Volume 6, Number 3, 2006 ISSN 1562-8353

NONLINEAR DYNAMICS AND SYSTEMS THEORY
An International Journal of Research and Surveys

Volume 6 Number 3 2006

CONTENTS

 ISSN 1562-8353 Printed in Ukraine
To receive contents and abstracts by e-mail, visit our Website at: **http://www.e-ndst.kiev.ua**

NONLINEAR DYNAMICS & SYSTEMS THEORY Volume 6, No. 3, 2006

Nonlinear Dynamics and Systems Theory

An International Journal of Research and Surveys

Academic Periodical Press

Nonlinear Dynamics and Systems Theory
An International Journal of Research and Surveys

NONLINEAR DYNAMICS AND SYSTEMS THEORY
An International Journal of Research and Surveys

INSTRUCTIONS FOR CONTRIBUTORS

(1) General. The Journal will publish original carefully refereed papers, brief notes and reviews on a wide range of nonlinear dynamics and systems theory problems. Contributions will be considered for publication in ND&ST if they have not been published previously. Before preparing your submission, it is essential that you consult our style guide; please visit our website: http://www.e-ndst.kiev.ua

(2) Manuscript and Correspondence. Contributions are welcome from all countries and should be written in English. Two copies of the manuscript, double spaced one column format, and the electronic version by AMSTEX, TEX or LATEX program (on diskette) should be sent directly to

Professor A.A. Martynyuk
Institute of Mechanics,
Nesterov str.3, 03057, MSP 680
Kiev-57, Ukraine
(e-mail: anmart@stability.kiev.ua).

or to one of the Editors or to a member of the Editorial Board.

The title of the article must include: author(s) name, name of institution, department, address, FAX, and e-mail; an Abstract of 50-100 words should not include any formulas and citations; key words, and AMS subject classifications number(s). The size for regular paper should be 10-14 pages, survey (up to 24 pages), short papers, letter to the editor and book reviews (2-3 pages).

(3) Tables, Graphs and Illustrations. All figures must be suitable for reproduction without being retouched or redrawn and must include a title. Line drawings should include all relevant details and should be drawn in black ink on plain white drawing paper. In addition to a hard copy of the artwork, it is necessary to attach a PC diskette with files of the artwork (preferably in PCX format).

(4) References. Each entry must be cited in the text by author(s) and number or by number alone. All references should be listed in their alphabetic order. Use please the following style:

Journal: [1] Poincaré, H. Title of the article. *Title of the Journal* **Vol.1**(No.1) (year) pages. [Language].

Book: [2] Liapunov, A.M. *Title of the book.* Name of the Publishers, Town, year.

Proceeding: [3] Bellman, R. Title of the article. In: *Title of the book.* (Eds.). Name of the Publishers, Town, year, pages. [Language].

(5) Proofs and Reprints. Proofs sent to authors should be returned to the Editor with corrections within three days after receipt. Acceptance of the paper entitles the author to 10 free reprints.

(6) Editorial Policy. Every paper is reviewed by the regional editor, and/or a referee, and it may be returned for revision or rejected if considered unsuitable for publication.

(7) Copyright Assignment. When a paper is accepted for publication, author(s) will be requested to sign a form assigning copyright to Informath Publishing Group. Failure to do it promptly may delay the publication.

NONLINEAR DYNAMICS AND SYSTEMS THEORY
An International Journal of Research and Surveys
Published since 2001

Volume 6 Number 3 2006

CONTENTS

Founded by A.A.Martynyuk in 2001.
Registered in Ukraine Number: KB №5267 / 04.07.2001.

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys

Nonlinear Dynamics and Systems Theory (ISSN 1562-8353 (Print), ISSN 1813-7385 (Online)) is an international journal published under the auspices of the S.P.Timoshenko Institute of Mechanics of National Academy of Sciences of Ukraine and the Laboratory for Industrial and Applied Mathematics (LIAM) at York University (Toronto, Canada). It is aimed at publishing high quality original scientific papers and surveys in area of nonlinear dynamics and systems theory and technical reports on solving practical problems. The scope of the journal is very broad covering:

SCOPE OF THE JOURNAL

Analysis of uncertain systems • Bifurcations and instability in dynamical behaviors • Celestial mechanics, variable mass processes, rockets • Control of chaotic systems • Controllability, observability, and structural properties • Deterministic and random vibrations • Differential games • Dynamical systems on manifolds • Dynamics of systems of particles • Hamilton and Lagrange equations • Hysteresis • Identification and adaptive control of stochastic systems • Modeling of real phenomena by ODE, FDE and PDE • Nonlinear boundary problems • Nonlinear control systems, guided systems • Nonlinear dynamics in biological systems • Nonlinear fluid dynamics • Nonlinear oscillations and waves • Nonlinear stability in continuum mechanics • Non-smooth dynamical systems with impacts or discontinuities • Numerical methods and simulation • Optimal control and applications • Qualitative analysis of systems with aftereffect • Robustness, sensitivity and disturbance rejection • Soft computing: artificial intelligence, neural networks, fuzzy logic, genetic algorithms, etc. • Stability of discrete systems • Stability of impulsive systems • Stability of large-scale power systems • Stability of linear and nonlinear control systems • Stochastic approximation and optimization • Symmetries and conservation laws

PUBLICATION AND SUBSCRIPTION INFORMATION

The **Nonlinear Dynamics and Systems Theory** is published four times per year in 2006. Base list subscription price per volume: US$149.00. This price is available only to individuals whose library subscribes to the journal OR who warrant that the Journal is for their own use and provide a home address for mailing. Separate rates apply to academic and corporate/government institutions. Our charge includes postage, packing, handling and airmail delivery of all issues. Mail order and inquires to: Department of Processes Stability, S.P.Timoshenko Institute of Mechanics NAS of Ukraine, Nesterov str.,3, 03057, Kiev-57, MSP 680, Ukraine, Tel: ++38-044-456-6140, Fax: ++38-044-456-0319, E-mail: anchern@stability.kiev.ua, http://www.sciencearea.com.ua; http://www.e-ndst.kiev.ua

ABSTRACTING AND INDEXING SERVICES

EBSCO Databases, Swets Information Services, Mathematical Reviews/MathSciNet, Zentralblatt MATH/Mathematics Abstracts.

Preface

Many theoretical results and methodologies developed for systems sciences and optimization are now found very useful in dealing with nonlinear dynamics and system theory as well as their high technology applications. These areas of research are interdisciplinary in nature with great potentials for high technology applications. In view of this the Guest Editors had made a call for high quality papers to be submitted to this special issue, where system science and optimization approaches are to be used in dealing with topics in nonlinear dynamics and system theory as well as their high technology applications. This is therefore the theme of this Special Issue:

System Science and Optimization Approaches to Nonlinear Dynamics and System Theory with High Technology Applications (1)

With this aim in mind, the goal of the special issue is to provide an international forum for scientists, researchers, and practitioners from both academia and industry to present their latest research findings and state-of-the-art solution methods in areas related to the theme of the Special Issue.

Scientists from many countries and regions — Australia, China, Greece, Hong Kong, Japan, India, Saudi Arabia, USA and Vietnam — accepted the invitation of the Guest Editors to submit papers for the Special Issue of the Journal. They all went through a rigorous refereeing process with at least two independent referees for each submitted paper. The number of the submitted papers exceed substantially the size of one issue, and we decided to publish two special issues. Topics included in these papers are modelling, design analysis, simulation, optimization, performance evaluation, intelligent information and technology, nonlinear stochastic systems, and optimal control. Applications involved include communication networks, engineering and management systems, computer and information technology, and knowledge management.

The completion of this volume would not have been possible without the assistance of many of our colleagues. We wish to express our sincere appreciation to all those who helped. We are deeply grateful to our referees who provided prompt and extensive reviews for all submissions. Their constructive comments contributed to the quality of the volume. In particular, we wish to thank Editor-in-Chief, Professor Anatolii A. Martynyuk for his kind cooperation and support. Our special thank also go to Mrs. Lisa Holling for her help during the editing process of this Special Issue. Last but not least, we wish to thank those authors who responded to our call for papers by submitting their papers to be considered for possible publication in this Special Issue.

Wuyi Yue[1] and Kok Lay Teo[2] – Guest Editors

[1]Department of Information Science and Systems Engineering, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan. E-mail: yue@konan-u.ac.jp

[2]Department of Mathematics and Statistics, Curtin University of Technology, GPO Box U1987, Perth, Western Australia 6845, Australia. E-mail: K.L.Teo@curtin.edu.au

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys
Published since 2001

System Science and Optimization Approaches to Nonlinear Dynamics and System Theory with High Technology Applications (2)

Special Issue

CONTENTS *)

*) The issue is in preparation.

Nonlinear Dynamics and Systems Theory, **6**(3) (2006) 211–224

Mixed Semidefinite and Second-Order Cone Optimization Approach for the Hankel Matrix Approximation Problem

Mohammed M. Alshahrani [1*] and Suliman S. Al-Homidan [2]

[1] *Department of Mathematics, Dammam Teachers' College, P.O. Box 14262, Dammam 31424, Saudi Arabia.*
[2] *Department of Mathematical Sciences, King Fahad University of Petroleum and Minerals, KFUPM Box 119, Dhahran 31261, Saudi Arabia.*

Received: April 1, 2005; Revised: July 15, 2006

Abstract: Approximating the nearest positive semidefinite Hankel matrix in the Frobenius norm to an arbitrary data covariance matrix is useful in many areas of engineering, including signal processing and control theory. In this paper, interior point primal-dual path-following method will be used to solve our problem after reformulating it into different forms, first as a semidefinite programming problem, then into the form of a mixed semidefintie and second-order cone optimization problem. Numerical results, comparing the performance of these methods with the modified alternating projection method will be reported.

Keywords: *Hankel matrix; primal-dual interior-point method; projection method; semidefinite programming.*

Mathematics Subject Classification (2000): 49J35, 49M99.

1 Introduction

In some application areas, such as digital signal processing and control theory, it is required to compute the closest, in some sense, positive semidefinite Hankel matrix, with no restriction on its rank, to a given data covariance matrix, computed from a data sequence. This problem was studied by Macinnes [19]. Similar problems involving structured covariance estimation were discussed in [16, 13, 24]. Related problems occur in many engineering and statistics applications [10].

* Corresponding author: mmogib@awalnet.net.sa

The problem was formulated as a nonlinear minimization problem with positive semidefinite Hankel matrix as constraints in [2] and then was solved by l_2 Sequential Quadratic Programming (l_2 SQP) method. Another approach to deal with this problem was to solve it as a smooth unconstrained minimization problem [1, 3]. Other methods to solve this problem or similar problems can be found in [19, 13, 16].

Our work is mainly casting the problem: first as a semidefinite programming problem and second as a mixed semidefinite and second-order cone optimization problem. A semidefinite programming (SDP) problem is to minimize a linear objective function subject to constraints over the cone of positive semidefinite matrices. It is a relatively new field of mathematical programming, and most of the papers on SDP were written in 1990s, although its roots can be traced back to a few decades earlier (see Bellman and Fan [8]). SDP problems are of great interest due to many reasons, e.g., SDP contains important classes of problems as special cases, such as linear and quadratic programming. Applications of SDP exist in combinatorial optimization, approximation theory, system and control theory, and mechanical and electrical engineering. SDP problems can be solved very efficiently in polynomial time by interior point algorithms [29, 32, 11, 6, 21].

The constraints in a mixed semidefinite and second-order cone optimization problem are constraints over the positive semidefinite and the second-order cones. Although the second-order cone constraints can be seen as positive semidefinite constraints, recent research has shown that it is more effecient to deal with mixed problems rather than the semidefinite programming problem. Nesterov et al. [21] can be considered as the first paper to deal with mixed semidefinite and second-order cone optimization problems. However, the area was really brought to life by Alizadeh et al. [5] with the introduction of SDPPack, a software package for solving optimization problems from this class. The practical importance of second-order programming was demonstrated by Lobo et al. [18] and many subsequent papers. In [22] Sturm presented implementational issues of interior point methods for mixed SDP and SOCP problems in a unified framework. One class of these interior point methods is the primal-dual path-following methods. These methods are considered the most successful interior point algorithms for linear programming. Their extension from linear to semidefinite and then mixed problems has followed the same trends. One of the successful implementation of primal-dual path-following methods is in the software SDPT3 by Toh et al. [28, 25].

Similar problems, such as the problem of minimizing the spectral norm of a matrix was first formulated as a semidefinite programming problem in [29, 26]. Then, these problems and some others were formulated as a mixed semidefinite and second-order cone optimization problems [18, 4, 23]. None of these formulations exploits the special structure of our problem. For the purpose of exploiting the Hankel structure of the variable in this problem we will introduce an isometry operator, **hvec**, taking $n \times n$ Hankel matrices into $2n-1$ vectors. We will see later that using this operator gives our formulations an advantage over the others.

Before we go any further, we should introduce some notations. Throughout this paper, we will denote the set of all $n \times n$ real symmetric matrices by $\mathcal{S}_n$, the cone of the $n \times n$ real symmetric positive semidefinite matrices by $\mathcal{S}_n^+$ and the second-order cone of dimension k by $\mathcal{Q}_k$, and is defined as

$$\mathcal{Q}_k = \{x \in R^k : \|x_{2:k}\|_2 \leq x_1\},$$

(also called Lorentz cone, ice cream cone or quadratic cone), where $\|.\|_2$ stands for the Euclidean distance norm defined as $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$, $\forall x \in R^n$ and $x_{2:k} =$

$[x_2, x_2, \dots, x_k]^T$. The set of all $n \times n$ real Hankel matrices will be denoted by $\mathcal{H}_n$. An $n \times n$ real Hankel matrix $H(h)$ has the following structure:

$$H(h) = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \\ h_2 & h_3 & \cdots & h_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_n & h_{n+1} & \cdots & h_{2n-1} \end{bmatrix}, \quad h \in R^{2n-1}.$$

It is clear that $\mathcal{H}_n \subset \mathcal{S}_n$. The Frobenius norm is defined on $\mathcal{S}_n$ as follows:

$$\|U\|_F = \sqrt{U \bullet U} = \|\mathbf{vec}^T(U)\mathbf{vec}(U)\|_2, \ \forall \ U \in \mathcal{S}_n. \tag{1.1}$$

Here $U \bullet U = \text{trace}(U \cdot U) = \sum_{i,j}^{n} U_{i,j}^2$, $\mathbf{vec}(U)$ stands for the vectorization operator found by stacking the columns of U together and $\mathbf{vec}^T$ is the transpose of $\mathbf{vec}$. The symbols $\succeq$ and $\geq_Q$ will be used to denote the partial orders induced by $\mathcal{S}_n^+$ and $\mathcal{Q}_k$ on $\mathcal{S}_n$ and R^k, respectively. That is,

$$U \succeq V \quad \Leftrightarrow \quad U - V \in \mathcal{S}_n^+, \ \ \forall \ U, \ V \in \mathcal{S}_n$$

and

$$u \geq_Q v \Leftrightarrow \quad u - v \in \ \mathcal{Q}_k, \ \ \forall \ u, \ v \ \in R^k.$$

The statement $x \geq 0$ for a vector $x \in R^n$ means that each component of x is nonnegative. We use I and 0 for the identity and zero matrices.

Our problem in mathematical notation can, now, be formulated as follows: Given a data matrix $F \in R^{n \times n}$, find the nearest positive semidefinite Hankel matrix $H(h)$ to F such that $\|F - H(h)\|_F$ is minimal. Thus, we have the following optimization problem:

$$\begin{array}{ll} \text{minimize} & \|F - H(h)\|_F \\ \text{subject to} & H(h) \in \mathcal{H}_n, \quad H(h) \succeq 0. \end{array} \tag{1.2}$$

We describe briefly the alternating projection method. Although the rate of convergence is slow, the method converges to the optimal solution globally, and provides us with accurate solutions against which we can compare the results obtained by this method with those of the interior point methods. Henc, we devote Section 2 to the projection method. A brief description of semidefinite and second-order cone optimization problems along with reformulations of problem (1.2) in the form of the respective class will be given in Sections 3 and 4, respectively. Numerical results, showing the performance of the projection method against the primal-dual path-following method acting on our formulations, will be reported in Section 5. Section 6 contains the paper's conclusions.

2 The Projection Method

The method of successive cyclic projections onto closed subspaces C_i's was first proposed by von Neumann [30] and independently by Wiener [31]. As a special case of their algorithm, we show that if C_1 and C_2 are subspaces and D is a given point, then the nearest point to D in $C_1 \cap C_2$ can be obtained by the following algorithm:

Alternating Projection Algorithm.

Let $X_1 = D$.

For $k = 1, 2, 3, \ldots$

$X_{k+1} = P_1(P_2(X_k))$.

X_k converges to the nearest point to D in $C_1 \cap C_2$, where P_1 and P_2 are the orthogonal projections on C_1 and C_2, respectively. Dykstra [12] and Boyle and Dykstra [9] modified von Neumann's algorithm to handle the situation when C_1 and C_2 are replaced by convex sets. Other proofs and connections to duality along with applications were given in Han [17]. These modifications were applied in [15] to find the nearest Euclidean distance matrix to a given data matrix. The modified Neumann's algorithm when applied to (1.2) yields the following algorithm, called the Modified Alternating Projection Algorithm: Given a data matrix F, we have:

Let $F_1 = F$.

For $j = 1, 2, 3, \ldots$

$F_{j+1} = F_j + [P_S(P_H(F_j)) - P_H(F_j)]$.

Then $\{P_H(F_j)\}$ and $P_S(P_H(F_j))$ converge in Frobenius norm to the solution. Here, $P_H(F)$ is the orthogonal projection onto the subspace of Hankel matrices $\mathcal{H}_n$. It is simply setting each antidiagonal to be the average of the corresponding antidiagonal of F. $P_S(F)$ is the projection of F onto the convex cone of positive semidefinite symmetric matrices. One finds $P_S(F)$ by finding a spectral decomposition of F and setting the negative eigenvalues to zero.

3 Semidefinite Programming Approach

The semidefinite programming (SDP) problem in primal standard form is:

$$\begin{array}{lll} (P) & \text{minimize}_X & C \bullet X \\ & \text{subject to} & A_i \bullet X = b_i, \quad i = 1, 2, \cdots, m, \quad X \succeq 0, \end{array} \tag{3.1}$$

where all $A_i,\ C \in \mathcal{S}_n,\ b \in R^m$ are given, and $X \in \mathcal{S}_n$ is the variable. This optimization problem (3.1) is a convex optimization problem since its objective and constraint are convex. The dual problem of (3.1) is

$$\begin{array}{lll} (D) & \text{maximize}_y & b^T y \\ & \text{subject to} & \sum_{i=1}^{m} y_i A_i \preceq C, \end{array} \tag{3.2}$$

where $y \in R^m$ is the variable. Problems (3.1) and (3.2) include many problems as special cases and have many applications, in particular, (1.2). The following theorem is useful in writing (1.2) in the form of (3.1).

Theorem 3.1 (Schur Complement) *If*

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where $A \in \mathcal{S}_n^+$ nonsingular matrix and $C \in \mathcal{S}_n$, then the matrix M is positive (semi)definite if and only if the matrix $C - B^T A^{-1} B$ is positive (semi)definite.

This matrix $C - B^T A^{-1} B$ is called the Schur complement of A in M. Letting $\|F - H(h)\|_F^2 \leq t$, t is a nonnegative real scalar and noting that:

$$\|F - H(h)\|_F^2 = \mathbf{vec}^T(F - H(h))\mathbf{vec}(F - H(h)),$$

we have:

$$\begin{aligned} & \mathbf{vec}^T(F - H(h))\mathbf{vec}(F - H(h)) \leq t \\ \Leftrightarrow \quad & t - \mathbf{vec}^T(F - H(h)) I \mathbf{vec}(F - H(h)) \geq 0 \\ \Leftrightarrow \quad & \begin{bmatrix} I & \mathbf{vec}(F - H(h)) \\ \mathbf{vec}^T(F - H(h)) & t \end{bmatrix} \succeq 0. \end{aligned}$$

The last equivalence is a direct application of Theorem 3.1. Thus, problem (1.2) can be rewritten as

$$\begin{aligned} (SDV) \qquad & \text{minimize} \quad t, \\ & \text{subject to} \quad \begin{bmatrix} t & 0 & 0 \\ 0 & H(h) & 0 \\ 0 & 0 & V \end{bmatrix} \succeq 0, \end{aligned} \tag{3.3}$$

where

$$V = \begin{bmatrix} I & \mathbf{vec}(F - H(h)) \\ \mathbf{vec}^T(F - H(h)) & t \end{bmatrix},$$

which is an SDP problem in the dual form (3.2) with dimensions $2n$ (number of variables) and n^2+n+2 (size of the matrices), SDP problem (3.3) is very large even for a small data matrix F. For example, a 50×50 matrix F will give rise to a problem with dimensions 100 and 2552, hence solving (1.2) using formulation (3.3) is not efficeint. Furthermore, we do not exploit the structure of $H(h)$ being Hankel. Another way of formulation that produces an SDP problem with reasonable dimensions and exploits the Hankel structure of $H(h)$ can be done by means of the following isometry operator.

Definition 3.1 Let $\mathbf{hvec} : \mathcal{H}_n \longrightarrow R^{2n-1}$ be defined as

$$\mathbf{hvec}(U) = [u_{1,1} \ \sqrt{2}u_{1,2} \ \cdots \ \sqrt{n-1}u_{1,n-1} \ \sqrt{n}u_{1,n} \ \sqrt{n-1}u_{2,n} \ \cdots \ \sqrt{2}u_{n-1,n} \ u_{n,n}]^T$$

for any $U \in \mathcal{H}_n$.

It is clear that **hvec** is a linear operator from the set of all $n \times n$ real Hankel matrices to R^{2n-1}. The following theorem gives us some characterizations of **hvec**.

Theorem 3.2 *For the operator* ***hvec****, defined in Definition 3.1, the following conditions hold: For any* $U, V \in \mathcal{H}_n$
1. $U \bullet U = \boldsymbol{hvec}^T(U)\boldsymbol{hvec}(U)$.
2. $\|U - V\|_F^2 = \boldsymbol{hvec}^T(U - V)\boldsymbol{hvec}(U - V)$.

Proof Part 1 is clear from the definition of the **hvec** operator. Part 2 is a consequence of part 1. □

Part 1 implies that **hvec** is an isometry. We cannot take any advantage of this theorem unless F is Hankel. Projecting F onto $\mathcal{H}_n$ using the orthogonal projection in Section 2 gets a Hankel matrix, say $\hat{F}$. The following proposition shows that the nearest Hankel positive semidefinite matrix to $\hat{F}$ is also the nearest to F.

Proposition 3.1 *Let $\hat{F}$ be the orthogonal projection of F onto $\mathcal{H}_n$ and let $H(h)$ be the nearest Hankel positive semidefinite matrix to $\hat{F}$, then $H(h)$ is so for F.*

Proof If $\hat{F}$ is positive semidefinite, then we are done. If not, then for any $T \in \mathcal{H}_n$, we have $(F-\hat{F})^T \bullet (\hat{F}-T) = 0$ since $\hat{F}$ is the orthogonal projection of F. Thus, $\|F-T\|_F^2 = \|F-\hat{F}\|_F^2 + \|\hat{F}-T\|_F^2$. □

As a consequence of this proposition, the following problem is equivalent to (1.2):

$$\begin{array}{ll} \text{minimize} & \|\hat{F}-H(h)\|_F \\ \text{subject to} & H(h) \in \mathcal{H}_n, \quad H(h) \succeq 0. \end{array} \tag{3.4}$$

3.1 Formulation I (SDH)

From Theorem 3.1, the following are equivalences (for $t \geq 0 \in R$):

$$\begin{array}{ll} & \|\hat{F}-H(h)\|_F^2 \leq t \\ \Leftrightarrow & \mathbf{hvec}^T(\hat{F}-H(h))\mathbf{hvec}(\hat{F}-H(h)) \leq t \quad by \ \ Theorem\ 3.2 \\ \Leftrightarrow & t - \mathbf{hvec}^T(\hat{F}-H(h))I\mathbf{hvec}(\hat{F}-H(h)) \leq 0 \\ \Leftrightarrow & \begin{bmatrix} I & \mathbf{hvec}(\hat{F}-H(h)) \\ \mathbf{hvec}^T(\hat{F}-H(h)) & t \end{bmatrix} \succeq 0 \quad by \ \ Theorem\ 3.1 \end{array}$$

Hence, we have the following SDP problem:

$$\begin{array}{lll} (SDH) & \text{minimize} & t, \\ & \text{subject to} & \begin{bmatrix} t & 0 & 0 \\ 0 & H(h) & 0 \\ 0 & 0 & \hat{V} \end{bmatrix} \succeq 0, \end{array} \tag{3.5}$$

where

$$\hat{V} = \begin{bmatrix} I & \mathbf{hvec}(\hat{F}-H(h)) \\ \mathbf{hvec}^T(\hat{F}-H(h)) & t \end{bmatrix}.$$

This SDP problem has dimensions $2n$ and $3n+1$ which is far better than (3.3).

3.2 Formulation II (SDQ)

Another way for formulating (1.2) is through the definition of the Frobenius norm:

$$\|F-H(h)\|_F^2 = y^T P y + 2q^T y + r,$$

where

$$y = [h_1 \ h_2 \ \cdots \ h_{2n-1}]^T, \quad P = \text{diag}([1 \ 2 \ \cdots \ n \ \cdots 2 \ 1]),$$
$$q_k = -\sum\nolimits^{n}_{\substack{i,j=1 \\ i+j=k+1}} F(i,j), \quad k = 1, 2, \cdots 2n-1 \ \text{ and } \ r = \|F\|_F^2.$$

Now, we have for a nonnegative real scalar t

$$\begin{aligned} & \|F-H\|_F^2 \le t \\ \Leftrightarrow\quad & y^T P y + 2q^T y + r \le t \\ \Leftrightarrow\quad & (P^{1/2}y)^T(P^{1/2}y) + 2q^T y + r \le t \\ \Leftrightarrow\quad & t - 2q^T y - r - (P^{1/2}y)^T I (P^{1/2}y) \ge 0 \\ \Leftrightarrow\quad & \begin{bmatrix} I & (P^{1/2}y) \\ (P^{1/2}y)^T & t-2q^T y - r \end{bmatrix} \succeq 0. \end{aligned}$$

Hence, we have the following SDP problem:

$$\begin{aligned} (SDQ) \qquad & \text{minimize} \quad t, \\ & \text{subject to} \quad \begin{bmatrix} t & 0 & 0 \\ 0 & H(h) & 0 \\ 0 & 0 & Q \end{bmatrix} \succeq 0, \end{aligned} \tag{3.6}$$

where

$$Q = \begin{bmatrix} I & (P^{1/2}y) \\ (P^{1/2}y)^T & t-2q^T y - r \end{bmatrix}.$$

This SDP problem is of dimenstions $2n$ and $3n+1$. Although problem (3.6) has the same dimentions as problem (3.5), it is less efficient to solve it over the positive semidefinite cone $\mathcal{S}_n^+$, especially when F is large in size. In practice, as we will see in Section 5, it has been found that the performance of this formulation is poor. The reason for that is the matrix P being of full rank and hence the system is badly conditioned. A more efficient interior point method for this formulation can be developed by using Nesterov and Nemirovsky technique [20] to reformulate it over the second-order cone as described in Section 4.

The last formulation seems to be straight forward, but it was found that using this formulation to solve similar problems was not a good idea. The reasons for that will be discussed in the following section when we talk about second-order cone programming. This fact about SDQ formulation will be clear in Section 5 when we use it to solve numerical examples with $n > 50$. The SDV formulation does not compete favorably with the other two SDH and SDQ formulations due to the amount of work per one iteration of interior-point methods that solve SDV fomulation is $\mathcal{O}(n^6)$, where n in the dimension of F and $\mathcal{O}(.)$ is the order of convergence. The SDV formulation is even slower than the projection method. Hence, using the SDV formulation to solve (1.2) is time consuming. This leaves us with SDH formulation from which we expect good performance, since it does not have the illness of SDQ nor the huge size of SDV.

4 Mixed Semidefinite and Second-Order Cone Approach

The primal mixed semidefinite, second-order and linear problem SQLP is of the form:

$$\begin{aligned} (P') \qquad & \text{minimize} \quad C_S \bullet X_S + C_Q^T X_Q + C_L^T X_L \\ & \text{subject to} \quad (A_S)_i \bullet X_S + (A_Q)_i^T X_Q + (A_L)_i^T X_L = b_i, \quad i = 1,2,\cdots,m \\ & \qquad\qquad\quad X_S \succeq 0, X_Q \ge_Q 0, X_L \ge 0, \end{aligned} \tag{4.1}$$

where $X_S \in \mathcal{S}_n,\ X_Q \in R^k$ and $X_L \in R^{n_L}$ are the variables. $C_S,\ (A_S)_i \in \mathcal{S}_n,\ \forall i$ $C_Q,\ (A_Q)_i \in R^k\ \forall i$ and $C_L,\ (A_L)_i \in R^{n_L}\ \forall i$ are given data. Each of the three inequalities has a different meaning: $X_S \succeq 0$ means $X_S \in \mathcal{S}_n^+$, $X_Q \geq_Q 0$ means that $X_Q \in \mathcal{Q}_k$ and $X_L \geq 0$ means that each component of X_L is nonnegative. It is possible that one or more of the three parts of (4.1) is not present. If the second-order part is not present, then (4.1) reduces to the ordinary SDP (3.1) and if the semidefinite part is not present, then (4.1) reduces to the so-called convex quadratically constrained linear programming problem.

The standard dual of (4.1) is:

$$
\begin{array}{lll}
(D') & \text{maximize} & b^T y \\
 & \text{subject to} & \displaystyle\sum_{i=1}^m y_i (A_S)_i \preceq C_S \\
 & & \displaystyle\sum_{i=1}^m y_i (A_Q)_i \leq_Q C_Q \\
 & & \displaystyle\sum_{i=1}^m y_i (A_L)_i \leq C_L.
\end{array} \tag{4.2}
$$

Here, $y \in R^m$ is the variable.

In our setting, we may drop the third part of the constraints in (4.1) and its dual (4.2), since we do not have explicit linear constraints. One natural claim can be made here: in (1.2) the objective function can be recast as a dual SQLP in three different ways.

4.1 Formulation III (SQV)

One way to define $\|F - H(h)\|_F$ is

$$\|F - H(h)\|_F = \|\mathbf{vec}(F - H(h))\|_2.$$

So, if we put $\|F - H(h)\|_F \leq t$ for $t \in R^+$, then by the definition of the second-order cone, we have

$$\begin{bmatrix} t \\ \mathbf{vec}(F - H(h)) \end{bmatrix} \in \mathcal{Q}_{1+n^2}.$$

Hence, we have the following reformulation of (1.2):

$$
\begin{array}{lll}
(SQV) & \text{minimize} & t, \\
 & \text{subject to} & \begin{bmatrix} t & 0 \\ 0 & H(h) \end{bmatrix} \succeq 0 \\
 & & \begin{bmatrix} t \\ \mathbf{vec}(F - H(h)) \end{bmatrix} \geq_Q 0.
\end{array} \tag{4.3}
$$

4.2 Formulation IV (SQQ)

The second definition is as introduced in Subsection 3.2, i.e.,

$$\|F - H(h)\|_F^2 \quad = y^T P y + 2q^T y + r.$$

Hence, we have the following equivalent problem to (1.2)

$$\begin{array}{ll} \text{minimize} & y^T P y + 2q^T y + r \\ \text{subject to} & H(h) \in \mathcal{H}_n, \quad H(h) \succeq 0. \end{array} \tag{4.4}$$

But

$$y^T P y + 2q^T y + r = \|P^{1/2}y + P^{-1/2}q\|_2^2 + r - q^T P^{-1} q.$$

Now, we minimize $\|F - H(h)\|_F^2$ by minimizing $\|P^{1/2}y + P^{-1/2}q\|_2$. Thus we have the following problem:

$$\begin{array}{lll} (SQQ) & \text{minimize} & t, \\ & \text{subject to} & \begin{bmatrix} t & 0 \\ 0 & H(h) \end{bmatrix} \succeq 0 \\ & & \begin{bmatrix} t \\ P^{1/2}y + P^{-1/2}q \end{bmatrix} \geq_Q 0, \end{array} \tag{4.5}$$

where $t \in R^+$. Again, this problem is in the form of problem (4.2). Here, the difference between this form and SQV is in the second-order cone constraint since the SDP part is the same as SQV. The dimension of the second-order cone in SQV is $1 + n^2$ and in SQQ is just $2n$, which makes us expect less efficiency in practice when we work with SQV. The optimal value of SQV is the same as that of problem (1.2), whereas the optimal values of SQQ (4.5) and (4.4) are equal up to a constant. The optimal value of (4.4) is equal to $(\rho^*)^2 + r - q^T P^{-1} q$, where ρ^* is the optimal value of (4.5). It may be noticed that we did not talk about the constraint of $H(h)$ being Hankel. This is because the Hankel structure of $H(h)$ is embedded in the other constraints.

4.3 Formulation V (SQH)

The last formulation will take advantage of the Hankel structure of $H(h)$ explicitly. The vectorization operator **hvec** on Hankel matrices, introduced in Section 3 will be used. From Theorem 3.2, we have the following:

$$\|\hat{F} - H(h)\|_F = \|\mathbf{hvec}(\hat{F} - H(h))\|_2,$$

where $\hat{F} = P_H(F)$ which leads to:

$$\begin{array}{lll} (SQH) & \text{minimize} & t, \\ & \text{subject to} & \begin{bmatrix} t & 0 \\ 0 & H(h) \end{bmatrix} \succeq 0 \\ & & \begin{bmatrix} t \\ \mathbf{hvec}(\hat{F} - H(h)) \end{bmatrix} \geq_Q 0. \end{array} \tag{4.6}$$

The dimension of the second-order cone in this form is $2n$, the same as that of SQQ. The optimal solutions of (4.6) and (1.2) are also identical.

Table 4.1 shows the dimensions of the semidefinite part (SD part) and the second-order cone part (SOC part) for each formulation. For the formulations SDH and SDQ, the second-order cone part is not applicable, so the cell in the table corresponding to that is left blank.

Formulation	SD part	SOC part
SDV	$2n \times (n^2 + n + 2)$	
SDH	$2n \times (3n + 1)$	
SDQ	$2n \times (3n + 1)$	
SQV	$2n \times (n + 1)$	$n^2 + 1$
SQQ	$2n \times (n + 1)$	$2n$
SQH	$2n \times (n + 1)$	$2n$

Table 4.1: Problem dimensions

In practise, we expect that the mixed formulations are more effecient than the SDP-only formulations, especially the SQQ and SQH which have second-order cone constraint of least dimension. The interior point methods for SOCP have better worst-case complexity than an SDP method. However, SDH has a less SDP dimension with no illness such as that SDQ has, which makes SDH a better choice among other SDP. This is due to the economical vectorization operator **hvec**. Practical experiments show a competitive behaviour of SDH to SQQ and SQH (see Section 5).

5 Numerical Results

We will now present some numerical results comparing the performance of the methods described in Sections 2, 3 and 4. The first is the projection method and the second is the interior-point primal-dual path-following method employing the NT-direction. The latter was used to solve five different formulations of the problem.

A Matlab code was written to implement the modified alternating projection method. The iteration is stopped when $\|P_S(P_H(F_j)) - P_H(F_j)\|_F \leq 10^{-8}$.

Size	Time (sec.)					
	Pro.	SDH	SDQ	SQH	SQQ	SQV
10	2	2	1	1	1	1
	9	1	1	1	1	1
30	11	5	4	3	4	2
	14	5	4	2	2	2
50	117	10	12	5	7	5
	30	11	11	4	3	5
100	61	53	64	28	20	28
	1003	48	42	22	25	21
200	16239	389	284	324	322	284
	4883	355	420	255	268	230
400	36556	4970	3913	3775	4098	2505

Table 5.1: Performance comparison (time) among the projection method and the path-following method with the formulations SDH, SDQ, SQH, SQQ and SQV.

For the other methods, the software SDPT3 ver. 3.0 [27, 25] was used because of its numerical stability [14] and its ability to exploit sparsity very efficiently. The default

Size	Iterations					
	Pro.	SDH	SDQ	SQH	SQQ	SQV
10	1253	16	18	14	14	11
	6629	18	17	14	14	11
30	1215	34	32	35	47	24
	1443	33	33	29	29	20
50	4849	32	41	25	36	24
	1295	32	42	22	18	26
100	504	34	45	27	19	26
	8310	33	28	23	26	20
200	22672	31	22	33	31	25
	6592	28	32	23	27	22
400	7870	28	25	26	26	18

Table 5.2: Performance comparison (number of iterations) among the projection method and the path-following method with the formulations SDH, SDQ, SQH, SQQ and SQV.

starting iterates in SDPT3 were used throughout with the NT-direction. The choice of the NT-direction came after some preliminary numerical results. The other direction is HKM-direction which we found less accurate, although, faster than the NT-direction. However, the difference between the two in speed is not of significant importance.

The problem was converted into the five formulations described in Sections 3 and 4. A Matlab code was written for each formulation. This code formulates the problem and passes it through to SDPT3 for a first time. A second run is done with the optimal iterate from the first run being the initial point. This process is repeated until no progress is detected. This is done when the relative gap:

$$\frac{|P-D|}{\max\{1,(|P|+|D|)/2\}}$$

of the current run is the same as the preceding one. (Here, P and D denote the optimal and the dual objective values, respectively).

Our numerical experiments were carried out on eleven randomly generated square matrices with different sizes, namely: 10, 30, 50, 100 and 200, two for each size and one of size 400. Each matrix is dense and its entries vary between -100 and 100 exclusively.

All numerical experiments in this section were executed in Matlab 6.1 on a 1.7GHz Pentium IV PC with 256 MB memory running MS-Windows 2000 Professional.

Table 5.1 compares the CPU time. We notice that the consumed time gets larger more rapidly in the projection method with the size of the data matrix F. An obvious remark is that the projection method is the slowest; it is at least seven times slower than the slowest of the five formulations of the path-following method. However, the difference in time between the five formulations is not big enough to have a significant importance.

Another clear advantage is in terms of number of iterations as shown in Table 5.2. Although the amount of work in each iteration is different for each method, it is still fair to consider it to be a comparison factor.

Table 5.3 shows how close, in Frobenius norm, the optimal solution of each method, $H(h)^*$, to the data matrix F. The projection and the path-following methods with the

Size	Norm					
	Pro.	SDH	SDQ	SQH	SQQ	SQV
10	96.6226	96.6226	96.6226	96.6226	96.6226	96.6226
	94.8320	94.8320	94.8320	94.8320	94.8320	94.8320
30	307.9339	307.9339	307.9406	307.9339	307.9339	307.9339
	327.6784	327.6784	327.6784	327.6784	327.6784	327.6784
50	494.3805	494.3805	494.5038	494.3805	494.3805	494.3805
	497.4383	497.4383	497.6330	497.4383	497.4383	497.4383
100	991.8832	991.8832	994.8612	991.8832	991.8832	991.8833
	997.4993	997.4993	998.8048	997.4993	997.4993	997.4994
200	1986.9397	1986.9398	1990.0924	1986.9402	1986.9402	1986.9414
	1994.8409	1994.8410	1998.6048	1994.8410	1994.8410	1994.8418
400	3998.4967	3998.5047	4001.9242	3998.5007	3998.5007	3998.6166

Table 5.3: Performance comparison (norm $\|H(h)^* - F\|_F$) among the projection method and the path-following method with the formulations SDH, SDQ, SQH, SQQ and SQV.

Size	Error				
	SDH	SDQ	SQH	SQQ	SQV
10	6.3×10^{-9}	3.4×10^{-9}	6.1×10^{-9}	6.1×10^{-9}	1.3×10^{-5}
	6.4×10^{-9}	3.2×10^{-8}	3.6×10^{-8}	3.6×10^{-8}	1.2×10^{-5}
30	7.5×10^{-10}	6.7×10^{-3}	2.6×10^{-8}	3.0×10^{-8}	9.7×10^{-8}
	1.6×10^{-9}	9.0×10^{-9}	2.0×10^{-9}	2.0×10^{-9}	1.2×10^{-8}
50	1.9×10^{-9}	1.2×10^{-1}	8.9×10^{-9}	9.0×10^{-9}	2.1×10^{-5}
	3.7×10^{-9}	0.2	7.8×10^{-9}	8.0×10^{-9}	2.1×10^{-5}
100	5.1×10^{-10}	3.0	1.8×10^{-8}	1.8×10^{-8}	1.0×10^{-4}
	9.2×10^{-10}	1.3	5.8×10^{-8}	5.8×10^{-8}	1.5×10^{-4}
200	6.6×10^{-5}	3.2	4.4×10^{-4}	4.2×10^{-4}	1.6×10^{-3}
	1.1×10^{-4}	3.8	9.1×10^{-5}	9.1×10^{-5}	9.3×10^{-4}
400	8.0×10^{-3}	3.4	4.0×10^{-3}	4.0×10^{-3}	1.2×10^{-1}

Table 5.4: Performance comparison (error).

formulation SDH, SQH and SQQ gave the same result to some extent. The formulation SDQ couldn't cope with the others as the problem size gets larger. The poor performance of this formulation is due to the matrix P being of full rank. The formulation SQV is less accurate than SDH, SQH and SQQ which is reasonable especially if we notice that the dimension of the second-order cone in this formualtion is $1 + n^2$ (see Table 4.1).

Table 5.4 gives a measure of how close the optimal solutions of SDH, SDQ, SQH, SQQ and SQV are from that of the projection method which is the most accurate. The error is computed simply by evaluating the difference between the norm $\|H(h)^* - F\|_F$ of the projection and the norm obtained by the different formulations of the path-following method.

6 Conclusions

The projection method, despite its accuracy, is very slow. On the other hands, the path-following method with SDH and SQQ formulations is very fast, sometimes more than 40 times faster than the projection method (see Table 5.1 when $n = 200$), and gives results of acceptable accuracy. The SQH, SQQ and SQV formulations did not gain any considerable advantage out of solving our problem as a mixed semidefinite and second-order cone problem. This can be seen clearly by noticing the good performance of the formulation SDH, which solves the problem as a semidefinite program. However, it is well known that positive definite Hankel matrices are extremely ill-conditioned; the optimal condition number for these matrices grows exponentially with the size of the matrix [7]. Therefore, computing the spectral decomposition (projection method) or solving the underlying linear systems (SDP/SOCP methods) might be numerically impractical.

References

[1] Al-Homidan, S. Hybrid methods for approximating Hankel matrix. *Numerical Algorithms* **32** (2003) 57–66.

[2] Al-Homidan, S. Combined methods for approximating Hankel matrix. *WSEAS Transactions on Systems* **1** (2002) 35–41.

[3] Al-Homidan, S. Structured Methods for Solving Hankel Matrix Approximation Problems. *Pacific journal of Optimization* **1** (3) (2005) 599–609.

[4] Alizadeh, F. and Goldfarb, D. Second-order cone programming. *Mathematical Programming* **95** (1) (2003) 3–15.

[5] Alizadeh, F. Haeberly, J. A. Nayakkanakuppann, M. V. Overton, M. and Schmieta, S. SDPPack, user's guide, 1997.

[6] Alizadeh, F. Haeberly, J. P. A. and Overton, M. L. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.* **8** (1998) 746–768.

[7] Beckermann, B. The condition number of real Vandermonde, Krylov and positive definite Hankel matrices. *Numer. Math.* **85** (2000) 553–577.

[8] Bellman, R. and Fan, K. On systems of linear inequalities in Hermitian matrix variables. In: *Convexity* (Klee, V. L., ed.). Symposia in Pure Mathematics, Amer. Math. Soc., Providence, RI, **7** (1963) 1–11.

[9] Boyle, J. P. and Dykstra, R. L. A method of finding projections onto the intersection of convex sets in Hilbert space. *Lecture Notes in Statistics* **37** (1986) 28–47.

[10] Burg, J.P. Luenberger, D. G. and Wenger, D. L. Estimation of structured covariance matrices. *Proc. IEEE* **70** (1982) 963–974.

[11] de Klerk, E. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Applied Optimization Series.* Kluwer Academic Publishers **65**, 2002.

[12] Dykstra, R. L. An algorithm for restricted least squares regression. *J. Amer. Stat.* **78** (1983) 839–842.

[13] Fang, W. and Yagle, A.E. Two methods of Toeplitz-plus-Hankel approximation to a data covariance matrix. *IEEE Trans. Signal Processing* **40** (1992) 1490–1498.

[14] Fujisawa, K. Fukuda, M. Kojima, M. and Nakata, K. Numerical evaluation of SDPA (semidefinite programming algorithm). In: *High Performance Optimization* (Frenk, H. Roos, K. Terlakey, T. and Zhang, S., eds.). Kluwer Academic Press, 2000, 267–301.

[15] Glunt, W. Hayden, L. Hong, S. and Wells, L. An alternating projection algorithm for computing the nearest Euclidean distance matrix. *SIAM J. Matrix Anal. Appl.* **11** (4) (1990) 589–600.

[16] Grigoriadis, K. M. Frazho, A. E. and Skelton, R. E. Application of alternating convex projection methods for computing of positive Toeplitz matrices. *IEEE Trans. Signal Processing* **42** (1994) 1873–1875.

[17] Han, S. P. A successive projection method. *Math. Programming* **40** (1988) 1–14.

[18] Lobo, M. Vandenberghe, L. Boyd, S. and Lebret, H. Applications of second-order cone programming. *Linear Algebra and Applications* **284** (1998) 193–228.

[19] Macinnes, C. S. The solution to a structured matrix approximation problem using Grassman coordinates. *SIAM J. Matrix Anal. Appl.* **211** (2) (1999) 446–453.

[20] Nesterov, Y. and Nemirovskii, A. *Interior Point Polynomial Methods in Convex Programming*. SIAM, Philadelphia, 1994.

[21] Nesterov, Y. E. and Todd, M. J. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.* **8** (1998) 324–364.

[22] Sturm, J. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. Technical report, August 2002.

[23] Sturm, J. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software* **14** (1999) 625–653.

[24] Suffridge, Y. J. and Hayden, T. L. Approximation by a Hermitian positive semidefinite Toeplitz matrix. *SIAM J. Matrix Analysis and Appl.* **14** (1993) 721–734.

[25] Tütüncü, R. Toh, K. and Todd, M. Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Program.* **95** (2003) 189–217.

[26] Todd, M. J. Semidefinite Optimization. *Acta Numerica* **10** (2001) 515–560.

[27] Todd, M. J. Toh, K. C. and Tütüncü, R. H. On the Nesterov-Todd direction in semidefinite programming. *SIAM J. Optim.* **8** (1998) 769–796.

[28] Todd, M. J. Toh, K. C. and Tütüncü, R. H. SDPT3 — a Matlab software package for semidefinite programming. *Optim. Methods Softw.* **11** (1999) 545–581.

[29] Vandenberghe L. and Boyd, S. Semidefinite programming. *SIAM Rev.* **38** (1) (1996) 49–95.

[30] von Neumman, J. *Functional Operators II, The geometry of orthogonal spaces.* Annals of Math. studies No.22, Princeton Univ. Press., 1950.

[31] Wiener, N. On factorization of matrices. *Comm. Math. Helv.* **29** (1955) 97–111.

[32] Wolkowicz, H. Saigal, R. and Vandenberghe, L. *Handbook of Semidefinite Programming*. Kluwer Academic Publishers Group, Boston-Dordrecht-London, 2000.

Nonlinear Dynamics and Systems Theory, 6(3) (2006) 225–244

Lagrangian Duality Algorithms for Finding a Global Optimal Solution to Mathematical Programs with Affine Equilibrium Constraints†

Pham Ngoc Anh[1] and Le Dung Muu[2*]

[1] *Posts and Telecommunications Institute of Technology, Hanoi, Vietnam*
[2] *Hanoi Institute of Mathematics, Vietnam*

Received: July 18, 2005; Revised: August 8, 2006

Abstract: Mathematical programs with equilibrium constraints, shortly MPEC, are optimization problems with parametric variational inequality constraints.MPEC include bilevel convex programming problems, mathematical programs with complementarity constraints, Nash-Cournot oligopolistic market models, as well as optimization over the efficient set of an affine fractional multicriteria program as special cases. MPEC are difficult global optimization ones, since their feasible domains, in general, are not convex even not connected. In this paper we consider linear programs with affine equilibrium constraints. We use the Lagrangian duality to compute lower bounds for a decomposition branch-and-bound procedure that allows approximating a global optimal solution of problems in this class of MPEC. Application to optimization over the efficient set of a multicriteria affine fractional program is discussed.

Keywords: *Equilibrium constraints; bilevel convex program; optimization over the Pareto set; Nash-Cournot model; branch-and-bound; global optimum.*

Mathematics Subject Classification (2000): 90C29.

1 Introduction

Mathematical programs with equilibrium (or variational inequality) constraints, shortly MPEC, are optimization problems whose constraints include parametric variational inequalities. For these problems we refer the readers to the comprehensive monograph [16] and the interesting bibliography paper [8]. MPEC play an important role, for example,

† This work was done during the visit of the second author at the University of Karlsruhe, Germany and at ICTP, Italy.

* Corresponding author: ldmuu@math.ac.vn

in the design of transportation networks, in economic models (see e.g. [13, 16]). These problems also include, as special cases, mathematical program with complementarity constraints, the bilevel convex programming problem, where some variables are restricted to be in the solution set of a parametric convex optimization problem, and the optimization over the efficient set of an affine fractional multicriteria program. Mathematically, for finite dimensional case, a mathematical program with equilibrium constraints can be written as

$$\begin{cases} \text{minimize } \ f(x,y) \\ \text{subject to } x \in X, y \in Y, (x,y) \in Z \\ \text{and } y \text{ solves the parametric variational inequality,} \\ \text{find } y \in C(x) \text{ such that } \langle F(x,y), v-y\rangle \geq 0 \ \forall v \in C(x), \end{cases}$$

where $X \subseteq R^n, Y \subseteq R^m, Z \subseteq R^n \times R^m$, are nonempty closed convex sets; $f : X \times Y \to R$, $F : X \times Y \to R^m$ and $C : R^n \to 2^{R^m}$ is a multivalued mapping. The MPEC which are known to be very difficult ones, being nonsmooth and nonconvex also under very favorable assumptions. Further, the computation of the (generalized) gradients of the constraints can be difficult, except special cases.

Several heuristic and deterministic methods were developed for finding local optimal solution to the MPEC. In [13] heuristic algorithms were suggested for solving some classes of MPEC. In [24] Outrata and Zowe converted a mathematical program with equilibrium constraints into an unconstrained nonsmooth Lipschitz optimization problem. Then one may use well developed nonsmooth optimization numerical methods for locally solving the converted problem. In [9] Facchinei et al applied known methods of nonlinear optimization to a regularized reformulation of the MPEC. Based on the study of subanalytic optimization problems and with the help of the theory of error bounds, some exact penalty results for the MPEC were proved by Lin and Fukushima in [15]. Recently, in [19] Mordukhovich discussed optimality conditions for the MPEC and EPEC (equilibrium problems with equilibrium constraints) by using tools of variational analysis.

Due to its nested structure, the feasible domain of a mathematical program with equilibrium constraints, even in the linear case, in general, is disconnected and may be neither open nor closed. Thus the MPEC are difficult global optimization ones and therefore it is less hope to develop an algorithm for finding global optimal solutions to general MPEC. In [21] a branch and bound algorithm based on a binary search is proposed for globally solving a class of mathematical programs with affine equilibrium constraints. The binary search method proposed in that paper works well for the case when the number of constraints defining the variational inequality-constraint is somewhat small, but it quickly becomes expensive when this number gets larger.

In this paper, we continue our work in [21] by using the Lagrangian duality bound to develop another branch-and-bound algorithm for globally solving a class of mathematical programs with affine equilibrium constraints. By contrast to the method in [21] the global optimization operation in this algorithm takes place in the y- space rather than the space of the Lagrangian variables. Thus it is expected that the proposed method works well when the number of the y-variables is somewhat small while the number of the constraints as well as the number of the x-variables may be much larger.

The rest of this paper will be organized as follows. In the next section we state the problem to be solved and list some of its special cases such as bilevel convex programming, optimization over the efficient set and Cournot-Nash oligopolistic market models. In the third section, first we show how to use the Lagrangian duality to compute lower bounds. Then we describe in detail a branch-and-bound algorithm when the vertices of the constrained set Y are known in advance. The last section is devoted to description of a relaxation algorithm that does not require prior knowledge of these vertices. Applications of the proposed algorithms to the optimization problem over the Pareto-efficient set of a multicriteria affine fractional program are also discussed in this section.

2 The Problem Statement and Examples

In what follows we restricted our attention to a special class of MPEC; namely, we consider the following affine MPEC that we call shortly AMPEC:

$$(\mathrm{P}) \quad \begin{cases} \text{minimize } \{f(x,y) := a^T x + b^T y\} \\ \text{subject to } \ x \in X, y \in Y, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \le 0, \} \\ \text{and } y \text{ solves the parametric variational inequality} \\ \text{find } y \in Y \text{ such that } \langle P(y)x + Qy + q, v - y\rangle \ge 0 \ \forall v \in Y \qquad \mathrm{VIP}(x) \end{cases}$$

where $X \subseteq R^n$, $Y \subseteq R^m$ are nonempty closed convex sets, $p \in R^l$, $a \in R^n, q, b \in R^m$ and for each $y \in Y$, $P(y)$, Q, M and N are given appropriate matrices.

Let $S(x)$ denote the solution-set of VIP(x). As usual we call a couple (x, y) such that $(x, y) \in Z$, $x \in X, y \in Y$, $y \in S(x)$ a feasible solution to Problem (P).

First we mention some important special cases of this problem.

Example 2.1 (Convex quadratic bilevel program). We consider the parametric variational inequality VIP(x), where $P(y) \equiv P$, and Q is a symmetric positive semidefinite matrix. In this case, since Y is convex, it is well-known that VIP(x) is equivalent to the convex programming problem

$$\min\left\{\frac{1}{2}y^TQy + (Px)^Ty + q^Ty \ : y \in Y\right\}.$$

Thus AMPEC problem (P) can be equivalently rewritten as a convex bilevel problem of the form

$$\min\left\{f(x,y) := a^Tx + b^Ty\right\}$$

subject to

$$(x,y) \in Z_1 := \{(x,y) \ : \ Mx + Ny + p \le 0, x \in X, y \in Y\},$$

where y solves the convex quadratic program

$$\min\left\{\frac{1}{2}y^TQy + (Px)^Ty + q^Ty \ : y \in Y\right\}. \qquad (C_x)$$

Example 2.2 (Optimization over the weakly efficient set). Other examples for the AMPEC are optimization problems over the efficient (Pareto) and weakly efficient sets of a multicriteria (vector) affine fractional program. These problems have been recently considered by some authors (see e.g. [17, 20, 22, 29]). The problems can be formulated in forms of AMPEC. To this end, consider the affine fractional vector optimization problem

$$\mathrm{vmin}\{F(v) := \Big(\frac{A_1^Tv + s_1}{B_1^Tv + t_1}, ..., \frac{A_\rho^Tv + s_\rho}{B_\rho^Tv + t_\rho}\Big) : \quad v \in V\}, \qquad (VP)$$

where $V \subset R^m$ is a bounded polyhedral convex set, A_i, B_i are m-dimensional vectors, s_i, t_i $(i = 1, ..., \rho)$ are real numbers. As usual we assume that $B_i^T v + t_i > 0$ for all $v \in V$ and all $i = 1, ..., \rho$. Thus F is continuous on V. We recall that a point $v \in V$ is said to be an (Pareto) efficient (resp. weakly efficient) solution of (VP) if there does not exist $w \in Y$ such that $F(w) \leq F(v)$, $F(w) \neq F(v)$ (resp. $F(w) < F(v)$), By $E(F, V)$ (resp. $WE(F, V)$) we will denote the set of all efficient (resp. weakly efficient) solutions of (VP). It is well-known (see e.g. [27]) that if V is compact, then the efficient set is nonempty. Hence so is the weakly efficient set, since $E(F, V) \subseteq WE(F, V)$. An optimization problem over the efficient set (resp. weakly efficient set) is the problem of optimizing (minimizing or maximizing) a real-valued function f over the efficient (resp. weakly efficient) set of (VP). These minimization problems can be written respectively as

$$\min\{f(v) \ : v \in E(F, V)\}, \tag{2.1}$$

$$\min\{f(v) \ : v \in WE(F, V)\}. \tag{2.2}$$

Note that, in general, both the efficient and weakly efficient sets are not convex. The weakly efficient set is closed but the efficient set may be neither closed nor open [7]. Thus these problems are difficult global optimization ones.In order to formulate these problems in the form of AMPEC we use the following theorem due to Malivert [17].

Theorem 2.1 ([17]) *A vector $v \in V$ is efficient (resp. weakly efficient) if and only if there exist real numbers $u_i > 0$ (resp. $u_i \geq 0$ not all zero) for all $i = 1, ..., \rho$ such that*

$$\sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \leq 0 \quad \forall w \in V.$$

In virtue of this theorem the problems (2.1) and (2.2) can be written as

$$\begin{cases} \min f(v) \ subjectto \ \ v \in V, \ u_i > 0 \ \text{n}otallzero \ \forall i = 1, ..., \rho, \\ \sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \leq 0 \quad \forall w \in V, \end{cases}$$

and

$$\begin{cases} \min f(v) \text{ subject to } \ v \in V, \ u_i \geq 0 \text{ not all zero } \forall i = 1, ..., \rho, \\ \sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \leq 0 \quad \forall w \in V \end{cases}$$

respectively.

Define the $(m \times \rho)$-matrix $P(v)$ by setting

$$P(v) := \Big\{(B_1^T v + t_1)A_1 - (A_1^T v + s_1)B_1, ..., (B_\rho^T v + t_\rho)A_\rho - (A_\rho^T v + s_\rho)B_\rho\Big\}.$$

Then we can rewrite these problems in the forms

$$\min\{f(v) : \quad v \in V, u > 0, \ \langle P(v)u, w - v\rangle \geq 0 \ \forall w \in V\} \tag{2.1a}$$

and

$$\min\{f(v) : \quad v \in V, u \geq 0, \ \langle P(v)u, w - v\rangle \geq 0 \ \forall w \in V\} \tag{2.2a}$$

respectively.

Clearly, the latter problem is of the form of AMPEC where v and u play the roles of y and x respectively, and $M \equiv 0$, $N \equiv 0$, $Q \equiv 0$, $p \equiv 0$, $q \equiv 0$. If in problem (2.1a) we replace the constraint $u > 0$ by $u \geq \delta$ with $\delta > 0$ sufficiently small as desired, we obtain an approximation problem to (2.1a) that is of the form of AMPEC.

In an important special case where $B_i = 0$ and $s_i = 1$ for all i, Problem (VP) becomes a linear vector program. In this case, it is well known [17, 27], both the efficient and weakly efficient sets are closed, but, in general, not convex. Thus Problems (2.1) and (2.2) remain global optimization ones, since there are local optimal solutions that are not global optimal ones. In this linear case due to a theorem of Philip [25] Problem (2.1a) can take the form of AMPEC as

$$\min\{f(v) : \ v \in V, u \geq \delta, \ \langle P(v)u, w - v\rangle \geq 0 \ \forall w \in V\},$$

where $\delta > 0$ is sufficiently small. Note that in this linear case

$$P(v) \equiv P := \Big(t_1 A_1, t_2 A_2, . \ \ . t_\rho A_\rho\Big)$$

is independent of v.

Example 2.3 (Nash-Cournot market model). The third section of AMPEC is a Nash-Cournot oligopolistic market model (see e.g. [10, 12]). The model can be described as follows.

Suppose that there are m-firms (sectors) that supply a homogeneous product whose price p_j at sector j $(j = 1, ..., m)$ depends on total producing quantity and is given by

$$p_j(\sum_{i=1}^{m} y_i) = \alpha_j - \beta_j \sum_{i=1}^{m} y_i,$$

where $\alpha_j \geq 0$, $\beta_j > 0$ are given constants whereas y_i is the quantity of goods supplied by firm i that we have to determine. Suppose further that to produce the goods the firms need n-different materials represented by a nonnegative vector $x \in R^n$. Let x_i $(i = 1, ..., n)$ be the quantity of material i needed to produce a unique of goods. Let $c_{ij} > 0$ denote the price of a unit material i for firm j $(i = 1, ..., n,\ j = 1, ..., m)$. Assume that the cost of firm j is given by

$$h_j(x, y_j) := y_j \sum_{i=1}^{n} c_{ij} x_i + \delta_j, \quad j = 1, ..., m,$$

where $\delta_j \geq 0$ is fixed charge cost at firm j. Then the utility function of firm j can be given by

$$u_j(x, y) := p_j(\sum_{i=1}^{m} y_i) y_j - h_j(x, y_j).$$

Let

$$X_i := \{t \ : \ 0 \leq t \leq \xi_i\} \ (i = 1, ..., n),$$

$$Y_j := \{\tau \ : \ 0 \leq \tau \leq \eta_j\} \ (j = 1, ..., m),$$

where ξ_i is the upper bound for material i, and η_j is the upper bound for the quantity of goods can be produced by firm j.

Let

$$X := X_1 \times X_2 ... \times X_n, \quad Y = Y_1 \times ... \times Y_m$$

be the feasible (strategy)-sets of the model.

Given $x \in X$ each firm j seeks to find its producing quantity y_j such that its benefit $u_j(x, y)$ is maximal. However, a maximal policy for all firms altogether, in general, does not exist. So they agree with an equilibrium point in the sense of Nash.

By definition, a vector $(y_1^*, ..., y_m^*) \in Y_1 \times Y_2 ... \times Y_m$ is said to be a (Nash) equilibrium point with respect to $x^* \in X$ if

$$\begin{cases} u_j(x^*, y_1^*, ..., y_{j-1}^*, y_j, y_{j+1}^*, ..., y_m^*) \\ \leq u_j(x^*, y_1^*, ..., y_{j-1}^*, y_j^*, y_{j+1}^*, ..., y_n^*) \ \forall y_j \in Y_j, \forall j. \end{cases} \tag{2.3}$$

We will refer to a pair (x^*, y^*) satisfying (2.3) as an *equilibrium pair* of the model.

Besides the utility function of each firm, there is another cost function (leader's objective function) $f(x, y)$ depending on x and the quantity y of the goods. The problem to be solved is of finding an equilibrium pair that minimizes leader's objective function over the set of all equilibrium pairs. This problem can be formulated as a mathematical program with affine equilibrium. To this end let

$$\begin{cases} H_j(x, y) := \nabla_{y_j} h_j(x, y_j) \ (j = 1, ..., m), \\ e := (1, ..., 1)^T, \ \sigma_y := \sum_{j=1}^m y_j. \end{cases} \tag{2.4}$$

Applying Proposition 3.2.6 in [12] we see that a point $(y_1, ..., y_m)$ is equilibrium with respect to x if and only if it is a solution to the variational inequality problem

$$\text{Find } y \in Y : \ \langle F(x, y), z - y \rangle \geq 0 \ \ \forall z \in Y,$$

where $F(x, y)$ is m-dimensional vector whose jth component is

$$F_j(x, y) := H_j(x, y) - p_j(\sigma_y)e - \nabla p_j(\sigma_y)y. \tag{2.5}$$

Using (2.4) and (2.5) we have

$$F(x, y) = \begin{pmatrix} \sum_{i=1}^n c_{i1}x_i - \alpha_1 + \beta_1 \sum_{j=1}^m y_j + \beta_1 y_1 \\ ...... \\ \sum_{i=1}^n c_{im}x_i - \alpha_m + \beta_m \sum_{j=1}^m y_j + \beta_m y_m \end{pmatrix}$$

$$= P(y)x + Qy + q,$$

where

$$Q = \begin{pmatrix} 2\beta_1 & \beta_1 & ... & \beta_1 \\ \beta_2 & 2\beta_2 & ... & \beta_2 \\ ... & ... & ... & ... \\ \beta_m & \beta_m & ... & 2\beta_m \end{pmatrix} \tag{2.6}$$

and $P(y)$ is the $n \times m$ matrix independent of y whose P_{ij} entry is

$$P_{ij} = c_{ij}, \ j = 1, ..., m, \ i = 1, ..., n, \tag{2.7}$$

and

$$q = (\delta_1 - \alpha_1, ..., \delta_m - \alpha_m)^T. \tag{2.8}$$

Thus problem to be solved can take the form

$$\begin{cases} \min f(x,y) \text{ subject to} \\ x \in X = X_1 \times ... \times X_n,\ y \in Y = Y_1 \times ... \times Y_n, \\ \text{where } y \text{ solves the parametric variational inequality} \\ \langle Px + Qy + q, z - y\rangle \geq 0 \quad \forall z \in Y, \end{cases}$$

with Q, P and q being given by (2.6), (2.7) and (2.8) respectively. Clearly, this problem is of form (P) with $M = 0, N = 0, p = 0$.

3 A Lagrangian Bounding Algorithm

The algorithm to be described in this section relies on the branch-and-bound strategy. Two main operations in a branch-and-bound algorithm are bounding and branching ones. The Lagrangian bound is widely used in global optimization as well as in discrete programming [6, 11, 14, 26]. To the algorithm we are going to describe for AMPEC problem (P) we also use the Lagrangian bounding operation.

3.1 The Lagrangian Bound

First, we consider the case when Y is a polytope and all of its vertices are known in advance. This case occurs frequently, for instant, in economics equilibrium models where Y is a simplex or a box. Let $y^1, y^2, ..., y^s$ be the vertices of polytope Y. It is easy to verify that

$$\langle P(y)x + Qy + q, z - y\rangle \geq 0 \ \forall z \in Y$$

if and only if

$$\langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s.$$

Thus, AMPEC problem (P) can be rewritten equivalently as

$$(P) \begin{cases} \text{minimize } \{f(x,y) := a^T x + b^T y\} \\ \text{subject to } x \in X, y \in Y, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \leq 0\}, \\ \text{and } y \in Y \text{ satisfying inequalitis} \\ \ \langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s. \end{cases}$$

Let

$$\hat{Y} := \{y \in Y : \exists x \in X \text{ such that } Mx + Ny + p \leq 0,$$
$$\langle P(y)x + Qy + q, y^k - y \geq 0 \ \forall y = 1, ..., s\}.$$

Note that if $\langle P(y)x, y\rangle$ is convex with respect to y, in particular when $X \subset R^n_+$ and $P(y) = P$ (see examples 2.1, 2.2 for linear case and 2.3), or when $P(y) = \text{Diag}(y)$, then $\hat{Y}$ is convex.

Define the function $\varphi : \hat{Y} \to I\!R$ by setting, for each $y \in \hat{Y}$,

$$\begin{cases} \varphi(y) := \min_x\{f(x,y) := a^T x + b^T y\} \\ \text{s. t.} \quad x \in X, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \leq 0\}, \\ \langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s. \end{cases} \qquad (P_y)$$

Then the master problem

$$\min\{\varphi(y) \ : \ y \in \hat{Y}\} \qquad (MP)$$

is equivalent to Problem (P) in the sense of the following proposition whose proof is obvious directly from the definitions.

Proposition 3.1 *A point (x^*, y^*) is optimal to Problem (P) if and only if y^* is optimizer to (MP) and $f(x^*, y^*) = \varphi(y^*)$.*

Note that, unlike global optimization problems having nonconvex feasible domains, feasible points of a MPEC problem can be computed by available methods of variational inequalities (see e. g. [1, 2, 10, 12] and the references therein). However for Problem (P), a feasible point can be obtained by solving a suitable linear program. In fact, if $y \in \hat{Y}$ is fixed and x^y is an optimal solution of the linear problem (P_y) then (x^y, y) is feasible for (P). So upper bounds for the optimal value w_* of (P) can be computed by solving a linear program. As the algorithm executes more feasible points can be found, and thereby upper bounds for w_* can be iteratively improved.

We now compute a tight lower bound for w_* by using Lagrangian duality. To be specific suppose that X is given explicitly as

$$X := \{x \in R^n : \ x \geq 0, Ax + d \leq 0\},$$

where $d \in R^l$ and A is $l \times n$-matrix. Let S be a fully dimensional simplex or a rectangle in y-space such that $S \cap \hat{Y} \neq \emptyset$. Consider Problem (P) restricted on $S \cap \hat{Y}$, i.e.,

$$w_*(S) = \min_{x,y} f(x, y) := a^T x + b^T y$$

subject to

$$\begin{cases} Mx + Ny + p \leq 0 \\ Ax + d \leq 0, \ x \geq 0, \ y \in S \cap \hat{Y} \\ \langle P(y)x + Qy + q, y - y^k \rangle \leq 0 \ \ \forall k = 1, ..., s. \end{cases} \qquad (P_S)$$

Let $L(x, y, \lambda, \mu, \xi)$ be the Lagrangian function of this problem associated with all constraints except the constraints $x \geq 0$, $y \in S \cap \hat{Y}$. That is

$$L(x, y, \lambda, \mu, \xi) := a^T x + b^T y + \sum_{k=1}^{s} \lambda_k \langle P(y)x + Qy + q, y - y^k \rangle + \mu^T (Mx + Ny + p) + \xi^T (Ax + d).$$

Define the function $m(y, \lambda, \mu, \xi)$ as

$$m(y, \lambda, \mu, \xi) := \inf_{x \geq 0} L(x, y, \lambda, \mu, \xi).$$

From the Lagrangian duality theorem we have

$$m(y, \lambda, \mu, \xi) \leq \varphi(y) \ \ \forall \lambda \geq 0, \mu \geq 0, \xi \geq 0, y \in S \cap \hat{Y}. \qquad (3.1)$$

Since for each fixed y the functions $Mx + Ny + p$, $Ax + d$, and $\langle P(y)x + Qy + q, y - y^k \rangle$ $\forall k = 1, ..., s$ are affine with respect to x, by the Lagrangian duality theorem, we have

$$\sup_{\lambda, \mu, \xi \geq 0,} m(y, \lambda, \mu, \xi) = \varphi(y) \ \ \forall y \in S \cap \hat{Y}.$$

Let

$$\gamma_S(\lambda, \mu, \xi) = \min_{y \in S \cap \hat{Y}} m(y, \lambda, \mu, \xi).$$

Then from (3.1) it follows that

$$\gamma_S(\lambda,\mu,\xi) \le \min_{y\in S\cap\hat{Y}} \varphi(y) = w_*(S) \quad \forall\lambda \ge 0, \mu \ge 0, \xi \ge 0.$$

Thus

$$\sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) \le w_*(S).$$

Hence

$$\beta(S) := \sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) \tag{3.2}$$

is a lower bound $\beta(S)$ for $w_*(S)$. The following lemma states that this lower bound can be computed by minimizing a certain convex function on $S \cap \hat{Y}$.

Lemma 3.1 *Suppose that Q is positive semidefinite matrix. Then*

$$\beta(S) = \min_{y\in S\cap\hat{Y}} g_S(y),$$

where

$$g_S(y) := \sup_{(\lambda,\mu,\xi)\in\Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\}$$

is convex on S and

$$\Omega(S) := \{(\lambda,\mu,\xi) : \ \lambda \ge 0, \mu \ge 0, \xi \ge 0, G_v(\lambda,\mu,\xi) \ge 0 \ \forall v \in S \cap \hat{Y}\}$$

with

$$G_v(\lambda,\mu,\xi) := a + \mu M + \xi A + \sum_{k=1}^{s} \lambda_k (P(v))^T (v - y^k).$$

Proof From (3.1) and the definition of $\nu_S(\lambda,\mu,\xi)$, it follows that

$$\beta(S) = \sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) = \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} m(y,\lambda,\mu,\xi).$$

Hence

$$\begin{cases} \beta(S) = \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \min_{x\ge 0} L(x,y,\lambda,\mu,\xi) = \\ \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \Big[\min_{x\ge 0} \{a^T x + b^T y + \sum_{k=1}^{s} \lambda_k \langle P(y)x + Qy + q, y - y^k\rangle \\ +\mu^T(Mx + Ny + p) + \xi^T(Ax + d)\}\Big] = \\ \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \Big[\{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\} \\ + \min_{x\ge 0}\{a^T x + \mu^T Mx + \xi^T Ax + \sum_{k=1}^{s} \lambda_k \langle P(y)x, y - y^k\rangle\}\Big]. \end{cases} \tag{3.3}$$

We now consider the last term of (3.3), that is

$$\min_{x\ge 0}\{a^T x + \mu^T Mx + \xi^T Ax + \sum_{k=1}^{s} \lambda_k \langle P(y)x, y - y^k\rangle\} =$$

$$\min_{x\geq 0}\{a^T x + \mu^T M x + \xi^T A x + \sum_{k=1}^{s} \lambda_k \langle x, P^T(y)(y-y^k)\rangle\} =$$

$$\min_{x\geq 0}\{a^T x + \mu^T M x + \xi^T A x + \langle x, \sum_{k=1}^{s} \lambda_k P^T(y)(y-y^k)\rangle\} =$$

$$\min_{x\geq 0}\langle x, a + M^T\mu + A^T\xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y-y^k)\rangle,$$

here and afterward, $P^T(y)$ denotes the transportation of the matrix $P(y)$.

If there exists $v \in S\cap \hat{Y}$ such that

$$a + M^T\mu + A^T\xi + \sum_{k=1}^{s} \lambda_k P^T(v)(v-y^k) \ngeq 0 \quad \forall \lambda \geq 0, \mu \geq 0, \xi \geq 0,$$

then

$$\min_{x\geq 0}\langle x, a + \mu M + \xi A + \sum_{k=1}^{s} \lambda_k P^T(v)(v-y^k)\rangle = -\infty.$$

So, the supremum in (3.3) can be taken over, all $\lambda \geq 0$, $\mu \geq 0$ and $\xi \geq 0$ satisfying

$$a + M^T\mu + A^T\xi + \sum_{k=1}^{s} \lambda_k P^T(v)(v-y^k) \geq 0 \quad \forall v \in S\cap \hat{Y}.$$

Clearly, under the condition

$$a + \mu M + \xi A + \sum_{k=1}^{s} \lambda_k P^T(y)(y-y^k) \geq 0 \quad \forall y \in S\cap \hat{Y},$$

one has

$$\min_{x\geq 0}\langle x, a + M^T\mu + A^T\xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y-y^k)\rangle = 0.$$

Thus we deduce from (3.3) that

$$\begin{cases} \beta(S) = \sup_{\lambda,\mu,\xi\geq 0} \min_{y\in S\cap\hat{Y}} \left[b^T y + \sum_{k=1}^{r} \lambda_k \langle Qy+q, y-y^k\rangle + \mu^T(Ny+p) + \xi^T d\right] \\ \text{subject to} \\ a + M^T\mu + A^T\xi + \sum_{k=1}^{s} \lambda_k P^T(v)(v-y^k) \geq 0 \quad \forall v \in S\cap\hat{Y}. \end{cases} \tag{3.4}$$

Let

$$\Omega(S) := \{(\lambda,\mu,\xi) : \ \lambda \geq 0, \mu \geq 0, \xi \geq 0, G_v(\lambda,\mu,\xi) \geq 0 \ \forall v \in S\cap\hat{Y}\}.$$

Then $\Omega(S)$ is a closed convex set and

$$\begin{cases} \beta(S) = \sup_{(\lambda,\mu,\xi)\in\Omega(S)} \min_{y\in S\cap\hat{Y}} \{b^T y + \sum_{k=1}^{r} \lambda_k \langle Qy+q, y-y^k\rangle + \\ \mu^T(Ny+p) + \xi^T d\}. \end{cases} \tag{3.5}$$

Since, by the assumption, Q is positive semidefinite, the function

$$b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d$$

is convex-linear on $(S \cap \hat{Y}) \times \Omega(S)$. Then, by the well known minimax theorem, we can interchange the supremum and infimum in (3.5) to obtain

$$\beta(S) = \min_{y \in S \cap \hat{Y}} \sup_{(\lambda,\mu,\xi) \in \Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}.$$

Note that, since

$$b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d$$

is convex with respect to y, the function

$$g_S(y) := \sup_{(\lambda,\mu,\xi) \in \Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}$$

is convex on $S \cap \hat{Y}$. Hence

$$\beta(S) = \min_{y \in S \cap \hat{Y}} g_S(y)$$

is a convex program. □

Remark 3.1 (a) If $P^T(y)y - P^T(y)y^k$ is concave on Y (in the sense that its every component is concave), then $\Omega(S)$ is a polyhedral convex set, since it can be represented by a finite affine inequalities. Indeed, since $P^T(y)y - P^T(y)s^k$ is concave, it is easy to verify that

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y - y^k) \geq 0 \quad \forall y \in Y \cap S$$

if and only if

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y - y^k) \geq 0 \quad \forall y \in V(Y \cap S),$$

where $V(Y \cap S)$ denotes the vertex-set of $Y \cap S$.

In the case $S \subseteq Y$ the last inequalities can be written as

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(v^j)(v^j - y^k) \geq 0 \quad \forall j = 1, 2, ..., r,$$

where $v^1, ..., v^r$ are the vertices of S.

(b) From the presentation of Section 2 we can see that

• For bilevel quadratic convex problem, $P(y) \equiv P$ (constant matrix).

• For optimization problem over the efficient set of an affine vector optimization program $P^T(y)y^k$ is affine and $P^T(y)y \equiv 0$.

3.2 Simplicial and Rectangular Bisections

At each iteration k of algorithm to be described below, a partition simplex (or rectangle) will be bisected into subsimplices (or subrectangles) such a way so that as the algorithm executes the obtained lower and upper bounds tend to the same limit. This can be achieved by using the following exhaustive simplicial (or rectangular) bisection that is commonly known in global optimization (see e. g. [14]).

Simplicial Bisection. We will use the following simplicial bisection [14].

Let S_k be a subsimplex of full dimension that we want to bisect at iteration k. Let v^k, w^k be two vertices of S_k such that the edge joining these vertices is longest. Let $u^k = t_k v^k + (1 - t_k) w^k$ with $0 < t_k < 1$. Let S_{k_1}, S_{k_2} be the subsimplices obtained from S_k by replacing v^k and w^k respectively by u^k. It is well known from [14] that $S_k = S_{k_1} \cup S_{k_2}$, and that if $\{S_k\}$ is an infinite sequence of nested simplices generated by this simplicial bisection process such that $0 < \delta_0 < t_k < \delta_1 < 1$ for every k, then the sequence $\{S_k\}$ shrinks to a singleton.

Rectangular Bisection. Suppose that the partition set is a rectangle given by

$$S_k := \{y = (y_1, ..., y_m) \in R^m : \ a_i \le y_i \le b_i \ i = 1, ..., m\}.$$

Let $[a_{i_k}, b_{i_k}]$ be a longest edge of S_k and $u_{i_k} = t_{i_k} a_{i_k} + (1 - i_k b_{i_k}$ with $0 < t_{i_k} < 1$. Then we bisect S_k into two rectangles S_{k_1} and S_{k_2} where

$$S_{k_1} = \{y \in R^m : \ a_i \le y_i \le b_i \ \forall i \ne i_k, \ a_{i_k} \le y_{i_k} \le u_{i_k}\},$$

and

$$S_{k_2} = \{y \in R^m : \ a_i \le y_i \le b_i \ \forall i \ne i_k, \ u_{i_k} \le y_{i_k} \le b_{i_k}\}.$$

As before, we have $S_k = S_{k_1} \cup S_{k_2}$, and that if $\{S_k\}$ is an infinite sequence of nested rectangles generated by this bisection process such that $0 < \delta_0 < t_k < \delta_1 < 1$ for every k, then the sequence $\{S_k\}$ shrinks to a singleton.

Now we are in a position to describe the algorithm for solving AMPEC problem (P), where Y is a polytope. We suppose that $\langle P(y)x, y\rangle$ is convex in Y with respect to y. In the sequel, as usual, we call (x, y) an ϵ-global optimal solution to (P) if it is feasible and $f(x, y) - w_* \le \epsilon(|f(x, y)| + 1)$ where w_* stands for its optimal value. Having the vertices $y^1, ..., y^s$ of Y we can describe the algorithm as follows.

Algorithm 1.

Initialization. Choose a tolerance $\epsilon > 0$ and a simplex or a rectangle S_0 containing Y. Compute the lower bound $\beta(S_0)$ by solving the convex program

$$\beta(S_0) := \min_{y \in \hat{Y}} \Big\{ g_0(y) := \sup_{(\lambda,\mu,\xi)\in\Omega(S_0)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\} \Big\}.$$

Let $y^0 \in S_0$ be the obtained solution.

Solve the linear program

$$\min_x \{f(x, y^0) := a^T x + b^T y^0\}$$

subject to

$$\begin{cases} Mx + Ny^0 + p \le 0 \\ Ax + d \le 0, x \ge 0 \\ \langle P(y^0)x + Qy^0 + q, y^0 - y^k\rangle \le 0 \ \ \forall k = 1, ..., s \end{cases}$$

to obtain x^0 (hence (x^0, y^0) is feasible). Let $\alpha_0 := f(x^0, y^0)$ (an upper bound for the optimal value w_*) and $\beta_0 := \beta(S_0)$ (a lower bound for w_*). Take

$$\Gamma_0 := \begin{cases} \{S_0\} & \text{if } \alpha_0 - \beta_0 > \epsilon(|\alpha_0| + 1), \\ \emptyset & \text{otherwise} \end{cases}$$

and go to iteration k with $k := 0$.

Iteration ($k = 0, 1 ...$). At the beginning of each iteration k we have family Γ_k of partition sets to each element $S \in \Gamma_k$ a real number $\beta(S)$ has been computed that serves as a lower bound for Problem (P) restricted in S. Moreover we have a lower bound β_k for w_*, a currently best feasible point (x^k, y^k) and an upper bound $\alpha_k = f(x^k, y^k)$ for w_*.

Step 1 (selection):
(i) If $\Gamma_k = \emptyset$ then terminate, (x^k, y^k) is an ϵ-global optimal solution and α_k is the ϵ-optimal value to Problem (P).
(ii) If $\Gamma_k \neq \emptyset$, then select $S_k \in \Gamma_k$ such that

$$\beta_k = \beta(S_k) = \min\{\beta(S) \ : \ S \in \Gamma_k\}.$$

Step 2 (bisection): Use the simplicial bisection (if S_k is simplicial) or use the rectangular bisection (if S_k is rectangular) to bisect S_k into two sets S_{k_1} and S_{k_2}.

Step 3 (bounding): For each newly generated sets S_{k_j} $(j = 1, 2)$ satisfying $S_{k_j} \cap \hat{Y} \neq \emptyset$, compute

$$\beta(S_{k_j}) := \min_{y \in S_{k_j} \cap \hat{Y}} g_{S_{k_j}}(y),$$

where

$$g_{S_{k_j}}(y) := \sup_{(\lambda, \mu, \xi) \in \Omega(S_{k_j})} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}.$$

Let y^{k_j} be the obtained solution.

Step 4 (updating upper bound): Solve the linear programs, one for each y^{k_j}, $j = 1, 2$

$$\min_x \{f(x, y^{k_j}) := a^T x + b^T y^{k_j}\}$$

subject to

$$\begin{cases} Mx + Ny^{k_j} + p \leq 0 \\ Ax + d \leq 0, x \geq 0 \\ \langle P(y^{k_j})x + Qy^{k_j} + q, y^{k_j} - y^k \rangle \leq 0 \quad \forall k = 1, ..., s \end{cases}$$

to obtain new feasible points. Use these feasible points to update the upper bound. Let (x^{k+1}, y^{k+1}) be the currently best feasible point among (x^k, y^k) and the newly generated feasible points. Set $\alpha_{k+1} := f(x^{k+1}, y^{k+1})$ and

$$\Gamma_{k+1} := \Big\{S \in (\Gamma_k \backslash \{S_k\}) \cup \{S_{k_1}, S_{k_2}\} \ : \ \alpha_{k+1} - \beta(S) > \epsilon(|\alpha_{k+1}| + 1)\Big\}.$$

Increase k by 1 and go to Step 1 of iteration k.

Theorem 3.1 *a) If Algorithm terminates at iteration k, then (x^k, y^k) is an ϵ-global optimal solution to Problem (P).*
b) If the algorithm does not terminate, then $\beta_k \nearrow w_*, \alpha_k \searrow w_*$ as $k \to +\infty$, and any cluster point of the sequence $\{(x^k, y^k)\}$ is a global optimal solution to (P).

Proof a) If the algorithm terminates at iteration k then $\Gamma_k = \emptyset$. This implies that $\alpha_k - \beta_k \leq \epsilon(|\alpha_k| + 1)$. Since $\beta_k \leq w_*$ and $\alpha_k = f(x^k, y^k) \geq w_*$, it follows that $f(x^k, y^k) - w_* \leq \epsilon(|f(x^k, y^k)| + 1)$. Hence (x^k, y^k) is an ϵ-global optimal solution.

b) Suppose that the algorithm does not terminate. First, note that since $S_k = S_{k_1} \cup S_{k_2}$, by the rule for computing lower bound $\beta(S_k)$ we have

$$\beta_k = \beta(S_k) \leq \beta(S_{k+1}) = \beta_{k+1} \quad \forall k.$$

Also, by definition of α_k, we have $\alpha_{k+1} \leq \alpha_k \ \forall k$. Thus, both $\beta_* = \lim \beta_k$ and $\alpha_* = \lim \alpha_k$ exist and satisfying

$$\beta_* \leq w_* \leq \alpha_*. \tag{3.6}$$

Since the algorithm does not terminate, it generates an infinite sequence of nested partition sets that, for simplicity of notation, we also denote by $\{S_k\}$. Since the subdivision is exhaustive, this sequence strinks to a singleton, say $y^* \in \hat{Y}$. By the rule for computing lower bound β_k we have

$$\beta_k = \sup_{\tau \geq 0} \min_{y \in S_k \cap \hat{Y}} m(y, \tau) \geq \min_{y \in S_k} m(y, \tau) \quad \forall \tau \equiv (\lambda, \mu, \xi) \geq 0.$$

Since the sequence $\{S_k\}$ shrinks to y^* as $k \to +\infty$, we obtain

$$\beta_* = \lim \beta_k \geq m(y^*, \tau) \quad \forall \tau \geq 0.$$

By definition, since $\varphi(y^k)$ is an upper bound determined at iteration k and α_{k+1} is the currently smallest upper bound obtained at this iteration, we can write

$$\alpha_{k+1} \leq \varphi(y^k) \quad \forall k.$$

From $y^k \to y^*$, it follows, by the continuity of φ (see e. g. [3, 5]), that

$$\alpha_* = \lim \alpha_k = \lim \alpha_{k+1} \leq \lim \varphi(y^k) = \varphi(y^*).$$

On the other hand, by Lagrangian duality theorem for the convex program determining $\varphi(y^*)$, we have

$$\sup_{\tau \geq 0} m(y^*, \tau) = \varphi(y^*).$$

Hence

$$\alpha_* \leq \varphi(y^*) \leq \beta_*$$

which together with (3.6) implies

$$\beta_* = w_* = \alpha_* = \varphi(y^*).$$

Let (x^*, y^*) be any cluster point of the sequence $\{(x^k, y^k)\}$. By the definition we have $\alpha_k = f(x^k, y^k)$. Since $\alpha_k \searrow w_*$, it follows from continuity of f that $w_* = f(x^*, y^*)$. Hence (x^*, y^*) is a globally optimal solution to Problem (P). □

Remark 3.2 Note that when Q is positive semidefinite, the function

$$g_S(y) := \max_{(\lambda,\mu,\xi)\in\Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\}$$

is convex and subdifferentiable, since it is the maximum of a family of convex functions. The subgradient of g_S at a point y can be obtained by taking the convex envelope of the gradients of those quadratic functions in the family at which $g_S(y)$ is attained [4].

3.3 Optimization over the Weakly Efficient Set

Now we return to the optimization over the weakly efficient set mentioned in the previous section. By using again the necessary and sufficient condition due to Malivert (see Theorem 2.1) the minimization problem over the weakly efficient set of a multicriteria affine fractional program can be written as

$$\min f(v)$$

subject to

$$v \in V,\ u \in U,$$

$$\sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle \geq 0 \ \ \forall w \in V,$$

where

$$U := \{u \in R^{\rho} :\ u \geq 0, \sum_{i=1}^{p} u_i = 1\}$$

is a simplex in the criteria space.

Since

$$\Big\langle \sum_{i=1}^{\rho} u_i B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle$$

is affine with respect to w, we can easily check that

$$\Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle \geq 0 \ \ \forall w \in V$$

if and only if

$$\Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v^j - v\Big\rangle \geq 0 \ \ \forall j = 1, ..., r,$$

where v^j $(j = 1, ..., r)$ are vertices of V. Thus we can write this problem as

$$\begin{cases} \min f(v) : \quad \text{s. t. } v \in V,\ u \in U, \\ \Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T + s_i)B_i], v^j - v\Big\rangle \geq 0 \ \ \forall j = 1, ..., r. \end{cases} \tag{3.7}$$

It is worth pointing out that, in contrast to the linear case, this problem does not necessarily attain its optimal solution among the vertices of V.

In this formulation we require that all vertices of V are known. This case appeared already in some applications in economics (see e. g. [18, 27, 28]) where each component v_j of the decision variable v represents the ratio of ith quantity to be determined. In such practical models, V is a simplex given by

$$V := \Big\{ v^T = (v_1, ..., v_m) \ : \sum_{i=1}^{m} v_i = 1, v_j \geq 0 \ \forall j = 1, ..., m \ \Big\},$$

whose vertices are easy to compute. Generally, let us assume that V is a polytope given explicitly as

$$V := \{ v \in R^m : \ v \geq 0, \ Gv - g \leq 0 \}.$$

For simplicity of notation, for each vertex v^j, we take

$$M_j(u, v) := \Big\langle \sum_{k=1}^{\rho} u_k [(B_k^T v + t_k) A_k - (A_k^T v + s_k) B_k], v - v^j \Big\rangle,$$

$$G_j(u) := \sum_{k=1}^{\rho} u_k [(B_k^T v^j + t_k) A_k^T - (A_k^T v^j + s_k) B_k^T],$$

$$g_j(u) := \sum_{k=1}^{\rho} u_k [t_k A_k^T - s_k B_k^T] v^j.$$

Let $G(u)$ denote the $(r \times m)$-matrix whose ith row is $G_j(u)$ $(j = 1, ..., r)$, and $g(u)$ denote the r-dimensional vector whose jth entry is $g_j(u)$. Now let

$$H(u) := \begin{pmatrix} G \\ G(u) \end{pmatrix}, \quad h(u) := \begin{pmatrix} g \\ g(u) \end{pmatrix}.$$

Under these notations we can write the problem (3.7) in the form

$$\min\{ f(v) : \ H(u)v - h(u) \leq 0, v \geq 0, u \in U \}. \tag{3.8}$$

To apply the Lagrangian duality we take the Lagrangian function for this problem with respect to the constraint $H(u)v - h(u) \leq 0$, that is

$$L(\lambda, u, v) := f(v) + \lambda^T \Big(H(u)v - h(u) \Big).$$

Using the fact that both $H(u)$ and $h(u)$ are affine, by a similar argument as in the proof of Lemma 2.1, we can compute lower bounds by solving linear programs as stated by the following lemma.

Lemma 3.2 *Suppose $f(v) = b^T v$. Let S be the subsimplex of the simplex U, and s^j $(j = 1, ..., \rho)$ be the vertices of S. Then $\beta(S) = \min\{\beta(s^j) \ : j = 1, ..., \rho\}$ where, for each fixed s^j, $\beta(s^j)$ is the optimal value of the linear program*

$$\beta(s^j) := \max\{ -h^T(s^j)u \ : H^T(s^j)u + b \geq 0 \}.$$

Having this lower bounding operation we can use Algorithm 1 with the exhaustive simplicial bisection taking over subsimplices of the simplex U to solve problem (3.8). Note that in this case, if (u^S, v^S) is a solution obtained by computing lower bound $\beta(S)$ according to Lemma 3.2, then v^S is weakly efficient. Hence (u^S, v^S) can serve for updating upper bound in the algorithm.

4 A Relaxation Algorithm

In the algorithm presented in the preceding section we required that all vertices of the polytope Y are known in advance. In the case computing all of these vertices is expensive, we recommend to use another algorithm that allows that these vertices can be computed iteratively. It is expected that the algorithm finds an approximate solution without computing all of these vertices. In order to present the algorithm, suppose that we know already some vertices $v^1, ..., v^r$ of Y. Having these vertices we form the relaxation problem

$$\begin{cases} r_* = \min f(x,y) := a^T x + b^T y \\ \text{s.t. } Mx + Ny + p \le 0 \\ Ax + d \le 0,\ x \ge 0,\ y \in Y \\ \langle P(y)x + Qy + q, v^k - y \rangle \ge 0 \ \ \forall k = 1, ..., r. \end{cases} \qquad (RP)$$

Clearly, the feasible domain of this problem contains that of Problem (P) presented in Section 2 with $X = \{x \ge 0 \ : \ Ax + d \le 0\}$. Applying Algorithm 1 to this problem we obtain an ϵ-global optimal solution of (RP). If this solution satisfies the variational inequality constraint, then it is also an ϵ-global solution to the original problem (P). Otherwise, it violates at least one constraint. Then we add one or more violated constraints to obtain new relaxation problem and repeat the process. The algorithm can be described in detail as follows.

Algorithm 2.

Step 1. Choose distinct vertices $v^1, ..., v^r$ of Y and a tolerance $\epsilon > 0$.

Step 2. Use Algorithm 1 to solve (RP) to obtain an ϵ-global optimal solution (x^r, y^r) to (RP).

Step 3. Solve the linear program

$$\min\{\langle P(y^r)x^r + Qy^r + q, y\rangle \ : y \in Y\} \qquad (L_r)$$

to obtain an basis (vertex) solution y^{r+1}.

(a) If

$$\langle P(y^r)x^r + Qy^r + q, y^r\rangle \le \langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle$$

(hence y^r also solves (L_x)), then terminate: (x^r, y^r) is an ϵ-global optimal solution to the original problem.

(b) Otherwise, take $v^{r+1} := y^{r+1}$. Add v^{r+1} to the list of known vertices of Y to form the new relaxation problem (RP) and go back to Step 2.

The following theorem shows validity and finiteness of this algorithm.

Theorem 4.1 *(i) If the algorithm terminates at case (a) of Step 3, then (x^r, y^r) is an ϵ-global optimal solution.*

(ii) The algorithm terminates after a finite number of Step 2 yielding an ϵ-global optimal solution to the original problem (P).

Proof (i) Note that Problem (P) can be written as

$$w_* := \min\{f(x,y) := a^T x + b^T y\}$$

subject to

$$\begin{cases} Mx + Ny + p \le 0 \\ Ax + d \le 0, x \ge 0, y \in Y \\ \langle P(y)x + Qy + q, v - y\rangle \ge 0 \ \ \forall v \in V(Y) \end{cases}$$

while the relaxation problem is

$$r_* := \min\{f(x,y) := a^T x + b^T y\}$$

subject to

$$\begin{cases} Mx + Ny + p \leq 0 \\ Ax + d \leq 0, x \geq 0,\ y \in Y \\ \langle P(y)x + Qy + q, v^k - y\rangle \geq 0 \quad \forall k = 1, ..., r \end{cases}$$

with v^k $(k = 1, ..., r)$ being some vertices of Y.

In the case (a) we have

$$\langle P(y^r)x^r + Qy^r + q, y^r\rangle \leq \langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle.$$

Since y^{r+1} is an optimal solution of (L_r), we have

$$\langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle \leq \langle P(y^r)x^r + Qy^r + q, y\rangle\ \forall y \in Y.$$

Thus

$$\langle P(y^r)x^r + Qy^r + q, y - y^r\rangle \geq 0\ \forall y \in Y.$$

Hence (x^r, y^r) is feasible for (P). But, since (x^r, y^r) is an ϵ-global optimal solution to the relaxed problem (RP), it must be an ϵ-global solution to (P).

(ii) Note that if $y^{r+1} = v^j$ for some $j \leq r$, then we have the case (a), and therefore, the algorithm terminates. Thus, if situation (a) is not the case, we have $v^{r+1} \neq v^j$ for all r and all $j \leq r$, since the number of the vertices of Y is finite, the algorithm must terminate with case (a). □

5 Conclusion

We have considered a class of mathematical programs with affine variational inequality constraints and presented some of its important special cases such as bilevel convex programming, optimization over the efficient set and Cournot-Nash oligopolistic market model. We have developed two decomposition branch-and-bound algorithms for globally solving this class of mathematical programs with affine equilibrium constraints. The proposed algorithms use the Lagrangian bound and exhaustive simplicial and rectangular subdivisions widely used in global optimization. The main subproblems needed to be solved in the algorithms are convex programs that can be solved by well developed methods of nonsmooth convex programming.

Acknowledgement

The second author wishes to thank the financial supports from the Alexander-von-Humboldt Foundation, Germany and from the Associate and Federation Schemes, ICTP, Italy.

References

[1] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J. On the contraction and non-expensiveness properties of the marginal mapping in generalized variational inequalities involving cocoercive operators. In: *Generalized Convexity, Generalized Monotonicity and Applications*. (Eberhard, A., Hadjisavvas, H. and Luc. D.T., eds.). Springer, Berlin, 2005, 89–111.

[2] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J. Using the Banach contraction principle to implement the proximal point method for solving multivalued monotone variational inequalities. *J. of Optimization Theory and Applications* **124** (2005) 285–306.

[3] Aubin, J.P. and Ekeland, I. *Applied Nonlinear Analysis*. Marcel Dekker, New York, 1984.

[4] Auslender, A. *Optimisation: Methodes Numeŕiques*. Masson, Paris, 1976.

[5] Berge, C. *Topological Spaces*. Macmillan, New York, 1968.

[6] Bental, A., Eiger, G. and Gershovitg, V. Global minimization by reducing the duality gap. *Mathematical Programming* **63** (1994) 193–212.

[7] Choo, E.U. and Atkins, D.H. Connectedness in multiple linear fractional programming. *Management Science* **29** (1983) 250–255.

[8] Dempe, S. Annotated Bibliography on Bilevel Programming and Mathematical Programs with Equilibrium Constraints. *Optimization* **52** (2003) 333–359.

[9] Facchinei, F., Jiang, H. and Qi, Q. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming* **85** (1999) 107–134.

[10] Facchinei, F. and Pang, J. S. *Finite Dimesional Variational Inequality and Complementarity Problems*. Springer, Berlin, 2002.

[11] Falk, J.E. Lagrange multipliers and nonconvex programs. *SIAM J. Control* **7** (1969) 534–545.

[12] Konnov, I.V. *Combined relaxation methods for variational inequalities*. Springer, Berlin, 2001.

[13] Friesz, T.L., Tobin, R.L., Cho, H.J. and Mehta, N.J. Sensitivity analysis based heuristic algorithm for mathematical programs with variational inequality constraints. *Mathematical Programming* **48** (1990) 265–284.

[14] Horst, R. and Tuy, H. *Global Optimization: Deterministic Approach*. Springer, Berlin, 2003.

[15] Lin, G.H. and Fukushima, M. Some exact penalty results for nonlinear programs and mathematical programs with equilibrium constraints. *J. Optimization Theory and Applications* **118** (2003) 67–80.

[16] Luo, Z.Q., Pang. J.S. and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[17] Malivert, C. Multicriteria fractional optimization. In: *The 2-nd Catalans Days on Appl.* Presses Universitaires de Perpignan, 1995, 189-198.

[18] Marcotte, P. Network design problem with congestion effects: A case of bilevel programming. *Mathematical Programming* **34** (1986) 142–162.

[19] Mordukhovich, B.S. Optimization and equilibrium problems with equilibrium constraints. *Omega* **33** (2005) 379–384.

[20] Muu, L.D. and Oettli, W. Optimization over equilibrium sets. *Optimization* **49** (2000) 179–189.

[21] Muu, L.D. and Quy, N.V. Method for finding global optimal solution to linear programs with equilibrium constraints. *Acta Mathematica Vietnamica* **26** (2001) 314–333.

[22] Muu,. L.D. and Tuyen, H.Q. Bilinear programming approach to optimization over the efficient set of a vector affine fractional problem. *ACTA Mathematica Vietnamica* **27** (2002) 119–139.

[23] Muu, L.D. and Quy, N.V. A global optimization method for solving convex quadratic bilevel programming problems. *J. of Global Optimization* **26** (2003) 199–219.

[24] Outrata, J.V. and Zowe, J. A numberical approach to optimization problems with variational inequality constraints. *Mathematical Programming* **68** (1995) 105–130.

[25] Philip, J. Algorithms for the vector maximization problem. *Mathematical Programming* **2** (1972) 207–228.

[26] Shor, N.Z. and Stetsyuk, P.I. Lagrangian bounds in multiextremal polynomial and discrete optimization problems. *J. of Global Optimization* **23** (2002) 1–41.

[27] R. E. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Applications*, John Willey and Sons, New York, 1986.

[28] Tobin, R.L. Uniqueness results and algorithms for Stackelberg-Cournot-Nash equilibria. In: *Hierarchical Optimization*. (Anandalingam, G. and Friesz, T., eds.), Ann. Oper. Res., 1992, 21–36.

[29] Tuyen, H.Q. and Muu, L.D. Biconvex programming approach to optimization over the weakly efficient set of a multiple objective affine fractional problem. *Operations Research Letters* **28** (2001) 81–92.

Nonlinear Dynamics and Systems Theory, **6**(3) (2006) 245–256

Thermal Stresses in a Hexagonal Region With an Elliptic Hole

Sukhwinder Kaur Bhullar *

Department of Mathematics, Panjab University Chandigarh–160014, India

Received: April 30, 2005; Revised: June 23, 2006

Abstract: Considering importance of stress concentration around holes and notches of arbitrary shape in a given elastic medium for modern engineering, a two dimensional model for a thermoelastic problem in an hexagon region with an elliptic hole is established. The expressions for the temperature distribution and thermal stresses which have their importance in nuclear engineering are obtained for the model. The five elementary function's method in plane thermoelasticity of multiply connected regions is used to obtain the solutions for temperature distribution and thermal stresses. Numerical calculations are computed assuming a central elliptic hole in the hexagonal region having thermally insulated outer boundary under uniform heat generation. The obtained results are depicted graphically.

Keywords: *Temperature; thermal stress; Lame's constants.*

Mathematics Subject Classification (2000): 74F05, 74A10, 74G50.

1 Introduction

The investigation of stress concentration around holes and notches of arbitrary shape in a given elastic medium is very important for modern engineering. The high stress concentration found at the edge of a hole is of great importance. The heat generating cylinder with a hole are used in the construction of the reactor. The circular cylinder with a square hole is an applicable problem in the construction of support of the bridge. Polygon region with an elliptic hole have been used in nuclear reactor. As an example holes in ships deck may be mentioned. When the hull of a ship is bent, tension or compression is produced in the decks and there is a high stress concentration at the holes. Under the cycles of stress produced by waves, fatigue of the metal at the over

* Corresponding author: sbhullar_2000@hotmail.com

stressed portions may result finally in fatigue cracks. It is often necessary to reduce the stress concentration at holes such as access holes in airplane wings and fuselages. The contribution of several authors in this field is in [1] − [5]. Takeouti et al. studied problem of a thick cylinder having a polygon hole, thermal stress distribution in a triangular, square, hexagonal and octagonal region with a circular hole and theoretical thermal stress distribution in square region with an elliptic hole in [6] − [11]. Deresiewicz [12] calculated thermal stresses in a plate due to disturbance of uniform heat flow by a hole of general shape. Florence et al. [13] studied the problem of an infinite plate under a steady-state temperature distribution with uniform heat due to presence of an insulated ovaloid hole. Chowdhury [14] obtained thermal stresses due to uniform temperature distributed over a band of the cylindrical hole in an infinite body. Verba [15] et al. discussed static problem of thermoelasticity for an infinite plate weakened by a rectangular hole. Pan [16] found stresses in an infinite elastic plate containing two unequal circular holes. Chao et al. [17] considered problems for an anisotropic thermoelastic body containing an elliptic hole boundary.

With above background in this paper, a basic analysis is presented for thermal stress analysis in multiply connected region and the solutions for the temperature and thermal stress in a hexagon regions with an elliptic hole are obtained in the form of the infinite series expressed by the elliptic co-ordinates. The unknown constants are determined so as to satisfy boundary conditions and as they become enormous, therefore, we use point matching technique, as an extension of five elementary function's method in plane thermoelasticity of multiply-connected regions [18], [19].

2 Formulation of problem

Consider a hexagon region as shown in Figure 2.1, with an elliptic hole at the centre.

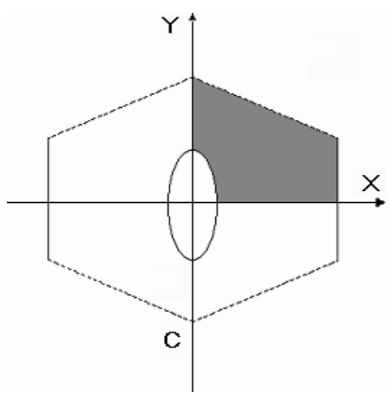


Figure 2.1: Geometry of the Problem.

Assume that the region is thermally insulated at the outer boundary with an internal convective boundary and is free from external forces. The region is made from an isotropic linear elastic material then following Takeuti [18], its behaviour under the influence of in-plane nonuniform temperature distribution which produces infinitesimal displacements is governed by the equation

$$\nabla\nabla\chi_\tau = -k\nabla\tau, \tag{1}$$

where τ is temperature change from reference state, $k = \varepsilon$ E for plane stress problem, ε is coefficient of linear thermal expansion, E is Young's modulus, χ is Airy's stress function, $\nabla = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$. The mean stresses in two dimensions are expressed in terms of stress function χ by the equation

$$\sigma_{ij} = (\nabla \delta_{ij} - \partial_i \partial_j)\chi, \tag{2}$$

where δ_{ij} is Kronecker delta, ∂_i is partial differential with respect to i $(i, j = 1, 2)$. Steady-state heat conduction with internal heat source is governed by the equation

$$-\lambda \nabla \tau = q, \tag{3}$$

where q is heat generation per unit volume per unit time, λ is thermal conductivity. As the region is multiply connected, the stress function χ can be expressed in terms of five elementary functions χ_τ, χ_0, χ_{1l}, χ_{2l} and χ_{3l} so that

$$\chi = \chi_\tau + \chi_0 + \sum_{h=1}^{3} \sum_{l=1}^{n} C_{hl} \chi_{hl}, \tag{4}$$

where C_{hl} are constants, $h = 1, 2, 3,$ $l = 1, 2, \ldots, n$ and h, l are not summed. Now, functions given in equation (4) should have to satisfy the following equations

$$\nabla \nabla \chi_\tau = -k \nabla \tau, \tag{5}$$

$$\nabla \nabla (\chi_0, \chi_{hl}) = 0. \tag{6}$$

Boundary conditions: Boundary conditions on the m-th boundary,

$$(\chi_\tau)_{p_m} = (\chi_{\tau,\nu})_{p_m} = 0, \tag{7}$$

$$(\chi_0)_{p_m} = -\int^{pm} dx_1 \int^{pm} \chi_{2\nu} ds + \int^{pm} dx_2 \int^{pm} \chi_{1\nu} ds, \tag{8}$$

$$(\chi_{0,\nu})_{p_m} = -(\nu_1)_{p_m} \int^{pm} \chi_{2\nu} ds + (\nu_2)_{p_m} \int \chi_{1\nu} ds, \tag{9}$$

$$(\chi_{hl})_{p_m} = [(x_h)_{p_m} (\delta_{1h} + \delta_{2h}) + \delta_{3h}] \delta_{lm}, \tag{10}$$

$$(\chi_{hl,\nu})_{p_m} = [(\nu_h) p_m (\delta_{1h} + \delta_{2h})] \delta_{lm}, \tag{11}$$

where ν is outward normal, $\nu_1 = \cos(x_1, \nu)$, $\nu_2 = \cos(x_2, \nu)$, p_m is an arbitrary point on m-th boundary, $m = 1$ corresponds to the elliptic boundary, m=0 corresponds to the hexagon boundary. In the multiply-connected bodies general equations expressed in stress components, are not sufficient for determining stresses and to get a complete solution an additional investigation of displacement is necessary. The first investigation of this kind was made by J.H. Michell [20] which are known as Michell's condition and given by

$$\int_{c_l} [\nabla(\chi_\tau + \chi_0 + \sum_{h=1}^{3}\sum_{l=1}^{n} C_{hl}\chi_{hl}) + k\tau] ds = 0, \tag{12}$$

$$\int_{c_l} [(x_2\partial_\nu - x_1\partial_s)(\chi_\tau + \chi_0 + \sum_{h=1}^{3}\sum_{l=1}^{n} C_{hl}\chi_{hl}) + k\tau] ds = 0, \tag{13}$$

$$\int_{c_l} [(x_1\partial_\nu + x_2\partial_s)(\chi_\tau + \chi_0 + \sum_{h=1}^{3}\sum_{l=1}^{n} C_{hl}\chi_{hl}) + k\tau] ds = 0. \tag{14}$$

The function given in equation (4) should have to satisfy the equations (7)-(14). We consider the resultant force and moment vanish on each boundary. Consequently, for pure thermal problem of zero traction on the boundary gives $\chi_0 = 0$, and we are taking $l = 1$. Thus (4) will take form as

$$\chi = \chi_\tau + \sum_{h=1}^{3} C_{hl}\chi_{hl} = 0. \tag{15}$$

Boundary conditions for the temperature

$$\tau = 0, \quad \textit{on the elliptic region}, \tag{16}$$

$$\tau_{,\nu} = 0, \quad \textit{on the hexagon region}. \tag{17}$$

To discuss thermal stresses and temperature distribution around the elliptic hole, use of elliptic co-ordinates is advantageous,therefore we are introducing the elliptic coordinates as (α, β) are defined for $0 \le \alpha \le \infty,\ 0 \le \beta \le 2\pi,\ x_1 = c\sinh\alpha\cosh\beta,\ x_2 = c\cosh\alpha\sinh\beta$,

$$\alpha = \sinh^{-1}\sqrt{\frac{x_1^2 + x_2^2 - c^2 + \sqrt{(x_1^2 + x_2^2 - c^2)^2 + 4x_1^2c^2}}{2c^2}},$$
$$\beta = \cosh^{-1}\sqrt{-\frac{x_1^2 + x_2^2 - c^2 - \sqrt{(x_1^2 + x_2^2 - c^2)^2 + 4x_1^2c^2}}{2c^2}}.$$

The coordinate α is constant, and $\alpha = \alpha_1$ on an ellipse of semi axes, $c\sinh\alpha_1$ and taking the semi axes as a and b. Hence c and α_1 are calculated as $c^2 = b^2 - a^2$ and $\alpha_1 = \tanh^{-1}\frac{a}{b}$, and

$$x_1 + \imath x_2 = c\sinh(\alpha + \imath\beta). \tag{18}$$

Now any complex quantity can be written in the form, $J\cos\theta + \imath J\sin\theta$, where J and θ are real. This together with equation (18) gives

$$J^2 = c^2(\cosh 2\alpha + \cos 2\beta), \quad \tan\theta = \tanh\alpha\tan\beta.$$

The expressions for thermal stress components given by equation (2), in the elliptic co-ordinates are

$$\sigma_{\alpha\alpha} = J^2\frac{\partial^2\chi}{\partial\beta^2} - J\frac{\partial J\partial\chi}{\partial\alpha\partial\alpha} + J\frac{\partial J\partial\chi}{\partial\beta\partial\beta}, \tag{19}$$

$$\sigma_{\beta\beta} = J^2\frac{\partial^2\chi}{\partial\alpha^2} - J\frac{\partial J\partial\chi}{\partial\alpha\partial\alpha} - J\frac{\partial J\partial\chi}{\partial\beta\partial\beta}, \tag{20}$$

$$\sigma_{\alpha\beta} = -J^2\frac{\partial^2\chi}{\partial\alpha\partial\beta} - J\frac{\partial J\partial\chi}{\partial\alpha\partial\beta} + J\frac{\partial J\partial\chi}{\partial\alpha\partial\beta}. \tag{21}$$

The expression for steady heat conduction with a constant heat generation given by (3) in elliptic coordinate will become

$$J^2\nabla^\star\tau = -\frac{q}{\lambda}. \tag{22}$$

The displacement equations given by (5)-(6) in elliptic co-ordinates will be written as

$$J^2\nabla^\star J^2\nabla^\star\chi_\tau = kJ^2\nabla^\star\tau. \tag{23}$$

$$J^2\nabla^\star J^2\nabla^\star\chi_{hl} = 0. \tag{24}$$

Boundary equations given by (7)-(11) in elliptic co-ordinates on the boundary (m=1,0) are

$$(\chi_\tau)_{p_m} = (\frac{\partial}{\partial n}\chi_\tau)_{p_m} = \frac{\partial}{\partial\alpha}\chi_\tau(\nabla\alpha\cdot n) = 0, \tag{25}$$

$$(\chi_{11},\chi_{21},\chi_{31})_{p_m} = (c\sinh\alpha\cos\beta, c\cosh\alpha\sin\beta, 1)\,\delta_{1m}, \tag{26}$$

$$(\frac{\partial}{\partial n}\chi_{11}, \frac{\partial}{\partial n}\chi_{21}, \frac{\partial}{\partial n}\chi_{31})_{p_m} = (c\cosh\alpha\cos\beta, c\sinh\alpha\sin\beta, 1)\,\delta_{1m}. \tag{27}$$

Michell's conditions given by (12)-(14) in elliptic co-ordinates will become

$$\int_{\alpha=\alpha_1}\frac{\partial}{\partial\alpha}[J^2\nabla^\star(\chi_\tau + \sum_{h=1}^{3}C_{hl}\chi_{hl}) + k\tau]d\beta = 0, \tag{28}$$

$$\int_{\alpha=\alpha_1}(\cosh\alpha\sin\beta\frac{\partial}{\partial\alpha} - \sinh\alpha\cos\beta\frac{\partial}{\partial\beta})[J^2\nabla^\star(\chi_\tau + \sum_{h=1}^{3}C_{hl}\chi_{hl}) + k\tau]d\beta = 0, \tag{29}$$

$$\int_{\alpha=\alpha_1}(\cosh\alpha\sin\beta\frac{\partial}{\partial\beta} + \sinh\alpha\cos\beta\frac{\partial}{\partial\alpha})[J^2\nabla^\star(\chi_\tau + \sum_{h=1}^{3}C_{hl}\chi_{hl}) + k\tau]d\beta = 0, \tag{30}$$

and boundary conditions given by (16)-(17) will become

$$\tau = 0, \quad on \quad the \quad elliptic\,(inner) \quad region, \tag{31}$$

$$\frac{\partial\tau}{\partial n} = \frac{\partial\tau}{\partial\alpha}(\nabla\alpha\cdot\hat{n}) = 0, \quad on \quad the \quad hexagon\,(outer) \quad region, \tag{32}$$

$$\nabla = J^2\nabla^\star, \qquad \nabla^\star = \frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\beta^2}.$$

3 Solution of the problem

Introducing a new variable τ_r and τ_s such that

$$\tau = \tau_r + \tau_s$$

together with equation (22) gives

$$\nabla^\star \tau_s = 0, \tag{33}$$

$$J^2 \nabla^\star \tau_r = -\frac{q}{\lambda}. \tag{34}$$

The general plane harmonic temperature distribution in elliptic co-ordinates is expressed in the following series

$$\phi = \bar{A}_0 + \bar{B}_0\alpha + \sum_{n=1}^{\infty}(\bar{A}_{2n}\cosh 2n\alpha + \bar{B}_{2n}\sinh 2n\alpha)\cos n\beta +$$

$$\sum_{n=1}^{\infty}(\bar{C}_{2n}\cosh 2n\alpha + \bar{D}_{2n}\sinh 2n\alpha)\sin n\beta. \tag{35}$$

Assuming symmetry of the region about x_1 and x_2-axis, the solution for τ_s is given as follows

$$\tau_s = \bar{A}_0 + \bar{B}_0\alpha + \sum_{n=1}^{\infty}(\bar{A}_{2n}\cosh 2n\alpha + \bar{B}_{2n}\sinh 2n\alpha)\cos n\beta, \tag{36}$$

where $\bar{A}_0, \bar{B}_0, \bar{A}_{2n}$ and $\bar{B}_{2n}$ are unknown constants. From equation (22) particular solution is

$$\tau_r = -\frac{q^2}{8\lambda}(\cosh 2\alpha - \cos 2\beta), \tag{37}$$

Therefore

$$\tau = -\frac{q^2}{8\lambda}(\cosh 2\alpha - \cos 2\beta) + \bar{A}_0 + \bar{B}_0\alpha + \sum_{n=1}^{\infty}(\bar{A}_{2n}\cosh 2n\alpha + \bar{B}_{2n}\sinh 2n\alpha)\cos 2n\beta. \tag{38}$$

From the consideration for Michell's conditions, the coefficients $\bar{A}_0$, $\bar{A}_{2n}$ and $\bar{B}_{2n}$ vanish in the integration of the equation as to continuity of the displacement on the boundary of the hole. These coefficients do not appear in the expressions of stress components. In our problem coefficients appearing in expression of temperature distribution are of less importance, except $\bar{B}_0$.

As the outer hexagon boundary is thermally insulated under the steady state conditions, the amount of heat generation must carry away by inner elliptic boundary. The condition of thermal insulation on outer boundary is

$$\lambda \int \frac{\partial \tau}{\partial n} ds = q \int (p \tan\frac{\pi}{p} - \pi ab), \tag{39}$$

where p represents the sides of polygon. Equation (39) together with (38) solved to obtain the value of $\bar{B}_0$ as follows

$$\bar{B}_0 = \frac{\sqrt{3}}{\lambda\pi} q. \tag{40}$$

Now to calculate stress function we are introducing a new stress function

$$\chi_\tau = \chi_{\tau r} + \chi_{\tau s}. \tag{41}$$

The equation (1) together with equation (41) will become

$$J^2 \nabla^\star J^2 \nabla^\star (\chi_{\tau r} + \chi_{\tau s}) = \frac{kq}{\lambda}. \tag{42}$$

By solving the equation (42), we get particular solution as

$$\chi_{\tau r} = \frac{kqc^4}{512\lambda}(\cosh 4\alpha + \cos 4\beta) \tag{43}$$

and

$$\nabla\nabla\chi_{\tau s} = 0. \tag{44}$$

We consider the symmetry of the region about x_1 and x_2- axis. The general solution for $\chi_{\tau s}$ is given by

$$\begin{aligned}\chi_{\tau s} &= A_{00} + B_{00}\alpha + C_{00}(\cosh 2\alpha - \cos 2\beta) + D_{00}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha) + \\ &\sum_{n=1}^{\infty} [A_{2n0} \cosh 2n\alpha \cos 2n\beta + B_{2n0} \sinh 2n\alpha \cos 2n\beta + C_{2n0}(\cosh(2n+2)\alpha \cos 2n\beta + \\ &B_{2n0} \cosh(2n+2)\alpha \cos 2n\beta) + D_{2n0}(\sinh(2n+2)\alpha \cos 2n\beta + \\ &B_{2n0} \sinh(2n+2)\alpha \cos 2n\beta)].\end{aligned} \tag{45}$$

Equations (43) and (45) together give the expression for the stress function χ_τ as

$$\begin{aligned}\chi_{\tau} &= A_{00} + B_{00}\alpha + C_{00}(\cosh 2\alpha - \cos 2\beta) + D_{00}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha) + \\ &\sum_{n=1}^{\infty} [A_{2n0} \cosh 2n\alpha \cos 2n\beta + B_{2n0} \sinh 2n\alpha \cos 2n\beta + C_{2n0}(\cosh(2n+2)\alpha \cos 2n\beta + \\ &B_{2n0} \cosh(2n+2)\alpha \cos 2n\beta) + D_{2n0}(\sinh(2n+2)\alpha \cos 2n\beta + \\ &B_{2n0} \sinh(2n+2)\alpha \cos 2n\beta)] + \frac{kqc^4}{512\lambda}(\cosh 4\alpha + \cos 4\beta).\end{aligned} \tag{46}$$

The remaining three constants C_{hl} in equation (4) are to be determined so as to satisfy the three relations of Michell's conditions. Symmetry of the region about x_1 and x_2 axis, temperature distribution of the body and Michell's conditions give

$$C_{11} = C_{12} = 0, \quad C_{31} = -\frac{8D_{00} + kB_0c^2}{8D_{03}}. \tag{47}$$

Similarly, symmetry of the region about x_1 and x_2 axis, the general solution for χ_{hl}, formed by the terms which satisfy the biharmonic equation in elliptic co-ordinates are given respectively as

$$\chi_{11} = A_{01} + B_{01} + C_{01}(\cosh 2\alpha - \cos 2\beta) + D_{01}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha)+$$
$$A_{11}(\cosh 3\alpha \cos \alpha + B_{11}\alpha \sinh \alpha \cos \beta) + D_{11}(\sinh 3\alpha \cos \beta - \alpha \sinh \alpha \cos 3\beta)+$$
$$\sum_{n=1}^{\infty}[A_{n1} \cosh n\alpha \cos n\beta + B_{n1} \sinh n\alpha \cos n\beta + C_{n1}(\cosh(n+2)\alpha \cos n\beta -$$
$$\cosh n\alpha \cos(n+2)\beta) + D_{n1}(\sinh(n+2)\alpha \cos n\beta - \sinh n\alpha \cos(n+2)\beta)], \tag{48}$$

$$\chi_{21} = E_{12}\alpha \sinh \alpha \sin \beta + F_{12}\alpha \cosh \alpha \sin \beta +$$
$$G_{12}(\cosh 3\alpha \sin \beta - \cosh \alpha \sin 3\beta) + H_{12}(\sinh 3\alpha \sin \beta - \alpha \sinh \alpha \sin 3\beta)+$$
$$\sum_{n=1}^{\infty}[E_{n2} \cosh n\alpha \sin n\beta + G_{n2}(\cosh(n+2)\alpha \sin n\beta - \cosh n\alpha \cos(n+2)\beta)+$$
$$H_{n2}(\sinh(n+2)\alpha \sin n\beta - \sinh n\alpha \sin(n+2)\beta)], \tag{49}$$

$$\chi_{31} = A_{03} + B_{03} + C_{03}(\cosh 2\alpha - \cos 2\beta) + D_{03}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha)+$$
$$\sum_{n=1}^{\infty}[A_{2n3} \cosh 2n\alpha \cos 2n\beta + C_{2n3}(\cosh(2n+2)\alpha \cos 2n\beta - \cosh 2n\alpha \cos(2n+2)\beta)+$$
$$D_{2n3}(\sinh(2n+2)\alpha \cos 2n\beta - \sinh 2n\alpha \cos(2n+2)\beta)]. \tag{50}$$

Therefore from relation (15), we have

$$\chi = \chi_\tau + C_{31}\chi_{31}, \tag{51}$$

which gives expression for χ as

$$\chi = A_{00} + B_{00}\alpha + C_{00}(\cosh 2\alpha - \cos 2\beta) + D_{00}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha)+$$
$$\sum_{n=1}^{\infty}[A_{2n0} \cosh 2n\alpha \cos 2n\beta + B_{2n0} \sinh 2n\alpha \cos 2n\beta + C_{2n0}(\cosh(2n+2)\alpha \cos 2n\beta +$$
$$B_{2n0} \cosh(2n+2)\alpha \cos 2n\beta) + D_{2n0}(\sinh(2n+2)\alpha \cos 2n\beta +$$
$$B_{2n0} \sinh(2n+2)\alpha \cos 2n\beta)] + \frac{kqc^4}{512\lambda}(\cosh 4\alpha + \cos 4\beta) - \frac{8D_{00} + kB_0c^2}{8D_{03}} \times$$
$$[A_{03} + B_{03} + C_{03}(\cosh 2\alpha - \cos 2\beta) + D_{03}(\alpha \cosh 2\alpha - \alpha \cos 2\beta - \sinh 2\alpha)+$$
$$\sum_{n=1}^{\infty}[A_{2n3} \cosh 2n\alpha \cos 2n\beta + C_{2n3}(\cosh(2n+2)\alpha \cos 2n\beta - \cosh 2n\alpha \cos(2n+2)\beta)+$$

$$D_{2n3}(\sinh(2n+2)\alpha\cos 2n\beta - \sinh 2n\alpha\cos(2n+2)\beta)]], \tag{52}$$

where $A_{ij}, B_{ij}, C_{ij}, D_{ij}$, $(i = 0,1,2,....n, j = 0,3)$, are constants appearing in thermal stress function. Substituting (45) and (48) into (25)-(27), we show that χ_{31} and χ_τ must satisfy the boundary conditions around elliptic hole (perimeter of elliptic hole) and outer edge of hexagon. As these functions have been derived in order to satisfy the requirements of symmetry of the region about both x_1 and x_2- axes it is only necessary to consider the conditions of one quadrant of the region and as these functions are expressed in the forms of infinite series, the conditional equations to get the unknown constants become infinite. For this purpose the numerical calculations performed to get the unknown constants $A_{ij}, B_{ij}, C_{ij}, D_{ij}$, $(i = 0,1,2,....n, j = 0,3)$ become enormous. Therefore, we use the point-matching technique to satisfy the boundary conditions. That is, if we replace $\sum_{n=1}^{\infty}$ in equation (45) and (48) by $\sum_{n=1}^{n}$ approximately, the temprature and stress functions contain $4(N+1)$ unknown constants. Hence we have to solve $4(N+1)$ simultaneous equations.

We have obtained numerical values for unknown constants as follows

$$A_{00} = -4.47137\times10^{15}, B_{00} = -4.47137\times10^{15}, C_{00} = 2.6966\times10^{14}, D_{00} = -9.43448\times10^{13},$$

$$A_{20} = 7.12448\times10^{13}, B_{20} = -7.1724\times10^{13}, C_{20} = -0.463993, D_{20} = 0.490715,$$

$$A_{03} = 1, B_{03} = C_{03} = D_{03} = A_{23} = B_{23} = C_{23} = D_{23} = 0.$$

The expressions for stress components $\sigma_{\alpha\alpha}$, $\sigma_{\beta\beta}$ and $\sigma_{\alpha\beta}$ are obtained substituting from (52) into (19)-(21).

4 Numerical calculations and conclusion

To analyze the results given here, we consider a numerical example. The results depict isothermals for the distributions of temperature and thermal stresses. For this purpose, we take steel as thermoelastic material. The values for the different physical parameters arising in the analysis in SI units are:

$Thermal\ \ Conductivity,\ \ \lambda = 19.5W/m^oC,$
$Specific\ \ heat\ \ at\ \ constant\ \ volume,\ \ q = 0.560\times10^3 J/kg^oC,$
$Linear\ \ thermal\ \ expansion,\ \ \varepsilon = 17.7\times10^{-6}\ \ {}^oC,$
$Young's\ \ modulus,\ \ E = 195\times10^9 Pa.$

Figure 4.2 exhibits the isothermals for the temperature distributions. The region $OABC$ in Figure 4.2, represents a quadrant of hexagon region with an elliptic hole. The distribution of temperature is shown around an elliptic hole of semi-axes $a = 0.495$ and $b = 0.505$. We see that contours are moving with the increase in distance.

Figure 4.3 depicts the variation of tangential stress $\sigma_{\alpha\beta}$ around elliptic hole with same semi-axes i.e. $a = 0.495$ and $b = 0.505$ and variation in thermal stresses in this case also occur with distance. It is observed contour lines are moving with the variation in distance.

Figure 4.4 depicts variation of principal stress $\sigma_{\alpha\alpha}$ with distance along x_1-axis while Figure 4.5 depicts the variation of principal stress $\sigma_{\beta\beta}$ with respect to distance along x_2-axis from elliptic hole. It can be seen from Figure 4.4 and Figure 4.5 that principal stresses increase with distance but in opposite fashion.

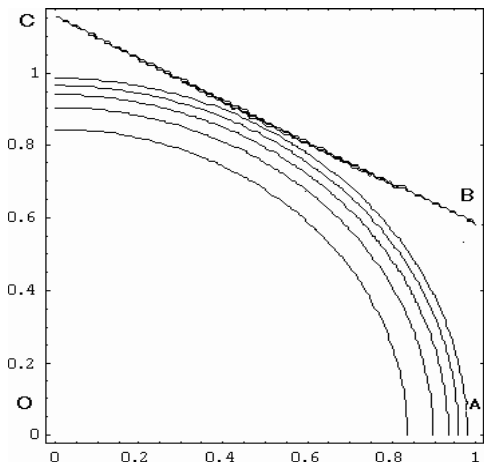


Figure 4.2: Isothermals for the elliptic hole.

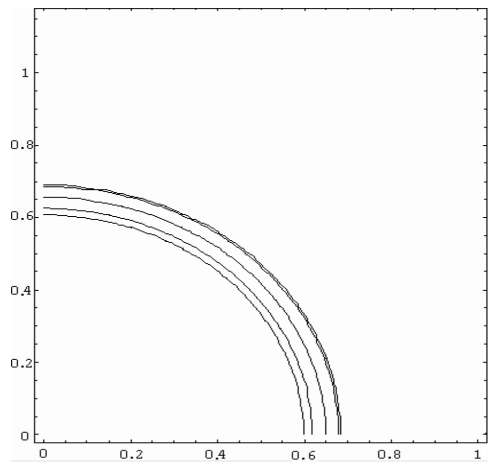


Figure 4.3: Variation of tangential stress $\sigma_{\alpha\beta}$.

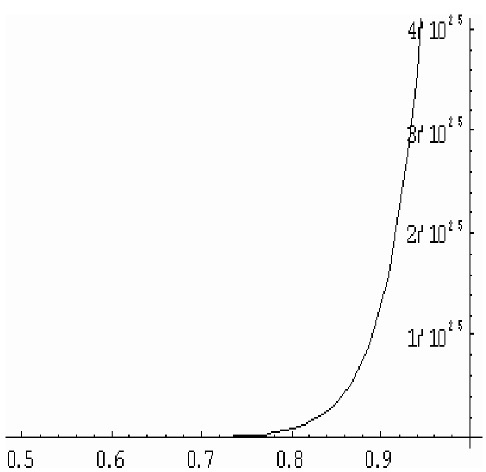


Figure 4.4: Variation of principal stress $\sigma_{\alpha\alpha}$.

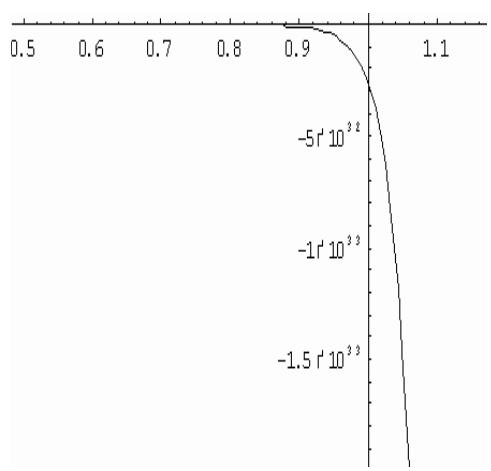


Figure 4.5: Variation of principal stress $\sigma_{\beta\beta}$.

We conclude that the isothermals of the temperature distribution around an elliptic hole within a hexagon region under constant heat generation shows that the variation in temperature occurs with distance and the pattern of variation is the same in the temperature and tangential stress $\sigma_{\alpha\beta}$ around elliptic hole. The variation in principal stress $\sigma_{\alpha\alpha}$ on x_1-axis follows the same behaviour as $\sigma_{\beta\beta}$ on x_2-axis but in opposite direction.

Acknowledgement

I express my sincere gratitude to Professor Harinder Singh, Panjab University Chandigarh, for guidance and encouragement during the preparation of this paper.

References

[1] Jeffery, G.B. Plane Stress and Strain in bipolar co-ordinates. *Trans. Roy. Soc., Series A* **221**(1921) p. 265.

[2] Neuber, H. *Kerbspannungslehre. 2. Aufl.* Springer, Berlin, 1958.

[3] Savin, G.N. *Spannungserhohung am Rande von Lochern*. VEB Verlag Technik, Berlin, 1956.

[4] Milne-Thomson, L.M. *Plane elastic systems.* Springer, Berlin, 1968.

[5] Buchwaid, V.T. and Davies, G.A.O. Plane elastostatic boundary value problems of doubly connected regions I, II. *Quart. Journ. Mech. and Appl. Math.* **17** (1964) 1-15 and 271–277.

[6] Takeouti, Y., Sumi, N. and Sugano, Y. Thermal stresses in a square region with an elliptic hole under a heat generation. In: *Structural Mechanics in Reactor Technology.* Proceedings of 2-nd International Conference, Berlin, Germany, Vol. V, Part L (1973) 10-14.

[7] Takeouti, Y. Thermal stresses in circular disc due to instantaneous line heat source. *ZAMM* **45:8** (1965) 177–184.

[8] Takeouti, Y. Thermal stresses in circular disc due to instantaneous point heat source. *Bulletin of JSME* **10** (1967) 423–429.

[9] Flynn, P.D. and Durelli, J. Discussion of stress distribution around hyderostatically loaded circular holesin the neighbourhood corners. *Trans. of ASME* **E 26** (1959) 150–152.

[10] Takeouti, Y. and Noda, N. Thermal stresses in heat generating cylinder containing a polygon hole. *ZAMM* **50**(1970) p. 587.

[11] Sekiya, T. and Sumi, N. The use of analytic continuation in solutions of plane thermoelastic problem on doubly connected regions. *Proceedings of the First International Conference on Structural Mechanics in Reactor Technology.* Berlin, M 7/3 (1971) p. 503.

[12] Deresiewicz, H. Thermal stress in a plate due to disturbance of uniform heat flow by a hole of general shape. *J. Appl. Mech.* **28** (1961) 147–149.

[13] Florence, A. L. and Goodier, J. N. Thermal stresses due to disturbance of uniform heat flow by an insulated ovaloid hole. *J. Appl. Mech.* **27** (1960) 635–639.

[14] Chowdhury, P. Thermal stresses due to uniform temperature distributed over a band of the cylindrical hole in an infinite body. *Appl. Sci.* **8** (1959) 474–478.

[15] Verba, I. I. and Kolyano, Yu. M. State of thermal stress and strain of a plate weakened by a rectangular hole. *J. Appl. Math. Mech.* **51**(3) (1987) 365–370.

[16] Pan, S. K. Thermal stresses in an infinite elastic plate containing two unequal circular holes the boundaries of which are kept at different temperatures. *Bull. Calcutta Math. Soc.* **55** (1963) 63–77.

[17] Chao, Ching-Kong, Gao, B. Mixed boundary-value problems of two-dimensional anisotropic thermoelasticity with elliptic boundaries. *Internat. J. Solids Structures* **38** (2001) 5975–5994.

[18] Takeouti, Y. and Tsuyoshi Sekiya. Thermal stresses in a polygon cylinder with a circular hole under internal heat generation. *ZAMM* **48**(1968) 237–246.

[19] Sekiya, T. and Takeouti, Y. Approximate solution of thermal stress problem in plane elasticity. *Proc. 8-th Japan Natl. Congr. App. Mech.* (1958) p. 119.

[20] Michell, J.H. *Proc. London Math Society* (1921) p. 305.

Nonlinear Dynamics and Systems Theory, **6**(3) (2006) 257–279

Duality in Distributed-Parameter Control of Nonconvex and Nonconservative Dynamical Systems with Applications

David Yang Gao *

Department of Mathematics & Department of Engineering Science and Mechancis,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061, USA.

Received: May 7, 2005; Revised: June 18, 2006

Abstract: Based on a newly developed canonical dual transformation methodology, this paper presents a potentially useful duality theory and method for solving fully nonlinear distributed-parameter control problems. The extended Lagrange duality and the interesting triality theory proposed recently in finite deformation theory are generalized into nonconvex dissipative Hamiltonian systems. It is shown that in canonical dual phase space, the solutions of chaotic systems form an invariant set. Thus, an important bifurcation criterion is proposed, which leads to an effective dual feedback control against chaotic vibrations. Applications are illustrated by a large deformation "smart" beam structure with both shear/damping actuators, and a dissipative Duffing system.

Keywords: *Duality; control theory; chaos; nonconvex analysis; Hamiltonian system.*

Mathematics Subject Classification (2000): 37K05, 37K45.

1 Problems and Motivations

We shall study a duality approach for solving the following very general abstract distributed parameter problem (($\mathcal{P}$) for short),

$$(\mathcal{P}): \quad \rho(u_{,tt} + \nu u_{,t}) + A(u, \mu) = 0 \;\; \forall u \in \mathcal{U}_k, \tag{1}$$

where the feasible space $\mathcal{U}_k$ is a convex, non-empty subset of a reflexive Banach space $\mathcal{U}$ over an open space-time domain $\Omega_t = \Omega \times (0, t_c) \subset R^n \times R^+$, in which certain essential boundary-initial conditions are prescribed. We assume that for a given distributed parameter control field $\mu(x,t)$ over Ω_t, the mapping $A(u, \mu)$ is a potential operator from

* Corresponding author: gao@vt.edu

$\mathcal{U}_k$ into its dual space $\mathcal{U}^*$, i.e., there exists a Gâteaux differentiable potential functional $P_\mu(u) = P(u;\mu)$, such that the directional derivative of P_μ at $\bar{u} \in \mathcal{U}_k$ in the direction δu can be written as

$$\delta P_\mu(\bar{u}; \delta u) = \langle DP_\mu(\bar{u}), \delta u \rangle \ \ \forall \delta u \in \mathcal{U}_k,$$

where the operator $DP_\mu(\bar{u}) = A(\bar{u}, \mu)$ is the Gâteaux derivative of P_μ at the point $\bar{u}$; the bilinear form $\langle \cdot, \cdot \rangle : \mathcal{U} \times \mathcal{U}^* \to R$ places $\mathcal{U}$ and $\mathcal{U}^*$ in duality. By nonlinear operator theory we know that the mapping $A : \mathcal{U}_k \to \mathcal{U}^*$ is monotone if P is convex on $\mathcal{U}_k$.

The problem $(\mathcal{P})$ is said to be *exactly controllable* if for certain given initial data $(u_0(x), v_0(x))$ in $\mathcal{U}_k$ and the final state $(\bar{u}_c(x), \bar{v}_c(x))$ there exists suitable control function $\mu(x,t)$ such that the solution $u(x,t)$ of the problem $(\mathcal{P})$ satisfies

$$u(x, t_c) = \bar{u}_c(x), \ \ u_{,t}(x, t_c) = \bar{v}_c(x) \ \ \forall x \in \Omega. \tag{2}$$

Dually, the problem $(\mathcal{P})$ is said to be *observable* if, for certain given input control $\mu(x,t)$, there exists an output function $h(u)$ such that the initial state $(u_0(x), v_0(x))$ can be uniquely determined from the output $h(u(x,t))$ over any interval $0 < t < t_c$. These dual concepts play a crucial role in many control system design methodologies that have evolved since the early 1960's, such as pole placement, LQG (H^2), H^∞ and minimum time optimization, realization theory, adaptive control, and system identification.

The abstract form of problem $(\mathcal{P})$ covers a great variety of situations. Very often, the total potential $P_\mu(u)$ can be written as

$$P_\mu(u) = \Phi_\mu(u, \Lambda(u)) = W_\mu(\Lambda(u)) - F_\mu(u),$$

where Λ is a Gâteaux differentiable operator from $\mathcal{U}$ into another Banach space $\mathcal{E}$; the functional $W_\mu(\xi)$ is the so-called stored (or internal) potential; while the functional $F_\mu(u)$ represents the external potential of the system.

In convex Hamiltonian systems, the total potential $P_\mu(u)$ is convex and its Gâteaux derivative $A(u,\mu) = DP_\mu(u)$ is usually an elliptic operator in conservative problems. In linear field theory of mathematical physics, Λ is usually a gradient-like operator, say $\Lambda = \text{grad}$, and $W_\mu(\xi)$ is a quadratic functional, for example,

$$P_\mu(u) = \int_\Omega \frac{1}{2} a(x) |\nabla u|^2 \, d\Omega - F_\mu(u),$$

where $a(x) > 0 \ \ \forall x \in \Omega$. In this case, the governing equation (1) reads

$$\rho(u_{,tt} + \nu u_{,t}) = \nabla \cdot (a(x) \nabla u) + DF_\mu(u) \ \ \forall (x,t) \in \Omega_t. \tag{3}$$

It is a linear wave equation if $F_\mu(u)$ is a linear functional, say $F_\mu(\mu) = \langle u \ , \ u^*(\mu) \rangle$, where $u^*(\mu)$ is a given function of the input control field $\mu(x,t)$. If $F_\mu(u)$ is nonlinear, then the governing equation (3) is semi-linear. Due to the efforts of more than thirty years research by many well-known mathematicians and scientists, the mathematical theory for distributed-parameter control systems have been well-established for convex Hamiltonian systems governed by partial differential equations with substantial applications in mechanics and structures (see, for examples, Lasiecka and Triggiani, 1999). In linear systems, there exists a very elegant duality relationship between the controllability and observability (see Dolecki and Russell, 1977).

Duality is a fundamental concept that underlies almost all natural phenomena. In classical optimization and calculus of variation, duality methods possess beautiful theoretical properties, potentially powerful alternative performances and wonderful relationships to many other fields. The associated theory and extremality principles have been well studied for convex static and Hamiltonian systems (cf. e.g., Toland, 1978, 1979; Auchmuty, 1983-2000; Strang, 1986; Rockafellar and Wets, 1997). There is a rapidly growing interest in studying and applications of convex duality theory in optimal control (cf., e.g., Mossino (1975), Chan and Ho (1979), Chan (1985), Chan and Yung (1987), Barron (1990), Tanimoto (1992), Lee and Yung (1997), Bergounioux *et al.* (1999), Arada and Raymond (1999) and many others). The interesting one-to-one analogy between the optimal control and engineering structural mechanics was discussed by Zhong *et al.* (1993, 1999). Recently, the so-called primal-dual interior-point (PDIP) method has been considered as a revolution in linear constrained optimization problems (cf. e.g., Gay *et al.*, 1998; Wright, 1998). It was shown by Helton *et al.* (1998) that the fundamental H^∞ optimization problem of control can be naturally treated with the PDIP methods.

However, the beautiful duality relationship in convex Hamiltonian systems is broken in nonconvex problems. In many applications of engineering and sciences, the total potential of system is usually nonconvex and even nonsmooth. The exact controllability and stability for nonconvex/nonsmooth systems are fundamentally difficult. For example, in the shear-damping control of large deformed beam structures, the actuators could be certain piezoelectric materials attached to the upper and lower beam surfaces, or distributed "smart" dampers (see Figure 1.1). The external signals effect changes of the properties of these actuators in such way that they produce shear forces $\mu^{\pm}(x,t)$ and damping force $\nu w_{,t}$. Thus, $\mu^{\pm}(x,t)$ and ν are, in effect, the applied distributed-control, and the composite beam/actuator system is then an instance of an active, or "smart" structure.

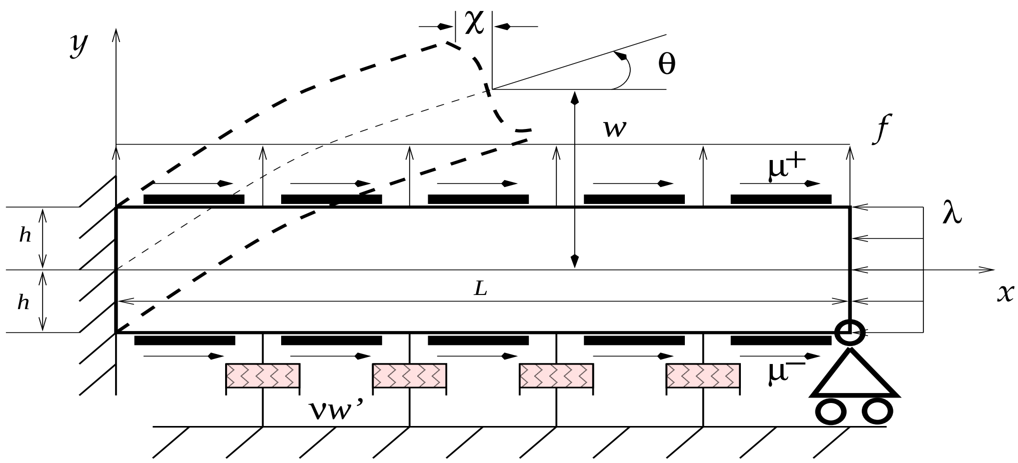


Figure 1.1. Large deformed beam with shear actuators and dampers.

Since the repeated operation of these actuator devices results large shear deformations, the traditional Timoshenko beam model can not be used to the study of these phenomena because it assumes that the shear deformation is a function of x and t alone and does not vary in the lateral beam direction. In order to study control problems of smart structures, several extended beams models have been proposed recently (see Gao *et al.*, 1997-2000), where the state variable space $\mathcal{U} = C^1(\Omega_t; R^2)$ is a displacement space over the space time domain $\Omega_t = (0,\ell) \times (-h,h) \times (0,t_c)$. The element

$u = \{\chi(x,y,t), w(x,t)\} \in \mathcal{U}$ is a continuous, differentiable vector in R^2 with domain Ω_t, where $\chi(x,y,t)$ measures the shear deformation of the beam at the point (x,y), while $w(x,t)$ is the deflection of the beam. In the case that the elastic beam subjected to the transverse load $f(x,t)$ undergone moderately large deformation, the total potential is a nonconvex functional (Gao, 2000a)

$$\begin{aligned} P_\mu(\chi, w) &= \frac{1}{2}\int_\Omega [(\chi_{,x}^2 + \frac{1}{2}\alpha w_{,x}^2 - \lambda)^2 + \beta(\chi_{,y} + w_{,x})^2]\, \mathrm{d}\Omega \\ &\quad - \int_0^\ell (\mu^+(x,t)\chi(x,h,t) + \mu^-(x,t)\chi(x,-h,t) + f(x,t)w)\, \mathrm{d}x. \end{aligned}$$

If the beam is clamped at $x = 0$, simply supported at $x = \ell$, and is subjected to a compressive load at $x = \ell$, the kinematical admissible space $\mathcal{U}_k \subset \mathcal{U}$ can be defined as

$$\mathcal{U}_k = \left\{ \begin{pmatrix} \chi \\ w \end{pmatrix} \in \mathcal{U} \;\middle|\; \begin{array}{l} w(0,t) = w(\ell,t) = 0,\ \chi(0,y,t) = \chi_{,x}(\ell,y,t) = 0; \\ (\chi, w) = (\chi_0, w_0),\ \ (\chi_{,t}, w_{,t}) = (\dot\chi_0, \dot w_0) \text{ at } t = 0 \end{array} \right\},$$

where (χ_0, w_0) and $(\dot\chi_0, \dot w_0)$ are initial conditions. In this case, the abstract governing equation (1) is a coupled nonlinear partial differential system

$$\begin{aligned} &\rho(w_{,tt} + \nu w_{,t}) = \left(\tfrac{3\alpha^2}{2} w_{,x}^2 + \beta - \lambda\alpha\right) w_{,xx} + \tfrac{\beta}{2h}|\chi_{,x}|_{\pm h} + f\ \forall (x,t) \in (0,\ell)\times(0,t_c), \\ &\chi_{,xx} + \beta\chi_{,yy} = 0,\ \ \forall (x,y,t) \in \Omega_t, \\ &\chi_{,y}(x,\pm h,t) + w_{,x}(x,t) = \pm\mu^\pm(x,t),\ \ \forall (x,t) \in (0,\ell)\times(0,t_c), \end{aligned} \tag{4}$$

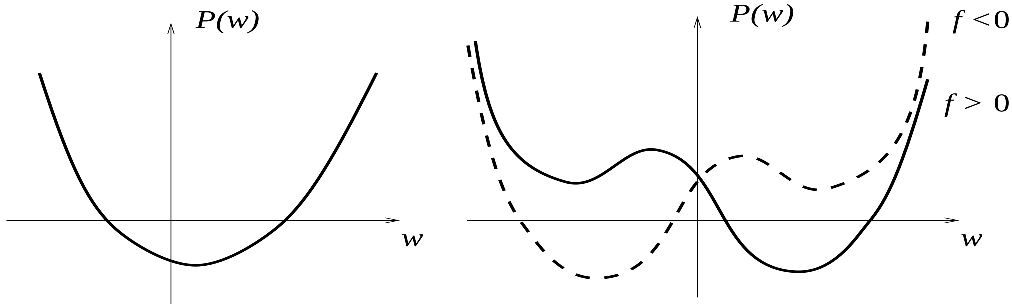
where $\alpha, \beta > 0$ are given material constants, $\lambda \in R$ represents the axial load, and $|\chi_{,x}|_{\pm h} = \chi_{,x}(x,h,t) - \chi_{,x}(x,-h,t)$ is the difference of the top and bottom shear displacements. This coupled nonlinear partial differential system is a typical example in finite deformation mechanics. Since the total potential of this system is nonconvex, the system is very sensitive to initial conditions, driving forces and numerical methods adopted. If the shear deformation can be ignored, the total potential can simply be written as

$$P(w) = \int_I \frac{1}{2}\left(\frac{1}{2}w_{,x}^2 - \lambda\right)^2\, \mathrm{d}x - \int_I fw\, \mathrm{d}x. \tag{5}$$

Clearly, if the beam is subjected to extension, then $\lambda < 0$ and the total potential $P(w)$ is strictly convex (see Figure 1.2 a). It possesses at most one global minimizer. In this case, the system is stable. However, for compressive axial load, $\lambda > 0$, the total potential $P(w)$ is a so-called double-well energy (see Figure 1.2 b). In static buckling problem, this nonconvex potential has three critical points: two local minimizers, corresponding to two possible stable buckled states, and one local maximizer, corresponding to an unstable buckled state. The global minimizer depends on the lateral load f.

If the compressed beam is subjected to a periodically dynamical load $f(x,t)$, the two local minimizers of P_μ become extremely unstable, and the beam is in dynamical post-buckling state. If the deflection $w(x,t)$ can be separated variables such that $w(x,t) = u(t)v(x)$, this post-buckling dynamical beam model leads to the well-known Duffing equation

$$u_{,tt} + \nu u_{,t} = au(\lambda - \frac{1}{2}u^2) + \mu(t), \tag{6}$$



(a) $\lambda < 0$ (b) $\lambda > 0$

Figure 1.2. Convex and double-well potentials.

where $a > 0$ is a parameter. It is known that this equation is extremely sensitive to the initial conditions. For certain given parameters λ, ν, and driving force $\mu(t)$, this equation possesses the so-called chaotic solutions. Figure 5.4 shows that even for the same given data, different numerical methods produce totally different results.

Control theory in finite deformation mechanics has emerged as the most challenging and active research field in recent years. Mathematically speaking, the total potentials of large deformed structures are generally nonconvex, or even nonsmooth. Very small perturbations of the system's initial conditions and parameters may lead the system to different operating points with significantly different performance characteristics. This is the one of main reasons why the traditional perturbation analysis, the direct approaches and many standard control techniques cannot successfully be applied to nonconvex systems. Based upon these observations and in order to handle the nonlinear problem, a school of new techniques has been developed (see, e.g., Fowler, 1989; Ott *et al.*, 1990; Chen and Dong, 1993; Ogorzalek, 1993; Antoniou *et al.*, 1996; Ghezzi and Piccardi, 1997; Koumboulis and Mertzios, 1996, 2000).

Duality theory in fully nonlinear variational problems was originally studied by Gao and Strang (1989) for large deformation nonsmooth mechanics. In order to recover the broken symmetry in fully nonlinear systems (see Definition 2.2), a so-called *complementary gap function* was introduced. It was realized in post-buckling analysis of nonlinear beam theory (Gao, 1997) that this function recovered the duality gap between the nonconvex primal problems and the Fenchel-Rockafellar dual problems. A self-contained comprehensive presentation of the mathematical theory in general nonconvex systems was given recently by Gao (2000d), wherein, a so-called *canonical dual transformation method* and associated triality theory have been proposed for solving nonconvex/nonsmooth variational-boundary value problems. Recent results show that certain very difficult constrained nonconvex problems in global optimization can be solved completely by this method (see Gao, 2003, 2005). Compared with the traditional analytic methods and direct approaches, the main advantages of this canonical dual transformation method are the following:

(1) it converts nonconvex/nonsmooth constrained variational problems into smooth unconstrained dual problems;

(2) it transforms certain fully nonlinear partial differential equations into algebraic

systems;

(3) it provides powerful and efficient primal-dual alternative approaches.

The aim of the present article is to generalize the author's previous results on nonconvex variational problems into nonconservative distributed-parameter control systems. The rest of this paper is divided into four main sections. The next section set up notations used in the paper. A general framework in fully nonlinear systems are discussed. Section 3 presents an extended Lagrangian critical point theorem and the associated triality theory in general nonconvex, nonconservative dynamical systems. The critical points in fully nonlinear systems are classified. Section 4 is devoted mainly to the construction of dual action in nonconvex dissipative Hamiltonian systems. The tri-duality proposed in static boundary value problems is generalized into control problems. Section 5 discusses application in dissipative Duffing system. A bifurcation criterion is proposed which can be used for feedback controlling against chaotic vibrations.

2 Framework for Canonical Systems and Classification

Let $\mathcal{U}$ and $\mathcal{U}^*$ be two real linear spaces, placed in duality by a bilinear form $\langle u, u^* \rangle : \mathcal{U} \times \mathcal{U}^* \to R$. Let $P : \mathcal{U}_s \to R$ be a given functional, well-defined on a convex domain $\mathcal{U}_s \subset \mathcal{U}$ such that for any given $u \in \mathcal{U}_s$, $P(u)$ is Gâteaux differentiable. Thus, the Gâteaux derivative DP of P at $u \in \mathcal{U}_s$ is a mapping from $\mathcal{U}_s$ into $\mathcal{U}^*$. Let $\mathcal{U}_s^* \subset \mathcal{U}^*$ be the range of the mapping $DP : \mathcal{U}_s \to \mathcal{U}^*$. If the relation $u^* = DP(u)$ is reversible on $\mathcal{U}_s$, then for any given $u^* \in \mathcal{U}_s^*$, the classical Legendre conjugate functional $P^* : \mathcal{U}_s^* \to R$ of $P(u)$ is defined by

$$P^*(u^*) = \{\langle u, u^* \rangle - P(u)| \ \ u^* = DP(u)\}.$$

The conjugate pair (u, u^*) is called the *canonical duality pair* on $\mathcal{U}_s \times \mathcal{U}_s^* \subset \mathcal{U} \times \mathcal{U}^*$ if and only if the equivalent relations

$$u^* = DP(u) \ \ \Leftrightarrow \ \ u = DP^*(u^*) \ \ \Leftrightarrow \ \ P(u) + P^*(u^*) = \langle u, u^* \rangle. \tag{7}$$

hold on $\mathcal{U}_s \times \mathcal{U}_s^*$.

The following notations and definitions, used in Gao (2000c,d), will be of convenience in nonconvex control problems.

Definition 2.1 The set of functionals $P : \mathcal{U} \to R$ which are either convex or concave is denoted by $\Gamma(\mathcal{U})$. In particular, let $\check{\Gamma}(\mathcal{U})$ denote the subset of functionals $P \in \Gamma(\mathcal{U})$ which are convex and $\hat{\Gamma}(\mathcal{U})$ the subset of $P \in \Gamma(\mathcal{U})$ which are concave.

The *canonical functional space* $\Gamma_G(\mathcal{U}_s)$ is a subset of functionals $P \in \Gamma(\mathcal{U}_s)$ which are Gâteaux differentiable on $\mathcal{U}_s \subset \mathcal{U}$, such that the relation $u^* = DP(u)$ is reversible for any given $u \in \mathcal{U}_s$. $\diamondsuit$

Clearly, if $P \in \Gamma_G(\mathcal{U}_s)$ and $\mathcal{U}_s^*$ is the range of the mapping $DP : \mathcal{U}_s \to \mathcal{U}^*$, then the canonical duality relations (7) hold on $\mathcal{U}_s \times \mathcal{U}_s^*$.

Let $(\mathcal{E}, \mathcal{E}^*)$ be an another pair of real linear spaces paired in duality by the second bilinear form $\langle \cdot \, ; \, \cdot \rangle : \mathcal{E} \times \mathcal{E}^* \to R$. The so-called *geometrical operator* $\Lambda : \mathcal{U} \to \mathcal{E}$ is a continuous, Gâteaux differentiable operator such that for any given $u \in \mathcal{U}_a \subset \mathcal{U}$, there exists an element $\xi \in \mathcal{E}_a \subset \mathcal{E}$ satisfying the *geometrical equation*

$$\xi = \Lambda(u).$$

The directional derivative of ξ at $\bar{u}$ in the direction $u \in \mathcal{U}$ is then defined by

$$\delta\xi(\bar{u}; u) := \lim_{\theta \to 0^+} \frac{\xi(\bar{u} + \theta u) - \xi(\bar{u})}{\theta} = \Lambda_t(\bar{u})u, \tag{8}$$

where $\Lambda_t(\bar{u}) = D\Lambda(\bar{u}) : \mathcal{U} \to \mathcal{E}$ denotes the Gâteaux derivative of the operator Λ at $\bar{u}$. For a given $\xi^* \in \mathcal{E}^*$, $G_\Lambda(u) = \langle \Lambda(u) \; ; \; \xi^* \rangle$ is a real-valued functional of u on $\mathcal{U}$. Its Gâteaux derivative at $\bar{u} \in \mathcal{U}$ in the direction $u \in \mathcal{U}$ reads

$$\delta G_\Lambda(\bar{u}; u) = \langle \Lambda_t(\bar{u})u \; ; \; \xi^* \rangle = \langle u \; , \; \Lambda_t^*(\bar{u})\xi^* \rangle,$$

where $\Lambda_t^*(\bar{u}) : \mathcal{E}^* \to \mathcal{U}^*$ is the adjoint operator of Λ_t associated with the two bilinear forms.

Let $\mathcal{V}$ and $\mathcal{V}^*$ be velocity and momentum spaces, respectively, placed in duality by the third bilinear form $\langle * \; , \; * \rangle : \mathcal{V} \times \mathcal{V}^* \to R$. For Newtonian systems, the kinetic energy $K : \mathcal{V} \to R$ and its Legendre conjugate $K^* : \mathcal{V}^* \to R$ are quadratic forms

$$K(v) = \int_\Omega \frac{1}{2}\rho v^2 \, \mathrm{d}\Omega, \quad K^*(p) = \int_\Omega \frac{1}{2}\rho^{-1} p^2 \, \mathrm{d}\Omega.$$

Thus the canonical physical relations between $\mathcal{V}$ and $\mathcal{V}^*$ are linear:

$$p = DK(v) = \rho v \quad \Leftrightarrow \quad v = DK^*(p) = \rho^{-1} p.$$

Let $\mathcal{V}_a \subset \mathcal{V}$ be an admissible velocity space, in which certain essential initial/boundary conditions are given, say

$$\mathcal{V}_a = \{v \in \mathcal{V} | \;\; v(x, 0) = v_0 \;\; \forall x \in \Omega\}. \tag{9}$$

Finally, we let $\mathcal{M}$ be an admissible control space over Ω_t. For any given $\mu \in \mathcal{M}$, we assume that there exists a Gâteaux differentiable functional $\Phi_\mu : \mathcal{U}_a \times \mathcal{E}_a \subset \mathcal{U} \times \mathcal{E} \to R$, such that the total potential $P(u; \mu)$ of the system can be written as

$$P_\mu(u) = P(u; \mu) = \Phi_\mu(u, \Lambda(u)). \tag{10}$$

Thus, for a dissipative dynamical system with linear damping, the total action of the system is a weighted nonconvex functional

$$\Pi_\mu(u) = \int_0^{t_c} e^{\nu t} \left[K(\partial_t u) - \Phi_\mu(u, \Lambda(u))\right] \mathrm{d}t, \tag{11}$$

which is well-defined on the feasible space $\mathcal{U}_k$ given by

$$\mathcal{U}_k = \{u \in \mathcal{U}_a | \;\; \Lambda(u) \in \mathcal{E}_a, \;\; \partial_t u \in \mathcal{V}_a\}. \tag{12}$$

For the linear time-differential operator $\partial_t = \partial/\partial t$, its formal adjoint associated with this weighted functional is an affine operator $\partial_t^* = -\partial/\partial t - \nu$ (see Gao (2000d)).

The following classification for distributed parameter control systems was originally introduced in nonlinear variational/boundary value problems by Gao (1998, 2000d,2000).

Definition 2.2 Suppose that for the problem $(\mathcal{P})$ given in (1), the associated total potential $P_\mu(u)$ is well-defined on its domain $\mathcal{U}_s \subset \mathcal{U}$. If the geometrical operator Λ :

$\mathcal{U} \to \mathcal{E}$ can be chosen such that $P_\mu(u) = \Phi_\mu(u, \Lambda(u))$, $\Phi_\mu \in \Gamma_G(\mathcal{U}_a) \times \Gamma_G(\mathcal{E}_a)$ and $\mathcal{U}_s = \{u \in \mathcal{U}_a | \ \Lambda(u) \in \mathcal{E}_a\}$, then

(1) the transformation $\{P; \mathcal{U}_s\} \to \{\Phi_\mu; \mathcal{U}_a \times \mathcal{E}_a\}$ is called the *canonical transformation*, and $\Phi_\mu : \mathcal{U}_a \times \mathcal{E}_a \to R$ is called the *canonical functional associated with* Λ;

(2) the problem $(\mathcal{P})$ is called *geometrically nonlinear (or linear)* if $\Lambda : \mathcal{U} \to \mathcal{E}$ is nonlinear (or linear); it is called *physically nonlinear* (resp. linear) if the duality mapping $D\Phi_\mu : \mathcal{U}_a \times \mathcal{E}_a \to \mathcal{U}_a^* \times \mathcal{E}_a^*$ is nonlinear (resp. linear); it is called *fully nonlinear* if it is both geometrically and physically nonlinear. ♢

The canonical transformation plays a fundamental role in duality theory of nonconvex systems. Clearly, if $\Phi_\mu \in \Gamma_G(\mathcal{U}_a) \times \Gamma_G(\mathcal{E}_a)$ is a canonical functional, the Gâteaux derivative $D\Phi_\mu : \mathcal{U}_a \times \mathcal{E}_a \to \mathcal{U}_a^* \times \mathcal{E}_a^* \subset \mathcal{U}^* \times \mathcal{E}^*$ is a monotone mapping, i.e., the duality relations

$$u^* = D_u \Phi_\mu(u, \xi), \quad \xi^* = D_\xi \Phi_\mu(u, \xi) \tag{13}$$

are reversible between the paired spaces $(\mathcal{U}_a, \mathcal{U}_a^*)$ and $(\mathcal{E}_a, \mathcal{E}_a^*)$, where $D_u\Phi_\mu$ and $D_\xi \Phi_\mu$ denote partial Gâteaux derivatives of Φ_μ with respect to u and ξ, respectively. Thus, on $\mathcal{U}_k$ the directional derivative of P_μ at $\bar{u}$ in the direction $u \in \mathcal{U}_k$ can be written as

$$\begin{aligned} \delta P_\mu(\bar{u}; u) &= \langle u \ , \ D_u\Phi_\mu(\bar{u}, \Lambda(\bar{u}))\rangle + \langle \Lambda_t(\bar{u}) u \ ; \ D_\xi \Phi_\mu(\bar{u}, \Lambda(\bar{u}))\rangle \\ &= \langle u \ , \ \bar{u}^* \rangle + \langle u \ ; \ \Lambda_t^*(\bar{u})\bar{\xi}^* \rangle \quad \forall u \in \mathcal{U}_k. \end{aligned}$$

In terms of canonical variables, the governing equation (1) for fully nonlinear problems can be written in the *tri-canonical forms*, namely,

$$\begin{array}{lll} \text{(1) geometrical equations:} & v = \partial_t u, \ \ \xi = \Lambda(u), \\ \text{(2) physical relations:} & p = \rho v, \ \ (u^*, \xi^*) = D\Phi_\mu(u, \xi), \\ \text{(3) balance equation:} & \partial_t^* p - u^* - \Lambda_t^*(u)\xi^* = 0. \end{array} \tag{14}$$

The framework for fully nonlinear systems is shown in Figure 2.1. Extensive illustrations of the canonical transformation and the tri-canonical forms in mathematical physics and variational analysis were given in the monograph by Gao (2000).

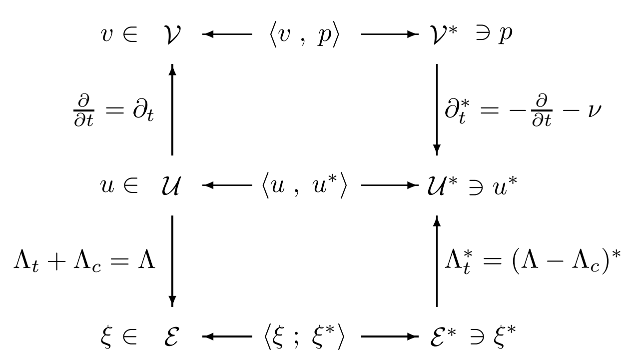


Figure 2.1. Framework in fully nonlinear Newtonian systems with linear damping.

In geometrically linear systems, where $\Lambda : \mathcal{U} \to \mathcal{E}$ is linear, we have $\Lambda = \Lambda_t$. For dynamical problems, if the total potential P_μ is convex, the total action associated with the problem $(\mathcal{P})$ is a *d.c. functional*, i.e., the difference of convex functionals:

$$\Pi_\mu(u) = \int_0^{t_c} e^{\nu t}[K(\partial_t u) - P_\mu(u)]\,\mathrm{d}t.$$

It was shown by Gao (2000d) that the critical point of Π_μ either minimizes or maximizes Π_μ over the kinetically admissible space. The classical Hamiltonian associated with this d.c. functional Π_μ is a convex functional on the phase space $\mathcal{U} \times \mathcal{V}^*$, i.e.

$$H(u,p) = K^*(p) + P_\mu(u), \tag{15}$$

The classical canonical forms for convex Hamilton systems are well-known

$$\partial_t u = D_p H(u,p), \quad \partial_t^* p = D_u H(u,p).$$

Furthermore, if the canonical functional Φ_μ can be written in the form $\Phi_\mu(u,\xi) = \frac{1}{2}\langle \xi\ ;\ C\xi\rangle - F_\mu(u)$, where $C : \mathcal{E} \to \mathcal{E}^*$ is a linear symmetrical operator, then the governing equations for linear system can be written as

$$\rho(u_{,tt} + \nu u_{,t}) + \Lambda^* C \Lambda u = DF_\mu(u).$$

In mathematical physics, the geometrical mapping Λ is usually a gradient-like operator. Then $A = \Lambda^* C \Lambda$ is an elliptic operator if C is positive-definite.

In geometrically nonlinear systems, $\Lambda \neq \Lambda_t$, and the total potential $P_\mu(u)$ is usually a nonconvex functional. In this case, we have the following operator decomposition

$$\Lambda(u) = \Lambda_t(u)u + \Lambda_c(u), \tag{16}$$

where $\Lambda_c : \mathcal{U} \to \mathcal{E}$ is called the complementary operator of the Gâteaux derivative operator Λ_t. By this decomposition, we have

$$\langle \Lambda(u)\ ;\ \xi^* \rangle = \langle u\ ,\ \Lambda_t^*(u)\xi^* \rangle - G(u,\xi^*), \tag{17}$$

where $G : \mathcal{U} \times \mathcal{E}^* \to R$ is so-called *complementary gap functional*, defined by

$$G(u,\xi^*) = \langle -\Lambda_c(u)\ ;\ \xi^* \rangle : \mathcal{U} \times \mathcal{E}^* \to R. \tag{18}$$

This functional was first introduced by Gao and Strang (1989) in finite deformation theory to recover a broken symmetry in geometrical nonlinear systems. It is now understood that this gap functional plays a key role in extremality analysis of nonconvex variational problems.

As a typical example in nonconvex dynamical systems, let us consider the following nonconvex variational problem over the domain $\Omega_t = (0,\ell) \times (0,t_c)$:

$$\Pi_\mu(u) = \int_{\Omega_t} e^{\nu t}\left[\frac{1}{2}\rho u_{,t}^2 - \frac{1}{2}a(\frac{1}{2}u_{,x}^2 - \mu)^2 + uf\right]\,\mathrm{d}x\,\mathrm{d}t \quad \to \text{sta} \ \ \forall u \in \mathcal{U}_k, \tag{19}$$

where $a,\ \mu$ are given positive constants. This nonconvex problem also appears very often in phase transitions and hysteresis.

First, we let $\Lambda = \partial/\partial x$ be a linear operator, and $P_\mu(u) = W_\mu(\Lambda u) - F_\mu(u)$ with

$$W_\mu(\epsilon) = \int_0^\ell \frac{1}{2}a(\frac{1}{2}\epsilon^2 - \mu)^2\ \mathrm{d}x, \quad F_\mu(u) = \int_0^\ell uf\ \mathrm{d}x.$$

Thus, $W_\mu(\epsilon)$ is the so-called van der Waals' double-well function of the linear "strain" $\epsilon = u_{,x}$. Since $W_\mu(\epsilon)$ is not a canonical functional, the constitutive equation $\epsilon^* = DW_\mu(\epsilon)$ is not one-to-one. Thus, the Legendre conjugate of $W_\mu(\epsilon)$ does not have a simple algebraic expression. The Fenchel conjugate $W_\mu^*(\epsilon^*)$ of the double-well energy $W_\mu(\epsilon)$, defined by

$$W_\mu^*(\epsilon^*) = \sup_\epsilon \{\langle \epsilon \, ; \, \epsilon^* \rangle - W_\mu(\epsilon)\},$$

is always a convex, lower semi-continuous functional. However, the well-known Fenchel-Young inequality

$$W_\mu(u_{,x}) \geq \langle u_{,x} \, ; \, \epsilon^* \rangle - W_\mu^*(\epsilon^*)$$

leads to a so-called duality gap between the primal problem and the Fenchel-Rockafellar dual problem (see Gao, 2000d). This nonzero duality gap indicates that the well-established Fenchel-Rockafellar duality theory can be used only for solving convex variational problems.

From the theory of continuum mechanics we know that in finite deformation problems, $\epsilon = u_{,x}$ is not a strain measure (it does not satisfy the *axiom of material frame-indifference* (cf. e.g., Gao, 2000d)). In order to recover this duality gap, we need to choose a suitable geometrical operator Λ, say, $\Lambda(u) = \frac{1}{2}u_{,x}^2 - \mu$, so that the nonconvex problem (19) can be put in our framework. In continuum mechanics, this quadratic measure $\xi = \Lambda(u)$ is a Cauchy-Green type strain. Thus, in terms of u and ξ, $\Phi_\mu(u, \xi) = W_\mu(\xi) - F_\mu(u) = \frac{1}{2}\langle \xi \, ; \, a\xi \rangle - \langle u \, , \, f \rangle$ is a canonical functional. The Legendre conjugate of the quadratic functional $W_\mu(\xi) = \frac{1}{2}\langle \xi \, ; \, a\xi \rangle$ is simply defined by $W^*(\xi^*) = \frac{1}{2}\langle a^{-1}\xi^* \, ; \, \xi^* \rangle$. The operator decomposition (16) for this quadratic operator reads

$$\Lambda(u) = \Lambda_t(u)u + \Lambda_c(u), \quad \Lambda_t(u)u = u_{,x}u_{,x}, \quad \Lambda_c(u) = -\frac{1}{2}u_{,x}^2 - \mu.$$

The complementary gap functional associated with this quadratic operator is a quadratic functional of u

$$G(u, \xi^*) = \langle -\Lambda_c(u) \, ; \, \xi^* \rangle = \int_0^\ell \frac{1}{2}u_{,x}^2 \xi^* \, \mathrm{d}x.$$

For homogeneous boundary conditions, we have

$$\langle \Lambda_t(u)u \, ; \, \xi^* \rangle = \int_0^\ell u_{,x}u_{,x}\xi^* \, \mathrm{d}x = -\int_0^\ell u(u_{,x}\xi^*)_{,x} \, \mathrm{d}x = \langle u \, , \, \Lambda_t^*(u)\xi^* \rangle,$$

which leads to the adjoint operator Λ_t^* of Λ_t. Thus, the tri-canonical equations for this nonconvex problem can be listed as the following.

$$v = \partial_t u, \quad \xi = \frac{1}{2}au_{,x}^2 - \mu,$$

$$p = \rho v, \quad \xi^* = DW_\mu(\xi) = a\xi, \quad u^* = DF_\mu(u) = f,$$

$$p_{,t} + \nu p = -\Lambda_t^*(u)\xi^* + u^* = (u_{,x}\xi^*)_{,x} + f.$$

Since the geometrical operator Λ is nonlinear, and the canonical constitutive equations are linear, the nonconvex problem (19) is a geometrically nonlinear system.

3 Extended Lagrangian and Triality Theory

The triality theory in nonconvex problems was originally proposed by the author (Gao, 1997, 1999, 2000) in static finite deformation theory and global optimization. In this section, we will generalize this interesting result into fully nonlinear dynamical systems. We assume that for a given fully nonlinear system, there exists a Gâteaux differentiable operator $\Lambda : \mathcal{U}_a \to \mathcal{E}_a$ such that the total potential of the system can be written as

$$P_\mu(u) = W_\mu(\Lambda(u)) - F_\mu(u), \tag{20}$$

where $W_\mu \in \check{\Gamma}_G(\mathcal{E}_a)$ is a convex canonical functional, while $F_\mu : \mathcal{U}_a \to R$ is a linear functional. Thus, the primal problem $(\mathcal{P})$ can be reformulated as the following.

Problem 3.1 (Primal Distributed-Parameter Control Problem) For a given primal feasible space $\mathcal{U}_k = \{u \in \mathcal{U}_a| \quad \partial_t u \in \mathcal{V}_a, \quad \Lambda(u) \in \mathcal{E}_a\}$ and the final state $(\bar{u}_c(x), \bar{v}_c(x))$, find the control field $\mu(x,t) \in \mathcal{M}$ such that the solution $\bar{u}(x,t)$ of the variational problem

$$(\mathcal{P}) : \quad \Pi_\mu(u) = \int_0^{t_c} e^{\nu t}[K(\partial_t u) - W_\mu(\Lambda(u)) + F_\mu(u)]\,\mathrm{d}t \to \text{sta} \ \ \forall u \in \mathcal{U}_k \tag{21}$$

satisfying the controllability condition

$$(\bar{u}(x,t_c), \bar{u}_{,t}(x,t_c)) = (\bar{u}_c(x), \bar{v}_c(x)) \ \ \forall x \in \Omega.$$

It is easy to check that the criticality condition $D\Pi_\mu(\bar{u}) = 0$ leads to the the canonical governing equation

$$\rho(\bar{u}_{,tt} + \nu \bar{u}_{,t}) = DF_\mu(\bar{u}) - \Lambda_t^*(\bar{u}) DW_\mu(\Lambda(\bar{u})). \tag{22}$$

By the Legendre-Fenchel transformation, the conjugate of $W_\mu(\xi)$ is defined by

$$W_\mu^*(\xi^*) = \sup_{\xi \in \mathcal{E}} \{\langle \xi \,;\, \xi^* \rangle - W_\mu(\xi)\}.$$

Since $W_\mu : \mathcal{E}_a \to R$ is a convex canonical functional, $W_\mu^*(\xi^*)$ is well-defined on the range $\mathcal{E}_a^*$ of the duality mapping $DW_\mu^* : \mathcal{E}_a \to \mathcal{E}^*$, the canonical duality relation

$$\xi^* = DW_\mu(\xi) \Leftrightarrow \xi = DW_\mu^*(\xi^*) \Leftrightarrow W_\mu(\xi) + W_\mu^*(\xi) = \langle \xi \,;\, \xi^* \rangle$$

holds on $\mathcal{E}_a \times \mathcal{E}_a^*$. Moreover, we have $W_\mu^{**}(\xi) = W_\mu(\xi)$ for all $\xi \in \mathcal{E}_a$. Let $\mathcal{Z} = \mathcal{U} \times \mathcal{V}^* \times \mathcal{E}^*$ be the so-called *extended canonical phase space.*

Definition 3.1 Suppose that for a given problem $(\mathcal{P})$, there exists a Gâteaux differentiable operator $\Lambda : \mathcal{U} \to \mathcal{E}$ and canonical functionals $W_\mu \in \Gamma(\mathcal{E})$, $F_\mu \in \Gamma(\mathcal{U})$ such that $P_\mu(u) = W_\mu(\Lambda(u)) - F_\mu(u)$. Then

(1) the functional $H_\mu : \mathcal{Z} \to R$ defined by

$$H_\mu(u, p, \xi^*) = K^*(p) - W_\mu^*(\xi^*) + F_\mu(u) \in \Gamma(\mathcal{U}) \times \Gamma(\mathcal{V}^*) \times \Gamma(\mathcal{E}^*) \tag{23}$$

is called *extended canonical Hamiltonian density* associated with Π_μ;

(2) the functional $L_\mu : \mathcal{Z} \to R$ defined by

$$L_\mu(u, p, \xi^*) = \langle \partial_t u \,,\, p \rangle - \langle \Lambda(u) \,;\, \xi^* \rangle - H_\mu(u, p, \xi^*) \tag{24}$$

is called *extended Lagrangian density* of $(\mathcal{P})$ associated with Λ;

(3) the functional $\Xi_\mu : \mathcal{Z} \to R$ defined by

$$\Xi_\mu(u,p,\xi^*) = \int_0^{t_c} e^{\nu t} L_\mu(u,p,\xi^*)\,\mathrm{d}t \tag{25}$$

is called *extended Lagrangian form* of $(\mathcal{P})$. It is called *canonical Lagrangian form* if $\Xi_\mu \in \Gamma(\mathcal{U}) \times \Gamma(\mathcal{V}^*) \times \Gamma(\mathcal{E}^*)$. $\diamondsuit$

A point $(\bar{u},\bar{p},\bar{\xi}^*) \in \mathcal{Z}$ is said to be a critical point of Ξ_μ if Ξ_μ is Gâteaux-differentiable at $(\bar{u},\bar{p},\bar{\xi}^*)$ and $D\Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*) = 0$. It is easy to find out that the criticality condition $D\Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*) = 0$ leads to *canonical Lagrange equations*

$$D\Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*) = 0 \;\;\Rightarrow \left\{ \begin{array}{l} \Lambda(\bar{u}) = D_{\xi^*} W^*_\mu(\bar{\xi}^*), \quad \partial_t \bar{u} = DK^*(\bar{p}), \\ \partial_t^* \bar{p} = \Lambda_t^*(\bar{u})\bar{\xi}^* - DF_\mu(\bar{u}). \end{array} \right. \tag{26}$$

By the fact that W_μ and F_μ are canonical functionals, we know that, by the Legendre duality theory, any critical point of Ξ_μ solves the variational problem $(\mathcal{P})$.

Since $F_\mu(u) : \mathcal{U}_a \to R$ is a linear functional, by the Riesz representation theory we know that there exists an element $\bar{u}^*(\mu) \in \mathcal{U}^*$ such that $F_\mu(u) = \langle u\ ,\ \bar{u}^*(\mu)\rangle$. Thus, the extended Lagrangian associated with $(\mathcal{P})$ can be written as

$$\Xi_\mu(u,p,\xi^*) = \int_0^{t_c} e^{\nu t} \left[\langle \partial_t u\ ,\ p\rangle - \langle \Lambda(u)\ ;\ \xi^*\rangle - K^*(p) + W^*(\xi^*) + \langle u\ ,\ \bar{u}^*(\mu)\rangle\right] \mathrm{d}t. \tag{27}$$

Note that $\Xi_\mu : \mathcal{V}_a^* \times \mathcal{E}_a^* \to R$ is a saddle functional for any given $u \in \mathcal{U}_a$, we have always the equality

$$\inf_{\xi^* \in \mathcal{E}_a^*} \sup_{p \in \mathcal{V}_a^*} \Xi_\mu(u,p,\xi^*) = \sup_{p \in \mathcal{V}_a^*} \inf_{\xi^* \in \mathcal{E}_a^*} \Xi_\mu(u,p,\xi^*) \;\; \forall u \in \mathcal{U}_a. \tag{28}$$

However, for any given $(p,\xi^*) \in \mathcal{V}_a^* \times \mathcal{E}_a^*$, the convexity of $\Xi_\mu(\cdot,p,\xi^*) \to R$ depends on the operator Λ. Let $\mathcal{L}_c \subset \mathcal{Z}_a = \mathcal{U}_a \times \mathcal{V}_a^* \times \mathcal{E}_a^*$ be a critical point set of Ξ_μ, i.e.

$$\mathcal{L}_c = \{(\bar{u},\bar{p},\bar{\xi}^*) \in \mathcal{Z}_a |\;\; \delta\Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*; u,p,\xi^*) = 0 \;\; \forall (u,p,\xi^*) \in \mathcal{Z}_a\}.$$

For any given critical point $(\bar{u},\bar{p},\bar{\xi}^*) \in \mathcal{L}_c$, we let $\mathcal{Z}_r = \mathcal{U}_r \times \mathcal{V}_r^* \times \mathcal{E}_r^* \subset \mathcal{Z}_a$ be its *neighborhood* such that $(\bar{u},\bar{p},\bar{\xi}^*)$ is the only critical point on $\mathcal{Z}_r$. The following triality theorem should play an important role in the stability analysis of nonlinear dynamical systems.

Theorem 3.1 (Triality Theorem) *Suppose that for a given control field $\mu(x,t)$ such that $(\bar{u},\bar{p},\bar{\xi}^*) \in \mathcal{L}_c$ is a critical point of Ξ_μ, and $\mathcal{Z}_r$ is a neighborhood of $(\bar{u},\bar{p},\bar{\xi}^*)$.*

If $\langle \Lambda(u)\ ;\ \bar{\xi}^*\rangle$ is concave on $\mathcal{U}_r$, then on $\mathcal{Z}_r$,

$$\Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*) = \min_u \max_p \min_{\xi^*} \Xi_\mu(u,p,\xi^*) = \max_p \min_u \min_{\xi^*} \Xi_\mu(u,p,\xi^*). \tag{29}$$

However, if $\langle \Lambda(u)\ ;\ \bar{\xi}^*\rangle$ is convex on $\mathcal{U}_r$, then on $\mathcal{Z}_r$ we have either

$$\begin{aligned} \Xi_\mu(\bar{u},\bar{p},\bar{\xi}^*) &= \min_u \max_p \min_{\xi^*} \Xi_\mu(u,p,\xi^*) = \min_p \max_u \min_{\xi^*} \Xi_\mu(u,p,\xi^*) \\ &= \min_{\xi^*,u} \max_p \Xi_\mu(u,p,\xi^*) = \min_{p,\xi^*} \max_u \Xi_\mu(u,p,\xi^*). \end{aligned} \tag{30}$$

or

$$\begin{aligned}\Xi_\mu(\bar u,\bar p,\bar\xi^*) &= \max_u \min_{\xi^*} \max_p \Xi_\mu(u,p,\xi^*) = \max_p \min_{\xi^*} \max_u \Xi_\mu(u,p,\xi^*)\\ &= \min_{\xi^*} \max_{u,p} \Xi_\mu(u,p,\xi^*) = \max_{u,p} \min_{\xi^*} \Xi_\mu(u,p,\xi^*). \qquad (31)\end{aligned}$$

Proof Since $W^*_\mu \in \check\Gamma(\mathcal{E}^*_a)$, $K^* \in \check\Gamma(\mathcal{V}^*_a)$, if $\langle \Lambda(u)\ ;\ \bar\xi^*\rangle$ is concave on $\mathcal{U}_r$, then for a given $\bar\xi^*$, $\Xi_\mu \in \check\Gamma(\mathcal{U}_r) \times \hat\Gamma(\mathcal{V}^*_a)$ is a saddle functional. Thus the equality (29) follows from the saddle-Lagrangian duality theorem (cf. e.g., Gao, 2000d). However, if $\langle\Lambda(u)\ ;\ \bar\xi^*\rangle$ is convex on $\mathcal{U}_r$, then for any given $\xi^* \in \mathcal{E}^*_r$, the extended Lagrangian $\Xi_\mu \in \hat\Gamma(\mathcal{U}_r) \times \hat\Gamma(\mathcal{V}^*_a)$ is a *super-critical functional* (see Gao, 2000d). By the *super-Lagrangian duality theorem* proved in Gao (2000d), we have either (30) or (31). □

4 Dual Action and Tri-Duality in Dissipative Systems

The goal of this section is to develop a dual approach for solving the distributed parameter control problem $(\mathcal{P})$. For any given $u \in \mathcal{U}_k$, the extended Lagrangian density $\Xi_\mu(u,p,\xi^*)$ is a saddle functional on $\mathcal{V}^* \times \mathcal{E}^*$, and we have

$$\Pi_\mu(u) = \sup_{p\in\mathcal{V}^*} \inf_{\xi^*\in\mathcal{E}^*} \Xi_\mu(u,p,\xi^*) \ \ \forall u \in \mathcal{U}_k. \qquad (32)$$

On the other hand, the dual action $\Pi^d_\mu : \mathcal{V}^*_a \times \mathcal{E}^*_a \to R$ can be defined by

$$\begin{aligned}\Pi^d_\mu(p,\xi^*) &= \operatorname{sta}\{\Xi_\mu(u,p,\xi^*) |\ \ \forall u \in \mathcal{U}_a\}\\ &= F^\Lambda_\mu(p,\xi^*) - \int_0^{t_c} [K^*(p) - W^*_\mu(\xi^*)]\,dt, \ \ \forall (p,\xi^*) \in \mathcal{V}^*_a \times \mathcal{E}^*_a, \qquad (33)\end{aligned}$$

where $F^\Lambda_\mu(p,\xi^*)$ is the so-called Λ-dual functional of $F_\mu(u)$ defined by

$$F^\Lambda_\mu(p,\xi^*) = \operatorname*{sta}_{u\in\mathcal{U}_a} \int_0^{t_c} e^{\nu t}[\langle \partial_t u\ ,\ p\rangle - \langle\Lambda(u)\ ;\ \xi^*\rangle + F_\mu(u)]\,dt. \qquad (34)$$

Since $F_\mu(u) = \langle u\ ,\ \bar u^*(\mu)\rangle$ is a linear functional, for any given $(p,\xi^*) \in \mathcal{V}^*_a \times \mathcal{E}^*_a$ and the applied control $\mu \in \mathcal{M}$, the solution $\bar u$ of this stationary problem (34) satisfies the balance equation

$$\partial^*_t p - \Lambda^*_t(\bar u)\xi^* + \bar u^*(\mu) = 0 \quad \text{in } \Omega_t. \qquad (35)$$

For geometrically linear conservative systems, where Λ is a linear operator, we have

$$F^\Lambda_\mu(p,\xi^*) = up|_{t=0}^{t=t_c}, \quad s.t.\ \ \Lambda^*\xi^* + p_{,t} = \bar u^*(\mu). \qquad (36)$$

In this case,

$$\Pi^d_\mu(p,\xi^*) = up|_{t=0}^{t=t_c} + \int_0^{t_c} e^{\nu t}[W^*_\mu(\xi^*) - K^*(p)]\,dt \qquad (37)$$

is the classical complementary action in linear engineering dynamical systems (see Tabarrok and Rimrott, 1994) defined on the dual feasible space

$$\mathcal{T}_s = \{(p,\xi^*) \in \mathcal{V}_a \times \mathcal{E}^*_a |\ \ p_{,t} + \Lambda^*\xi^* = \bar u^*(\mu)\}.$$

In fully nonlinear systems, we let $\mathcal{T}_s \subset \mathcal{V}_a^* \times \mathcal{E}_a^*$ be a subspace such that for any given $(p, \xi^*) \in \mathcal{T}_s$, the critical point $\bar{u}$ can be determined by (35) as $\bar{u} = \bar{u}(p, \xi^*)$ and the dual action Π_μ^d is well defined by (33). Thus, by the operator decomposition $\Lambda = \Lambda_t + \Lambda_c$, we have

$$F_\mu^\Lambda(p, \xi^*) = e^{\nu t} up|_{t=0}^{t=t_c} + \int_0^{t_c} e^{\nu t} G^d(p, \xi^*)\,\mathrm{d}t, \quad s.t.\ \ \partial_t^* p = \Lambda_t^*(\bar{u})\xi^* - u^*(\mu), \tag{38}$$

where $G^d(p, \xi^*) = \langle -\Lambda_c(\bar{u}) \ ; \ \xi^* \rangle$ is the so-called pure complementary gap functional. Then, the problem dual to the primal control problem $(\mathcal{P})$ can be proposed as the following.

Problem 4.1 (Dual Distributed-Parameter Control Problem) For a given dual feasible space $\mathcal{T}_s$ and the final state $(u_c(x), v_c(x))$, find the control field $\mu(x,t) \in \mathcal{M}$ such that the dual solution $(\bar{p}(x,t), \bar{\xi}^*(x,t))$ of the dual variational problem

$$(\mathcal{P}^d): \quad \Pi_\mu^d(p, \xi^*) \ \ \to \text{sta} \ \ \forall (p, \xi^*) \in \mathcal{T}_s \tag{39}$$

and the associated state $\bar{u}(x,t)$ satisfying the controllability condition

$$(\bar{u}(x, t_c), \rho^{-1}\bar{p}(x, t_c)) = (u_c(x), v_c(x)) \ \ \forall x \in \Omega. \tag{40}$$

Lemma 4.1 *Let $\Xi_\mu(u, p, \xi^*)$ be a given extended Lagrangian associated with $(\mathcal{P})$ and $\Pi_\mu^d(p, \xi^*)$ the dual action defined by (33). Suppose that $\mathcal{Z}_r = \mathcal{U}_r \times \mathcal{V}_r^* \times \mathcal{E}_r^*$ is an open subset of $\mathcal{Z}_a$ and $(\bar{u}, \bar{p}, \bar{\xi}^*) \in \mathcal{Z}_r$ is a critical point of Ξ_μ on $\mathcal{Z}_r$, Π_μ is Gâteaux differentiable at $\bar{u}$, and Π_μ^d is Gâteaux differentiable at $(\bar{p}, \bar{\xi}^*)$. Then $D\Pi_\mu(\bar{u}) = 0$, $D\Pi_\mu^d(\bar{p}, \bar{\xi}^*) = 0$, and*

$$\Pi_\mu(\bar{u}) = \Xi_\mu(\bar{u}, \bar{p}, \bar{\xi}^*) = \Pi_\mu^d(\bar{p}, \bar{\xi}^*). \tag{41}$$

The proof of this lemma was given by the author in parametrical variational analysis (Gao, 1998).

Lemma 4.1 shows that the critical points of the extended Lagrangian are also the critical points for both the primal and dual variational problems.

Theorem 4.1 (Tri-Duality Theorem) *Suppose that for a given control field $\mu(x,t)$ such that $(\bar{u}, \bar{p}, \bar{\xi}^*) \in \mathcal{L}_c$ is a critical point of Ξ_μ and $\mathcal{Z}_r = \mathcal{U}_r \times \mathcal{V}_r^* \times \mathcal{E}_r^*$ is a neighborhood of $(\bar{u}, \bar{p}, \bar{\xi}^*)$ such that $\mathcal{V}_r^* \times \mathcal{E}_r^* \subset \mathcal{T}_s$. If $\langle \Lambda(u) \ ; \ \bar{\xi}^* \rangle$ is concave on $\mathcal{U}_r$, then*

$$\Pi_\mu(\bar{u}) = \min_{u \in \mathcal{U}_r} \Pi_\mu(u) \quad \text{iff} \quad \Pi_\mu^d(\bar{p}, \bar{\xi}^*) = \max_{p \in \mathcal{V}_r^*} \min_{\xi^* \in \mathcal{E}_r^*} \Pi_\mu^d(p, \xi^*). \tag{42}$$

However, if $\langle \Lambda(u) \ ; \ \bar{\xi}^* \rangle$ is convex on $\mathcal{U}_r$, then

$$\Pi_\mu(\bar{u}) = \min_{u \in \mathcal{U}_r} \Pi_\mu(u) \quad \text{iff} \quad \Pi_\mu^d(\bar{p}, \bar{\xi}^*) = \min_{(p,\xi^*) \in \mathcal{T}_s} \Pi_\mu^d(p, \xi^*); \tag{43}$$

$$\Pi_\mu(\bar{u}) = \max_{u \in \mathcal{U}_r} \Pi_\mu(u) \quad \text{iff} \quad \Pi_\mu^d(\bar{p}, \bar{\xi}^*) = \max_{p \in \mathcal{V}_r^*} \min_{\xi^* \in \mathcal{E}_r^*} \Pi_\mu^d(p, \xi^*). \tag{44}$$

Proof This theorem can be proved by combining Lemma 4.1 and the triality theorem. □

5 Feedback Control Against Chaos in Dissipative Duffing System

As we have shown in the first section of this paper that the governing equations for shear/damping control of large deformed nonlinear beam structure are eventually equivalent to the well-known Duffing system. As a typical example, let us consider the very simple nonconvex dynamical problem over the time domain $I = (0, t_c)$

$$\Pi_\mu(u) = \int_I e^{\nu t}[\rho u'^2 - \frac{1}{2}a(\frac{1}{2}u^2 - \lambda)^2 + \mu u]\, \mathrm{d}t \quad \to \text{sta} \;\; \forall u \in \mathcal{U}_k. \tag{45}$$

For initial-value problem of this one-dimensional dynamical system, the kinematically admissible space $\mathcal{U}_k$ can simply be given as

$$\mathcal{U}_k = \{u \in \mathcal{L}^4(0,t_c) |\;\; u' \in \mathcal{L}^2(0,t_c),\;\; u(0) = u_0,\;\; u'(0) = v_0\}.$$

The criticality condition of Π_μ leads to the dissipative Duffing equation

$$\rho(u'' + \nu u') = au(\lambda - \frac{1}{2}u^2) + \mu(t), \;\; \forall t \in I, \;\; u \in \mathcal{U}_k. \tag{46}$$

In terms of the nonlinear canonical measure $\xi = \Lambda(u) = \frac{1}{2}u^2$, the energy density $W_\mu(\xi)$ and its conjugate $W^*_\mu(\varsigma)$ are convex functions:

$$W_\mu(\xi) = \frac{1}{2}a(\xi - \lambda)^2, \quad W^*_\mu(\varsigma) = \frac{1}{2a}\varsigma^2 + \lambda\varsigma.$$

The extended Lagrangian for this nonconvex system is

$$\Xi_\mu(u,p,\varsigma) = \int_I e^{\nu t}\left(pu' - \varsigma(\frac{1}{2}u^2 - \lambda) - \frac{1}{2\rho}p^2 + \frac{1}{2a}\varsigma^2 + \mu u\right) \mathrm{d}t. \tag{47}$$

The criticality condition $D_u\Xi_\mu(\bar{u},p,\varsigma) = 0$ leads to the equilibrium equation

$$p' + \nu p + \bar{u}\varsigma = \mu \;\; \forall t \in I.$$

Clearly, the critical point $\bar{u} = (\mu - p' - \nu p)/\varsigma$ is well-defined for any nonzero ς. Thus, the dual feasible space can be defined as

$$\mathcal{T}_s = \left\{ (p,\varsigma) \in \mathcal{C}^1(I) \;\middle|\; \begin{array}{l} p(0) = \rho v_0, \;\; -\lambda a \le \varsigma(t) < +\infty, \\ \varsigma(t) \neq 0 \;\; \forall t \in I, \;\; \varsigma(0) = a(\frac{1}{2}u_0^2 - \lambda) \end{array} \right\}.$$

Substituting $\bar{u} = (\mu - p' - \nu p)/\varsigma$ into Ξ^d_μ, the dual action is obtained as

$$\begin{aligned} &\Pi^d_\mu(p,\varsigma) = \underset{u\in\mathcal{U}_a}{\text{sta}}\; \Xi_\mu(u,p,\varsigma) \\ = \;\; & e^{\nu t_c}p(t_c)u(t_c) - \rho v_0 u_0 + \int_I e^{\nu t}[\frac{1}{2a}\varsigma^2 + \lambda\varsigma + \frac{(p' + \nu p - \mu)^2}{2\varsigma} - \frac{1}{2\rho}p^2]\, \mathrm{d}t, \end{aligned} \tag{48}$$

which is well defined on $\mathcal{T}_s$. The criticality condition for Π^d_μ leads to the *dual Duffing system* in the time domain $I \subset R$

$$\left(\frac{1}{\varsigma}(p' + \nu p - \mu)\right)' + \frac{1}{\rho}p = 0, \tag{49}$$

$$\varsigma^2 \left(\frac{1}{a} \varsigma + \lambda \right) = \frac{1}{2} (\mu - p' - \nu p)^2. \tag{50}$$

This system consists of the so-called *differential-algebraic equations* (DAE's), which arise naturally in many applications (cf. Brenan *et al*, 1996). Although the numerical solution of these types of systems has been the subject of intense research activity in the past few years, the solvability of each problem depends mainly on the so-called *index* of the system. Clearly, the algebraic equation (50) has zero solution $\varsigma = 0$ if and only if $g = (\mu - p' - \nu p) = 0$. Otherwise, for any nonzero $g(t) = \mu(t) - p'(t) - \nu p(t)$, the algebraic equation (50) has at most three real roots $\varsigma_i(t)$ $(i = 1, 2, 3)$, each of them leads to the state solution $u_i(t) = (\mu(t) - p_i'(t) - \nu p_i(t))/\varsigma_i(t)$.

Theorem 5.1 (Stability and Bifurcation Criteria) *For a given parameter $\lambda > 0$, initial data (u_0, v_0) and the input control $\mu(t)$, if at a certain time period $I_s \subset I = (0, t_c)$,*

$$\lambda_p(t) = \frac{3}{2} \left(\frac{\mu(t) - p'(t) - \nu p(t)}{a} \right)^{2/3} > \lambda, \quad t \in I_s \tag{51}$$

then the Duffing system possesses only one solution set $(\bar{u}(t), \bar{p}(t), \bar{\varsigma}(t))$ satisfying $\bar{\varsigma}(t) > 0 \ \ \forall t \in I_s$, and over the period I_s,

$$\Pi_\mu(\bar{u}) = \min \Pi_\mu(u) \quad \textit{iff} \quad \Pi^d_\mu(\bar{p}, \bar{\varsigma}) = \min \Pi^d_\mu(p, \varsigma), \tag{52}$$

$$\Pi_\mu(\bar{u}) = \max \Pi_\mu(u) \quad \textit{iff} \quad \Pi^d_\mu(\bar{p}, \bar{\varsigma}) = \max_p \min_\varsigma \Pi^d_\mu(p, \varsigma). \tag{53}$$

However, if at a certain time period $I_b \subset I = (0, t_c)$ we have $\lambda_p(t) < \lambda$, then, the system possesses three sets of different solutions $(\bar{u}_i, \bar{p}_i(t), \bar{\varsigma}_i(t))$, $i = 1, 2, 3$. In the case that the three solutions $\varsigma_i(t)$ are in the following ordering

$$-a\lambda \le \bar{\varsigma}_3(t) \le \bar{\varsigma}_2(t) \le 0 \le \bar{\varsigma}_1(t) \ \ \forall t \in I_b, \tag{54}$$

then over the period I_b, the solution set $(\bar{u}_1(t), \bar{p}_1(t), \bar{\varsigma}_1(t))$ satisfies either (52) or (53); while the solution sets $(\bar{u}_i(t), \bar{p}_i(t), \bar{\varsigma}_i(t))$ $(i = 2, 3)$ satisfy

$$\Pi_\mu(\bar{u}_i) = \min_u \Pi_\mu(u) = \max_p \min_\varsigma \Pi^d_\mu(p, \varsigma) = \Pi^d_\mu(\bar{p}_i, \bar{\varsigma}_i), \quad i = 2, 3. \tag{55}$$

This theorem can be proved by combining the theorem given by Gao (2000d, Theorem 3.4.4) and the triality theorem.

Remark 5.1 By Theorem 3.4.4 proved by the author (Gao, 2000d), for any given continuous function $g(t)$, if $\bar{\varsigma}_i(t)$ $(i = 1, 2, 3)$ are the three solutions of the dual Euler-Lagrange equation (50) in the order of (54), then the associated $\bar{u}_1(t)$ is a global minimizer of the total potential

$$P_\mu(u) = \int_I e^{\nu t} \left[\frac{1}{2} a (\frac{1}{2} u^2 - \lambda)^2 - g(t) u \right] \mathrm{d}t,$$

while $\bar{u}_2(t)$ is a local minimizer of P_μ and $\bar{u}_3(t)$ is a local maximizer of P_μ.

In *algebraic geometry*, the dual Euler-Lagrange equation (50) is the so-called *singular algebraic curve* in (ς, g)-space, i.e. $\varsigma = 0$ is on the curve (see Silverman & Tate, 1992,

p. 99). With a change of variables, the singular cubic curve (50) can be given by the well-known *Weierstrass equation*

$$y^2 = x^3 + \alpha x^2 + \beta x + \gamma,$$

where $\alpha,\ \beta,\ \gamma \in R$ are constants. If we let $\mathcal{C}_{ns}$ be a set consisting of non-singular points on the curve, then $\mathcal{C}_{ns}$ is an Abelian group. This fact in algebraic geometry is very important in understanding the stability of the nonconvex dynamical systems. Actually, from Figure 5.1 we can see clearly that for a given input control, if $\lambda_p(t) < \lambda$, the cubic algebraic equation (50) possesses three different real solutions for $\varsigma(t)$. The two negative solutions $\bar{\varsigma}(t)$ are the sources that lead to the chaotic motion of the system. Thus, the inequality (51) provides a *bifurcation (or chaotic) criterion* for the Duffing system. Figure 5.1 also shows that if the continuous function $g(t) = \mu(t) - p'(t) - \nu p(t)$ is one-signed on certain time interval $I_b \subset I = (0, t_c)$, each root $\bar{\varsigma}(t)$ of (50) is also one-signed on I_i.

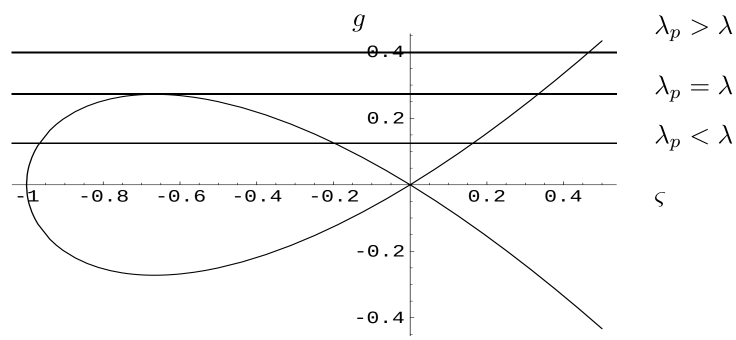


Figure 5.1. Invariant set of dual solutions and bifurcation criterion for Duffing equation (50).

Theoretically speaking, for the given same data, the Duffing equation (46) and its dual system (49-50) should have the same solution set. Numerically, the primal and dual Duffing problems give quite different results (see Figure 5.2 (a)). For the given data $a = 1, \lambda = 1.5, u_0 = 2, v_0 = 1.4$ and $\nu = 0$, Figures 5.2 and 5.3 show the numerical primal (solid line) and dual (dashed line) solutions. From the dual trajectories in the dual phase space ς-p-$p_{,t}$ (Figure 5.3 (c-d)) we can see that at the point $\varsigma_3(t) = -a\lambda$, if the function $g(t) = \mu(t) - p_{,t}(t) - \nu p(t)$ changes its sign, the state $u(t)$ crosses the t-axis and falls down to the another potential well in the phase space $\mathcal{Z} = \mathcal{U} \times \mathcal{V}^*$. The bifurcation is then occurred.

For the forced vibration with linear damping, the numerical results are extremely sensitive to the parameters. Figure 5.4 shows that the trajectories are chaotic in phase spaces q-p (Figure 5.4 (b)) and ς-p-g (Figure 5.4 (d)). However, trajectory in the dual phase space ς-g is an invariant (see Figure 5.4 (c)), which depends only on the parameters λ, a and the amplitude of the force $g(t)$.

As it is known that the nonconvex dynamical systems are very sensitive to both the parameters and numerical methods used. For the given periodic driving force $\mu(t) = 1.5\cos(2.75t)$ and $\nu = 0.1$, Figure 5.4 shows that different numerical solvers in MATLAB produce very different "chaotic results". However, solutions in dual phase space ς-g form an invariant set (Figure 5.4 (c)). This important fact shows that the triality theorem will play an important role in stability and bifurcation analysis of chaotic systems.

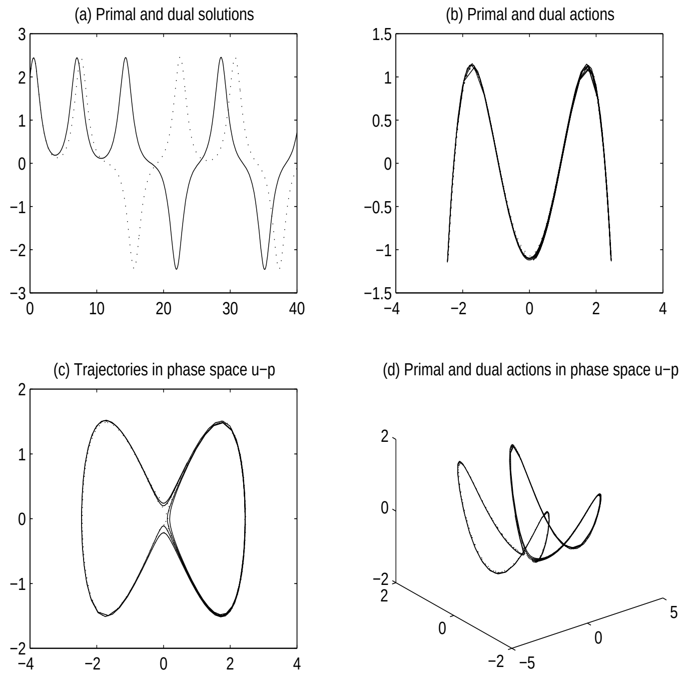


Figure 5.2. Primal and dual solutions in primal phase space.

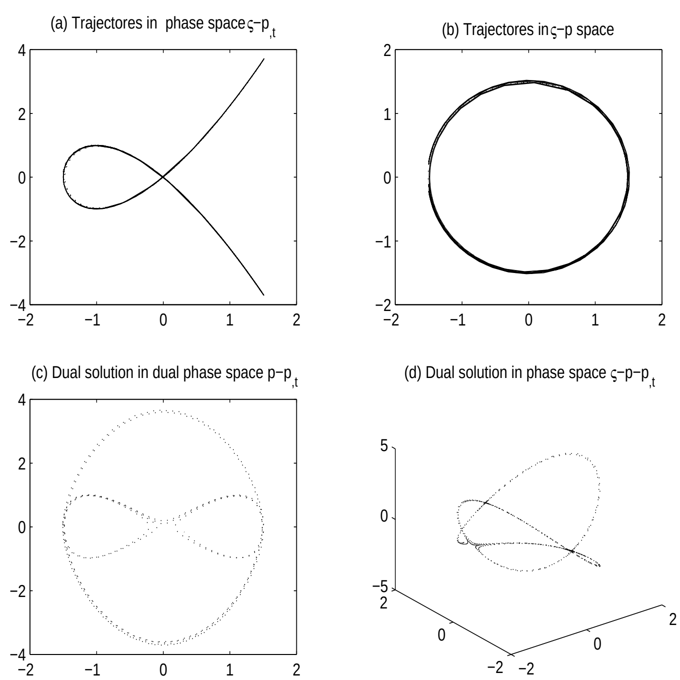
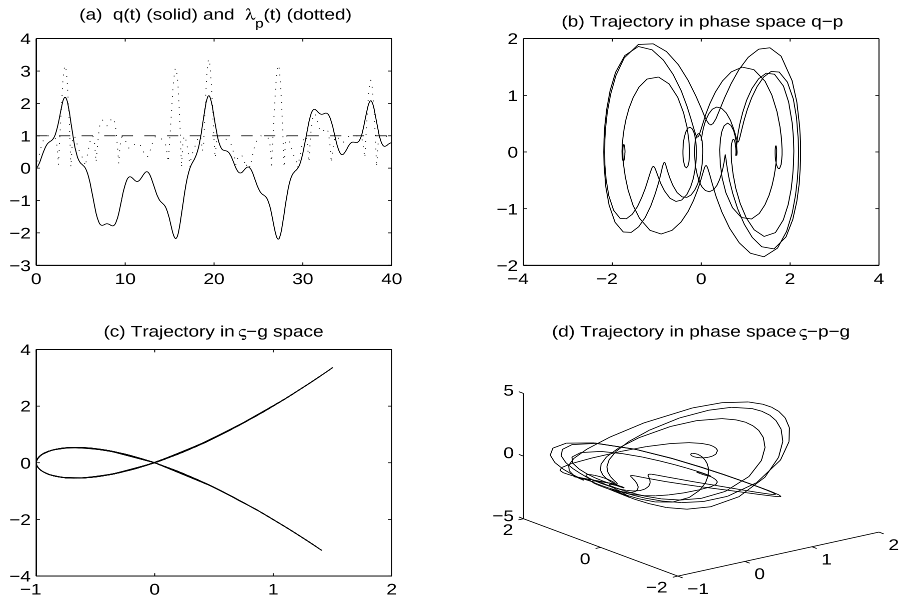
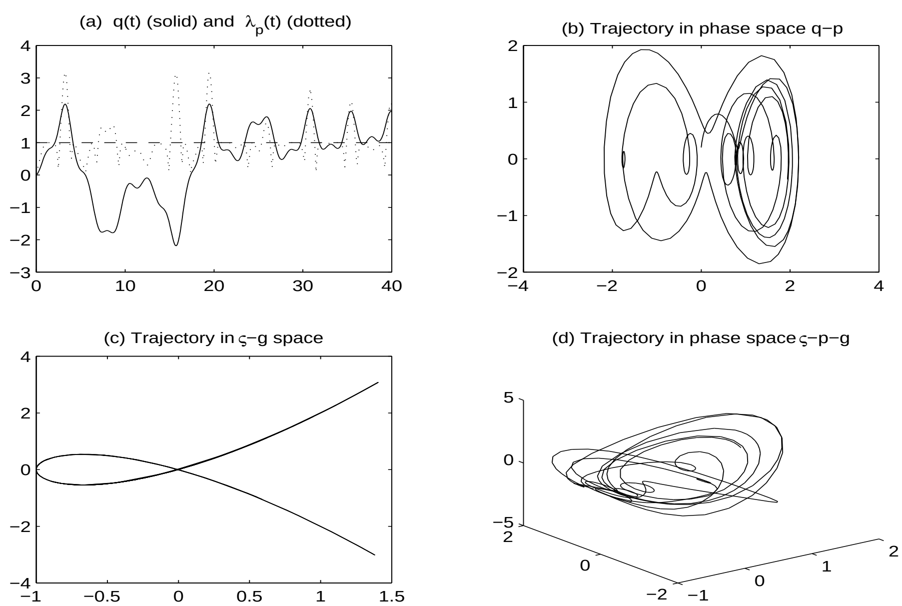


Figure 5.3. Duffing solutions in dual phase spaces.



(1) Numerical results computed by "ode23"



(2) Numerical results computed by "ode15s"

Figure 5.4. Chaos and invariant set: numerical results by two differential numerical methods in MATLAB.

Based on the canonical dual transformation method and theorems developed in this paper, the dual feedback control against the chaotic vibration of the Duffing system can be suggested as the following.

1. Periodic vibration on the whole phase plane.

Choosing the control parameters μ and ν such that the function $g(t) = \mu - p'(t) - \nu p(t)$ changes its sign at the point $\bar{\varsigma}_3(t) = -a\lambda$.

2. Unilateral vibrations on half phase planes (either $u(t) > 0$ or $u(t) < 0$).

There are two methods: (1) choosing the control parameters μ and ν such that the function $g(t) = \mu - p'(t) - \nu p(t)$ does not change its sign at the point $\bar{\varsigma}_3(t) = -a\lambda$; (2) choosing μ and ν such that

$$\lambda_p(t) = \frac{3}{2}\left(\frac{\mu(t) - p'(t) - \nu p(t)}{a}\right)^{2/3} > \lambda \ \ \forall t \in I. \tag{56}$$

By the bifurcation criterion (Theorem 5.1) we know that if $\lambda_p > \lambda$, the total potential of this dissipative Duffing equation is convex and the system is stable.

6 Concluding Remarks

The concept of duality is one of the most successful ideas in modern mathematics and science. The inner beauty of duality theory owes much to the fact that many different natural phenomena can be put in a unified trio-canonical framework (see Gao, 2000d, 2001). By the fact that the canonical physical variables appear always in pairs, the canonical dual transformation method can be used to solve many problems in natural systems. The associated extended Lagrange duality and triality theories have profound computational impacts. For any given nonlinear problem, as long as there exists a geometrical operator Λ such that the trio-canonical forms can be characterized correctly, the canonical dual transformation method and the associated triality principles can be used to establish nice theories and to develop powerful alternative algorithms for robust feedback control of chaotic systems. Actually, it has been shown that in global optimization many difficult nonconvex minimization problems in n-dimensional space can be converted into certain canonical dual problems (either convex minimization or concave maximization) in ONE-dimensional space, therefore, a class of problems have been solved completely, including the well-known quadratic minimization over a sphere (Gao, 2004), polynomial minimization (Gao, 2005), and quadratic programming with box constraints (Gao, 2006). In general n-dimensional distributed parameter systems, the dual algebraic equation (50) will be a tensor equation and the stability of the nonconvex system will depend on the eigenvalues of symmetrical canonical stress tensor field $\varsigma(x, t)$ (see Gao, 2001). The triality theory can be used for studying the controllability, observability and stability of distributed parameter control problems.

Acknowledgement

This research was supported by NSF grant CCF-0514768. Comments from two anonymous referees are gratefully acknowledged.

References

[1] Antoniou, I., Basios, V. and Bosco, F. Probabilistic control of chaos: The β-adic Renyi map under control. *Int. J. Bifurcation Chaos* **6** (1996) 1563–1573.

[2] Arada, N. and Raymond, J.-P. Minimax control of parabolic systems with state constraints. *SIAM J. Control Optim.* **38** (2000) 254–271.

[3] Auchmuty, G. Duality for non-convex variational principles. *J. Diff. Equations* **50** (1983) 80-145.

[4] Auchmuty, G. Duality algorithms for nonconvex variational principles. *Numer. Funct. Anal. and Optim.* **10** (1989) 211–264.

[5] Auchmuty, G. Variational principles for self-adjoint elliptic operators. In: *Nonsmooth/Nonconvex Mechanics: Modeling, Analysis and Numerical Methods.* (David Y. Gao, Ray W. Ogden and Georgios E. Stavroulakis, eds). Kluwer Academic Publishers, Boston, 2001, 15–42.

[6] Barron, E. N. The pontryagin maximum principle for minimax problems of optimal control, *Nonlinear Analysis* **15** (1990) 1155–1165.

[7] Bergounioux, M., Ito, K. and Kunisch, K. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.* **37** (1999) 1176–1194.

[8] Brenan, K.E., Camplell, S.L. and Pelzold, L.R. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations.* SIAM, Philadelphia, 1996.

[9] Chan, W. L. Duality in the optimal control of non-well-posed parabolic equations with positive controls. *J. Math. Analy. Appl.* **107** (1985) 509–519.

[10] Chan, W. L. and Yung, S. P. Duality theory for the linear-convex optimal control problem with delays. *IMA J. Math. Control Inform.* **4** (1987) 251–262.

[11] Chen, G., Lee, E.B., Littman, W. and Markus, L. (eds) *Distributed Parameter Control Systems.* Lecture Notes in Pure and Applied Math. Marcel Dekker, New York, 1991.

[12] Chen, G. and Dong, X. From chaos to order-perspectives and methodologies in controlling chaotic nonlinear dynamic systems. *Int. J. Bifurcation Chaos* **3** (1993) 1363–1409.

[13] Clarke, F.H. The dual action, optimal control, and generalized gradients, *Mathematical Control Theory.* Banach Center Publ. **14** (1985) PWN, Warsaw, 109–119.

[14] Dolecki, S. and Russell, D. L. A general theory of observation and control. *SIAM J. Control Opt.* **15** (1977) 185–220.

[15] Ekeland, I. *Convexity Methods in Hamiltonian Mechanics.* Springer-Verlag, 1990.

[16] Ekeland, I. and Temam, R. *Convex Analysis and Variational Problems.* North-Holland, 1976.

[17] Favini, A., Horn, M. A., Lasiecka, I. and Tataru, D. Global existence, uniqueness and regularity of solutions to a von Karman system with nonlinear boundary dissipation. *Diff. Integral Equations* **9** (1996) 267–294.

[18] Fowler, T. B. Application of stochastic control techniques to chaotic nonlinear systems. *IEEE Trans. Autom. Contr.* **34** (1989) 201–205.

[19] Gao, D. Y. Duality theory in nonlinear buckling analysis for von Karman equations. *Studies in Appl. Math.* **94** (1995) 423–444.

[20] Gao, D. Y. Dual extremum principles in finite deformation theory with applications in post-buckling analysis of nonlinear beam model. *Appl. Mech. Reviews. ASME* **50** (11) (1997) 64–71.

[21] Gao, D. Y. Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications. *IMA J. of Applied Math.* **61** (1998) 199–235.

[22] Gao, D. Y. General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. *Meccanica* **34** (1999) 169–198.

[23] Gao, D. Y. Finite deformation beam models and triality theory in dynamical post-buckling analysis. *Int. J. Non-Linear Mechanics* **35** (2000a) 103–131.

[24] Gao, D. Y. Analytic solutions and triality theory for nonconvex and nonsmooth variational problems with applications. *Nonlinear Analysis* **42** (2000b) 1161–1193.

[25] Gao, D. Y. Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. *J. Global Optimization* **17** (1/4) (2000c) 127–160.

[26] Gao, D. Y. *Duality Principles in Nonconvex Systems: Theory, Methods and Applications.* Kluwer Academic Publishers, Dordrecht / Boston / London, (2000d), xviii + 454 pp.

[27] Gao, D. Y. Complementary extremum principles and triality theory in non-smooth and nonconvex Hamilton mechanics. *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences* **359** (2001) 2347–2367.

[28] Gao, D. Y. Perfect duality theory and complete solutions to a class of global optimization problems. *Optimization* **52**(4-5) (2003) 467–493.

[29] Gao, D. Y. Canonical duality theory and solutions to constrained nonconvex quadratic programming. *Journal of Global Optimization* **29** (2004) 377–399.

[30] Gao, D. Y. Complete solutions to a class of polynomial minimization problems. *Journal of Global Optimization* **35** (2006) 131–143.

[31] Gao, D. Y. Complete solutions and optimality criteria for quadratic minimization with box constraints. Springer, Berlin, 2006. (to appear)

[32] Gao, D. Y. and Russell, D. L. An extended beam theory for smart materials applications, Part I. Extended beam theory, duality theory and finite element simulations. *Appl. Math. Optimization* **34** (1996) 279–298.

[33] Gao, D. Y. and Strang, G. Geometric nonlinearity: Potential energy, complementary energy, and the gap functional. *Quart. Appl. Math.* **XLVII**(3) (1989) 487–504.

[34] Gay, D. M., Overton, M. L. and Wright, M. H. A primal-dual interior method for nonconvex nonlinear programming. In: *Advances in nonlinear programming* (Beijing, 1996) 31–56, Appl. Optim., **14** Kluwer Acad. Publ., Dordrecht, 1998.

[35] Ghezzi, L. L. and Piccardi, C. PID control of a chaotic system: an application to an epidemiological model, *Automatica* **33** (1997) 181–191.

[36] Helton, J. W., Merino, O. and T. E. Walker. H^∞ optimization with plant uncertainty and semidefinite programming. *Int. J. Robust Nonlinear Control* **8**(9) (1998) 763–802.

[37] Horn, M. A. and Lasiecka, I. Global stabilization of a dynamic von Karman plate with nonlinear boundary feedback. *Appl. Math. Optimization* **31** (1995) 57–84.

[38] Koumboulis, F. N. and Mertzios, B. G. Decoupling of sigular systems via P-D feedback. In: *Proceedings of the International Conference on Control* **96** (1) (1996) 19–22.

[39] Koumboulis, F. N. and Mertzios, B. G. Feedback controlling against chaos. *Chaos, Solitons and Fractals* **11** (2000) 351–358.

[40] Lasiecka, I. and Triggiani, R. *Deterministic Control Theory for Infinite Dimensional Systems.* Vol. 1 and 2 *Encyclopedia of Mathematics.* Cambridge University Press, 1999.

[41] Lee, R. C. H. and Yung, S. P. Optimality conditions and duality for a non-linear time-delay control problem. *Optimal Control Appl. & Methods* **18** (1997) 327–340.

[42] Mossino, J. An application of duality to distributed optimal control problems with constraints on the control and state. *J. Math. Anal. Appl.* **50** (1975) 223–242.

[43] Motreanu, D. and Panagiotopoulos, P. D. *Minimax Theorems and Qualitative Properties of the Solutions of Hemivariational Inequalities.* Kluwer Academic Publishers, Dordrecht, 1999.

[44] Ogorzalek, M. J., Taming chaos. II: Control. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.* **40**(10) (1993) 700–706.

[45] Ott, E., Grebogi, C. and Yorke, J. A. Controlling chaos. *Phys. Rev. Lett.* **64** (1990) 1196–1199.

[46] Rockafellar, R. T. and Wets, R. J. B. *Variational analysis*. Springer, Berlin, New York, 1997.

[47] Strang, G. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, Wellesley, Massachusetts, 1986.

[48] Tanimoto, S. Duality in the optimal control of non-well-posed distributed systems. *J. Math. Analy. Appl.* **171** (1992) 277–287.

[49] Toland, J. F. Duality in nonconvex optimization. *J. Math. Anal. Appl.* **66** (1978) 399–415.

[50] Toland, J. F. A duality principle for non-convex optimization and the calculus of variations. *Arch. Rational Mech. Anal.* **71** (1979) 41–61.

[51] Wright, M. H. The interior-point revolution in constrained optimization. In: *High-Performance Algorithms and Software in Nonlinear Optimization* (R. DeLeone, A. Murli, P. M. Pardalos and G. Toraldo, eds.). Kluwer Academic Publishers, Dordrecht, 1998, 359–381.

[52] Zhong, W. X. Variational method and computation for H_∞ control. In: *APCOM-4* **2** (1999) 1037–1047.

[53] Zhong, W. X., Ouyang, H. J. and Deng, Z. C. *Computational Structural Mechanics and Optimal Control*. Press of Dalian University of Technology, Dalian, 1993. [Chinese].

Nonlinear Dynamics and Systems Theory, 6(3) (2006) 281–293

On a Class of Strongly Nonlinear Impulsive Differential Equation with Time Delay

W. Wei [1*], S.H. Hou [2] and K.L. Teo [3]

[1] *Department of Mathematics, University of Guizhou, Guiyang, Guizhou 550025, P.R. China*
[2] *Department of Applied Mathematics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[3] *Department of Mathematics and Statistics, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia*

Received: August 3, 2005; Revised: February 10, 2006

Abstract: In this paper, we prove the existence of solutions for a class of strongly nonlinear impulsive differential equations with time delay in infinite dimensional Banach spaces by means of a fixed point theorem due to Leray-Schauder.

Keywords: *Monotone operator; impulsive differential equation; delay; existence of solutions.*

Mathematics Subject Classification (2000): 47H05, 34A37, 34K30.

1 Introduction

In recent years there has been intensive research on systems governed by impulsive differential equations and impulsive functional equations in Banach spaces (see [1], [2], [3], [6], and the references therein). This is probably due to the fact that though a vast majority of physical systems are described by differential or difference equations, a more realistic model of a physical system can be constructed using differential equations with time delay and impulsive effects in describing the evolution and discrete events occurring in the system. In fact, many evolution processes in nature are characterized by the fact that there are inherently time delays and at certain moments of time experience an abrupt change of state. Most papers in the literature dealt with ordinary differential systems and semilinear differential equations. Their emphasis and advantage lie in the fact that solutions of these systems are being represented by means of integration formula via appropriate semigroup of operators. It seems that only a few papers discuss the strongly nonlinear impulsive functional differential system, which cover quasilinear partial differential equations with time delay.

* Corresponding author: e-mail:wei66@yahoo.com

Let $D = \{t_1 < t_2 < \cdots < t_m\}$ be fixed impulsive points in $(0, T)$. In this paper we study a class of strongly nonlinear impulsive differential equations with time delay in the form

$$\begin{cases} \dot{x}(t) + A(t, x) = f(t, x_t), \quad t \in [0, T] \setminus D, \\ x(t) = \phi(t), \quad t \in [-r, 0], \\ \triangle x(t_i) = G_i(x(t_i)), \quad i = 1, 2, \cdots, m. \end{cases} \tag{1.1}$$

Here A is a nonlinear monotone operator and f is a nonlinear nonmonotone perturbation, and G_i denotes the jump operator defined as

$$G_i(x(t_i)) = x(t_i+) - x(t_i-) = x(t_i+) - x(t_i),$$

$\phi \in PF$ and $x_t(\theta) = x(t + \theta)$, $\theta \in [-r, 0]$. The space PF will be introduced in Section 2.

We present here sufficient conditions for the existence of solution to this particular class of nonlinear impulsive functional equations in an appropriate infinite dimensional Banach space. The results are obtained by using the theory of nonlinear functional analysis and a fixed point theorem due to Leray-Schauder.

The rest of the paper is organized as follows. In Section 2, we introduce some basic notations. In Section 3, we prove the existence of solutions for a class of nonimpulsive delay differential equations in Banach spaces. In Section 4, we establish the new existence result for a class of nonlinear impulsive functional differential equation in Banach spaces. Finally, we conclude with an example to illustrate our results in Section 5.

2 Preliminaries

Let H be a separable Hilbert space and V be a dense subspace of H having the structure of a reflexive Banach space with continuous embedding, so that $V \hookrightarrow H \hookrightarrow V^*$ forms a Gelfand triple. We assume the injection $V \hookrightarrow H$ is continuous and compact. The system model considered here is based on this Gelfand triple (see [1] or Chapter 23 of [9]).

Let $I \equiv [0, T]$, $r > 0$, and $m > 0$ be given. The norm in any Banach space X will be denoted by $\| \cdot \|_X$. Let $PF(X) = \{\psi : [-r, 0] \to X;\ \ \psi$ is continuous everywhere except for a finite number of points $\widetilde{t}$ at which $\psi(\widetilde{t}-)$ and $\psi(\widetilde{t}+)$ exist and $\psi(\widetilde{t}-) = \psi(\widetilde{t})\}$ be endowed with the norm

$$\|\psi\|_{PF(X)} = sup\{\|\psi(\theta)\|_H, \theta \in [-r, 0]\}.$$

For any continuous function x defined on $[-r, T] \setminus D$ and $t \in [0, T]$, we denote by x_t the element of $PF \equiv PF(H)$ defined by

$$x_t(\theta) = x(t + \theta), \quad \theta \in [-r, 0].$$

Let $1 < q \le p < +\infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. The space $W_{pq}(I) \equiv W_{pq}$ is defined as follows:

$$W_{pq}(I) = \{x | x \in L_p(I, V), \dot{x} \in L_q(I, V^*)\}$$

with the norm

$$\|x\|^2_{W_{pq}} = \|x\|^2_{L_p(I,V)} + \|\dot{x}\|^2_{L_q(I,V^*)},$$

where $\dot{x}$ denotes the derivative of x in the generalized sense. $\{W_{pq}, \| \cdot \|_{W_{pq}}\}$ is a Banach space and the embedding $W_{pq} \hookrightarrow C(I, H)$ is continuous. If the embedding $V \hookrightarrow H$ is compact, the embedding $W_{pq} \hookrightarrow L_p(I, H)$ is also compact (see [9] and [1]). Similarly, we can define $W_{pq}([s,u])$ for $0 \le s < t < u \le T$. Furnished with the norm $\|\cdot\|_{W_{pq}([s,u])}$, the space $(W_{pq}([s,u]), \|\cdot\|_{W_{pq}([s,u])})$ becomes a Banach space which is clearly reflexive and separable. Moreover, the embedding $W_{pq}([s,u]) \hookrightarrow C([s,u],H)$ is continuous and the embedding $W_{pq}([s,u]) \hookrightarrow L_p((s,u),H)$ is also compact.

Set $PC(I,H) = \{x : x$ is a map from I into H such that $x(t)$ is continuous at $t \in I \backslash D$ and $x(t)$ is left continuous at $t \in D$ and the right limit $x(t_i+)$ exists for $i = 1, 2, \cdots m\}$, and

$$PW_{pq}(I) = \{x : x\,|_{[t_i,t_{i+1}]} \in W_{pq}([t_i,t_{i+1}]) \text{ for } i = 0, 1, \cdots, m\},$$

where $t_0 = 0$ and $t_{m+1} = T$. For $x \in PW_{pq}(I) \cap PC(I,H) \triangleq PWC$, define

$$\|x\|_{PWC} = \sum_{i=0}^{m} \|x\|_{W_{pq}[t_i,t_{i+1}]} + \sum_{i=1}^{m} \|x(t_i+) - x(t_i-)\|_H.$$

It is easy to show that PWC is a Banach space.

Let us consider the following nonlinear impulsive differential equation with time delay

$$\begin{cases} \dot{x}(t) + A(t,x) = f(t,x_t), & t \in [0,T] \setminus D, \\ x(t) = \phi(t), & t \in [-r,0], \\ \triangle_l x(t_i) = G_i(x(t_i)) & i = 1, 2, \cdots, m, \end{cases} \tag{2.1}$$

where A is a nonlinear monotone operator, f is a nonlinear nonmonotone perturbation, $G_i (i = 1, 2 \cdots m)$ are nonlinear maps. Here $\phi \in PF$, and $\triangle_l x(t_i) = x(t_i+) - x(t_i-) = x(t_i+) - x(t_i)$, which represents the jump in the state x at time t_i with G_i determining the size of the jump at time t_i.

We will impose the following hypotheses on problem (2.1).

(A1) $A : I \times V \to V^*$ is an operator such that

(i) $t \to A(t,x)$ is measurable.

(ii) $x \to A(t,x)$ is monotone and hemicontinuous; i.e., $\forall t \in I$,

$$\langle A(t,x_1) - A(t,x_2), x_1 - x_2 \rangle \ge 0 \qquad \forall x_1,\, x_2 \in V, \quad t \in I;$$
$$A(t, x + sy) \xrightarrow{W} A(t,x) \qquad \text{in } V^* \quad \text{as } s \to 0 \quad \forall x, y \in V.$$

(iii) There exist positive constants c_1, c_2, c_3 and a nonnegative function $c_4(\cdot) \in L_q(I)$ such that $\forall t \in I$,

$$\langle A(t,x), x \rangle \ge c_1 \| x \|_V^p - c_2, \qquad \text{for all } x \in V,$$
$$\|A(t,x)\|_{V^*} \le c_4(t) + c_3 \|x\|_V^{p-1} \text{ for all } x \in V.$$

(A2) $f : I \times PF \to H$ is an operator such that

(i) $t \to f(t,\xi)$ is measurable, and
$\xi \to f(t,\xi)$ is continuous.

(ii) There exist a constant $\alpha \ge 0$ and a nonnegative function $h(\cdot) \in L_2(I)$ such that

$$\|f(t,\xi)\|_H \le h(t) + \alpha \|\xi\|_{PF}^{\frac{2}{q}}, \forall t \in I, \xi \in PF.$$

(A3) For $i = 1, 2, \cdots, m$, $G_i : H \to H$ is a bounded map (i.e., G_i maps a bounded set to a bounded set).

To arrive at the main results of the paper, we need the following fixed point theorem due to Leray and Schauder [5].

Theorem 2.1 *Let B be a convex subset of a normed linear space E and $0 \in B$. Let $P : B \to B$ be a completely continuous operator and let*

$$\xi(P) = \{x \in B : x = \sigma P(x) \text{ for some } 0 < \sigma < 1\}.$$

Then either the set $\xi(P)$ is unbounded, or P has a fixed point.

3 Existence of solutions of functional differential equation

In this section we consider the following functional differential equation without impulsive effects:

$$\begin{cases} \dot{x}(t) + A(t,x) = f(t,x_t), & t \in [0,T], \\ x(t) = \phi(t), & t \in [-r,0], \end{cases} \tag{3.1}$$

Definition 3.1 *A function x is called a weak solution of (3.1) if $x|_{[0,T]} \in W_{pq}$ satisfies the equation in a weak sense and $x(t) = \phi(t)$ $\forall t \in [-r,0]$.*

Theorem 3.1 *Under assumptions (A1) and (A2), problem (3.1) has a solution in W_{pq}.*

Proof

Step 1: The proof will be given first for the case where $\phi(0) = 0$.

(1) Set

$$B = \{y | y \in C([0,T],H), y(0) = 0\}.$$

Obviously, B is a Banach space with the supremum norm. For any $x \in B$, we define $F : B \to L_2(I,H)$ by $F(x)(t) = f(t,\widehat{x}_t)$ with

$$\widehat{x}_t(s) = \begin{cases} \phi(t+s) & \text{for } t+s \in [-r,0), \\ x(t+s) & \text{for } t+s \in [0,T]. \end{cases}$$

The operator P is defined on B by letting $y = Px$ be the corresponding solution of the following Cauchy problem

$$\begin{cases} \dot{y}(t) + A(t,y(t)) = F(x)(t), & t \in I, \\ y(0) = 0. \end{cases}$$

Indeed, by assumption $(A2)$ and $1 < q \le p < +\infty$, $F(x)(t) = f(t,\widehat{x}_t)$ is measurable and

$$F(x)(\cdot) \in L_2(I,H) \subset L_q(I,V^*).$$

Thus, the above Cauchy problem has a unique solution $y \in W_{pq} \hookrightarrow C(I,H)$ (see Theorem 30.A of [9]). Hence P maps B into itself.

(2) $P : B \to B$ is continuous.

Suppose $x_n \longrightarrow x$ in B as $n \longrightarrow \infty$. This means

$$\sup_{0 \le t \le T} \|x_n(t) - x(t)\|_H \longrightarrow 0,$$

as $n \to +\infty$. Hence, there exists a constant $M > 0$ such that

$$\|\widehat{x}_n\|_{PC([-r,T],H)} \le M \quad \text{and} \quad \|\widehat{x}\|_{PC([-r,T],H)} \le M.$$

By virtue of assumption (A2), we have for $t \in I$,

$$F(x_n)(t) \longrightarrow F(x)(t) \text{ in } H$$

as $n \to \infty$ and there exists a constant $M_1 > 0$ such that

$$\|F(x_n)(t)\|_H \le h(t) + M_1 \quad \text{and} \quad \|F(x)(t)\|_H \le h(t) + M_1.$$

It follows from the majorized convergence principle that

$$F(x_n) \longrightarrow F(x) \text{ in } L_2(I,H)$$

as $n \to \infty$.

Let $y_n = Fx_n$ and $y = Fx$ satisfy the following equations respectively. For $t \in I$,

$$\begin{aligned} \dot{y}_n(t) + A(t, y_n(t)) &= F(x_n)(t), \quad y_n(0) = 0, \\ \dot{y}(t) + A(t, y(t)) &= F(x)(t), \quad y(0) = 0. \end{aligned}$$

Then

$$\begin{aligned} \frac{1}{2}\|y_n(t) - y(t)\|_H^2 &\le \|F(x_n) - F(x)\|_{L_2(0,t;H)}\|y_n - y\|_{L_2(0,t;H)} \\ &\le \frac{1}{2}\|F(x_n) - F(x)\|^2_{L_2(I,H)} + \frac{1}{2}\int_0^t \|y_n(\tau) - y(\tau)\|_H^2 d\tau. \end{aligned}$$

Thanks to Gronwall's lemma, it is easy to show that

$$y_n \longrightarrow y \quad \text{in } B \text{ as } n \to \infty.$$

(3) P is a compact operator on B.

Let $\{x_n\}$ be a bounded sequence in B. That is, there is a constant $M_2 > 0$ such that

$$\|x_n\|_{C(I,H)} \le M_2.$$

Again, by assumption $(A2)$, there exist constants $M_3, M_4 > 0$ such that

$$\|F(x_n)(t)\|_H \le h(t) + M_3 \quad \text{and} \quad \|F(x_n)\|_{L_2(I,H)} \le M_4.$$

Let $y_n = Px_n$ be a solution of the following equation

$$\begin{cases} \dot{y}_n(t) + A(t, y_n(t)) = F(x_n)(t). \\ y_n(0) = 0. \end{cases} \tag{3.2}$$

Integrating by parts in (3.2) and using assumption (A1), one can obtain

$$\frac{1}{2}\|y_n(t)\|_H^2 + C_1\|y_n\|^p_{L_p(0,t;V)} \le \|F(x_n)\|_{L_2(0,t;H)}\|y_n\|_{L_2(0,t;H)} + C_2.$$

It follows from the Cauchy inequality that there exist constants $\gamma > 0$ and $K > 0$ such that

$$\frac{1}{2}\|y_n(t)\|_H^2 + \gamma\|y_n\|^p_{L_p(0,t;V)} \le K\|F(x_n)\|^q_{L_2(I,H)} + C_2.$$

Hence $\{y_n\}$ is bounded in $C(I,H) \cap L_p(I,V)$. It follows from Eq.(3.2) that $\{\dot{y}_n\}$ is bounded in $L_q(I,V^*)$ and therefore $\{y_n\}$ is bounded in W_{pq}.

Since $W_{pq} \hookrightarrow L_p(I,H)$ is compact, there exists a subsequence, relabelled $\{y_n\}$, such that

$$y_n \longrightarrow y \quad \text{in } L_p(I,H) \quad \text{as } n \to \infty.$$

So $\{y_n\}$ is a Cauchy sequence in $L_p(I,H)$. Hence there exists a constant $M_5 > 0$ such that

$$\begin{aligned} \frac{1}{2}\|y_n(t) - y_m(t)\|_H^2 &\leq \|F(x_n) - F(x_m)\|_{L_q(I,H)}\|y_n - y_m\|_{L_p(I,H)} \\ &\leq M_5\|y_n - y_m\|_{L_p(I,H)}. \end{aligned}$$

This inequality implies that $\{y_n\}$ is a Cauchy sequence in B. Since B is closed, the sequence $\{y_n\}$ has a limit in B. This proves that P is compact.

(4) Boundedness of the set $\xi(P)$.

We will show that the set $\xi(P)$ is bounded. To this end, suppose $x \in B$ and $x = \sigma P x$ where $\sigma \in (0,1)$. This implies that x satisfies the following Cauchy problem:

$$\begin{cases} \frac{1}{\sigma}\dot{x}(t) + A(t, \frac{1}{\sigma}x(t)) = g(t, \widehat{x}_t), \quad t \in I, \\ x(0) = 0. \end{cases} \tag{3.3}$$

We will show that there exists a $Q > 0$ such that

$$\|x\|_{C(I,H)} \leq Q.$$

Using the same arguments and assumptions $(A1)$ and $(A2)$, we have

$$\begin{aligned} &\frac{1}{2\sigma}\|x(t)\|_H^2 + \frac{C_1}{\sigma^{p-1}}\|x\|_{L_p(0,t;V)}^p \leq \int_0^t \langle f(\tau, \widehat{x}_\tau), x(\tau)\rangle d\tau + C_2 \\ \leq\ & (\int_0^t \|f(\tau, \widehat{x}_\tau)\|_H^q d\tau)^{1/q}(\int_0^t \|x(\tau)\|_H^p d\tau)^{1/p} + C_2 \\ \leq\ & \frac{1}{q\varepsilon^q}\int_0^t \|f(\tau, \widehat{x}_\tau)\|_H^q d\tau + \frac{\varepsilon^p}{p}\|x\|_{L_p(0,t;H)}^p + C_2, \end{aligned}$$

for any constant $\varepsilon > 0$ and some constants $C_1 \geq 0$ and $C_2 \geq 0$. Hence

$$\frac{\sigma^{p-2}}{2}\|x(t)\|_H^2 + C_1\|x\|_{L_p(0,t;V)}^p \leq C_2\sigma^{p-1} + b_1\varepsilon^p\sigma^{p-1}\|x\|_{L_p(0,t;V)}^p + \frac{d_1\sigma^{p-1}}{q\varepsilon^q}\int_0^t \|\widehat{x}_\tau\|_{PF}^2 d\tau$$

where b_1 and d_1 are positive constants. So, we can choose $\varepsilon > 0$ small enough such that

$$\frac{\sigma^{p-2}}{2}\|x(t)\|_H^2 \leq a_1\sigma^{p-1} + b_2\sigma^{p-1}\int_0^t \|\widehat{x}_\tau\|_{PF}^2 d\tau$$

where a_1 and b_2 are positive constants. It follows from $0 < \sigma < 1$ that

$$\|x(t)\|_H^2 \leq a_2 + d_1\int_0^t \|\widehat{x}_\tau\|_{PF}^2 d\tau.$$

We denote

$$k(t) = a_2 + d_1\int_0^t \|\widehat{x}_\tau\|_{PF}^2 d\tau.$$

It is obvious that $k(t)$ is an increasing function. So,

$$\sup_{0\le\theta\le t}\|x(\theta)\|_H^2 \le a_2 + d_1 r\|\phi\|_{PF}^2 + \int_0^t \sup_{0\le\theta\le\tau}\|x(\theta)\|_H^2 d\tau \quad \text{for all } t\in[0,T]$$

Let $\omega(t) = \sup_{0\le\theta\le t}\|x(\theta)\|_H^2$. Then $\omega(t)$ is continuous and increasing since $x(t)$ is continuous. An application of the Gronwall lemma implies that

$$\|x\|_{C([0,T],H)} \le Q, \tag{3.4}$$

and so $\xi(P)$ is bounded.

By the Leray-Schauder fixed point theorem (Theorem 2.1), P has a fixed point x^* in B. Then x^* is a corresponding solution of (3.1).

Step 2: For the proof of the theorem, in general case where $\phi(0)\ne 0$, at first we assume that $\phi(0)\in V$, we use the transformation

$$y = x - \phi(0)$$

to reduce the problem (3.1) into the following problem:

$$\begin{cases} \dot y(t) + A(t, y+\phi(0)) = f(t, y_t + \phi(0)), & t\in I,\\ y(t) = \phi(t) - \phi(0), & t\in[-r,0], \end{cases} \tag{3.5}$$

We set $\hat A(t,y) = A(t, y+\phi(0))$ and $\hat f(t,y_t) = f(t, y_t+\phi(0))$. Then it is easy to see that $\hat A$ satisfies assumptions (A1)(i) and (ii). It follows from assumption (A1) (iii) that

$$\begin{aligned} \|\hat A(t,y)\|_{V^*} &= \|A(t,y+\phi(0))\|_{V^*} \le c_4(t) + c_3\|y+\phi(0)\|_V^{p-1}\\ &\le c_4(t) + c_3 2^{p-1}\|y\|_V^{p-1} + c_3 2^{p-1}\|\phi(0)\|_V^{p-1}. \end{aligned} \tag{3.6}$$

Let $m_4 = c_4(t) + c_3 2^{p-1}\|\phi(0)\|_V^{p-1}$ and $m_3 = c_3 2^{p-1}$. Then

$$\|\hat A(t,y)\|_{V^*} \le m_4(t) + m_3\|y\|_V^{p-1}$$

for all $y\in V$ and $t\in I$.

By assumption (A1)(iii), one can get

$$\begin{aligned} \langle \hat A(t,y), y\rangle &= \langle A(t,y+\phi(0)), y+\phi(0)\rangle - \langle A(t,y+\phi(0)),\phi(0)\rangle\\ &\ge c_1\|y+\phi(0)\|_V^p - c_2 - \|A(t,y+\phi(0))\|_{V^*}\cdot\|\phi(0)\|_V\\ &\ge c_1\|y+\phi(0)\|_V^p - c_2 - \frac{1}{p\epsilon^p}\|\phi(0)\|_V^p - \frac{\epsilon^q}{q}\|A(t,y+\phi(0))\|_{V^*}^q \end{aligned} \tag{3.7}$$

for any constant $\epsilon > 0$. Then, by (3.6), one can reduce (3.7) into

$$\langle \hat A(t,y), y\rangle \ge \left(c_1 - \frac{c_3 2^{q-1}}{q}\epsilon^q\right)\|y+\phi(0)\|_V^p - c_2 - \frac{1}{p\epsilon^p}\|\phi(0)\|_V^p - \frac{2^{q-1}c_4^q(t)}{q}\epsilon^q.$$

We can choose ϵ small enough such that $m_1 \equiv c_1 - \frac{c_3 2^{q-1}}{q}\epsilon^q > 0$ and note that the following inequality

$$\begin{aligned} \|y+\phi(0)\|_V^p + \|\phi(0)\|_V^p &\ge |\|y\|_V - \|\phi(0)\|_V|^p + \|\phi(0)\|_V\\ &\ge C\left(\|y\|_V - \|\phi(0)\|_V + \|\phi(0)\|_V\right)^p = C\|y\|_V^p \end{aligned}$$

holds for some constant $C > 0$.

One can obtain

$$\langle \hat{A}(t,y), y\rangle \geq m_1 C\|y\|_V^p - m_2$$

where $m_2 = c_2 - \frac{1}{p\epsilon^p}\|\phi(0)\|_V p - c_1\|\phi(0)\|_V^p - \frac{2^{q-1}c_4^q(t)}{q}\epsilon^q$. That is, $\hat{A}(t,y)$ satisfies assumption (A1).

For $\hat{f}(t,y) = f(t, y+\phi(0))$, one can easily verify that $\hat{f}(t,y)$ satisfying assumption (A2). Then the problem (3.5) has a solution from Step 1.

If $\phi(0) \in H$, there exists a sequence $\{\xi_n\} \subset V$, such that $\xi_n \to \phi(0)$ in H. Set

$$\phi_n(t) = \begin{cases} \phi(t) & \text{for } t \in [-r, 0), \\ \xi_n & \text{for } t = 0. \end{cases}$$

Then there exists $x_n \in W_{pq}$ such that

$$\begin{cases} \dot{x}_n(t) + A(t, x_n(t)) = f(t, (x_n)_t), & t \in I, \\ x_n(t) = \phi_n(t), & t \in [-r, 0]. \end{cases} \tag{3.8}$$

We define $\hat{A}(x)(t) = A(t, x(t))$ for $x \in L_p(I,V)$ and $t \in I$. Then $\hat{A} : L_p(I,V) \to L_q(I,V^*)$ is bounded, monotone, hemicontinuous, and coercive (see Theorem 30.A of [9]). It follows from (3.4) and assumption (A1) that

$$\|x_n\|_{W_{pq}} \leq M \qquad \text{and} \qquad \|A(x_n)\|_{L_q(I,V^*)} \leq M$$

for some constant $M > 0$. Then there exists a subsequence of $\{x_n\}$, denoted $\{x_n\}$ again, such that

$$\begin{aligned} x_n &\xrightarrow{W} x && \text{in} \quad L_p(I,V), \\ \dot{x}_n &\xrightarrow{W} \dot{x} && \text{in } L_q(I,V^*), \\ \hat{A}(x_n) &\xrightarrow{W} w && \text{in } L_q(I,V^*), \end{aligned}$$

as $n \to +\infty$. Since $W_{pq} \hookrightarrow L_p(I,H)$ is compact, we know that

$$\begin{aligned} x_n &\xrightarrow{W} x && \text{in } W_{pq}, \\ x_n &\xrightarrow{S} x && \text{in } L_p(I,H), \\ x_n(t) &\xrightarrow{S} x(t) && \text{a.e. on } I \text{ in } H. \end{aligned}$$

By assumption (A2) and using the similar method as in the proof of Lemma 1 of [8], it follows that

$$F(x_n) \xrightarrow{S} F(x) \qquad \text{in } L_q(I,H).$$

Hence

$$\begin{cases} \dot{x} + w = F(x), & t \in I, \\ x(t) = \phi(t), & t \in [-r, 0]. \end{cases} \tag{3.9}$$

Combining (3.8) and (3.9), we obtain

$$\langle \dot{x}_n - \dot{x}, x_n - x\rangle + \langle \hat{A}x_n - w, x_n - x\rangle = \langle F(x_n) - F(x), x_n - x\rangle.$$

Hence

$$\begin{aligned}\langle \hat{A}(x_n) - w, x_n\rangle &= \frac{1}{2}\|x_n(0) - x(0)\|_H^2 - \frac{1}{2}\|x_n(T) - x(T))\|_H^2 \\ &+ \langle \hat{A}(x_n) - w, x\rangle + \langle F(x_n) - F(x), x_n - x\rangle.\end{aligned}$$

So,

$$\limsup_{n\to+\infty}\langle Ax_n, x_n\rangle \le \langle w, x\rangle.$$

Note that $\hat{A} : L_p(I, V) \to L_q(I, V^*)$ is monotone, hemicontinuous, and so $\hat{A}$ satisfies the condition (M) (see p. 538 of [9]). We deduce that

$$w = \hat{A}(x).$$

Thus,

$$\begin{cases} \dot{x} + \hat{A}(x) = F(x), & t \in I, \\ x(t) = \phi(t), & t \in [-r, 0], \end{cases} \tag{3.10}$$

That is, x is a solution of (4.1).

The theorem is proved.

Remark 3.1 *It follows from the proof of Theorem 3.2 that if x is a solution of (3.1) then x is bounded in W_{pq}.*

Theorem 3.2 guarantees the existence of solutions for (3.1), but not the uniqueness of solutions. In order to obtain uniqueness, we have to impose a somewhat stronger assumption on f. Assume that

(A4) f is locally Lipschitz continuous with respect to ξ, i.e., for any $\rho > 0$, there exists a constant $L(\rho)$, such that

$$\|f(t, \xi_1) - f(t, \xi_2)\|_H \le L(\rho)(\|\xi_1 - \xi_2\|_{PF}), \quad \forall t \in I$$

and for all $\xi_1, \xi_2 \in PF(H)$ satisfying $\|\xi_1\|_{PF} \le \rho, \|\xi_2\|_{PF} \le \rho$.

Theorem 3.2 *(Uniqueness of solution) If assumption (A4) holds, then the problem (3.1) has at most one solution.*

Proof Let x_1 and x_2 be two solutions of problem (3.1). Then

$$\begin{aligned}\frac{1}{2}\|x_1(t) - x_2(t)\|_H^2 &\le \|F(x_1) - F(x_2)\|_{L_2(0,t;H)}\|x_1 - x_2\|_{L_2(0,t;H)} \\ &\le \frac{1}{2}\|F(x_1) - F(x_2)\|_{L_2(0,t;H)}^2 + \frac{1}{2}\|x_1 - x_2\|_{L_2(0,t;H)}^2.\end{aligned}$$

By assumption (A4), there exist constants $C_1^* > 0$ and $C_2^* > 0$ such that

$$\|x_1(t) - x_2(t)\|_H^2 \le C_1^* \int_0^t \|x_1(\tau) - x_2(\tau)\|_H^2 d\tau + C_2^* \int_0^t \|(x_1)_\tau - (x_2)_\tau\|_{PF}^2 d\tau.$$

Because $x_1(t) = x_2(t) = \phi(t), t \in [-r, 0]$ and the solution of (3.1) is continuous in $[0, T]$, one can modify x_1 and x_2 by setting $x_1(t) = x_2(t) \equiv \xi, \forall t \in [-r, 0]$. Then $x_1, x_2 \in C([-r, T]; H)$ such that

$$\|x_1(t) - x_2(t)\|_H^2 \le C_1^* \int_0^t \|x_1(\tau) - x_2(\tau)\|_H^2 d\tau + C_2^* \int_0^t \|(x_1)_\tau - (x_2)_\tau\|_C^2 d\tau \quad \forall t \in [0, T]$$

where $C = C([-r, 0], H)$ denotes all continuous maps from $[-r, 0]$ into H with the usual supremum norm. Thanks to Gronwall's lemma, it implies

$$x_1(t) = x_2(t) \text{ for all } t \in [0, T].$$

That is,

$$x_1 = x_2.$$

4 Existence of solutions for impulsive delay differential equations

In this section, we deal with the nonlinear impulsive differential equation (2.1) with time delay in Banach Space.

Definition 4.1 A function $x \in PWC$ is called a PWC solution of (2.1) if it satisfies the equation in a weak sense on every interval $[t_i, t_{i+1}]$ $(i = 0, 1, \cdots, m)$, $x(t) = \phi(t), \quad t \in [-r, 0]$, and the state jump at t_i $(i = 1, 2, \cdots, m)$.

Theorem 4.1 *Suppose assumptions (A1), (A2), and (A3) hold. Then, for each $\phi \in PF(H)$, the problem (2.1) has a solution $x \in PWC$. Moreover, there is a constant $M > 0$ such that*

$$\|x\|_{PWC} \leq M \quad \text{and} \quad \|x\|_{PC} \leq M.$$

Proof Define $I_i \equiv (t_i, t_{i+1}], i = 0, 1, \cdots, m$ with $t_0 = 0, t_{m+1} = T$. It follows from assumptions (A1), (A2), Theorem 3.2, and Theorem 3.4 that for each $\phi \in PF$ and $\phi(0) \in V$, the equation (2.1) has a unique solution $x^{(1)}$ where $x^{(1)} |_{I_0} \in W_{pq}(I_0) \bigcap C(I_0, H)$ and $x^{(1)} |_{[-r,0]} = \phi$. By assumption (A3), $x(t_1 + 0)$ is well defined and it is given by

$$x(t_1 + 0) = G_1(x^{(1)}(t_1)) + x^{(1)}(t_1) \equiv \xi^1.$$

Consider the following problem

$$\begin{cases} \dot{x}(t) + A(t, x) = f(t, x_t), \quad t \in I_1, \\ x(t) = x^{(1)}(t), \quad t \in [t_1 - r, t_1], \\ x(t_1) = \xi^1. \end{cases} \tag{4.1}$$

Using the same argument as in the proof of Theorem 3.2 and the fact $x^{(1)} \in PF([t_1 - r, t_1], H)$, one obtains that there is a unique solution $x^{(2)}$ in I_1.

We continue this process taking into account that $x^{(m+1)} := x |_{I_m}$ is a solution to the problem

$$\begin{cases} \dot{x}(t) + A(t, x) = f(t, x_t), \quad t \in (t_m, T], \\ x(t) = \phi(t), \quad t \in [t_m - r, t_m], \\ x(t_m + 0) = x^{(m)}(t_m) + G_m(x^{(m)}(t_m)). \end{cases} \tag{4.2}$$

The solution x of the problem (2.1) is defined by

$$x(t) = \begin{cases} x^{(1)}(t), & \text{if } t \in [-r, t_1], \\ x^{(2)}(t), & \text{if } t \in (t_1, t_2], \\ \vdots & \\ x^{(m+1)}(t), & \text{if } t \in (t_m, T]. \end{cases}$$

And it follows from Remark 3.3 that for $k = 0, 1, \cdots, m$

$$\|x^{(k+1)}\|_{W_{pq}(I_k)} \le M$$

for some constant $M > 0$. Hence x is a PWC solution of (2.1) and

$$\|x\|_{PWC} \le M_1 \quad \text{and} \quad \|x\|_{PC} \le M_1.$$

for some constant $M_1 > 0$.

5 Examples

In this section we present an example of delay evolution equations with impulse to which our general theory applies.

Let Ω be a bounded domain in R^n with smooth boundary $\partial\Omega$, $0 < t_1 < \cdots < t_k < T$ are given fixed points and $D \equiv \{t_1, t_2, \cdots, t_k\}$, $Q_T = (0,T) \setminus D \times \Omega$, $0 < T < \infty$. Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ be a multi-index with nonnegative integers $\{\alpha_i\}$, $i = 1, \ldots, n$, and $\|\alpha\| = \sum_{i=1}^n \alpha_i$. Let $p \ge 2$ and $q = p/(p-1)$and let $m > 0$ be an integer. $W^{m,p}(\Omega)$ denotes the standard Sobolev space with the usual norm:

$$\|\varphi\|_{W^{m,p}} = (\sum_{|\alpha| \le m} \|D^\alpha \varphi\|^p_{L^p(\Omega)})^{1/p}.$$

Let $W_0^{m,p}(\Omega) = \{\varphi \in W^{m,p} | D^\beta \varphi|_{\partial\Omega} = 0, \quad |\beta| \le m-1\}$. It is well known that $C_0^\infty(\Omega) \hookrightarrow W_0^{m,p}(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow W^{-m,p}(\Omega)$ and the embedding $W_0^{m,p}(\Omega) \hookrightarrow L^2(\Omega)$ is compact. Denote $V \equiv W_0^{m,p}(\Omega)$, $H \equiv L_2(\Omega)$, then $V^* \equiv W^{-m,q}(\Omega)$.

We consider the following initial-boundary impulsive value problem of $2m$-order quasi-linear delay parabolic equation:

$$\begin{cases} \frac{\partial}{\partial t} y(t,x) + \sum\limits_{|\alpha| \le m} (-1)^{|\alpha|} D^\alpha A_\alpha(t, x, \eta(y)(t,x)) = g(t, x, y(t-r, x)) \quad \text{on } Q_T, \\ D^\beta y(t,x) = 0 \quad \text{on } [0,T] \times \partial\Omega, \qquad \text{for all } \beta \text{ satisfying } |\beta| \le m-1, \\ y(s,x) = \phi(s,x) \quad \text{for } x \in \Omega \text{ and } -r \le s \le 0, \\ y(t_i+) = -y(t_i-), \quad i = 1, 2, \cdots, k, \end{cases} \tag{5.1}$$

where $\eta(y) \equiv \{(D^\gamma y), \quad |\gamma| \le m\}, \phi(t,x)$ is a given function, $\phi \in C([-r,0], L_2(\Omega))$, $\phi(0) \in W_0^{m,p}(\Omega)$, and $M = \frac{(n+m)!}{n!m!}$.

For $y_1, y_2 \in W_0^{m,p}(\Omega)$ and $t \in I$, we set

$$a(t, y_1, y_2) = \int_\Omega \sum_{|\alpha| \le m} A_\alpha(t, x, \eta(y_1)(t,x)) D^\alpha y_2 dx$$

and assume that for all α with $|\alpha| \le m$, the function $A_\alpha : Q_T \times R^M \to R$ satisfies the following properties.

(H1) (1)$(t,x) \to A_\alpha(t,x,\eta)$ is measurable on Q_T for $\eta \in R^M$, $\eta \to A_\alpha(t,x,\eta)$ is continuous on R^M for a.e. $(t,x) \in Q_T$;

(2) For $\eta = (\eta_\alpha) \in R^M, \widetilde{\eta} = (\widetilde{\eta}_\alpha) \in R^M$, there exist positive constants c, c_1, c_2, c_3, and c_4 such that

$$\sum_{|\alpha|\le m} (A_\alpha(t,x,\eta) - A_\alpha(t,x,\widetilde{\eta}))(\eta_\alpha - \widetilde{\eta}_\alpha) \ge 0,$$
$$\sum_{|\alpha|\le m} A_\alpha(t,x,\eta)\eta_\alpha \ge c_1 \sum_{|\gamma|\le m} |\eta_\gamma|^p - c_2,$$
$$|A_\alpha(t,x,\eta)| \le c_4 + c_3 \sum_{|\gamma|\le m} |\eta_\gamma|^{p-1}.$$

It is not difficult to verify that under the above assumption, for each $y_1 \in V$ and $t \in [0,T]$, $y_2 \to a(t, y_1, y_2)$ is a continuous linear form on V. Hence there exists an operator $A : I \times V \to V^*$ such that

$$\langle A(t, y_1),\ y_2\rangle_{V^*,V} = a(t, y_1, y_2).$$

Under the given assumption (H1), it is easy to see that A satisfies our assumption (A1) of Section 3.

Assume the function $f : Q_T \times R \to R$ satisfies the following properties.

(H2) (1) $(t,x) \to f(t,x,\eta)$ is measurable on Q_T for all $\eta \in R$;

(2) $\eta \to f(t,x,\eta)$ is continuous on R for almost all $(t,x) \in Q_T$;

(3) there exist constants $b_1 > 0$ and $b_2 > 0$ such that

$$|f(t,x,\eta)| \le b_1|\eta|^{2/q} + b_2(t,x)$$

for almost all $(t,x) \in Q_T$.

For $\phi_1 \in H$ and $t \in I$, set

$$b(t, \phi_1, \psi) = \int_\Omega f(t, x, \phi_1)\psi dx.$$

Then $\psi \to b(t, \phi_1)$ is a continuous linear form on H. Hence there exists an operator $F : [0,T] \times H \to H$ such that

$$b(t, \phi_1, \psi) = (F(t, \phi_1), \psi).$$

Noting that $y_t(\theta) = y_t(r)$ for all $-r \le \theta \le 0$ and using (H2), one can verify that F satisfies assumption (A2) of Section 3.

With the operators A and F as defined above, problem (5.1) can be written as the abstract evolution equation

$$\begin{cases} \dot{y}(t) + A(t, y(t)) = F(t, y(t-r)), \quad t \in [0,T] \setminus D, \\ y(t) = \phi(t), \quad t \in (-r, 0), \\ y(t_i) = -y(t_i), \quad 0 < t_1 < t_2 < \cdots < t_k < T. \end{cases}$$

Hence our result can be applied to this model to assert the existence of its solutions.

Acknowledgements

This work was done when the first author was visiting the Hong Kong Polytechnic University. This author would like to thank the staff in the Department of Applied Mathematics for their support and hospitality during her stay. The authors also thank anonymous referees for helpful comments.

The financial support of a research grant from the Australian Research Council is acknowledged.

References

[1] Ahmed, N.U., Teo, K.L., and Hou, S.H. Nonlinear impulsive systems on infinite dimensional spaces. *Nonlinear Analysis* **54** (5) (2003) 907–925.

[2] Benchohra, M., Henderson, J., Ntouyas, S.K., Ouahab A., Impulsive functional differential equations with variable times. *Comput. Math. Appl.* **47** (11) (2004) 1659–1665.

[3] Chang, Yong-Kui and Li, Wang-Tong, Existence results for second order impulsive functional differential inclusions. *J. Math. Anal. Appl.* **301** (2) (2005) 477–490.

[4] Hale, J.K. *Theory of Functional Differential Equations.* Springer-Verlag, New York, 1977.

[5] Granas A. and Dugundji J. *Fixed Point Theory.* Springer Verlag, New York, 2003.

[6] Sattayatham, P. Strongly nonlinear impulsive evolution equations and optimal control. *Nonlinear Analysis* **57** (7-8) (2004) 1005–1020.

[7] Xiang, X. and Kuang H. Delay Systems and Optimal Control. *ACTA Mathematicae Applicatae Sinica* **16** (4) (2000) 27–35.

[8] Wei, W.and Sattayatham, P. On Existence of optimal control governed by a class of periodic nonlinear evolution systems on Banach spaces. *ACTA Analysis Functionalis Applicata* (4) (2002) 124–136.

[9] Zeidler, E. *Nonlinear Functional Analysis and its Applications II.* Springer Verlag, New York, 1990.

Nonlinear Dynamics and Systems Theory, **6**(3) (2006) 295–308

The Matrix-Geometric Solution of the M/E$_k$/1 Queue with Balking and State-Dependent Service

Dequan Yue[1], Chunyan Li[1] and Wuyi Yue[2]

[1] *Department of Statistics, College of Sciences, Yanshan University Qinhuangdao 066004, China*
[2] *Department of Information Science and Systems Engineering Konan University, Kobe 658-8501, Japan*

Received: August 26, 2005; Revised: January 10, 2006

Abstract: In this paper, we present an analysis for an M/E$_k$/1 queuing system with balking and state-dependent service. Customers are served with two different rates depending on the number of customers in the system. If a customer on arrival finds other customers in the system, it either decides to enter the queue or balks with a constant probability. We first formulate the queuing model as a quasi-birth and death (QBD) process. Then, we obtain the equilibrium condition of the system. By using the matrix geometric solution method, we obtain the explicit expressions for steady-state probability vector via the rate matrix $\boldsymbol{R}$. The computation of the rate matrix $\boldsymbol{R}$ is also discussed. Then, we derive explicitly some performance measures of the system. Based on these performance analysis, we develop a cost model to determine numerically the optimal cost and optimal critical value. Finally, we perform sensitivity analysis through numerical experiments.

Keywords: *Balking; state-dependent; matrix geometric solution; steady-state probability.*

Mathematics Subject Classification (2000): 60K25, 68M20.

1 Introduction

We consider an M/E$_k$/1 queuing system with balking and state-dependent service. Customers are served with two different service rates depending on the number of customers in the system. If a customer on arrival finds other customers in the system, it either decides to enter the queue or balk (does not enter) with a constant probability. Balking is not only a common phenomenon in queues arising in daily activities, but also in communication systems, production line systems and in various machine interferences or repair models (see [1]-[4] and references therein).

The queuing systems with balking, or reneging, or both have been studied by many researchers. Haight [5] is the first person who considered an M/M/1 queue with balking.

An M/M/1 queue with customer reneging was also proposed by Haight [6]. The combined effects of balking and reneging in an M/M/1 queue with limited waiting room and with unlimited waiting room have been investigated by Ancker and Gafarian [7], [8]. They obtained the steady-state probabilities and some performance measures of the system such as the mean number in the queue, the mean number in the system and the mean rate of customer loss.

Abou-EI-Ata [9] extended the model in [7] to study the state-dependent M/M/1/N queue with reneging and a general balk function, where the server has two service rates depending on the number of customers in the system. Some of its variations have been studied by several authors including, for example, Abou-EI-Ata and Kotb [10], Abou-EI-Ata et al. [11] and Abou-EI-Ata and Shawky [12].

Recently, Drekic and Woolford [13] studied a preemptive priority Markovian queue with state-dependent service and lower priority balking customers. They formulated the queueing model as a quasi-birth and death (QBD) process. By using the method of generalized eigenvalues, they established an explicit representation for the so-called rate matrix. They also obtained the steady-state joint distribution of the number of high and low priority customers in the system.

The state-dependent M/M/1 queue with balking was studied by Al-seedy and Kotb [14]. They obtained the transient solution of the state probabilities. Al-seedy [15] extended the model proposed by Abou-EI-Ata [9] to the state-dependent M/E_k/1/N queue with balking. By solving the steady-state probability-difference equations, Al-seedy [15] obtained some iterative expressions of the steady-state probabilities. However, these iterative expressions are too complex to obtain explicit expressions of the steady-state probabilities in general cases, and they could not derive explicitly some performance measures such as the distribution of the queue length and the expected number of customers in the system and in the queue. Even for a special case when the waiting room is unlimited (i.e., $N \to \infty$), it is difficult to obtain the explicit expressions of the steady-state probabilities from the iterative expressions.

In this paper, we study a state-dependent M/E_k/1 queue with balking and an unlimited waiting room. The rest of the paper is organized as follows. In Section 2, we formulate the queuing model as a QBD process and obtain the equilibrium condition of the system. In Section 3, by using a matrix-geometric solution method, we derive the explicit expression for steady-state probability vector. Also, we derive explicitly some performance measures of the system such as the expected number of the customers in the system and in the queue and the mean balking rate of the system. Based on these analyses, we develop a cost model to determine numerically the optimal cost and optimal critical value. In Section 4, we perform sensitivity analysis through numerical experiments. Conclusions are given in Section 5.

2 System Model and Equilibrium Condition

In this section, we first describe the system model. Then, we derive an infinitesimal generator of a QBD process of the system. Finally, we provide an equilibrium condition of the system.

2.1 Model assumptions

In this paper, we consider an M/E_k/1 queuing system with balking and state-dependent service rate. The assumptions of the system model are as follows:

(a) There is only one server in the system, and the server can only serve one customer at the same time. The capacity of the system is infinite. It is assumed that the service is independent of the arrival of the customers.
(b) Customers arrive at the system one by one according to a Poisson process with rate λ ($\lambda > 0$).
(c) A customer on arrival decides to join the queue or balk. If a customer on arrival finds some customers in the system, then it joins the queue with probability β and balks with probability $1-\beta$. If a customer on arrival finds no customer in the system, then he joins the system and will be serviced immediately.
(d) The customers are served on a first-come, first served (FCFS) discipline. Once service commences it always proceeds to completion. The service times are assumed to be distributed according to an Erlang distribution with mean k/μ_n and stage parameter k. The Erlang type k distribution is made up of k independent and identical exponential stages, each with mean $1/\mu_n$, given by

$$\mu_n = \begin{cases} k\mu_1, & n = 1, 2, ..., r, \\ k\mu_2, & n = r+1, r+2, \dots . \end{cases}$$

This means that the server has two rates say called "slow and fast" depending on the number of customers n in the system. When the number of customers n in the system is less than or equal to the critical value r, the server has slow service rate μ_1; otherwise, the server has fast service rate μ_2 ($0 < \mu_1 < \mu_2$).

2.2 Infinitesimal generator of a QBD process

Let $N(t)$ denote the number of the customers in the system at time t, and $J(t)$ denote the service stage that the customer being served at time t ($t \geq 0$). A customer goes into the first stage of the service (say stage k), then progresses through the remaining stages and must complete the last stage (say stage 1). The state space of the two dimensional process $\{(N(t), J(t)); t \geq 0\}$ is given by

$$S = \{(i, j);\ i = 0, 1, ..., j = 1, 2, ..., k\}.$$

All states of this two dimensional process are labelled in the lexicographic order as follows:

$$(0, 0); (1, 1), (1, 2), ..., (1, k); (2, 1), (2, 2), ..., (2, k); \dots .$$

By the probability analysis, we have the following infinitesimal generator of the process $\{(N(t), J(t)); t \geq 0\}$.

$$\boldsymbol{Q} = \begin{pmatrix} B_0 & \boldsymbol{C}_0 & & & & & & \\ \boldsymbol{A}_1 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & & & & \\ & \boldsymbol{A}_2 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & & & \\ & & \cdots & \cdots & \cdots & & & \\ & & & \boldsymbol{A}_2 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & \\ & & & & \boldsymbol{A}_3 & \boldsymbol{B}_2 & \boldsymbol{C}_1 & \\ & & & & & \cdots & \cdots & \cdots \end{pmatrix} \begin{matrix} \cdots & 0 \\ \cdots & 1 \\ \cdots & 2 \\ & \vdots \\ \cdots & r \\ \cdots & r+1 \\ & \vdots \end{matrix}$$

where

$$\boldsymbol{A}_1 = \begin{pmatrix} k\mu_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{A}_2 = \begin{pmatrix} 0 & \cdots & 0 & k\mu_1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \quad \boldsymbol{A}_3 = \begin{pmatrix} 0 & \cdots & 0 & k\mu_2 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$B_0 = -\lambda, \quad \boldsymbol{B}_1 = \begin{pmatrix} -\beta\lambda - k\mu_1 & & & \\ k\mu_1 & -\beta\lambda - k\mu_1 & & \\ & \cdots & \cdots & \\ & & k\mu_1 & -\beta\lambda - k\mu_1 \end{pmatrix},$$

$$\boldsymbol{B}_2 = \begin{pmatrix} -\beta\lambda - k\mu_2 & & & \\ k\mu_2 & -\beta\lambda - k\mu_2 & & \\ & \cdots & \cdots & \\ & & k\mu_2 & -\beta\lambda - k\mu_2 \end{pmatrix},$$

$$\boldsymbol{C}_0 = \begin{pmatrix} 0 & \cdots & 0 & \lambda \end{pmatrix}, \quad \boldsymbol{C}_1 = \begin{pmatrix} \beta\lambda & & & \\ & \beta\lambda & & \\ & & \cdots & \\ & & & \beta\lambda \end{pmatrix},$$

where $\boldsymbol{C}_0$ is a matrix of order $1 \times k$, $\boldsymbol{A}_1$ is a matrix of order $k \times 1$, and other matrixes are square matrixes of order k.

From the book written by Neuts [16], we know that process $\{N(t), J(t); t \geq 0\}$ is a QBD process.

2.3 Equilibrium condition of the system

In the following, we provide a necessary and sufficient condition to ensure the existence for the stationary probability distribution of the process $\{N(t), J(t); t \geq 0\}$.

Let $\boldsymbol{H} = \boldsymbol{A}_3 + \boldsymbol{B}_2 + \boldsymbol{C}_1$, then $\boldsymbol{H}$ is given by

$$\boldsymbol{H} = \begin{pmatrix} -k\mu_2 & 0 & \cdots & 0 & k\mu_2 \\ k\mu_2 & -k\mu_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & k\mu_2 & -k\mu_2 \end{pmatrix}.$$

It is readily known that $\boldsymbol{H}$ is an irreducible generator. Let $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_k)$ be the steady-state probability vector of $\boldsymbol{H}$. Then, $\boldsymbol{\pi}$ satisfies the linear equations $\boldsymbol{\pi H} = 0$ and $\boldsymbol{\pi e} = 1$, where $\boldsymbol{e}$ is a column vector whose elements are all equal to 1. Solving the above linear equations, we get that

$$\pi_i = \frac{1}{k}, \quad i = 1, 2, ..., k. \tag{1}$$

By Theorem 3.1.1 in Chapter 3 of Neuts [16], the equilibrium condition of the system is given by

$$\boldsymbol{\pi A}_3 \boldsymbol{e} > \boldsymbol{\pi C}_1 \boldsymbol{e}.$$

Making substitution for $\boldsymbol{\pi}$ by Eq. (1), then we have the equilibrium condition of the system given by

$$\frac{\beta\lambda}{\mu_2} < 1. \tag{2}$$

Remark 2.1 We observe from the above condition that the equilibrium condition of the system is dependent with the fast service rate μ_2 and independent with the slow service rate μ_1. This is in agreement with the equilibrium condition obtained by Rao [17], where Rao considered an M/G/1 queueing system in which customers balk with a constant probability $1-\beta$ and renege according to a negative exponential distribution. It has been shown that as long as reneging is permitted, the steady states always exist, but when no reneging is permitted, the steady states exist only when $\lambda\beta\eta < 1$, where λ is the arrival rate of customers, and η is the mean service time of a customer.

3 Performance Measures and Cost Model

In this section, we first derive the explicit expression for the steady-state probability vector. Then, we give some useful performance measures of the system. Based on these performance measures, we develop a cost model to determine the optimal critical value r to minimize the total expected cost per unit time.

3.1 Steady-state probability vector

Let $\boldsymbol{X} = (X_0, \boldsymbol{X}_1, ..., \boldsymbol{X}_r, \boldsymbol{X}_{r+1}, ...)$, where X_0 is a number, $\boldsymbol{X}_i$ $(i = 1, 2, ...)$ is a vector of order k. By applying the matrix geometric solution method [16], the stationary probability vector is given by

$$\boldsymbol{X}_i = \boldsymbol{X}_r \boldsymbol{R}^{i-r}, \qquad i = r,\ r+1,\ ... \tag{3}$$

where $\boldsymbol{R}$ is the minimal nonnegative solution to the equation $\boldsymbol{R}^2\boldsymbol{A}_3 + \boldsymbol{R}\boldsymbol{B}_2 + \boldsymbol{C}_1 = 0$, and $X_0, \boldsymbol{X}_1, ..., \boldsymbol{X}_r$ are given by solving the following equations:

$$\begin{aligned}
X_0 B_0 + \boldsymbol{X}_1\boldsymbol{A}_1 &= 0, \\
X_0 \boldsymbol{C}_0 + \boldsymbol{X}_1\boldsymbol{B}_1 + \boldsymbol{X}_2\boldsymbol{A}_2 &= \boldsymbol{0}, \\
\boldsymbol{X}_i\boldsymbol{C}_1 + \boldsymbol{X}_{i+1}\boldsymbol{B}_1 + \boldsymbol{X}_{i+2}\boldsymbol{A}_2 &= \boldsymbol{0}, \qquad i = 1, 2, ..., r-2, \\
\boldsymbol{X}_{r-1}\boldsymbol{C}_1 + \boldsymbol{X}_r(\boldsymbol{B}_1 + \boldsymbol{R}\boldsymbol{A}_3) &= \boldsymbol{0}, \\
X_0 + \sum_{i=1}^{r-1} \boldsymbol{X}_i\boldsymbol{e} + \boldsymbol{X}_r(\boldsymbol{I} - \boldsymbol{R})^{-1}\boldsymbol{e} &= 1,
\end{aligned} \tag{4}$$

where $\boldsymbol{e}$ is a column vector of order k, and all its elements equal to 1.

In general, it is difficult to give an exact expression of $\boldsymbol{R}$ except for a few special cases. However, the matrix $\boldsymbol{R}$ can be approximately calculated by the following iterative procedure:

(a) $\boldsymbol{R}(0) = \boldsymbol{0}$,

(b) $\boldsymbol{R}(n+1) = -(\boldsymbol{C}_1 + \boldsymbol{R}^2(n)\boldsymbol{A}_3)\boldsymbol{B}_2^{-1}, \quad n \geq 0.$

This iterative algorithm is convergent, i.e. $\boldsymbol{R} = \lim_{n\to\infty} \boldsymbol{R}(n)$ (Section 1.9 of Chapter 1 of [16]).

Remark 3.1 The inverse of the matrix $\boldsymbol{B}_2^{-1}$ in the above algorithm exists, and can be explicitly given by

$$\boldsymbol{B}_2^{-1} = \frac{1}{a^k}\begin{pmatrix} a^{k-1} & 0 & 0 & \cdots & 0 \\ a^{k-2}(-b) & a^{k-1} & 0 & \cdots & 0 \\ a^{k-3}(-b)^2 & a^{k-2}(-b) & a^{k-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ (-b)^{k-1} & a(-b)^{k-2} & a^2(-b)^{k-3} & \cdots & a^{k-1} \end{pmatrix},$$

where $a = (-\beta\lambda - k\mu_2)$, $b = k\mu_2$.

For a special case of $k = 2$ and $r = 2$, we can readily obtain an explicit expression for matrix $\boldsymbol{R}$ given in the following theorem.

Theorem 3.1 *If $k = 2$ and $r = 2$, then the matrix $\boldsymbol{R}$ is explicitly given by*

$$\boldsymbol{R} = \frac{\beta\lambda}{4\mu_2^2}\begin{pmatrix} 2\mu_2 & \beta\lambda \\ 2\mu_2 & \beta\lambda + 2\mu_2 \end{pmatrix}. \tag{5}$$

Proof Let

$$\boldsymbol{R} = \begin{pmatrix} R_{00} & R_{01} \\ R_{10} & R_{11} \end{pmatrix},$$

then the equation $\boldsymbol{R}^2\boldsymbol{A}_3 + \boldsymbol{R}\boldsymbol{B}_2 + \boldsymbol{C}_1 = 0$ can be written as follows:

$$\begin{aligned} 2\mu_2(R_{00}^2 + R_{01}R_{10}) + (-\beta\lambda - 2\mu_2)R_{01} &= 0, \\ 2\mu_2(R_{00}R_{10} + R_{11}R_{10}) + (-\beta\lambda - 2\mu_2)R_{11} + \beta\lambda &= 0, \\ (-\beta\lambda - 2\mu_2)R_{00} + 2\mu_2R_{01} + \beta\lambda &= 0, \\ (-\beta\lambda - 2\mu_2)R_{10} + 2\mu_2R_{11} &= 0. \end{aligned} \tag{6}$$

Noting that $\boldsymbol{R}\boldsymbol{A}_3\boldsymbol{e} = \boldsymbol{C}_1\boldsymbol{e}$, (Eq. (3.1.6) in [16]), we obtain that

$$R_{00} = \frac{\beta\lambda}{2\mu_2}, \qquad R_{10} = \frac{\beta\lambda}{2\mu_2}. \tag{7}$$

Substituting Eq. (7) into Eq. (6), the other two elements R_{01} and R_{11} are readily obtained. Then, we obtained the matrix $\boldsymbol{R}$. □

In this special case, the vector $\boldsymbol{X}$ is given by

$$\boldsymbol{X}_i = \boldsymbol{X}_2\boldsymbol{R}^{i-2}, \qquad i = 2, 3, ..., \tag{8}$$

and the number X_0, the vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ satisfy the following equations:

$$\begin{aligned} X_0B_0 + \boldsymbol{X}_1\boldsymbol{A}_1 &= 0, \\ X_0\boldsymbol{C}_0 + \boldsymbol{X}_1\boldsymbol{B}_1 + \boldsymbol{X}_2\boldsymbol{A}_2 &= \boldsymbol{0}, \\ \boldsymbol{X}_1\boldsymbol{C}_1 + \boldsymbol{X}_2(\boldsymbol{B}_1 + \boldsymbol{R}\boldsymbol{A}_3) &= \boldsymbol{0}, \\ X_0 + \boldsymbol{X}_1\boldsymbol{e} + \boldsymbol{X}_2(\boldsymbol{I} - \boldsymbol{R})^{-1}\boldsymbol{e} &= 1, \end{aligned} \tag{9}$$

where

$$(\boldsymbol{I} - \boldsymbol{R})^{-1} = \frac{1}{4\mu_2\beta\lambda - 4\mu_2^2}\begin{pmatrix} \beta^2\lambda^2 + 2\mu_2\beta\lambda - 4\mu_2^2 & \beta^2\lambda^2 \\ (-2\mu_2)\beta\lambda & (-2\mu_2)(2\mu_2 - \beta\lambda) \end{pmatrix}. \tag{10}$$

Thus, X_0, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are readily obtained by solving Eq. (9). However, it is not simple to present it explicitly since the expressions are tediously long. Thus, their expressions are omitted.

3.2 Performance measures

Using the steady-state probability vector presented above, we can obtain some performance measures of the system. Let

$$\boldsymbol{X}_i = (x_{i1}, x_{i2}, ..., x_{ik}), \qquad i = 1, 2, ...,$$

then we have the following theorem.

Theorem 3.2 *(a) The expected number of customers in the system is given by*

$$E(N) = \sum_{n=1}^{r-1} n\boldsymbol{X}_n\boldsymbol{e} + \boldsymbol{X}_r \left[r(\boldsymbol{I} - \boldsymbol{R})^{-1} + \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{R})^{-2} \right] \boldsymbol{e}. \tag{11}$$

(b) The expected number of customers in the queue is given by

$$E(N_q) = \sum_{n=1}^{r-2} n\boldsymbol{X}_{n+1}\boldsymbol{e} + \boldsymbol{X}_r \left[(r-1)(\boldsymbol{I} - \boldsymbol{R})^{-1} + \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{R})^{-2} \right] \boldsymbol{e}. \tag{12}$$

(c) The mean balking rate of the system is given by

$$BR = (1 - \beta)\lambda(1 - X_0). \tag{13}$$

Proof The expected number of customers in the system is given by

$$E(N) = \sum_{i=1}^{k} \sum_{n=1}^{\infty} n x_{n,i} = \sum_{n=1}^{\infty} n\boldsymbol{X}_n\boldsymbol{e}.$$

From Eq. (3), we have

$$E(N) = \sum_{n=1}^{r-1} n\boldsymbol{X}_n\boldsymbol{e} + \sum_{n=r}^{\infty} n\boldsymbol{X}_r\boldsymbol{R}^{n-r}\boldsymbol{e}. \tag{14}$$

Hence, we obtain Eq. (11) by summation. Similarly, the expected number of customers in the queue is given by

$$E(N_q) = \sum_{i=1}^{k} \sum_{n=1}^{\infty} n x_{n+1,i} = \sum_{n=1}^{r-2} n\boldsymbol{X}_{n+1}\boldsymbol{e} + \sum_{n=r-1}^{\infty} n\boldsymbol{X}_r\boldsymbol{R}^{n-r+1}\boldsymbol{e}. \tag{15}$$

Hence, we obtain Eq. (12) by summation. Using the concept of Ancker and Gafarian [7], the mean balking rate of the system is given by

$$BR = \sum_{n=1}^{\infty} b_n\lambda\boldsymbol{X}_n\boldsymbol{e} = (1 - \beta)\lambda(1 - X_0). \tag{16}$$

Obviously, the probability that the server is busy is given by

$$P_B = 1 - X_0 \tag{17}$$

and the probability that the server is idle is given by

$$P_0 = X_0. \tag{18}$$

□

3.3 Cost model

In this subsection, we develop a steady-state expected cost function where the critical value r is a decision variable. Our objective is to determine the critical value r to minimize the total expected cost per unit time. Let

C_1=cost per unit time when customers are waiting for service,

C_2=cost per unit time when the server is busy,

C_3=loss cost per unit time when customers balk,

C_4=cost per unit time when the server is idle.

According to the definition of each cost of the parameters listed above, the total expected cost function per unit time is given by

$$F(r) = C_1 E(N_q) + C_2 P_B + C_3 BR + C_4 P_0, \tag{19}$$

where $E(N_q)$, BR, P_B, P_0 are given in Eqs. (12) and (13) and Eqs. (17) and (18). The first item of Eq. (19) is the cost incurred by the customer's waiting. The second and the last items of Eq. (19) are the costs incurred by the server. The third item of Eq. (19) is the cost incurred by the customer loss.

4 Sensitivity Analysis

In this section, we perform a sensitivity analysis on the optimal value r^* and its expected cost $F(r^*)$ based on changes in the values of the system parameters such as the arrival rate λ, the probability β, the slow service rate μ_1, the fast service rate μ_2 and cost parameters.

Let the service time follow a 2-stage Erlang distribution, and employ the cost parameters $C_1 = 100$, $C_2 = 150$, $C_3 = 300$ and $C_4 = 450$. The numerical results of the optimal critical value r^* and its expected minimum cost $F(r^*)$ are illustrated in Figures 4.1–4.4.

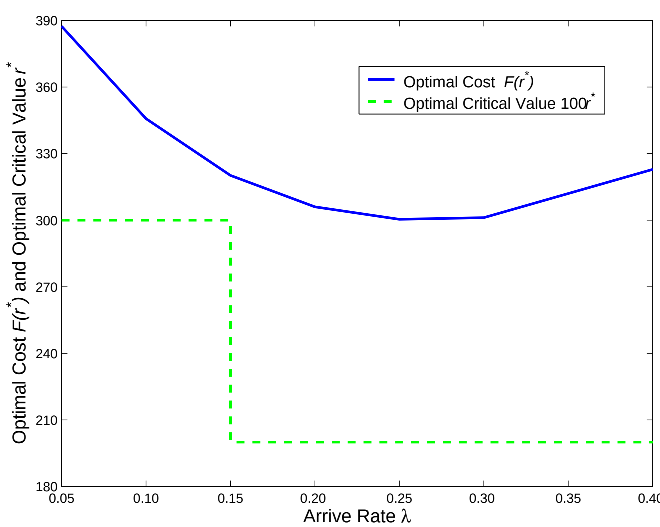


Figure 4.1: Optimal cost $F(r^*)$ and optimal critical value r^* versus arrive rate λ with $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$.

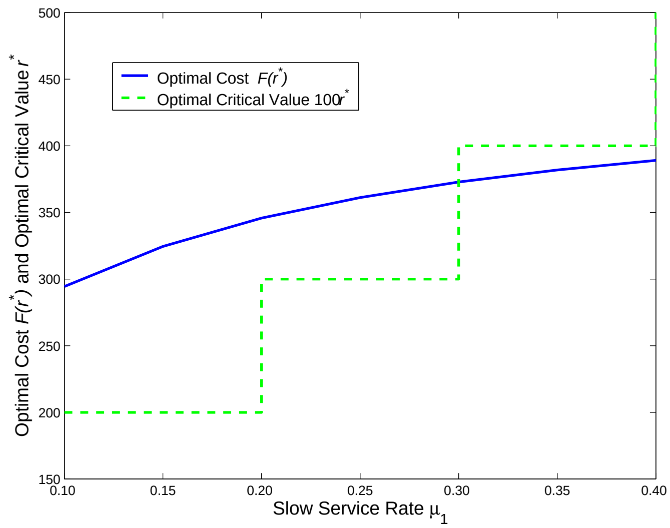


Figure 4.2: Optimal cost $F(r^*)$ and optimal critical value r^* versus slow service rate μ_1 with $\lambda = 0.1$, $\mu_2 = 0.8$ and $\beta = 0.5$.

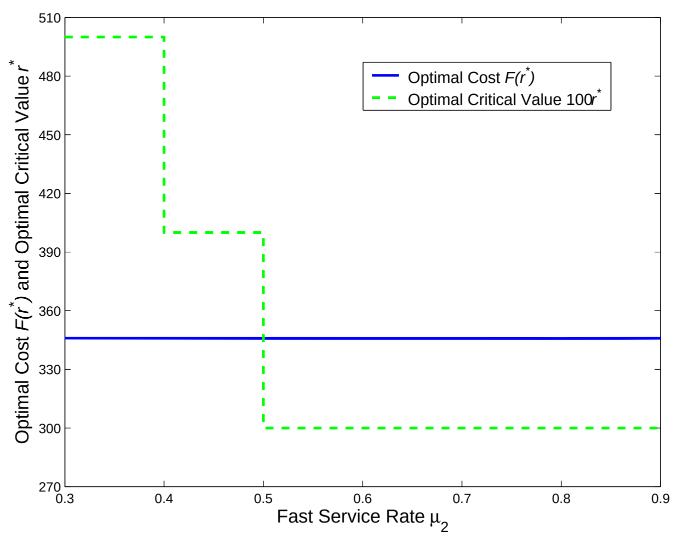


Figure 4.3: Optimal cost $F(r^*)$ and optimal critical value r^* versus fast service rate μ_2 with $\lambda = 0.1$, $\mu_1 = 0.2$ and $\beta = 0.5$.

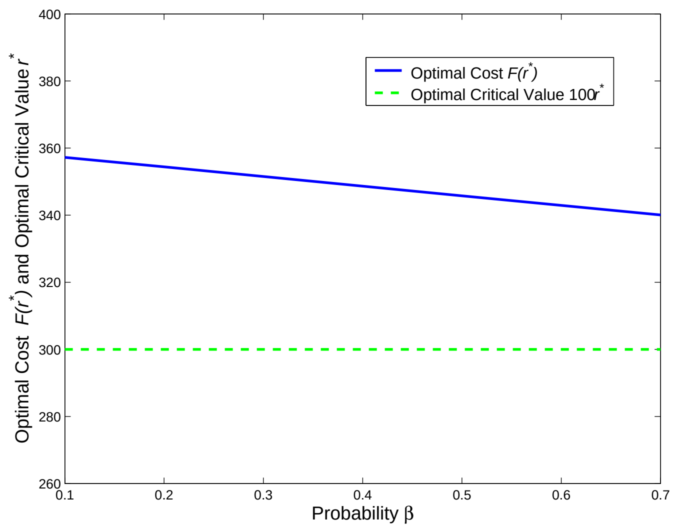


Figure 4.4: Optimal cost $F(r^*)$ and optimal critical value r^* versus probability β with $\lambda = 0.1$, $\mu_1 = 0.2$ and $\mu_2 = 0.8$.

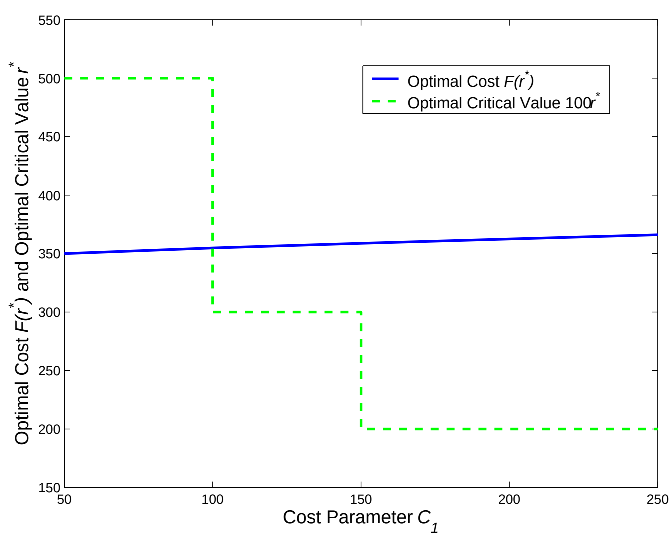


Figure 4.5: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_1 with $C_2 = 150$, $C_3 = 300$ and $C_4 = 450$.

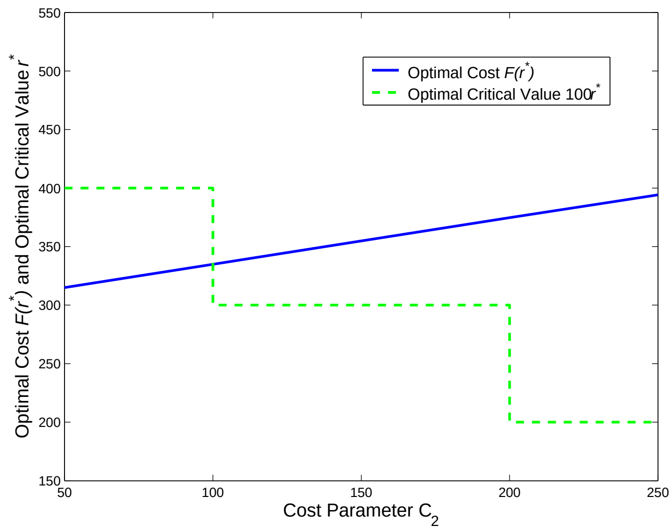


Figure 4.6: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_2 with $C_1 = 100$, $C_3 = 300$ and $C_4 = 450$.

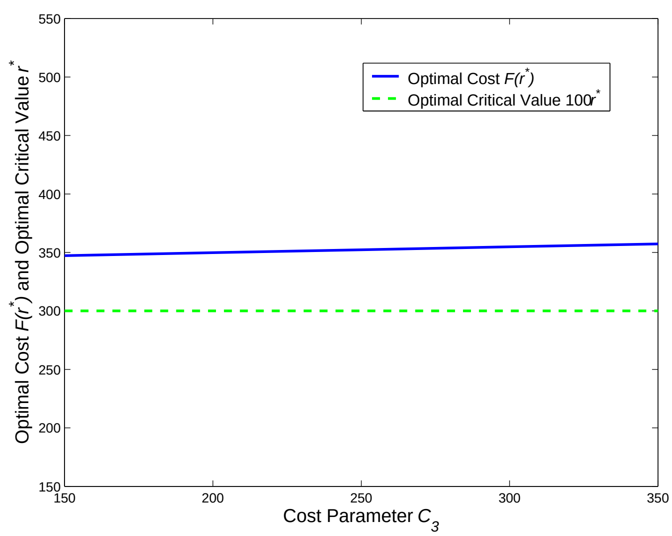


Figure 4.7: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_3 with $C_1 = 100$, $C_2 = 150$ and $C_4 = 450$.

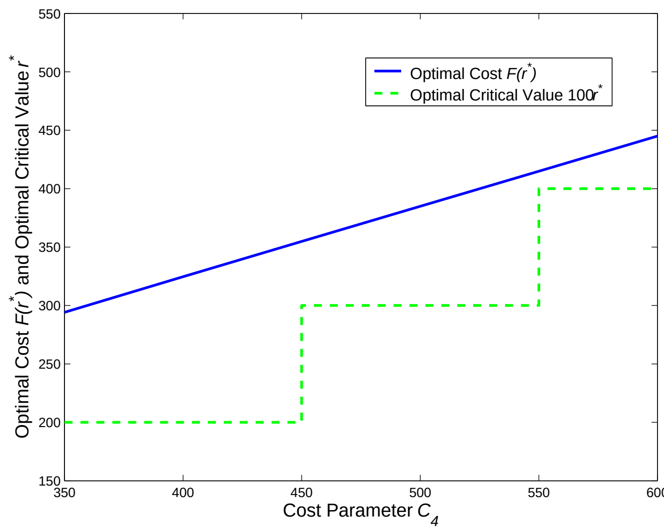


Figure 4.8: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_4 with $C_1 = 100$, $C_2 = 150$ and $C_3 = 300$.

In Figure 4.1, we fix $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the arrival rate λ. Figure 4.1 shows that: (i) The optimal critical value r^* decreases as λ increases from 0.05 to 0.15, while it does not change at all when λ varies from 0.15 to 0.4. (ii) The minimum expected cost $F(r^*)$ first decreases and then increases as λ increases. Intuitively, the optimal critical value r^* is insensitive to changes with λ.

In Figure 4.2, we fix $\lambda = 0.1$, $\mu_2 = 0.8$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the slow service rate μ_1. Figure 4.2 shows that the optimal critical value r^* and its minimum expected cost $F(r^*)$ increase as μ_1 increases.

In Figure 4.3, we fix $\lambda = 0.1$, $\mu_1 = 0.2$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the fast service rate μ_2. Figure 4.3 shows that: (i) The optimal critical value r^* decreases as μ_2 increases from 0.3 to 0.5, while it does not change at all when μ_2 varies from 0.5 to 0.9. (ii) The minimum expected cost $F(r^*)$ rarely changes when μ_2 varies from 0.3 to 0.9. Intuitively, the optimal critical value r^* and its minimum expected cost may be too insensitive to changes with μ_2.

In Figure 4.4, we fix $\lambda = 0.1$, $\mu_1 = 0.2$ and $\mu_2 = 0.8$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the probability β. Figure 4.4 shows that: (i) The optimal critical value r^* does not change at all when β varies from 0.1 to 0.7. (ii) The minimum expected cost $F(r^*)$ decreases slightly as β increases. Intuitively, the optimal critical value r^* and its expected minimum cost $F(r^*)$ are insensitive to changes with β.

It appears from Figures 4.1–4.4 that: (i) β does not affect r^*, but slightly affects $F(r^*)$. (ii) λ affects $F(r^*)$ and slightly affects r^*. (iii) μ_2 rarely affects r^* and $F(r^*)$. And (iv) μ_1 affects r^* and $F(r^*)$ significantly.

Furthermore, we perform a sensitivity analysis on the optimal value r^* and its expected cost $F(r^*)$ based on changes in values of the cost parameters C_1, C_2, C_3 and C_4.

Let the service time follow a 2-stage Erlang distribution, and employ the system parameters $\lambda = 0.1$, $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$. The numerical results of the optimal critical value r^* and its expected minimum cost $F(r^*)$ are illustrated in Figures 4.5–4.8.

It can be easily seen from Figures 4.5–4.8 that: (i) The cost C_1 slightly affects r^* and $F(r^*)$ when C_1 is larger than 150. (ii) The cost C_2 and C_4 affect r^* and $F(r^*)$ significantly. And (iii) the cost C_3 rarely affects r^* and $F(r^*)$.

5 Conclusions

We have considered an M/E$_k$/1 queuing system with balking and state-dependent service rate. By using the matrix geometric solution, we have obtained the matrix solution of the steady-state probability distribution and the explicit expressions of some performance measures of the system. Based on these performance measures, we have developed a cost model to determine the optimal critical value r^* to minimize the total expected cost per unit time. Furthermore, we have performed sensitivity analysis for the optimal critical value r^* and its expected minimum cost $F(r^*)$ with various parameters. Our numerical investigations indicate that: (i) The slow service rate μ_1, the cost parameters C_2 and C_4 affect the optimal critical value r^* and its expected minimum cost $F(r^*)$ significantly. (ii) The other parameters such as λ, μ_2, β and the cost parameters C_1 and C_3 rarely affect the optimal critical value r^* and its expected minimum cost $F(r^*)$.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (No. 10271102) and the Natural Science Foundation of Hebei Province (No. A2004000185), China, and was supported in part by GRANT-IN-AID FOR SCIENTIFIC RESEARCH (No. 16560350) and MEXT.ORC (2004-2008), Japan.

References

[1] Gaver, D. P. and Lehoczky, J. P. Gaussian approximations to service problems: a communication system example. *Journal of Applied Probability* **13** (1976) 768–780.

[2] Sharma, R. R. and Rai, R. C. and Mishra, A. Optimal bus services on express basis in the case of balking and reneging. *European Journal of Operational Research* **66** (1993) 113–123.

[3] Ackere, V. and Ninios, P. Simulation and queueing theory applied to a single-server queue with advertising and balking. *Journal of the Operational Research Society* **44** (1993) 407–414.

[4] Shawky, A. I. The single-server machine interference model with balking, reneging and an additional server for longer queues. *Microelectronic and Reliability* **37** (1997) 355–357.

[5] Haight, F. A. Queueing with balking. *Biometrika* **44** (1957) 360–369.

[6] Haight, F. A. Queueing with reneging. *Metrika* **2** (1959) 186–197.

[7] Ancker, Jr. C. J. and Gafarian, A. V. Some queueing problems with balking and reneging: I. *Operation Research* **11** (1963) 88–100.

[8] Ancker, Jr. C. J. and Gafarian, A. V. Some queueing problems with balking and reneging: II. *Operation Research* **11** (1963) 928–937.

[9] Abou-EI-Ata, M. O. The state-dependent queue: M/M/1/N with reneging and general balk functions. *Microelectronic and Reliability* **31** (1991) 1001–1007.

[10] Abou-EI-Ata, M. O. and Kotb, K. A. M. Linearly dependent service rate for the queue: M/M/1/N with general balk function, reflecting barrier, reneging and additional service for longer queue. *Microelectronic and Reliability* **32** (1992) 1693–1698.

[11] Abou-EI-Ata, M. O., Alseedy, R. O. and Kotb, K. A. M. Transient solution of the state-dependent queue: M/M/1/N with balking and reflecting barrier. *Microelectronic and Reliability*, **33** (1993) 681–688.

[12] Abou-EI-Ata, M. O. and Shawky, A. I. Single-server Markovian overflow queue with balking, reneging and an additional server for longer queues. *Microelectronic and Reliability* **32** (1992) 1389–1394.

[13] Drekic, S. and Woolford, D. G. A preemptive priority queue with balking. *European Journal of Operational Research* **164** (2005) 387–401.

[14] Al-Seedy, R. O. and Kotb, K. A. M. Transient solution of the state-dependent queue: M/M/1 with balking. *Advances in Modelling* & *Simulation* **35** (1991) 55–64.

[15] Al-Seedy, R. O. Analytical solution of the state-dependent Erlangian queue: $M/E_j/1/N$ with balking. *Microelectronic and Reliability* **36** (1996) 203–206.

[16] Neuts, M. F. *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore, 1981.

[17] Rao, S. S. Queuing with balking and reneging in M/G/1 systems. *Metrika* **12** (1968) 173–188.