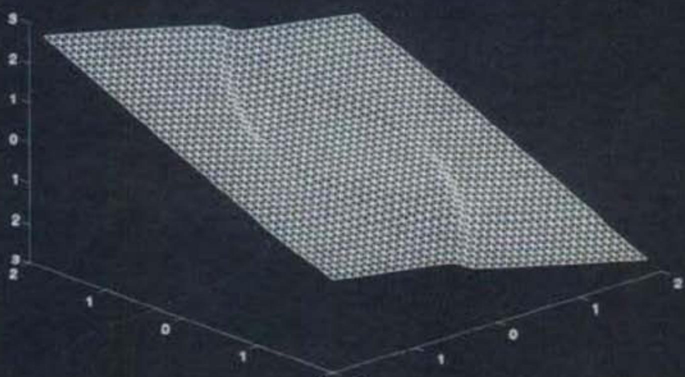
Systems & Control: Foundations & Applications

William M. McEneaney

Max-Plus Methods for Nonlinear Control and Estimation



Birkhäuser

Systems and Control: Foundations & Applications

William M. McEneaney

Max-Plus Methods for Nonlinear Control and Estimation

Birkhäuser
Boston • Basel • Berlin

William M. McEneaney
Department of Mechanical and Aerospace Engineering
and Department of Mathematics
University of California at San Diego
9500 Gilman Dr.
La Jolla, CA 92093-0112
USA

AMS Subject Classification: 35B37, 35F20, 47D99, 49Lxx, 60J60, 65M99, 65N99, 93-02, 93C10

Library of Congress Cataloging-in-Publication Data
McEneaney, William M.
Max-plus methods for nonlinear control and estimation / William M. McEneaney.
p. cm. – (Systems and control)
Includes bibliographical references and index.
ISBN 0-8176-3534-3 (alk. paper)
1. Nonlinear control theory. 2. Matrices. I. Title. II. Systems & control.

QA402.35.M44 2005
629.8'36–dc22 2005053046

ISBN-10 0-8176-3534-3 e-ISBN 0-8176-4453-9
ISBN-13 978-0-8176-3534-3

Printed on acid-free paper.

Printed in the United States of America. (IBT)

9 8 7 6 5 4 3 2 1

www.birkhauser.com

To the world's misfits

Preface

The control and estimation of continuous-time/continuous-space nonlinear systems continues to be a challenging problem, and this is one of the central foci of this book. A common approach is to use dynamic programming; this typically leads to solution of the control or estimation problem via the solution of a corresponding Hamilton–Jacobi (HJ) partial differential equation (PDE). This approach has the advantage of producing the “optimal” control. (The term “optimal” has a somewhat more complex meaning in the class of H_∞ problems. However, we will freely use the term for such controllers throughout, and this meaning will be made more precise when it is not obvious.) Thus, in solving the control/estimation problem, we will be solving some nonlinear HJ PDEs. One might note that a second focus of the book is the solution of a class of HJ PDEs whose viscosity solutions have interpretations as value functions of associated control problems. Note that we will briefly discuss the notion of viscosity solution of a nonlinear HJ PDE, and indicate that this solution has the property that it is the correct weak solution of the PDE. By correct weak solution in this context, we mean that it is the solution that is the value function of the associated control (or estimation) problem. The viscosity solution is also the correct weak solution in many PDE classes not considered here, and references to further literature on this subject will be given. Note that the value function is defined to be the optimal cost that can be achieved, typically as a function of the initial state of the system being controlled.

Perhaps at this point we should indicate the classes of nonlinear control and estimation problems that will be addressed here. One class is comprised of nonlinear optimal control problems (in the absence of disturbances). However, the classes of problems that were the primary motivation for the work were the nonlinear robust/H_∞ control and estimation problems.

Although H_∞ control was first formulated in the frequency domain, the state-space representation quickly appeared. In the case where one is computing the optimal H_∞ controller, the state-space form generally takes the form of a zero-sum differential game (although a subclass will be of special interest).

One player is the controller, and the other player is the disturbance process. The controller will be minimizing the same cost that the disturbance player will be maximizing. The corresponding PDE will be a Hamilton–Jacobi–Isaacs (HJI) PDE, although in many cases of interest, the PDE will belong to the subclass of Hamilton–Jacobi–Bellman (HJB) PDEs. We will more often be interested in the case where one is testing a potential feedback controller to determine whether it is indeed an H_∞ controller. In this case, as in the optimal control case, the value function is the viscosity solution of an HJB PDE. The robust/H_∞ estimation problems will also have associated PDEs which take the form of HJB PDEs.

In all of the above cases, any disturbance or control process will have finite energy over finite time periods. Consequently, we will be interested in solving first-order, nonlinear HJB PDEs. Our focus will be on solving control/estimation problems through solving associated HJB PDE problems, but one could also imagine that for others the primary motivation would be the solution of HJB PDEs with the control interpretation being ancillary.

The most common methods for solving HJB PDEs are finite element methods. With first-order PDEs there are certain conditions implied by the flow of information that must be met in order to obtain good performance in these algorithms. One approach to finite element methods where these conditions are met is through a Markov chain interpretation of the algorithm. A serious problem for solution of HJB PDEs associated with control problems is that the space dimension of the PDE is the dimension of the state-space of the control problem. Since the dimension of the state-space is generally relatively high (≥ 3) for most control systems of interest, this problem, referred to as the curse-of-dimensionality, is ubiquitous.

An alternate approach, which attempts to avoid this problem, is through the use of characteristics (as, for instance, in the Pontryagin Maximum Principle). In this case, one needs only to propagate the solution along (one-dimensional) paths to obtain the value and control at a desired point on such a path. Given a point in state-space, one can ideally propagate a single characteristic trajectory passing through that point to obtain the value and optimal control there. Of course, there are a number of technical difficulties with this approach as well. First of all, one needs to apply some shooting technique to find the correct initial conditions so that the path will pass through the desired point. Second, the viscosity solution is generally nonsmooth, and the projection of the characteristics onto the state-space often generates projected characteristics that cross. Further, there may be regions of state-space through which no such projected characteristic passes. (These last two difficulties are somewhat analogous to shocks and rarefaction waves in conservation laws.) This leads to the need to generate generalized characteristics, and there are associated "bookkeeping" problems which have, so far, proved to be somewhat of a deterrent to application of this method for large or complex problems.

The max-plus-based methods described in this monograph for solution of HJB PDEs belong to an entirely new class of methods for solutions of such

PDEs; they are not equivalent to either the finite element or characteristic approaches. As this is an entirely new class of numerical methods, the discipline is still in its infancy. It is hoped that this book can become a resource for those who would like to develop this new class of methods to a mature level. Through the judicious selection of chapters/sections, it can also provide a readable, concise description of the area for those who are only tangentially interested in this area of research.

Perhaps this is the correct place at which to broadly describe the main facets of this max-plus-based class of methods for first-order HJB PDEs. Nonlinear HJB PDEs have solutions that may be interpreted as the value functions for associated control problems. The semigroup that propagates the solution of such a PDE is identical to the dynamic programming principle (DPP) for the control problem. We assume that this control problem has a maximizing control. (If the control is minimizing, one uses the min-plus algebra rather than the max-plus algebra. The analysis is a mirror image of the maximizing case, and so we will deal mainly with the max-plus case.)

The semigroup associated with the nonlinear HJB PDE is a max-plus linear operator. This leads to what is essentially a max-plus analogue of spectral methods for linear (over the standard field) PDEs. In order to exploit this max-plus linearity, one needs to develop the notion of a vector space over the max-plus algebra (actually a semifield). By considering expansions of the desired solution in terms of the coefficients in a max-plus basis expansion, one obtains the first max-plus technique. In particular, propagation forward in time with a finite, truncated expansion reduces to max-plus matrix-vector multiplication. That is, the coefficients in the expansion one time-step into the future are obtained by max-plus multiplication of the current vector of coefficients by a matrix related to the problem dynamics. Steady-state PDE problems reduce to max-plus eigenvector problems, and steady-state variational inequalities reduce to affine max-plus algebraic problems.

This approach, solving a max-plus eigenvector problem to approximately solve the HJB PDE, has proven very effective. Additional, related approaches have also proven effective. In particular, it has been observed that, in applying this max-plus eigenvector approach, the computational cost of constructing the matrix is typically an order of magnitude higher than that of finding the eigenvector given the matrix. This motivated the development of methods where the computational cost of obtaining the matrix could be reduced. In particular, one can compute matrices for (standard-sense) linear problems, and then take max-plus linear combinations of these to generate the matrix corresponding to the HJB PDE that one wishes to solve.

However, this last concept can sometimes be taken much further. The matrix being referred to is actually a discretization of the kernel of the dual of the original max-plus linear operator. For linear problems, this kernel can be computed analytically (modulo solution of a Riccati equation). One can then take max-plus linear combinations of these analytically obtained kernels to generate the desired kernel. This concept leads to a method, which for a certain class of HJB PDEs avoids the curse-of-dimensionality. In particular, the computational complexity grows at a polynomial rate in state-space dimension, and is instead exponential in a certain measure of problem complexity.

We note that the max-plus algebra (more correctly the max-plus semifield) has been under intense research during the last decade. Much of this research has been directed toward the use of the max-plus algebra in discrete-event system problems. The max-plus additive and multiplicative operations, $\oplus$ and $\otimes$, are defined as $a \oplus b = \max\{a, b\}$ and $a \otimes b = a + b$, respectively. There is extensive literature on the research area. The contribution of this monograph is on the application of the max-plus algebra to solution of HJB PDEs.

Reiterating from above, this book is intended for researchers and graduate students with an interest in the solution of nonlinear control problems and/or HJ PDEs. It should be suitable both for those working in the field and for those who have only tangential interest.

The author would like to thank Wendell H. Fleming whose collaboration has been invaluable. The author also thanks Matthew R. James, J. William Helton and Peter M. Dower for helpful discussions.

Acknowledgments

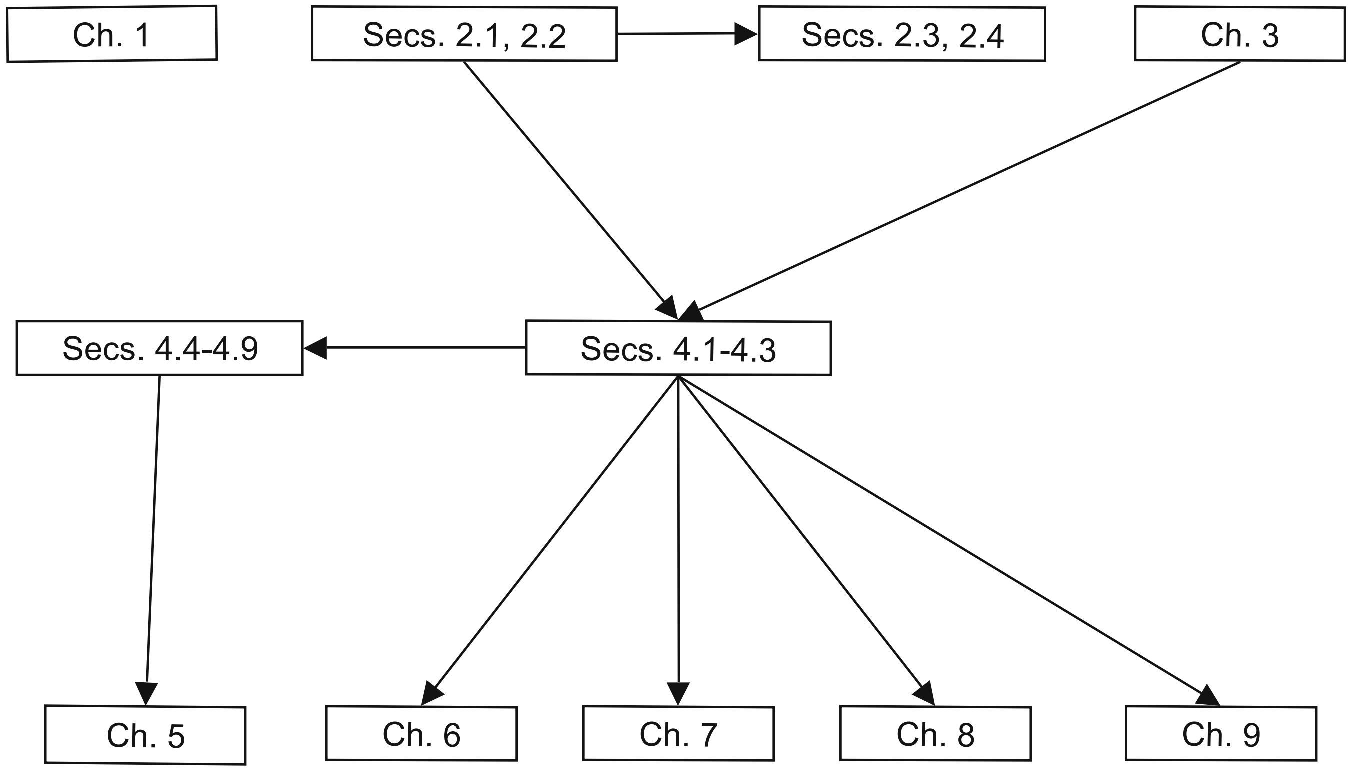
NSF, AFOSR, ONR, DARPA, the Australian Research Council and INRIA have contributed support for various components of the research that led to this book.

San Diego, California

William M. McEneaney
September 2005

Contents

Approximate Dependencies



1

Introduction

One of the chief aims of this book is the presentation of some entirely new classes of numerical methods for the solution of nonlinear control problems and, more generally, Hamilton–Jacobi–Bellman partial differential equations (HJB PDEs). These methods are based on the max-plus formulation of these problems. The development of the max-plus viewpoint on deterministic nonlinear control problems is a second major motivation. The max-plus algebra is a commutative semifield which has come under intense study in the last decade. This interest is due to the confluence of several factors. First of all, there is a rich mathematics including probability theory, analysis and geometry which can be built on the max-plus algebra. A second important factor is that the max-plus algebra is the most natural language for the study of many problems in discrete event systems. However, it was only more recently that the usefulness of the max-plus viewpoint for solution of nonlinear control problems in continuous space arose.

In the max-plus algebra, the addition operation is maximization, and the multiplication operation is what one usually refers to as addition. The additive identity is $-\infty$, and the multiplicative identity is 0. This algebra arises from the standard algebra through large deviations limits (c.f. [1], [84], [82], [97], [98]) and/or the quantization limit of quantum mechanics [69]. If one thinks of control as maximizing some accumulated cost (say an integral cost), then it is clear that the max-plus algebra is a natural framework for the study of these problems. Perhaps it should be noted that integral costs are multiplicative in max-plus and costs that involve suprema over time are additive in the max-plus sense. Both these multiplicative and additive forms are handled within the max-plus framework. It should be noted that when the optimization is of a minimizing form, the min-plus algebra is used instead.

One of the most generally applicable approaches to solving nonlinear control problems is dynamic programming (DP). The Dynamic Programming Principle (DPP) can be viewed as an operator mapping the value function (optimal cost as a function of system state) at one time to the value function at a later (or earlier) time. In the continuous-time case, if one takes the limit in the DPP as the time-interval goes to zero, the Dynamic Programming Equation (DPE) is obtained. For continuous-tme/continuous-space problems, this DPE takes the form of an HJB PDE. This PDE is a nonlinear, first-order PDE.

The typical approach is to obtain the HJB PDE corresponding to the control problem of interest, and then to apply a numerical method to solve the HJB PDE, thereby obtaining the value function. (Ignoring some technical issues, the value function can yield the optimal control in a feedback form.) The most common approach to solving the HJB PDE is to apply a technique for solution of general PDEs — the finite element method. There are some HJB-specific issues that arise in applying the finite element method to such PDEs (cf. [8], [66]).

Another approach is to consider the characteristic equations associated with the (first-order) HJB PDE. By solving these ordinary differential equations (ODEs), one can, in principle, obtain the value function along the projection of a path of a characteristic curve into the state space. Some serious issues arise due to nonsmoothness. In particular, the projections into state space can cross, or they may not cover the entire space. These roughly correspond to shock and rarefaction waves in conservation laws (cf. [27], [90], [93]).

Max-plus methods work directly with the DPP rather than the limit PDE. As noted above, the DPP corresponds to an operator, and in fact, the time-step parameterized family of DP operators for a problem is a semigroup whose generator is the HJB PDE. The naturalness of the max-plus framework for solution of deterministic control problems (and HJB PDEs) is reflected by the fact that these operators are max-plus linear operators. The solution of time-dependent problems reduces to propagation by the linear operator, while steady-state problems reduce to eigenfunction problems corresponding to the linear operator. This linearity was first noted by Maslov [71].

A key to the exploitation of this linearity is the development of function spaces over the max-plus algebra. (These are often referred to as *moduloids*; see [6], [20] and the references therein.) A well-known similar approach is when linear second-order PDEs are solved by considering Fourier series expansions of the solution. In order to utilize this same approach, one must define the appropriate max-plus spaces, and obtain bases for these spaces. Unfortunately, the lack of additive inverses in the max-plus algebra precludes a direct analogy with Hilbert spaces. However, there are nonetheless useful bases for these max-plus spaces.

If one truncates the basis expansions of elements of a max-plus space, the max-plus linear semigroups correspond to finite-dimensional matrices. Then, approximate propagation reduces to max-plus matrix-vector multiplication, and approximate solution of steady-state problems reduces to finite-dimensional max-plus eigenvector problems. The resulting class of numerical methods is completely new, and is not subsumed by any other existing class of methods. Further, these methods are specifically relevant to control problems and HJB PDEs. Additional classes of methods are obtained by taking max-plus linear combinations of simple, analytically obtainable, max-plus dual operators to approximately generate the max-plus dual operator for the problem of interest. As these classes are entirely new, a single book can only provide some initial foundations for these classes of methods. There is much more that can be done.

In addition to the numerical developments enabled by the max-plus algebra, the algebra provides a new language for formulation of the problems themselves. For instance, the supremum over L_2 functions (with an L_2 cost) can be represented as a max-plus expectation. Solution of deterministic control problems becomes propagation of max-plus conditional expectations. Zero-sum differential games become max-plus stochastic control problems (c.f. Fleming, [39], [40]). Suprema over time are max-plus integrals, and standard-sense integrals are max-plus multiplicative operators.

1.1 Some Control and Estimation Problems

We briefly indicate some control and estimation problems that will be addressed in this book. More rigorous formulations of these problems will appear further below. The chief purpose here is to give some idea of the classes of problems that will be addressable by the machinery to follow. Throughout, the state at time t will be denoted by ξ_t. The state space will be $\mathbf{R}^n$, and points in the state space will be denoted as $x \in \mathbf{R}^n$.

One obvious class of problems appropriate for max-plus analysis are the finite time-horizon optimal control problems. Let the dynamics and initial condition be

$$\dot{\xi} = f(\xi, u),$$
$$\xi_s = x,$$

where u_t is an input process, taking values in the set $U \subseteq \mathbf{R}^k$. This input process may be viewed either as a control input or as a disturbance input. The mathematics for solution of the associated HJB PDE will be the same regardless of the real-world interpretation. We will refer to the space of functions to which this input must belong as the control (or disturbance) space, denoted generically by $\mathcal{U}$. If the finite time-horizon problem is over time interval $[s,T]$ (where T is referred to as the terminal time), then the space of input functions might be designated as $u_\cdot \in \mathcal{U}_{[s,T]}$. For instance, one might have

$$\mathcal{U}_{[s,T]} = \{u : [s,t] \to U \,|\, u \text{ is measurable } \}$$

if U is compact. If U is not compact (e.g., $U = \mathbf{R}^k$), then one might take

$$\mathcal{U}_{[s,T]} = \left\{ u : [s,t] \to U \,\middle|\, \int_{[s,T]} |u_t|^2 \, dt < \infty \right\}.$$

It will generally be of interest to consider optimization of a payoff (or cost criterion) through choice of this input process. In the case where u represents a controller, we may interpret the problem as an optimal control problem with the given cost criterion. Such a criterion would typically take a form such as

$$J(s,x;u) \doteq \int_s^T L(\xi_t,u_t)\,dt + \phi(\xi_T)$$

or, somewhat more generally,

$$J(s,x;u) \doteq \int_s^{T\wedge\tau} L(\xi_t,u_t)\,dt + \phi(\xi_{T\wedge\tau}),$$

where $x \in G \subseteq \mathbf{R}^n$, G satisfies certain conditions, $\tau = \inf\{t \geq s \,|\, \xi_t \notin \overline{G}\}$ and $T \wedge \tau = \min\{T,\tau\}$. In this case, L would be referred to as the running cost, and ϕ would be referred to as a terminal cost or exit cost. Also, τ would be referred to as the exit time. The above payoff is max-plus multiplicative. (This will be discussed later.) Other problem formulations where one takes suprema or limit-suprema over time will also be of interest. If one desired to maximize the payoff, then the value function for this problem would be

$$V(s,x) = \sup_{u\in\mathcal{U}} J(s,x;u).$$

Problems where one maximizes over the control space will be amenable to max-plus methods. If instead, one wished to minimize, the min-plus algebra would be appropriate.

A particularly interesting class of problems will be the robust/H_∞ infinite time-horizon control problems. We will consider first the case where there is a fixed feedback control and an unknown disturbance process. The typical dynamics in such cases would take the form

$$\begin{aligned} \dot{\xi} &= f(\xi,v(\xi)) + \sigma(\xi)u = g(\xi) + \sigma(\xi)u, \\ \xi_0 &= x \in \mathbf{R}^n. \end{aligned} \tag{1.1}$$

Note that, in this case, $v(x)$ is the fixed feedback controller, and the first term in the dynamics, $g(\xi) = f(\xi,v(\xi))$ represents the nominal dynamics — the dynamics in the absence of a disturbance process. It will typically be assumed that $g(0) = 0$ so that the origin is an equilibrium point. Also, in this case, u represents an input disturbance process. If the range space of u were $\mathbf{R}^k$, then the disturbance space (also referred to as the control space for u) would typically be

$$\mathcal{U} = L_2^{loc}([0,\infty);\mathbf{R}^k) = \{u : [0,\infty) \to \mathbf{R}^k \,|\, u_{[0,T]} \in L_2(0,T)\ \forall T \in [0,\infty)\} \tag{1.2}$$

where $L_2(0,T)$ is the set of square-integrable functions over $[0,T]$ where the range space will not be included in the notation when there is no possibility of confusion. This system is said to satisfy an H_∞ criterion with disturbance attenuation parameter $\gamma \in (0,\infty)$ if there exists a locally bounded $\beta : \mathbf{R}^n \to [0,\infty)$ where $\beta(0) = 0$ such that

$$\int_0^T L(\xi_t)\,dt \le \beta(x) + \frac{\gamma^2}{2}\|u\|^2_{[0,T]} \doteq \beta(x) + \frac{\gamma^2}{2}\int_0^T |u_t|^2\,dt \tag{1.3}$$

for all $T \in [0,\infty)$ and all $u \in \mathcal{U}$. Here, $L(\cdot)$ is assumed nonnegative, and more specifically, *strictly positive definite* (i.e. $L(0) = 0$ and $L(x) > 0$ otherwise), and often takes the form of a quadratic. Note that a function is *locally bounded* if it is bounded on compact sets. The associated value function (known as the available storage [12], [52], [108]) is given by

$$\begin{aligned} W(x) &= \sup_{u\in\mathcal{U}} \sup_{T\in[0,\infty)} \int_0^T L(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \\ &= \sup_{T\in[0,\infty)} \sup_{u\in\mathcal{U}} \int_0^T L(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \\ &= \lim_{T\to\infty} \sup_{u\in\mathcal{U}} \int_0^T L(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt. \end{aligned} \tag{1.4}$$

Under the indicated assumptions, one has $W(0) = 0$ and $W(x) > 0$ if $x \neq 0$.

An H_∞ problem with *active* control which will be considered has dynamics

$$\begin{aligned} \dot\xi &= A(\xi) + B(\xi)v + \sigma(\xi)u, \\ \xi_0 &= x \in \mathbf{R}^n, \end{aligned}$$

where A is a vector-valued function of length n, B is an $n \times m$ matrix-valued function, and σ is an $n \times k$ matrix-valued function. The payoff is

$$J(x,T;v,u) = \tfrac{1}{2}\int_0^T \xi_t^T C\xi_t + v_t D v_t - \tfrac{\gamma^2}{2}|u_t|^2\,dt,$$

where C and D are symmetric positive definite matrices. The active control H_∞ problem is a zero-sum, differential game with this payoff. The value function can be expressed as an Elliott–Kalton [35] game value. In particular, the value of interest is

$$W(x) = \inf_{\theta\in\Theta} \sup_{u\in\mathcal{U}} \sup_{T\in[0,\infty)} J(x,T;\theta[u],u),$$

where Θ is the set of nonanticipative mappings from the disturbance space, $\mathcal{U}$, into the control space. Nonanticipativity and game values will be discussed at more length below. Under a certain condition, this problem will be addressable with techniques employing the min-plus algebra.

A class of problems that have come under study in recent years (c.f. [55], [56]) are ones with mixed L_∞/L_2 criteria. The dynamics may take the form

$\dot{\xi} = g(\xi) + \sigma(\xi)u$, $\xi_0 = x \in \mathbf{R}^n$. In this case, however, the criteria take a more general form such as

$$J(x,T,u) = \ell(\xi_T) + \int_0^T L(\xi_t) - \eta(|u_t|)\,dt,$$

and one may sometimes have an added constraint that $|u_t| \le M < \infty$ for all $t \ge 0$. The η function typically satisfies $\eta(0) = 0$ with η monotonically increasing, and one would have $L, \ell \ge 0$. One considers a problem,

$$W(x) \doteq \sup_{T\ge 0} \sup_{u\in\mathcal{U}} J(x,T,u).$$

Note that due to the presence of the ℓ term, the supremum over time may be achieved at some finite value (rather than as $T \to \infty$ in the above case), and so this is referred to as a stopping-time problem.

These same techniques can be applied to the robust/H_∞ filtering problem. In this problem, the dynamics are

$$\dot{\xi} = g(\xi) + \sigma(\xi)u,$$
$$\xi_s = x \in \mathbf{R}^n,$$

where x is unknown. The disturbance process, u_t, is unknown. There may be an observation process. If the observations occur at discrete times, t_n, then the observation process may be

$$y_n = h(\xi_{t_n}) + \rho(\xi_{t_n})w_n$$

where the w_n represent unknown disturbances in the observation process. One looks for an estimator, $\overline{x}_t$ such that

$$\int_s^t (\xi_r - \overline{x}_r)^T C(\xi_r - \overline{x}_r)\,dr \le \phi(\xi_s) + \int_s^t \frac{\gamma_1^2}{2}|u_r|^2\,dr + \sum_{n=o}^{N_t} \frac{\gamma_2^2}{2}|w_n|^2$$

where C is positive definite and $N_t = \max\{n|\, t_n \le t\}$. Then $\overline{x}_t$ will be a robust estimator. This will also be discussed more fully below.

1.2 Concepts of Max-Plus Methods

We outline the basic concepts behind the classes of max-plus methods. Consider the value function given by (1.4) with dynamics (1.1) and control/disturbance space (1.2). The value function, $W(x)$, is a function of the initial state. A common approach to solving such a problem is dynamic programming, or DP (see for instance [8], [46], [47] and [48]). DP methods generally lead to a dynamic programming equation (DPE) which in this case takes the form of a steady-state, nonlinear, first-order HJB PDE. In the case of this example, the HJB PDE is

$$\begin{aligned} 0 &= -\sup_{u\in\mathbf{R}^m}\left\{[g(x)+\sigma(x)u]^T\nabla W + L(x) - \tfrac{\gamma^2}{2}|u|^2\right\} \\ &= -\left[\tfrac{1}{2\gamma^2}(\nabla W)^T\sigma(x)\sigma^T(x)\nabla W + g^T(x)\nabla W + L(x)\right]. \end{aligned} \tag{1.5}$$

The boundary condition is that the solution is zero at the origin, that is

$$W(0) = 0, \tag{1.6}$$

where we note that 0 will be used to represent the origin in $\mathbf{R}^n$ for any $n \geq 1$ throughout the text. Since this is a first-order, nonlinear PDE, one cannot generally expect a smooth solution. Among the class of weak solutions, the most useful class appears to be that of viscosity solutions (c.f. [8], [22], [23], [47], [48], [57]). For most HJB PDEs, there exists a unique viscosity solution. However, for problems such as (1.5), (1.6), one may not have uniqueness among the class of viscosity solutions, and in such cases an additional condition is needed in order to uniquely specify the "correct" viscosity solution [88], [106]. Here, the "correct" solution is the value function, W, given by (1.4).

Once one has (1.5), a PDE solver is used to compute the solution. The most common approaches to numerical solution of PDEs are the finite element (FE) methods. A number of authors have developed FE methods specifically adapted to solution of first-order HJB PDEs (c.f. [8], [18], [32], [34], [66]). In all cases, the methods suffer from the curse-of-dimensionality. In particular, one must solve PDE (1.5) over $\mathbf{R}^n$ if the state takes values in $\mathbf{R}^n$. Consequently, the computational cost grows exponentially in the state-space dimension n — this computational growth being the aforementioned curse-of-dimensionality.

Other approaches have been applied. One class is that of characteristics (bicharacteristics, generalized characteristics). In this case, one computes $W(x)$ and the optimal feedback, $u^*(x)$, by solution of a $2n$-dimensional second-order ordinary differential equation (ODE) with two boundary conditions. Such a method may be appropriate for real-time computation, as well as for precomputation and storage. Unfortunately, the nonsmoothness property makes it difficult to propagate the "correct" solution of the ODE since the propagation across hypersurfaces of nonsmoothness can be quite technical. Further, one must apply a shooting approach to deal with the two-point boundary conditions. Lastly, for (1.5) in particular, there are typically both stable and unstable manifolds emerging from the origin, and shooting to a boundary condition such as $\lim_{t\to\infty}\zeta(t) = 0$ is extremely difficult. For discussion of such methods, one can see [27], [37], [49], [90], [72], [93].

With max-plus methods, one does not work with the HJB PDE, but with the semigroup associated with the HJB PDE. More specifically, instead of applying DP to derive the PDE, one works with the DPP prior to taking the infinitesimal limit in time. The DP associated with (1.4) is

$$W(x) = \sup_{u \in \mathcal{U}} \left[\int_0^\tau L(\xi_t) - \tfrac{\gamma^2}{2} |u_t|^2 \, dt + W(\xi_\tau) \right] \tag{1.7}$$

with dynamics given by (1.1). The right-hand side is an operator on the space that W lies in. (Such spaces will be considered in more detail later.) Suppose W is in the space of semiconvex functions, $\mathcal{S}$. For any $\tau \geq 0$, (1.7) takes the form

$$W = S_\tau[W],$$

where S_τ maps some domain $\mathcal{D}(S_\tau) \subseteq \mathcal{S}$ into $\mathcal{S}$, and for any $\phi \in \mathcal{D}(S_\tau)$,

$$S_\tau[\phi] \doteq \sup_{u \in \mathcal{U}} \left[\int_0^\tau L(\xi_t) - \frac{\gamma^2}{2} |u_t|^2 \, dt + \phi(\xi_\tau) \right].$$

Note that the operator S_τ is indexed by τ; that is, the S_τ form a one-parameter family of operators indexed by time, τ. Further, S_τ satisfies the semigroup properties [96]:

$$S_{\tau_1 + \tau_2} = S_{\tau_1} S_{\tau_2} \quad \forall \tau_1, \tau_2 > 0, \tag{1.8}$$

$$S_0 = I, \tag{1.9}$$

where I represents the identity operator. There are multiple classes of continuous semigroups [96] obtained by invoking various continuity requirements with respect to the time variable, but continuity issues will not be the focus here. We do note that the corresponding HJB PDE (1.5) is the generator of the semigroup.

Before proceeding with this introduction to max-plus methods, we need to briefly discuss the max-plus algebra. This will be discussed in detail in Chapter 2. Here, we include only the information needed to continue this informal discussion. The max-plus algebra is a commutative semifield over $\mathbf{R}^- \doteq \mathbf{R} \cup \{-\infty\}$. The addition and multiplication operations, $\oplus$ and $\otimes$ are defined for $a, b \in \mathbf{R}^-$ as

$$a \oplus b = \max\{a, b\}, \qquad a \otimes b = a + b.$$

The additive identity is $-\infty$, and the multiplicative identity is 0. The multiplicative inverse of $a > -\infty$ is $-a$. Note that one does not have additive inverses — this being the reason for the appellation "semifield" rather than field. As with standard vector spaces (i.e., spaces over the usual addition/multiplication operations), the max-plus addition of two functions (or two vectors) is done pointwise. Specifically, for functions $\phi_1(x)$ and $\phi_2(x)$, the max-plus sum is $[\phi_1 \oplus \phi_2](x) = \phi_1(x) \oplus \phi_2(x) = \max[\phi_1(x), \phi_2(x)]$ for all x. Multiplication by a scalar is analogous to the standard case as well. That is, for $a \in \mathbf{R}^-$ and function $\phi_1(x)$, $[a \otimes \phi_1](x) = a \otimes \phi_1(x) = a + \phi_1(x)$ for all x.

We will find that $\mathcal{S}$ is a max-plus vector space, and in particular, the above addition and multiplication operations of the previous paragraph will be well-defined for elements of $\mathcal{S}$. A fundamental issue will be that for $a_1, a_2 \in \mathbf{R}^-$ and $\phi_1, \phi_2 \in \mathcal{D}(S_\tau) \subseteq \mathcal{S}$,

$$S_\tau[a_1 \otimes \phi_1 \oplus a_2 \otimes \phi_2] = a_1 \otimes S_\tau[\phi_1] \oplus a_2 \otimes S_\tau[\phi_2]. \tag{1.10}$$

In other words, S_τ is a max-plus linear operator for each $\tau \geq 0$. The proof will appear in Chapter 4. The max-plus methods will seek to exploit this linearity.

As indicated above, max-plus methods may require basis expansions of the solutions (as elements of some max-plus vector space). Suppose $W(x) = \bigoplus_i e_i \otimes \psi_i(x)$ for all x where $\{\psi_i\}_{i=1}^\infty$ forms a max-plus basis (in a sense to be defined in Chapter 2) for $\mathcal{S}$. Then one has

$$\bigoplus_i e_i \otimes \psi_i = S_\tau\left[\bigoplus_i e_i \otimes \psi_i\right].$$

If only a finite number of basis functions, say N, were required, the max-plus linearity of S_τ would yield

$$\bigoplus_{i=1}^N e_i \otimes \psi_i = \bigoplus_{i=1}^N e_i \otimes S_\tau\left[\psi_i\right]. \tag{1.11}$$

The errors introduced by truncation to a finite number of elements in a basis expansion will be considered in detail. Equation (1.11) is typically equivalent to

$$0 \otimes \mathbf{e} = \mathbf{e} = B \otimes \mathbf{e}, \tag{1.12}$$

where $\mathbf{e}$ is the vector of length N with components e_i, and B is a matrix associated with S_τ. In other words, approximate solution of (1.5)/(1.4) reduces to solution of max-plus eigenvector problem (1.12). Similarly, propagation forward in time of time-dependent HJB PDEs takes the form

$$\mathbf{a}_{t+\tau} = B \otimes a_t$$

for time-step τ. Lastly, the mixed L_∞/L_2 problems lead to max-plus affine problems of the form

$$\mathbf{e} = B \otimes \mathbf{e} \oplus \mathbf{c}.$$

In this case, the problem data generate both matrix B and vector $\mathbf{c}$; one then solves this max-plus affine problem for $\mathbf{e}$. Note that this max-plus approach does not obviously remove the curse-of-dimensionality. If one expects the number of necessary basis vectors to grow exponentially in space dimension, n, then the computational costs will also grow exponentially in n. It should be noted that the entire matrix B is not generally needed — only an analogue to the terms in a banded matrix are needed. The expected advantage of such an approach follows from the expectation that exploitation of the linearity may lead to a lower computational cost and lower computational growth rate.

Some other max-plus methods will be mentioned. Specifically, some approaches related to semiconvex duality and the Legendre/Fenchel transform will be discussed. It should be noted that the Legendre/Fenchel transform plays a role similar to the Fourier and/or Laplace transform for standard-sense algebra. In fact, a particularly promising class of methods stems from taking max-plus linear combinations of transformed operators for simple problems to obtain approximate transformed operators for the problem of interest. If the simple problems are, say linear, then the transformed operators can be obtained analytically. This approach entirely avoids the curse-of-dimensionality for a class of HJB PDEs — which is a particularly exciting development.

From the above discussion, one sees several pieces of machinery that are needed for development of max-plus methods. The structures of max-plus spaces are discussed in Chapter 2. One also needs to prove that the value functions lie in appropriate max-plus vector spaces (or min-plus spaces in some cases). The ones that will be considered in the most detail will be spaces of convex and semiconvex functions. Secondly, one needs the DPP/semigroups and proofs of their max-plus linearity (the latter being rather trivial). In order to make rigorous statements about the quality of solutions to be computed with max-plus methods, one needs to understand convergence rates and error bounds for these methods. Exploration of semiconvex duality and Legendre/Fenchel transforms will obviously also prove useful.

The machinery needed for the various problem forms of Section 1.1 is somewhat uniform across problem classes. The most thorough discussion will be for the H_∞ problem (1.4), where in particular, a full error analysis will be provided. The machinery for other problem forms and solution methods will be discussed after the H_∞ problem form, and not all the technical issues will be explored in detail in those cases.

2

Max-Plus Analysis

In this chapter we introduce the max-plus algebra and some basics of functional analysis over the max-plus algebra. We will refer to our usual operations (addition and multiplication) on the real line as the standard field. The max-plus algebra is a commutative semifield over $\mathbf{R}^- = \mathbf{R} \cup \{-\infty\}$. The addition and multiplication operations are

$$\begin{aligned} a \oplus b &= \max\{a, b\}, \\ a \otimes b &= a + b. \end{aligned} \tag{2.1}$$

In particular we take

$$\begin{aligned} a \oplus -\infty &= a \quad \forall\, a \in \mathbf{R}^-, \\ a \otimes -\infty &= -\infty \quad \forall\, a \in \mathbf{R}^-. \end{aligned} \tag{2.2}$$

As evidenced by (2.2), the additive identity is $-\infty$. The multiplicative identity is 0 ($a \otimes 0 = a + 0 = a \;\; \forall a \in \mathbf{R}^-$). The multiplicative inverse of $a > -\infty$ is $-a$ (the standard additive inverse). The additive identity, $-\infty$, does not have a multiplicative inverse (which is analogous to the same property in the standard field). With the exception of the additive identity, $-\infty$, no elements have additive inverses. In [6], an approach to dealing with this lack of inverses in algebraic equations is resolved through the use of "balances."

We note that the commutative, associative and distributive properties hold:

$$\begin{aligned} & a \oplus b = b \oplus a, \\ & a \oplus (b \oplus c) = (a \oplus b) \oplus c, \\ & a \otimes b = b \otimes a, \\ & a \otimes (b \otimes c) = (a \otimes b) \otimes c, \\ & a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c). \end{aligned}$$

The usual precedence ordering between operations will be employed, that is

$$a \otimes b \oplus c \otimes d = (a \otimes b) \oplus (c \otimes d).$$

For more information on the max-plus algebra, the reader is referred to [6], [25], [63], [70], and the references therein.

We say that a sequence $\{a_n\}_{n=1}^{\infty} \subset \mathbf{R}^-$ converges to $a \in \mathbf{R}$ (denoted as $a_n \to a$) if $|a_n - a| \to 0$ where $|\cdot|$ represents the (standard) absolute value, and we say $a_n \to -\infty$ if given $M < \infty$, there exists $N_M < \infty$ such that $a_n < -M$ for all $n > N_M$. Equivalently, we say that $\mathbf{R}^-$ is equipped with its *usual topology* given by the metric $d^-(a,b) \doteq |e^a - e^b|$. This also relates to log-plus algebras and large deviations, c.f. [1], [39], [82], [97], and [98].

We now define a vector space over the max-plus algebra; this will generally be referred to as a max-plus vector space or a max-plus space. These are referred to as moduloids in [6]. We say $\mathcal{X}$ is a max-plus vector space (with zero element denoted by $\phi^0 \in \mathcal{X}$) if

$$\begin{aligned}
&a \otimes \phi \in \mathcal{X} \quad \forall a \in \mathbf{R}^-,\ \forall \phi \in \mathcal{X},\\
&\phi \oplus \psi = \psi \oplus \phi \in \mathcal{X} \quad \forall \phi, \psi \in \mathcal{X},\\
&(a \otimes b) \otimes \phi = a \otimes (b \otimes \phi) \quad \forall a, b \in \mathbf{R}^-,\ \forall \phi \in \mathcal{X},\\
&(a \oplus b) \otimes \phi = a \otimes \phi \oplus b \otimes \phi \quad \forall a, b \in \mathbf{R}^-,\ \forall \phi \in \mathcal{X},\\
&a \otimes (\phi \oplus \psi) = a \otimes \phi \oplus a \otimes \psi \quad \forall a \in \mathbf{R}^-,\ \forall \phi, \psi \in \mathcal{X}\\
&\phi \oplus \phi^0 = \phi,\ a \otimes \phi^0 = \phi^0,\ -\infty \otimes \phi = \phi^0,\ 0 \otimes \phi = \phi \quad \forall a \in \mathbf{R}^-,\ \forall \phi \in \mathcal{X}.
\end{aligned}$$

If $\mathcal{X}$ is defined as a set of vectors of elements of $\mathbf{R}^-$ indexed by $\lambda \in \Lambda$ for some index set Λ, then we may denote elements of $\mathcal{X}$ as $\phi = \{\phi_\lambda\}_{\lambda \in \Lambda}$, each element being denoted by ϕ_λ. In some cases the notation $\phi(\lambda)$ may be used in place of ϕ_λ.

We say the sequence $\{\phi^n\}_{n=1}^{\infty}$ converges to $\overline{\phi} \in \mathcal{X}$ (denoted as $\phi^n \to \overline{\phi}$) if $\phi^n_\lambda \to \overline{\phi}_\lambda$ for all $\lambda \in \Lambda$. Further, if given any $\varepsilon > 0$, there exists $N_\varepsilon < \infty$ such that $|\phi^n_\lambda - \overline{\phi}_\lambda| < \varepsilon$ for all $\lambda \in \Lambda$ and all $n \geq N_\varepsilon$, we say ϕ^n converges uniformly to $\overline{\phi}$. Let $-\infty.$ denote the element of $\mathcal{X}$ such that $\phi_\lambda = -\infty \in \mathbf{R}^-$ for all $\lambda \in \Lambda$. Note that since $-\infty \in \mathbf{R}^-$ is the additive identity, this is somewhat analogous to the origin in $\mathcal{X}$. In particular, ϕ^0 appearing in the above definition will be $-\infty.$. We say $\phi^n \to -\infty. \in \mathcal{X}$ uniformly if given $M < \infty$, there exists $N_M < \infty$ such that $\phi^n_\lambda \leq -M$ for all $\lambda \in \Lambda$ and all $n \geq N_M$.

Example 2.1. Let $B_R(0) \subset \mathbf{R}^n$ be the ball of radius R, that is $B_R(0) = \{y \in \mathbf{R}^n \,|\, |y| < R\}$ where $|\cdot|$ indicates Euclidean norm. The notation B_R will be used freely in place of $B_R(y)$ when y is the origin. Let $\mathcal{X}$ be the space of continuous functions mapping $B_R = B_R(0)$ into $\mathbf{R}$ appended by $-\infty.$. Then for any $\phi_1, \phi_2 \in \mathcal{X}$ and any $a \in \mathbf{R}^-$, $(a \otimes \phi)(y) = a + \phi(y)$ and $(\phi_1 \oplus \phi_2)(y) = \max\{\phi_1(y), \phi_2(y)\}$ for all $y \in B_R$. Clearly $\mathcal{X}$ is closed under these operations, and is a max-plus vector space. This is also true with $B_R(0)$ replaced by its closure, $\overline{B}_R(0)$.

Example 2.2. The natural definition of continuous functions with range in metric space $(\mathbf{R}^-, d^-)$ applies here. More specifically, $\phi : \mathbf{R}^n \to \mathbf{R}^-$ is continuous at $y_0 \in \mathbf{R}^n$ if given any sequence $y_n \to y_0$, $\phi(y_n) \to \phi(y_0)$. Then the space of continuous functions from $\mathbf{R}^n$ (or appropriate subset of $\mathbf{R}^n$) into $\mathbf{R}^-$ is also a max-plus vector space.

We say that $\{\phi^n\}_{n=1}^{\infty} \subset \mathcal{X}$ is Cauchy if given $\varepsilon > 0$, there exists $N_\varepsilon < \infty$ such that $|\phi^n_\lambda - \phi^m_\lambda| < \varepsilon$ for all $\lambda \in \Lambda$ and all $n, m \geq N_\varepsilon$. $\mathcal{X}$ is *standard-complete* if given any Cauchy sequence $\{\phi^n\}_{n=1}^{\infty}$, there exists $\overline{\phi} \in \mathcal{X}$ such that $\phi^n \to \overline{\phi}$.

Example 2.3. Let $\mathcal{X}_K$ be the space of real-valued functions over $B_R = B_R(0)$ which are all Lipschitz with the same constant, $K < \infty$, with $-\infty$. appended. Then $\mathcal{X}_K$ is a standard-complete max-plus vector space. To see that $\mathcal{X}_K$ is closed under the max-plus operations, let $\phi \doteq a \otimes \phi^1 \oplus b \otimes \phi^2$ where $a, b \in \mathbf{R}$ (the case of a or b being $-\infty$ being trivial) and $\phi^1, \phi^2 \in \mathcal{X}_K$. Fix $y_0 \in B_R$. Suppose $a \otimes \phi^1(y_0) \geq b \otimes \phi^2(y_0)$. Then for any $y \in B_R$,

$$\phi(y_0) - \phi(y) \leq [a + \phi^1(y_0)] - [a + \phi^1(y)] = \phi^1(y_0) - \phi^1(y) \leq K|y - y_0|.$$

Using symmetry, one sees that $\mathcal{X}_K$ is closed. Completeness follows in the usual way, and a proof is not included.

Let $C \subseteq \mathbf{R}^n$ be convex. (C is convex if $y^1, y^2 \in C$ and $\lambda \in [0, 1]$ imply $\lambda y^1 + (1 - \lambda) y^2 \in C$.) Recall that $\phi : C \to \mathbf{R}^-$ is convex if for any $y^1, y^2 \in C$ and any $\lambda \in [0, 1]$, $\phi(\lambda y^1 + (1 - \lambda) y^2) \leq \lambda \phi(y^1) + (1 - \lambda)\phi(y^2)$ where these operations are standard-sense [53], [101], [102]. Note that if $\phi(y^3) = -\infty$ for some $y^3 \in C$, then one must have $\phi \equiv -\infty$..

Example 2.4. Let $\mathcal{X}$ be the set of convex functions mapping the convex domain $C \subseteq \mathbf{R}^n$ into $\mathbf{R}$, with the function $-\infty$. appended. Then $\mathcal{X}$ is a max-plus space. If one adds the additional requirement that each $\phi \in \mathcal{X}$ be Lipschitz with some constant, $K < \infty$, $\mathcal{X}$ remains a max-plus space.

2.1 Spaces of Semiconvex Functions

Various spaces of semiconvex functions will be of particular interest. These will be max-plus vector spaces. In the following definition, the (standard-sense) closure of any set $G \subseteq \mathbf{R}^n$ is denoted by $\overline{G}$.

Definition 2.5. *A function, $\phi : \mathbf{R}^n \to \mathbf{R}^-$ is semiconvex (over $\mathbf{R}^n$) if given any $R < \infty$, there exists $c < \infty$ such that $\phi(y) + \frac{c}{2}|y|^2$ is convex over $B_R(0)$. Fix any $R > 0$, and consider $B_R(0)$ (or $\overline{B}_R(0)$). We say $\phi : B_R(0) \to \mathbf{R}^-$ ($\phi : \overline{B}_R(0) \to \mathbf{R}^-$) is semiconvex over $B_R(0)$ ($\overline{B}_R(0)$) with constant c if $\phi(y) + \frac{c}{2}|y|^2$ is convex over $B_R(0)$ ($\overline{B}_R(0)$).*

Suppose ϕ is semiconvex over $\mathbf{R}^n$. Fix $R \in (0,\infty)$. Let

$$c_R \doteq \inf\{c \in \mathbf{R} \,|\, \phi(y) + \tfrac{c}{2}|y|^2 \text{ is convex over } B_R(0)\}.$$

We will refer to c_R as the minimal semiconvexity constant for ϕ over B_R, and we will say ϕ is semiconvex over B_R with minimal semiconvexity constant c_R. In particular, $c_R = -\infty$ in the case where ϕ is the function $-\infty$..

Lemma 2.6. *Suppose ϕ is semiconvex over B_R with minimal semiconvexity constant c_R. Then $\phi(y) + \frac{c_R}{2}|y|^2$ is convex.*

Proof. Let $y_1, y_2 \in B_R$, $\alpha \in [0,1]$ and $\varepsilon > 0$. Let $c \in (c_R, c_R + \varepsilon/R^2)$. Then

$$\begin{aligned}
&\phi(\lambda y_1 + (1-\lambda)y_2) + \frac{c_R}{2}|\lambda y_1 + (1-\lambda)y_2|^2 \\
&\quad < \phi(\lambda y_1 + (1-\lambda)y_2) + \frac{c}{2}|\lambda y_1 + (1-\lambda)y_2|^2
\end{aligned}$$

which by convexity (of the entire right-hand side)

$$\begin{aligned}
&\le \lambda\phi(y_1) + (1-\lambda)\phi(y_2) + \lambda\frac{c}{2}|y_1|^2 + (1-\lambda)\frac{c}{2}|y_2|^2 \\
&< \lambda\phi(y_1) + (1-\lambda)\phi(y_2) + \lambda\frac{c_R}{2}|y_1|^2 + (1-\lambda)\frac{c_R}{2}|y_2|^2 + (c - c_R)R^2 \\
&< \lambda\left[\phi(y_1) + \frac{c_R}{2}|y_1|^2\right] + (1-\lambda)\left[\phi(y_2) + \frac{c_R}{2}|y_2|^2\right] + \varepsilon.
\end{aligned}$$

Since this is true for all $\varepsilon > 0$, one obtains the result. □

Let $\mathcal{S}$ denote the space of semiconvex functions mapping $\mathbf{R}^n$ into $\mathbf{R}^-$. Let $\mathcal{S}_R$ be the space of semiconvex functions mapping $B_R(0) \subset \mathbf{R}^n$ into $\mathbf{R}^-$. Let $\mathcal{S}_R^c$ be the space of semiconvex functions mapping $B_R \subset \mathbf{R}^n$ into $\mathbf{R}^-$ with semiconvexity constant, c. Let $\mathcal{S}_R^{cL}$ be the space of functions mapping $B_R \subset \mathbf{R}^n$ into $\mathbf{R}^-$ which are semiconvex with constant c and Lipschitz with constant L. Note that if ϕ is in $\mathcal{S}$, $\mathcal{S}_R$, $\mathcal{S}_R^c$ or $\mathcal{S}_R^{cL}$, and there exists a y_0 in the domain such that $\phi(y_0) = -\infty$, then $\phi(y) = -\infty$ for all y in the domain. We will also find it convenient to work with spaces of functions over the closed ball. Let the mnemonically handy notation $\mathcal{S}_{\bar{R}}$, $\mathcal{S}_{\bar{R}}^c$ and $\mathcal{S}_{\bar{R}}^{cL}$ represent the obvious analogues of the spaces $\mathcal{S}_R$, $\mathcal{S}_R^c$ and $\mathcal{S}_R^{cL}$ but where the domain is the closed ball, $\overline{B}_R = \overline{B}_R(0)$ rather than the open ball $B_R(0)$.

Theorem 2.7. *Let $R, c, L \in (0,\infty)$. Then $\mathcal{S}$, $\mathcal{S}_R$, $\mathcal{S}_R^c$, $\mathcal{S}_R^{cL}$, $\mathcal{S}_{\bar{R}}$, $\mathcal{S}_{\bar{R}}^c$ and $\mathcal{S}_{\bar{R}}^{cL}$ are max-plus vector spaces.*

Proof. We consider only the space $\mathcal{S}_R^{cL}$; proofs for the other spaces are similar. Let $\phi^1, \phi^2 \in \mathcal{S}_R^{cL}$, and $a_1, a_2 \in \mathbf{R}^-$. One must show $a_1 \otimes \phi^1 \oplus a_2 \otimes \phi^2 \in \mathcal{S}_R^{cL}$. The case with a_1 or a_2 being $-\infty$ is trivial; suppose not. Note that $[a_1 \otimes \phi^1 \oplus a_2 \otimes \phi^2](y) = \max\{a_1 + \phi^1(y), a_2 + \phi^2(y)\}$. One easily sees that adding a constant to a semiconvex or Lipschitz function does not affect the semiconvexity or Lipschitz constants. All that remains is to note that the pointwise maximum of two functions which are semiconvex with constant c and Lipschitz with constant L, is also semiconvex with constant c and Lipschitz with constant L. This is not difficult, and the proof is not included. □

It is useful when working with max-plus spaces (and more generally, idempotent spaces, that is, spaces over algebras for which $a \oplus a = a$) to use certain notions of completeness which are quite natural in this setting. Let $\leq$ denote the natural order on $\mathbf{R}^-$ (e.g., $7 \leq 18$). Let $A \subseteq \mathbf{R}^-$. We let the supremum of A, $\sup(A)$, be the standard definition, that is, the least upper bound. A semifield is *b-complete* if any nonempty set which is bounded above has a supremum. Clearly, $\mathbf{R}^-$ is b-complete. We say a semifield is *complete* (also known as *a-complete*) if any nonempty set has a supremum. The max-plus algebra, which we can denote here as $(\mathbf{R}^-, \oplus, \otimes)$ can be completed by the addition of a top element, $+\infty$, where $a \leq +\infty$ for all $a \in \mathbf{R}^-$., and we let $\mathbf{R}^{-+} \doteq \mathbf{R}^- \cup \{+\infty\}$. Then, $(\mathbf{R}^{-+}, \oplus, \otimes)$ is a complete semifield. We define $-\infty = \sup(\emptyset)$ where $\emptyset$ is the empty set. Note that since $-\infty$ is absorbing, one needs to set

$$-\infty \otimes +\infty = -\infty.$$

2.2 Bases

The max-plus numerical methods which have been developed for solution of HJB PDEs and nonlinear control rely on basis expansions in max-plus spaces and/or transforms between max-plus spaces. Heuristically, this approach uses max-plus analogues of spectral methods, Fourier/Laplace series and Fourier/Laplace transforms. The starting point is the development of basis expansions for max-plus spaces such as $\mathcal{S}^{cL}_{\bar{R}}$. Note that we will not have the notion of a complete orthonormal set which one has in Hilbert spaces. Instead we will use the notion of a "countable max-plus basis" (where this concept will be clarified in Definition 2.12). We begin the development by recalling a core result from convex analysis ([53], [101], [102]). This general result appears in a number of contexts, and is referred to as the Fenchel transform, convex duality, or the Legendre transform (it actually being essentially an extension of the classical Legendre transform). Let $\mathcal{C}^{L}_{\bar{R}}$ be the space of mappings from $\overline{B}_R$ into $\mathbf{R}^-$ which are convex and Lipschitz with Lipschitz constant $L < \infty$. Let $\overline{B}_L(0) \subset \mathbf{R}^n$. The following is a version of convex duality, but simplified due to the Lipschitz condition.

Theorem 2.8. *Let $\widehat{\phi} \in \mathcal{C}^{L}_{\bar{R}}$. Then for all $y \in \overline{B}_R$,*

$$\widehat{\phi}(y) = \max_{p \in \overline{B}_L} \left[p^T y + \widehat{\psi}(p) \right] = \max_{p \in G} \left[p^T y + \widehat{\psi}(p) \right] \tag{2.3}$$

for any $G \supseteq \overline{B}_L$, where

$$\widehat{\psi}(p) = -\max_{y\in\overline{B}_R}\left[y^T p - \widehat{\phi}(y)\right] \tag{2.4}$$

for all $p \in \mathbf{R}^n$*. If* $\widehat{\phi} = -\infty.$*, then* $\widehat{\psi} = -\infty.$*.*

The reader is referred to one of the many works on convex analysis for a proof and context. See, for instance, [100], Section 12; [101], Section 3; [53]; and/or [102]. A sketch of a proof is included in the appendix as an aid. Note that $-\widehat{\psi}$ is convex, and may be referred to as the convex conjugate, or sometimes the convex dual, of $\widehat{\phi}$. In particular, (2.3), (2.4) are more often written as

$$\widehat{\phi}(y) = \max_{p\in\overline{B}_L}\left[p^T y - \widetilde{\psi}(p)\right], \qquad \widetilde{\psi}(p) = \max_{y\in\overline{B}_R}\left[y^T p - \widehat{\phi}(y)\right], \tag{2.5}$$

where $\widetilde{\psi} = -\widehat{\psi}$.

This duality can be expanded to spaces of semiconvex functions, and this will form one of the fundamental building blocks for the max-plus numerical methods. Recall that $\mathcal{S}^{cL}_{\overline{R}}$ is the space of functions mapping $\overline{B}_R$ into $\mathbf{R}^-$ which are semiconvex and Lipschitz with (positive) constants, c and L, respectively. The following result from [78] is a minor variant of a result which first appeared (at least in this context) in [44].

Theorem 2.9. *Let* $\phi \in \mathcal{S}^{cL}_{\overline{R}}$*. Let* C *be a symmetric matrix such that* $C - cI \geq 0$ *(i.e., such that* $C - cI$ *is non-negative definite). Let* $D_R \geq |C||C^{-1}|R + |C^{-1}|L$ *where* $|C|$ *is the induced matrix norm of* C*. (In particular, one may take* $D_R = (|C|R + L)/c$*.) Then, for all* $y \in \overline{B}_R$*,*

$$\phi(y) = \max_{|C\overline{y}|\leq(L+|C|R)}\left[\psi(\overline{y}) - \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\right] \tag{2.6}$$

$$= \max_{\overline{y}\in\overline{B}_{D_R}}\left[\psi(\overline{y}) - \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\right] \tag{2.7}$$

$$= \max_{\overline{y}\in G}\left[\psi(\overline{y}) - \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\right] \tag{2.8}$$

$$\tag{2.9}$$

for any $G \supseteq \{y : |C\overline{y}| \leq (L + |C|R)\}$*, where*

$$\psi(\overline{y}) = -\max_{y\in\overline{B}_R}\left[-\phi(y) - \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\right] \tag{2.10}$$

for all $\overline{y} \in \mathbf{R}^n$*. Further, if* $C - cI > 0$ *(i.e., if* $C - cI$ *is positive definite), then* $\psi(\overline{y}) > -\infty$ *for all* $\overline{y}$*.*

Proof. The case $\phi = -\infty.$ (whence $\psi = -\infty.$) is trivial; we suppose $\phi \neq -\infty.$. Define $\widehat{\phi} : \overline{B}_R \to \mathbf{R}^-$ by

$$\widehat{\phi}(y) = \phi(y) + \tfrac{1}{2}y^T C y. \tag{2.11}$$

First it will be shown that $\widehat{\phi}$ is convex (i.e., $\widehat{\phi}(\lambda y_1 + (1-\lambda)y_2) < \lambda\widehat{\phi}(y_1) + (1-\lambda)\widehat{\phi}(y_2)$ for all $y_1, y_2 \in \overline{B}_R$, $y_1 \neq y_2$, $\lambda \in (0,1)$). Note that

$$\begin{aligned}\widehat{\phi}(\lambda y_1 + (1-\lambda)y_2) &= \phi(\lambda y_1 + (1-\lambda)y_2) + \tfrac{1}{2}c|\lambda y_1 + (1-\lambda)y_2|^2 \\ &\quad + \tfrac{1}{2}(\lambda y_1 + (1-\lambda)y_2)^T[C - cI](\lambda y_1 + (1-\lambda)y_2)\end{aligned}$$

which by the convexity of $\phi(\cdot) + \frac{1}{2}c|\cdot|^2$

$$\begin{aligned}&\le \lambda\phi(y_1) + (1-\lambda)\phi(y_2) + \tfrac{1}{2}\lambda c|y_1|^2 + \tfrac{1}{2}(1-\lambda)c|y_2|^2 \\ &\quad + \tfrac{1}{2}(\lambda y_1 + (1-\lambda)y_2)^T[C - cI](\lambda y_1 + (1-\lambda)y_2) \\ &= \lambda\widehat{\phi}(y_1) + (1-\lambda)\widehat{\phi}(y_2) + \tfrac{1}{2}\lambda[c|y_1|^2 - y_1^T C y_1] \\ &\quad + \tfrac{1}{2}(1-\lambda)[c|y_2|^2 - y_2^T C y_2] \\ &\quad + \tfrac{1}{2}(\lambda y_1 + (1-\lambda)y_2)^T[C - cI](\lambda y_1 + (1-\lambda)y_2),\end{aligned}$$

which, after some algebra,

$$\begin{aligned}&= \lambda\widehat{\phi}(y_1) + (1-\lambda)\widehat{\phi}(y_2) \\ &\quad - \lambda(1-\lambda)(y_1 - y_2)^T(C - cI)(y_1 - y_2),\end{aligned}$$

which by the non-negative definiteness of $C - cI$

$$\le \lambda\widehat{\phi}(y_1) + (1-\lambda)\widehat{\phi}(y_2).$$

(Note that if $C - cI > 0$, then one obtains strict convexity.)

Also note that

$$\begin{aligned}|\widehat{\phi}(y_1) - \widehat{\phi}(y_2)| &\le |\phi(y_1) - \phi(y_2)| + |\tfrac{1}{2}y_1^T C y_1 - \tfrac{1}{2}y_2^T C y_2| \\ &\le L|y_1 - y_2| + |(Cy_3)^T(y_1 - y_2)|\end{aligned}$$

for some $y_3 \in \overline{B}_R$

$$\le (L + |C|R)|y_1 - y_2|,$$

which implies that $\widehat{\phi}$ is Lipschitz with constant $(|C|R + L)$.

Then, by Theorem 2.8 for any $y \in \overline{B}_R$,

$$\begin{aligned}\phi(y) &= \widehat{\phi}(y) - \tfrac{1}{2}y^T C y \\ &= \max_{|p| \le (L+|C|R)} \Big[p^T y + \widehat{\psi}(p) - \tfrac{1}{2}y^T C y\Big],\end{aligned}$$

where $\widehat{\psi}$ is given by (2.4)

$$\begin{aligned}&= \max_{|p| \le (L+|C|R)} \Big[-\tfrac{1}{2}(p^T C^{-1} p - 2p^T y + y^T C y) + \widehat{\psi}(p) + \tfrac{1}{2}p^T C^{-1} p\Big] \\ &= \max_{|p| \le (L+|C|R)} \Big[-\tfrac{1}{2}(y - C^{-1}p)^T C (y - C^{-1}p) + \widehat{\psi}(p) + \tfrac{1}{2}p^T C^{-1} p\Big],\end{aligned}$$

and letting $\overline{y} \doteq C^{-1}p$, this is

$$= \max_{|C\overline{y}|\le(L+|C|R)} \Big[-\tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y}) + \widehat{\psi}(C\overline{y}) + \tfrac{1}{2}\overline{y}^T C\overline{y}\Big]. \tag{2.12}$$

Define

$$\psi(\overline{y}) \doteq \widehat{\psi}(C\overline{y}) + \tfrac{1}{2}\overline{y}C^T\overline{y}. \tag{2.13}$$

Combining (2.12) and (2.13), one has

$$\phi(y) = \max_{|C\overline{y}|\le(L+|C|R)} \Big[\psi(\overline{y}) - \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big].$$

Now, from (2.4),

$$\begin{aligned}\widehat{\psi}(p) &= -\max_{y\in\overline{B}_R}\Big[y^T p - \widehat{\phi}(y)\Big]\\ &= -\max_{y\in\overline{B}_R}\Big[y^T p - \phi(y) - \tfrac{1}{2}y^T C y\Big]\end{aligned} \tag{2.14}$$

(and, further, this is $> -\infty$ if $C - cI > 0$). Combining (2.13) and (2.14),

$$\begin{aligned}\psi(\overline{y}) &= -\max_{y\in\overline{B}_R}\Big[-\tfrac{1}{2}(y^T C y - 2y^T C\overline{y} + \overline{y}C\overline{y}) - \phi(y)\Big]\\ &= -\max_{y\in\overline{B}_R}\Big[-\tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y}) - \phi(y)\Big]\end{aligned}$$

(where, further, this is $> -\infty$ if $C - cI > 0$). This yields (2.6), (2.10). Equivalent forms (2.7) and (2.8) follow similarly using the second equality of (2.3) rather than the first. □

Remark 2.10. It may sometimes be useful to replace the scalar semiconvexity constant, with a matrix constant. Let C' be a symmetric, positive definite matrix. Let $\mathcal{S}_{\overline{R}}^{C'L}$ be the set of mappings, ϕ, from $\overline{B}_R$ into $\mathbf{R}^-$ which are Lipchitz with constant L, and such that $\phi(y) + \frac{1}{2}y^T C' y$ is convex. A result equivalent to Theorem 2.9 (i.e., (2.6)– (2.10)) holds where an assumption that $C - C' > 0$ replaces the assumption that $C - cI > 0$.

The above results can also be applied to spaces of functions that are strictly convex. This can be quite useful in numerical methods if one suspects the solution that one is searching for is convex to the extent required in the following theorem.

Theorem 2.11. *Suppose ϕ is Lipschitz with contant L over $\overline{B}_R$. Suppose that there exists $c > 0$ such that $\phi(y)-(c/2)|y|^2$ is convex on $\overline{B}_R$. Let C be a positive definite, symmetric matrix such that $cI - C \ge 0$. Let $D_R \ge |C||C^{-1}|R + |C^{-1}|L$. Then, for all $y \in \overline{B}_R$,*

$$\phi(y) = \max_{|C\overline{y}|\leq(L+|C|R)} \Big[\psi(\overline{y}) + \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big] \tag{2.15}$$

$$= \max_{\overline{y}\in\overline{B}_{D_R}} \Big[\psi(\overline{y}) + \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big] \tag{2.16}$$

$$= \max_{\overline{y}\in G} \Big[\psi(\overline{y}) + \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big] \tag{2.17}$$

for any $G \supseteq \{y : |C\overline{y}| \leq (L+|C|R)\}$ *where*

$$\psi(\overline{y}) = -\max_{y\in\overline{B}_R} \Big[-\phi(y) + \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big] \tag{2.18}$$

for all $\overline{y} \in \mathbf{R}^n$ *(and* $\psi(\overline{y}) > -\infty$ *for all* $\overline{y}$ *if* $cI - C > 0$*).*

Proof. The proof is a variation on the proof of Theorem 2.9, and consequently, only a sketch is provided. Let $\widehat{\phi}(y) = \phi(y) - \frac{1}{2}y^T C y$ with $\phi \neq -\infty$.. Then $\widehat{\phi}$ is strictly convex, and is Lipschitz with constant $(L+|C|R)$. Using Theorem 2.8, one finds that for any $y \in \overline{B}_R$,

$$\begin{aligned}\phi(y) &= \widehat{\phi}(y) + \tfrac{1}{2}y^T C y \\ &= \max_{|p|\leq(L+|C|R)} \Big[-p^T y + \widetilde{\psi}(p) + \tfrac{1}{2}y^T C y\Big],\end{aligned}$$

where $\widetilde{\psi}(p) \doteq \widehat{\psi}(-p)$ and $\widehat{\psi}$ is given by (2.4). Letting $\overline{y} \doteq C^{-1}p$, one finds

$$= \max_{|C\overline{y}|\leq(L+|C|R)} \Big[\tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y}) + \widetilde{\psi}(C\overline{y}) - \tfrac{1}{2}\overline{y}^T C\overline{y}\Big].$$

Letting $\psi(\overline{y}) = \widetilde{\psi}(C\overline{y}) - \frac{1}{2}\overline{y}^T C\overline{y}$, one has

$$\phi(y) = \max_{|C\overline{y}|\leq(L+|C|R)} \Big[\psi(\overline{y}) + \tfrac{1}{2}(y-\overline{y})^T C(y-\overline{y})\Big].$$

One then obtains (2.18) from the fact that $\psi(\overline{y}) = \widehat{\psi}(-C\overline{y})$ where $\widehat{\psi}$ is given by (2.4). □

Similar results can also be obtained for spaces of concave and semiconcave functions. In such spaces, the elements, $\phi(\cdot)$, will have representation in terms of minima of quadratics rather than maxima. For example, a semiconcave duality result for $\phi(\cdot)$ can be obtained directly from the corresponding semiconvex duality result for $-\phi(\cdot)$. In order to reach our goal more quickly, such results will be delayed until we need them.

The above duality results can be used to obtain bases for spaces of convex and semiconvex functions when these spaces are endowed with the max-plus operations rather than the standard field. Let $\mathcal{X}$ be a b-complete (a-complete) max-plus space. Let $\{\psi_i\}_{i=1}^{\infty} \subseteq \mathcal{X}$. The inifinite sum $\bigoplus_{i=1}^{\infty} \psi_i$ is defined to be $\sup[\{\psi_i\}_{i=1}^{\infty}]$, and this is guaranteed to exists if $\mathcal{X}$ is a-complete or if $\{\psi_i\}_{i=1}^{\infty}$ is bounded and $\mathcal{X}$ is b-complete.

Definition 2.12. *Let $\mathcal{X}$ be an a-complete or b-complete max-plus vector space. Let $\mathcal{A} \doteq \{\psi_i\}_{i=1}^{\infty} \subseteq \mathcal{X}$. Then $\mathcal{A}$ is a countable basis for $\mathcal{X}$ if given $\phi \in \mathcal{X}$, there exists $\{a_i\}_{i=1}^{\infty} \subset \mathbf{R}^-$ such that*

$$\phi = \bigoplus_{i=1}^{\infty} (a_i \otimes \psi_i).$$

Where no confusion arises, we will typically refer to a countable max-plus basis simply as a basis. The usage of the term basis is nonstandard here. However, we do not have a notion of "complete orthonormal set," and so we will work with countable basis expansions of elements of our max-plus space in place of expansions in terms of complete orthonormal sets.

Consider $\mathcal{S}_{\bar{R}}^{cL}$. (Note that $\mathcal{S}_{\bar{R}}^{cL}$ is b-complete; the a-completion of $\mathcal{S}_{\bar{R}}^{cL}$ is $\mathcal{S}_{\bar{R}}^{cL-+} = \mathcal{S}_{\bar{R}}^{cL} \cup \{+\infty.\}$ where $+\infty.$ is the function that is identically $+\infty.$ this last result is discussed more fully at Proposition 2.18.) Let C be a symmetric matrix such that $C - cI > 0$. Let $\mathcal{E} = \{\overline{y} \in \mathbf{R}^n : \overline{y}^T(C^2)\overline{y} \leq (L + |C|R)^2\}$. Let $\mathcal{A} \doteq \{y_i\}_{i=1}^{\infty}$ be a countable dense subset of $\mathcal{E}$. Let $\mathcal{N}$ denote the set of natural numbers. For each $i \in \mathcal{N}$, let $\xi_i(y) \doteq -\frac{1}{2}(y - y_i)^T C(y - y_i)$. Let $\phi \in \mathcal{S}_{\bar{R}}^{cL}$. By Theorem 2.9, for all $y \in \overline{B}_R$

$$\phi(y) = \max_{\overline{y} \in \mathcal{E}} \Big[\psi(\overline{y}) - \tfrac{1}{2}(y - \overline{y})^T C(y - \overline{y})\Big],$$

where

$$\psi(\overline{y}) = -\max_{y \in \overline{B}_R} \Big[-\phi(y) - \tfrac{1}{2}(y - \overline{y})^T C(y - \overline{y})\Big]$$

for all $\overline{y} \in \mathcal{E}$. Also let $a_i \doteq \psi(y_i)$ for all $i \in \mathcal{N}$. Then, by the denseness of $\{y_i\}_{i=1}^{\infty}$,

$$\begin{aligned} \phi(y) &= \sup_{i \in \mathcal{N}} \Big[\psi(y_i) - \tfrac{1}{2}(y - y_i)^T C(y - y_i)\Big] = \sup_{i \in \mathcal{N}} \Big[a_i + \xi_i(y)\Big] \\ &= \bigoplus_{i=1}^{\infty} \left[a_i \otimes \xi_i(y)\right], \end{aligned} \tag{2.19}$$

where

$$a_i = -\max_{y \in \overline{B}_R} \Big[-\phi(y) + \xi_i(y)\Big] \quad \forall\, i. \tag{2.20}$$

This yields the following:

Theorem 2.13. *Let $c, L, R > 0$, and let symmetric matrix C satisfy $C - cI > 0$. Let $\xi_i(y) \doteq -\frac{1}{2}(y - y_i)^T C(y - y_i)$ for all $i \in \mathcal{N}$ where the y_i form a countable dense subset of $\mathcal{E} = \{\overline{y} \in \mathbf{R}^n : \overline{y}^T(C^2)\overline{y} \leq (L + |C|R)^2\}$. Then, $\{\xi_i : i \in \mathcal{N}\} \subset \mathbf{R}$ is a countable basis for max-plus vector space $\mathcal{S}_{\bar{R}}^{cL}$. Further a choice of coefficients in the expansion of any $\phi \in \mathcal{S}_{\bar{R}}^{cL}$ is given by (2.20).*

We note that the a_i given by (2.20) may not form the only set of coefficients such that (2.19) holds. For example, suppose $L+|C|R > R$ (so that $\overline{B}_R \subset \mathcal{E}$), and let $\phi(y) = 0$ for all $y \in \overline{B}_R$. Then any set of coefficients such that $a_i = 0$ if $y_i \in \overline{B}_R$ and $a_i < 0$ otherwise, satisfies (2.19). Thus there are multiple expansions in such a case.

By similar approaches to that yielding the above basis for $\mathcal{S}_{\bar{R}}^{cL}$, one can obtain countable bases for other max-plus spaces such as those above. For instance, one may use Theorem 2.8 to obtain a basis for $\mathcal{C}_{\bar{R}}^{L}$; the basis would consist of a countable set of linear functionals. The reader can easily produce such bases.

Remark 2.14. As with the extension of Theorem 2.9 considered in Remark 2.10, Theorem 2.13 may be extended to matrix semiconvexity constants. In particular, let C' be a symmetric, positive definite matrix, and suppose symmetric C satisfies $C-C' > 0$. Then the set $\{\xi_i : i \in \mathcal{N}\}$ given in Theorem 2.13 is a countable basis for max-plus vector space $\mathcal{S}_{\bar{R}}^{C'L}$. Again the coefficients in the expansion of any $\phi \in \mathcal{S}_{\bar{R}}^{C'L}$ are given by (2.20).

2.3 Two-Parameter Families

Two-parameter families of basis functions may be of interest. There is an obvious analogy to wavelets here, but to the author's knowledge, this has not been extensively explored (see however, [2] and references therein). Consider $\mathcal{S}_{\bar{R}}^{cL}$. Let $\{c_j\}_{j\in\mathcal{N}}$ be a countable, unbounded subset of $[c,\infty)$. Let $\{y_i\}_{i\in\mathcal{N}}$ be a countable dense subset of $\overline{B}_{D_R}$ where $\widehat{c} \doteq \inf_j c_j$ and $D_R = R + L/\widehat{c}$. For each $i,j \in \mathcal{N}$, let $\xi_{i,j}(y) \doteq \frac{-c_j}{2}|y-y_i|^2$. Then $\{\xi_{i,j}|\, i,j \in \mathcal{N}\}$ forms a countable basis for $\mathcal{S}_{\bar{R}}^{cL}$. However, in this case there is little to be gained. One can also take $D_R^j = R + L/c_j$ for any j, and let $\{y_{i,j}\}_{i\in\mathcal{N}}$ be dense over the ball of radius D_R^j. Then, letting $\xi_{i,j}(y) \doteq \frac{-c_j}{2}|y-y_{i,j}|^2$ for each i,j, one still has that $\{\xi_{i,j}|\, i,j \in \mathcal{N}\}$ forms a countable basis for $\mathcal{S}_{\bar{R}}^{cL}$.

Two-parameter families may be more useful when one does not have a bound on the semiconvexity constant. For instance, consider $\mathcal{S}$. Let $\{y_i\}$ be a countable, dense set over $\mathbf{R}^n$, and let $\{c_j\}_{j\in\mathcal{N}}$ be a countable, unbounded subset of $(0,\infty)$. Again take $\xi_{i,j}(y) \doteq \frac{-c_j}{2}|y-y_i|^2$. Then let $\Sigma \doteq \{\xi_{i,j}|\, i,j \in \mathcal{N}\}$. Fix some $R < \infty$. Let $\phi \in \mathcal{S}$. Then (see [42] for instance), there exist $c, L < \infty$ such that the restriction of ϕ to $\overline{B}_R$, denoted ϕ_R, is in $\mathcal{S}_{\bar{R}}^{cL}$. There exists $\bar{j}$ such that $c_{\bar{j}} > c$. Let $D_R^{\bar{j}} > R + L/c_{\bar{j}}$. Let $\mathcal{I} \subseteq \mathcal{N}$ be such that $\{y_i\}_{i\in\mathcal{I}}$ is a dense subset of $\overline{B}_{D_R^{\bar{j}}}$. Then there exists $\{a_i\}_{i\in\mathcal{I}} \subset \mathbf{R}^-$ such that $\phi_R(y) = \bigoplus_{i\in\mathcal{I}} a_i \otimes \xi_{i,j}(y)$ for all $y \in \overline{B}_R$. In other words, any semiconvex function over any ball $\overline{B}_R$ can be represented as an infinite max-plus linear combination of elements of Σ.

2.4 Dual Spaces and Reflexivity

This section discusses some of the results for max-plus function spaces which would best be described as max-plus functional analysis. The results here are not directly necessary for the study of the numerical methods to follow, and so a reader with the goal of getting directly to the applicable material could skip this section on a first reading without risk of difficulty later.

The general theoretical material here is a condensation of results on idempotent spaces (spaces over semifields for which $b \oplus b = b$, of which max-plus is one example). In particular, some important references are [1], [20], [63] and [70]. The presentation here will closely follow the approach of [20]. This general theory is applied to an example in the space of convex functions.

Let $(\mathcal{X}, \oplus, \otimes)$ be a max-plus vector space with elements $\phi = \{\phi_y\}_{y \in \mathcal{Y}}$. Throughout, when we speak of a max-plus space, it will be assumed that we have such a representation for its elements. For $\phi^1, \phi^2 \in \mathcal{X}$, we define the partial order $\phi^1 \leq \phi^2$ if $\phi^1_y \leq \phi^2_y$ for all $y \in \mathcal{Y}$. Note that $\mathcal{X}$ is complete as a partially ordered set if any subset has a supremum (defined in the standard way, as a least upper bound according to the partial ordering). For any such partially ordered (or ordered) space, $\mathcal{X}$, and subset A, we define the greatest lower bound or infimum by

$$\inf(A) \doteq \sup\{b \in \mathcal{X} \,|\, b \leq a \;\; \forall a \in A\}.$$

We say $(\mathcal{X}, \oplus, \otimes)$ is a complete max-plus vector space if $\mathcal{X}$ is complete as a partially ordered set (not to be confused with completeness defined in terms of Cauchy sequences), and if the maps $a \mapsto \mathcal{X}$ given by $f_\phi(a) = a \otimes \phi$ and $\phi \mapsto \mathcal{X}$ given by $f_a(\phi) = a \otimes \phi$ are continuous for all $a \in \mathbf{R}^{-+}$ and all $\phi \in \mathcal{X}$. Here, f_ϕ is (max-plus sense) *continuous* if $f_\phi(\sup(A)) = \sup_{a \in A}[f_\phi(a)]$ for all $A \subset \mathbf{R}^{-+}$, and similarly, f_a is *continuous* if $f_a(\sup(G)) = \sup_{\phi \in G}[f_a(\phi)]$ for all $G \subset \mathcal{X}$. A max-plus space will be the *completion* of max-plus space $\mathcal{X}$ over $(\mathbf{R}^-, \oplus, \otimes)$ if it is the smallest complete max-plus space over $(\mathbf{R}^{-+}, \oplus, \otimes)$ containing the set $\mathcal{X}$.

Let $\mathcal{C}_R$ be the space of convex functions mapping $B_R(0) \subset \mathbf{R}^n$ into $\mathbf{R}^-$ over the max-plus algebra, $(\mathbf{R}^-, \oplus, \otimes)$.

Proposition 2.15. *The completion of $\mathcal{C}_R$ over $(\mathbf{R}^-, \oplus, \otimes)$ is the space of lower semicontinuous (lsc), convex functions mapping B_R into $\mathbf{R}^{-+}$, and this will be denoted by $\mathcal{C}_R^{-+}$. More specifically, the max-plus space is the set $\mathcal{C}_R^{-+}$ over the complete max-plus semifield $(\mathbf{R}^{-+}, \oplus, \otimes)$.*

Proof. We first verify that $\mathcal{C}_R^{-+}$ as a set is complete (a-complete) in the max-plus sense defined above. Note that this implies that the set $\mathcal{C}_R^{-+}$ contains the completion of $\mathcal{C}_R$. Let $G \subseteq \mathcal{C}_R^{-+}$, and index G as $G = \{\phi^\alpha\}_{\alpha \in \mathcal{A}}$ for some index set $\mathcal{A}$. For each $y \in B_R$, let $\widehat{\phi}_y \doteq \sup_{\alpha \in \mathcal{A}} \phi^\alpha_y$ (where we recall the notational equivalence $\phi_y = \phi(y)$). It is easy to show that the resulting function $\widehat{\phi} : B_R \to \mathbf{R}^{-+}$ is lsc (see, for instance, [104]). Now, let $y_1, y_2 \in B_R$ and $\lambda \in [0,1]$. Given $\varepsilon > 0$, there exists $\bar{\alpha} \in \mathcal{A}$ such that

$$\begin{aligned}\widehat{\phi}_{[\lambda y_1+(1-\lambda)y_2]} &\leq \phi^{\bar{\alpha}}_{[\lambda y_1+(1-\lambda)y_2]} + \varepsilon \\ &\leq \lambda\phi^{\bar{\alpha}}_{y_1} + (1-\lambda)\phi^{\bar{\alpha}}_{y_2} + \varepsilon \\ &\leq \lambda\widehat{\phi}_{y_1} + (1-\lambda)\widehat{\phi}_{y_2} + \varepsilon.\end{aligned}$$

Because this is true for all $\varepsilon > 0$, $\widehat{\phi}$ is convex. Consequently, one finds that $\widehat{\phi} \in \mathcal{C}_R^{-+}$, and so $\mathcal{C}_R^{-+}$, as a partially ordered set, is complete. Further, one can easily verify that this is complete as a max-plus vector space over $(\mathbf{R}^{-+}, \oplus, \otimes)$.

Next, one must show that any $\widehat{\phi} \in \mathcal{C}_R^{-+}$ takes the form $\widehat{\phi} = \sup\{G\}$ for some $G \in \mathcal{C}_R$. However, by [101], pp. 15–16, given $\widehat{\phi} \in \mathcal{C}_R^{-+}$, there exists a $\phi_v^* : \mathbf{R}^n \to \mathbf{R}^{-+}$ such that for any $y \in B_R$, $\widehat{\phi}_y = \widehat{\phi}_y^{**} \doteq \sup_{v\in\mathbf{R}^n}\{y \cdot v - \phi_v^*\}$. Since $y \cdot v - \phi_v^* \in \mathcal{C}_R$ for any $v \in \mathbf{R}^n$, one sees that $\widehat{\phi}$ has the form $\sup\{G\}$, $G \subseteq \mathcal{C}_R$. □

Let $\mathcal{S}_R^{c-+}$ be the space of lower semicontinuous functions, ϕ, mapping $B_R \subset \mathbf{R}^n$ into $\mathbf{R}^{-+}$ such that $\phi(y) + (c/2)|y|^2$ is convex. It is worth noting the following.

Lemma 2.16. *The mapping $\phi \mapsto \phi^c$ from $\mathcal{S}_R^{c-+}$ into $\mathcal{C}_R^{-+}$ given by $\phi_y^c = \phi_y + (c/2)|y|^2$ is a bijection.*

Proposition 2.17. *The completion of $\mathcal{S}_R^c$ (over $(\mathbf{R}^-, \oplus, \otimes)$) is $\mathcal{S}_R^{c-+}$ (over $(\mathbf{R}^{-+}, \oplus, \otimes)$).*

Proof. Let $G \subseteq \mathcal{S}_R^c$. Let $G_c = \{\phi^c \in \mathcal{C}_R \,|\, \exists \phi \in G \text{ such that } \phi_y^c = \phi_y + (c/2)|y|^2 \;\forall y \in B_R\}$, and denote this mapping from subsets of $\mathcal{S}_R^c$ into subsets of C_R by $G_c = F(G)$. Let $\bar{\phi}^c = \sup\{G_c\}$, and note that, by Proposition 2.15, $\bar{\phi}^c \in \mathcal{C}_R^{-+}$. Let $\bar{\phi}_y = \bar{\phi}_y^c - (c/2)|y|^2$ for all $y \in B_R$. Then, by Lemma 2.16, $\bar{\phi} \in \mathcal{S}_R^{c-+}$. It is not hard to show $\bar{\phi} = \sup\{G\}$. This implies that $\mathcal{S}_R^{c-+}$ is complete.

To demonstrate that it is the smallest complete set containing $\mathcal{S}_R^c$, let $\widetilde{\phi} \in \mathcal{S}_R^{c-+}$. Let $\widetilde{\phi}_y^c = \widetilde{\phi}_y + (c//2)|y|^2$, which implies $\widetilde{\phi}^c \in \mathcal{C}_R^{-+}$. From Proposition 2.15, there exists $G_c \subseteq \mathcal{C}_R$ such that $\widetilde{\phi}^c = \sup\{G_c\}$. Let $G = F^{-1}(G_c) \subseteq \mathcal{S}_R^c$. One can easily show $\widetilde{\phi} = \sup\{G_c\}$. □

Let $\mathcal{S}_R^{cL-+}$ be the space of functions mapping B_R into $\mathbf{R}^{-+}$ which are semiconvex and Lipschitz with constants c and L, respectively. Note that

$$\mathcal{S}_R^{cL-+} = \mathcal{S}_R^{cL} \cup \{+\infty.\}. \tag{2.21}$$

Proposition 2.18. *The completion of $\mathcal{S}_R^{cL}$ is $\mathcal{S}_R^{cL-+}$.*

Proof. Let $G \subseteq \mathcal{S}_R^{cL-+}$, and let $\widehat{\phi} = \sup\{G\}$. One must prove that $\widehat{\phi} \in \mathcal{S}_R^{cL-+}$. Suppose that $\widehat{\phi} \neq +\infty.$; otherwise this is trivial. The proof that $\widehat{\phi}$ is semiconvex with constant c is similar to the proof of convexity for Proposition 2.15, and so is not included. Let $y_1, y_2 \in B_R$. Given $\varepsilon > 0$, let $\phi^\varepsilon \in \mathcal{S}_R^{cL-+}$ be such that $\phi^\varepsilon_{y_1} \geq \widehat{\phi}_{y_1} + \varepsilon$. Then

$$\widehat{\phi}_{y_1} - \widehat{\phi}_{y_2} \leq \phi^\varepsilon_{y_1} - \phi^\varepsilon_{y_2} + \varepsilon \leq L|y_1 - y_2| + \varepsilon.$$

Because this is true for all $\varepsilon > 0$, $\widehat{\phi}_{y_1} - \widehat{\phi}_{y_2} \leq L|y_1 - y_2|$. Then, by symmetry, one has the Lipschitz condition $|\widehat{\phi}_{y_1} - \widehat{\phi}_{y_2}| \leq L|y_1 - y_2|$. Consequently, $\mathcal{S}_R^{cL-+}$ is complete.

Now, suppose $\widehat{\phi} \in \mathcal{S}_R^{cL-+}$. It must be shown that $\widehat{\phi} = \sup\{G\}$ for some $G \subseteq \mathcal{S}_R^{cL}$. Recall $\mathcal{S}_R^{cL-+} = \mathcal{S}_R^{cL} \cup \{+\infty.\}$. If $\widehat{\phi} \neq +\infty.$, let $G = \{\widehat{\phi}\} \subset \mathcal{S}_R^{cL}$; otherwise, let $G = \mathcal{S}_R^{cL}$. $\square$

In the development so far, the operations of addition of two elements of a max-plus space and multiplication of an element of a max-plus space by a scalar were inherited directly from the underlying max-plus algebra. Consequently, the max-plus vector space could be specified by the set of elements and the underlying max-plus semifield. In the following development, we will need to allow the operations on the space to be other than those directly inherited from the max-plus algebra. Consequently, for the remainder of this section we will specify a max-plus space with notation $(\mathcal{X}^*, \oplus^*, \otimes^*)$ over say $(\mathbf{R}^{-+}, \oplus, \otimes)$. When the operations are directly inherited from the underlying field, we may simply denote the space as $\mathcal{X}^*$ over say $(\mathbf{R}^{-+}, \oplus, \otimes)$.

Let $\mathcal{X}$ be a complete max-plus vector space over $(\mathbf{R}^{-+}, \oplus, \otimes)$ where vectors $\phi \in \mathcal{X}$ take the form $\{\phi_y\}_{y \in \mathcal{Y}}$. We will define the *opposite space,* $(\mathcal{X}^{op}, \oplus^{op}, \otimes^{op})$ over $(\mathbf{R}^{-+}, \oplus, \otimes)$ as follows. Let $\mathcal{X}^{op}$=$\mathcal{X}$ (as a set). Define multiplication of $\phi \in \mathcal{X}^{op}$ by a scalar $a \in \mathbf{R}^{-+}$ by

$$(a \otimes^{op} \phi)_y = a \between \phi_y \quad \forall\, y \in \mathcal{Y}, \tag{2.22}$$

where for $a, b \in \mathbf{R}^{-+}$,

$$a \between b \doteq \max\{c \in \mathbf{R}^{-+} \,|\, a \otimes c \leq b\} = \begin{cases} +\infty & \text{if } a = b = +\infty \\ & \text{or } a = b = -\infty, \\ b - a = b \otimes a^{-1} & \text{otherwise,} \end{cases} \tag{2.23}$$

where we use $\leq$ to indicate the original ordering on $\mathbf{R}^{-+}$, and the exponentiation a^{-1} is in the max-plus sense, that is, $a^{-1} \doteq -a$. (With regard to the above defintion, it is helpful to recall that $-\infty \otimes +\infty = -\infty$.) For $\phi^1, \phi^2 \in \mathcal{X}^{op}$, let

$$\phi^1 \oplus^{op} \phi^2 \doteq \min{}^*\{\phi^1, \phi^2\} \doteq \inf\{\phi^1, \phi^2\}$$

(which defines notation $\min^*$), and recall that this is

$$= \sup\{\widehat{\phi} \in \mathcal{X}^{op} = \mathcal{X} \,|\, \widehat{\phi} \leq \phi^1,\ \widehat{\phi} \leq \phi^2\}.$$

More generally, for any $\mathcal{A} \subseteq \mathcal{X}^{op}$, the possibly infinite sum is

$$\bigoplus_{\phi \in \mathcal{A}}^{op} \phi \doteq \inf{}^* \mathcal{A} \doteq \inf \mathcal{A}$$

(which defines the notation $\inf^*$), and recall that this is

$$= \sup\{\widehat{\phi} \in \mathcal{X}^{op} = \mathcal{X} \,|\, \widehat{\phi} \leq \phi \; \forall \, \phi \in \mathcal{A}\}.$$

Note that the additive identity is the element such that $\phi_y = +\infty$ for all $y \in \mathcal{Y}$, and denote this element as $+\infty$.. One should also note that the natural ordering on $\mathcal{X}^{op}$ is defined by $\phi^1 \leq^{op} \phi^2$ if $\phi^1 \oplus^{op} \phi^2 = \phi^2$.

Lemma 2.19. *If* $(\mathcal{X}, \oplus, \otimes)$ *is a complete max-plus vector space over semifield* $(\mathbf{R}^{-+}, \oplus, \otimes)$*, then* $(\mathcal{X}^{op}, \oplus^{op}, \otimes^{op})$ *is also a complete max-plus vector space over* $(\mathbf{R}^{-+}, \oplus, \otimes)$*. (Note that the underlying semifield is unchanged.)*

Proof. See [20]. □

Example 2.20. Let $\mathcal{X} = \mathcal{C}_R^{-+}$. Then, by Proposition 2.15 and Lemma 2.19, $(\mathcal{X}^{op}, \oplus^{op}, \otimes^{op}) = (\mathcal{C}_R^{-+}, \oplus^{op}, \otimes^{op})$ over $(\mathbf{R}^{-+}, \oplus, \otimes)$ (with $\oplus^{op}, \otimes^{op}$ defined as above), and this is a complete max-plus space. Specifically, note that for $\phi^1, \phi^2 \in \mathcal{X}^{op}$,

$$\phi^1 \oplus^{op} \phi^2 = \sup\{\widehat{\phi} \in \mathcal{C}_R^{-+} \,|\, \widehat{\phi} \leq \phi^1, \; \widehat{\phi} \leq \phi^2\}.$$

In other words, $\phi^1 \oplus^{op} \phi^2$ is the convexification of the minimum of ϕ^1 and ϕ^2 [53], [101], [102]. Consequently $\phi^1 \oplus^{op} \phi^2 \in \mathcal{C}_R^{-+}$. As above, for $a \in \mathbf{R}^{-+}$ and $\phi \in \mathcal{C}_R^{-+}$, one has

$$(a \otimes^{op} \phi)_y = \begin{cases} +\infty & \text{if } a = \phi_y = +\infty \text{ or } a = \phi_y = -\infty \\ \phi_y \otimes a^{-1} & \text{otherwise.} \end{cases}$$

Let us verify that $(\mathcal{C}_R^{-+}, \oplus^{op}, \otimes^{op})$ over $(\mathbf{R}^{-+}, \oplus, \otimes)$ is, in fact, a max-plus vector space. If $a, b \in (-\infty, +\infty)$ and $\phi_y \in (-\infty, +\infty)$, then

$$\begin{aligned} [(a \otimes b) \otimes^{op} \phi]_y &= \phi_y - (a + b) = (\phi_y - b) - a \\ &= (b \otimes^{op} \phi)_y - a = [a \otimes^{op} (b \otimes^{op} \phi)]_y. \end{aligned}$$

The other cases (where a, b and/or ϕ_y are $\pm\infty$) are easily checked, and the details are not included. Also,

$$\begin{aligned} [(a \oplus b) \otimes^{op} \phi]_y &= \phi_y - \max\{a, b\} = \phi_y + \min\{-a, -b\} \\ &= \min\{\phi_y - a, \phi_y - b\} = \min\{a \otimes^{op} \phi_y, b \otimes^{op} \phi_y\}. \end{aligned}$$

This implies

$$(a \oplus b) \otimes^{op} \phi = \min\{a \otimes^{op} \phi, b \otimes^{op} \phi\},$$

and since both functions on the right-hand side are convex,

$$= \min{}^*\{a \otimes^{op} \phi, b \otimes^{op} \phi\} = (a \otimes^{op} \phi) \oplus^{op} (b \otimes^{op} \phi).$$

Now, let $a \in \mathbf{R}^{-+}$ and $\phi^1, \phi^2 \in \mathcal{C}_R^{-+}$ (and note that we will continue skipping the special cases where $\pm\infty$ occur). Then

$$\begin{aligned}[a \otimes^{op} (\phi^1 \oplus^{op} \phi^2)]_y &= [\min{}^*\{\phi^1, \phi^2\} - a]_y \\ &= [\min{}^*\{\phi^1 - a, \phi^2 - a\}]_y = [(a \otimes^{op} \phi^1) \oplus^{op} (a \otimes^{op} \phi^2)]_y\end{aligned}$$

for all $y \in B_R$. Lastly, recall that $+\infty.$ is the zero function of $\mathcal{C}_R^{-+}$ while $-\infty$ is the zero element of $\mathbf{R}^{-+}$. One easliy sees that

$$-\infty \otimes^{op} \phi = +\infty., \; a \otimes^{op} (+\infty.) = +\infty., \; 0 \otimes^{op} \phi = \phi \quad \forall\, a \in \mathbf{R}^{-+}, \; \forall\, \phi \in \mathcal{C}_R^{-+}.$$

This completes the verification.

Example 2.21. Let $\mathcal{X} = \mathcal{S}_R^{c-+}$. Then, by Proposition 2.17 and Lemma 2.19, $(\mathcal{X}^{op}, \oplus^{op}, \otimes^{op}) = (\mathcal{S}_R^{c-+}, \oplus^{op}, \otimes^{op})$ over $(\mathbf{R}^{-+}, \oplus, \otimes)$, and this is a complete max-plus space. Note that

$$\phi^1 \oplus^{op} \phi^2 = \sup\{\widehat{\phi} \in \mathcal{S}_R^{c-+} \,|\, \widehat{\phi} \le \phi^1, \; \widehat{\phi} \le \phi^2\}.$$

Alternatively, noting the bijection given by Lemma 2.16, one has

$$\begin{aligned}\left(\phi^1 \oplus^{op} \phi^2\right)_y &= \Big(\sup\{\widehat{\phi}_y - (c/2)|y|^2 \;\big|\; \widehat{\phi} \in \mathcal{C}_R^{-+}, \; \widehat{\phi}_y \le \phi_y^1 + (c/2)|y|^2 \\ &\qquad\qquad \text{and } \widehat{\phi}_y \le \phi_y^1 + (c/2)|y|^2 \;\; \forall\, y \in \mathcal{Y}\}\Big)_y \\ &= \left([\phi^1 + (c/2)|\cdot|^2] \oplus_c^{op} [\phi^2 + (c/2)|\cdot|^2]\right)_y - (c/2)|y|^2\end{aligned}$$

where $\oplus_c^{op}$ indicates the opposite addition operation corresponding to the max-plus space $\mathcal{C}_R^{-+}$. In other words, $\phi^1 \oplus^{op} \phi^2$ can be obtained by adding $(c/2)|y|^2$ to ϕ^1, ϕ^2, taking the convexification of the minimum, and then subtracting $(c/2)|y|^2$.

It is useful to note the following [20]

Lemma 2.22. *If* $(\mathcal{X}, \oplus, \otimes)$ *is a complete max-plus space, then*

$$([\mathcal{X}^{op}]^{op}, [\oplus^{op}]^{op}, [\otimes^{op}]^{op}) = (\mathcal{X}, \oplus, \otimes).$$

Before proceeding to linear functionals and dual spaces, it is helpful to note the following technical result regarding the min* operation.

Proposition 2.23. *Let* $\mathcal{X}$ *be a complete max-plus space over* $(\mathbf{R}^{-+}, \oplus, \otimes)$. *Let* $(\mathcal{X}^{op}, \oplus^{op}, \otimes^{op})$ *denote the opposite space. Then for any* $\phi \in \mathcal{X}$ *and any* $\phi^1, \phi^2 \in \mathcal{X}^{op}$,

$$\sup_{y\in\mathcal{Y}} \left[\phi_y - (\min{}^*\{\phi^1,\phi^2\})_y\right] = \sup_{y\in\mathcal{Y}} \left[\phi_y - (\min\{\phi^1_y,\phi^2_y\})\right]. \tag{2.24}$$

More generally, for any $\mathcal{A} \subseteq \mathcal{X}^{op}$,

$$\sup_{y\in\mathcal{Y}} \left[\phi_y - (\inf{}^*\mathcal{A})_y\right] = \sup_{y\in\mathcal{Y}} \left[\phi_y - \inf\{\phi_y \,|\, \phi\in\mathcal{A}\}\right]. \tag{2.25}$$

Proof. We prove only the first assertion; the proof of the second is nearly identical. Let $\mathcal{A}_{12} \doteq \{\phi \in \mathcal{X}^{op} | \, \phi_y \le \min\{\phi^1_y,\phi^2_y\} \; \forall\, y\in\mathcal{Y}\}$. Then, by definition of $\mathcal{A}_{12}$, $(\sup_{\widehat{\phi}\in\mathcal{A}_{12}} \widehat{\phi})_y \le \min\{\phi^1_y,\phi^2_y\}$ for all y. By the definition of $\min^*$, this implies

$$\left(\min{}^*\{\phi^1,\phi^2\}\right)_y \le \min\{\phi^1_y,\phi^2_y\} \quad \forall\, y\in\mathcal{Y}.$$

This yields

$$\sup_{y\in\mathcal{Y}} \left[\phi_y - (\min{}^*\{\phi^1,\phi^2\})_y\right] \ge \sup_{y\in\mathcal{Y}} \left[\phi_y - (\min\{\phi^1_y,\phi^2_y\})\right]. \tag{2.26}$$

We now prove the reverse. Fix any $\widehat{\phi}\in\mathcal{X}$. Fix any $\overline{y}\in\mathcal{Y}$. Let

$$\beta \doteq (\min{}^*\{\phi^1,\phi^2\})_{\overline{y}} - \widehat{\phi}_{\overline{y}}. \tag{2.27}$$

Suppose there exists $\varepsilon > 0$ such that $\widehat{\phi}_y + \beta + \varepsilon < \min\{\phi^1_y,\phi^2_y\}$ for all $y\in\mathcal{Y}$. Then, by the definition of $\min^*$ and the fact that $\widehat{\phi}\in\mathcal{X}$,

$$(\min{}^*\{\phi^1,\phi^2\})_{\overline{y}} \ge \widehat{\phi}_{\overline{y}} + \beta + \varepsilon. \tag{2.28}$$

But this contradicts (2.27). Therefore, given $\varepsilon > 0$, there exists $y^\varepsilon\in\mathcal{Y}$ such that

$$\widehat{\phi}_{y^\varepsilon} + \beta + \varepsilon \ge \min\{\phi^1_{y^\varepsilon},\phi^2_{y^\varepsilon}\}. \tag{2.29}$$

By (2.27),

$$\widehat{\phi}_{\overline{y}} - (\min{}^*\{\phi^1,\phi^2\})_{\overline{y}} = -\beta,$$

which by (2.29),

$$\le \widehat{\phi}_{y^\varepsilon} + \varepsilon - \min\{\phi^1_{y^\varepsilon},\phi^2_{y^\varepsilon}\}$$
$$\le \sup_{y\in\mathcal{Y}} \left[\widehat{\phi}_y - \min\{\phi^1_y,\phi^2_y\}\right] + \varepsilon.$$

Because this is true for all $\varepsilon > 0$,

$$\widehat{\phi}_{\overline{y}} - (\min{}^*\{\phi^1,\phi^2\})_{\overline{y}} \le \sup_{y\in\mathcal{Y}} \left[\widehat{\phi}_y - \min\{\phi^1_y,\phi^2_y\}\right].$$

Then, because this is true for all $\overline{y}$,

$$\sup_{y\in\mathcal{Y}} \left[\widehat{\phi}_y - (\min{}^*\{\phi^1,\phi^2\})_y\right] \le \sup_{y\in\mathcal{Y}} \left[\widehat{\phi}_y - \min\{\phi^1_y,\phi^2_y\}\right]. \tag{2.30}$$

Combining (2.26) and (2.30) completes the proof. □

It is also useful to define the following operation. Let $\phi, \widehat{\phi}$ belong to complete max-plus space $\mathcal{X}$. Define

$$\phi \,\backslash\!\!\!\!\otimes\, \widehat{\phi} = \max\left\{\lambda \in \mathbf{R}^{-+} \,|\, \lambda \otimes \phi \leq \widehat{\phi}\right\} = \inf\left\{\widehat{\phi}_y - \phi_y \,|\, y \in \mathcal{Y}'\right\} \tag{2.31}$$

where $\mathcal{Y}' \doteq \{y \in \mathcal{Y}|$ neither $\widehat{\phi}_y = \phi_y = +\infty$ nor $\widehat{\phi}_y = \phi_y = -\infty\}$ where we recall $\inf \emptyset = +\infty$. Now, recalling that set-wise $\mathcal{X}^{op} = \mathcal{X}$, define the mapping from $\mathcal{X}^{op} \times \mathcal{X}$ into $\mathbf{R}^{-+}$ given by

$$\left\langle \widehat{\phi}, \phi \right\rangle \doteq -\left(\phi \,\backslash\!\!\!\!\otimes\, \widehat{\phi}\right) = -\inf_{y \in \mathcal{Y}'} \left[\widehat{\phi}_y - \phi_y\right] = \sup_{y \in \mathcal{Y}'} \left[\phi_y - \widehat{\phi}_y\right], \tag{2.32}$$

where $\sup \emptyset = -\infty$.

Theorem 2.24. $\langle \widehat{\phi}, \cdot \rangle : \mathcal{X} \to \mathbf{R}^{-+}$ *and* $\langle \cdot, \phi \rangle : \mathcal{X}^{op} \to \mathbf{R}^{-+}$ *are linear mappings.*

Proof. It is obvious that $\langle \widehat{\phi}, \cdot \rangle$ is linear for any $\widehat{\phi} \in \mathcal{X}^{op}$.

Let $\phi \in \mathcal{X}$. Consider $\langle \cdot, \phi \rangle$. Note that for $a \in \mathbf{R}^{-+}$ and $\widehat{\phi} \in \mathcal{X}^{op}$ (where we continue skipping special cases where $\pm\infty$ appear),

$$\begin{aligned} \langle a \otimes^{op} \widehat{\phi}, \phi \rangle &= \sup_{y \in \mathcal{Y}} \left[\phi_y - \left(a \otimes^{op} \widehat{\phi}\right)_y\right] \\ &= \sup_{y \in \mathcal{Y}} \left[\phi_y - \widehat{\phi}_y + a\right] = a + \sup_{y \in \mathcal{Y}} \left[\phi_y - \widehat{\phi}_y\right] = a \otimes \left\langle \widehat{\phi}, \phi \right\rangle. \end{aligned}$$

Now, note that for any $\phi^1, \phi^2 \in \mathcal{X}^{op}$,

$$\langle \phi^1 \oplus^{op} \phi^2, \phi \rangle = \sup_{y \in \mathcal{Y}} \left[\phi_y - (\min{}^*\{\phi^1, \phi^2\})_y\right]$$

which by Lemma 2.23,

$$\begin{aligned} &= \sup_{y \in \mathcal{Y}} \left[\phi_y - (\min\{\phi^1_y, \phi^2_y\})\right] \\ &= \sup_{y \in \mathcal{Y}} \left[\max\{\phi_y - \phi^1_y, \phi_y - \phi^2_y\}\right] \\ &= \max\left\{\sup_{y \in \mathcal{Y}}[\phi_y - \phi^1_y],\ \sup_{y \in \mathcal{Y}}[\phi_y - \phi^2_y]\right\} \\ &= \langle \phi^1, \phi \rangle \oplus \langle \phi^2, \phi \rangle . \square \end{aligned}$$

Let f map complete max-plus space $\mathcal{X}$ into complete max-plus space $\widehat{\mathcal{X}}$. The mapping, f, is *monotone* if $f(\phi^1) \leq f(\phi^2)$ whenever $\phi^1 \leq \phi^2$ (where the $\leq$ ordering is with respect to the natural ordering on the space in which the relation is being used). A monotone map, $f : \mathcal{X} \to \mathbf{R}^{-+}$ (where $\mathcal{X}$ is a complete max-plus space), is *continuous* if $f(\sup[\mathcal{A}]) = \sup[f(\mathcal{A})]$ for all $\mathcal{A} \subseteq \mathcal{X}$. When the range of a mapping is $\mathbf{R}^{-+}$ (or $\mathbf{R}^{-}$), the term "functional" will be freely used for the mapping.

Theorem 2.25. *Let $\mathcal{X}$ be a complete max-plus space. $\langle \widehat{\phi}, \cdot \rangle : \mathcal{X} \to \mathbf{R}^{-+}$ and $\langle \cdot, \phi \rangle : \mathcal{X}^{op} \to \mathbf{R}^{-+}$ are continuous linear functionals.*

Proof. The proof is similar to the proof of Theorem 2.24, but employing the second assertion of Lemma 2.23 rather than the first. □

When $\langle \widehat{\phi}, \cdot \rangle$ and $\langle \cdot, \phi \rangle$ are continuous linear functionals, they are referred to as a *predual pair* [20]. Let $\mathcal{X}$, $\widehat{\mathcal{X}}$ be complete max-plus spaces such that $\langle \widehat{\phi}, \phi \rangle : \widehat{\mathcal{X}} \times \mathcal{X} \to \mathbf{R}^{-+}$ is a predual pair. Then (see [20]) $\widehat{\mathcal{X}}$ *separates* $\mathcal{X}$ if

$$\langle \widehat{\phi}, \phi^1 \rangle = \langle \widehat{\phi}, \phi^2 \rangle \;\; \forall \, \widehat{\phi} \in \widehat{\mathcal{X}} \;\; \text{implies} \;\; \phi^1 = \phi^2,$$

and $\mathcal{X}$ *separates* $\widehat{\mathcal{X}}$ if

$$\langle \widehat{\phi}^1, \phi \rangle = \langle \widehat{\phi}^2, \phi \rangle \;\; \forall \, \phi \in \mathcal{X} \;\; \text{implies} \;\; \widehat{\phi}^1 = \widehat{\phi}^2.$$

A predual pair satisfying both separation conditions is a *dual pair* [20]. By Corollary 2.1, [20], if $(\mathcal{X}, \oplus, \otimes)$ is complete, then $\mathcal{X}^{op}$ separates $\mathcal{X}$ and vice versa. Consequently, $\langle \widehat{\phi}, \phi \rangle$ forms a dual pair. It is also shown (Corollary 2.2, [20]) that one has a Riesz representation theorem.

Theorem 2.26. *Any continuous linear functional, f, on complete max-plus space $(\mathcal{X}, \oplus, \otimes)$ has the representation*

$$f(\phi) = \langle \widehat{\phi}, \phi \rangle$$

for a unique $\widehat{\phi} \in \mathcal{X}^{op}$.

Combining this with Lemmas 2.19 and 2.22, one has

Corollary 2.27. *Let $\mathcal{X}$ be a complete max-plus space (which implies $\mathcal{X}^{op}$ complete). Any continuous linear functional, f, on $\mathcal{X}^{op}$ has the form*

$$f(\widehat{\phi}) = \langle \phi, \widehat{\phi} \rangle_{op}$$

for a unique $\phi \in (\mathcal{X}^{op})^{op} = \mathcal{X} = \mathcal{X}^{op}$.

Perhaps it should be noted here that

$$\begin{aligned} \langle \phi, \widehat{\phi} \rangle_{op} &= -\max\{\lambda \in \mathbf{R}^{-+} \,|\, \lambda \otimes^{op} \widehat{\phi} \leq^{op} \phi\} \\ &= -\max\{\lambda \in \mathbf{R}^{-+} \,|\, (\lambda \otimes^{op} \widehat{\phi})_y \geq \phi_y \;\; \forall \, y \in \mathcal{Y}\}, \end{aligned}$$

which after a small bit of work for the special cases

$$= -\inf\{\widehat{\phi}_y - \phi_y \,|\, y \in \mathcal{Y}'\},$$

where $\mathcal{Y}' \doteq \{y \in \mathcal{Y} |$ neither $\widehat{\phi}_y = \phi_y = +\infty$ nor $\widehat{\phi}_y = \phi_y = -\infty\}$

$$= \sup\{\phi_y - \widehat{\phi}_y \,|\, y \in \mathcal{Y}'\},$$

which by (2.32)

$$= \langle \widehat{\phi}, \phi \rangle. \tag{2.33}$$

Combining (2.33) with Corollary 2.27 yields the following representation result.

Corollary 2.28. *Let $\mathcal{X}$ be a complete max-plus space. Any continuous linear functional, f, on $\mathcal{X}^{op}$ has the form*

$$f(\widehat{\phi}) = \langle \widehat{\phi}, \phi \rangle$$

for a unique $\phi \in \mathcal{X}$.

It is natural to refer to this property as *reflexivity.* Recalling from Example 2.20 that $\mathcal{C}_R^{-+}$ is a complete max-plus space, we see that it is reflexive in this sense. The dual space is again $\mathcal{C}_R^{-+}$ but with the opposite operations being given there. (In particular, recall that $\phi^1 \oplus^{op} \phi^2 = \min^*\{\phi^1, \phi^2\}$ is the convexification of the minimum of ϕ^1 and ϕ^2.) The second dual is once again $\mathcal{C}_R^{-+}$ but with operations $(\oplus^{op})^{op} = \oplus$ and $(\otimes^{op})^{op} = \otimes$ being the original max-plus operations again. This is easily verified by noting that

$$\begin{aligned}
\left(\phi^1 (\oplus^{op})^{op} \phi^2\right)_y &\doteq \left(\inf\{\tilde{\phi} \in \mathcal{C}_R^{-+} \,|\, \tilde{\phi} \leq^{op} \phi^1,\, \tilde{\phi} \leq^{op} \phi^2\}\right)_y \\
&= \left(\inf\{\tilde{\phi} \in \mathcal{C}_R^{-+} \,|\, \tilde{\phi} \geq \phi^1,\, \tilde{\phi} \geq \phi^2\}\right)_y \\
&= \max\{\phi^1_y, \phi^2_y\} = (\phi^1 \oplus \phi^2)_y.
\end{aligned}$$

Similarly, $\mathcal{S}_R^{c-+}$ and $\mathcal{S}_R^{cL-+}$ (the completions of $\mathcal{S}_R^c$ and $\mathcal{S}_R^{cL}$) are reflexive.

3

Dynamic Programming and Viscosity Solutions

In this chapter, we present the theory of dynamic programming and viscosity solutions for some specific classes of continuous-time/continuous-space deterministic optimal control problems and games. Dynamic programming (DP) is a mature subject, and there are many excellent references. [8], [15], [36], [46], [47], [48] are a few among many such references. In the specific context of H_∞ and games, [12] and [13] are also excellent.

For the class of continuous-time/continuous-space problems, DP consists of two parts: The first part, the Dynamic Programming Principle (DPP), makes a statement about optimality when the time interval of the problem is broken into two segments. The second part, the Dynamic Programming Equation (DPE), will take the form of a Partial Differential Equation (PDE), and is obtained by taking an infinitesimal limit in the DPP. The DPP is a principle that holds in great generality. These results are well known. However, in the interest of making this book self-contained, some relevant DPP results will be proved.

On the other hand, the DPE/PDE part is much more problematic. When the DPE has a smooth (classical) solution, it is generally straightforward to demonstrate that there is a unique solution, and this solution is the value function of the originating control (or game) problem. Fleming–Rishel [46] is an excellent reference on this topic. However, the PDEs generated by deterministic control and game problems seldom have smooth solutions. Even seemingly innocuous problem statements lead to value functions that have discontinuities in the gradient. Further, in the cases where the DPE has an associated boundary condition (due to an exit condition in the control problem), this boundary condition may not be attained in the classical/pointwise sense. These difficulties were not cleanly resolved until the advent of the viscosity solution definition of solution of a PDE. Some early references are [10], [22], [23], [24], [38], [57], [58], and some useful books are [8], [11], [36], [47] and [48]. It should be noted that another, essentially equivalent, solution definition, "minimax solution," was obtained independently by Subbotin (see [107] and the references therein). However, viscosity solution theory is the dominant formulation.

If one drops the requirement on smooth/classical solutions, there are generally many weak-sense solutions (for instance, solutions that are smooth almost everywhere). The great benefit of viscosity solution theory is that it typically isolates the unique correct weak solution. By "correct," we mean the solution corresponding to the value function. Interestingly, this is also the limiting solution when one adds a diffusion term to the dynamics with a multiplicative small parameter on the diffusion, and lets the parameter go to zero. This has led to the name, "vanishing viscosity" solution, and later simply to viscosity solution.

It is useful to note that for certain infinite time-horizon problems, the standard H_∞ problem being one of them, the viscosity solution conditions are **not** sufficient to specify the unique correct solution. In this case, an additional condition is needed [88], [89], [106]. This issue is mentioned because this class of problems will be one of the main classes we use as a basis for development of the theory in this book.

3.1 Dynamic Programming Principle

Let us begin by deriving DPP results for the example problem classes we will be considering. The two most ubiquitous DPPs to appear here are the control and game problem DPPs where the cost function consists of an integral cost possibly plus a terminal cost. We will focus on these here; for problems with an exit cost, we refer the reader to [33], [47] and [48]. Further below we will consider a problem with an L_∞ payoff, and this will lead to a different form of DPP. However, we place that DPP result with the corresponding development in Chapter 9.

Consider first the control problem with dynamics and initial condition given by

$$\dot{\xi}_t = f(t, \xi_t, u_t), \tag{3.1}$$

$$\xi_s = x. \tag{3.2}$$

In other words, the state process here is ξ_t with initial condition $\xi_s = x \in \mathbf{R}^n$. The terminal time will be $T < \infty$. The control process is $u_\cdot$, and we let the control space (for controls over any time interval $[s, t)$) be

$$\mathcal{U}^U_{s,t} = \left\{ u : [s,t) \to U \subseteq \mathbf{R}^l \,\middle|\, \int_s^t |u_r|^2 \, dr < \infty \right\}.$$

If one simply has $U = \mathbf{R}^l$, then we use the notation $\mathcal{U}_{s,t}$. The choice of L_2 norm here is arbitrary, and is used since our examples will generally be from this class. We assume throughout that $f \in C(\mathbf{R} \times \mathbf{R}^n \times U; \mathbf{R}^l)$ and that there exists $K < \infty$ such that

$$|f(t,x,u) - f(t,y,v) \leq K(|x-y| + |u-v|) \ \ \forall t \in \mathbf{R}, \forall x, y \in \mathbf{R}^n; \quad \forall u, v \in U. \tag{A3.1G}$$

We note that with this assumption and the definition of $\mathcal{U}^U_{s,T}$, one is guaranteed existence and uniqueness of solutions to (3.1), (3.2). The payoff for the control problem will take the form

$$J(s,x;u) \doteq \int_s^T L(r,\xi_r,u_r)\,dr + \phi(\xi_T), \tag{3.3}$$

where ξ_r satisfies (3.1)–(3.2), L will be referred to as the running cost, and ϕ will be referred to as the terminal cost. We assume throughout that $L \in C(\mathbf{R} \times \mathbf{R}^n \times U; \mathbf{R})$ and that there exists $C < \infty$ such that

$$|L(t,x,u)| \leq C\big[\,|x|^2 + |u|^2\big] \quad \forall t \in (-\infty, T), \forall x \in \mathbf{R}^n, \forall u \in U. \tag{A3.2G}$$

Similarly, we assume throughout that $\phi \in C(\mathbf{R}^n \times U; \mathbf{R})$ and that there exists $C_\phi < \infty$ such that

$$|\phi(x)| \leq C_\phi |x|^2 \quad \forall x \in \mathbf{R}^n. \tag{A3.3G}$$

The value of the control problem is defined to be

$$V(s,x) = \sup_{u \in \mathcal{U}^U_{s,T}} J(s,x;u). \tag{3.4}$$

The DPP for this problem is given by the following theorem.

Theorem 3.1. *For any $-\infty < s \leq t < T < \infty$ and any $x \in \mathbf{R}^n$,*

$$V(s,x) = \sup_{u \in \mathcal{U}^U_{s,t}} \left\{ \int_s^t L(r,\xi_r,u_r)\,dr + V(t,\xi_t) \right\}. \tag{3.5}$$

Proof. By definition,

$$V(t,\xi_t) = \sup_{u^1 \in \mathcal{U}^U_{t,T}} \left[\int_t^T L(r,\xi^1_r,u^1_r)\,dr + \phi(\xi^1_T) \right],$$

where ξ^1 satisfies (3.1) with initial condition $\xi^1_t = \xi_t$ and input u^1. Substituting this into the right-hand side of (3.5) yields

$$\begin{aligned}
&\sup_{u \in \mathcal{U}^U_{s,t}} \left\{ \int_s^t L(r,\xi_r,u_r)\,dr + V(t,\xi_t) \right\} \\
&= \sup_{u \in \mathcal{U}^U_{s,t}} \left\{ \int_s^t L(r,\xi_r,u_r)\,dr + \sup_{u^1 \in \mathcal{U}^U_{t,T}} \left[\int_t^T L(r,\xi^1_r,u^1_r)\,dr + \phi(\xi^1_T) \right] \right\} \\
&= \sup_{u \in \mathcal{U}^U_{s,t}} \sup_{u^1 \in \mathcal{U}^U_{t,T}} \left\{ \int_s^t L(r,\xi_r,u_r)\,dr + \int_t^T L(r,\xi^1_r,u^1_r)\,dr + \phi(\xi^1_T) \right\},
\end{aligned}$$

and it is easy to show that this is

$$= \sup_{u \in \mathcal{U}^U_{s,T}} \left\{ \int_s^T L(r, \xi_r, u_r)\, dr + \phi(\xi_T) \right\}. \quad \square$$

The DPP for the infinite time-horizon problem we consider is equally simple. Suppose the problem again has dynamics (3.1), but now with no time dependence in the dynamics (i.e., $\dot{\xi} = f(x,u)$) and initial condition

$$\xi_0 = x \in \mathbf{R}^n \tag{3.6}$$

(i.e., with initial time always $t = 0$). Also take L to be time-independent (i.e., $L(s,x,u) = L(x,u)$). Again make assumptions (A3.1G), (A3.2G), (A3.3G), but now take

$$\mathcal{U}^U = L_2^{\text{loc}}([0,\infty); U) \doteq \left\{ u : [0,\infty) \to U \subseteq \mathbf{R}^l \;\middle|\; \int_0^T |u_r|^2\, dr < \infty \;\; \forall T \in [0,\infty) \right\},$$

and of course

$$\mathcal{U} = L_2^{\text{loc}}([0,\infty); \mathbf{R}^l).$$

Note that $L_2^{\text{loc}} \neq L_2[0,\infty)$. Note also that there exists a unique solution to (3.1), (3.6) for all time for any $x \in \mathbf{R}^n$.

Let the payoff be

$$J(x,T;u) \doteq \int_0^T L(\xi_r, u_r)\, dr, \tag{3.7}$$

and for the moment assume that $J(x,T;u)$ exists for all $x \in \mathbf{R}^n$, $T \in [0,\infty)$ and $u \in \mathcal{U}^U$. Let the value function be

$$W(x) = \sup_{u \in \mathcal{U}^U} \sup_{T \in [0,\infty)} J(x,T;u) \tag{3.8}$$

which we also assume exists for all $x \in \mathbf{R}^n$. The DPP then takes the following form.

Theorem 3.2. *For any $0 \leq t < \infty$ and any $x \in \mathbf{R}^n$,*

$$W(x) = \sup_{u \in \mathcal{U}^U_{0,t}} \left\{ \int_0^t L(\xi_r, u_r)\, dr + W(\xi_t) \right\}. \tag{3.9}$$

The proof is nearly identical to the proof of Theorem 3.1, and so we do not include it.

The above discussion was kept highly general to indicate that few assumptions are required for proof of the DPP. As noted earlier, we will have a few problem classes for which we will prove all our results. The results could be proved for more general problems, but then the assumptions would need to be more abstract; for instance, we assumed the existence of $J(x,T;u)$ for all x,T,u, and $W(x)$ for all x just above. By focusing on specific problem classes, we will be able to make the assumptions much more concrete and easily checkable. If one desires to prove the theory for another problem class, then one can make analogous assumptions and follow similar chains of theorems to obtain the results. We expect that the general max-plus numerical methods theory will hold generically, but that one would need to prove specific results for various problem classes. This is, of course, similar to other principles and general methods/theory.

Let us specialize the above general DPP results to specific classes which we will follow throughout the text. We will also sharpen the results for these problems as needed. First, we take the dynamics to have the specific form

$$\dot{\xi} = f(t,\xi) + \sigma(\xi)u \tag{3.10}$$

with initial condition

$$\xi_s = x \in \mathbf{R}^n \tag{3.11}$$

for the finite time-horizon case, and

$$\dot{\xi} = f(\xi) + \sigma(\xi)u \tag{3.12}$$

with initial condition

$$\xi_0 = x \in \mathbf{R}^n \tag{3.13}$$

for the infinite time-horizon case. For the finite time-horizon case, we assume

$$\begin{aligned} &f \in C(\mathbf{R}\times\mathbf{R}^n;\mathbf{R}), \\ &|f(t,x)-f(t,y)| \le K|x-y| \quad \forall t\in\mathbf{R},\, \forall x,y\in\mathbf{R}^n, \\ &|f(t,x)| \le K(1+|x|) \quad \forall t\in\mathbf{R},\, \forall x\in\mathbf{R}^n, \end{aligned} \tag{A3.1F}$$

for some $K<\infty$. (Although an inequality of the second type above follows from the first, we include it explicitly to indicate that we will be using the same constant, K, for both bounds.) We also assume

$$\begin{aligned} &\sigma \in C(\mathbf{R}^n;\mathcal{L}(\mathbf{R}^l,\mathbf{R}^n)), \\ &|\sigma(x)-\sigma(y)| \le K_\sigma|x-y| \quad \forall\, x,y\in\mathbf{R}^n, \\ &|\sigma(x)| \le m_\sigma \quad \forall\, x\in\mathbf{R}^n, \end{aligned} \tag{A3.2F}$$

for some $K_\sigma, m_\sigma < \infty$. We keep the same definitions of $\mathcal{U}^U_{s,T}$, $\mathcal{U}_{s,T}$, $\mathcal{U}^U$ and $\mathcal{U}$ as above. For the finite time-horizon problem, consider a payoff form

$$J(s,x;u) \doteq \int_s^T l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + \phi(\xi_T). \tag{3.14}$$

Remark 3.3. The more general running cost $l(r,\xi_r) - \frac{1}{2}u_r^T \Gamma^T \Gamma u_r\,dr$ with positive definite $\Gamma^T\Gamma$ is equivalent to the above problem by a simple change in σ, and so we are content to use form (3.14).

We replace the above assumption on L, (A3.2G), with an assumption that

$$\begin{aligned}
&l \in C(\mathbf{R}\times\mathbf{R}^n;\mathbf{R}),\\
&0 \le l(t,x) \le C_l\big[1+|x|^2\big] \quad \forall\, x\in\mathbf{R}^n,\ \forall t\in\mathbf{R},\\
&|l(t,x)-l(t,y)| \le C_l(1+|x|+|y|)|x-y| \ \ \forall x,y\in\mathbf{R}^n, \forall t\in\mathbf{R},
\end{aligned} \tag{A3.3F}$$

for some $C_l<\infty$. Similarly, we assume that there exists $C_\phi<\infty$ such that

$$\begin{aligned}
&\phi\in C(\mathbf{R}^n;\mathbf{R}),\\
&0\le\phi(x)\le C_\phi\big[1+|x|^2\big] \quad \forall\, x\in\mathbf{R}^n,\\
&|\phi(x)-\phi(y)|\le C_\phi(1+|x|+|y|)|x-y| \quad \forall\, x,y\in\mathbf{R}^n.
\end{aligned} \tag{A3.4F}$$

Let

$$V(s,x) = \sup_{u\in\mathcal{U}^U_{s,T}} J(s,x;u) \tag{3.15}$$

for some $U\subseteq\mathbf{R}^n$, $U\neq\emptyset$.

We will be obtaining a DPP where there is a bound on the L_2-norm of the u in the supremum. First, one must get a bound on the behavior of $V(s,x)$. Let $Q_T=[0,T]\times\mathbf{R}^n$.

Lemma 3.4. *$V(s,x)\ge 0$ for all $(s,x)\in Q_T$. Further, there exists $\gamma_2<\infty$ and $C_1<\infty$ such that for all $\gamma>\gamma_2$*

$$V(s,x) < C_1(1+|x|^2) \qquad \forall(s,x)\in Q_T.$$

Proof. Taking $u\equiv 0$ and noting that $l,\phi\ge 0$, yields the first assertion. We proceed to the second assertion. By (3.10) and Assumptions (A3.1F) and (A3.2F),

$$|\xi_t| \le |x| + K\int_s^t (1+|\xi_r|)\,dr + m_\sigma\int_s^t |u_r|\,dr.$$

By Gronwall's inequality, this implies there exists $C_2=C_2(T)<\infty$ such that

$$|\xi_t| \le C_2\left[1+|x|+\int_s^t |u_r|\,dr\right]. \tag{3.16}$$

Using Hölder's inequality, one finds there exists $C_3=C_3(T)<\infty$ such that

$$|\xi_t|^2 \le C_3\left[1+|x|^2+\int_s^t |u_r|^2\,dr\right]. \tag{3.17}$$

Now by assumptions (A3.3F) and (A3.4F),

$$\int_s^T \left[l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \right] dr + \phi(\xi_T)$$
$$\leq \int_s^T \left[C_l(1+|\xi_r|^2) - \frac{\gamma^2}{2}|u_r|^2 \right] dr + C_\phi(1+|\xi_T|^2)$$

which by (3.17) with $\widehat{C} = \max\{C_l, C_\phi\}$

$$\leq \widehat{C}(1+T)[1+C_3(1+|x|^2)] + \int_s^T \left[\widehat{C}C_3(1+T) - \frac{\gamma^2}{2} \right] |u_r|^2\, dr$$

which upon letting $\gamma_2^2 = 2\widehat{C}C_3(1+T)$, becomes

$$= \widehat{C}(1+T)[1+C_3(1+|x|^2)] + \int_s^T \left[\frac{\gamma_2^2}{2} - \frac{\gamma^2}{2} \right] |u_r|^2\, dr. \qquad (3.18)$$

From (3.18), one immediately obtains the second assertion. □

Lemma 3.5. *Let $\gamma > \gamma_2$ and $\epsilon \leq 1$. There exists $C_2, \widehat{C}_2 < \infty$ such that for any ε-optimal $\tilde{u}$ (i.e., such that $J(s,x;\tilde{u}) \geq V(s,x) - \varepsilon$), one has*

$$\int_s^T |\tilde{u}_r|^2\, dr \leq \frac{\widehat{C}_2}{\gamma^2 - \gamma_2^2}(1+|x|^2)$$

and

$$\int_{t_1}^{t_2} |\tilde{u}_r|\, dr \leq C_2(1+|x|)|t_2 - t_1|^{1/2} \quad \forall\, t_1, t_2 \in [s,T].$$

Proof. Let $\widetilde{\xi}$ satisfy (3.10), (3.11) with input $\tilde{u}$. By assumption and the first assertion of Lemma 3.4,

$$\frac{\gamma^2}{2} \int_s^T |\tilde{u}_r|^2\, dr \leq \int_s^T l(\widetilde{\xi}_r)\, dr + \phi(\widetilde{\xi}_T) + \epsilon,$$

which as in the proof of Lemma 3.4 yields (noting $\epsilon \leq 1$)

$$\frac{\gamma^2}{2} \int_s^T |\tilde{u}_r|^2\, dr \leq \frac{\widehat{C}_2}{2}(1+|x|^2) + \frac{\gamma_2^2}{2} \int_s^T |\tilde{u}_r|^2\, dr$$

for proper choice of $\widehat{C}_2, \gamma_2 \in (0,\infty)$. Consequently,

$$\int_s^T |\tilde{u}_r|^2\, dr \leq \frac{\widehat{C}_2}{\gamma^2 - \gamma_2^2}(1+|x|^2).$$

By Cauchy–Schwarz then for any t_1, t_2 such that $s \leq t_1 < t_2 \leq T$

$$\int_{t_1}^{t_2} |\tilde{u}_r|\, dr \leq \left[\frac{\widehat{C}_2}{\gamma^2 - \gamma_2^2}(1+|x|^2) \right]^{1/2} |t_2 - t_1|^{1/2} \leq C_2(1+|x|)|t_2 - t_1|^{1/2}$$

with $C_2 = [\widehat{C}_2/(\gamma^2 - \gamma_2^2)]^{1/2}$. □

Remark 3.6. Let $M_x = C_2(1+|x|)$, $M \in [M_x, \infty)$ and define

$$\mathcal{G}^M_{s,T} = \left\{ u \in \mathcal{U}^U_{s,T} : \int_{t_1}^{t_2} |u_r|\, dr \le M|t_2 - t_1|^{1/2} \quad \forall\, t_1, t_2 \in [s,T] \right\}.$$

We have shown that $V(s,x) = \sup_{u \in \mathcal{G}^M_{s,T}} J(s,x;u)$.

Lemma 3.7. *Let ξ_r be a solution of (3.10) on $[s,T]$ driven by $u \in \mathcal{G}^M_{s,T}$. Then there exist $B_1, B_2, B_3 < \infty$ such that*

$$|\xi_r - x| \le B_1(r-s) + B_2\sqrt{r-s} + B_3|x|(r-s) \quad \forall\, r \in [s,T].$$

Proof. By (3.10),

$$\xi_r - x = \int_s^r f(\xi_\rho) - f(x)\, d\rho + f(x)(r-s) + \int_s^r \sigma(\xi_\rho) u_\rho\, d\rho.$$

Consequently, using (A3.1F), (A3.2F) and the assumption on u, one finds

$$|\xi_r - x| \le K(1+|x|)(r-s) + M m_\sigma \sqrt{r-s} + K\int_s^r |\xi_\rho - x|\, d\rho.$$

Employing Gronwall's inequality then yields the result. □

We can now obtain some DPP results which will be more helpful for this specific class of problems than the general result of Theorem 3.1.

Theorem 3.8. *For any $0 \le s \le t \le T < \infty$ and any $x \in \mathbf{R}^n$,*

$$V(s,x) = \sup_{u \in \mathcal{G}^{M_x}_{s,t}} \left\{ \int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2\, dr + V(t,\xi_t) \right\}. \tag{3.19}$$

Proof. Let $t \in (s,T)$. For any $y \in \mathbf{R}^n$,

$$V(t,y) = \sup_{\hat{u} \in \mathcal{U}^U_{t,T}} \left\{ \int_t^T l(r,\hat{\xi}_r) - \frac{\gamma^2}{2}|\hat{u}_r|^2\, dr + \phi(\hat{\xi}_T) \right\}, \tag{3.20}$$

where $\hat{\xi}_\cdot$ is driven by $\hat{u}$ from initial state $\hat{\xi}_t = y$. Let the right-hand side of (3.19) be denoted by $R(s,x)$. Substituting (3.20) into the right-hand side of (3.19), one obtains

$$\begin{aligned} R(s,x) = \sup_{\tilde{u} \in \mathcal{G}^{M_x}_{s,t}} \sup_{\hat{u} \in \mathcal{U}^U_{t,T}} \bigg\{ & \int_s^t l(r,\tilde{\xi}_r) - \frac{\gamma^2}{2}|\tilde{u}_r|^2\, dr \\ & + \int_t^T l(r,\hat{\xi}_r) - \frac{\gamma^2}{2}|\hat{u}_r|^2\, dr + \phi(\hat{\xi}_T) \bigg\}, \end{aligned}$$

where $\hat{\xi}_t$ is driven by $\hat{u}$ with initial condition $\hat{\xi}_t = \tilde{\xi}_t$, and $\tilde{\xi}$ is driven by $\tilde{u}$ with initial condition $\xi_s = x$. Concatenating the two trajectory segments, and noting that the concatenation of $\tilde{u}$ with $\hat{u}$ is an element of $\mathcal{U}^U_{[s,T)}$, one has

$$R(s,x) \le \sup_{\hat{u}\in\mathcal{U}^U_{s,T}} \left\{ \int_s^T l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + \phi(\xi_T) \right\} = V(s,x).$$

The reverse direction has an analogous proof, and we do not include it. For a similar proof in a deterministic game context, see [91]. □

Remark 3.9. Minor modifications of the above proof also yield the DPPs

$$V(s,x) = \sup_{u\in\mathcal{G}^M_{s,t}} \left\{ \int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + V(t,\xi_t) \right\}$$

for any $M \ge M_x$ and

$$V(s,x) = \sup_{u\in\mathcal{U}^U_{s,t}} \left\{ \int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + V(t,\xi_t) \right\}.$$

We now turn to a DPP for a more specific infinite time-horizon problem than the general problem given above. Recall that the dynamics and initial condition are given by (3.12) and (3.13). The assumptions on the dynamics are modified as follows. We now assume

$$\begin{aligned}
&f \in C(\mathbf{R}^n;\mathbf{R}),\\
&|f(x)-f(y)| \le K|x-y| \quad \forall\, x,y\in\mathbf{R}^n,\\
&(x-y)^T[f(x)-f(y)] \le -c_f|x-y|^2 \quad \forall\, x,y\in\mathbf{R}^n,\\
&f(0)=0,
\end{aligned} \tag{A3.1I}$$

for some $K, c_f < \infty$. We note that this implies

$$x^T f(x) \le -c_f|x|^2$$

for all x. This last inequality implies exponential stability of the system when $u \equiv 0$. As before, we suppose

$$\begin{aligned}
&\sigma \in C(\mathbf{R}^n;\mathcal{L}(\mathbf{R}^l,\mathbf{R}^n)),\\
&|\sigma(x)-\sigma(y)| \le K_\sigma|x-y| \quad \forall\; x,y\in\mathbf{R}^n,\\
&|\sigma(x)| \le m_\sigma \quad \forall\; x\in\mathbf{R}^n,
\end{aligned} \tag{A3.2I}$$

for some $K_\sigma, m_\sigma < \infty$.

Consider the payoff

$$J(x,T,u) = \int_0^T l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr, \tag{3.21}$$

and value function (also known as available storage in this context [52], [108])

$$W(x) = \sup_{u \in \mathcal{U}^U} \sup_{T<\infty} J(x,T,u). \tag{3.22}$$

We assume that

$$\begin{aligned} &l \in C(\mathbf{R}^n;\mathbf{R}), \\ &|l(x)-l(y)| \le C_l(1+|x|+|y|)|x-y| \quad \forall\, x,y \in \mathbf{R}^n, \\ &0 \le l(x) \le \alpha_l |x|^2 \quad \forall\, x \in \mathbf{R}^n, \end{aligned} \tag{A3.3I}$$

for some $C_l, \alpha_l < \infty$. Lastly, we assume that

$$\frac{\gamma^2 c_f^2}{2m_\sigma^2 \alpha_l} > 1. \tag{A3.4I}$$

We remark that under these assumptions the value exists, and that for sufficiently small $\delta > 0$ (Section 3.2 below and [88])

$$0 \le W(x) \le c_f \frac{\gamma^2-\delta}{2m_\sigma^2}|x|^2 \quad \forall\, x \in \mathbf{R}^n. \tag{3.23}$$

We now indicate the more specific DPP that one can obtain in this context.

Theorem 3.10. *Let ξ satisfy (3.12), (3.13). Let $\delta > 0$ be sufficiently small such that (3.23) holds, and such that with $\widehat{\gamma}^2 \doteq \gamma^2 - \delta$ one still has the inequality $(\widehat{\gamma}^2 c_f^2)/(2m_\sigma^2\alpha_l) > 1$. Then for any $\varepsilon > 0$, for all $x \in \mathbf{R}^n$*

$$W(x) = \sup_{u \in \mathcal{U}_{0,T}^{U,\varepsilon,|x|}} \left\{ \int_0^T l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + W(\xi_T) \right\}, \tag{3.24}$$

where

$$\mathcal{U}_{0,T}^{U,\varepsilon,|x|} \doteq \left\{ u \in \mathcal{U}^U \;\middle|\; \frac{1}{2}\|u\|^2_{L_2(0,T)} \le \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{-c_f T} + \frac{\alpha_l}{c_f}\right]|x|^2 \right\} \tag{3.25}$$

Proof. The following proof is adapted from [88]. From Theorem 3.2 and (3.21),

$$W(x) = \sup_{u \in \mathcal{U}_{0,T}^U} \left\{ \int_0^T l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + W(\xi_T) \right\} \doteq \sup_{u \in \mathcal{U}_{0,T}^U} \widehat{J}(x,T,u). \tag{3.26}$$

Let

$$Q_T \doteq \int_0^T \alpha_l |\xi_t|^2\,dt + \frac{c_f\widehat{\gamma}^2}{2m_\sigma^2}|\xi_T|^2, \tag{3.27}$$

which by (A3.3I) and (3.23)

$$\ge \int_0^T l(\xi_t)\,dt + W(\xi_T). \tag{3.28}$$

Note that Q is absolutely continuous and the indefinite integral of its derivative. Then, almost everywhere,

$$\dot{Q}_T = \alpha_l \left[|x|^2 + \int_0^T 2\xi^T \dot{\xi}\, dt \right] + \frac{c_f \widehat{\gamma}^2}{m_\sigma^2} \xi_T^T \dot{\xi}_T,$$

which by (3.12), (A3.1I)

$$\begin{aligned} &\le \alpha_l |x|^2 - 2c_f \alpha_l \int_0^T |\xi|^2\, dt - c_f \frac{c_f \widehat{\gamma}^2}{m_\sigma^2} |\xi_T|^2 \\ &\quad + 2\alpha_l \int_0^T \xi^T \sigma(\xi) u\, dt + \frac{c_f \widehat{\gamma}^2}{m_\sigma^2} \xi_T^T \sigma(\xi_T) u_T \\ &= \alpha_l |x|^2 - 2c_f Q + 2\alpha_l \int_0^T \xi^T \sigma(\xi) u\, dt + \frac{c_f \widehat{\gamma}^2}{m_\sigma^2} \xi_T^T \sigma(\xi_T) u_T. \quad (3.29) \end{aligned}$$

Using the fact that $2a \cdot b \le c_f |a|^2 + \frac{1}{c_f} |b|^2$ on the last two terms of (3.29) yields

$$\dot{Q}_T \le \alpha_l |x|^2 - c_f Q + \frac{\alpha_l}{c_f} \int_0^T u^T \sigma^T(\xi) \sigma(\xi) u\, dt + \frac{\widehat{\gamma}^2}{2m_\sigma^2} u_T^T \sigma^T(\xi_T) \sigma(\xi_T) u_T$$

which by (A3.2I)

$$\le \alpha_l |x|^2 - c_f Q + \frac{\alpha_l m_\sigma^2}{c_f} \int_0^T |u|^2\, dt + \frac{\widehat{\gamma}^2}{2} |u_T|^2.$$

Solving this ordinary differential inequality with $Q_0 = \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} |x|^2$, we find

$$\begin{aligned} Q_T \le\ & \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} |x|^2 e^{-c_f T} + \frac{1 - e^{-c_f T}}{c_f} \alpha_l |x|^2 \\ & + \int_0^T \left[\frac{\alpha_l m_\sigma^2}{c_f} e^{c_f (t-T)} \int_0^t |u_r|^2\, dr + \frac{\widehat{\gamma}^2}{2} e^{c_f (t-T)} |u_t|^2 \right] dt. \end{aligned}$$

Applying integration by parts to the first term in the integral yields

$$\begin{aligned} Q_T &\le \left[\frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 + \int_0^T \left[\frac{\alpha_l m_\sigma^2}{c_f^2} (1 - e^{c_f (t-T)}) + \frac{\widehat{\gamma}^2}{2} e^{c_f (t-T)} \right] |u|^2\, dt \\ &= \left[\frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 + \int_0^T \left[\frac{\alpha_l m_\sigma^2}{c_f^2} - \frac{\widehat{\gamma}^2}{2} \right] (1 - e^{c_f (t-T)}) |u|^2\, dt \\ &\quad + \frac{\widehat{\gamma}^2}{2} \|u\|^2_{L_2[0,T]}. \end{aligned} \quad (3.30)$$

Let $\widetilde{\delta} \doteq \frac{\widehat{\gamma}^2}{2} - \frac{\alpha_l m_\sigma^2}{c_f^2}$, and note by (A3.4I) that $\widetilde{\delta} > 0$. We have

$$Q_T \le \left[\frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 - \widetilde{\delta} \int_0^T (1 - e^{c_f (t-T)}) |u|^2\, dt + \frac{\widehat{\gamma}^2}{2} \|u\|_{L_2[0,T]}. \quad (3.31)$$

Then, by (3.26), (3.28), (3.31) and the definition of $\widehat{\gamma}$,

$$\widehat{J}(x,T,u) \leq \left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{-c_fT} + \frac{\alpha_l}{c_f}\right]|x|^2 - \widetilde{\delta}\int_0^T (1-e^{c_f(t-T)})|u|^2\,dt - \frac{\delta}{2}\|u\|^2_{L_2[0,T]}$$
$$\leq \left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{-c_fT} + \frac{\alpha_l}{c_f}\right]|x|^2 - \frac{\delta}{2}\|u\|^2_{L_2[0,T]}. \tag{3.32}$$

On the other hand, by (A3.3I), (3.23) and (3.26), $\widehat{J}(x,T,0) \geq 0$, and so for ε-optimal u

$$\widehat{J}(x,T,u) \geq -\varepsilon. \tag{3.33}$$

Comparing (3.32) and(3.33), we see that for ε-optimal u, we have

$$\tfrac{1}{2}\|u\|^2_{L_2[0,T]} \leq \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{-c_fT} + \frac{\alpha_l}{c_f}\right]|x|^2. \tag{3.34}$$

This is an upper bound on the size of ε-optimal u which is independent of T (using $e^{-c_fT} \leq 1$). □

3.2 Viscosity Solutions

In the previous section, we concentrated on the relationship between the DPP and the control problem value function for the example problem classes we will concentrate on. As noted earlier, the DPE is obtained by an infinitesimal limit in the DPP, and takes the form of a nonlinear, first-order Hamilton–Jacobi–Bellman PDE (HJB PDE) in these problem classes. In the finite time-horizon problem, it is a time-dependent PDE over $(s,T) \times \mathbf{R}^n$ with terminal-time boundary data. In the infinite time-horizon case, it is a steady-state PDE over $\mathbf{R}^n$, with "boundary" data only at $x = 0$ ($W(0) = 0$) in the specific problem class we will concentrate on.

The value functions for such control problems are typically viscosity solutions of the corresponding HJB PDEs. Thus, when we use our numerical method to obtain the value function, we are also solving the nonlinear HJB PDE (for a viscosity solution). Alternatively, if one has a nonlinear, first-order HJB PDE, one may solve it via a numerical method for the associated control problem. PDEs have associated semigroups [96], and in the HJB case, this semigroup is equivalent to the DPP.

The relationship between the HJB PDE and the associated control problem is a classical topic. Some early references are [14], [46], [65], among many notable others. This early theory was restricted mainly to the case where the PDE had a classical solution (everywhere differentiable, and meeting the boundary conditions pointwise). However, it was well known that such control problems (unless containing an additional nondegenerate diffusion term in the dynamics) did not typically have classical solutions. (One of the very few counterexamples is the linear/quadratic case.)

A major advance came in the 1980s with the development of the notion of the *viscosity solution* of a PDE. A few early works are [10], [22], [23] and [24], and some books are [8], [11], [36], [47] and [48]. Again, the relevant literature is vast, and these represent only a tiny fraction of the work in this area. The viscosity solution of an HJB PDE need not be a classical solution, and typically has discontinuities in the gradient; in some problem classes it may even have discontinuities in the solution itself as well as not meeting the boundary conditions pointwise. Note that there are typically many weak solutions to a PDE which solve the PDE pointwise almost everywhere, and the viscosity solution conditions isolate the correct such weak solution. Specifically, there is generally a unique viscosity solution to the HJB PDE with given boundary data. A counterexample to this is our steady-state problem class example, where an additional condition is required to isolate the correct solution [88], [89], [106].

We will not actually need viscosity solution theory to compute the control problem value function, as we work, instead, directly with the DPP. Of course, viscosity solution theory is necessary to associate this value function with the solution of the associated HJB PDE. Consequently viscosity solution theory is not absolutely essential to the bulk of the theory developed after this chapter. Nonetheless, viscosity solution theory is a fundamental component in the way one thinks about nonlinear control, and so has a place in any book dealing with DP and nonlinear control. There are many excellent works on the theory of viscosity solutions (see above references), and we will not duplicate that material here. However, it will be useful to review the basics of viscosity solution theory here, and to indicate the relationship between the viscosity solution and the value function more completely for our example problem classes.

We start with a definition of viscosity solution for the finite time-horizon problem. (One can give a single definition for both the finite time-horizon and infinite time-horizon cases, but it appears that it will be simpler for our needs here to write separate definitions.) The following definition is for a continuous viscosity solution, as our assumptions will preclude problems with discontinuous viscosity solutions. We will also only need to consider cases where the boundary conditions are met pointwise. More general definitions can be found in the references.

The result for the finite time-horizon problem is quite well known. We consider the problem with dynamics (3.10), (3.11), payoff (3.14) and value (3.15) for initial times $s \in (0, T)$. We assume (A3.1F)–(A3.4F). The corresponding PDE problem is

$$0 = -V_s(s,x) + H(s,x,\nabla_x V(s,x)) \quad \forall\,(s,x) \in (0,T) \times \mathbf{R}^n, \tag{3.35}$$

$$V(T,x) = \phi(x) \quad \forall\, x \in \mathbf{R}^n, \tag{3.36}$$

where the Hamiltonian, H, is given by

$$H(s,x,p) \doteq -\max_{u \in U} \left\{ [f(s,x) + \sigma(x)u]^T p + l(s,x) - \frac{\gamma^2}{2}|u|^2 \right\}. \quad (3.37)$$

Here, V_s represents the partial derivative with respect to time variable, and $\nabla_x V$ represents the gradient with respect to the space variable. Where no confusion arises, we may often use ∇ in place of ∇_x to represent the gradient with respect to the space variable. It is helpful to note that in the common case where $U \equiv \mathbf{R}^l$, this Hamiltonian takes the form

$$H(s,x,p) = -\left[f^T(s,x)p + l(s,x) + \frac{1}{2\gamma^2} p^T \sigma(x)\sigma^T(x)p \right]. \quad (3.38)$$

There are several equivalent definitions of a continuous viscosity solution, particularly with regard to the set of test functions. We recall one definition of a continuous viscosity solution of (3.35) (c.f. [48], definition 2.4.1). Suppose $V \in C((0,T] \times \mathbf{R}^n)$. Further, suppose that for any $g \in C^1([0,T] \times \mathbf{R}^n)$ and $(s,x) \in (0,T) \times \mathbf{R}^n$ such that $V - g$ has a local maximum at (s,x), one has

$$-g_t(s,x) + H(s,x,\nabla_x g(s,x)) \le 0.$$

Then V is a *continuous viscosity subsolution* of (3.35). On the other hand, suppose that for any $g \in C^1([0,T] \times \mathbf{R}^n)$ and $(s,x) \in (0,T) \times \mathbf{R}^n$ such that $V - g$ has a local minimum at (s,x), one has

$$-g_t(s,x) + H(s,x,\nabla_x g(s,x)) \ge 0.$$

Then V is a *continuous viscosity supersolution* of (3.35). If V is both a continuous viscosity subsolution and a continuous viscosity supersolution, then it is a *continuous viscosity solution.*

We now state the viscosity solution result. This result is rather standard; the fact that we allow quadratic growth in the cost criteria moves it slightly beyond the realm of the most basic results. For completeness, we also provide a partial proof. See the references for a more complete discussion and extensions.

Theorem 3.11. *Under assumptions (A3.1F)–(A3.4F), and taking $\gamma > \gamma_2$ (see Lemma 3.4), the value function (3.15) is a continuous viscosity solution of (3.35) meeting the terminal condition (3.36) pointwise.*

Proof. As noted above, we provide a partial proof for completeness. First we address the continuity issue. Let ξ and η satisfy (3.10) with initial conditions $\xi_s = x$ and $\eta_s = y$, respectively, for some $s \in [0,T)$. In particular, fix some $R < \infty$, and let $|x|, |y| \le R$. Let $u \in \mathcal{U}^U_{s,T}$ be ε-optimal for either x or y with $\varepsilon \in (0,1]$. We have

$$\xi_t - \eta_t = \xi_s - \eta_s + \int_s^t [f(r,\xi_r) - f(r,\eta_r)]\,dr + \int_s^t [\sigma(\xi_r) - \sigma(\eta_r)]u_r\,dr.$$

Using Assumptions (A3.1F) and (A3.2F), this yields

$$|\xi_t - \eta_t| \le |x-y| + K\int_s^t |\xi_r - \eta_r|\,dr + K_\sigma \int_s^t |\xi_r - \eta_r||u_r|\,dr.$$

By Cauchy–Schwarz, there exists $C_1 < \infty$ (dependent on T) such that

$$|\xi_t - \eta_t|^2 \le C_1 \left\{ |x-y|^2 + \left[1 + \int_s^t |u_r|^2\,dr\right] \int_s^t |\xi_r - \eta_r|^2\,dr \right\}.$$

By the ε-optimality of u_r and Lemma 3.5, there exists $C_3 < \infty$ (dependent on R, T) such that

$$|\xi_t - \eta_t|^2 \le C_3 \left\{ |x-y|^2 + \int_s^t |\xi_r - \eta_r|^2\,dr \right\}.$$

Then, using Gronwall's Inequality, we find that there exists $C_4 = C_4(R,T) < \infty$ such that

$$|\xi_t - \eta_t| \le C_4|x-y| \quad \forall\, t \in [s,T],\ \forall x, y \in \overline{B}_R(0) \tag{3.39}$$

(where we recall $\overline{B}_R(0) = \{x \in \mathbf{R}^n |\, |x| \le R\}$).

Now, let us specifically take u_ε to be ε-optimal for initial condition $\xi_s = x$ (rather than for $\eta_s = y$). Then

$$\begin{aligned} V(s,x) - V(s,y) &\le J(s,x,u^\varepsilon) - J(s,y,u^\varepsilon) + \varepsilon \\ &\le \int_s^T [l(t,\xi_t) - l(t,\eta_t)]\,dt + \phi(\xi_T) - \phi(\eta_T) + \varepsilon, \end{aligned}$$

which by (A3.3F) and (A3.4F)

$$\begin{aligned} &\le C_l \int_s^T (1 + |\xi_t| + |\eta_t|)|\xi_t - \eta_t|\,dt \\ &\quad + C_\phi(1 + |\xi_T| + |\eta_T|)|\xi_T - \eta_T| + \varepsilon, \end{aligned}$$

which by (3.39)

$$\begin{aligned} &\le C_4|x-y| \left\{ C_l \int_s^T (1 + |\xi_t| + |\eta_t|)\,dt + C_\phi(1 + |\xi_T| + |\eta_T|) \right\} \\ &\quad + \varepsilon, \end{aligned}$$

which by (3.16) with the $C_2 = C_2(T)$ given there

$$\begin{aligned} &\le C_4|x-y| \left\{ C_l \left[T + 2C_2(1+R)T + 2C_2 \int_s^T \int_s^t |u_r|\,dr\,dt \right] \right. \\ &\quad \left. + C_\phi \left[1 + 2C_2(1+R) + 2C_2 \int_s^T |u_t|\,dt \right] \right\} + \varepsilon. \end{aligned}$$

Then, using Lemma 3.5, there exists $C_5 = C_5(R,T) < \infty$

$$\leq C_5|x-y|+\varepsilon.$$

Because this is true for all $\varepsilon > 0$,

$$V(s,x) - V(s,y) \leq C_5|x-y|$$

(with $|x|, |y| \leq R$). By symmetry, one obtains

$$|V(s,x) - V(s,y)| \leq C_5|x-y| \tag{3.40}$$

(with $|x|, |y| \leq R$). This is (local Lipschitz) continuity in the space variable. Continuity with respect to the time variable follows similarly, and we do not include the proof.

Now that we have continuity of value function V, we show that it satisfies the conditions for a viscosity solution of (3.35). Suppose (s,x) is a local maximum of $V-g$ where $g \in C^1$, and that the viscosity subsolution inequality is *not* satisfied at (s,x). Then there exists $\theta > 0$ such that

$$-g_t(s,x) + H(s,x,\nabla_x g(s,x)) \geq \theta > 0.$$

Then, by the definition of H and the assumptions, there exists $\delta > 0$ such that for $t \in [s, s+\delta)$, $y \in B_\delta(x)$,

$$-g_t(t,y) + H(t,x,\nabla_x g(t,y)) \geq \theta/2.$$

Let $\widetilde{u} \in \mathcal{G}^{M_x}_{s,T}$. Let ξ satisfy (3.10), (3.11) with control $\widetilde{u}$. By Lemma 3.7, there exists $\tilde{\delta} \in (0,\delta)$ such that for all $t \in [s, s+\tilde{\delta}]$, one has $\xi_t \in B_\delta(x)$. Consequently, for $t \in [s, s+\tilde{\delta}]$,

$$-g_t(t,\xi_t) + H(t,\xi_t,\nabla_x g(t,\xi_t)) \geq \theta/2,$$

and then by the definition of H, one has

$$-g_t(t,\xi_t) - \left\{ [f(t,\xi_t) + \sigma(\xi_t)\widetilde{u}_t]^T \nabla_x g(t,\xi_t) + l(t,\xi_t) - \frac{\gamma^2}{2}|\widetilde{u}_t|^2 \right\} \geq \theta/2.$$

Integrating (and multiplying by -1), one finds

$$\int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widetilde{u}_r|^2 + [f(r,\xi_r) + \sigma(\xi_r)\widetilde{u}_r]^T \nabla_x g(r,\xi_r)\, dr \leq \frac{-\theta(t-s)}{2},$$

which implies

$$\int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widetilde{u}_r|^2\, dr + g(t,\xi_t) - g(s,x) \leq \frac{-\theta(t-s)}{2}. \tag{3.41}$$

Recall that V satisfies the DPP of Theorem 3.8. Let $\varepsilon \in (0,1]$, and let $\widehat{u}$ be ε-optimal in (3.19), and let $\xi_\cdot$ again represent the corresponding trajectory. Then, by Theorem 3.8,

$$V(s,x) \le \int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widehat{u}_r|^2\,dr + V(t,\xi_t) + \varepsilon. \tag{3.42}$$

By continuity and the supposition that (s,x) is a local maximum of $V-g$, there exists $\hat{\delta} \in (0,\tilde{\delta}]$ such that

$$V(s,x) - g(s,x) \ge V(t,\xi_t) - g(t,\xi_t) \quad \forall\, t \in [s, s+\hat{\delta}]. \tag{3.43}$$

Combining (3.42) and (3.43), one has

$$-\varepsilon \le \int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widehat{u}_r|^2\,dr + g(t,\xi_t) - g(s,x).$$

Letting $\varepsilon < (\theta t)/2$, one has

$$\int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widehat{u}_r|^2\,dr + g(t,\xi_t) - g(s,x) > -\frac{\theta(t-s)}{2}. \tag{3.44}$$

On the other hand, taking $\widetilde{u} = \widehat{u}$ in (3.41) yields

$$\int_s^t l(r,\xi_r) - \frac{\gamma^2}{2}|\widehat{u}_r|^2\,dr + g(t,\xi_t) - g(s,x) \le -\frac{\theta(t-s)}{2}. \tag{3.45}$$

Clearly, (3.44), (3.45) is a contradiction, and so

$$-g_t(s,x) + H(s,x,\nabla_x g(s,x)) \le 0.$$

Consequently, value function V is a continuous viscosity subsolution of (3.35). The proof that the supersolution condition holds is similar, and we do not include it. □

Remark 3.12. Note that it has also been shown that V is locally Lipschitz continuous in the space variable, x.

It has now been demonstrated that the value function is a continuous viscosity solution of the associated HJB PDE for this problem class. However, this is of course only half of the story with regard to the relationship. The other half is the demonstration that the value is the only continuous viscosity solution (uniqueness).

Such uniqueness results have been a foundation of viscosity solution theory. As noted before, there are many excellent references on viscosity solution theory, and that is not a principal focus of this book, so we will not go into this in great detail. We will attempt to present sufficient detail so that the reader is comfortable with the area, but not so much as to detract from the principal focus of the book — max-plus based numerical methods. It *is* important to note that the max-plus results will be used to obtain solutions of DPPs, and therefore by the results of the previous section, control problem value functions. It is the one-to-one relationship between value functions (or

DPP solutions) and viscosity solutions of HJB PDEs that then allows us to say that our numerical methods are producing the unique solutions of the corresponding HJB PDEs.

Early uniqueness results for finite time-horizon HJB PDEs can be found in [10], [22], [23], [24], [57] among others, but the first results did not allow for Lipschitz dynamics/quadratic cost criteria (critical to much control theory which has a long history in studying linear/quadratic problems). A uniqueness result handling our problem here was obtained in [92] for this class of problems, but it required a constant σ. Closely related results appear in [9], [59]. In [26], a more general result (due to DaLio) for a game problem subsumes the uniqueness result needed here. Below, we provide a particularization of that result.

We first need to define the space in which uniqueness will be obtained. It is the set of continuous solutions satisfying a quadratic growth bound. This space, denoted by $\mathcal{K}$, is given by

$$D_x^+ U(t,x) = \left\{ p \in \mathbf{R}^n : \ \limsup_{y \to x} \frac{U(t,y) - U(t,x) - (y-x)\cdot p}{|x-y|} \le 0 \right\},$$

$$||U||_R := \sup\{|U(t,x)| + |p|, \ (t,x) \in [0,T] \times \overline{B}_R(0)\,, \, p \in D_x^+ U(t,x)\},$$

and

$$\mathcal{K} := \left\{ U \in C([0,T] \times \mathbf{R}^n) : \ U(t,x) \ge 0, \ ||U||_R < +\infty \quad \forall\, R > 0 \right.$$
$$\left. \text{and} \quad \sup_{(t,x)\in[0,T]\times\mathbf{R}^n} \frac{|U(t,x)|}{1+|x|^2} < +\infty \right\}.$$

Note that by Lemma 3.4 and Remark 3.12, the value function lies in $\mathcal{K}$.

Theorem 3.13. *Assume (A3.1F)–(A3.4F). If $V_1, V_2 \in \mathcal{K}$ are two continuous viscosity solutions in class $\mathcal{K}$, then $V_1 = V_2$. Further, if $\gamma > \gamma_2$ (see Lemma 3.4), then the value function is the unique continuous solution of (3.35), (3.36) in the class $\mathcal{K}$.*

In order to focus on the max-plus numerical methods, a simplification of Da Lio's proof to the case here is delayed to Appendix A.

We now turn to the HJB PDE for the infinite time-horizon problem class discussed in the previous section. As noted above, the viscosity solution of a PDE is typically unique (and is typically the "correct" solution as well). The infinite time-horizon problem with Lipschitz dynamics and quadratic (or higher) growth in the cost criterion violates this rule of thumb.

The HJB PDE corresponding to the infinite time-horizon problem (3.12), (3.13), (3.21), (3.22) is

$$0 = H(x, \nabla W(x)) \quad \forall\, x \in \mathbf{R}^n \setminus \{0\}, \tag{3.46}$$

$$W(0) = 0, \tag{3.47}$$

where the Hamiltonian, H, is given by

$$H(x,p) \doteq -\max_{u\in U}\left\{[f(x)+\sigma(x)u]^T p + l(x) - \frac{\gamma^2}{2}|u|^2\right\}. \tag{3.48}$$

Again, in the case where $U \equiv \mathbf{R}^l$, the Hamiltonian takes the form

$$H(x,p) = -\left[f^T(x)p + l(x) + \frac{1}{2\gamma^2}p^T\sigma(x)\sigma^T(x)p\right]. \tag{3.49}$$

As with the finite time-horizon case, we can restrict our attention to continous viscosity solutions meeting the boundary condition pointwise ($W(0)=0$ in this case). Let $W \in C(\mathbf{R}^n)$. W will be a *continuous viscosity solution* of (3.46) if it satisfies both of the following viscosity subsolution and viscosity supersolution conditions. Suppose that for any $g \in C^1(\mathbf{R}^n)$ and $x \in \mathbf{R}^n$ such that $W-g$ has a local maximum at x, one has

$$H(x,\nabla g(x)) \le 0.$$

Then W is a *continuous viscosity subsolution* of (3.46). One the other hand, suppose that for any $g \in C^1(\mathbf{R}^n)$ and $x \in \mathbf{R}^n$ such that $W-g$ has a local minimum at x, one has

$$H(x,\nabla g(x)) \ge 0.$$

Then W is a *continuous viscosity supersolution* of (3.46).

A simple, one-dimensional example indicating the lack of uniqueness is

$$0 = -\left[-xW_x + x^2 + \tfrac{1}{8}W_x^2\right].$$

There are two C^∞ solutions with $W(0)=0$:

$$W^1(x) = (2-\sqrt{2})x^2 \ \text{ and } \ W^2(x) = (2+\sqrt{2})x^2,$$

and an infinite number of viscosity solutions such as

$$W(x) = \begin{cases} (2-\sqrt{2})x^2 & \text{if } x \le 1 \\ (2+\sqrt{2})x^2 - 2\sqrt{2} & \text{if } x > 1. \end{cases}$$

We will show that there exists a continuous viscosity solution of (3.46), (3.47) satisfying growth condition (3.23), and that this solution is the value function (3.22). As with the finite time-horizon case, there will be two major parts to the proof of the above statement. Here, we will start with what is referred to as a verification theorem — which states that a solution of the PDE must be the value function. The second major component will be the proof that there exists a solution. The results presented here are distilled from a game problem discussed in [88] (see also [89]). The results in [88] rely heavily on the structure of our particular problem class. A related but different type of result, under weaker conditions, can be found in [106].

Suppose $\overline{W}$ is a continuous viscosity solution of (3.46), (3.47) satisfying condition (3.23) and locally Lipschitz in x. Then for any $T \in (0, \infty)$, $\overline{W}$ is also a (steady-state) solution of the Cauchy problem

$$0 = -V_s(s,x) + H(x, \nabla V(s,x)) \quad \forall (s,x) \in (-\infty, T) \times \mathbf{R}^n, \tag{3.50}$$

$$V(T,x) = \overline{W}(x) \quad \forall x \in \mathbf{R}^n, \tag{3.51}$$

where H is given by (3.48). However, (3.50), (3.51) is identical to our finite time-horizon PDE problem (3.35), (3.36). Consequently, by Theorem 3.13, we have

Lemma 3.14. *$\overline{W}$ is the unique continuous viscosity solution of (3.50), (3.51) in class $\mathcal{K}$.*

We also immediately find that if $\gamma > \gamma_2(T)$ (see Lemma 3.4), then $\overline{W}$ is the value function of the corresponding finite time-horizon control problem with dynamics (3.12) for any initial condition

$$\xi_s = x \tag{3.52}$$

with payoff and value

$$J(s,x,T,u) = \int_s^T l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + \overline{W}(\xi_T), \tag{3.53}$$

$$V(s,x) = V(s,x;T) = \sup_{u \in \mathcal{U}^U} J(s,x,T,u), \tag{3.54}$$

where we use the notation $V(s,x;T)$ to emphasize the dependence on the terminal time (actually dependence on $T - s$). However, this result is not quite sufficient for our needs. Using the stability implied by the third and fourth parts of Assumption (A3.1I), one finds

Lemma 3.15. *$\overline{W}$ is the value function of problem (3.12), (3.52), (3.53), (3.54) for any γ satisfying (A3.4I), i.e., $\overline{W}(x) = V(s,x)$ for all $-\infty < s \leq T < \infty$, for all $x \in \mathbf{R}^n$.*

The proof can be found in Appendix A.

Note that $\overline{W}$ is the value function for the finite time-horizon problem independent of s, T. We now use this representation of $\overline{W}$ to show that it must be the value of our infinite time-horizon problem. Specifically, we will take $s = 0$, and let $T \to \infty$. By showing that near optimal trajectories (of the dynamics) are such that $\xi_t \to 0$ (roughly speaking), the terminal cost satisfies $\overline{W}(\xi_T) \to 0$ (roughly speaking), and we obtain the result. Let us now fill this argument in a bit.

Lemma 3.16. *Let ξ satisfy (3.12), (3.52). Let $\delta > 0$ be sufficiently small such that (3.23) holds, and such that with $\widehat{\gamma}^2 \doteq \gamma^2 - \delta$ one still has inequality $(\widehat{\gamma}^2 c_f^2)/(2m_\sigma^2 \alpha_l) > 1$. Then for any $\varepsilon > 0$,*

$$\overline{W}(x) = V(0,x) = V(0,x;T) = \sup_{u \in \mathcal{U}_{0,T}^{U,\varepsilon,|x|}} \left\{ \int_0^T l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr + \overline{W}(\xi_T) \right\}$$

for all $T \in (0,\infty)$ and $x \in \mathbf{R}^n$ where $\mathcal{U}_{0,T}^{U,\varepsilon,|x|}$ is given by (3.25).

Note that Lemma 3.16 follows directly from the definition of $V(0,x) = V(0,x;T)$ and Theorem 3.10.

Lemma 3.17. *Let u be ε-optimal for problem (3.12), (3.52), (3.53), (3.54) and let ξ be the corresponding state process. Then*

$$\int_s^T |\xi_t|^2\,dt \le \frac{2\varepsilon}{\delta}\frac{m_\sigma^2}{c_f^2} + \left[\frac{m_\sigma^2}{\delta c_f}\left(\frac{2\alpha_l}{c_f^2} + \frac{\gamma^2}{m_\sigma^2}\right) + \frac{1}{c_f}\right]|x|^2. \tag{3.55}$$

Proof. Let $R_T \doteq \int_s^T |\xi_t|^2\,dt$. Then by (3.12) and (A3.1I),

$$\dot{R}_T \le -2c_f \int_s^T |\xi|^2\,dt + 2\int_s^T \xi^T \sigma(\xi) u\,dt + |x|^2, \tag{3.56}$$

which by the fact that $2a \cdot b \le c_f|a|^2 + (1/c_f)|b|^2$ for all a, b and (A3.2I)

$$\le -c_f R_T + \frac{m_\sigma^2}{c_f}\int_s^T |u|^2\,dt + |x|^2. \tag{3.57}$$

Solving this ODI, one finds

$$R_T \le \frac{m_\sigma^2}{c_f^2}\int_s^T (1 - e^{c_f(t-T)})|u|^2\,dt + \frac{1}{c_f}|x|^2.$$

Supposing that u is ε-optimal, and using Lemma 3.16 and the definition of $\mathcal{U}_{0,T}^{U,\varepsilon,|x|}$, yields the result. □

By (3.55) with $s = 0$, we see that for ε-optimal u there exists $\tau \in [T/2, T]$ such that

$$|\xi_\tau|^2 \le \frac{2}{T}\left\{\frac{2\varepsilon}{\delta}\frac{m_\sigma^2}{c_f^2} + \left[\frac{m_\sigma^2}{\delta c_f}\left(\frac{2\alpha_l}{c_f^2} + \frac{\gamma^2}{m_\sigma^2}\right) + \frac{1}{c_f}\right]|x|^2\right\} \tag{3.58}$$

for any $T < \infty$.

Now we need to assert that controls which are ε-optimal for the problem over $[0,T]$ are also ε-optimal for the problem over $[0,\tau]$ for any $\tau \in [0,T]$. Suppose that $\widehat{u}$ is ε-optimal over $[0,T]$, with corresponding state process $\widehat{\xi}$ given by (3.12), (3.13). Then

$$\int_0^T \left[l(\widehat{\xi}) - \frac{\gamma^2}{2}|\widehat{u}|^2\right]dt + \overline{W}(\widehat{\xi}_T) \ge \overline{W}(x) - \varepsilon. \tag{3.59}$$

Let $\tau \in [0, T]$ and suppose that $\widehat{u}$ is not ε-optimal over $[0, \tau]$. Then

$$\int_0^\tau \big[l(\widehat{\xi}) - \frac{\gamma^2}{2}|\widehat{u}|^2\big]\,dt + \overline{W}(\widehat{\xi}_\tau) < \overline{W}(x) - \varepsilon. \tag{3.60}$$

However, because $\overline{W}$ is also the value function for problem (3.12), (3.52), (3.53), (3.54) with $s = \tau$, (3.60) implies

$$\int_0^\tau \big[l(\widehat{\xi}) - \frac{\gamma^2}{2}|\widehat{u}|^2\big]\,dt + \int_\tau^T \big[l(\widehat{\xi}) - \frac{\gamma^2}{2}|\widehat{u}|^2\big]\,dt + \overline{W}(\widehat{\xi}_T) < \overline{W}(x) - \varepsilon$$

or

$$\int_0^T \big[l(\widehat{\xi}) - \frac{\gamma^2}{2}|\widehat{u}|^2\big]\,dt + \overline{W}(\widehat{\xi}_T) < \overline{W}(x) - \varepsilon,$$

which contradicts (3.59). Thus one has obtained the following lemma.

Lemma 3.18. *If u is ε-optimal for problem (3.12), (3.52), (3.53), (3.54) over interval $[0, T]$, then it is also ε-optimal for problem (3.12), (3.52), (3.53), (3.54) over any subinterval $[0, \tau]$, i.e.,*

$$\int_0^\tau \big[l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\big]\,dt + \overline{W}(\xi(\tau)) \geq \overline{W}(x) - \varepsilon. \tag{3.61}$$

By (3.61) and (3.23), if u is ε-optimal , then

$$\int_0^\tau \big[l(\xi) - \frac{\gamma^2}{2}|u|^2\big]\,dt + \frac{c_f\gamma^2}{2m_\sigma^2}|\xi(\tau)|^2 \geq \overline{W}(x) - \varepsilon. \tag{3.62}$$

However, by (3.58), given $\overline{\varepsilon} > 0$, we can choose T large enough so that $|\xi_\tau|^2 \leq \overline{\varepsilon}$. Therefore by (3.62), there exists $T < \infty$, $\tau \in [0, T]$ and ε-optimal u such that

$$\int_0^\tau \big[l(\xi) - \frac{\gamma^2}{2}|u|^2\big]\,dt \geq \overline{W}(x) - \varepsilon - \frac{c_f\gamma^2}{2m_\sigma^2}\overline{\varepsilon}. \tag{3.63}$$

Let

$$\widetilde{u}_t \doteq \begin{cases} u_t & \text{if } t \leq \tau \\ 0 & \text{if } t > \tau. \end{cases}$$

Then by Assumption (A3.3I) and (3.63)

$$\int_0^\infty \big[l(\widetilde{\xi}) - \frac{\gamma^2}{2}|\widetilde{u}|^2\big]\,dt \geq \overline{W}(x) - \varepsilon - \frac{c_f\gamma^2}{2m_\sigma^2}\overline{\varepsilon},$$

where $\widetilde{\xi}$ is the state process corresponding to $\widetilde{u}$. This implies that

$$\sup_{u\in\mathcal{U}^U} \int_0^\infty \big[l(\xi) - \frac{\gamma^2}{2}|u|^2\big]\,dt \geq \overline{W}(x) - \varepsilon - \frac{c_f\gamma^2}{2m_\sigma^2}\overline{\varepsilon}.$$

Because ε and $\overline{\varepsilon}$ were arbitrary, we have

$$\sup_{u\in\mathcal{U}^U}\int_0^\infty \left[l(\xi)-\frac{\gamma^2}{2}|u|^2\right]dt \geq \overline{W}(x). \tag{3.64}$$

By Assumption (A3.3I), this is equivalent to

$$\overline{W}(x) \leq \sup_{u\in\mathcal{U}^U}\sup_{T<\infty}\int_0^T \left[l(\xi)-\frac{\gamma^2}{2}|u|^2\right]dt. \tag{3.65}$$

On the other hand, because $\overline{W}$ is the value of (3.12), (3.52), (3.53), (3.54), we have

$$\overline{W}(x) \geq \int_0^T \left[l(\xi)-\frac{\gamma^2}{2}|u|^2\right]dt + \overline{W}(\xi_T) \geq \int_0^T \left[l(\xi)-\frac{\gamma^2}{2}|u|^2\right]dt \tag{3.66}$$

for all $u \in \mathcal{U}^U$ and $T < \infty$. By (3.65) and (3.66) we have the representation/uniqueness result.

Theorem 3.19. *Let $\overline{W}$ be a continuous viscosity solution of (3.46),(3.47) satisfying (3.23), and which is locally Lipschitz in x. Then*

$$\overline{W}(x) = \sup_{u\in\mathcal{U}^U}\sup_{T<\infty}\int_0^T \left[l(\xi_t)-\frac{\gamma^2}{2}|u_t|^2\right]dt. \tag{3.67}$$

This completes the verification theorem component of the relationship between the infinite time-horizon control problem and the corresponding HJB PDE. All that remains is to prove that there actually exists a continuous viscosity solution of (3.46), (3.47) satisfying (3.23) which is locally Lipschitz in x. In some cases, results such as this can be taken directly from the PDE literature. In this case here, we can obtain existence by using a control argument. In particular, we obtain the following.

Theorem 3.20. *There exists a continuous viscosity solution to (3.46), (3.47) satisfying (3.23), and which is locally Lipschitz in x.*

Proof. Consider the finite time-horizon problem given by dynamics and initial condition (3.12), (3.13) with payoff and value

$$\widehat{J}(x,T,u) = \int_0^T \left[l(\xi_t)-\frac{\gamma^2}{2}|u_t|^2\right]dt, \tag{3.68}$$

$$\widehat{V}(x,T) = \sup_{u\in\mathcal{U}^U}\widehat{J}(x,T,u). \tag{3.69}$$

Then by a proof similar to that of Lemma 3.15, $\widehat{V}$ is a continuous viscosity solution to

$$\begin{aligned}
&0 = V_T - \left[\frac{1}{2\gamma^2}\nabla V^T\sigma(x)\sigma^T(x)\nabla V + g^T(x)\nabla V + l(x)\right] \quad (x,T)\in\mathbf{R}^n\times(0,\infty)\\
&V(x,0) = 0 \qquad x\in\mathbf{R}^n
\end{aligned} \tag{3.70}$$

(where we note that T is the time variable in the PDE here).

Note that $\int_0^T l(\xi_t)\,dt$ is similar to the Q_T appearing in the proof of Theorem 3.10 but without the terminal cost. So by the same arguments as used for Q, one can show that

$$\widehat{J}(x,T,u) \leq \frac{\alpha_l}{c_f}|x|^2 - \frac{\widetilde{\delta}}{2}\int_0^T |u_t|^2\,dt \tag{3.71}$$

where $\frac{\widetilde{\delta}}{2} \doteq \frac{\gamma^2}{2} - \frac{\alpha_l m_\sigma^2}{c_f^2} > 0$. This implies

$$\widehat{J}(x,T,u) \;\leq \frac{\alpha_l}{c_f}|x|^2,$$

and so by (A3.4I), there exists $\widehat{\gamma} < \gamma$ such that

$$\widehat{J}(x,T,u) \;< \frac{c_f\widehat{\gamma}^2}{2m_\sigma^2}|x|^2,$$

which yields

$$\widehat{V}(x,T) < \frac{c_f\widehat{\gamma}^2}{2m_\sigma^2}|x|^2 \quad \forall\,(x,T) \in \mathbf{R}^n \times (0,\infty). \tag{3.72}$$

Let $T_2 > T_1$. Then by taking $u_t = 0$ for $t \in (T_1,T_2)$, it is easy to see that

$$\widehat{V}(x,T_2) \geq \widehat{V}(x,T_1).$$

Because $\widehat{V}$ is monotonically increasing and bounded above, we see that there exists $\widehat{W}$ such that

$$\widehat{V}(x,T) \uparrow \widehat{W}(x) \quad \forall\, x \in \mathbf{R}^n. \tag{3.73}$$

We now show that this convergence is uniform on compact sets. By (3.71) and (A3.3I), we see that for ε-optimal u for problem (3.69), we have

$$\frac{1}{2}\|u\|^2_{L_2[0,T]} \leq \frac{1}{\widetilde{\delta}}\left[\varepsilon + \frac{\alpha_l}{c_f}|x|^2\right]. \tag{3.74}$$

Let $x,y \in \overline{B}_R$. Let ξ and η satisfy (3.12) with different initial conditions, $\xi_0 = x$ and $\eta_0 = y$, but both with the same control, u, which is ε-optimal for initial condition x. Then by (A3.1I), (A3.2I)

$$\frac{d}{dt}|\xi_t - \eta_t|^2 \leq -2c_f|\xi_t - \eta_t|^2 + 2|\xi_t - \eta_t|\,(K_\sigma|\xi_t - \eta_t||u_t|)$$

and using inequality $2a\cdot b \leq c_f|a|^2 + |b|^2/c_f$ for all a,b, one finds

$$\leq -c_f|\xi_t - \eta_t|^2 + \frac{K_\sigma^2}{c_f}|\xi_t - \eta_t|^2|u_t|^2.$$

Solving this differential inequality yields

$$|\xi_t - \eta_t|^2 \le |x-y|^2 e^{-c_f t + \frac{K_\sigma^2}{c_f}\int_0^t |u_r|^2\,dr},$$

or by (3.74),

$$|\xi_t - \eta_t| \le |x-y| \exp\left\{ \frac{K_\sigma^2}{\widetilde{\delta} c_f}\left[\varepsilon + \frac{\alpha_l R^2}{c_f}\right]\right\} e^{\frac{-c_f t}{2}}. \tag{3.75}$$

One also has (using (A3.1I) and (A3.2I))

$$\frac{d}{dt}|\xi_t|^2 \le -c_f|\xi_i|^2 + \frac{m_\sigma^2}{c_f}|u_t|^2.$$

By similar estimates as used above, one finds

$$|\xi_t|^2 \le \frac{2\varepsilon m_\sigma^2}{\widetilde{\delta} c_f} + \left[1 + \frac{2\alpha_l m_\sigma^2}{\widetilde{\delta} c_f^2}\right] R^2, \tag{3.76}$$

and similarly for η. Then, by (3.75), (3.76) and (A3.3I)

$$\begin{aligned} \left|\int_0^T l(\xi_t) - l(\eta_t)\,dt\right| &\le \int_0^T C_l(1 + |\xi_t| + |\eta_t|)|\xi_t - \eta_t|\,dt \\ &\le F(R)|x-y| \end{aligned}$$

for some locally bounded function F independent of $T \in [0,\infty)$. This implies

$$\begin{aligned} \widehat{V}(x,T) - \widehat{V}(y,T) &\le \widehat{J}(T,x,u) + \varepsilon - \widehat{J}(T,y,u) \\ &\le \varepsilon + F(R)|x-y|. \end{aligned}$$

Using the fact that ε was arbitrary and symmetry, we obtain

$$|\widehat{V}(x,T) - \widehat{V}(y,T)| \le F(R)|x-y| \quad \forall\, |x|,|y| \le R,\ \forall\, R < \infty, \tag{3.77}$$

and note that this bound is independent of $T \in (0,\infty)$.

By (3.77) and the Ascoli–Arzela Theorem, we see that the convergence in (3.73) is uniform on compact sets, and that $\widehat{W}$ is continuous.

Define

$$V^{[T]}(t,x) \doteq \widehat{V}(x, t+T) \quad \forall -T \le t \le 0,\ \forall x \in \mathbf{R}^n,$$

so that for all $T > 1$, $V^{[T]}$ is a viscosity solution of

$$0 = V_t - \left[\frac{1}{2\gamma^2}\nabla V^T \sigma(x)\sigma^T(x)\nabla V + g^T(x)\nabla V + l(x)\right], \quad (t,x) \in (-1,0)\times\mathbf{R}^n. \tag{3.78}$$

Then, by the uniform convergence, (3.78) and Lemma 2.6.2 of [48], $\widehat{W}$ is a continuous viscosity solution of (3.78) as well. But since $\widehat{W}$ is independent of t, we see that $\widehat{W}$ is a continuous viscosity solution of

$$0 = -\left[\frac{1}{2\gamma^2}\nabla W^T \sigma(x)\sigma^T(x)\nabla W + g^T(x)\nabla W + l(x)\right] \qquad x \in \mathbf{R}^n,$$

and by (3.72)

$$0 \leq \widehat{W}(x) \leq \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}|x|^2. \quad \square$$

Remark 3.21. Note that by (3.73), we have also shown that the solution, W, is given by

$$W(x) = \lim_{T\to\infty} \widehat{V}(x,T) = \sup_{T>0} \widehat{V}(x,T) \quad \forall\, x \in \mathbf{R}^n. \tag{3.79}$$

This completes our discussion of viscosity solution theory. We have considered two example problem classes — a finite time-horizon problem class and an infinite time-horizon problem class. We have demonstrated equivalence between the control problem value functions and continuous viscosity solutions of their corresponding HJB PDEs. In the following sections, we will develop max-plus based numerical methods for solving these equivalent problems.

4

Max-Plus Eigenvector Method for the Infinite Time-Horizon Problem

We now (finally) begin development of max-plus based methods for solution of HJB PDE/nonlinear control problems. We will work first with an infinite time-horizon control problem (3.12), (3.13), (3.21), (3.22). This, of course, corresponds to the steady-state HJB PDE (3.46), (3.47). The infinite time-horizon problem has some facets that are more complex than the finite time-horizon problem. This is primarily due to the fact that the infinite time-horizon problem essentially corresponds to the limit of the finite time-horizon problem as the time-horizon goes to infinity. As we indicated in Chapter 3, there are nonuniqueness issues here as well. Also, for the infinite time-horizon problem, a fuller analysis of error sources and convergence rate for this first max-plus method exists than in the finite time-horizon case.

Several max-plus based methods for infinite time-horizon problems will be discussed in this book. In this chapter and the next, we will focus on a method where the solution is found as a max-plus basis expansion over a space of semiconvevx functions. The DPP, (3.24), for this problem takes the form $W = S_\tau[W]$ where S_τ (given by (4.14)) is max-plus linear. It will be demonstrated that W is an element of a (max-plus vector) space of semiconvex functions. Expanding W in terms of a max-plus basis over this space leads to a max-plus eigenvector problem. More exactly, the vector of coefficients in this max-plus expansion is the solution of a max-plus eigenvector problem with max-plus eigenvalue zero. Other max-plus numerical methods and a somewhat different min-plus problem will also be discussed later.

It makes a good deal of sense to break the development of this first numerical method into two parts. In the first part, we will blindly truncate the max-plus expansion of the value function (as an element of a space of semiconvex functions; see Chapter 2). This will allow some of the main concepts to come through. However, a successful application of the method will be greatly enabled by a full understanding of the error/convergence analysis (for the method), which is delayed to Chapter 5.

Before immersing ourselves in the development, we place this work in some context. As noted above, the approach relies on the max-plus linearity of the associated semigroup/DPP. To the author's knowledge, the first mention of the max-plus linearity appears in [71], and one might also note the discussion in [63]. The second key enabling concept is the notion of max-plus bases for spaces of semiconvex functions (as discussed in Chapter 2). This was first developed in [44]. The numerical method for the infinite time-horizon problem was developed in the series of papers [75], [78], [79], [80], [81], [85], [86]. Closely related papers are [50] and [76]. A somewhat different, but related approach is being developed in [3], [4] (currently for the finite time-horizon problem class).

4.1 Existence and Uniqueness

In this section, we recall the system and assumptions. We also provide some basic results that will be needed below.

In order to have our materials handy, let's recall the infinite time-horizon problem we will consider here. The dynamics and initial condition are

$$\dot{\xi} = f(\xi) + \sigma(\xi)u \tag{4.1}$$

with initial condition

$$\xi_0 = x \in \mathbf{R}^n \tag{4.2}$$

The payoff and value function (available storage) are

$$J(x,T,u) = \int_0^T l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr, \tag{4.3}$$

and

$$W(x) = \sup_{u\in\mathcal{U}^U}\sup_{T<\infty} J(x,T,u) \tag{4.4}$$

where the control u lies in

$$\mathcal{U}^U = L_2^{\mathrm{loc}}([0,\infty);U) \doteq \left\{ u:[0,\infty)\to U\subseteq\mathbf{R}^l \;\middle|\; \int_{[0,T)} |u_r|^2\,dr < \infty\;\forall T\in[0,\infty)\right\}. \tag{4.5}$$

Note that σ is an $n\times l$ matrix-valued multiplier on the control.

We will make the following assumptions which are similar (although not identical) to the assumptions for the infinite time-horizon problem in Chapter 3. These assumptions are not necessary but are sufficient for the results to follow. No attempt has been made at this point to formulate tight assumptions. In particular, in order to provide some clear sketches of proofs, we will assume that all the functions f, σ and l are smooth, although that is not required for the results. We assume there $K, c_f \in (0,\infty)$ such that

$$
\begin{aligned}
&|f_x(x)| \le K \quad \forall\, x \in \mathbf{R}^n, \\
&(x-y)^T[f(x)-f(y)] \le -c_f|x-y|^2 \quad \forall\, x,y \in \mathbf{R}^n, \\
&f(0)=0.
\end{aligned} \tag{A4.1I}
$$

We note that this implies

$$x^T f(x) \le -c_f|x|^2$$

for all x, which implies exponential stability of the system when $u \equiv 0$. We assume there exists $K_\sigma, m_\sigma \in (0,\infty)$ such that

$$
\begin{aligned}
&\mathrm{Range}(\sigma(x)) = \mathbf{R}^n \quad \forall\, x \in \mathbf{R}^n, \\
&|\sigma_x(x)| \le K_\sigma \quad \forall\, x \in \mathbf{R}^n, \\
&|\sigma(x)| \le m_\sigma \quad \forall\, x \in \mathbf{R}^n, \\
&|\sigma^{-1}(x)| \le m_\sigma \quad \forall\, x \in \mathbf{R}^n.
\end{aligned} \tag{A4.2I}
$$

Here, we of course use σ^{-1} to indicate the Moore–Penrose inverse (c.f. [54]), and it is implicit in the bound on $\sigma^{-1}(x)$ that σ is uniformly nondegenerate (i.e., there exists $\eta > 0$ such that $z^T\sigma(x)\sigma^T(x)z \ge \eta|z|^2$ for all $x, z \in \mathbf{R}^n$). We assume that there exist $C_l, \alpha_l \in (\infty)$ such that

$$
\begin{aligned}
&|l_{xx}(x)| \le C_l \quad \forall\, x \in \mathbf{R}^n, \\
&0 \le l(x) \le \alpha_l|x|^2 \quad \forall\, x \in \mathbf{R}^n.
\end{aligned} \tag{A4.3I}
$$

The system is said to satisfy an H_∞ attenuation bound (of γ) if there exists $\gamma < \infty$ and a locally bounded available storage function (i.e., the value function), $W(x)$, which is nonegative, zero at the origin, and such that

$$W(x) = \sup_{u\in\mathcal{U}^U} \sup_{T<\infty} \int_0^T l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \tag{4.6}$$

where ξ satisfies (4.1),(4.2). We reiterate that the corresponding HJB PDE is

$$0 = H(x, \nabla W(x)) \quad \forall\, x \in \mathbf{R}^n \setminus \{0\}, \tag{4.7}$$

$$W(0) = 0, \tag{4.8}$$

where the Hamiltonian, H, is given by

$$H(x,p) \doteq -\max_{v\in U}\left\{[f(x)+\sigma(x)v]^T p + l(x) - \frac{\gamma^2}{2}|v|^2\right\}, \tag{4.9}$$

which, in the case where $U \equiv \mathbf{R}^l$, takes the form

$$H(x,p) = -\left[f^T(x)p + l(x) + \frac{1}{2\gamma^2}p^T\sigma(x)\sigma^T(x)p\right]. \tag{4.10}$$

Because W itself (not its gradient) does not appear in (4.7), one can always scale by an additive constant. It will be assumed throughout that we are looking for a solution satisfying boundary condition (4.8) which fixes the additive constant. Lastly, we assume that

$$\frac{\gamma^2 c_f^2}{2m_\sigma^2 \alpha_l} > 1. \tag{A4.4I}$$

Then one has the following result (see Chapter 3).

Theorem 4.1. *There exists a unique continuous viscosity solution of (4.7), (4.8), locally Lipschitz in x, in any class*

$$0 \le \widetilde{W}(x) \le c_f \frac{\gamma^2 - \delta}{2m_\sigma^2}|x|^2 \tag{4.11}$$

such that $\delta > 0$ sufficiently small that (A4.4I) holds with γ^2 replaced by $\gamma^2 - \delta$.

We also have the following. Consider the finite time-horizon control problem with dynamics (4.1), initial condition (4.2), and payoff and value given by

$$\begin{aligned} J(x,T,u) &= \int_0^T l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt, \\ \widehat{V}(x,T) &= \sup\nolimits_{u\in\mathcal{U}^U} J(x,T,u). \end{aligned} \tag{4.12}$$

Theorem 4.2. *The unique continuous viscosity solution in the class such that (4.11) holds for some $\delta > 0$ sufficiently small that (A4.4I) holds with γ^2 replaced by $\gamma^2 - \delta$, is given by*

$$W(x) = \lim_{T\to\infty} \widehat{V}(x,T) = \sup_{T<\infty} \widehat{V}(x,T) \tag{4.13}$$

which is also equivalent to representation (4.6).

This was proved in Theorem 3.20 and Remark 3.21.

4.2 Max-Plus Linearity of the Semigroup

We will show that W is a fixed point of the corresponding semigroup operator. The semigroup is defined directly by the DPP for the finite time-horizon problem. We begin with the following lemma which is a statement of the DPP for this particular problem.

Lemma 4.3. *Let $\widehat{V}$ be given by (4.12). Then,*

$$\widehat{V}(x,T) = \sup_{u\in\mathcal{U}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + \widehat{V}(\xi_\tau, T-\tau) \right\}$$

(where ξ satisfies (4.1), (4.2)) for any $0 \le \tau \le T < \infty$ and any $x \in \mathbf{R}^n$.

Proof. Note that for any $s \in (0,T)$,

$$\widehat{V}(x,T-s) = \sup_{u\in\mathcal{U}^U} \left[\int_0^{T-s} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\, dr \right],$$

where ξ satisfies (4.1), (4.2). Since neither the dynamics nor the payoff is time-dependent, one can shift the time variable to obtain

$$\widehat{V}(x,T-s) = \sup_{u\in\mathcal{U}^U} \left[\int_s^{T} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\, dr \right],$$

where ξ satisfies (4.1), $\xi_s = x$, which recalling (3.15)

$$= V(s,x).$$

Consequently, by Theorem 3.1 for any $0 \le \tau \le T < \infty$

$$\begin{aligned}\widehat{V}(x,T) = V(0,x) &= \sup_{u\in\mathcal{U}^U} \left[\int_0^{\tau} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\, dr + V(\tau,\xi_\tau) \right] \\ &= \sup_{u\in\mathcal{U}^U} \left[\int_0^{\tau} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\, dr + \widehat{V}(\xi_\tau, T-\tau) \right]. \quad \square\end{aligned}$$

Define the semigroup $S_\tau[W]$ for W in the domain of S_τ, $\mathrm{Dom}[S_\tau] \doteq \{W \in \mathcal{S} :\ S_\tau[W(\cdot)](x) < \infty\ \forall x \in \mathbf{R}^n\}$, by

$$S_\tau[W(\cdot)](x) = \sup_{u\in\mathcal{U}^U} \left\{ \int_0^{\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\, dt + W(\xi_\tau) \right\}, \tag{4.14}$$

where ξ satisfies (4.1), (4.2). Note that the semigroup properties (1.8) and (1.9) are easily verified. Also note that $\mathrm{Dom}[S_\tau]$ includes all semiconvex functions satisfying (4.11).

Theorem 4.4. *For any $\tau \in [0,\infty)$, W given by (4.13) satisfies $S_\tau[W] = W$, and further, it is the unique solution in the class (4.11).*

Proof. We begin with the first assertion. From (4.13), one has

$$W(x) = \sup_{T<\infty} \widehat{V}(x,T),$$

which by Lemma 4.3

$$= \sup_{T<\infty} \sup_{u\in\mathcal{U}^U} \left\{ \int_0^{\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\, dt + \widehat{V}(\xi_\tau, T-\tau) \right\},$$

where ξ satisfies (4.1), (4.2)

$$= \sup_{u\in\mathcal{U}^U} \sup_{T<\infty} \left\{ \int_0^{\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\, dt + \widehat{V}(\xi_\tau, T-\tau) \right\},$$

which by (4.13) again

$$= \sup_{u \in \mathcal{U}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\, dt + W(\xi_\tau) \right\}$$
$$= S_\tau[W(\cdot)](x).$$

The proof of uniqueness is similar to the proof used to demonstrate uniqueness of the viscosity solution and value within the class (4.11) in Chapter 3. Let $\overline{W}$ satisfy $\overline{W} = S_\tau[\overline{W}]$. One shows that $\overline{W}(x) \le W(x)$ and $\overline{W}(x) \ge W(x)$ for all $x \in \mathbf{R}^n$. The first of these two inequalities is more technical, and we address that first.

Let m be any positive integer (which we eventually let go to ∞). Note that by the semigroup property,

$$\begin{aligned}\overline{W}(x) = S_\tau^m[\overline{W}](x) &= S_{m\tau}[\overline{W}](x) \\ &= \sup_{u \in \mathcal{U}^U} \left\{ \int_0^{m\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\, dt + \overline{W}(\xi_{m\tau}) \right\} \\ &\doteq \sup_{u \in \mathcal{U}^U} \overline{J}(x, m\tau, u).\end{aligned}$$

Let $\delta > 0$ be such that with $\widehat{\gamma}^2 \doteq \gamma^2 - \delta$, one has $\frac{\widehat{\gamma}^2 c_f^2}{2m_\sigma^2 \alpha_l} > 1$ and $0 \le \overline{W}(x) \le \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}|x|^2$. Fix any $x \in \mathbf{R}^n$. Let

$$Q_T \doteq \int_0^T \alpha_l |\xi_t|^2\, dt + \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}|\xi_T|^2 \tag{4.15}$$

where ξ satisfies (4.1), (4.2), and note that by (A4.3I) and (3.23)

$$\ge \int_0^T l(\xi_t)\, dt + W(\xi_T). \tag{4.16}$$

Note that this definition of Q_T is identical to (3.27), and following the same steps as found there (in the proof of Theorem 3.10. Specifically, letting $\widetilde{\delta} \doteq \frac{\widehat{\gamma}^2}{2} - \frac{\alpha_l m_\sigma^2}{c_f^2}$ and noting that by (A4.4I) $\widetilde{\delta} > 0$, one has

$$Q_T \le \left[\frac{c_f \widehat{\gamma}^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 - \widetilde{\delta} \int_0^T (1 - e^{c_f(t-T)})|u|^2\, dt + \tfrac{\widehat{\gamma}^2}{2}\|u\|_{L_2[0,T]}. \tag{4.17}$$

Consequently, by (4.15), (4.16) and (4.17)

$$\begin{aligned}\overline{J}(x,T,u) &\le \left[\frac{c_f \gamma^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 - \widetilde{\delta} \int_0^T (1 - e^{c_f(t-T)})|u|^2\, dt - \delta \|u\|^2_{L_2[0,T]} \\ &\le \left[\frac{c_f \gamma^2}{2m_\sigma^2} e^{-c_f T} + \frac{\alpha_l}{c_f} \right] |x|^2 - \delta \|u\|^2_{L_2[0,T]}.\end{aligned} \tag{4.18}$$

On the other hand, by (A4.3I) and (4.11), $\overline{J}(x, T, 0) \ge 0$, and so for ε-optimal u (where, as usual, we take $\varepsilon \in (0, 1]$)

$$\overline{J}(x,T,u) \geq -\varepsilon. \tag{4.19}$$

Comparing (4.18) and(4.19), and letting $T = m\tau$, we see that for ε-optimal u, we have

$$\begin{aligned}\|u\|^2_{L_2[0,m\tau]} &\leq \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{-c_f m\tau} + \frac{\alpha_l}{c_f}\right]|x|^2 \\ &\leq \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2} + \frac{\alpha_l}{c_f}\right]|x|^2. \end{aligned} \tag{4.20}$$

This is an upper bound on the size of ε-optimal u which is independent of m (for large m).

By a similar analysis as for Q_T, one easily shows that by (A4.1I) and (A4.2I),

$$|\xi_{(k+1)\tau}|^2 \leq e^{-c_f\tau}|\xi_{k\tau}|^2 + \frac{m_\sigma^2}{c_f}\|u\|^2_{L_2(k\tau,(k+1)\tau)}$$

for any $u \in \mathcal{U}^U$ and any nonegative integer $k < m$. Repeating this estimate, and combining terms, one finds that for any non-negative integers j, I such that $I + j \leq m$ and any $u \in \mathcal{U}^U$

$$\begin{aligned}|\xi_{(I+j)\tau}|^2 \leq{}& e^{-Ic_f\tau}e^{-jc_f\tau}|x|^2 + \frac{m_\sigma^2}{c_f}\Bigg\{e^{-jc_f\tau}\left[\sum_{i=1}^{I} e^{-ic_f\tau}\|u\|^2_{L_2((I-i)\tau,(I+1-i)\tau)}\right] \\ &+ \sum_{i=0}^{j-1} e^{-(j-1-i)c_f\tau}\|u\|^2_{L_2((I+i)\tau,(I+1+i)\tau)}\Bigg\} \\ \leq{}& e^{-Ic_f\tau}e^{-jc_f\tau}|x|^2 + \frac{m_\sigma^2}{c_f}\|u\|^2_{L_2(0,(I+j)\tau)}.\end{aligned}$$

For simplicity, we consider only m which are even. Adding subsequent estimates and taking $I = m/2$, one easily finds

$$\begin{aligned}\sum_{j=m/2}^{m} |\xi_{j\tau}|^2 &\leq e^{-mc_f\tau/2}\left(\sum_{j=0}^{m/2} e^{-jc_f\tau}\right)|x|^2 + \frac{m_\sigma^2}{c_f}\|u\|^2_{L_2(0,m\tau)} \\ &\leq e^{-mc_f\tau/2}\left(\frac{1}{1-e^{-c_f\tau}}\right)|x|^2 + \frac{m_\sigma^2}{c_f}\|u\|^2_{L_2(0,m\tau)}.\end{aligned}$$

Let u be ε-optimal (over $(0, m\tau)$). Then, substituting (4.20) into (4.21) yields

$$\sum_{j=m/2}^{m} |\xi_{j\tau}|^2 \leq e^{-mc_f\tau/2}\left(\frac{1}{1-e^{-c_f\tau}}\right)|x|^2 + \frac{m_\sigma^2}{c_f}\left\{\frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2} + \frac{\alpha_l}{c_f}\right]|x|^2\right\}$$

which taking $m \geq 2$

$$\leq \left(\frac{e^{-c_f \tau}}{1 - e^{-c_f \tau}} \right) |x|^2 + \frac{m_\sigma^2}{c_f} \left\{ \frac{\varepsilon}{\delta} + \frac{1}{\delta} \left[\frac{c_f \gamma^2}{2m_\sigma^2} + \frac{\alpha_l}{c_f} \right] |x|^2 \right\},$$

which for proper choice of K_1 and K_2 independent of m

$$\leq K_1 + K_2 |x|^2. \tag{4.21}$$

Fix some $\hat{\varepsilon} > 0$, and let $m > \frac{2}{\hat{\varepsilon}}(K_1 + K_2|x|^2)$. Suppose $|\xi_{j\tau}|^2 > m\hat{\varepsilon}/2$ for all $j \geq m/2$. Then by this choice of m,

$$K_1 + K_2|x|^2 < \frac{m\hat{\varepsilon}}{2} < \sum_{j=m/2}^{m} |\xi_{j\tau}|^2,$$

which by (4.21)

$$\leq K_1 + K_2|x|^2$$

which is a contradiction. Therefore, given any $\hat{\varepsilon} > 0$, there exists an $m < \infty$ and a corresponding ε-optimal u such that there exists integer $\hat{k} \in [m/2, m]$ such that

$$|\xi_{\hat{k}\tau}|^2 \leq \hat{\varepsilon}. \tag{4.22}$$

Suppose we have chosen such m and (ε-optimal) u, and let $\hat{k}$ be the above integer. By Lemma 3.18, this u is also optimal over for the problem over time interval $(0, \hat{k}\tau)$. Consequently,

$$\begin{aligned} \overline{W}(x) = S_\tau^{\hat{k}}[\overline{W}](x) &\leq \overline{J}(x, \hat{k}\tau, u) + \varepsilon \\ &= \int_0^{\hat{k}\tau} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr + \overline{W}(\xi_{\hat{k}\tau}) + \varepsilon. \end{aligned}$$

Using (4.11) and the choice of $\widehat{\gamma}$, this is

$$\begin{aligned} &\leq \int_0^{\hat{k}\tau} l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr + \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}\hat{\varepsilon} + \varepsilon \\ &\leq \widehat{V}(x, \hat{k}\tau) + \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}\hat{\varepsilon} + \varepsilon, \end{aligned}$$

which by (4.13)

$$\leq W(x) + \frac{c_f \widehat{\gamma}^2}{2m_\sigma^2}\hat{\varepsilon} + \varepsilon.$$

Because this is true for all $\varepsilon > 0$ and $\hat{\varepsilon} \in (0, 1]$, we have $\overline{W}(x) \leq W(x)$ which is one of the desired inequalities.

For the reverse inequality, note that for an arbitrarily large integer m,

$$\widehat{V}(x, m\tau) = \sup_{u \in \mathcal{U}^U} \left\{ \int_0^{m\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt \right\}$$

which since $\overline{W}$ is in the class specified at (4.11)

$$\begin{aligned} &\leq \sup_{u \in \mathcal{U}^U} \left\{ \int_0^{m\tau} l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + \overline{W}(\xi_\tau) \right\} \\ &= S_{m\tau}[\overline{W}](x) = S_\tau^m[\overline{W}](x) = \overline{W}(x). \end{aligned} \tag{4.23}$$

On the other hand, by Theorem 4.2,

$$\widehat{V}(x, T) \to W(x). \tag{4.24}$$

Then (4.23) and (4.24) imply $\overline{W}(x) \geq W(x)$ (otherwise one obtains a contradiction). Thus, one has the two inequalities which imply that any solution of $\overline{W} = S_\tau[\overline{W}]$ in the class given in (4.11) must be W. □

Note that W is a fixed point of S_τ for any τ, which provides some freedom in the choice of problem we wish to solve. This may be of interest in the actual construction of numerical algorithms.

Now we will demonstrate that S_τ is linear in the max-plus algebra. Let $\psi, \phi \in \mathrm{Dom}[S_\tau]$. For $a \in \mathbf{R}$, the proof that max-plus multiplication passes through the operator is trivial:

$$\begin{aligned} S_\tau[a \otimes \psi](x) &= S_\tau[a + \psi](x) \\ &= \sup_{u \in \mathcal{U}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + a + \psi(\xi_\tau) \right\} \\ &= a + S_\tau[\psi](x) = a \otimes S_\tau[\psi](x) \end{aligned} \tag{4.25}$$

for all $x \in \mathbf{R}^n$. In the case that $a = -\infty$, one has

$$\begin{aligned} S_\tau[-\infty \otimes \psi](x) &= \sup_{u \in \mathcal{U}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + \psi(\xi_\tau) - \infty \right\} \\ &= \sup_{u \in \mathcal{U}^U} \{-\infty\} = -\infty = -\infty \otimes S_\tau[\psi](x) \end{aligned}$$

for all $x \in \mathbf{R}^n$. Now considering max-plus addition, we note that

$$\begin{aligned} S_\tau[\phi \oplus \psi](x) &= \sup_{u \in \mathcal{U}^U} \left\{ \max[\phi(\xi_\tau), \psi(\xi_\tau)] + \int_0^\tau l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr \right\} \\ &= \sup_{u \in \mathcal{U}^U} \max\left\{ \phi(\xi_\tau) + \int_0^\tau l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr, \right. \\ &\qquad\qquad \left. \psi(\xi_\tau) + \int_0^\tau l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr \right\} \\ &= \max\left\{ \sup_{u \in \mathcal{U}^U} \left[\phi(\xi_\tau) + \int_0^\tau l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2 \, dr \right], \right. \end{aligned}$$

$$\begin{aligned}
&\sup_{u\in\mathcal{U}^U}\left[\psi(\xi_\tau)] + \int_0^\tau l(\xi_r) - \frac{\gamma^2}{2}|u_r|^2\,dr\right]\Bigg\} \\
&= \max\{S_\tau[\phi](x), S_\tau[\psi](x)\} \\
&= S_\tau[\phi](x) \oplus S_\tau[\psi](x) \\
&= \{S_\tau[\phi] \oplus S_\tau[\psi]\}(x). \qquad (4.26)
\end{aligned}$$

By (4.25) and (4.26), one has

Theorem 4.5. *The semigroup, S_τ, is linear in the max-plus algebra.*

Remark 4.6. Using the techniques of Chapter 3, it is not difficult to show that $S_\tau[\psi]$ is well-defined for all continuous ψ belonging to the class given by (4.11) for some $\delta > 0$ satisfying the condition there. Consequently one may take the domain of S_τ to be this set. One can further show that S_τ maps this domain into itself (a property without which the semigroup property (1.9) could not hold). The domain can be reduced to the subset of locally Lipschitz functions in this class, and in that case, one can show that S_τ again maps this domain into itself.

4.3 Semiconvexity and a Max-Plus Basis

In this section, we first show that the value, W, lies in the space of semiconvex functions, and then we describe the max-plus basis that we will use in the following sections. The following result is equivalent to Lemma 3.16, but introduces some new notation.

Lemma 4.7. *Given $\varepsilon \in (0,1)$, $R < \infty$ and $|x| < R$, let $u^{\varepsilon,T}$ be ε-optimal for problem (4.12). Then, there exists $M_R < \infty$ (where this is independent of ε, $|x| \le R$ and $T \in (0,\infty)$) such that*

$$\|u^{\varepsilon,T}\|_{L_2(0,T)} \le M_R.$$

Lemma 4.8. *Given $\varepsilon \in (0,1)$, $R < \infty$ and $|x| < R$, let $u^{\varepsilon,T}$ be ε-optimal for problem (4.12). Then, there exists $d_R < \infty$ (where this is independent of ε, $|x| \le R$ and $T \in (0,\infty)$) such that*

$$|\xi_t| \le d_R \quad \forall\, t \in [0,T].$$

Proof. The proof follows directly from Lemma 4.7 and the assumptions, particularly the contractivity of f. Specifically, let ξ satisfy (4.1) with initial $|\xi_0| = |x| \le R$ and ε-optimal disturbance, $u^{\varepsilon,T}$. Then,

$$\frac{d}{dt}\left[|\xi_t|^2\right] = 2\xi^T f(\xi) + 2\xi^T \sigma(\xi) u^{\varepsilon,T}$$

which by Assumptions (A4.1I) and (A4.2I)

$$\leq -2c_f|\xi|^2 + 2m_\sigma|\xi||u^{\varepsilon,T}|$$
$$\leq -c_f|\xi|^2 + \frac{m_\sigma^2}{c_f}|u^{\varepsilon,T}|^2.$$

Integrating, one finds

$$|\xi_t|^2 \leq |x|^2 e^{-c_f t} + \frac{m_\sigma^2}{c_f}\int_0^t e^{c_f(r-t)}|u_r^{\varepsilon,T}|^2\,dr$$

which by Lemma 4.7

$$\leq |x|^2 + \frac{m_\sigma^2}{c_f}M_R^2$$

which yields the result with $d_R \doteq \left[R^2 + (m_\sigma^2/c_f)M_R^2\right]^{\frac{1}{2}}$. □

The next result will be at the core of the methods, since we will be working in spaces of semiconvex functions.

Theorem 4.9. *$\widehat{V}(x,T)$ is semiconvex (in x) with constants independent of $T \in [0,\infty)$.*

Proof. It is sufficient to show that the second differences of $\widehat{V}(\cdot,T)$ are bounded from below on any $B_R(0)$ by some $-c_R$. Let $x \in B_R(0)$, $v \in \mathbf{R}^n$ and $|v| = 1$.

Let $\delta,\varepsilon \in (0,1)$. Let u be ε-optimal for (4.12) with corresponding $\xi^0_\cdot$ satisfying (4.1) with $\xi_0^0 = x$. Then

$$\widehat{V}(x-\delta v,T) - 2\widehat{V}(x,T) + \widehat{V}(x+\delta v,T) \geq \widehat{J}(T,x-\delta v,u+\Delta) - 2\widehat{J}(T,x,u) + \widehat{J}(T,x+\delta v,u) - 2\varepsilon \qquad (4.27)$$

where Δ will be given below. Let $\xi^{-\delta}$, ξ^δ satisfy the dynamics of (4.1) with initial conditions $\xi_0^{-\delta} = x - \delta v$ and $\xi_0^\delta = x + \delta v$. Also let the corresponding disturbance processes be $u + \Delta$ and u, respectively. Then,

$$\dot{\xi}^\delta - \dot{\xi}^0 = f(\xi^\delta) - f(\xi^0) + (\sigma(\xi^\delta) - \sigma(\xi^0))u,$$

and

$$\dot{\xi}^0 - \dot{\xi}^{-\delta} = f(\xi^0) - f(\xi^{-\delta}) + \sigma(\xi^0)u - \sigma(\xi^{-\delta})(u+\Delta).$$

Now we choose

$$\begin{aligned}\Delta &\doteq -\sigma^{-1}(\xi^{-\delta})\{f(\xi^\delta) - f(\xi^0) + (\sigma(\xi^\delta) - \sigma(\xi^0))u - (f(\xi^0) - f(\xi^{-\delta}))\\ &\quad -(\sigma(\xi^0) - \sigma(\xi^{-\delta}))u\}\\ &= -\sigma^{-1}(\xi^{-\delta})\{f(\xi^\delta) - 2f(\xi^0) + f(\xi^{-\delta})\\ &\quad +(\sigma(\xi^\delta) - 2\sigma(\xi^0) + \sigma(\xi^{-\delta}))u\}\end{aligned} \qquad (4.28)$$

(Although Δ is defined by the above feedback formula, the corresponding Δ as a function of t is used.) By substitution, we see

$$\dot{\xi}^\delta - \dot{\xi}^0 = \dot{\xi}^0 - \dot{\xi}^{-\delta},$$

and since $\xi_0^\delta - \xi_0^0 = \delta v$ and $\xi_0^0 - \xi_0^{-\delta} = \delta v$,

$$\xi_t^\delta - \xi_t^0 = \xi_t^0 - \xi_t^{-\delta} \quad \forall t \geq 0. \tag{4.29}$$

Consequently, by (4.27) and (4.29)

$$\begin{aligned}
&\widehat{V}(x-\delta v, T) - 2\widehat{V}(x,T) + \widehat{V}(x+\delta v, T)\\
&\geq \int_0^T l(\xi^{-\delta}) - 2l(\xi^0) + l(\xi^\delta)\,dt - \frac{\gamma^2}{2}\int_0^T |u+\Delta|^2 - 2|u|^2 + |u|^2\,dt - 2\varepsilon\\
&= \int_0^T l\big(\xi_t^0 - [\xi_t^\delta - \xi_t^0]\big) - 2l(\xi^0) + l\big(\xi_t^0 + [\xi_t^\delta - \xi_t^0]\big)\,dt\\
&\quad -\frac{\gamma^2}{2}\int_0^T 2u^T\Delta + |\Delta|^2\,dt - 2\varepsilon.
\end{aligned}$$

By Assumption (A4.3I), the second differences in l are bounded, and in fact, one has

$$\begin{aligned}
&\widehat{V}(x-\delta v, T) - 2\widehat{V}(x,T) + \widehat{V}(x+\delta v, T)\\
&\geq -\int_0^T 2C_l|\xi_t^\delta - \xi_t^0|^2\,dt - \frac{\gamma^2}{2}\int_0^T 2u^T\Delta + |\Delta|^2\,dt - 2\varepsilon.
\end{aligned} \tag{4.30}$$

Now,

$$\begin{aligned}
\frac{d}{dt}|\xi_t^\delta - \xi_t^0|^2 &= 2[\xi_t^\delta - \xi_t^0]^T[\dot{\xi}_t^\delta - \dot{\xi}_t^0]\\
&= 2[\xi_t^\delta - \xi_t^0]^T[f(\xi_t^\delta) - f(\xi_t^0) + (\sigma(\xi_t^\delta) - \sigma(\xi_t^0))u_t],
\end{aligned}$$

which by (A4.1I), (A4.2I) (and the general inequality $2ab \leq c_f a^2 + b^2/c_f$)

$$\begin{aligned}
&\leq -2c_f|\xi_t^\delta - \xi_t^0|^2 + 2K_\sigma|\xi_t^\delta - \xi_t^0|^2|u_t|\\
&\leq |\xi_t^\delta - \xi_t^0|^2\left[-c_f + \frac{K_\sigma^2}{c_f}|u_t|^2\right].
\end{aligned}$$

Using separation of variables (to solve the ordinary differential inequality) and Lemma 4.7, one obtains

$$\begin{aligned}
|\xi_t^\delta - \xi_t^0|^2 &\leq |(x+\delta v) - x|^2 e^{-c_f t + \frac{K_\sigma^2}{c_f}\int_0^t |u|^2\,dt}\\
&\leq \delta^2 e^{-c_f t} e^{\frac{K_\sigma^2}{c_f} {M_R}^2}.
\end{aligned}$$

Hence there exists $\overline{c}$ (independent of T) such that

$$|\xi_t^\delta - \xi_t^0|^2 \le \delta^2 \overline{c} e^{-c_f t}. \tag{4.31}$$

Substituting (4.31) into (4.30), we get

$$\widehat{V}(x-\delta v,T) - 2\widehat{V}(x,T) + \widehat{V}(x+\delta v,T) \ge -\int_0^T 2C_l \delta^2 \overline{c} e^{-c_f t}\,dt - \frac{\gamma^2}{2}\int_0^T 2u^T \Delta + |\Delta|^2\,dt - 2\varepsilon. \tag{4.32}$$

Now, by Lemma 4.8, $|\xi_t^0| \le d_R$ for all $t \ge 0$, and so by the smoothness of f, σ, there exist $Q_{f,R}, Q_{\sigma,R} < \infty$ such that $|f_{xx}(x)| \le Q_{f,R}$, $|\sigma_{xx}(x)| \le Q_{\sigma,R}$ on $|x| \le d_R + \delta$. Thus, by (4.28), (4.29) and (A4.3I),

$$|\Delta_t| \le m_\sigma \left\{ Q_{f,R} |\xi_t^\delta - \xi_t^0|^2 + Q_{\sigma,R} |\xi_t^\delta - \xi_t^0|^2 |u_t| \right\} \quad \forall t \in [0,T],$$

where the bound is independent of $T < \infty$. Employing (4.31), this yields

$$|\Delta_t| \le m_\sigma \tilde{c}_R \delta^2 \left[e^{-c_f t}(1+|u_t|) \right] \quad \forall\, 0 \le t \le T < \infty \tag{4.33}$$

for appropriate choice of $\tilde{c}_R$ dependent on R.

Substituting (4.33) into (4.32) yields

$$\begin{aligned}
&\widehat{V}(x-\delta v,T) - 2\widehat{V}(x,T) + \widehat{V}(x+\delta v,T) \\
&\ge -\int_0^T 2C_l \delta^2 \overline{c} e^{-c_f t}\,dt - \frac{\gamma^2}{2}\int_0^T \Big[2m_\sigma \tilde{c}_R \delta^2 (|u_t| + |u_t|^2) e^{-c_f t} \\
&\quad + 2m_\sigma^2 \tilde{c}_R^2 \delta^4 (1+|u_t|^2) e^{-2c_f t} \Big]\,dt - 2\varepsilon \\
&\ge -\widehat{M}_R \delta^2 - 2\varepsilon \quad \forall\, 0 \le t \le T < \infty
\end{aligned}$$

for appropriate choice of $\widehat{M}_R$. Since $\varepsilon > 0$ was arbitrary, this implies semiconvexity. □

Remark 4.10. Note that if one supposes uniformly bounded second derivatives of f and σ, then the use of Lemma 4.8 (which requires the contractivity of f) can be avoided. However, one still needs bounded $\|u\|$, and the proof of Lemma 4.7 also made use of the contractivity of f. Consequently, although the addition of assumptions of uniformly bounded second derivatives would shorten the proof, it is not clear that it would be useful **unless** one needed to obtain a tighter bound on the growth of $\widehat{M}_R$ as a function of R.

Corollary 4.11. *$W(x)$ is semiconvex.*

We will informally refer to the restriction of $W(x)$ to any closed ball, $\overline{B}_R(0) = \{x \in \mathbf{R}^n \,|\, |x| \le R\}$, as $W(x)$ as well. Recall that the semiconvexity of W implies that given any ball $\overline{B}_R(0)$, W is Lipschitz with some constant, L_R, over $\overline{B}_R(0)$ (c.f. [42]). Recall also, from Chapter 2, that $\mathcal{S}_R^{cL}$ denotes the space of semiconvex functions with (semiconvexity) constant c and Lipschitz constant L over $\overline{B}_R(0)$. Consequently, Corollary 4.11 immediately implies the following.

Corollary 4.12. *Given any* $R \in (0,\infty)$, *there exist* $c_R, L_R \in (0,\infty)$ *such that* $W \in \mathcal{S}_{\bar{R}}^{c_R L_R}$.

Finally, it is useful here to recall Theorem 2.13 which we paraphrase as follows.

Theorem 4.13. *Let* C *be a symmetric matrix such that* $C - c_R I > 0$. *Let* $\psi_i(x) \doteq -\frac{1}{2}(x-x_i)^T C(x-x_i)$ *for all* $i \in \mathcal{N}$ *where the* x_i *form a countable dense subset of* $\mathcal{E} = \{\overline{x} \in \mathbf{R}^n : \overline{x}^T (C^2)\overline{x} \le (L_R + |C|R)^2\}$. *Then,* $\{\psi_i : i \in \mathcal{N}\}$ *is a countable basis for max-plus vector space* $\mathcal{S}_{\bar{R}}^{c_R L_R}$. *In particular, for any* $\phi \in \mathcal{S}_{\bar{R}}^{c_R L_R}$,

$$\phi(x) = \sup_{i \in \mathcal{N}} [a_i + \psi_i(x)] = \bigoplus_{i=1}^{\infty} [a_i \otimes \psi_i(x)], \tag{4.34}$$

where

$$a_i = -\max_{x \in \overline{B}_R} [-\phi(x) + \psi_i(x)] \quad \forall\, i. \tag{4.35}$$

4.4 The Eigenvector Equation

In order to reduce complexity, we suppose throughout the next two sections that W has a max-plus basis expansion with a finite number of terms. Of course, one must consider the error introduced by truncating the expansion at a finite number of terms in the numerical computations. However, in order to focus on the theory underlying the solution of the eigenvector equation (to follow), we delay the error analysis to the next chapter.

Let $W(x) = \bigoplus_{i=1}^{\nu} a_i \otimes \psi_i$, $a^T = (a_1, a_2, \ldots, a_\nu)$, and B be the $\nu \times \nu$ matrix with entries

$$B_{i,j} = -\max_{x \in B_R} (\psi_i(x) - S_\tau[\psi_j](x)). \tag{4.36}$$

Here we are letting the finite number of terms in the basis expansion be ν. Note that B actually depends on τ, but for this section we fix any value $\tau \le \tau_R$, and suppress the dependence in the notation, where $S_\tau[\psi_i](x)$ is C^2 on $[0,\tau) \times B_R(0)$ for all i. (See, for instance [37] for existence of τ_R.)

We assume we may choose C such that the semiconvexity of the basis functions is increasing for small time; this ensures that the elements of B are finite. (A4.5I)

Remark 4.14. In practice, some informal checks are used to search for such a C. Consider the linear/quadratic case where $f(x) = Ax$, $\sigma(x) = \sigma$ (constant), and $l(x) = \frac{1}{2}x^T Dx$, where the matrices are such that the above assumptions are satisfied. The corresponding Riccati equation is $\dot{R} = D + A^T R + RA + \frac{1}{2\gamma^2} R\sigma\sigma^T R$ with $R_0 = -C$, and condition (A4.5I) is equivalent to choosing C such that

$$D - A^T C - CA + \frac{1}{2\gamma^2} C\sigma\sigma^T C > 0. \tag{4.37}$$

The current practice is to linearize the system at numerous points, and search for a C such that (4.37) holds for all the linearized systems.

By the semiconvexity preserving property of S_τ for the given choice of C, one notes that

$$S_\tau[\psi_j](x) = \bigoplus_{i=1}^{\nu} B_{i,j} \otimes \psi_i(x), \tag{4.38}$$

where again we truncate the expansion, and postpone analysis of the errors introduced by truncation to Chapter 5.

We will use the notation $B \otimes a$ for max-plus multiplication of matrix B and vector a.

Theorem 4.15.

1. *Suppose W is a solution to $S_\tau[W] = W$, and that any expansion $W(x) = \bigoplus_{i=1}^{\nu} a_i \otimes \psi_i(x)$ on $\overline{B}_R(0)$ requires $a_i > -\infty$ for $\forall i$. Then $a = B \otimes a$.*
2. *Conversely, suppose $a = B \otimes a$ and that $W(x) = \bigoplus_{i=1}^{\nu} a_i \otimes \psi_i(x)$ on $\overline{B}_R(0)$. Then $S_\tau[W] = W$ on $\overline{B}_R(0)$.*

Proof. We begin with the first assertion. We have

$$\bigoplus_{i=1}^{\nu} a_i \otimes \psi_i(x) = W(x)$$

which by assumption

$$= S_\tau[W(\cdot)](x) = S_\tau \left[\bigoplus_{j=1}^{\nu} a_j \otimes \psi_j(\cdot) \right](x),$$

and then by Theorem 4.5

$$= \bigoplus_{j=1}^{\nu} a_j \otimes S_\tau[\psi_j](x),$$

which by (4.38)

$$= \bigoplus_{j=1}^{\nu} \left\{ a_j \otimes \left[\bigoplus_{i=1}^{\nu} B_{i,j} \otimes \psi_i(x) \right] \right\}$$

$$= \bigoplus_{j=1}^{\nu} \left\{ \bigoplus_{i=1}^{\nu} (a_j \otimes B_{i,j} \otimes \psi_i(x)) \right\}$$

$$= \bigoplus_{i=1}^{\nu} \left\{ \left[\bigoplus_{j=1}^{\nu} (B_{i,j} \otimes a_j) \right] \otimes \psi_i(x) \right\}.$$

But this implies that

$$\max\{a_1 + \psi_1(x), \ldots, a_n + \psi_n(x)\} = \max\{(\max_j(B_{1,j} + a_j) + \psi_1(x)), \ldots$$
$$\ldots, (\max_j(B_{\nu,j} + a_j) + \psi_\nu(x))\}. \quad (4.39)$$

Because it is necessary that any particular $a_{i_0} > -\infty$ (i.e., that this term is needed in the expansion), there exists $y_{i_0} \in \overline{B}_R(0)$ such that $a_{i_0} + \psi_{i_0}(y_{i_0}) > a_j + \psi_j(y_{i_0})$ for all $j \neq i_0$. Then, by continuity, there exists a neighborhood $B_\delta(z) \subset \overline{B}_R(0)$ with $y_{i_0} \in B_\delta(z)$ such that the left-hand side of (4.39) is simply $a_{i_0} + \psi_{i_0}(x)$ for all $x \in B_\delta(z)$. Consequently,

$$a_{i_0} + \psi_{i_0}(x) = \max_i\{\max_j[B_{i,j} + a_j] + \psi_i(x)\} \quad \forall\, x \in B_\delta(z). \quad (4.40)$$

However, the only way in which a maximum over a finite set of quadratics can be exactly identical to another quadratic over an open neighborhood is if one element of the set of quadratics is exactly the quadratic being matched. Consequently, (4.40) yields

$$a_{i_0} = \max_j[B_{i_0,j} + a_j].$$

Since this must hold for any choice of i_0, one has

$$a = B \otimes a,$$

where a is the vector $[a_i]$ and B is the matrix $[B_{i,j}]$.

For the second assertion of the theorem, note that if $a = B \otimes a$, then

$$\begin{aligned}
\bigoplus_{i=1}^{\nu} a_i \otimes \psi_i(x) &= \bigoplus_{i=1}^{\nu} \left\{ \bigoplus_{j=1}^{\nu} (B_{i,j} \otimes a_j) \right\} \otimes \psi_i(x) \\
&= \bigoplus_{j=1}^{\nu} \left\{ a_j \otimes \bigoplus_{i=1}^{\nu} (B_{i,j} \otimes \psi_i(x)) \right\} \\
&= \bigoplus_{j=1}^{\nu} a_j \otimes S_\tau[\psi_j](x) \\
&= S_\tau \left[\bigoplus_{j=1}^{\nu} a_j \otimes \psi_j \right] (x).
\end{aligned}$$

Now since $W(x) = \bigoplus_{j=1}^{\nu} a_j \otimes \psi_j(x)$, one sees that this last expression is $W(x) = S_\tau[W](x)$. □

4.5 The Power Method

The goal then is to solve the eigenvector equation

$$e = B \otimes e, \quad \text{or, equivalently,} \quad 0 \otimes e = B \otimes e. \tag{4.41}$$

There are two main steps. The first is to compute (approximately) B, and the second is to solve the eigenvector equation given B. We will address the latter step in this section; the former step will be discussed in the next section.

Note that because the eigenvalue is known, the solution of (4.41) reduces to the solution of a linear system. One obvious approach would be to solve this linear system directly. However, without the expansion of the max-plus algebra in some way so as to compensate for the lack of additive inverses (see [6]), one cannot proceed in that manner. Consequently, we instead have used the power method to obtain the eigenvector. Not only does this yield a numerical algorithm for the solution of (4.41), but interestingly, it has also led to a proof of the fact that in this special case, there is a unique eigenvector (corresponding to the unique eigenvalue of 0 — see [6] for the uniqueness of max-plus eigenvalues). Of course, this uniqueness is in marked contrast to the state of affairs over the usual field. To be more specific regarding the power method, note that it will be shown that the eigenvector may be computed via the iteration

$$e = \lim_{m\to\infty} B^m \otimes 0$$

where 0 represents the (usual) zero vector, and B^m represents the max-plus product of B with itself repeated m times. Interestingly, this will terminate (i.e., the iterates will stop changing) in a finite number of steps. This method is referred to as the power method.

Let W be the available storage (given by (4.13)). For the remainder of the section, fix any $\tau \in (0,\infty)$. Define

$$H(x,y) \doteq S_\tau[W](x) - \sup_{u\in\mathcal{U}_y^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\}, \tag{4.42}$$

where $\xi_0 = x$ and $\mathcal{U}_y^U = \{u \in \mathcal{U}^U : \xi_\tau = y\}$.

Lemma 4.16. *There exist $M_{|x|}, M_{|x|,|y|} < \infty$, monotonically increasing as functions of the subscripts, such that*

$$S_\tau[W](x) = \sup_{u\in\mathcal{U}^U_{M_{|x|}}} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\}$$

and

$$\begin{aligned} &\sup_{u\in\mathcal{U}_y^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\} \\ &= \sup_{u\in\mathcal{U}^U_{y,M_{|x|,|y|}}} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\}, \end{aligned}$$

where

$$\mathcal{U}^U_{M_{|x|}} \doteq \{u \in \mathcal{U}^U : \|u\| \le M_{|x|}\} \ \text{ and } \ \mathcal{U}^U_{y,M_{|x|,|y|}} \doteq \{u \in \mathcal{U}^U_y : \|u\| \le M_{|x|,|y|}\}.$$

In other words, ε-optimal u (for $\varepsilon \le 1$) are bounded in L_2-norm by $M_{|x|}$ and $M_{|x|,|y|}$ for each of the above problems, respectively.

Proof. The proof of the first assertion is simply Theorem 3.10. For the second assertion, one has (because $u \in \mathcal{U}^U_y$)

$$\sup_{u\in\mathcal{U}^U_y}\left\{\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau)\right\} = \sup_{u\in\mathcal{U}^U_y}\left\{\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt\right\} + W(y).$$

After one has this form, the proof is a slight simplification of the proof of the first assertion. Somewhat more specifically, one finds (for proper choice of C_1, $C_{2,\gamma}$)

$$\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(y) \le C_1|x|^2 - C_{2,\gamma}\|u\|^2\,dt + W(y) \le -1 \quad (4.43)$$

if

$$\|u\|^2 \ge (1/C_{2,\gamma})[1 + C_1|x|^2 + W(y)] \doteq M_{|x|,|y|}. \quad (4.44)$$

On the other hand, taking $u^0 \doteq 0$ with corresponding trajectory $\xi^0(\cdot)$, one has

$$\int_0^\tau l(\xi^0_t) - \frac{\gamma^2}{2}|u^0_t|^2\,dt + W(y) = \int_0^\tau l(\xi^0_t)\,dt + W(y) \ge 0. \quad (4.45)$$

Comparing (4.43) and (4.45) yields the result. □

Lemma 4.17. *H is continuous.*

Proof. The proof follows easily from the definition of H and Lemma 4.16, and so we only sketch it.

The proof of continuity in y is clear. We consider only the continuity in x. Let $x_1, x_2 \in \mathbf{R}^n$, $|x_1 - x_2| \le 1$, and let $R \ge \max\{|x_1|, |x_2|\}$. Let $y \in \mathbf{R}^n$. Let $u^{1,\varepsilon}$ be ε-optimal for $S_\tau[W](x_1)$, and let $u^{2,\varepsilon}$ be ε-optimal for $\sup_{u\in\mathcal{U}^U_y}\left\{\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau)\right\}$ with $\xi_0 = x_2$. Let the trajectory generated by (3.12) with $\xi_0 = x_i$ and control $u^{j,\varepsilon}$ be denoted by $\xi^{i,j}$ for $i, j \in \{1, 2\}$. Then

$$\begin{aligned}
H(x_1,y) - H(x_2,y) \le& \int_0^\tau l(\xi^{1,1}) - \frac{\gamma^2}{2}|u^{1,\varepsilon}|^2\,dt + W(\xi^{1,1}_\tau) \\
&- \int_0^\tau l(\xi^{2,1}) - \frac{\gamma^2}{2}|u^{1,\varepsilon}|^2\,dt + W(\xi^{2,1}_\tau) + \varepsilon \\
&- \int_0^\tau l(\xi^{1,2}) - \frac{\gamma^2}{2}|u^{2,\varepsilon}|^2\,dt + \int_0^\tau l(\xi^{2,2}) - \frac{\gamma^2}{2}|u^{2,\varepsilon}|^2\,dt + \varepsilon \\
=& \int_0^\tau l(\xi^{1,1}) - l(\xi^{2,1})\,dt + W(\xi^{1,1}_\tau) - W(\xi^{2,1}_\tau) \\
&+ \int_0^\tau l(\xi^{2,2}) - l(\xi^{1,2})\,dt + 2\varepsilon.
\end{aligned}$$

However, $u^{1,\varepsilon} \in \mathcal{U}^U_{M_R}$ and $u^{2,\varepsilon} \in \mathcal{U}^U_{y,M_{R,R}}$. Consequently (as for instance in (4.31)), there exists $\delta > 0$ such that if $|x_1 - x_2| \le \delta$, then $|\xi^{1,1}_t - \xi^{1,2}_t| \le \varepsilon$, $|\xi^{2,1}_t - \xi^{2,2}_t| \le \varepsilon$ for all $t \in [0,\tau]$ (and, of course, there exists $\widehat{d}_R < \infty$ such that $|\xi^{1,1_t}|, |\xi^{2,1}_t|, |\xi^{1,2}_t|, |\xi^{2,2}_t| \le \widehat{d}_R$ for all $t \in [0,\tau)$. One then finds that, using Assumption (A4.3I) and Corollary 4.12, for proper choice of $\widetilde{M}_{R,\tau}$

$$H(x_1,y) - H(x_2,y) \le \widetilde{M}_{R,\tau}\varepsilon$$

if $|x_1 - x_2|$ are sufficiently small. A reverse inequality follows through symmetry. □

Lemma 4.18. *$H(0,0) = 0$, $H(x,y) \ge 0$ for all $x, y \in \mathbf{R}^n$, and $H(x,x) > 0$ if $x \ne 0$.*

Proof. Noting that

$$\begin{aligned} H(x,y) = &\sup_{u \in \mathcal{U}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\} \\ &- \sup_{u \in \mathcal{U}^U_y} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(\xi_\tau) \right\}, \end{aligned} \tag{4.46}$$

it is obvious that $H(x,y) \ge 0$ for any x, y. (It might be worthwhile to note that by the assumptions on the dynamics, $\mathcal{U}^U_y \ne \emptyset$ for any $y \in \mathbf{R}^n$, and so $H(x,y) < \infty$.)

We consider the case $x \ne 0$. To prove $H(x,x) > 0$, suppose $H(x,x) \le 0$. Then by (4.46) and Theorem 4.4, this implies

$$W(x) - \sup_{u \in \mathcal{U}^U_x} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + W(x) \right\} \le 0,$$

or

$$\sup_{u \in \mathcal{U}^U_x} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \right\} \ge 0. \tag{4.47}$$

As in the proof of Lemma 4.8, one has

$$\frac{d}{dt}|\xi_t|^2 \le -c_f|\xi_t|^2 + \frac{m_\sigma^2}{c_f}|u_t|^2,$$

which, using $\xi_0 = x$, and solving the differential inequality (as before) yields

$$|\xi_t|^2 \le |x|^2 e^{-c_f t} + \frac{m_\sigma^2}{c_f}\|u\|^2_{L_2(0,t)}.$$

Because we require $\xi_\tau = x$ (i.e., $u \in \mathcal{U}^U_x$), this implies

$$\|u\|^2_{L_2(0,\tau)} \ge \frac{c_f}{m_\sigma^2}|x|^2(1 - e^{-c_f\tau}). \tag{4.48}$$

Now, by (A4.4I),

$$\gamma^2 > \frac{2m_\sigma^2 \alpha_l}{c_f^2} \doteq \overline{\gamma}^2.$$

Let

$$\delta \doteq \gamma^2 - \overline{\gamma}^2, \tag{4.49}$$

and let

$$\varepsilon \doteq \frac{\delta c_f}{8m_\sigma^2}(1 - e^{-c_f \tau})|x|^2. \tag{4.50}$$

Let u^ε be ε-optimal for (4.47) with corresponding trajectory ξ^ε. Then $\xi_\tau^\varepsilon = x$ and

$$\int_0^\tau l(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2 \, dt \geq -\varepsilon. \tag{4.51}$$

Let

$$\tilde{\gamma}^2 \doteq \gamma^2 - \frac{\delta}{2} \tag{4.52}$$

(so that $\tilde{\gamma}^2 = \overline{\gamma}^2 + \delta/2 > \overline{\gamma}^2$). Then,

$$\int_0^\tau l(\xi_t^\varepsilon) - \frac{\tilde{\gamma}^2}{2}|u_t^\varepsilon|^2 \, dt = \int_0^\tau l(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2 \, dt + \int_0^\tau \Big(\frac{\gamma^2 - \tilde{\gamma}^2}{2}\Big)|u_t^\varepsilon|^2 \, dt,$$

which by (4.51) and (4.52)

$$\geq -\varepsilon + \frac{\delta}{4}\|u^\varepsilon\|_{L_2(0,\tau)}^2,$$

which by (4.48) and (4.50)

$$\begin{aligned} &\geq -\frac{\delta c_f}{8m_\sigma^2}(1 - e^{-c_f \tau})|x|^2 + \frac{\delta c_f}{4m_\sigma^2}(1 - e^{-c_f \tau})|x|^2 \\ &= \frac{\delta c_f}{8m_\sigma^2}(1 - e^{-c_f \tau})|x|^2. \end{aligned} \tag{4.53}$$

Then, because $\xi_\tau^\varepsilon = x$, one may loop over this trajectory repeatedly by employing the disturbance

$$\widetilde{u}_t \doteq u_{t-\bar{n}\tau}^\varepsilon \quad \forall\, t \in [\bar{n}\tau, (\bar{n}+1)\tau), \ \ \forall \bar{n} \in \{0, 1, 2, \ldots\}.$$

Letting $\widetilde{\xi}$ be the corresponding trajectory, and employing (4.53), this yields

$$\int_0^{\bar{n}\tau} l(\widetilde{\xi}_t) - \frac{\tilde{\gamma}^2}{2}|\widetilde{u}_t|^2 \, dt \geq \bar{n}\frac{\delta c_f}{8m_\sigma^2}(1 - e^{-c_f \tau})|x|^2.$$

This implies that

$$\lim_{T \to \infty} \sup_{u \in \mathcal{U}^U} \int_0^T l(\xi_t) - \frac{\tilde{\gamma}^2}{2}|u_t|^2 \, dt = +\infty,$$

which contradicts Theorems 4.1 and 4.2 (because $\tilde{\gamma} > \overline{\gamma}$). Therefore $H(x, x) > 0$.

Lastly, we turn to the proof that $H(0,0) = 0$. Because $H(0,0) \geq 0$ was noted above, it is only required to show $H(0,0) \leq 0$. However, using (4.46) and the fact that the first supremum in (4.46) is $W(x)$ where $W(0) = 0$, one has

$$H(0,0) = -\sup_{u \in \mathcal{U}_0^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt \right\},$$

where $\xi_0 = \xi_\tau = 0$. But taking $u \equiv 0$ (and using $l \geq 0$) implies that

$$\sup_{u \in \mathcal{U}_0^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt \right\} \geq 0.$$

Consequently $H(0,0) \leq 0$. □

The proof of convergence of the power method relies on the following two lemmas.

Lemma 4.19. *Let* $u \in \mathcal{U}^U$.

$$\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt \leq W(x) - W(\xi_\tau) - H(x, \xi_\tau), \tag{4.54}$$

where ξ *satisfies (4.1) with input* u *and initial condition* $\xi_0 = x$.

Proof. By (4.42) and the fact that $W = S_\tau[W]$,

$$W(x) = H(x, y) + \sup_{u \in \mathcal{U}_y^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + W(\xi_\tau) \right\}. \tag{4.55}$$

Fix any $\overline{u} \in \mathcal{U}^U$, and let $\overline{y} = \overline{\xi}_\tau$ where $\overline{\xi}_\cdot$ is the corresponding trajectory. Then

$$W(x) = H(x, \overline{\xi}_\tau) + \sup_{u \in \mathcal{U}_{\overline{y}}^U} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + W(\xi_\tau) \right\}.$$

But $\overline{u} \in \mathcal{U}_{\overline{y}}^U$, so the result follows. □

Now let the $\{x_j\}$ be such that $x_1 = 0$, that is,

$$\psi_1(x) = -\frac{1}{2} x^T C x.$$

Lemma 4.20. $B_{1,1} = 0$. *Also, there exists* $\delta > 0$ *such that for all* $j \neq 1$, $B_{j,j} \leq -\delta$.

Proof. We prove the second assertion first. Let

$$\overline{x}_j = \operatorname{argmax}\{\psi_j(x) - W(x)\}, \tag{4.56}$$

the existence and uniqueness of which follows from the choice of the $\{\psi_j\}$ — in particular, the quadratic growth condition $C - c_R I > 0$. In fact, there exists $\overline{K} > 0$ such that

$$W(\overline{x}_j) - W(x) \le \psi_j(\overline{x}_j) - \psi_j(x) - \overline{K}|x - \overline{x}_j|^2 \quad \forall x \in \mathbf{R}^n. \tag{4.57}$$

Let u^ε be ε-optimal for $S_\tau[\psi_j]$ with corresponding process ξ^ε. Then

$$S_\tau[\psi_j](\overline{x}_j) \le \int_0^\tau l(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + \psi_j(\xi_\tau^\varepsilon) + \varepsilon,$$

which by Lemma 4.19,

$$\le W(\overline{x}_j) - W(\xi_\tau^\varepsilon) + \psi_j(\xi_\tau^\varepsilon) - H(\overline{x}_j, \xi_\tau^\varepsilon) + \varepsilon,$$

which by (4.57),

$$\le \psi_j(\overline{x}_j) - \overline{K}|\xi_\tau^\varepsilon - \overline{x}_j|^2 - H(\overline{x}_j, \xi_\tau^\varepsilon) + \varepsilon \tag{4.58}$$

and so, by Lemmas 4.17 and 4.18, there exists $\delta_j > 0$ such that

$$\le \psi_j(\overline{x}_j) - \delta_j + \varepsilon$$

which yields the assertion (with $\delta \doteq \min_{j \le n} \delta_j$).

We now prove the first assertion. First we should perhaps note that by the definition of ψ_1 and (4.11), $\overline{x}_1 = 0$, the origin. Note that for disturbance process $u_t^0 \equiv 0$ and initial condition $x = 0$, the corresponding state process is $\xi_t^0 = 0$ for all t. Consequently, $\int_0^\tau l(\xi_t^0) - \frac{\gamma^2}{2}|u_t^0|^2\,dt + \psi_1(\xi_\tau^0) = 0$ Suppose there exists a disturbance process, $\widehat{u}$ and $\varepsilon > 0$ such that one has $\int_0^\tau l(\widehat{\xi}_t) - \frac{\gamma^2}{2}|\widehat{u}_t|^2\,dt + \psi_1(\widehat{\xi}_\tau) \ge \varepsilon > 0$ where $\widehat{\xi}$ is the state process corresponding to $\widehat{\xi}_0 = 0$ and disturbance $\widehat{u}$.

Suppose $\widehat{\xi}_\tau = 0$. In that case, let $\widehat{u}^\infty$ be an infinite repetition of $\widehat{u}$, i.e., let

$$\widehat{u}_t^\infty = \widehat{u}_{t - n_t\tau} \quad \forall t \ge 0,$$

where n_t is the largest integer less than or equal to t/τ. Then, letting $\widehat{\xi}^\infty$ be the corresponding state process one has (using Theorem 4.2),

$$\begin{aligned} W(0) &\ge \limsup_{\bar{n}\to\infty} \int_0^{\bar{n}\tau} l(\widehat{\xi}_t^\infty) - \frac{\gamma^2}{2}|\widehat{u}_t^\infty|^2\,dt \\ &\ge \lim_{\bar{n}\to\infty} \{\bar{n}\varepsilon\} = \infty, \end{aligned} \tag{4.59}$$

which contradicts the fact that W satisfies (4.11).

Now consider the case $\widehat{\xi}_\tau \neq 0$. This introduces an additional technical complication. Essentially, after $t = \tau$, one lets the system drift back arbitrarily close to the origin, then pushes it directly to the origin with the disturbance process. The integral over this whole path will be positive, and one again creates an infinite repetition of this loop to obtain the same contradiction. We now make this argument concrete. We first let $\overline{\xi}$ be a solution of (4.1) with $\overline{\xi}_0 = x \neq 0$ which will be driven to the origin in finite time as follows. Let $\bar{u}_t = \sigma^{-1}(\overline{\xi}_t)\Big[-f(\overline{\xi}_t) - \overline{\xi}_t/|\overline{\xi}_t|\Big]$ for $t \in [0, |x|]$ so that $\dot{\overline{\xi}} = -\overline{\xi}_t/|\overline{\xi}_t|$, and consequently, $\overline{\xi}_{|x|} = 0$. Note also that by Assumptions (A4.1I) and (A4.2I), $|\bar{u}_t| \leq m_\sigma[1 + K|\overline{\xi}_t|] \leq m_\sigma[1 + K|x|]$ for all $t \in [0, |x|]$. Consequently, if $|x| \leq \delta < 1$, one has $\int_0^{|x|} l(\overline{\xi}_t) - \frac{\gamma^2}{2}|\bar{u}_t|^2\, dt \geq -\frac{\gamma^2}{2} m_\sigma^2 (1 + K\delta)^2 \delta \geq -\gamma^2 m_\sigma^2 (1 + K^2)\delta$. Now, we begin to construct our loop trajectory. Let $\widehat{u}, \widehat{\xi}$ be as above where now $\widehat{\xi}_\tau \neq 0$. Let $\widetilde{u}_t = \widehat{u}_t$ for all $t \in [0, \tau)$, and so the corresponding trajectory satisfies $\widetilde{\xi}_t = \widehat{\xi}_t$ (for all $t \in [0, \tau]$). Let $\widetilde{u}_t = 0$ for all $t \in [\tau, \hat{\tau})$ where $\hat{\tau}$ is chosen such that $|\widetilde{\xi}_{\hat{\tau}}| \leq \delta$ where δ is yet to be specified and the existence of such a $\hat{\tau}$ follows directly from Assumption (A4.1I). We take $\delta < 1$ such that $\gamma^2 m_\sigma^2 (1 + K^2)\delta < \varepsilon/2$. Now, let $\widetilde{u}_t = \bar{u}_{t-\hat{\tau}}$ so that with $x = \overline{\xi}_0 = \widetilde{\xi}_{\hat{\tau}}$, one has $\widetilde{\xi}_t = \overline{\xi}_{t-\hat{\tau}}$ for all $t \in [\hat{\tau}, \hat{\tau} + |\widetilde{\xi}_{\hat{\tau}}|]$ and in particular

$$\widetilde{\xi}_{\hat{\tau} + |\widetilde{\xi}_{\hat{\tau}}|} = 0.$$

Further, $\int_{\hat{\tau}}^{\hat{\tau} + |\widetilde{\xi}_{\hat{\tau}}|} l(\widetilde{\xi}_t) - \frac{\gamma^2}{2}|\widetilde{u}_t|^2\, dt \geq -\gamma^2 m_\sigma^2 (1 + K^2)\delta > -\varepsilon/2$. Putting all three segments together, we see that $\int_0^{\hat{\tau} + |\widetilde{\xi}_{\hat{\tau}}|} l(\widetilde{\xi}_t) - \frac{\gamma^2}{2}|\widetilde{u}_t|^2\, dt \geq \varepsilon + 0 - \varepsilon/2 = \varepsilon/2$ and that $\widetilde{\xi}_{\hat{\tau} + |\widetilde{\xi}_{\hat{\tau}}|} = \widetilde{\xi}_0 = 0$. Then, as in the previous case, one repeats the disturbance process $\widetilde{u}$ an infinite number of times to obtain a contradiction of the form (4.59).

Thus, in either case, we find that there does not exist a disturbance process such that $\int_0^\tau l(\widehat{\xi}_t) - \frac{\gamma^2}{2}|\widehat{u}_t|^2\, dt + \psi_1(\widehat{\xi}_\tau) > 0$. Therefore, $u^0 \equiv 0$ is optimal, and so

$$S_\tau[\psi_1](0) = \int_0^\tau l(\xi_t^0) - \frac{\gamma^2}{2}|w_t^0|^2\, dt + \psi_1(\xi_\tau^0) = 0$$

or, equivalently, $S_\tau[\psi_1](\overline{x}_1) = \psi_1(\overline{x}_1)$. This implies

$$B_{1,1} \leq 0. \tag{4.60}$$

To get the reverse inequality, note that $W(0) = 0$ and $W(x) \geq 0$ for all $x \in \mathbf{R}^n$. Because W is smooth in some arbitrarily small neighborhood of the origin [108], one has $\nabla W(0) = 0$. Further, since $S_\tau[\psi_1](0) = 0$, $S_\tau[\psi_1](x) \leq S_\tau[W](x) = W(x)$, $W(0) = 0$, $\nabla W(0) = 0$ and $S_\tau[\psi_1]$ semiconvex, one has $\nabla S_\tau[\psi_1](0) = 0$. Then, by the semiconvexity preserving property of S_τ for our choice of C (i.e., Assumption (A4.5I)), and the fact that $\psi_1(x) = -\frac{1}{2}x^T C x$, we see that $\psi_1(x) - S_\tau[\psi_1](x) \leq 0$ for all x. Therefore

$$B_{1,1} \geq 0. \tag{4.61}$$

By (4.60) and (4.61), we are done. □

Theorem 4.21. *Let $N \in \{1, 2, \ldots, \nu\}$, $\{k_i\}_{i=1}^{i=N+1}$ such that $1 \leq k_i \leq \nu$ for all i and $k_{N+1} = k_1$. Suppose we are not in the case $k_i = 1$ for all i. Then*

$$\sum_{i=1}^{N} B_{k_i,k_{i+1}} \leq -\delta.$$

Proof. By Lemma 4.20, this is true for $N = 1$. We first prove the case $N = 2$. The proof of the general case will then be relatively clear, but we will also provide a sketch of the induction argument. First note the monotonicity of the semigroup in the sense that if $g_1(x) \leq g_2(x)$ for all x, then

$$S_\tau[g_1](x) \leq S_\tau[g_2](x) \quad \forall\, x \in \mathbf{R}^n. \tag{4.62}$$

Suppose either $i \neq 1$ or $j \neq 1$. Without loss of generality, suppose $i \neq 0$. By definition $\psi_j(x) + B_{j,i} \leq S_\tau[\psi_i](x)$ for all x. Using (4.62), this implies

$$S_\tau[\psi_j + B_{j,i}] \leq S_\tau\big[S_\tau[\psi_i]\big] = S_{2\tau}[\psi_i],$$

and by the max-plus linearity of S_τ, this yields

$$S_\tau[\psi_j] + B_{j,i} \leq S_{2\tau}[\psi_i],$$

which implies in particular that

$$S_\tau[\psi_j](\overline{x}_i) + B_{j,i} \leq S_{2\tau}[\psi_i](\overline{x}_i), \tag{4.63}$$

where $\overline{x}_i$ is the argmax as given in (4.56).

On the other hand, by (4.58) (with ξ^ε generated by an ε-optimal u^ε)

$$S_{2\tau}[\psi_i](\overline{x}_i) \leq \psi_i(\overline{x}_i) - \overline{K}|\xi^\varepsilon_{2\tau} - \overline{x}_i|^2 - H(\overline{x}_i, \xi^\varepsilon_{2\tau}) + \varepsilon,$$

and then by Lemmas 4.17, 4.18 (as in the proof of Lemma 4.20)

$$\leq \psi_i(\overline{x}_i) - \delta_i + \varepsilon. \tag{4.64}$$

(Note here that one may need to reduce the size of δ_i and δ to allow for the finite number of time horizons $\{\tau, 2\tau, \ldots, \nu\tau\}$.) Combining (4.63) and (4.64) yields

$$S_\tau[\psi_j](\overline{x}_i) - \psi_i(\overline{x}_i) + B_{j,i} \leq -\delta_i + \varepsilon \leq -\delta + \varepsilon. \tag{4.65}$$

Then, using the fact that $B_{i,j} = \min_x\{S_\tau[\psi_j](x) - \psi_i(x)\}$, and letting $\varepsilon \downarrow 0$ yields $B_{i,j} + B_{j,i} \leq -\delta$.

The case for $N > 2$ follows easily by induction, and we give only the main points. To obtain the result analogous to (4.63) for $N > 2$, suppose that for some $N \geq 2$, one has

$$S_\tau[\psi_j] + B_{j,k_1} + \sum_{l=1}^{N-2} B_{k_l,k_{l+1}} \le S_{N\tau}[\psi_{k_{N-1}}] \tag{4.66}$$

which we already have for $N = 2$, Then

$$\begin{aligned} S_\tau[\psi_j] + B_{j,k_1} + \sum_{l=1}^{N-2} B_{k_l,k_{l+1}} + B_{k_{N-1},i} &\le S_{N\tau}[\psi_{k_{N-1}}] + B_{k_{N-1},i} \\ &= S_{N\tau}[\psi_{k_{N-1}} + B_{k_{N-1},i}] \\ &\le S_{N\tau}[S_\tau[\psi_i]] = S_{(N+1)\tau}[\psi_i]. \end{aligned}$$

Consequently, an induction argument yields this for all N. The remainder of the proof follows analogously to that for the case of $N = 2$. □

A sufficient condition that a matrix has exactly one max-plus eigenvalue is that it not have any entries that are $-\infty$. This is demonstrated under a weaker sufficient condition in [6], Theorem 3.23. (See also [25].) By the above results, this eigenvalue must be zero (ignoring errors due to approximation).

Theorem 4.22. $\lim_{N\to\infty} B^N \otimes 0$ *exists, converges in a finite number of steps, and satisfies* $e = B \otimes e$.

Proof. Given the matrix B, one can associate a corresponding directed graph with ν nodes (recall B is $\nu \times \nu$) where the transition cost from node i to node j is $B_{i,j}$. (This is referred to as the precedence graph in [6], and more information on the relationship between the precedence graph and the matrix may be found there.) Let us refer to the node corresponding to ψ_1 as the origin node. The elements of $B^N \otimes 0$ correspond to the optimal costs of transition paths of length N, that is, $[B^N \otimes 0]_i$ is the maximum cost over all paths of length N starting at node i and ending at any node k. For instance,

$$[B^2 \otimes 0]_i = \bigoplus_j \bigoplus_k [B_{i,j} \otimes B_{j,k} \otimes 0] = \max_{j,k}[B_{i,j} + B_{j,k} + 0].$$

A loop will be defined to be any path (or segment of a larger path) that starts and ends at the same node. Because the graph has only ν nodes, any path of length greater than ν must contain at least one loop. Let l_i be the maximal cost of a path from node i to the origin node. By Theorem 4.21, any loop other than the trivial loop (the loop of length one from the origin node back to itself) must have negative cost to travel the loop. Thus, the path corresponding to cost l_i must not have any loops (except possibly trivial loops which we will ignore). Consequently, the length of the path corresponding to l_i might as well be assumed to be no longer than ν and have no loops. Similarly, let l^0 be the maximal cost of a path from the origin node to any other node. By the same analysis as for l_i, one may assume without loss of generality that the path has no loops and has length no longer than ν. Then $\bar{l}_i \doteq l_i + l^0$ is the maximal cost of any path from node i through the origin node to any node k, and the maximal cost path will have no loops (ignoring trivial loops). This path has length $\leq 2\nu$.

Let the maximum cost path from any node to any other node without loops be C_B.

Let $N > 2\nu$. Suppose a path from node i to some other node has length N. The path must contain at least one loop. In fact, defining M_N to be the largest integer less than N/ν, a path from node i to the origin node of length N must contain at least M_N loops. Thus, by Theorem 4.21, the path must have cost at most $C_B - M_N\delta$ unless at least one of the loops is the trivial loop. Suppose N is sufficiently large that $C_B - M_N\delta < \bar{l}_i$. Suppose the path does not contain the trivial loop. Then, the path yielding cost $\bar{l}_i$ with the addition of trivial loops until the length becomes N (recall the path passed through the origin node) has cost $\bar{l}_i > C_B - M_N\delta$, and so has larger cost. Alternatively, suppose the path does contain at least one trivial loop and at least one other loop. Let the other loop have length m. Then the path cost may be increased by at least δ by elimination of this loop followed by the addition of m trivial loops (so as to maintain total path length N). Thus, we find that any path of length sufficiently large so that $C_B - M_N/\delta < \bar{l}_i$ (and $N > 2\nu$) cannot be of maximal cost unless it is the loopless path of cost $\bar{l}_i$ with the addition of trivial loops until the length becomes N. Therefore, $[B^N \otimes 0]_i = \bar{l}_i$ for all N such that $C_B - M_N/\delta < \bar{l}_i$ and $N > 2\nu$.

Thus, we see that $B^N \otimes 0$ converges in a finite number of steps. Denote the limit as e. Note that

$$e = \lim_{N\to\infty} B^N \otimes 0 = B \otimes \lim_{N\to\infty} B^{N-1} \otimes 0$$
$$= B \otimes \lim_{M\to\infty} B^M \otimes 0 = B \otimes e,$$

and so we see that e is a max-plus eigenvector corresponding to max-plus eigenvalue 0. □

Not only is the max-plus eigenvalue unique [6], but one also obtains the following.

Corollary 4.23. *There is a unique max-plus eigenvector, e up to a max-plus multiplicative constant, and of course this is the output of the above power method.*

Proof. Let e be the eigenvector obtained by the above power method iteration (i.e., $e = \lim_{N\to\infty} B^N \otimes 0$). Consider $[B^N \otimes e']_i$ when e' is any eigenvector and N will be taken sufficiently large. This is the maximal cost of any path of length N starting at node i and ending at any node k with terminal cost e'_k. By a proof similar to that used in the proof of Theorem 4.22, for N sufficiently large, say $N \geq \overline{N}$, the maximal cost path will pass through the origin node, and the only loops will be trivial loops. Let l_i again be the maximal cost of a loopless path from node i to the origin node, and now let $\hat{l}^0_{e'}$ be the maximal cost of a loopless path from the origin node to any node, k, with terminal cost e_k. Then, the maximal cost of the path of length N will be given by $l_i + \hat{l}^0_{e'}$ independent of $N \geq \overline{N}$.

Because e' is an eigenvector, for each $i \in \{1, 2, \ldots, n\}$

$$e'_i = \left[\lim_{N\to\infty} B^N \otimes e'\right]_i = l_i + \hat{l}^0_{e'} = l_i + l^0 + (\hat{l}^0_{e'} - l^0)$$

(where l^0 is given in the proof of Theorem 4.22)

$$\begin{aligned} &= \left[\lim_{N\to\infty} B^N \otimes 0\right]_i + (\hat{l}^0_{e'} - l^0) \\ &= e_i + (\hat{l}^0_{e'} - l^0) \\ &= (\hat{l}^0_{e'} - l^0) \otimes e_i. \end{aligned}$$

Thus, $e' = (\hat{l}^0_{e'} - l^0) \otimes e$ which proves the result. □

4.6 Computing B: Initial Notes

The obvious first step in application of this method is the computation, more exactly numerical approximation, of B. A more complete discussion of the computation of B will appear in the next chapter. For completeness in this chapter, we indicate one approach (but see Section 5.3.1 for a more complete discussion).

Recall that,for any specific τ, (see definition (4.36))

$$B_{i,j} = -\max_{x\in B_R}\left\{\psi_i(x) - \sup_{u\in\mathcal{U}^U}\left[\int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt + \psi_j(\xi_\tau)\right]\right\}, \quad (4.67)$$

where we suppress the dependence of B on τ to reduce notation. Also, recall that (see, e.g., [37]) given any $R < \infty$, there exists a $\tau_R < \infty$ such that $\widehat{V}(x,\tau) = S_\tau[\psi_j](x)$ is a C^2 solution of (4.7) on $\overline{B}_R \times [0, \tau_R)$. Consequently, for $\tau \in (0, \tau_R)$, one can use a first-order Taylor approximation. That is,

$$B_{i,j} \simeq -\max_{x\in B_R}\left\{\psi_i(x) - \psi_j(x) - \tau\left[\frac{1}{2\gamma^2}\nabla\psi_j^T\sigma(x)\sigma^T(x)\nabla\psi_j + f^T(x)\nabla\psi_j + l(x)\right]\right\}.$$

Of course, one can proceed to higher-order terms.

Remark 4.24. It has been observed that one does not need to compute the entire B matrix. That is, there exists $\delta > 0$ such that one may take $\widetilde{B}_{i,j} = -\infty$ if $|x_i - x_j| > \delta$ with $\widetilde{B}_{i,j} = B_{i,j}$ otherwise, and then $\widetilde{B}$ has the same eigenvector as B. The value of δ appears dependent on τ. Some thoughts on the dependence of such a δ on τ can be found in [3], [4], [67].

Remark 4.25. Recall that $\tau > 0$ is free. Ideally, one obtains the same solution for all τ (ignoring error analysis — discussed in the next chapter). Interestingly, one may use different values of τ for different rows of the B matrix without altering the eigenvector. A discussion of the dependence of the error size on τ appears in Chapter 5.

4.7 Outline of Algorithm

For purposes of readability, we briefly outline the steps in the max-plus eigenvector algorithm for (approximate) computation of W over a ball, $\overline{B}_R$.

1. Choose a set of max-plus basis functions of the form $\psi_i(x) = -\frac{1}{2}(x - x_i)^T C(x-x_i)$ where the x_i lie in $\overline{B}_{D_R}$. (In practice however, a rectangular grid has been used.) Choose a "time-step," τ.
2. Compute (approximately) elements of the matrix B given by
$$B_{j,i} = -\max_{x\in\overline{B}_R} \{\psi_j(x) - S_\tau[\psi_i](x)\}.$$
A reasonably efficient means of computing B is important, and a Runge–Kutta based approach is indicated in Section 5.3.1.
3. Compute the max-plus eigenvector of B corresponding to max-plus eigenvalue $\lambda = 0$ (i.e., the solution of $e = B \otimes e$). This is obtained from the max-plus power method $a^{(k+1)} = B \otimes a^{(k)}$ starting from $a^{(0)}$ where $a^{(0)}$ is taken to be the zero vector. This converges exactly in a finite number of steps, say $\overline{n}$. Let the limit be $\overline{a} = a^{(\overline{n})}$.
4. Construct the solution approximation from $\widehat{W}(x) \doteq \bigoplus_{i=1}^{\nu} \overline{a}_i \otimes \psi_i(x)$ on $\overline{B}_R(0)$.

4.8 A Control Problem Without Nominal Stability and a Game

In this last section, we briefly indicate two extensions: one to a minimizing control problem where the dynamics are not nominally stable, and another to a game problem whose DPE is an HJB PDE. The material in this section may be skipped without creating any difficulty in the following chapters.

Consider the control problem with dynamics

$$\dot{\xi} = A(\xi) + Q(\xi)v, \tag{4.68}$$
$$\xi_0 = x, \tag{4.69}$$

where $v \in \mathcal{U} = L_2^{loc}([0,\infty); \mathbf{R}^l)$. Let the payoff and value be given by

$$J_2(x,T,v) = \int_0^T l(\xi_t) + \tfrac{1}{2}|v_t|^2 dt, \tag{4.70}$$

$$W_2(x) = \inf_{v\in\mathcal{U}} \sup_{T\in[0,\infty)} J_2(x,T,v). \tag{4.71}$$

Note that $v.$ is a minimizing controller.

Assume that Q is uniformly bounded and nondegenerate in the sense that there exist $\alpha_1, \alpha_2 \in (0,\infty)$ such that

$$\alpha_1|y|^2 \le \tfrac{1}{2} y^T Q(x) Q^T(x) y \le \alpha_2 |y|^2 \quad \forall x,y \in \mathbf{R}^n \tag{A4.1m}$$

Note that Assumption (A4.1m) implies

$$|Q(x)|, |Q^{-1}(x)| \le c_2 \quad \forall x \in \mathbf{R}^n \tag{4.72}$$

for some $c_2 \in (0,\infty)$.

We assume that $A, Q, l \in C^2(\mathbf{R}^n)$. We also assume that there exists $c_1 \in (0,\infty)$ such that

$$\begin{aligned} &|A_x(x)|,\ |A_{xx}(x)| \le c_1 \quad \forall x \in \mathbf{R}^n, \\ &A(0) = 0, \\ &|Q_x(x)|,\ |(Q^{-1})_x(x)|,\ |(Q^{-1})_{xx}(x)| \le c_1 \quad \forall x \in \mathbf{R}^n. \end{aligned} \tag{A4.2m}$$

We assume that there exist $c_3, c_4, c_5 \in (0,\infty)$ such that

$$|l_{xx}(x)| \le c_3 \quad \forall x \in \mathbf{R}^n, \tag{A4.3m}$$

and

$$c_4|x|^2 \le l(x) \le c_5|x|^2 \quad \forall x \in \mathbf{R}^n. \tag{A4.4m}$$

The quadratic bounds in (A4.4m) are chosen for convenience; it is expected that most of the results will hold for more general polynomial bounds (but which still enforce $l(0)=0$ and $l(x)>0$ for $x\neq 0$).

Note that we do not assume stability of the nominal dynamics; more exactly, we do not keep our typical assumption which in this context would be $x^T A(x) \le -k_A|x|^2$ for some k_A and all x. However, the facts that $l(0)=0$, $l(x)>0$ for $x \neq 0$ and the non-negativity of the $\frac{1}{2}|v_t|^2$ term imply that near-optimal controls must drive the system toward the origin. (This statement will be made more exact below.) Consequently, any near optimal control will stabilize the system to the origin.

Lastly, note that again by the non-negativity of the running cost

$$W_2(x) = \inf_{v\in\mathcal{U}} \left\{ \int_0^\infty l(\xi_t) + \tfrac{1}{2}|v_t|^2 dt \right\}. \tag{4.73}$$

We begin with a result that serves to produce some a priori bounds on the behavior of near-optimal trajectories of the system.

Lemma 4.26. *Given $R<\infty$, there exists $M_R<\infty$ such that $W_2(x)\le M_R$ for all $x\in B_R(0)$. Further, there exists $\widehat{M}_R<\infty$ such that for any $\varepsilon\in(0,1)$*

the following statements hold: If v^ε is ε-optimal for control problem (4.73), then $\|v^\varepsilon\|_{L_2[0,\infty)} \leq \widehat{M}_R$, and $\|\xi^\varepsilon\|_{L_2[0,\infty)} \leq \widehat{M}_R$ where ξ^ε satisfies (4.68), (4.69) with input v^ε. If v^ε is ε-optimal for finite horizon problem

$$V(x,T) = \inf_{v\in\mathcal{U}} \left\{ \int_0^T l(\xi_t) + \tfrac{1}{2}|v_t|^2 dt \right\}, \tag{4.74}$$

then $\|v^\varepsilon\|_{L_2[0,T]} \leq \widehat{M}_R$, and $\|\xi^\varepsilon\|_{L_2[0,T]} \leq \widehat{M}_R$.

Proof. Suppose $x = 0$. Note that with input $v^0 \equiv 0$, one has $\xi_t^0 = 0$ for all $t \geq 0$ by Assumption (A4.2m). Therefore, $J(0,T,v^0) = 0$. Since $l \geq 0$, we have $W_2(0) = V(0,T) = 0$.

Now suppose $x \neq 0$. Let

$$\tilde{v}_t \doteq \begin{cases} Q^{-1}(\xi_t)\left(-\frac{x}{|x|} - A(\xi_t)\right) & \text{if } 0 \leq t \leq |x|, \\ 0 & \text{if } t > |x|. \end{cases}$$

Then

$$\dot{\tilde{\xi}} = A(\tilde{\xi}) + Q(\tilde{\xi})\tilde{v} = \begin{cases} -\frac{x}{|x|} & \text{if } 0 \leq t \leq |x|, \\ A(\tilde{\xi}) & \text{if } t > |x| \end{cases},$$

which yields

$$\tilde{\xi}_t = \begin{cases} x\left(1 - \frac{t}{|x|}\right) & \text{if } t \leq |x| \\ 0 & \text{if } t > |x|, \end{cases}$$

and this implies $\tilde{v} \in C[0,\infty)$ with $\tilde{v}_t = 0$ for $t \geq |x|$. Therefore

$$J_2(x,T,\tilde{v}) = \int_0^{|x|} l\left(x - \frac{tx}{|x|}\right) + \tfrac{1}{2}|\tilde{v}_t|^2\, dt$$

for all $T > |x|$. Because $|Q|, |Q^{-1}|$ are bounded, $\tilde{v} \in \mathcal{U}$. Thus, there exists some $M_R < \infty$ (independent of $T > |x|$) such that

$$\begin{aligned} V(x,T) &= \inf_{v\in\mathcal{U}} J_2(x,T,v) \leq M_R \quad \forall T \geq |x|, \\ W_2(x) &= \inf_{v\in\mathcal{U}} \lim_{T\to\infty} J_2(x,T,v) \leq M_R \end{aligned} \tag{4.75}$$

for all $x \in B_R(0)$. Using this same $\tilde{v}$ for $T \in (0,|x|]$, one obtains (4.75) for all $T \in (0,\infty)$ and $x \in B_R(0)$.

If v^ε is ε-optimal for V, then

$$\int_0^T \tfrac{1}{2}|v_t^\varepsilon|^2\, dt \leq \int_0^T l(\xi_t^\varepsilon) + \tfrac{1}{2}|v_t^\varepsilon|^2\, dt < V(x,T) + \varepsilon \leq M_R + 1$$

and

$$\int_0^T c_4|\xi_t^\varepsilon|^2\, dt \leq \int_0^T l(\xi_t^\varepsilon) + \tfrac{1}{2}|v_t^\varepsilon|^2\, dt < V(x,T) + \varepsilon \leq M_R + 1.$$

Consequently,

$$\|v^\varepsilon\|_{L_2[0,T]} \le [2(M_R+1)]^{1/2}$$

and

$$\|\xi^\varepsilon\|_{L_2[0,T]} \le [(1/c_4)(M_R+1)]^{1/2}.$$

Similarly, if v^ε is ε-optimal for W_2, then $\|v^\varepsilon\|_{L_2[0,\infty)} \le [2(M_R+1)]^{1/2}$ and $\|\xi^\varepsilon\|_{L_2[0,\infty)} \le [(1/c_4)(M_R+1)]^{1/2}$. To complete the proof, simply let $\widehat{M}_R := \max\left\{[2(M_R+1)]^{1/2}, [(1/c_4)(M_R+1)]^{1/2}\right\}$. □

The corresponding HJB PDE problem is

$$0 = -\inf_{v\in\mathbf{R}^l}\left\{(A(x)+Q(x)v)\cdot\nabla W + l(x) + \tfrac{1}{2}|v|^2\right\} \tag{4.76}$$

$$= -\left[A(x)\cdot\nabla W + l(x) - \tfrac{1}{2}(\nabla W)^T Q(x)Q^T(x)\nabla W\right] \doteq H(x,\nabla W), \tag{4.77}$$

with boundary condition $W(0)=0$. The following two results are entirely standard, and in fact, simpler that the infinite time-horizon results of Chapter 3, due to the fact that one is minimizing a non-negative cost here. We do not include the proofs.

Theorem 4.27. *W_2 satisfies the DPP:*

$$W_2(x) = \inf_{v\in\mathcal{U}}\left\{\int_0^\tau l(\xi_t) + \tfrac{1}{2}|v_t|^2 dt + W_2(\xi_\tau)\right\} \tag{4.78}$$

for any $\tau\in(0,\infty)$.

Theorem 4.28. *W_2 is a continuous viscosity solution of HJB PDE (4.77).*

We now proceed to demonstrate that the value function is semiconcave. First we will need a technical lemma.

Lemma 4.29. *Given $\varepsilon\in(0,1]$, $R<\infty$ and $|x|<R$, let $v^{\varepsilon,T}$ be ε-optimal for finite time-horizon problem (4.74). Then there exists $d_R<\infty$ (independent of ε, x, and T) such that*

$$|\xi_t| \le d_R \quad \forall t\in[0,T].$$

Proof. For $T\le 2$, the result is obvious by the bound on $v^{\varepsilon,T}$ in Lemma 4.26 combined with (4.72) and (A4.2m).

We consider the case $T>2$. Let $t_1\in(0,T)$ and $s\in(-t_1,T-t_1)$. For ease of notation, consider only the case $s>0$ for the moment. One has

$$\left|\int_{t_1}^{t_1+s}\dot{\xi}_t\,dt\right| \le \int_{t_1}^{t_1+s} |A(\xi_t) + +|Q(\xi_t)|\,|v_t|\,dt$$

and by the assumptions, there exist $k_0, k_1 < \infty$ such that

$$\leq \int_{t_1}^{t_1+s} k_0 + k_1|\xi_t| + k_1|v_t|\,dt$$
$$\leq k_0 s + k_1' \left[\|\xi\|_{L_2(0,T)} + \|v\|_{L_2(0,T)}\right]\sqrt{s}$$

for proper choice of k_1'. Considering also the case $s < 0$, one then easily sees that for any $s \in (-t_1, T - t_1)$,

$$|\xi_{t_1+s}| \geq |\xi_{t_1}| - k_0|s| - k_1' \left[\|\xi\|_{L_2(0,T)} + \|v\|_{L_2(0,T)}\right]\sqrt{|s|}.$$

If v is ε-optimal with $\varepsilon \in (0,1]$, then combining this with Lemma 4.26, one finds

$$|\xi_{t_1+s}| \geq |\xi_{t_1}| - k_0|s| - 2k_1'\widehat{M}_R\sqrt{|s|}. \tag{4.79}$$

By (4.79),

$$\|\xi\|^2_{L_2(0,T)} \geq \int_0^T \left(|\xi_{t_1}| - k_0|r - t_1| - 2k_1'\widehat{M}_R\sqrt{|r - t_1|}\right)^2 dr. \tag{4.80}$$

However, there exists $\rho : [0,\infty) \to [0,\infty)$ such that

$$k_0|s| + 2k_1'\widehat{M}_R\sqrt{|s|} \leq \alpha/2 \quad \forall\, s \leq \rho(\alpha),$$

where $\rho(\alpha) \to \infty$ as $\alpha \to \infty$. Consequently,

$$\left(|\xi_{t_1}| - k_0|r - t_1| - 2k_1'\widehat{M}_R\sqrt{|r - t_1|}\right)^2 \geq \frac{|\xi_{t_1}|^2}{4} \tag{4.81}$$

if $|r - t_1| \leq \rho(\alpha)$.

Combining (4.80) and (4.81), one finds

$$\|\xi\|^2_{L_2(0,T)} \geq \frac{|\xi_{t_1}|^2}{4} \min\left\{T/2\,,\, \rho(|\xi_{t_1}|)\right\},$$

which since $T > 2$,

$$\geq \frac{|\xi_{t_1}|^2}{4} \min\left\{1\,,\, \rho(|\xi_{t_1}|)\right\}.$$

Combining this inequality with Lemma 4.26 yields

$$\max_{t_1 \in [0,T]} \left[\frac{|\xi_{t_1}|^2}{4} \min\left\{1\,,\, \rho(|\xi_{t_1}|)\right\}\right] \leq \widehat{M}_R^2,$$

which yields the result. □

The definition of semiconcavity is directly analogous to that of semiconvexity, and we do not include it. The semiconcavity of the value function will follow directly from the uniform-in-time semiconcavity of the finite time-horizon value function.

Theorem 4.30. *The value function for the finite time-horizon problem (4.74), is semiconcave in x, and the semiconcavity constant over any $\overline{B}_R$ is bounded uniformly in T.*

Proof. We will show that second differences of V in x are bounded from above. Let $R < \infty$ and $x \in B_R$. Recall

$$V(x,T) := \inf_{v \in \mathcal{U}} J_2(x,T,v). \tag{4.82}$$

Since we are interested in the long-term semiconcavity, we will assume $T \geq 1$. Let $\eta \in \mathbf{R}^n$ with $|\eta| = 1$. Let $\delta, \varepsilon \in (0,1)$. Let v^ε be ε-optimal for (4.74). Let

$$\begin{aligned}
\dot{\xi}^\varepsilon &= A(\xi^\varepsilon) + Q(\xi^\varepsilon)v^\varepsilon, \quad \xi^\varepsilon(0) = x,\\
\dot{\xi}^{\varepsilon,+} &= A(\xi^{\varepsilon,+}) + Q(\xi^{\varepsilon,+})v^{\varepsilon,+}, \quad \xi^{\varepsilon,+}(0) = x + \delta\eta,\\
\dot{\xi}^{\varepsilon,-} &= A(\xi^{\varepsilon,-}) + Q(\xi^{\varepsilon,-})v^{\varepsilon,-}, \quad \xi^{\varepsilon,-}(0) = x - \delta\eta,
\end{aligned}$$

where the controls $v^{\varepsilon,+}, v^{\varepsilon,-}$ are given as follows. Let

$$v_t^{\varepsilon,+} = \begin{cases} Q^{-1}(\xi_t^{\varepsilon,+})\left[A(\xi_t^\varepsilon) - A(\xi_t^{\varepsilon,+}) + Q(\xi_t^\varepsilon)v_t^\varepsilon - \delta\eta\right] & \text{if } t \leq 1\\ v_t^\varepsilon & \text{if } t > 1\end{cases}$$

and

$$v_t^{\varepsilon,-} = \begin{cases} Q^{-1}(\xi_t^{\varepsilon,-})\left[A(\xi_t^\varepsilon) - A(\xi_t^{\varepsilon,-}) + Q(\xi_t^\varepsilon)v_t^\varepsilon + \delta\eta\right] & \text{if } t \leq 1\\ v_t^\varepsilon & \text{if } t > 1.\end{cases}$$

With these choices of $v^{\varepsilon,+}, v^{\varepsilon,-}$, one has

$$\xi_t^{\varepsilon,+} - \xi_t^\varepsilon = \delta(1-t)\eta, \qquad \xi_t^{\varepsilon,-} - \xi_t^\varepsilon = -\delta(1-t)\eta \tag{4.83}$$

for $t < 1$ and

$$\xi_t^{\varepsilon,+} = \xi_t^\varepsilon = \xi_t^{\varepsilon,-}$$

for $t \geq 1$.

Now $V(x,T) \geq J_2(x,T,v^\varepsilon) - \varepsilon$ by the choice of v^ε, so that

$$\begin{aligned}
&V(T, x+\delta\eta) - 2V(t,x) + V(T, x-\delta\eta)\\
&\quad \leq J_2(x, T+\delta\eta, v^{\varepsilon,+}) - 2J_2(x,T,v^\varepsilon) + J_2(x, T-\delta\eta, v^{\varepsilon,-}) + 2\varepsilon\\
&\quad \leq \int_0^1 l(\xi_t^{\varepsilon,+}) - 2l(\xi_t^\varepsilon) + l(\xi_t^{\varepsilon,-})\,dt\\
&\qquad + \frac{1}{2}\int_0^1 |v_t^{\varepsilon,+}|^2 - 2|v_t^\varepsilon|^2 + |v_t^{\varepsilon,-}|^2\,dt + 2\varepsilon
\end{aligned}$$

which by (4.83) and (A4.3m)

$$\leq c_3\delta^2 + \frac{1}{2}\int_0^1 \left(|v_t^{\varepsilon,+}|^2 - |v_t^\varepsilon|^2\right) + \left(|v_t^{\varepsilon,-}|^2 - |v_t^\varepsilon|^2\right)dt + 2\varepsilon. \tag{4.84}$$

Considering the integral of the second difference in v^ε on the right-hand side of (4.84), (where, for simplicity of exposition, we drop the notation indicating the t-dependence), one has

$$\begin{aligned}
&\int_0^1 |v^{\varepsilon,+}|^2 - 2|v^\varepsilon|^2 + |v^{\varepsilon,-}|^2\,dt \\
&= \int_0^1 \big|Q^{-1}(\xi^{\varepsilon,+})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta + Q(\xi^\varepsilon)v^\varepsilon]\big|^2 - 2|v^\varepsilon|^2 \\
&\qquad + \big|Q^{-1}(\xi^{\varepsilon,-})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta + Q(\xi^\varepsilon)v^\varepsilon]\big|^2\,dt \\
&= \int_0^1 \big|Q^{-1}(\xi^{\varepsilon,+})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta]\big|^2 \\
&\qquad + \big|Q^{-1}(\xi^{\varepsilon,-})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta]\big|^2\,dt \\
&+ \int_0^1 2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta]^T Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+})Q(\xi^\varepsilon)v^\varepsilon \\
&\qquad + 2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta]^T Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})Q(\xi^\varepsilon)v^\varepsilon\,dt \\
&+ \int_0^1 \big|Q^{-1}(\xi^{\varepsilon,+})Q(\xi^\varepsilon)v^\varepsilon\big|^2 - 2|v^\varepsilon|^2 + \big|Q^{-1}(\xi^{\varepsilon,-})Q(\xi^\varepsilon)v^\varepsilon\big|^2\,dt. \qquad (4.85)
\end{aligned}$$

Considering the three integrals on the right-hand side of (4.85) separately, we first have, using $2a \cdot b \le |a|^2 + |b|^2$,

$$\begin{aligned}
&\int_0^1 \big|Q^{-1}(\xi^{\varepsilon,+})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta]\big|^2 \\
&\qquad + \big|Q^{-1}(\xi^{\varepsilon,-})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta]\big|^2\,dt \\
&\le 2\int_0^1 \big|Q^{-1}(\xi^{\varepsilon,+})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})]\big|^2 + \big|Q^{-1}(\xi^{\varepsilon,+})\delta\eta\big|^2 \\
&\qquad + \big|Q^{-1}(\xi^{\varepsilon,-})[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-})]\big|^2 + \big|Q^{-1}(\xi^{\varepsilon,-})\delta\eta\big|^2\,dt,
\end{aligned}$$

which using (4.72) and Assumption (A4.2m),

$$\begin{aligned}
&\le 2c_2^2 \int_0^1 \big|A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})\big|^2 + \big|A(\xi^\varepsilon) - A(\xi^{\varepsilon,-})\big|^2 + 2\delta^2|\eta|^2\,dt \\
&\le 2c_1^2c_2^2 \int_0^1 \big|\xi^\varepsilon - \xi^{\varepsilon,+}\big|^2 + \big|\xi^\varepsilon - \xi^{\varepsilon,-}\big|^2\,dt + 4c_2^2\delta^2
\end{aligned}$$

and by (4.83)

$$= 4c_1^2c_2^2 \int_0^1 \big|\delta(1-t)\eta\big|^2\,dt + 4c_2^2\delta^2 \le 4(c_1^2+1)c_2^2\delta^2. \qquad (4.86)$$

The second integral on the right-hand side of (4.85) can be bounded as follows:

$$
\begin{aligned}
&\int_0^1 2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta]^T Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+})Q(\xi^\varepsilon)v^\varepsilon \\
&\qquad +2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta]^T Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})Q(\xi^\varepsilon)v^\varepsilon \, dt \\
&= 2\int_0^1 \big\{[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})]^T Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) \\
&\qquad +[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-})]^T Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})\big\}Q(\xi^\varepsilon)v^\varepsilon \\
&\qquad -\delta\eta^T[Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) - Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})]Q(\xi^\varepsilon)v^\varepsilon \, dt \\
&= 2\int_0^1 \big\{[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})]^T [Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) - Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})] \\
&\qquad +[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) + A(\xi^\varepsilon) - A(\xi^{\varepsilon,-})]^T \\
&\qquad \cdot Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})\big\}Q(\xi^\varepsilon)v^\varepsilon \, dt \\
&\quad -2\int_0^1 \delta\eta^T[Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) - Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})]Q(\xi^\varepsilon)v^\varepsilon \, dt. \qquad (4.87)
\end{aligned}
$$

Combining Assumption (A4.2m) and (4.72) yields

$$
\big|(Q^{-T}(x)Q^{-1}(x))_x\big| \le 2c_1c_2 \qquad \forall\, x \in \mathbf{R}^n. \qquad (4.88)
$$

Therefore, the second integral on the right-hand side of (4.87) is

$$
\begin{aligned}
&\le 4c_1c_2\delta \int_0^1 |\eta|\, |\xi^{\varepsilon,+} - \xi^{\varepsilon,-}|\, |Q(\xi^\varepsilon)|\, |v^\varepsilon| \, dt \\
&\le 4c_1c_2^2\delta \int_0^1 |\eta|\, |2\delta(1-t)\eta|\, |v^\varepsilon| \, dt \le 8c_1c_2^2\widehat{M}_R\delta^2. \qquad (4.89)
\end{aligned}
$$

The first integral on the right-hand side of (4.87) becomes

$$
\begin{aligned}
&2\int_0^1 \big\{[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})]^T [Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) - Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})] \\
&\qquad -[A(\xi^{\varepsilon,+}) - 2A(\xi^\varepsilon) + A(\xi^{\varepsilon,-})]^T Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})\big\}Q(\xi^\varepsilon)v^\varepsilon \, dt \\
&\le 2\int_0^1 |A(\xi^\varepsilon) - A(\xi^{\varepsilon,+})| \\
&\qquad \cdot |Q^{-T}(\xi^{\varepsilon,+})Q^{-1}(\xi^{\varepsilon,+}) - Q^{-T}(\xi^{\varepsilon,-})Q^{-1}(\xi^{\varepsilon,-})|\, |Q(\xi^\varepsilon)|\, |v^\varepsilon| \, dt \\
&+2\int_0^1 |A(\xi^{\varepsilon,+}) - 2A(\xi^\varepsilon) + A(\xi^{\varepsilon,-})|\, |Q^{-T}(\xi^{\varepsilon,-})|\, |Q^{-1}(\xi^{\varepsilon,-})|\, |Q(\xi^\varepsilon)|\, |v^\varepsilon| \, dt.
\end{aligned}
$$

Employing (A4.2m), (4.72), and (4.88), one finds

$$
\le 4c_1^2c_2^2 \int_0^1 |\xi^\varepsilon - \xi^{\varepsilon,+}||\xi^{\varepsilon,+} - \xi^{\varepsilon,-}|\, |v^\varepsilon| + 2c_1c_2^3 \int_0^1 |\xi^{\varepsilon,+} - \xi^{\varepsilon,-}|^2 |v^\varepsilon| \, dt,
$$

which using (4.83) and Hölder's Inequality

$$\leq (4c_1^2 c_2^2 + 8c_1 c_2^3)\widehat{M}_R \delta^2. \tag{4.90}$$

Combining (4.87), (4.89) and (4.90), one finds that the second integral on the right-hand side of (4.85) satisfies

$$\begin{aligned}
&\int_0^1 2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,+}) - \delta\eta]^T Q^{-T}(\xi^{\varepsilon,+}) Q^{-1}(\xi^{\varepsilon,+}) Q(\xi^\varepsilon) v^\varepsilon \\
&\qquad +2[A(\xi^\varepsilon) - A(\xi^{\varepsilon,-}) + \delta\eta]^T Q^{-T}(\xi^{\varepsilon,-}) Q^{-1}(\xi^{\varepsilon,-}) Q(\xi^\varepsilon) v^\varepsilon \, dt \\
&\leq [8c_1 c_2^2 + 4c_1^2 c_2^2 + 8c_1 c_2^3]\widehat{M}_R \delta^2.
\end{aligned} \tag{4.91}$$

The third integral in (4.85) involves the second differences of v^ε.

$$\begin{aligned}
&\int_0^1 \left|Q^{-1}(\xi^{\varepsilon,+}) Q(\xi^\varepsilon) v^\varepsilon\right|^2 - 2|v^\varepsilon|^2 + \left|Q^{-1}(\xi^{\varepsilon,-}) Q(\xi^\varepsilon) v^\varepsilon\right|^2 dt \\
&= \int_0^1 v^{\varepsilon T} Q^T(\xi^\varepsilon) Q^{-T}(\xi^{\varepsilon,+}) Q^{-1}(\xi^{\varepsilon,+}) Q(\xi^\varepsilon) v^\varepsilon \\
&\qquad -2 v^{\varepsilon T} Q^T(\xi^\varepsilon) Q^{-T}(\xi^\varepsilon) Q^{-1}(\xi^\varepsilon) Q(\xi^\varepsilon) v^\varepsilon \\
&\qquad + v^{\varepsilon T} Q^T(\xi^\varepsilon) Q^{-T}(\xi^{\varepsilon,-}) Q^{-1}(\xi^{\varepsilon,-}) Q(\xi^\varepsilon) v^\varepsilon \, dt.
\end{aligned} \tag{4.92}$$

Consider any C^2 function $F : \mathbf{R}^n \to \mathbf{R}$ and $x \in \mathbf{R}^n$, and let $x_+ = x + \delta\eta$, $x_- = x - \delta\eta$. Denote the first and second directional derivatives in direction η by F_η and $F_{\eta\eta}$. Then

$$\begin{aligned}
F(x_+) - 2F(x) + F(x_-) &= \int_0^\delta F_\eta(x + r\eta) - F_\eta(x - \delta\eta + r\eta)\, dr \\
&= \int_0^\delta \int_0^\delta F_{\eta\eta}(x - (\delta + r)\eta + \rho\eta)\, d\rho\, dr \\
&\leq \max_{|r| \leq \delta} |F_{\eta\eta}(x + r\eta)| \delta^2.
\end{aligned} \tag{4.93}$$

Taking $F(x) = \zeta^T [Q^{-T}(x) Q^{-1}(x)] \zeta$ with $\zeta \in \mathbf{R}^n$, one has by (A4.2m) and (4.72)

$$|F_{\eta\eta}(x)| \leq |[Q^{-T}(x) Q^{-1}(x)]_{xx}|\, |\zeta|^2 \leq 2(c_1 c_2 + c_1^2)|\zeta|^2. \tag{4.94}$$

Combining (4.92) with (4.93) and (4.94) (for $x = \xi^\varepsilon$, $x_+ = \xi^{\varepsilon,+}$, $x_- = \xi^{\varepsilon,-}$ and $\zeta = Q(\xi^\varepsilon) v^\varepsilon$) yields

$$\begin{aligned}
&\int_0^1 \left|Q^{-1}(\xi^{\varepsilon,+}) Q(\xi^\varepsilon) v^\varepsilon\right|^2 - 2|v^\varepsilon|^2 + \left|Q^{-1}(\xi^{\varepsilon,-}) Q(\xi^\varepsilon) v^\varepsilon\right|^2 dt \\
&\leq \int_0^1 2(c_1 c_2 + c_1^2) |Q(\xi^\varepsilon) v^\varepsilon|^2 \delta^2 \, dt \\
&\leq 2c_2^2 c_1 (c_2 + c_1)\delta^2 \int_0^1 |v^\varepsilon|^2 \, dt \leq 2\widehat{M}_R^2 c_2^2 c_1 (c_2 + c_1)\delta^2.
\end{aligned} \tag{4.95}$$

Combining (4.86), (4.91) and (4.95) bounds the right-hand side of (4.85); therefore there exists $c_5 < \infty$ such that

$$\int_0^1 |v^{\varepsilon,+}|^2 - 2|v^\varepsilon|^2 + |v^{\varepsilon,-}|^2 \, dt \le c_5 \delta^2.$$

Substituting this into (4.84) yields

$$V(T, x - \delta\eta) - 2V(t, x) + V(T, x + \delta\eta) \le (c_3 + c_5)\delta^2 + 2\varepsilon. \qquad (4.96)$$

Because $\delta, \varepsilon \in (0, 1)$ were arbitrary, this implies semiconcavity. □

Lemma 4.31. $V(x, T) \to W_2(x)$ *uniformly on compact sets.*

Proof. Note that $V(x, T)$ is monotonically increasing in T and bounded above by $W_2(x)$. The proof that, in fact, $V(x, T) \to W_2(x)$ uniformly on compact sets is similar to that appearing in the proof of Theorem 3.20. □

Combining Theorem 4.30 and Lemma 4.31 yields the following.

Corollary 4.32. W_2 *is semiconcave.*

One can now use a semiconcave space structure analogous to the semiconvex case. First, in analogy with the max-plus algebra, one defines the min-plus algebra (commutative semifield) over $\mathbf{R}^+ = \mathbf{R} \cup \{+\infty\}$ with operations

$$a \oplus^- b \doteq \min\{a, b\}, \quad a \otimes^- b \doteq a + b.$$

In direct analogy with max-plus spaces, a variety of spaces of concave and semiconcave functions are min-plus vector spaces (or moduloids [6]). One would then proceed in a manner analogous to that in the previous sections of this chapter to obtain an approximate computation of the value function (4.71) via a min-plus eigenvector problem with eigenvalue zero. We leave the development to future researchers.

4.8.1 A Game Problem

There is a class of game problems that are closely related to control problems such as these. These game problems encompass a subclass of nonlinear H_∞ control problems (via their state-space game representations).

Consider the game problem with dynamics

$$\dot{\xi} = A(\xi) + D(\xi)u + \sigma(\xi)w \qquad (4.97)$$

with initial condition

$$\xi_0 = x, \in \mathbf{R}^n, \qquad (4.98)$$

where here $u.$ will be the control for the *minimizing* player and $w.$ will be the control for the *maximizing* player. We suppose $u \in \mathcal{U} = L_2^{loc}([0, \infty); \mathbf{R}^l)$ and

$w \in \mathcal{W} = L_2^{loc}([0,\infty); \mathbf{R}^k)$. Here D and σ are $n \times l$ and $n \times k$ matrix-valued, respectively. Consider the payoff

$$J_1(x,T,u,w) = \int_0^T l(\xi_t) + \frac{\eta^2}{2}|u_t|^2 - \frac{\gamma^2}{2}|w_t|^2\,dt. \tag{4.99}$$

We will be using the Elliott–Kalton definition of value (c.f. [8], [35], [38]). A nonanticipative strategy for the minimizing player is a mapping $\lambda : \mathcal{W} \to \mathcal{U}$ such that if $w_r = \hat{w}_r$ for almost every $r \in [0,t]$, then $\lambda_r[w] = \lambda_r[\hat{w}]$ for almost every $r \in [0,t]$, for all $t > 0$. Let the set of nonanticipative strategies for the minimizing player be denoted by Λ. A nonanticipative strategy for the maximizing player is a mapping $\theta : \mathcal{U} \to \mathcal{W}$ such that if $u_r = \hat{u}_r$ for almost every $r \in [0,t]$, then $\theta_r[u] = \theta_r[\hat{u}]$ for almost every $r \in [0,t]$, for all $t > 0$. Let the set of nonanticipative strategies for the maximizing player be denoted by Θ. The lower and upper (Elliott–Kalton) values are

$$\underline{W}_1 = \inf_{\lambda\in\Lambda}\sup_{w\in\mathcal{W}}\sup_{T\in[0,\infty)} J_1(x,T,\lambda[w],w) \tag{4.100}$$

and

$$\overline{W}_1 = \sup_{\theta\in\Theta}\inf_{u\in\mathcal{U}}\sup_{T\in[0,\infty)} J_1(x,T,u,\theta[u]).$$

(Note that the existence of these upper and lower values follows from the Soravia [106] result discussed below.) If $\underline{W}_1 = \overline{W}_1$, then the game is said to have value (in the Elliott–Kalton sense), and we say that $W_1 \doteq \underline{W}_1 = \overline{W}_1$ is the value of the game.

The DPE corresponding to a game is referred to as a Hamilton–Jacobi–Isaacs (HJI) PDE. The HJI PDE corresponding to the lower value is (c.f. [8], [38])

$$\begin{aligned}
0 &= -\max_{w\in\mathbf{R}^k}\min_{u\in\mathbf{R}^l}\Big\{[A(x) + D(x)u + \sigma(x)w]\cdot\nabla W + l(x) \\
&\qquad\qquad + \frac{\eta^2}{2}|u|^2 - \frac{\gamma^2}{2}|w|^2\Big\} &(4.101)\\
&= -A(x)\cdot\nabla W - l(x) - \min_{u\in\mathbf{R}^l}\left\{(D(x)u)\cdot\nabla W + \frac{\eta^2}{2}|u|^2\right\} \\
&\quad - \max_{w\in\mathbf{R}^k}\left\{(\sigma(x)w)\cdot\nabla W - \frac{\gamma^2}{2}|w|^2\right\} &(4.102)\\
&= -\Bigg[A(x)\cdot\nabla W + l(x) \\
&\qquad -(\nabla W)^T\left(\frac{D(x)D^T(x)}{2\eta^2} - \frac{\sigma(x)\sigma^T(x)}{2\gamma^2}\right)\nabla W\Bigg] &(4.103)\\
&= -\min_{u\in\mathbf{R}^l}\max_{w\in\mathbf{R}^k}\Big\{[A(x) + D(x)u + \sigma(x)w]\cdot\nabla W + l(x)
\end{aligned}$$

$$+\frac{\eta^2}{2}|u|^2 - \frac{\gamma^2}{2}|w|^2\Big\}$$

with boundary condition $W(0) = 0$. Note that the right-hand side for the last equation corresponds to the HJI PDE for the upper value. In this case, we say that the Isaacs condition is satisfied, and one generally expects the game to have value, the value being the unique (correct, as discussed in Chapter 3) viscosity solution of the HJI PDE.

We consider only the case where the controller dominates the disturbance in the sense that

$$\frac{QQ^T}{2} := \frac{DD^T}{2\eta^2} - \frac{\sigma\sigma^T}{2\gamma^2}$$

is uniformly bounded and nondegenerate in the sense of Assumption (A4.1m). In this case we see that the HJI PDE for this game, i.e., (4.103), coincides with the HJB PDE (4.77) for the control problem (4.68)–(4.71).

Now note that, by Soravia [106], one also has that the game value function, W_1, and the control value function, W_2, are both minimal, non-negative, continuous viscosity supersolutions of (4.103)(equivalently (4.77)). Consequently, $W_1 \equiv W_2$, and of course then by Theorem 4.28 and Corollary 4.32, $\underline{W}_1$ is a semiconcave viscosity solution of (4.103). Therefore, the min-plus algorithm, outlined above, to compute control problem value W_2 also obtains the value of the game problem, W_1.

4.9 An Example

As an example, we consider the two-dimensional min-plus problem with $\xi = (\xi_1, \xi_2)^T$ and dynamics

$$\frac{d}{dt}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} = \begin{pmatrix}\xi_2\\ (3/4)\arctan(2\xi_1)\end{pmatrix} + \begin{pmatrix}0\\ 1\end{pmatrix}u + \begin{pmatrix}0\\ 1\end{pmatrix}w. \tag{4.104}$$

Let γ, η be such that the reduction of the previous section yields a matrix in the quadratic term in the gradient given by

$$\frac{QQ^T}{2} = \begin{bmatrix}0 & 0\\ 0 & 1/2\end{bmatrix}.$$

Note that this example was chosen so as to represent a second-order system of the form $\ddot{y} = (3/4)\arctan(2y) + u + w$. The running cost was simply $l(x) = x_1^2 + x_2^2$. One might note that the assumption of uniform nondegeneracy, (A4.1m), is violated in this example (although one still has controllability). The corresponding PDE is

$$0 = -\left\{x_2 W_{x_1} + \tfrac{3}{4}\arctan(2x_1)W_{x_2} + \tfrac{1}{2}(x_1^2 + x_2^2) - \tfrac{1}{2}W_{x_2}^2\right\}$$

with usual boundary condition $W(0) = 0$.

The computations were run at a coarseness level that allowed computation in less than 10 seconds on a 2000 Sun Ultra 10. The left-hand plot in Figure 4.1 depicts W. The right-hand plot in Figure 4.1 and the left-hand plot in Figure 4.2 depict the partials of W; note the nonsmoothness. The right-hand plot in Figure 4.2 depicts the approximate backsubstitution error — computed by taking first-order differences on the grid to approximate ∇W, and then substituting this back into the PDE. The resulting error values were scaled by dividing by $1 + |x|^2$.

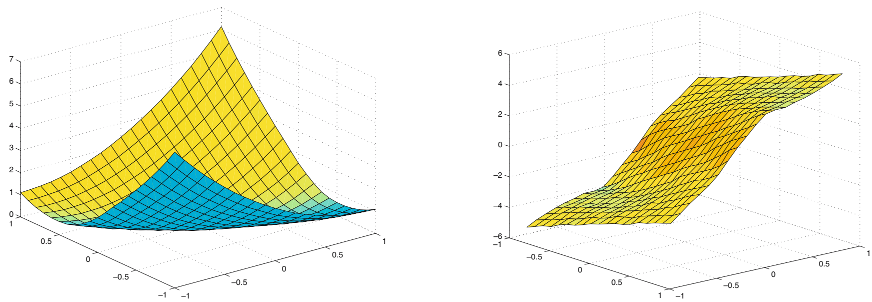


Fig. 4.1. Value and partial with respect to x_1, 2-D problem

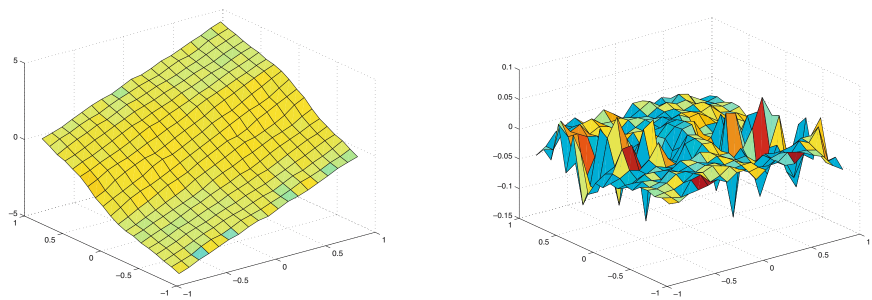


Fig. 4.2. Partial with respect to x_2 and scaled backsubstitution error, 2-D problem

5

Max-Plus Eigenvector Method Error Analysis

In the preceding chapter, the basics of a max-plus eigenvector method for an infinite time-horizon control problem were laid out. In that development, we assumed the extremely "unlikely" case that the value function had a finite max-plus expansion. Proceeding in that way allowed us to analyze the power method for solution of the eigenvector without great delay. We now return to the problem, and remove this assumption. The mechanics developed in the previous chapter will continue to be used for solution of the finite-dimensional eigenvector problem. However, we will now analyze the errors induced by the truncation of the infinite-dimensional max-plus eigenvector (eigenfunction) to a finite number of terms. In particular, the truncation will be achieved by taking the centers, x_i, of the basis functions, $\psi_i = -\frac{1}{2}(x - x_i)^T C(x - x_i)$, on a uniform grid over the ball of radius D_R (see Theorems 2.11 and 2.13). At the same time, we will consider the errors induced by the fact that we cannot compute the elements of the (finite) B matrix exactly, but only approximately. We will find that the convergence of the approximate solution to the exact solution will depend on the relative rates at which the distance between the basis function centers and the errors in the computation of the elements of the corresponding B matrices go to zero.

The problem dynamics continue to be given by (3.12) with initial condition (3.13). The payoff and value, W, continue to be (3.21) and (4.4). Consequently, the HJB PDE and boundary condition remain (4.7) and (4.8). We retain assumptions (A4.1I)–(A4.4I). The value function is the unique, "correct" solution of the eigenvector problem $0 \otimes W = S_\tau[W]$ and/or HJB PDE in the appropriate class (as given in Theorems 4.1 and 4.4).

Let c_R be the semiconvexity constant discussed in Section 4.3. In this chapter, we modify Assumption (A4.5I) to allow matrix semiconvexity coefficients (see Remarks 2.10 and 2.14).

> We assume that one may choose symmetric matrix C such that $C - c_R I > 0$ and such that $S_\tau[\psi_i] \in \mathcal{S}_{\overline{R}}^{C'L'}$ for all i where $C - C' > 0$, $|C||C^{-1}|R + |C^{-1}|L' \le D_R$ and $\psi_i(x) = -\frac{1}{2}(x - x_i)^T C(x - x_i)$. (A5.5I)

This assures us that each $S_\tau[\psi_i]$ has a max-plus basis expansion in terms of the basis $\{\psi_j\}$. We remark that in the linear/quadratic case with $f(x) = Ax$, σ constant, and $l(x) = \frac{1}{2}x^T Dx$ with A, σ, D, γ such that Assumptions (A4.1I)–(A4.4I) are met, condition (A5.5I) is satisfied. This follows from the fact that matrix expression $D - A^T C - CA + \frac{1}{2\gamma^2} C^T \sigma\sigma^T C$ is positive definite for $C = cI$ with sufficiently large c. We note that this assumption will need to be replaced by a slightly stricter assumption (A5.6I) in Section 5.2 for the results there and beyond.

Sections 5.1 to 5.3 will provide development of the error estimates. The details will be somewhat technical. For those desiring a preview, we note that a summary of these results appears in Section 5.4.

5.1 Allowable Errors in Computation of B

In this section, we obtain a bound on the maximum allowable errors in the computation of B. If the errors are below this bound, then we can guarantee convergence of the power method to the unique eigenvector. In particular, the guaranteed convergence of the power method relies on Lemma 4.20 and Theorem 4.21 because these imply a certain structure to a directed graph associated with B (see [78], [85], and for a more general discussion, [6]). If there was a sequence $\{k_i\}_{i=1}^{i=N+1}$ such that $1 \le k_i \le \nu$ for all i and $k_{N+1} = k_1$ such that one does not have $k_i = 1$ for all i, and such that

$$\sum_{i=1}^{N} B_{k_i, k_{i+1}} \ge 0$$

then there would be no guarantee of convergence of the power method (nor the ensuing uniqueness result for that matter). In order to determine more exactly, the allowable errors in the computation of the elements of B, we first need to obtain a more exact expression for the δ that appears in Lemma 4.20 and Theorem 4.21, and this will appear in Theorems 5.4 and 5.5. That will be followed by results indicating the allowable error bounds. To begin, one needs the following lemma.

Lemma 5.1. *Let ξ satisfy (3.12) with initial state $\xi_0 = x \in \mathbf{R}^n$. Let $K, \tau \in (0, \infty)$, and let $u \in L_2[0, \tau]$. Suppose $\delta > 0$ sufficiently small so that*

$$\delta \le \frac{K m_\sigma^2}{c_f(1 - e^{-c_f \tau})} \tag{5.1}$$

where c_f, m_σ are given in Assumptions (A4.1I), (A4.2I). Then

$$K|\xi_\tau - x|^2 + \delta \|u\|^2_{L_2[0,\tau]} \ge \frac{\delta c_f}{8 m_\sigma^2} |x|^2 \left(1 - e^{-c_f \tau}\right)^4.$$

Remark 5.2. It may be of interest to note that the assumption on the size of δ may not be necessary. At one point in the proof to follow, this assumption is used in order to eliminate a case that would lead to a more complex expression on the right-hand side in the result in the lemma statement. If some later technique benefited from not having such an assumption, the lemma proof could be revisited.

Remark 5.3. It is perhaps also worth indicating the intuition behind the inequality obtained in Lemma 5.1. Essentially, it states that, due to the nature of the dynamics of the system, the only way that $|\xi_\tau - x|^2$ can be kept small is through input disturbance energy $\|u\|^2$, and so their weighted sum is bounded from below. The dependence on $|x|$ on the right-hand side is indicative of the fact that $|f(x)|$ goes to zero at the origin.

Proof. Note that by (3.12) and Assumptions (A4.1I) and (A4.2I),

$$\frac{d}{dt}|\xi|^2 \le -2c_f|\xi|^2 + 2m_\sigma|\xi||u| \tag{5.2}$$

$$\le -c_f|\xi|^2 + \frac{m_\sigma^2}{c_f}|u|^2. \tag{5.3}$$

Consequently, for any $t \in [0,\tau]$,

$$|\xi_t|^2 \le e^{-c_f t}|x|^2 + \frac{m_\sigma^2}{c_f}\int_0^t |u_r|^2\,dr$$

and so

$$\|u\|^2_{L_2(0,t)} \ge \frac{c_f}{m_\sigma^2}\left[|\xi_t|^2 - |x|^2\right] \quad \forall t \in [0,\tau]. \tag{5.4}$$

We may suppose

$$|\xi_t| \le \sqrt{1 + (1-e^{-c_f\tau})^4/2}\,|x| \quad \forall t \in [0,\tau]. \tag{5.5}$$

Otherwise by (5.4) and the reverse of (5.5), there exists $t \in [0,\tau]$ such that

$$K|\xi_\tau - x|^2 + \delta\|u\|^2_{L_2[0,\tau]} \ge \delta\|u\|^2_{L_2[0,t]} \ge \frac{\delta c_f}{2m_\sigma^2}(1-e^{-c_f\tau})^4|x|^2 \tag{5.6}$$

in which case one already has the desired result.

Define $\overline{K} \doteq \sqrt{1+(1-e^{-c_f\tau})^4/2}$. Recalling (5.2), and applying (5.5), one has

$$\frac{d}{dt}|\xi_t|^2 \le -2c_f|\xi_t|^2 + 2m_\sigma\overline{K}|x||u_t|.$$

Solving this ODI for $|\xi_t|^2$, and using the Hölder inequality, yields the bound

$$|\xi_\tau|^2 \le |x|^2 e^{-2c_f\tau} + \frac{m_\sigma\overline{K}|x|\|u\|}{\sqrt{c_f}}\left(1-e^{-4c_f\tau}\right)^{1/2}. \tag{5.7}$$

This implies

$$|\xi_\tau| \leq |x|e^{-c_f\tau} + \frac{1}{c_f^{1/4}}\sqrt{m_\sigma \overline{K}|x|\|u\|}\left(1 - e^{-4c_f\tau}\right)^{1/4}. \tag{5.8}$$

We consider two cases separately. First we consider the case where $|\xi_\tau| \leq |x|$. Then, by (5.8)

$$\begin{aligned} |\xi_\tau - x| &\geq |x| - |\xi_\tau| \\ &\geq |x|(1 - e^{-c_f\tau}) - \frac{1}{c_f^{1/4}}\sqrt{m_\sigma \overline{K}|x|\|u\|}\left(1 - e^{-4c_f\tau}\right)^{1/4}. \end{aligned} \tag{5.9}$$

Now note that for general $a, b, c \in [0, \infty)$, $a + c \geq b$ implies

$$a^2 \geq \frac{b^2}{2} - c^2. \tag{5.10}$$

By (5.9) and (5.10) (and noting the non-negativity of the norm),

$$|\xi_\tau - x|^2 \geq \max\left\{\frac{1}{2}|x|^2(1 - e^{-c_f\tau})^2 - \frac{m_\sigma \overline{K}}{\sqrt{c_f}}|x|\|u\|\left(1 - e^{-4c_f\tau}\right)^{1/2}, 0\right\}$$

which implies

$$\begin{aligned} &K|\xi_\tau - x|^2 + \delta\|u\|^2 \\ &\geq \max\left\{\frac{K}{2}|x|^2(1 - e^{-c_f\tau})^2 - \frac{K m_\sigma \overline{K}}{\sqrt{c_f}}|x|\|u\|\left(1 - e^{-4c_f\tau}\right)^{1/2} + \delta\|u\|^2, \right. \\ &\qquad\qquad \left. \delta\|u\|^2\right\}. \end{aligned} \tag{5.11}$$

The right-hand side of (5.11) is a maximum of two convex quadratic functions of $\|u\|$. The second is monotonically increasing, while the first is positive at $\|u\| = 0$ and initially decreasing. This implies that there are two possibilities for the location of the minimum of the maximum of the two functions. If the minimum of the first function is to the left of the point where the two functions intersect, then the minimum occurs at the minimum of the first function; alternatively it occurs where the two functions intersect. The minimum of the first function occurs at $\|u\|_{\min}$ (where we are abusing notation here, using the *min* subscript on the norm to indicate the value of $\|u\|$ at which the minimum occurs), and this is given by

$$\|u\|_{\min} = \frac{K m_\sigma \overline{K}|x|(1 - e^{-4c_f\tau})^{1/2}}{2\sqrt{c_f}\delta}. \tag{5.12}$$

The point of intersection of the two functions occurs at

$$\|u\|_{\text{int}} = \frac{\sqrt{c_f}|x|(1-e^{-c_f\tau})^2}{2m_\sigma \overline{K}(1-e^{-4c_f\tau})^{1/2}}. \tag{5.13}$$

The two points coincide when

$$\delta = \frac{Km_\sigma^2\overline{K}^2(1-e^{-4c_f\tau})}{c_f(1-e^{-c_f\tau})^2} = \frac{Km_\sigma^2[1+(1-e^{-c_f\tau})^4/2](1-e^{-4c_f\tau})}{c_f(1-e^{-c_f\tau})^2},$$

and $\|u\|_{\text{int}}$ occurs to the left of $\|u\|_{\text{min}}$ for δ less than this. It is easy to see that assumption (5.1) implies that δ is less than the value at which the points coincide, and consequently, the minimum of the right-hand side of (5.11) occurs at $\|u\|_{\text{int}}$.

Using the value of the right-hand side of (5.11) corresponding to $\|u\|_{\text{int}}$, we find that for any disturbance, u,

$$K|\xi_\tau - x|^2 + \delta\|u\|^2 \geq \frac{\delta c_f|x|^2}{4m_\sigma^2\overline{K}^2}\frac{(1-e^{-c_f\tau})^4}{(1-e^{-4c_f\tau})},$$

which, using definition of $\overline{K}$

$$\begin{aligned} , &= \frac{\delta c_f|x|^2}{4m_\sigma^2}\frac{(1-e^{-c_f\tau})^4}{(1-e^{-4c_f\tau})[1+(1-e^{-c_f\tau})^4/2]} \\ &\geq \frac{\delta c_f|x|^2}{8m_\sigma^2}(1-e^{-c_f\tau})^4. \end{aligned} \tag{5.14}$$

Now we turn to the second case,

$$|\xi_\tau| > |x|. \tag{5.15}$$

In this case, (5.15) and (5.8) yield

$$|x|e^{-c_f\tau} + \frac{1}{c_f^{1/4}}\sqrt{m_\sigma\overline{K}|x|\|u\|}\left(1-e^{-4c_f\tau}\right)^{1/4} > |x|. \tag{5.16}$$

Upon rearrangement, (5.16) yields

$$\|u\| > \frac{\sqrt{c_f}|x|}{m_\sigma\overline{K}}\frac{(1-e^{-c_f\tau})^2}{(1-e^{-4c_f\tau})^{1/2}}.$$

Consequently, using the definition of $\overline{K}$ and some simple manipulations,

$$\begin{aligned} K|\xi_\tau - x|^2 + \delta\|u\|^2 &\geq \frac{\delta c_f|x|^2(1-e^{-c_f\tau})^4}{m_\sigma^2(1-e^{-4c_f\tau})[1+(1-e^{-c_f\tau})^4/2]} \\ &\geq \frac{\delta c_f|x|^2}{2m_\sigma^2}(1-e^{-c_f\tau})^4. \end{aligned} \tag{5.17}$$

Combining (5.14) and (5.17) completes the proof. □

Now we turn to how Lemma 5.1 can be used to obtain a more detailed replacement for the δ that appears in 4.20 and Theorem 4.21. Fix $\tau > 0$. Let

$$\widehat{\gamma}_0^2 \in \left(\frac{2m_\sigma^2 \alpha_l}{c_f^2}, \gamma^2 \right), \tag{5.18}$$

and in particular, let $\widehat{\gamma}_0^2 = \gamma^2 - \delta$ where δ is sufficiently small so that

$$\delta < \gamma^2 - \frac{2m_\sigma^2 \alpha_l}{c_f^2}. \tag{5.19}$$

Then all results of Chapter 4 for W hold with γ^2 replaced by $\widehat{\gamma}_0^2$, and we denote the corresponding value by $W^{\widehat{\gamma}_0}$. In particular, by Theorem 2.13, for any $R < \infty$ there exists semiconvexity constant $c_R^0 < \infty$ for $W^{\widehat{\gamma}_0}$ over $\overline{B}_R$, and a Lipschitz constant, L_R^0 also for $W^{\widehat{\gamma}_0}$ over $\overline{B}_R$,. Note that the required constants satisfy $c_R^0 < c_R$ (see proof of Theorem 2.13). If $L_R^0 > L_R$ sufficiently so that $|C||C^{-1}|R + |C^{-1}|L_R^0 > D_R$, we modify our basis to be dense over $\overline{B}_{D_R^0}$ where $D_R^0 \geq |C||C^{-1}|R + |C^{-1}|L_R^0$ (and redefine $D_R \doteq D_R^0$ in that case). Then, as before, the set $\{\psi_i\}$ forms a max–plus basis for the space of semiconvex functions over $\overline{B}_R$ with semiconvexity constant, c_R^0, i.e., $\mathcal{S}_{\bar{R}}^{c_R^0, L_R^0}$.

For any j, let

$$\overline{x}_j \in \operatorname*{argmax}_{|x| \leq R} \{\psi_j(x) - W^{\hat{\gamma}_0}(x)\}. \tag{5.20}$$

Then for any $x \in \overline{B}_R$,

$$\psi_j(x) - \psi_j(\overline{x}_j) \leq W^{\hat{\gamma}_0}(x) - W^{\hat{\gamma}_0}(\overline{x}_j) - K_0|x - \overline{x}_j|^2, \tag{5.21}$$

where $2K_0 > 0$ is the minimum eigenvalue of $C - c_R^0 I > 0$. Note that K_0 depends on $\widehat{\gamma}_0$.

Theorem 5.4. *Let $\widehat{\gamma}_0$ satisfy (5.18). Let $K = K_0$ satisfy (5.21) (where we may take $K_0 > 0$ to be the minimum eigenvalue of $C - c_R^0 I > 0$ if desired). Let $\delta > 0$ satisfy $\delta \leq \frac{\gamma^2}{2} - \frac{\widehat{\gamma}_0^2}{2}$ and (5.1). Then, for any $j \neq 1$,*

$$B_{j,j} \leq \frac{-\delta c_f |\overline{x}_j|^2}{8m_\sigma^2} \left(1 - e^{-c_f \tau}\right)^4.$$

(Recall that by the choice of ψ_1 as the basis function centered at the origin, $B_{1,1} = 0$; see Lemma 4.20.)

Proof. Let K_0, τ, δ satisfy the assumptions (i.e., (5.1), (5.19), (5.21)). Then

$$S_\tau[\psi_j](\overline{x}_j) - \psi_j(\overline{x}_j) = \sup_{u \in L_2} \left\{ \int_0^\tau l(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + \psi_j(\xi_\tau) - \psi_j(\overline{x}_j) \right\}, \tag{5.22}$$

where ξ satisfies (3.12) with $\xi_0 = \overline{x}_j$. Let $\varepsilon > 0$, and u^ε be ε-optimal . Then this implies

$$S_\tau[\psi_j](\overline{x}_j) - \psi_j(\overline{x}_j) \le \int_0^\tau l(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + \psi_j(\xi_\tau^\varepsilon) - \psi_j(\overline{x}_j) + \varepsilon,$$

and by (5.21) and the definition of $\widehat{\gamma}_0$

$$\le \int_0^\tau l(\xi_t^\varepsilon) - \frac{\widehat{\gamma}_0^2}{2}|u_t^\varepsilon|^2 - \delta|u_t^\varepsilon|^2\,dt + W^{\hat{\gamma}_0}(\xi_\tau^\varepsilon) - W^{\hat{\gamma}_0}(\overline{x}_j)$$
$$-K_0|\xi_\tau^\varepsilon - \overline{x}_j|^2 + \varepsilon$$

and by Theorem 4.4 (for $W^{\hat{\gamma}_0}$),

$$\le -\delta\|u^\varepsilon\|^2 - K_0|\xi_\tau^\varepsilon - \overline{x}_j|^2 + \varepsilon.$$

Combining this with Lemma 5.1 yields

$$S_\tau[\psi_j](\overline{x}_j) - \psi_j(\overline{x}_j) \le \frac{-\delta c_f|\overline{x}_j|^2}{8m_\sigma^2}\left(1 - e^{-c_f\tau}\right)^4 + \varepsilon.$$

Because this is true for all $\varepsilon > 0$, one has

$$S_\tau[\psi_j](\overline{x}_j) - \psi_j(\overline{x}_j) \le \frac{-\delta c_f|\overline{x}_j|^2}{8m_\sigma^2}\left(1 - e^{-c_f\tau}\right)^4. \tag{5.23}$$

But

$$B_{j,j} = \min_{|x|\le R}\{S_\tau[\psi_j](x) - \psi_j(x)\} \tag{5.24}$$

which by (5.23)

$$\le \frac{-\delta c_f|\overline{x}_j|^2}{8m_\sigma^2}\left(1 - e^{-c_f\tau}\right)^4. \quad \square$$

Theorem 5.5. *Let $\widehat{\gamma}_0$ satisfy (5.18). Let K_0 be as in (5.21), and let $\delta > 0$ be given by*

$$\delta = \min\left\{\frac{K_0 m_\sigma^2}{c_f}, \frac{\gamma^2}{2} - \frac{\widehat{\gamma}_0^2}{2}\right\} \tag{5.25}$$

(which is somewhat tighter than the requirement in the previous theorem). Let $N \in \mathcal{N}$, $\{k_i\}_{i=1}^{i=N+1}$ such that $1 \le k_i \le \nu$ for all i and $k_{N+1} = k_1$. Suppose we are not in the case $k_i = 1$ for all i. Then

$$\sum_{i=1}^N B_{k_i,k_{i+1}} \le -\max_{k_i}|\overline{x}_{k_i}|^2\frac{\delta c_f}{8m_\sigma^2}\left(1 - e^{-c_f N\tau}\right)^4.$$

Proof. By Theorem 5.4, this is true for $N = 1$. We prove the case $N = 2$. The proof of the general case will then be obvious. First note the monotonicity of the semigroup in the sense that if $g_1(x) \le g_2(x)$ for all x, then

$$S_\tau[g_1](x) \le S_\tau[g_2](x) \quad \forall x \in \mathbf{R}^n. \tag{5.26}$$

Suppose either $i \ne 1$ or $j \ne 1$. By definition, $\psi_j(x) + B_{j,i} \le S_\tau[\psi_i](x)$ for all $x \in \mathbf{R}^n$. Using (5.26) and the max-plus linearity of the semigroup yields

$$S_\tau[\psi_j](x) + B_{j,i} \le S_{2\tau}[\psi_i](x) \quad \forall x,$$

which implies in particular that

$$S_\tau[\psi_j](\overline{x}_i) + B_{j,i} \le S_{2\tau}[\psi_i](\overline{x}_i). \tag{5.27}$$

Now, employing the same proof as that of Theorem 5.4, but with τ replaced by 2τ (noting that condition (5.1) is satisfied with 2τ replacing τ by our assumption (5.25)), one has as in (5.23)

$$S_{2\tau}[\psi_i](\overline{x}_i) - \psi_i(\overline{x}_i) \le \frac{-\delta c_f |\overline{x}_i|^2}{8m_\sigma^2}\left(1 - e^{-2c_f\tau}\right)^4. \tag{5.28}$$

Combining (5.27) and (5.28) yields

$$\Big[S_\tau[\psi_j](\overline{x}_i) - \psi_i(\overline{x}_i)\Big] + B_{j,i} \le \frac{-\delta c_f |\overline{x}_i|^2}{8m_\sigma^2}\left(1 - e^{-2c_f\tau}\right)^4.$$

Using the definition of $B_{i,j}$, this implies

$$B_{i,j} + B_{j,i} \le \frac{-\delta c_f |\overline{x}_i|^2}{8m_\sigma^2}\left(1 - e^{-2c_f\tau}\right)^4. \tag{5.29}$$

By symmetry, one also has

$$B_{i,j} + B_{j,i} \le \frac{-\delta c_f |\overline{x}_j|^2}{8m_\sigma^2}\left(1 - e^{-2c_f\tau}\right)^4. \tag{5.30}$$

Combining (5.29) and (5.30) yields

$$B_{i,j} + B_{j,i} \le -\max\Big\{|\overline{x}_i|^2, |\overline{x}_j|^2\Big\}\frac{\delta c_f}{8m_\sigma^2}\left(1 - e^{-2c_f\tau}\right)^4. \quad \square$$

The convergence of the power method (see Chapter 4) relied on a certain structure of B ($B_{1,1} = 0$ and strictly negative loop sums as described in the assumptions of Theorem 4.21). Combining this with the above result on the size of loop sums, one can obtain a condition that guarantees convergence of the power method to a unique eigenvector corresponding to eigenvalue zero. This is given in the next theorem.

Theorem 5.6. *Let B be given by $B_{j,i} = -\max_{x \in \overline{B}_R}\{\psi_j(x) - S_\tau[\psi_i](x)\}$ for all $i, j \le \nu$, and let $\widetilde{B}$ be an approximation of B with $\widetilde{B}_{1,1} = 0$ and such that there exists $\varepsilon > 0$ such that for all $(i, j) \ne (1, 1)$,*

$$|\widetilde{B}_{i,j} - B_{i,j}| \leq \max\{|\overline{x}_i|^2, |\overline{x}_j|^2\} \left(\frac{\delta c_f}{8m_\sigma^2}\right) \frac{\left(1 - e^{-c_f \tau}\right)^4}{\nu^2} - \varepsilon, \qquad (5.31)$$

where

$$\delta = \min\left\{\frac{K_0 m_\sigma^2}{c_f}, \frac{\gamma^2}{2} - \frac{\widehat{\gamma}_0^2}{2}\right\}. \qquad (5.32)$$

Then the power method applied to $\widetilde{B}$ *converges in a finite number of steps to the unique eigenvector* $\widetilde{e}$ *corresponding to eigenvalue zero, that is*

$$\widetilde{e} = \widetilde{B} \otimes \widetilde{e}.$$

Proof. Let $N \in \mathcal{N}$, and consider a sequence of nodes $\{k_i\}_{i=1}^{N+1}$ with $k_1 = k_{N+1}$. We must show that if we are not in the case $k_i = 1$ for all i, then

$$\sum_{i=1}^{N} \widetilde{B}_{k_i,k_{i+1}} < 0.$$

Suppose $N > \nu^2$. Then any sequence $\{k_i\}_{i=1}^{N+1}$ with $k_1 = k_{N+1}$ must be composed of subloops of length no greater than ν^2. Therefore, it is sufficient to prove the result for $N \leq \nu^2$. Note that by the assumptions and Theorem 5.5,

$$\begin{aligned}\sum_{i=1}^{N} \widetilde{B}_{k_i,k_{i+1}} &\leq \sum_{i=1}^{N} B_{k_i,k_{i+1}} + \sum_{i=1}^{N} |\widetilde{B}_{k_i,k_{i+1}} - B_{k_i,k_{i+1}}| \\ &\leq -\max_{k_i} |\overline{x}_{k_i}|^2 \frac{\delta c_f}{8m_\sigma^2} \left(1 - e^{-c_f N\tau}\right)^4 \\ &\quad + \max_{k_i} |\overline{x}_{k_i}|^2 \frac{\delta c_f}{8m_\sigma^2} \left(1 - e^{-c_f N\tau}\right)^4 (N/\nu^2) - \varepsilon \\ &\leq -\varepsilon.\end{aligned}$$

Then by the same proofs as for Theorem 4.22, and Corollary 4.23, the result follows. □

Theorem 5.6 will be useful later when we analyze the size of errors introduced by our computational approximation to the elements of B.

If the conditions of Theorem 5.6 are met, then one can ask what the size of the errors in the corresponding eigenvector are. Specifically, if eigenvector $\tilde{e}$ is computed using approximation $\widetilde{B}$, what is a bound on the size of the difference between e (the eigenvector of B) and $\tilde{e}$? The following theorem gives a rough, but easily obtained, bound.

Theorem 5.7. *Let* B *be given by* $B_{i,j} = -\max_{x \in \overline{B}_R} \big(\psi_j(x) - S_\tau[\psi_i](x)\big)$ *for all* $i, j \leq \nu$*, and let* $\widetilde{B}$ *be an approximation of* B *with* $\widetilde{B}_{1,1} = 0$ *and such that there exists* $\varepsilon > 0$ *such that*

$$|\widetilde{B}_{i,j} - B_{i,j}| \leq \max\{|\overline{x}_i|^2, |\overline{x}_j|^2\} \left(\frac{\delta c_f}{8m_\sigma^2}\right) \frac{\left(1-e^{-c_f\tau}\right)^4}{\nu^\mu} - \varepsilon \quad \forall i,j \tag{5.33}$$

where $\mu \in \{2,3,4,\ldots\}$ *and* δ *is given by (5.32). Then the power method will yield the unique eigenvectors* e *and* $\tilde{e}$ *of* B *and* $\widetilde{B}$*, respectively, in finite numbers of steps, and*

$$\|e-\tilde{e}\| \doteq \max_i |e_i - \tilde{e}_i| \leq \left(D_R\right)^2 \left(\frac{\delta c_f}{8m_\sigma^2}\right) \frac{\left(1-e^{-c_f\tau}\right)^4}{\nu^{\mu-2}} - \varepsilon.$$

Proof. By Theorem 5.6, one may use the power method to compute $\tilde{e}$, and so one has that for any $j \leq \nu^2$,

$$\tilde{e}_j = \left[\widetilde{B}^{\nu^2} \otimes 0\right]_j = \max_{m\leq\nu^2} \left[\widetilde{B}^m \otimes 0\right]_j = \max_{m\leq\nu^2} \max_{\{k_l\}_{l=1}^m,\, k_1=j} \sum_{l=1}^m \widetilde{B}_{k_l,k_{l+1}}$$

where the exponents on $\widetilde{B}$ represent max-plus exponentiation and the bound $m \leq \nu^2$ follows from the fact that under the assumption, the sum around any loop other than that of the trivial loop, $\widetilde{B}_{1,1} = 0$, are strictly negative. Therefore,

$$\tilde{e}_j \leq \max_{m\leq\nu^2} \max_{\{k_l\}_{l=1}^m,\, k_1=j} \left[\sum_{l=1}^m \left|\widetilde{B}_{k_l,k_{l+1}} - B_{k_l,k_{l+1}}\right| + \sum_{l=1}^m B_{k_l,k_{l+1}}\right]$$

which by the assumption (5.33) and the fact that e is the eigenvector of B,

$$\leq \left(D_R\right)^2 \left(\frac{\delta c_f}{8m_\sigma^2}\right) \frac{\left(1-e^{-c_f\tau}\right)^4}{\nu^{\mu-2}} - \varepsilon + e_j.$$

Using symmetry, one obtains

$$|\tilde{e}_j - e_j| \leq \left(D_R\right)^2 \left(\frac{\delta c_f}{8m_\sigma^2}\right) \frac{\left(1-e^{-c_f\tau}\right)^4}{\nu^{\mu-2}} - \varepsilon. \quad \square$$

We remark that by taking ε sufficiently small, and noting that $1-e^{-c_f\tau} \leq c_f\tau$ for non-negative τ, Theorem 5.7 implies (under its assumptions)

$$\|e-\tilde{e}\| = \max_i |e_i - \tilde{e}_i| \leq \left(D_R\right)^2 \left(\frac{\delta c_f^5}{8m_\sigma^2}\right) \frac{\tau^4}{\nu^{\mu-2}}. \tag{5.34}$$

Also note that aside from the case $i=j=1$ (recall $B_{1,1}=0$), one has

$$\min_{i\neq 1}\{|\overline{x}_i|^2\} \leq \max\{|\overline{x}_i|^2, |\overline{x}_j|^2\} \quad \forall i,j.$$

Using this, and choosing $\varepsilon > 0$ appropriately, one has the following theorem (where we note the condition on the errors in B is uniform but potentially significantly stricter). The proof is nearly identical to that for Theorem 5.7

Theorem 5.8. *Let B be as in Theorem 5.7, and let $\widetilde{B}$ be an approximation of B with $\widetilde{B}_{1,1} = 0$ and such that*

$$|\widetilde{B}_{i,j} - B_{i,j}| \le \min_{i\neq 1}\{|\overline{x}_i|^2\} \left(\frac{\delta c_f}{9m_\sigma^2}\right) \frac{\left(1 - e^{-c_f\tau}\right)^4}{\nu^\mu} \quad \forall\, i,j, \tag{5.35}$$

where $\mu \in \{2,3,4,\dots\}$ and δ is given by (5.32). Then the power method will yield the unique eigenvectors e and $\tilde{e}$ of B and $\widetilde{B}$, respectively, in finite numbers of steps, and

$$\|e - \tilde{e}\| \le \min_{i\neq 1}\{|\overline{x}_i|^2\} \left(\frac{\delta c_f}{9m_\sigma^2}\right) \frac{\left(1 - e^{-c_f\tau}\right)^4}{\nu^{\mu-2}}.$$

A simpler variant on this result may be worth using. Note that for $\tau \in [0, 1/c_f]$, one has $1 - e^{-c_f\tau} \ge (c_f/2)\tau$. Then by a proof again nearly identical to that of Theorem 5.7, one has:

Theorem 5.9. *Suppose $\tau \le 1/c_f$. Let B be as in Theorem 5.7, and let $\widetilde{B}$ be an approximation of B with $\widetilde{B}_{1,1} = 0$ and such that*

$$|\widetilde{B}_{i,j} - B_{i,j}| \le \min_{i\neq 1}\{|\overline{x}_i|^2\} \left(\frac{\delta c_f^5}{9(16)m_\sigma^2}\right) \frac{\tau^4}{\nu^\mu} \quad \forall\, i,j \tag{5.36}$$

where $\mu \in \{2,3,4,\dots\}$ and δ is given by (5.32). Then the power method will yield the unique eigenvectors e and $\tilde{e}$ of B and $\widetilde{B}$, respectively, in finite numbers of steps, and

$$\|e - \tilde{e}\| \le \min_{i\neq 1}\{|\overline{x}_i|^2\} \left(\frac{\delta c_f^5}{9(16)m_\sigma^2}\right) \frac{\tau^4}{\nu^{\mu-2}}.$$

This variant is included because the simpler right-hand sides might simplify analysis.

5.2 Convergence and Truncation Errors

In this section we consider the approximation due to using only a finite number of functions in the max-plus basis expansion. It will be shown that as the number of functions increases (in a reasonable way), the approximate solution obtained by the eigenvector computation of Chapter 4 converges from below to the value function, W. Error bounds will also be obtained. These error bounds will be useful in the error summary to follow in Section 5.4.

5.2.1 Convergence

This subsection contains a quick proof that the errors due to truncation of the basis go to zero as the number of basis functions increases (more exactly, as the distance between basis function centers decreases). No specific error bounds are obtained; those require the more complex analysis of the next subsection.

Note that in this subsection, a slightly different notation for the indexing and numbers of basis functions in the sets of basis functions is used. This will make the proof simpler. *This alternate notation appears only in this subsection.* Specifically, let us have the sets of basis functions indexed by ν, that is the *sets* are indexed by ν. Let the cardinality of the ν^{th} set be $\mathcal{I}^{(\nu)}$. For each ν, let $\mathcal{X}^{(\nu)} \doteq \{x_i^{(\nu)}\}_{i=1}^{\mathcal{I}^{(\nu)}}$ and $\mathcal{X}^{(\nu)} \subset \mathcal{X}^{(\nu+1)}$. For instance, in the one-dimensional case, one might have $\mathcal{X}^{(1)} = \{0\}$, $\mathcal{X}^{(2)} = \{-1/2, 0, 1/2\}$, $\mathcal{X}^{(3)} = \{-3/4, -1/2, -1/4, 0, 1/4, 1/2, 3/4\}$, and so on. Further, we will let the basis functions be given by $\psi_i^{(\nu)} \doteq \frac{-1}{2}(x - x_i^{(\nu)})^T C(x - x_i^{(\nu)})$, and consider the sets of basis functions $\Psi^{(\nu)} \doteq \{\psi_i^{(\nu)} : i \in \mathcal{I}^{(\nu)}\}$. Then define the approximations to the semigroup operator, S_τ by

$$S_\tau^{(\nu)}[\phi](x) \doteq \bigoplus_{i=1}^{\mathcal{I}^\nu} a_i^{(\nu)} \otimes \psi_i^{(\nu)}(x), \tag{5.37}$$

where

$$a_i^{(\nu)} \doteq -\max_x \left[\psi_i^{(\nu)}(x) - S_\tau[\phi](x)\right]. \tag{5.38}$$

In other words, $S_\tau^{(\nu)}$ is the result of the application of the S_τ followed by the truncation due to a finite number of basis functions. More specifically, if one defines $\mathcal{T}^{(\nu)}[\phi](x) = \bigoplus_{i=1}^{\mathcal{I}^\nu} a_i^{(\nu)} \otimes \psi_i^{(\nu)}(x)$ with the $a_i^{(\nu)}$ given by (5.38), then $S_\tau^{(\nu)}[\phi] = \mathcal{T}^{(\nu)} \circ S_\tau[\phi]$. Also, let $\mathcal{Y}^{(\nu)} = \{\phi : \overline{B}_R(0) \to \mathbf{R} \mid \exists \{a_i^{(\nu)}\}$ such that $\phi(x) = \bigoplus_{i=1}^{\mathcal{I}^\nu} a_i^{(\nu)} \otimes \psi_i^{(\nu)}(x) \quad \forall x \in \overline{B}_R(0)\}$. Then note that for $\phi \in \mathcal{Y}^{(\nu)}$, one has

$$S_\tau^{(\nu)}[\phi](x) = \bigoplus_{i=1}^{\mathcal{I}^\nu} \left[\bigoplus_{j=1}^{\mathcal{I}^\nu} B_{i,j}^{(\nu)} \otimes a_j^{(\nu)}\right] \otimes \psi_i^{(\nu)}(x) \tag{5.39}$$

where $B_{i,j}^{(\nu)}$ corresponds to $S_\tau^{(\nu)}$ (i.e., $B_{i,j}^{(\nu)} = -\max_x \{\psi_i^{(\nu)}(x) - S_\tau^{(\nu)}[\psi_j^{(\nu)}](x)\}$).

Lastly, we use the notation ${S_\tau}^N$ to indicate repeated application of S_τ N times. (Of course, by the semigroup property, ${S_\tau}^N = S_{N\tau}$.) Correspondingly, we use the notation ${S_\tau^{(\nu)}}^N$ to indicate the application of $S_\tau^{(\nu)}$ N times.

Define $\phi_0(x) \equiv 0$ and

$$\phi_0^{(\nu)}(x) \doteq \bigoplus_{i=1}^{\mathcal{I}^\nu} {a^0}_i^{(\nu)} \otimes \psi_i^{(\nu)}(x), \qquad {a^0}_i^{(\nu)} \doteq -\max_x \left[\psi_i^{(\nu)}(x) - \phi_0(x)\right]. \tag{5.40}$$

From Chapter 3 (among many other sources) one has that

$$\lim_{N\to\infty} {S_\tau}^N[\phi_0] = W. \tag{5.41}$$

Also, note that since $\mathcal{X}^{(\nu)} \subset \mathcal{X}^{(\nu+1)}$, one has

$${S_\tau^{(\nu)}}^N[\phi_0^{(\nu)}](x) \le {S_\tau}^{(\nu+1)^N}[\phi_0^{(\nu+1)}](x) \le {S_\tau}^N[\phi_0](x) \tag{5.42}$$

for all $x \in B_R$.

Note that by (5.40), and the definition of ϕ_0, the corresponding coefficients, ${a^0}_i^{(\nu)}$, satisfy ${a^0}_i^{(\nu)} = 0$ for all i. Combining this with Theorem 4.22 and (5.39), one finds that for each ν, there exists $\overline{N}(\nu)$ such that

$${S_\tau^{(\nu)}}^N[\phi_0^{(\nu)}] = {S_\tau^{(\nu)}}^{\overline{N}(\nu)}[\phi_0^{(\nu)}] \quad \forall\, N \ge \overline{N}(\nu). \tag{5.43}$$

Defining

$${W^{(\nu)}}^\infty \doteq {S_\tau^{(\nu)}}^{\overline{N}(\nu)}[\phi_0^{(\nu)}], \tag{5.44}$$

we further find that the limit is the fixed point. That is,

$$S_\tau^{(\nu)}[{W^{(\nu)}}^\infty] = {W^{(\nu)}}^\infty. \tag{5.45}$$

Then, by (5.41), (5.42) and (5.44), we find that

$${W^{(\nu)}}^\infty \text{ is monotonically increasing in } \nu \tag{5.46}$$

and

$${W^{(\nu)}}^\infty \le W. \tag{5.47}$$

Therefore, there exists $W^{\infty\infty} \le W$ such that

$${W^{(\nu)}}^\infty \uparrow W^{\infty\infty}, \tag{5.48}$$

and in fact, one can demonstrate equicontinuity of the ${W^{(\nu)}}^\infty$ on $\overline{B}_R$ given the assumptions (and consequently uniform convergence).

Under Assumption (A5.5I), one can show (see for instance Lemma 5.12 below, although this is more specific than what is is needed, or Theorem 3.3 in [81]) that given $\varepsilon > 0$, there exists $\nu_\varepsilon < \infty$ such that

$${W^{(\nu)}}^\infty(x) = S_\tau^{(\nu)}[{W^{(\nu)}}^\infty](x) \ge S_\tau[{W^{(\nu)}}^\infty](x) - \varepsilon$$

for all $x \in B_R$ for any $\nu \ge \nu_\varepsilon$. On the other hand, one always has

$$S_\tau^{(\nu)}[\phi] \le S_\tau[\phi].$$

Combining these last two inequalities, one obtains

$${W^{(\nu)}}^\infty = S_\tau^{(\nu)}[{W^{(\nu)}}^\infty] \le S_\tau[{W^{(\nu)}}^\infty] \le S_\tau^{(\nu)}[{W^{(\nu)}}^\infty] + \varepsilon = {W^{(\nu)}}^\infty + \varepsilon. \tag{5.49}$$

Combining this with (5.48), one finds

Theorem 5.10.

$$W^{\infty\infty} = S_\tau[W^{\infty\infty}], \tag{5.50}$$

or in other words, $W^{\infty\infty}$ *is a fixed point of* S_τ*.*

Then, with some more work (see [81], Theorem 3.2), one obtains a convergence theorem.

Theorem 5.11.

$$W^{\infty\infty}(x) = W(x) \quad \forall x \in \overline{B}_R.$$

5.2.2 Truncation Error Estimate

Theorem 5.11 demonstrates convergence of the algorithm to the value function as the basis function density increases. Here we outline one approach to obtaining specific error estimates. The estimates may be rather conservative due to the form of the truncation error bound used; this issue will become clearer below. The main results are in Theorem 5.14 and Remark 5.15. Note that these are only the errors due to truncation to a finite number of basis functions; as noted above, analysis of the errors due to approximation of the entries in the B matrix is discussed further below.

Recall that we choose the basis functions throughout such that $x_1^{(\nu)} = 0$, or in other words, $\psi_1^{(\nu)}(x) = \frac{-1}{2}x^T Cx$, for all ν. (Note that we return here to a notation where the (ν) superscript corresponds to the number of basis functions — as opposed to the more complex notation with cardinality $\mathcal{I}^{(\nu)}$, which was used in the previous subsection only.) Also, we will use the notation

$$W_{N,\tau}^{(\nu)}(x) \doteq {S_\tau^{(\nu)}}^N[\phi_0^{(\nu)}](x)$$

and we reiterate that the N superscript indicates repeated application of the operator N times. Also, $\phi_0^{(\nu)}$ is the finite basis expansion of ϕ_0 (with ν basis fucntions).

To specifically set C, we will replace Assumption (A5.5I) with the following:

> We assume throughout the remainder of the chapter that one may choose symmetric C such that $C - c_R I > 0$ and $\delta' \in (0,1)$ such that with $C' \doteq (1-\delta')C$, one has $S_\tau[\psi_i] \in \mathcal{S}_R^{C'L'}$ for all i where $|C||C^{-1}|R + |C^{-1}|L' \leq D_R$. (A5.6I)

Note that one could be more general, allowing C' to be a more general positive definite symmetric matrix such that $C - C' > 0$, but we will not generalize to that here. Finally, it should be noted that δ' would depend on τ; as $\tau \downarrow 0$, one would need to take $\delta' \downarrow 0$. Since δ' will appear in the denominator of the error bound of the next lemma (as well as implicitly in the denominator of the quotient on the right-hand side of the error bound in Theorem 5.14), this implies that one does not want to simply take $\tau \downarrow 0$ as the means for reducing the errors. This will be discussed further in the next section.

The following lemma is a general result about the errors due to truncation when using the above max-plus basis expansion. The proof is long and rather technical.

Lemma 5.12. *Let δ', C', L' be as in Assumption (A5.6I), and let $\phi \in \mathcal{S}_{\overline{R}}^{C'L'}$ with $\phi(0) = 0$, ϕ differentiable at zero with $\nabla_x \phi(0) = 0$, and $-\frac{1}{2}x^T C' x \leq \phi(x) \leq \frac{1}{2}\widehat{\mathcal{M}}|x|^2$ for all x for some $\widehat{\mathcal{M}} < \infty$. Let $\{\psi_i\}_{i=1}^{\nu}$ consist of basis functions with matrix C and centers $\{x_i\} \subseteq \overline{B}_{D_R}$ such that $C - C' > 0$, and let $\Delta \doteq \max_{x \in \overline{B}_{D_R}} \min_i |x - x_i|$. Let*

$$\phi^{\Delta}(x) = \max_i [a_i + \psi_i(x)] \quad \forall\, x \in \overline{B}_R$$

where

$$a_i = -\max_{x \in \overline{B}_R} [\psi_i(x) - \phi(x)] \quad \forall\, i.$$

Then

$$0 \leq \phi(x) - \phi^{\Delta}(x) \leq \begin{cases} |C|\Big[2\widehat{\beta} + 1 + |C|/(\delta' c_R)\Big]|x|\Delta & \text{if } |x| \geq \Delta \\ \frac{1}{2}\big[\widehat{\mathcal{M}} + |C|\big]|x|\Delta & \text{otherwise,} \end{cases}$$

where $\widehat{\beta}$ is specified in the proof.

Proof. Recall that

$$\phi(x) = \max_{\widetilde{x} \in \overline{B}_{D_R}} [a(\widetilde{x}) + \psi_{\widetilde{x}}(x)] \quad \forall\, x \in \overline{B}_R,$$

where

$$a(\widetilde{x}) = -\max_{x \in \overline{B}_R} [\psi_{\widetilde{x}}(x) - \phi(x)] \quad \forall\, \widetilde{x} \in \overline{B}_{D_R}$$

and

$$\psi_{\widetilde{x}}(x) \doteq -\tfrac{1}{2}(x - \widetilde{x})^T C (x - \widetilde{x}) \quad \forall\, x \in \overline{B}_R, \widetilde{x} \in \overline{B}_{D_R}.$$

(There is obviously a slight conflict in notation between such a $\psi_{\widetilde{x}}$ where the subscript $\widetilde{x} \in \mathbf{R}^n$ and ψ_i where the subscript is an index of $x_i \in \mathbf{R}^n$, but this should not lead to confusion, and seems the best compromise.) It is obvious that $0 \leq \phi(x) - \phi^{\Delta}(x)$, and so we prove the other bound.

Consider any $\overline{x} \in \overline{B}_R$. Then

$$\phi(\overline{x}) = a(\widetilde{x}) + \psi_{\widetilde{x}}(\overline{x}) \tag{5.51}$$

if and only if

$$C(\overline{x} - \widetilde{x}) \in -D_x^- \phi(\overline{x}),$$

where

$$D_x^- \phi(x) = \left\{ p \in \mathbf{R}^n : \liminf_{|y-x|\to 0} \frac{\phi(y) - \phi(x) - (y-x)\cdot p}{|y-x|} \geq 0 \right\}.$$

We denote such an $\widetilde{x}$ corresponding to $\overline{x}$ (in (5.51)) as $\overline{\widetilde{x}}$. By the Lipschitz nature of ϕ, one can easily establish that

$$|\overline{\widetilde{x}} - \overline{x}| \leq |C^{-1}|L'. \tag{5.52}$$

However, it will be desirable to have a bound where the right-hand side depends linearly on $|\overline{x}|$. (Actually, this may only be necessary for small $\overline{x}$, while (5.52) may be a smaller bound for large $\overline{x}$, but we will obtain it for general $\overline{x}$.) Using (5.51), and noting that $\phi \geq -\frac{1}{2}x^T C' x \geq -\frac{1}{2}x^T C x$, one has

$$\tfrac{1}{2}(\overline{x} - \overline{\widetilde{x}})^T C(\overline{x} - \overline{\widetilde{x}}) \leq a(\overline{\widetilde{x}}) + \tfrac{1}{2}\overline{x}^T C \overline{x}.$$

Also, because $a(\overline{\widetilde{x}}) + \psi_{\overline{\widetilde{x}}}(\cdot)$ touches ϕ from below at $\overline{x}$, one must have

$$\begin{aligned}
&\tfrac{1}{2}(\overline{x} - \overline{\widetilde{x}})^T C(\overline{x} - \overline{\widetilde{x}}) - \tfrac{1}{2}(x - \overline{\widetilde{x}})^T C(x - \overline{\widetilde{x}}) \\
&\leq a(\overline{\widetilde{x}}) + \tfrac{1}{2}\overline{x}^T C\overline{x} - \tfrac{1}{2}(x - \overline{\widetilde{x}})^T C(x - \overline{\widetilde{x}}) \\
&\leq \phi(x) + \tfrac{1}{2}\overline{x}^T C\overline{x} \leq \tfrac{1}{2}\widehat{\mathcal{M}}|x|^2 + \tfrac{1}{2}\overline{x}^T C\overline{x}
\end{aligned}$$

for all $x \in \overline{B}_R$ where the last inequality is by assumption. Define

$$F(x) \doteq \tfrac{1}{2}(\overline{x} - \overline{\widetilde{x}})^T C(\overline{x} - \overline{\widetilde{x}}) - \tfrac{1}{2}(x - \overline{\widetilde{x}})^T C(x - \overline{\widetilde{x}}) - \tfrac{1}{2}\widehat{\mathcal{M}}|x|^2,$$

and we see that we require $F(x) \leq \frac{1}{2}\overline{x}^T C\overline{x}$ for all $x \in \overline{B}_R$. Taking the derivative, we find the maximum of F at $\widehat{x}$ given by

$$\widehat{x} = (C + \widehat{\mathcal{M}}I)^{-1} C\overline{\widetilde{x}} \tag{5.53}$$

and so

$$\widehat{x} - \overline{\widetilde{x}} = -\widehat{\mathcal{M}}(C + \widehat{\mathcal{M}}I)^{-1}\overline{\widetilde{x}}. \tag{5.54}$$

(In the interest of readability, we ignore the detail of the case where $\widehat{x} \notin B_R(0)$ here.) Therefore, $F(\widehat{x}) \leq \frac{1}{2}\overline{x}^T C\overline{x}$ implies

$$\begin{aligned}
(\overline{x} - \overline{\widetilde{x}})^T C(\overline{x} - \overline{\widetilde{x}}) &\leq \overline{\widetilde{x}}^T \widehat{\mathcal{M}}(C + \widehat{\mathcal{M}}I)^{-1} C(C + \widehat{\mathcal{M}}I)^{-1}\widehat{\mathcal{M}}\overline{\widetilde{x}} \\
&\quad + \overline{\widetilde{x}}^T C(C + \widehat{\mathcal{M}}I)^{-1}\widehat{\mathcal{M}}(C + \widehat{\mathcal{M}}I)^{-1} C\overline{\widetilde{x}} + \overline{x}^T C\overline{x} \\
&= \widehat{\mathcal{M}}\overline{\widetilde{x}}^T \Big[\widehat{\mathcal{M}}(C + \widehat{\mathcal{M}}I)^{-1} C(C + \widehat{\mathcal{M}}I)^{-1} \\
&\qquad + \widehat{\mathcal{M}}(C + \widehat{\mathcal{M}}I)^{-1}\widehat{\mathcal{M}}I(C + \widehat{\mathcal{M}}I)^{-1} \\
&\qquad - \widehat{\mathcal{M}}^2(C + \widehat{\mathcal{M}}I)^{-2} + C(C + \widehat{\mathcal{M}}I)^{-2} C\Big]\overline{\widetilde{x}} + \overline{x}^T C\overline{x} \\
&= \widehat{\mathcal{M}}\overline{\widetilde{x}}^T \Big[\widehat{\mathcal{M}} C(C + \widehat{\mathcal{M}}I)^{-2} + C(C + \widehat{\mathcal{M}}I)^{-2} C\Big]\overline{\widetilde{x}} + \overline{x}^T C\overline{x} \\
&= \overline{\widetilde{x}}^T \widehat{\mathcal{M}} C(C + \widehat{\mathcal{M}}I)^{-1}\overline{\widetilde{x}} + \overline{x}^T C\overline{x}.
\end{aligned} \tag{5.55}$$

Noting that C is positive definite symmetric, and writing it as $C = \sqrt{C}\sqrt{C}^T$ where $\sqrt{C} = S\sqrt{\Lambda}$ with S unitary and Λ the matrix of eigenvalues, one may rewrite the first term on the right-hand side of (5.55) as

$$\bar{\bar{x}}^T \widehat{\mathcal{M}} C(C + \widehat{\mathcal{M}} I)^{-1}\bar{\bar{x}} = \bar{\bar{x}}^T \widehat{\mathcal{M}} \tfrac{1}{2}\Big[C(C + \widehat{\mathcal{M}} I)^{-1} + (C + \widehat{\mathcal{M}} I)^{-1} C\Big]\bar{\bar{x}}$$
$$= \bar{\bar{x}}^T \sqrt{C} Q \sqrt{C}^T \bar{\bar{x}}$$

where

$$Q \doteq \tfrac{1}{2}\widehat{\mathcal{M}}\Big[\sqrt{C}^T (C + \widehat{\mathcal{M}} I)^{-1}\sqrt{C}^{-T} + \sqrt{C}^{-1}(C + \widehat{\mathcal{M}} I)^{-1}\sqrt{C}\Big].$$

Making the change of variables $y = \sqrt{C}^T x$, (5.55) becomes

$$|\bar{y} - \bar{\bar{y}}|^2 \le \bar{\bar{y}}^T Q \bar{\bar{y}} + |\bar{y}|^2.$$

Noting that $\sqrt{C}^T (C + \widehat{\mathcal{M}} I)^{-1}\sqrt{C}^{-T}$ is a similarity transform of $(C + \widehat{\mathcal{M}} I)^{-1}$, one sees that the eigenvalues of Q are the eigenvalues of $(C + \widehat{\mathcal{M}} I)^{-1}$. Now, since $(C + \widehat{\mathcal{M}} I)$ is positive definite,

$$(C + \widehat{\mathcal{M}} I) = \overline{S}\,\overline{\Lambda}\,\overline{S}^{-1}$$

with $\overline{\Lambda}$ the diagonal matrix of eigenvalues and $\overline{S}$ the unitary matrix of eigenvectors. Therefore, $\widehat{\mathcal{M}}(C + \widehat{\mathcal{M}} I)^{-1} = \overline{S}(\widehat{\mathcal{M}\overline{\Lambda}^{-1}})\overline{S}^{-1}$, and note that $\beta \doteq \max_i\{\widehat{\mathcal{M}\overline{\lambda}_i^{-1}}\} < 1$ where the $\overline{\lambda}_i$ are the diagonal elements of $\overline{\Lambda}$. Consequently,

$$|\bar{y} - \bar{\bar{y}}|^2 \le \beta|\bar{\bar{y}}|^2 + |\bar{y}|^2 \tag{5.56}$$

where $\beta \in (0, 1)$. This implies

$$\begin{aligned} |\bar{\bar{y}} - \bar{y}|^2 &\le \beta|\bar{\bar{y}} - \bar{y} + \bar{y}|^2 + |\bar{y}|^2 \\ &= \beta\Big[|\bar{\bar{y}} - \bar{y}|^2 + |\bar{y}|^2 + 2(\bar{\bar{y}} - \bar{y})\cdot\bar{y}\Big] + |\bar{y}|^2 \\ &\le \beta|\bar{\bar{y}} - \bar{y}|^2 + (\beta + 1)|\bar{y}|^2 \\ &\quad + \beta\Big[\frac{(1-\beta)/2}{\beta}|\bar{\bar{y}} - \bar{y}|^2 + \frac{\beta}{(1-\beta)/2}|\bar{y}|^2\Big], \end{aligned}$$

which after some rearrangement, yields

$$|\bar{\bar{y}} - \bar{y}|^2 \le \frac{2(1+\beta^2)}{(1-\beta)^2}|\bar{y}|^2 \tag{5.57}$$

or equivalently,

$$(\bar{x} - \bar{\bar{x}})^T C(\bar{x} - \bar{\bar{x}}) \le \left[\frac{2(1+\beta^2)}{(1-\beta)^2}\right]\bar{x}^T C\bar{x}.$$

Consequently, there exists $\widehat{\beta} < \infty$ (i.e., $\widehat{\beta} = [\|\sqrt{C}|/\sqrt{c_R}][\sqrt{2(1+\beta^2)}/(1-\beta)]$) such that

$$|\overline{x} - \overline{\overline{x}}| \leq \widehat{\beta}|\overline{x}|. \tag{5.58}$$

Given $\overline{\overline{x}}$, let $\bar{i} \in \operatorname{argmin}_i |x_i - \overline{\overline{x}}|$, and note that

$$|x_{\bar{i}} - \overline{\overline{x}}| \leq \Delta. \tag{5.59}$$

It is easy to see that

$$\begin{aligned}
|\psi_{\overline{\overline{x}}}(x) - \psi_{\bar{i}}(x)| &\leq \tfrac{1}{2}|(x - \overline{\overline{x}})^T C(x - \overline{\overline{x}}) - (x - \overline{\overline{x}})^T C(x - x_{\bar{i}})| \\
&\quad + \tfrac{1}{2}|(x - \overline{\overline{x}})^T C(x - x_{\bar{i}}) - (x - x_{\bar{i}})^T C(x - x_{\bar{i}})| \\
&\leq \tfrac{1}{2}|C|\Big[|\overline{\overline{x}} - x_{\bar{i}}||x - \overline{\overline{x}}| + |\overline{\overline{x}} - x_{\bar{i}}||x - x_{\bar{i}}|\Big] \\
&\leq \tfrac{1}{2}|C|\Big[|\overline{\overline{x}} - x_{\bar{i}}||x - \overline{\overline{x}}| + |\overline{\overline{x}} - x_{\bar{i}}|(|x - \overline{\overline{x}}| + |\overline{\overline{x}} - x_{\bar{i}}|)\Big],
\end{aligned}$$

which by (5.59)

$$\leq |C|\Big[|x - \overline{\overline{x}}|\Delta + \tfrac{1}{2}\Delta^2\Big]. \tag{5.60}$$

Combining (5.58) and (5.60), one finds

$$|\psi_{\overline{\overline{x}}}(\overline{x}) - \psi_{\bar{i}}(\overline{x})| \leq |C|\Big[\widehat{\beta}|\overline{x}|\Delta + \tfrac{1}{2}\Delta^2\Big]. \tag{5.61}$$

Now note that

$$\phi(\overline{x}) - \phi^{\Delta}(\overline{x}) \leq a(\overline{\overline{x}}) + \psi_{\overline{\overline{x}}}(\overline{x}) - [a_{\bar{i}} + \psi_{\bar{i}}(\overline{x})]$$

which by (5.61)

$$\leq |C|\Big[\widehat{\beta}|\overline{x}|\Delta + \tfrac{1}{2}\Delta^2\Big] + a(\overline{\overline{x}}) - a_{\bar{i}}. \tag{5.62}$$

We now deal with the last two terms in this bound. Let

$$\overline{x}_{\bar{i}} \doteq \operatorname*{argmax}_{x \in \overline{B}_R}[\psi_{\bar{i}}(x) - \phi(x)].$$

(Note that we will also skip the technical details of the additional case where $\overline{x}_{\bar{i}}$ lies on the boundary of $\overline{B}_R$, and consider only the case where the argmax in the interior, B_R.) Then,

$$-C(\overline{x}_{\bar{i}} - x_{\bar{i}}) \in D^-\phi(\overline{x}_{\bar{i}})$$

and

$$-C(\overline{x} - \overline{\overline{x}}) \in D^-\phi(\overline{x}).$$

By the semiconvexity, one has the general result that $p \in D^-\phi(x)$, $q \in D^-\phi(y)$ implies

$$(p - q) \cdot (x - y) \geq -(x - y)^T C'(x - y).$$

Consequently,

$$-(\overline{x}_{\bar{i}} - x_{\bar{i}} + \overline{\overline{x}} - \overline{x})^T C(\overline{x}_{\bar{i}} - \overline{x}) \geq -(\overline{x}_{\bar{i}} - \overline{x})^T C'(\overline{x}_{\bar{i}} - \overline{x}).$$

Recalling that $C' = (1-\delta')C$, we see that this implies

$$\begin{aligned} -(\overline{x}_{\bar{i}} - \overline{x})^T C(\overline{x}_{\bar{i}} - \overline{x}) + (1-\delta')(\overline{x}_{\bar{i}} - \overline{x})^T C(\overline{x}_{\bar{i}} - \overline{x}) &\geq -|C|\,|\overline{x}_{\bar{i}} - \overline{x}|\,|x_{\bar{i}} - \overline{\overline{x}}| \\ &\geq -|C|\,|\overline{x}_{\bar{i}} - \overline{x}|\varDelta, \end{aligned}$$

or

$$\delta'(\overline{x}_{\bar{i}} - \overline{x})^T C(\overline{x}_{\bar{i}} - \overline{x}) \leq |C|\,|\overline{x}_{\bar{i}} - \overline{x}|\varDelta.$$

Noting that $C - c_R I > 0$, this implies

$$|\overline{x}_{\bar{i}} - \overline{x}| \leq \frac{|C|}{\delta' c_R}\varDelta. \tag{5.63}$$

Now,

$$\begin{aligned} \overline{\overline{a}} - a_{\bar{i}} &\leq \psi_{\bar{i}}(\overline{x}_{\bar{i}}) - \psi_{\overline{\overline{x}}}(\overline{x}_{\bar{i}}) \\ &= \psi_{\bar{i}}(\overline{x}) - \psi_{\overline{\overline{x}}}(\overline{x}) + [\psi_{\bar{i}}(\overline{x}_{\bar{i}}) - \psi_{\overline{\overline{x}}}(\overline{x}_{\bar{i}})] - [\psi_{\bar{i}}(\overline{x}) - \psi_{\overline{\overline{x}}}(\overline{x})], \end{aligned}$$

which, after cancellation,

$$\begin{aligned} &= \psi_{\bar{i}}(\overline{x}) - \psi_{\overline{\overline{x}}}(\overline{x}) - (\overline{x} - \overline{x}_{\bar{i}})C(x_{\bar{i}} - \overline{\overline{x}}) \\ &\leq |\psi_{\bar{i}}(\overline{x}) - \psi_{\overline{\overline{x}}}(\overline{x})| + |C|\varDelta|\overline{x} - \overline{x}_{\bar{i}}|, \end{aligned}$$

which by (5.61) and (5.63)

$$\leq |C|\Big[\widehat{\beta}|\overline{x}| + \big(\tfrac{1}{2} + |C|/(\delta' c_R)\big)\varDelta\Big]\varDelta. \tag{5.64}$$

Combining (5.62) and (5.64) yields

$$\phi(\overline{x}) - \phi^{\varDelta}(\overline{x}) \leq |C|\Big[2\widehat{\beta}|\overline{x}| + \big(1 + |C|/(\delta' c_R)\big)\varDelta\Big]\varDelta. \tag{5.65}$$

Suppose $|\overline{x}| \geq \varDelta$. Then, (5.65) implies

$$\phi(\overline{x}) - \phi^{\varDelta}(\overline{x}) \leq |C|\Big[2\widehat{\beta} + 1 + |C|/(\delta' c_R)\Big]|\overline{x}|\varDelta, \tag{5.66}$$

which is the first case in right-hand side of the assertion.

Lastly, suppose $|\overline{x}| < \varDelta$. By assumption, there exists $\widehat{\mathcal{M}} < \infty$ such that $\phi(x) \leq \frac{1}{2}\widehat{\mathcal{M}}|x|^2$. Therefore,

$$\phi(\overline{x}) - \phi^{\varDelta}(\overline{x}) \leq \tfrac{1}{2}(\widehat{\mathcal{M}} + |C|)|\overline{x}|^2 \leq \tfrac{1}{2}(\widehat{\mathcal{M}} + |C|)|\overline{x}|\varDelta$$

which completes the proof. □

The above lemma is a general result about the errors due to truncation with the above max-plus basis expansion. In order to apply this to the problem at hand, one must consider the effect of repeated application of the truncated operator $S_\tau^{(\nu)}$. Note that $S_\tau^{(\nu)}$ may be written as the composition of S_τ and a truncation operator, $\mathcal{T}^{(\nu)}$ where we have

$$\mathcal{T}^{(\nu)}[\phi] = \phi^\Delta$$

in the notation of the previous lemma, where in particular, ϕ^Δ was given by

$$\phi^\Delta(x) = \max_i [a_i + \psi_i(x)] \quad \forall\, x \in \overline{B}_R,$$

where

$$a_i = -\max_{x\in\overline{B}_R(0)} [\psi_i(x) - \phi(x)] \quad \forall\, i.$$

In other words, one has the following equivalence of notation

$$S_\tau^{(\nu)}[\phi] = \{\mathcal{T}^{(\nu)} \circ S_\tau\}[\phi] = \{S_\tau[\phi]\}^\Delta, \tag{5.67}$$

which we shall use freely throughout the remainder of this chapter.

We now proceed to consider how truncation errors accumulate. In order to simplify the analysis, we simply let

$$\mathcal{M}_{C'} \doteq \max\Big\{|C|\big[2\widehat{\beta} + 1 + |C|/(\delta' c_R)\big],\, \tfrac{1}{2}\big[\widehat{\mathcal{M}} + |C|\big]\Big\}. \tag{5.68}$$

Fix Δ. We suppose that we have ν sufficiently large (with properly distributed basis function centers) so that

$$\max_{x\in\overline{B}_{D_R}} \min_i |x - x_i| \le \Delta.$$

Let ϕ_0 satisfy the conditions on ϕ in Lemma 5.12. (One can simply take $\phi_0 \equiv 0$.) Then, by Lemma 5.12,

$$\phi_0(x) - \mathcal{M}_{C'}|x|\Delta \le \phi_0^{(\nu)}(x) \le \phi_0(x) \quad \forall\, x \in \overline{B}_R(0). \tag{5.69}$$

Now, for any $x \in \overline{B}_R$, let $u^{1,\overline{\varepsilon},x}$ be $\overline{\varepsilon}/2$-optimal for $S_\tau[\phi_0](x)$, and let $\xi^{1,\overline{\varepsilon},x}$ be the corresponding trajectory. Then,

$$\begin{aligned} 0 &\le S_\tau[\phi_0](x) - S_\tau[\phi_0^{(\nu)}](x) \\ &\le \phi_0(\xi_\tau^{1,\overline{\varepsilon},x}) - \phi_0^{(\nu)}(\xi_\tau^{1,\overline{\varepsilon},x}) + \frac{\overline{\varepsilon}}{2}, \end{aligned}$$

which by (5.69)

$$\le \mathcal{M}_{C'}|\xi_\tau^{1,\overline{\varepsilon},x}|\Delta + \frac{\overline{\varepsilon}}{2}. \tag{5.70}$$

Proceeding along, one then finds

$$\begin{aligned}0 &\le S_\tau[\phi_0](x) - S_\tau^{(\nu)}[\phi_0^{(\nu)}](x)\\ &= S_\tau[\phi_0](x) - S_\tau[\phi_0^{(\nu)}](x) + S_\tau[\phi_0^{(\nu)}](x) - S_\tau^{(\nu)}[\phi_0^{(\nu)}](x),\end{aligned}$$

which by Lemma 5.12, the fact that $S_\tau[\phi_0^{(\nu)}] \in \mathcal{S}_{\overline{R}}^{C'L'}$ (by Assumption (A5.6I)), and (5.70)

$$\le \mathcal{M}_{C'}|\xi_\tau^{1,\overline{\varepsilon},x}|\Delta + \mathcal{M}_{C'}|x|\Delta + \frac{\overline{\varepsilon}}{2}. \tag{5.71}$$

Let us proceed one more step with this approach. For any $x \in \overline{B}_R$, let $u^{2,\overline{\varepsilon},x}$ be $\overline{\varepsilon}/4$-optimal for $S_\tau[S_\tau[\phi_0]](x)$ (that is, $\overline{\varepsilon}/4$-optimal for problem S_τ with terminal cost $S_\tau[\phi_0]$), and let $\xi^{2,\overline{\varepsilon},x}$ be the corresponding trajectory. Then, as before,

$$\begin{aligned}0 &\le S_{2\tau}[\phi_0](x) - S_\tau[S_\tau^{(\nu)}\phi_0^{(\nu)}](x)\\ &= S_\tau[S_\tau[\phi_0]](x) - S_\tau[S_\tau^{(\nu)}[\phi_0^{(\nu)}]](x)\\ &\le S_\tau[\phi_0](\xi_\tau^{2,\overline{\varepsilon},x}) - S_\tau^{(\nu)}[\phi_0^{(\nu)}](\xi_\tau^{2,\overline{\varepsilon},x}) + \frac{\overline{\varepsilon}}{4}.\end{aligned} \tag{5.72}$$

Now let

$$u_t^{2,\overline{\varepsilon}} \doteq \begin{cases} u_t^{2,\overline{\varepsilon},x} & \text{if } t \in [0,\tau]\\ u_{(t-\tau)}^{1,\overline{\varepsilon},\xi_\tau^{2,\overline{\varepsilon},x}} & \text{if } t \in (\tau, 2\tau],\end{cases}$$

and let $\overline{\xi}^{2,\overline{\varepsilon},x}$ be the corresponding trajectory. Then combining (5.71) and (5.72), one has

$$\begin{aligned}0 &\le S_{2\tau}[\phi_0](x) - S_\tau[S_\tau^{(\nu)}\phi_0^{(\nu)}](x)\\ &\le \mathcal{M}_{C'}|\overline{\xi}_{2\tau}^{2,\overline{\varepsilon},x}|\Delta + \mathcal{M}_{C'}|\overline{\xi}_\tau^{2,\overline{\varepsilon},x}|\Delta + \frac{\overline{\varepsilon}}{2} + \frac{\overline{\varepsilon}}{4}.\end{aligned} \tag{5.73}$$

Applying Lemma 5.12 again, but now using (5.73), one has

$$\begin{aligned}0 &\le S_{2\tau}[\phi_0](x) - S_\tau^{(\nu)}[S_\tau^{(\nu)}[\phi_0^{(\nu)}]](x)\\ &= S_\tau[S_\tau[\phi_0]](x) - S_\tau[S_\tau^{(\nu)}[\phi_0^{(\nu)}]](x) + S_\tau[S_\tau^{(\nu)}[\phi_0^{(\nu)}]](x) - S_\tau^{(\nu)}[S_\tau^{(\nu)}[\phi_0^{(\nu)}]](x)\\ &\le \mathcal{M}_{C'}|\overline{\xi}_{2\tau}^{2,\overline{\varepsilon},x}|\Delta + \mathcal{M}_{C'}|\overline{\xi}_\tau^{2,\overline{\varepsilon},x}|\Delta + \mathcal{M}_{C'}|x|\Delta + \frac{\overline{\varepsilon}}{2} + \frac{\overline{\varepsilon}}{4}\\ &= \mathcal{M}_{C'}\Delta\sum_{i=0}^{2}|\overline{\xi}_{i\tau}^{2,\overline{\varepsilon},x}| + \sum_{i=1}^{2}\frac{\overline{\varepsilon}}{2^i}.\end{aligned} \tag{5.74}$$

It is then clear that, by induction, one obtains

Lemma 5.13.

$$0 \le S_{N\tau}[\phi_0](x) - {S_\tau^{(\nu)}}^N[\phi_0](x) \le \mathcal{M}_{C'}\Delta\sum_{i=0}^{N}|\overline{\xi}_{i\tau}^{N,\overline{\varepsilon},x}| + \sum_{i=1}^{N}\frac{\overline{\varepsilon}}{2^i}, \tag{5.75}$$

where the construction of $\overline{\varepsilon}$-optimal $\overline{\xi}_\cdot^{N,\overline{\varepsilon},x}$ by induction follows in the obvious way as above.

Theorem 5.14. *Let $\{\psi_i\}_{i=1}^{\nu}$, C' and Δ be as in Lemma 5.12. Then, there exists $\overline{m}, \overline{\lambda} \in (0,\infty)$ such that*

$$0 \le W(x) - W^{(\nu)^\infty}(x) \le \mathcal{M}_{C'}\left(\frac{e^{\overline{m}}}{1-e^{-\overline{\lambda}\tau}}\right)|x|\Delta \quad \forall x \in \overline{B}_R$$

where $\mathcal{M}_{C'}$ is given in (5.68).

Remark 5.15. By Theorem 4.22, there exists $N = N(\nu) < \infty$ such that

$$W^{(\nu)^\infty}(x) = W^{(\nu)^N}(x) \quad \forall x \in \overline{B}_R,$$

and so Theorem 5.14 also implies

$$0 \le W(x) - W^{(\nu)^N}(x) \le \mathcal{M}_{C'}\left(\frac{e^{\overline{m}}}{1-e^{-\overline{\lambda}\tau}}\right)|x|\Delta \quad \forall x \in \overline{B}_R$$

for $N \ge N(\nu)$.

Proof. Let $\overline{\varepsilon} \in (0,1)$. Fix ϕ_0 and x. For each $N < \infty$, construct $u_\cdot^{N,\overline{\varepsilon}}$ as above along with the corresponding $\overline{\xi}^{N,\overline{\varepsilon},x}$. Let $u_t^{\infty,\overline{\varepsilon}} = u_t^{N,\overline{\varepsilon}}$ if $t \in [0,N\tau]$, and similarly, $\overline{\xi}_t^{\infty,\overline{\varepsilon},x} = \overline{\xi}_t^{N,\overline{\varepsilon},x}$ if $t \in [0,N\tau]$. Then, by the results of Chapter 3, there exists $\widetilde{K} < \infty$ (independent of $\overline{\varepsilon} \in (0,1)$) such that

$$\|u^{\infty,\overline{\varepsilon}}\|_{L_2(0,N\tau)} \le \widetilde{K}(1+|x|^2)$$

for all $N < \infty$. Consequently, using Assumptions (A4.1I) and (A4.2I), there exist $\overline{m}, \overline{\lambda} \in (0,\infty)$ such that

$$|\overline{\xi}_t^{\infty,\overline{\varepsilon},x}| \le |x|e^{\overline{m}-\overline{\lambda}t} \quad \forall t \in [0,\infty). \tag{5.76}$$

(We remark that $\overline{m}$ and λ may depend on the particular $u^{\infty,\overline{\varepsilon}}$ constructed above.) Then, by Lemma 5.13 and (5.76),

$$\begin{aligned} 0 \le S_{N\tau}[\phi_0](x) - S_\tau^{(\nu)^N}[\phi_0](x) &\le \mathcal{M}_{C'}\Delta|x|e^{\overline{m}}\sum_{i=0}^{N} e^{-\overline{\lambda}i\tau} + \sum_{i=1}^{N}\frac{\overline{\varepsilon}}{2^i} \\ &\le \mathcal{M}_{C'}\Delta|x|\frac{e^{\overline{m}}}{1-e^{-\overline{\lambda}\tau}} + \overline{\varepsilon}. \end{aligned}$$

Because this is true for all $N \in \mathcal{N}$, and $S_\tau^{(\nu)^N}[\phi_0](x) = S_\tau^{(\nu)^{N+1}}[\phi_0](x) = W^{(\nu)^\infty}(x)$ for all $N \ge N(\nu)$, one obtains the result by taking the limit as $N \to \infty$ and then as $\overline{\varepsilon} \downarrow 0$. □

Lastly, we note that for τ sufficiently small, where

$$\tau \le 1/\overline{\lambda} \tag{5.77}$$

is sufficient (so that $\overline{\lambda}\tau/2 \le (1-e^{-\overline{\lambda}\tau})$), one has

$$0 \le W(x) - W^{(\nu)^\infty}(x) \le \mathcal{M}_{C'}\Delta\left(\frac{e^{\overline{m}}}{1-e^{-\overline{\lambda}\tau}}\right)|x| \le K_{1,\tau}|x|(\Delta/\tau) \tag{5.78}$$

with

$$K_{1,\tau} \doteq 2\mathcal{M}_{C'}e^{\overline{m}}/\overline{\lambda}. \tag{5.79}$$

We note that through the dependency of $K_{1,\tau}$ on $\mathcal{M}_{C'}$, and in turn the dependency of $\mathcal{M}_{C'}$ on δ' (see (5.68)), $K_{1,\tau} \to \infty$ as $\tau \downarrow 0$ (see (A5.6I) and the discussion following it).

5.3 Errors in the Approximation of B

In the previous section, we considered the errors due to truncation while assuming that B and consequently, the eigenvector e were computed exactly. Of course, as discussed in Section 5.1, there is an allowable upper limit for errors in the elements of B, below which one can guarantee the convergence of the power method. The errors in B also translate into errors in the eigenvector and consequently the approximate solution as discussed in Sections 5.1 and 5.4. In this section, we consider a power series (in t) for $\overline{V}(x,t) \doteq S_t[\psi_i](x)$ where we recall $B_{j,i} = -\max_{x\in\overline{B}_R(0)}[-\psi_j(x) - S_\tau[\psi_i](x)]$. With the power series for $\overline{V}(x,t) = S_t[\psi_i](x)$ truncated at some level, $t^{\nu'-1}$ (for each i), we obtain a relationship between ν', τ and basis function density which guarantees that the errors in B do not exceed the allowable bounds obtained in Section 5.1. In addition to the errors incurred by truncation of the power series, there may be errors in the computation of the terms in the series themselves. In Subsection 5.3.1, one particular method for computing the power series terms to sufficient accuracy is given.

As noted above, one approach to the computation of B is a Taylor series (in t) approximation to $S_t[\psi_i](x)$. More specifically, letting $\overline{V}(x,t) = S_t[\psi_i](x)$, so that $\overline{V}$ satisfies

$$\begin{aligned} V_t &= f\cdot\nabla V + l + \tfrac{1}{2\gamma^2}\nabla V^T\sigma\sigma^T\nabla V, \\ V(x,0) &= \psi_i(x), \end{aligned} \tag{5.80}$$

one may approximate $\overline{V}$ as

$$\overline{V}(x,t) = V_0(x) + V_1(x)t + \frac{1}{2}V_2(x)t^2 + \cdots. \tag{5.81}$$

Here $V_0(x) = \psi_i(x)$ and V_1 is the right-hand side of (5.80) with ψ_i replacing $\overline{V}$. Specifically,

$$V_1(x) = f\psi_{i_x} + l + \frac{1}{2\gamma^2}\psi_{i_x}a\psi_{i_x},$$

where $a = \sigma\sigma^T$ and we drop the gradient/vector notation for simplification here and below. The higher-order terms are computed by differentiating (5.80) at $t = 0$. Of course this process requires some smoothness for V. The following is well known, and so we only sketch a proof.

Theorem 5.16. *Given* $R' < \infty$ *and* $\nu' \in \mathcal{N}$, *there exists* $\tau' > 0$ *such that* $\overline{V} \in C^{\nu'}(B_{R'} \times (0,\tau'))$.

Proof. The result for C^2 can be found, for instance, in [37] as well as many earlier works (see the references in [37] as well as [47] and [48]). In order to obtain continuity of higher derivatives, one simply differentiates (5.80), and applies the same technique. For example, the partial $\overline{V}_{x_l}(x,t)$ satisfies

$$\begin{aligned} U_t &= \Big[f_{x_l}\overline{V}_x + l_{x_l} + \overline{V}_x a_{x_l} \overline{V}_x\Big] + \Big[f + 2\overline{V}_x a\Big] U_x, \\ U(x,0) &= \psi_{i_{x_l}}(x). \end{aligned}$$

Note that τ' may depend on ν'. □

Fix some $R', \nu' < \infty$. Let τ' be given by Theorem 5.16. We assume $\tau < \min\{\tau', 1, 1/c\}$ (where the motivation for the bounds of 1 and $1/c$ appear in (5.87) and (5.90) below) and $\widetilde{R} < R'$. Then we may approximate $\overline{V}$ over $\overline{B}_{\widetilde{R}} \times (0,\tau)$ by

$$\widetilde{V}(x,t) = V_0(x) + V_1(x)t + V_2(x)\frac{t^2}{2} + \cdots + V_{\nu'-1}(x)\frac{t^{\nu'-1}}{(\nu'-1)!}. \tag{5.82}$$

Letting

$$M_{R',\nu'} \doteq \max_{(x,t)\in \overline{B}_{\widetilde{R}} \times [0,\tau]} |\overline{V}_{t^{(\nu')}}(x,t)|,$$

one has

$$|\overline{V}(x,t) - \widetilde{V}(x,t)| \le M_{R',\nu'} \frac{\tau^{\nu'}}{(\nu')!} \quad \forall\, (x,t) \in \overline{B}_{\widetilde{R}} \times [0,\tau]. \tag{5.83}$$

Now define the corresponding approximation to B by

$$\widetilde{B}_{j,i} = -\max_{x \in \overline{B}_{\widetilde{R}}} \Big\{ \psi_j(x) - \widetilde{V}(x,\tau) \Big\}. \tag{5.84}$$

By (5.83) and (5.84), one has

$$|B_{j,i} - \widetilde{B}_{j,i}| \le M_{R',\nu'} \frac{\tau^{\nu'}}{(\nu')!}. \tag{5.85}$$

Comparing (5.85) with Theorem 5.9, one finds that a sufficient condition for the convergence of the power method (using $\widetilde{B}$ computed from approximation $\widetilde{V}$) is that $\tau \le 1/c_f$ and that for some $\mu \in \{2,3,4,\ldots\}$ ($\mu = 2$ is the weakest condition)

$$M_{R',\nu'} \frac{\tau^{\nu'}}{(\nu')!} \le \Big[\min_{i \ne 1} |\overline{x}_i|^2\Big] \left(\frac{\delta c_f^5}{9(16) m_\sigma^2}\right) \frac{\tau^4}{\nu^\mu}.$$

Note that the $\tau \leq 1/c_f$ condition can be removed by using Theorems 5.7 and 5.8 instead of 5.9.

Because computation of $\widetilde{B}_{j,i}$ requires the maximization operation, below we will introduce an approximation for $\widetilde{B}_{j,i}$, to be denoted by $\widehat{B}_{j,i}$ (where the maximum may only be computed approximately rather than exactly). Suppose further that

$$|\widetilde{B}_{j,i} - \widehat{B}_{j,i}| \leq M_{R',\nu'} \frac{\tau^{\nu'}}{(\nu')!}. \tag{5.86}$$

Then, by (5.86), (5.85) with Theorem 5.9, one finds that a sufficient condition for the convergence of the power method (using $\widehat{B}$) is that $\tau \leq 1/c_f$ and that for some $\mu \in \{2, 3, 4, \ldots\}$ ($\mu = 2$ is the weakest condition)

$$2M_{R',\nu'} \frac{\tau^{\nu'}}{(\nu')!} \leq \left[\min_{i \neq 1} |\overline{x}_i|^2\right] \left(\frac{\delta c_f^5}{9(16) m_\sigma^2}\right) \frac{\tau^4}{\nu^\mu}, \tag{5.87}$$

and so, alternatively, a sufficient condition is

$$\tau^{\nu'-4} \leq \left[\min_{i \neq 1} |\overline{x}_i|^2\right] \left(\frac{\delta c_f^5 (\nu')!}{9(32) m_\sigma^2 M_{R',\nu'}}\right) \frac{1}{\nu^\mu}. \tag{5.88}$$

Suppose a rectangular grid of evenly spaced basis function centers with N_D centerpoints per dimension, and recall that ψ_1 is centered at the origin which implies N_D is odd. (Perhaps it should be noted that this is conservative in that we are considering a rectangular grid encompassing $\overline{B}_{D_R}$ rather than just those basis functions centered in the sphere itself.) This implies $\min_{i \neq 1} |\overline{x}_i|^2 = 4D_R^2/(N_D - 1)^2$, and (5.88) becomes

$$\tau^{\nu'-4} \leq \left(\frac{D_R^2 \delta c_f^5 (\nu')!}{9(8) m_\sigma^2 M_{R',\nu'}}\right) \left(\frac{1}{N_D}\right)^{n\mu} \left(\frac{1}{N_D - 1}\right)^2,$$

which implies a sufficient condition is

$$\tau^{\nu'-4} \leq \left(\frac{D_R^2 \delta c_f^5 (\nu')!}{9(8) m_\sigma^2 M_{R',\nu'}}\right) \left(\frac{1}{N_D}\right)^{n\mu+2} \doteq \widetilde{\mathcal{M}}_{R',\nu'} \left(\frac{1}{N_D}\right)^{n\mu+2}, \tag{5.89}$$

where we recall that n is the dimension of the state space.

Therefore, if one fixes $\tau < \min\{1, 1/c_f\}$, then it is sufficient that

$$\nu' \geq 4 + \frac{\log \widetilde{\mathcal{M}}_{R',\nu'} + (n\mu + 2) \log (1/N_D)}{\log \tau}. \tag{5.90}$$

Alternatively, one may, without loss of generality, require $\widetilde{\mathcal{M}}_{R',\nu'} \geq 1$ in which case (noting that $\log \tau < 0$ because $\tau < 1$) (5.90) yields the sufficient condition

$$\nu' \geq 4 + \frac{(n\mu + 2) \log (1/N_D)}{\log \tau} \tag{5.91}$$

in which case the lower bound on ν' scales like $\log (1/N_D)$. We remark that this sufficient condition may be quite conservative.

5.3.1 A Method for Computing B

As noted above, one would not typically have a closed-form expression for the $B_{j,i}$ or even the $\widetilde{B}_{j,i}$ terms, and we denote the approximation of $\widetilde{B}$ by $\widehat{B}$. In this subsection, we indicate some specifics of a numerical method for the approximation. This is not essential to the error analysis, but it seems appropriate to sketch an approximation technique so as to concretely indicate one approach to this subproblem.

This particular approach requires one to define

$$\widetilde{X}_{j,i}(t) \doteq \operatorname{argmax}\{\psi_j(x) - \widetilde{V}(x,t)\},$$

where $\widetilde{V}$ is given by (5.82) (i.e., the truncated power series expansion of $S_t[\psi_i](x)$), and then to propagate $\widetilde{X}_{j,i}$ as the solution of an ODE forward from $t = 0$ to τ via a Runge–Kutta method. One difficulty is that $\widetilde{X}_{j,i}(t)$ diverges as $t \downarrow 0$. In order to remedy this, and also remedy unbounded derivatives as $t \downarrow 0$, we replace $\psi_j(x)$ by $\psi^\tau_{j,i}(x,t)$ where

$$\psi^\tau_{j,i}(x,t) \doteq -\frac{1}{2}(x-\zeta_t)^T[(C + \overline{\delta}(1-t/\tau))I](x-\zeta_t), \tag{5.92}$$

where

$$\zeta_t \doteq x_i + (t/\tau)(x_j - x_i), \tag{5.93}$$

and $\overline{\delta} > 0$. Then one may define

$$\widetilde{X}^\tau_{j,i}(t) \doteq \operatorname*{argmax}_x\{\psi^\tau_{j,i}(x,t) - \widetilde{V}(x,t)\}, \tag{5.94}$$

and note that

$$\widetilde{X}^\tau_{j,i}(\tau) = \widetilde{X}_{j,i}(\tau) = \operatorname{argmax}\{\psi_j(x) - \widetilde{V}(x,\tau)\}.$$

Because $\widetilde{X}^\tau_{j,i}(t)$ is the argmax at each time $t \in [0,\tau]$, this implies

$$[\psi^\tau_{j,i}]_x(\widetilde{X}^\tau_{j,i}(t),t) - \widetilde{V}_x(\widetilde{X}^\tau_{j,i}(t),t) = 0$$

for all $t \in [0,\tau]$. Differentiating with respect to time, implies

$$\begin{aligned}\Big[[\psi^\tau_{j,i}]_{xx}(\widetilde{X}^\tau_{j,i}(t),t) - \widetilde{V}_{xx}(\widetilde{X}^\tau_{j,i}(t),t)\Big]\dot{\widetilde{X}}^\tau_{j,i}(t)&\\ +\Big[[\psi^\tau_{j,i}]_{tx}(\widetilde{X}^\tau_{j,i}(t),t) - \widetilde{V}_{tx}(\widetilde{X}^\tau_{j,i}(t),t)\Big] &= 0,\end{aligned}$$

or

$$\begin{aligned}\dot{\widetilde{X}}^\tau_{j,i}(t) = {}&\Big[[\psi^\tau_{j,i}]_{xx}(\widetilde{X}^\tau_{j,i}(t),t) - \widetilde{V}_{xx}(\widetilde{X}^\tau_{j,i}(t),t)\Big]^{-1}\\ &\Big[[\psi^\tau_{j,i}]_{tx}(\widetilde{X}^\tau_{j,i}(t),t) - \widetilde{V}_{tx}(\widetilde{X}^\tau_{j,i}(t),t)\Big].\end{aligned} \tag{5.95}$$

The initial state for (5.95) is

$$\begin{aligned}\widetilde{X}^{\tau}_{j,i}(0) &= \operatorname*{argmax}_x \left\{\psi^{\tau}_{j,i}(x,0) - \widetilde{V}(x,0)\right\} \\ &= \operatorname*{argmax}_x \left\{-\tfrac{1}{2}(x-x_i)^T(C+\overline{\delta}I)(x-x_i) - \psi_i(x)\right\} = x_i.\end{aligned}$$

Note that

$$\left[[\psi^{\tau}_{j,i}]_{xx}(x,0) - \widetilde{V}_{xx}(x,0)\right] = -[C+\overline{\delta}I] + C = -\overline{\delta}I, \tag{5.96}$$

which is negative definite, and

$$\left[[\psi^{\tau}_{j,i}]_{xx}(x,\tau) - \widetilde{V}_{xx}(x,\tau)\right] = -C - \widetilde{V}_{xx}(x,\tau) \tag{5.97}$$

would be negative definite on $\overline{B}_R$ by Assumption (A5.6I) if approximation $\widetilde{V}(\cdot,\tau)$ were replaced by $S_\tau[\psi_i]$. Also,

$$\widetilde{X}^{\tau}_{j,i}(0) = x_i \in \overline{B}_{D_R} \quad \text{and} \quad \widetilde{X}^{\tau}_{j,i}(\tau) \in \overline{B}_R \tag{5.98}$$

if approximation $\widetilde{V}(\tau,\cdot)$ is replaced by $S_\tau[\psi_i]$. This suggests the following assumption (which is only used for this approach to computing B). Suppose there exists $\hat{\delta} > 0$ such that

$$\left[[\psi^{\tau}_{j,i}]_{xx}(x,t) - \widetilde{V}_{xx}(x,t)\right] + \hat{\delta}I < 0 \quad \forall\, |x| \le \hat{g}(t),\ \forall t \in [0,\tau]$$

and

$$|\widetilde{X}^{\tau}_{j,i}(t)| \le \hat{g}(t) \quad \forall\, t \in [0,\tau] \tag{5.99}$$

where $g : [0,\tau] \to \mathbf{R}$ is any function such that $\hat{g}(0) = D_R$, $\hat{g}(\tau) = R$ and $\hat{g}$ is monotonically decreasing. Note that, by (5.96)–(5.98), the conditions are satisfied at both endpoints ($t = 0$ and $t = \tau$) when $\widetilde{V}(\tau,\cdot)$ is replaced by $S_\tau[\psi_i]$. Consequently, this may not be significantly more restrictive than the general assumptions, and for the purposes of sketching this particular approach to computing B, let us assume (5.99). Note that this guarantees the existence of the inverse in (5.95), and further that $\widetilde{X}^{\tau}_{j,i}(\tau)$ is the unique maximizer in $\overline{B}_R$.

Analytical expressions for the right-hand side of (5.95) can be obtained from (5.82) and (5.92). (These can be used to generate sufficient conditions that guarantee (5.99), but these are likely much too conservative.) Thus, one merely needs to propagate the n-dimensional ODE (5.95) forward to time τ. A Runge–Kutta method may be used for this, and the resulting approximate solution denoted by $\widehat{X}^{\tau}_{j,i}$. The approximation of the elements of $\widehat{B}$ are then given by

$$\begin{aligned}\widehat{B}_{j,i} &= -\left\{\psi^{\tau}_{j,i}(\widehat{X}^{\tau}(\tau)) - \widetilde{V}(\tau,\widehat{X}^{\tau}(\tau))\right\} \\ &= -\left\{\psi_j(\widehat{X}^{\tau}(\tau)) - \widetilde{V}(\tau,\widehat{X}^{\tau}(\tau))\right\}.\end{aligned} \tag{5.100}$$

Note the following:

The number of steps in the Runge–Kutta algorithm must be controlled so that (5.86) is satisfied. (5.101)

5.4 Error Summary

The error analyses of the previous three sections will now be combined. In particular, the errors due to truncation and the errors in computation of B will be combined to produce overall error bounds (5.109), (5.110). A condition required for the algorithm to work (assuming one uses the power series of Section 5.3 for computation of B) is also obtained.

Theorems 5.6 to 5.9 provided sufficient conditions for the power method step to converge to the max-plus eigenvector. Employing the simplest condition (but also the strictest), that of Theorem 5.9, convergence of the power method with approximation $\widehat{B}$ to B is guaranteed if

$$|\widehat{B}_{i,j} - B_{i,j}| \le \min_{i\neq 1}\{|\overline{x}_i|^2\}\left(\frac{\delta c_f^5}{9(16)m_\sigma^2}\right)\frac{\tau^4}{\nu^\mu} \quad \forall\, i,j, \tag{5.102}$$

where $\mu \in \{2,3,4,\dots\}$ and δ is given by (5.32). Note that we are assuming $\tau \le \min\{1, 1/c_f, \tau'\}$ as in Section 5.3 (as well as all assumptions including (A5.6I) and technical conditions (5.1), (5.19) which appear in Section 5.1). Then, Theorem 5.9 implies that a resulting error bound for the max-plus eigenvector given by

$$\|e-\hat{e}\| \doteq \max_i |e_i - \hat{e}_i| \le \min_{i\neq 1}\{|\overline{x}_i|^2\}\left(\frac{\delta c_f^5}{9(16)m_\sigma^2}\right)\frac{\tau^4}{\nu^{\mu-2}}, \tag{5.103}$$

where $\hat{e}$ corresponds to $\widehat{B}$. We remark that slightly different error estimates (under slightly different conditions) are also given in Theorems 5.7 and 5.8.

Suppose we adopt the notation $\widehat{W}(x) \doteq \bigoplus_{i=1}^{\nu} \hat{e}_i \otimes \psi_i(x)$ and $W^f(x) \doteq \bigoplus_{i=1}^{\nu} e_i \otimes \psi_i(x)$ so that W^f corresponds to the finite expansion with zero error in the computation/approximation of B. Then, by (5.103),

$$\begin{aligned}\|\widehat{W} - W^f\| &\doteq \max_{|x|\le R} |\widehat{W}(x) - W^f(x)| \le \min_{i\neq 1}\{|\overline{x}_i|^2\}\left(\frac{\delta c_f^5}{9(16)m_\sigma^2}\right)\frac{\tau^4}{\nu^{\mu-2}} \\ &= \min_{i\neq 1}\{|\overline{x}_i|^2\}\left(\frac{\delta c_f^5\tau^4}{9(16)m_\sigma^2}\right)\left(\frac{1}{N_D}\right)^{n(\mu-2)},\end{aligned} \tag{5.104}$$

where again, N_D is the number of centers of basis functions per dimension of the state space with a rectangular, evenly spaced grid of centers. It should be recalled that the basis functions are such that ψ_1 is centered at the origin ($\overline{x}_1 = 0$), and so N_D is odd. (Perhaps one should note that we are being sloppy here by using the number of basis functions corresponding to covering the entire rectangle which encloses the sphere $\overline{B}_{D_R}$, although only those with centers covering the sphere itself are required for the bound. Consequently, the above bound is a bit more conservative.) Also, with the evenly spaced basis function centers, (5.104) can be written as

$$\|\widehat{W} - W^f\| \leq \left(\frac{D_R^2 \delta c_f^5 \tau^4}{9(4)m_\sigma^2}\right)\left(\frac{1}{N_D}\right)^{n(\mu-2)}\left(\frac{1}{N_D - 1}\right)^2. \tag{5.105}$$

Using the approach of Section 5.3, (5.102) is satisfied if

$$\tau^{\nu'-4} \leq \widetilde{\mathcal{M}}_{R',\nu'}\left(\frac{1}{N_D}\right)^{n\mu+2} \tag{5.106}$$

where $\widetilde{\mathcal{M}}_{R',\nu'}$ is given by (5.89) and ν' is the number of terms (including zeroth order) in the Taylor series, **and** if (5.101) is satisfied.

This does not account for the truncation errors induced by using only a finite number of basis functions. Let W be the true value function. Then, by (5.78),

$$|W(x) - W^f(x)| \leq K_{1,\tau}\frac{|x|}{\tau}\left(\frac{2D_R}{N_D - 1}\right) \quad \forall x \in \overline{B}_R, \tag{5.107}$$

where $K_{1,\tau}$ is given by (5.79) (and we recall $K_{1,\tau} \to \infty$ as $\tau \downarrow 0$), $2D_R/(N_D - 1) = \varDelta$, and τ satisfies (5.77); τ now must satisfy

$$\tau \leq \min\{1, 1/c_f, 1/\overline{\lambda}, \tau'\}. \tag{5.108}$$

The error bound (5.107) is not without drawbacks. In particular, τ appears in the denominator. However, it does not seem possible with the techniques presented here to remove that term. This is the reason for concentrating in Section 5.3 on fixed τ with increasing ν' as the means for reducing errors.

Combining (5.105) and (5.107), the total error bound (assuming convergence of the power method — for which (5.106) and (5.101) form a sufficient condition — and $\tau \leq \min\{1, 1/c_f, 1/\overline{\lambda}, \tau'\}$) is given by

$$|W(x) - \widehat{W}(x)| \leq \left(\frac{D_R^2 \delta c_f^5 \tau^4}{9(4)m_\sigma^2}\right)\left(\frac{1}{N_D}\right)^{n(\mu-2)}\left(\frac{1}{N_D - 1}\right)^2 + K_{1,\tau}\frac{|x|}{\tau}\left(\frac{2D_R}{N_D - 1}\right),$$

which for $N_D \geq 3$

$$\leq \left(\frac{D_R^2 \delta c_f^5 \tau^4}{18 m_\sigma^2}\right)\left(\frac{1}{N_D}\right)^{n(\mu-2)+2} + K_{1,\tau}\frac{|x|}{\tau}\left(\frac{2D_R}{N_D}\right). \tag{5.109}$$

Because the best error rate is limited by the $1/N_D$ in the last term, we take $\mu = 2$, and find in that case

$$|W(x) - \widehat{W}(x)| \leq \left[\frac{D_R^2 \delta c_f^5 \tau^4}{18 m_\sigma^2} + 2K_{1,\tau} D_R \frac{|x|}{\tau} \right] \left(\frac{1}{N_D} \right). \tag{5.110}$$

That is, the total error goes down linearly in $(1/N_D)$. Note that this rate is constrained by the fact that the solutions are only viscosity solutions — which may have discontinuous first derivatives. It is conjectured that with smooth solutions, the rate would instead be $(1/N_D)^2$.

This assumes that conditions (5.106) and (5.101) are met as well as (5.108). Also, as in Section 5.3, one may prefer to write (5.106) as

$$\nu' \geq 4 + \frac{\log \widetilde{\mathcal{M}}_{R',\nu'} + (n\mu + 2)\log(1/N_D)}{\log \tau},$$

or assuming without loss of generality that $\widetilde{\mathcal{M}}_{R',\nu'} \geq 1$, one has the less tight but clearer bound of

$$\nu' \geq 4 + \frac{(n\mu + 2)\log(1/N_D)}{\log \tau}$$

in which case the lower bound on ν' scales like $\log(1/N_D)$. From this, one sees, for instance, that doubling N_D would typically imply the addition of

$$\lceil \left(\frac{(2n+2)\log(1/2)}{\log \tau} \right) = \lceil \left(\frac{(2n+2)\log 2}{\log(1/\tau)} \right)$$

to ν' where $\lceil(z)$ indicates the smallest integer greater than or equal to z. Again, this assumes the use of the Taylor series/Runge–Kutta approach of Section 5.3 toward the approximation of B. Alternate approaches may yield different conditions.

Remark 5.17. All error bounds are actually conceived as the errors that may be achieved with given computer effort. A key underlying assumption of this chapter is that all the elements of B are computed. This requires substantial effort because the number of terms in B is the square of the number of basis functions. In practice, it has been observed that elements of B, $B_{i,j}$, corresponding to basis function pairs where $|x_i - x_j|$ is large generally do not contribute *at all* to the resulting eigenvector (recall that this is the max-plus algebra). By not computing these terms, one can greatly reduce the computations. All examples computed by the author have used this computation reduction method. This is a question for further study.

5.5 Example of Convergence Rate

As an example, we consider the problem (4.1)–(4.4) in two dimensions with dynamics

$$\frac{d}{dt}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} -2\xi_1[1 + \frac{1}{2}\arctan(3\xi_2^2/2)] \\ \frac{1}{2}\xi_1 - 3\xi_2(e^{-\xi_1/3}) \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \tag{5.111}$$

The running cost is $l(x) = \frac{1}{2}|x|^2$ and $\gamma^2 = 1$. The corresponding HJB PDE is

$$0 = -\Big\{ -2x_1\left[1 + \tfrac{1}{2}\arctan(3x_2^2/2)\right] W_{x_1} + \left[x_1/2 - 3x_2 e^{-x_1/3}\right] W_{x_2} + \tfrac{1}{2}(x_1^2 + x_2^2) - \tfrac{1}{2}(W_{x_1}^2 + W_{x_2}^2)\Big\}$$

with usual boundary condition $W(0) = 0$.

The backsubstitution errors (obtained by substituting the approximated partials back into the PDE) for a basis function center spacing of 0.07 is depicted in the left plot of Figure 5.1. (The partials are approximated by simple first-order differences — possibly too coarse a method for examination of the errors.) This is a much higher density of basis functions than we normally use, and the code took roughly 15 seconds to run on a 2000 Sony Vaio laptop. The right plot of Figure 5.1 depicts the convergence as the spacing between basis function centers goes to zero (i.e., as the number of basis functions increases). In the left plot of Figure 5.1, one can see the effect of having a region where none of the basis functions in our supply of basis functions is the one that would achieve the solution there; this is evidenced by the sharp rise in errors in the right-hand corner. There are two error convergence plots in the right-hand graph. The lower one is the maximum backsubstitution error magnitude over the plotting region when one excises that region not properly covered by our basis function set. The upper curve includes this region not properly covered in its computation of the maximum error. Considering the lower curve, one sees that the convergence actually appears to be closer to second order than first order as predicted in Section 5.4. As noted above, one actually expects convergence to be second order in C^1 regions, and this is the effect we appear to see in the plot.

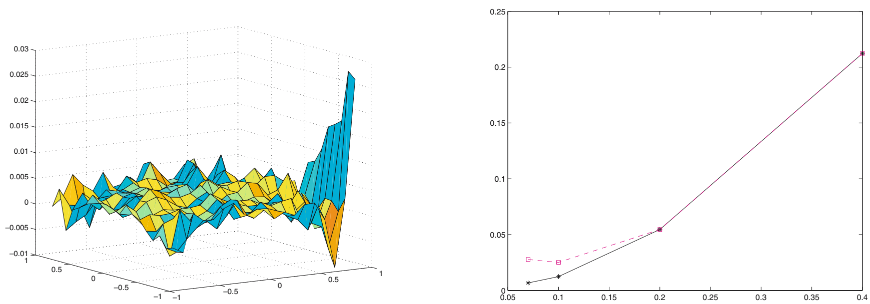


Fig. 5.1. Backsubstitution errors for one case and error convergence rate

6

A Semigroup Construction Method

In this chapter, we develop a different max-plus based approach to the solution of nonlinear control problems. We will again consider infinite time-horizon problems. In particular, we will use the same problem class considered in Chapters 4 and 5 as a vehicle for the development of the approach. Of course the approach is not limited to this particular class, but it will be convenient to work with the same class.

We continue to express the solution/value function in terms of a max-plus basis, and again the coefficients in this expansion will satisfy an eigenvector equation, $0 \otimes e = B \otimes e$. Recall that the main steps in the eigenvector algorithm (Chapters 4 and 5) were the computation of B and the solution of the eigenvector problem given B. It has been observed that the first step was typically an order of magnitude more computationally expensive than the second. (This was the case even when the technique for not computing the entire matrix B, indicated in Remarks 4.24 and 5.17, was used.) This motivates a search for alternative methods for computation of B.

Recall that $B = B_\tau$ operates on the discretization of the semiconvex dual of the value, W where W satisfies $W = S_\tau[W]$ (or $0 \otimes W = S_\tau[W]$). Suppose we have an indexed set of operators $B^m = B_\tau^m$ for $m \in \mathcal{M} \doteq \{1, 2, \ldots, M\}$, and let $\bar{B} = \bigoplus_{m\in\mathcal{M}} B^m$. (More generally one can let $\bar{B}$ be a general max-plus linear combination as $\bar{B} = \bigoplus_{m\in\mathcal{M}} c^m \otimes B^m$ for some coefficients $c^m \in \mathbf{R}^-$.) If each B^m is relatively easy to compute, then $\bar{B}$ is relatively easy to construct. We will see that $\bar{B}$ is related to $\bar{S}_\tau = \bigoplus_{m\in\mathcal{M}} S_\tau^m$. Suppose each S_τ^m corresponds to an HJB PDE problem $0 = H^m(x, \nabla W)$, $W(0) = 0$. Let $\widetilde{H}(x,p) \doteq \max_{m\in\mathcal{M}} H^m(x,p)$, and let the semigroup corresponding to $\widetilde{H}$ be $\widetilde{S}_\tau$. Then, for small $\tau > 0$, $\widetilde{S}_\tau$ may be closely approximated by $\bar{S}_\tau$. Thus solution of the eigenvector problem for $\bar{B}_\tau$ may closely resemble solution of $W = \widetilde{S}_\tau[W]$, and consequently, solution of HJB PDE problem $0 = \widetilde{H}(x, \nabla W)$, $W(0) = 0$.

The theory in support of the above discussion will be developed in this chapter. We will also indicate application with a simple example. A class of problems where this method is especially handy is the class of problems with a set of linear dynamics coupled to a switching controller. That is, where one has $\dot{\xi} = A^{\mu_t}\xi_t + \sigma^{\mu_t}u_t$ where the μ_t process will take values in some finite set, say $\{\mu^m | \, m \in \mathcal{M}\}$, while u_t will be an L_2 control process. In such problems, the HJB PDE is nonlinear, and solution of the problem cannot typically be obtained by piecewise pasting together of solutions of the constituent problems. Of course, one could apply this approach to a wider class of problems as well.

As in Chapter 4, we will make the unwarranted assumption that the value has a finite max-plus basis expansion (i.e., requiring only a finite number of nonzero terms). This will again allow us to quickly obtain the algorithm. However, in this case, we will not generate an error analysis analogous to that in Chapter 5 as inclusion of such analyses for each max-plus algorithm and/or application would be excessively long.

6.1 Constituent Problems

We will be solving a nonlinear HJB PDE problem

$$\begin{aligned} 0 &= \widetilde{H}(x, \nabla W), \\ W(0) &= 0, \end{aligned} \tag{6.1}$$

where, as always, $W(0) = 0$ is the boundary data indicating that W is zero at the origin. We will be interested in $\widetilde{H}$ that take the form

$$\begin{aligned} \widetilde{H}(x, \nabla\phi) &\doteq \max_{m\in\mathcal{M}} \sup_{u\in\mathbf{R}^l} \left[(f^m(x) + \sigma^m(x)u)^T \nabla\phi + l^m(x) - \frac{\gamma^2}{2}|u|^2 \right] \\ &= \max_{m\in\mathcal{M}} (f^m(x))^T \nabla\phi + l^m(x) + \frac{1}{2\gamma^2} \nabla\phi^T \sigma^T(x)\sigma(x)\nabla\phi \end{aligned} \tag{6.2}$$

where $\mathcal{M} = \{1, 2, \ldots, M\}$ and $M < \infty$. One may also want to consider $\widetilde{H}$ that are only *approximated* by a finite maximization of this form, but that is a problem for future research. Define the constituent Hamiltonians as

$$\begin{aligned} H^m(x, \nabla\phi) &\doteq \sup_{u\in\mathbf{R}^l} \left[(f^m(x) + \sigma^m(x)u)^T \nabla\phi + l^m(x) - \frac{\gamma^2}{2}|u|^2 \right] \\ &= (f^m(x))^T \nabla\phi + l^m(x) + \frac{1}{2\gamma^2} \nabla\phi^T \sigma^T(x)\sigma(x)\nabla\phi \end{aligned}$$

so that

$$\widetilde{H} = \max_{m\in\mathcal{M}} H^m.$$

All of the constituent H^m will correspond to constituent control problems, and will also have associated semigroups S_τ^m and their duals $B^m = B_\tau^m$. These constituent problems should be ones such that the B^m are relatively easily computable. An obvious class consists of linear/quadratic problems.

Remark 6.1. Perhaps we should remark here that it is the interplay of the max-plus and standard algebras that makes this approach (as well as that of Chapter 7) possible. Linear problems are particularly simple in the standard algebra. However, standard-sense linear combinations of linear problems only yield linear problems. However, max-plus linear combinations of standard-sense linear problems yields a larger class of problems.

Although the obvious class of constituent problems are the linear/quadratic problems, the analysis will be kept at a more general level. In particular, consider a set of contituent system dynamics indexed by $m \in \mathcal{M}$ and given by

$$\begin{aligned} \dot{\xi}^m &= f^m(\xi^m) + \sigma^m(\xi^m)u, \\ \xi_0^m &= x \in \mathbf{R}^n, \end{aligned} \tag{6.3}$$

where we note that all the systems have the same initial condition, and $u \in \mathcal{U}$ will be a (payoff-maximizing) input. To be specific, we continue to take $\mathcal{U} = L_2^{\mathrm{loc}}([0,\infty);\mathbf{R}^l)$. As noted above, the underlying concept being developed in this chapter could be applied to a wide range of systems. However, to be specific we will use the same class as considered in Chapters 4 and 5. Consequently, we assume the following.

Each of the f^m and σ^m satisfy (A4.1I) and (A4.2I), respectively. In particular, the constants K, c_f, K_σ and m_σ are independent of m. (A7.1I)

We consider constituent payoffs and values given by

$$J_\tau^m(x,u) \doteq \int_0^\tau l^m(\xi_t^m) - \frac{\gamma^2}{2}|u_t|^2\,dt \tag{6.4}$$

$$\begin{aligned} W^m(x) &\doteq \sup_{\tau\in(0,\infty)} \sup_{u\in L_2(0,\tau)} J_\tau^m(x,u) = \lim_{\tau\to\infty} \sup_{u\in L_2(0,\tau)} J_\tau^m(x,u) \\ &= \sup_{u\in\mathcal{U}} \sup_{\tau\in(0,\infty)} J_\tau^m(x,u), \end{aligned} \tag{6.5}$$

where ξ^m satisfies (6.3). We also assume that

Each of the l^m satisfy (A4.3I) where the constants C_l and α_l are independent of m. Further, the constants satisfy (A4.4I). (A7.2I)

The corresponding (max-plus linear) semigroups are again given by

$$S_\tau^m[\phi](x) \doteq \sup_{u\in\mathcal{U}} \left\{ \int_0^\tau l^m(\xi_t^m) - \frac{\gamma^2}{2}|u_t|^2\,dt + \phi(\xi_\tau^m) \right\}(x), \tag{6.6}$$

where ξ^m satisfies (6.3). The corresponding max-plus eigenvector problems are

$$0 \otimes \phi = S_\tau^m[\phi], \tag{6.7}$$

where it is implicit that we are looking for solutions satisfying $\phi(0) = 0$ (which of course eliminates the nonuniqueness due to max-plus multiplication by a constant). The corresponding HJB PDE problems are given by

$$0 = H^m(x, \nabla\phi), \tag{6.8}$$
$$\phi(0) = 0. \tag{6.9}$$

Let

$$\mathcal{C}_\delta \doteq \left\{ \text{semiconvex } \phi : \mathbf{R}^n \to \mathbf{R}^- \,\middle|\, 0 \le \phi(x) \le c_f \frac{(\gamma-\delta)^2}{m_\sigma^2} |x|^2 \;\forall x \in \mathbf{R}^n \right\}.$$

From Chapters 3 and 4 (see Theorem 4.1, Theorem 4.4, and Corollary 4.11), we have the following.

Theorem 6.2. *Suppose Assumptions (A7.1I), (A7.2I) hold. For fixed* $\delta > 0$ *sufficiently small, each value function* W^m *(given by (6.5)) is the unique solution in* $\mathcal{C}_\delta$ *of the corresponding HJB PDE problem (6.8), (6.9). Further, each* W^m *is also the unique solution in* $\mathcal{C}_\delta$ *of the corresponding fixed-point/max-plus eigenvector problem (6.7).*

We will be interested in solving our original problem (with Hamiltonian $\widetilde{H}$) over some ball centered at the origin. Fix any $R > 0$. Then there exist $c', L \in (0, \infty)$ (independent of m) such that

$$W^m \in \mathcal{S}_{\bar{R}}^{c'L} \quad \forall m \in \mathcal{M}. \tag{6.10}$$

(We will adjust the values of c', L, δ below, but the presentation is improved by delaying this point.) Note that we abuse notation by using the notation W^m for both the value function and its restriction to $\overline{B}_R$. Let C be a symmetric matrix satisfying $C - c'I > 0$. For $x_i \in \mathbf{R}^n$, define

$$\psi_i(x) \doteq -\tfrac{1}{2}(x - x_i)^T C (x - x_i).$$

Recall from Theorem 2.13 that if a set $\{x_i : i \in \mathcal{N}\}$ form a countable dense subset of $\mathcal{E} = \{\overline{x} \in \mathbf{R}^n : \overline{x}^T (C^2) \overline{x} \le (L + |C|R)^2\}$, Then, the set $\{\psi_i : i \in \mathcal{N}\}$ is a countable basis for max-plus vector space $\mathcal{S}_{\bar{R}}^{c'L}$. Further, the coefficients in the expansion of any $\phi \in \mathcal{S}_{\bar{R}}^{c'L}$ are given by (2.20). That is,

$$\phi(x) = \bigoplus_{i=1}^{\infty} [a_i \otimes \psi_i(x)],$$

where

$$a_i = -\max_{x \in \overline{B}_R} [\psi_i(x) - \phi(x)] \quad \forall i.$$

Throughout this chapter, we will assume that such available storage functions actually have *finite* max-plus expansions, i.e., that

$$W^m(x) = \bigoplus_{i=1}^{\nu} [a_i^m \otimes \psi_i(x)] \quad \forall x \in \overline{B}_R \tag{6.11}$$

for some $\nu < \infty$. Note that we used this same assumption in Chapter 4. However, as noted above, an error analysis analogous to that in Chapter 5 will not be included.

6.2 Operating on the Transformed Operators

We are not concerned here with a direct numerical method for solution of HJB PDEs based on the max-plus eigenvector problem solution, but rather on the construction of B matrices for max-plus eigenvector problems from other matrices whose max-plus eigenvector problems are analytically tractable, and the relationship of the constructed matrices to corresponding HJB PDEs. The following theorem makes a critical connection between the problems over the corresponding domains.

Theorem 6.3. *Let S_τ^m be defined by (6.6) for each m in some finite set $\mathcal{M}$. Suppose that for each $j \in \{1, 2, \ldots, \nu\}$ and each $m \in \mathcal{M}$, there exists a finite basis expansion of $S_\tau^m[\psi_j]$, i.e., that*

$$S_\tau^m[\psi_j](x) = \bigoplus_{i=1}^{\nu} B_{i,j}^m \otimes \psi_i(x) \quad \forall\, x \in \overline{B}_R. \tag{6.12}$$

Define $\bar{S}_\tau[\phi]$ for any ϕ in the domain (to be specified for specific problems below) by

$$\bar{S}_\tau[\phi](x) = \sup_{u \in \mathcal{U}} \left\{ \max_{m \in \mathcal{M}} \left[\int_0^\tau l^m(\xi_t^m) - \frac{\gamma^2}{2}|u_t|^2\, dt + \phi(\xi_\tau^m) \right] \right\} \tag{6.13}$$

for all $x \in \overline{B}_R$ where ξ^m satisfies (6.3). Then

$$\bar{S}_\tau[\psi_j](x) = \bigoplus_{i=1}^{\nu} \bar{B}_{i,j} \otimes \psi_i(x) \tag{6.14}$$

for all $x \in \overline{B}_R$ where

$$\bar{B}_{i,j} = \max_{m \in \mathcal{M}} B_{i,j}^m = \bigoplus_{m \in \mathcal{M}} B_{i,j}^m \tag{6.15}$$

for all $i, j \in \{1, 2, \ldots, \nu\}$.

Proof. The proof is a simple manipulation given by

$$\begin{aligned}
\bar{S}_\tau[\psi_j](x) &= \sup_{u \in \mathcal{U}} \left\{ \max_{m \in \mathcal{M}} \left[J_\tau^m(x, u) + \psi_j(\xi_\tau^m) \right] \right\} \\
&= \max_{m \in \mathcal{M}} S_\tau^m[\psi_j](x) = \max_{m \in \mathcal{M}} \max_{i \in \{1,2,\ldots,\nu\}} B_{i,j}^m \otimes \psi_i(x) \\
&= \max_{i \in \{1,2,\ldots,\nu\}} \left[\max_{m \in \mathcal{M}} B_{i,j}^m \right] \otimes \psi_i(x) = \bigoplus_{i=1}^{\nu} \bar{B}_{i,j} \otimes \psi_i(x). \ \square
\end{aligned}$$

By simple modifications of the proofs of Theorem 4.4 and Corollary 4.11, one also has the following.

Theorem 6.4. *Suppose Assumptions (A7.1I), (A7.2I) hold. For sufficiently small $\delta > 0$, Then there is a unique solution in $\mathcal{C}_\delta$ of $\overline{W} = \bar{S}_\tau[\overline{W}]$.*

Remark 6.5. If necessary, we adjust the δ of Theorem 6.2 so that Theorem 6.4 also holds. Further, if necessary we adjust c', L so that $\overline{W} \in \mathcal{S}_{\overline{R}}^{c'L}$.

Corollary 6.6. *Suppose Assumptions (A7.1I), (A7.2I) hold. Further, assume that the expansion for $\overline{W}$ is finite with ν coefficients which we denote as $\overline{W} = \bigoplus_{i=1}^{\nu} \bar{e}_i \otimes \psi_i$. Also assume that each ψ_i is active in the sense that $\bigoplus_{i \neq j} \bar{e}_i \otimes \psi_i \neq \bigoplus_{i=1}^{\nu} \bar{e}_i \otimes \psi_i$ for any $j \leq \nu$. Then the vector of coefficients, $\bar{e}$ is the solution of the max-plus eigenvector equation $\bar{e} = \bar{B} \otimes \bar{e}$ where $\bar{B}_{i,j} = \bigoplus_{m \in \mathcal{M}} B_{i,j}^m$ for all i, j.*

Proof. By assumption, for all $x \in \overline{B}_R$

$$\begin{aligned}\bigoplus_{i=1}^{\nu} \bar{e}_i \otimes \psi_i(x) &= \overline{W}(x) \\ &= \bar{S}_\tau[\overline{W}](x) = \bar{S}_\tau\left[\bigoplus_{j=1}^{\nu} \bar{e}_j \otimes \psi_j\right](x) = \bigoplus_{j=1}^{\nu} \bar{e}_j \otimes \bar{S}_\tau[\psi_j](x),\end{aligned}$$

which by Theroem 6.3

$$= \bigoplus_{j=1}^{\nu} \bar{e}_j \otimes \left[\bigoplus_{i=1}^{\nu} \bar{B}_{i,j} \otimes \psi_i(x)\right] = \bigoplus_{i=1}^{\nu}\left[\bigoplus_{j=1}^{\nu} \bar{B}_{i,j} \otimes \bar{e}_j\right] \otimes \psi_i(x).$$

Using the assumption that all the ψ_i are active, this implies that $\bar{e}_i = \bigoplus_{j=1}^{\nu} \bar{B}_{i,j} \otimes \bar{e}_j$ $\forall i$, or equivalently, $\bar{e} = \bar{B} \otimes \bar{e}$. □

6.3 The HJB PDE Limit Problems

Now suppose that instead of desiring to solve for fixed points of the semigroups, one desires to solve related HJB PDEs. Consider the sets of measurable processes with values in $\mathcal{M}$ given by

$$\mathcal{D}_\infty = \{\mu : [0, \infty) \to \mathcal{M} \mid \text{measurable}\}$$

and

$$\mathcal{D}_T = \{\mu : [0, T) \to \mathcal{M} \mid \text{measurable}\}.$$

Then by standard dynamic programming results under typical assumptions (cf. [78], [79], [86], [88]), one obtains the following theorem. A specific example of a class of dynamics, cost and set $\mathcal{C}$ is given in the remark just below the theorem statement.

Again by simple modifications of the proofs of Theorem 4.1, Theorem 4.4, and Corollary 4.11, we have the following. Note that we are implicitly adjusting the values of δ, c', L if necessary.

Theorem 6.7. *Suppose Assumptions (A7.1I), (A7.2I) hold. There exists a unique solution in $\mathcal{C}_\delta$ of PDE (6.1), and this viscosity solution is also the unique solution in $\mathcal{C}_\delta$ of*

$$\widetilde{W} = \widetilde{S}_\tau[\widetilde{W}] \tag{6.16}$$

where

$$\widetilde{S}_\tau[\phi](x) \doteq \sup_{\mu\in\mathcal{D}_\tau} \sup_{u\in\mathcal{U}} \left\{ \int_0^\tau l^{\mu_t}(\widetilde{\xi}_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + \phi(\widetilde{\xi}_\tau) \right\}, \tag{6.17}$$

$$\dot{\widetilde{\xi}} = \tilde{f}(\widetilde{\xi}, u, \mu) \doteq f^{\mu_t}(\widetilde{\xi}_t) + \sigma^{\mu_t}(u_t). \tag{6.18}$$

This solution is given by

$$\widetilde{W}(x) = \sup_{\mu\in\mathcal{D}_\infty} \sup_{u\in\mathcal{U}} \sup_{T<\infty} \int_0^T l^{\mu_t}(\widetilde{\xi}_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \tag{6.19}$$

where $\widetilde{\xi}$ *satisfies (6.18).*

Note that the operators $\bar{S}_\tau$ do not necessarily form a semigroup, although they do form a sub-semigroup (i.e., $\bar{S}_{\tau_1+\tau_2}[\phi](x) \le \bar{S}_{\tau_1}\bar{S}_{\tau_2}[\phi](x)$ for all x and all ϕ in the domain). Further, it is easily seen that $S_\tau^m \le \bar{S}_\tau \le \widetilde{S}_\tau$ for all m.

With τ acting as a time-discretization step-size, let

$$\mathcal{D}_\infty^\tau = \Big\{ \mu : [0,\infty) \to \mathcal{M} \,\Big|\, \text{for each } n \in \mathbf{N}\cup\{0\}, \text{ there exists } m_n \in \mathcal{M} \\ \text{such that } \mu(t) = m_n \text{ for } t \in [n\tau, (n+1)\tau) \Big\},$$

and for $T = \bar{n}\tau$ with $\bar{n} \in \mathbf{N}$ define $\mathcal{D}_T^\tau$ similarly but with domain $[0, T)$ rather than $[0,\infty)$. Let $\mathcal{M}^N$ denote the outer product of $\mathcal{M}$, N times. Let $T = \bar{n}\tau$, and define

$$\bar{\bar{S}}_T^\tau[\phi](x) = \max_{\{m_k\}_{k=0}^{\bar{n}-1}\in\mathcal{M}^N} \left\{ \prod_{k=0}^{\bar{n}-1} S_\tau^{m_k} \right\} [\phi](x)$$

where the $\prod$ notation indicates operator composition.

Roughly speaking the following theorem simply states that any nearly optimal (worst case) $u \in \mathcal{D}_T$ can be arbitrarily closely approximated (in terms of the cost) by a piecewise constant $u \in \mathcal{D}_T^\tau$ for some small τ.

Theorem 6.8. *Suppose that for any $x \in \mathbf{R}^n$, the origin lies in the interior of the convex hull of the set $\{f^m(x)\}_{m\in\mathcal{M}}$ and that Assumptions (A7.1I), (A7.2I) hold. Given $T < \infty$, $R < \infty$ and $\varepsilon > 0$, there exists $N \in \mathbf{N}$ sufficiently large such that letting $\tau = T/N$, one has $\widetilde{S}_T[W^m](x) - \varepsilon \le \bar{\bar{S}}_T^\tau[W^m](x)$ for all $x \in \overline{B}_R$ and all $m \in \mathcal{M}$.*

Sketch of Proof. The proof is heuristically clear, but technically complex; we present only a sketch of the main points. We note that the first assumption is essentially some sort of technical controllability assumption which is sufficient for the proof, but may not be necessary. The first step is to work with simple integrals. Consider some $g^m \in \mathbf{R}^n$ for all $m \in \mathcal{M}$, and suppose $0 \in \langle \{g^m\}_{m\in\mathcal{M}}\rangle^\circ$ (i.e., that the origin is in the interior of the convex hull of the set of the g^m where the $\circ$ superscript denotes interior and the angle brackets denote convex hull). This guarantees that for any $\widehat{T} < \infty$ and $\Delta_0 \in \mathbf{R}^n$, there exists $L < \infty$ and $\{\lambda^0_m\}_{m\in\mathcal{M}}$ such that $\lambda^0_m \in [0,1]$ for all m, $\sum_m \lambda^0_m = 1$ and $\Delta_0 = L\widehat{T}\sum_{m\in\mathcal{M}} \lambda^0_m g^m$. Then, given $\varepsilon > 0$, $\widehat{T} \in (0,\infty)$, $\mu \in \mathcal{D}_T$ and $\Delta_0, \{g^m\}, L$ as above, there exist $N < \infty$, $\tau = \widehat{T}/N$ and $\overline{\mu} \in \mathcal{D}^\tau_{\widehat{T}}$ such that

$$\left|\int_0^{\widehat{T}} g^{\overline{\mu}_t}\,dt - \left[\Delta_0 + \int_0^{\widehat{T}} g^{\mu_t}\,dt\right]\right| < \varepsilon + \frac{L}{L+1}|\Delta_0|. \tag{6.20}$$

In particular, one has $\int_0^{\widehat{T}} g^{\mu_t}\,dt = \widehat{T}\sum_{m\in\mathcal{M}} \lambda^1_m g^m$ for appropriate coefficients λ^1_m (where $\lambda^1_m \in [0,1]$ for all m, $\sum_m \lambda^1_m = 1$). The $g^{\overline{\mu}_t}$ process is created by setting the time-steps where $\overline{\mu}_t = m$ for each m to approximate the fraction of time needed according to the $\{\lambda^0_m\}$ and $\{\lambda^1_m\}$ allocations, and this yields (6.20). One then approximates (6.18) over $[0,T)$ by holding the $\widetilde{\xi}$ terms on the right-hand side constant over each $[\widetilde{N}\widehat{T}, (\widetilde{N}+1)\widehat{T})$ for $\widetilde{N} \in \{0,1,\ldots,\widehat{N}-1\}$ where $\widehat{N}\widehat{T} = T$ as

$$\dot{\widehat{\widetilde{\xi}}}_t = f^{\mu_t}(\widetilde{\xi}_{\widetilde{N}\widehat{T}}) + \sigma^{\mu_t}(\widetilde{\xi}_{\widetilde{N}\widehat{T}})u_t. \tag{6.21}$$

For $\widehat{N}$ sufficiently large, $\max_{\widetilde{N}} |\widehat{\widetilde{\xi}}_{N\widehat{T}} - \widetilde{\xi}_{N\widehat{T}}|$ can be made arbitrarily small. (Because there is no a priori bound on $\|u\|$ in this formulation, an L_2 bound on near-optimal u processes which holds under Assumptions (A7.1I), (A7.2I) is used here; see Chapter 3.) A similar discretization approximation is employed with the state driven by the $\overline{\mu}$ process. Note that (6.21) takes the form $\dot{\widehat{\widetilde{\xi}}}_t = g_1^{\mu_t} + g_2^{\mu_t}u_t$. The final discretization of $[0,T]$ is then with time-step $\tau = (T/\widehat{N})/N$. $\square$

Now note that because $W^m, \overline{W} \in \mathcal{C}$, one has $\forall\, m$ (see proof of Theorem 3.20)

$$\lim_{T\to\infty} \widetilde{S}_T[W^m] = \widetilde{W}, \quad \lim_{T\to\infty} \widetilde{S}_T[\overline{W}] = \widetilde{W}. \tag{6.22}$$

Also, for all $T < \infty$,

$$\widetilde{W} = \widetilde{S}_T[\widetilde{W}] = \lim_{T\to\infty} \widetilde{S}_T[\overline{W}] \tag{6.23}$$

uniformly on compact sets. By (6.22) and (6.23), given $R < \infty$ and $\varepsilon > 0$, there exists $\widehat{T} < \infty$ such that for all $T \ge \widehat{T}$ and all $m \in \mathcal{M}$,

$$\widetilde{S}_T[\widetilde{W}](x) - \varepsilon \le \widetilde{S}_T[W^m](x) \quad \forall\, x \in \overline{B}_R. \tag{6.24}$$

Also note that

$$\overline{W} = \bar{S}_\tau[\overline{W}] = \prod_{k=0}^{n-1} \bar{S}_\tau[\overline{W}] \ge \prod_{k=0}^{n-1} S^m_\tau[\overline{W}].$$

which using the fact that $\overline{W} \ge 0$ from Theorem 6.4 and the monotonicity of $S^m_\tau[\cdot]$

$$\geq \prod_{k=0}^{n-1} S_\tau^m[0]$$

where 0 represents the function identically equal to zero. Because this is true for all n, one has (using Theorem 2.6, [88])

$$\overline{W} \geq \lim_{n\to\infty} S_{n\tau}^m[0] = W^m \tag{6.25}$$

for any $m \in \mathcal{M}$. On the other hand,

$$\overline{W} = \prod_{k=0}^{n-1} \bar{S}_\tau[\overline{W}] \leq \prod_{k=0}^{n-1} \widetilde{S}_\tau[\overline{W}] = \widetilde{S}_{n\tau}[\overline{W}]$$

which implies (using (6.22))

$$\overline{W} \leq \lim_{T\to\infty} \widetilde{S}_T[\overline{W}] = \widetilde{W}. \tag{6.26}$$

Combining (6.25) and (6.26), one has

$$W^m \leq \overline{W} \leq \widetilde{W} \quad \forall\, m \in \mathcal{M}. \tag{6.27}$$

Also, by definition it is obvious that

$$\bar{\bar{S}}_T^\tau[\phi] \leq \widetilde{S}_T[\phi] \quad \forall\, \phi \in \mathcal{C}. \tag{6.28}$$

Now, by Theorem 6.8 and (6.24), given $R < \infty$ and $\varepsilon > 0$, there exist $T < \infty$ and $\bar{n} < \infty$ such that with $\tau = T/\bar{n}$, one has

$$\widetilde{W}(x) - 2\varepsilon = \widetilde{S}_T[\widetilde{W}](x) - 2\varepsilon \leq \bar{\bar{S}}_T^\tau[W^m](x)$$

which by (6.27), (6.28) and the monotonicity of $\widetilde{S}_T[\cdot]$

$$\leq \bar{\bar{S}}_T^\tau[\overline{W}](x) \leq \widetilde{S}_T[\overline{W}](x) \leq \widetilde{S}_T[\widetilde{W}](x) = \widetilde{W}(x)$$

$\forall x \in \overline{B}_R$. Because $\overline{W}(x) = \bar{S}_\tau[\overline{W}](x) = (\bar{S}_\tau)^{\bar{n}}[\overline{W}](x) = \bar{\bar{S}}_T^\tau[\overline{W}](x)$ on $\overline{B}_R$, this implies

Theorem 6.9. *Given $R < \infty$ and $\varepsilon > 0$, there exists $\tau > 0$ such that*

$$\widetilde{W}(x) - 2\varepsilon \leq \overline{W}(x) \leq \widetilde{W}(x) \quad \forall\, x \in \overline{B}_R \tag{6.29}$$

where $\widetilde{W}$ and $\overline{W}$ satisfy $\widetilde{W} = \widetilde{S}_\tau[\widetilde{W}]$ and $\overline{W} = \bar{S}_\tau[\overline{W}]$.

Recall from Corollary 6.6, that under the conditions given there,

$$\overline{W}(x) = \bigoplus_{j=1}^{\nu} \bar{e}_j \otimes \psi_j(x) \quad \forall\, x \in \overline{B}_R,$$

where the vector of coefficients, $\bar{e}$, is the solution of the max-plus eigenvector equation $\bar{e} = \bar{B} \otimes \bar{e}$ with $\bar{B} = \bigoplus_{m\in\mathcal{M}} B^m$. Thus Theorem 6.9 implies that one can (approximately) solve HJB PDE (6.2) by solution of this eigenvector equation. If the B^m are such that they are easily computed (say by Riccati equations), then one has a method for computation of (approximate) solutions of nonlinear HJB PDEs of the form (6.2) (or those that can be closely approximated by HJB PDEs of that form) where the most difficult portion of the computation, that of $\bar{B}$, can be greatly simplified by representation of $\bar{B}$ as a max-plus sum of the B^m.

6.4 A Simple Example

As indicated above, a useful direction for application of this transform approach is as follows. One can solve simple linear/quadratic control problems through solution of the corresponding Riccati equations. Given solutions of the Riccati equations, one can construct the transformed operator (typically the discretized version thereof), B, analytically with little effort. Thus, it is natural to consider HJB PDEs that can be represented or approximated by maxima of HJB PDEs corresponding to linear/quadratic problems. The simplest example of this approach will be discussed here to give the reader some flavor of a transformed operator construction approach.

Consider the HJB PDE over $\mathbf{R}^n$ given by

$$0 = \max_{m\in\mathcal{M}} [(A^m x)^T \nabla W] + \tfrac{1}{2} x^T D x + \tfrac{1}{2} \nabla W^T \Sigma \nabla W, \tag{6.30}$$

where $\Sigma \doteq \frac{1}{\gamma^2}\sigma\sigma^T$, and we assume D symmetric, positive definite, and Assumptions (A7.1I), (A7.2I) with $f^m(x) = A^m x$, $\sigma^m(x) = \sigma$, $l^m(x) = \frac{1}{2}x^T D x$ for all m. (Also, as above, we assume that $\mathcal{M}$ is a finite set.) The goal here is to demonstrate the mechanics of a procedure for solution of (6.30). HJB PDE (6.30) corresponds to a control problem which has both an unknown L_2 disturbance process and a switching disturbance process given by

$$\dot{\xi}_t = A^{\mu_t}\xi_t + \sigma u_t, \qquad \xi_0 = x \in \mathbf{R}^n \tag{6.31}$$

$$W(x) = \sup_{\tau\in[0,\infty)} \sup_{u\in L_2} \sup_{\mu\in\mathcal{D}_\infty} \int_0^\tau \left[\tfrac{1}{2}\xi_t^T D \xi_t - \frac{1}{2\gamma^2}|u_t|^2 \right] dt. \tag{6.32}$$

For each $m \in \mathcal{M}$, the corresponding HJB PDE and semigroups are

$$0 = (A^m x)^T \nabla W + \tfrac{1}{2} x^T D x + \tfrac{1}{2} \nabla W^T \Sigma \nabla W \tag{6.33}$$

and

$$S_\tau^m[\psi_i](x) = \sup_{u\in L_2} \left\{ \int_0^\tau [\tfrac{1}{2}\xi_t^{mT} D \xi_t^m - \frac{1}{2\gamma^2}|u_t|^2]\, dt + \psi_i(\xi_\tau^m) \right\}$$

$$\dot{\xi}_t^m = A^m \xi_t^m + \sigma u_t, \qquad \xi_0 = x. \tag{6.34}$$

The solutions will be denoted by $V_i^m(0,x) = S_\tau^m[\psi_i](x)$ where the $V_i^m : [0,\tau]\times \mathbf{R}^n$. Letting $\psi_i(x) = -\frac{1}{2}(x-x_i)^T C(x-x_i)$, one may assume (without loss of generality) that the V_i^m take the form $V_i^m(t,x) = \frac{1}{2}(x-\Lambda_t^m x_i)^T Q_t^m (x - \Lambda_t^m x_i) + \frac{1}{2}x_i^T R_t^m x_i$. One then finds terminal conditions

$$Q_\tau^m = C, \quad \Lambda_\tau^m = I, \quad R_\tau^m = 0$$

and ordinary differential equations (ODEs)

$$\dot{Q}^m = \Big[D + Q^m \Sigma Q^m - \big((A^m)^T Q^m + Q^m A^m\big)\Big]$$

$$\dot{\Lambda}^m = \big(A^m - Q^{m-1}D\big)\Lambda, \qquad \dot{R}^m = -\Lambda^{mT} D \Lambda^m.$$

It is important to note that none of these ODEs depend on x_i. They only need to be solved once for each $m \in \mathcal{M}$. Noting that $S_\tau^m[\psi_i](x) = V_i^m(0,x)$, and that one then has $B_{i,j}^m = -\max_{x\in B_R}\big\{\psi_i(x) - V_j^m(0,x)\big\}$, one can show that

$$B_{i,j}^m = q_i^1 + q_j^2 + \gamma_i^T l_j^1 - \alpha_i^T l_j^2, \tag{6.35}$$

where we drop the m superscripts for notational simplicity, and where

$$q_i^1 = \tfrac{1}{2}(\alpha_i^T C\alpha_i - \gamma_i^T Q_0 \gamma_i^T), \qquad q_i^1 = \tfrac{1}{2}(\beta_j^T C\beta_j - \delta_j^T Q_0 \delta_j^T) + \tfrac{1}{2}x_j^T R_0 x_j$$

$$l_j^1 = q_0\delta_j, \qquad l_i^2 = C\beta_i$$

$$\alpha_i = [(C-Q_0)^{-1}C - I]x_i, \qquad \gamma_i = (C-Q_0)^{-1}Cx_i$$

$$\beta_j = (C-Q_0)^{-1}Q_0\Lambda_0 x_j, \qquad \delta_j = [(C-Q_0)^{-1}Q_0 + I]x_j.$$

Note that the only computation that needs to be done for all pairs (i,j) is (6.35). This implies that only $4n+1$ floating point operations need be performed for each of the ν^2 *pairs* (i,j) where ν is the number of basis functions in the truncated expansion. The other operations above are performed only once for each of the ν *single indices* i. (In practice, it has been observed that one does not need to compute $B_{i,j}^m$ for all pairs (i,j). The solution obtained by computing only those $B_{i,j}^m$ such that $x_i - x_j$ is relatively small is identical to the solution obtained by computing the entire matrix. That the solution is identically the same rather than merely "close" is a typical property in idempotent algebras. This is also noted in Remarks 4.24 and 5.17.) One then obtains

$$\bar{B}_{i,j} = \max_{m\in\mathcal{M}} B_{i,j}^m \quad \forall\, i,j.$$

Lastly, one solves $\bar{e} = \bar{B} \otimes \bar{e}$. This max-plus eigenvector problem may be solved via the power method (see Chapter 4). This converges exactly in a finite number of steps to the unique eigenvector $\bar{e}$.

A simple example with $\mathcal{M} = \{1,2\}$ has been included. The computation takes about 5 seconds on a standard 2001 desktop PC. In the example,

$$A^1 = \begin{bmatrix} -1 & 0.5 \\ 0.1 & -1 \end{bmatrix}, \quad A^2 = \begin{bmatrix} -1 & 0.1 \\ 0.5 & -1 \end{bmatrix},$$

$$D = \begin{bmatrix} 1.5 & 0.2 \\ 0.2 & 1.5 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 0.4 & -0.1 \\ -0.1 & 0.4 \end{bmatrix}$$

Perhaps we should note that the solution is not the piecewise combination of the solutions of the constituent linear/quadratic problems. The value function, its two partials and a backsubstitution error are plotted in Figures 6.1, and 6.2. (The backsubstitution error is computed through approximation of the gradient via simple (perhaps overly simple) first-order differencing, and substitution into the HJB PDE.) The sharp cleft in the error plot is due to a discontinuity in the gradient of the value function. The rise at the corners opposite this cleft indicates that some additional basis functions should have been added to cover this region; thus the user can determine when one needs to extend this set.

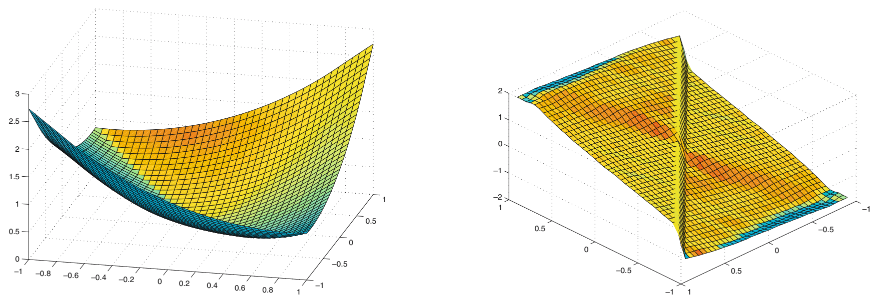


Fig. 6.1. Value function and first partial

It is worth noting that when one uses HJB PDEs that are quadratic functions of x and ∇W, the corresponding transformed operators, B^m are quadratic functions. In particular, for the very simple example HJB PDE above (with no linear or zeroth order terms), one finds that the B^m take the simple quadratic form

$$B^m_{i,j} = \tfrac{1}{2}(x_i^T, x_j^T) G^m (x_i^T, x_j^T)^T \tag{6.36}$$

for a matrix G^m which is easily computed.

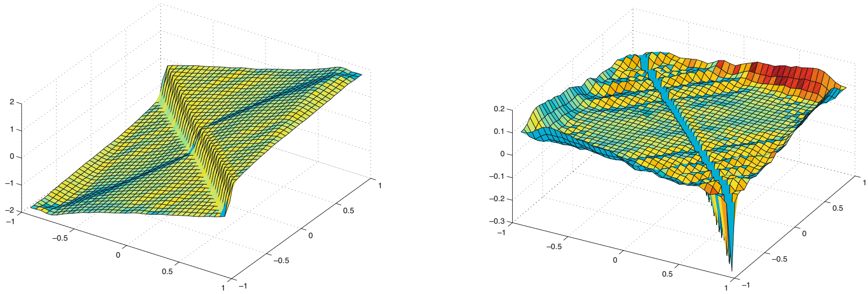


Fig. 6.2. Second partial and backsubstitution error

7

Curse-of-Dimensionality-Free Method

In Chapter 6, we moved away from the *direct* eigenvector method. In particular, we considered problems where the semigroups could be constructed (or possibly approximated) as max-plus sums of constituent semigroups, say $\bar{S}_\tau = \bigoplus_{m\in\mathcal{M}} S^m_\tau$. Equivalently, we considered eigenvector problems with matrices $\overline{B} = \overline{B}_\tau = \bigoplus_{m\in\mathcal{M}} B^m_\tau$. If the constituent B^m were easily computed, then $\overline{B}$ would be also. Because, with the direct eigenvector method, the cost of computing B dominates the cost of computing the eigenvector given B, this could provide a superior method when B takes the form of a max-plus sum of simple constituent B^m (or, possibly, is well approximated as such).

However, the number of basis functions required still typically grows exponentially with space dimension. For instance, one might use only 25 basis functions per space dimension. Yet in such a case, the computational cost would still grow at a rate of $(25k)^n$ for some constant k where n continues to denote the space dimension. We see that one still has the curse-of-dimensionality. With the max-plus methods, the "time-step" tends to be much larger than what can be used in finite element methods (because it encapsulates the action of the semigroup propagation on each basis function), and so these methods can be quite fast on small problems. However, even with the max-plus approach, the curse-of-dimensionality growth is so fast that one cannot expect to solve general problems of more than say dimension 4 or 5 on current machinery, and again the computing machinery speed increases expected in the foreseeable future cannot do much to raise this.

With the construction approach of Chapter 6, the computational cost of computing B can be hugely reduced for some problems. However, one still needs to compute the elements of B, and then the eigenvector, e, of dimension K^n where K is the number of basis functions per dimension and n remains the dimension of the state space. (We remark again that it appears one does not need to compute the entire matrix, B, but only the elements corresponding to closely spaced pairs of basis functions where the definition of "closely spaced" needs clarification, and appears related to τ.) Consequently, even with a computational cost reduction of an order of magnitude, one can only expect to gain roughly one additional space dimension over the direct method.

Many researchers have noticed that the introduction of even a single simple nonlinearity into an otherwise linear control problem of high dimensionality, n, has disastrous computational repercussions. Specifically, one goes from the solution of an n-dimensional Riccati equation to the solution of a grid-based (e.g., finite element) or max-plus method over a space of dimension n. While the Riccati equation may be "relatively" easily solved for large n, the above max-plus methods would not likely be computationally feasible for $n > 6$ in the best cases (without further advances in the algorithms). Of course, grid-based methods would not be computationally feasible either. This has been a frustrating, counterintuitive situation.

This chapter discusses an approach to certain nonlinear HJB PDEs which is not subject to the curse-of-dimensionality. In fact, the computational growth in state-space dimension is on the order of n^3. There is, of course, no "free lunch," and there is exponential computational growth in a certain measure of complexity of the Hamiltonian. Under this measure, the minimal complexity Hamiltonian is the linear/quadratic Hamiltonian — corresponding to solution by a Riccati equation. If the Hamiltonian is given as (or approximated by) a maximum or minimum of M linear/quadratic Hamiltonians, then one could say the complexity of the Hamiltonian (or the approximation of the Hamiltonian) is M.

The elimination of the curse-of-dimensionality requires the elimination of the max-plus basis function expansion. One works instead with the Legendre/Fenchel transforms. The B matrices are replaced by the kernels of max-plus integral operators on the transform/dual space. When the constituent problems are linear/quadratic, the constituent kernels, B^m, are obtained analytically. The discretization comes through repeated application of max-plus sums of these operators. Let us give some more detail on this concept.

As in the previous chapter, we will be concerned with HJB PDEs given or approximated as

$$\widetilde{H}(x, \nabla W) = \max_{m \in \{1,2,\dots,M\}} \{H^m(x, \nabla W)\} \tag{7.1}$$

or

$$\widetilde{H}(x, \nabla W) = \min_{m \in \{1,2,\dots,M\}} \{H^m(x, \nabla W)\}.$$

In order to make the problem tractable, we will concentrate on a single class of HJB PDEs of form (7.1). However, the theory can obviously be expanded to a much larger class.

To give an idea of the proposed method, recall that the solution of (7.1) is the eigenfunction of the corresponding semigroup, that is,

$$0 \otimes W = W = \widetilde{S}_\tau[W],$$

where we recall that $\widetilde{S}_\tau$ is max-plus linear. The Legendre/Fenchel transform maps this to the dual-space eigenfunction problem

$$0 \otimes e = \widetilde{\mathcal{B}}_\tau \odot e$$

where we use the $\odot$ notation to indicate $\widetilde{\mathcal{B}}_\tau \odot e \doteq \int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(x,y) \otimes e(y)\,dy = \sup_{y\in\mathbf{R}^n}\left[\widetilde{\mathcal{B}}_\tau(x,y) + e(y)\right]$ where $\int^{\oplus}$ denotes max-plus integration (maximization), c.f. [63]. Then one approximates $\widetilde{\mathcal{B}}_\tau \simeq \bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m$ where $\mathcal{M} \doteq \{1,2,\dots,M\}$ and the $\mathcal{B}_\tau^m$ correspond to the H^m. The power method for finite-size matrices (see Chapter 4) suggests that the solution is given by

$$e \simeq \lim_{N\to\infty}\left[\bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m\right]^N \odot 0$$

where the N superscript denotes the $\odot$ operation N times, and 0 represents the zero function. Given linear/quadratic forms for each of the H^m, the $\mathcal{B}_\tau^m$ are obtained by solving Riccati equations for the coefficients in the quadratic forms. Let $e_N \doteq \left[\bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m\right]^N \odot 0$. Note that

$$\begin{aligned}
e_1 &= \bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m \odot 0 \\
e_2 &= \bigoplus_{(m_1,m_2)\in\mathcal{M}\times\mathcal{M}} \mathcal{B}_\tau^{m_1,m_2} \odot 0 \doteq \left[\bigoplus_{m_2\in\mathcal{M}} \mathcal{B}_\tau^{m_2}\right] \odot \left[\bigoplus_{m_1\in\mathcal{M}} \mathcal{B}_\tau^{m_1}\right] \odot 0 \\
e_3 &= \bigoplus_{(m_1,m_2,m_3)\in\mathcal{M}\times\mathcal{M}\times\mathcal{M}} \mathcal{B}_\tau^{m_1,m_2,m_3} \odot 0 \\
&\doteq \left[\bigoplus_{m_3\in\mathcal{M}} \mathcal{B}_\tau^{m_3}\right] \odot \left[\bigoplus_{(m_1,m_2)\in\mathcal{M}\times\mathcal{M}} \mathcal{B}_\tau^{m_1,m_2}\right] \odot 0
\end{aligned}$$

and so on. Then $e_N \to e$. The convergence rate does not depend on space dimension, but on the dynamics of the problem. There is no curse-of-dimensionality. The exponential computational growth is in $M = \#\mathcal{M}$. (However, we remark that those $\mathcal{B}_\tau^{m_1,m_2,\dots,m_N} \odot 0$ which are dominated by others can be deleted from the list of such objects without consequence, which is important in mollifying the growth.) The computation of each $\mathcal{B}_\tau^{\{m_i\}_{i=1}^N}$ is analytical given the solution of the Riccati equations for the H^m.

It should be remarked that, although only the case of Hamiltonians which are maxima of linear/quadratic forms will be considered, much of the theory is applicable to a much larger class of problems. In particular, the concepts of Legendre/Fenchel transforms and kernels of max-plus linear operators on the dual space could be applied in a wider setting. However, our interest here will be focused on a certain computational approach which is greatly enabled in the special case of Hamiltonians which are pointwise maxima of linear/quadratic constituent Hamiltonians.

7.1 DP for the Constituent and Originating Problems

There are certain conditions that must be satisfied for solutions to exist and the method to apply. In order that the assumptions are not completely abstract, we will work with a specific problem class — the infinite time-horizon H_∞ problem with fixed feedback. This is a problem class where we have already developed a great deal of machinery in the earlier chapters, and so less analysis will be required for application of the new method. Of course the concept is applicable to a much wider class.

As indicated above, we suppose the individual H^m are linear/quadratic Hamiltonians. Consequently, consider a finite set of linear systems

$$\begin{aligned}\dot{\xi}^m &= A^m \xi^m + \sigma^m u,\\ \xi^m_0 &= x \in \mathbf{R}^n.\end{aligned} \tag{7.2}$$

Again let $u \in \mathcal{U} \doteq L_2^{loc}([0,\infty);\mathbf{R}^l)$. Let the cost functionals be

$$\widehat{J}^m(x,T;u) \doteq \int_0^T \tfrac{1}{2}(\xi^m_t)^T D^m \xi^m_t - \frac{\gamma^2}{2}|u_t|^2\,dt, \tag{7.3}$$

and let the value function be

$$W^m(x) = \sup_{u\in\mathcal{U}} \sup_{T<\infty} \widehat{J}^m(x,T;u) = \lim_{T\to\infty} \sup_{u\in\mathcal{U}} \widehat{J}^m(x,T;u), \tag{7.4}$$

where use of the limit over T is justified in Chapter 3. We remark that a generalization of the second term in the integrand of the cost functional to $\frac{1}{2}u^T(\Gamma^m)^T(\Gamma^m)u$ with $(\Gamma^m)^T\Gamma^m$ positive definite is not needed because this is equivalent to a change in σ^m in the dynamics (7.2). Obviously $\widehat{J}^m$ and W^m require some assumptions in order to guarantee their existence. The assumptions will hold throughout the chapter. Because these assumptions only appear together, we will refer to this entire set of assumptions as Assumption Block (A7.1I), and this is:

Assume that there exists $c_A \in (0,\infty)$ such that

$$x^T A^m x \le -c_A|x|^2 \quad \forall\, x \in \mathbf{R}^n,\ \forall m \in \mathcal{M}.$$

Assume that there exists $m_\sigma < \infty$ such that

$$|\sigma^m| \le m_\sigma \quad \forall\, m \in \mathcal{M}.$$

(A7.1I)

Assume that all D^m are positive definite, symmetric, and let c_D be such that

$$x^T D^m x \le c_D|x|^2 \quad \forall\, x \in \mathbf{R}^n,\ \forall m \in \mathcal{M}$$

(which is obviously equivalent to all eigenvalues of the D^m being no greater than c_D). Lastly, assume that $\gamma^2/m_\sigma^2 > c_D/c_A^2$.

These assumptions are obviously similar to (A4.1I)–(A4.4I), but with the above linear systems notation. Note also that these assumptions guarantee the existence of the W^m as locally bounded functions which are zero at the origin (see Chapter 3). In fact, the specific linear/quadratic structure of the above assumptions implies that these W^m will be quadratic.

The corresponding HJB PDEs are

$$\begin{aligned} 0 &= -H^m(x, \nabla W) \\ &= -\left\{\frac{1}{2}x^T D^m x + (A^m x)^T \nabla W + \max_{u\in\mathbf{R}^l}[(\sigma^m u)^T \nabla W - \frac{\gamma^2}{2}|u|^2]\right\} \\ &= -\left\{\frac{1}{2}x^T D^m x + (A^m x)^T \nabla W + \frac{1}{2}\nabla W^T \Sigma^m \nabla W\right\} \\ W(0) &= 0, \end{aligned} \tag{7.5}$$

where $\Sigma^m \doteq \frac{1}{\gamma^2}\sigma^m(\sigma^m)^T$. Let $\mathcal{C}_\delta$ be the subset of semiconvex functions on $\mathbf{R}^n$ such that $0 \le W(x) \le \frac{c_A(\gamma^2-\delta)}{2m_\sigma^2}|x|^2$ for all x. From Chapter 3, Theorem 3.19 and Chapter 4, Corollary 4.11 (undoubtedly among many other works on linear systems),

Theorem 7.1. *Each value function (7.4) is the unique viscosity solution of its corresponding HJB PDE (7.5) in the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$.*

Defining

$$\widehat{V}^m(x,T) = \sup_{u\in\mathcal{U}} \widehat{J}^m(x,T;u),$$

we have

$$W^m(x) = \lim_{T\to\infty} \widehat{V}^m(x,T), \tag{7.6}$$

where $\widehat{V}^m$ is also the unique continuous viscosity solution of

$$\begin{aligned} 0 &= V_T - H^m(x, \nabla V), \\ V(0,x) &= 0 \end{aligned} \tag{7.7}$$

(see Chapter 3). It is easy to see that these solutions have the form $\widehat{V}^m(x,t) = \frac{1}{2}x^T P_t^{m,f} x$ where each (symmetric) $P^{m,f}$ satisfies the differential Riccati equation

$$\begin{aligned} \dot{P}^{m,f} &= (A^m)^T P^{m,f} + P^{m,f} A^m + D^m + P^{m,f}\Sigma^m P^{m,f}, \\ P_0^{m,f} &= 0. \end{aligned} \tag{7.8}$$

By (7.6) and (7.8), the W^m take the form $W^m(x) = \frac{1}{2}x^T P^m x$ where $P^m = \lim_{t\to\infty} P_t^{m,f}$. With this form, and (7.5) (or (7.8)), we see that the P^m satisfy the algebraic Riccati equations

$$0 = (A^m)^T P^m + P^m A^m + D^m + P^m \Sigma^m P^m. \tag{7.9}$$

Combining this with Theorem 7.1, one has:

Theorem 7.2. *Each value function (7.4) is the unique* classical *solution of its corresponding HJB PDE (7.5) in the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$. Further, $W^m(x) = \frac{1}{2}x^T P^m x$ where P^m is the smallest symmetric, positive definite solution of (7.9)*

Corollary 7.3. *Each W^m is strictly convex. Further, there exists symmetric, positive definite $\overline{C}$ and $\overline{\varepsilon} > 0$ such that $W^m(x) - \frac{1}{2}x^T \overline{C} x$ is convex, and in fact, bounded below by $(\overline{\varepsilon}/2)|x|^2$, for all $m \in \mathcal{M}$.*

For each m define the semigroup

$$S_T^m[\phi](x) \doteq \sup_{u \in \mathcal{U}} \left[\int_0^T \tfrac{1}{2}(\xi_t^m)^T D^m \xi_t^m - \frac{\gamma^2}{2}|u_t|^2 \, dt + \phi(\xi_T^m) \right] \tag{7.10}$$

where ξ^m satisfies (7.2). From Chapters 3 and 4, the domain of S_T^m includes $\mathcal{C}_\delta$ for all $\delta > 0$.

Theorem 7.4. *Fix any $T > 0$. Each value function, W^m, is the unique smooth solution of*

$$W = S_T^m[W]$$

in the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$. Further, given any $W \in \mathcal{C}_\delta$, $\lim_{T\to\infty} S_T^m[W](x) = W^m(x)$ for all $x \in \mathbf{R}^n$ (uniformly on compact sets).

Proof. (Sketch of proof.) Neglecting the smoothness, the first statement is equivalent to Theorem 4.4 and Corollary 4.11. The smoothness follows from the quadratic form. The proof of the second statement is nearly identical to the bulk of the proof of Theorem 4.4. In particular, note that the right-hand side of (4.15) is greater than $\int_0^T \alpha_l |\xi_t|^2 \, dt + W(\xi_T)$ for proper choice of δ in (the definition of $\widehat{\gamma}$) in (4.15). The remainder of the proof follows similarly to the remainder of the proof of Theorem 4.4. □

Recall that the HJB PDE of interest is

$$\begin{aligned} 0 &= -\widetilde{H}(x, \nabla W) \doteq -\max_{m \in \mathcal{M}} H^m(x, \nabla W), \\ W(x) &= 0. \end{aligned} \tag{7.11}$$

The corresponding value function is

$$\begin{aligned} \widetilde{W}(x) &= \sup_{u \in \mathcal{U}} \sup_{\mu \in \mathcal{D}_\infty} \widetilde{J}(x, u, \mu) \\ &\doteq \sup_{u \in \mathcal{U}} \sup_{\mu \in \mathcal{D}_\infty} \sup_{T < \infty} \int_0^T l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt, \end{aligned} \tag{7.12}$$

where

$$l^{\mu_t}(x) = \tfrac{1}{2}x^T D^{\mu_t} x,$$
$$\mathcal{D}_\infty = \{\mu : [0,\infty) \to \mathcal{M} \ : \ \text{measurable}\},$$

and ξ satisfies

$$\begin{aligned} \dot{\xi} &= A^{\mu_t}\xi + \sigma^{\mu_t} u_t, \\ \xi_0 &= x. \end{aligned} \tag{7.13}$$

Theorem 7.5. *Value function $\widetilde{W}$ is the unique viscosity solution to (7.11) in the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$.*

Remark 7.6. The proof of Theorem 7.5 is identical to the proofs in Chapter 3 with only trivial changes, and so is not included. In particular, rather than choosing any $u \in \mathcal{U}$, one chooses both any $u \in \mathcal{U}$ and any $\mu \in \mathcal{D}_\infty$. Also, the finite time-horizon PDEs now include maximization over $m \in \mathcal{M}$. In particular, (3.70) now becomes

$$0 = \widetilde{V}^f_T - \max_{m \in \mathcal{M}} \left[\tfrac{1}{2}x^T D^m x + (A^m x)^T \nabla \widetilde{V}^f + \tfrac{1}{2}(\nabla \widetilde{V}^f)^T \Sigma^m \nabla \widetilde{V}^f \right]$$
$$\widetilde{V}^f(x,0) = 0,$$

where previously there was no maximization over m.

Define the semigroup

$$\widetilde{S}_T[\phi] = \sup_{u \in \mathcal{U}} \sup_{\mu \in \mathcal{D}_T} \left[\int_0^T l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2 \, dt + \phi(\xi_T) \right], \tag{7.14}$$

where

$$\mathcal{D}_T = \{\mu : [0,T) \to \mathcal{M} \ : \ \text{measurable}\}. \tag{7.15}$$

In analogy with Theorem 7.4, one has the following.

Theorem 7.7. *Fix any $T > 0$. Value function $\widetilde{W}$ is the unique continuous solution of*

$$W = \widetilde{S}_T[W]$$

in the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$. Further, given any $W \in \mathcal{C}_\delta$, $\lim_{T\to\infty} \widetilde{S}_T[W](x) = \widetilde{W}(x)$ for all $x \in \mathbf{R}^n$ (uniformly on compact sets).

The proof is nearly identical to the proof of Theorem 7.4, and so is not included. In particular, the only change is the addition of the supremum over $\mathcal{D}_T$ — which makes no substantive change in the proof.

Importantly, we also have the following.

Theorem 7.8. *Value function $\widetilde{W}$ is strictly convex. Further, there exists $c_W > 0$ such that $\widetilde{W}(x) - \frac{1}{2}c_W|x|^2$ is strictly convex.*

Proof. Fix any $x, \eta \in \mathbf{R}^n$ with $|\eta| = 1$ and any $\delta > 0$. Let $\varepsilon > 0$. Given x, let $u^\varepsilon \in \mathcal{U}$, $\mu^\varepsilon \in \mathcal{D}_\infty$ be ε-optimal for $\widetilde{W}(x)$ (i.e., so that $\widetilde{J}(x, u^\varepsilon, \mu^\varepsilon) \geq \widetilde{W}(x) - \varepsilon$). Then

$$\begin{aligned}&\widetilde{W}(x - \delta\eta) - 2\widetilde{W}(x) + \widetilde{W}(x + \delta\eta) \\ &\geq \widetilde{J}(x - \delta\eta, u^\varepsilon, \mu^\varepsilon) - 2\widetilde{J}(x, u^\varepsilon, \mu^\varepsilon) + \widetilde{J}(x + \delta\eta, u^\varepsilon, \mu^\varepsilon) - 2\varepsilon. \end{aligned} \tag{7.16}$$

Let $\xi^\delta, \xi^0, \xi^{-\delta}$ be solutions of dynamics (7.13) with initial conditions $\xi_0^\delta = x + \delta\eta$, $\xi_0^0 = x$ and $\xi_0^{-\delta} = x - \delta\eta$, respectively, where the inputs are u^ε and μ^ε for all three processes. Then

$$\dot{\xi}^\delta - \dot{\xi}^0 = A^{\mu_t^\varepsilon}[\xi^\delta - \xi^0] \ \text{ and } \ \dot{\xi}^0 - \dot{\xi}^{-\delta} = A^{\mu_t^\varepsilon}[\xi^0 - \xi^{-\delta}]. \tag{7.17}$$

Letting $\Delta_t^+ \doteq \xi_t^\delta - \xi_t^0$, one also has $\xi_t^0 - \xi_t^{-\delta} = \Delta_t^+$, and by linearity one finds $\dot{\Delta}^+ = A^{\mu_t^\varepsilon}\Delta^+$. Also, using (7.16) and (7.12)

$$\begin{aligned}&\widetilde{W}(x - \delta\eta) - 2\widetilde{W}(x) + \widetilde{W}(x + \delta\eta) \\ &\geq \tfrac{1}{2}\int_0^\infty \left[(\xi_t^\delta)^T D^{\mu_t^\varepsilon}\xi_t^\delta - 2(\xi_t^0)^T D^{\mu_t^\varepsilon}\xi_t^0 + (\xi_t^{-\delta})^T D^{\mu_t^\varepsilon}\xi_t^{-\delta}\right] dt - 2\varepsilon \\ &= \int_0^\infty (\Delta^+)^T D^{\mu_t^\varepsilon}\Delta^+ \, dt - 2\varepsilon. \end{aligned} \tag{7.18}$$

Also, by the finiteness of $\mathcal{M}$, there exists $K < \infty$ such that

$$\frac{d}{dt}|\Delta^+|^2 = 2(\Delta^+)^T A^{\mu_t^\varepsilon}\Delta^+ \geq -K|\Delta^+|^2,$$

which implies

$$|\Delta^+|^2 \geq e^{-Kt}\delta^2 \quad \forall t \geq 0. \tag{7.19}$$

Let $\lambda_D \doteq \min\{\lambda \in \mathbf{R} : \lambda \text{ is an eigenvalue of a } D^m\}$. By the positive definiteness of the D^m and finiteness of $\mathcal{M}$, $\lambda_D > 0$. Then, by (7.18)

$$\widetilde{W}(x - \delta\eta) - 2\widetilde{W}(x) + \widetilde{W}(x + \delta\eta) \geq \int_0^\infty \lambda_D|\Delta^+|^2 \, dt - 2\varepsilon,$$

which by (7.19)

$$\geq \frac{\lambda_D}{K}\delta^2 - 2\varepsilon.$$

Because $\varepsilon > 0$ and $|\eta| = 1$ were arbitrary, one obtains the result. □

7.2 Max-Plus Spaces and Dual Operators

Let $\mathcal{S}^\beta = \mathcal{S}^\beta(\mathbf{R}^n)$ be the max-plus vector space of functions mapping $\mathbf{R}^n$ into $\mathbf{R}^-$ which are *uniformly semiconvex* with constant β (where ϕ is uniformly semiconvex over $\mathbf{R}^n$ with constant β if $\phi(x) + (\beta/2)|x|^2$ is convex on $\mathbf{R}^n$). Note that we will now be generalizing to $\beta \in \mathbf{R}$, whereas we previously always implicitly assumed that the semiconvexity constant was positive. (A negative semiconvexity constant corresponds to functions from which subtracting the appropriate convex quadratic still yields a convex function.) Combining Corollary 7.3 and Theorem 7.8, we have the following.

Theorem 7.9. *There exists $\overline{\beta} \in \mathbf{R}$ such that given any $\beta > \overline{\beta}$, $\widetilde{W} \in \mathcal{S}^\beta$ and $W^m \in \mathcal{S}^\beta$ for all $m \in \mathcal{M}$. Further, one may take $\beta < 0$ (i.e., $\widetilde{W}, W^m$ are convex).*

We will be returning to using semiconvex duality again in this chapter, as opposed to max-plus basis expansions. For simplicity, we use a scalar coefficient. Define $\psi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ as

$$\psi(x,z) \doteq (c/2)|x-z|^2, \tag{7.20}$$

where $c \in \mathbf{R}$. Note that because we are allowing $\beta < 0$ in our class of semiconvex spaces here, it is convenient to define ψ in (7.20) without the usual minus sign. It is easy to check that the form of the semiconvex duality result, Theorem 2.11, is essentially unchanged. In particular, we have the following. (See also [101], [102].)

Theorem 7.10. *Let $\phi \in \mathcal{S}^\beta$. Let $c \in \mathbf{R}$, $c \neq 0$ such that $-c > \beta$. Let ψ be as in (7.20). Then, for all $x \in \mathbf{R}^n$,*

$$\phi(x) = \max_{z \in \mathbf{R}^n} [\psi(x,z) + a(z)] \tag{7.21}$$

$$= \int_{\mathbf{R}^n}^{\oplus} \psi(x,z) \otimes a(z)\, dz = \psi(x,\cdot) \odot a(\cdot), \tag{7.22}$$

where for all $z \in \mathbf{R}^n$

$$a(z) = -\max_{x \in \mathbf{R}^n} [\psi(x,z) - \phi(x)] \tag{7.23}$$

$$= -\int_{\mathbf{R}^n}^{\oplus} \psi(x,z) \otimes [-\phi(x)]\, dx = -\left\{\psi(\cdot,z) \odot [-\phi(\cdot)]\right\}, \tag{7.24}$$

which using the notation of [20]

$$= \left\{\psi(\cdot,z) \odot [\phi^-(\cdot)]\right\}^- . \tag{7.25}$$

Remark 7.11. Recall that $\phi \in \mathcal{S}^\beta$ implies that ϕ is locally Lipschitz (c.f. Chapter 2 and [42]). We also note that if $\phi \in \mathcal{S}^\beta$ and if there is any $x \in \mathbf{R}^n$ such that $\phi(x) = -\infty$, then $\phi \equiv -\infty$. Henceforth, we will ignore the special case of $\phi \equiv -\infty$, and assume that all functions are real-valued.

Semiconcavity is the obvious analogue of semiconvexity. In particular, a function, $\phi : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, is *uniformly semiconcave* with constant β if

$\phi(x) - \frac{\beta}{2}|x|^2$ is concave over $\mathbf{R}^n$. Let $\mathcal{S}_-^\beta$ be the set of functions mapping $\mathbf{R}^n$ into $\mathbf{R} \cup \{+\infty\}$ which are uniformly semiconcave with constant β. The next lemma is an obvious result of Theorem 7.10.

Lemma 7.12. *Let $\phi \in \mathcal{S}^\beta$ (still with $-c > \beta$), and let a be the semiconvex dual of ϕ. Then $a \in \mathcal{S}_-^d$ for some $d < -c$.*

Proof. A proof only in the case $\phi \in C^2$ is provided; the more general proof would be more technical. Without loss of generality, one may assume $x \in \mathbf{R}$; otherwise, one considers restrictions to lines through the domain, and proves convexity of the restrictions.

Noting that $\phi \in \mathcal{S}^\beta$ and $-c > \beta$, there exists a unique minimizer,

$$\overline{x}(z) = \operatorname*{argmin}_{x \in \mathbf{R}}[\phi(x) - \psi(x,z)],$$

and one has

$$a(z) = \phi(\overline{x}(z)) - \psi(\overline{x}(z), z). \tag{7.26}$$

Differentiating, and using the fact that $\phi_x(\overline{x}(z)) - \psi_x(\overline{x}(z), z) = 0$, one finds

$$\frac{da}{dz} = -\psi_z(\overline{x}(z), z).$$

Differentiating again, one finds

$$\frac{d^2a}{dz^2} = -\psi_{zz}(\overline{x}(z), z) - \psi_{zx}(\overline{x}(z), z)\frac{d\overline{x}}{dz}. \tag{7.27}$$

However, using the definition, one finds

$$\frac{d\overline{x}}{dz} = [\phi_{xx}(\overline{x}(z) - \psi_{xx}(\overline{x}(z), z)]^{-1} \psi_{zx}(\overline{x}(z), z). \tag{7.28}$$

Combining (7.27) and (7.28), one obtains

$$\begin{aligned}\frac{d^2a}{dz^2} &= -\psi_{zz}(\overline{x}(z), z) - [\phi_{xx}(\overline{x}(z) - \psi_{xx}(\overline{x}(z), z)]^{-1} \psi_{zx}^2(\overline{x}(z), z) \\ &= -c - \frac{c^2}{\phi_{xx}(\overline{x}(z)) - c}.\end{aligned} \tag{7.29}$$

However, $\phi_{xx}(x) > -\beta > c$ for all $x \in \mathbf{R}$, and consequently, (7.29) yields

$$\frac{d^2a}{dz^2} < -c + \frac{c^2}{\beta + c}$$

Letting $d \doteq -c + c^2/(\beta + c)$, one has $a \in \mathcal{S}^d$ and $d < -c$. □

Lemma 7.13. *Let $\phi \in \mathcal{S}^\beta$ with semiconvex dual a. Suppose $b \in \mathcal{S}_-^d$ with $d < -c$ is such that $\phi = \psi(x,\cdot) \odot b(\cdot)$. Then $b = a$.*

Proof. Note that $-b \in \mathcal{S}^d$. Therefore, for all $y \in \mathbf{R}^n$

$$-b(y) = \max_{\zeta \in \mathbf{R}^n} [\psi(y,\zeta) + \alpha(\zeta)]$$

or equivalently,

$$b(y) = -\max_{\zeta \in \mathbf{R}^n} [\psi(y,\zeta) + \alpha(\zeta)], \tag{7.30}$$

where for all $\zeta \in \mathbf{R}^n$

$$\alpha(\zeta) = -\max_{y \in \mathbf{R}^n} [\psi(y,\zeta) + b(y)],$$

which by assumption

$$= -\phi(\zeta). \tag{7.31}$$

Combining (7.30) and (7.31), and then using (7.23), one obtains

$$b(y) = -\max_{\zeta \in \mathbf{R}^n} [\psi(y,\zeta) - \phi(\zeta)] = a(y) \quad \forall\, y \in \mathbf{R}^n. \quad \square$$

It will be critical to the method that the functions obtained by application of the semigroups to the $\psi(\cdot,z)$ be semiconvex with less concavity than the $\psi(\cdot,z)$ themselves. In other words, we will want for instance $\widetilde{S}_\tau[\psi(\cdot,z)] \in \mathcal{S}^{-(c+\varepsilon)}$. This is the subject of the next theorem. Also, in order to keep the theorem statement clean, we will first make some definitions. Define

$$\lambda_D \doteq \min\{\lambda \in \mathbf{R} :\ \lambda \text{ is an eigenvalue of } D^m,\ m \in \mathcal{M}\}.$$

Note that the finiteness of $\mathcal{M}$ implies that $\lambda_D > 0$. Let

$$\overline{K} \doteq \max_{m \in \mathcal{M},\ x \neq 0} \frac{|x^T A^m x|}{|x|^2}.$$

We define the interval

$$I_{\overline{K}} \doteq \left(\frac{-\lambda_D}{2\overline{K}}, \frac{\lambda_D}{2\overline{K}} \right).$$

Theorem 7.14. *Let $c \in I_{\overline{K}}$, $c \neq 0$. Then there exists $\overline{\tau} > 0$ and $\nu > 0$ such that for all $\tau \in [0,\overline{\tau}]$*

$$\widetilde{S}_\tau[\psi(\cdot,z)],\ S_\tau^m[\psi(\cdot,z)] \in \mathcal{S}^{-(c+\nu\tau)}.$$

Remark 7.15. From the proof to follow, one can obtain feasible values for $\overline{\tau},\nu$. For instance, if $c > 0$, $c \in I_{\overline{K}}$, then one may take $\nu = \frac{1}{2}\lambda_D - \overline{K}c$ and $\overline{\tau}$ such that $e^{-2\overline{K}\overline{\tau}} = \frac{1}{2}$. However, in practice, such a $\overline{\tau}$ tends to be highly conservative. Because these estimates are also quite technical, we do not give explicit values.

Proof. We prove the result only for $\widetilde{S}_\tau$. The proof for S_τ^m is nearly identical and slightly simpler.

The first portion of the proof is similar to the proof of Theorem 7.8. Again, fix any $x, \eta \in \mathbf{R}^n$ with $|\eta| = 1$ and any $\delta > 0$. Fix $\tau > 0$, and let $\varepsilon > 0$. Given x, let $u^\varepsilon, \mu^\varepsilon$ be ε-optimal for $\widetilde{S}_\tau[\psi(\cdot,z)](x)$. Specifically, suppose $\widehat{\mathcal{I}}^\psi(x,\tau,u^\varepsilon,\mu^\varepsilon) \geq \widetilde{S}_\tau[\psi(\cdot,z)](x) - \varepsilon$ where

$$\widehat{\mathcal{I}}^\psi(x,\tau,u,\mu) \doteq \int_0^\tau l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + \psi(\xi_T,z) \tag{7.32}$$

and ξ_t satisfies (7.13). For simplicity of notation, let $\widehat{V}^{\tau,\psi} = \widetilde{S}_\tau[\psi(\cdot,z)]$. Then

$$\begin{aligned}&\widehat{V}^{\tau,\psi}(x-\delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x+\delta\eta)\\ &\geq \widehat{\mathcal{I}}^\psi(x-\delta\eta,\tau,u^\varepsilon,\mu^\varepsilon) - 2\widehat{\mathcal{I}}^\psi(x,\tau,u^\varepsilon,\mu^\varepsilon) + \widehat{\mathcal{I}}^\psi(x+\delta\eta,\tau,u^\varepsilon,\mu^\varepsilon)\\ &\quad -2\varepsilon.\end{aligned} \tag{7.33}$$

Let $\xi^\delta, \xi^0, \xi^{-\delta}, \Delta^+$ be as given in the proof of Theorem 7.8. Note that

$$\psi(\xi^\delta_\tau,z) - 2\psi(\xi^0_\tau,z) + \psi(\xi^{-\delta}_\tau,z) = c|\Delta^+_\tau|^2. \tag{7.34}$$

Note also that as in the proof of Theorem 7.8,

$$\tfrac{1}{2}\left[\xi^\delta_t D^{\mu^\varepsilon_t}\xi^\delta_t - 2\xi^0_t D^{\mu^\varepsilon_t}\xi^0_t + \xi^{-\delta}_t D^{\mu^\varepsilon_t}\xi^{-\delta}_t\right] = (\Delta^+_t)^T D^{\mu^\varepsilon_t}\Delta^+_t. \tag{7.35}$$

Combining (7.32), (7.33), (7.34) and (7.35), one obtains

$$\begin{aligned}&\widehat{V}^{\tau,\psi}(x-\delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x+\delta\eta)\\ &\quad\geq \int_0^\tau (\Delta^+_t)^T D^{\mu^\varepsilon_t}\Delta^+_t\,dt + c|\Delta^+_\tau|^2 - 2\varepsilon\\ &\quad\geq \int_0^\tau \lambda_D|\Delta^+_t|^2\,dt + c|\Delta^+_\tau|^2 - 2\varepsilon.\end{aligned} \tag{7.36}$$

Further, noting again that $\dot{\Delta}^+ = A^{\mu^\varepsilon_t}\Delta^+$, one has

$$\frac{d}{dt}|\Delta^+|^2 = 2(\Delta^+)^T A^{\mu^\varepsilon_t}\Delta^+.$$

Consequently, using the definition of $\overline{K}$,

$$-2\overline{K}|\Delta^+|^2 \leq \frac{d}{dt}|\Delta^+|^2 \leq 2\overline{K}|\Delta^+|^2,$$

and so

$$\delta^2 e^{-2\overline{K}t} \leq |\Delta^+_t|^2 \leq \delta^2 e^{2\overline{K}t}. \tag{7.37}$$

Suppose $c > 0$. Then by (7.36) and (7.37),

$$\begin{aligned}\widehat{V}^{\tau,\psi}(x-\delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x+\delta\eta) &\geq \lambda_D\delta^2\int_0^\tau e^{-2\overline{K}t}\,dt + c\delta^2 e^{-2\overline{K}\tau}\\ &\quad -2\varepsilon\\ &= \delta^2 f(\tau) - 2\varepsilon\end{aligned}$$

where

$$f(\tau) \doteq \lambda_D \frac{1 - e^{-2\overline{K}\tau}}{2\overline{K}} + c e^{-2\overline{K}\tau}.$$

Note that $f(0) = c$ and $f'(\tau) = (\lambda_D - 2\overline{K}c)e^{-2\overline{K}\tau}$. Then $f'(0) = \lambda_D - 2\overline{K}c$, and we suppose $\lambda_D - 2\overline{K}c > 0$. Letting $\nu \doteq \frac{1}{2}(\lambda_D - 2\overline{K}c)$, one sees that there exists $\overline{\tau} > 0$ such that

$$\widehat{V}^{\tau,\psi}(x - \delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x + \delta\eta) \geq \delta^2[f(0) + \nu\tau] - 2\varepsilon \quad \forall \tau \in [0, \overline{\tau}].$$

Because this is true for all $\varepsilon > 0$,

$$\widehat{V}^{\tau,\psi}(x - \delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x + \delta\eta) \geq \delta^2[c + \nu\tau] \quad \forall \tau \in [0, \overline{\tau}]. \tag{7.38}$$

Now suppose $c < 0$. Then by (7.36) and (7.37),

$$\widehat{V}^{\tau,\psi}(x - \delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x + \delta\eta) \geq \delta^2 \hat{f}(\tau) - 2\varepsilon \tag{7.39}$$

where

$$\hat{f}(\tau) \doteq \lambda_D \frac{1 - e^{-2\overline{K}\tau}}{2\overline{K}} + c e^{2\overline{K}\tau}.$$

Note that $\hat{f}(0) = c$ and $\hat{f}'(\tau) = \lambda_D e^{-2\overline{K}\tau} + 2c e^{2\overline{K}\tau}$. Then $\hat{f}'(0) = \lambda_D + 2\overline{K}c$, and we suppose $\lambda_D + 2\overline{K}c > 0$. Letting $\nu \doteq \frac{1}{2}(\lambda_D + 2\overline{K}c)$, one sees that there exists $\overline{\tau} > 0$ such that

$$\widehat{V}^{\tau,\psi}(x - \delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x + \delta\eta) \geq \delta^2[\hat{f}(0) + \nu\tau] - 2\varepsilon \quad \forall \tau \in [0, \overline{\tau}].$$

Because this is true for all $\varepsilon > 0$,

$$\widehat{V}^{\tau,\psi}(x - \delta\eta) - 2\widehat{V}^{\tau,\psi}(x) + \widehat{V}^{\tau,\psi}(x + \delta\eta) \geq \delta^2[c + \nu\tau] \quad \forall \tau \in [0, \overline{\tau}]. \tag{7.40}$$

Combining (7.38) and (7.40) yields the result if the two conditions $\lambda_D - 2\overline{K}c > 0$ when $c > 0$ and $\lambda_D + 2\overline{K}c > 0$ when $c < 0$ are met. The reader can check that these conditions are met if $c \in I_{\overline{K}}$. □

Corollary 7.16. *We may choose $c \in \mathbf{R}$ such that $\widetilde{W}, W^m \in \mathcal{S}^{-c}$, and such that with $\psi, \overline{\tau}, \nu$ as in the statement of Theorem 7.14,*

$$\widetilde{S}_\tau[\psi(\cdot, z)], S^m_\tau[\psi(\cdot, z)] \in \mathcal{S}^{-(c+\nu\tau)} \quad \forall \tau \in [0, \overline{\tau}].$$

Henceforth, we suppose c chosen so that the results of Corollary 7.16 hold, and take $\psi(x, z) = \frac{c}{2}|x - z|^2$. We also suppose τ, ν chosen according to the corollary as well.

Now for each $z \in \mathbf{R}^n$, $\widetilde{S}_\tau[\psi(\cdot, z)] \in \mathcal{S}^{-(c+\nu\tau)}$. Therefore, by Theorem 7.10, for all $x, z \in \mathbf{R}^n$

$$\widetilde{S}_\tau[\psi(\cdot, z)](x) = \int_{\mathbf{R}^n}^{\oplus} \psi(x, y) \otimes \widetilde{\mathcal{B}}_\tau(y, z)\, dy = \psi(x, \cdot) \odot \widetilde{\mathcal{B}}_\tau(\cdot, z), \tag{7.41}$$

where for all $y \in \mathbf{R}^n$

$$\begin{aligned}\widetilde{\mathcal{B}}_\tau(y,z) &= -\int_{\mathbf{R}^n}^{\oplus} \psi(x,y)\otimes\left\{-\widetilde{S}_\tau[\psi(\cdot,z)](x)\right\}dx\\ &= \left\{\psi(\cdot,y)\odot\left[\widetilde{S}_\tau[\psi(\cdot,z)](\cdot)\right]^-\right\}^- \end{aligned} \tag{7.42}$$

It is handy to define the max-plus linear operator with "kernel" $\widetilde{\mathcal{B}}_\tau$ (where we do not rigorously define the term kernel as it will not be needed here) as $\widehat{\widetilde{\mathcal{B}}}_\tau[\alpha](z) \doteq \widetilde{\mathcal{B}}_\tau(z,\cdot)\odot\alpha(\cdot)$ for all $\alpha\in\mathcal{S}^{-c}$.

Proposition 7.17. *Let $\phi\in\mathcal{S}^{-c}$ with semiconvex dual denoted by a. Define $\phi^1=\widetilde{S}_\tau[\phi]$. Then $\phi^1\in\mathcal{S}^{-(c+\nu\tau)}$, and*

$$\phi^1(x)=\psi(x,\cdot)\odot a^1(\cdot),$$

where

$$a^1(x)=\widetilde{\mathcal{B}}_\tau(x,\cdot)\odot a(\cdot).$$

Proof. The proof that $\phi^1\in\mathcal{S}^{-(c+\nu\tau)}$ is similar to the proof in Theorem 7.14. Consequently, we prove only the second assertion.

$$\begin{aligned}\phi^1(x) &= \sup_{u\in\mathcal{U}}\sup_{\mu\in\mathcal{D}_\infty}\left[\int_0^\tau l^{\mu_t}(\xi_t)-\frac{\gamma^2}{2}|u_t|^2\,dt+\phi(\xi_\tau)\right]\\ &= \sup_{u\in\mathcal{U}}\sup_{\mu\in\mathcal{D}_\infty}\max_{z\in\mathbf{R}^n}\left[\int_0^\tau l^{\mu_t}(\xi_t)-\frac{\gamma^2}{2}|u_t|^2\,dt+\psi(\xi_\tau,z)+a(z)\right]\\ &= \max_{z\in\mathbf{R}^n}\left\{\widetilde{S}_\tau[\psi(\cdot,z)](x)+a(z)\right\},\end{aligned}$$

which by (7.41)

$$\begin{aligned}&= \int_{y\in\mathbf{R}^n}^{\oplus}\int_{z\in\mathbf{R}^n}^{\oplus}\widetilde{\mathcal{B}}_\tau(y,z)\otimes a(z)\,dz\otimes\psi(x,y)\,dy\\ &= \int_{y\in\mathbf{R}^n}^{\oplus} a^1(x)\otimes\psi(x,y)\,dy. \quad\square\end{aligned}$$

Theorem 7.18. *Let $W\in\mathcal{S}^{-c}$, and let a be its semiconvex dual (with respect to ψ). Then*

$$W=\widetilde{S}_\tau[W]$$

if and only if

$$a(z)=\max_{y\in\mathbf{R}^n}\left[\widetilde{\mathcal{B}}_\tau(z,y)+a(y)\right],$$

which of course

$$=\int_{\mathbf{R}^n}^{\oplus}\widetilde{\mathcal{B}}_\tau(z,y)\otimes a(y)\,dy=\widetilde{\mathcal{B}}_\tau(z,\cdot)\odot a(\cdot)=\widehat{\widetilde{\mathcal{B}}}_\tau[a](z)\quad\forall\, z\in\mathbf{R}^n.$$

Proof. Because a is the semiconvex dual of W, for all $x \in \mathbf{R}^n$,

$$\begin{aligned}
\psi(x,\cdot) \odot a(\cdot) &= W(x) = \widetilde{S}_\tau[W](x) \\
&= \widetilde{S}_\tau \left[\max_{z \in \mathbf{R}^n} \{\psi(\cdot,z) + a(z)\} \right](x) \\
&= \max_{z \in \mathbf{R}^n} \left\{ a(z) + \widetilde{S}_\tau[\psi(\cdot,z)](x) \right\} \\
&= \int_{\mathbf{R}^n}^{\oplus} a(z) \otimes \widetilde{S}_\tau[\psi(\cdot,z)](x)\, dz,
\end{aligned}$$

which by (7.41)

$$\begin{aligned}
&= \int_{\mathbf{R}^n}^{\oplus} a(z) \otimes \int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(y,z) \otimes \psi(x,y)\, dy\, dz \\
&= \int_{\mathbf{R}^n}^{\oplus} \int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(y,z) \otimes a(z) \otimes \psi(x,y)\, dy\, dz \\
&= \int_{\mathbf{R}^n}^{\oplus} \left[\int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(y,z) \otimes a(z)\, dz \right] \otimes \psi(x,y)\, dy \\
&= \left[\int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(\cdot,z) \otimes a(z)\, dz \right] \odot \psi(x,\cdot).
\end{aligned}$$

Combining this with Lemma 7.13, one has

$$a(y) = \int_{\mathbf{R}^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(\cdot,z) \otimes a(z)\, dz = \widetilde{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot) \quad \forall\, y \in \mathbf{R}^n.$$

The reverse direction follows by supposing $a(\cdot) = \widetilde{\mathcal{B}}_\tau(z,\cdot) \odot a(\cdot)$ and reordering the above argument. □

Corollary 7.19. *Value function $\widetilde{W}$ is given by $\widetilde{W}(x) = \psi(x,\cdot) \odot \widetilde{a}(\cdot)$ where $\widetilde{a}$ is the unique solution of*

$$\widetilde{a}(y) = \widetilde{\mathcal{B}}_\tau(y,\cdot) \odot \widetilde{a}(\cdot) \quad \forall\, y \in= \mathbf{R}^n$$

or equivalently, $\widetilde{a} = \widehat{\widetilde{\mathcal{B}}}_\tau[\widetilde{a}]$.

Proof. Combining Theorem 7.7 and Theorem 7.18 yields the assertion that $\widetilde{W}$ has this representation. The uniqueness follows from the uniqueness assertion of Theorem 7.7 and Lemma 7.13. □

Similarly, for each $m \in \mathcal{M}$ and $z \in \mathbf{R}^n$, $S_\tau^m[\psi(\cdot,z)] \in \mathcal{S}^{-(c+\nu\tau)}$ and

$$S_\tau^m[\psi(\cdot,z)](x) = \psi(x,\cdot) \odot \mathcal{B}_\tau^m(\cdot,z) \quad \forall\, x \in \mathbf{R}^n,$$

where

$$\mathcal{B}_\tau^m(y,z) = \left\{ \psi(\cdot,y) \odot \left[S_\tau^m[\psi(\cdot,z)] \right]^-(\cdot) \right\}^- \quad \forall\, y \in \mathbf{R}^n.$$

As before, it will be handy to define the max-plus linear operator with "kernel" $\mathcal{B}_\tau^m$ as $\widehat{\mathcal{B}}_\tau^m[a](z) \doteq \mathcal{B}_\tau^m(z,\cdot) \odot a(\cdot)$ for all $a \in \mathcal{S}^{-c}$. Further, one also obtains analogous results (by similar proofs). In particular, one has the following

Theorem 7.20. *Let $W \in \mathcal{S}^{-c}$, and let a be its semiconvex dual (with respect to ψ). Then*

$$W = S_\tau^m[W]$$

if and only if

$$a(z) = \mathcal{B}_\tau^m(z,\cdot) \odot a(\cdot) \quad \forall\, z \in \mathbf{R}^n.$$

Corollary 7.21. *Each value function W^m is given by $W^m(x) = \psi(x,\cdot) \odot a^m(\cdot)$ where each a^m is the unique solution of*

$$a^m(y) = \mathcal{B}_\tau^m(y,\cdot) \odot a^m(\cdot) \quad \forall\, y \in= \mathbf{R}^n.$$

7.3 Discrete Time Approximation

The method developed here will not involve any discretization over space. Of course this is obvious because otherwise one could not avoid the curse-of-dimensionality. The discretization will be over time where approximate μ processes will be constant over the length of each time-step. This is similar to a technique used in Chapter 6.

We define the operator $\bar{S}_\tau$ on $\mathcal{C}_\delta$ by

$$\begin{aligned}\bar{S}_\tau[\phi](x) &= \sup_{u\in\mathcal{U}} \max_{m\in\mathcal{M}} \left[\int_0^\tau l^m(\xi_t^m) - \frac{\gamma^2}{2}|u_t|^2\, dt + \phi(\xi_\tau^m) \right](x) \\ &= \max_{m\in\mathcal{M}} S_\tau^m[\phi](x),\end{aligned}$$

where ξ^m satisfies (7.2). Let

$$\overline{\mathcal{B}}_\tau(y,z) \doteq \max_{m\in\mathcal{M}} \mathcal{B}_\tau^m(y,z) = \bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m(y,z) \quad \forall\, y,z \in \mathbf{R}^n.$$

The corresponding max-plus linear operator is

$$\widehat{\overline{\mathcal{B}}}_\tau = \bigoplus_{m\in\mathcal{M}} \widehat{\mathcal{B}}_\tau^m.$$

Lemma 7.22. *For all $z \in \mathbf{R}^n$, $\bar{S}_\tau[\psi(\cdot,z)] \in \mathcal{S}^{-(c+\nu\tau)}$. Further,*

$$\bar{S}_\tau[\psi(\cdot,z)](x) = \psi(x,\cdot) \odot \overline{\mathcal{B}}_\tau(\cdot,z) \quad \forall\, x \in \mathbf{R}^n.$$

Proof. We provide the proof of the last statement, and this is as follows.

$$\begin{aligned}\bar{S}_\tau[\psi(\cdot,z)](x) &= \max_{m\in\mathcal{M}} S_\tau^m[\psi(\cdot,z)](x) = \bigoplus_{m\in\mathcal{M}} \psi(x,\cdot) \odot \mathcal{B}_\tau^m(\cdot,z) \\ &= \bigoplus_{m\in\mathcal{M}} \int_{\mathbf{R}^n}^{\oplus} \psi(x,y) \otimes \mathcal{B}_\tau^m(y,z)\, dy \\ &= \int_{\mathbf{R}^n}^{\oplus} \psi(x,y) \otimes \left[\bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m(y,z) \right] dy \\ &= \psi(x,\cdot) \odot \left[\max_{m\in\mathcal{M}} \mathcal{B}_\tau^m(\cdot,z) \right]. \quad \square\end{aligned}$$

We remark that, parameterized by τ, the operators $\bar{S}_\tau$ do not necessarily form a semigroup, although they do form a sub-semigroup (i.e., $\bar{S}_{\tau_1+\tau_2}[\phi](x) \le \bar{S}_{\tau_1}\bar{S}_{\tau_2}[\phi](x)$ for all $x \in \mathbf{R}^n$ and all $\phi \in \mathcal{S}^{-c}$). In spite of this, one does have $S_\tau^m \le \bar{S}_\tau \le \widetilde{S}_\tau$ for all $m \in \mathcal{M}$.

With τ acting as a time-discretization step-size, as in Chapter 6, we let

$$\mathcal{D}_\infty^\tau = \Big\{\mu : [0,\infty) \to \mathcal{M} \mid \text{for each } n \in \mathbf{N} \cup \{0\}, \text{ there exists } m_n \in \mathcal{M} \\ \text{such that } \mu(t) = m_n \;\forall\; t \in [n\tau, (n+1)\tau)\Big\},$$

and for $T = \bar{n}\tau$ with $\bar{n} \in \mathbf{N}$ define $\mathcal{D}_T^\tau$ similarly but with function domain being $[0,T)$ rather than $[0,\infty)$. Let $\mathcal{M}^{\bar{n}}$ denote the outer product of $\mathcal{M}$, $\bar{n}$ times. Let $T = \bar{n}\tau$, and define

$$\bar{\bar{S}}_T^\tau[\phi](x) = \max_{\{m_k\}_{k=0}^{\bar{n}-1} \in \mathcal{M}^{\bar{n}}} \left\{ \prod_{k=0}^{\bar{n}-1} S_\tau^{m_k} \right\} [\phi](x) = (\bar{S}_\tau)^{\bar{n}}[\phi](x)$$

where the $\prod$ notation indicates operator composition, and the superscript in the last expression indicates repeated application of $\bar{S}_\tau$, $\bar{n}$ times.

We will be approximating $\widetilde{W}$ by solving $W = \bar{S}_\tau[W]$ via its dual problem $a = \widehat{\bar{\mathcal{B}}}_\tau[a]$ for small τ. Consequently, we will need to show that there exists a solution to $W = \bar{S}_\tau[W]$, that the solution is unique, and that it can be found by solving the dual problem. We begin with existence.

Theorem 7.23. *Let*

$$\overline{W}(x) \doteq \lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[0](x) \tag{7.43}$$

for all $x \in \mathbf{R}^n$ where 0 represents the zero-function. Then, $\overline{W}$ satisfies

$$\begin{aligned} W &= \bar{S}_\tau[W], \\ W(0) &= 0. \end{aligned} \tag{7.44}$$

Further, $0 \le W^m \le \overline{W} \le \widetilde{W}$ for all $m \in \mathcal{M}$, and consequently, $\overline{W} \in \mathcal{C}_\delta$.

Proof. Note that for any $m \in \mathcal{M}$ (see Theorem 7.4),

$$\begin{aligned} W^m(x) &= \lim_{N\to\infty} S_{N\tau}^m[0](x) \le \limsup_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[0](x) \\ &\le \lim_{N\to\infty} \widetilde{S}_{N\tau}[0](x) = \widetilde{W}(x) \quad \forall\, x \in \mathbf{R}^n. \end{aligned} \tag{7.45}$$

Also,

$$\begin{aligned} \bar{\bar{S}}_{(N+1)\tau}^\tau[0](x) &= \bar{\bar{S}}_{N\tau}^\tau[\bar{S}_\tau[0](\cdot)](x) \qquad (7.46) \\ &= \sup_{\widehat{u}\in\mathcal{U}} \sup_{\hat{\mu}\in\mathcal{D}_{N\tau}} \int_0^{N\tau} l^{\hat{\mu}_t}(\xi_t) - \frac{\gamma^2}{2}|\widehat{u}_t|^2\, dt \end{aligned}$$

$$+\sup_{u\in\mathcal{U}}\max_{m\in\mathcal{M}}\int_{N\tau}^{(N+1)\tau} l^m(\xi_t)-\frac{\gamma^2}{2}|u_t|^2\,dt,$$

which by taking $u\equiv 0$

$$\geq \sup_{\widehat{u}\in\mathcal{U}}\sup_{\hat{\mu}\in\mathcal{D}_{N\tau}}\int_0^{N\tau} l^{\hat{\mu}_t}(\xi_t)-\frac{\gamma^2}{2}|\widehat{u}_t|^2\,dt=\bar{\bar{S}}^{\tau}_{N\tau}[0](x), \quad (7.47)$$

which implies that $\bar{\bar{S}}^{\tau}_{N\tau}[0](x)$ is a monotonically increasing function of N. Because it is also bounded from above (by (7.45)), one finds

$$W^m(x)\leq \lim_{N\to\infty}\bar{\bar{S}}^{\tau}_{N\tau}[0](x)\leq \widetilde{W}(x)\quad \forall x\in\mathbf{R}^n, \quad (7.48)$$

which also justifies the use of the limit definition of $\overline{W}$ in the statement of the theorem. In particular, one has $0\leq W^m\leq\overline{W}\leq\widetilde{W}$, and so $\overline{W}\in\mathcal{C}_\delta$.

Fix any $x\in\mathbf{R}^n$, and suppose there exists $\delta>0$ such that

$$\overline{W}(x)\leq \bar{S}_\tau[\overline{W}](x)-\delta. \quad (7.49)$$

However, by the definition of $\overline{W}$, given any $y\in\mathbf{R}^n$, there exists $N_\delta<\infty$ such that for all $N\geq N_\delta$

$$\overline{W}(y)\leq \bar{\bar{S}}^{\tau}_{N\tau}[0](y)+\delta/4. \quad (7.50)$$

Combining (7.49) and (7.50), one finds after a small bit of work that

$$\overline{W}(x)\leq \bar{S}_\tau\big[\bar{\bar{S}}^{\tau}_{N\tau}[0]+\delta/2\big](x)-\delta,$$

which using the max-plus linearity of $\bar{S}_\tau$

$$=\bar{\bar{S}}^{\tau}_{(N+1)\tau}[0](x)-\delta/2$$

for all $N\geq N_\delta$. Consequently, $\overline{W}(x)\leq \lim_{N\to\infty}\bar{\bar{S}}^{\tau}_{N\tau}[0](x)-\delta/2$ which is a contradiction. Therefore, $\overline{W}(x)\geq\bar{S}_\tau[\overline{W}](x)$ for all $x\in\mathbf{R}^n$. The reverse inequality follows in a similar way. Specifically, fix $x\in\mathbf{R}^n$ and suppose there exists $\delta>0$ such that

$$\overline{W}(x)\geq \bar{S}_\tau[\overline{W}](x)+\delta. \quad (7.51)$$

By the monotonicity of $\bar{\bar{S}}^{\tau}_{N\tau}$ with respect to N, for any $N<\infty$,

$$\overline{W}(x)\geq \bar{\bar{S}}^{\tau}_{N\tau}[0](x)\quad\forall x\in\mathbf{R}^n.$$

By the monotonicity of $\bar{S}_\tau$ with respect to its argument (i.e., $\phi_1(x)\leq\phi_2(x)$ for all x implying $\bar{S}_\tau[\phi_1](x)\leq\bar{S}_\tau[\phi_2](x)$ for all x), this implies

$$\bar{S}_\tau[\overline{W}]\geq \bar{\bar{S}}^{\tau}_{(N+1)\tau}[0]\quad\forall x\in\mathbf{R}^n. \quad (7.52)$$

Combining (7.51) and (7.52) yields

$$\overline{W}(x)\geq \bar{\bar{S}}^{\tau}_{(N+1)\tau}[0](x)+\delta.$$

Letting $N\to\infty$ yields a contradiction, and so $\overline{W}\leq\bar{S}_\tau[\overline{W}]$. □

The following result is immediate.

Theorem 7.24.

$$\overline{W}(x) = \sup_{\mu\in\mathcal{D}^\tau_\infty} \sup_{u\in\mathcal{U}} \sup_{T\in[0,\infty)} \left[\int_0^T l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \right],$$

where ξ_t satisfies (7.13).

Theorem 7.25. $\overline{W}(x) - \frac{1}{2}c_W|x|^2$ *is convex.*

Proof. The proof is identical to the proof of Theorem 7.8 with the exception that μ^ε is chosen from $\mathcal{D}^\tau_\infty$ instead of $\mathcal{D}_\infty$. □

We now address the uniqueness issue. Similar techniques to those used for W^m and $\widetilde{W}$ will prove uniqueness for (7.44) within $\mathcal{C}_\delta$. A slightly weaker type of result under weaker assumptions will be obtained first; this result is similar in form to that of [106].

Suppose $\overline{V}' \neq \overline{W}$, $\overline{V}' \in \mathcal{C}_\delta$ satisfies (7.44). This implies that for all $x \in \mathbf{R}^n$ and all $N < \infty$

$$\begin{aligned}\overline{V}'(x) &= \bar{\bar{S}}^\tau_{N\tau}[\overline{V}'](x) \\ &= \sup_{u\in\mathcal{U}} \sup_{\mu\in\mathcal{D}^\tau_\infty} \left\{ \int_0^{N\tau} l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + \overline{V}'(\xi_{N\tau}) \right\},\end{aligned}$$

which by taking $u^0 \equiv 0$ (with corresponding trajectory denoted by ξ^0)

$$\geq \overline{V}'(\xi^0_{N\tau}). \tag{7.53}$$

However, by (7.13), one has $\dot{\xi}^0 = A^{\mu_t}\xi^0$, and so $|\xi^0_t| \leq e^{-c_A t}|x|$ for all $t \geq 0$ which implies that $|\xi^0_{N\tau}| \to 0$ as $N \to \infty$. Consequently

$$\lim_{N\to\infty} \overline{V}'(\xi^0_{N\tau}) = 0. \tag{7.54}$$

Combining (7.53) and (7.54), one has

$$\overline{V}'(x) \geq 0 \quad \forall x \in \mathbf{R}^n. \tag{7.55}$$

Also, by (7.44)

$$\overline{V}'(x) = \lim_{N\to\infty} \bar{\bar{S}}^\tau_{N\tau}[\overline{V}'](x) \quad \forall x \in \mathbf{R}^n.$$

By (7.55) and the monotonicity of $\bar{\bar{S}}^\tau_{N\tau}$ with respect to its argument, this is

$$\geq \lim_{N\to\infty} \bar{\bar{S}}^\tau_{N\tau}[0](x) = \overline{W}(x). \tag{7.56}$$

By (7.55), (7.56), one has the uniqueness result analogous to [106], which is as follows.

Theorem 7.26. *$\overline{W}$ is the unique minimal, non-negative solution to (7.44).*

The stronger uniqueness statement (making use of the quadratic bound on $l^{\mu_t}(x)$) is as follows. As with W^m and $\widetilde{V}$, the proof is similar to those in Chapters 3 and 4. However, in this case, there is a small difference in the proof, and this difference requires another lemma. Due to this difference in the case of $\overline{W}$, we include a sketch of the proof (but with the new lemma in full) in Appendix A.

Theorem 7.27. *$\overline{W}$ is the unique solution of (7.44) within the class $\mathcal{C}_\delta$ for sufficiently small $\delta > 0$. Further, given any $W \in \mathcal{C}_\delta$, $\lim_{N\to\infty} \bar{\bar{S}}^{\tau}_{N\tau}[W](x) = \overline{W}(x)$ for all $x \in \mathbf{R}^n$ (uniformly on compact sets).*

Henceforth, we let $\delta > 0$ be sufficiently small such that $W^m, \widetilde{W}, \overline{W} \in \mathcal{C}_\delta$ for all $m \in \mathcal{M}$.

Theorem 7.28. *Let $W \in \mathcal{S}^{-c}$, and let a be its semiconvex dual. Then*

$$W = \bar{S}_\tau[W]$$

if and only if

$$a(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot) \quad \forall\, y \in \mathbf{R}^n.$$

Proof. By the semiconvex duality

$$\begin{aligned} \psi(x,\cdot) \odot a(\cdot) &= W(x) = \bar{S}_\tau[W](x) \qquad (7.57)\\ &= \bar{S}_\tau\Big[\max_{z\in\mathbf{R}^n}\{\psi(\cdot,z) + a(z)\}\Big](x), \end{aligned}$$

which as in the first part of the proof of Theorem 7.18

$$= \int_{\mathbf{R}^n}^{\oplus} a(z) \otimes \bar{S}_\tau[\psi(\cdot,z)](x)\, dz,$$

which by Lemma 7.22

$$= \int_{\mathbf{R}^n}^{\oplus} a(z) \otimes \int_{\mathbf{R}^n}^{\oplus} \psi(x,y) \otimes \overline{\mathcal{B}}_\tau(y,z)\, dy\, dz,$$

which as in the latter part of the proof of Theorem 7.18

$$= \left[\int_{\mathbf{R}^n}^{\oplus} \overline{\mathcal{B}}_\tau(\cdot,z) \otimes a(z)\, dz\right] \odot \psi(x,\cdot). \qquad (7.58)$$

By Lemma 7.13, this implies

$$a(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot) \quad \forall\, y \in \mathbf{R}^n.$$

Alternatively, if $a(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot)$ for all y, then

$$W(x) = \psi(x,\cdot) \odot a(\cdot) = \left[\int_{\mathbf{R}^n}^{\oplus} \overline{\mathcal{B}}_\tau(\cdot,z) \otimes a(z)\,dz\right] \odot \psi(x,\cdot) \quad \forall\, x \in \mathbf{R}^n,$$

which by (7.57)–(7.58) yields $W = \bar{S}_\tau[W]$. □

Corollary 7.29. *Value function $\overline{W}$ given by (7.43) is in $\mathcal{S}^{-c}$, and has representation $\overline{W}(x) = \psi(x,\cdot) \odot \overline{a}(\cdot)$ where $\overline{a}$ is the unique solution in $\mathcal{S}^{-c}$ of*

$$\overline{a}(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot \overline{a}(\cdot) \quad \forall\, y \in \mathbf{R}^n \tag{7.59}$$

or equivalently, $\overline{a} = \widehat{\overline{\mathcal{B}}}_\tau[\overline{a}]$.

Proof. The representation follows from Theorems 7.23 and 7.28. The uniqueness follows from Theorem 7.27 and Lemma 7.13. □

The following result on propagation of the semiconvex dual will also come in handy. The proof is similar to the proof of Proposition 7.17, and so is not included.

Proposition 7.30. *Let $\phi \in \mathcal{S}^{-c}$ with semiconvex dual denoted by a. Define $\phi^1 = \bar{S}_\tau[\phi]$. Then $\phi^1 \in \mathcal{S}^{-(c+\nu\tau)}$, and*

$$\phi^1(x) = \psi(x,\cdot) \odot a^1(\cdot),$$

where

$$a^1(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot) \quad \forall\, y \in \mathbf{R}^n.$$

We now show that one may approximate $\widetilde{W}$, the solution of $W = \widetilde{S}_\tau[W]$, to as accurate a level as one desires by solving $W = \bar{S}_\tau[W]$ for sufficiently small τ. Recall that if $W = \bar{S}_\tau[W]$, then it satisfies $W = \bar{\bar{S}}^{\tau}_{N\tau}[W]$ for all $N > 0$ (while $\widetilde{W}$ satisfies $W = \widetilde{S}_{N\tau}[W]$), and so this is essentially equivalent to introducing a discrete-time $\overline{\mu} \in \mathcal{D}^{\tau}_{N\tau}$ approximation to the μ process in $\widetilde{S}_{N\tau}$. The result will follow easily from the following technical lemma. The lemma uses the particular structure of our example class of problems as given by Assumption (A7.1I). As the proof of the lemma is technical but long, it is delayed to Appendix A. We also note that a similar result under different assumptions appears as Theorem 6.9.

Lemma 7.31. *Given $\hat{\varepsilon} \in (0,1]$, $\overline{T} < \infty$, there exist $T \in [\overline{T}/2, \overline{T}]$ and $\tau > 0$ such that*

$$\widetilde{S}_T[W^m](x) - \bar{\bar{S}}^{\tau}_T[W^m](x) \le \hat{\varepsilon}(1 + |x|^2) \quad \forall\, x \in \mathbf{R}^n, \forall m \in \mathcal{M}.$$

We now obtain the main approximation result.

Theorem 7.32. *Given $\overline{\varepsilon} > 0$ and $R < \infty$, there exists $\tau > 0$ such that*

$$\widetilde{W}(x) - \overline{\varepsilon} \le \overline{W}(x) \le \widetilde{W}(x) \quad \forall\, x \in \overline{B}_R(0).$$

Proof. From Theorem 7.23, we have

$$0 \le W^m(x) \le \overline{W}(x) \le \widetilde{W}(x) \le \frac{c_A(\gamma-\delta)^2}{m_\sigma^2}|x|^2 \quad \forall\, x \in \mathbf{R}^n. \tag{7.60}$$

Also, with $T = N\tau$ for any positive integer N,

$$\bar{\bar{S}}^{\tau}_{N\tau}[\phi] \le \widetilde{S}_T[\phi] \quad \forall\, \phi \in \mathcal{C}_\delta. \tag{7.61}$$

Further, by Theorem 7.7, given $\varepsilon > 0$ and $R < \infty$, there exists $\widehat{T} < \infty$ such that for all $T > \widehat{T}$ and all $m \in \mathcal{M}$

$$\widetilde{S}_T[\widetilde{W}](x) - \varepsilon/2 \le \widetilde{S}_T[W^m](x) \quad \forall\, x \in \overline{B}_R(0). \tag{7.62}$$

By (7.62) and Lemma 7.31, given $\bar{\varepsilon} > 0$ and $R < \infty$, there exists $T \in [0,\infty)$, $\tau \in [0,T]$ where $T = N\tau$ for some integer N such that for all $|x| \le R$

$$\begin{aligned} \widetilde{W}(x) - \bar{\varepsilon} &= \widetilde{S}_T[\widetilde{W}](x) - \bar{\varepsilon} \\ &\le \widetilde{S}_T[W^m](x) - \bar{\varepsilon}/2 \\ &\le \bar{\bar{S}}^{\tau}_T[W^m](x), \end{aligned}$$

where $\hat{\varepsilon}(1+R^2) = \bar{\varepsilon}/2$, and which by (7.60) and the monotonicity of $\bar{\bar{S}}^{\tau}_T[\cdot]$,

$$\le \bar{\bar{S}}^{\tau}_T[\overline{W}](x),$$

which by (7.61)

$$\le \widetilde{S}_T[\overline{W}](x),$$

which by the monotonicity of $\widetilde{S}_T[\cdot]$

$$\le \widetilde{S}_T[\widetilde{W}](x) = \widetilde{W}(x).$$

Noting (from Theorem 7.27) that $\overline{W} = \bar{\bar{S}}^{\tau}_T[\overline{W}]$ completes the proof. □

Remark 7.33. For this class of systems (defined by Assumption Block (A7.1I)), we expect this result could be sharpened to

$$\widetilde{W}(x) \le -\hat{\varepsilon}(1+|x|^2) \le \overline{W}(x) \le \widetilde{W}(x) \quad \forall\, x \in \mathbf{R}^n$$

by sharpening Theorem 7.7. However, this type of result might only be valid for limited classes of systems, and it has not yet been pursued.

7.4 The Algorithm

We now begin discussion of the actual algorithm.

Let $\psi(x,z) = \frac{c_W}{4}|x-z|^2$ and $\overline{W}^0(x) = \frac{c_W}{2}|x|^2$. By Theorem 7.23, $\overline{W} = \lim_{N\to\infty} \bar{\bar{S}}^{\tau}_{N\tau}[\overline{W}^0]$. Given $\overline{W}^k$, let

$$\overline{W}^{k+1} \doteq \bar{S}_\tau[\overline{W}^k]$$

so that $\overline{W}^k = \bar{\bar{S}}^{\tau}_{k\tau}[0]$ for all $k \geq 1$.

Let $\overline{a}^k$ be the semiconvex dual of $\overline{W}^k$ for all k. Because $\overline{W}^0 \equiv \frac{c_W}{2}|x|^2$, one easily finds $\overline{a}^0(y)$ for all $y \in \mathbf{R}^n$. Note also that by Proposition 7.30,

$$\overline{a}^{k+1} = \overline{\mathcal{B}}_\tau(x,\cdot) \odot \overline{a}^k(\cdot) = \widehat{\overline{\mathcal{B}}}_\tau[\overline{a}^k]$$

for all $n \geq 0$.

Recall that

$$\begin{aligned}
\overline{\mathcal{B}}_\tau(x,\cdot) \odot \overline{a}^k(\cdot) &= \int_{\mathbf{R}^n}^{\oplus} \overline{\mathcal{B}}_\tau(x,y) \otimes \overline{a}^k(y)\,dy = \int_{\mathbf{R}^n}^{\oplus} \bigoplus_{m\in\mathcal{M}} \mathcal{B}^m_\tau(x,y) \otimes \overline{a}^k(y)\,dy \\
&= \bigoplus_{m\in\mathcal{M}} \int_{\mathbf{R}^n}^{\oplus} \mathcal{B}^m_\tau(x,y) \otimes \overline{a}^k(y)\,dy \\
&= \bigoplus_{m\in\mathcal{M}} \left[\mathcal{B}^m_\tau(x,\cdot) \odot \overline{a}^k(\cdot)\right]. \qquad (7.63)
\end{aligned}$$

By (7.63),

$$\overline{a}^1(x) = \bigoplus_{m\in\mathcal{M}} \widehat{a}^1_m(x)$$

where

$$\widehat{a}^1_m(x) \doteq \mathcal{B}^m_\tau(x,\cdot) \odot \overline{a}^0(\cdot) \quad \forall\, m. \qquad (7.64)$$

By (7.63) and (7.64),

$$\begin{aligned}
\overline{a}^2(x) &= \bigoplus_{m_2\in\mathcal{M}} \int_{\mathbf{R}^n}^{\oplus} \mathcal{B}^{m_2}_\tau(x,y) \otimes \left[\bigoplus_{m_1\in\mathcal{M}} \widehat{a}^1_{m_1}(y)\right] dy \\
&= \bigoplus_{\{m_1,m_2\}\in\mathcal{M}\times\mathcal{M}} \int_{\mathbf{R}^n}^{\oplus} \mathcal{B}^{m_2}_\tau(x,y) \otimes \widehat{a}^1_{m_1}(y)\,dy.
\end{aligned}$$

Consequently,

$$\overline{a}^2(x) = \bigoplus_{\{m_1,m_2\}\in\mathcal{M}^2} \widehat{a}^2_{\{m_1,m_2\}}(x)$$

where

$$\widehat{a}^2_{\{m_1,m_2\}}(x) \doteq \mathcal{B}^{m_2}_\tau(x,\cdot) \odot \widehat{a}^1_{m_1}(\cdot) \quad \forall\, m_1, m_2 \qquad (7.65)$$

and $\mathcal{M}^2$ represents the outer product $\mathcal{M}\times\mathcal{M}$. Proceeding with this, one finds that in general,

$$\overline{a}^k(x) = \bigoplus_{\{m_i\}_{i=1}^k\in\mathcal{M}^k} \widehat{a}^k_{\{m_i\}_{i=1}^k}(x),$$

where

$$\widehat{a}^k_{\{m_i\}_{i=1}^k}(x) \doteq \mathcal{B}_\tau^{m_k}(x,\cdot)\odot\widehat{a}^{k-1}_{\{m_i\}_{i=1}^{k-1}}(\cdot) \quad \forall\,\{m_i\}_{i=1}^k\in\mathcal{M}^k. \tag{7.66}$$

Of course, one can obtain $\overline{W}^k$ from its dual as

$$\begin{aligned}\overline{W}^k(x) &= \max_{y\in\mathbf{R}^n}[\psi(x,y)+\overline{a}^k(y)]\\ &= \int_{\mathbf{R}^n}^{\oplus}\left[\psi(x,y)\otimes\bigoplus_{\{m_i\}_{i=1}^k\in\mathcal{M}^k}\widehat{a}^k_{\{m_i\}_{i=1}^k}(y)\right]dy\\ &= \bigoplus_{\{m_i\}_{i=1}^k\in\mathcal{M}^k}\left\{\int_{\mathbf{R}^n}^{\oplus}\psi(x,y)\otimes\widehat{a}^k_{\{m_i\}_{i=1}^k}(y)\,dy\right\}\\ &\doteq \bigoplus_{\{m_i\}_{i=1}^k\in\mathcal{M}^k}\widehat{W}^k_{\{m_i\}_{i=1}^k}(x),\end{aligned} \tag{7.67}$$

where

$$\begin{aligned}\widehat{W}^k_{\{m_i\}_{i=1}^k} &= \int_{\mathbf{R}^n}^{\oplus}\psi(x,y)\otimes\widehat{a}^k_{\{m_i\}_{i=1}^k}(y)\,dy\\ &= \max_{y\in\mathbf{R}^n}[\psi(x,y)+\widehat{a}^k_{\{m_i\}_{i=1}^k}(y)].\end{aligned} \tag{7.68}$$

The algorithm will consist of the forward propagation of the $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ (according to (7.66)) from $k=0$ to some termination step $k=N$, followed by construction of the value as $\widehat{W}^k_{\{m_i\}_{i=1}^k}$ (according to (7.68)).

It is important to note that the computation of each $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ is analytical. We will indicate the actual analytical computations.

By the linear/quadratic nature of the m-indexed systems, we find that the $S_\tau^m[\psi(\cdot,z)]$ take the form

$$S_\tau^m[\psi(\cdot,z)](x) = \tfrac{1}{2}(x-\Lambda_\tau^m z)^T P_\tau^m(x-\Lambda_\tau^m z)+\tfrac{1}{2}z^T R_\tau^m z,$$

where the time-dependent $n\times n$ matrices P_t^m, Λ_t^m and R_t^m satisfy $P_0^m=cI$, $\Lambda_0^m=I$, $R_0^m=0$,

$$\begin{aligned}\dot{P}^m &= (A^m)^T P^m + P^m A^m + [D^m + P^m\Sigma^m P^m],\\ \dot{\Lambda}^m &= -\left[(P^m)^{-1}D^m - A^m\right]\Lambda^m,\\ \dot{R}^m &= (\Lambda^m)^T D^m \Lambda^m.\end{aligned}$$

We note that each of the $P_\tau^m, \Lambda_\tau^m, R_\tau^m$ need only be computed once.

Next one computes each quadratic function $\mathcal{B}_\tau^m(x,z)$ (one time only) as follows. One has

$$\mathcal{B}_\tau^m = -\max_{y\in\mathbf{R}^n}\{\psi(y,x) - S_\tau^m[\psi(\cdot,z)](y)\}$$

which by the above with $c = c_W/2$

$$= \min_{y\in\mathbf{R}^n}\Big\{\tfrac{c}{2}(y-x)^T(y-x) + \tfrac{1}{2}(y-\Lambda_\tau^m z)^T P_\tau^m (y-\Lambda_\tau^m z) + \tfrac{1}{2}z^T R_\tau^m z\Big\}. \tag{7.69}$$

Recall that by Theorem 7.14, this has a finite minimum ($P^m - (c+\nu\tau)I$ positive definite). Taking the minimum in (7.69), one has

$$\mathcal{B}_\tau^m(x,z) = \tfrac{1}{2}\left[x^T M_{1,1}^m x + x^T M_{1,2}^m z + z^T (M_{1,2}^m)^T x + z^T M_{2,2}^m z\right]$$

where with shorthand notation $C \doteq cI$ and $D_\tau \doteq (P_\tau^m - cI)^{-1}$,

$$\begin{aligned}
M_{1,1}^m &= \left[C D_\tau^{-1} P_\tau^m D_\tau^{-1} C - (D_\tau^{-1}C + I)^T C (D_\tau^{-1}C + I)\right],\\
M_{1,2}^m &= \left[(D_\tau^{-1}C + I)^T C D_\tau^{-1} P_\tau^m - C D_\tau^{-1} P_\tau^m (D_\tau^{-1} P_\tau^m - I)\right]\Lambda_\tau^m,\\
M_{2,2}^m &= (\Lambda_\tau^m)^T\left[(D_\tau^{-1}P_\tau^m - I)^T P_\tau^m (D_\tau^{-1}P_\tau^m - I) - P_\tau^m D_\tau^{-1} C D_\tau^{-1} P_\tau^m\right]\Lambda_\tau^m\\
&\quad + R_\tau^m.
\end{aligned}$$

Note that given the $P_\tau^m, \Lambda_\tau^m, R_\tau^m$, the $\mathcal{B}_\tau^m$ are quadratic functions with analytical expressions for their coefficients. Also note that all the matrices in the definition of $\mathcal{B}_\tau^m$ may be precomputed.

Now let us write the (quadratic) $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ in the form

$$\widehat{a}^k_{\{m_i\}_{i=1}^k}(x) = \tfrac{1}{2}(x - \widehat{z}^k_{\{m_i\}_{i=1}^k})^T \widehat{Q}^k_{\{m_i\}_{i=1}^k}(x - \widehat{z}^k_{\{m_i\}_{i=1}^k}) + \widehat{r}^k_{\{m_i\}_{i=1}^k}.$$

Then, for each m_{k+1},

$$\begin{aligned}
\widehat{a}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= \max_{z\in\mathbf{R}^n}\Big\{\mathcal{B}_\tau^{m_{k+1}}(x,z) + \widehat{a}^k_{\{m_i\}_{i=1}^k}(z)\Big\}\\
&= \max_{z\in\mathbf{R}^n}\Big\{\tfrac{1}{2}\left[x^T M_{1,1}^m x + x^T M_{1,2}^m z + z^T (M_{1,2}^m)^T x + z^T M_{2,2}^m z\right]\\
&\qquad + \tfrac{1}{2}(x - \widehat{z}^k_{\{m_i\}_{i=1}^k})^T \widehat{Q}^k_{\{m_i\}_{i=1}^k}(x - \widehat{z}^k_{\{m_i\}_{i=1}^k}) + \widehat{r}^k_{\{m_i\}_{i=1}^k}\Big\}\\
&= \tfrac{1}{2}(x - \widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}})^T \widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}}(x - \widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}}) + \widehat{r}^{k+1}_{\{m_i\}_{i=1}^{k+1}}
\end{aligned} \tag{7.70}$$

where

$$\begin{aligned}
\widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= M_{1,1}^{m_{k+1}} - M_{1,2}^{m_{k+1}}\widehat{D}\left(M_{1,2}^{m_{k+1}}\right)^T,\\
\widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= -\left(\widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}}\right)^{-1} M_{1,2}^{m_{k+1}}\widehat{E},
\end{aligned}$$

$$\begin{aligned}
\widehat{r}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= \widehat{r}^{k}_{\{m_i\}_{i=1}^{k}} + \tfrac{1}{2}\widehat{E}^T M^m_{2,2}\widehat{z}^{k}_{\{m_i\}_{i=1}^{k}} - \tfrac{1}{2}\left(\widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}}\right)^T \widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}}\widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}},\\
\widehat{D} &= \left(M^{m_{k+1}}_{2,2} + \widehat{Q}^{k}_{\{m_i\}_{i=1}^{k}}\right)^{-1},\\
\widehat{E} &= \widehat{D}\widehat{Q}^{k}_{\{m_i\}_{i=1}^{k}}\widehat{z}^{k}_{\{m_i\}_{i=1}^{k}}.
\end{aligned}$$

Thus we have the analytical expression for the propagation of each (quadratic) $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ function. Specifically, we see that the propagation of each $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ amounts to a set of matrix multiplications (and an inverse). One might note that for purely quadratic constituent Hamiltonians (without terms that are linear or constant in the state and gradient variables), one would expect that, without loss of generality, one could take the $\widehat{z}^k_{\{m_i\}_{i=1}^k}$ and $\widehat{r}^k_{\{m_i\}_{i=1}^k}$ to be zero. However, we will also be considering Hamiltonians with linear and constant terms below.

At each step, k, the semiconvex dual of $\overline{W}^k$, $\overline{a}^k$, is represented as the finite set of functions

$$\widehat{\mathcal{A}}_k \doteq \left\{\widehat{a}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M}\ \forall i \in \{1,2,\ldots,k\}\right\},$$

where this is equivalently represented as the set of triples

$$\widehat{\mathcal{Q}}_k \doteq \left\{\left(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k}\right) \,|\, m_i \in \mathcal{M}\ \forall i \in \{1,2,\ldots,k\}\right\}.$$

At any desired stopping time, one can recover a representation of $\overline{W}^k$ as

$$\widehat{\mathcal{U}}_k \doteq \left\{\widehat{V}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M}\ \forall i \in \{1,2,\ldots,k\}\right\},$$

where these $\widehat{V}^k_{\{m_i\}_{i=1}^k}$ are also quadratics. In fact, recall

$$\begin{aligned}
\overline{W}^k(x) &= \max_{z\in\mathbf{R}^n}[\overline{a}^k(z) + \psi(x,z)]\\
&= \max_{\{m_i\}_{i=1}^k}\max_{z\in\mathbf{R}^n}\Big[\tfrac{1}{2}(z-\widehat{z}^k_{\{m_i\}_{i=1}^k})^T\widehat{Q}^k_{\{m_i\}_{i=1}^k}(z-\widehat{z}^k_{\{m_i\}_{i=1}^k}) + \widehat{r}^k_{\{m_i\}_{i=1}^k}\\
&\qquad\qquad + \tfrac{c}{2}|x-z|^2\Big]\\
&\doteq \max_{\{m_i\}_{i=1}^k}\tfrac{1}{2}(x-\widehat{x}^k_{\{m_i\}_{i=1}^k})^T\widehat{P}^k_{\{m_i\}_{i=1}^k}(x-\widehat{x}^k_{\{m_i\}_{i=1}^k}) + \widehat{\rho}^k_{\{m_i\}_{i=1}^k}\\
&\doteq \bigoplus_{\{m_i\}_{i=1}^k}\widehat{V}^k_{\{m_i\}_{i=1}^k}(x),
\end{aligned}$$

where with $C \doteq cI$

$$\begin{aligned}
\widehat{P}^k_{\{m_i\}_{i=1}^k} &= C\widehat{F}\widehat{Q}^k_{\{m_i\}_{i=1}^k}\widehat{F}C + (\widehat{F}C-I)^TC(\widehat{F}C-I),\\
\widehat{x}^k_{\{m_i\}_{i=1}^k} &= -(\widehat{P}^k_{\{m_i\}_{i=1}^k})^{-1}\Big[C\widehat{F}\widehat{Q}^k_{\{m_i\}_{i=1}^k}\widehat{G} + (\widehat{F}C-I)^TC\widehat{F}\widehat{Q}^k_{\{m_i\}_{i=1}^k}\Big]\widehat{z}^k_{\{m_i\}_{i=1}^k},
\end{aligned}$$

$$\widehat{\rho}^k_{\{m_i\}_{i=1}^k} = \widehat{r}^k_{\{m_i\}_{i=1}^k} + \tfrac{1}{2}(\widehat{z}^k_{\{m_i\}_{i=1}^k})^T \Big[\widehat{G}^T \widehat{Q}^k_{\{m_i\}_{i=1}^k} \widehat{G} + \widehat{Q}^k_{\{m_i\}_{i=1}^k} \widehat{F} C \widehat{F} \widehat{Q}^k_{\{m_i\}_{i=1}^k}\Big] \widehat{z}^k_{\{m_i\}_{i=1}^k},$$

$$\widehat{F} \doteq (\widehat{Q}^k_{\{m_i\}_{i=1}^k} + C)^{-1},$$

and

$$\widehat{G} \doteq (\widehat{F}\widehat{Q}^k_{\{m_i\}_{i=1}^k} - I).$$

Thus, $\overline{W}^k$ has the representation as the set of triples

$$\mathcal{P}_k \doteq \left\{ \left(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k}\right) \,|\, m_i \in \mathcal{M} \ \forall i \in \{1,2,\ldots,k\} \right\}. \tag{7.71}$$

We note that the triples that comprise $\mathcal{P}_k$ are analytically obtained from the triples $(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k})$ by matrix multiplications and an inverse. The transference from

$$(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k})$$

to

$$(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k})$$

need only be done once, which is at the termination of the algorithm propagation. We note that (7.71) is our approximate solution of the original control problem/HJB PDE.

The errors are due to our approximation of $\widetilde{W}$ by $\overline{W}$ (see Theorem 7.32 and Remark 7.33), and to the approximation of $\overline{W}$ by the prelimit $\overline{W}^N$ for termination time $k = N$. Neither of these errors are related to the space dimension. The errors in $|\widetilde{W} - \overline{W}|$ are dependent on the step size τ. The errors in $|\overline{W}^N - \overline{W}| = |\bar{\bar{S}}^\tau_{N\tau}[0] - \overline{W}|$ are due to premature termination in the limit $\overline{W} = \lim_{N\to\infty} \bar{\bar{S}}^\tau_{N\tau}[0]$. The computation of each triple $(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k})$ grows like the cube of the space dimension (due to the matrix operations). Thus one avoids the curse-of-dimensionality. Of course, if one then chooses to compute $\overline{W}^N(x)$ for all x on some grid over say a rectangular region in $\mathbf{R}^n$, then by definition one has exponential growth in this computation as the space dimension increases. The point is that one does not need to compute $\overline{W}^N \simeq \widetilde{W}$ at each such point.

However, the curse-of-dimensionality is replaced by another type of rapid computational cost growth. Here, we refer to this as the curse-of-complexity. If $\#\mathcal{M} = 1$, then all the computations of our algorithm (excepting the solution of the Riccati equation) are unnecessary, and we *informally* refer to this as complexity one. When there are $M = \#\mathcal{M}$ such quadratics in the Hamiltonian, $\widetilde{H}$, we say it has complexity M. Note that

$$\#\left\{ \widehat{V}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M} \ \forall i \in \{1,2,\ldots,k\} \right\} \sim M^k.$$

For large k, this is indeed a large number. (We very briefly discuss means for reducing this in the next section.) Nevertheless, for small values of M, we obtain a very rapid solution of such nonlinear HJB PDEs, as will be indicated in the examples to follow. Further, the computational cost growth in space dimension n is limited to cubic growth.

7.5 Practical Issues

The bulk of this chapter develops an algorithm that avoids the curse-of-dimensionality. However, the curse-of-complexity is also a formidable barrier. The purpose of the chapter is to bring the existence of this class of algorithms to light. Considering the long development of finite element methods, it is clear that the development of highly efficient methods from this new class could be a further substantial achievement. (Nevertheless, some impressive computational times are already indicated in the next section.) In this section, we briefly indicate some practical heuristics that have been helpful.

7.5.1 Pruning

The number of quadratics in $\widehat{\mathcal{Q}}_k$ grows exponentially in k. However, in practice (for the cases we have tried) we have found that relatively few of these actually contribute to $\overline{W}^k$. Thus it would be very useful to prune the set.

Note that if

$$\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}(x) \le \bigoplus_{\{m_i\}_{i=1}^k \neq \{\hat{m}_i\}_{i=1}^k} \widehat{a}^k_{\{m_i\}_{i=1}^k}(x) \quad \forall\, x \in \mathbf{R}^n, \tag{7.72}$$

then

$$\int_{\mathbf{R}^n}^{\oplus} \overline{\mathcal{B}}_\tau(x,z) \otimes \overline{a}^k(z)\, dz = \int_{\mathbf{R}^n}^{\oplus} \overline{\mathcal{B}}_\tau(x,z) \otimes \left[\bigoplus_{\{m_i\}_{i=1}^k \neq \{\hat{m}_i\}_{i=1}^k} \widehat{a}^k_{\{m_i\}_{i=1}^k}(z) \right] dz.$$

That is, $\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}$ will play no role whatsoever in the computation of $\overline{W}^k$. Further, it is easy to show that the progeny of $\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}$ (i.e., those $\widehat{a}^{k+j}_{\{m_i\}_{i=1}^{k+j}}$ for which $\{m_i\}_{i=1}^k = \{\hat{m}_i\}_{i=1}^k$) never contribute either. Thus, one may prune such $\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}$ without any loss of accuracy. This shrinks not only the current $\widehat{\mathcal{Q}}_k$, but also the growth of the future $\widehat{\mathcal{Q}}_{k+j}$.

In the examples to follow, we pruned $\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}$ if there existed a single sequence $\{\tilde{m}_i\}_{i=1}^k$ such that $\widehat{a}^k_{\{\hat{m}_i\}_{i=1}^k}(x) \le \widehat{a}^k_{\{\tilde{m}_i\}_{i=1}^k}(x)$ for all x. This significantly reduced the growth in the size of $\widehat{\mathcal{Q}}_k$. However, it clearly failed to prune anywhere near the number of elements that could be pruned according to condition (7.72), and thus much greater computational reduction might be possible. This would require an ability to determine when a quadratic was dominated by the maximum of a set of other quadratic functions.

Also in the examples to follow, an additional heuristic pruning technique was applied employed for a number of iterations to delay hitting the curse-of-complexity growth rate. A function $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ was pruned if it did not dominate at at least one of the corners of the unit cube. Specifically, let $\mathcal{C} = \{x^j\}$ be the corners of the unit cube. The set of functions was pruned down to a subset of $L \leq 2^n$ functions, $\{\widehat{a}^k_{\{\hat{m}_i^l\}_{i=1}^k} \mid l \leq L\}$, such that $\overline{a}^k(x^j) = \max_{l \leq L} \widehat{a}^k_{\{\hat{m}_i^l\}_{i=1}^k}(x^j)$ for all $x^j \in \mathcal{C}$. This introduces a component of the calculations which is subject to curse-of-dimensionality growth, but in the examples run to date, it reduced the computational times over what was needed without the heuristic. (Also, the curse-of-dimensionality growth due to this heuristic is 2^n rather than on the order of 100^n for grid-based methods.)

7.5.2 Initialization

It is also easy to see that one may initialize with an arbitrary quadratic function less than an $\overline{a}^k(x)$ rather than with $\overline{a}^0 \equiv 0$. Significant savings are obtained by initializing with a set of $M = \#\mathcal{M}$ quadratics, $\{a^m(x)\}$ where the a^m are the convex duals of the W^m (which are each obtained by solution of the corresponding Riccati equation). With $\overline{a}^0(z) \doteq \bigoplus_{m \in \mathcal{M}} a^m(z)$, one starts much closer to the final solution, and so the number of steps where one is encountering the curse-of-complexity is greatly reduced. For the more general quadratics of Section 7.7 below, determining how to make improvements through initialization may be less trivial.

7.6 Examples

A number of examples have so far been tested. In these tests, the computational speeds were very great. This is due to the fact that $M = \#\mathcal{M}$ was small. The algorithm as described above was coded in MATLAB. This includes the very simple pruning technique and initialization discussed in the previous section. The quoted computational times were obtained with a standard 2001 PC. The times correspond to the time to compute $\overline{W}^N$ or, equivalently, $\mathcal{P}_N$. The plots below require one to compute the value function and/or gradients pointwise on planes in the state space. These plotting computations are not included in the quoted computational times.

We will briefly indicate the results of three similar examples with state space dimensions of 2, 3, and 4. The number of constituent linear/quadratic Hamiltonians for each of them is three. The structures of the dynamics are similar for each of them so as to focus on the change in dimension.

The first example has constituent Hamiltonians with the A^m given by

$$A^1 = \begin{bmatrix} -1.0 & 0.5 \\ 0.1 & -1.0 \end{bmatrix}, \quad A^2 = (A^1)^T, \quad A^3 = \begin{bmatrix} -1.0 & 0.5 \\ 0.5 & -1.9 \end{bmatrix}.$$

The D^m and Σ^m were simply

$$D^1 = D^2 = D^3 = \begin{bmatrix} 1.5 & 0.2 \\ 0.2 & 1.5 \end{bmatrix},$$

and

$$\Sigma^1 = \Sigma^2 = \Sigma^3 = \begin{bmatrix} 0.27 & -0.01 \\ -0.01 & 0.27 \end{bmatrix}.$$

Figure 7.1 depicts the value function and first partial derivative (computed by a simple first-difference on the grid points) over the region $[-1,1] \times [-1,1]$. Note the discontinuity in the first partial along one of the diagonals. Figure 7.2 depicts the second partial and a backsubstitution error over the same region. The second partial also has a discontinuity along the same diagonal as the first. The error plot has been rotated for better viewing due to the high error along the discontinuity in the gradient. The backsubstitution error is computed by taking these approximate partials and substituting them back into the original HJB PDE. Consequently the depicted errors contain components due to the approximate gradient dotted in with the dynamics and the term with the square in the gradient in the Hamiltonian. Perhaps it should be noted that the solutions of such problems *cannot* be obtained by patching together the quadratic functions corresponding to solutions of the corresponding algebraic Riccati equations. The computations required slightly less than 10 seconds.

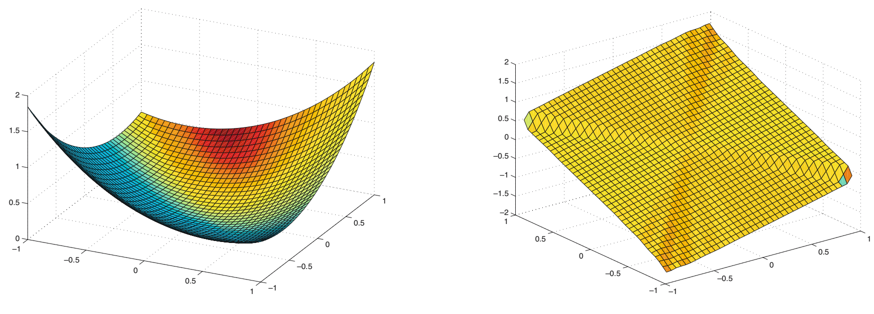


Fig. 7.1. Value function and first partial (2-D case)

The second example has constituent Hamiltonians with the A^m given by

$$A^1 = \begin{bmatrix} -1.0 & 0.5 & 0.0 \\ 0.1 & -1.0 & 0.2 \\ 0.2 & 0.0 & -1.5 \end{bmatrix}, \quad A^2 = (A^1)^T, \quad A^3 = \begin{bmatrix} -1.0 & 0.5 & 0.0 \\ 0.1 & -1.0 & 0.2 \\ 0.2 & 0.0 & -1.5 \end{bmatrix}.$$

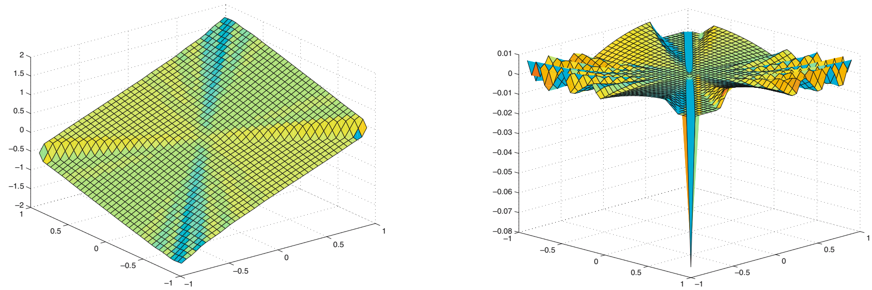


Fig. 7.2. Value function and first partial (2-D case)

The D^m were

$$D^1 = \begin{bmatrix} 1.5 & 0.2 & 0.1 \\ 0.2 & 1.5 & 0.0 \\ 0.1 & 0.0 & 1.5 \end{bmatrix}, \quad D^2 = \begin{bmatrix} 1.6 & 0.2 & 0.1 \\ 0.2 & 1.6 & 0.0 \\ 0.1 & 0.0 & 1.6 \end{bmatrix}, \quad D^3 = D^1.$$

The Σ^m were

$$\Sigma^1 = \begin{bmatrix} 0.2 & -0.01 & 0.02 \\ -0.01 & 0.2 & 0.0 \\ 0.02 & 0.0 & 0.25 \end{bmatrix}, \quad \Sigma^2 = \begin{bmatrix} 0.16 & -0.005 & 0.015 \\ -0.005 & 0.16 & 0.0 \\ 0.015 & 0.0 & 0.2 \end{bmatrix},$$
$$\Sigma^3 = \Sigma^1.$$

The results of this three-dimensional example appear in Figures 7.3–7.5. In this case, the results have been plotted over the region of the affine plane $x_3 = 3$ given by $x_1 \in [-10, 10]$ and $x_2 \in [-10, 10]$. The backsubstitution error has been scaled by dividing by $|x|^2 + 10^{-5}$. Note that the scaled backsubstitution errors (away from the discontinuity in the gradient) grow only slowly or are possibly bounded with increasing $|x|$. (Recall that the approximate solution is obtained over the whole space.) Because the gradient errors are multiplied by the nominal dynamics in one component of this term (as well as being squared in another), this indicates that the errors in the gradient itself likely grow only linearly (or nearly linearly) with increasing $|x|$. The computations required approximately 13 seconds.

The third example has constituent Hamiltonians with the A^m given by

$$A^1 = \begin{bmatrix} -1.0 & 0.5 & 0.0 & 0.1 \\ 0.1 & -1.0 & 0.2 & 0.0 \\ 0.2 & 0.0 & -1.5 & 0.1 \\ 0.0 & -0.1 & 0.0 & -1.5 \end{bmatrix}, \quad A^2 = (A^1)^T,$$

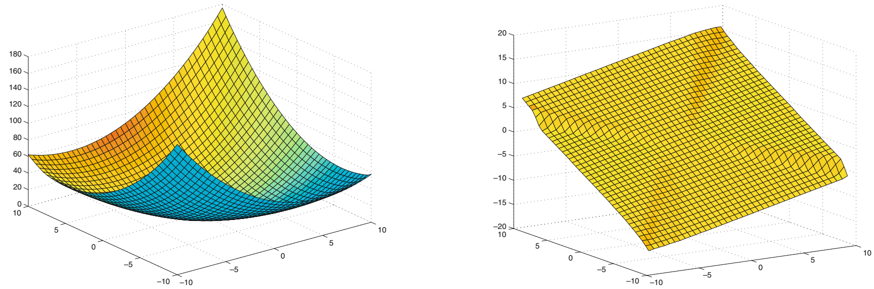


Fig. 7.3. Value function and first partial (3-D case)

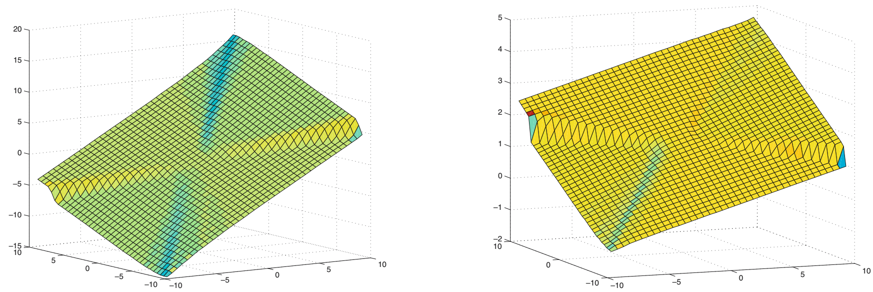


Fig. 7.4. Partials with respect to second and third variables (3-D case)

$$A^3 = \begin{bmatrix} -1.0 & 0.5 & 0.0 & 0.1 \\ 0.1 & -1.0 & 0.2 & 0.0 \\ 0.2 & 0.0 & -1.6 & -0.1 \\ 0.0 & -0.05 & 0.1 & -1.5 \end{bmatrix}.$$

The D^m and Σ^m were simply

$$D^1 = D^2 = D^3 = \begin{bmatrix} 1.5 & 0.2 & 0.1 & 0.0 \\ 0.2 & 1.5 & 0.0 & 0.1 \\ 0.1 & 0.0 & 1.5 & 0.0 \\ 0.0 & 0.1 & 0.0 & 1.5 \end{bmatrix},$$

and

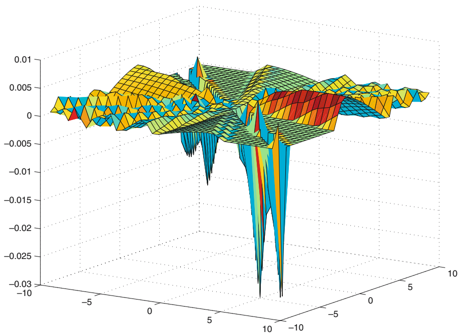


Fig. 7.5. Scaled backsubstitution error (3-D case)

$$\Sigma^1 = \Sigma^2 = \Sigma^3 = \begin{bmatrix} 0.2 & -0.01 & 0.02 & 0.01 \\ -0.01 & 0.2 & 0.0 & 0.0 \\ 0.02 & 0.0 & 0.25 & 0.0 \\ 0.01 & 0.0 & 0.0 & 0.25 \end{bmatrix}.$$

The results of this four-dimensional example appear in Figures 7.6–7.8. In this case, the results have been plotted over the region of the affine plane $x_3 = 3$, $x_4 = -0.5$ given by $x_1 \in [-10, 10]$ and $x_2 \in [-10, 10]$. The backsubstitution error has again been scaled by dividing by $|x|^2 + 10^{-5}$. The computations required approximately 40 seconds. We remark that one cannot change dimension independent of dynamics (except in the trivial case where each component of the system has exactly the same dynamics of the other components with no interdependence), and so one cannot directly compare the computation times of these three examples. However, it is easy to see that the computation time increases are on the order of square to cubic in space dimension, rather than being subject to curse-of-dimensionality type growth.

7.7 More General Quadratic Constituents

The examples given in the previous section all possessed similar structures where the partial derivatives were linear along straight lines passing through the origin. This was due to the fact that the constituent Hamiltonians all had the structure

$$H^m(x, p) = \tfrac{1}{2} x^T D^m x + \tfrac{1}{2} p^T \Sigma^m p + (A^m x)^T p. \tag{7.73}$$

This implied that all the iterates had the form

$$\overline{W}^k(x) = \max_{\{m_i\}_{i=1}^k \in \mathcal{M}^k} \tfrac{1}{2} x^T \widehat{P}^k_{\{m_i\}_{i=1}^k} x,$$

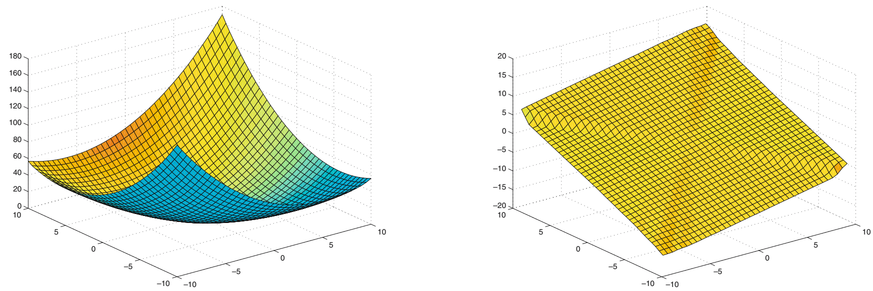


Fig. 7.6. Value function and first partial (4-D case)

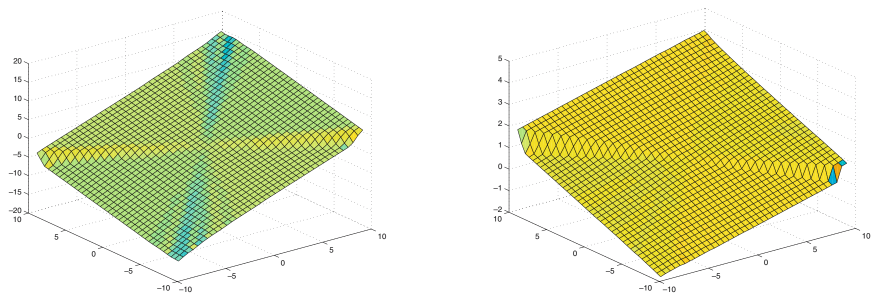


Fig. 7.7. Partials with respect to second and third variables (4-D case)

that is, the linear and constant terms were zero. Thus the $\overline{W}^k$ were quadratic along lines $\ell_v : \ \{x = cv \,|\, c \in \mathbf{R}\}$ where $v \in \mathbf{R}^n$. On the other hand, the spaces of semiconvex functions are quite large, and any semiconvex function can be expanded as a pointwise maximum of quadratic forms. Thus, we now expand our class of constituent Hamiltonians to be of the form

$$H^m(x,p) = \tfrac{1}{2}x^T D^m x + \tfrac{1}{2}p^T \Sigma^m p + (A^m x)^T p + (l_1^m)^T x + (l_2^m)^T p + \alpha^m, \quad (7.74)$$

where $l_1^m, l_2^m \in \mathbf{R}^n$ and $\alpha^m \in \mathbf{R}$. From Chapter 2, we know that any semiconvex Hamiltonian, say H^{sc}, can be arbitrarily well-approximated on a ball, $B_R \subset \mathbf{R}^n \times \mathbf{R}^n$, by some finite maximum of quadratic constituent Hamiltonians, i.e., $H^{sc}(x,p) \simeq \max_m H^m(x,p)$ with H^m of the form (7.74). Consequently, we now expand our class of HJB PDEs to

$$0 = -\widetilde{H}(x, \nabla W), \qquad W(0) = 0 \qquad (7.75)$$

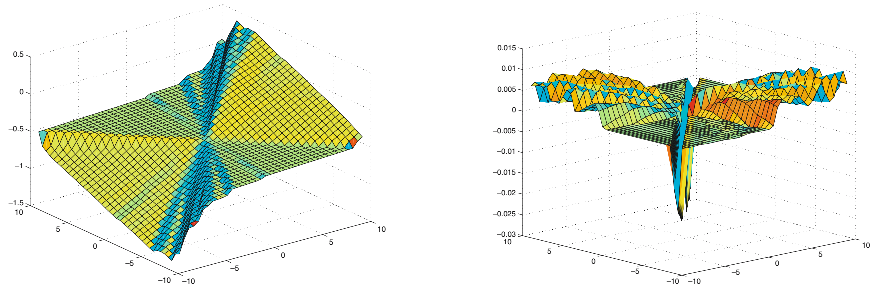


Fig. 7.8. Fourth partial and scaled backsubstitution error (4-D case)

with

$$\widetilde{H}(x,p) = \max_{m \in \mathcal{M}} H^m(x,p),$$

where the H^m have the form (7.74), and $\mathcal{M}$ is a finite set of indices.

We will not provide the theory, analogous to that in Sections 7.1–7.4, in this case. We do note that it appears that existence of a constituent Hamiltonian of pure quadratic form (7.73) satisfying our usual conditions (plus existence of a solution to problem (7.75) of course) may be sufficient to guarantee that the above approach will work in this wider class (when $\widetilde{W}$ exists of course). Instead, our goal here will only be to indicate some of the wider range of behaviors that can be captured within this larger class. We present two simple examples where a constant term has been added to one of the Hamiltonians.

For the first example, consider our standard problem, $0 = -\widetilde{H}(x, \nabla W)$, $W(0) = 0$, where

$$\widetilde{H}(x,p) = \max\{H^1(x,p),\ \widehat{H}^2(x,p)\}$$

with

$$H^1(x,p) = \frac{1}{2}x^T C x + (Ax)^T p + \frac{1}{2}p^T \Sigma p$$
$$\widehat{H}^2(x,p) = \frac{1+\delta}{2}x^T C x + (Ax)^T p + \frac{1}{2}p^T \Sigma p - \frac{\alpha^2}{2}$$

with specific coefficients

$$A = \begin{bmatrix} -1 & 0.5 \\ 0.1 & -1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1.5 & 0.2 \\ 0.2 & 1.5 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 0.216 & -0.008 \\ -0.008 & 0.216 \end{bmatrix}$$
$$\delta = 0.4, \quad \text{and} \quad \alpha^2 = 0.4.$$

Let $\mathcal{R}_1 \doteq \{(x,p) \in \mathbf{R}^4 |\, H^1(x,p) > \widehat{H}^2(x,p)\}$. Setting $H^1 = \widehat{H}^2$, one finds $\partial\mathcal{R}_1$ as the cylindrical ellipsoid

$$\partial\mathcal{R}_1 = \{(x,p)|\, x^T C x = \alpha^2/\delta\}.$$

The boundary, $x = 0$, lies inside $\mathcal{R}_1$, and we can denote the solution of $0 = -H^1(x, \nabla W(x))$, $W(0) = 0$ on $\mathcal{R}_1$ as W_1. The astute observer will note that one must verify that the characteristic flow of $\widetilde{H}$ must be such that the characteristics are flowing outward through $\partial\mathcal{R}_1$ in order to claim that W_1 on $\mathcal{R}_1$ will be identical to the solution of the $\widetilde{H}$ problem restricted to that region.

It is interesting to note that constituent problem $-\widehat{H}^2(x, \nabla W(x)) = 0$ on the complement, $\mathcal{R}_1^c$ with boundary condition $W(x) = W_1(x)$ on $\partial\mathcal{R}_1$ is equivalent to

$$0 = -H^2(x, \nabla w) = \frac{1}{2} x^T C x + \frac{1}{1+\delta}(Ax)^T \nabla W + \frac{1}{2(1+\delta)} \nabla W^T \Sigma \nabla W$$

with boundary condition $W(x) = W_1(x)$ on $\partial\mathcal{R}_1$. As before, with $\Sigma = \sigma\sigma^T$, it is easily seen that this is equivalent to

$$0 = -\max_{u \in \mathbf{R}^l} \left\{ \left[\frac{1}{1+\delta} Ax + \sigma u \right]^T \nabla W + \frac{1}{2} x^T C x - \frac{1+\delta}{2} |u|^2 - \frac{\alpha^2}{2} \right\}$$

with boundary condition $W(x) = W_1(x)$ on $\partial\mathcal{R}_1$. Of course, one also has

$$H^1(x, \nabla W) = \max_{u \in \mathbf{R}^l} \left\{ [Ax + \sigma u]^T \nabla W + \frac{1}{2} x^T C x - \frac{1}{2} |u|^2 \right\}.$$

Therefore, solving $0 = -\widetilde{H}(x, \nabla W)$ with $W(0) = 0$ is equivalent to determining the value function of control problem with dynamics

$$\dot{\xi} = \begin{cases} A\xi + \sigma u & \text{if } \xi \in \mathcal{R}_1 \\ \frac{1}{1+\delta} A\xi + \sigma u & \text{otherwise,} \end{cases}$$

and payoff and value given by

$$J(x,T,u) = \int_0^T \widetilde{L}(\xi_t, u_t)\, dt,$$
$$V(x) = \sup_{u \in \mathcal{U}} \sup_{T < \infty} J(x,T,u),$$

where

$$\widetilde{L}(x,u) = \begin{cases} \frac{1}{2} x^T C x - \frac{1}{2} |u|^2 & \text{if } \xi \in \mathcal{R}_1 \\ \frac{1}{2} x^T C x - \frac{1+\delta}{2} |u|^2 - \frac{\alpha^2}{2} & \text{otherwise.} \end{cases}$$

Thus there is a change in the dynamics across $\partial\mathcal{R}_1$ with a corresponding change in the cost criterion.

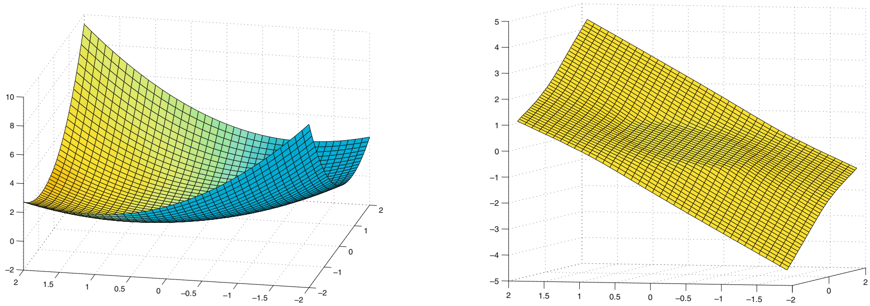


Fig. 7.9. Value function and partial with respect to first variable

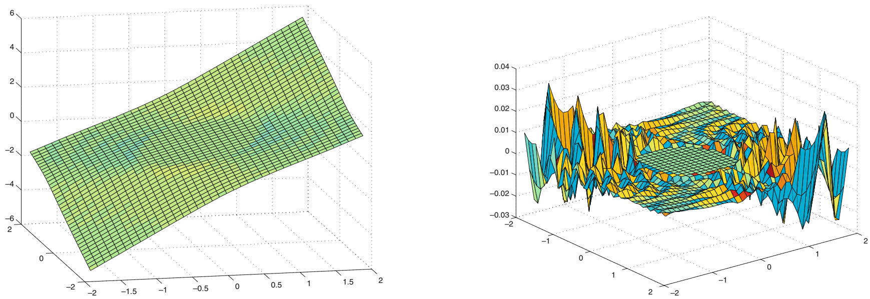


Fig. 7.10. Partial with respect to second variable and backsubstitution errors

The value, partial derivatives with respect to first and second variables, and the backsubstitution error are depicted in Figures 7.9 and 7.10. The discontinuity in the second derivative along the boundary of the ellipse $x^T Cx = \alpha^2/\delta$ is clearly evident.

As another example, we make a small change in the coefficients of the above problem, and see a change in the structure of the solutions. In particular, we consider the same $\widetilde{H}$ as in the previous example with the exception that we change matrices A and C to be

$$A = \begin{bmatrix} -2 & 1.6 \\ -1.6 & -0.4 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1.4 & 0.25 \\ 0.25 & 1.6 \end{bmatrix}.$$

The value, partial derivatives and backsubstitution errors appear in Figures 7.11 and 7.12. Note the rotational effects induced by the change in A.

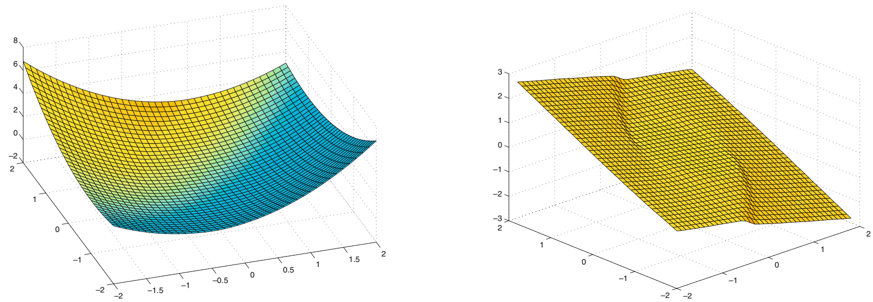


Fig. 7.11. Value function and partial with respect to first variable

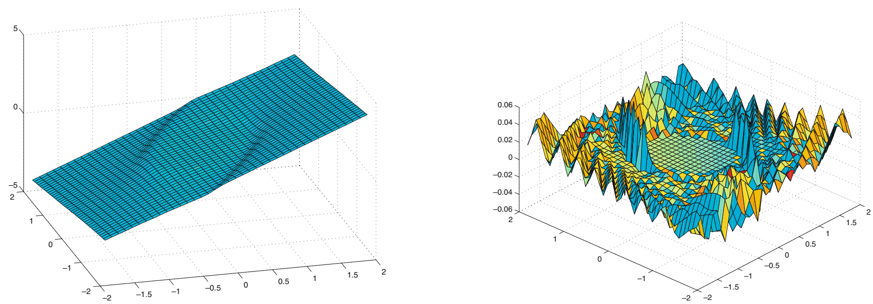


Fig. 7.12. Partial with respect to second variable and backsubstitution errors

7.8 Future Directions

Pruning

In order to make these methods more practical, algorithms need to be developed for determining when a quadratic function is dominated by the function that is the pointwise maximum of a set of quadratic functions. This has the potential for greatly reducing the effects of the curse-of-complexity, and consequently greatly decreasing computational times.

Wider Classes of Constituent Hamiltonians

The theory for an instantiation of this class of methods was developed here only for the very particular type of Hamiltonian, $\widetilde{H}(x,p) = \max_m\{H^m(x,p)\}$, where the H^m corresponded to a very specific type of quadratic problem (see (7.5)). As indicated in the previous section, a much richer class of problems may be considered through constituent Hamiltonians of the form (7.74). The theory for this wider class is not yet fully understood. Further, in the work here, the H^m corresponded to linear/quadratic problems with maximizing controllers/disturbances. It is not known whether the constituent linear/quadratic problems need to be constricted in this way either. For instance, could some or all of the H^m correspond to say game problems?

Convergence/Error Analysis

Only convergence of the approximation to the solution has so far been obtained. Estimates of error size and convergence rate need to be determined. For instance, it was hypothesized (and roughly observed in the examples) that one obtains the solution over the whole state space with linear growth rate in the errors in the gradient.

Other Nonlinearities

The model problem in this chapter considered only the case of a nonlinearity due to taking the maximum of a set of Hamiltonians for linear/quadratic problems. An obvious question is how well this approach might work for other classes of nonlinearities. What classes of nonlinear HJB PDEs could be best approximated by maxima over reasonably small numbers of linear/quadratic HJB PDEs? Perhaps a single nonlinearity in only one variable (possibly appearing in multiple places) would be the most tractable.

8

Finite Time-Horizon Application: Nonlinear Filtering

In this chapter we consider a finite time-horizon application. An obvious example where this technology may be applicable is in nonlinear filtering — specifically robust/H_∞ nonlinear filtering. If the dynamics are time-invariant, then between each successive pair of observations, one uses the same HJB PDE to propagate the information state (to be discussed below) forward in time. If the observations are at times $t_i = t_0 + i\delta_t$, then one is repeatedly solving the same HJB PDE for periods δ_t between observations; only the "initial conditions" at the start of each time period change. With the max-plus approach, the real-time solution of the PDE can be replaced by a max-plus matrix-vector multiplication where the matrix is fixed, independent of time-step. Thus, the matrix may be precomputed in advance rather than in real-time. We will use this filtering application as the vehicle for motivation and development of a max-plus approach to finite time-horizon problems. With respect to the filtering aspects, some standard details will only be sketched; further material and a continuous-time observation case are discussed in [44], [87]. It should be noted that the approach in this chapter is analogous to results (over the standard algebra) for estimation of nonlinear stochastic systems as obtained by Rozovskii et al. (c.f. [103]) In fact, the existence of such results in the stochastic case provided motivation for the developments presented here.

We consider a system with dynamics

$$\dot{\xi} = f(\xi) + \sigma(\xi)w_t \tag{8.1}$$

and unknown initial condition, ξ_0. We wish to estimate the state, ξ_t at some time $t \geq 0$ using data up to that time. The measure of estimate quality will be defined below. As usual, the state will take values in $\mathbf{R}^n$, and we let the disturbance process $w.$ take values in $\mathcal{W} \doteq L_2^{loc}([0,\infty);\mathbf{R}^l)$. Note that we will not observe $w.$ directly. As usual, σ will be $n \times l$ matrix-valued. Let the observations occur at times $t_i = i\delta_t$ for $i = 1, 2, \ldots$. The observation at time t_i, y_i, will be given by

$$y_i = h(\xi_{t_i}) + \rho(\xi_{t_i})v_i \tag{8.2}$$

where $y_i \in \mathbf{R}^k$. The $v_i \in \mathbf{R}^\kappa$ are unknown and ρ is $k \times \kappa$ matrix-valued. We will assume that

f, σ, h and ρ are all smooth, and that f, σ and h are globally Lipschitz in x. (A8.1)

We also assume that there exists $m_\sigma < \infty$ such that

$$|\sigma(x)| \le m_\sigma \quad \forall x \in \mathbf{R}^n. \tag{A8.2}$$

We will assume that

Range$(\rho(x)) = \mathbf{R}^k$ for all $x \in \mathbf{R}^n$ (which guarantees that for any y_i, ξ_{t_i} there exists some v_i satisfying (8.2)), that $\rho^{-1} \in C^2$, and that there exist $m_\rho < \infty$ such that $|\rho^{-1}(x)b| \le m_\rho |b|$ for all $x \in \mathbf{R}^n$ and all $b \in \mathbf{R}^k$, (A8.3)

where we use the Moore–Penrose inverse [54]

$$\rho^{-1}(x)b = \operatorname{argmin}\{|v| : \rho(x)v = b\}. \tag{8.3}$$

We will also assume that

Range$(\sigma(x)) = \mathbf{R}^n$ for all $x \in \mathbf{R}^n$, that $\sigma^{-1} \in C^2$, and that $|\sigma^{-1}(x)b| \le M_\sigma |b|$ for all $x \in \mathbf{R}^n$ and all $b \in \mathbf{R}^n$. (A8.4)

where we are again using the Moore–Penrose inverse. For notational simplicity, we let $a \doteq \sigma\sigma^T$.

Note also that these assumptions imply that if we view the integral version of (8.1)

$$x_T \doteq \xi_T = \xi_0 + \int_0^T f(\xi_t) + \sigma(\xi_t) w_t \, dt \tag{8.4}$$

as a mapping from ξ_0 to x_T then this mapping is one-to-one and onto for any $w \in L_2$. Let $\phi : \mathbf{R}^n \to \mathbf{R}$. Assume there exists $m_\phi < \infty$ such that

$\phi(x) \le m_\phi(1 + |x|^2) \quad \forall x \in \mathbf{R}^n$, and that ϕ is locally Lipschitz. (A8.5)

In order to design the nonlinear robust/H_∞ filter, we define the following payoff

$$J_f(T, x_T, w) = -\tfrac{\zeta^2}{2}\phi(\xi_0) - \frac{\gamma^2}{2}\int_0^T |w(t)|^2 \, dt - \tfrac{\eta^2}{2} \sum_{i=1}^{N_T} |v_i|^2, \tag{8.5}$$

which, using Assumption (A8.3),

$$\begin{aligned} &= -\tfrac{\zeta^2}{2}\phi(\xi_0) - \frac{\gamma^2}{2}\int_0^T |w(t)|^2 \, dt \\ &\quad -\tfrac{\eta^2}{2} \sum_{i=1}^{N_T} |\rho^{-1}(\xi_{t_i})[y_i - h(\xi_{t_i})]|^2 \end{aligned} \tag{8.6}$$

where ξ_0 is given by (8.4) for any particular $\xi_T = x_T$ and w, and $N_T \doteq \max\{i \in \mathcal{N} \,|\, t_i \le T\}$. Further, define the value function

$$P(T,x_T) = \sup_{w\in\mathcal{W}} J_f(T,x_T,w). \tag{8.7}$$

This value function will be referred to as an *information state* (see, e.g., [12], [52], [61] for more detail). In particular, at any time $T \geq 0$, $P(T,\cdot)$ will contain sufficient information to estimate the state ξ_T. Note that J and P depend on the observation sequence, $\{y_i\}$ and initial $\phi(\cdot)$, although this will be suppressed. The filtering goal will be to obtain a robust state estimate given a specific observation sequence (up to the current time) and initial information $\phi(\cdot)$.

There are multiple estimators that have been defined using this information state. One estimator, a "minimum energy" estimator, was developed by Mortensen in the 1960s. In that case, the estimate of ξ_t is given by $\operatorname{argmax}_{x\in\mathbf{R}^n} P(T,x)$ (see [41], [95]). However, we are interested in a filter estimate that has an interpretation in a robust/H_∞ sense. This estimator is also the natural estimator in that it is essentially the maximum likelihood estimator in a certain max-plus probabilistic sense. (Max-plus probability is discussed in [39], [40], [84], [97], [98] among others.) Further, it is the limit of a risk-sensitive estimator [43]. In the state-space interpretation of H_∞ control (c.f. [12], [52], [108]), one develops a controller where for all (finite-energy) disturbance inputs, the cost (typically L_2) is bounded by a multiple of a measure of the disturbance energy (where we use the term "energy" in a generalized sense); see Chapter 1. In analogy with that, this robust/H_∞ estimator is chosen so that the squared estimate error is bounded by a measure of disturbance energy. More specifically, at any time, T, one wants the estimate, $\hat{e}_T$, to satisfy (for some $\zeta,\gamma,\eta \in \mathbf{R}$)

$$|\hat{e}_T - \xi_T|^2 \leq \left[\tfrac{\zeta^2}{2}\phi(\xi_0) + \frac{\gamma^2}{2}\int_0^T |w_t|^2\,dt + \tfrac{\eta^2}{2}\sum_{i=1}^{N_T}|v_i|^2\right], \tag{8.8}$$

where ξ satisfies (8.1) for any possible $\xi_0, w_\cdot, v_\cdot$. (Note that these are not independent quantities given an observed $y_\cdot$ process.) The existence of such an estimator is guaranteed given that there exists $\bar{x} \in \mathbf{R}^n$ such that $\phi(x) \geq |x-\bar{x}|^2$ and that ζ,γ,η are sufficiently large. Further, under this condition (and the above assumptions), there exist $\zeta,\gamma,\eta \in \mathbf{R}$ such that (8.8) is satisfied by the estimator

$$\hat{e}_T \doteq \operatorname*{argmin}_e \max_x \left[|x-e|^2 + P(T,x)\right] \tag{8.9}$$

(where the existence of unique minimizing e and maximizing x are guaranteed). Proofs of these claims can be found in Appendix A. See also [41], [87]. Some readers may note that the above information state differs from some others which include an integral running cost term. Such other information states are particularly suited to the problem of H_∞ control under partial information as opposed the problem of robust state estimation considered here; see [12], [61].

Clearly the propagation of the robust/H_∞ filter estimates requires the forward propagation of the information state, P. We will propagate $P(T,\cdot)$ forward in time (approximately) by max-plus matrix-vector multiplication operating on the (time-dependent) vector of coefficients in the max-plus expansion of $P(T,\cdot)$. We begin, as always, with the DPP which we present without proof (but see Chapter 3 for general DPP arguments for a finite time-horizon problem).

Theorem 8.1. *Let $T, T-\delta \in [i\delta_t, (i+1)\delta_t)$ for some integer $i \geq 0$. Then for any $x \in \mathbf{R}^n$,*

$$P(T,x) = \sup_{w\in\mathcal{W}} \left\{ P(T-\delta, \xi_{T-\delta}) - \frac{\gamma^2}{2}\int_{T-\delta}^{T} |w_t|^2\, dt \right\},$$

where ξ_t satisfies (8.1) with terminal condition $\xi_T = x$. Further, letting $P^-(t_{i+1},x) = \lim_{t\uparrow t_{i+1}} P(t,x)$ for all non-negative integers i, one has

$$P^-(t_{i+1},x) = \sup_{w\in\mathcal{W}} \left\{ P(T-\delta, \xi_{T-\delta}) - \frac{\gamma^2}{2}\int_{T-\delta}^{t_{i+1}} |w_t|^2\, dt \right\},$$

where ξ_t satisfies (8.1) with terminal condition $\xi_{t_{i+1}} = x$.

Although we will not be using the viscosity solution representation, combining Lemma A.10 in the appendix and the machinery from Chapter 3, one can show that the information state is given as a viscosity solution of the associated HJB PDE between observation times.

Theorem 8.2. *Over any time interval $[t_i, t_{i+1})$ (for non-negative integer i), the information state, P is the unique viscosity solution of*

$$\begin{aligned} 0 &= \mathcal{P}_T - H(x, \nabla_x \mathcal{P}) \\ &\doteq \mathcal{P}_T - \left[\frac{\gamma^2}{2} \nabla_x \mathcal{P}^T a(x) \nabla_x \mathcal{P} - f^T(x)\nabla_x\mathcal{P} \right], \\ \mathcal{P}(t_i,x) &= P(t_i,x). \end{aligned} \tag{8.10}$$

As usual, we use the DPP as a guide for defining a semigroup operator which will be max-plus linear. Let

$$S_\tau[\psi](x) \doteq \sup_{w\in\mathcal{W}} \left\{ \psi(\xi_0) - \frac{\gamma^2}{2}\int_0^\tau |w_t|^2\, dt \right\}, \tag{8.11}$$

where ξ_t satisfies (8.1) with terminal condition $\xi_\tau = x$, and the domain is implicit (but will include classes of semiconvex functions). Note that time is reversed in this definition relative to our usual semigroup, but it is not difficult to verify that the semigroup property remains valid. A result that is, by now, obvious is the following:

Theorem 8.3. *S_τ is a max-plus linear operator for any $\tau > 0$.*

Using Theorem 8.1, we find that for $\tau \in [0, \delta_t)$ and any $i \geq 0$,

$$P(t_i + \tau, x) = S_\tau[P(t_i, \cdot)](x), \tag{8.12}$$

and that

$$P(t_{i+1}, x) = P^-(t_{i+1}, x) - \tfrac{\eta^2}{2}|\rho^{-1}(x)[y_{i+1} - h(x)]|^2, \tag{8.13}$$

where

$$P^-(t_{i+1}, x) = S_{\delta_t}[P(t_i, \cdot)](x) \tag{8.14}$$

for all $x \in \mathbf{R}^n$. Note that (8.12)–(8.14) define the forward propagation of the information state.

8.1 Semiconvexity

In order to make use of this propagation via the max-plus linear operator, as usual we need to specify appropriate max-plus spaces of functions and max-plus bases for these spaces. The appropriate spaces are of course spaces of semiconvex functions. We first demonstrate that the information state is indeed semiconvex.

In [44], two proofs of semiconvexity under slightly differing assumptions are given. The following proof of semiconvexity is adapted from a proof appearing there.

Let us introduce the following function, $\mathcal{V}^{fs}(T, x_0, x_T)$, which we refer to as a *fundamental solution* of (8.10). For $x_0, x_T \in \mathbf{R}^n$, $T > 0$, let

$$\mathcal{V}^{fs}(T, x_0, x_T) = \sup_{w \in L_2((0,T);\mathbf{R}^l)} \left\{ -\frac{\gamma^2}{2} \int_0^T |w(t)|^2 \, dt : \, \xi_0 = x_0, \, \xi_T = x_T \right\}, \tag{8.15}$$

where $\xi.$ satisfies (8.1). From (8.11),

$$S_T[\phi](x) = \sup_{x_0 \in \mathbf{R}^n} \Big\{ \mathcal{V}^{fs}(T, x_0, x) + \phi(x_0) \Big\}. \tag{8.16}$$

We can rewrite $\mathcal{V}^{fs}$ in terms of the following calculus of variations problem with fixed initial and terminal conditions. Let

$$L(\xi, \dot{\xi}) = \frac{\gamma^2}{2}|\sigma^{-1}(\xi)(\dot{\xi} - f(\xi))|^2,$$

$$I(T, x_0, x_T; \xi.) = -\int_0^T L(\xi_t, \dot{\xi}_t) \, dt.$$

Then

$$\mathcal{V}^{fs}(T, x_0, x_T) = \sup\{I(T, x_0, x_T; \xi.) : \, \xi_0 = x_0, \, \xi_T = x_T\}$$

with $\xi_\cdot$ satisfying (8.1). Equivalence of (8.15) and this calculus of variations version follows easily by noting that for each path, $\xi_\cdot$, there is a corresponding unique minimal-norm $w_\cdot$ given in feedback form with the Moore–Penrose definition of σ^{-1}.

Lemma 8.4. *Given $\delta_t, R, R_0 \in (0,\infty)$, there exists $c = c(\delta_t, R, R_0)$ such that $\mathcal{V}^{fs}(T, x_0, x) + (c/2)|x|^2$ is convex on the closed ball $\overline{B}_{RT}(\widetilde{x})$ where $\widetilde{x} \doteq x_0 + f(x_0)T$ for all $T \in (0, \delta_t]$ and all $|x_0| \le R_0$.*

Proof. Suppose $T \in (0, \delta_t]$, $|x_0| \le R_0$, and

$$|x - [x_0 + f(x_0)T]| \le RT. \tag{8.17}$$

Note that (8.17) implies

$$|x - x_0| \le RT + |f(x_0)|T. \tag{8.18}$$

By Assumption (A8.1), there exists $K_f < \infty$ such that

$$|f(x_1) - f(x_1)| \le K_f |x_1 - x_2| \quad \forall x_1, x_2 \in \mathbf{R}^n \tag{8.19}$$

$$|f(x_1)| \le K_f(1 + |x_1|) \quad \forall x_1 \in \mathbf{R}^n \tag{8.20}$$

By (8.18) and (8.20),

$$|x - x_0| \le (R + K_f + K_f R_0)T \tag{8.21}$$

Let

$$\overline{\xi}_t \doteq x_0 + \left(\frac{t}{T}\right)(x - x_0).$$

Then

$$\left|\dot{\overline{\xi}} - f(\overline{\xi})\right| = \left|\tfrac{1}{T}[x - (x_0 + f(x_0)T)] + f(x_0) - f\left(x_0 + \left[\frac{t}{T}\right](x - x_0)\right)\right|,$$

which by (8.17)

$$\le R + \left|f(x_0) - f\left(x_0 + \left[\frac{t}{T}\right](x - x_0)\right)\right|,$$

which by (8.19)

$$\le R + K_f \left|\left[\frac{t}{T}\right](x - x_0)\right|,$$

which by (8.21)

$$\le R + K_f t(R + K_f + K_f R_0)$$
$$\le R + K_f \delta_t (R + K_f + K_f R_0) \doteq C_1.$$

for all $t \in [0, \delta_t]$. Combining this bound with the definition of L and (A8.4), one finds

$$-L(\bar{\xi}_t, \dot{\bar{\xi}}_t) \geq -\frac{\gamma^2}{2} M_\sigma^2 C_1^2 \quad \forall t \in [0, \delta_t]. \tag{8.22}$$

On the other hand, given x_0, x, there exists optimal $\xi^*_\cdot$ (c.f. [46]) such that $\xi^*_{t_i} = x_0$, $\xi^*_T = x$ and

$$\mathcal{V}^{fs}(T, x_0, x) = I(T, x_0, x; \xi^*_\cdot).$$

Then by the optimality of ξ^* and (8.22),

$$I(T, x_0, x; \xi^*_\cdot) \geq I(T, x_0, x, \bar{\xi}_\cdot) \geq -\frac{\gamma^2}{2} M_\sigma^2 C_1^2 \delta_t,$$

and using (A8.2), this yields

$$\int_0^T |\dot{\xi}^* - f(\xi^*(t))|^2 \, dt \leq m_\sigma^2 M_\sigma^2 C_1^2 \delta_t. \tag{8.23}$$

Now, noting (8.20), one has

$$\begin{aligned}
\frac{d}{dt}|\xi^*|^2 &= 2(\xi^*)^T (f(\xi^*) + (\dot{\xi}^* - f(\xi^*))) \\
&\leq 2K_f |\xi^*|^2 + 2K_f |\xi^*| + 2|\xi^*||\dot{\xi}^* - f(\xi^*)| \\
&\leq 3K_f |\xi^*|^2 + K_f + 2|\xi^*||\dot{\xi}^* - f(\xi^*)| \\
&\leq (3K_f + 1)|\xi^*|^2 + K_f + |\dot{\xi}^* - f(\xi^*)|^2.
\end{aligned}$$

Integrating, and using (8.23), this yields

$$|\xi_t^*|^2 \leq R_0^2 + (K_f + m_\sigma^2 M_\sigma^2 C_1^2)\delta_t + \int_0^t (3K_f + 1)|\xi_r^*|^2 \, dr.$$

Then, employing Gronwall's inequality, one finds that there exist $R_1, \overline{M} < \infty$ (depending on R_0, R, δ_t) such that $|\xi_t^*| \leq R_1$ for all $t \in [0, \delta_t]$ and $\|\dot{\xi}^*_\cdot\|_{L_2(0,\delta_t)} \leq \overline{M}$.

Next, consider any direction $v \in \mathbf{R}^n$ ($|v| = 1$) and scalar $h \in [-1, 1]$. Let

$$\xi_t^h \doteq \xi_t^* + \frac{t}{T} hv.$$

Then $\xi^h_{t_i} = x_0$ and $\xi^h_T = x + hv$. Moreover,

$$\frac{d^2}{dh^2} I(T, x_0, x + hv; \xi^h_\cdot) = v^T \Lambda^h v,$$

where

$$\Lambda^h \doteq \frac{1}{T^2} \int_0^T \left[t^2 L_{1,1}(\xi^h, \dot{\xi}^h) + 2t L_{1,2}(\xi^h, \dot{\xi}^h) + L_{2,2}(\xi^h, \dot{\xi}^h) \right] dt,$$

where $L_{i,j}$ represents the $n \times n$ matrix second partial of L with respect to its i^{th} and j^{th} arguments. Because $|\xi_t^h|$ and $\|\dot{\xi}_\cdot^h\|_{L_2(0,T)}$ are uniformly bounded,

$$\frac{d^2}{dh^2} I(T, x_0, x + hv; \xi_\cdot^h) \geq -c$$

for some constant c depending on δ_t, R, R_0. Then $I(T, x_0, x+hv; \xi_\cdot^h) + (c/2)h^2$ is a convex function of h for $|h| \leq 1$. Consequently,

$$\begin{aligned}
&\mathcal{V}^{fs}(T, x_0, x + hv) + \mathcal{V}^{fs}(T, x_0, x - hv) - 2\mathcal{V}^{fs}(T, x_0, x) \\
&\quad \geq I(T, x_0, x + hv; \xi_\cdot^h) + I(T, x_0, x - hv; \xi_\cdot^{-h}) - 2I(T, x_0, x; \xi_\cdot^*) \\
&\quad \geq -(c/2)h^2.
\end{aligned}$$

This implies that $\mathcal{V}^{fs}(T, x_0, x) + (c/2)|x|^2$ (as a function of x) is convex on the closed ball $\overline{B}_{(R+K_f+K_f R_0)T}(\widetilde{x})$. □

Lemma 8.5. *Let $\varepsilon \in (0, 1]$, and $\delta_t, R_1 \in (0, \infty)$. There exists $C_2 = C_2(\delta_t, R_1)$ such that for all $x \in \overline{B}_{R_1}(0)$ and all $T \in [\delta_t/2, \delta_t]$,*

$$\begin{aligned}
S_T[\phi](x) &= \sup_{x_0 \in \{|x_0 - [x + f(x)T]| \leq C_2 T\}} \left\{ \phi(x_0) + \mathcal{V}^{fs}(T, x_0, x) \right\} \\
&= \sup_{x_0 \in G} \left\{ \phi(x_0) + \mathcal{V}^{fs}(T, x_0, x) \right\}
\end{aligned}$$

for any $G \subseteq \mathbf{R}^n$ such that $G \supseteq \overline{B}_{C_2 T}(x - f(x)T)$.

Proof. By (8.15) and Lemma A.10

$$\sup_{x_0 \in \mathbf{R}^n} \{\phi(x_0) + \mathcal{V}^{fs}(T, x_0, x)\} = \sup_{\|w\|^2 \leq M_{\delta_t}^w (1 + R_1^2)} \left\{ \phi(\xi_0) - \frac{\gamma^2}{2} \int_0^T |w_t|^2 \, dt \right\} \tag{8.24}$$

where ξ_0 is given by (8.1) with terminal condition $\xi_T = x$, driven by $w_\cdot$.

Let $\xi_\cdot^\varepsilon$ satisfy (8.1) with $\xi_T^\varepsilon = x$ and $\|w^\varepsilon\|^2 \leq M_{\delta_t}^w(1 + R_1^2)$. Let $\widetilde{\xi}_t \doteq x + (t - T)f(x)$ for $t \in [0, T]$. Then

$$\frac{d}{dt}|\xi_t^\varepsilon - \widetilde{\xi}_t|^2 = 2(\xi_t^\varepsilon - \widetilde{\xi}_t)^T [f(\xi_t^\varepsilon) - f(x) + \sigma(\xi_t^\varepsilon) w_t^\varepsilon],$$

which using Assumption (A8.1)

$$\leq 2K_f |\xi_t^\varepsilon - \widetilde{\xi}_t|^2 + 2|\xi_t^\varepsilon - \widetilde{\xi}_t| \, |\sigma(\xi_t^\varepsilon) w_t^\varepsilon|,$$

and then using Assumption (A8.2)

$$\leq (2K_f + 1)|\xi_t^\varepsilon - \widetilde{\xi}_t|^2 + m_\sigma^2 |w_t^\varepsilon|^2.$$

Integrating, one finds

$$\begin{aligned}|\xi_t^\varepsilon - \widetilde{\xi}_t|^2 &\le m_\sigma^2 \|w^\varepsilon\|_{L_2(t,T)}^2 + (2K_f+1)\int_t^T |\xi_r^\varepsilon - \widetilde{\xi}_r|^2\,dr \\ &\le m_\sigma^2 M_{\delta_t}^w (1+R_1^2) + (2K_f+1)\int_t^T |\xi_r^\varepsilon - \widetilde{\xi}_r|^2\,dr.\end{aligned}$$

Using Gronwall's inequality, one can show this implies, for $T \in (\delta_t/2, \delta_t)$, there exists $C_2 = C_2(\delta_t, R_1)$ such that

$$|\xi_0^\varepsilon - \widetilde{\xi}_0| \le C_2 T$$

where we skip the technical details. Therefore, if w^ε is ε-optimal with $\varepsilon \in (0,1]$, one has $|\xi_0^\varepsilon - [x - f(x)T]| \le C_2 T$. Therefore, by (8.24) and (8.15), one has the result. □

Lemma 8.6. *Let $\delta_t, R_1, d \in (0,\infty)$. Let $|x| \le R_1$ and $T \in (0, \delta_t]$. There exists $R \in (0,\infty)$ independent of x, T such that*

$$\overline{B}_{dT}(x) \subseteq \bigcap_{z \in \overline{B}_{dT}(x)} \; \bigcap_{x_0 \in \overline{B}_{C_2T}(z+f(z)T)} \overline{B}_{RT}(\widetilde{x}(x_0))$$

where $\widetilde{x}(x_0) \doteq x_0 + f(x_0)T$. Further, there exists $R_0 \in (0,\infty)$ such that

$$\bigcup_{z \in \overline{B}_{dT}(x)} \overline{B}_{C_2T}(z - f(z)T) \subseteq \overline{B}_{R_0}(0).$$

Proof. The second assertion is obvious. The first assertion will be proved if one finds $R < \infty$ such that if $|z - x| \le dT$, $|x_0 - [z - f(z)T]| \le C_2 T$ and $|y - x| \le dT$, then $|y - [x_0 + f(x_0)T]| \le RT$ (i.e., $y \in \overline{B}_{RT}(\widetilde{x}(x_0))$. To prove this, we note

$$\begin{aligned}|y - [x_0 + f(x_0)]| &= |y - x + x - z + z - [x_0 + f(x_0)T]| \\ &\le 2dT + |z - f(z)T - x_0| + |f(z)T - f(x_0)T]| \\ &\le 2dT + C_2T + K_f|z - x_0|T \\ &\le (2d + C_2)T + K_f|z - f(z)T - x_0|T + K_f|f(z)T|T \\ &\le (2d + C_2)T + K_f C_2 T^2 + K_f^2(1 + R_1 + d)T^2 \\ &\le [2d + C_2 + K_f C_2 \delta_t + K_f^2(1 + R_1 + d)\delta_t]T.\end{aligned}$$

The result follows if one takes $R \doteq 2d + C_2 + K_f C_2 \delta_t + K_f^2(1 + R_1 + d)\delta_t$. □

Theorem 8.7. *Let $R_1, d, \delta_t \in (0,\infty)$, and let $T \in (\delta_t/2, \delta_t)$. There exists $c = c(R_1, d, \delta_t)$ such that $S_T[\phi](x) + (c/2)|x|^2$ is convex or $\overline{B}_{R_1+dT}(0)$.*

Proof. Let R, R_0 be as given by Lemma 8.6. Let $x \in \overline{B}_{R_1}(0)$. Let

$$G \doteq \bigcup_{\hat{z} \in \overline{B}_{dT}(x)} \overline{B}_{C_2T}(\hat{z} - f(\hat{z})T).$$

Let $z \in \overline{B}_{dT}(x)$. By Lemma 8.5,

$$S_T[\phi](z) + (c/2)|z|^2 = \sup_{x_0 \in G} \left[\phi(x_0) + \mathcal{V}^{fs}(T, x_0, z) + (c/2)|z|^2\right]. \qquad (8.25)$$

By Lemma 8.6 and $z \in \overline{B}_{dT}(x)$, for any $x_0 \in G$, one has $z \in \overline{B}_{RT}(\widetilde{x}(x_0)) \cap \overline{B}_{R_0}(0)$. Consequently, by Lemma 8.4, $\mathcal{V}^{fs}(T, x_0, z) + (c/2)|z|^2$ is convex over $z \in \overline{B}_{dT}(x)$. Therefore, the right-hand side of (8.25) is a supremum of convex functions over $\overline{B}_{dT}(x)$. Therefore, $S_T[\phi](z) + (c/2)|z|^2$ is convex on $\overline{B}_{dT}(x)$ for any $x \in \overline{B}_{R_1}(0)$. Further, because R_0 depends only on R_1, d, δ_t and c depends only on R_0, R_1, δ_t, we see that c is independent of $x \in \overline{B}_{R_1}(0)$. Consequently, $S_T[\phi](z) + (c/2)|z|^2$ is convex on $\overline{B}_{R_1+dT}$. □

Theorem 8.7 implies that the information state (in the absence of measurement updates) will be semiconvex on $[t_i + \delta_t/2, t_{i+1}]$ for all $i \in \{0, 1, 2, \ldots\}$. To complete this discussion, one also needs to show that the measurement updates (8.13) maintain semiconvexity.

Lemma 8.8. *Suppose ψ is semiconvex. Then $\psi(x) - \frac{1}{2}|\rho^{-1}(x)[y - h(x)]|^2$ is semiconvex for any $y \in \mathbf{R}^k$. In particular, if $\psi(x) + (c/2)|x|^2$ is convex over $\overline{B}_{R_1}(0)$ and $y \in \overline{B}_{R_y}(0)$, then there exists $\widehat{C} = \widehat{C}(c, R_1, R_y)$ such that $\psi(x) - \frac{1}{2}|\rho^{-1}(x)[y - h(x)]|^2 + (\widehat{C}/2)|x|^2$ is convex over $\overline{B}_{R_1}(0)$.*

Proof. Let $G(x) \doteq |\rho^{-1}(x)[y - h(x)]|^2$. Fix $R < \infty$. Then, because ψ is semiconvex, there exists $C_R < \infty$ such that $\psi(x) + \frac{C_R}{2}|x|^2$ is convex. But by (A8.1) and (A8.3), there exists $C_{R,y} < \infty$ such that $|G_{xx}(x)| \leq C_{R,y}$ for all $x \in \overline{B}_R$. Consequently $\psi(x) - \frac{1}{2}|\rho^{-1}(x)[y - h(x)]|^2 + \frac{C_R + C_{R,y}}{2}|x|^2$ is convex. □

By Theorem 8.1, Theorem 8.7 and Lemma 8.8, we see that the information state, $P(t, \cdot)$, will be semiconvex for all $t \in [t_i + \delta_t/2, t_{i+1}]$ for all $i \in \{0, 1, 2, \ldots\}$.

It is interesting to note that there exists a semiconvexity constant over $\overline{B}_{R_1}(0)$ which is uniform over $\bigcup_i (t_i + \delta/2, t_{i+1})$ such that the semiconvexity constant depends on R_y where $\{y_i\} \subset \overline{B}_{R_y}(0)$, but independent of the y_i.

8.2 Max-Plus Propagation

One would like to calculate $P(t, x)$ in a region of interest around $\operatorname{argmax}_x \left[|x - \hat{e}_T|^2 + P(t, x)\right]$. Suppose that the region of interest remains in some $\overline{B}_{R_1}(0)$ for the period of interest say $t \in [0, \bar{T}] = [0, \bar{N}\delta_t]$. By the above results, we see that there exists $\bar{c} < \infty$ such that $P(t, \cdot)$ is semiconvex on $\overline{B}_{R_1+1}(0)$ with constant $\bar{c}$ for all $t \in \bigcup_{i<\bar{N}} [t_i + \delta_t/2, t_{i+1}]$, and that $P^-(t_i, \cdot)$ is also semiconvex with the constant for $i \leq \bar{N}$. Suppose ϕ is also semiconvex with constant $\bar{c}$ over $\overline{B}_{R_1+1}(0)$.

Recall that semiconvexity of $P(t, \cdot)$ (or $P^-(t_i, \cdot)$) implies that given any ball $\overline{B}_R(0)$, $P(t, \cdot)$ (or $P^-(t_i, \cdot)$) is Lipschitz with some constant, L, over $\overline{B}_R(0)$

(c.f. [42]). Recall also, from Chapter 2, that $\mathcal{S}^{\bar{C}\bar{L}}_{\bar{R}_1}$ denotes the space of semiconvex functions with (semiconvexity) constant $\bar{c}$ and Lipschitz constant $\bar{L}$ over $\overline{B}_{R_1}(0)$. We view the restrictions of $P(t,\cdot)$ (over appropriate time intervals as given above) and $P^-(t_i,\cdot)$ to $\overline{B}_{R_1}(0)$ as elements of $\mathcal{S}^{\bar{C}\bar{L}}_{\bar{R}_1}$.

Let $\widetilde{C}$ be an $n \times n$ symmetric matrix such that $\widetilde{C} - \bar{c}I > 0$. Recall (see Chapters 2 and 4) that the set of functions

$$\psi_i(x) \doteq -\frac{1}{2}(x - x_i)^T \widetilde{C}(x - x_i),$$

where the x_i form a countable dense set over $\mathcal{E} = \{\overline{x} \in \mathbf{R}^n : \overline{x}^T(\widetilde{C})^2)\overline{x} \leq (\bar{L} + |\widetilde{C}|R_1)\}$, comprise a countable basis for max-plus vector space $\mathcal{S}^{\bar{C}\bar{L}}_{\bar{R}_1}$. In particular, for any $P(t,\cdot) \in \mathcal{S}^{\bar{C}\bar{L}}_{\bar{R}_1}$,

$$P(t,x) = \sup_{i \in \mathcal{N}} [a_i + \psi_i(x)] = \bigoplus_{i=1}^{\infty} [a_i \otimes \psi_i(x)], \tag{8.26}$$

where

$$a_i = -\max_{x \in \overline{B}_{R_1}} [-P(t,x) + \psi_i(x)] \quad \forall\, i. \tag{8.27}$$

As in Chapter 4, we make the unwarranted assumption that $P(t_k,\cdot)$ and $P^-(t_k,\cdot)$ have *finite* max-plus basis expansions which we write as

$$P(t_k,x) = \bigoplus_{i=1}^{N} a_i^k \otimes \psi(x) \quad \forall\, x \in \overline{B}_{R_1},$$
$$P^-(t_k,x) = \bigoplus_{i=1}^{N} a_i^{-,k} \otimes \psi(x) \quad \forall\, x \in \overline{B}_{R_1}.$$

Let a^k and $a^{-,k}$ be the vectors of length N with elements a_i^k and $a_i^{-,k}$, respectively. Also, let $\boldsymbol{\psi}(x)$ denote the vector function of length N with elements $\psi_i(x)$. We will not perform a convergence analysis with regard to the errors introduced by such a truncation of expansions; an error analysis for an infinite time-horizon control problem class appears in Chapter 5.

Recall that

$$P^-(t_{k+1},x) = S_{\delta_t}[P(t_k,\cdot)](x),$$

where S_{δ_t} is max-plus linear. As in Chapter 4, we also suppose that the $S_{\delta_t}[\psi_j]$ have finite max-plus basis expansions in terms of the same $\{\psi_i\}_{i=1}^N$, and define $N \times N$ matrix B componentwise by

$$B_{i,j} = -\max_{x \in \overline{B}_{R_1}} (\psi_i(x) - S_{\delta_t}[\psi_j](x)),$$

in which case

$$S_\tau[\psi_j](x) = \bigoplus_{i=1}^{N} B_{i,j} \otimes \psi_i(x) \quad \forall x \in \overline{B}_{R_1}$$

for all $j \leq N$. Then, with $\odot$ representing the max-plus dot product here,

$$\bigoplus_{i=1}^{N} a_i^{-,k+1} \otimes \psi_i(x) = \left[a^{-,k+1} \odot \boldsymbol{\psi}\right](x) = P^-(t_{k+1}, x) = S_{\delta_t}[P(t_k, \cdot)](x)$$

$$= S_{\delta_t}[a^k \odot \boldsymbol{\psi}(\cdot)](x)$$

which by the max-plus linearity of S_{δ_t}

$$= \bigoplus_{j=1}^{N} a_j^k \otimes S_{\delta_t}[\psi_j](x)$$

$$= \bigoplus_{j=1}^{N} a_j^k \otimes \left[\bigoplus_{i=1}^{N} B_{i,j} \otimes \psi_i(x)\right]$$

$$= \bigoplus_{i=1}^{N} \left[\bigoplus_{j=1}^{N} B_{i,j} \otimes a_j^k\right] \otimes \psi_i(x).$$

This holds if we take

$$a_i^{-,k+1} = \bigoplus_{j=1}^{N} B_{i,j} \otimes a_j^k$$

for all i, or in other words, if

$$a^{-,k+1} = B \otimes a^k.$$

This leads us to the following algorithm for the propagation of the robust/H_∞ filter information state.

1. Let $k = 0$ and compute $a_i^0 = -\max_{x \in \overline{B}_{R_1}} [\psi_i(x) - \phi(x)]$ for all $i \leq N$.
2. Propagate the dynamics forward in time by $a^{-,k+1} = B \otimes a^k$.
3. Transform back to state space by

$$P^-(t_{k+1}, x) = \bigoplus_{i=1}^{N} a_i^{-,k+1} \otimes \psi_i(x) = a^{-,k+1} \odot \boldsymbol{\psi}(x)$$

   to obtain $P^-(t_{k+1}, x)$ over $\overline{B}_{R_1}$.
4. Perform the measurement update via

$$P(t_{k+1}, x) = P^-(t_{k+1}, x) - \frac{1}{2}|\rho^{-1}(x)[y_{k+1} - h(x)]|^2$$

   for $x \in \overline{B}_{R_1}$.
5. Compute $a_i^{k+1} = -\max_{x \in \overline{B}_{R_1}} [-P(t_{k+1}, x) + \psi_i(x)]$ for all $i \leq N$.

6. Increment k and return to Step 2.

Remark 8.9. The above algorithm requires the precomputation of B, which then allows one to bypass the computationally difficult propagation of the HJB PDE forward in time. This leaves the transformations in Steps 3 and 5 as the most demanding portions of the computations.

Remark 8.10. Note that there exists a semiconvexity constant, $\bar{c}$, such that the $P^{-}(t_k, \cdot)$ and $P^{(}t_k, \cdot)$ are all semiconvex with this constant over $\overline{B}_{R_1}(0)$. However, an approximation generated by *forward propagation* of a truncated max-plus expansion, with say $\widetilde{C}$ (where $\widetilde{C} - \bar{c}I > 0$) as the quadratic growth constant in the basis functions, may have a higher semiconvexity constant than $\widetilde{C}$. Employing the same basis functions to approximate this propagated approximation, can induce additional errors in the propagation. One approach to dealing with this issue is to search for a basis expansion such that the basis functions do not become more concave with propagation over a time-step.

9

Mixed L_∞/L_2 Criteria

In previous chapters we considered problems with cost criteria in integral form. In this chapter we will introduce L_∞-type criteria as well. Recall that suprema over time correspond to max-plus integrals over time. This leads one to refer to such criteria as max-plus additive criteria, and to standard-sense integral criteria as max-plus multiplicative. In a sense, this is a very natural class of cost criteria for max-plus analysis.

With the exception of Chapter 8, the approach so far has been to use an infinite time-horizon problem as the basis for development of max-plus methods. We continue that here. With the max-plus multiplicative criteria, infinite time-horizon problems led to max-plus eigenvector problems, say $e = B \otimes e$. Here, the addition of L_∞ (max-plus additive) components to our criteria will lead to problems of the form $a = c \oplus B \otimes a$ where now the problem data yield both the matrix B and the vector c. One then solves the problem by obtaining a. This approach was first studied in McEneaney and Dower [76]. Interestingly, a generalization of the power method from Chapter 4 can be used here. Curse-of-dimensionality-free methods have not yet been explored in this class. We will not perform an error analysis, but refer to Chapter 5 as an example of how such an analysis might proceed.

9.1 Mixed L_∞/L_2 Problem Formulation

We return here to the familiar infinite time-horizon dynamic model that was considered in Chapters 4–7. More specifically, we consider dynamics

$$\dot{\xi} = f(\xi) + \sigma(\xi)u \tag{9.1}$$

with initial condition

$$\xi_0 = x \in \mathbf{R}^n. \tag{9.2}$$

We again assume (A4.1I) and (A4.2I), and also that $f, \sigma \in C^2$. However, the theory to follow will hold for a much wider class of systems; see [76].

As before, u_t will take values in $\mathbf{R}^l$. We will consider two classes of control spaces: $\mathcal{U}^M_{[s,T)}$ for $M \in (0,\infty)$ and $0 \le s \le T < \infty$, and $\mathcal{U}^M_{[s,T)}$ for $M = \infty$ and $0 \le s \le T < \infty$, where specifically, we let

$$\mathcal{U}^M_{[s,T)} \doteq \begin{cases} \{u:\,[s,T) \to \mathbf{R}^l \mid \|u\|_{L_\infty(s,T)} < \infty, \int_s^T \eta(|u_t|)\,dt < \infty\} & \text{if } M = \infty, \\ \{u:\,[s,T) \to \mathbf{R}^l \mid \|u\|_{L_\infty(s,T)} \le M, \int_s^T \eta(|u_t|)\,dt < \infty\} & \text{if } M < \infty, \end{cases} \tag{9.3}$$

where $\eta : [0,\infty) \to [0,\infty)$ is continuous and strictly increasing with $\eta(0) = 0$.

Let $\ell, L \in C^2(\mathbf{R}^n)$ and $\eta \in C^2((0,\infty)) \cap C([0,\infty))$. Fix any $M \in (0,\infty]$. One says that system (9.1), (9.2) satisfies the following mixed L_∞/L_2-gain property if there exists a locally bounded nonnegative function $\beta : \mathbf{R}^n \to [0,\infty)$ such that

$$\ell(\xi_T) + \int_0^T L(\xi_t) - \eta(|u_t|)\,dt \;\le\; \beta(x) \tag{9.4}$$

for all $x \in \mathbf{R}^n$, $u \in \mathcal{U}^M_{[0,T)}$ and $T \in [0,\infty)$, where we assume throughout that we use $M = \infty$ if and only if $\ell \equiv 0$. This class of problems is closely related to the stability notions (such as ISS — input-to-state stability) of Sontag et al. (c.f. [105]). More complete discussions of this class of problems and their relation to stability issues can be found in [55], [56], [62].

We assume there exist $C_\ell, C_L, k_1, k_2 \in (0,\infty)$ such that

$$\begin{aligned} &L(x), \ell(x) \ge 0 \quad \forall x \in \mathbf{R}^n, \\ &\ell_{xx}(x) + C_\ell \ge 0 \quad \forall x \in \mathbf{R}^n, \\ &L_{xx}(x) + C_L \ge 0 \quad \forall x \in \mathbf{R}^n, \\ &\eta(|v_1+v_2|) - \eta(|v_1|) \le k_1|v_1||v_2| + k_2|v_2|^2 \quad \forall v_1, v_2 \in \mathbf{R}^l. \end{aligned} \tag{A9.1I}$$

where we note that the last assumption (on η) holds if the second derivative of η is bounded. We will now make two rather broad assumptions. These are not specific in terms of say a Lipschitz bound on some function, but are instead given in terms of the general behavior of the system. These assumptions are left broad so as to cover the wide variety of problems being considered in this chapter. For a specific problem form, one would need to verify whether these following conditions are met. We assume:

There exists $c_1 < \infty$ and $\beta : \mathbf{R}^n \to \mathbf{R}$ which is bounded on compact sets, such that

$$|\xi_T|^2 \le \beta(x) + c_1 T + \int_0^T \eta(|u_t|)\,dt \tag{A9.2I}$$

for all $x \in \mathbf{R}^n$, $u \in \mathcal{U}^M_{[0,T)}$ and $T \in (0,\infty)$.

There exists $c_2 < \infty$, $\alpha \in (0,1)$ and $\hat{\beta} : \mathbf{R}^n \to \mathbf{R}$ which is bounded on compact sets, such that

$$\ell(\xi_T) + \int_0^T L(\xi_t)\,dt \le \hat{\beta}(x) + c_2 T + \alpha \int_0^T \eta(|u_t|)\,dt \tag{A9.3I}$$

for all $x \in \mathbf{R}^n$, $u \in \mathcal{U}^M_{[0,T)}$ and $T \in (0,\infty)$.

For purposes of concise presentation, define

$$I(s,t,x,u) \doteq \int_s^t L(\xi_r) - \eta(|u_r|)\,dr, \tag{9.5}$$

$$J(s,t,x,u) \doteq \ell(\xi_t) + I(s,t,x,u), \tag{9.6}$$

where $\xi_\cdot$ satisfies (9.1) with $\xi_s = x$. We will be primarily interested in value function

$$W(x) \doteq \sup_{T<\infty} \sup_{u \in \mathcal{U}^M_{[0,T)}} J(0,T,x,u). \tag{9.7}$$

A closely related value function is given by

$$V(x) \doteq \limsup_{T\to\infty} \sup_{u \in \mathcal{U}^M_{[0,T)}} J(0,T,x,u). \tag{9.8}$$

These will be referred to as the *stopping-time problem* value function and the *infinite time-horizon problem* value function, respectively. We will see that the solution of the infinite time-horizon problem plays a role in the solution of the stopping-time problem. The max-plus analysis of this problem first appeared in McEneaney and Dower [76], and we follow that here. Note that the presence of the $\ell(\xi_T)$ term implies that $\sup_{u \in \mathcal{U}^M_{[0,T)}} J(0,T,x,u)$ may not be monotonically increasing as a function of T, and so W may not be equal to V. This is distinct from the infinite time-horizon problem formulations in Chapters 4–7. However, with $\ell \equiv 0$, W is identical to the value function considered there if one takes $\eta(s) = \frac{\gamma^2}{2}s^2$. Further, in that case ($\ell \equiv 0$, $\eta(s) = \frac{\gamma^2}{2}s^2$) one sees that, by considering $u \equiv 0$,

$$\sup_{u \in \mathcal{U}^M_{[0,T)}} J(0,T,x,u) = \sup_{u \in \mathcal{U}^M_{[0,T)}} \left\{ \int_0^T L(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt \right\}$$

is monotonically increasing in T. Consequently in that case, $W = V$. Also note that in that case

$$W(x) = \sup_{u \in \mathcal{U}^{M,loc}} \sup_{T<\infty} J(0,T,x,u),$$

where

$$\mathcal{U}^{M,\mathrm{loc}} \doteq \left\{ u : [0,\infty) \to \mathbf{R}^l \mid u_{[0,T)} \in \mathcal{U}^M_{[0,T)}\ \forall T \in (0,\infty) \right\}$$

and $u_{[s,T)}$ denotes the restriction of u to domain $[s,T)$. We note that $\sup_{T<\infty} J(0,T,x,u)$ contains an L_∞-type component. In particular, if $L \equiv 0$, one has

$$\sup_{T<\infty} J(0,T,x,u) = \sup_{T<\infty} \left\{ \ell(\xi_T) - \int_0^T \eta(|u_t|)\,dt \right\}.$$

One can view this as a max-plus integral. Further, one can view the $\sup_{u\in\mathcal{U}^{M,\mathrm{loc}}}$ as an expectation with respect to a max-plus probability measure specified by η (c.f. [1], [5], [28], [39], [40], [97], [98], [99]).

In keeping with the broad discussion of multiple problem forms in this chapter, we assume directly that

$$V(x), W(x) \in \mathbf{R} \text{ for all } x \in \mathbf{R}^n \text{ (i.e., } V, W \text{ are finite).} \tag{A9.4I}$$

9.2 Dynamic Programming

We again begin with DPPs for these value functions. The DPP for the infinite time-horizon problem will yield, as usual, a max-plus linear semigroup. However, the DPP for the stopping-time problem will yield a max-plus linear sub-semigroup.

9.2.1 Dynamic Programming Principles

It is not difficult to show that the stopping-time value W satisfies the following DPP, and we will do so.

Theorem 9.1. *Let $\tau \in (0,\infty)$. W given by (9.7) satisfies for all $x \in \mathbf{R}^n$,*

$$W(x) = \max\Big\{ \sup_{T\in[0,\tau)} \sup_{u\in\mathcal{U}^M_{[0,T)}} J(0,T,x,u),\ \sup_{u\in\mathcal{U}^M_{[0,\tau)}} I(0,\tau,x,u) + W(\xi_\tau) \Big\}. \tag{9.9}$$

Proof. Fix $\tau > 0$. The supremum over $T \in [0,\infty)$ in (9.7) is equivalent to maximum of the suprema over $[0,\tau)$ and $[\tau,\infty)$. That is,

$$W(x) = \max\Bigg(\sup_{T\in[0,\tau)} \sup_{u\in\mathcal{U}^M_{[0,T)}} J(0,T,x,u),\ \underbrace{\sup_{T\in[\tau,\infty)} \sup_{u\in\mathcal{U}^M_{[0,T)}} \{\ell(\xi_T) + I(0,T,x,u) : \xi_0 = x\}}_{=:\zeta_\tau(x)} \Bigg)$$

where ξ satisfies (9.1) with $\xi_0 = x$ (as indicated) in the second term on the right. Considering the $\zeta_\tau(x)$ term only (i.e., the second term in the maximum), one finds

$$\zeta_\tau(x) = \sup_{T\in[\tau,\infty)} \sup_{u\in\mathcal{U}^M_{[0,T)}} \left\{ I(0,\tau,x,u) + \ell(\xi_T) + I(\tau,T,\xi_\tau,u) \,:\, \xi_0 = x \right\}$$

$$= \sup_{T\in[\tau,\infty)} \sup_{u^1\in\mathcal{U}^M_{[0,\tau)}} \sup_{u^2\in\mathcal{U}^M_{[\tau,T)}} \Big\{ I(0,\tau,x,u^1) + \ell(\xi^2_T) + I(\tau,T,\xi^1_\tau,u^2) \,:\, \xi^1_0 = x,\, \xi^2_\tau = \xi^1_\tau \Big\}$$

where ξ^1 (over $[0,\tau)$) satisfies (9.1),(9.2) with input u^1, and ξ^2 (over $[\tau,T]$) satisfies (9.1) with input u^2 and initial condition $\xi^2_\tau = \xi^1_\tau$. This is

$$= \sup_{u^1\in\mathcal{U}^M_{[0,\tau)}} \Big\{ I(0,\tau,x,u^1) + \sup_{T\in[\tau,\infty)} \sup_{u^2\in\mathcal{U}^M_{[\tau,T)}} \left[\ell(\xi^2_T) + I(\tau,T,\xi^1_\tau,u^2) \right] \,:\, \xi^1_0 = x,\, \xi^2_\tau = \xi^1_\tau \Big\}$$

$$= \sup_{u^1\in\mathcal{U}^M_{[0,\tau)}} \left\{ I(0,\tau,x,u^1) + W(\xi^1_\tau) \,:\, \xi^1_0 = x \right\}. \quad \square$$

In order to prove a dynamic programming principle for the infinite time-horizon value V, it is useful to consider the following auxiliary finite horizon optimization problem:

$$\widetilde{W}(x,T) = \sup_{u\in\mathcal{U}^M_{[0,T)}} J(0,T,x,u) \tag{9.10}$$

It follows immediately from (9.7), (9.8) and (9.10) that

$$W(x) = \sup_{T<\infty} \widetilde{W}(x,T), \tag{9.11}$$

$$V(x) = \limsup_{T\to\infty} \widetilde{W}(x,T). \tag{9.12}$$

Identity (9.12) will be used to prove the DPP for V. The following lemma is essentially identical to a DPP from Remark 3.9, but with a change of variable in the time component.

Lemma 9.2. *$\widetilde{W}$ given by (9.10) satisfies*

$$\widetilde{W}(x,T) = \sup_{u\in\mathcal{U}^M_{[0,\tau)}} \left\{ I(0,\tau,x,u) + \widetilde{W}(\xi_\tau, T-\tau) \,:\, \xi_0 = x \right\} \tag{9.13}$$

for all $x \in \mathbf{R}^n$, $0 \le \tau \le T < \infty$.

Note that DPP (9.9) may also be proved using (9.13) by first rewriting (9.13) so that it holds for all $\tau \in [0,\infty)$; see [76].

We now continue by using Lemma 9.2 to prove the DPP for the infinite horizon value function V. First we need the following technical lemma.

Lemma 9.3. *Fix $\rho \in [0,\infty)$, $\overline{\varepsilon} \in [0,\infty)$, and $0 \le \tau \le T < \infty$. Let $\varepsilon \in (0,\overline{\varepsilon}]$. There exists $R = R(\rho,\tau,T,\overline{\varepsilon}) < \infty$ such that the following holds: Given any $x \in \overline{B}_\rho$ and ε-optimal $u^{x,T,\varepsilon}$ for $\widetilde{W}(x,T)$ (i.e., such that $J(0,T,x,u^{x,T,\varepsilon}) \ge \widetilde{W}(x,T) - \varepsilon$), one has $\xi^{x,T,\varepsilon}_\tau \in \overline{B}_R$ where $\xi^{x,T,\varepsilon}$ satisfies (9.1) with $\xi^{x,T,\varepsilon}_0 = x$ and input $u^{x,T,\varepsilon}$.*

Proof. In the case $M < \infty$, the proof is immediate. We turn to the case $M = \infty$. Fix $T \in (0,\infty)$, $\rho \in (0,\infty)$ and $x \in \overline{B}_\rho$. let $\varepsilon \in (0,1]$. Let u^ε be ε-optimal for $\widetilde{W}(x,T)$. By the non-negativity of L, ℓ and the fact that $\eta(0) = 0$,

$$-\varepsilon \le \ell(\xi_T^\varepsilon) + \int_0^T L(\xi_t^\varepsilon) - \eta(|u_t^\varepsilon|)\,dt,$$

where ξ^ε is driven by u^ε with $\xi_0^\varepsilon = x$. By Assumption (A9.3I), this is

$$\le (\alpha - 1)\int_0^T \eta(|u_t^\varepsilon|)\,dt + c_2 T + \hat{\beta}(x).$$

Rearranging this, one sees that

$$\int_0^T \eta(|u_t^\varepsilon|)\,dt \le \frac{1}{1-\alpha}[1 + c_2 T + \hat{\beta}(x)] \le C_1^{\rho,T}$$

for proper choice of $C_1^{\rho,T}$. Then, using Assumption (A9.2I) and the non-negativity of η, we see that this implies that for any $\tau \in [0,T]$,

$$|\xi_\tau^\varepsilon|^2 \le \beta(x) + c_1\tau + \int_0^\tau \eta(|u_t^\varepsilon|)\,dt \le \beta(x) + c_1 T + C_1^{\rho,T}.$$

Because β is bounded on compact sets, one has the desired bound. □

Definition 9.4. *The limit supremum in (9.12) (equivalently, (9.8)) is* attained uniformly on compact sets from above *if the following condition holds: Given any compact set $\mathcal{X} \subset \mathbf{R}^n$ and $\varepsilon > 0$, there exists $\overline{T}_{\varepsilon,\mathcal{X}}$ such that*

$$\widetilde{W}(x,T) \le V(x) + \varepsilon \quad \forall T \ge \overline{T}_{\varepsilon,\mathcal{X}},\ \forall x \in \mathcal{X}. \tag{9.14}$$

Definition 9.5. *The supremum in (9.11) (equivalently, (9.7)) is* attained uniformly on compact sets from below *if the following condition holds: Given any compact set $\mathcal{X} \subset \mathbf{R}^n$ and $\varepsilon > 0$, there exists $\underline{T}_{\varepsilon,\mathcal{X}}$ such that given any $x \in \mathcal{X}$, there exists $T_x \in [0, \underline{T}_{\varepsilon,\mathcal{X}}]$ such that*

$$\widetilde{W}(x,T_x) \ge W(x) - \varepsilon. \tag{9.15}$$

Further, the supremum in (9.11) (equivalently, (9.7)) is attained uniformly on compact sets from below with a lower time *if the following condition holds: Given any compact set $\mathcal{X} \subset \mathbf{R}^n$, $T_0 \ge 0$ and $\varepsilon > 0$, there exists $\underline{T}_{\varepsilon,\mathcal{X},T_0}$ such that given any $x \in \mathcal{X}$, there exists $T_x \in [T_0, \underline{T}_{\varepsilon,\mathcal{X},T_0}]$ such that inequality (9.15) holds.*

Note that the second item in Definition 9.5 poses a stronger requirement than the first. For the remainder of this chapter, we assume that

the limit supremum in (9.12) is attained uniformly on compact sets from above, and that the supremum in (9.11) is attained uniformly on compact sets from below with a lower time. (A9.5I)

Lemma 9.6. *V given by (9.8) satisfies*

$$V(x) = \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \{I(0,\tau,x,u) + V(\xi_\tau) : \xi_0 = x\} \tag{9.16}$$

$\forall x \in \mathbf{R}^n$, $\tau \in [0,\infty)$.

Proof. Fix $x \in \mathbf{R}^n$, $\tau \in [0,\infty)$. Let the notation $a \wedge b$ denote $\min(a,b)$ for any $a, b \in \mathbf{R}$. Then, (9.13) implies that

$$\widetilde{W}(x,T) = \sup_{u \in \mathcal{U}^M_{[0,\tau\wedge T)}} \left\{ I(0,\tau \wedge T, x, u) + \widetilde{W}(\xi_{(\tau\wedge T)}, T - (\tau \wedge T)) \right\} \tag{9.17}$$

for all $\tau \in [0,\infty)$ where $\xi_0 = x$. By (9.12) and (9.17),

$$V(x) = \limsup_{T\to\infty} \sup_{u \in \mathcal{U}^M_{[0,\tau\wedge T)}} \left\{ I(0,\tau \wedge T, x, u) + \widetilde{W}(\xi_{(\tau\wedge T)}, T - (\tau \wedge T)) \right\},$$

where $\xi_0 = x$, and this is

$$\begin{aligned} &\geq \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \limsup_{T\to\infty} \left\{ I(0,\tau,x,u) + \widetilde{W}(\xi_\tau, T-\tau) \right\} \\ &= \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \left\{ I(0,\tau,x,u) + \limsup_{T\to\infty} \widetilde{W}(\xi_\tau, T-\tau) \right\} \\ &= \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \{ I(0,\tau,x,u) + V(\xi_\tau) \}, \end{aligned} \tag{9.18}$$

which proves the inequality in one direction.

We now prove the reverse inequality. Fix $\tau \in [0,\infty)$. Fix $\rho \in (0,\infty)$ and $x \in \overline{B}_\rho \subset \mathbf{R}^n$. From the definition of V, we know that there exists a strictly monotonically increasing sequence, $\{T_i\}$, such that

$$V(x) = \lim_{i\to\infty} \sup_{u \in \mathcal{U}^M_{[0,T_i)}} J(0,T_i,x,u).$$

Consequently, given $\varepsilon > 0$, there exists $i(x)$ with $T_{i(x)} \geq \tau$ such that

$$V(x) \leq \sup_{u \in \mathcal{U}^M_{[0,T_i)}} J(0,T_i,x,u) + \varepsilon \quad \forall\, i \geq i(x).$$

Therefore, for any $i \geq i(x)$, there exists $u^{\varepsilon,i} \in \mathcal{U}^M_{[0,T_i)}$ such that

$$\begin{aligned} V(x) &\leq J(0,T_i,x,u^{\varepsilon,i}) + 2\varepsilon \\ &= I(0,\tau,x,u^{\varepsilon,i}) + J(\tau,T_i,\xi^{\varepsilon,i}_\tau,u^{\varepsilon,i}) + 2\varepsilon \end{aligned}$$

where $\xi^{\varepsilon,i}$ is driven by $u^{\varepsilon,i}$ with initial condition $\xi^{\varepsilon,i}_0 = x$, and we note the slight abuse of notation incurred by the use of $u^{\varepsilon,i}$ where we should actually use the restriction of that process to the relevant time domain. By the definition of $\widetilde{W}$, this obviously implies

$$V(x) \le I(0,\tau,x,u^{\varepsilon,i}) + \widetilde{W}(\xi_\tau^{\varepsilon,i}, T_i - \tau) + 2\varepsilon. \tag{9.19}$$

Now by Lemma 9.3, there exists $R = R(\rho,\tau)$ (independent of $x, u^{\varepsilon,i}$) such that

$$\xi_\tau^{\varepsilon,i} \in \overline{B}_R. \tag{9.20}$$

Further, by the assumption that limit supremum in (9.8) is attained uniformly on compact sets from above, one then finds that there exists $\overline{T}_{\varepsilon,R}$ (independent of $\xi_\tau^{\varepsilon,i}$) such that

$$\widetilde{W}(x,T) \le V(x) + \varepsilon \quad \forall x \in \overline{B}_R,\ \forall T \ge \overline{T}_{\varepsilon,R}. \tag{9.21}$$

Combining (9.19), (9.20) and (9.21), one finds that for $i \ge i(x)$ sufficiently large such that $T_i - \tau \ge \overline{T}_{\varepsilon,R}$,

$$V(x) \le I(0,\tau,x,u^{\varepsilon,i}) + V(\xi_\tau^{\varepsilon,i}) + 3\varepsilon.$$

Consequently,

$$V(x) \le \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \{I(0,\tau,x,u) + V(\xi_\tau)\} + 3\varepsilon.$$

Because this is true for all $\varepsilon > 0$, one obtains the reverse inequality, and so the proof is complete. □

9.2.2 Dynamic Programming Equations

By considering the DPPs (9.9) and (9.16) in the limit as $\tau \downarrow 0$, one would typically find that W and V satisfy respectively a variational inequality (VI) and an HJB PDE. Because the focus here is on max-plus methods rather than PDE/VI–based methods, we only indicate the corresponding VI and PDE problems.

Define the Hamiltonian

$$H(x,p) \doteq L(x) + \sup_{v \in \mathbf{R}^l} \{f(x,v)\cdot p - \eta(|v|)\}.$$

One expects W to be a viscosity solution of the VI (c.f. [8])

$$-\max\left(\ell(x) - W(x), H(x, \nabla_x W(x))\right) = 0. \tag{9.22}$$

One expects V to be a viscosity solution of the HJB PDE (see Chapter 3)

$$-H(x, \nabla_x V(x)) = 0. \tag{9.23}$$

9.3 Max-Plus Representations and Semiconvexity

Define the time-indexed operator

$$S_\tau[\phi](x) = \max\Big\{ \sup_{t\in[0,\tau)} \sup_{u\in\mathcal{U}^M_{[0,\tau)}} J(0,t,x,u), \sup_{u\in\mathcal{U}^M_{[0,\tau)}} \big[I(0,\tau,x,u) + \phi(\xi_\tau)\big] \Big\}$$

and the semigroup

$$\mathcal{L}_\tau[\phi] = \sup_{u\in\mathcal{U}^M_{[0,\tau)}} \big[I(0,\tau,x,u) + \phi(\xi_\tau)\big]$$

where the domains are implicit. Then DPP (9.9) can be rewritten $\forall\, x \in \mathbf{R}^n$ as

$$W(x) = S_\tau[W](x) = \max\Big\{ \sup_{t\in[0,\tau)} \sup_{u\in\mathcal{U}^M_{[0,\tau)}} J(0,t,x,u), \mathcal{L}_\tau[W](x) \Big\}$$

and the DPP for V, (9.16) can be rewritten $\forall\, x \in \mathbf{R}^n$ as

$$V(x) = \mathcal{L}_\tau[V](x)$$

One can easily show that $\mathcal{L}_\tau$ is a max-plus linear operator. Define $c_\tau : \mathbf{R}^n \to \mathbf{R}$ by

$$c_\tau(x) = \sup_{t\in[0,\tau)} \sup_{u\in\mathcal{U}^M_{[0,\tau)}} J(0,t,x,u) \quad \forall\, x \in \mathbf{R}^n. \tag{9.24}$$

Then, for all ϕ in the domain of $\mathcal{L}_\tau$,

$$S_\tau[\phi](x) = \max\big\{c_\tau(x), \mathcal{L}_\tau[\phi](x)\big\} = \{c_\tau \oplus \mathcal{L}_\tau[\phi]\}(x) \quad \forall\, x \in \mathbf{R}^n, \tag{9.25}$$

and consequently, S_τ is a max-plus affine operator. Note that our DPP for W, (9.9), now takes the form

$$W = c_\tau \oplus \mathcal{L}_\tau[W]. \tag{9.26}$$

Similarly, one easily has (as in Chapter 4) that our DPP for V, (9.16), takes the max-plus eigenvector form

$$0 \otimes V = \mathcal{L}_\tau[V]. \tag{9.27}$$

We will focus in this chapter on methods similar to those used in Chapters 4–5; methods similar to those of Chapters 6 and 7 have not yet been developed for this class of problems. Consequently, we will again be interested in spaces of the form $\mathcal{S}^{cL}_{\bar{R}}$ where c is the semiconvexity constant over $\overline{B}_R$ and L is the Lipschitz constant over $\overline{B}_R$. With this in mind, we see that it will be useful to prove that the value functions, V and W are semiconvex. We again note that this chapter is intended to be rather general in terms of the scope of problems. In particular, we are allowing the cost criterion and control spaces to take a variety of forms. Consequently, and in contradistinction to say Chapter 4, a single set of assumptions on f, σ, L, ℓ, η under which one could guarantee semiconvexity of W, V for all cases would be far too restrictive. Instead, we demonstrate semiconvexity for two specific subclasses of problems, and then assume semiconvexity of value for the remainder of the chapter without reference to a specific set of assumptions yielding the semiconvexity.

Theorem 9.7. *Suppose* $\ell \equiv 0$, $\eta(s) = \frac{\gamma^2}{2}s^2$, $M = \infty$ *and (A4.3I), (A4.4I). If (A4.4I) holds with* γ^2 *replaced by* $\gamma^2 - \delta$ *where* $0 \le V(x) \le c_f \frac{\gamma^2-\delta}{2m_\sigma^2}|x|^2$ *for all* $x \in \mathbf{R}^n$ *(where we note* $V = W$ *in this case), then* V *and* W *are semiconvex, and consequently, given* $R < \infty$*, there exists* $c, L < \infty$ *such that* $W, V \in \mathcal{S}_R^{cL}$.

Proof. This follows by Theorem 4.1, Theorem 4.2 and Corollary 4.11. □

The above result is really appropriate only to that subclass of problems without an L_∞ component. In keeping with the spirit of this chapter, we now include a proof of semiconvexity under more general conditions which would need to be verified for any specific mixed L_∞/L_2 problem.

Theorem 9.8. *Suppose* $M < \infty$*, and that* $W, V : \mathbf{R}^n \to \mathbf{R}$*. Assume that the supremum in (9.11) (equivalently (9.7)) is attained uniformly on compact sets from below (see Definition 9.5). Then* W *is semiconvex. Suppose, that the limit supremum in (9.12) (equivalently (9.8)) is attained uniformly on compact sets from above (see Definition 9.4), and that the supremum in (9.11) (equivalently (9.7)) is attained uniformly on compact sets from below with a lower time (see Definition 9.5). Then* V *is semiconvex.*

Note that some assumptions which are already being assumed throughout this chapter, are specifically called out in the above theorem statement in order to be clear regarding exactly which ones are needed for each of the two assertions. Prior to the proof we note the usual result that semiconvexity implies the local Lipschitz property (c.f. [42]). This yields the following corollary.

Corollary 9.9. *Under the conditions of either Theorem 9.7 or Theorem 9.8,* V *and* W *are locally Lipschitz.*

Proof. (proof of Theorem 9.8) Fix $R < \infty$. Let $x \in \overline{B}_R \subset \mathbf{R}^n$ and $v \in \mathbf{R}^n$ with $|v| = 1$. As in Chapter 4, we will prove semiconvexity by showing that second differences of W and V, such as $W(x+\delta v) - 2W(x) + W(x-\delta v)$, are bounded from below for sufficiently small $\delta > 0$. We will take $\delta \in (0, 1]$.

Let $\varepsilon > 0$. As the supremum in (9.11) is attained uniformly on compact sets from below, we see that there exists $T, \underline{T} = \underline{T}_{\overline{B}_R,\varepsilon}$ with $T \le \underline{T}$ such that

$$\widetilde{W}(x,T) \ge W(x) - \varepsilon \quad \forall x \in \overline{B}_R,\ \forall T \ge \underline{T}. \tag{9.28}$$

Let $u^\varepsilon \in \mathcal{U}^M_{[0,T)}$ be such that

$$J(0,T,x,u^\varepsilon) \geq \widetilde{W}(x,T) - \varepsilon. \tag{9.29}$$

Now define $x^+ = x + \delta v$, $x^- = x - \delta v$. Let trajectories $\xi^{\varepsilon,+}$, $\xi^{\varepsilon,-}$ satisfy (9.1) with initial conditions $\xi_0^{\varepsilon,+} = x^+$, $\xi_0^{\varepsilon,-} = x^-$ and inputs $u_t^+ = u_t^\varepsilon$, $u_t^- = u_t^\varepsilon + \Delta_t$. One has

$$\begin{aligned}
\dot{\xi}^{\varepsilon,+} - \dot{\xi}^\varepsilon &= f(\xi^{\varepsilon,+}) - f(\xi^\varepsilon) + [\sigma(\xi^{\varepsilon,+}) - \sigma(\xi^\varepsilon)]u^\varepsilon, \\
\dot{\xi}^\varepsilon - \dot{\xi}^{\varepsilon,-} &= f(\xi^\varepsilon) - f(\xi^{\varepsilon,-}) + \sigma(\xi^\varepsilon)u^\varepsilon - \sigma(\xi^{\varepsilon,-})[u^\varepsilon + \Delta].
\end{aligned}$$

Then, as in the proof of Theorem 4.9, one chooses

$$\begin{aligned}
\Delta_t \doteq -\sigma^{-1}(\xi_t^{\varepsilon,-})\{f(\xi_t^{\varepsilon,+}) - 2f(\xi_t^\varepsilon) + f(\xi_t^{\varepsilon,-}) \\
+[\sigma(\xi_t^{\varepsilon,+}) - 2\sigma(\xi_t^\varepsilon) + \sigma(\xi_t^{\varepsilon,-})]u_t^\varepsilon\}
\end{aligned} \tag{9.30}$$

(As before, although Δ_t is defined by the above feedback formula, the corresponding Δ_t as a function of t is used.) Consequently, one again obtains

$$\xi_t^{\varepsilon,+} - \xi_t^\varepsilon = \xi_t^\varepsilon - \xi_t^{\varepsilon,-} \quad \forall t \in [0,T]. \tag{9.31}$$

Thus, one has

$$\begin{aligned}
&\widetilde{W}(x+\delta v,T) - 2\widetilde{W}(x,T) + \widetilde{W}(x-\delta v,T) \\
&\quad \geq \ell(\xi_T^\varepsilon + [\xi_T^{\varepsilon,+} - \xi_T^\varepsilon]) - 2\ell(\xi_T^\varepsilon) + \ell(\xi_T^\varepsilon - [\xi_T^{\varepsilon,+} - \xi_T^\varepsilon]) \\
&\qquad + \int_0^T L(\xi_t^\varepsilon + [\xi_t^{\varepsilon,+} - \xi_t^\varepsilon]) - 2L(\xi_t^\varepsilon) + L(\xi_t^\varepsilon - [\xi_t^{\varepsilon,+} - \xi_t^\varepsilon])\,dt \\
&\qquad - \int_0^T \eta(|u_t^\varepsilon + \Delta_t|) - \eta(|u_t^\varepsilon|)\,dt - 2\varepsilon.
\end{aligned} \tag{9.32}$$

Using Assumption (A9.1I), this yields

$$\begin{aligned}
&\widetilde{W}(x+\delta v,T) - 2\widetilde{W}(x,T) + \widetilde{W}(x-\delta v,T) \\
&\quad \geq -2C_\ell|\xi_T^{\varepsilon,+} - \xi_T^\varepsilon|^2 - 2C_l\int_0^T |\xi_t^{\varepsilon,+} - \xi_t^\varepsilon|^2\,dt \\
&\qquad - \int_0^T k_1|u_t^\varepsilon||\Delta_t| + k_2|\Delta_t|^2\,dt - 2\varepsilon
\end{aligned} \tag{9.33}$$

Using Assumptions (A4.1I) and (A4.2I), one has

$$\begin{aligned}
\frac{d}{dt}|\xi^{\varepsilon,+} - \xi^\varepsilon|^2 &\leq 2K|\xi^{\varepsilon,+} - \xi^\varepsilon|^2 + 2K_\sigma|\xi^{\varepsilon,+} - \xi^\varepsilon|^2|u^\varepsilon| \\
&\leq 2(K + K_\sigma M)|\xi^{\varepsilon,+} - \xi^\varepsilon|^2 \doteq C_1|\xi^{\varepsilon,+} - \xi^\varepsilon|^2
\end{aligned}$$

which implies

$$|\xi_t^{\varepsilon,+} - \xi_t^\varepsilon|^2 \leq \delta^2 e^{C_1 T} \leq \delta^2 e^{C_1 \underline{T}} \quad \forall t \in [0,T]. \tag{9.34}$$

Sustituting (9.34) into (9.33) yields

$$\begin{aligned}&\widetilde{W}(x+\delta v,T)-2\widetilde{W}(x,T)+\widetilde{W}(x-\delta v,T)\\&\geq -\left[2C_\ell e^{C_1\underline{T}}+2C_l(\underline{T}/C_1)e^{C_1\underline{T}}\right]\delta^2-\int_0^T k_1|u_t^\varepsilon||\Delta_t|+k_2|\Delta_t|^2\,dt-2\varepsilon\\&\doteq C_2\delta^2-\int_0^T k_1|u_t^\varepsilon||\Delta_t|+k_2|\Delta_t|^2\,dt-2\varepsilon.\end{aligned}\tag{9.35}$$

Noting that

$$\begin{aligned}\frac{d}{dt}|\xi^\varepsilon|^2&\leq 2|\xi^\varepsilon|\left[|f(0)|+K|\xi^\varepsilon-0|+m_\sigma M\right]\\&\leq |f(0)|^2+(m_\sigma M)^2+2(K+1)|\xi^\varepsilon|^2,\end{aligned}$$

it is easy to show that there exists $D_{R,\underline{T}}<\infty$ such that

$$|\xi_t^\varepsilon|^2\leq D_{R,\underline{T}}\quad\forall t\in[0,T],\ \forall x\in\overline{B}_R.\tag{9.36}$$

By the smoothness of f and σ, there exist $Q^f_{R,\underline{T}},Q^\sigma_{R,\underline{T}}<\infty$ such that $|f_{xx}(x)|\leq Q^f_{R,\underline{T}}$ and $|\sigma_{xx}(x)|\leq Q^\sigma_{R,\underline{T}}$ for all $x\in\overline{B}_{D_{R,\underline{T}}}$. Using this, (9.30) and Assumption (A4.2I), one has

$$|\Delta_t|\leq m_\sigma[2Q^f_{R,\underline{T}}\delta^2+2Q^\sigma_{R,\underline{T}}M\delta^2]\doteq C_{3,R,\underline{T}}\delta^2\quad\forall t\in[0,\underline{T}]\tag{9.37}$$

for any initial $x\in\overline{B}_R$. Combining (9.35) and (9.37) yields

$$\begin{aligned}&\widetilde{W}(x+\delta v,T)-2\widetilde{W}(x,T)+\widetilde{W}(x-\delta v,T)\\&\geq -C_2\delta^2-\underline{T}k_1MC_{3,R,\underline{T}}\delta^2-\underline{T}k_2(C_{3,R,\underline{T}})^2\delta^4-2\varepsilon\\&\geq -C_2\delta^2-\underline{T}k_1MC_{3,R,\underline{T}}\delta^2-\underline{T}k_2(C_{3,R,\underline{T}})^2\delta^2-2\varepsilon\doteq -C_{4,R,\underline{T}}\delta^2-2\varepsilon.\end{aligned}\tag{9.38}$$

Combining the fact that W is a supremum over T of $\widetilde{W}$ and (9.28) with (9.38) yields

$$W(x+\delta v)-2W(x)+W(x-\delta v)\geq -C_{4,R,\underline{T}}\delta^2-4\varepsilon.$$

Using the fact that this is true for all $\varepsilon>0$, one finds that W is semiconvex.

Now we proceed to prove that V is semiconvex. By the assumptions that the limit supremum in (9.12) is attained uniformly on compact sets from above, and that the supremum in (9.11) is attained uniformly on compact sets from below with a lower time, one finds that there exist $\overline{T},\underline{T}\in(0,\infty)$ (dependent only on R) such that for any $x\in\overline{B}_R$, there exists $T\in[\overline{T},\underline{T}]$ (where T itself may depend on x) such that

$$\widetilde{W}(x,T)\geq W(x)-\varepsilon\geq V(x)-\varepsilon,\tag{9.39}$$

$$\widetilde{W}(x+\delta v,T)\leq V(x+\delta v)+\varepsilon,\tag{9.40}$$

and

$$\widetilde{W}(x-\delta v,T) \le V(x-\delta v)+\varepsilon. \tag{9.41}$$

By (9.39)–(9.41),

$$V(x+\delta v)-2V(x)+V(x-\delta v) \ge \widetilde{W}(x+\delta v,T)-2\widetilde{W}(x,T)+\widetilde{W}(x-\delta v,T) \\ -4\varepsilon,$$

which by (9.38)

$$\ge -C_{4,R,\underline{T}}\delta^2-6\varepsilon.$$

Because this is also true for all $\varepsilon>0$, we see that V is semiconvex. □

9.4 Max-Plus Numerical Methods

We now have the structure needed to construct max-plus based methods. In particular, we have semiconvexity of the value functions, and max-plus linear and max-plus affine representations of the DPPs. Chapters 4–7 discussed max-plus numerical methods for steady-state problems. In Chapters 4 and 5, we discussed a max-plus eigenvector approach. In Chapters 6 and 7, we considered methods based on dual-space operator construction including the curse-of-dimensionality-free method. The purpose of this chapter is to give an overview for a wider class (mixed L_∞/L_2 problems), with a correspondingly less specific study of detailed assumptions and calculations. In keeping with this broad approach, we will indicate only the main points of a generalization of the max-plus eigenvector approach. Error analyses will not be provided. Operator construction/curse-of-dimensionality-free methods have not yet been developed for this wider class of problems, but of course the prospect is promising. We will, however, discuss an interesting max-plus algebraic representation for a certain lack of uniqueness in the VI.

Suppose we wish to compute V,W over $\overline{B}_R$. We will assume throughout this section that there are sufficient conditions such that one of the semiconvexity results of the previous section holds. Then, based on this semiconvexity, one has that given $R<\infty$, $V,W\in\mathcal{S}^{cL}_{\overline{R}}$ for some $c\in\mathbf{R}$ and $L\in[0,\infty)$. Recall Theorem 4.13. In particular, let C be a symmetric matrix such that $C-cI>0$. Let $\psi_i(x)\doteq-\frac{1}{2}(x-x_i)^TC(x-x_i)$ where the $x_i\in\mathcal{N}$ form a countable dense subset of $\mathcal{E}=\{\overline{x}\in\mathbf{R}^n : \overline{x}^T(C^2)\overline{x}\le(L+|C|R)^2\}$. Then, for any $\phi\in\mathcal{S}^{cL}_{\overline{R}}$,

$$\phi(x) = \sup_{i\in\mathcal{N}}\left[a_i+\psi_i(x)\right] = \bigoplus_{i=1}^{\infty}\left[a_i\otimes\psi_i(x)\right], \tag{9.42}$$

where

$$a_i = -\max_{x\in\overline{B}_R}\left[-\phi(x)+\psi_i(x)\right] \quad \forall\, i. \tag{9.43}$$

Now, as in Chapters 4–6, we will suppose that the value function, $W(x)$ has a max-plus expansion with a specific, *finite* number of max-plus basis functions. (Again we remark that an error analysis for a particular L_2-gain case appears in Chapter 5, and we do not perform a similar analysis here.) Suppose

$$W(x) = \left\{\bigoplus_{i=1}^{\nu} [a_i \otimes \psi_i]\right\}(x) \tag{9.44}$$

with the a_i given in (9.43), with W replacing ϕ. Similarly, suppose that for each $i \in \{1, 2, \ldots, \nu\}$ one has a finite max-plus expansion of $\mathcal{L}_\tau[\psi_i]$ which we denote by

$$\mathcal{L}_\tau[\psi_i] = \bigoplus_{j=1}^{\nu} [B_{j,i} \otimes \psi_j], \tag{9.45}$$

where $B_{j,i} = -\max_x\{\psi_j(x) - \mathcal{L}_\tau[\psi_i](x)\}$, and also that

$$c_\tau = \bigoplus_{j=1}^{\nu} [c_j \otimes \psi_j], \tag{9.46}$$

where $c_j = -\max_x\{\psi_j(x) - c_\tau(x)\}$. Then, by (9.26) and (9.44)

$$\bigoplus_{j=1}^{\nu} a_j \otimes \psi_j = c_\tau \oplus \mathcal{L}_\tau\left[\bigoplus_{i=1}^{\nu} a_i \otimes \psi_i\right],$$

which by (9.45), (9.46) and the max-plus linearity of $\mathcal{L}_\tau$,

$$\begin{aligned} &= \left[\bigoplus_{j=1}^{\nu} c_j \otimes \psi_j\right] \oplus \left\{\bigoplus_{i=1}^{\nu} a_i \otimes \left[\bigoplus_{j=1}^{\nu} B_{j,i} \otimes \psi_j\right]\right\} \\ &= \left[\bigoplus_{j=1}^{\nu} c_j \otimes \psi_j\right] \oplus \left\{\bigoplus_{j=1}^{\nu}\left[\bigoplus_{i=1}^{\nu} a_i \otimes B_{j,i}\right] \otimes \psi_j\right\} \\ &= \bigoplus_{j=1}^{\nu}\left[c_j \oplus \bigoplus_{i=1}^{\nu} B_{j,i} \otimes a_i\right] \otimes \psi_j \end{aligned} \tag{9.47}$$

Under an assumption that each basis function is active (see Theorem 4.15), (9.47) implies that the vector of coefficients a_i, denoted simply by a, satisfies the affine equation

$$a = c \oplus \left[B \otimes a\right] \tag{9.48}$$

where c is the vector of coefficients c_i, and B is the $\nu \times \nu$ matrix of elements $B_{j,i}$. In summary, one finds the following.

Theorem 9.10. *The solution of DPP (9.9) is given by* $W = \bigoplus_{i=1}^{\nu} a_i \otimes \psi_i$ *where the vector of coefficients satisfies max-plus affine equation (9.48).*

Similarly, suppose for now that V has the finite expansion (but see Chapter 5 for discussion of the associated errors in a simpler case with $\ell \equiv 0$) with all basis functions active

$$V(x) = \left\{ \bigoplus_{i=1}^{\nu} [e_i \otimes \psi_i] \right\}(x) \tag{9.49}$$

with the e_i given by

$$e_i = -\max_x \{\psi_i(x) - V(x)\}. \tag{9.50}$$

Then one has (see Theorem 4.15) the following:

Theorem 9.11. *The solution of DPP (9.16) is given by* $V = \bigoplus_{i=1}^{\nu} e_i \otimes \psi_i$ *where the vector of coefficients satisfies max-plus eigenvector equation*

$$0 \otimes e = B \otimes e. \tag{9.51}$$

9.4.1 Nonuniqueness for the Max-Plus Affine Equation

There are serious nonuniqueness issues for the DPPs for both W and V. It will be simpler to quantify this lack of uniqueness with the technology below. Note that this nonuniqueness also appears in the above PDE and VI forms. Some (although not all) of these nonuniqueness issues also appear in the max-plus algebraic forms of these equations, (9.48) and (9.51).

In the case of V, the max-plus equation (9.51) is simply an eigenvector problem for eigenvalue zero. The following property can be shown to hold for some problem forms. In particular, it is shown to hold for the L_2-gain/H_∞ problem form under reasonable conditions on the dynamics and cost (see Chapter 4). We will assume throughout the remainder of the chapter that it holds for the B matrix obtained above.

Max-Plus Matrix Dissipation Property: *Let* $x_1 = 0$. $B_{1,1} = 0$, *and there exists* $N_B < \infty$, $\varepsilon > 0$ *such that for all* $N \geq N_B$ *and all* $\{k_i\}_{i=1}^N$ *such that* $k_1 = k_N$ *and not* $k_i = 1$ *for all* i, *one has* $\sum_{i=1}^{N-1} B_{k_i,k_{i+1}} < -\varepsilon$.

We will also suppose that $B_{j,i} \neq -\infty$ for all j,i, and this holds under reasonable conditions on the dynamics and choice of C in the basis functions. In particular, this has also been shown to hold in the problem of Chapters 4 and 5. The condition $B_{j,i} \neq -\infty$ for all j,i is sufficient (although not necessary) to guarantee that B has exactly one max-plus eigenvalue [6]. Further, under the additional condition of the Max-Plus Matrix Dissipation Property, there is a unique eigenvector (modulo max-plus multiplication by a scalar, of course — see Chapter 4 and [6]).

Now, consider our max-plus affine problem (9.48). Suppose this problem has solution a^0. Also suppose that the eigenvector problem, (9.51), has solution e^0. Let $a^1 \doteq a^0 \oplus e^0$. Then

$$a^1 = a^0 \oplus e^0$$

which, because a^0 is a solution of (9.48), and e^0 is a solution of (9.51),

$$\begin{aligned}
&= [c \oplus (B \otimes a^0)] \oplus (B \otimes e^0) \\
&= c \oplus [(B \otimes a^0) \oplus (B \otimes e^0)] \\
&= c \oplus [B \otimes (a^0 \oplus e^0)] = c \oplus (B \otimes a^1).
\end{aligned}$$

Therefore, one has the following.

Theorem 9.12. *Solutions of (9.48) are at most unique modulo max-plus addition by a max-plus eigenvector corresponding to max-plus eigenvalue zero.*

This nonuniqueness also holds for the original problem (before truncation of the basis expansion to finite length).

Theorem 9.13. *Solutions of affine problem (9.26) are at most unique modulo max-plus addition by a solution of eigenvector problem (9.27).*

Proof. Let W^0 satisfy (9.26), and V^0 satisfy (9.27). Let $W^1 \doteq W^0 \oplus V^0$. Then

$$\begin{aligned}
W^1 &= W^0 \oplus V^0 \\
&= \left(c_\tau \oplus \mathcal{L}_\tau[W^0]\right) \oplus \mathcal{L}_\tau[V^0]
\end{aligned}$$

which by the max-plus linearity of $\mathcal{L}_\tau$

$$\begin{aligned}
&= c_\tau \oplus \mathcal{L}_\tau[W^0 \oplus V^0] \\
&= c_\tau \oplus \mathcal{L}_\tau[W^1],
\end{aligned}$$

and so W^1 is a solution of (9.26). □

This also yields a way to view nonuniqueness in the originating DPP and VI. More specifically, if W is a solution of the DPP or VI, and if V is a solution of the corresponding DPP or PDE for the infinite time-horizon problem, then the pointwise maximum of W and V (i.e., max-plus sum of W and V) yields another solution of the DPP or VI for W.

9.4.2 The Affine Power Method

Given this lack of uniqueness in the DPP and variational inequality for W, and the corresponding lack of uniqueness in the max-plus affine equation (9.48), one could question how one would know that the solution computed for any of these characterizations was the correct solution (the value function). Interestingly, there is a method for solution of the max-plus affine equation (9.48) that yields this correct solution. The underlying reason that it yields the correct solution is that it corresponds to forward propagation of the original control problem. One particularly nice property of the solution method is that it converges exactly in a finite number of steps.

Let $F : (\mathbf{R}^-)^n \to (\mathbf{R}^-)^n$ be defined by

$$F[e] \doteq c \oplus (B \otimes e). \tag{9.52}$$

The max-plus affine power method will simply be the repeated applicaton of F until one has convergence.

Prior to discussing the max-plus affine power method in greater detail, we recall the results of Chapter 4 for the solution of (9.51). In particular, Theorem 4.22 and Corollary 4.23 indicate that (given the Max-Plus Matrix Dissipation Property) $\lim_{M\to\infty} B^M \otimes 0$ (where, here, 0 indicates the vector of length ν consisting entirely of zeros) exists and converges in a finite number of steps to the unique solution of eigenvector problem (9.51).

We will denote the initial vector for the affine power method as a^0. The following requires only a very minor modification of the proof of Theorem 4.22.

Lemma 9.14. *Given any $a^0, c \in (\mathbf{R}^-)^n$, and B satisfying the Max-Plus Matrix Dissipation Property, there exists $\widehat{K} < \infty$ such that*

$$B^k \otimes a^0 = B^{\widehat{K}} \otimes a^0 \;\; \textit{and} \quad B^k \otimes c = B^{\widehat{K}} \otimes c$$

for all $k \geq \widehat{K}$.

Now note that for any $k \geq \widehat{K}$, one has

$$F^{k+1}[a^0] = \left[\bigoplus_{i=0}^{k} \left(B^i \otimes c\right) \right] \oplus \left(B^{k+1} \otimes a^0\right),$$

which by the assumptions and that $d \oplus d = d$ for any d (idempotency),

$$= \left[\bigoplus_{i=0}^{\widehat{K}} \left(B^i \otimes c\right) \right] \oplus \left(B^{\widehat{K}+1} \otimes a^0\right) = F^{\widehat{K}+1}[a^0].$$

Let

$$a^* \doteq \lim_{k\to\infty} F^k[a^0]. \tag{9.53}$$

Then

$$a^* = F^{\widehat{K}+1}[a^0]. \tag{9.54}$$

Further,

$$F[a^*] = F^{\widehat{K}+2}[a^0] = F^{\widehat{K}+1}[a^0] = a^*.$$

Consequently, one has the following:

Theorem 9.15. *For any initial a^0, a^* given by (9.53) is a solution of (9.48).*

Not only is the limit a solution of (9.48) and achieved in a finite number of steps, we will see that it is also the correct solution of (9.48) in that it is the solution corresponding to the value function. Let the k^{th} iterate be $a^k = F^k[a^0]$. Also, define the corresponding k^{th} approximation of the solution to be

$W^k(x) \doteq \bigoplus_{i=1}^{\nu} a_i^k \otimes \psi_i(x)$. Note that from the above, one has $W^k = W^{\widehat{K}+1}$ for all $k \geq \widehat{K}+1$. Let $W^*(x) \doteq \bigoplus_{i=1}^{\nu} a_i^* \otimes \psi_i(x) = \bigoplus_{i=1}^{\nu} a_i^{\widehat{K}+1} \otimes \psi_i(x)$. We will choose a^0 such that

$$W^0(x) \leq W(x) \quad \forall\, x \in \mathbf{R}^n, \tag{9.55}$$

where W is the value function. Note that this is an assumption on the choice of the a_i^0 coefficients. Under the above assumptions (specifically non-negativity of L, ℓ and $\eta(0) = 0$), one has $W(x) \geq 0$ for all x. Consequently, if the basis functions are of the form $\psi(x) = -\frac{1}{2}(x-x_i)^T C(x-x_i)$ with C positive definite, then one only needs to take the $a_i^0 \leq 0$ in order to satisfy (9.55).

Theorem 9.16. *Suppose W^0 satisfies (9.55). W^* given by the above algorithm is the correct solution of the DPP (i.e., the value function of the original control problem).*

Proof. The result will follow by showing that repeated application of the F operator corresponds to forward propagation of the value function of a finite time-horizon problem. We only sketch the main points of the proof. First note that

$$\begin{aligned} W^1(x) &= \bigoplus_{i=1}^{\nu} \Big\{ \Big[c_i \oplus (B \otimes a^0)_i \Big] \otimes \psi_i(x) \Big\} \\ &= \Big[\bigoplus_{i=1}^{\nu} c_i \otimes \psi_i(x) \Big] \oplus \Big[\bigoplus_{i=1}^{\nu} (B \otimes a^0)_i \otimes \psi_i(x) \Big], \end{aligned}$$

which by the definitions of c and B,

$$\begin{aligned} &= \max\Big\{ c_\tau(x), \mathcal{L}_\tau[W^0](x) \Big\} \\ &= \max\Big\{ \sup_{t \in [0,\tau)} \sup_{u \in \mathcal{U}^M_{[0,\tau)}} J(0,t,x,u), \\ &\qquad \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \Big[I(0,\tau,x,u) + W^0(\xi_\tau) \Big] \Big\}. \end{aligned} \tag{9.56}$$

Similarly,

$$\begin{aligned} W^2(x) &= \bigoplus_{i=1}^{\nu} \Big\{ \Big[c_i \oplus (B \otimes a^1)_i \Big] \otimes \psi_i(x) \Big\} \\ &= \max\Big\{ \sup_{t \in [0,\tau)} \sup_{u \in \mathcal{U}^M_{[0,\tau)}} J(0,t,x,u), \ \sup_{u \in \mathcal{U}^M_{[0,\tau)}} \Big[I(0,\tau,x,u) + W^1(\xi_\tau) \Big] \Big\} \end{aligned}$$

and using (9.56),

$$= \max\left\{ \sup_{t\in[0,2\tau)} \sup_{u\in\mathcal{U}^M_{[0,2\tau)}} J(0,t,x,u), \sup_{u\in\mathcal{U}^M_{[0,2\tau)}} \left[I(0,2\tau,x,u) + W^0(\xi_{2\tau}) \right] \right\}.$$

By induction, one finds

$$W^k(x) = \max\Big\{ \sup\nolimits_{t\in[0,k\tau)} \sup\nolimits_{u\in\mathcal{U}^M[0,k\tau)} J(0,t,x,u), \\ \sup\nolimits_{u\in\mathcal{U}^M_{[0,k\tau)}} \left[I(0,k\tau,x,u) + W^0(\xi_{k\tau}) \right] \Big\}. \tag{9.57}$$

The next step is to note that given $\varepsilon > 0$, there exists $K_{\varepsilon,x} < \infty$ such that (by the definition of W)

$$W(x) \le \sup_{t\in[0,k\tau)} \sup_{u\in\mathcal{U}^M_{[0,k\tau)}} J(0,t,x,u) + \varepsilon$$

for any $k \ge K_{\varepsilon,x}$. Consequently, using (9.57), one has

$$W(x) \le W^k(x) + \varepsilon \tag{9.58}$$

for any $k \ge K_{\varepsilon,x}$. On the other hand, by the DPP of Theorem 9.1 one has

$$W(x) = \max\left\{ \sup_{t\in[0,k\tau)} \sup_{u\in\mathcal{U}^M_{[0,k\tau)}} J(0,t,x,u), \sup_{u\in\mathcal{U}^M_{[0,k\tau)}} \left[I(0,k\tau,x,u) + W(\xi_{k\tau}) \right] \right\}$$

which by the condition (9.55),

$$\ge \max\left\{ \sup_{t\in[0,k\tau)} \sup_{u\in\mathcal{U}^M_{[0,k\tau)}} J(0,t,x,u), \sup_{u\in\mathcal{U}^M_{[0,k\tau)}} \left[I(0,k\tau,x,u) + W^0(\xi_{k\tau}) \right] \right\}$$

$$= W^k(x). \tag{9.59}$$

Combining (9.58) and (9.59) leads to the result. □

A

Miscellaneous Proofs

A.0.1 Sketch of Proof of Theorem 2.8

A short sketch of a proof of Theorem 2.8 follows. For a more general and complete treatment see,[53]; [100], [101]; and/or [102].

Let $\widehat{\psi}$ be given by (2.4) (i.e., $\widehat{\psi}(p) = -\max_{x\in\overline{B}_R}[x^T p - \widehat{\phi}(x)]$). This implies that

$$-\widehat{\psi}(p) \geq x^T p - \widehat{\phi}(x) \quad \forall\, x \in \overline{B}_R,\ \forall\, p \in \overline{B}_L.$$

or

$$\widehat{\phi}(x) \geq p^T x + \widehat{\psi}(p) \quad \forall\, x \in \overline{B}_R,\ \forall\, p \in \overline{B}_L$$

which implies

$$\widehat{\phi}(x) \geq \max_{p\in\overline{B}_L}[p^T x + \widehat{\psi}(p)] \quad \forall\, x \in \overline{B}_R. \tag{A.1}$$

On the other hand, by the convexity and Lipschitz conditions, one can show that given any $\overline{x} \in \overline{B}_R$, there exists $\overline{p} \in \overline{B}_L$ and $\overline{r} \in \mathbf{R}$ such that

$$\widehat{\phi}(\overline{x}) = \overline{p}^T \overline{x} + \overline{r}, \tag{A.2}$$
$$\widehat{\phi}(x) \geq \overline{p}^T x + \overline{r} \quad \forall\, x \in \overline{B}_R. \tag{A.3}$$

By (A.3),

$$\overline{r} \leq \min_{x\in\overline{B}_R}[\widehat{\phi}(x) - \overline{p}^T x] = -\max_{x\in\overline{B}_R}[\overline{p}^T x - \widehat{\phi}(x)] = \widehat{\psi}(p) \tag{A.4}$$

Substituting this into (A.2), one finds

$$\widehat{\phi}(\overline{x}) \leq \overline{p}^T \overline{x} + \widehat{\psi}(p) \leq \max_{p\in\overline{B}_L}[\overline{x}^T p + \widehat{\psi}(p)]. \tag{A.5}$$

Comparing (A.1) and (A.5) yields the result. □

A.0.2 Proof of Theorem 3.13

The following proof is a reduction of Da Lio's uniqueness proof from [26] to the case needed here. The main difference is that [26] concerned a game problem, whereas we are concerned with a control problem here.

Recall that the HJB PDE problem of interest here is

$$0 = -V_s(s,x) + H(x, \nabla_x V(s,x)) \quad \forall\, (s,x) \in (0,T) \times \mathbf{R}^n, \tag{A.6}$$

$$V(T,x) = \phi(x) \quad \forall\, x \in \mathbf{R}^n \tag{A.7}$$

where the Hamiltonian, H, is given by

$$H(x,p) \doteq -\max_{u \in U} \left\{ [f(x) + \sigma(x)u]^T p + l(x) - \frac{\gamma^2}{2}|u|^2 \right\} \tag{A.8}$$

where $U \subseteq \mathbf{R}^l$. We weaken the assumption that $|\sigma(x)| \le m_\sigma$ to $|\sigma(x)| \le m_\sigma(1+|x|)$ for all x for the purposes of maintaining a bit more of the generality of the original. (Note that such a linear-growth bound is implied by (A3.1F), but now we drop the condition $|\sigma(x)| \le m_\sigma$ and use the notation m_σ for the constant in the linear-growth bound, $|\sigma(x)| \le m_\sigma(1+|x|)$ instead.)

Let $V_1, V_2 \in \mathcal{K}$ be viscosity solutions of (A.6) meeting boundary condition (A.7). We first note that if (x,p) belongs to a compact set $A_c \subseteq \mathbf{R}^{2n}$, then there exists $\widehat{R} > 0$, depending on A_c, such that

$$\begin{aligned} H(x,p) &= H_{\widetilde{R}}(x,p) \\ &\doteq -\max_{u \in U, |u| \le \widehat{R}} \left\{ [f(x) + \sigma(x)u]^T p + l(x) - \frac{\gamma^2}{2}|u|^2 \right\} \quad \forall\, (x,p) \in A_c. \end{aligned} \tag{A.9}$$

Now fix $r > 0$ and choose $R > r$. Because V_1, $V_2 \in \mathcal{K}$ and (A.9) holds, there exists $\widetilde{R} > 0$ such that

$$H(x,p) = H_{\widetilde{R}}(x,p) \quad \forall\, p \in D_x^+ V_1(t,x) \cup D_x^+ V_2(t,x),\ (t,x) \in [0,T] \times \overline{B}_R(0).$$

Thus V_1, V_2 are viscosity solutions of

$$\begin{cases} -V_s(s,x) + H_{\widetilde{R}}(x, \nabla_x V(s,x)) = 0 & \forall (s,x) \in (0,T) \times B_R(0), \\ V(x,T) = \phi(x) & \forall x \in B_R(0). \end{cases} \tag{A.10}$$

Let $\widetilde{U} \doteq U \cap \overline{B}_{\widetilde{R}}(0)$. The continuous viscosity solutions V of (A.10) satisfy the *optimality principle* (see, e.g., Propositions 2.1 and 2.2 in [68]); namely, for all $x \in B_r(0)$, $s \in [0,T]$ and $0 \le \rho \le T - s$, the following estimate holds:

$$V(s,x) = \sup_{u \in \mathcal{U}^{\widetilde{U}}} \left\{ \mathcal{I}(s, (s+\rho) \wedge t_x^R(u)) + \ V((s+\rho) \wedge t_x^R(u), \xi_{(s+\rho)\wedge t_x^R(u)}) \right\}, \tag{A.11}$$

where for all $s,t \in [0,T]$, $\mathcal{I}(s,t) := \int_s^t l(\xi_\tau) - \frac{\gamma^2}{2}|u_\tau|^2\,d\tau$, and $t_x^R(u)$ is the time of first exit of ξ from $B_R(0)$ where ξ satisfies the finite time-horizon dynamics with $\xi_s = x \in B_r(0)$, i.e.,

$$t_x^R(u) := \inf\{t \geq s \,|\, |\xi_t| \geq R\}.$$

In particular, (A.11) is satisfied by V_1, V_2.

Let $\lambda \doteq \inf_{v\in U} |u|^2$. Fix any $\varepsilon \in (0,1]$. We observe that, because $V_1, V_2 \geq 0$, the supremum in (A.11) corresponding to $\rho = T - s$, may be confined to the controls u such that for $i = 1,2$, $\mathcal{I}(s, T \wedge t_x^R(u)) + V_i(T \wedge t_x^R(u), \xi_{T\wedge t_x^R(u)}) \geq -\Lambda$ where $\Lambda \doteq -\frac{\gamma^2}{2}(\lambda + \varepsilon)T$. Define the set

$$\mathcal{A} \doteq \left\{u \in \mathcal{U}^U \,:\, \mathcal{I}(s, T \wedge t_x^R(u)) + V_i(T \wedge t_x^R(u), \xi_{T\wedge t_x^R(u)}) \geq -\Lambda \;\; \forall i \in \{1,2\}\right\}.$$

Let $v^\varepsilon \in U$ be such that $|v^\varepsilon|^2 \leq \lambda + \varepsilon$, and $u^\varepsilon_\cdot \in \mathcal{U}^u$ be given by $u^\varepsilon_\tau \equiv v^\varepsilon$ for all τ (which implies $-\frac{\gamma^2}{2}\|u\|^2_{L_2(s,T)} \geq -\Lambda$). We observe that $u^\varepsilon_\cdot \in \mathcal{A}$, and consequently, $\mathcal{A} \neq \emptyset$.

Now we suppose that

$$T \leq \delta(\gamma, r, R) \doteq \min\left\{1,\ \frac{R-r}{K(1+R)}, \frac{\gamma^2(R-r)^2}{2m_\sigma^2(1+R)^2[\overline{C}(1+R^2)+\Lambda] + 2\gamma^2 K(R-r)(1+R)}\right\}, \tag{A.12}$$

where

$$\overline{C} = 2\max\left\{C_l, C_\phi, \sup_{(t,x)\in[0,T]\times\mathbf{R}^n} \frac{|V_1(t,x)|}{1+|x|^2}, \sup_{(t,x)\in[0,T]\times\mathbf{R}^n} \frac{|V_2(t,x)|}{1+|x|^2}\right\}.$$

We claim that if $t_x^R(u) < T$, then $u \notin \mathcal{A}$. In fact, suppose that $\xi_{\bar{t}} \in \partial B_R(0)$, for some $\bar{t} \in [s,T)$, and $\xi_t \in B_R(0)$ for all $t \in [s,\bar{t})$. Then we have the following estimate:

$$\begin{aligned} R - r \leq R - |x| &= |\xi_{\bar{t}}| - |x| \leq \int_s^{\bar{t}} K(1+|\xi_\tau|)\,d\tau + \int_s^{\bar{t}} m_\sigma(1+|\xi_\tau|)|u_\tau|\,d\tau \\ &\leq K(1+R)(\bar{t}-s) + m_\sigma(1+R)\|u\|_{L^2(s,\bar{t})}(\bar{t}-s)^{1/2}. \end{aligned}$$

Thus $\|u\|_{L^2(s,\bar{t})} \geq \chi[r,R](\bar{t})$, where

$$\chi[r,R](\bar{t}) \doteq \frac{R - r - K(1+R)(\bar{t}-s)}{m_\sigma(1+R)(\bar{t}-s)^{1/2}}.$$

We observe that by supposition (A.12), $\chi[r,R](t)$ is positive for all $t \in [s,T]$. Hence we have

$$\begin{aligned}
\mathcal{I}(s,\bar{t}) + V_i(\bar{t},\xi_{\bar{t}}) &\le \int_s^{\bar{t}} C_l(1+|\xi_t|^2) - \frac{\gamma^2}{2}|u_t|^2\,dt + \widetilde{C}(1+|\xi_{\bar{t}}|^2) \\
&\le C_l(1+R^2)(\bar{t}-s) - \frac{\gamma^2}{2}|\chi[r,R](\bar{t})|^2 + \widetilde{C}(1+R^2) \\
&\le \overline{C}(1+R^2) + \gamma^2\frac{K(R-r)(1+R)}{m_\sigma^2(1+R)^2} - \frac{\gamma^2}{2}\frac{(R-r)^2}{m_\sigma^2(1+R)^2(\bar{t}-s)} \\
&\quad - \frac{\gamma^2}{2}\frac{K^2(\bar{t}-s)}{m_\sigma^2} \\
&\le \overline{C}(1+R^2) + \gamma^2\frac{K(R-r)(1+R)}{m_\sigma^2(1+R)^2} - \frac{\gamma^2}{2}\frac{(R-r)^2}{m_\sigma^2(1+R)^2(T-s)},
\end{aligned}$$

where $\widetilde{C} \doteq \max(\sup_{(t,x)\in[0,T]\times\mathbf{R}^n}\frac{|V_1(t,x)|}{1+|x|^2}, \sup_{(t,x)\in[0,T]\times\mathbf{R}^n}\frac{|V_2(t,x)|}{1+|x|^2})$. Combining this with condition (A.12) and a little algebra, we have $\mathcal{I}(s,\bar{t}) + V(\bar{t},\xi_{\bar{t}}) < -\Lambda$, and this proves the claim. Therefore for all $(s,x) \in [0,T]\times\overline{B}_r(0)$, each V_i $(i=1,2)$ satisfies

$$V_i(s,x) = \sup_{u\in\mathcal{U}^{\widetilde{U}}}\{\mathcal{I}(s,T) + \phi(\xi_T)\}, \tag{A.13}$$

and we can conclude that $V_1 = V_2$ in $[0,T]\times\overline{B}_r(0)$. In the case of $T > \delta(\gamma,r,R)$, we can divide the interval $[0,T]$ into subintervals whose length is less than $\delta(\gamma,r,R)$. Let $0 = t_0, t_1, \ldots, t_n = T$ be the points of such a division, and for any $k = 1,\ldots,n$ let us consider the following Cauchy problem:

$$\begin{aligned}
&-V_s(s,x) + H_{\widetilde{R}}(x,\nabla V(s,x)) = 0 \quad \forall\,(s,x)\in(t_{k-1},t_k)\times B_R(0), \\
&V(t_k,x) = \phi_k(x) \quad \forall\,x\in B_R(0),
\end{aligned}$$

where the terminal value $\phi_k(x)$ can coincide either with $V_1(t_k,x)$ or with $V_2(t_k,x)$. We start with $k=n$. Because $|t_n - t_{n-1}| < \delta(\gamma,r,R)$, we can argue as above and obtain that the V_i's coincide in $[t_{n-1},t_n]\times\overline{B}_r(0)$. Then, by proceeding backward in the time variable, we obtain that for all $k\in\{0,\ldots,n\}$, $V_1 = V_2$ in $[t_{k-1},t_k]\times\overline{B}_r(0)$, and this completes the proof. □

A.0.3 Proof of Lemma 3.15

From Lemma 3.14, we know that $\overline{V}$ is the *unique* continuous viscosity solution of (3.46), (3.47). If we can prove that the value function of the indicated problem is a continuous viscosity solution, then the uniqueness implies that the value function and $\overline{V}$ are one in the same.

The first steps involve proving that the value is bounded above by a quadratic function, and that consequently, there is an L_2 bound on ε-optimal controls u. In fact, we already have the quadratic bound. In particular, by

(3.32) (using s rather than 0 as the starting time, and indicating this in the argument list of J) we have

$$J(s,x,T,u) \le \left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{c_f(s-T)} + \frac{\alpha_l}{c_f}\right]|x|^2 - \delta\|u\|^2_{L_2[s,T]} \tag{A.14}$$

for all $u \in \mathcal{U}^U$, and by (A3.3I), (3.21)

$$J(s,x,T,0) \ge 0. \tag{A.15}$$

(A.14) and (A.15) imply the value satisfies

$$0 \le V(s,x) \le \left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{c_f(s-T)} + \frac{\alpha_l}{c_f}\right]|x|^2. \tag{A.16}$$

Also, for ε-optimal disturbances u, we already have from Theorem 3.10

$$\frac{1}{2}\|u\|^2_{L_2[s,T]} \le \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{c_f(s-T)} + \frac{\alpha_l}{c_f}\right]|x|^2. \tag{A.17}$$

We will use this to show that the state process satisfies a certain continuity bound near the initial time for all ε-optimal controls u. This result ((A.18) below) will be similar to the continuity proof in Theorem 3.11, and will be obtained in the same manner. Let u be ε-optimal . By (3.12)

$$|\xi_t - x| \le \int_s^t \left[|f(\xi_r)| + |\sigma(\xi_r)||u_r|\right] dr,$$

which by (A3.1I), (A3.2I)

$$\le K\int_s^t |\xi_r - x|\,dr + K|x|(t-s) + m_\sigma \int_s^t |u_r|\,dr$$

$$\le K\int_s^t |\xi_r - x|\,dr + K|x|(t-s) + m_\sigma \left[\int_s^t |u_r|^2\,dr\right]^{\frac{1}{2}} \sqrt{t-s},$$

which by (A.17)

$$\le K\int_s^t |\xi_r - x|\,dr + K|x|(t-s)$$

$$+m_\sigma\left[\frac{2\varepsilon}{\delta} + \frac{2}{\delta}\left[\frac{c_f\gamma^2}{2m_\sigma^2}e^{c_f(s-T)} + \frac{\alpha_l}{c_f}\right]|x|^2\right]^{\frac{1}{2}}\sqrt{t-s}.$$

Employing Gronwall's inequality yields

$$|\xi_t - x| \le [C_1 + C_2|x|][(t-s) + \sqrt{t-s}] \tag{A.18}$$

for all $t \in [s, s + (T - s)/2]$ for sufficiently large C_1 and C_2. One also obtains continuity with respect to initial conditions similarly. In particular, if one lets ξ and η satisfy (3.12) with $\xi_s = x$ and $\eta_s = y$ and the same control u which is ε-optimal for x, one obtains

$$\frac{d}{dt}|\xi - \eta|^2 \leq -2c_f|\xi - \eta|^2 + 2K_\sigma|\xi - \eta|^2|u|,$$

which leads to

$$|\xi_t - \eta_t| \leq \exp\left\{-c_f t + 2K_\sigma \left[\frac{\varepsilon}{\delta} + \frac{1}{\delta}\left(\frac{c_f\gamma^2}{2m_\sigma^2}e^{c_f(s-T)} + \frac{\alpha_l}{c_f}\right)|x|^2\right]\right\}. \tag{A.19}$$

The next step is the DPP. The proof is standard, so is not included. (However, see the DPP results in Chapter 3.) The final steps in proving that $\widetilde{V}$ is a continuous viscosity solution of the HJB PDE are similar to those in the proof of Theorem 3.11 with for instance (A.19) replacing (3.39). The details are not included. □

A.0.4 Sketch of Proof of Theorem 7.27

Fix $\delta > 0$ (used in the definition of $\mathcal{C}_\delta$). Suppose $\overline{V}' \in \mathcal{C}_\delta$ satisfies (7.44). Then,

$$\begin{aligned}\overline{V}'(x) &= \bar{\bar{S}}^\tau_{N\tau}[\overline{V}'](x) \\ &= \sup_{u\in\mathcal{U}} \sup_{\mu\in\mathcal{D}^\tau_\infty} \left\{\int_0^{N\tau} l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|u_t|^2\,dt + \overline{V}'(\xi_{N\tau})\right\} \qquad \forall x \in \mathbf{R}^n\end{aligned}$$

where ξ satisfies (7.13). Fix $x \in \mathbf{R}^n$, and let $\mu^\varepsilon \in \mathcal{D}^\tau_\infty$, $u^\varepsilon \in \mathcal{U}$ be ε–optimal, that is,

$$\overline{V}'(x) \leq \int_0^{N\tau} l^{\mu^\varepsilon_t}(\xi^\varepsilon_t) - \frac{\gamma^2}{2}|u^\varepsilon_t|^2\,dt + \overline{V}'(\xi^\varepsilon_{N\tau}) + \varepsilon,$$

where ξ^ε satisfies (7.13) with inputs $\mu^\varepsilon, u^\varepsilon$.

Following the same steps as in Chapter 3, one obtains essentially the same lemmas:

Lemma A.1 *For any $N < \infty$,*

$$\|u^\varepsilon\|^2_{L_2(0,N\tau)} \leq \frac{\varepsilon}{\delta} + \frac{1}{\delta}\left[\frac{c_A\gamma^2}{m_\sigma^2}e^{-c_A N\tau} + \frac{c_D}{c_A}\right]|x|^2.$$

Lemma A.2 *For any $N < \infty$,*

$$\int_0^{N\tau} |\xi^\varepsilon_t|^2\,dt \leq \frac{\varepsilon}{\delta}\frac{m_\sigma^2}{c_A} + \frac{m_\sigma^2}{\delta}\left[\left(\frac{c_D}{c_A^2} + \frac{\gamma^2}{m_\sigma^2}\right) + \frac{1}{c_A}\right]|x|^2.$$

Lemma A.3 *If $u^\varepsilon, \mu^\varepsilon$ are ε-optimal over $[0, N\tau)$, then they are also ε-optimal over $[0, n\tau)$ for all $n \le N$, that is,*

$$\int_0^{n\tau} l^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\, dt + \overline{V}'(\xi_{n\tau}^\varepsilon) \ge \overline{V}'(x) - \varepsilon.$$

The independence of the above bounds with respect to N is important. Specifically, because there is a finite bound on the energy (the bound on u^ε) coming in to the trajectories, roughly speaking the ξ^ε "tend" toward the origin.

Now we need a lemma which will replace equation (3.58).

Lemma A.4 *For any $N < \infty$,*

$$\sum_{n=1}^N |\xi_{n\tau}^\varepsilon|^2 \le \frac{1}{1 - e^{-c_A \tau}} \left[|x|^2 + (m_\sigma / c_A^2) \|u^\varepsilon\|_{L_2(0,N\tau)}^2 \right].$$

Proof. Note that

$$\frac{d}{dt}|\xi^\varepsilon|^2 \le -2c_A|\xi^\varepsilon|^2 + 2m_\sigma|\xi^\varepsilon||u^\varepsilon| \le -c_A|\xi^\varepsilon| + \hat{d}|u^\varepsilon|^2,$$

with $\hat{d} = m_\sigma^2/c_A$. Solving this on interval $[n\tau, (n+1)\tau)$ implies that

$$|\xi_t^\varepsilon|^2 \le |\xi_{(n\tau)}^\varepsilon|^2 e^{-c_A\tau} + \hat{d}\|u^\varepsilon\|_{L_2(n\tau,(n+1)\tau)}^2 \quad \forall\, t \in [n\tau, (n+1)\tau].$$

In particular, one has

$$|\xi_\tau^\varepsilon|^2 \le |x|^2 e^{-c_A\tau} + \hat{d}\|u^\varepsilon\|_{L_2(0,\tau)}^2,$$
$$|\xi_{2\tau}^\varepsilon|^2 \le |\xi_\tau^\varepsilon|^2 e^{-c_A\tau} + \hat{d}\|u^\varepsilon\|_{L_2(\tau,2\tau)}^2,$$

and these two inequalities imply

$$|\xi_\tau^\varepsilon|^2 + |\xi_{2\tau}^\varepsilon|^2 \le (e^{-c_A\tau} + e^{-2c_A\tau})|x|^2 + \hat{d}(1 + e^{-c_A\tau})\|u^\varepsilon\|_{L_2(0,\tau)}^2 + \hat{d}\|u^\varepsilon\|_{L_2(\tau,2\tau)}^2.$$

Continuing this process, one finds

$$\sum_{n=1}^N |\xi_{n\tau}^\varepsilon|^2 \le \left(\sum_{n=1}^N e^{-nc_A\tau}\right)|x|^2 + \hat{d}\sum_{n=1}^N \left[\left(\sum_{j=0}^{N-n} e^{-jc_A\tau}\right)\|u^\varepsilon\|_{L_2((n-1)\tau,n\tau)}^2\right].$$

Using the standard geometric series limit yields the result. □

Combining Lemmas A.2 and A.4, one obtains a bound on $\sum_{n=1}^N |\xi_{n\tau}^\varepsilon|^2$ which is independent of N. Consequently, at least some of the $|\xi_{n\tau}^\varepsilon|$ can be guaranteed to be arbitrarily small for large N. The remainder of the proof (of Theorem 7.27) then follows as in equations (3.62) to (3.66), but with $N\tau$ replacing T, and $n\tau$ replacing τ. This completes the sketch of the proof. □

A.0.5 Sketch of Proof of Lemma 7.31

Fix $\hat{\varepsilon}, \delta > 0$ (where δ is used in the definition of $\mathcal{C}_\delta$). Fix $m \in \mathcal{M}$. Fix any $T < \infty$ and $x \in \mathbf{R}^n$. Let $\varepsilon = (\hat{\varepsilon}/2)(1+|x|^2)$. Let $u^\varepsilon \in \mathcal{W}$, $\mu^\varepsilon \in \mathcal{D}_\infty$ be ε-optimal for $\widetilde{S}_T[W^m](x)$, i.e.,

$$\widetilde{S}_T[W^m](x) - \left[\int_0^T l^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\xi_T^\varepsilon)\right] \le \varepsilon = \frac{\hat{\varepsilon}}{2}(1+|x|^2) \quad \text{(A.20)}$$

where ξ^ε satisfies (7.13) with inputs u^ε, μ^ε.

We will let $\overline{\xi}^\varepsilon$ satisfy (7.13) with inputs u^ε and a $\overline{\mu}^\varepsilon \in \mathcal{D}_\infty^\tau$ (where τ has yet to be chosen). Solving (7.13), one has

$$\begin{aligned} \xi_t^\varepsilon &= \exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right] x + \int_0^t \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\right] \sigma^{\mu_r^\varepsilon} u_r^\varepsilon\,dr \\ \overline{\xi}_t^\varepsilon &= \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\right] x + \int_0^t \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\right] \sigma^{\overline{\mu}_r^\varepsilon} u_r^\varepsilon\,dr. \end{aligned}$$

Consequently,

$$\begin{aligned} |\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon| \le{} & \left|\exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right] - \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\right]\right| |x| \qquad \text{(A.21)} \\ & + \left\{\int_0^t \left|\exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\right]\sigma^{\mu_r^\varepsilon} - \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\right]\sigma^{\overline{\mu}_r^\varepsilon}\right|^2 dr\right\}^{1/2} \|u^\varepsilon\|_{L_2(0,t)}. \end{aligned}$$

We now simply show that this can be made arbitrarily small by taking τ small. We will use the boundedness of $\|u^\varepsilon\|$ and $\|\xi^\varepsilon\|$ which is independent of t for this class of systems (Chapter 3).

Consider the first term on the right in (A.21). Note that

$$\begin{aligned} & \left|\exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right] - \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\right]\right| \\ & = \left|\exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right]\right| \left|1 - \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\right]\right|. \qquad \text{(A.22)} \end{aligned}$$

Fix $\tau > 0$. For any subset of $\mathbf{R}$, $\mathcal{I}$, let $\mathcal{L}(\mathcal{I})$ be the Lebesgue measure of $\mathcal{I}$. Let N be the largest integer such that $N\tau \le t$. Given $m \in \mathcal{M}$, let

$$\begin{aligned} \mathcal{I}^m &= \{r \in [0, N\tau) \,|\, A^{\mu_r^\varepsilon} = A^m\}, \\ \lambda^m &= \mathcal{L}(\mathcal{I}^m). \end{aligned}$$

Let $n_0 = 0$. For $1 \le k < M = \#\mathcal{M}$, let n_k be the largest integer such that $n_k\tau \le \lambda^k + n_{k-1}\tau$. For $m < M$ let

$$\overline{\mu}_r^\varepsilon = m \quad \forall t \in [n_{m-1}\tau, n_m\tau).$$

Let $\overline{\mu}^\varepsilon{}_r = M$ for all $t \in [n_{M-1}\tau, t) = [n_{M-1}\tau, N\tau) \cup [N\tau, t)$. With this choice of $\overline{\mu}^\varepsilon$, one finds

$$\left| 1 - \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\, dr - \int_0^t A^{\mu_r^\varepsilon}\, dr \right] \right| < \beta_\tau^1 \tag{A.23}$$

where $\beta_\tau^1 \to 0$ as $\tau \to 0$ independent of t. We skip the details.

Let $y \in \mathbf{R}^n$. Define $F_t = \exp\left[\int_0^t A^{\mu_r^\varepsilon}\, dr\right]$. Then,

$$\begin{aligned} \frac{d}{dt}\left[y^T F_t^T F_t y\right] &= y^T \left[F_t^T \dot{F}_t + \dot{F}_t^T F_t\right] y = 2y^T \left[F_t^T A^{\mu_t^\varepsilon} F_t\right] y \\ &= 2(F_t y)^T A^{\mu_t^\varepsilon} (F_t y), \end{aligned}$$

which by Assumption Block (A7.1I)

$$\le -2c_A |F_t y|^2 = -2c_A \left[y^T F_t^T F_t y\right].$$

Solving this ordinary differential inequality, one finds

$$\left[y^T F_t^T F_t y\right] \le |y|^2 e^{-2c_A t}.$$

Because this is true for all $y \in \mathbf{R}^n$, we have

$$\left| \exp\left[\int_0^t A^{\mu_r^\varepsilon}\, dr \right] \right| \le e^{-c_A t} \quad \forall t \ge 0. \tag{A.24}$$

By (A.22), (A.23) and (A.24)

$$\left| \exp\left[\int_0^t A^{\mu_r^\varepsilon}\, dr \right] - \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\, dr \right] \right| \le \beta_\tau^1 e^{-c_A t} \quad \forall t \ge 0. \tag{A.25}$$

We now turn to the second term on the right-hand side of (A.21). Note that

$$\begin{aligned} &\left\{ \int_0^t \left| \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\, d\rho \right] \sigma^{\mu_r^\varepsilon} - \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\, d\rho \right] \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr \right\}^{1/2} \\ &\le \left\{ 2 \int_0^t \left| \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\, d\rho \right] \right|^2 \left| \sigma^{\mu_r^\varepsilon} - \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr \right. \\ &\qquad \left. + 2 \int_0^t \left| \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\, d\rho \right] - \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\, d\rho \right] \right|^2 \left| \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr \right\}^{1/2} \end{aligned}$$

and proceeding as above

$$\le \left\{ 2 \int_0^t e^{-2c_A(t-r)} \left| \sigma^{\mu_r^\varepsilon} - \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr + 2\beta_\tau^1 \int_0^t e^{-2c_A(t-r)} \left| \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr \right\}^{1/2}$$

$$\leq \left\{ 2 \left[\int_0^t e^{-4c_A(t-r)} \, dr \right]^{1/2} \left[\int_0^t \left| \sigma^{\mu_r^\varepsilon} - \sigma^{\overline{\mu}_r^\varepsilon} \right|^4 \, dr \right]^{1/2} \right.$$
$$\left. +2\beta_\tau^1 m_\sigma^2 \int_0^t e^{-2c_A(t-r)} \, dr \right\}^{1/2}.$$

Further, there exists β_τ^2 such that $[\int_0^t |\sigma^{\mu_r^\varepsilon} - \sigma^{\overline{\mu}_r^\varepsilon}|^4 \, dr]^{1/2} \leq \beta_\tau^2$ where $\beta_\tau^2 \to 0$ as $\tau \to 0$, and we skip the obvious but technical details. Consequently,

$$\left\{ \int_0^t \left| \exp\left[\int_r^t A^{\mu_\rho^\varepsilon} \, d\rho \right] \sigma^{\mu_r^\varepsilon} - \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon} \, d\rho \right] \sigma^{\overline{\mu}_r^\varepsilon} \right|^2 dr \right\}^{1/2}$$
$$\leq \left\{ 2\beta_\tau^2 (4c_A)^{-1/2} + 2\beta_\tau^1 m_\sigma^2 (2c_A)^{-1} \right\}^{1/2} \leq \beta_\tau^3 \tag{A.26}$$

where $\beta_\tau^3 \to 0$ as $\tau \to 0$ (independent of t).

Combining (A.21), (A.25) and (A.26), one has

$$|\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon| \leq \beta_\tau^1 e^{-c_A t} |x| + \beta_\tau^3 \|u^\varepsilon\|_{L_2(0,t)}. \tag{A.27}$$

Now, by the system structure given by Assumption (A7.1I) and by the fact that the W^m are in $\mathcal{C}_\delta$, one obtains the following lemmas exactly as in Chapter 3. These are also analogous to their counterparts in the proof of Theorem 7.27 above.

Lemma A.5 *For any $t < \infty$,*

$$\|u^\varepsilon\|_{L_2(0,t)}^2 \leq \frac{\varepsilon}{\delta} + \frac{1}{\delta} \left[\frac{c_A \gamma^2}{m_\sigma^2} e^{-c_A N \tau} + \frac{c_D}{c_A} \right] |x|^2.$$

Lemma A.6 *For any $t < \infty$,*

$$\int_0^t |\xi_r^\varepsilon|^2 \, dt \leq \frac{\varepsilon}{\delta} \frac{m_\sigma^2}{c_A} + \frac{m_\sigma^2}{\delta} \left[\left(\frac{c_D}{c_A^2} + \frac{\gamma^2}{m_\sigma^2} \right) + \frac{1}{c_A} \right] |x|^2.$$

Let $c_1 \doteq \varepsilon/\delta$ and $c_2 \doteq \frac{1}{\delta}[\frac{c_A\gamma^2}{m_\sigma^2} + \frac{c_D}{c_A}]$. By Lemma A.5 and (A.27), for all $t < \infty$ one has

$$|\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon| \leq \beta_\tau^1 e^{-c_A t} |x| + \beta_\tau^3 (c_1 + c_2 |x|^2)^{1/2}$$

and by proper choice of β_τ^4,

$$\leq \beta_\tau^4 (1 + |x|) \tag{A.28}$$

where $\beta_\tau^4 \to 0$ as $\tau \to 0$ (independent of $t > 0$).

Now,

$$\int_0^T l^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\xi_T^\varepsilon)$$
$$- \int_0^T l^{\overline{\mu}_t^\varepsilon}(\overline{\xi}_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\overline{\xi}_T^\varepsilon)$$
$$= \int_0^T \left[\xi_t^\varepsilon D^{\mu_t^\varepsilon}\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon D^{\overline{\mu}_t^\varepsilon}\overline{\xi}_t^\varepsilon\right]\,dt + (\xi_T^\varepsilon)^T P^m \xi_T^\varepsilon - (\overline{\xi}_T^\varepsilon)^T P^m \overline{\xi}_T^\varepsilon. \quad \text{(A.29)}$$

Note that the integral term on the right-hand side in (A.29) is

$$\int_0^T (\xi_t^\varepsilon)^T D^{\mu_t^\varepsilon}(\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon) + (\xi_t^\varepsilon)^T\left(D^{\mu_t^\varepsilon} - D^{\overline{\mu}_t^\varepsilon}\right)\overline{\xi}_t^\varepsilon + (\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon)^T D^{\overline{\mu}_t^\varepsilon}\overline{\xi}_t^\varepsilon\;dt$$
$$\leq \beta_\tau^4(1+|x|)\int_0^T \left(|D^{\mu_t^\varepsilon}|\,|\xi_t^\varepsilon| + |D^{\overline{\mu}_t^\varepsilon}|\,|\overline{\xi}_t^\varepsilon|\right)\,dt + \beta_\tau^5\int_0^T |\xi_t^\varepsilon|\,|\overline{\xi}_t^\varepsilon|\,dt$$

for appropriate $\beta_\tau^5 \to 0$ as $\tau \to 0$, which after some work

$$\leq \beta_\tau^6(1+|x|^2)(1+\sqrt{T}) \qquad \text{(A.30)}$$

for appropriate choice of $\beta_\tau^6 \to 0$ as $\tau \to 0$ (independent of T).

Similarly, the last two terms on the right-hand side in (A.29) are

$$\begin{aligned}{\xi_T^\varepsilon}^T P^m \xi_T^\varepsilon - {\overline{\xi}_T^\varepsilon}^T P^m \overline{\xi}_T^\varepsilon &= (\xi_T^\varepsilon + \overline{\xi}_T^\varepsilon)^T P^m(\xi_T^\varepsilon - \overline{\xi}_T^\varepsilon)\\ &\leq |P^m|\left[|\xi_T^\varepsilon - \overline{\xi}_T^\varepsilon|^2 + 2|\xi_T^\varepsilon|\,|\xi_T^\varepsilon - \overline{\xi}_T^\varepsilon|\right],\end{aligned}$$

which by (A.28)

$$\leq \beta_\tau^7(1+|x|^2) + \beta_\tau^8|\xi_T^\varepsilon|(1+|x|) \qquad \text{(A.31)}$$

for some $\beta_\tau^7, \beta_\tau^8 \to 0$ as $\tau \to 0$.

We also need the following lemma which is obtained in Chapter 3 as equation (3.58).

Lemma A.7 *Given $\overline{T} < \infty$, there exist $T \in [\overline{T}/2, \overline{T}]$ and ε-optimal $u^\varepsilon \in \mathcal{W}$, $\mu^\varepsilon \in \mathcal{D}_\infty$ for $\widetilde{S}_T[W^m]$ such that*

$$|\xi_T^\varepsilon|^2 \leq \frac{1}{\overline{T}}\left\{\frac{\varepsilon}{\delta}\frac{m_\sigma^2}{c_A} + \frac{m_\sigma^2}{\delta}\left[\left(\frac{c_D}{c_A^2} + \frac{\gamma^2}{m_\sigma^2}\right) + \frac{1}{c_A}\right]|x|^2\right\}.$$

Combining (A.31) and Lemma A.7, one finds that

$${\xi_T^\varepsilon}^T P^m \xi_T^\varepsilon - {\overline{\xi}_T^\varepsilon}^T P^m \overline{\xi}_T^\varepsilon \leq \beta_\tau^9(1+|x|^2) \qquad \text{(A.32)}$$

for some $\beta_\tau^9 \to 0$ as $\tau \to 0$ (independent of T).

Combining (A.29), (A.30) and (A.32),

$$\int_0^T l^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\xi_T^\varepsilon) - \int_0^T l^{\overline{\mu}_t^\varepsilon}(\overline{\xi}_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\overline{\xi}_T^\varepsilon)$$
$$\leq \beta_\tau^{10}(1+|x|^2)(1+\sqrt{T}) \tag{A.33}$$

for some $\beta_\tau^{10} \to 0$ as $\tau \to 0$ (independent of T).

Combining (A.20) and (A.33), one has

$$\widetilde{S}_T[W^m](x) - \int_0^T l^{\overline{\mu}_t^\varepsilon}(\overline{\xi}_t^\varepsilon) - \frac{\gamma^2}{2}|u_t^\varepsilon|^2\,dt + W^m(\overline{\xi}_T^\varepsilon)$$
$$\leq \frac{\hat{\varepsilon}}{2}(1+|x|^2) + \beta_\tau^{10}(1+|x|^2)(1+\sqrt{T}),$$

which for τ sufficiently small (depending on T now),

$$\leq \hat{\varepsilon}(1+|x|^2).$$

This completes the proof of Lemma 7.31. □

A.0.6 Existence of Robust/H$_\infty$ Estimator and a Disturbance Bound

The following theorem (with accompanying proof) supports the existence and representation claim for the robust/H$_\infty$ estimator of Chapter 8.

Theorem A.8 *Let state ξ dynamics be given by (8.1), and let observation process y. be given by (8.2). Assume (A8.1)–(A8.5), and suppose that there exists $\overline{x} \in \mathbf{R}^n$ such that*

$$\phi(x) \geq |x - \overline{x}|^2 \quad \forall\, x. \tag{A.34}$$

Let $T \in (0,\infty)$. There exist γ, ζ, η such that there exists an estimator $\hat{e}_T$ such that (8.8) holds for all possible $w.$, $v.$ and initial conditions ξ_0. Further, such an estimator is given by (8.9).

The proof of Theorem A.8 will begin with a lemma. We suppose throughout this section that ϕ satisfies (A.34).

Lemma A.9 *Let P be given by (8.6), (8.7). Let $T \in [0,\infty)$. There exists $\overline{x}_T \in \mathbf{R}^n$ such that*

$$P(T, x_T) \leq -M_T^P |x_T - \overline{x}_T|^2 \quad \forall\, x_T \in \mathbf{R}^n,$$

where

$$M_T^P = \frac{\zeta^2}{2(C_T^1)^2(1 + m_\sigma^2 T\zeta^2/\gamma^2)}.$$

Proof. Let the global Lipschitz constant for f again be K_f. Let $\overline{x}_T \in \mathbf{R}^n$ be such that $\bar{\xi}_0 = \overline{x}$ (with $\overline{x}$ given in (A.34)) when $\bar{\xi}_\cdot$ satisfies (8.1) with $w \equiv 0$ and terminal condition $\bar{\xi}_T = \overline{x}_T$. The existence of $\overline{x}_T$ follows from the fact that (8.4) as a mapping between x_T and ξ_0 for any specific $w \in L_2$, including $w \equiv 0$, is a bijection. Let $\xi_\cdot$ satisfy (8.1) with any disturbance, $w \in \mathcal{W}$, and any terminal condition $\xi_T = x_T$. Let $x_0 \doteq \xi_0$. Then

$$\begin{aligned}|\xi_t - \bar{\xi}_t| &= \left|(x_0 - \overline{x}) + \int_0^t [f(\xi_r) - f(\bar{\xi}_r) + \sigma(\xi_r)w_r]\,dr\right| \\ &\leq |x_0 - \overline{x}| + \int_0^t [K_f|\xi_r - \bar{\xi}_r| + m_\sigma|w_r|]\,dr,\end{aligned}$$

which by Cauchy–Schwarz

$$\leq |x_0 - \overline{x}| + m_\sigma\sqrt{T}\|w\|_{L_2} + \int_0^t K_f|\xi_r - \bar{\xi}_r|\,dr. \tag{A.35}$$

Employing (A.35) and Gronwall's inequality yields

$$|x_T - \overline{x}_T| \leq (|x_0 - \overline{x}| + m_\sigma\sqrt{T}\|w\|_{L_2})C_T^1,$$

where $C_T^1 \doteq 1 + K_f T e^{K_f T}$. Consequently,

$$\frac{\zeta^2}{2}|x_0 - \overline{x}_0|^2 + \frac{\gamma^2}{2}\|w\|_{L_2}^2 \geq \frac{\zeta^2}{2}\left[\frac{|x_T - \overline{x}_T|}{C_T^1} - m_\sigma\sqrt{T}\|w\|_{L_2}\right]^2 + \frac{\gamma^2}{2}\|w\|_{L_2}^2,$$

which after some calculations one can show

$$\geq \frac{\zeta^2}{2(C_T^1)^2(1 + m_\sigma^2 T\zeta^2/\gamma^2)}|x_T - \overline{x}_T|^2.$$

Therefore, by (8.6)

$$J_f(T, x_T, w) \leq -\frac{\zeta^2}{2(C_T^1)^2(1 + m_\sigma^2 T\zeta^2/\gamma^2)}|x_T - \overline{x}_T|^2.$$

Because this is true for all $w \in \mathcal{W}$, one has the desired result. □

We now proceed to prove Theorem A.8.

Proof. Define

$$I(T, e, x_T, w) = J_f(T, x_T, w) + |x_T - e|^2,$$

which by (8.5)

$$= -\frac{\zeta^2}{2}\phi(\xi_0) - \frac{\gamma^2}{2}\int_0^T |w(t)|^2\,dt - \frac{\eta^2}{2}\sum_{i=1}^{N_T}|v_i|^2 + |x_T - e|^2, \tag{A.36}$$

where ξ_0 is given by (8.4). Let

$$W(T,e) = \sup_{x_T \in \mathbf{R}^n} \sup_{w \in \mathcal{W}} I(T,e,x_T,w), \tag{A.37}$$

and note that

$$W(T,e) = \sup_{x_T \in \mathbf{R}^n} \left[P(T,x_T) + |x_T - e|^2 \right],$$

which by Lemma A.9

$$\leq \sup_{x_T \in \mathbf{R}^n} \left[-M_T^P |x_T - \overline{x}_T|^2 + |x_T - e|^2 \right].$$

Note that one can choose $\zeta, \eta, \gamma < \infty$ such that $M_T^P > 1$, and consequently the supremum on the right-hand side is finite and given by $(M_T^P/(M_T^P - 1))|\overline{x}_T - e|^2$. Thus

$$W(T,e) \leq \frac{M_T^P}{M_T^P - 1} |\overline{x}_T - e|^2. \tag{A.38}$$

Note that, as a function of e, $W(T,e)$ is a supremum of strictly convex functions, and so $W(T,\cdot)$ is strictly convex (as well as going to infinity as $|e| \to \infty$). Consequently, there exists a unique minimizer. Let

$$\hat{e}_T \doteq \operatorname*{argmin}_{e \in \mathbf{R}^n} W(T,e),$$

which is exactly the estimator given by (8.9). Now,

$$W(T,\hat{e}_T) \leq W(T,\overline{x}_T),$$

which by (A.38)

$$= 0.$$

Combining this with (A.37) and (A.36) yields

$$\sup_{x_T \in \mathbf{R}^n} \sup_{w \in \mathcal{W}} \left\{ -\frac{\zeta^2}{2} \phi(\xi_0) - \frac{\gamma^2}{2} \int_0^T |w(t)|^2 \, dt - \frac{\eta^2}{2} \sum_{i=1}^{N_T} |v_i|^2 + |x_T - e|^2 \right\} \leq 0$$

which implies (8.8). □

The following lemma can be used with the machinery of Chapter 3 to show that information state P is the unique viscosity solution of the corresponding HJB PDE.

Lemma A.10 *Let P be given by (8.6), (8.7). Let $\varepsilon \in (0,1]$, $T \in (0,\infty)$ and $x_T \in \mathbf{R}^n$. There exists $M_T^w < \infty$ such that any ε-optimal w satisfies*

$$\|w\|^2 \leq M_T^w (1 + |x_T|^2). \tag{A.39}$$

Proof. Let $w^0(t) = 0$ for all $t \in [0,T]$. Let ξ^0 satisfy (8.1) with disturbance w^0 and terminal condition $\xi^0_T = x_T$. By Assumption (A8.1), there exists $K_f < \infty$ such that $|f(x)| \le K_f(1+|x|)$ for all x. Consequently,

$$\frac{d}{dt}|\xi^0_t|^2 \ge -2K_f|\xi^0|^2 - 2K_f|\xi^0| \ge -4K_f|xi^0|^2 - (K_f/2),$$

which implies

$$|\xi^0_t|^2 \le e^{4K_fT}|x_T|^2 + \frac{1}{8}e^{4K_fT} \quad \forall\, t \in [0,T]. \tag{A.40}$$

This implies

$$J_f(T,x_T,w^0) \ge -\frac{\zeta^2}{2}m_\phi\left[1+|\xi^0_0|^2\right] - \eta^2 m_\rho^2 \sum_{i=1}^{N_T}\left[|y_i|^2 + |h(\xi^0_{t_i})|^2\right],$$

which noting that, by Assumption (A8.1), there exists $K_h < \infty$ such that $|h(x)| \le K_h(1+|x|)$ for all x

$$\begin{aligned} &\ge -\frac{\zeta^2}{2}m_\phi\left[1+|\xi^0_0|^2\right] - \eta^2 m_\rho^2 \sum_{i=1}^{N_T}\left[|y_i|^2 + 2K_h^2|\xi^0_{t_i}|^2\right] \\ &\quad -2\eta^2 m_\rho^2 K_h^2 N_T. \end{aligned} \tag{A.41}$$

Combining (A.40) and (A.41), one sees that there exists $C_T < \infty$ such that

$$J_f(T,x_T,w^0) \ge -C_T(1+|x_T|^2). \tag{A.42}$$

On the other hand, for any $w \in \mathcal{W}$,

$$J_f(T,x_T,w) \le -\frac{\gamma^2}{2}\|w\|^2. \tag{A.43}$$

Combining (A.42) and (A.43) yields the result. □

References

1. M. Akian, Densities of idempotent measures and large deviations, Trans. Amer. Math. Soc. **351** (1999) 4515–4543.
2. M. Akian, S. Gaubert, and V. Kolokoltsov, Set coverings and invertibility of functional Galois connections, *Idempotent Mathematics and Mathematical Physics,* G.L. Litvinov and V.P. Maslov (eds.), Contemporary Math., **377** (2005) 19–52.
3. M. Akian, S. Gaubert, and A. Lakhoua, A max-plus finite element method for solving finite horizon determinsitic optimal control problems, Proc. Math. Theory of Networks and Systems (2004).
4. M. Akian, S. Gaubert, and A. Lakhoua, The max-plus finite element method for optimal control problems: further approximation results, Proc. IEEE Conf. Dec. Control/European Control Conf., Seville 2005 (to appear).
5. M. Akian, J.-P. Quadrat, and M. Viot, Bellman processes, In: G. Cohen and J.-P. Quadrat (eds.), *Lecture Notes in Control and Information Science, Vol. 199,* Springer-Verlag, New York (1994).
6. F.L. Baccelli, G. Cohen, G.J. Olsder, and J.-P. Quadrat, *Synchronization and Linearity,* John Wiley, New York (1992).
7. J.A. Ball and J.W. Helton, H_∞ control for nonlinear plants: connections with differential games, Proc. 28^{th} IEEE Conf. Dec. Control (1989) 956–962.
8. M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations,* Birkhäuser, Boston (1997).
9. M. Bardi and F. Da Lio, On the Bellman equation for some unbounded control problems, Nonlinear Differential Equations Appl., **4** (1997) 491–510.
10. G. Barles and B. Perthame, Exit time problems in optimal control and vanishing viscosity solutions of Hamilton–Jacobi equations, SIAM J. Control Optim., **26** (1988) 1133–1148.
11. G. Barles, *Solutions de Viscosité des Équations de Hamilton–Jacobi,* Mathématiques et Applications de la SMAI, **17**, Springer-Verlag, New York (1994).
12. T. Basar and P. Bernhard, *H_∞ -Optimal Control and Related Minimax Design Problems,* Birkhäuser, Boston (1991).
13. T. Basar and G.J. Olsder, *Dynamic Noncooperative Game Theory,* Academic Press, Boston (1982). Reissued by SIAM, Philadelphia (1999).
14. R. Bellman, *Dynamic Programming,* Princeton University Press, Princeton (1957).

15. D.P. Bertsekas, *Dynamic Programming, Deterministic and Stochastic Models,* Prentice-Hall, Englewood Cliffs, NJ (1987).
16. R.K. Boel, M.R. James, and I.R. Petersen, Robustness and risk sensitive filtering, IEEE Trans. Auto. Control, **47** (2002) 451–461.
17. P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control,* Birkhäuser, Boston (2004).
18. E. Carlini, M. Falcone, and R. Ferretti, An efficient algorithm for Hamilton–Jacobi equations in high dimensions, Computing and Visualization Science, **7** (2004) 15–29.
19. F.H. Clarke, *Optimization and Nonsmooth Analysis,* John Wiley, NewYork (1983).
20. G. Cohen, S. Gaubert, and J.-P. Quadrat, Duality and separation theorems in idempotent semimodules, Linear Algebra and Applications, **379** (2004) 395–422.
21. G. Collins and W.M. McEneaney, Min-plus eigenvector methods for nonlinear H_∞ problems with active control, Optimal Control, Stabilization and Nonsmooth Analysis, In: M.S. de Queiroz, M. Malisoff, P. Wolenski, Peter (eds.), *Lecture Notes in Control and Information Sciences, Vol. 301*, Springer-Verlag, New York (2004) 101–120.
22. M.G. Crandall, H. Ishii, and P.-L. Lions, Uniqueness of viscosity solutions of Hamilton–Jacobi equations revisited, J. Math. Society Japan, **39** (1987) 581–595.
23. M.G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations, Trans. of the Amer. Math. Soc., **277** (1983) 1–42.
24. M.G. Crandall and P.-L. Lions, On existence and uniqueness of solutions of Hamilton–Jacobi equations, Nonlinear Anal. Theory and Appl., **10** (1986) 353–370.
25. R.A. Cuninghame-Green, *Minimax Algebra, Lecture Notes in Economics and Mathematical Systems, Vol. 166,* Springer-Verlag, New York (1979).
26. F. Da Lio and W.M. McEneaney, Finite time-horizon risk-sensitive control and the robust limit under a quadratic growth assumption, SIAM J. Control and Optim. **40** (2001) 1628–1661.
27. M.V. Day, On Lagrange manifolds and viscosity solutions, J. Math. Systems, Estimation and Control, **8** (1998).
28. P. Del Moral and M. Doisy, Maslov idempotent probability calculus, Theory Probability Appl. **43** (1999) 562–576.
29. G. Didinsky, T. Basar, and P. Bernhard, Structural properties of minimax policies of a class of differential games arising in nonlinear H_∞ -control and filtering, Proc. 32^{nd} IEEE CDC (1993).
30. P.M. Dower and M.R. James, Dissipativity and nonlinear systems with finite power gain, Int. J. Robust and Nonlinear Contr., **8** (1998) 699–724.
31. P.M. Dower and M.R. James, Worst case power generating capabilities of nonlinear systems, Math. Contr. Sig. Sys., **15** (2002) 13–41.
32. P. Dupuis and M. Boué, Markov chain approximations for deterministic control problems with affine dynamics and quadratic cost in the control, SIAM J. on Numerical Analysis, **36** (1999) 667–695.
33. P. Dupuis and W. M. McEneaney, Risk-sensitive and robust escape criteria, SIAM J. Control and Optim., **35** (1997) 2021–2049.

34. P. Dupuis and A. Szpiro, Convergence of the optimal feedback policies in a numerical method for a class of deterministic optimal control problems, SIAM J. Control and Optim. **40** (2001) 393–420.
35. R. J. Elliott and N. J. Kalton, The existence of value in differential games, Memoirs of the Amer. Math. Society, **126** (1972).
36. R.J Elliott *Viscosity Solutions and Optimal Control,* John Wiley, New York (1987).
37. L.C. Evans, *Partial Differential Equations,* Springer-Verlag, New York (1998).
38. L.C. Evans and P.E. Souganidis, Differential games and representation formulas for solutions of Hamilton–Jacobi–Isaacs equations, Indiana Univ. Math. J., **33** (1984) 773–797.
39. W.H. Fleming, Max-plus stochastic processes, Applied Math. and Optim. **48** (2004) 159–181.
40. W.H. Fleming, Max-plus stochastic control, In: B. Pasik-Duncan (ed.), *Stochastic Theory and Control, Lecture Notes in Control and Information Sciences, Vol. 280,* Springer-Verlag, New York (2002).
41. W.H. Fleming, Deterministic nonlinear filtering, Annali Scuola Normale Superiore Pisa, Cl. Scienze Fisiche e Matematiche, Ser. IV, **25** (1997) 435–454.
42. W.H. Fleming, *Functions of Several Variables,* Springer-Verlag, New York (1977).
43. W.H. Fleming and W.M. McEneaney, Robust limits of risk sensitive filters, Math. Control, Signals and Systems, **14** (2001) 109–142.
44. W.H. Fleming and W.M. McEneaney, A max-plus based algorithm for an HJB equation of nonlinear filtering, SIAM J. Control and Optim., **38** (2000) 683–710.
45. W.H. Fleming and W.M. McEneaney, Risk sensitive and robust nonlinear filtering, 36^{th} IEEE Conf. on Decision and Control (1997) 1088–1093.
46. W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control,* Springer-Verlag, New York (1975).
47. W.H. Fleming and H.M. Soner, *Controlled Markov Process and Viscosity Solutions, Second Edition,* Springer-Verlag, New York (2006).
48. W.H. Fleming and H.M. Soner, *Controlled Markov Process and Viscosity Solutions,* Springer-Verlag, New York (1993).
49. F. John, *Partial Differential Equations,* Springer-Verlag, New York (1978).
50. E. Gallestey, M.R. James, and W.M. McEneaney, Max-plus approximation methods in partially observed H_∞ control, 38^{th} IEEE Conf. on Decision and Control, 3011–3016.
51. M. Hardt, J.W. Helton, and K. Kreutz-Delgado, Numerical solution of nonlinear H_2 and H_∞ control problems with application to jet engine compressors, Control Systems Technology (2000).
52. J.W. Helton and M.R. James, *Extending H_∞ Control to Nonlinear Systems,* SIAM, Philadelphia (1999).
53. J.-B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms I: Fundamentals, Second Edition* Springer-Verlag, New York (1993).
54. R.A. Horn and C.R. Johnson, *Matrix Analysis,* Cambridge Univ. Press, Cambridge (1985).
55. S. Huang, M.R. James, D. Nesic, and P.M. Dower, Analysis of input to state stability for discrete time nonlinear systems via dynamic programming, Automatica (to appear).
56. S. Huang, M.R. James, D. Nesic, and P.M. Dower, A unified approach to controller design for achieving ISS and related properties, IEEE Trans. Auto. Control (to appear).

57. H. Ishii, Uniqueness of unbounded viscosity solution of Hamilton–Jacobi equations, Indiana Univ. Math. J., **33** (1984) 721–748.
58. H. Ishii, Representation of solutions of Hamilton–Jacobi equations, Nonlinear Anal. Theory and Appl., **12** (1988) 121–146.
59. H. Ishii, Comparison results for Hamilton–Jacobi equations without growth condition on solutions from above, Appl. Anal., **67** (1997) 357–372.
60. M.R. James, A partial differential inequality for dissipative nonlinear systems, Systems and Control Letters, **21** (1993) 315–320.
61. M.R. James and J.S. Baras, Partially observed differential games, infinite-dimensional Hamilton–Jacobi–Isaacs equations, and nonlinear H_∞ control, SIAM J. Control and Optim., **34** (1996) 1342–1364.
62. M.R. James and S. Huang, L_∞ bounded robustness for nonlinear systems: analysis and synthesis, Submitted to IEEE Trans. Autom. Control (2002).
63. V.N. Kolokoltsov and V.P. Maslov, *Idempotent Analysis and Its Applications,* Kluwer (1997).
64. A.J. Krener, Necessary and sufficient conditions for worst case (H_∞) control and estimation, J. Math., Systems, Estimation and Control, **4** (1994) 485–488.
65. N.V. Krylov, *Controlled Diffusion Processes,* Springer-Verlag, New York (1980).
66. H.J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time,* Springer-Verlag, New York (1992).
67. A. Lakhoua, Résolution numérique de problèmes de commande optimale déterministe et algèbre max-plus, Rapport de stage DEA, INRIA, Rocquencort (2003).
68. P.-L. Lions and P.E. Souganidis, Differential games, optimal control and directional derivates of viscosity solutions of Bellman's and Isaacs' equations, SIAM J. Control Optim., **23** (1985) 566–583.
69. G.L. Litvinov and V.P. Maslov, Correspondence principle for idempotent calculus and some computer applications, *Idempotency,* J. Gunawardena ed., Publ. Newton Inst. **11**, Cambridge Univ. Press, Cambridge (1998) 420–443.
70. G.L. Litvinov, V.P. Maslov and G.B. Shpiz, Idempotent functional analysis: an algebraic approach, Mathematical Notes, **69**, No. 5 (2001) 696–729.
71. V.P. Maslov, On a new principle of superposition for optimization problems, Russian Math. Surveys, **42** (1987) 43–54.
72. D. McCaffrey and S.P. Banks, Lagrangian manifolds, viscosity solutions and Maslov index, J. Convex Analysis, **9**, No. 1 (2002) 185–224.
73. W.M. McEneaney, Legendre transforms on max-plus spaces as a tool for nonlinear control problems, Proc. Math. Theory of Networks and Systems (2004).
74. W.M. McEneaney, Max-plus summation of Fenchel-transformed semigroups for solution of nonlinear Bellman equations, Submitted to Systems and Control Letters.
75. W.M. McEneaney, Max-plus eigenvector representations for solution of nonlinear H_∞ problems: error analysis, SIAM J. Control and Optim., **43** (2004) 379–412.
76. W.M. McEneaney and P.M. Dower, A max-plus affine power method for approximation of a class of mixed l_∞/l_2 value functions, Proc. 42^{nd} IEEE Conf. on Dec. and Control, Maui (2003) 2573–2578.
77. W.M. McEneaney, Max-plus methods for nonlinear H_∞ control: operating in the transform space, Proc. 41^{st} IFAC Symposium on Robust Control Design, Milan (2003).

78. W.M. McEneaney, Max-plus eigenvector representations for solution of nonlinear H_∞ problems: basic concepts, IEEE Trans. Auto. Control, **48** (2003) 1150–1163.
79. W.M. McEneaney Error analysis of a max-plus algorithm for a first-order HJB equation, In: B. Pasik-Duncan (ed.), *Stochastic Theory and Control, Lecture Notes in Control and Information Sciences, Vol. 280,* Springer-Verlag, New York (2002).
80. W.M. McEneaney, Error analysis of a max-plus algorithm for a first-order HJB equation, Proc. Workshop On Max-Plus Algebras and Their Applications to Discrete-event Systems, Theoretical Computer Science, and Optimization, Prague, 27–29, August 2001.
81. W.M. McEneaney, Convergence and error analysis for a max-plus algorithm, Proc. 39^{th} IEEE Conf. on Decision and Control (2000) 1194–1199.
82. W.M. McEneaney and C.D. Charalambous, Large deviations theory, induced log-plus and max-plus measures and their applications, Proc. Math. Theory of Networks and Systems (2000).
83. W.M. McEneaney, Robust/game-theoretic methods in filtering and estimation, Proc. Symposium on Advances in Enterprise Control (1999) 3–10.
84. W.M. McEneaney, Exactly linearizing algebras for risk-sensitive filtering, 38^{th} IEEE Conf. on Decision and Control (1999) 137–142.
85. W.M. McEneaney and M. Horton, Computation of max-plus eigenvector representations for nonlinear H_∞ value functions, Proc. American Control Conference (1999) 1400–1404.
86. W.M. McEneaney and M. Horton, Max-plus eigenvector representations for nonlinear H_∞ value functions, Proc. 37^{th} IEEE Conf. on Decision and Control (1998) 3506–3511.
87. W.M. McEneaney, Robust/H_∞ filtering for nonlinear systems, Systems and Control Letters, **33** (1998) 315–325.
88. W.M. McEneaney, A uniqueness result for the Isaacs equation corresponding to nonlinear H_∞ control, Math. Controls, Signals and Systems, **11** (1998) 303–334.
89. W.M. McEneaney, Elimination of troublesome disturbances with application to representation results for H_∞ control DPEs, Proc. Seventh International Symposium on Dynamic Games and Applications, Shonan Village, Japan (1996) 662–671.
90. W.M. McEneaney and M.V. Day, Characteristic characterization of viscosity supersolutions corresponding to nonlinear H_∞ control, Proc. 13^{th} World Congress of International Federation of Automatic Control (1996) 401–406.
91. W.M. McEneaney, Robust control and differential games on a finite time horizon, Math. Controls, Signals and Systems, **8** (1995) 138–166.
92. W.M. McEneaney, Uniqueness for viscosity solutions of nonstationary Hamilton–Jacobi–Bellman equations under some a priori conditions (with applications), SIAM J. Control Optim., **33** (1995) 1560–1576.
93. A.A. Melikyan, *Generalized characteristics of first-order PDEs: Applications in Optimal Control and Differential Games,* Birkhäuser, Boston (1998).
94. R.K. Miller and A.N. Michel *Ordinary Differential Equations,* Academic Press, Boston (1982).
95. R.E. Mortensen, Maximum likelihood recursive nonlinear filtering, J. Optim. Theory Appl., **2** (1968) 386–394.
96. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations,* Springer-Verlag, New York (1983).

97. A.A. Puhalskii, Large deviations of semimartingales: a maxingale approach, Stochastic and Stochastics Reports. Part I, **61** (1997) 141–243; Part II, **68** (1999) 65–143.
98. A.A. Puhalskii, *Large Deviations and Idempotent Probability,* Chapman and Hall/CRC Press, Boca Raton (2001).
99. J.-P. Quadrat, Min-plus probability calculus, Actes 26 eme École de Printemps d'Informatique Theorique, Noirmoutier (1998).
100. R.T. Rockafellar, *Convex Analysis,* Princeton University Press, Princeton (1970).
101. R.T. Rockafellar, *Conjugate Duality and Optimization,* Regional conference Series in Applied Mathematics 16, SIAM (1974).
102. R.T. Rockafellar and R.J. Wets, *Variational Analysis,* Springer-Verlag, New York (1997).
103. B.L. Rozovskii, S. Lototsky, and R. Mikulevicius, Nonlinear filtering revisited: a spectral approach, SIAM J. Control Optim., **35** (1997) 435–461.
104. H.L. Royden, *Real Analysis,* 2nd Ed., Macmillan (1968).
105. E.D. Sontag and Y. Wang, New characterisations of input to state stability, IEEE Trans. Autom. Contr., **41** (1996) 1283–1294.
106. P. Soravia, H_∞ control of nonlinear systems: differential games and viscosity solutions, SIAM J. Control and Optim., **34** (1996) 1071–1097.
107. A.I. Subbotin, *Generalized Solutions of First-Order PDE's: The Dynamical Optimization Perspective,* Birkhäuser, Boston (1995).
108. A.J. van der Schaft, L_2–Gain analysis of nonlinear systems and nonlinear state feedback H_∞ control, IEEE Trans. Auto. Control, **37** (1992) 770–784.
109. A.J. van der Schaft, L_2*-Gain and Passitivity Techniques in Nonlinear Control,* Springer-Verlag, New York (1996).

Index

Systems & Control: Foundations & Applications

Series Editor
Tamer Başar
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
1308 W. Main St.
Urbana, IL 61801-2307
U.S.A.

Systems & Control: Foundations & Applications

Aims and Scope

The aim of this series is to publish top quality state-of-the art books and research monographs at the graduate and post-graduate levels in systems, control, and related fields. Both foundations and applications will be covered, with the latter spanning the gamut of areas from information technology (particularly communication networks) to biotechnology (particularly mathematical biology) and economics.

Readership

The books in this series are intended primarily for mathematically oriented engineers, scientists, and economists at the graduate and post-graduate levels.

Types of Books

Advanced books, graduate-level textbooks, and research monographs on current and emerging topics in systems, control and related fields.

Preparation of manuscripts is preferable in LaTeX. The publisher will supply a macro package and examples of implementation for all types of manuscripts.

Proposals should be sent directly to the editor or to: Birkhäuser Boston, 675 Massachusetts Avenue, Cambridge, MA 02139, U.S.A. or to Birkhäuser Publishers, 40-44 Viadukstrasse, CH-4051 Basel, Switzerland

A Partial Listing of Books Published in the Series

Identification and Stochastic Adaptive Control
Han-Fu Chen and Lei Guo

Viability Theory
Jean-Pierre Aubin

Representation and Control of Infinite Dimensional Systems, Vol. I
A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter

Representation and Control of Infinite Dimensional Systems, Vol. II
A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter

Mathematical Control Theory: An Introduction
Jerzy Zabczyk

H_∞-Control for Distributed Parameter Systems: A State-Space Approach
Bert van Keulen

Disease Dynamics
Alexander Asachenkov, Guri Marchuk, Ronald Mohler, and Serge Zuev

Theory of Chattering Control with Applications to Astronautics, Robotics, Economics, and Engineering
Michail I. Zelikin and Vladimir F. Borisov

Modeling, Analysis and Control of Dynamic Elastic Multi-Link Structures
J. E. Lagnese, Günter Leugering, and E. J. P. G. Schmidt

First-Order Representations of Linear Systems
Margreet Kuijper

Hierarchical Decision Making in Stochastic Manufacturing Systems
Suresh P. Sethi and Qing Zhang

Optimal Control Theory for Infinite Dimensional Systems
Xunjing Li and Jiongmin Yong

Generalized Solutions of First-Order PDEs: The Dynamical Optimization Perspective
Andreĭ I. Subbotin

Finite Horizon H_∞ and Related Control Problems
M. B. Subrahmanyam

Control Under Lack of Information
A. N. Krasovskii and N. N. Krasovskii

H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach
Tamer Başar and Pierre Bernhard

Control of Uncertain Sampled-Data Systems
Geir E. Dullerud

Robust Nonlinear Control Design: State-Space and Lyapunov Techniques
Randy A. Freeman and Petar V. Kokotović

Adaptive Systems: An Introduction
Iven Mareels and Jan Willem Polderman

Sampling in Digital Signal Processing and Control
Arie Feuer and Graham C. Goodwin

Ellipsoidal Calculus for Estimation and Control
Alexander Kurzhanski and István Vályi

Minimum Entropy Control for Time-Varying Systems
Marc A. Peters and Pablo A. Iglesias

Chain-Scattering Approach to H^∞-Control
Hidenori Kimura

Output Regulation of Uncertain Nonlinear Systems
Christopher I. Byrnes, Francesco Delli Priscoli, and Alberto Isidori

High Performance Control
Teng-Tiow Tay, Iven Mareels, and John B. Moore

Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations
Martino Bardi and Italo Capuzzo-Dolcetta

Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming
William M. McEneaney, G. George Yin, and Qing Zhang, Editors

Mutational and Morphological Analysis: Tools for Shape Evolution and Morphogenesis
Jean-Pierre Aubin

Stabilization of Linear Systems
Vasile Dragan and Aristide Halanay

The Dynamics of Control
Fritz Colonius and Wolfgang Kliemann

Optimal Control
Richard Vinter

Advances in Mathematical Systems Theory:
A Volume in Honor of Diederich Hinrichsen
Fritz Colonius, Uwe Helmke, Dieter Prätzel-Wolters,
and Fabian Wirth, Editors

Nonlinear and Robust Control of PDE Systems:
Methods and Applications to Transport-Reaction Processes
Panagiotis D. Christofides

Foundations of Deterministic and Stochastic Control
Jon H. Davis

Partially Observable Linear Systems Under Dependent Noises
Agamirza E. Bashirov

Switching in Systems and Control
Daniel Liberzon

Matrix Riccati Equations in Control and Systems Theory
Hisham Abou-Kandil, Gerhard Freiling, Vlad Ionescu, and Gerhard Jank

The Mathematics of Internet Congestion Control
Rayadurgam Srirkant

H^∞ Engineering and Amplifier Optimization
Jeffery C. Allen

Advances in Control, Communication Networks, and Transportation Systems:
In Honor of Pravin Varaiya
Eyad H. Abed

Convex Functional Analysis
Andrew J. Kurdila and Michael Zabarankin

Max-Plus Methods for Nonlinear Control and Estimation
William M. McEneaney

Uniform Output Regulation of Nonlinear Systems:
A Convergent Dynamics Approach
Alexey Pavlov, Nathan van de Wouw, and Henk Nijmeijer

William M. McEneaney

Max-Plus Methods for Nonlinear Control and Estimation

The central focus of this book is the control of continuous-time/continuous-space nonlinear systems. Using new techniques that employ the max-plus algebra, the author addresses several classes of nonlinear control problems, including nonlinear optimal control problems and nonlinear robust/H_∞ control and estimation problems. Several numerical techniques are employed, including a max-plus eigenvector approach and an approach that avoids the curse-of-dimensionality.

Well-known dynamic programming arguments show there is a direct relationship between the solution of a control problem and the solution of a corresponding Hamilton–Jacobi–Bellman (HJB) partial differential equation (PDE). The max-plus-based methods examined in this monograph belong to an entirely new class of numerical methods for the solution of nonlinear control problems and their associated HJB PDEs; they are not equivalent to either of the more commonly used finite element or characteristic approaches. The potential advantages of the max-plus-based approaches lie in the fact that solution operators for nonlinear HJB problems are linear over the max-plus algebra, and this linearity is exploited in the construction of algorithms.

The book will be of interest to applied mathematicians, engineers, and graduate students interested in the control of nonlinear systems through the implementation of recently developed numerical methods. Researchers and practitioners tangentially interested in this area will also find a readable, concise discussion of the subject through a careful selection of specific chapters and sections. Basic knowledge of control theory for systems with dynamics governed by differential equations is required.

ISBN 0-8176-3534-3
EAN
9 780817 635343

Birkhäuser
ISBN 0-8176-3534-3
www.birkhauser.com