\mu,z)\xi(d\mu,dz) \qquad \forall h(\cdot) \in C(U \times Y \times Z), \tag{4.3.4}$$

where $\tilde{h}(\mu,z) = \int_{U\times Y} h(u,y,z)\mu(du,dy)$ (this definition is legitimate since the right-hand side of the above expression defines a linear continuous functional on $C(U\times Y\times Z)$, the latter being associated with an element of $\mathcal{P}(U \times Y \times Z)$ that makes the equality (4.3.4) valid). Note that the map $\Phi(\cdot) : \mathcal{P}(F) \to \mathcal{P}(U \times Y \times Z)$ is linear and it is continuous in the sense that

$$\lim_{\xi_l \to \xi} \Phi(\xi_l) = \Phi(\xi), \tag{4.3.5}$$

with ξ_l converging to ξ in the weak* topology of $\mathcal{P}(F)$ and $\Phi(\xi_l)$ converging to $\Phi(\xi)$ in the weak* topology of $\mathcal{P}(U \times Y \times Z)$ (see Lemma 4.3 in [56]).

The following result establishes that the averaged IDLP problem (4.2.24) is equivalent to the augmented reduced IDLP problem (4.1.17).

Proposition 4.3.2 *The averaged and the augmented reduced IDLP problems are equivalent in the sense that*

$$W^{\mathcal{A}} = \Phi(\tilde{W}), \tag{4.3.6}$$

$$G^{\mathcal{A}} = \tilde{G}^*. \tag{4.3.7}$$

Also, $\gamma = \Phi(\xi)$ is an optimal solution of the augmented reduced IDLP problem (4.1.17) if and only if ξ is an optimal solution of the averaged IDLP problem (4.2.24).

Proof. To prove (4.3.6), let us first prove that the inclusion

$$\Phi(\tilde{W}) \subset W^{\mathcal{A}} \tag{4.3.8}$$

is valid. Take an arbitrary $\gamma \in \Phi(\tilde{W})$. That is, $\gamma = \Phi(\xi)$ for some $\xi \in \tilde{W}$. By (4.3.4),

$$\int_{U\times Y\times Z} [\ \psi(z)(\nabla\phi(y))^T f(u,y,z)\ ]\Phi(\xi)(du,dy,dz)$$

$$= \int_F [\ \psi(z) \int_{U\times Y} (\nabla\phi(y))^T f(u,y,z)\mu(du,dy)\ ]\xi(d\mu,dz).$$

By definition of F (see (4.2.25)),

$$\int_{U\times Y} (\nabla\phi(y))^T f(u,y,z)\mu(du,dy) = 0 \qquad \forall(\mu,z)\in F.$$

Consequently (see (4.1.12)),

$$\int_{U\times Y\times Z} [\ \psi(z)(\nabla\phi(y))^T f(u,y,z)\ ]\Phi(\xi)(du,dy,dz) = 0 \quad \Rightarrow \quad \Phi(\xi)\in W. \tag{4.3.9}$$

Also, from (4.3.4) and from the fact that $\xi\in\tilde{W}$ it follows that

$$\int_{U\times Y\times Z} [\ (\nabla\phi(z))^T g(u,y,z)\ ]\Phi(\xi)(du,dy,dz)$$
$$= \int_F [\ (\nabla\phi(z))^T \tilde{g}(\mu,z)\ ]\xi(d\mu,dz) \ = \ 0 \quad \Rightarrow \quad \Phi(\xi)\in\mathcal{A}.$$

Thus, $\gamma = \Phi(\xi)\subset W\cap\mathcal{A}$. This proves (4.3.8). Let us now show that the converse inclusion

$$\Phi(\tilde{W})\supset W^{\mathcal{A}} \tag{4.3.10}$$

is valid. To this end, take $\gamma\in W^{\mathcal{A}}$ and show that $\gamma\in\Phi(\tilde{W})$. Due to (4.3.2), γ can be presented in the form (4.3.1) with $\mu(du,dy|z)\in W(z)$ for ν almost all $z\in Z$. Changing values of μ on a subset of Z having the ν measure 0, one can come to the conclusion that γ can be presented in the form (4.3.1) with

$$\mu(du,dy|z)\in W(z) \qquad \forall z\in Z. \tag{4.3.11}$$

Let $\mathcal{L}$ be a subspace of $C(F)$ defined by the equation

$$\mathcal{L}\stackrel{\text{def}}{=}\{\tilde{h}(\cdot,\cdot)\ :\ \tilde{h}(\mu,z) = \int_{U\times Y} h(u,y,z)\mu(du,dy),\quad h\in C(U\times Y\times Z)\}. \tag{4.3.12}$$

For every $\tilde{h}\in\mathcal{L}$, let $\xi_{\mathcal{L}}\tilde{h}) : \mathcal{L}\to I\!R^1$ be defined by the equation

$$\xi_{\mathcal{L}}(\tilde{h})\stackrel{\text{def}}{=}\int_{z\in Z}\tilde{h}(\mu(\cdot|z),z)\nu(dz) = \int_{z\in Z}\Big[\int_{U\times Y} h(u,y,z)\mu(du,dy|z)\Big]\nu(dz) \tag{4.3.13}$$
$$= \int_{U\times Y} h(u,y,z)\gamma(du,dy,dz).$$

Note that $\xi_{\mathcal{L}}$ is a positive linear functional on $\mathcal{L}$. That is, if $\tilde{h}_1(\mu,z)\le\tilde{h}_2(\mu,z)\ \forall(\mu,z)\in F$, then $\xi_{\mathcal{L}}(\tilde{h}_1)\le\xi_{\mathcal{L}}(\tilde{h}_2)$. Note also that $1\in\mathcal{L}$. Hence, by Kantorovich theorem (see, e.g., [1], p. 330), ξ_F can be extended to a positive linear functional ξ on the whole $C(F)$, with

$$\xi(\tilde{h}) = \xi_{\mathcal{L}}(\tilde{h}) \quad \forall \tilde{h} \in \mathcal{L}. \tag{4.3.14}$$

Due to the fact that ξ is positive, one obtains that

$$\sup_{\beta(\cdot)\in\bar{B}} \xi(\beta(\cdot)) \leq \sup_{\beta(\cdot)\in\bar{B}} \xi(|\beta(\cdot)|) \leq \xi(1) = 1, \tag{4.3.15}$$

where $\bar{B}$ is the closed unit ball in $C(F)$ (that is, $\bar{B} \stackrel{\text{def}}{=} \{\beta(\cdot) \in C(F) \ : \ \max_{(\mu,z)\in F} |\beta(\mu, z)| \leq 1\}$). Thus, $\xi \in C(F)$, and, moreover, $||\xi|| = \xi(1) = 1$. This implies that there exists a unique probability measure $\xi(d\mu, dz) \in \mathcal{P}(F)$ such that, for any $\beta(\mu, z) \in C(F)$,

$$\xi(\beta) = \int_F \beta(\mu, z)\xi(d\mu, dz) \tag{4.3.16}$$

(see, e.g., Theorem 5.8 on page 38 in [91]). Using this relationship for $\beta(\mu, z) = \tilde{h}(\mu, z) \in F$, one obtains (see (4.3.13) and (4.3.14)) that

$$\int_F \tilde{h}(\mu, z)\xi(d\mu, dz) = \int_{U\times Y} h(u, y, z)\gamma(du, dy, dz). \tag{4.3.17}$$

Since the latter is valid for any $h(u, y, z) \in C(U \times Y \times Z)$, it follows that

$$\gamma = \Phi(\xi). \tag{4.3.18}$$

Considering now (4.3.17) with $h(u, y, z) = \nabla\psi(z)^T g(u, y, z)$, and taking into account that, in this case, $\tilde{h}(\mu, z) = \nabla\psi(z)^T \tilde{g}(\mu, z)$ one obtains that

$$\int_F \Big[(\nabla\psi(z))^T \tilde{g}(\mu, z)\Big]\xi(d\mu, dz)$$
$$= \int_{U\times Y\times Z} \Big[(\nabla\psi(z))^T g(u, y, z)\Big]\gamma(du, dy, dz) = 0,$$

where the equality to zero follows from the fact that $\gamma \in \mathcal{A}$ (see (4.1.15)). This implies that $\xi \in \tilde{W}$. Hence, by (4.3.18), $\gamma \in \Phi(\tilde{W})$. This proves (4.3.6).

The validity of (4.3.7) as well as the fact that $\gamma = \Phi(\xi)$ is optimal in (4.1.17) if and only if ξ is optimal on (4.2.24) follow from (4.3.6) and the definition of the map $\Phi(\cdot)$ (see (4.3.4)). $\square$

Proposition 4.3.3 *The following relationships are valid*

$$\liminf_{\epsilon\to 0} V^*(\epsilon) \geq \liminf_{\epsilon\to 0} G^*(\epsilon) \geq \tilde{G}^*. \tag{4.3.19}$$

If the averaged system uniformly approximates the SP system (see Definition 4.2.3) and if (4.2.28) is valid, then

$$\lim_{\epsilon\to 0} V^*(\epsilon) = \lim_{\epsilon\to 0} G^*(\epsilon) = \tilde{G}^*. \tag{4.3.20}$$

Proof. Note that the first inequality in (4.3.19) follows from (4.1.9).The validity of (4.3.20) follows from (4.2.18), (4.2.28) and the second inequality in (4.3.19).

Thus, to prove the proposition, it is sufficient to establish the validity of the second inequality in (4.3.19), which can be proved on the basis of Proposition 4.1.2 and Proposition 4.3.2. More specifically, by (4.1.19)

$$\liminf_{\epsilon\to 0} G^*(\epsilon) \geq \min_{\gamma\in W^{\mathcal{A}}} \int_{U\times Y\times Z} q(u,y,z)\gamma(du,dy,dz). \tag{4.3.21}$$

Also, by (4.3.6),

$$\begin{gathered}\min_{\gamma\in W^{\mathcal{A}}} \int_{U\times Y\times Z} q(u,y,z)\gamma(du,dy,dz) = \\ \min_{\zeta\in \tilde{W}} \int_{U\times Y\times Z} q(u,y,z)\Phi(\xi)(du,dy,dz) = \tilde{G}^*.\end{gathered} \tag{4.3.22}$$

By comparing (4.3.21) and (4.3.22), one obtains the second inequality in (4.3.19). □

4.4 Additional comments for Chapter 4

The consideration of this chapter is based on the results obtained in [61] and [64]. Note that the paper [64] is devoted to the development of LP approach to singularly perturbed optimal control problems with time discounting. Many results obtained in this paper are formulated in terms of IDLP problems. These results are straightforwardly extendable to the IDLP problems related to singularly perturbed LRAOC problems. Examples of such extensions are Proposition 4.1.2, Proposition 4.3.1 and Proposition 4.3.2. These can be proved by equating the discount rate to zero in the corresponding statements of [64]. We, however, do not follow this path and we give the full proofs of the aforementioned propositions (for the sake of completeness).

Note that a most common approach to problems of optimal control of singular perturbed systems is based on the idea of approximating the slow z-components of the solutions of the SP system (4.1.1)-(4.1.2) by the solutions of the so-called reduced system

$$z'(t) = g(u(t), s(u(t), z(t)), z(t)), \tag{4.4.1}$$

which is obtained from (4.1.1) via the replacement of $y(t)$ by $s(u(t), z(t))$, with $s(u, z)$ being the root of the equation

$$f(u, y, z) = 0. \tag{4.4.2}$$

The equation (4.4.2) can be obtained by formally equating ϵ to zero in (4.1.1). Being very efficient in dealing with many important classes of optimal control problems (see, e.g., [22], [39], [74], [76], [78], [88], [90], [102], [105]), this approach may not be applicable in the general case (see examples in [8], [51], [52], [82]).

The validity of the assertion that the system (4.4.1) can be used for finding a near optimal control of the SP system (4.1.1)-(4.1.2) is related to the validity of the hypothesis that the optimal control of the latter is in some sense slow and that (in the optimal or near optimal regime) the fast state variables converge rapidly to their quasi steady states defined by the root of (4.4.2) and remain in a neighborhood of this root, while the slow variables are changing in accordance with (4.4.1). While the validity of such a hypothesis has been established under natural stability conditions by famous Tikhonov's theorem in the case of uncontrolled dynamics (see [89] and [101]), this hypothesis may not be valid in the control setting if the dynamics is nonlinear and/or the objective function is non-convex, the reason for this being the fact that the use of rapidly oscillating controls may lead to significant (not tending to zero with ϵ) improvements of the performance indexes. The approach that we are developing in the thesis allows the construction of such near optimal rapidly oscillating controls.

5

Average control generating (ACG) families

In this chapter, we introduce the concept of an average control generating (ACG) family and we use duality results for the IDLP problems involved and their semi-infinite approximations to characterize and construct optimal and near optimal ACG families. The chapter consists of six sections. In Section 5.1, the definitions of an ACG family and of optimal/near optimal ACG families are given. In Section 5.2, averaged and associated dual problems are introduced. Also in this section, sufficient and necessary optimality conditions for an ACG family to be optimal is established. In Section 5.3, approximating averaged semi-infinite dimensional linear programming (SILP) problem and the corresponding approximating averaged and associated dual problems are introduced. In Section 5.4, it is proved that solutions of these approximating dual problems exist under natural controllability conditions. In Sections 5.5 and 5.6, it is established that solutions of the approximating averaged and associated dual problems can be used for construction of near optimal ACG families.

5.1 Average control generating families.

The validity of the representation (4.3.2) for the set $W^{\mathcal{A}}$ motivates the definition of the *average control generating family* given below. For any $z \in Z$, let $(u_z(\cdot), y_z(\cdot))$

be an admissible pair of the associated system (4.2.1) and $\mu(du, dy|z)$ be the occupational measure generated by this pair on $[0, \infty)$ (see (1.1.20)), with the integral $\int_{U\times Y} h(u, y, z)\mu(du, dy|z)$ being a measurable function of z and

$$|\frac{1}{S}\int_0^S h(u_z(\tau), y_z(\tau), z)d\tau - \int_{U\times Y} h(u, y, z)\mu(du, dy|z)| \leq \delta_h(S) \quad \forall z \in Z,$$

$$\lim_{S\to\infty} \delta_h(S) = 0, \tag{5.1.1}$$

for any continuous $h(u, y, z) : U \times Y \times Z \to I\!R^1$. Note that the estimate (5.1.1) is valid if $(u_z(\cdot), y_z(\cdot))$ is T_z-periodic, with T_z being uniformly bounded on Z. Note that due to (4.2.4),

$$\mu(du, dy|z) \in W(z) \qquad \forall z \in Z. \tag{5.1.2}$$

Definition 5.1.1 *The family $(u_z(\cdot), y_z(\cdot))$ will be called average control generating (ACG) if the system*

$$z'(t) = \tilde{g}_\mu(z(t)), \qquad z(0) = z_0, \tag{5.1.3}$$

where

$$\tilde{g}_\mu(z) \stackrel{\text{def}}{=} \tilde{g}(\mu(\cdot|z), z) = \int_{U\times Y} g(u, y, z)\mu(du, dy|z), \tag{5.1.4}$$

has a unique solution $z(t) \in Z$ $\forall t \in [0, \infty)$ and, for any continuous function $\tilde{h}(\mu, z) : F \to I\!R^1$, there exists a limit

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t))dt, \tag{5.1.5}$$

where $\mu(t) \stackrel{\text{def}}{=} \mu(du, dy|z(t))$.

Note that, according to this definition, if $(u_z(\cdot), y_z(\cdot))$ is an ACG family, with $\mu(du, dy|z)$ being the family of occupational measures generated by this family, and if $z(\cdot)$ is the corresponding solution of (5.1.3), then the pair $(\mu(\cdot), z(\cdot))$, where $\mu(t) \stackrel{\text{def}}{=} \mu(du, dy|z(t))$, is an admissible pair of the averaged system (for convenience, this admissible pair will also be referred to as one generated by the ACG family). From the fact that the limit (5.1.5) exists for any continuous $\tilde{h}(\mu, z)$ it follows that the pair $(\mu(\cdot), z(\cdot))$ generates the occupational measure $\xi \in \mathcal{P}(F)$ defined by the equation

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t))dt = \int_F \tilde{h}(\mu, z)\xi(d\mu, dz) \quad \forall \tilde{h}(\mu, z) \in C(F). \tag{5.1.6}$$

Also note that, the state trajectory $z(\cdot)$ generates the occupational measure $\nu \in \mathcal{P}(Z)$ defined by the equation

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(z(t))dt = \int_Z h(z)\nu(dz) \qquad \forall h(z) \in C(Z). \tag{5.1.7}$$

Proposition 5.1.2 *Let $(u_z(\cdot), y_z(\cdot))$ be an ACG family and let $\mu(du, dy|z)$ and $(\mu(\cdot), z(\cdot))$ be, respectively, the family of occupational measures and the admissible pair of the averaged system generated by this family. Let also, ξ be the occupational measure generated by $(\mu(\cdot), z(\cdot))$ and ν be the occupational measure generated by $z(\cdot)$ (in accordance with (5.1.6) and (5.1.7) respectively). Then*

$$\xi \in \tilde{W} \tag{5.1.8}$$

and

$$\Phi(\xi) = \mu(du, dy|z)\nu(dz), \tag{5.1.9}$$

where $\Phi(\cdot)$ is defined by (4.3.4).

Proof. For an arbitrary $\psi(\cdot) \in C^1(\mathbb{R}^n)$,

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} (\nabla\psi(z(t)))^T \tilde{g}(\mu(t), z(t))dt = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}}\Big(\psi(z(\mathcal{T})) - \psi(z(0))\Big) = 0.$$

Hence, by (5.1.6),

$$\int_F (\nabla\psi(z))^T \tilde{g}(\mu, z)\xi(d\mu, dz) = 0 \qquad\qquad \psi(\cdot) \in C^1(\mathbb{R}^n).$$

The latter implies (5.1.8). To prove (5.1.9), note that, for an arbitrary continuous function $h(u, y, z)$ and $\tilde{h}(\mu, z)$ defined in accordance with (4.2.15), one can write down

$$\int_F \tilde{h}(\mu, z)\xi(d\mu, dz) = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t))dt =$$

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \Big[\int_{U\times Y} h(u, y, z)\mu(du, dy|z(t))\Big]dt = \int_Z \Big(\int_{U\times Y} h(u, y, z)\mu(du, dy|z)\Big)\nu(dz). \tag{5.1.10}$$

By the definition of $\Phi(\cdot)$ (see (4.3.4)), the latter implies (5.1.9). $\square$

Definition 5.1.3 *An ACG family* $(u_z(\cdot), y_z(\cdot))$ *will be called optimal if the admissible pair* $(\mu(\cdot), z(\cdot))$ *generated by this family is optimal in the averaged problem (4.2.16). That is,*

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(t), z(t))dt = \tilde{V}^*. \tag{5.1.11}$$

An ACG family $(u_z(\cdot), y_z(\cdot))$ *will be called* α*-near optimal* ($\alpha > 0$) *if*

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(t), z(t))dt \leq \tilde{V}^* + \alpha. \tag{5.1.12}$$

Corollary 5.1.4 *Let the equality (4.2.28) be valid. An ACG family* $(u_z(\cdot), y_z(\cdot))$ *generating the admissible pair* $(\mu(\cdot), z(\cdot))$ *will be optimal (near optimal) if and only if the occupational measure* ξ *generated by this pair (according to (5.1.6)) is an optimal (near optimal) solution of the averaged IDLP problem (4.2.24).*

Remark 5.1.5 *The "near optimal" solutions are defined by the value of the objective function. To measure the error of the approximation in terms of solutions would require imposing stronger regularity conditions.These are not considered in the thesis.*

5.2 Averaged and associated dual problems. Sufficient and necessary optimality condition.

Let $\tilde{H}(p, z)$ be the Hamiltonian of the averaged system

$$\tilde{H}(p, z) \stackrel{\text{def}}{=} \min_{\mu \in W(z)} \{\tilde{q}(\mu, z) + p^T \tilde{g}(\mu, z)\}, \tag{5.2.1}$$

where $\tilde{g}(\mu, z)$ and $\tilde{q}(\mu, z)$ are defined by (4.2.7) and (4.2.17).

Consider the problem

$$\sup_{\zeta(\cdot) \in C^1(R^n)} \{\tilde{d} : \tilde{d} \leq \tilde{H}(\nabla\zeta(z), z) \quad \forall z \in Z\} = \tilde{G}^*, \tag{5.2.2}$$

where *sup* is sought over all continuously differentiable functions $\zeta(\cdot) : I\!R^n \to I\!R^1$. Note that the optimal value of the problem (5.2.2) is equal to the optimal value of the averaged IDLP problem (4.2.24). The former is in fact dual with respect to the later, the equality of the optimal values being one of the duality relationships between the two (see Theorem 1.3.1). For brevity, (5.2.2) will be referred to as just *averaged dual problem.* Note that the averaged dual problem can be equivalently rewritten in the form

$$\sup_{\zeta(\cdot)\in C^1(\mathbb{R}^n)} \{\tilde{d} : \tilde{d} \leq \tilde{q}(\mu, z) + (\nabla\zeta(z))^T \tilde{g}(\mu, z) \ \ \forall(\mu, z) \in F\} = \tilde{G}^*, \tag{5.2.3}$$

where F is the graph of $W(\cdot)$ (see (4.2.25)). A function $\zeta^*(\cdot) \in C^1(\mathbb{R}^n)$ will be called a *solution of the averaged dual problem* if

$$\tilde{G}^* \leq \tilde{H}(\nabla\zeta^*(z), z) \ \ \forall z \in Z \ , \tag{5.2.4}$$

or, equivalently, if

$$\tilde{G}^* \leq \tilde{q}(\mu, z) + (\nabla\zeta^*(z))^T \tilde{g}(\mu, z) \ \ \forall(\mu, z) \in F \ . \tag{5.2.5}$$

Note that, if $\zeta^*(\cdot) \in C^1(\mathbb{R}^n)$ satisfies (5.2.4), then $\zeta^*(\cdot) + const$ satisfies (5.2.4) as well.

Assume that a solution of the averaged dual problem (that is, a functions $\zeta^*(\cdot)$ satisfying (5.2.4)) exists and consider the problem in the right hand side of (5.2.1) with $p = \nabla\zeta^*(z)$ rewriting it in the form

$$\min_{\mu\in W(z)} \left\{ \int_{U\times Y} \Big[q(u, y, z) + (\nabla\zeta^*(z))^T g(u, y, z)\Big]\mu(du, dz)\right\} = \tilde{H}(\nabla\zeta^*(z), z). \tag{5.2.6}$$

The latter is an IDLP problem, with the dual of it having the form

$$\sup_{\eta(\cdot)\in C^1(\mathbb{R}^m)} \Big\{d : d \leq q(u, y, z) + (\nabla\zeta^*(z))^T g(u, y, z) + (\nabla\eta(y))^T f(u, y, z) \tag{5.2.7}$$
$$\forall(u, y) \in U \times Y\Big\} \stackrel{\text{def}}{=} D^*(z),$$

where *sup* is sought over all continuously differentiable functions $\eta(\cdot) : \mathbb{R}^m \to \mathbb{R}^1$. The optimal values of the problems (5.2.6) and (5.2.7) are equal. That is, $\tilde{H}(\nabla\zeta^*(z), z) = D^*(z)$, this being one of the duality relationships between these two problems (see Theorem 1.3.1). The problem (5.2.7) will be referred to as *associated dual problem.* A function $\eta_z^*(\cdot) \in C^1(\mathbb{R}^m)$ will be called a *solution of the problem* (5.2.7) if $\forall(u, y) \in U \times Y$

$$\tilde{H}(\nabla\zeta^*(z), z) \leq q(u, y, z) + (\nabla\zeta^*(z))^T g(u, y, z) + (\nabla\eta_z^*(y))^T f(u, y, z). \tag{5.2.8}$$

Note that from (5.2.4) and from (5.2.8) it follows that $\forall(u, y, z) \in U \times Y \times Z$

$$q(u, y, z) + (\nabla\zeta^*(z))^T g(u, y, z) + (\nabla\eta_z^*(y))^T f(u, y, z) \geq \tilde{G}^*. \tag{5.2.9}$$

The following result gives sufficient and also (under additional periodicity assumptions) necessary condition for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (4.2.28) to be valid.

Proposition 5.2.1 *Let a solution $\zeta^*(z)$ of the averaged dual problem exists and a solution $\eta_z^*(y)$ of the associated dual problem exists for any $z \in Z$. Then an ACG family $(u_z(\cdot), y_z(\cdot))$ generating the admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ is optimal and the equality (4.2.28) is valid if*

$$q(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))$$

$$+(\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) = \tilde{G}^* \qquad \forall\ \tau \in P_t, \quad \forall t \in A, \tag{5.2.10}$$

for some $P_t \subset \mathbb{R}^1$ and $A \subset \mathbb{R}^1$ such that

$$meas\{\mathbb{R}^1 \setminus P_t\} = 0 \quad \forall t \in A \quad and \qquad meas\{\mathbb{R}^1 \setminus A\} = 0. \tag{5.2.11}$$

Under the additional assumption that an ACG family $(u_z(\cdot), y_z(\cdot))$ is periodic, that is,

$$(u_z(\tau), y_z(\tau)) = (u_z(\tau + T_z), y_z(\tau + T_z)) \qquad \forall \tau \geq 0 \tag{5.2.12}$$

for some $T_z > 0$ and that the admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ generated by this family is periodic as well, that is,

$$(\mu(t), z(t)) = (\mu(t + \tilde{T}), z(t + \tilde{T})) \qquad \forall t \geq 0 \tag{5.2.13}$$

for some $\tilde{T} > 0$, the fulfillment of (5.2.10) is also necessary for $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (4.2.28) to be valid.

Proof. Assume (5.2.10) is true. Then

$$\lim_{S\to\infty} \frac{1}{S} \int_0^S [q(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))$$

$$+(\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))]d\tau$$

$$= \tilde{q}(\mu(t), z(t)) + (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t)) = \tilde{G}^* \quad \forall t \in A, \tag{5.2.14}$$

where it has been taken into account that

$$\lim_{S\to\infty} \frac{1}{S} \int_0^S (\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))d\tau$$

$$= \lim_{S\to\infty} \frac{1}{S}[\eta^*_{z(t)}(y_{z(t)}(S)) - \eta^*_{z(t)}(y_{z(t)}(0))] = 0.$$

Since

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t))dt = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}}[\zeta^*(z(\mathcal{T})) - \zeta^*(z(0))] = 0,$$

from (5.2.14) it follows that

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{q}(\mu(t), z(t))dt = \tilde{G}^*.$$

By (4.2.27), the latter implies that $(u_z(\cdot), y_z(\cdot))$ is optimal and that the equality (4.2.28) is valid.

Let us now prove (assuming that (5.2.12) and (5.2.13) are true) that the fulfillment of (5.2.10) is necessary for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (4.2.28) to be valid. In fact, let an ACG family $(u_z(\cdot), y_z(\cdot))$ be optimal and let (4.2.28) be true. Then

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} \tilde{q}(\mu(t), z(t))dt = \tilde{G}^*.$$

Since (by (5.2.13))

$$\int_0^{\tilde{T}} (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t))dt = \zeta^*(z(\tilde{T})) - \zeta^*(z(0)) = 0,$$

it follows that

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} [\tilde{q}(\mu(t), z(t)) + (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t)) - \tilde{G}^*]dt = 0$$

and, hence, by (5.2.5),

$$\tilde{q}(\mu(t), z(t)) + (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t)) = \tilde{G}^* \tag{5.2.15}$$

for almost all $t \in [0, \tilde{T}]$. Note that (due to periodicity condition (5.2.13)) the equality above is also valid for almost all $t \in [0, \infty]$.

Let the set A ($meas\{I\!R^1 \setminus A\} = 0$) be such that the equality (5.2.15) is valid and let $t \in A$. Due to the periodicity condition (5.2.12), to prove the required statement it is sufficient to show that the equality (5.2.10) is satisfied for almost all $\tau \in [0, T_{z(t)}]$.

Assume it is not the case and there exists a set $Q_t \subset [0, T_{z(t)}]$, with $meas\{Q_t\} > 0$, on which (5.2.10) is not satisfied, the latter implying (due to (5.2.9)) that

$$q(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))$$

$$+(\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \;>\; \tilde{G}^* \quad \forall\tau \in Q_t.$$

From the above inequality and from (5.2.9) it follows that

$$\frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} \Big[q(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))$$

$$+(\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))\Big] d\tau \;>\; \tilde{G}^*. \tag{5.2.16}$$

By (5.2.12),

$$\int_0^{T_{z(t)}} (\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) d\tau$$

$$= \eta^*_{z(t)}(y_{z(t)}(T_{z(t)})) - \eta^*_{z(t)}(y_{z(t)}(0)) = 0. \tag{5.2.17}$$

Hence, from (5.2.16) it follows that

$$\frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} \Big[q(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))\Big] d\tau \;>\; \tilde{G}^*,$$

which is equivalent to

$$\tilde{q}(\mu(t), z(t)) + (\nabla\zeta^*(z(t)))^T \tilde{g}(\mu(t), z(t)) \;>\; \tilde{G}^*.$$

This contradicts to the fact that t was chosen to belong to the set A on which (5.2.15) is satisfied. This completes the proof of the proposition. □

Remark 5.2.2 *Note that, due to (5.2.9), the validity of (5.2.10) implies the validity of the inclusion*

$$(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \in Argmin_{(u,y,z)\in U\times Y\times Z}\Big\{q(u, y, z) + (\nabla\zeta^*(z))^T g(u, y, z)$$

$$+(\nabla\eta^*_z(y))^T f(u, y, z)\Big\} \qquad \forall\tau \in P_t, \quad \forall t \in A \tag{5.2.18}$$

which, in turn, implies

$$u_{z(t)}(\tau) \in Argmin_{u\in U}\Big\{q(u, y_{z(t)}(\tau), z(t)) + (\nabla\zeta^*(z(t)))^T g(u, y_{z(t)}(\tau), z(t))$$

$$+(\nabla\eta^*_{z(t)}(y_{z(t)}(\tau)))^T f(u,y_{z(t)}(\tau),z(t))\Big\} \quad \forall\tau\in P_t, \quad \forall t\in A. \tag{5.2.19}$$

That is, if the equality (4.2.28) is valid, then for an ACG family $(u_z(\cdot),y_z(\cdot))$ satisfying the periodicity conditions (5.2.12) and (5.2.13) to be optimal, it is necessary that the inclusion (5.2.19) is satisfied.

5.3 Approximating averaged IDLP problem and approximating averaged/associated dual problems.

Let $\psi_i(\cdot)\in C^1(\mathbb{R}^n)$, $i=1,2,...$, be a sequence of functions such that any $\zeta(\cdot)\in C^1(\mathbb{R}^n)$ and its gradient are simultaneously approximated by a linear combination of $\psi_i(\cdot)$ and their gradients. Also, let $\phi_i(\cdot)\in C^1(\mathbb{R}^m)$, $i=1,2,...$, be a sequence of functions such that any $\eta(\cdot)\in C^1(\mathbb{R}^m)$ and its gradient are simultaneously approximated by a linear combination of $\phi_i(\cdot)$ and their gradients. Examples of such sequences are monomials $z_1^{i_1}...z_n^{i_n}$, $i_1,...,i_n=0,1,...$ and, respectively, $y_1^{i_1}...y_m^{i_m}$, $i_1,...,i_m=0,1,...$, with z_k, and y_l standing for the components of z and y (see, e.g., [84]).

Let us introduce the following notations:

$$W_M(z)\stackrel{\text{def}}{=}\{\mu\in\mathcal{P}(U\times Y)\ :\ \int_{U\times Y}(\nabla\phi_i(y))^T f(u,y,z)\mu(du,dy)=0, \quad i=1,...,M\}, \tag{5.3.1}$$

$$F_M\stackrel{\text{def}}{=}\{(\mu,z)\ :\ \mu\in W_M(z), \quad z\in Z\}\ \subset\mathcal{P}(U\times Y)\times Z\ , \tag{5.3.2}$$

and

$$\tilde{W}_{N,M}\stackrel{\text{def}}{=}\{\xi\in\mathcal{P}(F_M):\int_{F_M}(\nabla\psi_i(z))^T\tilde{g}(\mu,z)\xi(d\mu,dz)=0, \quad i=1,...,N\}, \tag{5.3.3}$$

(compare with (4.2.3), (4.2.25) and (4.2.26), respectively) and let us consider the following SILP problem (compare with (4.2.24))

$$\min_{\xi\in\tilde{W}_{N,M}}\int_{F_M}\tilde{q}(\mu,z)\xi(d\mu,dz)\stackrel{\text{def}}{=}\tilde{G}^{N,M}. \tag{5.3.4}$$

This problem will be referred to as (N,M)-*approximating averaged problem.* It is obvious that

$$W_1(z) \supset W_2(z) \supset ... \supset W_M(z) \supset ... \supset W(z)$$

$$\Rightarrow \quad F_1 \supset F_2 \supset ... \supset F_M \supset ... \supset F. \tag{5.3.5}$$

Defining the set $\tilde{W}_N$ by the equation

$$\tilde{W}_N \stackrel{\text{def}}{=} \{\xi \in \mathcal{P}(F) : \int_F (\nabla\psi_i(z))^T \tilde{g}(\mu, z)\xi(d\mu, dz) = 0, \quad i = 1, ..., N\}, \tag{5.3.6}$$

one can also see that

$$\tilde{W}_{N,M} \supset \tilde{W}_N \supset \tilde{W} \qquad \forall\ N, M = 1, 2, ... \tag{5.3.7}$$

(with $\tilde{W}_{N,M}$, $\tilde{W}_N$ and $\tilde{W}$ being considered as subsets of $\mathcal{P}(\mathcal{P}(U \times Y) \times Z)$), the latter implying, in particular, that

$$\tilde{G}^{N,M} \leq \tilde{G}^* \quad \forall\ N, M = 1, 2, ... \tag{5.3.8}$$

It can be readily verified that (see the proof of Proposition 2.3.1 above) that

$$\lim_{M\to\infty} W_M(z) = W(z), \qquad \lim_{M\to\infty} F_M = F, \tag{5.3.9}$$

where, in the first case, the convergence is in the Hausdorff metric generated by the weak* convergence in $\mathcal{P}(U \times Y)$ and, in the second, it is in the Hausdorff metric generated by the weak* convergence in $\mathcal{P}(U \times Y)$ and the convergence in Z.

Proposition 5.3.1 *The following relationships are valid:*

$$\lim_{M\to\infty} \tilde{W}_{N,M} = \tilde{W}_N, \tag{5.3.10}$$

$$\lim_{N\to\infty} \tilde{W}_N = \tilde{W}, \tag{5.3.11}$$

where the convergence in both cases is in Hausdorff metric generated by the weak* convergence in $\mathcal{P}(\mathcal{P}(U \times Y) \times Z)$. Also,

$$\lim_{N\to\infty} \lim_{M\to\infty} \tilde{G}^{N,M} = \tilde{G}^*. \tag{5.3.12}$$

If the optimal solution ξ^* of the averaged IDLP problem (4.2.24) is unique, then, for *an arbitrary optimal solution* $\xi^{N,M}$ *of the* (N, M)*-approximating problem (5.3.4),*

$$\lim_{N\to\infty} \limsup_{M\to\infty} \rho(\xi^{N,M}, \xi^*) = 0. \tag{5.3.13}$$

Proof. In order to prove (5.3.10), due to (5.3.7), it is sufficient to show that, if $\xi^{M_l} \in W_{N,M_l}$ and if $\lim\limits_{M_l\to\infty} \xi^{M_l} = \xi$, then $\xi \in W_N$. By (5.3.9), for any $\delta > 0$,

$$F_M \in F + \delta\bar{B}$$

if M is large enough, where $\bar{B}$ is the closed unit ball in $\mathcal{P}(U \times Y) \times Z$. Hence, from the fact that $supp(\xi^{M_l}) \in F_{M_l}$ it follows that

$$\xi^{M_l}(F + \delta\bar{B}) = 1$$

for M_l large enough. Since

$$\overline{\lim}_{M_l\to\infty} \xi^{M_l}(F + \delta\bar{B}) \le \xi(F + \delta\bar{B}),$$

one obtains

$$\xi(F + \delta\bar{B}) = 1,$$

the latter being valid for any $\delta > 0$. This implies the equality

$$\xi(F) = 1 \qquad \Rightarrow \qquad supp(\xi) \in F. \tag{5.3.14}$$

From the fact that $\xi^{M_l} \in W_{N,M_l}$ it follows that

$$0 = \int_{F_{M_l}} (\nabla\psi_i(z))^T \tilde{g}(\mu, z)\xi^{M_l}(d\mu, dz) =$$

$$\int_{F+\delta\bar{B}} (\nabla\psi_i(z))^T \tilde{g}(\mu, z)\xi^{M_l}(d\mu, dz), \qquad i = 1, ..., N. \tag{5.3.15}$$

Passing to the limit when $M_l \to \infty$, one obtains

$$\int_{F+\delta\bar{B}} (\nabla\psi_i(z))^T \tilde{g}(\mu, z)\xi(d\mu, dz) = 0, \qquad i = 1, ..., N.$$

Due to (5.3.14), the expression above implies that

$$\int_{F} (\nabla\psi_i(z))^T \tilde{g}(\mu, z)\xi(d\mu, dz) = 0, \qquad i = 1, ..., N,$$

which, in turn, implies that $\xi \in W_N$.

The proof of (5.3.11) is straightforward (it is analogous to the proof of Proposition 2.3.1). From (5.3.10) it follows that

$$\lim_{M\to\infty} \tilde{G}^{N,M} = \min_{\xi\in\tilde{W}_N} \int_F \tilde{q}(\mu,z)\xi(d\mu,dz) \stackrel{\text{def}}{=} \tilde{G}^N, \tag{5.3.16}$$

and from (5.3.11) it follows that

$$\lim_{N\to\infty} \tilde{G}^N = \tilde{G}^*. \tag{5.3.17}$$

The above two relationships imply (5.3.12). If the optimal solution ξ^* of the averaged IDLP problem (4.2.24) is unique, then, by (5.3.17), for any solution ξ^N of the problem in the right-hand side of (5.3.16) there exists the limit

$$\lim_{N\to\infty} \xi^N = \xi^*. \tag{5.3.18}$$

Also, if for an arbitrary optimal solution $\xi^{N,M}$ of the (N,M) approximating problem (5.3.4) and for some $M' \to \infty$, there exists $\lim\limits_{M'\to\infty} \xi^{N,M'}$, then this limit is an optimal solution of the problem in the right-hand side of (5.3.16). This proves (5.3.13). $\square$

Define the finite dimensional space $\tilde{\Omega}_N \subset C^1(I\!R^n)$ by the equation

$$\tilde{\Omega}_N \stackrel{\text{def}}{=} \{\zeta(\cdot) \in C^1(I\!R^n) : \zeta(z) = \sum_{i=1}^{N} \lambda_i \psi_i(z), \quad \lambda = (\lambda_i) \in I\!R^N\} \tag{5.3.19}$$

and consider the following problem

$$\sup_{\zeta(\cdot)\in\tilde{\Omega}_N} \{\tilde{d} : \tilde{d} \le \tilde{q}(\mu,z) + (\nabla\zeta(z))^T \tilde{g}(\mu,z) \quad \forall(\mu,z) \in F_M\} \stackrel{\text{def}}{=} \tilde{D}^{N,M}. \tag{5.3.20}$$

This problem is dual with respect to the problem (5.3.4), its optimal value is equal to the optimal value of the later. That is,

$$\tilde{D}^{N,M} = \tilde{G}^{N,M}. \tag{5.3.21}$$

Note that the problem (5.3.20) looks similar to the averaged dual problem (5.2.3). However, in contrast to the latter, *sup* in (5.3.20) is sought over the finite dimensional subspace $\tilde{\Omega}_N$ of $C^1(I\!R^n)$ and F_M is used instead of F. The problem (5.3.20) will be referred to as (N,M)-*approximating averaged dual problem.* A function $\zeta^{N,M}(\cdot) \in \tilde{\Omega}_N$,

$$\zeta^{N,M}(z) = \sum_{i=1}^{N} \lambda_i^{N,M} \psi_i(z), \tag{5.3.22}$$

will be called a *solution of the* (N,M)-*approximating averaged dual problem* if

$$\tilde{G}^{N,M} \leq \tilde{q}(\mu, z) + (\nabla \zeta^{N,M}(z))^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M. \tag{5.3.23}$$

Define the finite dimensional space $\Omega_M \subset C^1(I\!R^m)$ by the equation

$$\Omega_M \stackrel{\text{def}}{=} \{\eta(\cdot) \in C^1(I\!R^m) : \eta(y) = \sum_{i=1}^{M} \alpha_i \phi_i(y), \quad \alpha = (\alpha_i) \in I\!R^M\} \tag{5.3.24}$$

and, assuming that a solution $\zeta^{N,M}(z)$ of the (N,M)-approximating averaged dual problem exists, consider the following problem

$$\sup_{\eta(\cdot) \in \Omega_M} \{d : d \leq q(u,y,z) + (\nabla \zeta^{N,M}(z))^T g(u,y,z) + (\nabla \eta(y))^T f(u,y,z)$$

$$\forall (u,y) \in U \times Y\} \stackrel{\text{def}}{=} D^{N,M}(z). \tag{5.3.25}$$

While the problem (5.3.25) looks similar to the associated dual problem (5.2.7), it differs from the latter, firstly, by that *sup* is sought over the finite dimensional subspace Ω_M of $C^1(I\!R^m)$ and, secondly, by that a solution $\zeta^{N,M}(z)$ of (5.3.20) is used instead of a solution $\zeta^*(z)$ of (5.2.2) (the later may not exist). The problem (5.3.25) will be referred to as (N,M)-*approximating associated dual problem.* It can be shown that it is, indeed, dual with respect to the SILP problem

$$\min_{\mu \in W_M(z)} \{\int_{U \times Y} [q(u,y,z) + (\nabla \zeta^{N,M}(z))^T g(u,y,z)] \mu(du,dy)\} = D^{N,M}(z), \tag{5.3.26}$$

the duality relationships including the equality of the optimal values (see Theorem 1.3.1 and also Theorem 5.2(ii) in [48]). A function $\eta_z^{NM}(\cdot) \in \Omega_M$,

$$\eta_z^{N,M}(y) = \sum_{i=1}^{M} \alpha_{z,i}^{N,M} \phi_i(y), \tag{5.3.27}$$

will be called a *solution of the* (N,M)-*approximating associated dual problem* if

$\forall (u, y) \in U \times Y$

$$D^{N,M}(z) \leq q(u, y, z) + (\nabla \zeta^{N,M}(z))^T g(u, y, z) + (\nabla \eta_z^{NM}(y))^T f(u, y, z). \qquad (5.3.28)$$

In the next section, we show that solutions of the (N, M)-approximating averaged and associated dual problems exist under natural local controllability conditions.

5.4 Controllability conditions sufficient for the existence of solutions of approximating averaged and associated dual problems.

In what follows it is assumed that, for any $N = 1, 2, ...,$ and $M = 1, 2, ...,$ the gradients $\nabla \psi_i(z),\ i = 1, 2, ...N$, and $\nabla \phi_i(y),\ i = 1, 2, ...M$, are linearly independent on any open subset of $I\!R^N$ and, respectively, $I\!R^M$. That is, if $\mathcal{Z}$ is an open subset of $I\!R^N$, then the equality

$$\sum_{i=1}^{N} v_i \nabla \psi_i(z) = 0 \quad \forall z \in \mathcal{Z}$$

is valid only if $v_i = 0,\ i = 1, ..., N$. Similarly, if $\mathcal{Y}$ is an open subset of $I\!R^M$, then the equality

$$\sum_{i=1}^{M} \bar{v}_i \nabla \phi_i(y) = 0 \ \forall y \in \mathcal{Y}$$

is valid only if $\bar{v}_i = 0,\ i = 1, ..., M$.

The existence of a solution of the approximating averaged dual problem can be guaranteed under the following controllability type assumption about the averaged system.

Assumption 5.4.1 *Let the averaged system be locally approximately controllable on Z. That is, there exists a set $Z^0 \subset Z$, such that any two points in Z^0 can be connected by an admissible trajectory of the averaged system (that is, for any $z', z'' \in Z^0$), there exists an admissible pair $(\mu(\cdot), z(\cdot))$ of the averaged system defined on some interval $[0, \mathcal{T}]$ such that $z(0) = z'$ and $z(\mathcal{T}) = z''$) and such that the closure of Z^0 has a nonempty interior ($int(cl\mathcal{Z}^0) \neq \emptyset$).*

Proposition 5.4.2 *If Assumption 5.4.1 is satisfied, then a solution of the (N, M)-approximating averaged dual problem exists for any N and M.*

Proof. The proof is given at the end of this section (its idea being similar to that of the proof of Proposition 3.2 in [63]). $\square$

The existence of a solution of the approximating associated dual problem is guaranteed by the following assumption about controllability properties of the associated system.

Assumption 5.4.3 *Let the associated system be locally approximately controllable on Y. That is, there exists a set $Y^0(z) \subset Y$ such that any two points in $Y^0(z)$ can be connected by an admissible trajectory of the associated system (that is, for any $y', y'' \in Y^0(z)$, there exists an admissible pair $(u(\cdot), y(\cdot))$ of the associated system defined on some interval $[0, S]$ such that $y(0) = y'$ and $y(S) = y''$) and such that the closure of $Y^0(z)$ has a nonempty interior ($int(clY^0(z)) \neq \emptyset$).*

Proposition 5.4.4 *If Assumption 5.4.3 is satisfied for any $z \in Z$, then a solution of the (N, M)-approximating associated dual problem exists for any N and M, and for any $z \in Z$.*

Proof. The proof is given at the end of this section. $\square$

The proofs of Propositions 5.4.2 and 5.4.4 are based on the following lemma.

Lemma 5.4.5 *Let X be a compact metric space and let $h_i(\cdot) : X \to I\!R^1,\ i = 0, 1, ..., K$, be continuous functional on X. Let*

$$\bar{D}^* \stackrel{\text{def}}{=} \sup_{\{\lambda_i\}}\{\theta \ : \ \theta \leq h_0(x) + \sum_{i=1}^{K} \lambda_i h_i(x) \ \forall x \in X\}, \tag{5.4.1}$$

where sup is sought over $\lambda \stackrel{\text{def}}{=} \{\lambda_i\} \in I\!R^K$. A solution of the problem (5.4.1), that is $\lambda^* \stackrel{\text{def}}{=} \{\lambda_i^*\} \in I\!R^K$ such that

$$\bar{D}^* \leq h_0(x) + \sum_{i=1}^{K} \lambda_i^* h_i(x) \quad \forall x \in X \tag{5.4.2}$$

exists if the inequality

$$0 \leq \sum_{i=1}^{K} v_i h_i(x) \ \forall x \in X \tag{5.4.3}$$

is valid only with $v_i = 0,\ i = 1, ..., K$.

Proof. Assume that the inequality (5.4.3) implies that $v_i = 0,\ i = 1, ..., K$. Note that from this assumption it immediately follows that $\bar{D}^*$ is bounded (since, otherwise, (5.4.1) would imply that there exist $\{\lambda_i\}$ such that $\sum_{i=1}^{K} \lambda_i h_i(x) > 0 \ \forall x \in X$). For any $k = 1, 2, ...$, let $\lambda^k = (\lambda_i^k) \in I\!R^K$ be such that

$$\bar{D}^* - \frac{1}{k} \leq h_0(x) + \sum_{i=1}^{K} \lambda_i^k h_i(x) \quad \forall x \in X \tag{5.4.4}$$

Let us show that the sequence λ^k , $k = 1, 2, ...,$ is bounded. That is, there exists $\beta > 0$ such that

$$||\lambda^k|| \leq \beta \ , \qquad k = 1, 2, ... \ . \tag{5.4.5}$$

Assume that the sequence λ^k , $k = 1, 2, ...,$ is not bounded. Then there exists a subsequence $\lambda^{k'}$ such that

$$\lim_{k' \to \infty} ||\lambda^{k'}|| = \infty \ , \qquad \lim_{k' \to \infty} \frac{\lambda^{k'}}{||\lambda^{k'}||} \stackrel{\text{def}}{=} v, \qquad ||v|| = 1 \ . \tag{5.4.6}$$

Dividing (5.4.4) by $||\lambda^k||$ and passing to the limit along the subsequence $\{k'\}$, one can obtain that

$$0 \leq \sum_{i=1}^{K} v_i^k h_i(x) \quad \forall x \in X,$$

which, by our assumption, implies that $v = (v_i) = 0$. The latter contradicts (5.4.6). Thus, the validity of (5.4.5) is established. Due to (5.4.5), there exists a subsequence $\{k'\}$ such that there exists a limit

$$\lim_{k' \to \infty} \lambda^{k'} \stackrel{\text{def}}{=} \lambda^* \ . \tag{5.4.7}$$

Passing to the limit in (5.4.4) along this subsequence, one proves (5.4.2). $\Box$

Proof of Proposition 5.4.2. By Lemma 5.4.5, to prove the proposition, it is sufficient to show that, under Assumption 5.4.1, the inequality

$$0 \leq \sum_{i=1}^{N} v_i (\nabla \psi_i(z))^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M \tag{5.4.8}$$

can be valid only with $v_i = 0, \ i = 1, ..., N$. Let us assume that (5.4.8) is valid and let us rewrite it in the form

$$0 \leq (\nabla \psi(z))^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M, \qquad \text{where} \qquad \psi(z) \stackrel{\text{def}}{=} \sum_{i=1}^{N} v_i \psi_i(z). \tag{5.4.9}$$

Let $z', z'' \in Z^0(z)$ and let an admissible pair $(\mu(\cdot), z(\cdot))$ of the associated system be such that $z(0) = z'$ and $z(\mathcal{T}) = z''$ for some $\mathcal{T} > 0$. Then, by (5.4.9),

$$\psi(z'') - \psi(z') = \int_0^{\mathcal{T}} (\nabla\psi(z(t)))^T \tilde{g}(\mu(t), z(t))dt \geq 0 \qquad \Rightarrow \qquad \psi(z'') \geq \psi(z').$$

Since z', z'' can be arbitrary points in Z^0, it follows that

$$\psi(z) = \text{const} \ \ \forall z \in Z^0 \qquad \Rightarrow \qquad \psi(z) = \text{const} \ \ \forall z \in clZ^0.$$

The latter implies that

$$\nabla\psi(z) = \sum_{i=1}^{N} v_i \nabla\psi_i(z) = 0 \quad \forall z \in int(clZ^0),$$

which, in turn, implies that $v_i = 0, \ i = 1, ..., N$ (due to linear independence of $\nabla\psi_i(\cdot)$).
□

Proof of Proposition 5.4.4. By Lemma 5.4.5, to prove the proposition, it is sufficient to show that, under Assumption 5.4.3, the inequality

$$0 \leq \sum_{i=1}^{M} \bar{v}_i [\nabla\phi_i(y)^T f(u, y, z)] \ \ \forall (u, y) \in U \times Y \tag{5.4.10}$$

can be valid only with $\bar{v}_i = 0, \ i = 1, ..., M$ (remind that z = constant). Let us assume that (5.4.10) is valid and let us rewrite it in the form

$$0 \leq ((\nabla\phi(y)))^T f(u, y, z) \ \ \forall (u, y) \in U \times Y, \qquad \text{where} \qquad \phi(y) \stackrel{\text{def}}{=} \sum_{i=1}^{M} \bar{v}_i \phi_i(y). \tag{5.4.11}$$

Let $y', y'' \in Y^0(z)$ and let an admissible pair $(u(\cdot), y(\cdot))$ of the associated system be such that $y(0) = y'$ and $y(S) = y''$ for some $S > 0$. Then, by (5.4.11),

$$\phi(y'') - \phi(y') = \int_0^S (\nabla\phi(y(\tau)))^T f(u(\tau), y(\tau), z)d\tau \geq 0 \qquad \Rightarrow \qquad \phi(y'') \geq \phi(y').$$

Since y', y'' can be arbitrary points in $Y^0(z)$, it follows that

$$\phi(y) = \text{const} \ \ \forall y \in Y^0(z) \qquad \Rightarrow \qquad \phi(y) = \text{const} \ \ \forall y \in clY^0(z).$$

The latter implies that

$$\nabla\phi(y) = \sum_{i=1}^{M} \bar{v}_i \nabla\phi_i(y) = 0 \qquad \forall y \in int(clY^0(z)),$$

which, in turn, implies that $\bar{v}_i = 0,\ i = 1, ..., M$ (due to linear independence of $\nabla\phi_i(\cdot)$). $\square$

Remark 5.4.6 *Note that the proof of Propositions 5.4.2 and 5.4.4 is similar to the proof of Proposition 2.2.8.*

5.5 Construction of near optimal ACG families.

Let us assume that, for any N and M, a solution $\zeta^{N,M}(z)$ of the (N, M)-approximating averaged dual problem exists and a solution $\eta_z^{N,M}(y)$ of the (N, M)-approximating associated dual problem exists for any $z \in Z$ (as follows from Propositions 5.4.2 and 5.4.4 these exist if Assumptions 5.4.1 and 5.4.3 are satisfied).

Define a control $u^{N,M}(y, z)$ as an optimal solution of the problem

$$\min_{u\in U}\{q(u, y, z) + (\nabla\zeta^{N,M}(z))^T g(u, y, z) + (\nabla\eta_z^{N,M}(y))^T f(u, y, z)\}. \tag{5.5.1}$$

That is,

$$\begin{aligned} u^{N,M}(y, z) = argmin_{u\in U}\{q(u, y, z) + (\nabla\zeta^{N,M}(z))^T g(u, y, z) \\ +(\nabla\eta_z^{N,M}(y))^T f(u, y, z)\}. \end{aligned} \tag{5.5.2}$$

Assume that the system

$$y_z'(\tau) = f(u^{N,M}(y_z(\tau), z), y_z(\tau), z), \qquad y_z(0) = y \in Y, \tag{5.5.3}$$

has a unique solution $y_z^{N,M}(\tau) \in Y$. Below, we introduce assumptions under which it will be established that $(u_z^{N,M}(\cdot), y_z^{N,M}(\cdot))$, where $u_z^{N,M}(\tau) \stackrel{\text{def}}{=} u_z^{N,M}(y_z^{N,M}(\tau), z)$, is a near optimal ACG family (see Theorem 5.5.8).

Assumption 5.5.1 *The following conditions are satisfied:*

(i) the optimal solution ξ^* of the IDLP problem (4.2.24) is unique, and the equality (4.2.28) is valid;

(ii) the optimal solution of the averaged problem (4.2.16) (that is, an admissible pair $(\mu^*(\cdot), z^*(\cdot))$ that delivers minimum in (4.2.16)) exists and, for any continuous function $\tilde{h}(\mu, z) : F \to I\!R^1$, there exists a limit

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu^*(t), z^*(t))dt; \tag{5.5.4}$$

(iii) for almost all $t \in [0, \infty)$ and any $r > 0$, the ξ^* -measure of the set

$$\mathcal{B}_r(\mu^*(t), z^*(t)) \stackrel{\text{def}}{=} \{(\mu, z) : \ \rho(\mu, \mu^*(t)) + ||z - z^*(t)|| < r\}$$

is not zero. That is,

$$\xi^*(\mathcal{B}_r(\mu^*(t), z^*(t))) > 0. \tag{5.5.5}$$

Note that from Assumption 5.5.1 (ii) it follows that the pair $(\mu^*(\cdot), z^*(\cdot))$ generates an occupational measure and from Assumption 5.5.1 (i) it follows that this measure coincides with ξ^* (see Corollary 5.1.4). That is,

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu^*(t), z^*(t))dt = \int_F \tilde{h}(\mu, z)\xi^*(d\mu, dz). \tag{5.5.6}$$

The following statement gives sufficient conditions for the validity of Assumption 5.5.1 (iii).

Proposition 5.5.2 *Let Assumptions 5.5.1 (i) and 5.5.1 (ii) be satisfied. Then Assumption 5.5.1 (iii) will be satisfied if the pair $(\mu^*(\cdot), z^*(\cdot))$ is $\tilde{\mathcal{T}}$-periodic ($\tilde{\mathcal{T}}$ is some positive number) and if $\mu^*(\cdot)$ is piecewise continuous on $[0, \tilde{\mathcal{T}}]$.*

Proof. Let t be a continuous point of $\mu^*(\cdot)$. Due to the assumed periodicity of the pair $(\mu^*(\cdot), z^*(\cdot))$,

$$\frac{1}{\tilde{\mathcal{T}}} \int_0^{\tilde{\mathcal{T}}} \tilde{h}(\mu^*(t), z^*(t))dt = \int_F \tilde{h}(\mu, z)\xi^*(d\mu, dz)$$

and

$$\xi^*(\mathcal{B}_r(\mu^*(t), z^*(t))) = \frac{1}{\tilde{\mathcal{T}}} meas\{t : \ t \in [0, \tilde{\mathcal{T}}], \ (\mu^*(t), z^*(t)) \in \mathcal{B}_r(\mu^*(t), z^*(t))\}. \tag{5.5.7}$$

Since t is a continuous point of $\mu^*(\cdot)$ and since $z^*(\cdot)$ is continuous, there exists $\alpha > 0$ such that $(\mu^*(t'), z^*(t')) \in \mathcal{B}_r(\mu^*(t), z^*(t)) \ \ \forall t' \in [t - \alpha, t + \alpha]$. Hence, the right-hand-side in (5.5.7) is greater than $\frac{2\alpha}{\tilde{\mathcal{T}}}$. This proves the required statement as the number of discontinuity points of $\mu^*(\cdot)$ is finite (due to the assumed piecewise continuity). $\square$

Assumption 5.5.3 *The following conditions are satisfied:*

(i) for almost all $t \in [0,\infty)$, there exists an admissible pair $(u_t^*(\tau), y_t^*(\tau))$ of the associated system (considered with $z = z^*(t)$) such that $\mu^*(t)$ is the occupational measure generated by this pair on the interval $[0,\infty)$. That is, for any continuous $h(u,y) : U \times Y \to I\!R^1$,

$$\lim_{S\to\infty} \frac{1}{S}\int_0^S h(u_t^*(\tau), y_t^*(\tau))d\tau = \int_{U\times Y} h(u,y)\mu^*(t)(du,dy); \tag{5.5.8}$$

(ii) for almost all $t \in [0,\infty)$, for almost $\tau \in [0,\infty)$ and for any $r > 0$, the $\mu^*(t)$-measure of the set

$$B_r(u_t^*(\tau), y_t^*(\tau)) \stackrel{\text{def}}{=} \{(u,y) : \ ||u - u_t^*(\tau)|| + ||y - y_t^*(\tau)|| < r\}$$

is not zero. That is,

$$\mu^*(t)(B_r(u_t^*(\tau), y_t^*(\tau))) > 0. \tag{5.5.9}$$

The following proposition gives sufficient conditions for the validity of Assumption 5.5.3 (ii).

Proposition 5.5.4 *Let Assumption 5.5.3 (i) be valid. Then Assumption 5.5.3 (ii) will be satisfied if, for almost all $t \in [0,\infty)$, the pair $(u_t^*(\tau), y_t^*(\tau))$ is T_t-periodic (T_t is some positive number) and if $u^*(\cdot)$ is piecewise continuous on $[0, T_t]$.*

Proof. The proof is similar to that of Proposition 5.5.2. $\square$

Assumption 5.5.5 *The following conditions are satisfied:*

(i) the pair $(u_z^{N,M}(\tau), y_z^{N,M}(\tau))$, where $y_z^{N,M}(\tau)$ is the solution of (5.5.3) and $u_z^{N,M}(\tau) = u_z^{N,M}(y_z^{N,M}(\tau), z)$ is an ACG family that generates the occupational measure $\mu^{N,M}(du,dy|z)$ on the interval $[0,\infty)$, the latter being independent of the initial conditions $y_z^{N,M}(0) = y$ for y in a neighbourhood of $y_t^*(\cdot)$. Also, for any continuous $h(u,y,z) : U \times Y \times Z \to I\!R^1$,

$$\left| \frac{1}{S}\int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau), z)d\tau - \int_{U\times Y} h(u,y,z)\mu^{N,M}(du,dy|z) \right| \le \delta_h(S)$$

$$\forall z \in Z, \qquad \lim_{S\to\infty} \delta_h(S) = 0; \tag{5.5.10}$$

(ii) the admissible pair of the averaged system $(\mu^{N,M}(\cdot), z^{N,M}(\cdot))$ *generated by* $(u_z^{N,M}(\cdot),$ $y_z^{N,M}(\cdot))$ *generates the occupational measure* $\bar{\xi}^{N,M} \in \mathcal{P}(F)$*, the latter being independent of the initial conditions* $z^{N,M}(0) = z$ *for* z *in a neighbourhood of* $z^*(\cdot)$*. Also, for any continuous function* $\tilde{h}(\mu, z) : F \to I\!R^1$,

$$| \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t))dt - \int_F \tilde{h}(\mu, z)\bar{\xi}^{N,M}(d\mu, dz) \mid \leq \delta_{\tilde{h}}(\mathcal{T}),$$

$$\lim_{\mathcal{T}\to\infty} \delta_{\tilde{h}}(\mathcal{T}) = 0. \tag{5.5.11}$$

To state our next assumption, let us re-denote the occupational measure $\mu^{N,M}(du, dy|z)$ (introduced in Assumption 5.5.5 above) as $\mu^{N,M}(z)$ (that is, $\mu^{N,M}(du, dy|z) = \mu^{N,M}(z)$).

Assumption 5.5.6 *For almost all* $t \in [0, \infty)$*, there exists an open ball* $\mathcal{Z}_t \subset I\!R^n$ *centered at* $z^*(t)$ *such that:*

(i) the occupational measure $\mu^{N,M}(z)$ *is continuous on* $\mathcal{Z}_t$*. Namely, for any* $z', z'' \in \mathcal{Z}_t$,

$$\rho(\mu^{N,M}(z'), \mu^{N,M}(z'')) \leq \kappa(||z' - z''||), \tag{5.5.12}$$

where $\kappa(\theta)$ *is a function tending to zero when* θ *tends to zero* ($\lim_{\theta\to 0} \kappa(\theta) = 0$)*. Also, for any* $z', z'' \in \mathcal{Z}_t$,

$$|| \int_{U\times Y} g(u, y, z')\mu^{N,M}(z')(du, dy) - \int_{U\times Y} g(u, y, z'')\mu^{N,M}(z'')(du, dy)|| \leq L||z' - z''||, \tag{5.5.13}$$

where L *is a constant;*

(ii) let $z^{N,M}(\cdot)$ *be the solution of the system*

$$z'(t) = \tilde{g}(\mu^{N,M}(z(t)), z(t)) \ , \qquad z(0) = z_0. \tag{5.5.14}$$

We assume that, for any $t > 0$,

$$\lim_{N\to\infty} \limsup_{M\to\infty} meas\{A_t(N, M)\} = 0, \tag{5.5.15}$$

where

$$A_t(N, M) \stackrel{\text{def}}{=} \{t' \in [0, t] \ : \ z^{N,M}(t') \notin \mathcal{Z}_{t'}\} \tag{5.5.16}$$

and $meas\{\cdot\}$ *stands for the Lebesgue measure of the corresponding set.*

In addition to assumptions above, let us also introduce

Assumption 5.5.7 *For each* $t \in [0,\infty)$ *such that* $\mathcal{Z}_t \neq \emptyset$, *the following conditions are satisfied:*

(i) for almost all $\tau \in [0,\infty)$, *there exists an open ball* $\mathcal{Y}_{t,\tau} \subset I\!R^m$ *centered at* $y_t^*(\tau)$ *such* $u^{N,M}(y,z)$ *is uniquely defined (the problem (5.5.1) has a unique solution) for* $(y,z) \in \mathcal{Y}_{t,\tau} \times \mathcal{Z}_t$;

(ii) the function $u^{N,M}(y,z)$ *satisfies Lipschitz conditions on* $\mathcal{Y}_{t,\tau} \times \mathcal{Z}_t$. *That is,*

$$||u^{N,M}(y',z') - u^{N,M}(y'',z'')|| \leq L(||y'-y''|| + ||z'-z''||)$$

$$\forall (y',z'),(y'',z'') \in \mathcal{Y}_{t,\tau} \times \mathcal{Z}_t, \tag{5.5.17}$$

where L *is a constant;*

(iii) let $y_t^{N,M}(\tau) \stackrel{\text{def}}{=} y_{z^*(t)}^{N,M}(\tau)$ *be the solution of the system (5.5.3) considered with* $z = z^*(t)$ *and with the initial condition* $y_z(0) = y_t^*(0)$. *We assume that, for any* $\tau > 0$,

$$\lim_{N\to\infty} \limsup_{M\to\infty} meas\{P_{t,\tau}(N,M)\} = 0, \tag{5.5.18}$$

where

$$P_{t,\tau}(N,M) \stackrel{\text{def}}{=} \{\tau' \in [0,\tau] \ : \ y_t^{N,M}(\tau') \notin \mathcal{Y}_{t,\tau'}\}. \tag{5.5.19}$$

Theorem 5.5.8 *Let Assumptions 5.5.1, 5.5.3, 5.5.5, 5.5.6 and 5.5.7 be satisfied. Then the family* $(u_z^{N,M}(\cdot), y_z^{N,M}(\cdot))$ *introduced in Assumption 5.5.5 (i) is a* $\beta(N,M)$*- near optimal ACG family, where*

$$\lim_{N\to\infty} \limsup_{M\to\infty} \beta(N,M) = 0. \tag{5.5.20}$$

Also,

$$\lim_{N\to\infty} \limsup_{M\to\infty} \rho(\bar{\xi}^{N,M}, \xi^*) = 0, \tag{5.5.21}$$

where $\bar{\xi}^{N,M}$ *is defined by (5.5.11).*

Proof. The proof is given in Section 5.6. It is based on Lemma 5.5.9 stated at the end of this section. Note that in the process of the proof of the theorem it is established that

$$\lim_{N\to\infty} \limsup_{M\to\infty} \max_{t'\in[0,t]} ||z^{N,M}(t') - z^*(t')|| = 0 \quad \forall t \in [0,\infty), \tag{5.5.22}$$

where $z^{N,M}(\cdot)$ is the solution of (5.5.14). Also, it is shown that

$$\lim_{N\to\infty}\limsup_{M\to\infty}\rho(\mu^{N,M}(z^{N,M}(t)),\mu^*(t))=0 \tag{5.5.23}$$

for almost all $t\in[0,\infty)$, and

$$\lim_{N\to\infty}\limsup_{M\to\infty}|\tilde{V}^{N,M}-\tilde{G}^*|=0, \tag{5.5.24}$$

where

$$\tilde{V}^{N,M}\stackrel{\text{def}}{=}\lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}}\tilde{q}(\mu^{N,M}(z^{N,M}(t)),z^{N,M}(t))dt. \tag{5.5.25}$$

The relationship (5.5.24) implies the statement of the theorem with

$$\beta(N,M)\stackrel{\text{def}}{=}\tilde{V}^{N,M}-\tilde{G}^* \tag{5.5.26}$$

(see Definition 5.1.3).

Lemma 5.5.9 *Let the assumptions of Theorem 5.5.8 be satisfied and let $t\in[0,\infty)$ be such that $\mathcal{Z}_t\neq\emptyset$. Then,*

$$\lim_{N\to\infty}\limsup_{M\to\infty}\max_{\tau'\in[0,\tau]}|y_t^{N,M}(\tau')-y_t^*(\tau')||=0\quad\forall\tau\in[0,\infty). \tag{5.5.27}$$

Also,

$$\lim_{N\to\infty}\limsup_{M\to\infty}||u^{N,M}(y_t^{N,M}(\tau),z^*(t))-u_t^*(\tau)||=0 \tag{5.5.28}$$

for almost all $\tau\in[0,\infty)$.

Proof. The proof is given in Section 5.6.□

Remark 5.5.10 *Note that from (5.3.13) and (5.5.21) it follows that*

$$\lim_{N\to\infty}\limsup_{M\to\infty}\rho(\bar{\xi}^{N,M},\xi^{N,M})=0, \tag{5.5.29}$$

where $\xi^{N,M}$ is an arbitrary optimal solution of the (N,M)-approximating averaged problem (5.3.4).

5.6 Proof of Theorem 5.5.8

Note, first of all, that there exists an optimal solution $\xi^{N,M}$ of the problem (5.3.4) which is presented as a convex combination of (no more than $N+1$) Dirac measures (see, e.g., Theorems A.4 and A.5 in [95]). That is,

$$\xi^{N,M} = \sum_{k=1}^{K^{N,M}} \xi_k^{N,M} \delta_{(\mu_k^{N,M}, z_k^{N,M})}, \tag{5.6.1}$$

where $\delta_{(\mu_k^{N,M}, z_k^{N,M})}$ is the Dirac measure concentrated at $(\mu_k^{N,M}, z_k^{N,M})$ and

$$(\mu_k^{N,M}, z_k^{N,M}) \in F_M, \quad \xi_k^{N,M} > 0, \ \ k = 1, ..., K^{N,M} \leq N+1; \qquad \sum_{k=1}^{K^{N,M}} \xi_k^{N,M} = 1. \tag{5.6.2}$$

Lemma 5.6.1 *For any $k = 1, ..., K^{N,M}$,*

$$\mu_k^{N,M} = \operatorname{argmin}_{\mu \in W_M(z_k^{N,M})} \{\tilde{q}(\mu, z_k^{N,M}) + (\nabla \zeta^{N,M}(z_k^{N,M}))^T \tilde{g}(\mu, z_k^{N,M})\}. \tag{5.6.3}$$

That is, $\mu_k^{N,M}$ is a minimizer of the problem

$$\min_{\mu \in W_M(z_k^{N,M})} \{\tilde{q}(\mu, z_k^{N,M}) + (\nabla \zeta^{N,M}(z_k^{N,M}))^T \tilde{g}(\mu, z_k^{N,M}) \ \}. \tag{5.6.4}$$

Proof. From (5.3.20) and (5.3.23) it follows that

$$\tilde{G}^{N,M} = \min_{(\mu,z) \in F_M} \{\tilde{q}(\mu, z) + (\nabla \zeta^{N,M}(z))^T \tilde{g}(\mu, z) \ \}. \tag{5.6.5}$$

Also, for any $\xi \in \tilde{W}_{N,M}$,

$$\int_{F_M} \tilde{q}(\mu, z)\xi(d\mu, dz)$$
$$= \int_{F_M} [\tilde{q}(\mu, z) + (\nabla \zeta^{N,M}(z))^T \tilde{g}(\mu, z)]\xi(d\mu, dz).$$

Consequently, for $\xi = \xi^{N,M}$,

$$\tilde{G}^{N,M} = \int_{F_M} \tilde{q}(\mu, z)\xi^{N,M}(d\mu, dz)$$
$$= \int_{F_M} [\tilde{q}(\mu, z) + (\nabla \zeta^{N,M}(z))^T \tilde{g}(\mu, z)]\xi^{N,M}(d\mu, dz)$$
$$= \sum_{k=1}^{K^{N,M}} \xi_k^{N,M} [\tilde{q}(\mu_k^{N,M}, z_k^{N,M}) + (\nabla \zeta^{N,M}(z_k^{N,M}))^T \tilde{g}(\mu_k^{N,M}, z_k^{N,M})].$$

Since $(\mu_k^{N,M}, z_k^{N,M}) \in F_M$, from the equalities above and from (5.6.5) it follows that

$$\tilde{q}(\mu_k^{N,M}, z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T \tilde{g}(\mu_k^{N,M}, z_k^{N,M})$$

$$= \min_{(\mu,z)\in F_M} \{\tilde{q}(\mu, z) + (\nabla\zeta^{N,M}(z))^T \tilde{g}(\mu, z)\ \}, \quad k = 1, ..., K^{N,M}.$$

That is, for $k = 1, ..., K^{N,M}$,

$$(\mu_k^{N,M}, z_k^{N,M}) = \text{argmin}_{(\mu,z)\in F_M} \{\tilde{q}(\mu, z) + (\nabla\zeta^{N,M}(z))^T \tilde{g}(\mu, z)\ \}.$$

The latter imply (5.6.3).□

Lemma 5.6.2 *In the presentation (5.6.1) of an optimal solution $\xi^{N,M}$ of the problem (5.3.4), $\mu_k^{N,M}$ can be chosen as follows:*

$$\mu_k^{N,M} = \sum_{j=1}^{J^{N,M,k}} b_j^{N,M,k} \delta_{(u_j^{N,M,k}, y_j^{N,M,k})}, \qquad k = 1, ..., K^{N,M}, \tag{5.6.6}$$

where

$$b_j^{N,M,k} > 0, \quad j = 1, ..., J^{N,M,k}, \qquad \sum_{j=1}^{J^{N,M,k}} b_j^{N,M,k} = 1, \tag{5.6.7}$$

and

$$J^{N,M,k} \leq N + M + 2. \tag{5.6.8}$$

In (5.6.6), $\delta_{(u_j^{N,M,k}, y_j^{N,M,k})} \in \mathcal{P}(U \times Y)$ are the Dirac measures concentrated at $(u_j^{N,M,k}, y_j^{N,M,k}) \in U \times Y, \quad j = 1, ..., J^{N,M,k}$, with

$$u_j^{N,M,k} = argmin_{u\in U}\{q(u, y_j^{N,M,k}, z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u, y_j^{N,M,k}, z_k^{N,M})$$

$$+(\nabla\eta^{N,M}(y_j^{N,M,k}))^T f(u, y_j^{N,M,k}, z_k^{N,M})\}. \tag{5.6.9}$$

Proof. Assume that $\xi_k^{N,M}$, $k = 1, ..., K^{N,M}$, in (5.6.1) are fixed. Then $\mu_k^{N,M}$, $k = 1, ..., K^{N,M}$, form an optimal solution of the following problem

$$\min_{\{\mu_k\}}\{ \sum_{k=1}^{K^{N,M}} \xi_k^{N,M} \int_{U\times Y} q(u, y, z_k^{N,M})\mu_k(du, dy)\ \}, \tag{5.6.10}$$

where minimization is over $\mu_k \in \mathcal{P}(U \times Y)$, $k = 1, ..., K^{N,M}$, that satisfy the following constraints

$$\sum_{k=1}^{K^{N,M}} \xi_k^{N,M} \int_{U\times Y} (\nabla\psi_i(z_k^{N,M}))^T g(u,y,z_k^{N,M})\mu_k(du,dy) = 0, \quad i=1,...,N, \tag{5.6.11}$$

$$\int_{U\times Y} (\nabla\phi_j(y))^T \ f(u,y,z_k^{N,M})\mu_k(du,dy) = 0, \quad j=1,...,M, \qquad k=1,...,K^{N,M}. \tag{5.6.12}$$

In fact, if $\mu_k^{N,M}$, $k=1,...,K^{N,M}$ is an optimal solution of the problem (5.6.10)-(5.6.12), then $\hat{\xi}^{N,M} = \sum_{k=1}^{K^{N,M}} \xi_k^{N,M} \delta_{(\mu_k^{N,M}, z_k^{N,M})}$ will be an optimal solution of the problem (5.3.4). Let us show that the former has an optimal solution that can be presented as the sum in the right-hand side of (5.6.6). To this end, note that the problem (5.6.10)-(5.6.12) can be rewritten in the following equivalent form

$$\min_{\{\omega_0^k,\omega_i^k,\upsilon_j^k\}} \{ \sum_{k=1}^{K^{N,M}} \xi_k^{N,M}\omega_0^k\}, \tag{5.6.13}$$

where minimization is over $\omega_0^k,\omega_i^k,\upsilon_j^k$, $i=1,...,N,\ j=1,...,M,\ k=1,...,K^{N,M}$, such that

$$\sum_{k=1}^{K^{N,M}} \xi_k^{N,M}\omega_i^k = 0, \quad i=1,...,N, \tag{5.6.14}$$

$$\upsilon_j^k = 0, \quad j=1,...,M, \qquad k=1,...,K^{N,M} \tag{5.6.15}$$

and such that

$$\{\omega_0^k,\omega_i^k,\upsilon_j^k, \quad i=1,...,N, \quad j=1,...,M, \ \} \in \bar{co}V_k, \quad k=1,...,K^{N,M} \tag{5.6.16}$$

where

$$V_k = \{\omega_0^k,\omega_i^k,\upsilon_j^k \ : \ \omega_0^k = q(u,y,z_k^{N,M}), \ \ \omega_i^k = (\nabla\psi_i(z_k^{N,M}))^T g(u,y,z_k^{N,M}),$$

$$\upsilon_j^k = (\nabla\phi_j(y))^T \ f(u,y,z_k^{N,M}), \ \ i=1,...,N,\ j=1,...,M;\ (u,y)\in U\times Y\}.$$

By Caratheodory's theorem,

$$\bar{co}V_k = \cup_{\{b_l\}}\{b_1V_k + ... + b_{N+M+2}V_k\ \},$$

where the union is taken over all $b_l \geq 0,\ l=1,...,N+M+2$, such that $\sum_{l=1}^{N+M+2} b_l = 1$.

Thus, an optimal solution of the problem (5.6.14)-(5.6.16) can be presented in the form

$$\bar{\omega}_0^k = \sum_{l=1}^{N+M+2} \bar{b}_l^k q(u_l^k, y_l^k, z_k^{N,M}), \quad \bar{\omega}_i^k = \sum_{l=1}^{N+M+2} \bar{b}_l^k (\nabla\psi_i(z_k^{N,M}))^T g(u_l^k, y_l^k, z_k^{N,M}),$$
$$\bar{\upsilon}_j^k = \sum_{l=1}^{N+M+2} \bar{b}_l^k (\nabla\phi_j(y_l^k))^T \ f(u_l^k, y_l^k, z_k^{N,M}), \qquad i = 1, ..., N, \ j = 1, ..., M.$$

The latter implies that there exists an optimal solution of the problem (5.6.10)-(5.6.12) that is presentable in the form (5.6.6).

Let us now show that the relationships (5.6.9) are valid. Note, firstly, that from (5.3.25) and (5.3.28) it follows that

$$D^{N,M}(z) = \min_{(u,y)\in U\times Y} \{q(u,y,z) + (\nabla\zeta^{N,M}(z))^T g(u,y,z) + (\nabla\eta^{N,M}(y))^T f(u,y,z)\}. \tag{5.6.17}$$

By Lemma 5.6.1, $\mu_k^{N,M}$ is an optimal solution of the problem (5.6.4). That is,

$$\int_{U\times Y} [q(u,y,z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u,y,z_k^{N,M})]\mu_k^{N,M}(du,dy)$$
$$= \min_{\mu\in W_M(z_k^{N,M})} \int_{U\times Y} [q(u,y,z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u,y,z_k^{N,M})]\mu(du,dy)$$
$$= D^{N,M}(z_k^{N,M}),$$

the latter equality being due to the duality relationships between the problem (5.3.25) and (5.3.26). Since $\mu_k^{N,M} \in W_M(z_k^{N,M})$,

$$\int_{U\times Y} \Big[q(u,y,z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u,y,z_k^{N,M})\Big]\mu_k^{N,M}(du,dy)$$
$$= \int_{U\times Y} \Big[q(u,y,z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u,y,z_k^{N,M})$$
$$+(\nabla\eta^{N,M}(y))^T f(u,y,z_k^{N,M})\Big]\mu_k^{N,M}(du,dy).$$

Consequently,

$$\int_{U\times Y} \Big[q(u,y,z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u,y,z_k^{N,M})$$
$$+(\nabla\eta^{N,M}(y))^T f(u,y,z_k^{N,M})\Big]\mu_k^{N,M}(du,dy) = D^{N,M}(z_k^{N,M}).$$

After the substitution of (5.6.6) into the equality above and taking into account (5.6.7), one can obtain

$$\sum_{j=1}^{J^{N,M,k}} b_j^{N,M,k}\Big[q(u_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M})$$

$$+(\nabla\eta^{N,M}(y_j^{N,M,k}))^T f(_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M}) - D^{N,M}(z_k^{N,M})\Big] = 0. \quad (5.6.18)$$

By (5.6.17), from (5.6.18) it follows that

$$q(u_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M})$$

$$+(\nabla\eta^{N,M}(y_j^{N,M,k}))^T f(u_j^{N,M,k}, y_j^{N,M,k}, z_k^{N,M}) = D^{N,M}(z_k^{N,M}) \quad \forall j = 1, ..., J^{M,N,k}.$$

Also by (5.6.17), the latter implies

$$(u_j^{N,M,k}, y_j^{N,M,k}) = \mathrm{argmin}_{(u,y)\in U\times Y}\{q(u, y, z_k^{N,M}) + (\nabla\zeta^{N,M}(z_k^{N,M}))^T g(u, y, z_k^{N,M})$$

$$+ (\nabla\eta^{N,M}(y))^T f(u, y, z_k^{N,M})\}, \quad (5.6.19)$$

which, in turn, implies (5.6.9). □

Lemma 5.6.3 *For any $t \in [0,\infty)$ such that (5.5.5) is satisfied, there exists a sequence*

$$(\mu_{k^{N,M}}^{N,M}, z_{k^{N,M}}^{N,M}) \in \{(\mu_k^{N,M}, z_k^{N,M}),\ k = 1, ..., K^{N,M}\}, \quad N = 1, 2, ..., \quad M = 1, 2, ..., \quad (5.6.20)$$

(with $\{(\mu_k^{N,M}, z_k^{N,M}),\ k = 1, ..., K^{N,M}\}$ being the set of concentration points of the Dirac measures in (5.6.1)) such that

$$\lim_{N\to\infty} \limsup_{M\to\infty} \Big(\rho(\mu^*(t), \mu_{k^{N,M}}^{N,M}) + ||z^*(t) - z_{k^{N,M}}^{N,M}||\Big) = 0. \quad (5.6.21)$$

Let t be such that (5.6.21) is valid and let $(\mu_{k^{N,M}}^{N,M}, z_{k^{N,M}}^{N,M})$ be as in (5.6.21), then for any $\tau \in [0,\infty)$ such that (5.5.9) is satisfied, there exists a sequence

$$(u_{j^{N,M}}^{N,M,k^{N,M}}, y_{j^{N,M}}^{N,M,k^{N,M}}) \in \{(u_j^{N,M,k^{N,M}}, y_j^{N,M,k^{N,M}}), j = 1, ..., J^{N,M,k^{N,M}}\},$$
$$N = 1, 2, ..., \quad M = 1, 2, ..., \quad (5.6.22)$$

($\{(u_j^{N,M,k^{N,M}}, y_j^{N,M,k^{N,M}}),\ j = 1, ..., J^{M,N,k^{N,M}}\}$ being the set of concentration points of

the Dirac measures in (5.6.6) taken with $k = k^{N,M}$) such that

$$\lim_{N\to\infty} \limsup_{M\to\infty} \left(||u_t^*(\tau) - u_{j^{N,M}}^{N,M,k^{N,M}}|| + ||y_t^*(\tau)) - y_{j^{N,M}}^{N,M,k^{N,M}}|| \right) = 0. \qquad (5.6.23)$$

Proof. Assume that (5.6.21) is not true. Then there exists a number $r > 0$ and sequences N_i, $M_{i,j}$ with $i = 1, 2, ...,$ $j = 1, 2, ...$ and with $\lim_{i\to\infty} N_i = \infty$, $\lim_{j\to\infty} M_{i,j} = \infty$ such that

$$dist((\mu^*(t), z^*(t)), \Theta^{N_i,M_{i,j}}) \geq r, \qquad (5.6.24)$$

where $\Theta^{N,M}$ is the set of the concentration points of the Dirac measures in (5.6.1), that is,

$$\Theta^{N,M} \stackrel{\text{def}}{=} \{(\mu_k^{N,M}, z_k^{N,M}), \quad k = 1, ..., K^{N,M}\},$$

taken with $N = N_i$ and $M = M_{i,j}$, and where

$$dist((\mu, z), \Theta^{N,M}) \stackrel{\text{def}}{=} \min_{(\mu',z')\in\Theta^{N,M}} \{\rho(\mu, \mu') + ||z - z'||\}.$$

Hence,

$$(\mu_k^{N_i,M_{i,j}}, z_k^{N_i,M_{i,j}}) \notin \mathcal{B}_r(\bar{\mu}, \bar{z}), \quad k = 1, ..., K^{N_i,M_{i,j}}, \quad j = 1, 2, ... \,.$$

The latter implies that

$$\xi^{N_i,M_{i,j}}(\mathcal{B}_r(\bar{\mu}, \bar{z})) = 0, \qquad j = 1, 2, ... \,, \qquad (5.6.25)$$

where $\xi^{N,N}$ is defined by (5.6.1). Due to the fact that the optimal solution ξ^* of the IDLP problem (4.2.24) is unique (Assumption 5.5.1(i)), the relationship (5.3.13) is valid. Consequently,

$$\lim_{i\to\infty} \limsup_{j\to\infty} \rho(\xi^{N_i,M_{i,j}}, \xi^*) = 0. \qquad (5.6.26)$$

From (5.6.25) and (5.6.26) it follows that

$$\xi^*(\mathcal{B}_r(\bar{\mu}, \bar{z})) \leq \lim_{i\to\infty} \limsup_{j\to\infty} \xi^{N_i,M_{i,j}}(\mathcal{B}_r(\bar{\mu}, \bar{z})) = 0.$$

The latter contradicts to (5.5.5). Thus, (5.6.21) is proved.

Let us now prove the validity of (5.6.23). Assume it is not valid. Then there exists $r > 0$ and sequences N_i , $M_{i,j}$ with $i = 1, 2, ...$, $j = 1, 2, ...$, and with $\lim_{i\to\infty} N_i = \infty$, $\lim_{j\to\infty} M_{i,j} = \infty$ such that

$$dist((u_t^*(\tau), y_t^*(\tau)), \theta^{N_i,M_{i,j}}) \geq r, \qquad (5.6.27)$$

where $\theta^{N,M}$ is the set of the concentration points of the Dirac measures in (5.6.6),

$$\theta^{N,M} \stackrel{\text{def}}{=} \{(u_j^{N,M,k^{N,M}}, y_j^{N,M,k^{N,M}}), \quad j = 1, ..., J^{N,M,k^{N,M}}\},$$

taken with $k = k^{N,M}$ and with $N = N_i$, $M = M_{i,j}$, and where

$$dist((u,y), \theta^{N,M}) \stackrel{\text{def}}{=} \min_{(u',y') \in \theta^{N,M}} \{||u - u'|| + ||y - y'||\}.$$

From (5.6.27) it follows that

$$dist((u_t^*(\tau), y_t^*(\tau)), \theta^{N_i, M_{i,j}}) \geq r, \quad i, j = 1, 2, ... \ . \tag{5.6.28}$$

Hence,

$$(u_j^{N_i,M_{i,j},k^{N_i,M_{i,j}}}, y_j^{N_i,M_{i,j},k^{N_i,M_{i,j}}}) \notin B_r(\bar{u}, \bar{y}), \quad j = 1, ..., J^{N_i,M_{i,j},k^{N_i,M_{i,j}}}, \quad i, j = 1, 2, ... \ .$$

The latter implies that

$$\mu_{k^{N_i,M_{i,j}}}^{N_i,M_{i,j}}(B_r(\bar{u}, \bar{y})) = 0, \quad i, j = 1, 2, ... \ , \tag{5.6.29}$$

where $\mu_k^{N,M}$ is defined by (5.6.6) (taken with $k = k^{N_i,M_{i,j}}$, $N = N_i$ and $M = M_{i,j}$).

From (5.6.21) it follows, in particular, that

$$\lim_{N\to\infty} \limsup_{M\to\infty} \rho(\mu^*(t), \mu_{k^{N,M}}^{N,M}) = 0 \qquad \Rightarrow \qquad \lim_{i\to\infty} \limsup_{j\to\infty} \rho(\mu^*(t), \mu_{k^{N_i,M_{i,j}}}^{N_i,M_{i,j}}) = 0. \tag{5.6.30}$$

The later and (5.6.29) lead to

$$\mu^*(t)(B_r(\bar{u}, \bar{y})) \leq \lim_{i\to\infty} \limsup_{j\to\infty} \mu_{k^{N_i,M_{i,j}}}^{N_i,M_{i,j}}(B_r(\bar{u}, \bar{y})) = 0,$$

which contradicts to (5.5.9). Thus (5.6.23) is proved. □

Lemma 5.6.4 *For any $t \in [0, \infty)$ such that $\mathcal{Z}_t$ is not empty and (5.5.5) is valid for an arbitrary $r > 0$, and for any $\tau \in [0, \infty)$ such that $\mathcal{Y}_{t,\tau}$ is not empty and (5.5.9) is valid for an arbitrary $r > 0$,*

$$\lim_{N\to\infty} \limsup_{M\to\infty} ||u_t^*(\tau) - u^{N,M}(y_t^*(\tau), z^*(t))|| = 0. \tag{5.6.31}$$

Proof. By Lemma 5.6.3, there exist $(\mu_{k^{N,M}}^{N,M}, z_{k^{N,M}}^{N,M})$ such that (5.6.21) is satisfied and there exist $(u_{j^{N,M}}^{N,M,k^{N,M}}, y_{j^{N,M}}^{N,M,k^{N,M}})$ such that (5.6.23) is satisfied.

Note that, due to (5.6.9),

$$u_{j^{N,M}}^{N,M,k^{N,M}} = u(y_{j^{N,M}}^{N,M,k^{N,M}}, z_{k^{N,M}}^{N,M}), \tag{5.6.32}$$

where $u(y, z)$ is as in (5.5.2). From (5.6.21) and (5.6.23) it follows that $z_{k^{N,M}}^{N,M} \in \mathcal{Z}_t$ and $y_{j^{N,M}}^{N,M,k^{N,M}} \in \mathcal{Y}_{t,\tau}$ for N and M large enough. Hence, one can use (5.5.17) to obtain

$$||u_t^*(\tau) - u^{N,M}(y_t^*(\tau), z^*(t))|| \leq ||u_t^*(\tau) - u_{j^{N,M}}^{N,M,k^{N,M}}||$$

$$+||u(y_{j^{N,M}}^{N,M,k^{N,M}}, z_{k^{N,M}}^{N,M}) - u^{N,M}(y_t^*(\tau), z^*(t))|| \leq ||u_t^*(\tau) - u_{j^{N,M}}^{N,M,k^{N,M}}||$$

$$+L(||y_t^*(\tau) - y_{j^{N,M}}^{N,M,k^{N,M}}|| + ||z^*(t) - z_{k^{N,M}}^{N,M}||).$$

By (5.6.21) and (5.6.23), the latter implies (5.6.31). Since $\mathcal{Y}_{t,\tau}$ is not empty for almost all $\tau \in [0, \infty)$ (Assumption 5.5.7 (i)), the convergence (5.6.31) takes place for almost all $\tau \in [0, \infty)$. $\square$

Proof of Lemma 5.5.9. Let $t \in [0, \infty)$ be such that $\mathcal{Z}_t$ is not empty and (5.5.5) is satisfied for an arbitrary $r > 0$. Note that, from the assumptions made, it follows that (5.6.31) is valid for almost all $\tau \in [0, \infty)$.

Take an arbitrary $\tau \in [0, \infty)$ and subtract the equation

$$y_t^*(\tau) = y_t^*(0) + \int_0^\tau f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))d\tau' \tag{5.6.33}$$

from the equation

$$y_t^{N,M}(\tau) = y_t^*(0) + \int_0^\tau f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t))d\tau'. \tag{5.6.34}$$

We will obtain

$$||y_t^{N,M}(\tau) - y_t^*(\tau)|| \leq \int_0^\tau ||f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t))$$

$$-f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))||d\tau'$$

$$\leq \int_0^\tau ||f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t)) - f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))||d\tau'$$

$$+ \int_0^\tau ||f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))||d\tau'. \tag{5.6.35}$$

Using Assumption 5.5.7 (ii),(iii), one can derive that

$$\int_0^{\tau} ||f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t)) - f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))||d\tau'$$

$$\leq \int_{\tau' \notin P_{t,\tau}(N,M)} ||f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t))$$

$$-f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))||d\tau'$$

$$+ \int_{\tau' \in P_{t,\tau}(N,M)} (||f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_{z^*(t)}^{N,M}(\tau'), z^*(t))||$$

$$+||f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))||)d\tau'$$

$$\leq L_1 \int_0^{\tau} ||y_t^{N,M}(\tau') - y_t^*(\tau')||d\tau' + L_2 meas\{P_{t,\tau}(N,M)\}, \tag{5.6.36}$$

where L_1 is a constant defined (in an obvious way) by Lipschitz constants of $f(\cdot)$ and $u^{N,M}(\cdot)$, and $L_2 \stackrel{\text{def}}{=} 2\ max_{(u,y,z)\in U\times Y\times Z}\{||f(u,y,z)||\}$.

Also, due to (5.6.31) and the dominated convergence theorem (see, e.g., p. 49 in [15])

$$\lim_{N\to\infty} \limsup_{M\to\infty} \int_0^{\tau} ||f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))||d\tau' = 0. \tag{5.6.37}$$

Let us introduce the notation

$$\kappa_{t,\tau}(N,M) \stackrel{\text{def}}{=} L_2\ meas\{P_{t,\tau}(N,M)\}$$

$$+ \int_0^{\tau} ||f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))||d\tau'$$

and rewrite the inequality (5.6.35) in the form

$$||y_t^{N,M}(\tau) - y_t^*(\tau)|| \leq L_1 \int_0^{\tau} ||y_t^{N,M}(\tau') - y_t^*(\tau')||d\tau' + \kappa_{t,\tau}(N,M). \tag{5.6.38}$$

By Gronwall-Bellman lemma (see, e.g., p.218 in [20]), it follows that

$$\max_{\tau'\in[0,\tau]} ||y_t^{N,M}(\tau') - y_t^*(\tau')|| \leq \kappa_{t,\tau}(N,M)e^{L_1\tau}. \tag{5.6.39}$$

Since, by (5.5.18) and (5.6.37),

$$\lim_{N\to\infty} \limsup_{M\to\infty} \kappa_{t,\tau}(N,M) = 0, \tag{5.6.40}$$

the inequality (5.6.39) implies (5.5.27).

By (5.5.27), $y_t^{N,M}(\tau) \in \mathcal{Y}_{t,\tau}$ for N and M large enough (for $\tau \in [0,\infty)$ such that the ball $\mathcal{Y}_{t,\tau}$ is not empty). Hence,

$$||u^{N,M}(y_t^{N,M}(\tau), z^*(t)) - u_t^*(\tau)|| \leq ||u^{N,M}(y_t^{N,M}(\tau), z^*(t)) - u^{N,M}(y_t^*(\tau), z^*(t))||$$

$$+||u^{N,M}(y_t^*(\tau), z^*(t)) - u_t^*(\tau)|| \leq L||y_{z^*(t)}^{N,M}(\tau) - y_t^*(\tau)|| + ||u^{N,M}(y_t^*(\tau), z^*(t)) - u_t^*(\tau)||.$$

The latter implies (5.5.28) (by (5.5.27) and (5.6.31)).□

Proof of Theorem 5.5.8. Let $t \in [0,\infty)$ be such that $\mathcal{Z}_t$ is not empty and (5.5.5) is satisfied for an arbitrary $r > 0$. By (5.5.8) and (5.5.10), for any continuous $h(u,y)$ and for an arbitrary small $\alpha > 0$, there exists $S > 0$ such that

$$|\frac{1}{S}\int_0^S h(u_t^*(\tau), y_t^*(\tau))d\tau - \int_{U\times Y} h(u,y)\mu^*(t)(du,dy)| \leq \frac{\alpha}{2} \tag{5.6.41}$$

and

$$|\frac{1}{S}\int_0^S h(u^{N,M}(y_t^{N,M}(\tau), z^*(t)), y_t^{N,M}(\tau))d\tau - \int_{U\times Y} h(u,y)\mu^{N,M}(z^*(t))(du,dy)| \leq \frac{\alpha}{2}. \tag{5.6.42}$$

Using (5.6.42) and (5.6.41), one can obtain

$$|\int_{U\times Y} h(u,y)\mu^{N,M}(z^*(t))(du,dy) - \int_{U\times Y} h(u,y)\mu^*(t)(du,dy)|$$

$$\leq \ |\frac{1}{S}\int_0^S h(u^{N,M}(y_t^{N,M}(\tau), z^*(t)), y_t^{N,M}(\tau))d\tau - \frac{1}{S}\int_0^S h(u_t^*(\tau), y_t^*(\tau))d\tau| + \alpha.$$

Due to Lemma 5.5.9, the latter implies the following inequality

$$\lim_{N\to\infty}\limsup_{M\to\infty} |\int_{U\times Y} h(u,y)\mu^{N,M}(z^*(t))(du,dy) - \int_{U\times Y} h(u,y)\mu^*(t)(du,dy)| \leq \alpha,$$

which, in turn, implies

$$\lim_{N\to\infty}\limsup_{M\to\infty} |\int_{U\times Y} h(u,y)\mu^{N,M}(z^*(t))(du,dy) - \int_{U\times Y} h(u,y)\mu^*(t)(du,dy)| = 0 \tag{5.6.43}$$

(due to the fact that α can be arbitrary small). Since $h(u,y)$ is an arbitrary continuous function, from (5.6.43) it follows that

$$\lim_{N\to\infty}\limsup_{M\to\infty} \rho(\mu^{N,M}(z^*(t)), \mu^*(t)) = 0. \tag{5.6.44}$$

Note that from the assumptions made it follows that $\mathcal{Z}_t$ is not empty and (5.5.5) is satisfied for an arbitrary $r > 0$ for almost all $t \in [0, \infty)$. Hence, (5.6.44) is valid for almost all $t \in [0, \infty)$.

Taking an arbitrary $t \in [0, \infty)$ and subtracting the equation

$$z^*(t) = z_0 + \int_0^t \tilde{g}(\mu^*(t'), z^*(t'))dt' \tag{5.6.45}$$

from the equation

$$z^{N,M}(t) = z_0 + \int_0^t \tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t'))dt', \tag{5.6.46}$$

one obtains

$$\begin{aligned} ||z^{N,M}(t) - z^*(t)|| &\leq \int_0^t ||\tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t')) - \tilde{g}(\mu^*(t'), z^*(t'))||dt' \\ &\leq \int_0^t ||\tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t')) - \tilde{g}(\mu^{N,M}(z^*(t')), z^*(t'))||dt' \\ &\quad + \int_0^t ||\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))||dt'. \end{aligned} \tag{5.6.47}$$

From (5.5.13) and from the definition of the set $A_t(N, M)$ (see (5.5.16)), it follows that

$$\begin{aligned} &\int_0^t ||\tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t')) - \tilde{g}(\mu^{N,M}(z^*(t')), z^*(t'))||dt' \\ &\leq \int_{t' \notin A_t(N,M)} ||\tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t')) - \tilde{g}(\mu^{N,M}(z^*(t')), z^*(t'))||dt' \\ &+ \int_{t' \in A_t(N,M)} (||\tilde{g}(\mu^{N,M}(z^{N,M}(t')), z^{N,M}(t'))|| + ||\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t'))||\ )dt' \\ &\leq L \int_0^t ||z^{N,M}(t') - z^*(t')||dt' + 2L_g meas\{A_t(N, M)\}, \end{aligned} \tag{5.6.48}$$

where $L_g \stackrel{\text{def}}{=} \max_{(u,y,z) \in U \times Y \times Z} ||g(u, y, z)||$. This and (5.6.47) allows one to obtain the inequality

$$||z^{N,M}(t) - z^*(t)|| \leq L \int_0^t ||z^{N,M}(t') - z^*(t')||dt' + \kappa_t(N, M), \tag{5.6.49}$$

where

$$\kappa_t(N, M) \stackrel{\text{def}}{=} 2L_g meas\{A_t(N, M)\} + \int_0^t ||\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))||dt'.$$

Note that, by (5.6.44),

$$\lim_{N\to\infty}\limsup_{M\to\infty}\int_0^t ||\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))||dt' = 0, \tag{5.6.50}$$

which, along with (5.5.15), imply that

$$\lim_{N\to\infty}\limsup_{M\to\infty}\kappa_t(N,M) = 0. \tag{5.6.51}$$

By Gronwall-Bellman lemma, from (5.6.49) it follows that

$$max_{t'\in[0,t]}||z^{N,M}(t') - z^*(t')|| \leq \kappa_t(N,M)e^{Lt}.$$

The latter along with (5.6.51) imply (5.5.22).

Let us now establish the validity of (5.5.23). Let $t \in [0,\infty)$ be such that the ball $\mathcal{Z}_t$ introduced in Assumption 5.5.6 is not empty. By triangle inequality,

$$\rho(\mu^{N,M}(z^{N,M}(t)), \mu^*(t)) \leq \rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) + \rho(\mu^{N,M}(z^*(t)), \mu^*(t)). \tag{5.6.52}$$

Due to (5.5.22), $z^{N,M}(t) \in \mathcal{Z}_t$ for M and N large enough. Hence, by (5.5.12),

$$\rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) \leq \kappa(||z^{N,M}(t') - z^*(t')||),$$

which implies that

$$\lim_{N\to\infty}\limsup_{M\to\infty}\rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) = 0.$$

The latter, along with (5.6.44) and (5.6.52), imply (5.5.23).

Finally, let us prove (5.5.24). By (5.5.6) and (5.5.11), for any continuous function $\tilde{h}(\mu,z) : \mathcal{P}(\mathcal{P}(U\times Y)\times Z) \to I\!R^1$, and for an arbitrary small $\alpha > 0$, there exists $\tilde{\mathcal{T}} > 0$ such that

$$|\frac{1}{\tilde{\mathcal{T}}}\int_0^{\tilde{\mathcal{T}}} \tilde{h}(\mu^*(t), z^*(t))dt - \int_F \tilde{h}(\mu,z)\xi^*(d\mu,dz)| \leq \frac{\alpha}{2} \tag{5.6.53}$$

and

$$|\frac{1}{\tilde{\mathcal{T}}}\int_0^{\tilde{\mathcal{T}}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t))dt - \int_F \tilde{h}(\mu,z)\bar{\xi}^{N,M}(d\mu,dz)| \leq \frac{\alpha}{2}. \tag{5.6.54}$$

Using (5.6.54) and (5.6.53), one can obtain

$$
\begin{aligned}
& |\int_F \tilde{h}(\mu, z)\bar{\xi}^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z)\xi^*(d\mu, dz)| \\
\leq \ & |\frac{1}{\tilde{\mathcal{T}}}\int_0^{\tilde{\mathcal{T}}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\tilde{\mathcal{T}}}\int_0^{\tilde{\mathcal{T}}} \tilde{h}(\mu^*(t), z^*(t))dt| + \alpha.
\end{aligned}
$$

Due to (5.5.22) and (5.5.23), the latter implies the following inequality

$$
\lim_{N\to\infty} \limsup_{M\to\infty} |\int_F \tilde{h}(\mu, z)\bar{\xi}^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z)\xi^*(d\mu, dz)| \leq \alpha,
$$

which, in turn, implies

$$
\lim_{N\to\infty} \limsup_{M\to\infty} |\int_F \tilde{h}(\mu, z)\bar{\xi}^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z)\xi^*(d\mu, dz)| = 0 \qquad (5.6.55)
$$

(due to the fact that α can be arbitrary small). This proves (5.5.21). Taking now $\tilde{h}(\mu, z) = \tilde{q}(\mu, z)$ in (5.6.55) and having in mind that

$$
\int_F \tilde{q}(\mu, z)\bar{\xi}^{N,M}(d\mu, dz) = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt
$$

(see (5.5.11)) and that

$$
\int_F \tilde{q}(\mu, z)\xi^*(d\mu, dz) = \tilde{G}^*,
$$

one proves the validity of (5.5.24). This completes the proof of the theorem. $\square$

5.7 Additional comments for Chapter 5

The concept of an ACG family was introduced for singularly perturbed problems with time discounting criteria of optimality and it was extended to LRAOC problems in [61]. The consideration of this chapter is based on results of the paper [61].

6

Asymptotically near optimal controls of the singularly perturbed problem

In this chapter, we discuss a construction of an asymptotically optimal (near optimal) controls of singularly perturbed (SP) problems with long run time average criteria. Namely, in Section 6.1, we indicate a way how asymptotically near optimal controls of the SP problems can be constructed on the basis of near optimal ACG families. In Section 6.2, a linear programming based algorithm allowing one to find solutions of approximating averaged problem and solutions of the corresponding approximating (averaged and associated) dual problems numerically is discussed. In Section 6.3, we consider an example of SP optimal control problem, the near optimal solution of which is obtained with the proposed technique.

6.1 Construction of asymptotically optimal/near optimal controls of the singularly perturbed problem.

In this section, we describe a way how an asymptotically optimal (near optimal) control of the SP optimal control problem (4.1.5) can be constructed given that an asymptotically optimal (near optimal) ACG family is known (a way of construction of the latter has been discussed in Section 5.5).

Definition 6.1.1 *A control $u_\epsilon(\cdot)$ will be called asymptotically α-near optimal ($\alpha > 0$) in the SP problem (4.1.5) if the solution $(y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the system (4.1.1)-(4.1.2)obtained with this control satisfies (4.1.4) (that is, the triplet $(u_\epsilon(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ is admissible) and if*

$$\lim_{\epsilon\to 0}\liminf_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u_\epsilon(t), y_\epsilon(t), z_\epsilon(t))dt \leq \liminf_{\epsilon\to 0} V^*(\epsilon) + \alpha. \tag{6.1.1}$$

For simplicity, we will be dealing with a special case when $f(u, y, z) = f(u, y)$. That is, the right hand side in (4.1.1) is independent of z (the SP systems that have such a property are called "weakly coupled"). Note that in this case the set $W(z)$ defined in (4.2.3) does not depend on z too. That is, $W(z) = W$.

Let us also introduce the following assumptions about the functions $f(u, y)$ and $\tilde{g}(\mu, z)$.

Assumption 6.1.2 *(i) There exists a positive definite matrix A_1 such that its eigenvalues are greater than a positive constant and such that*

$$(f(u, y') - f(u, y''))^T A_1(y' - y'') \tag{6.1.2}$$

$$\leq -(y' - y'')^T(y' - y'') \quad \forall y', y'' \in I\!R^m\ ,\ \forall u \in U.$$

(ii) There exists a positive definite matrix A_2 such that its eigenvalues are greater than a positive constant and such that

$$(\tilde{g}(\mu, z') - \tilde{g}(\mu, z'')^T A_2(z' - z'') \tag{6.1.3}$$

$$\leq -(z' - z'')^T(z' - z'') \quad \forall z', z'' \in I\!R^n\ ,\ \forall \mu \in W.$$

Note that these are Liapunov type stability conditions and, as has been established in [56], their fulfillment is sufficient for the validity of the statement that the SP system is uniformly approximated by the averaged system (see Definition 4.2.3). Also, as can be readily verified, Assumption 6.1.2(i) implies that the solutions $y(\tau, u(\cdot), y_1)$ and $y(\tau, u(\cdot), y_2)$ of the associated system (4.2.1) obtained with an arbitrary control $u(\cdot)$ and with initial values $y(0) = y_1$ and $y(0) = y_2$ (y_1 and y_2 being arbitrary vectors in Y) satisfy the inequality

$$||y(\tau, u(\cdot), y_1) - y(\tau, u(\cdot), y_2)|| \leq c_1 e^{-c_2 \tau} ||y_1 - y_2||, \tag{6.1.4}$$

where c_1, c_2 are some positive constants. Similarly, Assumption 6.1.2(ii) implies that the solutions $z(t, \mu(\cdot), z_1)$ and $z(t, \mu(\cdot), z_2)$ of the averaged system (4.2.8) obtained with an arbitrary control $\mu(\cdot)$ and with initial values $z(0) = z_1$ and $z(0) = z_2$ (z_1 and z_2 being arbitrary vectors in Z) satisfy the inequality

$$||z(t, \mu(\cdot), z_1) - z(t, \mu(\cdot), z_2)|| \leq c_3 e^{-c_4 t} ||z_1 - z_2||, \tag{6.1.5}$$

where c_3, c_4 are some positive constants.

From the validity of (6.1.4) and (6.1.5) it follows that the associated system (4.2.1) and the averaged system (4.2.8) have unique forward invariant sets which also are global attractors for the solutions of these systems (see Theorem 3.1(ii) in [52]). For simplicity, we will assume that Y and Z are these sets.

Let $(u_z^{N,M}(\tau), y_z^{N,M}(\tau))$ be the ACG family introduced in Assumptions 5.5.5(i) and let $\mu^{N,M}(du, dy|z) = \mu^{N,M}(z)$, $z^{N,M}(t)$ and $\mu^{N,M}(z^{N,M}(t))$ be generated by this family as assumed in Section 5.5 (all the assumptions made in that section are supposed to be satisfied in the consideration below). Let $y_z^{N,M}(\tau, y)$ stand for the solution of the associated system (4.2.1) obtained with the control $u_z^{N,M}(\tau)$ and with the initial condition $y_z^{N,M}(0, y) = y \in Y$. From (6.1.4) it follows that

$$||y_z^{N,M}(\tau, y) - y_z^{N,M}(\tau)|| \leq c_1 e^{-c_2 \tau} \max_{y', y'' \in Y} ||y' - y''||.$$

The latter implies that, for any Lipschitz continuous function $h(u, y, z)$, there exists $\bar{\delta}_h(S)$, $\lim_{S \to \infty} \bar{\delta}_h(S) = 0$, such that

$$\left|\frac{1}{S} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau, y), z) d\tau - \frac{1}{S} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau), z) d\tau\right| \leq \bar{\delta}_h(S), \tag{6.1.6}$$

which, due to (5.5.10), implies that

$$|\frac{1}{S}\int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau,y), z)d\tau - \int_{U\times Y} h(u,y,z)\mu^{N,M}(du,dy|z))| \tag{6.1.7}$$

$$\leq \delta_h(S) + \bar{\delta}_h(S) \stackrel{\text{def}}{=} \bar{\bar{\delta}}_h(S),$$

with $\lim_{S\to\infty} \bar{\bar{\delta}}_h(S) = 0$. Hence,

$$\lim_{S\to\infty} \rho(\mu^{N,M}(S,y), \mu^{N,M}(z)) \ \leq \ \delta(S), \qquad \lim_{S\to\infty} \delta(S) = 0, \tag{6.1.8}$$

where $\mu^{N,M}(S,y)$ is the occupational measure generated by the pair $(u_z^{N,M}(\tau), y_z^{N,M}(\tau,y))$ on the interval $[0,S]$. That is, the family of measures $\mu^{N,M}(z)$ is uniformly attainable by the associated system with the use of the control $u_z^{N,M}(\tau)$ (see Definition 4.3 in [64]).

Partition the interval $[0,\infty)$ by the points

$$t_l = l\Delta(\epsilon), \quad l = 0,1,\ldots\ , \tag{6.1.9}$$

where $\Delta(\epsilon) > 0$ is such that

$$\lim_{\epsilon\to 0} \Delta(\epsilon) = 0, \qquad \lim_{\epsilon\to 0} \frac{\Delta(\epsilon)}{\epsilon} = \infty. \tag{6.1.10}$$

Define the control $u_\epsilon^{N,M}(t)$ by the equation

$$u_\epsilon^{N,M}(t) \stackrel{\text{def}}{=} u^{N,M}_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}) \quad \forall t \in [t_l, t_{1+1}), \quad l = 0,1,\ldots\ . \tag{6.1.11}$$

Theorem 6.1.3 *Let the assumptions of Theorem 5.5.8 be satisfied and let the function*

$$\mu^{N,M}(t) \stackrel{\text{def}}{=} \mu^{N,M}(z^{N,M}(t)) \tag{6.1.12}$$

has the following piecewise continuity property: for any $\mathcal{T} > 0$, there may exist no more than a finite number of points $\mathcal{T}_i \in (0,\mathcal{T})$, $i = 1,\ldots k,$, with

$$k \leq c\mathcal{T}, \qquad c = \text{const}, \tag{6.1.13}$$

such that, for any $t \neq \mathcal{T}_i$,

$$\max\{||\tilde{g}(\mu^{N,M}(t'), z) - \tilde{g}(\mu^{N,M}(t), z)||, \ |\tilde{q}(\mu^{N,M}(t'), z) - \tilde{q}(\mu^{N,M}(t), z)|\} \leq \nu(t - t') \tag{6.1.14}$$

$$\forall t' \in (t - a_t, t + a_t)$$

where $\nu(\cdot)$ *is monotone decreasing, with* $\lim_{\theta \to 0} \nu(\theta) = 0$, *and where* $a_t > 0$, *with* r_δ,

$$r_\delta \stackrel{\text{def}}{=} \inf\{a_t \ : \ t \notin \cup_{i=1}^k (\mathcal{T}_i - \delta, \mathcal{T}_i + \delta)\},$$

being a positive continuous function of δ *(which may tend to zero when* δ *tends to zero). Let also Assumption 6.1.2 be valid and the solution* $(y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ *of the system (4.1.1)-(4.1.2) obtained with the control* $u_\epsilon^{N,M}(\cdot)$ *and with the initial conditions* $(y_\epsilon^{N,M}(0), z_\epsilon^{N,M}(0)) = (y^{N,M}(0), z^{N,M}(0))$ *satisfies the inclusion (4.1.4). Then the control* $u_\epsilon^{N,M}(\cdot)$ *is* $\beta(N,M)$*-asymptotically near optimal in the problem (4.1.5), where* $\beta(N,M)$ *is defined in (5.5.26). Also,*

$$\lim_{\epsilon \to 0} \sup_{t \in [0,\infty)} ||z_\epsilon^{N,M}(t) - z^{N,M}(t)|| \ = \ 0 \tag{6.1.15}$$

and, if the triplet $(u_\epsilon^{N,M}(\cdot), y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ *generates the occupational measure* $\gamma_\epsilon^{N,M}$ *on the interval* $[0,\infty)$ *(see (1.1.18)), then*

$$\rho(\gamma_\epsilon^{N,M}, \Phi(\bar{\xi}^{N,M})) \ \leq \ \kappa(\epsilon) \quad \text{ where } \quad \lim_{\epsilon \to 0} \kappa(\epsilon) = 0, \tag{6.1.16}$$

with $\bar{\xi}^{N,M}$ *being the occupational measure generated by* $(\mu^{N,M}(\cdot), z^{N,M}(\cdot))$ *(see (5.5.11)) and the map* $\Phi(\cdot)$ *being defined by (4.3.4).*

Remark 6.1.4 *Note that, in case when the function* $\mu^{N,M}(t)$ *is periodic (as in the numerical example considered in Section 6.3 below) the inequality (6.1.13) will be satisfied if* $\mu^{N,M}(t)$ *is picewise continuous.*

Proof. Denote by $z(t)$ the solution of the differential equation

$$z'(t) = \tilde{g}(\mu^{N,M}(t), z(t)) \tag{6.1.17}$$

considered on the interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ that satisfies the initial condition

$$z(\mathcal{T}_0) = z \in Z. \tag{6.1.18}$$

Also, denote by $\bar{z}(t)$ the solution of the differential equation

$$z'(t) = \tilde{g}(\bar{\mu}^{N,M}(t), z(t)) \tag{6.1.19}$$

considered on the same interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ and satisfying the same initial condition (6.1.18), where $\bar{\mu}^{N,M}(t)$ is the piecewise constant function defined as follows

$$\bar{\mu}^{N,M}(t) \stackrel{\text{def}}{=} \mu^{N,M}(t_l) \quad \forall t \in [t_l, t_{l+1}), \quad l = 0, 1, \dots \, . \tag{6.1.20}$$

Using the piecewise continuity property (6.1.14), it can be readily established (using a standard argument, see, e.g., the proof of Theorem 4.5 in [64]) that

$$\max_{t \in [\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]} ||\bar{z}(t) - z(t)|| \ \leq \ \kappa_1(\epsilon, \mathcal{T}), \qquad \text{where} \qquad \lim_{\epsilon \to 0} \kappa_1(\epsilon, \mathcal{T}) = 0. \tag{6.1.21}$$

The latter implies, in particular,

$$\max_{t \in [0, \mathcal{T}_0]} ||z^{N,M}(t) - \bar{z}^{N,M}(t)|| \leq \kappa_1(\epsilon, \mathcal{T}_0), \tag{6.1.22}$$

where $\bar{z}^{N,M}(t)$ is the solution of (6.1.19) that satisfies the initial condition $\bar{z}^{N,M}(0) = z^{N,M}(0)$.

Choose now $\mathcal{T}_0$ in such a way that

$$c_3 e^{-c_4 \mathcal{T}_0} \stackrel{\text{def}}{=} a < 1 \tag{6.1.23}$$

and denote by $\bar{z}_1^{N,M}(t)$ the solution of the system (6.1.19) considered on the interval $[\mathcal{T}_0, 2\mathcal{T}_0]$ with the initial condition $\bar{z}_1^{N,M}(\mathcal{T}_0) = z^{N,M}(\mathcal{T}_0)$. From (6.1.5) and (6.1.22) it follows that

$$||\bar{z}_1^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)|| \leq a||z^{N,M}(\mathcal{T}_0) - \bar{z}^{N,M}(\mathcal{T}_0)|| \leq a\kappa_1(\epsilon, \mathcal{T}_0). \tag{6.1.24}$$

Also, taking into account the validity of (6.1.21), one can write down

$$||z^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)|| \ \leq \ ||z^{N,M}(2\mathcal{T}_0) - \bar{z}_1^{N,M}(2\mathcal{T}_0)|| \ + \ ||\bar{z}_1^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)||$$

$$\leq \ \kappa_1(\epsilon, \mathcal{T}_0) \ + \ a\kappa_1(\epsilon, \mathcal{T}_0) \ \leq \ \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1 - a}.$$

By continuing in a similar way, one can prove that, for any $k = 1, 2, ...$,

$$||z^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)|| \leq (1 + a + ... + a^{k-1})\kappa_1(\epsilon, \mathcal{T}_0) \leq \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a}.$$

Hence, by (6.1.5),

$$\max_{t\in[k\mathcal{T}_0,(k+1)\mathcal{T}_0]} ||z^{N,M}(t) - \bar{z}^{N,M}(t)|| \leq c_3 \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a} \quad \forall k = 0, 1, ...$$

and, consequently,

$$\sup_{t\in[0,\infty)} ||z^{N,M}(t) - \bar{z}^{N,M}(t)|| \leq c_3 \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a} \stackrel{\text{def}}{=} \kappa_2(\epsilon), \qquad \lim_{\epsilon\to 0} \kappa_2(\epsilon) = 0. \qquad (6.1.25)$$

Using (6.1.25), one can obtain

$$\begin{aligned} &|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt| \\ &\leq \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} |\tilde{q}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{q}(\bar{\mu}^{N,M}(t), z^{N,M}(t))|dt + L\kappa_2(\epsilon) \\ &\leq \frac{1}{\mathcal{T}}\sum_{l=0}^{\lfloor \frac{\mathcal{T}}{\Delta(\epsilon)}\rfloor} \int_{t_l}^{t_{l+1}} |\tilde{q}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{q}(\mu^{N,M}(t_l), z^{N,M}(t))|dt + L\kappa_2(\epsilon) + 2K\Delta(\epsilon) \end{aligned} \qquad (6.1.26)$$

$$\forall \mathcal{T} \geq 1,$$

where $\lfloor\cdot\rfloor$ stands for the floor function ($\lfloor x\rfloor$ is the maximal integer number that is less or equal than x), L is a Lipschitz constant (for simplicity it is assumed that $q(u, y, z)$ is Lipschitz continuous in z) and $K \stackrel{\text{def}}{=} \max_{(u,y,z)\in U\times Y\times Z} |q(u, y, z)|$.

Without loss of generality, one may assume that r_δ is decreasing with δ and that $r_\delta \leq \delta$ (the later can be achieved by replacing r_δ with $\min\{\delta, r_\delta\}$ if necessary). Having this in mind, define $\delta(\epsilon)$ as the solution of the problem

$$\min\{\delta \ : \ r_\delta \geq \Delta^{\frac{1}{2}}(\epsilon)\}. \qquad (6.1.27)$$

That is,

$$r_{\delta(\epsilon)} = \Delta^{\frac{1}{2}}(\epsilon). \qquad (6.1.28)$$

Note that, by construction,

$$\lim_{\epsilon\to 0} \delta(\epsilon) = 0, \qquad \delta(\epsilon) \geq \Delta^{\frac{1}{2}}(\epsilon). \qquad (6.1.29)$$

By (6.1.14), if $t_l \notin \cup_{i=1}^{k}(\mathcal{T}_i - \delta(\epsilon), \mathcal{T}_i + \delta(\epsilon))$

$$\int_{t_l}^{t_{l+1}} |\tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt - \tilde{q}(\mu^{N,M}(t_l), z^{N,M}(t))|dt \leq \Delta(\epsilon)\nu(\Delta(\epsilon)). \qquad (6.1.30)$$

Also, if $t_l \in \cup_{i=1}^{k}(\mathcal{T}_i - \delta(\epsilon), \mathcal{T}_i + \delta(\epsilon))$

$$\int_{t_l}^{t_{l+1}} |\tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt - \tilde{q}(\mu^{N,M}(t_l), z^{N,M}(t))|dt \leq 2K\Delta(\epsilon). \qquad (6.1.31)$$

Taking (6.1.13), (6.1.30) and (6.1.31) into account, one can use (6.1.26) to obtain the following estimate

$$\Big|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt\Big|$$
$$\leq \frac{1}{\mathcal{T}}\lfloor\frac{\mathcal{T}}{\Delta(\epsilon)}\rfloor\Delta(\epsilon)\nu(\Delta(\epsilon)) + \frac{1}{\mathcal{T}}(c\mathcal{T})\Big[\frac{2\delta(\epsilon)}{\Delta(\epsilon)} + 2\Big]\Big(2K\Delta(\epsilon)\Big) + L\kappa_2(\epsilon) + 2K\Delta(\epsilon)$$
$$\stackrel{\text{def}}{=} \kappa_3(\epsilon) \qquad \forall \mathcal{T} \geq 1.$$

Thus,

$$\sup_{\mathcal{T}\geq 1}|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt| \leq \kappa_3(\epsilon), \qquad (6.1.32)$$

$$\lim_{\epsilon\to 0}\kappa_3(\epsilon) = 0.$$

Denote by $\bar{\bar{z}}(t)$ the solution of the differential equation (6.1.19) considered on the interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ and satisfying the initial condition $\bar{\bar{z}}(\mathcal{T}_0) = z_\epsilon^{N,M}(\mathcal{T}_0)$, where $\mathcal{T}_0 \stackrel{\text{def}}{=} l_0\Delta(\epsilon)$ for some $l_0 \geq 0$. Subtracting the equation

$$\bar{\bar{z}}(t_{l+1}) = \bar{\bar{z}}(t_l) + \int_{t_l}^{t_{l+1}} \tilde{g}(\mu^{N,M}(t_l), \bar{\bar{z}}(t))dt, \qquad l \geq l_0,$$

from the equation

$$z_\epsilon^{N,M}(t_{l+1}) = z_\epsilon^{N,M}(t_l) + \int_{t_l}^{t_{l+1}} g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt, \qquad l \geq l_0,$$

one can obtain

$$||z_\epsilon^{N,M}(t_{l+1}) - \bar{\bar{z}}(t_{l+1})|| \leq ||z_\epsilon^{N,M}(t_l) - \bar{\bar{z}}(t_l)||$$
$$+ \int_{t_l}^{t_{l+1}} ||g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt - g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t_l))||dt$$

$$+ || \int_{t_l}^{t_{l+1}} g(u^{N,M}_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}),\ y^{N,M}_{\epsilon}(t), z^{N,M}_{\epsilon}(t_l))dt \ - \ \Delta(\epsilon)\tilde{g}(\mu^{N,M}(t_l), z^{N,M}_{\epsilon}(t_l))||$$

$$+ \int_{t_l}^{t_{l+1}} ||\tilde{g}(\mu^{N,M}(t_l), z^{N,M}_{\epsilon}(t_l)) - \tilde{g}(\mu^{N,M}(t_l), \bar{\bar{z}}(t_l))||dt$$

$$\leq \ ||z^{N,M}_{\epsilon}(t_l) - \bar{\bar{z}}(t_l)|| + L_1\Delta(\epsilon)||z^{N,M}_{\epsilon}(t_l) \ - \ \bar{\bar{z}}(t_l)|| + L_2\Delta^2(\epsilon) + \Delta(\epsilon)\bar{\bar{\delta}}_g(\frac{\Delta(\epsilon)}{\epsilon}), \quad (6.1.33)$$

where $L_i, \ i = 1, 2,$ are positive constants and $\bar{\bar{\delta}}_g(\cdot)$ is defined in (6.1.7). Note that, in order to obtain the estimate above, one needs to take into account the fact that

$$\max\{ \max_{t\in[t_l,t_{l+1}]} \{||z^{N,M}_{\epsilon}(t) - z^{N,M}_{\epsilon}(t_l)||\}, \quad \max_{t\in[t_l,t_{l+1}]} \{||\bar{\bar{z}}(t) - \bar{\bar{z}}(t_l)||\} \ \} \ \leq \ L_3\Delta(\epsilon), \qquad L_3 > 0 \quad (6.1.34)$$

as well as the fact that (see (6.1.7))

$$|| \int_{t_l}^{t_{l+1}} g(u^{N,M}_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}),\ y^{N,M}_{\epsilon}(t),\ z^{N,M}_{\epsilon}(t_l))dt \ - \ \Delta(\epsilon)\tilde{g}(\mu^{N,M}(t_l), z^{N,M}_{\epsilon}(t_l))||$$

$$= \Delta(\epsilon)[\ (\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_0^{\frac{\Delta(\epsilon)}{\epsilon}} g(u^{N,M}_{z^{N,M}(t_l)}(\tau),\ y^{N,M}_{z^{N,M}(t_l)}(\tau, y^{N,M}_{\epsilon}(t_l)),\ z^{N,M}_{\epsilon}(t_l))d\tau$$

$$- \ \tilde{g}(\mu^{N,M}(t_l), z^{N,M}_{\epsilon}(t_l))] \leq \ \Delta(\epsilon)\bar{\bar{\delta}}_g(\frac{\Delta(\epsilon)}{\epsilon}), \quad (6.1.35)$$

where $\tau = \frac{t-t_l}{\epsilon}$ and $y^{N,M}_{z^{N,M}(t_l)}(\tau, y^{N,M}_{\epsilon}(t_l)) = y^{N,M}_{\epsilon}(t_l + \epsilon\tau)$. From (6.1.33) it follows (see Proposition 5.1 in [52])

$$||z^{N,M}_{\epsilon}(t_l) - \bar{\bar{z}}(t_l)|| \leq \kappa_4(\epsilon, \mathcal{T}), \quad l = l_0,\ l_0+1, ...,\ l_0 + \lfloor \frac{\mathcal{T}}{\Delta(\epsilon)} \rfloor, \qquad \lim_{\epsilon\to 0} \kappa_4(\epsilon) = 0.$$

This (due to (6.1.34)) leads to

$$\max_{t\in[\mathcal{T}_0, \mathcal{T}_0+\mathcal{T}]} ||z^{N,M}_{\epsilon}(t) - \bar{\bar{z}}(t)|| \ \leq \ \kappa_5(\epsilon, \mathcal{T}), \qquad \text{where} \qquad \lim_{\epsilon\to 0} \kappa_5(\epsilon, \mathcal{T}) = 0, \quad (6.1.36)$$

and, in particular, to

$$\max_{t\in[0,\mathcal{T}_0]} ||z^{N,M}_{\epsilon}(t) - \bar{z}^{N,M}(t)|| \ \leq \ \kappa_5(\epsilon, \mathcal{T}_0) \quad (6.1.37)$$

(since, by definition, $z^{N,M}_{\epsilon}(0) = z^{N,M}(0)$ and $\bar{z}^{N,M}(0) = z^{N,M}(0)$). Assume that $\mathcal{T}_0$ is chosen in such a way that (6.1.23) is satisfied and denote by $\bar{\bar{z}}_1(t)$ the solution of the system (6.1.19) considered on the interval $[\mathcal{T}_0, 2\mathcal{T}_0]$ with the initial condition

$\bar{\bar{z}}_1(\mathcal{T}_0) = z_\epsilon^{N,M}(\mathcal{T}_0)$. By (6.1.5) and (6.1.37),

$$||\bar{\bar{z}}_1(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)|| \leq a||z_\epsilon^{N,M}(\mathcal{T}_0) - \bar{z}^{N,M}(\mathcal{T}_0)|| \leq a\kappa_5(\epsilon,\mathcal{T}_0). \tag{6.1.38}$$

Also, by (6.1.36),

$$||z_\epsilon^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)|| \;\leq\; ||z_\epsilon^{N,M}(2\mathcal{T}_0) - \bar{\bar{z}}_1(2\mathcal{T}_0)|| \;+\; ||\bar{\bar{z}}_1(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)||$$

$$\leq\; \kappa_5(\epsilon,\mathcal{T}_0) \;+\; a\kappa_5(\epsilon,\mathcal{T}_0) \;\leq\; \frac{\kappa_5(\epsilon,\mathcal{T}_0)}{1-a}.$$

Continuing in a similar way, one can prove that, for any $k = 1, 2, \ldots$,

$$||z_\epsilon^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)|| \;\leq\; (1 + a + \ldots + a^{k-1})\kappa_5(\epsilon,\mathcal{T}_0) \;\leq\; \frac{\kappa_5(\epsilon,\mathcal{T}_0)}{1-a}. \tag{6.1.39}$$

Denote by $\bar{\bar{z}}_k(t)$ the solution of the system (6.1.19) considered on the interval $[k\mathcal{T}_0, (k+1)\mathcal{T}_0]$ with the initial condition $\bar{\bar{z}}_k(k\mathcal{T}_0) = z_\epsilon^{N,M}(k\mathcal{T}_0)$. By (6.1.5) and (6.1.37),

$$\max_{t\in[k\mathcal{T}_0,(k+1)\mathcal{T}_0]} ||z_\epsilon^{N,M}(t) - \bar{z}^{N,M}(t)|| \;\leq\; \max_{t\in[k\mathcal{T}_0,(k+1)\mathcal{T}_0]} ||z_\epsilon^{N,M}(t) - \bar{\bar{z}}_k(t)||$$

$$+ \max_{t\in[k\mathcal{T}_0,(k+1)\mathcal{T}_0]} ||\bar{\bar{z}}_k(t) - \bar{z}^{N,M}(t)|| \leq\; \kappa_5(\epsilon,\mathcal{T}_0) \;+ c_3||z_\epsilon^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)||.$$

Thus, by (6.1.39),

$$\sup_{t\in[0,\infty)} ||z_\epsilon^{N,M}(t) - \bar{z}^{N,M}(t)|| \;\leq\; \kappa_5(\epsilon,\mathcal{T}_0) \;+ c_3\frac{\kappa_5(\epsilon,\mathcal{T}_0)}{1-a} \stackrel{\text{def}}{=} \kappa_6(\epsilon), \qquad \lim_{\epsilon\to 0}\kappa_6(\epsilon) = 0. \tag{6.1.40}$$

From (6.1.25) and (6.1.40) it also follows that $\forall \mathcal{T} \geq 1$

$$|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt|$$

$$\leq\; |\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), z^{N,M}(t))dt|$$

$$+\; L(\kappa_2(\epsilon) + \kappa_6(\epsilon)) \leq\; \frac{1}{\mathcal{T}}\sum_{l=0}^{\lfloor\frac{\mathcal{T}}{\Delta(\epsilon)}\rfloor} |\int_{t_l}^{t_{l+1}} q(u_{z^{N,M}(t_l)}^{N,M}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z^{N,M}(t_l))dt$$

$$-\int_{t_l}^{t_{l+1}} \tilde{q}(\mu^{N,M}(t_l), z^{N,M}(t_l))dt| \;+\; L(\kappa_2(\epsilon)+\kappa_6(\epsilon)) \;+\; 2K\Delta(\epsilon) \;+\; 2LL_3\Delta(\epsilon), \tag{6.1.41}$$

where L and K are as in (6.1.26) and it has been taking into account that $\max_{t\in[t_l,t_{l+1}]} ||z^{N,M}(t) - z^{N,M}(t_l)|| \leq\; L_3\Delta(\epsilon)$, with L_3 being the same constant as in (6.1.34).

Similarly to (6.1.35), one can obtain (using (6.1.7))

$$\Big|\Big| \int_{t_l}^{t_{l+1}} q\Big(u^{N,M}_{z^{N,M}(t_l)}\Big(\frac{t-t_l}{\epsilon}\Big),\ y^{N,M}_\epsilon(t),\ z^{N,M}(t_l)\Big)dt\ -\ \Delta(\epsilon)\ \tilde{q}\Big(\mu^{N,M}(t_l), z^{N,M}(t_l)\Big)\Big|\Big|$$

$$= \Delta(\epsilon)\Big[\ \Big(\frac{\Delta(\epsilon)}{\epsilon}\Big)^{-1}\quad \int_0^{\frac{\Delta(\epsilon)}{\epsilon}} q\Big(u^{N,M}_{z^{N,M}(t_l)}(\tau),\ y^{N,M}_{z^{N,M}(t_l)}(\tau, y^{N,M}_\epsilon(t_l)),\ z^{N,M}(t_l)\Big)dt$$

$$- \tilde{q}\Big(\mu^{N,M}(t_l), z^{N,M}(t_l)\Big)\Big] \leq\ \Delta(\epsilon)\bar{\bar{\delta}}_q\Big(\frac{\Delta(\epsilon)}{\epsilon}\Big), \qquad .$$

The latter along with (6.1.41) imply that $\forall \mathcal{T} \geq 1$

$$|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u^{N,M}_\epsilon(t), y^{N,M}_\epsilon(t), z^{N,M}_\epsilon(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt|\ \leq\ \kappa_7(\epsilon), \tag{6.1.42}$$

where

$$\kappa_7(\epsilon) \stackrel{\text{def}}{=}\ L(\kappa_2(\epsilon)+\kappa_6(\epsilon))\ +\ 2K\Delta(\epsilon)\ +\ 2LL_3\Delta(\epsilon)\ +\ \bar{\bar{\delta}}_q(\frac{\Delta(\epsilon)}{\epsilon}), \qquad \lim_{\epsilon\to 0}\kappa_7(\epsilon) = 0.$$

Hence, by (6.1.32), $\forall \mathcal{T} \geq 1$

$$|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u^{N,M}_\epsilon(t), y^{N,M}_\epsilon(t), z^{N,M}_\epsilon(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{q}(\mu^{N,M}(t), z^{N,M}(t))dt|\ \leq\ \kappa_3(\epsilon)\ +\ \kappa_7(\epsilon), \tag{6.1.43}$$

and, consequently,

$$|\liminf_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u^{N,M}_\epsilon(t), y^{N,M}_\epsilon(t), z^{N,M}_\epsilon(t))dt - \tilde{V}^{N,M}|\ \leq\ \kappa_3(\epsilon)\ +\ \kappa_7(\epsilon),$$

$$\Rightarrow\qquad \lim_{\epsilon\to 0}\liminf_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} q(u^{N,M}_\epsilon(t), y^{N,M}_\epsilon(t), z^{N,M}_\epsilon(t))dt = \tilde{V}^{N,M} = \tilde{G}^* + \beta(N,M)$$

(see (5.5.25) and (5.5.26)). Due to (4.3.19), the latter proves the $\beta(N,M)$-asymptotic near optimality of the control $u^{N,M}_\epsilon(\cdot)$. Also, the estimate (6.1.15) follows from (6.1.25) and (6.1.40).

Using an arbitrary Lipschitz continuous function $h(u,y,z)$ instead of $q(u,y,z)$, one can obtain (similarly to (6.1.43)), $\forall \mathcal{T} \geq 1$

$$|\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} h(u^{N,M}_\epsilon(t), y^{N,M}_\epsilon(t), z^{N,M}_\epsilon(t))dt - \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t))dt|\ \leq\ \kappa_8(\epsilon), \tag{6.1.44}$$

where $\lim_{\epsilon\to\infty} \kappa_8(\epsilon) = 0$. If the triplets $(u_\epsilon^{N,M}(\cdot), y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ generates the occupational measure $\gamma_\epsilon^{N,M}$, then (see (1.1.20))

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt = \int_{U\times Y\times Z} h(u,y,z)\gamma_\epsilon^{N,M}(du,dy,dz).$$

Hence, passing to the limit in (6.1.44) with $\mathcal{T}\to\infty$ and taking into account (5.5.11), one obtains

$$|\int_{U\times Y\times Z} h(u,y,z)\gamma_\epsilon^{N,M}(du,dy,dz) \ - \ \int_F \tilde{h}(\mu,z)\bar{\xi}^{N,M}(d\mu,dz)| \ \leq \ \kappa_8(\epsilon).$$

By the definition of the map $\Phi(\cdot)$ (see (4.3.4)), the latter implies that

$$|\int_{U\times Y\times Z} h(u,y,z)\gamma_\epsilon^{N,M}(du,dy,dz) \ - \ \int_{U\times Y\times Z} h(u,y,z)\Phi(\bar{\xi}^{N,M})(du,dy,dz)| \ \leq \ \kappa_8(\epsilon),$$

which, in turn, implies (6.1.16). This completes the proof. $\square$

Note that from (5.5.29) and (6.1.16) it follows (due to continuity of $\Phi(\cdot)$) that

$$\rho(\gamma_\epsilon^{N,M}, \Phi(\xi^{N,M})) \ \leq \ \kappa(\epsilon) \ + \ \theta(N,M), \quad \text{where} \quad \lim_{N\to\infty} \limsup_{M\to\infty} \theta(N,M) = 0 \tag{6.1.45}$$

and where $\xi^{N,M}$ is an arbitrary optimal solution of the (N,M)-approximating averaged problem (5.3.4). This problem always has an optimal solution that can be presented in the form (see Section 5.6)

$$\xi^{N,M} \stackrel{\text{def}}{=} \sum_{k=1}^{K} \xi_k \delta_{(\mu_k,z_k)}, \qquad \sum_{k=1}^{K} \xi_k = 1, \qquad \xi_k > 0, \ k = 1, ..., K, \tag{6.1.46}$$

where

$$\mu_k = \sum_{j=1}^{J_k} b_j^k \delta_{(u_j^k,y_j^k)}, \qquad \sum_{j=1}^{J_k} b_j^k = 1, \qquad b_j^k > 0, \ j = 1, ..., J_k, \tag{6.1.47}$$

$\delta_{(u_j^k,y_j^k)}$ being the Dirac measure concentrated at $(u_j^k, y_j^k) \in U\times Y$ $(j = 1, ..., J_k)$ and $\delta_{(\mu_k,z_k)}$ being the Dirac measure concentrated at $(\mu_k, z_k) \in \mathcal{P}(U\times Y)\times Z$ $(k = 1, ...K)$. As can be readily verified,

$$\Phi(\sum_{k=1}^{K} \xi_k \delta_{(\mu_k,z_k)}) = \sum_{k=1}^{K} \sum_{j=1}^{J_k} \xi_k b_j^k \delta_{(u_j^k,y_j^k,z_k)}, \tag{6.1.48}$$

where $\delta_{(u_j^k, y_j^k, z_k)}$ is the Dirac measure concentrated at (u_j^k, y_j^k, z_k). Thus, by (6.1.45),

$$\rho(\gamma_\epsilon^{N,M}, \sum_{k=1}^{K}\sum_{j=1}^{J_k} \xi_k b_j^k \delta_{(u_j^k, y_j^k, z_k)}) \ \leq \ \kappa(\epsilon) \ + \ \theta(N, M). \tag{6.1.49}$$

That is, for N, M large enough and ϵ small enough, the occupational measure $\gamma_\epsilon^{N,M}$ is approximated by a convex combination of the Dirac measures, which implies, in particular, that the state trajectory $(y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ spends a non-zero proportion of time in a vicinity of each of the points (u_j^k, y_j^k, z_k).

6.2 LP based algorithm for solving (N, M) - approximating problems.

For convenience, let us outline the algorithm for solving SILP problems discussed above (see Section 3.2) using slightly different notations. To this end, consider the problem

$$\min_{p\in\Omega}\{\int_X h_0(x)p(dx)\} \stackrel{\text{def}}{=} \sigma^*, \tag{6.2.1}$$

where $$\Omega \stackrel{\text{def}}{=} \{p \in \mathcal{P}(X) \ : \ \int_X h_i(x)p(dx) = 0, \quad i = 1, ..., K\}, \tag{6.2.2}$$

with X being a non-empty compact metric space and with $h_i(\cdot) : X \to I\!R^1$, $i = 0, 1, ..., K$, being continuous functional on X. Note that the problem dual with respect to (6.2.1) is the problem (5.4.1), and we assume that the inequality (5.4.3) is valid only with $v_i = 0$, $i = 1, ..., K$ (which, by Lemma 5.4.5, ensures the existence of a solution of the problem (5.4.1)).

It is known (see, e.g., Theorems A.4 and A.5 in [95]) that among the optimal solutions of the problem (6.2.1) there exists one that is presented in the form

$$p^* = \sum_{l=1}^{K+1} p_l^* \delta_{x_l^*}, \qquad \text{where} \qquad p_l^* \geq 0, \qquad \sum_{l=1}^{K+1} p_l^* = 1,$$

where $\delta_{x_l^*}$ are Dirac measures concentrated at $x_l^* \in X$, $l = 1, ..., K+1$. Having in mind this presentation, let us consider the following algorithm for finding optimal concentration points $\{x_l^*\}$ and optimal weights $\{p_l^*\}$. Let points $\{x_l \in X \ , \ l = 1, ..., L\}$ $(L \geq K+1)$ be chosen to define an initial grid $\mathcal{X}_0$ on X

$$\mathcal{X}_0 = \{x_l \in X \ , \ l = 1, ..., L\}.$$

At every iteration a new point is defined and added to this set. Assume that after J iterations the points $x_{L+1}, ..., x_{L+J}$ have been defined and the set $\mathcal{X}_J$ has been constructed. Namely,

$$\mathcal{X}_J = \{x_l \in X \ , \ l = 1, ..., L + J\}.$$

The iteration $J + 1$ ($J = 0, 1, ...$) is described as follows:

(i) Find a basic optimal solution $p^J = \{p_l^J\}$ of the LP problem

$$\min_{p \in \Omega_J} \{\sum_{l=1}^{L+J} p_l h_0(x_l)\} \stackrel{\text{def}}{=} \sigma^J, \tag{6.2.3}$$

where

$$\Omega_J \stackrel{\text{def}}{=} \{p \ : \ p = \{p_l\} \geq 0 \ , \ \sum_{l=1}^{L+J} p_l = 1, \quad \sum_{l=1}^{L+J} p_l h_i(x_l) = 0 \ , \ i = 1, ..., K\}. \tag{6.2.4}$$

Note that no more than $K + 1$ components of p^J are positive, these being called basic components. Also, find an optimal solution $\lambda^J = \{\lambda_0^J, \ \lambda_i^J, i = 1, ..., K\}$ of the problem dual with respect to (6.2.3). The latter being of the form

$$\max\{\lambda_0 \ : \ \lambda_0 \leq h_0(x_l) + \sum_{i=1}^{K} \lambda_i h_i(x_l) \ \ \forall \ l = 1, ..., K + J\}; \tag{6.2.5}$$

(ii) Find an optimal solution x_{L+J+1} of the problem

$$\min_{x \in X} \{h_0(x) + \sum_{i=1}^{K} \lambda_i^J h_i(x)\} \stackrel{\text{def}}{=} a^J; \tag{6.2.6}$$

(iii) Define the set $\mathcal{X}_{J+1}$ by the equation

$$\mathcal{X}_{J+1} = \mathcal{X}_J \cup \{x_{L+J+1}\}.$$

As has been established in Section 3.3, if $a^J \geq \lambda_0^J$, then $\sigma^J = \sigma^*$ and the measure $\sum_{l \in I^J} p_l^J \delta_{x_l}$ (where I^J stands for the index set of basic components of p^J) is an optimal solution of the problem (6.2.1), with $\lambda^J \stackrel{\text{def}}{=} \{\lambda_i^J, \ i = 1, ..., K\}$ being an optimal solution of the problem (5.4.1). If $a^J < \lambda_0^J$, for $J = 1, 2, ...$, then, under some non-degeneracy assumptions, $\lim_{J \to \infty} \sigma^J = \sigma^*$, and any cluster (limit) point of the set of measures $\{\sum_{l\in I^J} p_l^J \delta_{x_l},\ J = 1, 2, ...\}$ is an optimal solution of the problem (6.2.1), while any cluster (limit) point of the set $\{\lambda^J,\ J = 1, 2, ...\}$ is an optimal solution of the problem (5.4.1) (see Theorem 3.3.1).

The (N, M) approximating problem (5.3.4) is a special case of the problem (6.2.1) with an obvious correspondence between the notations:

$$x = (\mu, z), \qquad X = F_M, \quad p = \xi, \quad \Omega = \tilde{W}^{N,M}, \quad K = N,$$

$$h_0(x) = \tilde{q}(\mu, z), \qquad h_i(x) = (\nabla\psi_i(z))^T \tilde{g}(\mu, z), \quad i = 1, ..., N.$$

Assume that the set

$$\mathcal{X}_J = \{(\mu_l, z_l) \in F_M\ ,\ l = 1, ..., L + J\} \tag{6.2.7}$$

has been constructed. The LP problem (6.2.3) takes in this case the form

$$\min_{\xi\in\tilde{W}_J^{N,M}} \{\sum_{l=1}^{L+J} \xi_l \tilde{q}(\mu_l, z_l)\} \stackrel{\text{def}}{=} \tilde{G}^{N,M,J}, \tag{6.2.8}$$

where

$$\tilde{W}_J^{N,M} \stackrel{\text{def}}{=} \{\xi\ :\ \xi = \{\xi_l\} \geq 0\ , \quad \sum_{l=1}^{L+J} \xi_l = 1, \qquad \sum_{l=1}^{L+J} \xi_l[(\nabla\psi_i(z))^T \tilde{g}(\mu_l, z_l)] = 0\ ,$$

$$i = 1, ..., N\},$$

with the corresponding dual being of the form

$$\max\{\lambda_0\ :\ \lambda_0 \leq \tilde{q}(\mu_l, z_l) + \sum_{i=1}^{N} \lambda_i[(\nabla\psi_i(z))^T \tilde{g}(\mu_l, z_l)] \quad \forall\ l = 1, ..., K + J\}. \tag{6.2.9}$$

Denote by $\xi^{N,M,J} = \{\xi_l^{N,M,J}\}$ an optimal basic solution of the problem (6.2.8) and by $\{\lambda_0^{N,M,J},\ \lambda_i^{N,M,J}, i = 1, ..., N\}$ an optimal solution of the dual problem (6.2.9). The problem (6.2.6) identifying the point to be added to the set $\mathcal{X}_J$ takes the following form

$$\min_{(\mu,z)\in F_M} \{\tilde{q}(\mu, z) + \sum_{i=1}^{N} \lambda_i^{N,M,J} (\nabla\psi_i(z))^T \tilde{g}(\mu, z)]\} = \min_{z\in Z}\{\tilde{\mathcal{G}}^{N,M,J}(z)\}, \tag{6.2.10}$$

where

$$\tilde{\mathcal{G}}^{N,M,J}(z) \stackrel{\text{def}}{=} \min_{\mu \in W_M(z)} \{ \int_{U \times Y} [q(u,y,z) + \sum_{i=1}^{N} \lambda_i^{N,M,J} (\nabla \psi_i(z))^T g(u,y,z)] \mu(du,dy) \}. \tag{6.2.11}$$

Note that the problem (6.2.11) is also a special case of the problem (6.2.1) with

$$x = (u,y), \qquad X = U \times Y, \qquad p = \mu, \quad \Omega = W_M(z), \qquad K = M,$$

$$h_0(x) = q(u,y,z) + \sum_{i=1}^{N} \lambda_i^{N,M,J} (\nabla \psi_i(z))^T g(u,y,z),$$

$$h_i(x) = (\nabla \phi_i(y))^T f(u,y,z), \quad i = 1, ..., M.$$

Its optimal solution as well as an optimal solution of the corresponding dual problem can be found with the help of the same approach. Denote the latter as $\mu_z^{N,M,J}$ and $\{\alpha_{z,0}^{N,M,J}, \alpha_{z,i}^{N,M,J}, \ i = 1, ..., M\}$, respectively. By adding the point $(\mu_{z^*}^{N,M,J}, z^*)$ to the set $\mathcal{X}_J$ (z^* being an optimal solution of the problem in the right-hand side of (6.2.10)), one can define the set $\mathcal{X}_{J+1}$ and then proceed to the next iteration.

Under the controllability conditions introduced in Section 5.4 (see Assumptions 5.4.1 and 5.4.3) and under additional (simplex method related) non-degeneracy conditions similar to those used in Section 3.3, it can be proved that the optimal value of the problem (6.2.8) converges to the optimal value of the (N, M)-approximating averaged problem

$$\lim_{J \to \infty} \tilde{G}^{N,M,J} = \tilde{G}^{N,M} \tag{6.2.12}$$

and that, if $\lambda^{N,M} = \{\lambda_i^{N,M}, i = 1, ..., N\}$ is a cluster (limit) point of the set of optimal solutions $\lambda^{N,M,J} = \{\lambda_i^{N,M,J}, i = 1, ..., N\}$ of the problem (6.2.9) considered with $J = 1, 2, ...,$ then

$$\zeta^{N,M}(z) \stackrel{\text{def}}{=} \sum_{i=1}^{N} \lambda_i^{N,M} \psi_i(z)$$

is an optimal solution of the (N, M)-approximating averaged dual problem (5.3.20). In addition to this, it can be shown that, if $\alpha_z^{N,M} = \{\alpha_{z,i}^{N,M}, i = 1, ..., M\}$ is a cluster (limit) point of the set of optimal solutions $\alpha_z^{N,M,J} = \{\alpha_{z,i}^{N,M,J}, i = 1, ..., M\}$ of the problem dual to (6.2.11) considered with $J = 1, 2, ...,$ then

$$\eta_z^{N,M}(y) \stackrel{\text{def}}{=} \sum_{i=1}^{M} \alpha_{z,i}^{N,M} \phi_i(y)$$

is an optimal solution of the (N, M)-approximating associated dual problem (5.3.25).

S. Rossomakhine developed a software that implements this algorithm on the basis of the IBM ILOG CPLEX LP solver and global nonlinear optimization routines designed by A. Bagirov and M. Mammadov has been developed (with the CPLEX solver being used for finding optimal solutions of the LP problems involved and Bagirov's and Mammadov's routines being used for finding optimizers in (6.2.10) and in problems similar to (6.2.6) that arise when solving (6.2.11)). A numerical solution of the SP optimal control problem introduced in Section 6.3 was obtained with the help of this software.

Remark 6.2.1 *The decomposition of the problem (6.2.6), an optimal solution of which identifies the point to be added to the set $\mathcal{X}_J$, into problems (6.2.10) and (6.2.11) resembles the column generating technique of generalized linear programming (see [37]).*

6.3 Numerical example.

To illustrate the construction of asymptotically near optimal controls, let us consider the optimal control problem

$$\inf_{(u(\cdot),y_\epsilon(\cdot),z_\epsilon(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} (0.1u_1^2(t) + 0.1u_2^2(t) - z_1^2(t))dt = V^*(\epsilon), \tag{6.3.1}$$

where minimization is over the controls $u(\cdot) = (u_1(\cdot), u_2(\cdot))$,

$$(u_1(t), u_2(t)) \in U \stackrel{\text{def}}{=} \{(u_1, u_2) \ : \ |u_i| \leq 1, \ i = 1, 2\}, \tag{6.3.2}$$

and the corresponding solutions $y_\epsilon(\cdot) = (y_{1,\epsilon}(\cdot), y_{2,\epsilon}(\cdot))$ and $z_\epsilon(\cdot) = (z_{1,\epsilon}(\cdot), z_{2,\epsilon}(\cdot))$) of the SP system

$$\epsilon y_i'(t) = -y_i(t) + u_i(t), \quad i = 1, 2, \tag{6.3.3}$$

$$z_1'(t) = z_2(t), \qquad z_2'(t) = -4z_1(t) - 0.3z_2(t) - y_1(t)u_2(t) + y_2(t)u_1(t), \tag{6.3.4}$$

with

$$(y_1(t), y_2(t)) \in Y \stackrel{\text{def}}{=} \{(y_1, y_2) \ : \ |y_i| \leq 1, \ i = 1, 2\}$$

and with

$$(z_1(t), z_2(t)) \in Z \stackrel{\text{def}}{=} \{(z_1, z_2) \ : \ |z_1| \leq 2.5, \quad |z_2| \leq 4.5\}.$$

The averaged system (4.2.8) takes in this case the form

$$z_1'(t) = z_2(t), \qquad z_2'(t) = -4z_1(t) + \int_{U\times Y} (-y_1u_2 + y_2u_1)\mu(t)(du, dy), \tag{6.3.5}$$

where

$$\mu(t) \in W \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U\times Y) \ : \ \int_{U\times Y} \Big[\frac{\partial\phi(y)}{\partial y_1}(-y_1+u_1) + \frac{\partial\phi(y)}{\partial y_2}(-y_2+u_2)\Big]\mu(du,dy) = 0$$

$$\forall\phi(\cdot) \in C^1(I\!\!R^2)\}, \tag{6.3.6}$$

Note that, as can be readily verified, the function $f(u,y) = (-y_1 + u_1,\ -y_2 + u_2)$ satisfies Assumption 6.1.2(i) and the function $\tilde{g}(\mu,z) = (z_2,\ -4z_1 + \int_{U\times Y}(-y_1u_2 + y_2u_1)\mu(du,dy)\ )$ satisfies Assumption 6.1.2(ii), with A_1 and A_2 being equal to the identity matrix.

The (N,M)-approximating averaged problem (5.3.4) was constructed in this example with the use of the monomials $z_1^{j_1}z_2^{j_2}$ $(1 \leq j_1 + j_2 \leq 5)$ as the test functions in (5.3.3) and the monomials $y_1^{i_1}y_2^{i_2}$ $(1 \leq i_1 + i_2 \leq 5)$ as the test functions in (5.3.1). Note that $N, M = 35$ in this case (recall that N stands for the number of constraints in (5.3.3) and M stands for the number of constraints in (5.3.1)). This problem was solved numerically with the help of the linear programming based algorithm described in the previous section, its output including the optimal value of the problem, an optimal solution of the problem and solutions of the corresponding averaged and associated dual problems.

The optimal value of the problem was obtained to be approximately equal to -1.186:

$$\tilde{G}^{35,35} \approx -1.186. \tag{6.3.7}$$

Along with the optimal value, the points

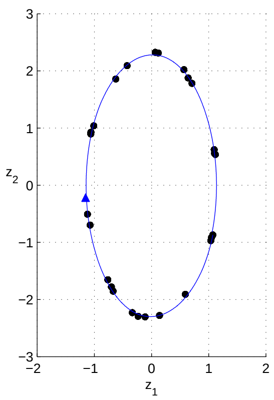
$$z_k = (z_{1,k}, z_{2,k}) \in Z, \qquad k = 1, ..., K, \tag{6.3.8}$$

and weights $\{\xi_k\}$ that enter the expansion (6.1.46) as well as the points

$$u_j^k = (u_{1,j}^k, u_{2,j}^k) \in U, \qquad y_j^k = (y_{1,j}^k, y_{2,j}^k) \in Y, \qquad j = 1, ..., J_k, \qquad k = 1, ..., K, \tag{6.3.9}$$

and the corresponding weights $\{q_j^k\}$ that enter the expansions (6.1.47) were numerically found. Below in Figure 1, the points $\{z_k\}$ that enter the expansion (6.1.46) are marked with dotes on the "z-plane". Corresponding to each such a point z_k, there are points $\{y_j^k\}$ that enter the expansion (6.1.47). These points are marked with dots on the "y-plane" in Figure 2 for $z_k \approx (1.07, -0.87)$ (which is one of the points marked in Fig.1; for other points marked in Fig. 1, the configurations of the corresponding $\{y_j^k\}$ points look similar).

Graphs of $z_\epsilon^{35,35}(t)$ *and* $y_z^{35,35}(\tau)$



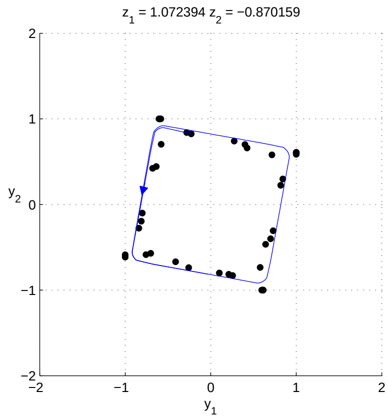


Fig. 1: $z_\epsilon^{35,35}(t) = (z_{1\epsilon}^{35,35}(t), z_{2,\epsilon}^{35,35}(t))$ Fig.2: $y_z^{35,35}(\tau) = (y_{1,z}^{35,35}(\tau), y_{2,z}^{35,35}(\tau))$

The expansions (5.3.22) and (5.3.27) that define solutions of the (N, M) - approximating averaged and (N, M) - approximating associated dual problems take the form

$$\zeta^{35,35}(z) = \sum_{1\le j_1+j_2\le 5} \lambda_{j_1,j_2}^{35,35} z_1^{j_1} z_2^{j_2}, \qquad \eta_z^{35,35}(y) = \sum_{1\le i_1+i_2\le 5} \alpha_{z,i_1,i_2}^{35,35}\, y_1^{i_1} y_2^{i_2}, \tag{6.3.10}$$

where the coefficients $\{\lambda_{j_1,j_2}^{35,35}\}$ and $\{\alpha_{z,i_1,i_2}^{35,35}\}$ are obtained as a part of the solution with the above mentioned algorithm. Using $\zeta^{35,35}(z)$ and $\eta_z^{35,35}(y)$, one can compose the problem (5.5.1):

$$\min_{u_i\in[-1,1]}\{0.1u_1^2 + 0.1u_2^2 - z_1^2 + \frac{\partial\zeta^{35,35}(z)}{\partial z_1} z_2 + \frac{\partial\zeta^{35,35}(z)}{\partial z_2}(-4z_1 - 0.3z_2 - y_1u_2 + y_2u_1) +$$
$$\frac{\partial\eta_z^{35,35}(y)}{\partial y_1}(-y_1 + u_1) + \frac{\partial\eta_z^{35,35}(y)}{\partial y_2}(-y_2 + u_2)\}, \tag{6.3.11}$$

the solution of which is written in the form

$$u_i^{35,35}(y,z) = \left\{\begin{array}{rcr} -5b_i^{35,35}(y,z) & if & |5b_i^{35,35}(y,z)| \le 1, \\ -1 & if & -5b_i^{35,35}(y,z) < -1, \\ 1 & if & -5b_i^{35,35}(y,z) > 1, \end{array}\right\}, \quad i = 1,2, \tag{6.3.12}$$

where $b_1^{35,35}(y,z) \stackrel{\text{def}}{=} \frac{\partial\zeta^{35,35}(z)}{\partial z_2} y_2 + \frac{\partial\eta_z^{35,35}(y)}{\partial y_1}$ and $b_2^{35,35}(y,z) \stackrel{\text{def}}{=} -\frac{\partial\zeta^{35,35}(z)}{\partial z_2} y_1 + \frac{\partial\eta_z^{35,35}(y)}{\partial y_2}$.

Using the feedback controls $u_i^{35,35}(y,z)$, $i = 1,2$, with fixed $z = z_k \approx (1.07, -0.87)$ and integrating the associated system with MATLAB from the initial conditions defined by one of the points marked in Figure 2, one obtains a periodic solution $y_z^{35,35}(\tau) =$

$(y_{1,z}^{35,35}(\tau), y_{2,z}^{35,35}(\tau))$. The corresponding square like state trajectory of the associated system is also depicted in Figure 2. Note that this trajectory is located in a close vicinity of the marked points, this being consistent with the comments made after the statement of Theorem 6.1.3.

Using the same controls $u_i^{35,35}(y,z)$, $i = 1, 2$, in the SP system (6.3.3)-(6.3.4) and integrating the latter (taken with $\epsilon = 0.01$ and $\epsilon = 0.001$) with MATLAB from the initial conditions defined by one of the points marked in Figure 1 and one of the points marked in Figure 2, one obtains visibly periodic solutions, the images of which are depicted in Figures 3 and 4, with the state trajectory of the slow dynamics $z_\epsilon^{35,35}(t) = (z_{1\epsilon}^{35,35}(t), z_{2,\epsilon}^{35,35}(t))$ being also depicted in Figure 1. The slow z-components appear to be moving periodically along an ellipse like figure on the plane (z_1, z_2), with the period being approximately equal to 3.16. Note that this figure and the period appear to be the same for $\epsilon = 0.01$ and for $\epsilon = 0.001$, with the marked points being located on or very close to the ellipse like figure in Fig. 1 (which again is consistent with the comments made after Theorem 6.1.3). In Figures 3 and 4, the fast y-components are moving along square like figures (similar to that in Fig. 2) centered around the points on the "ellipse", with about 50 rounds for the case $\epsilon = 0.01$ (Fig. 3) and about 500 rounds for the case $\epsilon = 0.001$ (Fig. 4). The values of the objective functions obtained for these two cases are approximately the same and ≈ -1.177, the latter being close to the value of $\tilde{G}^{35,35}$ (see (6.3.7)). Due to (4.3.19) and due to (5.3.8), this indicates that the found solution is close to the optimal one.

Images of the state trajectories of the SP system for $\epsilon = 0.01$ *and* $\epsilon = 0.001$

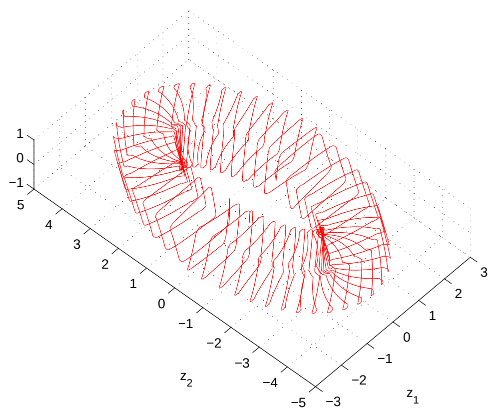


Fig. 3: $(y_\epsilon^{35,35}(t), z_\epsilon^{35,35}(t))$ for $\epsilon = 0.01$

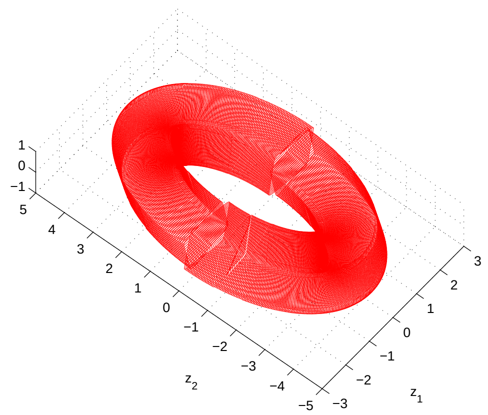


Fig.4: $(y_\epsilon^{35,35}(t), z_\epsilon^{35,35}(t))$ for $\epsilon = 0.001$

Note, in conclusion, that by taking $\epsilon = 0$ in (6.3.3), one obtains $y_i(t) = u_i(t),\ i = 1, 2$, and, thus, arrives at the equality

$$-y_1(t)u_2(t) + y_2(t)u_1(t) = 0 \quad \forall t,$$

which makes the slow dynamics uncontrolled and leads to the optimality of the "trivial" steady state regime: $u_1(t) = u_2(t) = y_1(t) = y_2(t) = z_1(t) = z_2(t) = 0 \ \ \forall t$ implying that $V^*(0) = 0$. Thus, in the present example

$$\lim_{\epsilon \to 0} V^*(\epsilon) \approx -1.177 \ < \ 0 \ = \ V^*(0).$$

6.4 Additional comments for Chapter 6

Consideration of Sections 6.1 and 6.3 is based on [61]. The LP based algorithm of Section 6.2 was originally described in [64] for SP optimal control problems with time discounting. Note that problems of optimal control of singularly perturbed systems can be very difficult to tackle with the help of traditional optimization techniques (due to the fact that they may be very ill-conditioned for small values of the singular perturbation parameter ϵ). The proposed numerical technique deals with "limit" problems that are independent of ϵ. This allows one to overcome difficulties caused by the ill-conditioning.

Part III

Perturbations of semi-infinite dimensional linear programming problems

7

Regularly and singularly perturbed semi-infinite dimensional linear programming problems

In Part III and, in particular, in this chapter we consider a family of semi-infinite dimensional linear programming (SILP) problems depending on a small parameter ϵ. Recall that the *perturbed problems* (that is, the problems that depends on small parameter ϵ) can be of two types: regularly perturbed and singularly perturbed. The family of SILP problems is called *regularly perturbed* if its optimal value is continuous at $\epsilon = 0$. The family is called *singularly perturbed* if its optimal value is discontinuous at $\epsilon = 0$. The chapter consists of two sections, Sections 7.1 and 7.2, devoted to regularly and singularly perturbed SILP problems respectively. In Section 7.1, we state some regularity condition under which it is establishes that the family is regularly perturbed.

7.1 Regularly perturbed SILP problems.

Consider a family of semi-infinite dimensional linear programming (SILP) problems depending on a small parameter ϵ. Namely

$$\min_{\gamma \in W_N(\epsilon)} \int_X q(x)\gamma(dx) \stackrel{def}{=} G^N(\epsilon), \tag{7.1.1}$$

where

$$W_N(\epsilon) \stackrel{def}{=} \Big\{\gamma : \gamma \in \mathcal{P}(X), \int_X [h_i^0(x) + \epsilon h_i^1(x)]\gamma(dx) = 0, \quad i = 1, ..., N\Big\}, \tag{7.1.2}$$

where X is a compact metric spaces and $q(x), h_i^0(x), h_i^1(x)$ are continuous functions on X.

Consider also, a SILP problem obtained from the above with $\epsilon = 0$. That is,

$$\min_{\gamma \in W_N(0)} \int_X q(x)\gamma(dx) \stackrel{def}{=} G^N(0), \tag{7.1.3}$$

where

$$W_N(0) \stackrel{def}{=} \Big\{\gamma : \gamma \in \mathcal{P}(X), \quad \int_X h_i^0(x)\gamma(dx) = 0, \quad i = 1, 2, ..., N\Big\}. \tag{7.1.4}$$

The family of problems (7.1.1) is called *perturbed problem.* The problem obtained from the perturbed problem by equating ϵ to zero (problem (7.1.3)) is called *reduced problem.*

Definition 7.1.1 *The family of problems (7.1.1) will be referred to as regularly perturbed, if*

$$\lim_{\epsilon \to 0} G^N(\epsilon) = G^N(0),$$

and, it will be called singularly perturbed, if

$$\lim_{\epsilon \to 0} G^N(\epsilon) \neq G^N(0).$$

Let us define the problem dual to the perturbed problem (7.1.1) by the equation

$$D^N(\epsilon) \stackrel{def}{=} \max_{(v,d)} \Big\{d \leq q(x) + \sum_{i=1}^{N} v_i\Big(h_i^0(x) + \epsilon h_i^1(x)\Big), \quad v = (v_i) \in I\!R^N, \forall x \in X\Big\}. \tag{7.1.5}$$

Similarly, we define the dual to the reduced problem (7.1.3). Namely,

$$D^N(0) \stackrel{def}{=} \max_{(v,d)} \Big\{ d \le q(x) + \sum_{i=1}^{N} v_i h_i^0(x), \quad v = (v_i) \in I\!R^N, \forall x \in X \Big\}. \tag{7.1.6}$$

The problems (7.1.5) and (7.1.6) will be referred to as *perturbed* and *reduced* dual problems respectively.

Note that the relationships between the SILP problem (7.1.3) and its dual (7.1.6) are similar to those established in Theorem 2.3.3 (see Section 2.3). Namely, the following results are valid.

Theorem 7.1.2 *(i) The optimal value of the dual reduced problem is bounded (that is $D^N(0) < \infty$) if and only if the set $W_N(0)$ is not empty;*

(ii) If the optimal value of the dual reduced problem is bounded, then

$$D^N(0) = G^N(0). \tag{7.1.7}$$

Theorem 7.1.3 *(i) The optimal value of the dual perturbed problem is bounded (that is $D^N(\epsilon) < \infty$) if and only if the set $W_N(\epsilon)$ is not empty;*

(ii) If the optimal value of the dual perturbed problemis bounded, then

$$D^N(\epsilon) = G^N(\epsilon). \tag{7.1.8}$$

Proof. The proofs of these theorems are similar to the proof of the corresponding parts of Theorem 2.3.3. □

A vector $v^* = (v_i^*)$ $i = 1, 2, ..., N$, will be called an optimal solution of the problem (7.1.6) if

$$G^N(0) = \min_{x \in X} \Big\{ q(x) + \sum_{i=1}^{N} v_i^* h_i^0(x) \Big\} \tag{7.1.9}$$

and, a vector $v^*(\epsilon) = (v_i^*(\epsilon))$ $i = 1, 2, ..., N$, will be called an optimal solution of the problem (7.1.5) if

$$G^N(\epsilon) = \min_{x \in X} \Big\{ q(x) + \sum_{i=1}^{N} v_i^*(\epsilon) \Big(h_i^0(x) + \epsilon h_i^1(x) \Big) \Big\}. \tag{7.1.10}$$

Definition 7.1.4 *The perturbed problem will be said to satisfy the regularity condition* *if the inequality*

$$\sum_{i=1}^{N} v_i \Big(h_i^0(x) + \epsilon h_i^1(x) \Big) \geq 0, \qquad \forall x \in X, \tag{7.1.11}$$

can be valid only with $v_i = 0, \quad i = 1, ..., N.$

Definition 7.1.5 *The reduced problem will be said to satisfy the regularity condition if the inequality*

$$\sum_{i=1}^{N} v_i h_i^0(x) \geq 0, \qquad \forall x \in X, \tag{7.1.12}$$

can be valid only with $v_i = 0, \quad i = 1, ..., N.$

Proposition 7.1.6 *If the reduced problem satisfies the regularity condition, then the perturbed problem also satisfies the regularity condition for any $\epsilon > 0$ small enough.*

Proof. Assume that the perturbed problem does not satisfy the regularity condition. Then, there exists a sequence ϵ_l, $\lim\limits_{l \to \infty} \epsilon_l = 0$, such that $v(\epsilon_l) = (v_i(\epsilon_l)),\ i = 1, ..., N$ satisfies the inequality

$$\sum_{i=1}^{N} v_i(\epsilon_l) \Big(h_i^0(x) + \epsilon_l h_i^1(x) \Big) \geq 0 \ \text{ and } \ ||v(\epsilon_l)|| > 0. \tag{7.1.13}$$

Note that, without loss of generality, one may assume that

$$\lim_{\epsilon_l \to 0} \frac{v(\epsilon_l)}{||v(\epsilon_l)||} \stackrel{\text{def}}{=} \tilde{v}, \qquad ||\tilde{v}|| = 1. \tag{7.1.14}$$

Dividing (7.1.13) by $||v(\epsilon_l)||$ and passing to the limit along the sequence ϵ_l, one can obtain

$$\sum_{i=1}^{N} \tilde{v}_i h_i^0(x) \geq 0 \qquad \forall x \in X. \tag{7.1.15}$$

Hence, by the assumption of the proposition $\tilde{v} = (\tilde{v}_i) = 0$, which is in contradiction with (7.1.14). Thus, the perturbed problem satisfies the regularity condition for $\epsilon > 0$ small enough. $\square$

Proposition 7.1.7 *If the reduced problem satisfies the regularity condition, then both optimal solutions of the perturbed dual and the reduced dual problems exist.*

Proof. The proof below is the same for both perturbed and reduced dual problems and therefore we will use the following notations $h_i(x) \stackrel{\text{def}}{=} h_i^0(x) + \epsilon h_i^1(x)$ (valid for both $\epsilon > 0$ and for $\epsilon = 0$).

Let us define the set $\mathcal{V}_N$ as follows

$$\mathcal{V}_N \overset{def}{=} \{v = (v_i) : D^N \leq q(x) + \sum_{i=1}^{N} v_i \; h_i(x), \quad \forall x \in X\}. \tag{7.1.16}$$

We have to show that $\mathcal{V}_N \neq \emptyset$.

Let $v^k = (v_i^k) \in I\!R^N$ be such that

$$D^N - \frac{1}{k} \leq q(x) + \sum_{i=1}^{N} v_i^k \; h_i(x), \quad \forall x \in X, \quad k = 1, 2, \dots \, . \tag{7.1.17}$$

Show that the sequence v^k is bounded. That is,

$$\| \, v^k \, \| \leq c = const, \quad k = 1, 2, \dots \, . \tag{7.1.18}$$

In fact, if v^k, $k = 1, 2, \dots$, were not bounded, then there would exist a sequence $v^{k'}$ such that

$$\lim_{k' \to \infty} \| \, v^{k'} \, \| = \infty. \tag{7.1.19}$$

Also note,

$$\lim_{k' \to \infty} \frac{v^{k'}}{\| \, v^{k'} \, \|} \overset{def}{=} \tilde{v} \quad and \quad \| \, \tilde{v} \, \| = 1. \tag{7.1.20}$$

Dividing (7.1.17) by $\| \, v^{k'} \, \|$ and passing to the limit over the subsequence $\{k'\}$, one would obtain the following inequality:

$$0 \leq \sum_{i=1}^{N} \tilde{v} h_i(x), \quad \forall x \in X. \tag{7.1.21}$$

Due to regularity condition, the fact that (7.1.21) is valid implies that $\tilde{v} = (\tilde{v}_i) = 0$, which is in contradiction to (7.1.20). Thus (7.1.18) is true.

Due to (7.1.18), there exists a subsequence $\{k'\}$ such that there exists a limit

$$\lim_{k' \to \infty} v^{k'} \overset{def}{=} v. \tag{7.1.22}$$

Passing over this subsequence to the limit in (7.1.17), one can obtain

$$D^N \leq q(x) + \sum_{i=1}^{N} v_i h_i(x), \quad \forall x \in X \quad \Rightarrow \quad v = (v_i) \in \mathcal{V}_N. \tag{7.1.23}$$

The latter proves that $\mathcal{V}_N$ is not empty. $\square$

Proposition 7.1.8 *If the reduced problem satisfies the regularity condition then the set* $\mathcal{V}_N(\epsilon) = \{v = (v_i) : D^N(\epsilon) \le q(x) + \sum_{i=1}^N v_i(h_i^0(x) + \epsilon h_i^1(x))\}$ *of optimal solutions of the perturbed dual problem (7.1.5) is bounded. That is,*

$$\sup\{\| v \| \mid v \in \mathcal{V}_N(\epsilon)\} \le k, \quad k = const \tag{7.1.24}$$

for ϵ *small enough.*

Proof. Assume that the statement of the proposition is not true, then there exist a sequence $\epsilon_l \to 0$, such that $v_i(\epsilon_l)$ satisfies the inequality

$$D^N(\epsilon_l) \le q(x) + \sum_{i=1}^N v_i(\epsilon_l)(h_i^0(x) + h_i^1(x)), \quad \forall x \in X. \tag{7.1.25}$$

and

$$\lim_{\epsilon_l \to 0} \| v(\epsilon_l) \| = \infty. \tag{7.1.26}$$

Also, note that, without loss of generality, one may assume that

$$\lim_{\epsilon_l \to 0} \frac{v(\epsilon_l)}{\| v(\epsilon_l) \|} \stackrel{def}{=} \tilde{v} \quad and \quad \| \tilde{v} \| = 1. \tag{7.1.27}$$

Dividing (7.1.25) by $\| v(\epsilon_l) \|$ and passing to the limit as $\epsilon_l \to 0$ one can obtain

$$0 \le \sum_{i=1}^N \tilde{v}_i h_i^0(x), \quad \forall x \in X. \tag{7.1.28}$$

Due to fulfilment of the regularity condition of the reduced problem, the fact that (7.1.28) is valid implies that $\tilde{v} = (\tilde{v}_i) = 0$, which is in contradiction to (7.1.27). Thus, the validity of (7.1.24) is established. □

Proposition 7.1.9 *If the reduced problem satisfies regularity condition, then*

$$\lim_{\epsilon \to 0} G^N(\epsilon) = G^N(0) \tag{7.1.29}$$

and

$$\lim_{\epsilon \to 0} \rho_H(W_N(\epsilon), W_N(0)) = 0. \tag{7.1.30}$$

Proof. Let us divide the proof of (7.1.29) in two parts. First of all, we will prove that

(i) $\overline{\lim}_{\epsilon\to 0} G^N(\epsilon) \le G^N(0)$

and then,

(ii) $\underline{\lim}_{\epsilon\to 0} G^N(\epsilon) \ge G^N(0)$.

(i) Let a sequence $\{\epsilon_l\}$ be such that

$$\lim_{\epsilon_l\to 0} G^N(\epsilon_l) = \overline{\lim_{\epsilon\to 0}} G^N(\epsilon). \tag{7.1.31}$$

Due to Proposition 7.1.8 passing to the limit as $\epsilon_l \to 0$ in (7.1.25) and taking in consideration (7.1.8) one can obtain

$$\lim_{\epsilon_l\to 0} G^N(\epsilon_l) = \lim_{\epsilon_l\to 0} D^N(\epsilon_l) \le q(x) + \sum_{i=1}^{N} \bar{v}_i h_i^0(x), \quad \forall x \in X. \tag{7.1.32}$$

Hence,

$$\overline{\lim_{\epsilon\to 0}} G^N(\epsilon) \le q(x) + \sum_{i=1}^{N} \bar{v}_i h_i^0(x), \quad \forall x \in X. \tag{7.1.33}$$

Consequently (see 7.1.6),

$$\overline{\lim_{\epsilon\to 0}} G^N(\epsilon) \le \sup_{v} \min_{x\in X} \{q(x) + \sum_{i=1}^{N} v_i h_i^0(x)\} = D^N(0). \tag{7.1.34}$$

Thus, the first part of the proof is established and

$$\overline{\lim_{\epsilon\to 0}} G^N(\epsilon) \le G^N(0). \tag{7.1.35}$$

ii) Let a sequence $\{\epsilon_l\}$ be such that

$$\underline{\lim}_{\epsilon\to 0} G^N(\epsilon) = \lim_{l\to\infty} G^N(\epsilon_l), \tag{7.1.36}$$

and let $\gamma^*_{\epsilon_l}$ be an optimal solution of the perturbed problem (7.1.1) (it exists due to compactness of $W_N(\epsilon)$). That is,

$$G^N(\epsilon_l) \overset{def}{=} \int_X q(x) \gamma^*_{\epsilon_l}(dx), \tag{7.1.37}$$

where $\gamma^*_{\epsilon_l} \in W_N(\epsilon_l)$, that is it satisfies the equations

$$\int_X \Big(h_i^0(x) + \epsilon_l h_i^1(x) \Big) \gamma^*_{\epsilon_l}(dx) = 0. \tag{7.1.38}$$

Due to compactness of $\mathcal{P}(X)$, without loss of generality, one may assume that

$$\lim_{l\to\infty} \gamma^*_{\epsilon_l} \stackrel{def}{=} \tilde{\gamma}. \tag{7.1.39}$$

Hence, by (7.1.37) and (7.1.38)

$$\lim_{l\to\infty} G^N(\epsilon_l) = \int_X q(x)\tilde{\gamma}(dx) \tag{7.1.40}$$

and

$$\lim_{l\to\infty} \int_X \Big(h_i^0(x) + \epsilon_l h_i^1(x)\Big)\gamma^*_{\epsilon_l}(dx) = \int_X h_i^0(x)\tilde{\gamma}(dx) = 0 \tag{7.1.41}$$

and thus $\tilde{\gamma} \in W_N(0)$. Consequently,

$$\int_X q(x)\tilde{\gamma}(dx) \geq G^N(0). \tag{7.1.42}$$

Due to (7.1.36), (7.1.40) and (7.1.42)

$$\underline{\lim}_{\epsilon\to 0} G^N(\epsilon) \geq G^N(0). \tag{7.1.43}$$

From (7.1.35) and (7.1.43) follows that

$$\underline{\lim}_{\epsilon\to 0} G^N(\epsilon) = \overline{\lim}_{\epsilon\to 0} G^N(\epsilon) = G^N(0). \tag{7.1.44}$$

Hence,

$$\lim_{\epsilon\to 0} G^N(\epsilon) = G^N(0). \tag{7.1.45}$$

Since (7.1.45) is valid for any continuous functions $q(x)$ (used as integrand in (7.1.1)) the validity of (7.1.30) follows from (7.1.45)(see, e.g., [59]). $\square$

7.2 Singularly perturbed SILP problems.

Let us assume that the reduced problem does not satisfy the regularity condition. That is, there exist $v_1, v_2, ..., v_N$ such that

$$\sum_{i=1}^{N} v_i h_i^0(x) \geq 0, \qquad \forall x \in X \tag{7.2.1}$$

and

$$\|v\| \neq 0. \tag{7.2.2}$$

Assumption 7.2.1 *Assume that the following Slater Type condition is satisfied:* $\exists \gamma \in W_N(0)$ *such that for any open ball* $Q \subset I\!R^m$, $Q \bigcap X \neq \emptyset$

$$\gamma(Q \cap X) > 0. \tag{7.2.3}$$

Lemma 7.2.2 *Let Assumption 7.2.1 is satisfied. If* $v_1, v_2, ..., v_N$ *are such that (7.2.1) is true, then*

$$\sum_{i=1}^{N} v_i h_i^0(x) = 0 \quad \forall x \in X. \tag{7.2.4}$$

Proof. Assume it is not true, then there exists $\bar{x} \in X$ such that

$$\sum_{i=1}^{N} v_i h_i^0(\bar{x}) > 0. \tag{7.2.5}$$

Hence,

$$\sum_{i=1}^{N} v_i h_i^0(x) > 0, \quad \forall x \in B_r(\bar{x}), \tag{7.2.6}$$

where $B_r(\bar{x})$ is an open ball of radius r centered at $\bar{x}$ and $r > 0$ is small enough. Thus,

$$\sum_{i=1}^{N} v_i h_i^0(x) > 0, \quad \forall x \in B_r(\bar{x}) \cap X. \tag{7.2.7}$$

One can observe that the constraints in (7.1.4) can be rewritten as follows

$$\int_{B_r(\bar{x})\cap X} \sum_{i=1}^{N} v_i h_i^0(x)\gamma(dx) + \int_{X\setminus\left(B_r(\bar{x})\cap X\right)} \sum_{i=1}^{N} v_i h_i^0(x)\gamma(dx) = 0. \tag{7.2.8}$$

Note that, the integrant in the second term is greater than or equal to zero (due to (7.2.1)), whereas the integrant in the first term is strictly positive. The sum can be equal to zero only if the measure of the set $B_r(\bar{x}) \cap X$ is zero, which contradicts to Assumption 7.2.1. Thus, the statement of the Lemma 7.2.2 is proved. $\Box$

Assume that the reduced problem does not satisfy the regularity condition, but Assumption 7.2.1 is valid. Then, by Lemma 7.2.2, there exists a vector $v^1 = (v_i^1)$, $i = 1, 2, ...N$, $\|v^1\| \neq 0$ such that

$$\sum_{i=1}^{N} v_i^1 h_i^0(x) = 0, \quad \forall (x) \in X. \tag{7.2.9}$$

Observe that, by multiplying the constraints (7.1.2) by v_i^1 and summing up the resulted equations over $i = 1, 2, ..., N$ one can obtain

$$\int_X [\sum_{i=1}^N v_i^1 h_i^0(x) + \epsilon \sum_{i=1}^N v_i^1 h_i^1(x)] \gamma(dx) = 0. \tag{7.2.10}$$

Due to (7.2.9), the first term in (7.2.10) is zero. Hence, for any $\gamma \in W_N(\epsilon)$

$$\epsilon \int_X \sum_{i=1}^N v_i^1 h_i^1(x) \gamma(dx) = 0. \tag{7.2.11}$$

Since ϵ is positive it can be reduced in (7.2.11). That is,

$$\int_X \sum_{i=1}^N v_i^1 h_i^1(x) \gamma(dx) = 0. \tag{7.2.12}$$

The above equality defines the new constraint that can be added to the set of constraints $W_N(\epsilon)$ without changing the set. That is,

$$W_N(\epsilon) = W_N^{\mathcal{A}}(\epsilon) \stackrel{\text{def}}{=} \Big\{ \gamma : \gamma \in \mathcal{P}(X), \int_X \Big(h_i^0(x) + \epsilon h_i^1(x) \Big) \gamma(dx) = 0, \quad i = 1, 2, ..., N, \\ \int_X \sum_{i=1}^N v_i^1 h_i^1(x) \gamma(dx) = 0 \Big\}. \tag{7.2.13}$$

Let us assume that there exist k linearly independent vectors $v^j, j = 1, ..., k$, such that

$$\sum_{i=1}^N v_i^1 h_i^0(x) = 0, \quad \sum_{i=1}^N v_i^2 h_i^0(x) = 0, \quad \ldots, \quad \sum_{i=1}^N v_i^k h_i^0(x) = 0, \quad \forall x \in X. \tag{7.2.14}$$

By following the procedure described above, one can augment the set of constraints defining the set $W_N(\epsilon)$ (without changing the latter) as follows

$$W_N(\epsilon) = W_N^{\mathcal{A}}(\epsilon) \stackrel{\text{def}}{=} \Big\{ \gamma : \gamma \in \mathcal{P}(X), \int_X \Big(h_i^0(x) + \epsilon h_i^1(x) \Big) \gamma(dx) = 0, \\ \int_X \sum_{i=1}^N v_i^j h_i^1(x) \gamma(dx) = 0, \ i = 1, 2, ..., N, \ j = 1, 2, ..., k \Big\}. \tag{7.2.15}$$

By taking $\epsilon = 0$ in (7.2.15), one obtains the set $W_N^{\mathcal{A}}(0)$ defined as follows

$$W_N^{\mathcal{A}}(0) \stackrel{\text{def}}{=} \Big\{\gamma : \gamma \in \mathcal{P}(X), \int_X h_i^0(x)\gamma(dx) = 0,$$
$$\int_X \sum_{i=1}^{N} v_i^j h_i^1(x)\gamma(dx) = 0, \quad i = 1, 2, ..., N, \; j = 1, 2, ..., k\Big\}. \tag{7.2.16}$$

Note that, in the general case

$$W_N^{\mathcal{A}}(0) \neq W_N(0).$$

Also note that, due to (7.2.14), k constraints in the definition (7.2.16) of $W_N^{\mathcal{A}}(0)$ are redundant, and we assume that these are removed.

Consider the problem

$$\min_{\gamma \in W_N^{\mathcal{A}}(0)} \int_X q(x)\gamma(dx) \stackrel{\text{def}}{=} G_N^{\mathcal{A}}(0). \tag{7.2.17}$$

Proposition 7.2.3 *If the regularity condition is satisfied for $W_N^{\mathcal{A}}(0)$ (with the redundant constraints being removed), then*

1. $\lim_{\epsilon \to 0} W_N(\epsilon) = W_N^{\mathcal{A}}(0)$; *and*
2. $\lim_{\epsilon \to 0} G^N(\epsilon) = G_N^{\mathcal{A}}(0)$.

Proof. The proof follows from Proposition 7.1.9. $\square$

7.3 Additional comments for Chapter 7

A possibility of the presence of implicit constraints in families of finite-dimensional LP problems depending on a small parameter was noted in [92], where such families were called singularly perturbed. In [92] it has been also shown that, under certain conditions, the "true limits" of the optimal value and of the optimal solutions set of such SP families of LP problems can be obtained by adding these implicit constraints to the set of constrains defining the feasible set with the zero value of the parameter (see Theorem 2.3, p. 149 in [92] and also more recent results in [17]). This chapter extends the aforementioned earlier results to the SILP setting.

8

Perturbations of semi-infinite dimensional linear programming problems related to long run average optimal control problems

The purpose of this chapter is to establish some interconnections between regularly and singularly perturbed SILP problems and the corresponding perturbed LRAOC problems. The chapter consists of two sections. In Section 8.1, we deal with regularly perturbed semi-infinite dimensional linear programming (SILP) and related perturbed LRAOC problems. In Section 8.2, we consider singularly perturbed SILP related to singularly perturbed LRAOC problems.

8.1 Regularly perturbed SILP problems related to long run average optimal control problems.

Let us consider the perturbed control system

$$y'(t) = f_0(u(t), y(t)) + \epsilon f_1(u(t), y(t)), \tag{8.1.1}$$

where $f_0(u,y) : U \times I\!R^m \to I\!R^m$, $f_1(u,y) : U \times I\!R^m \to I\!R^m$ are continuous in (u,y) and satisfy Lipschitz conditions in y. The controls $u(\cdot)$ are assumed to be Lebesgue measurable and take values in a given compact metric space U.

Also, consider the long run average optimal control problem defined on the trajectory of the system (8.1.1). Namely,

$$\inf_{(u(\cdot),y(\cdot))} \liminf_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} q(u(t),y(t))dt, \tag{8.1.2}$$

where $q(u,y) : U \times I\!R^m \to I\!R^1$ is a given continuous function and *inf* is sought over all admissible pairs of the system (8.1.1) (that is ones that satisfy the inclusions $u(t) \in U,\ y(t) \in Y$).

The problem (8.1.2) can be reformulated as the IDLP problem of the form (see Section 1.2)

$$\min_{\gamma\in W(\epsilon)} \int_{U\times Y} q(u,y)\gamma(du,dy), \tag{8.1.3}$$

where

$$W(\epsilon) = \Big\{\gamma | \gamma \in \mathcal{P}(U\times Y), \int_{U\times Y} (\nabla\phi(y))^T \Big[f_0(u,y) + \epsilon f_1(u,y)\Big]\gamma(du,dy) = 0,$$

$$\forall \phi \in C^1\Big\}. \tag{8.1.4}$$

The above problem can be approximated by the following SILP problem (see Section 2.3 and also Section 7.1)

$$\min_{\gamma\in W_N(\epsilon)} \int_{U\times Y} q(u,y)\gamma(du,dy) = G^N(\epsilon), \tag{8.1.5}$$

where

$$W_N(\epsilon) = \Big\{\gamma | \gamma \in \mathcal{P}(U\times Y), \int_{U\times Y} (\nabla\phi_i(y))^T \Big[f_0(u,y) + \epsilon f_1(u,y)\Big]\gamma(du,dy) = 0,$$

$$i = 1,2,...,N\Big\} \tag{8.1.6}$$

and $\nabla\phi_i(y)$ being assumed to be linear independent (see 2.2.5). By taking $\epsilon = 0$ in the problem defined above, one can obtain the reduced problem

$$\min_{\gamma\in W_N(0)} \int_{U\times Y} q(u,y)\gamma(du,dy) = G^N(0), \tag{8.1.7}$$

where

$$W_N(0) = \Big\{\gamma|\gamma \in \mathcal{P}(U\times Y), \int_{U\times Y} (\nabla\phi_i(y))^T f_0(u,y)\gamma(du,dy) = 0, \quad i = 1,2,...,N\Big\}. \tag{8.1.8}$$

The dual problems of the perturbed and reduced SILP problems (8.1.5) and (8.1.7) are as defined in Section 7.1. Namely,

$$D^N(\epsilon) \stackrel{def}{=} \max_{(v,d)} \Big\{ d \le q(u,y) + \sum_{i=1}^{N} v_i \Big((\nabla\phi_i(y))^T f_i^0(u,y) + \epsilon(\nabla\phi_i(y))^T f_i^1(u,y)\Big), \tag{8.1.9}$$

$$v = (v_i) \in I\!R^N \Big\}.$$

and

$$D^N(0) \stackrel{def}{=} \max_{(v,d)} \Big\{ d \le q(u,y) + \sum_{i=1}^{N} v_i (\nabla\phi_i(y))^T f_i^0(u,y), \quad v = (v_i) \in I\!R^N \Big\}. \tag{8.1.10}$$

The duality relationships between the perturbed and reduced SILP problems and the corresponding duals have been established in Section 7.1.

Note that with the change of the notations

$$(u,y) = x, \qquad U\times Y = X, \qquad \gamma(du,dy) = \gamma(dx),$$

$$\nabla\phi_i(y) f_0(u,y) = h_i^0(x), \quad \nabla\phi_i(y) f_1(u,y) = h_i^1(x),$$

the problems (8.1.5) and (8.1.7) take the form (7.1.1) and (7.1.3).

Theorem 8.1.1 *If the system (8.1.1), taken with $\epsilon = 0$, is locally approximately controllable on Y (see Definition 2.2.7), then the reduced problem (8.1.7) satisfies the regularity condition of Definition 7.1.5 and, hence (by Proposition 7.1.9),*

$$\lim_{\epsilon\to 0} G^N(\epsilon) = G^N(0). \tag{8.1.11}$$

Proof. The proof is similar to that of Lemma 2.2.9. □

8.2 Singularly perturbed SILP problems related to long run average optimal control problems.

Consider the LRAOC problem (4.1.5). The IDLP problem related to the problem (4.1.5) is of the form (4.1.7), where $W(\epsilon)$ is defined by (4.1.8). The IDLP problem (4.1.7) is approximated by the following SILP problem

$$\min_{\gamma \in W_{N,M}(\epsilon)} \int_{U\times Y\times Z} q(u,y,z)\gamma(du,dy,dz) \stackrel{\text{def}}{=} G^{N,M}(\epsilon), \tag{8.2.1}$$

$$W_{N,M}(\epsilon) = \Big\{\gamma | \gamma \in \mathcal{P}(U\times Y\times Z),$$

$$\int_{U\times Y\times Z} \Big[\psi_j(z)(\nabla\phi_i(y))^T f(u,y,z) + \epsilon\phi_i(y)(\nabla\psi_j(z))^T \; g(u,y,z)\Big]\gamma(du,dy,dz) = 0,$$

$$i = 0,1,2,...,N, \quad j = 0,1,2,...,M\Big\}, \tag{8.2.2}$$

where $\nabla\psi_j(z), \quad j = 1,...,M$ and $\nabla\phi_i(y), \quad i = 1,...,N,$ are assumed to be linearly independent (as in Section 5.4), and where $\psi_0(z) \stackrel{\text{def}}{=} 1$ and $\phi_0(y) \stackrel{\text{def}}{=} 1$.

The reduced problem is obtained from (8.2.1) and (8.2.2) by taking $\epsilon = 0$:

$$\min_{\gamma \in W_{N,M}(0)} \int_{U\times Y\times Z} q(u,y,z)\gamma(du,dy,dz) \stackrel{\text{def}}{=} G^{N,M}(0), \tag{8.2.3}$$

$$W_{N,M}(0) = \Big\{\gamma | \gamma \in \mathcal{P}(U\times Y\times Z), \int_{U\times Y\times Z} \psi_j(z)(\nabla\phi_i(y))^T f(u,y,z)\gamma(du,dy,dz) = 0,$$

$$i = 0,1,2,...,N, \quad j = 0,1,2,...,M\Big\}. \tag{8.2.4}$$

Consider the group of constraints in (8.2.2) corresponding to $i = 0$. Recalling that $\phi_0(y) = 1$, one can obtain

$$\epsilon \int_{U\times Y\times Z} (\nabla\psi_j(z))^T g(u,y,z)\gamma(du,dy,dz) = 0. \tag{8.2.5}$$

Since ϵ is positive, the equalities (8.2.5) are equivalent to

$$\int_{U\times Y\times Z} (\nabla\psi_j(z))^T g(u,y,z)\gamma(du,dy,dz) = 0, \quad j = 1,2,...,M. \tag{8.2.6}$$

By adding these constraints to the set of constraints in $W_{N,M}(\epsilon)$, one can define the set

$$W^{\mathcal{A}}_{N,M}(\epsilon) \stackrel{\text{def}}{=} \Big\{ \gamma | \gamma \in \mathcal{P}(U \times Y \times Z),$$

$$\int_{U\times Y\times Z} \Big[\psi_j(z)(\nabla\phi_i(y))^T f(u,y,z) + \epsilon\phi_i(y)(\nabla\psi_j(z))^T \; g(u,y,z) \Big] \gamma(du,dy,dz) = 0,$$

$$\int_{U\times Y\times Z} (\nabla\psi_j(z))^T g(u,y,z)\gamma(du,dy,dz) = 0, \qquad i = 1,2,...,N, \;\; j = 0,1,2,...,M \Big\}. \tag{8.2.7}$$

By taking $\epsilon = 0$ in (8.2.7), one obtains the set (compare with the derivation of Section 4.1)

$$W^{\mathcal{A}}_{N,M}(0) = \Big\{ \gamma | \gamma \in \mathcal{P}(U \times Y \times Z), \int_{U\times Y\times Z} \psi_j(z)(\nabla\phi_i(y))^T \; f(u,y,z)\gamma(du,dy,dz) = 0,$$

$$\int_{U\times Y\times Z} (\nabla\psi_j(z))^T g(u,y,z)\gamma(du,dy,dz) = 0, \qquad i = 1,2,...,N, \;\; j = 0,1,2,...,M \Big\}. \tag{8.2.8}$$

Note that, from the augmentation process described above, it follows that $W^{\mathcal{A}}_{N,M}(\epsilon) = W_{N,M}(\epsilon)$. However, in the general case, $W^{\mathcal{A}}_{N,M}(0) \neq W_{N,M}(0)$.

Consider the problem

$$\min_{\gamma\in W^{\mathcal{A}}_{N,M}(0)} \int_X q(x)\gamma(dx) \stackrel{\text{def}}{=} G^{\mathcal{A}}_{N,M}(0). \tag{8.2.9}$$

Recall some definitions and notations from Section 4.1. The associated and the averaged systems are defined by the equation

$$y'(\tau) = f(u(\tau), y(\tau), z), \qquad z = const, \tag{8.2.10}$$

and, respectively, by the equations

$$z'(\tau) = \epsilon\tilde{g}(\mu(\tau), z(\tau)), \tag{8.2.11}$$

where

$$\tilde{g}(\mu, z) \stackrel{def}{=} \int_{U\times Y\times Z} g(u,y,z)\mu(du,dy) \tag{8.2.12}$$

and

$$\mu(\tau) \in W(z), \tag{8.2.13}$$

$$W(z) \stackrel{def}{=} \Big\{ \mu : \mu \in \mathcal{P}(U \times Y \times Z), \int_{U\times Y\times Z} (\nabla\phi(y))^T f(u,y,z)\mu(du,dy,dz) = 0 \Big\}. \tag{8.2.14}$$

Let us assume that the associated system (8.2.10) and the averaged system (8.2.11) satisfy the local controllability condition on Y and Z, respectively (see Assumptions 5.4.3 and 5.4.1).

Theorem 8.2.1 *If the associated system (8.2.10) and the averaged system (8.2.11) satisfy the local controllability condition in Y and Z (that is, the Assumptions 5.4.3 and 5.4.1 are satisfied), then the optimal value of the perturbed problem (8.2.1) converges to the optimal value of the augmented reduced problem $G^{\mathcal{A}}_{N,M}(0)$. That is,*

$$\lim_{\epsilon \to 0} G^{N,M}(\epsilon) = G^{\mathcal{A}}_{N,M}(0). \tag{8.2.15}$$

Proof. By Proposition 7.2.3, the converges (8.2.15) will follow if one shows that the constraints defining $W^{\mathcal{A}}_{N,M}(0)$ satisfy the regularity condition. That is, it is enough to show that $\forall (u,y,z) \in U \times Y \times Z$ the inequality

$$\sum_{i=1}^{N}\sum_{j=0}^{M} v_{i,j}\psi_j(z)(\nabla\phi_i(y))^T f(u,y,z) + \sum_{j=1}^{M} \tilde{v}_j(\nabla\psi_j(z))^T g(u,y,z) \geq 0 \tag{8.2.16}$$

is valid only if $v_{i,j} = 0 \;\; i = 1,2,...,N, \;\; j = 0,1,2,...,M$ and $\tilde{v}_j = 0 \;\; j = 1,2,...,M$.

Take $\mu \in W(z)$. By integrating (8.2.16) over $\mu(du,dy)$, one can obtain

$$\begin{aligned}\int_{U\times Y} \sum_{i=1}^{N}\sum_{j=0}^{M} v_{i,j}\psi_j(z)(\nabla\phi_i(y))^T f(u,y,z)\mu(du,dy) + \\ \int_{U\times Y} \sum_{j=1}^{M} \tilde{v}_j(\nabla\psi_j(z))^T g(u,y,z)\mu(du,dy) \geq 0, \quad \forall (u,y,z) \in U \times Y \times Z.\end{aligned} \tag{8.2.17}$$

Note that the first integral is zero due to (8.2.13). Hence,

$$\int_{U\times Y} \sum_{j=1}^{M} \tilde{v}_j(\nabla\psi_j(z))^T g(u,y,z)\mu(du,dy) \geq 0. \tag{8.2.18}$$

By (8.2.12), from (8.2.18) it follows that

$$\sum_{j=1}^{M} \tilde{v}_j(\nabla\psi_j(z))^T \tilde{g}(\mu,z) \geq 0. \tag{8.2.19}$$

Let $(\mu(\tau), z(\tau))$ be a solution of the averaged system (8.2.11). From (8.2.19),

$$\sum_{j=1}^{M} \tilde{v}_j (\nabla \psi_j(z(\tau)))^T \tilde{g}(\mu(\tau), z(\tau)) \geq 0. \tag{8.2.20}$$

Denote,

$$\Psi(z) \stackrel{\text{def}}{=} \sum_{j=1}^{M} \tilde{v}_j \psi_j(z). \tag{8.2.21}$$

From (8.2.19) it follows that

$$\Psi(z_2) - \Psi(z_1) = \int_0^T \Big(\nabla\Psi(z(\tau))\Big)^T \tilde{g}(\mu(\tau), z(\tau)) d\tau \geq 0 \quad \Rightarrow \quad \Psi(z_2) \geq \Psi(z_1). \tag{8.2.22}$$

Since z_1, z_2 are arbitrary points in Z^0, the above inequality allows one to conclude that

$$\Psi(z) = const \ \ \forall z \in Z^0 \quad \Rightarrow \quad \Psi(z) = const \quad \forall z \in cl(Z^0).$$

The latter implies that $\nabla\Psi(z) = 0, \quad \forall z \in int(clZ^0)$. Hence, due to the linear independence of $\nabla\psi_j(z)$

$$\tilde{v}_j = 0. \quad j = 1, 2, ..., M. \tag{8.2.23}$$

Let us show that $v_{i,j} = 0$. By (8.2.23), from (8.2.16) it follows that

$$\sum_{i=1}^{N} \sum_{j=0}^{M} v_{i,j} \psi_j(z) (\nabla\phi_i(y))^T f(u, y, z) \geq 0. \tag{8.2.24}$$

Denote,

$$v_i(z) \stackrel{\text{def}}{=} \sum_{j=0}^{M} v_{i,j} \psi_j(z). \tag{8.2.25}$$

Then, the inequality (8.2.24) will take the form

$$\sum_{i=1}^{N} v_i(z) (\nabla\phi_i(y))^T f(u, y, z) \geq 0. \tag{8.2.26}$$

By following the same procedure as above, it can be shown that the inequality (8.2.26) is valid only if $v_i(z) = 0, \forall z \in Z$. Thus, due to (8.2.25)

$$\sum_{j=0}^{M} v_{i,j} \psi_j(z) = 0 \quad \forall z \in Z \qquad \Rightarrow \qquad \sum_{j=1}^{M} v_{i,j} \nabla\psi_j(z) = 0 \quad \forall z \in int(Z). \tag{8.2.27}$$

The latter implies that $v_{i,j} = 0 \quad j = 1, 2, ..., M, \quad i = 1, 2, ..., N$, (due to the linear independence of $\nabla \psi_j(z)$).

Thus, to finalise the proof, we need to show that $v_{i,0} = 0, \quad \forall i = 1, 2, ..., N$.

From (8.2.16), we obtain

$$\sum_{i=1}^{N} v_{i,0} (\nabla \phi_i(y))^T f(u, y, z) \geq 0. \tag{8.2.28}$$

This implies that $v_{i,0} = 0$ (due to the local controllability condition of the associated system on Y; see the proof of Proposition 5.4.4). $\square$

8.3 Additional comments for Chapter 8

The results of this chapter are related to the results of the Sections 5.3 and 5.4. The difference is that in Sections 5.3 and 5.4 we dealt with SILP approximation of the averaged IDLP problem and the corresponding associated dual problems. In this chapter, however, we consider SILP problems approximating the IDLP problem related to the singularly perturbed LRAOC problem.

Conclusions and suggestions for further research

We have developed techniques for analysis and construction of near optimal solutions of optimal control problems with long run average optimality criteria (LRAOC). Our main results can be summarised as follows.

In Part I, we have shown (based on results about relationships between the LRAOC problem and the corresponding IDLP problem), that necessary and sufficient optimality conditions for the LRAOC problem can be stated in terms of a solution of the HJB inequality (see Proposition 2.1.1 and Proposition 2.1.4), which is equivalent to the problem dual with respect to the IDLP problem. Note that the difference of Propositions 2.1.1 and 2.1.4 from "classic" sufficient and necessary conditions of optimality is that a solution of the HJB inequality (rather than that of the HJB equation) is used. The dual to the IDLP problem is a max-min type variational problem on the space of continuously differentiable functions. This dual problem is approximated by max-min problems on finite dimensional subspaces of the space of continuously differentiable functions, which are dual to the semi-infinite dimensional linear programming (SILP) problems approximating the IDLP problem. We have given conditions under which solutions of these "semi-infinite" duals exist and can be used for construction of near optimal solutions of the LRAOC problem (see Proposition 2.2.8 and Theorem 3.1.5). One of the results obtained in Part I is stated in the form of an algorithm, the convergence of which is proved (see Theorem 3.3.1) and which is illustrated with numerical example (see Section 3.4).

In Part II, we extend the consideration of Part I to singularly perturbed LRAOC problems. The key concepts introduced and dealt with, in this part, are those of optimal and near optimal average control generating (ACG) families (see Definitions 5.1.1 and 5.1.3). Sufficient and necessary optimality conditions for an ACG family to be optimal (based on the assumption that solutions of the averaged and associated dual problems exist) are established (see Proposition 5.2.1). Sufficient conditions for existence of solutions of approximating averaged and associated dual problem have been given (see Propositions 5.4.2 and 5.4.4), the latter being used for construction of near optimal ACG families (see Theorem 5.5.8). Also, a linear programming based algorithm allowing one to find solutions of approximating averaged problem and solutions of the corresponding approximating (averaged and associated) dual problems numerically is outlined. A way how an asymptotically near optimal control of a singularly perturbed LRAOC problem can be constructed on the basis of a near optimal ACG family is indicated (see Theorem 6.1.3), the construction being illustrated with a numerical example (see Section 6.3).

In Part III, we have studied families of SILP problems depending on a small parameter. We introduced a regularity condition and we showed that if it is fulfilled, then the family of SILP problems is regularly perturbed and if it is not fulfilled, then the family is likely to be singularly perturbed (see Proposition 7.1.9 and Section 7.2). We have shown that the phenomenon of discontinuity of the optimal value in the SILP setting can be explained by the presence of some implicit constraints that disappear with equating of the small parameter to zero. By adding these constraints, we constructed the problem the optimal value of which defines the "true limit" of the optimal value of the singularly perturbed family (see Proposition 7.2.3). Also, we showed, how the obtained results can be used in dealing with a family of SILP problems related to perturbed LRAOC problems (see Sections 8.1 and 8.2).

Many of the results obtained are readily extendable to other classes of optimal control problems and some of the ideas that we exploited can be used in the dynamic games setting. This, however, will be the subject of future research.

Appendix

A Construction of the dual problem

Let $\{\phi_i(\cdot) \in C^1, i = 1, 2, ...\}$ be the sequence of functions introduced in Section 2.2. Observe that due to approximating property (see (2.2.4)) of this sequence of functions, the set W (see (1.2.1)) can be presented in the form of a countable system of equations. That is,

$$W = \Big\{\gamma \in \mathcal{P}(U \times Y) : \int_{U\times Y} (\nabla\phi_i(y))^T f(u,y)\gamma(du,dy) = 0, \quad i = 1, 2, ...\Big\}, \qquad (A.1)$$

where, without loss of generality, one may assume that the functions $\phi_i(\cdot)$ satisfy the following normalization conditions:

$$\max_{y\in\hat{B}}\{|\phi_i(y)|,\ ||\nabla\phi_i(y)||, ||\nabla^2\phi_i(y)\} \leq \frac{1}{2^i}, \quad i = 1, 2, ... \,. \qquad (A.2)$$

In the above expression, $||\nabla\phi_i(y)||$ is a norm of $\nabla\phi_i(y)$ in $I\!R^m$, $||\nabla^2\phi_i(y)||$ is a norm of the Hessian (the matrix of second derivatives of $\phi_i(y)$) in $I\!R^m \times I\!R^m$, and $\hat{B}$ is closed ball in $I\!R^m$ that contains Y in its interior.

Let l_1 and l_∞ stand for the Banach spaces of infinite sequences such that, for any $x = (x_1, x_2, ...) \in l_1$,

$$||x||_{l_1} \stackrel{\text{def}}{=} \sum_i |x_i| < \infty$$

and, for any $\lambda = (\lambda_1, \lambda_2, ...) \in l_\infty$,

$$||\lambda||_{l_\infty} \stackrel{\text{def}}{=} \sup_i |\lambda_i| < \infty.$$

It easy to see that, given an element $\lambda \in l_\infty$ one can define a linear continuous functional

$\lambda(\cdot) : l_1 \to R^1$ by the equation

$$a\lambda(x) = \sum_i \lambda_i x_i \qquad \forall x \in l_1, \qquad\qquad ||\lambda(\cdot)|| = ||\lambda||_{l_\infty}. \tag{A.3}$$

It is also known (see, e.g., [96], p.86) that any continuous linear functional $\lambda(\cdot) : l_1 \to I\!R^1$ can be presented in the form (A.3) with some $\lambda \in l_\infty$.

Note that from (A.2) it follows that $(\phi_1(y), \phi_2(y), ...) \in l_\infty$ and $(\frac{\partial \phi_1}{\partial y_j}, \frac{\partial \phi_2}{\partial y_j}, ...) \in l_1$ for any $y \in Y$. Hence, the function $\eta_\lambda(y)$,

$$\eta_\lambda(y) \stackrel{\text{def}}{=} \sum_i \lambda_i \phi_i(y), \qquad \lambda = (\lambda_1, \lambda_2, ...) \in l_\infty, \tag{A.4}$$

is continuous differentiable, with $\nabla \eta_\lambda(y) = \sum_i \lambda_i \nabla \phi_i(y)$.

Let us now rewrite problem (1.2.11) in a "standard" LP form by using the representation (A.1). Let $\mathcal{M}(U \times Y)$ (respectively, $\mathcal{M}_+(U \times Y)$) stand for the space of all (respectively, all nonnegative) measures with bounded variations defined on Borel subsets of $U \times Y$, and let $\mathcal{A}(\cdot) : \mathcal{M}(U \times Y) \mapsto I\!R^1 \times l_1$ stand for the linear operator defined for any $\gamma \in \mathcal{M}(U \times Y)$ by the equation

$$\mathcal{A}(\gamma) \stackrel{\text{def}}{=} \Big(\int_{U \times Y} 1_{U \times Y}(u, y) \gamma(du, dy), \ \int_{U \times Y} (\nabla \phi_i(y))^T f(u, y) \gamma(du, dy), \quad i = 1, 2, ... \Big).$$

In this notation problem (1.2.11) takes the form

$$\min_\gamma \{ \langle q, \gamma \rangle | \ \mathcal{A}(\gamma) = (1, 0), \ \gamma \in \mathcal{M}_+ \}, \tag{A.5}$$

where 0 is the zero element of l_1, and $\langle \cdot, \gamma \rangle$, here and in what follows, stands for the integral of the corresponding function over γ.

Define now the linear operator $\mathcal{A}^*(\cdot) : I\!R^1 \times l_\infty \mapsto C(U \times Y) \subset \mathcal{M}^*(U \times Y)$ by the equation

$$\mathcal{A}^*(d, \gamma) \stackrel{\text{def}}{=} d + (\nabla \eta_\lambda(\cdot))^T f(\cdot, \cdot) \qquad \forall d \in I\!R^1, \ \forall \lambda = (\lambda_i) \in l_\infty, \tag{A.6}$$

where $\eta_\lambda(\cdot)$ is as defined in (A.4). Note that from (A.6) it follows that, for any $\gamma \in \mathcal{M}(U \times Y)$,

$$\langle \mathcal{A}^*(d, \lambda), \gamma \rangle = \int_{U \times Y} (d \ 1_{U \times Y}(u, y) + (\nabla \eta_\lambda(y))^T f(u, y) \gamma(du, dy) \stackrel{\text{def}}{=} \langle (d, \lambda), \mathcal{A}(\lambda) \rangle. \tag{A.7}$$

That is, the operator $\mathcal{A}^*(\cdot)$ is the adjoint of $\mathcal{A}(\cdot)$, and, hence, the problem dual to (A.5) can be written in the form (see [5], p.39)

$$\sup_{(d,\lambda)\in R^1\times l_\infty} \{d| \ -\mathcal{A}^*(d,\lambda)+q(\cdot)\geq 0\} \tag{A.8}$$

and, by (A.6), is equivalent to

$$\sup_{(d,\lambda)\in R^1\times l_\infty} \{d| \ -d-(\nabla\eta_\lambda(y))^T f(u,y)+q(u,y)\geq 0 \quad \forall (u,y)\in U\times Y\}. \tag{A.9}$$

Due to the approximation property (2.2.4), the optimal value in (A.9) will be the same as in the problem

$$\sup_{(d,\eta(\cdot))\in R^1\times C^1} \{d| \ -d-(\nabla\eta(y))^T f(u,y)+q(u,y)\geq 0 \quad \forall (u,y)\in U\times Y\}$$

$$\Rightarrow \qquad \sup_{(d,\eta(\cdot))\in R^1\times C^1} \{d| \ d\leq(\nabla\eta(y))^T f(u,y)+q(u,y) \quad \forall (u,y)\in U\times Y\}=D^*,$$

the latter being equivalent to (1.3.1).

B Proof of Theorem 1.3.1

Proof of Theorem 1.3.1 (iii). If the function $\eta(\cdot)$ satisfying (1.3.5) exists, then

$$\min_{(u,y)\in U\times Y}(-\nabla\eta(y))^T f(u,y)>0$$

and, hence,

$$\lim_{\alpha\to\infty}\min_{(u,y)\in U\times Y}\{q(u,y)+\alpha(-\nabla\eta(y))^T f(u,y)\}=\infty. \tag{B.10}$$

This implies that the optimal value of the dual problem is unbounded ($D^*=\infty$).

Assume now that the optimal value of the dual problem is unbounded. That is, there exists a sequence $(d_k,\eta_k(\cdot))$ such that

$$d_k\leq q(u,y)+(\nabla\eta_k(y))^T f(u,y) \qquad \forall (u,y)\in U\times Y, \qquad \lim_{k\to\infty} d_k=\infty \tag{B.11}$$

$$\Rightarrow \qquad 1\leq\frac{1}{d_k}q(u,y)+\frac{1}{d_k}(\nabla\eta_k(y))^T f(u,y) \qquad \forall (u,y)\in U\times Y. \tag{B.12}$$

For k large enough,

$$\frac{1}{d_k}|q(u,y)|\leq\frac{1}{2} \qquad \forall (u,y)\in U\times Y.$$

Hence,

$$\frac{1}{2} \leq \frac{1}{d_k}(\nabla \eta_k(y))^T f(u,y) \qquad \forall (u,y) \in U \times Y. \tag{B.13}$$

That is, the function $\eta(y) \stackrel{\text{def}}{=} -\frac{1}{d_k}\eta_k(y)$ satisfies (1.3.5). $\square$

Proof of Theorem 1.3.1 (i). From (1.3.3) it follows that, if W is not empty, then the optimal value of the dual problem is bounded.

Conversely, let us assume that the optimal value D^* of the dual problem is bounded and let us establish that W is not empty. Assume that this is not true and W is empty. Define the set Q by the equation

$$Q \stackrel{\text{def}}{=} \Big\{ x = (x_1, x_2, ...) : \;\; x_i = \int_{U \times Y} (\nabla \phi_i(y))^T f(u,y) \gamma(du, dy), \;\; \gamma \in \mathcal{P}(U \times Y) \Big\}. \tag{B.14}$$

It is easy to see that the set Q is a convex and compact subset of l_1 (the fact that Q is relatively compact in l_1 is implied by (A.2); the fact that it is closed follows from that $\mathcal{P}(U \times Y)$ is compact in weak convergence topology).

By (A.1), the assumption that W is empty is equivalent to the assumption that the set Q does not contain the "zero element" ($0 \notin Q$). Hence, by a separation theorem (see, e.g., [96], p.59), there exists $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, ...) \in l_\infty$ such that

$$0 = \bar{\lambda}(0) > \max_{x \in Q} \sum_i \bar{\lambda}_i x_i = \max_{\gamma \in \mathcal{P}(U \times Y)} \int_{U \times Y} (\nabla \eta_{\bar{\lambda}}(y))^T f(u,y) \gamma(du, dy)$$

$$= \max_{(u,y) \in U \times Y} (\nabla \eta_{\bar{\lambda}}(y))^T f(u,y),$$

where $\eta_{\bar{\lambda}}(y) = \sum_i \bar{\lambda}_i \phi_i(y)$ (see A.4). This implies that the function $\eta(y) \stackrel{\text{def}}{=} \eta_{\bar{\lambda}}(y)$ satisfies (1.3.5), and, by Theorem 1.3.1 (iii), D^* is unbounded. Thus, we have obtained a contradiction that proves that W is not empty. $\square$

Proof of Theorem 1.3.1 (ii). By Theorem 1.3.1 (i), if the optimal value of the dual problem (1.3.1) is bounded, then W is not empty and, hence, a solution of the problem (1.2.11) exists.

Define the set $\hat{Q} \subset I\!R^1 \times l_1$ by the equation

$$\hat{Q} \stackrel{\text{def}}{=} \Big\{ (\theta, x) : \theta \geq \int_{U \times Y} q(u,y) \gamma(du, dy), \tag{B.15}$$

$$x = (x_1, x_2, ...), \; x_i = \int_{U \times Y} (\nabla \phi_i(y))^T f(u,y) \gamma(du, dy), \; \gamma \in \mathcal{P}(U \times Y) \Big\}. \tag{B.16}$$

The set $\hat{Q}$ is convex an closed. Also, for any $j = 1, 2, ...$, the point $(\theta_j, 0) \notin \hat{Q}$, where

$\theta_j \stackrel{\text{def}}{=} G^* - \frac{1}{j}$ and 0 is the zero element of l_1. On the basis of a separation theorem (see [96] p.59), one may conclude that there exists a sequence $(k^j, \lambda^j) \in I\!R^1 \times l_\infty,\ j = 1, 2, ...$ (with $\lambda^j \stackrel{\text{def}}{=} (\lambda_1^j, \lambda_2^j, ...)$) such that

$$k^j(G^* - \frac{1}{j}) + \delta^j \leq \inf_{(\theta,x)\in\hat{Q}} \Big\{k^j\theta + \sum_i \lambda_i^j x_i\Big\} =$$

$$\inf_{\gamma\in\mathcal{P}(U\times Y)} \Big\{k^j\theta \ + \int_{U\times Y} (\nabla\eta_{\lambda^j}(y))^T f(u,y)\gamma(du,dy) \quad s.t.\ \ \theta \geq \int_{U\times Y} q(u,y)\gamma(du,dy)\Big\}, \tag{B.17}$$

where $\delta^j > 0$ for all j and $\eta_{\lambda^j} = \sum_i \lambda_i^j \phi_i(y)$. From (B.17) it immediately follows that $k^j \geq 0$. Let us show that $k^j > 0$. In fact, if this were not the case, one would obtain that

$$0 < \delta^j \leq \min_{\gamma\in\mathcal{P}(U\times Y)} \int_{U\times Y} (\nabla\eta_{\lambda^j}(y))^T f(u,y)\gamma(du,dy) = \min_{(u,y)\in U\times Y} \{(\nabla\eta_{\lambda^j}(y))^T f(u,y)\}$$

$$\Rightarrow \qquad \max_{(u,y)\in U\times Y} \{(-\nabla\eta_{\lambda^j}(y))^T f(u,y)\} \leq \ - \delta^j < 0.$$

The latter would lead to the validity of the inequality (1.3.5) with $\eta(y) = -\eta_{\lambda^j}(y)$, which, by Theorem 1.3.1 (iii), would imply that the optimal value of the dual problem is unbounded. Thus, $k^j > 0$.

Dividing (B.17) by k^j one can obtain that

$$G^* - \frac{1}{j} < (G^* - \frac{1}{j}) + \frac{\delta^j}{k^j} \leq$$

$$\min_{\gamma\in\mathcal{P}(U\times Y)} \Big\{ \int_{U\times Y} \Big(q(u,y) \ + \frac{1}{\delta^j}(\nabla\eta_{\lambda^j}(y))^T f(u,y)\Big)\gamma(du,dy)\Big\} =$$

$$\min_{(u,y)\in U\times Y} \Big\{q(u,y) + \frac{1}{\delta^j}(\nabla\eta_{\lambda^j}(y))^T f(u,y)\Big\} \leq D^*$$

$$\Rightarrow \qquad G^* \leq D^*.$$

The latter and (1.3.3) prove (1.3.4). $\Box$

Note that the proof of the Theorem 2.3.3, that defines the duality type relationships of the SILP problem (2.3.4) and the N-approximating max-min problem (2.2.7), follows the same steps as those in the proof of Theorem 1.3.1.

Namely, the proofs of statements (i) and (ii) of the theorem are based on separation theorem in finite-dimensional spaces and follow the argument used in the proofs of Theorem 1.3.1 (i) and Theorem 1.3.1 (ii), with the replacement of the set Q defined in

(B.14) by the set $Q' \subset I\!\!R^N$,

$$Q' \stackrel{\text{def}}{=} \Big\{ h = (h_1, ..., h_N): \ h_i = \int_{U\times Y} (\nabla \eta_i(y))^T f(u,y)\gamma(du,dy), \ \gamma \in \mathcal{P}(U\times Y) \Big\}, \quad \text{(B.18)}$$

and with the replacement of the set $\hat{Q}$ defined in (B.15) by the set $\hat{Q}' \subset I\!\!R^1 \times I\!\!R^N$,

$$\begin{aligned} \hat{Q}' \stackrel{\text{def}}{=} \Big\{ (\theta, h) : \theta \geq \int_{U\times Y} q(u,y)\gamma(du,dy), \quad h = (h_1, ..., h_N) : \\ h_i = \textstyle\int_{U\times Y} (\nabla \eta_i(y))^T f(u,y)\gamma(du,dy). \ \gamma \in \mathcal{P}(U \times Y) \Big\}. \end{aligned} \quad \text{(B.19)}$$

The proof of the statement (iii) of the theorem follows the argument used in the proof of Theorem 7.1.3 (iii), with the replacement of $\eta(y)$ in (B.10) by

$$\eta_v(y) = \sum_{i=1}^{N} v_i \phi_i(y), \ v = (v_i) \in I\!\!R^N,$$

and with the replacement of $\eta_k(y)$ in (B.11), (B.12) by

$$\eta_{v^k}(y) = \sum_{i=1}^{N} v_i^k \phi_i(y), \ v^k = (v_i^k) \in I\!\!R^N.$$

References

[1] C.D. Aliprantis and K.C. Border, "Infinite Dimensional Analysis: A Hitchhiker's Guide", 2nd Edition, Springer, Berlin, 1999.

[2] O. Alvarez and M. Bardi, "Viscosity Solutions Methods for Singular Perturbations in Deterministic and Stochastic Control", *SIAM J. Control and Optimization*, 40 (2001): 4, pp. 1159-1188.

[3] O. Alvarez and M. Bardi, "Singular Perturbations of Nonlinear Degenerate Parabolic PDEs: a General Convergence result", *Arch. Ration. Mech. Anal.*, 170 (2003), pp. 17-61.

[4] B.D.O. Anderson and P.V.Kokotovic, "Optimal Control Problems over Large Time Intervals", *Automatica*, 23 (1987), no 3, pp 355-363.

[5] E.J.Anderson and P.Nash, "Linear Programming in Infinite-Dimensional Spaces", Wiley, Chichester, 1987.

[6] M.Arisawa, H.Ishii and P.L.Lions, "A Characterization of the Existence of Solutions for Hamilton-Jacobi Equations in Ergodic Control Problems with Applications", *Applied Mathematics and Optimization*, 42 (2000), pp 35-50.

[7] Z. Artstein and V. Gaitsgory, "Tracking Fast Trajectories Along a Slow Dynamics: A Singular Perturbations Approach", *SIAM J. Control and Optimization*, 35 (1997): 5, pp. 1487-1507.

[8] Z. Artstein, "Invariant Measures of Differential Inclusions Applied to Singular Perturbations", *J. Differential Equations*, 152 (1999), pp. 289-307.

[9] Z. Artstein, "The Chattering Limit of Singularly Perturbed Optimal Control Problems", *Proceedings of CDC-2000, Control and Decision Conference*, Sydney, 2000, pp. 564-569.

[10] Z. Artstein, “An Occupational Measures Solution to a Singularly Perturbed Optimal Control Problem”, *Control and Cybernetics*, 31 (2002), pp. 623-642.

[11] Z. Artstein, “Invariant Measures and Their Projections in Nonautonomous Dynamical Systems”, *Stochastics and Dynamics*, 4 (2004): 3, pp 439-459.

[12] Z. Artstein and I. Bright, “Periodic Optimization Suffices for Infinite Horizon Planar Optimal Control”, *SIAM J. Control and Optimization*, 48 (2010): 8, pp. 4963-4986

[13] Z. Arstein and A.Vigodner, “Singularly perturbed ordinary differential equations with dynamic limits”, *Proceedings of the Royal Society of Edinburgh*, 126A (1996), 541-569.

[14] R. B. Ash, “Real Analysis and Probability”, Academic Press, New York, 1972

[15] R.B. Ash, “Measure, Integration and Functional Analysis”, Academic Press, New York, 1972.

[16] J.P.Aubin, “Viability Theory”, Birkhauser, Boston, 1991.

[17] K. Avrachenkov, R. S. Burachik, J. A. Filar, V. Gaitsgory, “Constraint augmentation in pseudo-singularly perturbed linear programs”, *Mathematical Programming*, 132 (2012): 1-2, pp. 179-208.

[18] K. Avrachenkov, J. Filar and M. Haviv, “Singular perturbations of Markov chains and decision processes. A Survey”, A chapter in “Handbook of Markov Decision Processes: Methods and Applications”, E.A. Feinberg, A. Shwartz (eds), *International Series in Operations Research and Management Science*, 40 (2002), pp.113-153, Kluwer Academic Publishers.

[19] A. Bagirov, N. Karmitsa and M. Makela, “Introduction to Nonsmooth Optimization: Theory, Practice and Software”, Springer, 2014.

[20] M. Bardi and I. Capuzzo-Dolcetta, “Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations”, Birkhauser, Boston, 1997.

[21] G.K. Basak, V.S. Borkar and M.K. Ghosh, “Ergodic Control of Degenerate Diffusions”, *Stochastic Analysis and Applications*, 15 (1997), pp. 1-17.

[22] A. Bensoussan, “Perturbation Methods in Optimal Control”, John Wiley, New York, 1989.

[23] D.P. Bertsekas and S.E. Shreve, "Stochastic Optimal Control : The Discrete Time Case", Academic Press, New York, 1978.

[24] A.G. Bhatt and V.S. Borkar, "Occupation Measures for Controlled Markov Processes: Characterization and Optimality", *Annals of Probability*, 24 (1996), pp. 1531-1562.

[25] P.Billingsley, "Convvergence of Probability Measures", John Wiley & Sons, New York, 1968.

[26] V.S. Borkar, "Ergodic Control of Diffusion Processes", Springer, Berlin, Forthcoming

[27] V.S. Borkar and V. Gaitsgory, "On Existence of Limit Occupational Measures Set of a Controlled Stochastic Differential Equation", *SIAM J. on Control and Optimization*, 44 (2005/2006): 4, pp. 1436-1473.

[28] V.S. Borkar and V. Gaitsgory, "On Averaging of Singularly Perturbed Controlled Stochastic Differential Equations", *Applied Mathematics and Optimization*, 56 (2007), pp. 169-209.

[29] V.S. Borkar and V. Gaitsgory, "Singular Perturbations in Ergodic Control of Diffusions", *SIAM J. on Control and Optimization*, 46 (2007): 5, pp. 1562-1577.

[30] R. Buckdahn, D. Goreac, and M. Quincampoix, "Stochastic optimal control and linear programming approach", *Applied Mathematics and Optimization*, 63 (2011): 2, pp. 257-276.

[31] D.A. Carlson, A.B. Haurie and A.Leizarowitz, "Optimal Control on Infinite Time Horizon", 2nd Edition, Springer-Verlag, Berlin, 1991.

[32] I. Capuzzo-Dolcetta, P.-L. Lions, "Hamilton-Jacoby equations with state constraints", *Trans. Amer. Math. Soc.* 318 (1990): 2, pp. 643-683.

[33] F. Colonius, "Optimal Periodic Control", Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1988.

[34] F. Colonius and W. Kliemann, "Infinite Time Optimal Control and Periodicity", *Applied Mathematics and Optimization*, 20 (1989), pp. 113-130.

[35] F. Colonius and W. Kliemann, "Some aspects of control systems as dynamical systems", *J.Dynamics and Differential Equations*, 5 (1993), pp. 468-494.

[36] F. Colonius and R. Fabri, "Controllability for Systems with Slowly Varying Parameters", *ESAIM: Control, Optimisation and Calculus of Variations*, 9 (2003), pp. 207-216.

[37] G.B. Dantzig, "Linear Programming and Extensions", Princeton University Press, Princeton, 1963.

[38] M.G. Dmitriev and G.A. Kurina, "Singular Perturbations in Control Problems (Review)", *Automation and Remote Control*, 67 (2006): 1, pp. 1-43

[39] A. Dontchev, T. Donchev and I. Slavov, "A Tichonov -type Theorem for Singularly Perturbed Differential Inclusion", *Nonlinear Analysis*, 26 (1996), pp. 1547-1554.

[40] T.D. Donchev and A.L. Dontchev, "Singular Perturbations in Infinite-Dimensional Control Systems", *SIAM J. Control and Optimization*, 42 (2003), pp. 1795-1812.

[41] T.D. Donchev, "Singularly Perturbed Evolution Inclusions", *SIAM J. Control and Optimization*, 48 (2010): 7, pp. 4572-4590.

[42] T. Dontchev and I. Slavov, "Averaging Method for One Sided Lipschitz Differential Inclusions", *SIAM J. Control and Optimization*, 37 (1999), pp. 1600-1613.

[43] L.C. Evans and D. Gomes, "Linear Programming Interpretations of Mather's Variational Principle", *ESAIM: Control Optimisation and Calculus of Variations*, 8 (2002), pp. 693-702.

[44] M. Falcone, "Numerical Solution of Dynamic Programming Equations", Appendix A in [20].

[45] J.A. Filar, V. Gaitsgory and A. Haurie, "Control of Singularly Perturbed Hybrid Stochastic Systems", *IEEE Transactions on Automatic Control*, 46 (2001): 2, pp. 179-190.

[46] O.P. Filatov and M.M. Hapaev, "Averaging of Systems of Differential Inclusions", Moscow University Publishing House, Moscow, 1998 (in Russian).

[47] L. Finlay, V. Gaitsgory, and I. Lebedev, "Linear Programming Solutions of Periodic Optimization Problems: Approximation of the Optimal Control", *Journal of Industrial and Management Optimization*, 3 (2007): 2, pp. 399-413.

[48] L. Finlay, V.Gaitsgory and I. Lebedev, “Duality In Linear Programming Problems Related to Long Run Average Problems of Optimal Control”, *SIAM J. Control and Optimization*, 47 (2008), pp. 1667-1700.

[49] W.H. Fleming and H.M. Soner, “Controlled Markov Processes and Viscosity Solutions”, Springer-Verlag, New York, 1991.

[50] W. H. Fleming and D. Vermes, “Convex duality approach to the optimal control of diffusions”, *SIAM J. Control Optimization*, 27 (1989): 5, pp. 1136-1155.

[51] V.Gaitsgory, “Control of Systems with Slow and Fast Motions: Averaging and Asymptotic Analysis”, Nauka, Moscow, 1991 (in Rusian).

[52] V.Gaitsgory, “Suboptimization of Singularly Perturbed Control Problems”, *SIAM Journal on Control and Optimization*, 30 (1992), pp. 1228-1249.

[53] V. Gaitsgory, “Suboptimization Control of singularly perturbed systems and periodic optimization”, *IEEE Transactions on Automatic Control*, 38 (1993), pp. 888-903.

[54] V. Gaitsgory, “Limit Hamilton-Jacobi-Isaacs equations for singularly perturbed zero-sum differential games”, *Journal of mathematical analysis and applications*, 202 (1996): 3, pp.862-899.

[55] V. Gaitsgory and A.Leizarowitz, “Limit Occupational Measures Set for a Control System and Averaging of Singularly Perturbed Control Systems”, *Journal of mathematical analysis and applications*, 233 (1999), pp. 461-475.

[56] V. Gaitsgory and M.T. Nguyen, “Multiscale Singularly Perturbed Control Systems: Limit Occupational Measures Sets and Averaging”, *SIAM J. Control and Optimization*, 41 (2002): 3, pp. 954-974.

[57] V. Gaitsgory, “On Representation of the Limit Occupational Measures Set of a Control Systems with Applications to Singularly Perturbed Control Systems”, *SIAM J. Control and Optimization*, 43 (2004), No 1, pp 325-340.

[58] V. Gaitsgory and S. Rossomakhine, “Linear Programming Approach to Deterministic Long Run Average Problems of Optimal Control”, *SIAM J. Control and Optimization*, 44 (2005/2006), No 6, pp. 2006-2037.

[59] V. Gaitsgory and M. Quincampoix, “Linear Programming Approach to Deterministic Infinite Horizon Optimal Control Problems with Discounting”, *SIAM J. of Control and Optimization*, 48 (2009): 4, pp. 2480-2512.

[60] V. Gaitsgory and L. Manic, “Use of Approximations of Hamilton-Jacobi-Bellman Inequality for Solving Periodic Optimization Problems”, *Springer Proceedings in Mathematics and Statistics*, 86 (2014), pp.91-114.

[61] V. Gaitsgory, L. Manic and S.Rossomakhine, “On Average Control Generating Families for Singularly Perturbed Optimal Control Problems with Long Run Average Optimality Criteria”, *Set-Valued and Variational Analysis: Theory and Applications*, 23 (2015): 1, pp.87-131.

[62] V. Gaitsgory and M. Quincampoix, “On sets of occupational measures generated by a deterministic control system on an infinite time horizon”, *Nonlinear Analysis Series A: Theory, Methods & Applications*, 88 (2013), pp 27-41.

[63] V. Gaitsgory, S. Rossomakhine and N. Thatcher, “Approximate Solutions of the HJB Inequality Related to the Infinite Horizon Optimall Control Problem with Discounting”, *Dynamics of Continuous and Impulsive Systems Series B: Applications and Algorithms*, 19 (2012), pp. 65-92.

[64] V. Gaitsgory and S. Rossomakhine, “Averaging and Linear Programming in Some Singularly Perturbed Problems of Optimal Control”, *Journal of Applied Mathematics and Optimization*, Published online in June 2014, DOI 10.1007/s00245-014-9257-1 (arXiv:1309.3734v2 [math.OC])

[65] E. G. Gilbert, “Optimal periodic control: A general theory of necessary conditions”, *SIAM J. of Control and Optimization*, 5 (1977), pp. 717-746.

[66] D.Goreac and C.Ivaşcu “Discontinuous control problems with state constraints: Linear formulations and dynamic programming principles”, *Journal of Mathematical Analysis and Applications*, 402 (2013): 2, pp. 635–647.

[67] D.Goreac and O.S.Serea, “Min–max control problems via occupational measures”, *Optimal Control Applications and Methods*, 35 (2014): 3, pp. 340–360.

[68] D. Goreac and O.S.Serea, “Linearization techniques for L^∞ - control problems and dynamic programming principles in classical and L^∞ control problems”, *ESAIM: Control, Optimization and Calculus of Variations*, 18 (2012), pp. 836-859.

[69] G. Grammel, "Averaging of Singularly Perturbed Systems", *Nonlinear Analysis*, 28 (1997), 1851-1865.

[70] G. Grammel, "On Nonlinear Control Systems with Multiple Time Scales", *Journal of Dynamical and Control Systems*, 10 (2004), pp. 11-28.

[71] G. Guardabassi, A. Locatelli and S. Rinaldi, "Status of periodic optimization of dynamical systems", *Journal of Optimization Theory and Applications*, 14 (1974): 1, pp. 1-20.

[72] K. Helmes and R.H.Stockbridge, "Numerical comparison of controls and verification of optimality for stochastic control problems", *Journal of Optimization Theory and Applications*, 106 (2000): 1, pp. 107-127.

[73] D. Hernandez-Hernandez, O.Hernandez-Lerma, M. Taksar, "The Linear Programming Approach to Deterministic Optimal Control Problems", *Applicationes Mathematicae*, 24 (1996): 1, pp 17-33.

[74] Y. Kabanov and S. Pergamenshchikov, "Two-Scale Stochastic Systems", Springer-Verlag, Berlin-Heidelberg, 2003.

[75] R. Kasimbeyli and M. A. Mammadov, "On Weak Subdifferentials, Directional Derivatives and Radial Epiderivatives for Nonconvex Functions", *SIAM Journal on Optimization*, 20 (2009): 2, pp. 841-855.

[76] P.V.Kokotovic, H.K.Khalil and J.O'Reilly, "Singular Perturbation Methods in Control: Analysis and Design", 2nd Edition, SIAM Classics in Applied Mathematics, 25, 1999.

[77] T.G. Kurtz and R.H. Stockbridge, "Existence of Markov Controls and Characterization of Optimal Markov Controls", *SIAM J. on Control and Optimization*, 36 (1998): 2, pp. 609-653.

[78] H.J. Kushner, "Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems", Birkhauser, Boston, 1990.

[79] H.J. Kushner and P.G. Dupuis, "Numerical Methods for Stochastic Control Problems in Continuous Time", 2nd Edition (revised), Springer-Verlag, New York, 2002.

[80] J.-M. Lasry, A. Bensoussan and J.L. Lions, "Controle stationnaire asymptotique", *Control theory, numerical methods and computer systems modelling: Lecture notes in Econom. and Math. Systems*, Springer, Berlin, 107 (1975), pp. 296-313.

[81] J.B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, "Nonlinear optimal control via occupation measures and LMI-relaxations", *SIAM J. of Control and Optimization*, 47 (2008), pp. 1643-1666.

[82] A. Leizarowitz, "Order Reduction is Invalid for Singularly Perturbed Control Problems with Vector Fast Variables", *Math. Control Signals and Systems*, 15 (2002), pp. 101-119.

[83] P.-L. Lions, "Neumann type boundary conditions for Hamilton-Jacobi equations", *Duke Math. J.*, 52 (1985): 4, pp. 793-820.

[84] J.G. Llavona, "Approximation of Continuously Differentiable Functions", Mathematics Studies, 130, North Holland, Amsterdam, 1986.

[85] H. Maurer, C. Bskens, J.H.R. Kim and C.Y. Kaya, "Optimization methods for the verification of second order sufficient conditions for bang-bang controls", *Optimal Control Applications and Methods*, 26 (2005): 3, pp. 129-156.

[86] H. Maurer and H.J. Pesch, "Direct optimization methods for solving a complex state-constrained optimal control problem in microeconomics", in Special Issue on New Approaches in Dynamic Optimization to Assessment of Economics and Environmental Systems, *Applied Mathematics and Computation*, 204 (2008): 2, pp. 568-579.

[87] M.S. Mendiondo and R.H. Stockbridge, "Approximation of infinite-dimensional linear programming problems which arise in stochastic control", *SIAM J. on Control and Optimization*, 36 (1998), pp. 1448-1472.

[88] S.D. Naidu, "Singular Perturbations and Time Scales in Control Theory and Applications: An Overview", *Dynamics of Continuous Discrete and Impulsive Systems, Series B: Applications and Algorithms*, 9 (2002), pp. 233-278.

[89] R.E. O'Malley, Jr. "Introduction to Singular Perturbations", Academic Press, New York, 1974.

[90] R.E. O'Malley, Jr. "Singular Perturbations and Optimal Control", *Mathematical Control Theory*, Lecture Notes in Mathematics, 680 (1978), pp. 170-218.

[91] K.R Parthasarathy, "Probability Measures on Metric Spaces", Academic Press, New York, 1967.

[92] A.A.Pervozvanskii and V.G. Gaitsgori, "Theory of suboptimal decisions: Decomposition and aggregation", Kluwer, Dordrecht, 1988.

[93] V.A. Plotnikov, "Differential Equations with Multivalued Right-Hand Sides: Asymptotic Methods", AstroPrint, Odessa, 1999 (in Russian).

[94] M. Quincampoix and F. Watbled, "Averaging method for discontinuous Mayer's problem of singularly perturbed control systems", *Nonlinear Analysis: Theory, Methods & Applications*, 54 (2003), pp. 819-837.

[95] J.E. Rubio, "Control and Optimization. The Linear Treatment of Nonlinear Problems", Manchester University Press, Manchester, 1985.

[96] W.Rudin, "Functional Analysis", 2nd Edition, McGraw-Hill, New York, 1991.

[97] O.S. Serea, "Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems", *Journal of Differential Equations*, 255 (2013): 11, pp. 4226-4243.

[98] R.H. Stockbridge, "Time-Average Control of a Martingale Problem. Existence of a Stationary Solution", *Annals of Probability*, 18 (1990), pp. 190-205.

[99] R.H. Stockbridge, "Time-Average Control of a Martingale Problem: A Linear Programming Formulation", *Annals of Probability*, 18 (1990), pp 206-217.

[100] K.L.Teo and L.S.Jennings, "Nonlinear optimal control problems with continuous state inequality constraints", *Journal of Optimization Theory and Applications*, 63 (1989): 1, pp. 1-22.

[101] Vasil'eva A.B., Butuzov V.F., "Asymptotic Expansions of Solutions of Singularly Perturbed Equations", Nauka, Moscow, 1973 (in Russian).

[102] V.Veliov, "A generalization of Tikhonov Theorem for Singularly Perturbed Differential Inclusions", *Journal of Dynamical and Control Systems*, 3 (1997), pp. 1-28.

[103] A.Vigodner, “Limits of Singularly Perturbed Control Problems: Dynamical Systems Approach”, Thesis for the Degree of Doctor of Philosophy, The Weizmann Institute of Science, Rehovot, Israel, 1995.

[104] R. Vinter, “Convex Duality and Nonlinear Optimal Control”, *SIAM J. on Control and Optimization*, 31 (1993): 2, pp. 518-538.

[105] G.G. Yin and Q. Zhang, “Continuous-Time Markov Chains and Applications. A Singular Perturbation Approach”, Springer, New York, 1997.