Volume 18, Number 4, 2018 ISSN 1562-8353

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys

Volume 18 Number 4 2018

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 18, No. 4, 2018**

# Nonlinear Dynamics and Systems Theory

## An International Journal of Research and Surveys



**InforMath Publishing Group**
**http://www.e-ndst.kiev.ua**

# Nonlinear Dynamics and Systems Theory

An International Journal of Research and Surveys






## INSTRUCTIONS FOR CONTRIBUTORS

**(1) General.** Nonlinear Dynamics and Systems Theory (ND&ST) is an international journal devoted to publishing peer-refereed, high quality, original papers, brief notes and review articles focusing on nonlinear dynamics and systems theory and their practical applications in engineering, physical and life sciences. Submission of a manuscript is a representation that the submission has been approved by all of the authors and by the institution where the work was carried out. It also represents that the manuscript has not been previously published, has not been copyrighted, is not being submitted for publication elsewhere, and that the authors have agreed that the copyright in the article shall be assigned exclusively to InforMath Publishing Group by signing a transfer of copyright form. Before submission, the authors should visit the website:

**http://www.e-ndst.kiev.ua**

for information on the preparation of accepted manuscripts. Please download the archive Sample_NDST.zip containing example of article file (you can edit only the file Samplefilename.tex).

**(2) Manuscript and Correspondence.** Manuscripts should be in English and must meet common standards of usage and grammar. To submit a paper, send by e-mail a file in PDF format directly to

*Professor A.A. Martynyuk,* Institute of Mechanics,
Nesterov str.3, 03057, MSP 680, Kiev-57, Ukraine
e-mail: journalndst@gmail.com

or to one of the Regional Editors or to a member of the Editorial Board. Final version of the manuscript must typeset using LaTex program which is prepared in accordance with the style file of the Journal. Manuscript texts should contain the title of the article, name(s) of the author(s) and complete affiliations. Each article requires an abstract not exceeding 150 words. Formulas and citations should not be included in the abstract. AMS subject classifications and key words must be included in all accepted papers. Each article requires a running head (abbreviated form of the title) of no more than 30 characters. The sizes for regular papers, survey articles, brief notes, letters to editors and book reviews are: (i) 10-14 pages for regular papers, (ii) up to 24 pages for survey articles, and (iii) 2-3 pages for brief notes, letters to the editor and book reviews.

**(3) Tables, Graphs and Illustrations.** Each figure must be of a quality suitable for direct reproduction and must include a caption. Drawings should include all relevant details and should be drawn professionally in black ink on plain white drawing paper. In addition to a hard copy of the artwork, it is necessary to attach the electronic file 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] Poincare, H. Title of the article. *Title of the Journal* **volume** (issue) (year) pages. [Language]

Book: [2] Lyapunov, 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 Sample Copy.** Proofs sent to authors should be returned to the Editorial Office with corrections within three days after receipt. The corresponding author will receive a sample copy of the issue of the Journal for which his/her paper is published.

**(6) Editorial Policy.** Every submission will undergo a stringent peer review process. An editor will be assigned to handle the review process of the paper. He/she will secure at least two reviewers' reports. The decision on acceptance, rejection or acceptance subject to revision will be made based on these reviewers' reports and the editor's own reading of the paper.

**NONLINEAR DYNAMICS AND SYSTEMS THEORY**
An International Journal of Research and Surveys
Published by InforMath Publishing Group since 2001

Volume 18 Number 4 2018

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

*Impact Factor from SCOPUS for 2017: SNIP – 0.707, SJR – 0.316*


**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 Curtin University of Technology (Perth, Australia). It aims to publish high quality original scientific papers and surveys in areas of nonlinear dynamics and systems theory and their real world applications.

## AIMS AND SCOPE

**Nonlinear Dynamics and Systems Theory** is a multidisciplinary journal. It publishes papers focusing on proofs of important theorems as well as papers presenting new ideas and new theory, conjectures, numerical algorithms and physical experiments in areas related to nonlinear dynamics and systems theory. Papers that deal with theoretical aspects of nonlinear dynamics and/or systems theory should contain significant mathematical results with an indication of their possible applications. Papers that emphasize applications should contain new mathematical models of real world phenomena and/or description of engineering problems. They should include rigorous analysis of data used and results obtained. Papers that integrate and interrelate ideas and methods of nonlinear dynamics and systems theory will be particularly welcomed. This journal and the individual contributions published therein are protected under the copyright by International InforMath Publishing Group.

## PUBLICATION AND SUBSCRIPTION INFORMATION


**Nonlinear Dynamics and Systems Theory** will have 4 issues in 2019, printed in hard copy (ISSN 1562–8353) and available online (ISSN 1813–7385), by InforMath Publishing Group, Nesterov str., 3, Institute of Mechanics, Kiev, MSP 680, Ukraine, 03057. Subscription prices are available upon request from the Publisher, EBSCO Information Services (`mailto:journals@ebsco.com`), or website of the Journal: `http://e-ndst.kiev.ua`. Subscriptions are accepted on a calendar year basis. Issues are sent by airmail to all countries of the world. Claims for missing issues should be made within six months of the date of dispatch.


## ABSTRACTING AND INDEXING SERVICES

Papers published in this journal are indexed or abstracted in: Mathematical Reviews / MathSciNet, Zentralblatt MATH / Mathematics Abstracts, PASCAL database (INIST–CNRS) and SCOPUS.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 319–330

# Existence and Approximation of Solutions for Systems of First Order Differential Equations

K. Al-Khaled *

*Department of Mathematics and Statistics, Jordan University of Science and Technology, P.O. Box(3030), Irbid 22110, Jordan*

Received: March 8, 2018; Revised: October 6, 2018

**Abstract:** The purpose of this paper is to present some general results concerning the existence of solutions for systems of differential equations. The existence results to be presented will be based on an effective procedure for constructing approximate solution. Namely, a numerical scheme using the Sinc function, in which it is shown that the solution converges exponentially. Furthermore, a numerical example and comparisons are presented to prove the validity of the suggested method.

**Keywords:** *fixed-point theory; numerical solutions; systems of differential equations; existence of solutions.*

**Mathematics Subject Classification (2010):** 26A33, 35F25, 35C10.

## 1 Introduction

In addition to its intrinsic mathematical interest, the theory of ordinary differential equations has extensive applications in many general fields, for instance, physics, chemistry, biology, economics and engineering. The existence and uniqueness of a solution to a first-order differential equation, given a set of initial conditions, is one of the most fundamental results of ordinary differential equations.

In this paper, we shall confine our discussion to systems of first order differential equations of the form

$$\begin{aligned} \tfrac{dx_1}{dt} &= F_1(x_1, x_2, ..., x_n, t), \\ \tfrac{dx_2}{dt} &= F_2(x_1, x_2, ..., x_n, t), \\ &\vdots \\ \tfrac{dx_n}{dt} &= F_n(x_1, x_2, ..., x_n, t), \end{aligned} \tag{1}$$

* Corresponding author: mailto:kamel@just.edu.jo

that are often encountered in many mathematical models such as the Neutron flow, electrical networks, residential segregation. Also the first order differential equations are shown to be adequate models for various physical phenomena in the areas like damping laws, and diffusion processes. If $\mathbf{x} = (x_1, x_2, ..., x_n)^t$ denotes an element in $R^n$, the system of equations (1) can be written in a more compact form

$$\mathbf{x}' = F(t)\mathbf{x} + g(t, \mathbf{x}), \tag{2}$$

where $F(t)$ is a continuous $n \times n$ matrix function, its norm is defined as $\parallel F \parallel = \sup_{\|x\|=1} \parallel F\mathbf{x} \parallel$ .

Consider a special case of the system of first order differential equations

$$\begin{aligned} \tfrac{du_1}{dt} &= b_{11}(t)u_1 + b_{12}(t)u_2 + f_1(t, u_1, u_2), \\ \tfrac{du_2}{dt} &= b_{21}(t)u_1 + b_{22}(t)u_2 + f_2(t, u_1, u_2) \end{aligned} \tag{3}$$

on the interval $[a, T]$ with the conditions

$$u_1(a) = u_1^0, \;\; u_2(a) = u_2^0. \tag{4}$$

We write equation (3) in a vector form as

$$\frac{d\vec{u}}{dt} = B(t)\vec{u} + \vec{f}(t, \vec{u}), \tag{5}$$

where $B$ is the matrix $[b_{ij}(t)]_{2\times 2},\ i, j = 1, 2$ and $\vec{u} = (u_1, u_2)^t,\ \vec{f} = (f_1, f_2)^t$.

Many other authors have studied, under some conditions, the existence and uniqueness of solutions for systems of first-order differential equations. For example, in [8] Fransis and Miller examined fundamental and general existence theorems along with the Picard iterations. The author in [12] extended the version of Caratheodory's existence theorem for ordinary differential equations. While in [9], representation and approximation of the solutions to linear equations are studied. The authors in [14] used a recent Schauder-type result for discontinuous operators to study the existence of absolutely continuous solutions of first order initial value problems. Existence theorems for iterative differential equations as well as convergence theorems for a fixed point iteration designed to approximate the solutions are proved in [5]. Some applications of the fixed point theory to a nonlinear differential equations are presented in [7]. The fractional derivatives accurately describe natural phenomena that occur in general physical problems, existence and uniqueness of solutions for coupled systems of fractional differential equations are studied in [1]. The authors in [11], studied sufficient conditions for the existence of optimal controls for system of functional-differential equations.

The objectives of this paper are twofold. Firstly, we use the Schauder fixed point theorem to develop an existence theory for a general class of systems of first order differential equations (1), subject to the linear constraint

$$\ell\mathbf{x} = r, \tag{6}$$

where $r \in R^n$. Mainly, we will redo the first section from [2]. Secondly, we aim to implement the Sinc methodology to find approximate solutions for systems of differential equations (5).

## 2 Preliminary Results

Let us review some notations and facts that will be used in this paper. Let $C[\alpha,\beta]$ denote the Banach space of continuous functions such that $x(t)$ is a map that transforms the closed interval $[\alpha,\beta]$ into $R^n$, where the norm is defined as

$$\parallel x \parallel \equiv \max_{[\alpha,\beta]} \parallel x(t) \parallel . \tag{7}$$

We also define the domain $\mathbf{D} = \{(t,\mathbf{x}), t \in [\alpha,\beta], \mathbf{x} \in R^n\}$, then we assume the following three hypotheses:

**H1.** $F(t)$ is a square matrix of size $n$ in which each entry is a continuous function on $[\alpha,\beta]$.

**H2.** $g(t,\mathbf{x})$ is continuous on the domain $\mathbf{D}$ defined above.

**H3.** $\mathbf{M} = \mathbf{M}(\alpha,\beta)$ is a bounded linear mapping from $C[\alpha,\beta]$ into $R^n$ with bound $\parallel \mathbf{M} \parallel \equiv \sup_{\parallel\mathbf{x}\parallel=1} \parallel \mathbf{Mx} \parallel$ .

A function $\mathbf{x}(t) \in C[\alpha,\beta]$ which has a continuous derivative that satisfies equation (2) on $[\alpha,\beta]$ is called the solution to (2). We consider the following mappings that are defined from the space $C[\alpha,\beta]$ into $R^n$

$$\begin{aligned} &\mathbf{M_0x} \equiv \mathbf{x}(\alpha),\\ &\mathbf{M_1x} \equiv (\mathbf{x}_1(t_1), \mathbf{x}_2(t_2), ..., \mathbf{x}_n(t_n)),\\ &\mathbf{M_2x} \equiv \textstyle\int_\alpha^\beta \mathbf{x}(\tau)d\tau. \end{aligned} \tag{8}$$

In order to prove our main result in this paper, we do need a sequence of lemmas.

**Lemma 2.1** *The above defined mappings* $\mathbf{M_0}, \mathbf{M_1}$, *and* $\mathbf{M_2}$ *satisfy* $\parallel \mathbf{M_0} \parallel = 1$, $\parallel \mathbf{M_1} \parallel \leq n$ *and* $\parallel \mathbf{M_2} \parallel = \beta - \alpha$.

**Proof:** Linearity of the mappings is obvious. For bounds, we have

$$\parallel \mathbf{M_0}(\mathbf{x}) \parallel = \parallel \mathbf{x}(\alpha) \parallel \leq \max_{[\alpha,\beta]} \parallel \mathbf{x}(\mathbf{t}) \parallel .$$

Using Equation (7), we arrive at $\parallel \mathbf{M_0} \parallel \leq \parallel \mathbf{x} \parallel$. For $\mathbf{x}(t) \equiv c \in R^n$ we obtain the equality $\parallel \mathbf{M_0}(\mathbf{x}) \parallel = \mathbf{1}$. For the mapping $\mathbf{M_1}(\mathbf{x})$, we have

$$\parallel \mathbf{M_1}(\mathbf{x}) \parallel = \sum_{\mathbf{i=1}}^{\mathbf{n}} |\mathbf{x}(\mathbf{t_i})| \leq \mathbf{n} \parallel \mathbf{x} \parallel,$$

which implies that $\parallel \mathbf{M_1} \parallel \leq n$. Finally,

$$\parallel \mathbf{M_2}(\mathbf{x}) \parallel = \parallel \int_\alpha^\beta \mathbf{x}(\tau)\mathbf{d}\tau \parallel \leq \int_\alpha^\beta \parallel \mathbf{x}(\tau) \parallel \mathbf{d}\tau \leq (\beta-\alpha) \parallel \mathbf{x} \parallel .$$

For equality, let $\mathbf{x}(t) \equiv c \in R^n$, that is, $\|\mathbf{M_2}\| = (\beta - \alpha)$.

**Lemma 2.2** *(Schauder, see [6]) Let $B$ be a convex compact subset of a normal linear space $X$, then any continuous mapping $L$ from $B$ into $B$ has a fixed point in $B$.*

**Lemma 2.3** *If $B$ is a closed convex subset of the Banach space $X$, and $L$ is a continuous mapping of $B$ into $B$ such that $L(B)$ is relatively compact, then $L$ has a fixed point in $B$.*

**Proof:** As $L(B)$ is a subset of the closed set $B$, then $cl(L(B))$ is also a subset of $B$. Let $\hat{B}$ denote the closed convex hull of $cl(L(B))$. By Lemma 2.2, $\hat{B}$ is compact and is also the smallest closed convex set containing $cl(L(B))$, this would imply that $\hat{B}$ is subset of $B$. Therefore $L(\hat{B}) \subset L(B) \subset cl(L(B)) \subset \hat{B}$. By Lemma 2.2, $L$ has a fixed point in $\hat{B}$. We would like to impose one more hypothesis in addition to **H1**, **H2** and **H3** mentioned above:

**H4** The initial value problem

$$\mathbf{x}' = F(t)\mathbf{x}, \quad \ell\mathbf{x} = r, \tag{9}$$

for any $r \in R^n$, the problem in (9) has a unique solution. Therefore, for the $n-$dimensional space of solutions, call it $\mathcal{F}$, to the problem $\mathbf{x}' = F(t)\mathbf{x}$, $(\ell|\mathcal{F})^{-1}$ exists, i.e., the null space of $\ell|\mathcal{F}$ is $\{0\}$.

For the application of Lemma 2.3, we define an appropriate mapping together with the following lemma.

**Lemma 2.4** *If $F(t)$ satisfies* **H1**, *then the problem*

$$\mathbf{x}' = \mathbf{F}(t)\mathbf{x} + z(t), \ \mathbf{x}(t_0) = r \tag{10}$$

*has a unique solution $\omega(t)$ on $[\alpha, \beta]$ for any $t_0 \in [\alpha, \beta)$, $z \in C[\alpha, \beta]$, and $r \in R^n$. Moreover, for any $t$, $t_0$ in $[\alpha, \beta]$,*

$$\|\omega(t)\| \leq \|\omega(t_0)\| \exp\Big|\int_{t_0}^{t} \|F(\tau)\| d\tau\Big| + \Big|\int_{t_0}^{t} \exp\Big|\int_{\tau}^{t} \|F(s)\| ds\Big| \|z(\tau) d\tau\|\Big|. \tag{11}$$

The existence and uniqueness of solutions to such initial value problems is well known [10].

**Lemma 2.5** *If* **H1**, **H3** *and* **H4** *are satisfied, then the problem*

$$\mathbf{x}' = \mathbf{F}(t)\mathbf{x} + z(t), \ \ell\mathbf{x} = r \tag{12}$$

*has a unique solution $\omega(t)$ for any $r \in R^n$ and $z \in C[\alpha, \beta]$.*

**Proof:** Let

$$\omega(t) \equiv \omega_0(t) + (\ell|\mathcal{F})^{-1}(-\ell\omega_0) + (\ell|\mathcal{F})^{-1} r, \tag{13}$$

where $\omega_0$ is the unique solution to

$$\mathbf{x}' = \mathbf{F}(t)\mathbf{x} + z(t), \ \mathbf{x}(\alpha) = 0. \tag{14}$$

It is easily verified by differentiation that $\omega(t)$ is a solution to (12). If $\omega_1(t), \omega_2(t)$ are two solutions to (12), then $\omega_1(t) - \omega_2(t)$ is the unique solution to

$$\mathbf{x}' = \mathbf{F}(t)\mathbf{x}, \quad \ell\mathbf{x} = 0. \tag{15}$$

Hence, by property **H4**, $\omega_1(t) \equiv \omega_2(t)$, that is, $\omega(t)$ is the unique solution to (12).

**Lemma 2.6** *Suppose* **H1** *and* **H4** *are satisfied. If $\hat{\omega}_0(t)$ is the unique solution to (12) with $r = 0$, then*

$$\|\hat{\omega}_0(t)\| \le \int_\alpha^\beta \Big( \exp \int_\tau^\beta \|F(s)\| ds \Big) \|z(\tau)\| d\tau \Big( 1 + \|(\ell|\mathcal{F})^{-1}\ell\| \Big). \tag{16}$$

*In particular,*

$$\|\hat{\omega}_0\| \le K_1 \|z\|, \tag{17}$$

*where $K_1 = (\beta - \alpha) \exp \int_\alpha^\beta \|F(s)\| ds (1 + \|(\ell|\mathcal{F})^{-1}\ell\|)$.*

**Proof:** We have from (2.7) that

$$\hat{\omega}_0(t) = \omega_0(t) + (\ell|\mathcal{F})^{-1}(-\ell\omega_0), \tag{18}$$

where $\omega_0(t)$ is the unique solution to the initial value problem

$$\mathbf{x}' = F(t)\mathbf{x} + z(t), \quad \mathbf{x}(\alpha) = 0. \tag{19}$$

Thus

$$\|\hat{\omega_0}\| \le \|\omega_0\| + \|(\ell|\mathcal{F})^{-1}\ell\| \, \|\omega_0\| \le \|\omega_0\|(1 + \|(\ell|\mathcal{F})^{-1}\ell)\|. \tag{20}$$

From Lemma 2.4 with $t_0 = \alpha$ and $r = 0$, we have

$$\|\omega_0\| \le \int_\alpha^\beta \Big( \exp \int_\tau^\beta \|F(s)\| ds \Big) \|z(\tau)\| d\tau \tag{21}$$

and so (16) holds.

The following notation will be used in this paper. Let $H$ be a positive number, $\vec{H} = (H_1, H_2, ..., H_n)$, where each $H_i$ is positive, and $H(t)$ be a continuous positive function for $\alpha \le t \le \beta$. Let

$$\begin{aligned}
&\psi(r) = \psi(t, r, \alpha, \beta) \equiv (\ell(\alpha, \beta)|\mathcal{F})^{-1} r,\\
&C(H) = C(H, r, \alpha, \beta) \equiv \{y \in C[\alpha, \beta] : \|y - \psi(r)\| \le H\},\\
&D(H) = D(H, r, \alpha, \beta) \equiv \{(t, y) \in D(\alpha, \beta) : \|y - \psi(t; r)\| \le H\},\\
&D(H, t) = D(H, r, t) \equiv \{y \in R^n : \|y - \psi(t; r)\| \le H\},\\
&C(H(t)) = C(H(t), r, \alpha, \beta) \equiv \{y \in C[\alpha, \beta] : \|y(t) - \psi(t; r)\| \le H(t), \forall t \in [\alpha, \beta]\},\\
&C(\vec{H}) = C(\vec{H}(t), r, \alpha, \beta) \equiv \{y \in C[\alpha, \beta] : \|y_i(t) - \psi_i(t; r)\| \le H_i, \ i = 1, ..., n; \ t \in [\alpha, \beta]\}.
\end{aligned} \tag{22}$$

Note that $C(H), C(H(t))$, and $C(\vec{H})$ are closed, convex subsets of $C[\alpha, \beta]$. Note also that if $y \in C(H)$, then $(t, y(t)) \in D(H)$ and $y(t) \in C(H, t)$ for $t \in [\alpha, \beta]$. If **H2** is satisfied and $y \in C[\alpha, \beta]$, then $g(t, y(t))$ is continuous on $[\alpha, \beta]$. By Lemma 11, the problem

$$\mathbf{x}' = F(t)\mathbf{x} + g(t, y(t)), \ \ell\mathbf{x} = r, \tag{23}$$

has a unique solution. We denote this solution by $u(r, y) = u(t; r, y)$. Note that

$$u(r, y) = \psi(r) + u(0, y). \tag{24}$$

We now define a mapping $L$ on $C[\alpha, \beta]$ by

$$L(y) \equiv u(r, y). \tag{25}$$

Note that if $Lx = x$, then $x$ is a solution to the problem (2), (6).

**Lemma 2.7** *If* **H1** *to* **H4** *are satisfied, then* $L(C(H))$ *is relatively compact in* $C[\alpha,\beta]$.

**Proof:** By Lemma 12 and equation (24), for $y \in C(H)$, we have

$$\begin{aligned} \|Ly\| &= \|u(r,y)\| \le \|\psi(r)\| + \|u(0,y)\| \le \|(\ell|\mathcal{F})^{-1}\|\|r\| + K_1 \max_{[\alpha,\beta]} \|g(t,y(t))\| \\ &\le \|(\ell|\mathcal{F})^{-1}\|\|r\| + K_1 \max_{D(H)} \|g(t,z)\| \equiv B_1, \end{aligned}$$

that is, $L(C(H))$ is bounded by $B_1$. By Ascoli's theorem, it is sufficient to show that $L(C(H))$ is equicontinuous. For $y \in C(H), Ly = u(r,y)$ and

$$\|u'(t;r,y)\| = \|F(t)u(t;r,y) + g(t,y(t))\| \le \|F\|B_1 + \max_{D(H)} \|g(t,z)\| \equiv B_2. \tag{26}$$

By the mean value theorem, for $t_1, t_2 \in [\alpha,\beta]$,

$$\|u(t_2;r,y) - u(t_1;r,y)\| \le B_2|t_2 - t_1|. \tag{27}$$

Thus $L(C(H))$ is equicontinuous.

**Lemma 2.8** *If* **H1** *to* **H4** *are satisfied, then* $L$ *is continuous on* $C(H)$.

**Proof:** Let $\epsilon > 0$ be given. Since $D(H)$ is compact, $g$ is uniformly continuous on $D(H)$. There exists $\delta > 0$ such that if $(t_1,x_1)$ and $(t_2,x_2)$ are in $D(H)$ and $|t_1 - t_2| + \|x_1 - x_2\| < \delta$, then $\|g(t_1,x_1) - g(t_2,x_2)\| \le \epsilon/K_1$, where $K_1$ is defined in Lemma 12. If $y_1, y_2 \in C(H)$, then $Ly_1 - Ly_2$ is the solution to

$$\mathbf{x}' = F(t)\mathbf{x} + g(t,y_1(t)) - g(t,y_2(t)),\ \ell\mathbf{x} = 0. \tag{28}$$

By Lemma 12, we have

$$\|Ly_1 - Ly_2\| \le K_1 \max_{[\alpha,\beta]} \|g(t,y_1(t)) - g(t,y_2(t))\|. \tag{29}$$

If $\|y_1 - y_2\| < \delta$, that is, $\|y_1(t) - y_2(t)\| < \delta$ for $t \in [\alpha,\beta]$, then

$$\max_{[\alpha,\beta]} \|g(t,y_1(t)) - g(t,y_2(t))\| < \frac{\epsilon}{K_1},\ \ \|Ly_1 - Ly_2\| < K_1\frac{\epsilon}{K_1} = \epsilon. \tag{30}$$

Hence, $L$ is continuous on $C(H)$.

## 3 Existence Results

With the aid of the preceding lemmas we can prove our main results.

**Theorem 3.1** *Suppose* **H1** *to* **H4** *are satisfied. If there exists* $H > 0$ *such that*

$$M(H) = M(H,r,\alpha,\beta) \equiv \sup_{y \in C(H)} \|u(0,y)\| \le H, \tag{31}$$

*then problem (2), (6) has a solution* $\mathbf{x(t)} \in C(H)$.

**Proof:** From (24) we have $Ly - \psi(r) = u(r, y) - \psi(r) = u(0, y)$. Thus, for $y \in C(H)$, (31) yields

$$\|Ly - \psi(r)\| = \|u(0, y)\| \leq H, \tag{32}$$

that is, $Ly \in C(H)$, and $L(C(H)) \subset C(H)$. Since $C(H)$ is closed and convex, we can conclude from Lemmas 2.7 and 2.8 and the Schauder theorem in the form of Lemma 2.3 that $L$ has a fixed point $x \in C(H)$; that is, problem (2),(6) has a solution $\mathbf{x} \in C(H)$. □

We may easily obtain some natural generalization of Theorem 31.

**Theorem 3.2** *Suppose* **H1** *to* **H4** *are satisfied. If there exists a positive, continuous function $H(t)$ on $[\alpha, \beta]$ such that*

$$M(t, H(t)) \equiv \sup_{y \in C(H(t))} \|u(t; 0, y)\| \leq H(t), \tag{33}$$

*for $t \in [\alpha, \beta]$, then the problem (2), (6) has a solution $\mathbf{x} \in C(H(t))$.*

**Proof:** We have $Ly(t) - \psi(t; r) = u(t; r, y) - \psi(t; r) = u(t; 0, y)$. The condition (33) implies that, for $y \in C(H(t))$, (31) yields

$$\|Ly(t) - \psi(t; r)\| = \|u(t; 0, y)\| \leq H(t), \tag{34}$$

for $t \in [\alpha, \beta]$. Thus $Ly \in C(H(t))$, that is, $L(C(H(T))) \subset C(H(t))$. If $H \equiv \max_{[\alpha,\beta]} H(t)$, then $C(H(t)) \subset C(H)$. Moreover, $L(C(H(t))) \subset L(C(H))$. By Lemma 2.7, $L(C(H))$ is a relatively compact subset of $C[\alpha, \beta]$; hence, $L(C(H(t)))$ is relatively compact. By Lemma 2.8, $L$ is continuous on $C(H)$; hence $L$ is continuous on $C(H(t))$. Since $C(H(t))$ is a closed, convex subset of $C[\alpha, \beta]$, we may conclude from Lemma 2.3 that $L$ has a fixed point $\mathbf{x}$ in $C(H(t))$, that is, problem (2),(6) has a solution $\mathbf{x} \in C(H(t))$.

In what follows, we use the Sinc methodology to find a numerical solution for equation (5).

## 4 Description of the Sinc Approximation

The goal of this section is to recall notations and definitions of the Sinc function that will be used in this paper. These are discussed in [3,15]. The Sinc function is defined on the whole real line $R$ by

$$sinc(x) = \frac{\sin(\pi x)}{\pi x}, \quad x \in R. \tag{35}$$

Recall that a radial basis function is a function whose value depends only on the distance of its input to a central point. For a series of nodes equally spaced $h$ apart, the Sinc function can be written as a radial basis function:

$$S(j, h)(kh) = \delta_{jk}^{(0)} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases} \tag{36}$$

Let

$$\delta_{kj}^{(-1)} = \frac{1}{2} + \delta_{kj} = \int_0^{k-j} \frac{\sin(\pi t)}{\pi t} dt.$$

We define a matrix $I^{(-1)}$ whose $(k, j)th$ entry is given by $\delta_{kj}^{(-1)}$. If a function $f(x)$ is defined on the real line, then for $h > 0$, the series

$$C(f, h)(x) = \sum_{j=-\infty}^{\infty} f(jh) sinc\Big(\frac{x - jh}{h}\Big), \quad j = 0, \mp 1, ...$$

is called the Whittaker cardinal expansion, which has been extensively studied in [13,15]. In practice, we need to use a finite number of terms in the above series, say $j = -N, ..., N$, where $N$ is the number of Sinc grid points. For a restricted class of functions known as the Paly-Weiner class, which are entire functions, the Sinc interpolation and quadrature formulae are exact [15]. A less restricted class of functions that are analytic only on an infinite strip containing the real line, and that allow specific growth restriction has exponentially decaying absolute errors in the Sinc approximation.

**Definition 4.1** Let $\mathcal{D}_d$ denote the infinite strip domain of width $2d$, $d > 0$, given by

$$\mathcal{D}_d = \{w = u + iv : |v| < d \leq \pi/2\}.$$

To construct an approximation on the interval $\Gamma = (a, T)$, which is our space interval in this paper, we consider the conformal map $\phi(x) = \ln(\frac{x-a}{T-x})$, the map $\phi$ carries the eye-shaped region

$$\mathcal{D} = \Big\{z = x + iy : \Big|\arg\Big(\frac{z-a}{T-z}\Big)\Big| < d \leq \pi/2\Big\}$$

onto the infinite strip $\mathcal{D}_d$. For the Sinc method, the basis functions on the interval $\Gamma$ at $z \in \mathcal{D}$ are derived from the composite translated Sinc functions

$$S_j(z) = S(j, h) \circ \phi(z) = sinc\Big(\frac{\phi(z) - jh}{h}\Big).$$

The function $z = \phi^{-1}(w) = \frac{a+T\exp(w)}{1+\exp(w)}$ is an inverse mapping of the $w = \phi$. We define the range of $\phi^{-1}$ on the real line as

$$\Gamma = \{\phi^{-1}(y) \in \mathcal{D} : -\infty < y < \infty\} = (0, T).$$

The Sinc grid points $z_k \in \Gamma$ in $\mathcal{D}$ will be denoted by $x_k$, because they are real, and are given by

$$x_k = \phi^{-1}(kh) = \frac{a + T\exp(kh)}{1 + \exp(kh)}, \quad k = 0, \mp 1, \mp 2, ...$$

To further explain the Sinc method, an important class of functions is denoted by $\mathbf{L}_\alpha(\mathcal{D})$. The properties of the functions in $\mathbf{L}_\alpha(\mathcal{D})$ and detailed discussion are given in [15]. We recall the following definition followed by two theorems for our purpose.

**Definition 4.2** Let $\mathbf{L}_\alpha(\mathcal{D})$ be the class of all analytic functions $f$ in $\mathcal{D}$, for which there is a number $C_0$ such that, for $\rho(z) = \exp(\phi(z))$, we have

$$|f(z)| \leq C_0 \frac{|\rho(z)|^\alpha}{[1 + |\rho(z)|]^{2\alpha}}, \ \forall z \in \mathcal{D}.$$

The class $\mathbf{L}_\alpha(\mathcal{D})$ is important in Sinc methodology since it guarantees the rapid convergence of Sinc approximations. In the next theorem, we shall give a general formula for approximating the integral $\int_a^\nu F(u)du, \nu \in \Gamma$. To this end, we state the following result, which we will use to approximate the obtained integral equation.

**Theorem 4.1** *Let $\frac{F(t)}{\phi'(t)} \in \mathbf{L}_\alpha(\mathcal{D})$, with $0 < \alpha \leq 1$, $\delta_{jk}^{(-1)}$ be defined as above, $N$ be a positive integer, and $h$ be selected as $h = \sqrt{\pi d/(\alpha N)}$, then there exists a positive constant $K$ independent of $N$, such that*

$$\Big|\int_a^{t_k} F(t)dt - h\sum_{j=-N}^{N} \delta_{jk}^{(-1)} \frac{F(t_k)}{\phi'(t_k)}\Big| \leq K\exp(-\sqrt{\pi d\alpha N}).$$

It is convenient, for deriving an approximate solution for the system (5) by the Sinc-Galerkin method, to start with the scalar first order differential equation

$$\frac{du}{dt} = B(t)u(t) + f(t),\ t \in (a, T) \tag{37}$$

subject to the initial condition

$$u(a) = u^0. \tag{38}$$

Integrating with respect to $t$, and using the initial condition we arrive at the integral equation

$$u(t) = \int_a^t \Big[B(\tau)u(\tau) + f(\tau)\Big] d\tau + u^0. \tag{39}$$

To approximate over $\Gamma = (a, T)$, we make use of the conformal mapping $\phi(x)$ mentioned above. We also assume that both $B/\phi'$ and $f$ belong to the class of functions $\mathbf{L}_\alpha(\mathcal{D})$. If $B$ is a matrix, this shall imply that all the components of $B/\phi'$ (or $f$ as a vector) are in the class $\mathbf{L}_\alpha(\mathcal{D})$. Now in equation (39) we collocate via the use of the indefinite integration formula (as in Theorem 4.1). We use the notation $\mathcal{D}(1/\phi'(t_i)) = diag[1/\phi'(t_{-N}), ..., 1/\phi'(t_N)]$, then equation (39) can be written as a system of $m = 2N + 1$ linear equations

$$U = hI_m^{(-1)}\mathcal{D}(B/\phi'(t_i))U + hI_m^{(-1)}\mathcal{D}(1/\phi'(t_i))F + U^0, \tag{40}$$

where $U = [u_{-N}, ..., u_N]^t$, $F = [f_{-N}, ..., f_N]^t$ with the nodes $t_i = \phi^{-1}(ih)$ for $i = -N, ..., N$ where $h = \sqrt{\pi d/\alpha N}$, and $U^0$ denotes the vector of $2N + 1$ constant values $U^0 = [u^0_{z_{-N}}, ..., u^0_{z_N}]^t$. Define the matrices $A$ and $E$ by $A = hI_m^{(-1)}\mathcal{D}(B/\phi'(t_i))$, $E = hI_m^{(-1)}\mathcal{D}(1/\phi'(t_i))$. Then equation (40) can be written as

$$U = AU + EF + U^0. \tag{41}$$

To prove convergence of the method, we evaluate the integral in (39) at the nodes $t_i$, where $i = -N, ..., N$, to get

$$u(t_i) = \int_a^{t_i} \Big[B(\tau)u(\tau) + f(\tau)\Big] d\tau + u^0$$

with the same matrices $A, E$ and $U^0$ as mentioned above, and using the approximation in Theorem 4.1 we get, in matrix form, the approximation

$$U + AU + EF + U^0 + \tilde{K}\exp(-\sqrt{\pi d\alpha N}),$$

where the constant $\tilde{K}$ is a vector such that each entry is bounded by the constant $K$ in Theorem 4.1. So, the error ERR can be bounded as

$$\|ERR\| \le \|U - (AU + EF + U^0)\| \le \tilde{K}\exp(-\sqrt{\pi d\alpha N}),$$

i.e., the discretization error that arises when a differential equation is replaced by a discrete system of algebraic equations is exponentially small. With the notation as above, we just proved the following theorem.

**Theorem 4.2** *Let $B/\phi', f \in \mathbf{L}_\alpha(\mathcal{D})$, let the function $u(t)$ be defined as in (39), and let the matrix $U$ be defined as in (41). Then for $h = \sqrt{\pi d/(\alpha N)}$ there exists a constant $K$ independent of $N$ such that*

$$\sup \|[u(t_i)] - U\| \leq K \exp(-\sqrt{\pi d \alpha N}).$$

Now, we may attempt to solve the linear systems of equations (41) by successive approximations, that is, by means of the iterative scheme:

$$U^{(n+1)} = AU^{(n)} + EF + U^0. \tag{42}$$

It is easy to show that the convergence of the scheme depends on the $\ell^\infty$ norm of $B$, as noted in the following theorem.

**Theorem 4.3** *The sequence $U^{(n)}$ defined in (42) converges, for all $N$ being sufficiently large, to the exact solution provided that $(T-a) < 11/(10\|B\|_\infty)$.*

**Proof:** Recall that by definition of $\delta^{(-1)}$ as defined in Section 2, it satisfies the inequality [15, p. 172] $\delta^{(-1)} \leq 11/10$, we have

$$\begin{aligned}
\|B\|_\infty &= \|hI_m^{(-1)}\mathcal{D}(1/\phi'(t_i))\| = |\max_i \sum_{j=-N}^{N} h\delta_{i-j}^{(-1)}(B(z_j)/\phi'(z_j))| \\
&\leq \frac{11}{10} h \sum_{j=-N}^{N} (B(z_j)/\phi'(z_j)) \approx \frac{11}{10}\int_a^T |B(t)|dt \leq \frac{11}{10}(T-a) \sup_{t\in(a,T)} |B(t)| \\
&\leq \frac{11}{10}(T-a)\|B\|_\infty,
\end{aligned}$$

where in the third inequality we used Theorem 4.1, with the fact that $B/\phi' \in \mathbf{L}_\alpha(\mathcal{D})$. For the iteration scheme to converge we require that $\|B\|_\infty < 1$. Therefore we can achieve convergence of the scheme (42) by choosing $(T-a) < \frac{11}{10\|B\|_\infty}$.

It remains to show that the approximate solution $U^*$ of node values of equation (41) converges to the node values of the exact solution $U$ (see, [4]). For that end, choose a constant $R$ so that $U$ and $U*$ belong to the ball $\mathcal{B} = \{X : \|X\|_\infty < \frac{11}{10}\|B\|_\infty < R/2\}$. It is enough to show that $|U - U^*|$ is small. If $U^*$ is the approximate solution and satisfies equation (41), then $(U - U^*) - A(U - U^*) = Error$, or

$$\|(U - U^*)\| \leq \|A(U - U^*)\| + \|Error\|. \tag{43}$$

Now we can find a small constant $r$, that is $0 < r < 1$ such that the Jacobian of the matrix $A$ is less than $r$, so by the mean-value theorem, we obtain $\|U - AU\| \leq r\|U - U^*\|$, so equation (43) reduces to

$$\|(U - U^*)\| \leq \frac{1}{1-r}\|Error\|. \tag{44}$$

This shows that the approximate solution is sufficiently close to the exact solution. With the above notations, we have proved the following theorem.

**Theorem 4.4** *For a constant $R > 0$, with $\|B\|_\infty < 1$, the solution in equation (5) with the iteration scheme (41) converges to the unique solution.*

| $t$ | Errors $x(t)$ | Errors $y(t)$ |
|---|---|---|
| 0.1 | 1.048 E -07 | 2.098 E -08 |
| 0.3 | 3.549 E -07 | 1.560 E -08 |
| 0.6 | 2.963 E -06 | 3.905 E -08 |
| 0.9 | 4.998 E -06 | 4.848 E -08 |
| 1.2 | 1.009 E -05 | 3.555 E -08 |
| 1.5 | 2.286 E -05 | 9.098 E -08 |
| 0.8 | 5.201 E -05 | 8.948 E -08 |
| 1.8 | 7.579 E -05 | 8.011 E -08 |

**Table 1**: Numerical results for the example given in 45. Comparison between the Sinc solution and the exact solution.

## 5 Test Example

Consider the initial value problem in equation (2) of the form

$$\mathbf{X}'(\mathbf{t}) = \begin{pmatrix} 2 & -3 \\ 1 & -2 \end{pmatrix} \mathbf{X}(\mathbf{t}) + \begin{pmatrix} e^{2t} \\ 1 \end{pmatrix}, \quad \mathbf{X}(\mathbf{0}) = \begin{pmatrix} -1 \\ 0, \end{pmatrix}, \tag{45}$$

that is, we consider equation (2) with $n = 2$, $F(t) = \begin{pmatrix} 2 & -3 \\ 1 & -2 \end{pmatrix}$, $g(t, X) = \begin{pmatrix} e^{2t} \\ 1 \end{pmatrix}$ and $r = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. Since $F(t)$ and $g(t, \mathbf{X})$ are continuous, **H1** and **H2** are satisfied. Assumption **H3** can be established by Lemma (8). While assumption **H4** is an immediate consequence of Lemma (34). To prove the existence of the solution for (45), with the fact that in our case $u(t; r; y)$ is given by $u(t; r, y) = re^t + \int_0^t e^{t-s} \begin{pmatrix} e^{2s} \\ 1 \end{pmatrix} ds$, it is easy to manipulate the steps of Example 4.2 in [2]. To show the efficiency of the Sinc method in comparison with the exact solution of the given equation, which is known to be

$$\mathbf{X}(\mathbf{t}) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} -\frac{9}{2}e^t - \frac{5}{6}e^{-t} + \frac{4}{3}e^{2t} - 3 \\ \\ -\frac{3}{2}e^t - \frac{5}{6}e^{-t} + \frac{1}{3}e^{2t} + 2 \end{pmatrix},$$

we use the Sinc method to solve the problem in (45) with the parameters $d = \frac{1}{2}, \alpha = 1$ and $N = 32$. In Table 1, the comparison of the numerical results demonstrates the accuracy of this approach.

## Conclusions

The study of systems of ODEs is still a very active area of research due to its application in modeling various physical, chemical, biological, engineering and social systems. This paper mainly focused on the application of the Schauder fixed point theorem to study the existence of solutions for systems of ODE. On the other hand, a numerical scheme using Sinc functions is developed to approximate the solution of a $2 \times 2$ system of first order differential equations. The numerical results demonstrate the reliability and efficiency of using the Sinc method to solve such problems. The error in the numerical solution is shown to converge exponentially.

## References

[1] Ahmed, B. and Alsadi, A. Existence and Uniqueness of Solutions for Coupled Systems of Higher-Order Nonlinear Fractional Differential Equations. *Fixed Point Theory and Applications*, Vol. 2010 (2010) ID: 364560.

[2] Al-Khaled, K. On the existence of solutions for a class of first order differential equations. *Intern. J. Math. Math. Sci.* **25** (1) (2001) 1–10.

[3] Al-Khaled, K. Sinc numerical solution for solitons and solitary waves. *Journal of Computational and Applied Mathematics* **130** (1-2) (2001) 283–292.

[4] Al-Khaled, K. Theory and computations for systems modeled by first order differential equations. *Qatar Univ. Sci. J.* **20** (2000) 21–24.

[5] Berinde, V. Existence and Approximation of Solutions. *Miskole Mathematical Notes* **11** (1) (2010) 13–26.

[6] Dunford, N. and Schwartz, J.I. *Linear Operators*. Interscience Publishers, Inc., New York, 1958.

[7] Farajzadeh, A.P., Kaewcharoen, A. and Plubtieng, S. An application of fixed point theory to a nonlinear differential equation. *Abstract and Applied Analysis*, Vol. 2014 (2014) ID: 605405.

[8] Francis, J. and Miller, K.S. *Existence Theorems for Ordinary Differential Equations*. Dover, 2007.

[9] Gavrilyuk, I.P. and Makarov, V.L. Representation and approximation of the solution of an initial value problem for a first order differential equation in Banach spaces. *Z. Anal. Anwendungen* **15** (2) (1996) 495–527.

[10] Hartman, P. *Ordinary Differential Equations*. John Wiley & Sons, Inc., New York, London, Sydney, 1964.

[11] Kichmarenko, O. and Stanzhytskyi, O. Sufficient Conditions for the Existence of Optimal Controls for Some Classes of Functional-Differential Equations. *Nonlinear Dynamics and Systems Theory* **18** (1) (2018) 196–211.

[12] Persson, J. A generalization of Caratheodory's existence theorem for ordinary differential equations. *J. of Math. Analy. Appli.* **49** (1975) 496–503.

[13] Ralph, B.K. A sinc approximation for the indefinite integral, *Math. of Comput.* **41** (164) (1983) 559–572.

[14] Sestelo, R.F. and Pouso, R.L. Existence of solutions of first-order differential equations via a fixed point theorem for discontinuous operators. *Fixed Point Theory and Applications*, 2015. DOI 10.1186/s13663-015-0472-5.

[15] Stenger, F. *Numerical Methods Based on Sinc and Analytic Functions*. Springer-Verlag, 1993.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 331–341

# Solvability Criterion for Integro-Differential Equations with Degenerate Kernel in Banach Spaces

A.A. Boichuk [1]* and V.F. Zhuravlev [2]

[1] *Institute of Mathematics of the National Academy of Sciences of Ukraine, 3, Tereshchenkivska Str., Kyiv, 01601, Ukraine*
[2] *Zhytomyr National Agricultural-Ecological University Ukraine, 7, Staryi Boulevard, Zhytomyr, 10008, Ukraine*

Received: June 26, 2018; Revised: September 27, 2018

**Abstract:** By means of the theory of generalized inversion of operators in Banach spaces, a solvability criterion and a general form of solutions for integro-differential equations with a degenerate kernel in Banach spaces have been established. The obtained results have been illustrated by examples.

**Keywords:** *integro-differential equation; degenerate kernel; Banach space; generalized invertible operator; general form of solutions.*

**Mathematics Subject Classification (2010):** 34K30, 45J05, 47G10, 47G20.

## 1 Introduction

The investigation of the solvability of integro-differential equations is a problem the specific nature of which lies in the fact that the integro-differential operator has no inverse. Such equations in Euclidean spaces were considered in [1–4] and others.

Sufficient conditions for the existence and uniqueness of piecewise-continuous mild solutions of fractional integro-differential equations in a Banach space with non instantaneous impulses were obtained in [5]. In paper [6] V. Gupta and J. Dabas established the existence and uniqueness of solution for a class of impulsive fractional integro-differential equations with nonlocal boundary conditions.

In this paper, we propose a somewhat different approach to the study of integro-differential equations in Banach spaces. In its realization, the theory of generalized inversion of operators in Banach spaces is effectively used [7, 8].

The proposed approach can be used in the study of the phenomena of energy transfer and diffusion of neutrons, viscoelastic oscillations various systems and structures, in nuclear physics and the mathematical theory of biological populations (see [9–11]).

* Corresponding author: mailto:boichuk.aa@gmail.com

## 2 Formulation of the Problem

Consider the integro-differential equation

$$z(t) - M(t) \int_a^b \big[W(s)z(s) + V(s)\dot{z}(s)\big] ds = f(t), \tag{1}$$

where the operator-valued function $M(t)$ acts from the Banach space $\mathbf{B}_2$ into the Banach space $\mathbf{B}_1$ and is strongly continuous with the norm $|||M||| = \sup\limits_{t\in\mathcal{I}} \|M(t)\|_{\mathbf{B}_2} = M_0 < \infty$, and the operator-valued functions $W(t)$ and $V(t)$ act from the Banach space $\mathbf{B}_1$ into the Banach space $\mathbf{B}_2$ and are strongly continuous with the norms $|||W||| = \sup\limits_{t\in\mathcal{I}} \|W(t)\|_{\mathbf{B}_1} = W_0 < \infty$ and $|||V||| = \sup\limits_{t\in\mathcal{I}} \|V(t)\|_{\mathbf{B}_1} = V_0 < \infty$, the vector- function $f(t)$ acts from the interval $\mathcal{I}$ into the Banach space $\mathbf{B}_1 : f(t) \in \mathbf{C}(\mathcal{I}, \mathbf{B}_1) := \big\{f(\cdot) : \mathcal{I} \to \mathbf{B}_1,\ |||f||| = \sup_{t\in\mathcal{I}} \|f(t)\|\big\}$, $\mathbf{C}(\mathcal{I}, \mathbf{B}_1)$ is the Banach space of vector-functions continuous on $\mathcal{I}$ with values in $\mathbf{B}_1$.

By the solution $z(t)$ of the operator equation (1) we mean vector-functions such that $z(t) \in \mathbf{C}(\mathcal{I}, \mathbf{B}_1)$, $\dot{z}(t) \in \mathbf{C}^1(\mathcal{I}, \mathbf{B}_1)$, where $\mathbf{C}^1(\mathcal{I}, \mathbf{B}_1)$ is the Banach space of continuously differentiable vector-functions with the norm $|||z||| = \sum_{k=0}^1 \sup_{t\in\mathcal{I}} \|z^{(k)}(t)\|\big\}$, where $z^{(k)}(t)$ is the $k$-th derivative $z(t)$. The derivative $\dot{z}(t)$ is understood in the sense of [12, p. 140].

The problem is to obtain a solvability criterion and to find the structure of solutions for the integro-differential equation (1).

## 3 Preliminary Information

Consider the linear integral Fredholm equation with a degenerate kernel

$$z(t) - M(t) \int_a^b N(s)z(s)ds = f(t), \tag{2}$$

where the operator-valued function $N(t)$ acts from the Banach space $\mathbf{B}_1$ into the Banach space $\mathbf{B}_2$ and is strongly continuous with the norm $|||N||| = \sup\limits_{t\in\mathcal{I}} \|N(t)\|_{\mathbf{B}_1} = N_0 < \infty$.

Denote: $D = I_{\mathbf{B}_2} - A$, $\quad A = \int\limits_a^b N(s)M(s)\,ds$, $\quad D : \mathbf{B}_2 \to \mathbf{B}_2$. In [8] it is shown that if $D$ is a bounded generalized invertible operator, then the integral operator $L$ is generalized invertible.

In this case, there exist bounded projections $\mathcal{P}_{N(D)}$, $\mathcal{P}_{Y_D}$ onto the null space $N(D)$ and the subspace $Y_D = I_{\mathbf{B}_2} \ominus R(D)$ of the operator $D$, respectively [13] and the bounded generalized inverse operator $D^-$ to the operator $D$ [7].

The following theorem holds for the integral equation (2).

**Theorem 3.1** *[14] Let $D : \mathbf{B}_2 \to \mathbf{B}_2$. Then the homogeneous (f(t)=0) integral equation (2) has a family of solutions*

$$z(t) = M(t)\mathcal{P}_{N(D)}c,$$

*where $c$ is an arbitrary element of the Banach space $\mathbf{B}_2$.*

*Under and only under the condition*

$$\mathcal{P}_{Y_D} \int_a^b N(s)f(s)ds = 0$$

*the nonhomogeneous integral equation (2) has a family of solutions*

$$z(t) = M(t)\mathcal{P}_{N(D)}c + f(t) + M(t)D^- \int_a^b N(s)f(s)ds.$$

## 4 The Main Result

**1.** We obtain the solvability conditions for the general form of solutions of the equation (1).

We make the substitution $\dot{z}(t) = y(t)$ in (1), then

$$z(t) = \int_a^t y(s)ds + c_0, \quad c_0 \in \mathbf{B}_1. \tag{3}$$

Putting (3) in (1), we obtain the integral equation

$$y(t) - M(t) \int_a^b \Big[W(s) \int_a^s y(\tau)d\tau + V(s)y(s)\Big] ds = f(t) + M(t)W_0c_0, \tag{4}$$

where $W_0 = \int_a^b W(s)ds$, $W_0 : \mathbf{B}_1 \to \mathbf{B}_2$.

Changing the order of integration in the integral $\int_a^b W(s) \int_a^s y(\tau)d\tau ds$, we obtain from (4)

$$y(t) - M(t) \int_a^b N(s)y(s)ds = g(t), \tag{5}$$

where

$$N(s) = \int_s^b W(\tau)d\tau + V(s),$$

$$g(t) = f(t) + M(t)W_0c_0. \tag{6}$$

By Theorem 3.1, under and only under the condition

$$\mathcal{P}_{Y_D} \int_a^b N(s)g(s)ds = 0 \tag{7}$$

the integral equation (5) has a family of solutions

$$y(t) = M(t)\mathcal{P}_{N(D)}c + g(t) + M(t)D^{-}\int_a^b N(s)g(s)ds, \tag{8}$$

where $c$ is an arbitrary element of the Banach space $\mathbf{B}_2$.

From the solvability condition (7) we find the value of $c_0 \in \mathbf{B}_1$, for which the integral equation (5) has a solution. We put (6) in (7)

$$\mathcal{P}_{Y_D}\int_a^b N(s)\Big[f(s) + M(s)W_0c_0\Big]ds = 0.$$

After the transformations, we obtain the operator equation

$$Sc_0 = b_0, \tag{9}$$

where

$$b_0 = -\mathcal{P}_{Y_D}\int_a^b N(s)f(s)ds.$$

$$S = \mathcal{P}_{Y_D}\int_a^b N(s)M(s)W_0ds = \mathcal{P}_{Y_D}AW_0 = \mathcal{P}_{Y_D}\big[I - D\big]W_0 = \mathcal{P}_{Y_D}W_0,$$

because $\mathcal{P}_{Y_D}D = 0$.

Let the operator $S : \mathbf{B}_1 \to \mathbf{B}_2$ be generalized invertible. Then there exist bounded projectors $\mathcal{P}_{N(S)} : \mathbf{B}_1 \to \mathbf{B}_1$ and $\mathcal{P}_{Y_S} : \mathbf{B}_2 \to \mathbf{B}_2$ and a bounded generalized inverse operator $S^{-} : \mathbf{B}_2 \to \mathbf{B}_1$ to the operator $S$. The operator equation (9) is solvable under and only under the condition [7]

$$\mathcal{P}_{Y_S}b_0 = \mathcal{P}_{Y_S}\mathcal{P}_{Y_D}\int_a^b N(s)f(s)ds = 0, \tag{10}$$

and, under this condition, the equation (9) has a family of solutions

$$c_0 = \mathcal{P}_{N(S)}\tilde{c} + S^{-}b_0,$$

where $\tilde{c}$ is an arbitrary element of the Banach space $\mathbf{B}_1$.

Then $g(t)$ takes the form

$$g(t) = f(t) + M(t)W_0\Big[\mathcal{P}_{N(S)}\tilde{c} + S^{-}b_0\Big].$$

We put $g(s)$ in the solution (8) of the integral equation (5)

$$\begin{aligned} y(t) = M(t)\mathcal{P}_{N(D)}c + f(t) + M(t)W_0\Big[\mathcal{P}_{N(S)}\tilde{c} + S^{-}b_0\Big] + \\ + M(t)D^{-}\int_a^b N(s)\Big\{f(s) + M(s)W_0\Big[\mathcal{P}_{N(S)}\tilde{c} + S^{-}b_0\Big]\Big\}ds. \end{aligned} \tag{11}$$

Denoting $\widetilde{D} = (I_{\mathbf{B}_1} + D^{-}A)W_0$, after the transformations we obtain the general solution of the equation (5)

$$y(t) = M(t)\Big[\mathcal{P}_{N(D)},\ \ \widetilde{D}\mathcal{P}_{N(S)}\Big]\left[\begin{array}{c} c \\ \tilde{c} \end{array}\right] + f(t) +$$

$$+M(t)D^{-}\int\limits_a^b N(s)f(s)ds - M(t)\widetilde{D}S^{-}\mathcal{P}_{Y_D}\int\limits_a^b N(s)f(s)ds =$$

$$= M(t)\Big[\mathcal{P}_{N(D)},\ \ \widetilde{D}\mathcal{P}_{N(S)}\Big]\left[\begin{array}{c} c \\ \tilde{c} \end{array}\right] + f(t) +$$

$$+M(t)\Big[D^{-} - \widetilde{D}S^{-}\mathcal{P}_{Y_D}\Big]\int\limits_a^b N(s)f(s)ds,$$

where $c \in \mathbf{B}_2$, $\tilde{c} \in \mathbf{B}_1$ are arbitrary constants.

Putting the obtained $y(t)$ in (2), we obtain the general solution of the integro-differential equation (1)

$$z(t) = \Big[\widetilde{M}(t)\mathcal{P}_{N(D)},\ \ \widetilde{M}(t)\big(\widetilde{D}\mathcal{P}_{N(S)} + \mathcal{P}_{N(S)}\big)\Big]\left[\begin{array}{c} c \\ \tilde{c} \end{array}\right] + \tilde{f}(t) + F(t),$$

where

$$\begin{gathered} \widetilde{M}(t) = \int\limits_a^t M(s)ds, \ \ \tilde{f}(t) = \int\limits_a^t f(s)ds, \\ F(t) = \Big\{\widetilde{M}(t)\Big[D^{-} - \widetilde{D}S^{-}\mathcal{P}_{Y_D}\Big] - S^{-}\mathcal{P}_{Y_D}\Big\}\int\limits_a^b N(s)f(s)ds. \end{gathered} \tag{12}$$

Thus, the following theorem holds for the integro-differential equation (1).

**Theorem 4.1** *Let the operators $D : \mathbf{B}_2 \to \mathbf{B}_2$ and $S : \mathbf{B}_1 \to \mathbf{B}_2$ be generalized invertible. Then the integro-differential equation (1) is solvable for those and only those $f(t) \in \mathbf{C}([a,b], \mathbf{B}_1)$, that satisfy the condition*

$$\mathcal{P}_{Y_S}\mathcal{P}_{Y_D}\int\limits_a^b N(s)f(s)ds = 0$$

*and has a family of solutions*

$$z(t) = \Big[\widetilde{M}(t)\mathcal{P}_{N(D)},\ \ \big(\widetilde{M}(t)\widetilde{D}\mathcal{P}_{N(S)} + \mathcal{P}_{N(S)}\big)\Big]\left[\begin{array}{c} c \\ \tilde{c} \end{array}\right] + \tilde{f}(t) + F(t).$$

**Remark 4.1** As shown in [3] the integro-differential equation

$$(Lz)(t) := \dot{z}(t) + H(t)z(t) - M(t)\int\limits_a^b \big[W(s)z(s) + V(s)\dot{z}(s)\big]ds = f(t),$$

where the operator-valued function $H(t)$ acts from the Banach space $\mathbf{B}_1$ to the Banach space $\mathbf{B}_1$ and is strongly continuous with the norm $|||H||| = \sup\limits_{t\in\mathcal{I}}\|H(t)\|_{\mathbf{B}_1} = H_0 < \infty$, with the help of substitution $z(t) = X(t)y(t)$, where $X(t)$ is the fundamental operator [12, p. 148] $\dot{z}(t) = -H(t)z(t)$, is reduced to an equation of the form (1).

**Remark 4.2** The integro-differential equation

$$(Lz)(t) := \dot{z}(t) - \sum_{i=1}^{q} M_i(t) \int_a^b \big[W_i(s)z(s) + V_i(s)\dot{z}(s)\big]ds = f(t)$$

is reduced to an equation of the form (1), if we denote the operator matrices $M(t) = \big[M_1(t), M_2(t), \ldots, M_q(t)\big]$, $W(t) = \mathrm{col}\big[W_1(t), W_2(t), \ldots, W_q(t)\big]$, $V(t) = \mathrm{col}\big[V_1(t), V_2(t), \ldots, V_q(t)\big]$.

**2.** In the case when the integro-differential equation is considered in Euclidean spaces, the proposed method of investigation can be refined and concretized.

Consider the equation (1) under the assumption that $M(t)$ is an $(n \times m)$-dimensional matrix, $W(t)$ and $V(t)$ are $(m \times n)$-dimensional matrices, $f(t)$ is an $(n \times 1)$-dimensional matrix whose elements belong to the space $\mathbf{L}_2[a, b]$. The solution will be sought in the class of functions $z(t) \in \mathbf{D}_2^n[a, b]$, $\dot{z}(t) \in \mathbf{L}_2^n[a, b]$.

In this case, the operator $D = I_m - A$, $\quad A = \int_a^b N(s)M(s)\, ds$ and orthoprojectors $P_{N(D)}$, $P_{N(D^*)}$ [15, 16] are $(m \times m)$-dimensional matrices.

Let rank$D = n_1$. Denote an $(m \times r)$-dimensional matrix by $P_{N_r(D)}$, which is composed of $r = m - n_1$ linearly independent columns of the orthoprojector matrix $P_{N(D)}$, and an $(r \times m)$-dimensional matrix by $P_{N_r(D^*)}$, which is composed of $r$ linearly independent rows of the orthoprojector matrix $P_{N(D^*)}$.

Then by Theorem 3.1, under and only under $r$ linearly independent conditions

$$P_{N_r(D^*)} \int_a^b N(s)g(s)ds = 0 \tag{13}$$

the integral equation (5) has $r$ linearly independent solutions

$$y(t) = M(t)P_{N_r(D)}c_r + g(t) + M(t)D^+ \int_a^b N(s)g(s)ds, \tag{14}$$

where $c_r$ is an arbitrary element of the Euclidean space $\mathbf{R}^r$, $D^+$ is the Moor-Penrose pseudoinverse matrix to the matrix $D$ [15, 16].

From the condition (13) we obtain an algebraic system with respect to the vector $c_0 \in \mathbf{R}^n$

$$Sc_0 = b_0, \tag{15}$$

where $S = P_{N_r(D^*)}W_0$ is an $(r \times n)$-dimensional matrix, $b_0 = -P_{N_r(D^*)} \int_a^b N(s)f(s)ds$.

Let rank$S = n_2$. Denote an $(n \times k)$-dimensional matrix by $P_{N_k(S)}$, which is composed of $k = n - n_2$ linearly independent columns of the orthoprojector matrix $P_{N(S)}$, and an $(d \times r)$-dimensional matrix by $P_{N_d(S^*)}$, which is composed of $d = r - n_2$ linearly independent rows of the orthoprojector matrix $P_{N(S^*)}$.

The system (15) is solvable if and only if the vector $b_0$ satisfies the condition

$$P_{N_d(S^*)}b_0 = P_{N_d(S^*)}P_{N_r(D^*)} \int_a^b N(s)f(s)ds = 0, \tag{16}$$

under which the equation (15) has a family of solutions

$$c_0 = P_{N_k(S)}\tilde{c}_k + S^+ b_0,$$

where $\tilde{c}_k$ is an arbitrary element of the Euclidean space $\mathbf{R}^k$, $S^+$ is the Moor-Penrose pseudoinverse matrix to the matrix $S$ [15,16].

The condition (16) consists of $d$ linearly independent conditions. Indeed, since the matrices $P_{N_d(S^*)}$, $P_{N_r(D^*)}$ are of the full rank: $\mathrm{rank}P_{N_d(S^*)} = d$, $\mathrm{rank}P_{N_r(D^*)} = r$ and $d \leq r$, we have from the Sylvester inequality [17, p. 31] that

$$\mathrm{rank}P_{N_d(S^*)} + \mathrm{rank}P_{N_r(D^*)} - r \leq \mathrm{rank}\big(P_{N_d(S^*)}P_{N_r(D^*)}\big) \leq$$

$$\leq \min\big(\mathrm{rank}P_{N_d(S^*)}, \mathrm{rank}P_{N_r(D^*)}\big)$$

or

$$d + r - r \leq \mathrm{rank}\big(P_{N_d(S^*)}P_{N_r(D^*)}\big) \leq d.$$

It follows that $\mathrm{rank}\big(P_{N_d(S^*)}P_{N_r(D^*)}\big) = d$.

Then the following theorem holds for the integro-differential equation (1).

**Theorem 4.2** *Let* $\mathrm{rank}D = n_1$, *and* $\mathrm{rank}S = n_2$.

*Then the integro-differential equation (1) is solvable for those and only those* $f(t) \in \mathbf{R}^n$, *that satisfy* $d = r - n_2$ *linearly independent conditions*

$$P_{N_d(S^*)}P_{N_r(D^*)}\int_a^b N(s)f(s)ds = 0,$$

*and at the same time it has an* $(r+k)$*-parametric family of linearly independent solutions*

$$z(t) = \Big[\widetilde{M}(t)P_{N_r(D)},\ \ \big(\widetilde{M}(t)\widetilde{D}P_{N_k(S)} + P_{N_k(S)}\big)\Big]\left[\begin{array}{c} c_r \\ \tilde{c}_k \end{array}\right] + \tilde{f}(t) + F(t),$$

*where* $c_r \in \mathbf{R}_r$, $\tilde{c}_k \in \mathbf{R}_k$ *are arbitrary constants;* $\widetilde{D} = (I_m + D^+A)W_0$; $\widetilde{M}(t), \tilde{f}(t)$ *have the form (12);*

$$F(t) = \Big\{\widetilde{M}(t)\Big[D^+ - \widetilde{D}S^+P_{N_r(D^*)}\Big] - S^+P_{N_r(D^*)}\Big\}\int_a^b N(s)f(s)ds.$$

**Example 4.1** Consider the integro-differential equation

$$(Lz)(t) := \dot{z}(t) - M(t)\int_0^2 \left[W(s)z(s) + V(s)\dot{z}(s)\right] ds = f(t), \tag{17}$$

where

$$M(t) = \mathrm{diag}\left\{\left[\begin{array}{ccc} 0 & t-1 & 0 \\ 1 & 0 & 3t \end{array}\right], \left[\begin{array}{ccc} 0 & t-1 & 0 \\ 1 & 0 & 3t \end{array}\right], \ldots\right\},$$

$$W(s) = \mathrm{diag}\left\{\left[\begin{array}{cc} 0 & s-\frac{3}{2} \\ -\frac{3}{2} & 0 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & s-\frac{3}{2} \\ -\frac{3}{2} & 0 \\ 1 & 0 \end{array}\right], \ldots\right\},$$

$$V(s) = \operatorname{diag}\left\{ \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ s-1 & \frac{s-1}{2} \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ s-1 & \frac{s-1}{2} \end{bmatrix}, \dots \right\}.$$

Let the vector-function $f(t)$ act from the interval $[0,2]$ into the Banach space $\mathbf{c}$ of all convergent numerical sequences: $f(t) \in \mathbf{C}([0,2],\mathbf{c}) := \{f(\cdot) : [0,2] \to \mathbf{c}\}$, the operator-valued functions $M(t)$, $W(t)$ and $V(t)$ act from the Banach space $\mathbf{C}([0,2],\mathbf{c})$ to itself with the norms $|||M|||_{\mathbf{C}([0,2],\mathbf{c})} = \sup_{t\in[0,2]}\|M(t)\|_{\mathbf{c}}$, $|||W|||_{\mathbf{C}([0,2],\mathbf{c})} = \sup_{t\in[0,2]}\|W(t)\|_{\mathbf{c}}$, $|||V|||_{\mathbf{C}([0,2],\mathbf{c})} = \sup_{t\in[0,2]}\|V(t)\|_{\mathbf{c}}$.

It is obvious that the operator $L$ is a linear bounded operator acting from the Banach space of continuously differentiable functions $\mathbf{C}^1([0,2],\mathbf{c})$ on the interval $[0,2]$ into the Banach space of continuous functions $\mathbf{C}([0,2],\mathbf{c})$.

For this equation we have:

$$N(s) = \int_s^2 W(s)ds + V(s) =$$

$$= \operatorname{diag}\left\{ \begin{bmatrix} 1 & -1-\frac{s^2}{2}+\frac{3s}{2} \\ -4+\frac{3s}{2} & 0 \\ 1 & \frac{s-1}{2} \end{bmatrix}, \begin{bmatrix} 1 & -1-\frac{s^2}{2}+\frac{3s}{2} \\ -4+\frac{3s}{2} & 0 \\ 1 & \frac{s-1}{2} \end{bmatrix}, \dots \right\},$$

$$W_0 = \int_0^2 W(s)ds = \operatorname{diag}\left\{ \begin{bmatrix} 0 & -1 \\ -3 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -3 & 0 \\ 2 & 0 \end{bmatrix}, \dots \right\}.$$

Then

$$D = I - A = I - \int_0^2 N(s)M(s)ds = \operatorname{diag}\left\{ \begin{bmatrix} \frac{4}{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{4}{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \dots \right\},$$

$$P_{N(D)} = \mathcal{P}_{Y_D} = \operatorname{diag}\left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \dots \right\},$$

$$D^- = \operatorname{diag}\left\{ \begin{bmatrix} \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \dots \right\}.$$

To find the solvability condition, we compute the operator

$$S = \mathcal{P}_{Y_D} W_0 = \operatorname{diag}\left\{ \begin{bmatrix} 0 & 0 \\ -3 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ -3 & 0 \\ 2 & 0 \end{bmatrix}, \dots \right\},$$

$$\mathcal{P}_{N(S)} = \operatorname{diag}\left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \dots \right\}, \mathcal{P}_{Y_S} = \operatorname{diag}\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}, \dots \right\}.$$

Then the solvability condition for the equation (17) takes the form

$$\mathcal{P}_{Y_S}\mathcal{P}_{Y_D} \int_0^2 N(s)f(s)ds = \operatorname{diag}\left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}, \dots \right\} \times$$

$$\times \int_0^2 \operatorname{diag}\left\{ \begin{bmatrix} 1 & -1-\frac{s^2}{2}+\frac{3s}{2} \\ -4+\frac{3s}{2} & 0 \\ 1 & \frac{s-1}{2} \end{bmatrix}, \begin{bmatrix} 1 & -1-\frac{s^2}{2}+\frac{3s}{2} \\ -4+\frac{3s}{2} & 0 \\ 1 & \frac{s-1}{2} \end{bmatrix}, \dots \right\} f(s)ds =$$

$$= \int_0^2 \operatorname{diag}\left\{ \begin{bmatrix} 0 & 0 \\ \frac{3s-5}{2} & \frac{3(s-1)}{4} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ \frac{3s-5}{2} & \frac{3(s-1)}{4} \\ 0 & 0 \end{bmatrix}, \dots \right\} f(s)ds = 0.$$

After the transformations, we obtain the results which show that the components of the vector-function $f(t) = \operatorname{col}\big(f_1(t),\ f_2(t), \dots\big)$ must satisfy the conditions

$$\int_0^2 \Big[2(3s-5)f_{2k-1}(s) + 3(s-1)f_{2k}(s)\Big]ds = 0, \quad k = 1, 2, 3, \dots .$$

These conditions are satisfied, for example, by the vector $f(t) = \operatorname{col}(0, 1, 0, 1, 0, 1, \dots)$.

For this vector, the solution of the equation will have the form

$$z(t) = \operatorname{col}\left[ \begin{bmatrix} \frac{t^2}{2} - t)c_2 \\ \frac{4-3t}{4}\tilde{c}_2 + \frac{3t^2}{2}c_3 \end{bmatrix}, \begin{bmatrix} \frac{t^2}{2} - t)c_4 \\ \frac{4-3t}{4}\tilde{c}_4 + \frac{3t^2}{2}c_5 \end{bmatrix}, \dots, \right].$$

**Example 4.2** We find the conditions for the solvability of the integro-differential equation, which is considered in the finite-dimensional Euclidean space [1, 3]

$$z(t) - M(t)\int_0^{2\pi} W(s)z(s)ds = g(t),$$

where

$$M(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}, \quad W(s) = \frac{1}{2\pi}\begin{bmatrix} 0 & 0 \\ \cos s & \sin s \end{bmatrix},$$

$$V(s) = 0, \quad g(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} f(t).$$

For this equation we have

$$W(s) = \int_s^{2\pi} W(s)ds = \frac{1}{2\pi}\begin{bmatrix} 0 & 0 \\ \sin s & \cos s - 1 \end{bmatrix}, \quad W_0 = \int_0^{2\pi} W(s)ds = \frac{1}{2\pi}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$D = I_2 - A = I_2 - \int_0^{2\pi} W(s)M(s)ds = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad P_{N(D)} = \mathcal{P}_{Y_D} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The matrix $S = \mathcal{P}_{Y_D}W_0$ is zero, so $\mathcal{P}_{Y_S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and the solvability condition (10) will have the form

$$\mathcal{P}_{Y_S}\mathcal{P}_{Y_D}\int_0^{2\pi} W(s)g(s)ds =$$

$$= \frac{1}{2\pi}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\int\limits_0^{2\pi}\begin{bmatrix} 0 & 0 \\ \sin s & \cos s - 1 \end{bmatrix}\begin{bmatrix} \cos s & \sin s \\ -\sin s & \cos s \end{bmatrix} f(s).$$

After the transformations, we obtain the condition

$$\int\limits_0^{2\pi}\big[f_1(s)\sin s + f_2(s)(1-\cos s)\big]ds,$$

which completely coincides with the conditions from [1, 3], obtained by other methods.

The proposed research method can be used to study the solvability conditions for integro-differential systems of the Volterra type equations [18] in the case when the system is not everywhere solvable.

## References


[1] Lando, Yu.K. On the solvability of an integro-differential equation. *Differential equations* **3**(4) (1967) 695–697. [Russian]

[2] Lando, Yu.K. On the index and the normal solvability of integro-differential operators. *Differential equations* **4**(6) (1968) 1112–1126. [Russian]

[3] Samoilenko, A.M., Boichuk, O.A. and Krivosheya, S.A. Boundary-value problems for systems of linear integro-differential equations with a degenerate kernel. *Ukrainian Mathematical Journal* **48**(11) (1996) 1576–1579 [Ukrainian]

[4] Boichuk, O.A. and Golovatska, I.A. Boundary-value problems for systems of integro-differential equations. *Nonlinear oscillations* **16**(4) (2013) 460–474. [Ukrainian]

[5] Kumar, P., Haloi, R., Bahuguna, D. and Pandey, D. N. Existence of Solutions to a New Class of Abstract Non-Instantaneous Impulsive Fractional Integro-Differential Equations. *Nonlinear Dynamics and Systems Theory* **16** (1) (2016) 73–85.

[6] Gupta, V. and Dabas, J. Existence Results for a Fractional Integro-Differential Equation with Nonlocal Boundary Conditions and Fractional Impulsive Conditions. *Nonlinear Dynamics and Systems Theory* **15** (4) (2015) 370–382.

[7] Boichuk, A.A., Zhuravlev, V.F. and Pokutnyi, A.A. Normally solvable operator equations in Banach space. *Ukrainian Mathematical Journal* **65**(2) (2013) 163–174. [Russian]

[8] Zhuravl'ov, V.P. Generalized Inversion of Fredholm Integral Operators with Degenerate Kernels in Banach Spaces. *Jornal of Mathematical Sciences* **212**(3) (2016) 275–289.

[9] Diekmann, O. Thresholds and traveling for the geographical spread of infection. *J. Math. Biol.* **6** (2) (1978) 109-130.

[10] Sahu, P.K. and S. Saha Ray. Comparative experiment on the numerical solutions of Hammerstein integral equation arising from chemical phenomenon. *Journal of Computational and Applied Mathematics* **291** (2016) 402-409.

[11] Sokolnikoff, I.S. *Mathematical Theory of Elasticity.* 2nd edition, McGraw-Hill Book Company, Inc., New York-Toronto-London, 1956.

[12] Daletskyi, Yu.L. and Krein, M.G. *Stability of Solutions of Differential Equations in Banach Space.* Moscow: Nauka, 1970. [Russian]

[13] Popov M. M. Addition spaces and some problems of modern geometry of Banach spaces. *Mathematics today'07* **13** (2007) 78–116. [Ukrainian]

[14] Zhuravlev V.F. Linear boundary-value problems for integral Fredholm equations with a degenerate kernel in Banach spaces. *Bukovynsky Mathematical Journal* **2**(4) (2014) 57–66. [Ukrainian]

[15] Boichuk, A.A., Zhuravlev, V.F. and Samoilenko, A.M. *Generalized Inverse Operators and Noetherian Boundary-Value Problems.* Kyiv: Publishing house of IM NASU, 1995. [Russian]

[16] Boichuk, A.A., Samoilenko, A.M. *Generalized Inverse Operators and Fredholm Boundary-Value Problems.* Berlin: De Gruyter, 2-nd Edition, 2016.

[17] Voevodin, V.V. and Kuznetsov, Yu.A. *Matrices and Calculations.* Moscow: Nauka, 1984. [Russian]

[18] Jafari, H. and Azad, A. A computational Method for Solving a System of Volterra Integro-differential Equations. *Nonlinear Dynamics and Systems Theory* **12** (4) (2012) 389–396.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 342–358

# Krasnoselskii's Theorem, Integral Equations, Open Mappings, and Non-Uniqueness

T.A. Burton [1*] and I.K. Purnaras [2]

[1] *Northwest Research Institute, 732 Caroline St., Port Angeles, WA 98362,*
[2] *Department of Mathematics, University of Ioannina,*
*P.O. Box 1186, 451 10, Ioannina, Greece*

Received: June 6, 2018; Revised: October 12, 2018

**Abstract:** We study Krasnoselskii's fixed point theorem on the sum of two operators restricted to the Banach space of continuous functions with the supremum norm. The work is based on "open mappings" in the sense that our mapping $P$ maps a closed bounded convex set $M$ into its interior $M^o$. We show that any fixed point of a mapping of the whole space must reside in $M^o$. This is very informative in case of non-uniqueness. We also extend a known transformation to hold for integral equations being the sum of a contraction and a compact map where the "forcing function" is the contraction. Several examples are given showing the construction of the unusually simple mapping sets.

**Keywords:** *Krasnoselskii's fixed points; open mappings; transformations; uniqueness.*

**Mathematics Subject Classification (2010):** 34A08, 34A12, 45D05, 45G05, 4H09.

## 1 Introduction

Much has been written about fixed point mappings which are either contractions or compact. But around 1954 Krasnoselskii studied a paper by Schauder on differential equations and concluded a variant of the idea that the inversion of a perturbed differential operator yields the sum of a contraction and a compact map. Embodied in that theorem are both Banach's contraction mapping principle and Schauder's second fixed point theorem. All three of these are conveniently found in the monograph by Smart [15]. Accordingly, Krasnoselskii offered the following fixed point theorem [15, p. 31 ].

* Corresponding author: mailto:taburton@olypen.com

**Theorem 1.1** *Let $M$ be a closed convex non-empty subset of a Banach space $(\mathcal{B}, \|\cdot\|)$. Suppose that $A$ and $B$ map $M$ into $\mathcal{B}$ and that*
*(i) $x, y \in M \implies Ax + By \in M$,*
*(ii) $A$ is continuous and maps bounded sets into compact sets.*
*(iii) $B$ is a contraction mapping with constant $\alpha < 1$.*
*Then there exists $y \in M$ such that $Ay + By = y$.*

We can get the idea from the survey paper by Park [14] that this has generated much interest. In fact, the literature on it is so large that we make no effort to survey it here.

It is noteworthy that investigators focus on the very general form of (i), but gloss over the common assumption that $(A + B) : M \to M$. Moreover, nothing is said about uniqueness or where another fixed point might reside. In real-world problems it can be a complete disaster if there is another fixed point and if it is in a set having fundamental features very different from those of points in $M$. We discuss these later.

We loosen (i) to just the mapping $(A+B)$ and then tighten it to $(A+B) : M \to M^o$, the interior of $M$. Then we restrict the space to continuous functions. This allows us to prove the result and be certain that any fixed point will reside in $M$. Hence, we will know that the fixed point has the general properties displayed in $M$. The next paragraph justifies the restriction of the space.

Normed spaces and even Banach spaces are very general and contain examples which have almost none of our intuitive properties in common. There is a classical example showing how wrong we can be when we see a counterexample to a conjecture when the counterexample involves sequence spaces, but our main interest lies in spaces of continuous functions with the supremum norm. A 1980 article in Smart [15, p. 39] states that there appears to be an open conjecture that [every shrinking mapping of the closed unit ball in a Banach space has a fixed point]. In fact, it is true for important spaces of continuous functions, but was shown to be false for sequence spaces in 1967. See MR3695827.

We believe that the same thing is at work here and it involves the property that a continuous function starting inside a Jordan curve cannot pass from the inside to the outside without explicitly crossing the curve. That is used frequently in stability theory of ordinary differential equations, but fails completely for corresponding results for difference equations in which the solution jumps over the boundary without touching it.

It is for these reasons that we believe that Krasnoselskii's theorem may be simple for Banach spaces of continuous functions $\phi : [0, E] \to \Re$ with the supremum norm denoted by $(\mathcal{B}, \|\cdot\|_{[0,E]})$ which typically concern a general class of integral equations represented as

$$x(t) = g(t, x(t)) + \int_0^t a(t, s) f(s, x(s))\, ds, \quad t \geq 0, \tag{$\mathcal{E}$}$$

in the aforementioned space. A special case of $(\mathcal{E})$ is

$$x(t) = g(t, x(t)) - \int_0^t C(t - s) f(s, x(s)) ds. \tag{1}$$

In the context of Krasnoselskii's theorem, $g(t, x)$ is $B$ - the contraction, while the integral is $A$ - the compact map. The natural mapping is then $P = B + A$ and we seek an appropriate closed bounded convex subset of $\mathcal{B}$. While we have greatly simplified (i), among other things we ask that $P : M \to M^o$ the interior of $M$ and we also have asked that $M$ be bounded. But this will yield far more than just a fixed point. With these conditions in hand we will use an old combination of Schaefer's theorem with Krasnoselskii's theorem and obtain a fixed point. With reference now to future remarks on non-uniqueness, we can be assured that any other solution will also reside in $M$ which tells us also that all solutions are bounded.

**Non-anticipative**: The entire paper rests on the following two paragraphs.

We will assume throughout that we are dealing with Volterra operators [8, p. 84]. For any pair of functions $x, y \in \mathcal{B}$ if $x(s) = y(s)$ on $0 \leq s \leq t \leq E$ then the operator $V$ satisfies $(Vx)(t) = (Vy)(t)$. When the operator $P$ is the natural operator defined by $(\mathcal{E})$ and the integral has, at $t = 0$, constant value for any function $x \in \mathcal{B}$, say zero, then $(Px)(0) = g(0, x(0))$ for any function $x \in \mathcal{B}$. In particular, if $\phi \in \mathcal{B}$ is a fixed point so that $P\phi = \phi$, then it is true that

$$(P\phi)(0) = \phi(0) = g(0, \phi(0)),$$

and the last equality is an algebraic relation in $\phi(0)$. Since $g$ is a contraction, this has a unique solution $\phi(0)$ independent of which fixed point of $P$ is under discussion. As we have asked no smoothness conditions on $f$, there may be many fixed points of $P$ but they all start at this single $\phi(0)$. Frequently we can solve that algebraic relation explicitly for $\phi(0)$.

Looking ahead, we will be finding a closed convex bounded set $M$ with the property that $P$ maps $M$ into its interior. If $\psi$ belongs to $M$ and has the property that $\phi(0) = \psi(0)$, then $P\psi$ is in the interior of $M$, so the distance from $\phi(0)$ to the complement of $M$ is positive.

## 2 The Location of a Fixed Point

Let $(\mathcal{B}, \|\cdot\|_{[0,E]})$ be the Banach space of continuous $\phi : [0, E] \to \Re$ with the supremum norm. If $B, A : \mathcal{B} \to \mathcal{B}$ are the operators defined, respectively, by $g$ and the integral in $(\mathcal{E})$, then we want $A$ to be compact in the sense that it is continuous and maps bounded subsets of $\mathcal{B}$ into compact sets, while $B$ is a contraction. The central idea here is that if $P_\lambda$ is the non-anticipative mapping $P_\lambda : \mathcal{B} \to \mathcal{B}$ defined by

$$P_\lambda \phi := \lambda B(\phi/\lambda) + \lambda A\phi, \quad 0 < \lambda \leq 1, \tag{2}$$

then conditions of the following theorem can be verified if for every $\phi \in \mathcal{B}$ the number $(P_\lambda \phi)(0)$ is completely known. This is readily seen in $(\mathcal{E})$ which we later explain in detail.

Throughout the paper, for a positive number $K$ we denote by $M_K$ the closed ball of center 0 and radius $K$ in the Banach space $\mathcal{B}$, i.e.,

$$M_K := \{x \in \mathcal{B} : ||\phi|| \leq K\}.$$

**Theorem 2.1** *Let the conditions on A, B, and $P_\lambda$ of (2) hold. Assume that there is a $K > 0$ such that the unique fixed point $x_0$ of $g(0, x) = x$ belongs to $(-K, K)$, and, for each $\lambda \in (0, 1]$ the mapping $P_\lambda$ of (2) satisfies $P_\lambda M_K \subset M_K^o := (M_K)^o$. Then $P\phi = \phi$ has a solution in $M_K^o$. Moreover, any solution $\phi$ of $P\phi = \phi$ resides in $M_K^o$.*

***Proof.*** We firstly prove that all fixed points of $P$ (if any) reside in $M_K^o$. Recall that any fixed point of $P$ starts from the unique solution of the equation $g(0, x) = x$. We now come to the idea introduced in [5]. Let $\phi$ be a fixed point of $P$ which does not reside in $M_K^o$. As $\phi(0) = x_0 \in (-K, K)$, we see that there exists $T \in (0, E]$ such that either $\phi(T) = K$ or $\phi(T) = -K$, while $-K < \phi(t) < K$, $t \in [0, T)$, say it holds $\phi(T) = K$. Then for the function

$$\phi_T(t) := \begin{cases} \phi(t), & t \in [0, T), \\ K, & T \le t \le E, \end{cases}$$

we see that $\phi_T \in M_K$ hence $P\phi_T \in M_K^o := \{x \in \mathcal{B} : ||\phi|| < K\}$, and so $||P\phi_T|| < K$. But $\phi_T(t) = \phi(t)$ for $t \in [0, T]$, thus, as $P$ is non-anticipative, we take $P\phi_T(T) = P\phi(T) = K$ which implies $||P\phi_T|| \ge K$, a contradiction to $P\phi_T \in M_K^o$. Thus, all fixed points of $P$ (if any) reside in $M_K^o$. Then, in view of $P = P_1$, existence of a solution of the equation $P\phi = \phi$ is yielded by the following result of Burton and Kirk [4] applied on the Banach space $\mathcal{B}$ and the fact that, by assumption, for any $\lambda \in (0, 1)$ we have $P_\lambda M \subset M_K^o$ which excludes (ii) in that theorem.

**Theorem 2.2** *Let $(\mathcal{B}, \|\cdot\|)$ be a Banach space, $A, B : \mathcal{B} \to \mathcal{B}$, $B$ be a contraction with constant $\alpha < 1$, and $A$ be continuous with $A$ mapping bounded sets into compact sets. Either*

*(i) $x = \lambda B(x/\lambda) + \lambda Ax$ has a solution in $\mathcal{B}$ for $\lambda = 1$, or*

*(ii) the set of all solutions $\lambda \in (0, 1)$ (if any), is unbounded.*

We now want to employ Theorem 2.1 to look at solutions to the equation

$$x(t) = g(t, x(t)) + \int_0^t a(t, s) f(s, x(s)) ds, \quad t \ge 0. \tag{$\mathcal{E}$}$$

We assume that $g$, $f : [0, \infty) \times \Re \to \Re$ are continuous with $g(t, z)$ being a contraction in its second variable $z$, i.e., there exists an $\alpha \in (0, 1)$ such that

$$|g(t, z_1) - g(t, z_2)| \le \alpha |z_1 - z_2|, \quad t \ge 0, \ \ z_1, z_2 \in \Re,$$

and that the kernel $a(t, s) : \{(t, s) : 0 < s < t\} \to \Re$ is absolutely integrable with respect to the second variable on $[0, t]$, for $t > 0$, and such that the function $\widetilde{a}$ defined by

$$\widetilde{a}(t) := \int_0^t |a(t, s)| ds, \quad t \ge 0,$$

is well defined and continuous for $t \ge 0$. Note that $a(t, s)$ may not be defined for $t = 0$ but we ask that $\underset{t\to 0+}{lim} \int_0^t |a(t, s)| ds \in \Re$, in other words mild singularities of $a$ are allowed. Clearly, if $a$ is continuous on the closed triangle sets $\{(t, s) : 0 \le s \le t\}$, then $a$ is absolutely integrable and $\underset{t\to 0+}{lim} \int_0^t |a(t, s)| ds = 0$.

In each one of the following three theorems we focus on giving conditions ensuring the existence of a ball $M$ in the Banach space $\mathcal{B}$ so that for each $\lambda \in (0, 1]$ the corresponding mapping $P_\lambda$ defined in (2) satisfies $P_\lambda M \subset M^o$. Recall that we have also assumed that $g$ is a contraction and that $A$ is compact. Then Theorem 2.1 not only yields existence of (at least one) solution to ($\mathcal{E}$), but also ensures that all solutions reside in $M^o$. As $E$ may vary and the set $M$ in Theorem 2.1 is taken to be a ball, for the sake of clarity we adopt the following notation:

For an $E > 0$ we denote by $\mathcal{B}_E$ the Banach space of bounded continuous functions $\phi : [0, E] \to \mathbb{R}$ equipped with the usual supremum norm

$$\|\phi\|_E := \sup_{t\in[0,E]} |\phi(t)|,$$

(or, simply $||\phi||$ when $E$ is fixed), and set

$$M_K := \{\phi \in \mathcal{B}_E : \|\phi\| \le K\},$$

where $K$ is a positive number. As $g(t,0)$ is continuous on $[0,E]$ for any $E>0$, we set

$$L_E := \sup_{t\in[0,E]} |g(t,0)|.$$

Clearly, $L_E$ is a well defined nonnegative real number for any $E>0$.

For $\lambda \in (0,1]$, we consider the mapping $P_\lambda$ on $\mathcal{B}_E$ defined by $\phi \in \mathcal{B}_E$ implies

$$(P_\lambda\phi)(t) := \lambda g\left(t, \frac{\phi(t)}{\lambda}\right) + \lambda \int_0^t A(t,s) f(s,\phi(s))\,ds, \quad t \in [0,E]. \tag{3}$$

Then $P_\lambda\phi$ is continuous on $[0,E]$, so $P_\lambda\phi : \mathcal{B}_E \to \mathcal{B}_E$.

Our first result presents a limit condition posed on the kernel $a(t,s)$ ensuring that $P_\lambda M_K \subset M_K^o$ for properly chosen positive numbers $E$ and $K$.

**Theorem 2.3** *Assume that*

$$\lim_{t\to 0+} \int_0^t |a(t,s)|\,ds = 0. \tag{4}$$

*Then there always exist $E>0$ and $K>0$ such that for the mapping $P_\lambda$ defined by (3) we have $P_\lambda : M_K \to M_K^o$ for any $\lambda \in (0,1]$.*

***Proof.*** Let $\lambda \in (0,1]$ and consider an arbitrary $E_1>0$. As $L_{E_1} := \sup_{t\in[0,E_1]} |g(t,0)|$ is a nonnegative real number, we may consider a $K>0$ such that

$$\frac{L_{E_1}+1}{1-\alpha} < K. \tag{5}$$

By continuity of $f$ on the compact set $S_1 := [0,E_1]\times[-K,K]$ we see that

$$m := \sup_{(t,z)\in S_1} |f(t,z)| \in \Re.$$

In view of (4), if necessary, we may choose an $E \in (0,E_1]$ such that

$$m\int_0^t |a(t,s)|\,ds < 1, \quad t\in[0,E]. \tag{6}$$

With the real numbers $E$ and $K$ defined above, we consider the Banach space $\mathcal{B}_E$ and the ball

$$M_K := \{\phi \in \mathcal{B}_E : \|\phi\| \le K\},$$

and note that as $[0,E]\subset[0,E_1]$, we have

$$S := [0,E]\times[-K,K] \subset [0,E_1]\times[-K,K] = S_1,$$

so

$$|f(t,z)| \le m \quad \text{for all} \quad (t,z)\in[0,E]\times[-K,K],$$

and

$$|f(t,\phi(t))| \leq m \quad \text{for all} \quad t \in [0,E], \quad \phi \in M_K.$$

Now for $\phi \in M_K$ and $\lambda \in (0,1]$ we have for $t \in [0,E]$

$$\left| g\left(t, \frac{\phi(t)}{\lambda}\right) - g(t,0) \right| \leq \alpha \left| \frac{\phi(t)}{\lambda} \right| \Longrightarrow \left| g\left(t, \frac{\phi(t)}{\lambda}\right) \right| \leq |g(t,0)| + \alpha \left| \frac{\phi(t)}{\lambda} \right|,$$

so

$$\begin{aligned} |(P_\lambda \phi)(t)| &\leq \left| \lambda g\left(t, \frac{\phi(t)}{\lambda}\right) \right| + \lambda \int_0^t |a(t,s)| \, |f(s,\phi(s))| \, ds \\ &\leq \lambda \left[ |g(t,0)| + \alpha \left| \frac{\phi(t)}{\lambda} \right| \right] + \lambda \int_0^t |a(t,s)| \, |f(s,\phi(s))| \, ds \\ &\leq L_{E_1} + \alpha\lambda \frac{|\phi(t)|}{\lambda} + \int_0^t |a(t,s)| \, m ds \\ &\leq L_{E_1} + \alpha \|\phi\| + m \int_0^t |a(t,s)| \, ds, \end{aligned}$$

from which, by the use of (6) we find

$$|(P_\lambda \phi)(t)| < L_{E_1} + \alpha K + 1, \ \ t \in [0,E], \quad \phi \in M_K.$$

Consequently, in view of (5) we have

$$\|P_\lambda \phi\| \leq L_{E_1} + \alpha K + 1 := K_0 < K, \quad \phi \in M_K,$$

and hence for any $\phi \in M_K$ it holds

$$P_\lambda \phi \in \left\{ \phi \in \mathcal{B}_E : \|\phi\| < K_0 + \frac{K - K_0}{2} \right\} \subset M_K^o,$$

i.e., $P_\lambda(M_K) \subset M_K^o$.

As already mentioned, when $a$ is continuous on the closed triangle sets $\{(s,t) : 0 \leq s \leq t\}, t > 0$, then (4) is automatically satisfied. However, if the kernel $a$ has singularity at $t = 0$, the limit condition (4) may not be satisfied. Indeed, the function $a(t,s) : \{0 < s < t\}, \to \mathbb{R}$, defined by $a(t,s) = \frac{1}{\sqrt{t}}(t-s)^{-1/2}, 0 < s < t$, is absolutely integrable with respect to $s$ on $[0,t]$ for any $t > 0$, also the function $\widetilde{a}(t) = \int_0^t |a(t,s)| \, ds$ may be well defined and continuous on $[0,t]$ by setting

$$\widetilde{a}(0) = \lim_{t \to 0+} \int_0^t a(t,s) \, ds = \lim_{t \to 0+} \frac{1}{\sqrt{t}} \int_0^t (t-s)^{-1/2} \, ds = \lim_{t \to 0+} \frac{1}{\sqrt{t}} 2\sqrt{t} = 2,$$

however, the limit condition (4) is not satisfied.

**Example 2.1** Fractional kernels of the type

$$a(t,s) = (t-s)^{q-1}, \quad 0 < s < t,$$

with $q \in (0,1)$ do satisfy condition (4) since

$$\lim_{t \to 0+} \int_0^t (t-s)^{q-1} \, ds = \lim_{t \to 0+} \frac{t^q}{q} = 0,$$

and so Theorem 2.3 may be applied on the equation with fractional kernel

$$x(t) = g(t, x(t)) + \int_0^t (t-s)^{q-1} f(s, x(s))\, ds, \quad t \geq 0,$$

with $q \in (0,1)$, $g$, $f$ continuous, and $g$ being a contraction. In particular, Caputo type fractional equations can be considered as special cases of the above equation.

Theorem 2.3 yields the existence of a (sufficiently small) interval $[0, E]$ and a corresponding $K > 0$ so that it holds $P_\lambda M_K \subset M_K^o$. When $f$ is bounded we may always find a $K_E > 0$ so that $P_\lambda M_{K_E} \subset M_{K_E}^o$ for any (arbitrarily large) $E > 0$, yet without assuming (4). In view of continuity of the function $\widetilde{a}(t) := \int_0^t |a(t,s)|\, ds$, we set

$$a_T := \sup_{0 \leq t \leq T} \widetilde{a}(t) = \sup_{0 \leq t \leq T} \int_0^t |a(t,s)|\, ds, \quad T > 0.$$

**Theorem 2.4** *Assume that there exists an $m_f \geq 0$ with*

$$|f(t,z)| \leq m_f, \quad (t,z) \in [0,\infty) \times \Re.$$

*Then for an arbitrary $E > 0$, there always exists a $K_E > 0$ so that for the mapping $P_\lambda$ defined by (3) we have $P_\lambda M_{K_E} \subset M_{K_E}^o$ for any $\lambda \in (0,1]$.*

***Proof.*** Let $\lambda \in (0,1]$ and consider an arbitrary $E > 0$. Take $K_E > 0$ with

$$\frac{L_E + m_f a_E}{1-\alpha} < K_E. \tag{7}$$

Then for $\phi \in M_{K_E}$, $t \in [0,E]$, we have

$$\begin{aligned}
|(P_\lambda \phi)(t)| &\leq \left| \lambda g\left(t, \frac{\phi(t)}{\lambda}\right) \right| + \lambda \int_0^t |a(t,s)|\, |f(s,\phi(s))|\, ds \\
&\leq L_E + \alpha |\phi(t)| + m_f \int_0^t |a(t,s)|\, ds \\
&\leq L_E + \alpha \|\phi\| + m_f \sup_{0 \leq t \leq E} \int_0^t |a(t,s)|\, ds \\
&\leq L_E + \alpha K_E + m_f a_E := \widehat{K}_E,
\end{aligned}$$

so by (7) we take $\|(P_\lambda \phi)\| \leq \widehat{K}_E < K_E, \quad \phi \in M_{K_E}$, from which it follows that $P_\lambda M_{K_E} \subset M_{K_E}^o$.

**Example 2.2** In relation to Example 2.1, as the function $f(t,x) = \sin^3 x + \frac{x}{x^2+1}$ is bounded, one can easily see that Theorem 2.4 applies to the equation

$$x(t) = g(t, x(t)) + \int_0^t (t-s)^{q-1} \left[ \sin^3 x(s) + \frac{x(s)}{x^2(s)+1} \right] ds, \quad t \geq 0,$$

with $q \in (0,1)$, $g$ continuous and contraction in $x$.

In the next result the strict condition of boundedness of $f$ is removed and sufficient conditions yielding that for a given $E > 0$ there exists a set $M_E \subset B_E$ so that $P_\lambda M_E \subset M_E^o$ are given.

**Theorem 2.5** *Assume that for a given $E > 0$, there exists $K_E > 0$ such that*

$$L_E + a_E \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right] < (1-\alpha) K_E. \tag{8}$$

*Then for the mapping $P_\lambda$ defined by (3) we have $P_\lambda M_{K_E}) \subset M_{K_E}^o$ for any $\lambda \in (0,1]$.*

***Proof.*** Let $E > 0$ be given and $K_E > 0$ be such that (8) holds true, i.e.,

$$\frac{L_E}{(1-\alpha) K_E} + \frac{a_E}{(1-\alpha) K_E} \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right] < 1. \tag{9}$$

Then for $\lambda \in (0,1]$, $\phi \in M_{K_E}$, $t \in [0,E]$, we have

$$\begin{aligned} |(P_\lambda \phi)(t)| &\le \lambda \left| g\left(t, \frac{\phi(t)}{\lambda}\right) \right| + \lambda \int_0^t |a(t,s)| \, |f(s,\phi(s))| \, ds \\ &\le \lambda L_E + \lambda\alpha \left| \frac{\phi(t)}{\lambda} \right| + \lambda \int_0^t |a(t,s)| \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right] ds \\ &\le L_E + \alpha K_E + a_E \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right], \end{aligned}$$

thus

$$\|P_\lambda \phi\| \le L_E + \alpha K_E + a_E \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right] := \widetilde{K}_E.$$

In view of (9) we have $\widetilde{K}_E < K_E$ and so

$$P_\lambda \phi \in \left\{ \phi \in \mathcal{B}_E : \|\phi\| \le \widetilde{K}_E \right\} \subset (M_{K_E})^o,$$

i.e., $PM_{K_E} \subset M_{K_E}^o$.

**Example 2.3** Consider the equation

$$\begin{aligned} x(t) &= g_0(t) + \frac{t}{2t+1} \sin x(t) \\ &+ \frac{1}{2(11^2 t + 1)} \int_0^t (t-s)^{-1/2} \left( x^2(s) + \sqrt{|x(s)|} \right) \sin s ds, \quad t \ge 0 \end{aligned}$$

with $g_0$ continuous and bounded by 1. Here we have

$$g(t,x) = g_0(t) + \frac{t}{2t+1} \sin x, \quad t \ge 0, \quad x \in \Re,$$

and

$$|g(t,x) - g(t,y)| = \frac{t}{2t+1} |\sin x - \sin y| \le \frac{1}{2} |x-y|,$$

so $g$ is a contraction with $\alpha = 1/2$. Setting

$$a(t,s) = \frac{(t-s)^{-1/2}}{2(11^2 t + 1)}, \quad 0 < s < t,$$

we may see that the function $\widetilde{a}(t)$ is continuous and for any $t \in [0,1]$ we have

$$\int_0^t |a(t,s)|\, ds = \frac{1}{2(11^2 t + 1)} \int_0^t (t-s)^{-1/2}\, ds = \frac{2\sqrt{t}}{2(11^2 t + 1)} \leq \frac{1}{2 \cdot 11},$$

thus, for any $T \geq 0$ it holds

$$a_T := \sup_{0 \leq t \leq T} \int_0^t |a(t,s)|\, ds \leq \frac{1}{2 \cdot 11}.$$

Next, we let $f(s,x) = \left(x^2 + \sqrt{|x|}\right) \sin s$, $s \geq 0$, $x \in \Re$ and so, taking $K_E = 3$ for any $E > 0$, we have

$$\sup_{(s,z) \in [0,E] \times [-K_E, K_E]} |f(s,z)| \leq \sup_{z \in [-3,3]} \left(z^2 + \sqrt{|z|}\right) = 9 + \sqrt{3},$$

and

$$\begin{aligned} L_E + A_E \left[ \sup_{(s,z) \in [0,E] \times [-K_E, K_E]} |f(s,z)| \right] &\leq 1 + \frac{1}{2 \cdot 11} \left(9 + \sqrt{3}\right) \\ &< 1 + \frac{1}{2 \cdot 11}(9 + 2) \\ &= \frac{3}{2} = \left(1 - \frac{1}{2}\right) 3 \\ &= (1 - \alpha) K_E, \end{aligned}$$

i.e., condition (8) is satisfied with $K_E = 3$ for any $E > 0$, so Theorem 2.5 is applied. We conclude that for any $\lambda \in (0,1]$, if $P_\lambda$ is the mapping defined by (3), then $P_\lambda M_3 \subset M_3^o$ for any $E > 0$. In particular, for any $\phi \in M_3$ we have

$$\|P_\lambda \phi\| \leq \frac{31 + \sqrt{3}}{11} < 3.$$

To get a more convenient condition than (8), for a fixed (but arbitrary) $z \in \Re$ we set

$$f_E(z) := \sup_{s \in [0,E]} |f(s,z)|, \quad z \in \Re,$$

and consider $\widehat{f}_E : \Re \to \Re^+$ with

$$\widehat{f}_E(z) := \sup_{t \in [-z,z]} f_E(t) \quad z \in \Re.$$

From Theorem 2.5 we have the following corollary.

**Corollary 2.1** *If for some $E > 0$ it holds*

$$a_E \liminf_{z \to \infty} \frac{\widehat{f}_E(z)}{z} < 1 - \alpha, \tag{10}$$

*then there always exists an $M_E \subset B_E$ such that for the mapping $P_\lambda$ defined by (3) we have $PM_E \subset M_E)^o$, for any $\lambda \in (0, 1]$.*

***Proof.*** Let $E > 0$ be such that (10) holds. Since $\underset{z\to\infty}{lim} \frac{L_E}{z} = 0$, by (10) we may find a $z_0 > 0$ so large that

$$\frac{L_E}{z} + a_E \liminf_{z \to \infty} \frac{\widehat{f}_E(z)}{z} < 1 - \alpha, \quad z > z_0.$$

It follows that we may choose a $K_E > z_0$ so that

$$\frac{L_E}{K_E} + a_E \frac{\widehat{f}_E(z)}{K_E} < 1 - \alpha,$$

i.e.,

$$L_E + a_E \widehat{f}_E(z) < K_E(1 - \alpha), \tag{11}$$

or

$$L_E + a_E \left[ \sup_{(s,z)\in[0,E]\times[-K_E,K_E]} |f(s,z)| \right] < K_E(1 - \alpha),$$

which is (8), so Theorem 2.5 is applied.

## 3 Non-uniqueness and Examples

If $M$ is bounded, then the conclusion that any solution resides in $M$ can be far more important than Krasnoselskii's theorem itself for it can be a suitable substitute for uniqueness, a property that neither Krasnoselskii's nor Schauder's theorem yield. This brings us to another main idea. It was Kneser [12] in 1923 who jolted us with the idea that non-uniqueness is the father of disaster, as the example $x' = x^{1/3}, x(0) = 0$, shows by having both a bounded solution, namely $x(t)$ identically zero, and an infinite collection of unbounded solutions. In such cases what possible good can come from the information that there is a solution in the form of a fixed point of a natural mapping?

That information was one of the great motivating factors in the study of stability theory showing that solutions of differential equations starting near each other will stay near each other.

Our focus here will be on the mapping into $M^o$ and on uniqueness which we believe is a new project.

**Example 3.1** We consider the scalar integral equation

$$x(t) = (1/2)x(t)\sin t + \int_0^t a(t-s)x^{1/3}(s)ds$$

in which $a : (0, \infty) \to [0, \infty)$ is continuous and there is an $E > 0$ with

$$\int_0^E a(s)ds \le 1.$$

The mapping set is

$$M = \{\phi : [0, E] \to \Re \colon |\phi(t)| \leq 8\}$$

and our Banach space is the continuous functions on $[0, E]$ with the supremum norm.

The natural mapping $P : M \to M$ is defined by $\phi \in M$ implies

$$(P\phi)(t) = (1/2)\phi(t)\sin t + \int_0^t a(t-s)\phi^{1/3}(s)ds$$

so that by inserting $\lambda$ as in (3) and then taking norms we obtain

$$\|P\phi\| \leq (1/2)\|\phi\| + \|\phi\|^{1/3}\int_0^E a(s)ds \leq (1/2)8 + 8^{1/3} = 6 < 8$$

and so $P : M \to M^o$ for every $\lambda \in (0, 1]$. A result in [10] shows the continuity and the compactness of the integral maps. The conditions of Theorem 2.1 are satisfied and there is a fixed point. Moreover, every fixed point resides in $M$. More general results on compact mappings by integrals are found in [7] and [6].

Our next example will show that the mapping into $M$ but not into $M^o$ results in a fixed point in $M$ and one which is not in $M$. This problem starts from a differential equation about which we know a great deal. This enables us to see what the fixed point theorem can do and cannot do.

**Example 3.2** Consider the initial value problem

$$x' = x^{1/3}, \quad x(0) = 0$$

which has one solution $x(t) \equiv 0$. Separation of variables yields

$$x^{-1/3}dx = dt \implies (3/2)x^{2/3} = (3/2)x(0)^{2/3} + t$$

and

$$x^{2/3} = (2/3)t \implies x = \pm\left(\frac{2}{3}t\right)^{3/2}$$

as second and third solutions which are unbounded and hence are of a very different type than the first solution.

This problem occurs in elementary text books, but it is enormously complicated. Kneser's theorem tells us that there is a continuum of solutions between

$$x = 0 \textit{ and } x = (2t/3)^{3/2}$$

as well as between

$$x = 0 \textit{ and } x = -(2t/3)^{3/2}.$$

If we applied Schauder's fixed point theorem it would tell us that there is a solution, but would not suggest which of the three types it might be. But things can get worse. Do we really know that these are the only kinds which might arise? We are going to find out.

It is a routine matter to convert our initial value problem into an integral equation which is

$$x(t) = \int_0^t e^{-(t-s)}(x(s) + x^{1/3}(s))ds, \tag{12}$$

still retaining the solution $x(t) \equiv 0$.

Let $E > 0$ be given and find

$$\int_0^E e^{-s} ds = 1 - e^{-E}.$$

Next, find $a > 0$ so that

$$(a + a^{1/3})\left(1 - e^{-E}\right) < a. \tag{13}$$

This is possible since

$$\frac{a + a^{1/3}}{a} \to 1$$

as $a \to \infty$.

Let

$$M = \{\phi : [0, E] \to \Re\colon 0 \leq \phi(t) \leq a\}$$

and define $P : M \to M$ by $\phi \in M$ implies

$$0 \leq (P\phi)(t) = \int_0^t e^{-(t-s)} \left[\phi(s) + \phi^{1/3}(s)\right] ds \leq [a + a^{1/3}] \left(1 - e^{-E}\right) < a.$$

Notice that $PM$ is not in $M^o$ since $\phi(t) \geq 0$. It is shown in [10] that $P$ is a compact map, so by Schauder's theorem $P$ has a fixed point. As this is not an open map, we cannot be sure that all fixed points are in $M$, as we already know.

First, there is a parallel mapping with

$$M = \{\phi : [0, E] \to \Re\colon 0 \geq \phi(t) \geq -a\}$$

which is mapped into itself, but not into its interior as it still contains the zero function.

We now combine the two and take

$$M = \{\phi : [0, E] \to \Re\colon \|\phi\| \leq a\}$$

with $a$ and $E$ generated as before. This time when we take the natural mapping we find that $M$ is mapped into $M^o$ and we now know that all solutions reside in this set.

Note that Theorem 2.5 can be applied here with $g \equiv 0, a(t, s) = e^{-(t-s)}, f(s, x) = x + x^{1/3}$. We find

$$L_E = 0, \quad a_E = 1 - e^{-E}, \quad \sup_{(s,z)\in[0,E]\times[-x,x]} |f(s, z)| = x + x^{1/3}.$$

Since $g \equiv 0$ may be considered as a contraction with contraction constant any $\alpha \in (0, 1)$, condition (8) reduces to asking for a constant $K_E$ with

$$a_E \left[K_E + (K_E)^{1/3}\right] < (1 - \alpha)K_E \tag{14}$$

and $\alpha$ being any convenient number in $(0, 1)$. This is equivalent to asking for a $K_E > 0$ with

$$a_E \left[K_E + (K_E)^{1/3}\right] < K_E. \tag{15}$$

Indeed, if a $K_E > 0$ satisfying (15) exists, then (14) is satisfied by taking $0 < \alpha < 1 - \frac{a_E\left[K_E+(K_E)^{1/3}\right]}{K_E}$. Clearly, (15) is (13) with $a = K_E$.

## 4 The Transformation: Large Kernels

The reader has noticed that some of our examples involve small kernels, quite unlike the vast majority of real-world problems. That can be remedied in an interesting and useful way provided that the kernel $a(t)$ satisfies a set of conditions ensuring the existence of a resolvent, $R(t)$, having a small integral. The conditions are as follows and they include kernels typified by $(t-s)^{q-1}$, $0 < q < 1$, occurring in heat problems and fractional differential equations of both Riemann-Liouville and Caputo type, as well as many others. That class is discussed in depth by Miller [13, p. 209] with consequences on pp. 212–213 and Gripenberg [11]. They are defined as follows:

(A1) $a(t) \in C(0,\infty) \cap L^1(0,1)$.

(A2) $a(t)$ is positive and non-increasing for $t > 0$.

(A3) For each $T > 0$ the function $a(t)/a(t+T)$ is non-increasing in $t$ for $0 < t < \infty$.

In those references it is shown that when $a$ has an infinite integral then the resolvent equation is

$$R(t) = a(t) - \int_0^t a(t-s)R(s)ds \tag{16}$$

and that

$$0 < R(t) \le a(t), \quad \int_0^\infty R(t)dt = 1. \tag{17}$$

When

$$\int_0^\infty a(t)dt = \alpha < \infty$$

then

$$\int_0^\infty R(t)dt = \frac{\alpha}{1+\alpha}$$

and that can simplify some of the calculations we previously saw with $E$.

In the work to follow, notice that if $J$ is a positive constant, then $Ja(t)$ still satisfies (A1)–(A3).

In a sequence of papers we showed the advantages of transforming an integral equation

$$x(t) = b(t) - \int_0^t a(t-s)f(s,x(s))ds \tag{18}$$

using a variation of parameters formula of Miller [13, pp. 191-192] into

$$x(t) = z(t) + \int_0^t R(t-s)\left[x(s) - \frac{f(s,x(s))}{J}\right]ds, \tag{19}$$

with

$$z(t) = b(t) - \int_0^t R(t-s)b(s)ds. \tag{20}$$

Here are the steps. Starting with (18) and $b(t)$ continuous on $[0,\infty)$ while $a$ satisfies (A1)–(A3) we have

$$\begin{aligned} x(t) &= b(t) - \int_0^t a(t-s)[Jx(s) - Jx(s) + f(s,x(s))ds] \\ &= b(t) - \int_0^t Ja(t-s)x(s)ds + \int_0^t Ja(t-s)\left[x(s) - \frac{f(s,x(s))}{J}\right]ds. \end{aligned}$$

The linear part is

$$z(t) = b(t) - \int_0^t Ja(t-s)z(s)ds \tag{21}$$

and the resolvent equation is

$$R(t) = Ja(t) - \int_0^t Ja(t-s)R(s)ds \tag{22}$$

so that by the linear variation-of-parameters formula we have

$$z(t) = b(t) - \int_0^t R(t-s)b(s)ds \tag{23}$$

and the non-linear variation of parameters formula then yields

$$x(t) = z(t) + \int_0^t R(t-s)\left[x(s) - \frac{f(s, x(s))}{J}\right]ds. \tag{24}$$

Miller [13, p. 192] points out that all steps are reversible and that will be important for our next step. This transformation was first given in [3] for integral equations of Caputo type and has since been worked out in a variety of situations. Many fine details are found in [1].

## 5 Extending the Transformation

Our transformation involves

$$x(t) = b(t) - \int_0^t a(t-s)f(s, x(s))ds \tag{25}$$

and that does not cover the sum of two operators as in

$$x(t) = g(t, x(t)) - \int_0^t a(t-s)f(s, x(s))ds \tag{26}$$

with $g$ a contraction and the last term a compact mapping. But we are reminded of an old trick that can be found in Bellman [2, p. 35]. The idea is to say: if there is a solution $x(t)$, then we can identify $g(t, x(t))$ as the forcing function $b(t)$ and perform the transformation. Recall that Miller [13, p. 192] points out that all steps are reversible and that will be important for our next step. According to our transformation, this equation is transformed into

$$x(t) = b(t) - \int_0^t R(t-s)b(s)ds + \int_0^t R(t-s)\left[x(s) - \frac{f(s, x(s))}{J}\right]ds. \tag{27}$$

If we can show that this equation has a solution, then reversing the steps we have $b(t) = g(t, x(t))$ so that this last equation is now

$$x(t) = g(t, x(t)) + \int_0^t R(t-s)\left[x(s) - g(s, x(s)) - \frac{f(s, x(s))}{J}\right]ds. \tag{28}$$

Once again we are looking at the sum of two operators, one is the same original function, while the integral is the compact map.

**Strategy** In this form, the equation is more algebraic than it is differential. With the integral of $R$ being 1, when we take the norm of both sides the integrand slips out as the sup of $g$ and $f$, multiplied by the integral of $R$ which is bounded by 1. Now in the derivation of the transformation the idea is to work inside a bounded mapping set in which we can take $J$ so large that $x(s)$ dominates $f(s,x(s))/J$. But now we have $g(s,x(s))$ in the integrand but it is a contraction, so $x(s)$ can still dominate it AND $f(s,x(s))/J$. In many problems from applied mathematics $f$ satisfies the "spring condition", $f$ has the sign of $x$.

What this means is that we can very often get a mapping set as simply

$$M = \{\phi \colon \|\phi\| \le \text{constant}\}$$

and that is going to be a set where $P : M \to M^o$. The problems become almost entirely algebraic.

**Example 5.1** We will go through the details for an integral equation of the type of (1), namely

$$x(t) = g(t,x(t)) - \int_0^t C(t-s)f(s,x(s))ds. \tag{1}$$

Using the transformation and the nonlinear variation of parameters formula, we write (1) as

$$x(t) = g(t,x(t)) + \int_0^t R(t-s)\left[x(s) - g(s,x(s)) - \frac{f(s,x(s))}{J}\right]ds \tag{29}$$

where $J$ is an arbitrary positive constant and

$$0 < R(t) \le C(t), \quad \int_0^\infty R(s)ds = 1. \tag{30}$$

Compactness: There are many known conditions under which $\int_0^t R(t-s)h(s,\phi(s))ds$ maps bounded sets into compact sets. The prime example is $C(t) = t^{q-1}, 0 < q < 1$, which includes heat transfer problems as well as fractional differential equations of both Caputo and Riemann-Liouville type. See, for example, [13, pp. 207-213] and [9].

To get a bound on solutions we ask the "spring conditions"

$$x \ne 0 \implies xg(t,x) > 0, \quad xf(t,x) \ge 0. \tag{31}$$

In view of Krasnoselskii's theorem we ask that $g(t,0) = 0$ and that there exist $0 < \alpha < 1$ with

$$|g(t,x) - g(t,y)| \le \alpha|x-y|, \tag{32}$$

for all $x, y \in \Re$, $t \ge 0$. Concerning $f$ we ask that there exist a $K > 0$ and a $J > 0$ such that $|x| \le K$ and $t \ge 0$ imply that

$$1 - \alpha \le \frac{f(t,x)}{Jx} \le 1, \quad x \ne 0. \tag{33}$$

**Proposition 5.1** *For $h$ continuous and bounded for $\phi$ bounded, let $\int_0^t R(t-s)h(s,\phi(s))ds$ map bounded sets in $(\mathcal{B}, \|\cdot\|_{[0,E]})$ into compact sets. Under conditions (30)–(33) with $\alpha \leq 1/2$, $K > 0$ fixed and*

$$M = \{\phi \in \mathcal{B} \colon \|\phi\| \leq K\},$$

*there is a solution of (29) in $M$ and if there is any other solution of (29), then it also resides in $M$. As $M$ is bounded, all possible solutions of (29) share that bound.*

***Proof.*** We verify the conditions of Theorem 2.1. For a fixed $K > 0$ let $J > 0$ be such that (32) and (33) hold. Following (2) we change (29) to

$$x(t) = \lambda g\biggl(t, \frac{x(t)}{\lambda}\biggr) + \lambda \int_0^t R(t-s)\biggl[x(s) - g(s,x(s)) - \frac{f(s,x(s))}{J}\biggr]ds, \tag{34}$$

for $0 < \lambda \leq 1$, and define $P_\lambda : M \to \mathcal{B}$ by $\phi \in M$ implies

$$(P_\lambda\phi)(t) = \lambda g(t, \phi(t)/\lambda) + \lambda \int_0^t R(t-s)\biggl[\phi(s) - g(s,\phi(s)) - \frac{f(s,\phi(s))}{J}\biggr]ds, t \geq 0.$$

Then, in view of (32) we have $|g(s,x)| \leq a|x|$ and

$$1 - \alpha \leq \frac{g(s,x)}{x} + \frac{f(s,x)}{Jx} \leq 1 + \alpha$$

or

$$-\alpha \leq \left[\frac{g(s,x)}{x} + \frac{f(s,x)}{Jx}\right] - 1 \leq \alpha,$$

thus

$$\left|1 - \left(\frac{g(s,x)}{x} + \frac{f(s,x)}{Jx}\right)\right| \leq \alpha.$$

It follows that for any $\phi \in M$ we take for $t \in [0,E]$

$$\begin{aligned}
|(P_\lambda\phi)(t)| \leq &\lambda|g(t,\phi(t)/\lambda)| + \int_0^t R(t-s)|\phi(s)|\left|1 - \left(\frac{g(s,\phi)}{\phi} + \frac{f(s,\phi)}{J\phi}\right)\right|ds \\
&\leq \lambda\alpha|\phi(t)|/\lambda + \int_0^t R(t-s)\|\phi\|\alpha ds \\
&\leq \|\phi\|\left[\alpha + \alpha\int_0^t R(t-s)ds\right] \leq \alpha\left(1 + \int_0^E R(t-s)ds\right)\|\phi\|
\end{aligned}$$

so, by $\alpha \leq 1/2$ and in view of $\int_0^E R(t-s)ds < 1$ we conclude that $P_\lambda M \subset M^o$.

The conditions of Theorem 2.1 hold, so there is a solution in $M$ which also contains any other solution of (29). This proves Proposition 5.1.

We may easily see that if there exist positive numbers $m_1, m_2$ with $1 - \alpha \leq \frac{m_1}{m_2}$ and such that

$$m_1|x| \leq |f(t,x)| \leq m_2|x|, \quad t \in [0,E], |x| \leq K,$$

then condition (33) is satisfied by taking $J = m_2$.

We note that the requirement that the contraction constant satisfies $\alpha \leq 1/2$ may be relaxed to $\alpha \leq 2/3$ by replacing (33) in Proposition 5.1 by the assumption that there exists a $J > 0$ such that $f$ satisfies

$$\alpha \leq \frac{f(t,x)}{Jx} \leq 2(1-\alpha), \quad t \in [0,E], 0 \neq |x| \leq K. \tag{35}$$

Indeed, by (32) and (35) we take

$$\alpha - 1 \leq \frac{g(s,x)}{x} + \frac{f(s,x)}{Jx} - 1 \leq 1 - \alpha,$$

so

$$\left| \frac{g(s,x)}{x} + \frac{f(s,x)}{Jx} - 1 \right| \leq 1 - \alpha,$$

and the result is obtained following the arguments in the proof of Proposition 5.1.

## References


[1] Becker, L.C., Burton, T.A., and Purnaras, I.K. An inversion of a fractional differential equation and fixed points. *Nonlinear Dynamics and Systems Theory* **15** (4) (2015) 242–271.

[2] Bellman, Richard. *Stability Theory of Differential Equations*, McGraw-Hill, New York, 1953.

[3] Burton, T.A. Fractional differential equations and Lyapunov functionals. *Nonlinear Anal.: TMA* **74** (2011) 5648–5662.

[4] Burton, T.A. and Kirk, Colleen. A fixed point theorem of Krasnoselskii-Schaefer type. *Math. Nachr.* **189** (1998) 23–31.

[5] Burton, T. A. and Purnaras, I.K. Equivalence of differential, fractional differential, and integral equations: Fixed points by open mappings *Mathematics in Engineering, Science and Aerospace* **8**(3) (2017) 293–305.

[6] Burton, T.A. and Zhang, Bo. Fixed points and fractional differential equations: Examples. *Fixed Point Theory* **14** (2) (2013) 313–326.

[7] Burton, T.A. and Zhang, Bo. A NASC for equicontinuous maps for integral equations. *Nonlinear Dynamics and Systems Theory* **17** (3) (2017) 247–265.

[8] Corduneanu, C. *Integral Equations and Applications*, Cambridge, 1991.

[9] Diethelm, Kai. *The Analysis of Fractional Differential Equations*. Springer, Berlin Heidelberg, 2010

[10] Dwiggins, D.P. Fixed point theory and integral equations. *Dynamics of Continuous, Discrete and Impulsive Systems Series A: Mathematical Analysis* **23** (2016) 47–57.

[11] Gripenberg, G. On positive, nonincreasing resolvent of Volterra equations. *J. Differential Equations* **30** (1978) 380–390.

[12] Kneser, H. Uber die Lösungen eines Systems gewöhnlicher Differential gleichungen das der Lipschitzschen Bedingung nicht genğt. *S. B. Preuss, Akad. Wiss. Phys.-Math. Kl.* (II4) (1923) 171–174.

[13] Miller, R.K. *Nonlinear Volterra Integral Equations*. Benjamin, Menlo Park, CA, 1971.

[14] Park, Sehie. Generalizations of Krasnosleskii fixed point theorem. *Nonlinear Analysis* **67** (2007) 3401–3410.

[15] Smart, D.R. *Fixed Point Theorems*. Cambridge, 1980.

Nonlinear Dynamics and Systems Theory, 18 (4) (2018) 359–371

# Dual Star Induction Motor Supplied with Double Photovoltaic Panels Based on Fuzzy Logic Type-2

F. Hamidia [1], A. Abbadi [1], A. Tlemçani [1*] and M.S. Boucherit [2]

[1] *LREA Laboratory, Yahia Feres University, Medea, Algeria*
[2] *LCP Laboratory, Ecole Nationale Polytechnique, ENP, Algiers, Algeria*

Received: March 24, 2018; Revised: October 10, 2018

**Abstract:** Production of electrical energy is carried out mainly from fossil fuels and nuclear fuel. The excessive consumption of these energies during the 20th century has led to an aggravated pollution of the atmosphere. Since this form of energy covers most of the current energy production, it is necessary to find alternative solutions. The constraint is therefore to have an economical and sustainable source of energy, since environmental protection has also become a very important point. Several studies have been carried out in the field of renewable energies, such as photovoltaic energy, it has gained a lot of attention in recent years because it is environmentally friendly and sustainable compared to traditional energy sources. We can consider also the direct torque control (DTC) as an alternative to conventional methods of control by pulse width modulation (PWM) and by field oriented control (FOC), the direct torque control (DTC) found by Takahashi offers high performance in terms of simplicity in control and fast electromagnetic torque response. With dominant characteristics, the direct torque control for AC electric motor drive supplied by a solar energy is alternative in industrial applications. This paper discusses and presents the application of direct torque control (DTC) in open and closed loop, using voltage source inverter to control motor torque and flux with maximum power point tracking in weather conditions and load variation.

The P&O MPPT algorithm is mostly used, due to its ease of implementation, however, this MPPT algorithm gives us more torque ripples mainly with load variation. To resolve this problem, we will propose in this paper a fuzzy logic type-2 technique to replace the first one (P&O).

**Keywords:** *DTC; DSIM; Photovoltaic (PV) array; MPPT; DC/DC converter; P&O; FLC type-2.*

**Mathematics Subject Classification (2010):** 03B52, 93C42, 94D05.

* Corresponding author: mailto:h_tlemcani@yahoo.fr

## 1 Introduction

At present, photovoltaic (PV) cell arrays are a promising source of renewable energy since solar energy is free, abundant and readily available in many locations [1]. It is durable, clean and environmentally friendly. Solar energy is an attractive alternative solution to the energy of fossil fuels. Photovoltaic energy develops very quickly. It is multidisciplinary in nature, involving mechanics, power electronics, control theory, and other fields.

In order to control the electrical power delivered by PV panel, various methods are used: the action on the physicochemical properties of the cells, the action on the mechanical trackers for automatic orientation of the solar panels, and the action on the interface power electronics that connects the PV generator with its load. This last action is commonly called the electrical control of PV systems. It consists in the development of topologies of static converters and the development of MPPT (Maximum Power Point Tracking) control algorithms for the best capture of maximum power. Several studies have dealt with the problem of finding the operating point to draw the maximum energy from the PV modules using different MPPT methods. However, the non-linearity of the characteristic of the PV modules, their dependence on temperature, sunlight and the level of degradation of the characteristic make the implementation of these methods very complex. In the case of variations in weather conditions, these methods also exhibit poor convergence or oscillation around the optimum power point under normal operating conditions. If the transfer of power between the renewable energy sources and the load is not optimal, the overall efficiency of the system will be greatly affected. The research is going on to make these methods more effective.

The perturbation and observation (P & O) and incontinence (IncCond) techniques are widely used in the literature, but they fail under a rapid variation of weather conditions. This is why many researchers have made changes to these algorithms to improve their performance. Kook Soon et al. proposed an improvement of the IncCond algorithm to attenuate inaccurate responses during abrupt changes in the level of sunlight. There are also other techniques such as the method based on short-circuit current measurement, the method based on open-circuit voltage measurement, the method based on artificial neural networks and the method based on fuzzy logic [2–5].

Among the techniques mentioned above, the MPPT method based on fuzzy logic type-2 is presented in this paper. The application of this method allows to adapt the load to the PV modules and to follow the PPM whatever the variations of the meteorological conditions. In recent years, the field of the AC machines application is significantly expanded with the development of power electronics, the Microelectronics and computer engineering. Indeed, technological developments have allowed alternative machines, especially the double stator machine DSM, to acquire the control flexibility and dynamic performance naturally obtained. Since the late 1920s, stator machines with two three-phase windings were introduced to increase the power of synchronous generators of high-power. Multi-phase machines have subsequently been a growing interest, especially the dual star induction machine (DSIM), which has many advantages compare to asynchronous machines. Indeed, multiphase drives have several advantages over conventional three-phase machine, such as power segmentation, minimizing torque ripples and rotor losses, reduction of harmonic currents, high reliability and high power etc. Multi-phase machines are used much more in high power applications. Among these applications there are pumps, fans, compressors, compressor mills, cement mill. It uses a power electronics devices which allow a higher commutation frequency if compared to the simple three phases machines. Usually, high performance motor drive systems used in robotics, rolling mills, machine tools etc. require fast and accurate speed response, quick recovery of speed from any disturbance and uncertainties [6]. The introductions of new control such as DTC are very promising. In PV systems, they achieve great performance, fast responses and less fluctuations in steady state, for rapid irradiance and/or temperature variation improving the amount of energy effectively extracted from PV generator [7], the first idea of direct torque control was developed in 1986. This method of controlling that has progressed during past decade, provides a fast torque response and also it is robust against machine parameter variations [8–10].

In this paper, we study and discuss the application of Fuzzy logic type-2 for DTC-DSIM which is supplied with photovoltaic energy.

## 2 PV Module

A practical PV array consists of a collection of solar cells connected in series and/or parallel. An equivalent circuit model for a solar cell is shown in Fig. 1. The model consists of a current source, a diode, a shunt resistor $R_P$ and a series resistance $R_S$.

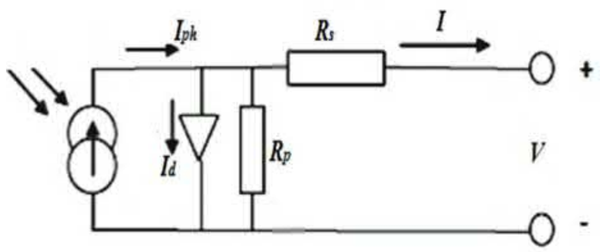


**Figure 1**: PV cell model.

The topology of boost converter is shown in Fig. 2. For this converter, the output voltage is always higher than the input PV voltage. Power flow is controlled by the on/off duty cycle of the switching transistor.

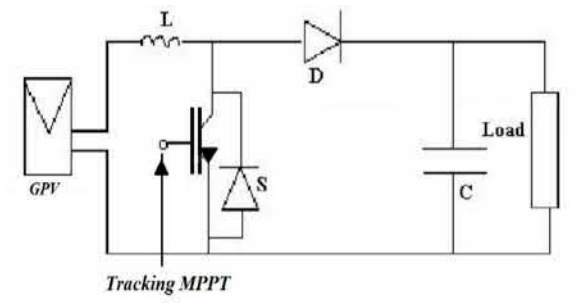


**Figure 2**: Boost converter DC/DC.

For sizing a photovoltaic system, we have to specify first the motor consumption in order to define the energy need for the load. Second we must take into account the obtained results and also the meteorological data as the input parameters of the photovoltaic installation of the input program. The sizing of the photovoltaic system is carried out according to the algorithm [11].

## 3 MPPT Control

To optimize the power provided by the generator, a static converter which operates as an adapter must be added as cited below. It exploits the MPPT technique, there are many algorithms that are used to control the MPPT, classical as P and O or artificial intelligent technique as ANN.

The commercialized solar modules are formed generally by a number of cells assembled in parallel $N_P$ or /and in series $N_S$. In addition, a data sheet is provided and includes the following main information about the product presented in Table 1. 12 panels are assembled in series and parallel to generate a DC voltage range in a MPP operation under different load changes.

| | |
|---|---|
| The nominal open-circuit voltage | 42.1V |
| The nominal short-circuit current | 3.87A |
| The voltage at the MPP | 33.7V |
| The maximum experimental peak output power | 120W |
| The current at the MPP | 3.56A |
| Parallel resistance Rs | 0.473Ω |
| Serie resistance Rp | 1367Ω |

**Table 1**: Data sheet information on a PV panel BP MSX120.

The algorithms that are most commonly used are the perturbation and observation method (P&O), the dynamic approach method and the incremental conductance algorithm. The P&O method is used because of its simplicity [11]. The Perturbation and observation (P&O) method has a simple feedback structure and fewer measured parameters. The P&O method is the most widely applied method in PV industry. It is based on the idea of introducing perturbation to the operating voltage and observing whether the power increases or decreases [12].

However, when we use this method on the panel related to the DTC-DSIM we face a problem to couple the inverter and the boost chopper, which appears in high torque ripples (Fig. 8) to resolve this problem, we propose in this paper to replace P&O by fuzzy logic type-2 approach (Fig. 3).

## 4 MPPT With Fuzzy Logic Type-2

Fuzzy logic is an approach for easy and flexible modeling of complex systems whose modeling is difficult and in some cases impossible by mathematics or classic modeling methods [10, 13–15].

In this section, we will replace the most algorithms of P&O (used to get the maximum of power provided by the PVG), by fuzzy logic type-2. The Fig. 4 presents the structure

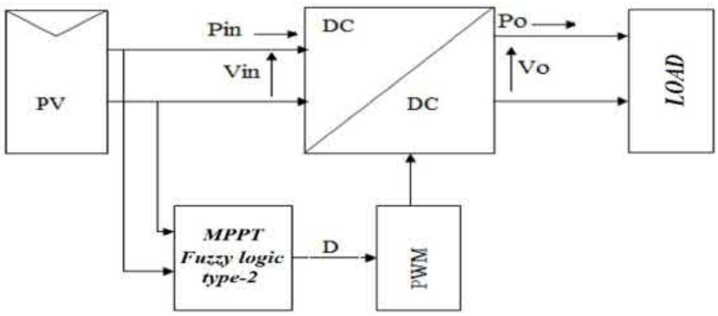


**Figure 3**: Boost converter DC/DC using MPPT-FL2.

of fuzzy logic type-2.

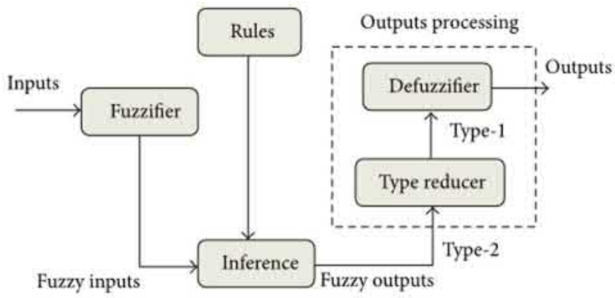


**Figure 4**: Structure of type-2 fuzzy logic system.

Fuzzy controller contains the five Gaussian membership function forms. The FL2 membership function has lower and upper bounds, their rules are almost similar to those for the conventional type-1.

The Reduction type used in this paper is centroid. The input and the output present respectively the power error and the duty cycle of the switching transistor.

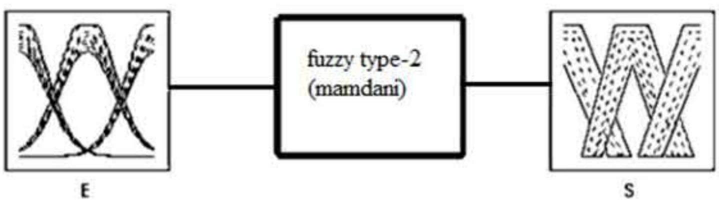


**Figure 5**: Fuzzy logic-2 MPPT.

## 5 Dual-Star Induction Motor

The most used example of multi-phase motors is the dual-star induction motor. In the conventional configuration, the double stator induction machine needs a double three-phase supply; the two stars share the same stator and are shifted by an electrical angle $\alpha$(in this work $\alpha = 30^o$) and a rotor is either wound or a squirrel cage. To simplify the study, we consider the electrical circuits of the rotor as equivalent to a three-phase short-circuit winding.

The position of the winding axes of the nine phases constituting the machine is shown in Fig. 6. There are six phases for the stator and three phases for the rotor. The six-stator phases are divided into two wyes-connected three phase sets labeled $A_{S1}$, $B_{S1}$, $C_{S1}$, and $A_{S2}$, $B_{S2}$, $C_{S2}$, The windings of each three phase set are uniformly distributed with their axes spaced by $2\pi/3$ in the space. The three phase rotor windings $A_r$, $B_r$, $C_r$ are also sinusoidal distributed and have axes that are displaced apart by $2\pi/3$ [17].

The following assumptions are made [18]:

- Motor windings are sinusoidal distributed.
- The two stars have same parameters.
- Symmetrical construction machine.
- The magnetic saturation, the mutual leakage inductances and the core losses are negligible.
- Resistances of the windings do not vary with the temperature and neglect the effect of skin (skin effect).
- Flux path is linear.

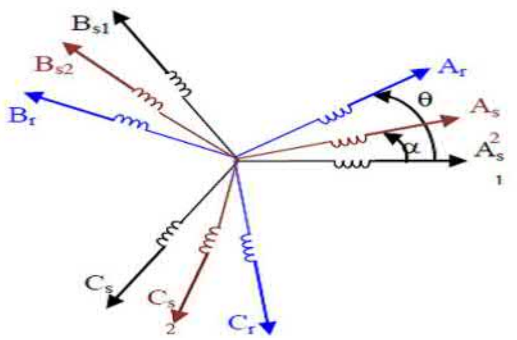


**Figure 6**: Windings of the six phase induction machine [18].

The equations for the stator are calculated in the reference frame related to the rotor.

$$\begin{aligned} \frac{d\varphi_{sd1}}{dt} &= V_{sd1} - R_{s1} I_{sd1} + \omega_s \varphi_{sq1}, \\ \frac{d\varphi_{sq1}}{dt} &= V_{sq1} - R_{s1} I_{sq1} - \omega_s \varphi_{sd1}, \\ \frac{d\varphi_{sd2}}{dt} &= V_{sd2} - R_{s2} I_{sd2} + \omega_s \varphi_{sq2}, \\ \frac{d\varphi_{sq2}}{dt} &= V_{sq2} - R_{s2} I_{sq2} - \omega_s \varphi_{sd2}, \\ \frac{d\varphi_{rd}}{dt} &= V_{rd} - R_r I_{rd} + (\omega_s - \omega_r)\varphi_{rq}, \\ \frac{d\varphi_{rq}}{dt} &= V_{rq} - R_r I_{rq} + (\omega_s - \omega_r)\varphi_{rd}, \end{aligned} \tag{1}$$

where the expressions for the stator and rotor flux are:

$$\begin{aligned} \varphi_{sd1} &= L_{s1} I_{sd1} + L_m (I_{sd1} + I_{sd2} + I_{rd}), \\ \varphi_{sq1} &= L_{s1} I_{sq1} + L_m (I_{sq1} + I_{sq2} + I_{rq}), \\ \varphi_{sd2} &= L_{s2} I_{sd2} + L_m (I_{sd1} + I_{sd2} + I_{rd}), \\ \varphi_{sq2} &= L_{s2} I_{sq2} + L_m (I_{sq1} + I_{sq2} + I_{rq}), \\ \varphi_{rd} &= L_r I_{rd} + L_m (I_{sd1} + I_{sd2} + I_{rd}), \\ \varphi_{rq} &= L_r I_{rq} + L_m (I_{sq1} + I_{sq2} + I_{rq}). \end{aligned} \tag{2}$$

The electromagnetic torque is obtained by:

$$T_e = \frac{P L_m}{L_r + L_m} [(\varphi_{rd}(I_{sq1} + I_{sq2}) - \varphi_{rq}(I_{sq1} + I_{sq2})]. \tag{3}$$

## 6 Direct Torque Control (DTC)

With the development of power electronics components, numerous studies of AC drive control have been proposed.One of most known, presents the vector control, which was in latest years the most important field research and most suitable to industrial requirements; however, this structure suffers from some drawbacks because it need establishment of a mechanical sensor and it is very sensitive to parameter variations.To resolve these problems,a new technique called control DTC (the Direct Torque Control) has been applied as an alternative.

The basic idea of DTC focuses on the flux orientation, using the instantaneous values of voltage vector. An inverter provides eight voltage vectors, among which two are zeros [19]. These vectors are chosen from a switching table according to the flux and torque errors as well as the stator flux vector position. In this technique, we do not need a modulator and a mechanical sensor to ensure feedback of speed or position [18].

To determine the electromagnetic state of the motor in order to determine the control of the inverter switches, a suitable model of the machine must be available. From the measurements of the DC voltage at the input of the inverter and of the stator currents, the model gives at each moment:

- actual stator flux of the machine,
- the real torque that it develops,
- its rotor speed.

Measurement of the rotor speed is not necessary, which is a great advantage of these methods. Which is independent of the rotor machine parameters, it provides a faster torque response and it has a simpler configuration. The structure of the DTC-DSIM fed by double PV panel is shown in Fig 7, the basic idea of DTC control is to choose the best stator voltage vector, which is taken from a switching table according to the stator flux vector position, the flux and torque errors. During this rotation, the amplitude of the flux rests in a pre-defined band [20], [21]. In this technique, we do not need the rotor position in order to choose the voltage vector. This particularity defines the DTC as an adapted control technique of AC machines and is inherently a motion sensor-less control method [19].

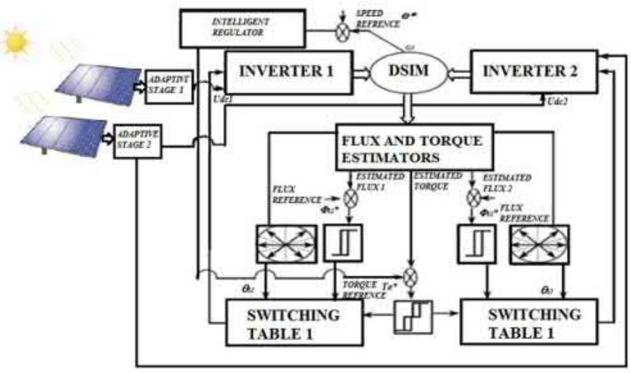


**Figure 7**: Schematic diagram of DTC-DSIM supplied by double PVG.

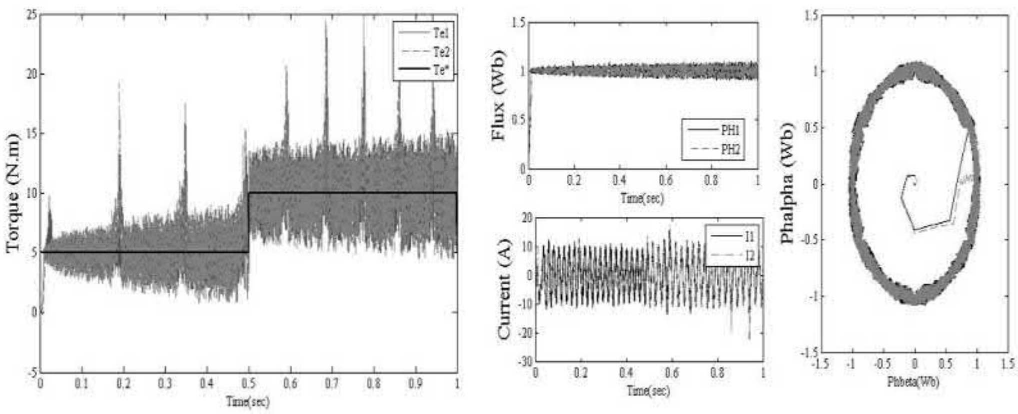


**Figure 8**: Torque, flux and current responses using DTC-DSIM in open loop with P&O method and load variation (G=1000W/m $^2$, T=298K ).

The stator flux vector can be estimated using the measured current and voltage vectors [20]:

$$\frac{d\varphi_s}{dt} = V_s - R_s I_s \tag{4}$$

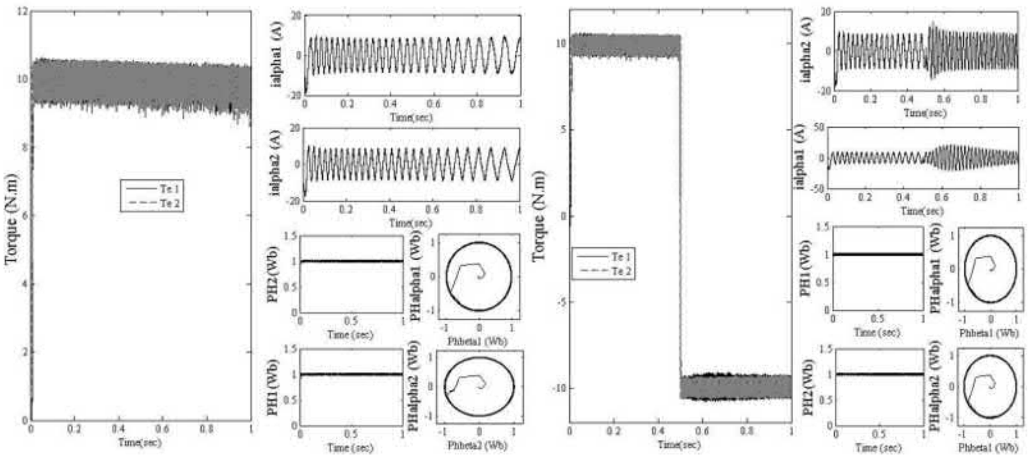


**Figure 9**: Torque, flux and current responses using DTC-DSIM in open loop by using Fuzzy logic type-2 with T=298K°, G=1000W/m $^2$.

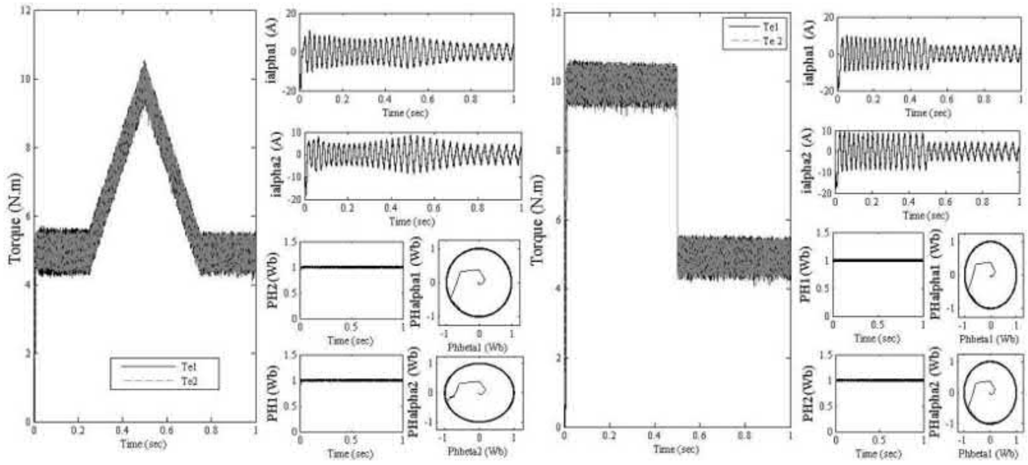


**Figure 10**: Torque, flux and current responses using DTC-DSIM in open using Fuzzy logic type-2 with T=298K°, G=1000W/m $^2$.

or

$$\varphi_s = \int (V_{s-} R_s I_s)\, dt. \tag{5}$$

## 7 Digital Simulation

Firstly, our system has been simulated in the same conditions to compare the performances of dynamic torque control with and without speed control. To verify the effectiveness of the proposed techniques, simulations are performed in this section by using MATLAB/SIMULINK. In this simulation of dual-star asynchronous machine, the nom-

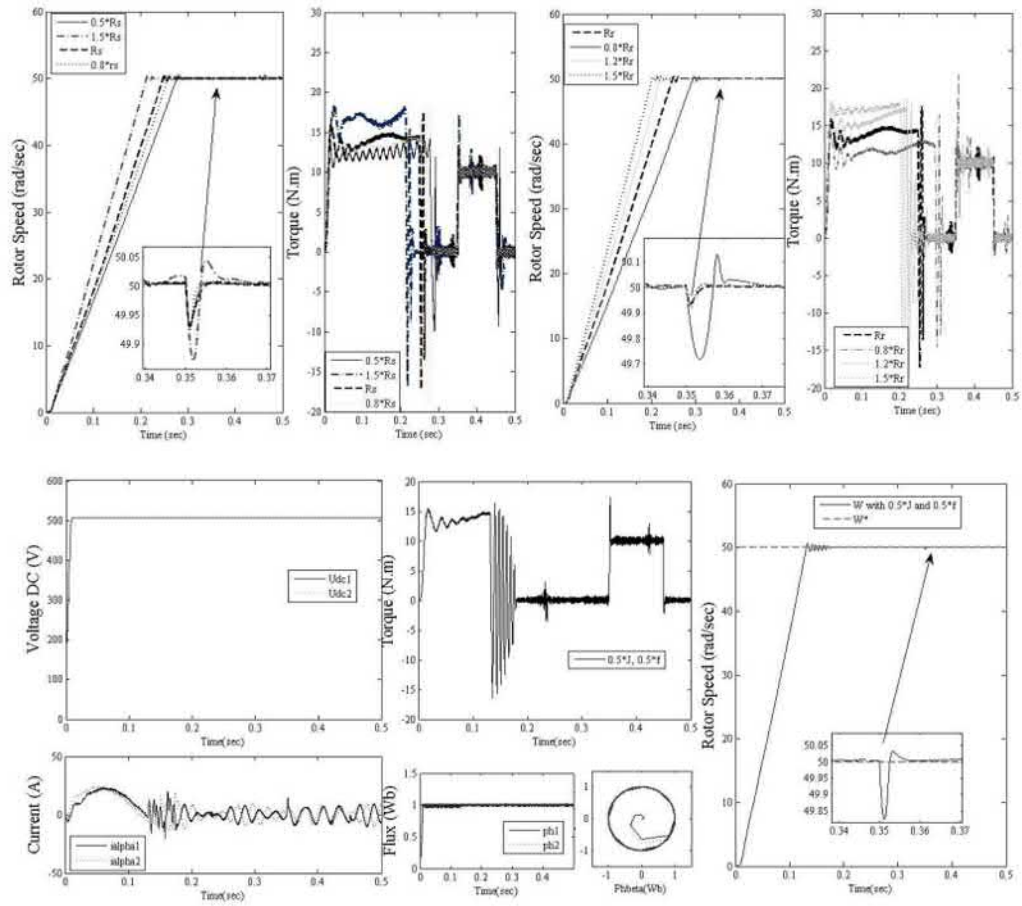


**Figure 11**: Torque and rotor speed responses using DTC-DSIM in closed loop using Fuzzy logic type-2 with T=298K°, G=1000W/m2 and load parameters variation (stator and rotor resistance, friction f and moment of inertia J).

inal voltage is 220V, nominal power is 4.5kw, stator resistances are 3.72 Ohm, rotor resistance is 2.12Ohm, mutual inductance is 0.3672H, rotor inductance is 0.006H, moment of inertia is 0.0662kg.m $^2$ and friction coefficient is 0.001.

## 8 Discussion of Results

As a first step, we have presented simulation results of DTC-DSIM without speed regulation (in open loop) as shown Figs. 8,9 and 10.

It can be seen that the estimated values follow their references by applying a load torque 10N, with a changing load torque applied from 10Nm to -10Nm (at 0.5sec) and constant command flux of 1Wb.

It can be noticed that these results obtained by using P&O algorithm give us higher torque and flux ripples (Fig 8) than the results obtained by using fuzzy logic type-2 (Fig.

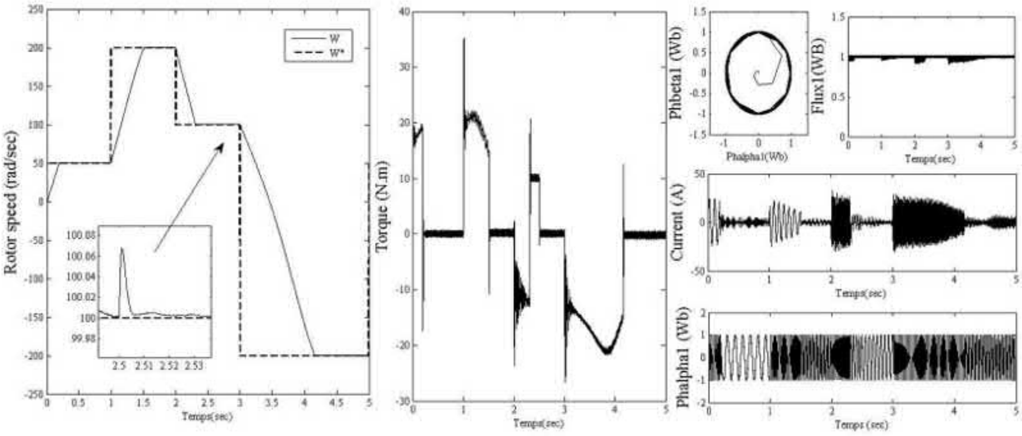


**Figure 12**: Torque, flux, current and rotor speed responses using DTC-DSIM (in closed loop) with load torque, speed variation.

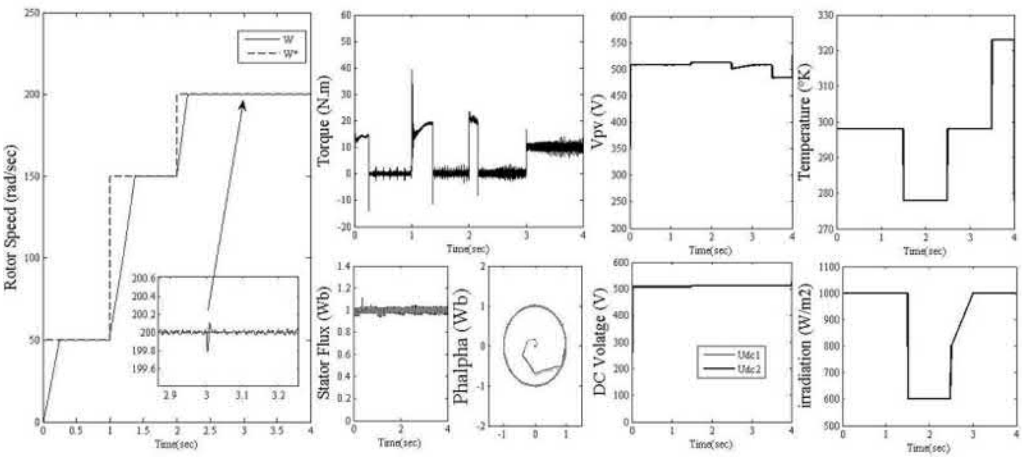


**Figure 13**: Torque and rotor speed responses using DTC-DSIM (in closed loop) with load torque, speed, irradiation and temperature variation.

9 and Fig. 10).

In order to verified the robustness of the proposed approach under load parameters variation (stator and rotor resistance, friction f and moment of inertia J) we carried out a test for DTC-DSIM in closed loop as is shown in Fig. 11.

Fig. 12 depicts, during the starting up with no load,the system simulated with changing speed reference; (50rad/sec changed to 200 and 100 rad/sec, at 3sec the rotor speed is inversed to -200rad/sec). The speed reaches quickly its reference value without overshoot, however, when the nominal load is applied at 2.5sec, a little overtaking is noticed and the command reject rapidly the disturbance.

Fig. 13 presents also the same responses as the first one but simulated under different conditions, that means by changing the load (electromagnetic torque) and weather conditions such the temperature and irradiation. From this figure, it can be found that our system has satisfactory performance.

## 9 Conclusion

Direct torque controlling dual star induction motor with and without speed regulation has been discussed in this paper. As can be analyzed from current waveforms,it shows that it is nearly sinusoidal; stator flux and electromagnetic torque track their references. So, the obtained results were very successful and confirm the validity of the proposed technique. It has been possible to obtain satisfactory results using DSIM supplied by double PVG which alone fed power to the AC motor and maintained the output voltage at a predetermined value. Summing up the results, it can be concluded that this study show the feasibility of fuzzy logic type-2 approaches and the improved dynamic performance of the machine in the case of variations in weather conditions or load. This can offer a very interesting solution to the conventional controller, and give us a control of the AC motor in a friendly and clean environment.

## References


[1] Sami Fadali, M. Jafarzadeh, S. and Nafeh, A. Fuzzy TKS Approximation Using Type-2 Fuzzy Logic Systems and Its Application to Modeling a Photovolaic Array. In: *American Control Conference Marriott Waterfront*. Baltimore, MD, USA, 2010, 6454–6459.

[2] Kok Soon Tey and Mekhilef, S. Modified incremental conductance MPPT algorithm to mitigate inaccurate responses under fast-changing solar irradiation level. *Solar Energy*, Elsevier, **101** (2014) 333–342.

[3] Belkaid, A., Gaubert, J.P, Gherbi, A. and Rahmani, L. Maximum Power Point Tracking for Photovoltaic Systems with Boost Converter Sliding Mode Control. In: *IEEE 23$^{rd}$ International Symposium on Industrial Electronics*, 2014, 556–561.

[4] Bhatnagar, A.P. and Nema, B.R.K. Conventional and global maximum power point tracking techniques in photovoltaic applications: A review. *J. Renew. Sustainable Energy* **5** (2013) 1–23.

[5] Ghassami, A.A., Sadeghzadeh, S.M. and Soleimani, A. A high performance maximum power point tracker for PV Systems. *Electr Power Energy Syst* **53** (2013) 237–243.

[6] Uddin, M.N. Development of a Neuro-Fuzzy controller for Induction Motor, In: *Canadian Conference on Electrical and Computer Engineeirng*, 2004.

[7] Lara, D., Merino, G. and Salazar, L. Power Converter with maximum power point tracking MPPT for small wind-electric pumping systems. *Energy Conversion and Management* (2015) 53–62.

[8] Kianinezhad, R., Alcharea, R. and Nahid, B. A novel Direct Torque Control (DTC) for six-phase induction motors with common neutrals. In: *International Symposium on Power Electronics, Electrical Drives, Automation and Motion*, SPEEDAM, 2008, 107–112.

[9] Gh. Seifossadat. A direct torque control (DTC) method for dual three phase induction motors using a fuzzy inference system. In: *International Symposium on Power Electronics Electrical Drives Automation and Motion.* 2008.

[10] Hamidia, F. and Larabi, A., Tlemcani, A. and Boucherit, M.S. AIDTC Techniques for Induction Motors. *Nonlinear Dynamics and Systems Theory* **13** (2) (2013) 147–156.

[11] Barazane, L. A sliding mode control associated to the field-oriented control of asynchronous motor supplied by photovoltaic solar energy. *Revue des Energies Renouvelables* **11** (2) (2008) 317–327.

[12] Hadjaissa, A., Ameur, K., Ait Cheikh, M.S. and Essounbouli, N., A PSO-based Optimization of a Fuzzy-based MPPT Controller for a Photovoltaic Pumping System used for Irragation of greenhouse. *Iranian Journal of Fuzzy Systems* **13** (3) (2016) 1–18.

[13] Asgharpour, H.A., Alinejad-Beromi, Y. and Yaghobi, H. A Fuzzy-Based Speed Controller For Improvement of Induction Motors Drive Performance, *Iranian Journal of Fuzzy Systems* **13** (2) (2016) 61–70.

[14] Tlemçani, A., Sebaa, K. and Henini, N. Indirect adaptive fuzzy control of multivariable nonlinear systems class with unknown parameters. *Nonlinear Dynamics and Systems Theory* **14** (2) (2014) 162–174.

[15] Bentouati, S., Tlemçani, A., Boucherit, M.S. and Barazane, L. A DTC Neurofuzzy Speed Regulation Concept for a Permanent Magnet Synchronous Machine. *Nonlinear Dynamics and Systems Theory* **13** (4) (2013) 344–358.

[16] Loukal, K. and Benalia, L. Type-2 Fuzzy Logic Controller of a Doubly Fed Induction Machine. *Advances in Fuzzy Systems* **16** (2016) 1–10.

[17] Sadouni, R., Meroufel, A. and Djriou, S. Study and Simulation of Direct Torque Control (DTC) for a Six Phase Induction Machine (SPIM). *International Journal of Energy* **2** (7) (2013) 31–37.

[18] Sadouni, R. and Meroufel, A. Performances comparative study of Field Oriented Control (FOC) and Direct Torque Control (DTC) of Dual Three Phase Induction Motor (DTPIM). *International Journal of Circuits, Systems and Signal Processing* **2** (6) (2012) 163–170.

[19] Benalia, L., Chachi, C. and Abdessemed, R. Control and simulation of a Double Star Induction Machine Using Direct Torque Control DTC. *Journal of Electrical Engineering* (2006) 1–6.

[20] Kaboli, S., Zolghadri, M.R. and Emadi, A. A fast flux search controller for DTC based induction motor drives. In: *IEEE 36$^{th}$ Conference on Power Electronics Specialists* (2005) 739–744.

[21] Kaboli, S., Zolghadri, M.R., Roye, D., Guiraud, J. and Crebier, J.C. Design and Implementation of a Predictive Controller for Reducing the Torque Ripple in Direct Torque Control Based Induction Motor Drives, In: *IEEE 35$^{th}$ Annual IEEE Power Electronics Specialists Conference*. Aachen, Germany, 2004, 476–481.

[22] Dhar, S., Sridhar, R. and Mathew, G. Implementation of pv cell based standalone solar power system employing incremental conductance mppt algorithm, In: *IEEE International Conference on Circuits, Power and Computing Technologies.* Nagercoil, India, 2013, 356–361.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 372–391

# Bayesian Approach for Multi-Mode Kalman Filter for Abnormal Estimation

Ali Hussein Hasan [1*] Alia Muhsin Mnati [2] and Ali Obied [1]

[1] *University of Sumer, College of Computer Science and Information Technology*
[2] *University of Thiqar, College of Engineering*

Received: April 5, 2018; Revised: October 14, 2018

**Abstract:** The paper deals with Bayesian approach for multi-mode Kalman filter estimation for the states $x(k)$ from the set of successive observations $Y_k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations is driven. Abnormal operations may be related to fault in one of system components; sudden internal thermal noise or even missing the input signal and can be extended to the maneuver target tracking case. Whenever the abnormal operation is detected, we can start tracking the states in this mode of operation. So the main problem may be reformulated to be detection of the starting point of the abnormal operation. The numerical simulation for fault estimation of phosphor furnace in different conditions are used to show the effectiveness of the proposed approach.

**Keywords:** *Bayesian estimation; multi mode operation; interactive multiple model; Kalman filter fault estimation.*

**Mathematics Subject Classification (2010):** 93E10.

## 1 Introduction

Fault detection and isolation problems have many significant applications during the past three decades, such as parity space Eigen structure assignment, $H_\infty$ filtering, $H_\infty$ optimization, and unknown input observer [1]. It is known that multiple model systems (known as hybrid systems) are an important class of combination filtering, which are mostly used in many practical engineering and industrial fields such as maneuver target tracking systems and fault detection systems, ets. In general, the multiple model systems combine hierarchically discrete or continuous state spaces, and each state (which is called the mode) is associated with a dynamic process [2].

* Corresponding author: mailto:ali.husain@uos.edu.iq

Many algorithms have been proposed for solving the problem of hybrid systems, such as the generalized pseudo Bays [3], interacting multiple model [4], expectation propagation algorithm [5]. These algorithms estimate the states with low computational costs by approximating of the posterior state distributions in finite mixture models. The main drawback of these algorithms is that, whenever the system contains nonlinear or non-Gaussian modes, the mixture models can not approximate the distribution accurately, that leads to the estimation fail. The interactive multiple model (IMM) algorithm, see [4] and [9], may be considered one of the most important approaches for solving the switching systems with the Gaussian linear states. It is basic operation by applying the Kalman filter for each estimation mode under consideration that this mode is a correct one at this operational instant. Then, the weighted combination of state estimations by all filters is calculated to produce the final Gaussian mean and covariance [7]. This mixed estimation is used for the next estimation. The weights are calculated according to the probabilities of the models.

In this paper, a Bayesian approach for multi-mode Kalman filter estimation for the states $x(k)$ from the set of successive observations $Y_k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations is developed. The rest of the paper is organized as follows: Statement of the problem is described in Section 2. Multi-mode Kalman filter approach is described in Section 3. Proposed multi model estimation is described in Section 4. Simulation Results is described in Section 5. Conclusion is described in Section 6.

## 2 Statement of The Problem

Consider the classical discrete-time state space problems in case of linear model of the form:

$$x(k+1) = F(k+1,k)x(k) + BU(k) + \Gamma(k)v(k), \tag{1}$$

$$y(k) = Hx(k) + w(k), \tag{2}$$

where $x(k)$ is the discrete time state, $F$ is a state transition matrix, $B$ is an input matrix, the vector $U(k)$ is the input and assumed to be known at perspective times, $y(k)$ is a vector including the measurements at time $k$, $H$ is the associated observation matrix, $v(k)$ is a state noise process (or an input noise), and $w(k)$ is the measurement noise; both sequences $w(k)$ and $v(k)$ are assumed to be uncorrelated with the white Gaussian noise sequence with zero means and the covariance matrices $Q(k)$ and $R(k)$, and $v(k){\sim}N(0,Q(k)), w(k){\sim}N(0,R(k))$, $[cov(v(k),v(j)) = E[v(k)v(j)^T]\delta(k,j) = Q_k\delta(k,j)]$, $\delta(k,j)$ is the Kronecker delta, $cov\,(w(k),w(j))] = E\left[w(k)w(j)^T\right]\delta(k,j) = R_k\delta(k,j)$.

It is also assumed that $w$ and $v$ are uncorrelated and $cov\,(w(k),v(k)) = 0$ for all $k,j$. The initial value of $x$ represents a random variable with an average $\mu_x(0)$ and variance $V_x(0)$ such that $E\,[x(0)] = \mu_x(0)$ and $var\,[x(0)] = V_x(0)$. Assume that the observation noise $w(k)$ is uncorrelated with the system discrete-time state $x(k)$, such that $E\left[x(k)w(k)^T\right] = 0\ \forall\ k \geq 0$.

Now our problem is to estimate the states $x(k)$ from the set of successive observations $Y^k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations. Abnormal operations are related to the fault in one of the system components; sudden internal thermal noise or even missing the input signal and can be extended to the maneuver target tracking case. Whenever the abnormal operation is detected, we can start tracking the states in this mode of operation. So the main problem may be reformulated to be: detection of the starting point of the abnormal operation.

## 3 Multi-Mode Kalman Filter Approach

Let $\overset{\wedge}{x}(k/k)$ be the state estimation, then the estimation error will be:

$$\tilde{x} = x(k) - \overset{\wedge}{x}(k/k), \tag{3}$$

The estimation will be conditionally and unconditionally unbiased, such that

$$E\left\{\overset{\wedge}{x}(k/k)\,|y(k)\right\} = E\left\{x(k)|y(k)\right\}, \tag{4}$$

which leads to

$$E\left\{\overset{\wedge}{x}(k/k)\right\} = E\left\{x(k)\right\}, \tag{5}$$

which is a linear function of the observations $y(k)$. According to the linear unbiased estimation algorithms, we choose only the one that gives the minimal variance of error, i.e., $V_{\tilde{x}}(k/k) = var\left\{\tilde{x}(k/k)\right\}$ or simply $var\left\{\tilde{x}(k/k)|y(k)\right\}$ is as minimum as possible. For a given set of observations $Y^k$, the estimate based on the minimum of the mean-square error coincides with the conditional mean value of $x$, which is based on linear Kalman filter process.
1- Extrapolation process:

$$\widehat{x}(k/k-1) = F(k,k-1)\widehat{x}(k-1/k-1). \tag{6}$$

2- Estimation:

$$\widehat{x}(k/k) = \widehat{x}(k/k-1) + K(k)\left[y(k) - H(k)\widehat{x}(k/k-1)\right]. \tag{7}$$

3- Coefficient gain:

$$K(k) = P(k/k-1)H^T(k)\left[H(k)P(k/k-1)H^T(k) + R(k)\right]^{-1}. \tag{8}$$

4- Covariance extrapolation:

$$P(k+1/k) = F(k+1,k)P(k/k)F^T(k+1,k) + Q(k). \tag{9}$$

5- Covariance filtration:

$$P(k/k) = \left[1 - K(k)H(k)\right]P(k/k-1). \tag{10}$$

**Definition 3.1** Let the innovation process noise $\vartheta(k)$ be [8]

$$\vartheta(k) = y(k) - H(k)\widehat{x}(k/k-1). \tag{11}$$

Assume the innovation process noise $\vartheta(k)$ is the white Gaussian noise with zero mean expectation in the normal operation and, also, the white Gaussian noise but with non-zero mean expectation in the abnormal operation:

$$E\left[\vartheta(k)\right] = \begin{cases} 0, \ \textit{at the normal operation}, \\ g(k-k_0,\alpha), \ \textit{at the abnormal operation}, \end{cases} \tag{12}$$

where $g$ is a deterministic function, $k_0+1$ is the starting of abnormal operation and $\alpha$ is an intensity vector.

**Definition 3.2** Let us define the covariance of the additional intensity vector $\alpha$ as

$$s(k) = H(k)P(k/k-1)H^T(k) + R(k). \tag{13}$$

**Definition 3.3** Let the density of the probability distribution of the innovation process noise $\vartheta(k)$ be $w(\vartheta)$:

$$w(\vartheta) = \frac{1}{\sqrt{(2\pi)^n \det(s)}} e\left(\frac{-1}{2}\vartheta^T s^{-1}\vartheta\right). \tag{14}$$

The following equation is a Riccati equation

$$\begin{aligned} &P(k+1/k) = F(k+1,k) \\ &\left[P(k/k-1) - P(k/k-1)H^T(k)s(k)^{-1}H(k)P(k/k-1)\right] F(k+1/k)^T + Q(k). \end{aligned} \tag{15}$$

**Note**: The Riccati equation will converge to the constant matrix $P$ as $k \to \infty$ in case of the system we are dealing with time invariant. As a result, the gain coefficient $K$ converges to a constant small value and estimates of the parameters practically independent of the observed data. The result, is that, for small intensity $\alpha$, will not be taken into account in any way.

The problem of detection the starting point of the abnormal operation $k_o$ and estimation of its intensity vector $\alpha$ is a problem of detection a useful signal $g(k-k_o,\alpha)$ from the white noise and estimation of its parameters. It is known that, when solving the problem of detection, various optimality criteria (Bayesian or non-Bayesian) lead to a general decisive rule - the formation of the likelihood ratio and comparison with the threshold. The difference lies in the choice of the detection threshold. Here we assume that the abnormal operation intensity is a nonrandom process. To synthesize the meter of the intensity of the abnormal operation, it is convenient to use the maximum likelihood criterion. In this subsection we consider a general approach to the synthesis of the algorithm for simultaneous detection of the moment of the beginning and estimation of the abnormal operation. Let

$$U(k) = \begin{cases} u(k), & \text{at normal operation,} \\ u(k)+\alpha, & \text{at abnormal operation.} \end{cases} \tag{16}$$

It is assumed that the system starts the abnormal operation in the time between the moments $k_0$ and $k_0+1$, therefore the time moment, which is considered the beginning of the abnormal operation, is $k_0+1$. Let $\vartheta^0(k)$ and $\vartheta^1(k)$ be the updating processes corresponding to the absence and presence of an abnormal operation, then the problem of finding an abnormal operation consists of choosing one of the two alternative hypotheses.

$$H_0 : \vartheta(k) = \vartheta^0(k) \textit{ at the normal operation,} \tag{17}$$

$$H_1 : \vartheta(k) = \vartheta^1(k) = \vartheta^0(k) + g(k-k_0,\alpha) \textit{ at the abnormal operation,} \tag{18}$$

where $g(k-k_0,\alpha)$ is a useful signal that should be detected, whenever the abnormal operation is started, $g(k-k_0,\alpha)$ is introduced as an offset or bias in the innovation process at the time $k-k_0$. Since $\vartheta^0(k)$ is a random process with the white Gaussian noise with zero expectation and covariance and $\vartheta^1(k)$ is the mixing of the useful signal and the white noise $\vartheta^0(k)$, with the mathematical expectation $g(k-k_0,\alpha)$.

**Theorem 3.1** *The task is detecting a vector of deterministic signal $g(k-k_0,\alpha)$ with unknown parameters (the intensity $\alpha$ and the moment of the onset of the abnormal operation $k_0+1$) from a background of the white noise $\vartheta^0(k)$.*

***Proof.*** The optimal procedure for detection the of intensity vector $\alpha$ may be reduced to the formation of the likelihood ratio and compared with a predetermined threshold value $\lambda$,

$$L(k) = \frac{p\left[\vartheta(k-m+1),...,\vartheta(k)|H_1\right]}{p\left[\vartheta(k-m+1),...,\vartheta(k)|H_0\right]} \geq \lambda, \tag{19}$$

where $m = k - k_0 = 1...M$, $M$ is the number of samples. Assume that the samples of the white Gaussian noises are statistically independent, then we have

$$p\left[\vartheta(k-m+1),...,\vartheta(k)|H_0\right] = \prod_{n=k-m+1}^{k} \frac{1}{\sqrt{(2\pi)^l \det(s(n))}} e\left(\frac{-1}{2}\vartheta^T(n)s^{-1}\vartheta(n)\right), \tag{20}$$

and

$$p\left[\vartheta(k-m+1),...,\vartheta(k)|H_1\right] = \prod_{n=k-m+1}^{k} \frac{1}{\sqrt{(2\pi)^l \det(s(n))}} e\left(\frac{-1}{2}\left[\vartheta(n)-g(n-k_0,\alpha)\right]^T s^{-1}(n)\left[\vartheta(n)-g(n-k_0,\alpha)\right]\right), \tag{21}$$

where $l$ is the order of the vector $\vartheta(k)$ and $k$ is the current instant of time. From the equations (19)–(21) after algebraic transformations, taking into account that $s$ is a symmetric matrix, $\vartheta$ and $g$ are columns vectors,

$$\vartheta^T s^{-1} g = g^T s^{-1} \vartheta, \tag{22}$$

we get

$$L(k) = e\left\{\sum_{n=k-m+1}^{k}\left[g^T(n-k_0,\alpha)s^{-1}(n)\vartheta(n) - \frac{1}{2}g^T(n-k_0,\alpha)s^{-1}(n)g(n-k_0,\alpha)\right]\right\}, \tag{23}$$

Taking the logarithm of both sides of (23), we get

$$\ln L(k) = \sum_{n=k-m+1}^{k}\left[g^T(n-k_0,\alpha)s^{-1}(n)\vartheta(n) - \frac{1}{2}g^T(n-k_0,\alpha)s^{-1}(n)g(n-k_0,\alpha)\right], \tag{24}$$

□

Let $g(m,\alpha) = \vartheta^1(k) - \vartheta^0(k) = \vartheta^1(k_0+m) - \vartheta^0(k_0+m)$. Then $\vartheta^1(k_0+m)$ is the updating process at the time $k = k_0+m$ under the condition of correspondence between the real state (presence of the abnormal operation) of the system and the model of system used by the abnormal operation (model equation with $U(k) = u(k)+\alpha$), and $\vartheta^0(k_0+m)$ is the updating process at time $k = k_0+m$ under the condition of a discrepancy between the real state (the presence of the abnormal operation) of the system and the model of system used by the abnormal operation (model equation 1 with $U(k) = u(k)$). Then

$$\begin{aligned}\widehat{g}\ (m,\alpha) &= \left[y(k_0+m) - H\ \widehat{x}^1\ (k_0+m/k_0+m-1)\right] \\ &\quad - \left[y(k_0+m) - H\ \widehat{x}^0\ (k_0+m/k_0+m-1)\right] \\ &= H\left[\widehat{x}^0\ (k_0+m/k_0+m-1) - \widehat{x}^1\ (k_0+m/k_0+m-1)\right] \\ &= HF\left[\widehat{x}^0\ (k_0+m/k_0+m-1) - \widehat{x}^1\ (k_0+m/k_0+m-1)\right].\end{aligned} \tag{25}$$

Proof: The state estimation equation (1) at the normal and abnormal operation will be:

1. At time step 1

$$\begin{aligned}\widehat{x}^0\ (k_0+1/k_0+1) &= [F - K(k_0+1)HF]\ \widehat{x}\ (k_0/k_0) \\ &\quad + K(k_0+1)y(k_0+1) + [B - K(k_0+1)HB]\,u(k),\end{aligned} \tag{26}$$

2. At time step 2

$$\begin{gathered}\widehat{x}^0\ (k_0+2/k_0+2) = \\ [F-K(k_0+2)HF]\,[F-K(k_0+1)HF]\ \widehat{x}\ (k_0/k_0)\,[F-K(k_0+2)HF]\,K(k_0+1)y(k_0+1) \\ +K(k_0+2)y(k_0+2)\ \{[F-K(k_0+2)HF]\,[B-K(k_0+1)HB]+[B-K(k_0+2)HB]\} \\ u(k),\end{gathered} \tag{27}$$

3. At time step m

$$\begin{aligned}\widehat{x}^0\ (k_0+m/k_0+m) &= \prod_{i=0}^{m-1} [F - K(k_0+m-i)HF]\ \widehat{x}\ (k_0/k_0) \\ &+ \sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F-K(k_0+m-i)HF]\right]K(k_0+m-j-1)y(k_0+m-j-1)\right\} \\ &+ K(k_0+m)y(k_0+m) \\ &+ \sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F-K(k_0+m-i)HF]\right][B-K(k_0+m-j-1)HB]\right\} \\ &+ [B-K(k_0+m)HB]\,u(k),\end{aligned} \tag{28}$$

from the above equations (25)–(28) and after the substitution $m = k - k_0$, we obtain the expression for the "useful signal" $g(m,\alpha)$ in the following form

$$\begin{aligned}g(m,\alpha) &= HF\left\{\sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F-K(k-i)HF]\right][B-K(k-j-1)HB]\right\}\right. \\ &\qquad\qquad \left. + [B-K(k)HB]\right\}u(k) \\ &= G(m)u(k),\end{aligned} \tag{29}$$

where

$$G(m) = HF... \left\{ \sum_{j=0}^{m-2} \left\{ \left[ \prod_{i=0}^{j} [F - K(k-i)HF] \right] [B - K(k-j-1)HB] \right\} + [B - K(k)HB] \right\}, \tag{30}$$

$m \geq 1$. When $m = 1$, the term $\sum_{j=0}^{m-2} \left\{ \left[ \prod_{i=0}^{j} [F - K(k-i)HF] \right] [B - K(k-j-1)HB] \right\}$ will be equal to zero, and the matrix $G(m)$ becomes square matrix with dimensions $l \times l$, $l$ is the order of the observation vector. Substituting equation (29) into (24), we obtain

$$\ln L(k, \alpha) = \sum_{n=k-m+1}^{k} \left[ \alpha^T G^T(n - k_0) s^{-1}(n) \vartheta(n) - \frac{1}{2} \alpha^T G^T(n - k_0) s^{-1}(n) G(n - k_0) \alpha \right], \tag{31}$$

then

$$\begin{aligned} \ln L(k, \alpha) = \alpha^T & \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) \vartheta(n) \right] \right] \\ & - \frac{1}{2} \alpha^T \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) G^T(n - k_0) \right] \right] \alpha. \end{aligned} \tag{32}$$

From equation (32), by the criterion of maximum likelihood ratio, we find the estimate of the intensity vector,

$$\begin{aligned} \frac{\partial \ln L(k, \alpha)}{\partial \alpha} &= \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) \vartheta(n) \right] \\ &\quad - \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) G^T(n - k_0) \right] \right] \alpha \\ &= 0. \end{aligned} \tag{33}$$

Then

$$\widehat{\alpha}_m = \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) G(n - k_0) \right] \right]^{-1} \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) \vartheta(n) \right] \right]. \tag{34}$$

Taking the second order derivative, we get

$$\frac{\partial^2 \ln L(k, \alpha)}{\partial \alpha \partial \alpha^T} = \left[ \sum_{n=k-m+1}^{k} \left[ G^T(n - k_0) s^{-1}(n) G^T(n - k_0) \right] \right] \prec 0. \tag{35}$$

The function $\ln L(k)$ will reach its maximum value at the point $\alpha = \widehat{\alpha}_m$.

**Theorem 3.2** *$var\{Xy\} = X\,[var\{y\}]\,X^T$.*

***Proof.*** Let $X$ be an m × n matrix and $y$ be a n ×1 random vector. Then

$$var\{Xy\} = E\left[(Xy - XE(y))(Xy - XE(y))^T\right] = E\left[X(y - E(y))(y - E(y))^T X^T,\right]$$

$$= XE\left[(y - E(y))(y - E(y))^T\right] X^T = X\left[var\{y\}\right] X^T.$$

□

If all values are scaled by a constant, the variance is scaled by the square of that constant. Then, the covariance matrix of the estimation is

$$\begin{aligned} var\left\{\widehat{\alpha}_m\right\} = &\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]^{-1} \ldots \\ &\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)var\{\vartheta(n)\}\left[G^T(n-k_0)s^{-1}\right]^T\right]\right] \times \ldots \\ &\left[\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}\right]^T. \end{aligned} \tag{36}$$

Let the covariance matrix of the innovation process be $var\{\vartheta(n)\} = s(n)$. Since $s^{-1}(n)$ is a symmetrical matrix, we have that

$$\begin{aligned} &\left[\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}\right]^T \\ &= \left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}, \end{aligned} \tag{37}$$

yields the covariance matrix of this estimate as follows

$$var\left\{\widehat{\alpha}_m\right\} = \left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]^{-1}, \tag{38}$$

Substituting equation (33) into (32), we get

$$\ln L(k,\widehat{\alpha}_m) = \frac{1}{2}\sum_{n=k-m+1}^{k}\vartheta^T s^{-1}G\left[\sum_{m=k-m+1}^{k}G^T s^{-1}G\right]^{-1}\sum_{m=k-m+1}^{k}G^T s^{-1}\vartheta \geq \tilde{\lambda}_m, \tag{39}$$

It is obvious that for every value of $m$, there exists an estimate $\widehat{\alpha}_m$. Therefore, the simultaneous detection of the start of the abnormal operation and estimating the abnormal operation is the abnormal operation in multichannel (M channels). Therefore the optimal procedure for detecting the moment of the beginning of the abnormal operation has the form

$$\max_{m=1,M}\ln L(k,\widehat{\alpha}_m) \geq \tilde{\lambda}_m, \tag{40}$$

and the moment of the beginning of the abnormal operation is the difference$(k-m+1)$, where $m$ is the channel index, in which $\ln L(k, \widehat{\alpha}_m)$takes the maximum value, and $k$ is the current time moment.

It follows that (equations (39)–(40) ) the optimal procedure for detecting the moment of the onset of the abnormal operation is reduced to linear accumulation of the values of the quadratic forms (equation (39)) of the innovation process in $m$ adjacent cycles from $(k_0+1)$ to $k=(k_0+m)$th moments of time. Then, the maximum value is selected from the set of $M$ different values and is compared with a given threshold $\tilde{\lambda}_m$.

**Note:** The different channels accumulate different values, so to stabilize the probability of false alarm, it is necessary to compare the maximum value of the likelihood ratio (equation (40)) with different thresholds $\tilde{\lambda}_m$, chosen from the given probability of false alarm $P_{FA}$. Using the decision rule ( equation (40)), we can detect and evaluate, both the abnormal operation and the moment of the beginning of the abnormal operation.
To calculate the probability of a false alarm, it is necessary to know the law of distribution of the quadratic form $\ln L(k, \widehat{\alpha}_m)$ in the absence of an abnormal operation. It is known that [8] if $\vartheta(n)$ is an $l$ -dimensional vector with independent normally distributed components, each of which has the variance $\sigma_{iq}^2(q=1,2,...,l)$ and the mathematical expectation $E[\vartheta(n)]=0$, then the probability distribution density of the quadratic form $\ln L(k, \widehat{\alpha}_m)$ is the central $\chi^2$-distribution with $l \times m$ degrees of freedom. The corresponding density of the probability distribution is written in the form

$$\chi^2\left[\ln L(k, \widehat{\alpha}_m)\right] = \left[2^{\frac{lm}{2}}\Gamma(\frac{lm}{2})\right]^{-1}\left[\ln L(k, \widehat{\alpha}_m)\right]^{\frac{lm}{2}-1}\exp\left(\frac{1}{2}\ln L(k, \widehat{\alpha}_m)\right). \tag{41}$$

Then the probability of false alarm is given by

$$\int_{\tilde{\lambda}_m}^{\infty}\chi^2\left[\ln L(k, \widehat{\alpha}_m)\right]d\left[\ln L(k, \widehat{\alpha}_m)\right], \tag{42}$$

after detecting the beginning of the abnormal operation, either the Kalman filter parameters are adjusted (usually the gain factors or the covariance matrix Q (k)), or their structure is changed by using more complex models of state change taking into account the abnormal operation.

The general scheme that realizes this algorithm is shown in Figure 1 below.

In the first case, parameters adjustment of the Kalman filter will be changed according to the following formula

$$if\ \ln L(k, \widehat{\alpha}_1) \geq \tilde{\lambda}_1\ then\ Q^0(k) \to Q^1(k), \tag{43}$$

where $Q^0(k)$ and $Q^1(k)$ are chosen on the basis of the experiment so that to better reflect the true estimate of the state, both in the absence and in the presence of an abnormal operation. According to this, the elements of the matrix $Q^0(k)$ must take small values corresponding to rectilinear and uniform tracking with weak perturbations and $Q^1(k)$ are large values corresponding to tracking with abnormal operation. In general, the value of $Q^1(k)$ is selected on the basis of the possible maximum abnormal operation intensity of the system [10], for example, one may choose $Q^0(k) = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$ and $Q^1(k) = \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}$. The structure of the Kalman filter is changing by take the following

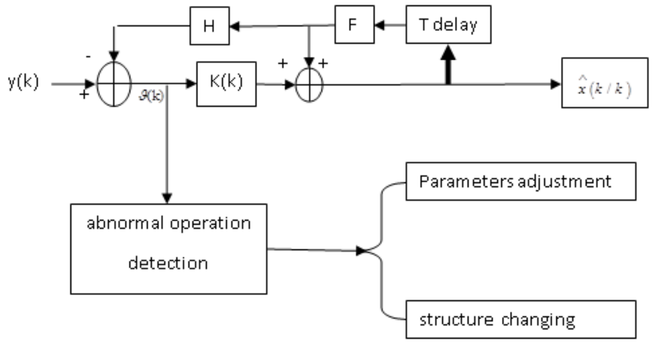


**Figure 1**: Kalman filter operation schema for abnormal estimation.

formula in the second case:

$$if\ \ \ln L(k,\widehat{\alpha}_1) \geq \tilde{\lambda}_1\ then\ model\ U(k) = u(k)\ converted\ to\ model\ U(k) = u(k) + \alpha. \quad (44)$$

Bayes adaptive approach [9] takes into account the possibility that the adopted model is inadequate, in other words, all possible variants of the parameters or hypotheses concerning the model are taken into account.

## 4 Proposed Multi Model Estimation

In the problem of filtering tracking of the states, the a priori uncertainty of the statistical characteristics of the system leads to an uncertainty in the statistical characteristics of the filtered state vector. According to this, which is unknown to the observer in advance, leads to a mismatch between the real state and the model used in the estimation devices. Let $\zeta$ be a vector containing all the indeterminate parameters that represent all the undefined events associated with the hypotheses about the model.

According to this, in the general case, the equations of state of the system and the observations take the form

$$x(k+1) = F(\zeta_{k+1})x(k) + \Gamma(\zeta_{k+1})v(k), \quad (45)$$

$$y(k) = Hx(k) + w(k). \quad (46)$$

Then, in this section, we have to estimate the conditional mathematical expectation

$$\overset{\wedge}{x}(k/k) = E\left[\, x(k) \middle|\, Y^k \right]. \quad (47)$$

According to the Bayesian approach, this operation is performed on the basis of the total probability theorem:

$$p\left[x(k)\left|Y^k\right.\right] = \int_\Upsilon p\left[x(k)\left|\zeta, Y^k\right.\right] p\left[\zeta\left|Y^k\right.\right] d\zeta, \quad (48)$$

where $\Upsilon$ is the set of all possible values of $\zeta$, $p\left[x(k)\left|Y^k\right.\right]$ is the conditional probability distribution of states $x(k)$ for given observation, $p\left[x(k)\left|\zeta, Y^k\right.\right]$ is the conditional probability distribution of states $x(k)$ for given observation and accepted value of $\zeta$ for undefined parameters and events, $p\left[\zeta\left|Y^k\right.\right]$ is the posteriori distribution of vector $\zeta$ for given observations. From equations (47) and (48) the optimal mean square estimate of $\overset{\wedge}{x}(k/k)$ and its covariance matrix $\overset{\wedge}{P}(k/k)$ are

$$\overset{\wedge}{x}(k/k) = \int_{\Upsilon} \widehat{x}^{\zeta}(k/k) p\left[\zeta\left|Y^k\right.\right] d\zeta, \tag{49}$$

$$\hat{P}(k/k) = \int_{\Upsilon} \left\{ \hat{P}^{\zeta}(k/k) + \left[\hat{x}^{\zeta}(k/k) - \hat{x}(k/k)\right]\left[\hat{x}^{\zeta}(k/k) - \hat{x}(k/k)\right]^T \right\} p\left[\zeta|Y^k\right] d\zeta, \tag{50}$$

where $\hat{x}^{\zeta}(k/k)$ and $\hat{P}^{\zeta}(k/k)$ are the partial estimate and its covariance obtained for a given value of $\zeta$, respectively, and equal to

$$\hat{x}^{\alpha}(k/k) = E\left\{x(k)|\alpha, Y^k\right\}, \tag{51}$$

$$\hat{P}^{\zeta}(k/k) = E\left\{ \left[\hat{x}^{\zeta}(k/k) - x(k)\right]\left[\hat{x}^{\zeta}(k/k) - x(k)\right]^T \left|\zeta, Y^k\right. \right\}, \tag{52}$$

Therefore, the obtained estimate can be represented as a linear combination of partial estimates $\hat{x}^{\zeta}(k/k)$, each of which is obtained under a certain hypothesis with respect to the model. The weight coefficients of this linear combination expressing the total estimate are determined by the probabilities of each hypothesis under consideration. A complete covariance matrix is calculated similarly, as a linear combination (with the same weights) of conditional matrices $\hat{P}^{\zeta}(k/k)$. Therefore, in principle, by the equations (49)–(52) we can obtain the optimal estimate $\hat{x}(k/k)$. In addition, as follows from equation (48), $p\left[x(k)|Y^k\right]$ is a linear combination of Gaussian probabilities. So, $\{\zeta_k = i\}$ or $\left\{\zeta_k^i\right\}$ the event consists the state in accordance with the state $i$ at the $k^{th}$ instant of time; $\Upsilon(k) = \{\Upsilon_1, \Upsilon_2, ..., \Upsilon_k\} = \left\{\{\zeta_1^i\}, \left\{\zeta_2^j\right\}, ..., \{\zeta_k^s\}\right\}, i, j, ..., s = \overline{0, M-1}$, is the set of all possible values that the discrete-random process $\zeta_k$ can take from the initial value to the $k^{th}$ time moment. And $\Upsilon(k, l_k) = \{\Upsilon(k) \cap l_k\}$ $l_k = \overline{1, M^k}$ is a subset containing $k$-value. Figure 2 describes the set of possible values of $\Upsilon$.

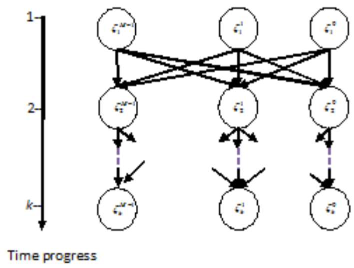


**Figure 2**: The set of possible values of $\Upsilon$.

From Figure 2; $\Upsilon(k, l_k)$ is the ${l_k}^{th}$ branch with the duration $k$. Hence the number of different branches in $\Upsilon(k)$ is equal to $M^k$. In this case, according to the Bayes formula, we obtain the posterior probability density of the state vector as follows:

$$p\left[x(k)\left|Y^k\right.\right] = \sum_{l_k=1}^{M^k} p\left[x(k)|\Upsilon(k, l_k), Y^k\right] P(\Upsilon(k, l_k)|Y^k), \tag{53}$$

the indices $i(l_k), j(l_k), ..., s(l_k)$ are varying between the limits [0,M-1] and the index $l_k \in 1, M^k$. Let the a posteriori probability density of the $l_k$-th branch $\Upsilon(k, l_k)$ be

$$P(\Upsilon(k,l_k)|Y^k) = P(\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|y(k), Y^{k-1}) = \frac{p\left[y(k)|\zeta_1^i, \zeta_2^j, ..., \zeta_k^s, Y^{k-1}\right] P\left[\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|Y^{k-1}\right]}{p\left[y(k)|Y^{k-1}\right]}, \tag{54}$$

then equation (53) in base of indices will be

$$p\left[x\left(k\right)\left|Y^k\right.\right] = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} ... \sum_{s=0}^{M-1} p\left[x(k)|\zeta_1^i, \zeta_2^j, ..., \zeta_k^s, Y^k\right] P(\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|Y^k), \tag{55}$$

To reduce possible hypotheses, it is assumed that [4] the unknown parameter $\zeta_k$ is a simply-connected Markov chain, with a matrix of transition probabilities $P = [p_{ts}]_{t,s=0,M-1}$, where $p_{ts} = P\left(\zeta_k^s|\zeta_{k-1}^t\right)$. Since the algorithm starts working at the $k^{th}$ time point, at the k$^{th}$ instant of time equation (55) will be

$$p\left[x(k)|Y^k\right] = \sum_{s=0}^{M-1} p\left[x(k)|\zeta_k^s, Y^k\right] P(\zeta_k^s|Z^k). \tag{56}$$

The Bayes description of the expression $p\left[x(k)|\zeta_k^s, Y^k\right]$ is written in the form

$$p\left[x(k)|\zeta_k^s, Y^k\right] = p\left[x(k)|\zeta_k^s, y(k), Y^{k-1}\right] = \frac{p\left[y(k)|\zeta_k^s, x(k), Y^{k-1}\right] p\left[x(k)|\zeta_k^s, Y^{k-1}\right]}{p\left[y(k)|\zeta_k^s, Y^{k-1}\right]}. \tag{57}$$

This leads to

$$p\left[x(k)|\zeta_k^s, y(k), Y^{k-1}\right] = \sum_{i=0}^{M-1} p\left[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}\right] P\left(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}\right). \tag{58}$$

Equation (58) proves that the term $p\left[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}\right]$ is the Gaussian apriori probability density of the state vector corresponding to the values $\left\{\zeta_{k-1}^i, \zeta_k^s\right\}$ between two adjacent moments. Thus

$$\begin{aligned} p[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}] = N\Big[&x(k); F(k/k-1, \zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1), \\ &F(k/k-1, \zeta_k^s)P^i(k-1/k-1)F(k/k-1, \zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T\Big]. \end{aligned} \tag{59}$$

If it is assumed that the probability density $p\left[x(k)|\zeta_k^s, Y^{k-1}\right]$of the state vector corresponding to the value $\zeta_k^s$ is Gaussian, then taking into account (58)–(59) we obtain

$$\begin{aligned} N[x(k); \stackrel{\frown}{x}^s (k/k-1), \stackrel{\frown}{P}^s (k/k-1)] \simeq N\Bigg[&x(k); \sum_{i=0}^{M-1} F(\zeta_k^s) \stackrel{\frown}{x}^i (k/k-1) P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \\ &\sum_{i=0}^{M-1} F(\zeta_k^s) \stackrel{\frown}{P}^i (k-1/k-1) F(\zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T + [F(\zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1) \\ &- \hat{x}^s(k/k-1)][[F(\zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1) - \hat{x}^s(k/k-1)]^T] P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1})\Bigg], \end{aligned} \tag{60}$$

equating both sides of this expression, we obtain

$$\widehat{x}^{s}\ (k/k-1) = F(\zeta_k^s) \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1) P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \tag{61}$$

or

$$\widehat{x}^{s}\ (k/k-1) = F(\zeta_k^s)\hat{x}^{0s}(k-1/k-1), \tag{62}$$

where

$$\hat{x}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1) P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \tag{63}$$

and

$$\hat{P}^s(k/k-1) = \sum_{i=0}^{M-1} F(\zeta_k^s)\ \widehat{P}^{i}\ (k-1/k-1)F(\zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T +$$
$$\left[F(\zeta_k^s)\ \widehat{x}^{i}\ (k-1/k-1) - \hat{x}^s(k/k-1)\right]\left[\left[F(\zeta_k^s)\ \widehat{x}^{i}\ (k-1/k-1) - \hat{x}^s(k/k-1)\right]^T\right]$$
$$P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}). \tag{64}$$

Substituting (63) into (64), we obtain

$$[\hat{P}^s(k/k-1) = F(\zeta_k^s)\hat{P}^{0s}(k/k-1)F(\zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T, \tag{65}$$

where

$$\hat{P}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1})...$$
$$\left[\hat{P}^i(k-1/k-1) + \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right) \times \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right)^T\right] \frac{\partial^2\Omega}{\partial v^2}, \tag{66}$$

and

$$[P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}) = \frac{P(\zeta_k^s|\zeta_{k-1}^i)P(\zeta_{k-1}^i|Y^{k-1})}{\sum_{i=0}^{M-1} P(\zeta_k^s|\zeta_{k-1}^i)P(\zeta_{k-1}^i|Y^{k-1})}. \tag{67}$$

The output estimates is calculated from equations (59) and (60)

$$[\widehat{x}\ (k/k) = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\cdots\sum_{s=0}^{M-1} \hat{x}^{ij\ldots s}(k/k)P(\zeta_1^i, \zeta_2^j, \cdots, \zeta_k^s|Y^k), \tag{68}$$

and its covariance matrix:

$$\hat{P}(k/k) = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\cdot\cdot$$
$$\cdot\sum_{s=0}^{M-1}\left[\hat{P}^{ij\ldots s}(k/k) + \left[\hat{x}^{ij\ldots s}(k/k) - \hat{x}(k/k)\right]\left[\hat{x}^{ij\ldots s}(k/k) - \hat{x}(k/k)\right]^T\right] P(\zeta_1^i, \zeta_2^j, \cdots, \zeta_k^s|Y^{k)}, \tag{69}$$

where estimation

$$\hat{x}^{ij\dots s}(k/k) = \hat{x}^{\zeta_1^i,\zeta_2^j,\dots,\zeta_k^s}(k/k), \tag{70}$$

and covariance matrix

$$\hat{P}^{ij\dots s}(k/k) = \hat{P}^{\zeta_1^i,\zeta_2^j,\dots,\zeta_k^s}(k/k), \tag{71}$$

the corresponding ${l_k}^{th}$ branch of the algorithm M parallel Kalman filters Algorithm. The algorithm consists of M parallel Kalman filters, each of them is tuned to a single value $\zeta^i$, as described below, Figure 3 shows the realization diagram of the algorithm.

1. Initialization:

   (a) Set the initial value of the M – estimates of the state vector $\hat{x}^i(k/k)$:

$$\hat{x}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1)P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \tag{72}$$

   (b) Set the initial value of the M – covariance $\hat{P}^i(k-1/k-1)$:

$$\hat{P}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1})$$
$$\left[\hat{P}^i(k-1/k-1) + \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right) \times \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right)^T\right]$$
$$\frac{\partial^2\Omega}{\partial v^2}, \tag{73}$$

2. Estimation process:

   (a) Extrapolation of the state vectors:

$$\widehat{x}^s(k/k-1) = F(\zeta_k^s)\hat{x}^{0s}(k-1/k-1), \tag{74}$$

   (b) Extrapolation of correlation matrices:

$$\hat{P}^s(k/k-1) = F(\zeta_k^s)\hat{P}^{0s}(k/k-1)F(\zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T. \tag{75}$$

   (c) Coefficient gain:

$$K^s(k) = \hat{P}^s(k/k-1)H\left[H\hat{P}^s(k/k-1)H^T + R(k)\right]^{-1}. \tag{76}$$

   (d) Estimation:

$$\hat{x}^s(k/k) = \hat{x}^s(k/k-1) + K^s(k)\left[y(k) - H\hat{x}^s(k/k-1)\right]. \tag{77}$$

   (e) Covariance filtration

$$P^s(k/k) = \left[I - K^s(k)H(k)\right]\hat{P}^s(k/k-1). \tag{78}$$

3. Calculate the conditional probability density of observation, corresponding to the states of $\zeta_k^s$, we obtain

$$\begin{aligned} L^s(k) &= p[y(k)|\zeta_k^s, Y^{k-1}] = p\left[y(k)|\zeta_k^s, H\hat{x}^s(k/k-1), \hat{P}^s(k/k-1)\right] \\ &= \left[(2\pi)^s \det(s^s)(k)\right]^{-1/2} \exp\left[-0.5\vartheta^s(k)^T s^s(k)^{-1}\vartheta^j(k)\right] \end{aligned}, \tag{79}$$

where

$$\vartheta^s(k) = y(k) - H(k)\widehat{x}^s(k/k-1), \tag{80}$$

and

$$s^s(k) = H\hat{P}^s(k/k-1)H^T(k) + R(k). \tag{81}$$

4. Calculate the a posteriori probabilities of the state $\zeta_k^s$, according to the Bayes formula

$$P\left[\zeta_k^s|Y^k\right] = P\left[\zeta_k^s|y(k), Y^{k-1}\right] = \frac{p\left[y(k)|\zeta_k^s, Y^{k-1}\right] P\left[\zeta_k^s|Y^{k-1}\right]}{\sum_{s=0}^{M-1} p\left[y(k)|\zeta_k^s, Y^{k-1}\right] P\left[\zeta_k^s|Y^{k-1}\right]}, \tag{82}$$

where

$$P\left[\zeta_k^s|Y^k\right] = \frac{L^s(k)\sum_{i=o}^{M-1} p_{is} P\left[\zeta_{k-1}^i|Y^{k-1}\right]}{\sum_{s=0}^{M-1} L^s(k)\sum_{i=o}^{M-1} p_{is} P\left[\zeta_{k-1}^i|Y^{k-1}\right]}. \tag{83}$$

5. The output estimate of the target state vector and its covariance were obtained on the basis of a linear combination of a posteriori probabilities

$$\widehat{x}\ (k/k) = \sum_{s=0}^{M-1} \hat{x}^s(k/k) P(\zeta_k^s|Y^k), \tag{84}$$

$$\hat{P}(k/k) = \sum_{s=0}^{M-1} \left[\hat{P}^s(k/k) + [\hat{x}^s(k/k) - \hat{x}(k/k)]\,[\hat{x}^s(k/k) - \hat{x}(k/k)]^T\right] P\left[\zeta_k^s|Y^k\right]. \tag{85}$$

## 5 Simulation Results

The proposed algorithm was applied to the phosphorus furnace type RKZ-80F. The system has the following specifications: a linear model of indirect control of electro-thermal processes in a three-electrode. This model is based on the well-known band structure of electric furnace baths and takes into account its geometric symmetry with respect to the three electrodes. Therefore, all the variables used in the model below will have indices i = 1,2,3; they relate to one of the three electrodes or the corresponding near-electrode region. The index 0 of the variable will indicate its relation to the whole bath of the furnace as a whole. The vector of state space $x(k) \in R^n$ ($n = 7$) has the following states:

$$x^T(k) = \begin{bmatrix} x_{len}^1(k) & x_{vol}^1(k) & x_{len}^2(k) & x_{vol}^2(k) & x_{len}^3(k) & x_{vol}^3(k) & x_{high}^0(k) \end{bmatrix},$$

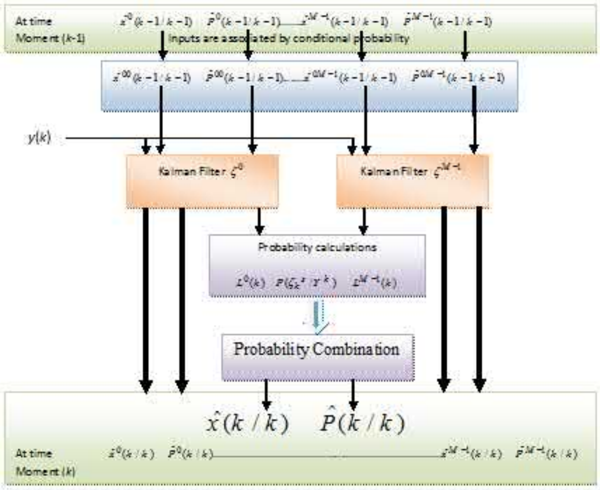


**Figure 3**: Proposed algorithm implementation.

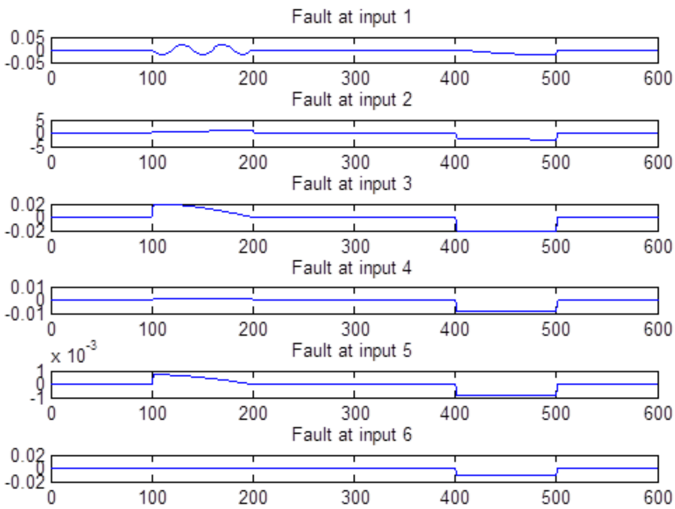


**Figure 4**: Fault signals forms.

where $x_{len}^{i}(k)$ is the length of the i-th electrode, m; $x_{vol}^{i}(k)$ is the volume of the crucible of the i−th near-electrode region, $m^3$; $x_{high}^{0}(k)$ is the height of the total working (carbon) zone in the furnace, m. The control vector $u(k) \in R^p$ $(p = 7)$ has the following structure:

$$u^T(k) = \begin{bmatrix} u_{byp}^{1}(k) & u_{pa}^{1}(k) & u_{bup}^{2}(k) & u_{pa}^{2}(k) & u_{len}^{3}(k) & u_{byp}^{3}(k) & u_{c}^{0}(k) \end{bmatrix},$$

$$u^T(k) = \left[u_{?5@}^{1}(k)\ u_{P}^{1}(k)\ u_{?5@}^{2}(k)\ u_{P}^{2}(k)\ u_{?5@}^{3}(k)\ u_{P}^{3}(k)\ u_{C}^{0}(k)\right].$$

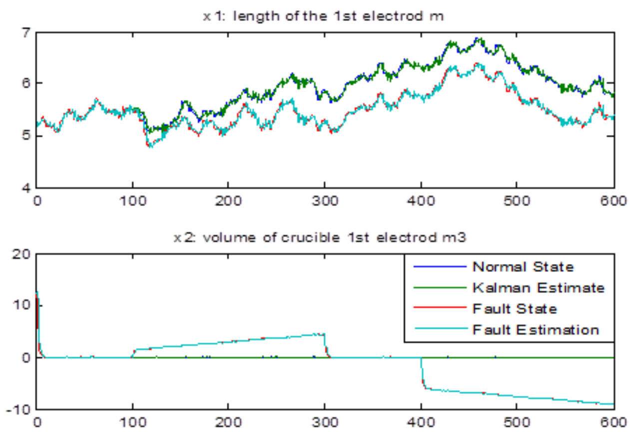


**Figure 5**: Estimations of the length and volume of the first electrode.

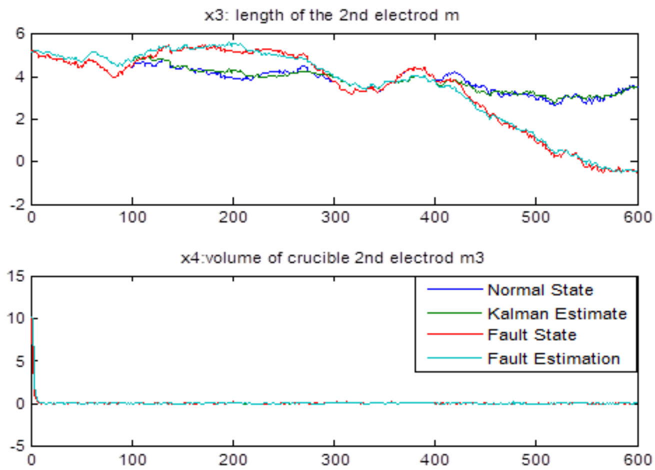


**Figure 6**: Estimations of the length and volume of the second electrode.

$u^i_{byp}(k$ is the bypass of the ith electrode, m;
$u^i_{pa}(k)$ is the average useful active power, in the i-th near-electrode region, MW;
$u^0_c(k)$ is the the amount of carbon entering the furnace with the charge, T.
The vector of observation $y(k) \in R^m$ $(m = 15)$ has the following structure:

$$y^T(k) = \left[y^1_1(k)y^1_2(k)y^1_3(k)y^1_4(k)y^2_1(k)y^2_2(k)y^2_3(k)y^2_4(k)y^3_1(k)y^3_2(k)y^3_3(k)y^3_4(k)y^0_1(k)y^0_2(k)y^0_3(k)\right],$$

where
$y^i_1(k)$ is the position of the electrode holder of the i-th electrode relative to the slag tap, m;
$y^i_2(k)$ is the active resistance of the i-th phase, m$\Omega$;
$y^i_3(k)$ is the temperature under the arch at the i-th electrode,$^\circ C$;

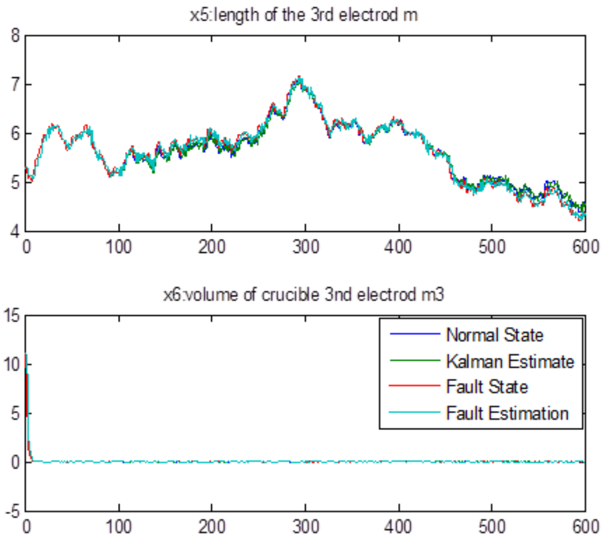


**Figure 7**: Estimations of the length and volume of the third electrode.

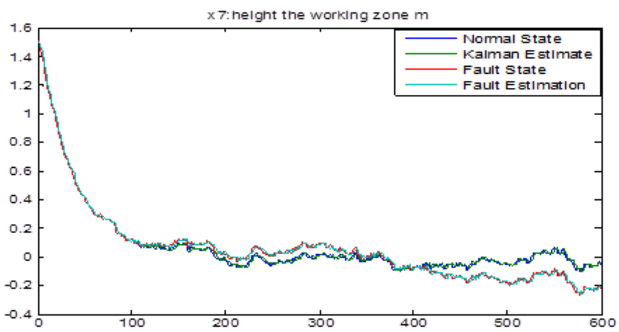


**Figure 8**: Estimations of the high of the working area.

$y_4^i(k)$ is the gathering of charge for the i-th electrode, T;
$y_1^0(k)$ is the the average temperature under the roof of the entire furnace, $°C$;
$y_2^0(k)$ is the current furnace capacity, T;
$y_3^0(k)$ is the relative recoverability of the product, (%$P_2O_5$ in the batch /% $P_2O_5$ in the slag).

The matrices parameters $\Phi$, $\Gamma$ and $H$ are as follows

$$\Phi = \begin{bmatrix} \Phi_1 & 0 & 0 & f_{01} \\ 0 & \Phi_1 & 0 & f_{01} \\ 0 & 0 & \Phi_1 & f_{01} \\ f_{10} & f_{10} & f_{10} & f_{00} \end{bmatrix}, \Gamma = \begin{bmatrix} \Gamma_1 & 0 & 0 & 0 \\ 0 & \Gamma_1 & 0 & 0 \\ 0 & 0 & \Gamma_1 & 0 \\ 0 & 0 & 0 & g_{00} \end{bmatrix}, H = \begin{bmatrix} H_1 & 0 & 0 & h_{01} \\ 0 & H_1 & 0 & h_{01} \\ 0 & 0 & H_1 & h_{01} \\ H_{10} & H_{10} & H_{10} & h_{00} \end{bmatrix},$$

where $\Phi_1 = \begin{bmatrix} a_1 & -a_2 \\ 0 & a_4 \end{bmatrix}$, $f_{01} = \begin{bmatrix} \frac{a_3}{a_5} \end{bmatrix}$, $f_{10} = \begin{bmatrix} 0 & a_6 \end{bmatrix}$, $f_{00} = a_7, g_{00} = \beta_2$, $\Gamma_1 = \begin{bmatrix} 1 & 0 \\ 0 & \beta_1 \end{bmatrix}$, $H_1 = \begin{bmatrix} 1 & 0 \\ \gamma_2 & 0 \\ 0 & \gamma_4 \\ 0 & \gamma_6 \end{bmatrix}$, $h_{01} = \begin{bmatrix} \gamma_1 \\ \gamma_3 \\ \gamma_5 \\ 0 \end{bmatrix}$, $H_{10} = \begin{bmatrix} 0 & \gamma_7 \\ 0 & \gamma_9 \\ 0 & 0 \end{bmatrix}$, $h_{00} = \begin{bmatrix} \gamma_8 \\ 0 \\ \gamma_{10} \end{bmatrix}$,

$\theta = [a_1, ..., a_7, \beta_1, \beta_2, \gamma_1, ..., \gamma_{10}]$. The parameters values are described in Table below:

| No. | Parameter | Physical dimension | Admissible values | Recommended initial approximation |
|---|---|---|---|---|
| 1 | $\alpha_1$ | $-$ | $0.95 - 1.0$ | 0.999 |
| 2 | $\alpha_2$ | $1/M^2$ | $0.0002 - 0.002$ | 0.001 |
| 3 | $\alpha_3$ | $-$ | $0.0001 - 0.002$ | 0.001 |
| 4 | $\alpha_4$ | $-$ | $0.0 - 1.0$ | 0.5 |
| 5 | $\alpha_5$ | $M^2$ | $0.0 - 1.0$ | 0.5 |
| 6 | $\alpha_6$ | $1/M^2$ | $0.0001 - 0.005$ | 0.001 |
| 7 | $\alpha_7$ | $-$ | $0.8 - 1.0$ | 0.9 |
| 8 | $\beta_1$ | $M^3/MW$ | $0.2 - 0.4$ | 0.28 |
| 9 | $\beta_2$ | $M/T$ | $0.004 - 0.0045$ | 0.0042 |
| 10 | $\gamma_1$ | - | $0.5 - 1.0$ | 0.75 |
| 11 | $\gamma_2$ | $m\Omega/M$ | $0.05 - 0.5$ | 0.1 |
| 12 | $\gamma_3$ | $m\Omega/M$ | $0.2 - 0.8$ | 0.5 |
| 13 | $\gamma_4$ | $^\circ C/M^3$ | $1.0 - 10.0$ | 5.0 |
| 14 | $\gamma_5$ | $^\circ C/M$ | $300 - 500$ | 400 |
| 15 | $\gamma_6$ | $t/M^3$ | $1.0 - 2.0$ | 1.5 |
| 16 | $\gamma_7$ | $^\circ C/M^3$ | $1.0 - 10.0$ | 5 |
| 17 | $\gamma_8$ | $^\circ C/M$ | $250 - 400$ | 300 |
| 18 | $\gamma_9$ | $T/M^3$ | $0.05 - 0.2$ | 0.1 |
| 19 | $\gamma_{10}$ | $1/M$ | $20 - 40$ | 30 |

The initial values of states are: x = $[5.25\ 25\ 5.18\ 20\ 5.36\ 22\ 1.5]^T$; the number of samples is 600 samples; the input vector is generated as random signal in the interval [-0.05, 0.05]; different faults scenarios are added to the input vectors. The simulation results are shown in figures below, in which Figure 4 represents the fault signals forms. Figure 5 shows that the algorithm starts tracking in the abnormal state in sample time number 100 for both states estimations and could not return to the normal operation even the input signal return without fault. Figures 6 to 8 show the switching between modes to track the changes in states directly after the abnormal operation is detected to start the fault tracking.

## 6 Conclusion

In this paper we proposed a multi-operational mode based on Bayes approach to select the best Kalman filter estimator. The estimation was focused on internal state estimation during the fault operation which may be caused by many different resources. The simulation fault testing inputs signals failure, which were lost by composed signals. The fault scenarios started at time sample 100 with different shapes and values for each input and ended at time sample 200. Another faults signals were inserted at time sample 400 and ended at time sample 500. From the simulation results of states estimations, we noted that the estimation of the lengths of the electrodes were very satisfied in which most of estimations values were very near from the estimations in normal operation. The volume of the first electrode during the fault was not estimated correctly where a negative values was estimated. While the volumes of the second and third electrodes during the fault was estimated correctly. Finally, estimations of the high of the working area may be accepted with some error percentage. As a result the proposed algorithm provide a good results with some acceptable range of error during the faults operations.

## References


[1] Maiying Zhong, Hao Ye, Steven X. Ding, Guizeng Wang, and Zhou Donghua. Fault detection filter for linear time-delay systems. *Nonlinear Dynamics and Systems Theory* **5** (3) (2005) 273–284.

[2] Wu, H. and Noe, F. A flat Dirichlet process switching model for Bayesian estimation of hybrid systems. *IEEE Transaction on Automatic Control* **15** (1) (2002) 10–17.

[3] Ackerson, G. and Fu, K. On state estimation in switching environments. *Procedia Computer Science* (4) (2011) 1393–1402.

[4] Blom, H. and Bar-Shalom, Y. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transaction on Automatic Control* (17) (2001) 362–369.

[5] Minka, T. Expectation propagation for approximate Bayesian inference. *Proceedings of Uncertainty in Artificial Intelligence* **33** (8) (2002) 780–783.

[6] Bar-Shalom, Y., Kirubasrajan, T. and Li, X.R. *Estimation with Applications to Tracking and Navigation*. Wiley, New York, 2002.

[7] Evgenia Suzdaleva and Ivan Nagy. Recursive state estimation for hybrid systems. *Applied Mathematical Modelling* (36) (2012) 1347–1358.

[8] F. van der Heijden, Duin, R.P.W., D. de Ridder and Tax, D.M.J. *Classification, Parameter Estimation and State Estimation*. John Wiley & Sons, 2004.

[9] Bar-Shalom, Y. and Fortmann, T.E. *Tracking Data Association*. Academic Press, 1988.

[10] Ali Hussein Hasan and Grachev, A.N. Adaptive $\alpha - \beta-$filter for target tracking using real time genetic algorithm. *Journal of Electrical and Control Engineering* **3** (4) (203) 32–38.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 392–408

# A Self-Diffusion Mathematical Model to Describe the Toxin Effect on the Zooplankton-Phytoplankton Dynamics

Hamidou Ouedraogo, Wendkouni Ouedraogo and Boureima Sangaré *

*University Nazi Boni, Department of Mathematics and Informatics, UFR/ST, 01 BP 1091 Bobo Dsso 01, Burkina Faso*

Received: June 3, 2018; Revised: October 11, 2018

**Abstract:** The main goal of this work is the mathematical formulation, the analysis and the numerical simulation of a prey-predator model by taking into account the toxin produced by the phytoplankton species. The mathematical study of the model leads us to have an idea on the existence of solution, the existence of equilibria and the stability of the stationary equilibria. These results are obtained through the principle of comparison. Finally, the numerical simulations in two-dimensional allowed us to establish the formation of spatial patterns and a threshold of release of the toxin, above which we talk about the phytoplankton blooms.

**Keywords:** *toxin effect; populations dynamics; predator-prey model; reaction-diffusion; pattern formation.*

**Mathematics Subject Classification (2010):** 65L12, 65M20, 65N40.

## 1 Introduction

Ecology and harmful toxic release in marine environment are major fields of study in their own right, but there are some common features of these systems. It is interesting and important from biological viewpoints to study ecological systems under the influence of the toxic substance release factors. However, this goal remains difficult to attain due to the complexity of natural systems, especially in the aquatic environment where many processes of all types interact with living organisms. The fundamental basis of all aquatic food chains is plankton, and phytoplankton in particular occupies the first trophic level and the fluctuations in its abundance determine the production of a whole marine

* Corresponding author: mailto:mazou1979@yahoo.fr

biological output. The dynamics of rapid (massive) increase or almost equally decrease of phytoplankton populations is a common feature in marine plankton ecology and is known as bloom. This phenomenon can occur in a matter of days and can disappear just as rapidly.

Several authors have argued that there has been a global increase in harmful phytoplankton blooms in recent decades, see [7, 14, 24]. The rapid massive growth of phytoplankton is, generally, caused by high nutrient levels and favorable conditions (water temperature, salinity levels, etc.). Herbivore grazing takes an important role in the bloom dynamics, see [7,24,26]. Toxic substances produced by phytoplankton species reduce the growth of zooplankton by decreasing grazing pressure and this is one of the important common phenomena in plankton ecology, see [10, 17, 22].

Within the broad perspective drawn above, the present paper explores and compares the coupled dynamics of phytoplankton and zooplankton in a number of mathematical models. The system of phytoplankton-zooplankton has attracted considerable attention from various fields of research, see [10, 21, 25]. It is an important issue in mathematical ecology. The literature abounds in models focusing on various aspects of the problem. Recently, the attention has been focused on the role of the space in explaining heterogeneity and the distribution of the species and the influence of the spatial structure on their abundance, [10, 17, 29]. However, the very question of the interactions between phytoplankton and zooplankton depending on space is far from being fully elucidated.

As part of our work, we will highlight a threshold value of the toxin released by phytoplankton below which the effect of the toxin influences less the dynamics of the zooplankton-fish system. The proposed model consists of two interactive component: zooplankton and toxin-producing phytoplankton that reduces the growth of zooplankton population. The model studied here is of the reaction-diffusion type describing the dynamics of the phytoplankton-zooplankton system in the sense of the works of F. Courchamp [8, 28].

The paper is organized as follows. As far as Section 3 is concerned, we will establish mathematical results such as the existence of a solution, stability of equilibria, persistence, relating to the constructed model in Section 2. Section 4 will be devoted to numerical experiments to illustrate the mathematical results. Finally, Section 5 is devoted to the conclusion and perspectives.

## 2 Mathematical Model

In this section, we propose a model to describe the dynamics of the phytoplankton-zooplankton system in the presence of toxin. We begin our modeling by a general model describing the dynamics of the prey-predator system, based on the equations with ordinary derivatives. And then we transform this model into a model of reaction-diffusion type while remaining in the logic of the work of F. Courchamp [8] and Bendahmane et al [5]. The aim is to take into account the effect of the toxin on the zooplankton-phytoplankton dynamics.

### 2.1 Original model formulation

Let $P$ be the density of the prey population and $Z$ be the density of the predator population. According to [5, 8, 18, 28], the general model at any time $T > 0$ is written as follows:

$$\begin{cases} \dfrac{dP}{dT} = \phi_1(P) - g_2(P,Z)Z, \\ \dfrac{dZ}{dT} = g_3(P,Z)Z - g_4(P,Z)Z, \end{cases} \tag{1}$$

where

- $\phi_1, g_2, g_3, g_4$ are positive functions and $C^\infty$,
- $\phi_1(P)$ is the growth function of the prey population,
- $g_2(P,Z)$ is the amount of prey consumed by a predator per time unit,
- $g_3(P,Z)$ represents the rate of conversion of the prey into predator,
- $g_4(P,Z)$ is the predator mortality rate due to harmful prey consumption.

We continue our modeling by fixing the expressions of the functions intervening in the model (1), see [5, 8]. The dynamics of the system can be represented by the following figure:

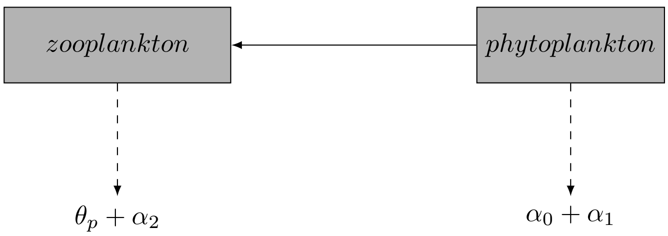


**Figure 1**: The compartmental model of the zooplankton-phytoplankton system.

According to Figure 1, at any time $T > 0$, the dynamics of the phytoplankton (prey)-zooplankton (predator) system is governed by the following ODE system:

$$\begin{cases} \dfrac{dP}{dT} = r_p P - \alpha_0 P^2 - \dfrac{\alpha_1 Z}{P + \gamma_1} P, \quad P(0) = P_0 \geq 0, \\ \dfrac{dZ}{dT} = r_z Z - \theta_p Z - \dfrac{\alpha_2 Z^2}{P + \gamma_2}, \quad Z(0) = Z_0 \geq 0, \end{cases} \tag{2}$$

where

- $r_p$ denotes the phytoplankton growth rates,
- $\alpha_0$ denotes the mortality rate due to competition between the individuals of the phytoplankton population,
- $\alpha_1$ is the maximum value that the reduction rate per individual phytoplankton can reach,
- $\alpha_2$ is the maximum value that the reduction rate per individual zooplankton can reach,

- $r_z$ denotes the zooplankton growth rates,
- $\theta_p$ is the rate of toxic phytoplankton consumed by zooplankton,
- $\gamma_1$ is the protection of prey $P$ from the environment,
- $\gamma_2$ is the protection of predator $Z$ from the environment.

### 2.2 A spatially structured model

Here, our aim is to introduce the notion of spatial structuring in the model. By considering the relationship between the climate and the diffusion of species and the fact of the existence of diffusion in population, system (1) is developed into a spatial system with diffusion. We expect to explore the effect of climate change on the plankton population by studying the spatial dynamics of the diffusion system. We will introduce the concept of spatial structure in the model, that is to say that the population densities depend now on the time and space. Diffusion models are a simple and reasonable choice for modeling dispersion of populations on a spatial domain, see [3, 12, 18, 25]. Indeed, let $\delta_0(x)$ and $\delta_1(x)$ be respectively the diffusion terms of $P$ and $Z$. Based on the results established in [5, 8, 28], the reaction-diffusion model associated with the model (1) can be modeled for $x \in \Omega$ as follows:

$$\begin{cases} \partial_T P - div(\delta_0(x)\nabla P) = \phi_1(P) - g_2(P,Z)P, \\ \partial_T Z - div(\delta_1(x)\nabla Z) = g_3(P,Z)Z - g_4(P,Z)Z, \end{cases} \tag{3}$$

where $\Omega \subset \mathbb{R}^n (n \geq 1)$ is the spatial domain in which species occur. We consider the zero-flux boundary condition

$$\delta_i(x)\nabla Q(x,t).\nu(x) = 0, \qquad i = 0,1, \qquad x \in \partial\Omega, \qquad T > 0$$

for $Q = P, Z$, where $\nu$ is the unit normal vector to $\partial\Omega$ on $\Omega$, and the nonnegative initial and bounded conditions

$$Q(x,0) = Q_0(x) > 0, \qquad Q = P, Z \qquad x \in \Omega.$$

We make the following assumptions:

$(H_1)$ : all demographic parameters of the system (2) are positive constants,

$(H_2)$ : the diffusion coefficients of the system (3) are independent of the spatial variable.

By considering $\delta_0(x) = \delta_0$, $\delta_1(x) = \delta_1$, and taking into account the hypotheses $(H_1)$ and $(H_2)$, the model (3) obtained previously is written as:

$$\begin{cases} \dfrac{\partial P}{\partial T} = \left(r_p - \alpha_0 P - \dfrac{\alpha_1 Z}{P+\gamma_1}\right)P + \delta_0 \Delta P, \\ \dfrac{\partial Z}{\partial T} = \left(r_z - \theta_p - \dfrac{\alpha_2 Z}{P+\gamma_2}\right)Z + \delta_1 \Delta Z, \\ Q(x,0) = Q_0(x) > 0, \quad Q = P, Z, \quad x \in \Omega. \end{cases} \tag{4}$$

## 3 Mathematical Results

In this section, we aim to establish the mathematical results of the system (4). The mathematical results are based on the works [5, 6, 25].

### 3.1 Reduction of model parameters

To simplify the writing, we will make changes of variables in the following way:

$$r = r_z - \theta_p, \qquad t = r_p T, \qquad U_1(t) = \frac{\alpha_0}{r_p} P(T), \qquad U_2(t) = \frac{\alpha_0 \alpha_2}{r_p r} Z(T),$$
$$a = \frac{r\alpha_1}{r_p \alpha_2}, \qquad b = \frac{r}{r_p}\ , \qquad d_1 = \frac{\alpha_0 \gamma_1}{r_p}, \qquad d_2 = \frac{\alpha_0 \gamma_2}{r_p},$$
$$x = X\left(\frac{r_p}{\delta_0}\right)^{\frac{1}{2}}, \qquad y = Y\left(\frac{r_p}{\delta_0}\right)^{\frac{1}{2}}, \qquad d_z = \frac{\delta_1}{\delta_0}.$$

Thus, the systems (2) and (4) can be written respectively as follows:

$$\begin{cases} \dfrac{dE_1}{dt} = \left(1 - E_1 - \dfrac{aE_2}{E_1 + d_1}\right) E_1 = f(E_1, E_2), \\ \dfrac{dE_2}{dt} = b\left(1 - \dfrac{E_2}{E_1 + d_2}\right) E_2 = g(E_1, E_2), \end{cases} \tag{5}$$

and

$$\begin{cases} \dfrac{\partial U_1}{\partial t} = \left(1 - U_1 - \dfrac{aU_2}{U_1 + d_1}\right) U_1 + \Delta U_1 = f(U_1, U_2) + \Delta U_1 \\ \dfrac{\partial U_2}{\partial t} = b\left(1 - \dfrac{U_2}{U_1 + d_2}\right) U_2 + d_z \triangle U_2 = g(U_1, U_2) + d_z \Delta U_2. \end{cases} \tag{6}$$

### 3.2 Existence and boundedness of solution

Before stating the boundedness of the solution, we give the local existence of the solution. The following theorem ensures the existence and uniqueness of the local solution of the system (4).

**Theorem 3.1** *[1, 2] The system* (4) *has a unique local solution* $(U_1(.,t), U_2(.,t))$ *under the condition* $0 \le t < T_{max}$*, where* $T_{max}$ *depends on nonnegative initial data* $U_{01}(x)$ *and* $U_{02}(x)$*.*

The following theorem ensures the global solution existence for the system (6).

**Theorem 3.2** *For any regular positive functions* $U_{01}(x) \le 1$ *and* $U_{02}(x)$*, the system* (6) *admits a global solution* $(U_1(.,t), U_2(.,t))$ *for any* $t > 0$*.*

***Proof:*** Indeed:

- on the one hand, we have $U_1(x,t) \ge 0$ and $U_2(x,t) \ge 0$ because 0 is the sub-solution of each equation of the system (6).
- on the other hand, $U_1$ satisfies the following problem:

$$\begin{cases} \dfrac{\partial U_1(x,t)}{\partial t} \le U_1(1 - U_1) + \Delta U_1, \\ \dfrac{\partial U_1}{\partial \nu} = 0\ \ ,\ t > 0, \\ U_1(x,0) = U_{01}(x) \le \ \ U_{01} \equiv \ \ max_{\overline{\Omega}}\ \ U_{01}(x), x \in \Omega. \end{cases} \tag{7}$$

According to the principle of comparison, we have $U_1(x,t) \leq U(t) \leq 1$, where $U(t) = \dfrac{U_0}{U_0 + (1-U_0)e^{-t}}$ is the solution of the problem

$$\frac{\partial U}{\partial t} = U(1-U), \quad U(0) = U_0 \leq 1. \tag{8}$$

In the same order $U_2$ satisfies

$$\frac{\partial U_2}{\partial t} = b\left(1 - \frac{U_2}{U_1 + d_2}\right)U_2 + d_z \Delta U_2,$$

and we obtain the following inequality:

$$\frac{\partial U_2}{\partial t} \leq \frac{dE_2}{dt},$$

where $E_2$ is a solution of the second equation of the system (5) with the initial condition $E_2(0) = max_{\overline{\Omega}}\ U_{02}(x)$.

$$\frac{\partial U_2}{\partial t} \leq \frac{dE_2}{dt} + \frac{dE_1}{dt}.$$

Let us denote by $\sigma = E_2 + E_1$, we deduce that

$$\frac{\partial U_2}{\partial t} \leq \frac{d\sigma}{dt} \leq \frac{5}{4} + \frac{(1+b)^2(1+d_2)}{4b} - \sigma.$$

Using the Gronwall lemma, see [10, 18, 29], we deduce that $U_2 \leq \dfrac{5}{4} + \dfrac{(1+b)^2(1+d_2)}{4b}$. Thus, the solutions $U_1$ and $U_2$ are bounded.

The following theorem ensures the boundedness of the system (6).

**Theorem 3.3** *The domain $\mathbb{R}^+ \times \mathbb{R}^+$ is positively invariant for the system (6). Furthermore, any solution of the system (6) whose initial condition is in $\mathbb{R}^+ \times \mathbb{R}^+$ converges to the set defined by $S_1 = [0,1] \times \left[0, \dfrac{5}{4} + \dfrac{(1+b)^2(1+d_2)}{4b}\right]$.*

***Proof:*** For the initial condition $(U_{01}(x), U_{02}(x))$ of the system (6), we have

$$0 \leq U_1 \leq E_1, \quad E_1(0) = max_{\overline{\Omega}}\ U_{01}(x),$$

$$0 \leq U_2 \leq E_2, \quad E_2(0) = max_{\overline{\Omega}}\ U_{02}(x).$$

On the other hand, according to [3, 13], we have

$$\overline{\lim}_{t \longrightarrow +\infty} E_1(t) \leq 1, \quad \overline{\lim}_{t \longrightarrow +\infty}(E_1(t) + E_2(t)) \leq \frac{5}{4} + \frac{(1+b)^2(1+d_2)}{4b}. \qquad \square$$

### 3.3 Analysis of stationary solutions

Now, we study the existence of positive equilibrium states of the system (6). Let us consider the following system:

$$\begin{cases} \left(1 - U_1 - \dfrac{aU_2}{U_1 + d_1}\right) U_1 + \Delta U_1 = 0, & x \in \Omega, \\ b\left(1 - \dfrac{U_2}{U_1 + d_2}\right) U_2 + d_z \Delta U_2 = 0, & x \in \Omega, \\ \dfrac{\partial U_1}{\partial \nu} = \dfrac{\partial U_2}{\partial \nu} = 0. \end{cases} \tag{9}$$

Then, $(U_1, U_2)$ is a positive equilibrium state of the system (6) if it satisfies the system (9).

**Remark 3.1** Let $V_1(x) = (\overline{U}_1(x), \overline{U}_2(x))$ and $V_2(x) = (\underline{U}_1(x), \underline{U}_2(x))$. According to [4], $V_1(x)$ is an over-solution and $V_2(x)$ is a sub-solution of the system (9) if we have

$$\frac{\overline{U}_1}{\partial \nu} \geq 0 \geq \frac{\underline{U}_1}{\partial \nu} \quad on \;\; \partial\, \Omega, \qquad \frac{\overline{U}_2}{\partial \nu} \geq 0 \geq \frac{\underline{U}_2}{\partial \nu} \quad on \;\; \partial\, \Omega,$$

and

$$-\Delta \overline{U}_1 - \overline{U}_1 \left(1 - \overline{U}_1 - \frac{a\underline{U}_2}{\overline{U}_1 + d_1}\right) \geq 0 \geq -\Delta \underline{U}_1 - \underline{U}_1 \left(1 - \underline{U}_1 - \frac{a\overline{U}_2}{\underline{U}_1 + d_2}\right), \tag{10}$$

$$-\Delta \overline{U}_2 - b\overline{U}_2 \left(1 - \frac{U_2}{\overline{U}_1 + d_2}\right) \geq 0 \geq -\Delta \underline{U}_2 - b\underline{U}_2 \left(1 - \frac{U_2}{\underline{U}_1 + d_2}\right). \tag{11}$$

Let us consider the following conditions

$$a < 1, \quad d_1 - a + 1 - A > 0, \tag{12}$$

where $A$ is an over-solution of the second equation of the system (9).

**Theorem 3.4** *[6,12] If the conditions* (12) *are satisfied, then the system* (9) *admits at least one positive solution* $(U_1(x), U_2(x))$.

***Proof:*** We write the system (9) as follows:

$$\begin{cases} -\Delta U_1 = \left(1 - U_1 + \dfrac{aU_2}{U_1 + d_1}\right) U_1 = f(U_1, U_2), & x \in \Omega, \\ -d_z \Delta U_2 = \left(1 - \dfrac{U_2}{U_1 + d_2}\right) U_2 = g(U_1, U_2) & x \in \Omega, \\ \dfrac{\partial U_1}{\partial \nu} = \dfrac{\partial U_2}{\partial \nu} = 0. \end{cases} \tag{13}$$

If $U_1 \geq 0$, $U_2 \geq 0$, we obtain

$$\frac{\partial f}{\partial U_2} = \frac{-aU_1}{U_1 + d_1} \leq 0, \quad \frac{\partial g}{\partial U_2} = \frac{bU_2^2}{(U_1 + d_2)^2} \geq 0.$$

This means that the function $f$ is quasi-monotone decreasing and the function $g$ is quasi-monotone increasing. The system (9) is then called a quasi-monotonic mixed system.

We will now construct a pair of over-solution and sub-solution of the system (9) that we denote respectively by $V_1(x) = (\overline{U}_1(x), \overline{U}_2(x))$ and $V_2(x) = (\underline{U}_1(x), \underline{U}_2(x))$.

Let $\overline{U}_1(x) = 1$, then for every $\overline{U}_2 \geq 0$, the first inequality of (10) is satisfied. By fixing $A$ such that $A \geq \frac{5}{4} + \frac{(1+b)^2(1+d_2)}{4b}$ and by considering $\overline{U}_2(x) = A$, the inequality of (11) is satisfied. If we consider that $\overline{U}_2(x) = A$, the second inequality of (10) becomes

$$-\Delta \underline{U}_1 - \underline{U}_1 \left(1 - \underline{U}_1 - \frac{aA}{\underline{U}_1 + d_2}\right) \leq 0.$$

Let $\underline{U}_1(x)$ be the strictly positive solution of the following system:

$$\begin{cases} \Delta \underline{U}_1 - \underline{U}_1 \left(1 - \underline{U}_1 - \dfrac{aA}{\underline{U}_1 + d_2}\right) = 0, \\ \dfrac{\partial \underline{U}_1}{\partial \nu} = 0. \end{cases} \tag{14}$$

We will show that if the conditions $a < 1$ and $d_2 - a + 1 - M > 0$ are satisfied, then the system (14) will admit a positive solution. If $a < 1$ and $d_2 - a + 1 - A > 0$, then one can easily verify that $(1; 1-a)$ is a pair of over-solution and sub-solution of the equation (14). This equation admits a positive solution $\underline{U}_1(x)$ which checks $1 - a \leq \underline{U}_1(x) \leq 1$. Obviously, we have $\overline{U}_1(x) \geq \underline{U}_1(x)$.

We take arbitrarily $\overline{U}_1(x)$, $\underline{U}_1(x)$ and $\overline{U}_2(x)$. If $d_2 - a + 1 - A > 0$, then we can choose $\underline{U}_2(x)$ constant positive and small enough so that the following inequality is satisfied:

$$-\Delta \underline{U}_2 - b\underline{U}_2 \left(1 - \frac{\underline{U}_2}{\underline{U}_1 + d_2}\right) \leq 0.$$

Note that this inequality is satisfied as soon as $0 < \underline{U}_2 < 1 - a - d_2$. Thus we build a pair of over-solution and sub-solution $(\overline{U}_1(x), \overline{U}_2(x))$ and $(\underline{U}_1(x), \underline{U}_2(x))$ of the system (9).

Thus, the system (9) admits at least one solution $(U_1(x), U_2(x))$ which satisfies

$$\underline{U}_1(x) \leq U_1(x) \leq \overline{U}_1(x), \quad \underline{U}_2(x) \leq U_2(x) \leq \overline{U}_2(x). \qquad \square$$

### 3.4 Stability of homogeneous stationary solutions

The following result gives the stationary states and their stability condition for the system (6).

**Proposition 3.1** *[3, 13, 29]*

*(i)* $E_0 = (0, 0)$ *is the trivial state. This equilibrium is unstable.*

*(ii)* $E_1 = (1, 0)$ *is an equilibrium point. This equilibrium is unstable.*

*(iii)* $E_2 = (0, d_2)$ *is an equilibrium point. This state is unstable if* $d_1 > ad_2$ *and stable if* $d_1 < ad_2$

For the proof of the local stability of $E_i$, we consider the eigenvalue problem of the corresponding linearized operator, see [3,4,13,28,29]. In fact, we consider $(U_1(x,t), U_2(x,t))$ the solution of the system (6), then, we have

$$(U_1(x,t), U_2(x,t)) = E_i + W(x,t) = E_i + (w_1(x,t), w_2(x,t)).$$

We will make the following hypothesis:

$$(H_3): \quad a \geq \frac{1}{2} \text{ and } 0 < d_1 < \overline{d}_1 \text{ with } \overline{d}_1 = -(a+1)+\sqrt{\xi},\ \xi = (a+1)^2+2a(1+2a)-1.$$

**Theorem 3.5** *[6, 12, 27, 29] The interior equilibrium point $E_3 = (U_1^*, U_2^*)$ of the system $(6)$ is stable if the hypothesis $(H_3)$ is satisfied.*

With regard to the analysis of the global stability of the interior equilibrium state, we will make the following hypothesis:
$(H_4): \quad 1 \leq d_1 \leq d_2$.

**Theorem 3.6** *Suppose that $(H_4)$ is satisfied, then the equilibrium $(U_1^*, U_2^*)$ of the system $(9)$ is globally asymptotically stable.*

***Proof:*** Let us consider the functions $l$ and $L$ defined by

$$l(U_1,U_2) = \int_{U_1^*}^{U_1} \frac{(\eta - U_1^*)(\eta + d_1)}{a\eta(\eta + d_2)} d\eta + \frac{U_1^* + d_2}{bU_2^*} \int_{U_2^*}^{U_2} \frac{\eta - U_2^*}{\eta} d\eta,$$

$$L(U_1,U_2) = \int_\Omega l(U_1,U_2)dx = \int_\Omega \left( \int_{U_1^*}^{U_1} \frac{(\eta - U_1^*)(\eta + d_1)}{a\eta(\eta + d_2)} d\eta + \frac{U_1^* + d_2}{bU_2^*} \int_{U_2^*}^{U_2} \frac{\eta - U_2^*}{\eta} d\eta \right) dx.$$

Our goal is to show that $L$ is a Lyapunov function, with a negative orbital derivative. For any solution $(U_1, U_2)$ of $(6)$ whose initial condition $(U_{01}(x), U_{02}(x))$ is in the positive quadrant, $L(U_1, U_2)$ is positive. Moreover, $L(U_1, U_2) = 0$ if and only if $(U_1, U_2) = (U_1^*, U_2^*)$. It remains to prove the following inequality $\dfrac{dL}{dt} < 0$.

Indeed,

$$\begin{aligned}
\frac{dL}{dt} &= \int_\Omega \left( \frac{(U_1 - U_1^*)(U_1 + d_1)}{aU_1(U_1 + d_2)} \right) \left( \Delta U_1 + U_1(1 - U_1) - \frac{aU_1U_2}{U_1 + d_1} \right) dx \\
&+ \int_\Omega \frac{U_1^* + d_2}{bU_2^*} \frac{U_2 - U_2^*}{U_2} \left( d_z \Delta U_2 + b\left(1 - \frac{U_2}{U_1 + d_2}\right) U_2 \right) dx \\
&= \int_\Omega \left( \frac{(U_1 - U_1^*)(U_1 + d_1)}{aU_1(U_1 + d_2)} \right) \left( U_1(1 - U_1) - \frac{aU_1U_2}{U_1 + d_1} \right) dx \qquad (15) \\
&+ \int_\Omega \frac{U_1^* + d_2}{bU_2^*} \frac{U_2 - U_2^*}{U_2} bU_2 \left( 1 - \frac{U_2}{U_1 + d_2} \right) dx \\
&+ \int_\Omega \left( \Delta U_1 \frac{(U_1 - U_1^*)(U_1 + d_1)}{aU_1(U_1 + d_2)} + d_z \Delta U_2 \frac{U_1^* + d_2}{bU_2^*} \frac{U_2 - U_2^*}{U_2} \right) dx.
\end{aligned}$$

Let us denote by $T_1$ the first two right terms of the equality (15) and by $T_2$ the last term on the right. After a simple calculation followed by a reduction, $T_1$ becomes:

$$T_1 = -\int_\Omega \left( (U_1 + U_1^* + d_1 - 1)\frac{(U_1 - U_1^*)^2}{a(U_1 + d_2)} + \frac{(U_2 - U_2^*)^2}{U_1 + d_2} \right) dx \tag{16}$$

By the green formula, $T_2$ becomes

$$\begin{aligned} T_2 = &-\int_\Omega \left( \mid \nabla U_1 \mid^2 \frac{d}{dU_1} \left( \frac{(U_1 - U_1^*)(U_1 + d_1)}{aU_1(U_1 + d_2)} \right) \right. \\ &\left. + d_z \frac{U_1^* + d_2}{bU_2^*} \mid \nabla U_2 \mid^2 \frac{d}{dU_2} \left( \frac{U_2 - U_2^*}{U_2} \right) \right) dx \\ = &-\int_\Omega \left( \mid \nabla U_1 \mid^2 \left( \frac{d_2 - d_1 + 1 + U_1^*}{a(U_1 + d_2)^2} + \frac{U_1^* d_1(2U_1 + d_2)}{a(U_1^2 + d_2U_1)^2} \right) \right. \\ &\left. + d_z \frac{U_1^* + d_2}{bU_2^*} \mid \nabla U_2 \mid^2 \frac{U_2^*}{U_2^2} \right) dx \end{aligned} \tag{17}$$

According to the expressions of $T_1$ and $T_2$, we have $\frac{dL}{dt}(U_1, U_2) < 0$.

Therefore, according to the LaSalle's theorem [7], the equilibrium point $(U_1^*, U_2^*)$ is globally asymptotically stable.

## 4 Numerical Results

In this section, we perform extensive numerical simulations of the spatial model system (9) in two dimensional space using the forward finite difference method. The set of parameter values used for the numerical simulation is given in Table 1, see [5, 9, 11, 12]. Here, the system is studied on a spatial domain $\Omega = [0, 50] \times [0, 50]$. It is assumed that the zooplankton and phytoplankton populations are spread over the whole domain at the beginning. These results show that for every strictly positive initial condition, under the assumptions $(H1) - (H4)$, the non-homogeneous equilibrium is always globally asymptotically stable.

| Param | Description | Values | Refs |
|---|---|---|---|
| $r_p$ | the natural growth-rate of phytoplankton | 1.58 | [9, 20] |
| $\delta_0$ | diffusivity coefficient of $P$ | 5 | [9, 20] |
| $\delta_1$ | diffusivity coefficient of $Z$ | 600.5 | [5] |
| $\alpha_0$ | mortality rate due to competition between the individuals of $P$ | 0.30 | [9] |
| $\gamma_1$ | the protection of prey $P$ from the environment | 0.00661 | [5] |
| $r_z$ | the zooplankton growth rates | 0.25 | [9] |
| $\alpha_2$ | maximum value of the reduction rate per individual of $Z$ | 0.26 | [9] |
| $\gamma_2$ | the protection of predator $Z$ from the environment | 0.231 | [15] |
| $P_0$ | initial condition of the phytoplankton | 150 | [16] |
| $F_0$ | initial condition of the zooplankton | 100 | [15] |

**Table 1**: Parameters values for the numerical simulation of the system.

### 4.1 Pattern formation

Here, we will illustrate the mathematical predictions, by numerical simulations, concerning the behavior of the dynamics under the hypotheses $(H_1) - (H_4)$. The qualitative results of different pattern formations due to the variation of $\alpha_1$ are shown. We consider the value of released toxin $\theta_p = 0.06$. These numerical results show that for every strictly positive initial condition, under the assumptions $(H_1)-(H_4)$, the non-homogeneous equilibrium is always globally asymptotically stable. Figure $2-7$ show the spatial structures formation for the two species described in (9). This numerical results confirm the mathematical results for the existence of positive equilibrium and its stability according to the values of $\alpha_1$. In this case, we will speak of a subsistence phenomenon of the zooplankton population.

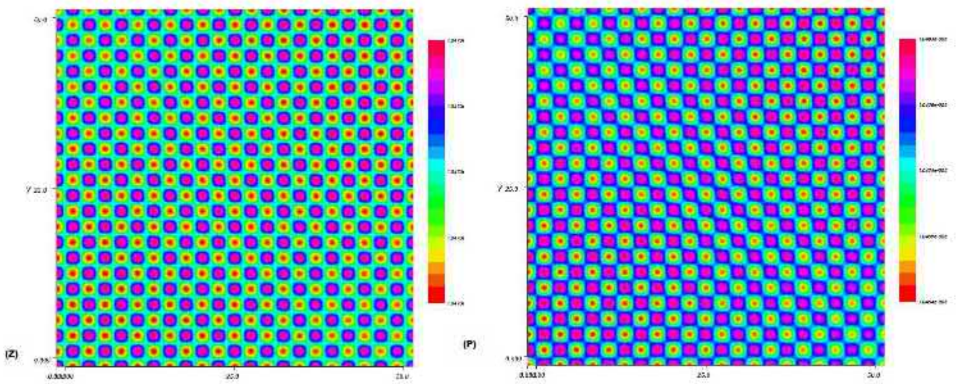


**Figure 2**: Spatial distribution of the two species, zooplankton and phytoplankton, if we consider in the system $\alpha_1 = 0.125$ and $d_z = 120.1$.

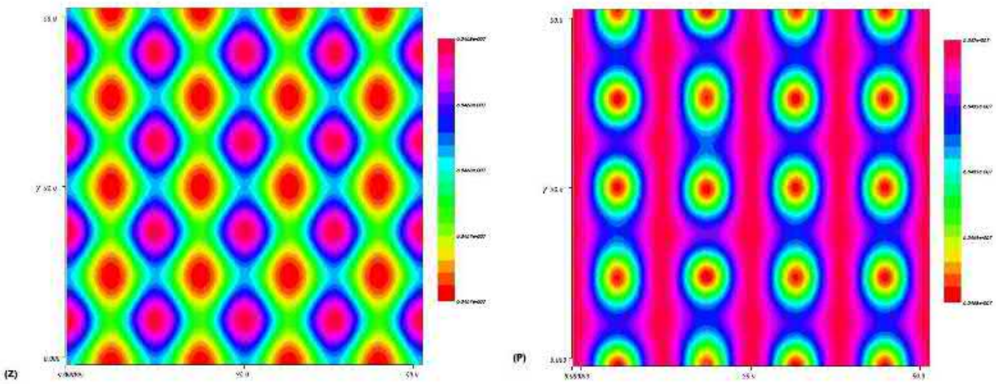


**Figure 3**: Spatial distribution of the two species, zooplankton and phytoplankton, if we consider in the system $\alpha_1 = 0.195$ and $d_z = 120.1$.

**Remark 4.1** From a biological point of view, these results (Figure $2-7$) show that there is coexistence between the two populations despite the release of the toxin into the aquatic environment. This means that despite the harmful effects of the toxin released by phytoplankton, the zooplankton population persists.

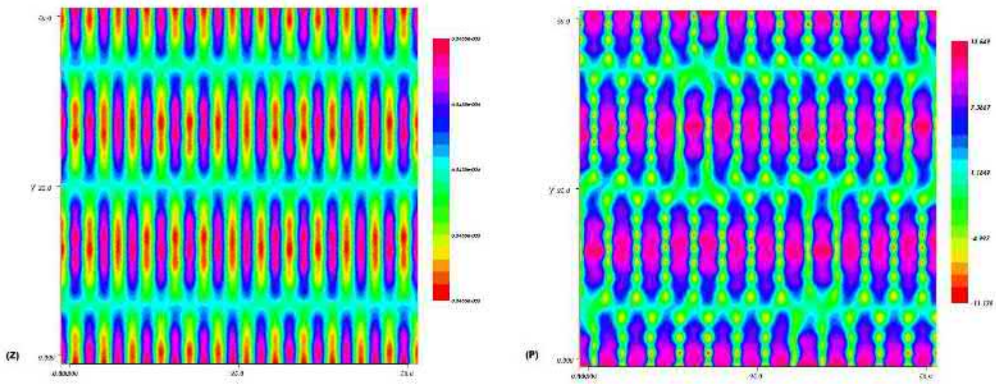

**Figure 4**: Spatial distribution of the two species, zooplankton and phytoplankton, if we consider in the system $\alpha_1 = 0.198$ and $d_z = 120.1$.

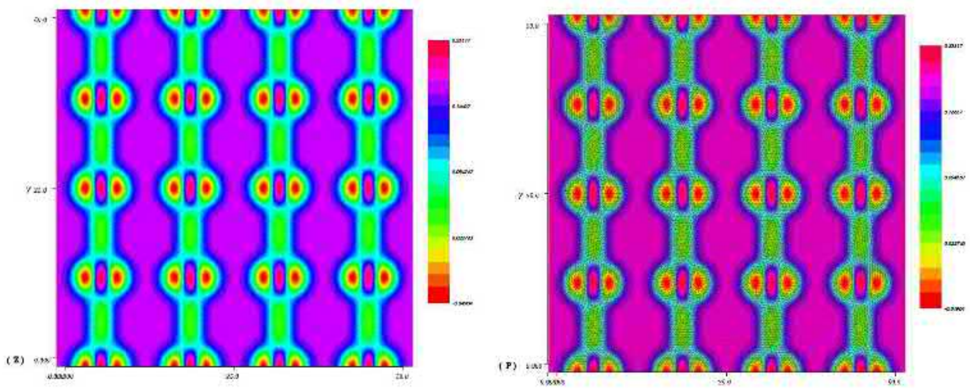

**Figure 5**: Spatial distribution of the two species, zooplankton and phytoplankton if we consider in the system $\alpha_1 = 0.205$ and $d_z = 120.1$.

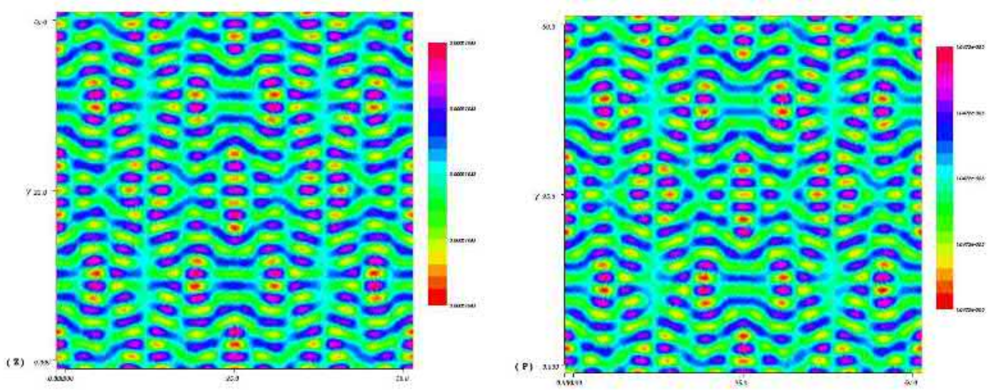

**Figure 6**: Spatial distribution of the two species, zooplankton and phytoplankton, if we consider in the system $\alpha_1 = 0.23$ and $d_z = 120.1$.

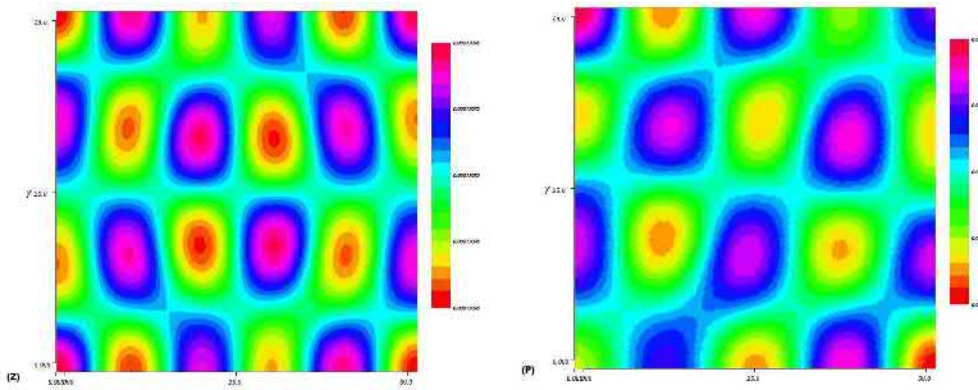

**Figure 7**: Spatial distribution of the two species, zooplankton and phytoplankton, if we consider in the system $\alpha_1 = 0.26$ and $d_z = 120.1$.

### 4.2 Analysis of the dynamics behavior with toxin effect

We continue our numerical study in this sub-section to look at the dynamics behavior of the system by considering different values of the toxin parameter. Here, we consider that $\alpha_1 = 0.25$. The numerical simulations show that after a transitional phase, the equilibrium can be established with coexistence of the two populations. Figure $8-11$ show the behavior of the two populations. As a biological interpretation we can say that if the toxin is released below this value the impact is not significant on the zooplankton population (Figure $8-10$). In fact, the effect does not disrupt the survival of other species. Figure 11 shows the spatial distributions of the two populations. A less dense distribution of the zooplankton population than the previous one was observed. This explains the considerable decrease of these species due to the increase in the number of toxic phytoplankton. There is a strong distribution of the phytoplankton population. Since the distribution is high, this explains the release of the toxin in large quantities by this population. This period corresponds to the phytoplankton bloom.

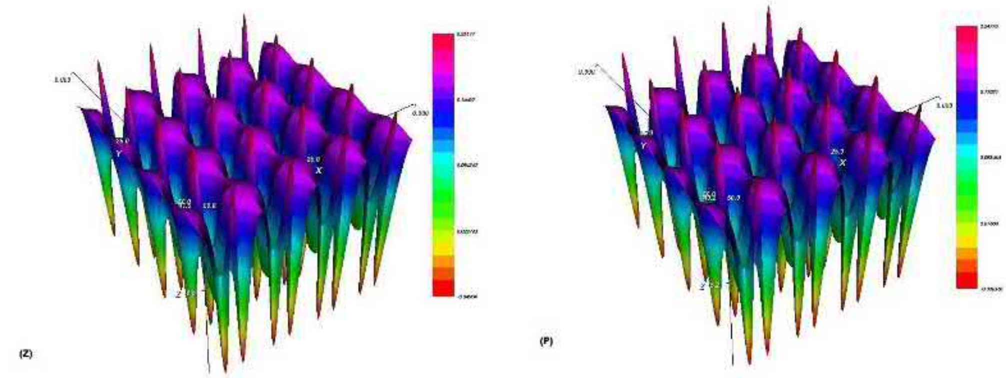

**Figure 8**: Dynamics behavior of the two species with $\alpha_1 = 0.26$, $d_z = 120.1$ and $\theta_p = 0.2$

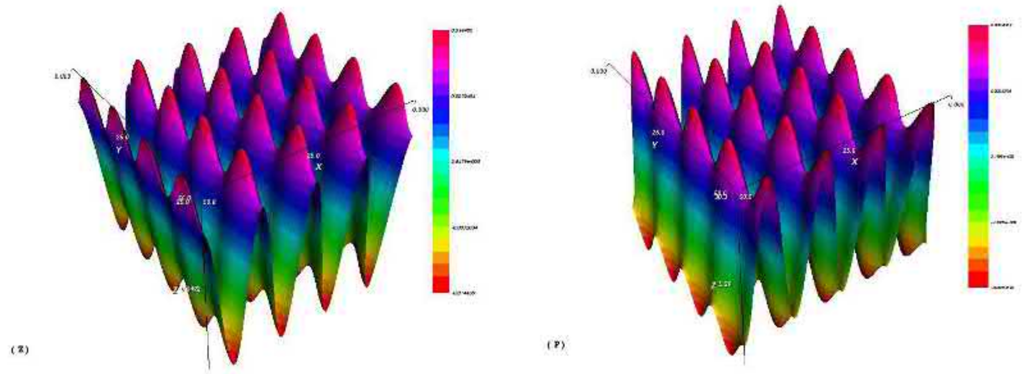

**Figure 9**: Dynamics behavior of the two species with $\alpha_1 = 0.26, \quad d_z = 120.1$ and $\theta_p = 0.35$.

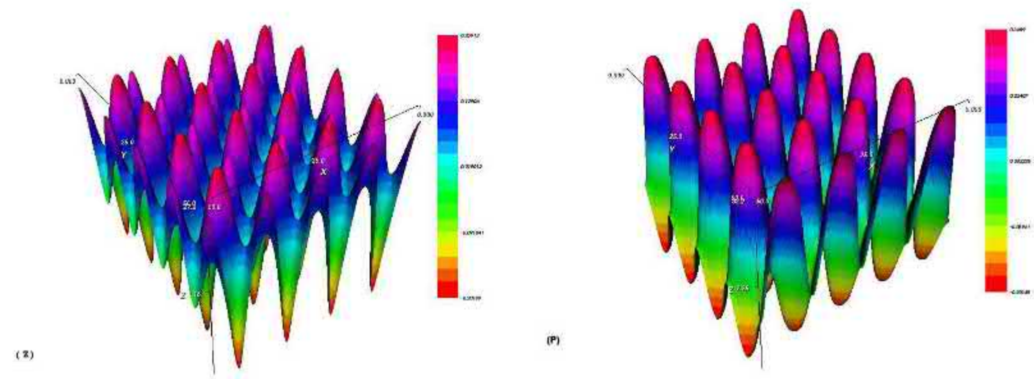

**Figure 10**: Dynamics behavior of the two species with $\alpha_1 = 0.26, \quad d_z = 120.1$ and $\theta_p = 0.504$.

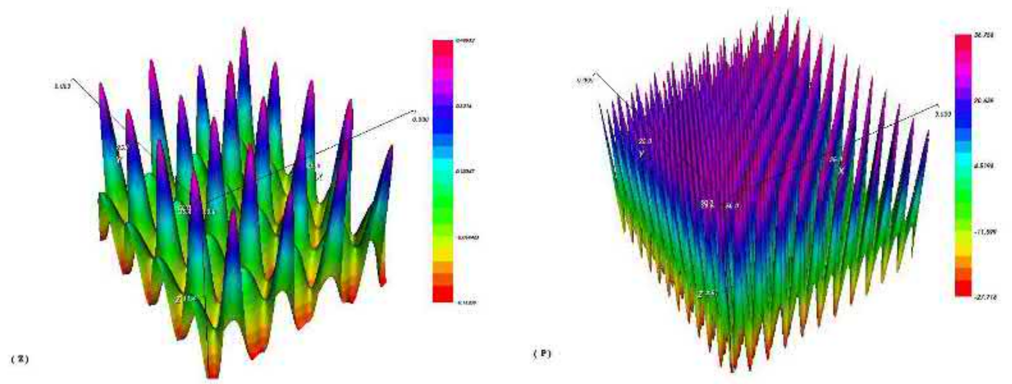

**Figure 11**: Dynamics behavior of the two species with $\alpha_1 = 0.26, \quad d_z = 120.1$ and $\theta_p = 1.4$.

## 5 Conclusion

In this paper, our interest is the formulation of a reaction-diffusion model to represent the dynamics of zooplankton and phytoplankton population by taking into account the effect of the toxin. The model construction is derived from an ODE system by considering an isotropic distribution as in [5, 8]. It should be noted that we consider diffusion independently of the spatial variable in the construction of the reaction-diffusion model. The mathematical results allowed us to establish conditions of existence of equilibrium which depend on the demographic parameters. We also gave some results about the stability of the stationary equilibria and we established the conditions on the non existence of the equilibrium with strictly positive components.

We continued our study through numerical experiments in order to confirm our mathematical results. The numerical results have yielded interesting results on the effect of the toxin on the dynamics. This is why we are led to conclude that the release of the toxin under certain conditions, in the aquatic environment contributes to the regulation of the system. The phytoplankton bloom was observed during our simulations and is in perfect agreement with the biological observations.

Despite of important results on this dynamics, in order to further our study, we consider, for our future work, to clearly subdivide the phytoplankton population into toxic phytoplankton and non-toxic phytoplankton to extend our results to these types of cross-diffusion system.

## Acknowledgements

We would like to thank the referees for their careful reading and their useful remarks.

## References

[1] Amann, H. Dynamics theory of quasilinear parabolic equations-I. Abstract evolution equations. *Nonlinear Anal.* **(9)** (12) (1988) 895–919.

[2] Aronson, D.G. and Weinberger, H.F. Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. *Lecture Notes in Math.* **446** (1975) 5–49.

[3] Ainseba, B.E., Heiser, F. and Langlais, M. A mathematical analysis of a predator-prey system in a highly heterogeneous environment. *Differential and Integral Equations* **15** (2002) 385–404.

[4] Aziz-Alaoui, M. and Daher Okiye, M. Boundedness and global stability for a predator-prey model with modified Leslie-Gower and Holling type II shemes. *Applied Mathematical Letters* **16**(7) (2003) 1069–1075.

[5] Andreianov, B., Bendahmane, M. and Ruiz-Baier, R. Analysis of a finite volume method for a cross-diffusion model in population dynamics. *Mathematical Models and Methods in Applied Sciences* **21** (2) (2011) 307–344.

[6] Camara, B.I. *Complexité de Dynamiques de Modèles Proie-prédateur Avec Diffusion et Applications.* Mathematics, Université du Havre, 2009. French.

[7] Briand, E. Contribution à la compréhention du determinisme de la mise en place des proliférations du cyanobacteries et leur production de toxine. *Thèse de doctorat*, Université Paris 7.

[8] Courchamp, F., Langlais, M. and Sugihara, G. Rabbits Killing birds: modelling the hyperpredation process. *Journal of Animal Ecology* **69** (2000) 154–164.

[9] Dockery, J., Hutson, V., Mischaikow, K. and Pernarowski, M. The evolution of slow dispersal rates : a reaction diffusion model. *Journal of Mathematical Biology* **37** (1) (1998) 61–83.

[10] Edward, A.M. and Brindley, J. Zooplankton Mortality and the Dynamical Behaviour of Plankton Population Models. *Bulletin of Mathematical Biology* **61** (1999) 303–339.

[11] Galiano, G., Garzon, M.L. and Jungel, A. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numerical Mathematics* **93** (4) (2003) 655–673.

[12] Gaucel, S., Langlais, M. and Pontier, D. Invading introduced species in insular heterogeneous environments. *Ecology Modelling* **188** (2005) 62–75.

[13] Kohlmeier, C. and Ebenhh, W. The stabilizing role of cannibalism in a predator-prey system. *Bulletin of Mathematical Biology* **57** (1995) 401–411.

[14] Koutou, O., Traoré, B. and Sangaré B. Mathematical modeling of malaria transmission global dynamics: taking into account the immature stages of the vectors. *Advance in Difference Equation* **2018** (220) (2018) 1–34.

[15] Kot, M. *Elements of Mathematical Ecology. Cambridge University Press*, 2001.

[16] Kreisselmeir, G. and Engel, R. Nonlinear observer autonomous Lipschitz continuous systems. *IEEE Trans. on Automatic control.* **48** (2003) 451–464.

[17] Mathews, L. and Brindley, J. Patchiness in plankton populations. *Dynamics and Stability of Systems* **12** (1996) 39–59.

[18] Narayan Guin, L. and Acharya, S. Dynamic behaviour of a reaction-diffusion predator-prey model with both refuge and harvesting. *Nonlinear Dynamical*, DOI 10.1007/s11071-016-3326-8.

[19] Ouedraogo, W., Sangaré, B. and Traoré, S. Some mathematical problems arising in biological models: a predator-prey model fish-plankton. *Journal of Applied Mathematic and Bioinformatics* **5** (4) (2015) 1–27.

[20] Ouedraogo, W., Ouedraogo, H. and Sangaré, B. A Reaction Diffusion Model to Describe the Toxin Effect on the Fish-Plankton Population. *Journal of Mathematics* **2018** (2018), Article ID 2037093, 15 pages.

[21] Ouedraogo, W., Sangaré, B. and Traoré S. A mathematical study of cannibalism in the fish-plankton model by taking into account the catching effect. *AMO – Advanced Modeling and Optimization* **18** (2) (2016) 197–216.

[22] Parandyk, W., Lewandowski, D. and Awrejcewicz, J. Mathematical Modeling of the Hydro-Mechanical Fluid Flow System on the Basis of the Human Circulatory System. *Nonlinear Dynamics and Systems Theory* **15** (1) (2015) 50–62.

[23] Pousin, J. and Rappaz, J. Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problem. *Numerical Mathematics* **69** (1994) 213–232.

[24] Rosenzweig, M. and MacArthur, R. Graphical representation and stability conditions of predator-prey interaction. *American Naturalist* **97** (1963) 209–223.

[25] Sambatha, M. and Sahadevanb, R. Hopf bifurcation analysis of a diffusive predator-prey model with Monod-Haldane response. *Journal of Mathematical Modeling* **5** (2) (2017) 119–136.

[26] Traoré, B., Sangaré, B. and Traoré, S. A Mathematical Model of Malaria Transmission with Structured Vector Population and Seasonality. *Journal of Applied Mathematics* **2017** (2017), Article ID 6754097, 15 pages.

[27] Traoré, B., Sangaré, B. and Traoré, S. A mathematical model of malaria transmission in a periodic environment. *Journal of Biological Dynamics* **12** (1) (2018) 400–432.

[28] Xiaona, G., Wang, W. and Cai, Y. Spatio-temporal dynamics of a Leslie-Gower predator-prey model incorporating a prey refuge. *Nonlinear Analysis: Real World Applications* **12** (2011) 2385–2395.

[29] Zhang, L.Y. Hopf bifurcation analysis in a Monod-Haldane predator-prey model with delays and diffusion. *Applied Mathematical Modelling* **39** (2015) 1369–1382.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 409–417

# Boundedness Results for a New Hyperchaotic System and Their Application in Chaos Synchronization

Samia Rezzag *

*Department of Mathematics and Informatics, University Larbi Ben M'hidi, 04000, Oum-El-Bouaghi, Algeria*

Received: June 19, 2018; Revised: October 5, 2018

**Abstract:** In this paper, we attempt to investigate the boundedness of a new hyperchaotic system using the combination of the Lyapunov stability theory with the comparison principle method. Furthermore, explicit estimation of the two-dimensional parabolic ultimate bound with respect to x-z is established. Finally, a linear feedback approach with one input is used to realize the global synchronization of two four-dimesional hyperchaotic systems. Some numerical simulations are also used to verify the effectiveness and correctness of the proposed scheme.

**Keywords:** *4D hyperchaotic system; boundedness of solutions; Lyapunov stability; chaos synchronization; comparison principle method.*

**Mathematics Subject Classification (2010):** 65P20, 65P30, 65P40.

## 1 Introduction

Hyperchaotic systems are dissipative nonlinear dynamical systems with more than one positive Lyapunov exponent. The Lyapunov exponent of a chaotic system is a measure of the divergence of points which are initially very close and this can be used to quantify chaotic systems. So, the hyperchaos may be more useful in some fields such as communication encryption, and so forth.

An important paradigm of a 3-D chaotic system was discovered by Lorenz [7] while he was studying a 3-D weather model. Subsequently, many chaotic systems have attracted tremendous research interest, and many chaotic and hyperchaotic systems have been presented.

Chaotic systems are ultimately bounded. Thus, the phase portraits of the systems will be ultimately trapped in some compact sets. The ultimate boundedness of a chaotic

* Corresponding author: mailto:rezzag.samia@gmail.com

system is very important for the study of the qualitative behavior of a chaotic system. In fact, except for the stability property, boundedness is also one of the foundational concepts of dynamical systems, which plays an important role in investigating the uniqueness of equilibrium, global asymptotic stability, global exponential stability, the existence of the periodic solution, its control and synchronization and so on. Furthermore, it can be applied in estimating the fractal dimensions of chaotic attractors, such as the Hausdorff dimension and the Lyapunov dimension of chaotic attractors [3].

Ultimate bound estimation of chaotic systems is a difficult yet interesting mathematical question. At present, several works on this topic were realized for some 3D and 4D dynamical systems, see [2, 4–6, 8–12, 14, 15].

Recently, Chen Hai-tao, Chen Di-yi and Ma Xiao-yi [1] introduced the following new system

$$\begin{cases} x' = \alpha\left(y - x\right), \\ y' = \gamma x - xz - y, \\ z' = xy - \beta z, \\ w' = -x - \alpha w, \end{cases} \tag{1}$$

where $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ is a vector parameter. When $\alpha = 5$, $\beta = 0.7$, $\gamma = 26$, system (1) has a hyperchaotic attractor. Fig. 1. shows the phase portraits of system (1).

In this paper, we firstly investigate the boundedness for this new hyperchaotic system using a combination of Lyapunov stability theory with the comparison principle method. In addition, the two-dimensional parabolic ultimate bound with respect to $x - z$ is established.

Synchronization of chaotic systems has become an important topic in nonlinear science not only for its importance in theory but also for its potential applications in various areas, for example, secure communication, chemical and biomedical science, life science, electromechanical engineering and so on. During the last decades, many methods have been successfully applied to chaos synchronization such as PC method, linear feedback control, adaptive control, backstepping design, active control and nonlinear control, etc.

In this paper, based on the bounds previously obtained, we use linear feedback control with one input to realize global synchronization between two identical hyperchaotic systems.

The rest of this paper is organized as follows. In Section 2, we study the boundedness of the hyperchaotic systems (1). In Section 3, the two-dimensional bound estimation with respect to $x - z$ is established. In Section 4, our outcomes are applied between the master system and the slave system to the study of completely chaos synchronization. In Section 5, numerical simulation is presented to show the effectiveness of our results. Section 6 is the conclusion of the paper.

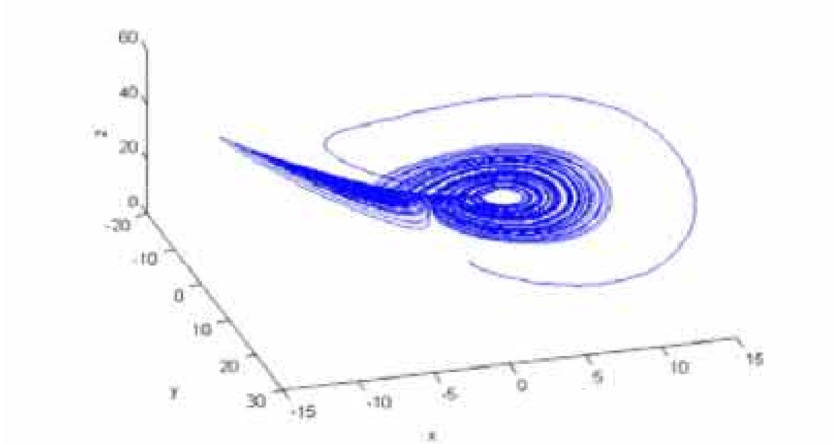


**Fig. 1**: Phase portrait of the system (1) in the $x-y-z$ space with parameters $\alpha = 5$, $\beta = 0.7$, $\gamma = 26$.

## 2 Bounds for Solutions of the New Hyperchaotic System

**Lemma 2.1** *[5] Define a set*

$$\Gamma = \left\{ (x, y, z) / \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{(z-c)^2}{c^2} = 1, \ a > 0, \ b > 0, \ c > 0 \right\} \tag{2}$$

*and* $G = x^2 + y^2 + z^2$, $H = x^2 + y^2 + (z - 2c)^2$, $(x, y, z) \in \Gamma$. *Then we have*

$$\max_{(x,y,z)\in\Gamma} G = \max_{(x,y,z)\in\Gamma} H = \begin{cases} \dfrac{a^4}{a^2 - c^2}, \ a \geq b, \ a \geq \sqrt{2}c, \\ 4c^2, \ a < \sqrt{2}c, \ b < \sqrt{2}c, \\ \dfrac{b^4}{b^2 - c^2}, \ b > a, \ b \geq \sqrt{2}c. \end{cases} \tag{3}$$

**Theorem 2.1** *For* $\alpha > 0$, $\beta > 0$, $\gamma > 0$ *the following set*

$$\Omega = \left\{ (x, y, z, w) / x^2 + y^2 + (z - \alpha - \gamma)^2 \leq R^2, \ w^2 \leq \frac{R^2}{\alpha^2} \right\} \tag{4}$$

*is the bound for system* (1)*, where*

$$R^2 = \begin{cases} \dfrac{\beta^2 (\alpha + \gamma)^2}{4\alpha (\beta - \alpha)} \ , \ \text{if } \beta \geq 2\alpha, \ \alpha \leq 1, \\ (\alpha + \gamma)^2 \ , \ \text{if } \beta < 2\alpha, \ \beta < 2, \\ \dfrac{\beta^2 (\alpha + \gamma)^2}{4 (\beta - 1)} \ , \ \text{if } \alpha > 1, \ \beta \geq 2. \end{cases} \tag{5}$$

***Proof.*** Construct the following Lyapunov function

$$V(x, y, z) = x^2 + y^2 + (z - \alpha - \gamma)^2 . \tag{6}$$

Then, its time derivative along the orbits of system (1) is

$$\dot{V} = 2x\dot{x} + 2y\dot{y} + 2\left(z - \alpha - \gamma\right)\dot{z} = -2\alpha x^2 - 2y^2 - 2\beta\left(z - \frac{\alpha+\gamma}{2}\right)^2 + \beta\frac{\left(\alpha+\gamma\right)^2}{2}. \quad (7)$$

Therefore, $\dot{V} = 0$, that means, the surface

$$\Gamma = \left\{ \left(x, y, z\right) / \frac{x^2}{\frac{\beta\left(\alpha+\gamma\right)^2}{4\alpha}} + \frac{y^2}{\frac{\beta\left(\alpha+\gamma\right)^2}{4}} + \frac{\left(z - \frac{\alpha+\gamma}{2}\right)^2}{\frac{\left(\alpha+\gamma\right)^2}{4}} = 1 \right\} \quad (8)$$

is an ellipsoid in three-dimensional space. Outside $\Gamma$, $\dot{V} < 0$, while inside $\Gamma$, $\dot{V} > 0$. Since the function $V = x^2 + y^2 + \left(z - \alpha - \gamma\right)^2$ is continuous on the closed set $\Gamma$, $V$ can reach its maximum on the surface $\Gamma$. Next, we use Lemma 1 and obtain the optimal value of $V$ on $\Gamma$.

$$V \leq \max_{(x,y,z)\in\Gamma} V = R^2 = \begin{cases} \dfrac{\beta^2\left(\alpha+\gamma\right)^2}{4\alpha\left(\beta-\alpha\right)}, \text{ if } \beta \geq 2\alpha,\ \alpha \leq 1, \\ \left(\alpha+\gamma\right)^2, \text{ if } \beta < 2\alpha,\ \beta < 2, \\ \dfrac{\beta^2\left(\alpha+\gamma\right)^2}{4\left(\beta-1\right)}, \text{ if } \alpha > 1,\ \beta \geq 2. \end{cases} \quad (9)$$

Thus, we have

$$|x| \leq R \quad (10)$$

and

$$w' = -x - \alpha w \leq -\alpha w + R. \quad (11)$$

By the comparison principle, we obtain

$$w\left(t\right) \leq \frac{R}{\alpha} + \left(w\left(t_0\right) - \frac{R}{\alpha}\right) e^{-\alpha\left(t-t_0\right)} \quad (12)$$

and

$$\lim_{t\to+\infty} w\left(t\right) \leq \frac{R}{\alpha}. \quad (13)$$

Consequently, we get $w^2 \leq \dfrac{R^2}{\alpha^2}$ as $t \to +\infty$. Summarizing the above, we have the main result that

$$\Omega = \left\{ \left(x, y, z, w\right) / x^2 + y^2 + \left(z - \alpha - \gamma\right)^2 \leq R^2,\ w^2 \leq \frac{R^2}{\alpha^2} \right\} \quad (14)$$

is the bound for the hyperchaotic systems (1). This completes the proof.

## 3 Estimate of the Two-Dimensional Parabolic Ultimate Bound with Respect to x–z

**Theorem 3.1** *When $\beta < 2\alpha$, the system* (1) *has the following two-dimensional parabolic ultimate bound*

$$z \geq \frac{x^2}{2\alpha}. \quad (15)$$

***Proof.*** Define

$$V(t) = \frac{1}{2\alpha}x^2(t) - z(t).$$

Then, its time derivative along the orbits of system (1) is

$$\dot{V} = \frac{1}{\alpha}x\dot{x} - \dot{z} = -x^2 + \beta z.$$

Thus,

$$\dot{V} + \beta V = -x^2 + \beta z + \frac{\beta}{2\alpha}x^2 - \beta z = \left(\frac{\beta}{2\alpha} - 1\right)x^2.$$

When $\beta < 2\alpha$, we have

$$\dot{V} + \beta V \leq 0.$$

For any initial value$V(t_0) = V_0$, according to the comparison theorem, we have

$$V(t) \leq V_0 e^{-\beta(t-t_0)} \to 0\,(t \to \infty).$$

Thus,

$$\lim_{t\to\infty} V(t) = \lim_{t\to\infty}\left[\frac{1}{2\alpha}x^2(t) - z(t)\right] \leq 0.$$

So, we get that system orbits satisfy the parabolic ultimate bound

$$z \geq \frac{x^2}{2\alpha}.$$

This completes the proof.

## 4 The Application in Chaos Synchronization

In this section, we will use the results obtained in Section 2 to study chaos synchronization via linear feedback. For the master system (1), we construct another system called the slave system, which is designed as

$$\begin{cases} \dot{x}_1 = \alpha(y_1 - x_1), \\ \dot{y}_1 = \gamma x_1 - x_1 z_1 - y_1 - k(y_1 - y), \\ \dot{z}_1 = x_1 y_1 - \beta z_1, \\ \dot{w}_1 = -x_1 - \alpha w_1, \end{cases} \tag{16}$$

where $x_1$, $y_1$, $z_1$, $w_1$ are the state variables and $k > 0$ is the control. From Theorem 2.1, we obtain

$$|y| \leq R, \;\; |z| \leq R + \alpha + \gamma. \tag{17}$$

**Theorem 4.1** *Systems* (1) *and* (16) *are globally and asymptotically synchronized when*

$$k > \frac{\beta\left(\alpha(\sigma+1) + R + 2\gamma\right)^2}{4\alpha\beta\sigma - R^2} - 1. \;\left(\sigma > \frac{R^2}{4\alpha\beta} > 0\right). \tag{18}$$

***Proof.*** The complete synchronization error is defined by $e_1 = x_1 - x$, $e_2 = y_1 - y$, $e_3 = z_1 - z$, $e_4 = w_1 - w$. Then, the error dynamics is obtained as

$$\begin{cases} \dot{e}_1 = \alpha (e_2 - e_1), \\ \dot{e}_2 = (\gamma - z) e_1 - x e_3 - e_1 e_3 - (k+1) e_2, \\ \dot{e}_3 = y e_1 + x e_2 + e_1 e_2 - \beta e_3, \\ \dot{e}_4 = -e_1 - \alpha e_4. \end{cases} \tag{19}$$

Define the following Lyapunov function

$$V(e_1, e_2, e_3) = \sigma e_1^2 + e_2^2 + e_3^2,$$

where $\sigma$ is a positive constant and $\sigma > \dfrac{R^2}{4\alpha\beta} > 0$. Then, its time derivative along the system (19) is

$$\begin{aligned} \frac{1}{2}\dot{V} &= \sigma e_1 \dot{e}_1 + e_2 \dot{e}_2 + e_3 \dot{e}_3 \\ &= \sigma e_1 (\alpha e_2 - \alpha e_1) + e_2 ((\gamma - z) e_1 - x e_3 - e_1 e_3 - (k+1) e_2) \\ &+ e_3 (y e_1 + x e_2 + e_1 e_2 - \beta e_3) \\ &= -\sigma\alpha e_1^2 - (k+1) e_2^2 - \beta e_3^2 + (\sigma\alpha + \gamma - z) e_1 e_2 + y e_1 e_3 \\ &\leq -\sigma\alpha e_1^2 - (k+1) e_2^2 - \beta e_3^2 + (\alpha(\sigma+1) + R + 2\gamma) |e_1| |e_2| + R |e_1| |e_3| \\ &= -E^T P E, \end{aligned}$$

where

$$E = [|e_1|, |e_2|, |e_3|]^T, \ P = \begin{bmatrix} \sigma\alpha & -\dfrac{\alpha(\sigma+1) + R + 2\gamma}{2} & -\dfrac{R}{2} \\ -\dfrac{\alpha(\sigma+1) + R + 2\gamma}{2} & k+1 & 0 \\ -\dfrac{R}{2} & 0 & \beta \end{bmatrix},$$

which is positive definite when

$$\sigma > \frac{R^2}{4\alpha\beta} > 0, \ k > \frac{\beta (\alpha(\sigma+1) + R + 2\gamma)^2}{4\alpha\beta\sigma - R^2} - 1.$$

Thus, according to the Lyapunov function theory, it follows that

$$\lim_{t\to+\infty} |e_1| = 0, \lim_{t\to+\infty} |e_2| = 0, \lim_{t\to+\infty} |e_3| = 0. \tag{20}$$

In the following, we will prove $\lim_{t\to+\infty} e_4 = 0$. From (20), we have $\lim_{t\to+\infty} e_1 = 0$. Therefore, for any $\varepsilon > 0$, there is a sufficiently large $T > t_0$ such that, when $t \geq T$, we have $\left|\dfrac{e_1}{\alpha}\right| < \varepsilon$. So, for any $\varepsilon > 0$, when $t \geq T$, from (19), we have

$$\begin{aligned} e_4(t) &= e_4(t_0) e^{-\alpha(t-t_0)} + e^{-\alpha t} \int_{t_0}^{t} (-e_1) e^{\alpha\tau} d\tau \\ &\leq e_4(t_0) e^{-\alpha(t-t_0)} + e^{-\alpha t} \int_{t_0}^{t} \alpha\varepsilon e^{\alpha\tau} d\tau \\ &= (e_4(t_0) - \varepsilon) e^{-\alpha(t-t_0)} + \varepsilon. \end{aligned}$$

Thus, if the initial value $e_4(t_0) > \varepsilon$ and $t \to +\infty$, we obtain

$$e_4(t) - \varepsilon \leq (e_4(t_0) - \varepsilon)\, e^{-\alpha(t-t_0)} \to 0.$$

Also, we have

$$\begin{aligned} e_4(t) &= e_4(t_0)e^{-\alpha(t-t_0)} + e^{-\alpha t}\int_{t_0}^{t} (-e_1)\, e^{\alpha\tau} d\tau \\ &\geq e_4(t_0)e^{-\alpha(t-t_0)} - e^{-\alpha t}\int_{t_0}^{t} \alpha\varepsilon e^{\alpha\tau} d\tau \\ &= (e_4(t_0) + \varepsilon)\, e^{-\alpha(t-t_0)} - \varepsilon. \end{aligned}$$

Thus, if the initial value $e_4(t_0) < -\varepsilon$ and $t \to +\infty$, we get

$$e_4(t) + \varepsilon \leq (e_4(t_0) + \varepsilon)\, e^{-\alpha(t-t_0)} \to 0.$$

Consequently, when the initial value $|e_4(t_0)| > \varepsilon$ and $t \to +\infty$, we have the distance $d\,(e_4(t), I) \to 0$, where $I = [-\varepsilon, \varepsilon]$. So, for any sufficiently small $\varepsilon > 0$, there is a sufficiently large $T > t_0$ such that, when $t > T$, we have $|e_4(t)| < \varepsilon$. By the definition of limit, we obtain

$$\lim_{t\to+\infty} e_4(t) = 0. \tag{21}$$

Summarizing the above, we have

$$\lim_{t\to+\infty} |e_1| = 0, \lim_{t\to+\infty} |e_2| = 0, \lim_{t\to+\infty} |e_3| = 0, \lim_{t\to+\infty} |e_4| = 0.$$

Finally, we conclude that the master system (1) and the slave system (16) are globally synchronized. This completes the proof.

## 5 Simulation Studies

In this section, using the MATLAB 7.4, some numerical simulations are presented. As initial conditions for the master and slave systems, we take $(1, -0.5, 3, 4)$ and $(-8, -1, -4, -1)$, respectively. When $\alpha = 5$, $\beta = 0.7$, $\gamma = 26$, it is easy to obtain $R = 31$, $\sigma > \dfrac{R^2}{4\alpha\beta} = 68.64$, according to Theorems 2.1 and 4.1. Choose $\sigma = 69$, $k = 26248$, then Fig. 2. shows the complete synchronization between systems (1) and (16).

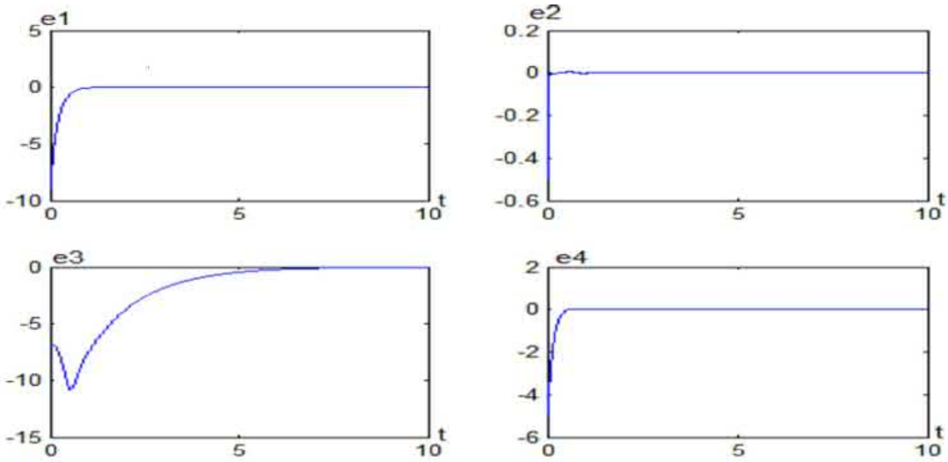


**Fig. 2**: Time evolution of synchronization errors $e_1$, $e_2$, $e_3$ and $e_4$ between the master system (1) and the slave system (16).

## 6 Conclusion

In this research work, the boundedness of a new hyperchaotic system has been investigated. Furthermore, the two-dimensional parabolic estimate with respect to $x-z$ for the new system is established. Finally, the result is applied to the chaos synchronization and numerical simulations are presented to show the effectiveness of the proposed scheme.

## References


[1] Chen, H. and Chen, D. and Ma, X. The Recursive Inversion of Nonlinear Control of a Four-dimensional Hyperchaotic System. In: *International Conference on Intelligent Computing and Integrated Systems.* IEEE, Guilin, China, 2010, 367–370.

[2] Elhadj, Z. and Sprott, J.C. About the boundedness of 3D continuous time quadratic systems. *Nonlinear Osci.* **13** (2-3) (2010) 515–521.

[3] Kuznetsov, N., Mokaev, T. and Vasilyev, P. Numerical justification of Leonov conjecture on Lyapunov dimension of Rossler attractor. *Commun. Nonlinear Sci. Numer. Simul.* **19** (4) (2014) 1027–1034.

[4] Leonov, G., Bunin, A. and Koksch, N. Attractor localisation of the Lorenz system. *Z. Angew. Math. Mech.* **67** (1987) 649–656.

[5] Li, D., Lu, JA., Wu, X. and Chen, G. Estimating the bounds for the Lorenz family of chaotic systems. *Chaos, Solitons Fract.* **23** (2005) 529–534.

[6] Li, D., Wu, X. and Lu, J. Estimating the ultimating bound and positively invariant set for the hyperchaotic Lorenz-Haken system. *Chaos, Solitons Fract.* **39** (2009) 1290–1296.

[7] Lorenz, E.N. Deterministic non-periods flows. *J. Atmos. Sci.* **20** (1963) 130–141.

[8] Pogromsky, AY., Santoboni, G. and Nijmeijer, H. An ultimate bound on the trajectories of the Lorenz systems and its applications. *Non-linearity* **16** (2003) 1597–1605.

[9] Rezzag, S. Solution bounds of the hyper-chaotic Rabinovich system. *Nonlinear Studies* **24**(4) (2017) 903–909.

[10] Rezzag, S. Boundedness of the new modified hyperchaotic Pan System. Nonlinear Dyn. Syst. Theo **17**(3) (2017) 402–408.

[11] Rezzag, S., Zehrour, O, and Aliouche, A. Estimating the bounds for the general 4-D hyperchaotic system. *Nonlinear Studies* **22**(1) (2015) 41–48.

[12] Rezzag, S., Zehrour, O. and Aliouche, A. Estimating the Bounds for the General 4-D Continuous-Time Autonomous System. *Nonlinear Dyn. Syst. Theory* **15**(3) (2015) 313–320.

[13] Shu, Y., Xu, H. and Zhao, Y. Estimating the ultimate bound and positively invariant set for a new chaotic system and its application in chaos synchronization. *Chaos Solitons Fract.* **42**(5) (2009) 2852–2857.

[14] Sun, YJ. Solution bounds of generalized Lorenz chaotic system. *Chaos Soliton Fract.* **40** (2009) 691–696.

[15] Wang, P., Li, D. and Hu, Q. Bounds of the hyper-chaotic Lorenz-Stenflo system. *Commun. Nonlinear Sci. Numer. Simul.* **15** (2010) 2514–2520.

Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 418–425

# On the Stabilization of Infinite Dimensional Bilinear Systems with Unbounded Control Operator

E. Zerrik *, A. Ait Aadi and R. Larhrissi

*MACS Laboratory, Department of Mathematics, University of Moulay Ismail, Meknes, Morocco*

Received: December 25, 2017; Revised: October 5, 2018

**Abstract:** In this paper, we deal with regional stabilization of infinite dimensional bilinear system evolving in a spatial domain $\Omega$ with unbounded control operator. It consists in studying the asymptotic behaviour of such a system in a subregion $\omega$ of $\Omega$. Hence, we give sufficient conditions to obtain weak and strong stabilization on $\omega$. An example and simulations are presented.

**Keywords:** *infinite dimensional bilinear systems; unbounded control operator; regional stabilization.*

**Mathematics Subject Classification (2010):** 93C15, 93C10, 49J20.

## 1 Introduction

Bilinear systems constitute an important subclass of nonlinear systems. The nonlinearity in mathematical models appears in the multiplication of state and control in the dynamical process. Bilinear systems model several phenomena in nature and in industry, e.g. the mass action law in chemistry, the transfer of heat by conduction convection in energetic systems, the generation of cells via cellular division, and the dynamics of the blood's organs in biology [4]. Yet, the modeling may give rise to an unbounded control operator which allows us to describe some interesting phenomena when the control is acting in regions or on a boundary or when the measure is taken at some sensing point. The problem of feedback stabilization of distributed systems has been studied in many works along with various types of controls [1–3, 5].

The question of regional stabilization for linear systems was tackled and developed by Zerrik and Ouzahra [7], and consists in studying the asymptotic behaviour of a distributed

* Corresponding author: mailto:zerrik3@yahoo.fr

system only within a subregion $\omega$ interior or in the boundary of its evolution domain $\Omega$. The principal reason for introducing this notion is that it makes sense for the usual concept of stabilization taking into account the spacial variable and then it becomes closer to real world problems, where one wishes to stabilize a system on a critical subregion of its geometrical domain. Regional stabilization of bilinear systems with bounded control operator has been considered by Zerrik and Ouzahra [6]. Many approaches were used to characterize different kinds of stabilization, and mainly the control which achieves the stabilization minimizing a given functional cost.

In this paper, we examine regional stabilization of infinite dimensional bilinear systems with unbounded control operator. The paper is organized as follows : in Section 2, we discuss different kinds of regional stabilization, and we give sufficient conditions to achieve weak and strong stabilization of such a system on $\omega$. Finally, an example and simulations are given to illustrate the efficiency of the obtained results.

## 2 Regional Stabilization

### 2.1 Considered system and notations

Define an open, regular set $\Omega \subset \mathbb{R}^n$ $(n = 1, 2, ...)$ and consider the bilinear system

$$\begin{cases} \dot{z}(t) = Az(t) + v(t)Bz(t), \\ z(0) = z_0, \end{cases} \tag{1}$$

where $A : \mathcal{D}(A) \subset H \to H$ generates a strongly continuous semigroup of contractions $S(t)_{t\geq 0}$ on a Hilbert space $H := L^2(\Omega)$, endowed with norm and scalar product denoted, respectively, by $\|.\|$ and $\langle ., . \rangle$, $v(.) \in L^2(0, \infty)$ denotes the control function and $B$ is an unbounded linear operator from $H$ into itself, positive and self-adjoint but bounded from a subspace $V \subset H$ to some large space $V'$ such that $H \subset V'$. Identify $H$ with its dual $H'$ so that $V \hookrightarrow H \hookrightarrow V'$, and $\langle h, w\rangle_{V',V} = \langle h, w\rangle$ $\forall h, w \in H$.
We suppose that the state $z(t) \in V$. The solution of system (1) is a solution of the equation

$$z(t) = S(t)z_0 + \int_0^t v(s)S(t-s)Bz(s)ds. \tag{2}$$

Let $\omega$ be an open subregion of $\Omega$ and Lebesgue non-null measure, $\chi_\omega : L^2(\Omega) \to L^2(\omega)$ the restriction operator to $\omega$, while $\chi^*_\omega$ denotes the adjoint operator given by

$$\chi^*_\omega y(x) = \begin{cases} y(x), & \text{if } x \in \omega, \\ 0, & \text{if } x \in \Omega\backslash\omega. \end{cases}$$

Denote $i_\omega = \chi^*_\omega\chi_\omega$ and suppose that
$(H_1)$ $\langle i_\omega Ay, y\rangle \leq 0$, $\forall y \in \mathcal{D}(A)$;
$(H_2)$ $By \in H$, $\forall y \in V$;
$(H_3)$ $\langle i_\omega By, y\rangle_{V',V} \geq 0$, $\forall y \in V$.

**Definition 2.1** System (1) is said to be
1. weakly stabilizable on $\omega$, if $\chi_\omega z(t)$ tends to 0 weakly, as $t \to \infty$.
2. strongly stabilizable on $\omega$, if $\chi_\omega z(t)$ tends to 0 strongly, as $t \to \infty$.

**Remark 2.1** It is clear that we are only interested in the behaviour of system (1) on a subregion $\omega$ without constraints on the residual part $\Omega\backslash\omega$, and when $\omega = \Omega$ we retrieve the classical definition of stabilization.

## 2.2 Regional weak stabilization

The following result gives sufficient conditions for weak stabilization of system (1) on $\omega$.

**Theorem 2.1** *Let $A$ generate a semigroup $S(t)$ of contractions on $H$, the assumptions $(H_1), (H_2)$ and $(H_3)$ hold, $B$ be compact from $V$ to $V'$, and if the condition*

$$\langle i_\omega BS(t)y, S(t)y\rangle_{V',V} = 0,\ \forall t \geq 0 \implies \chi_\omega y = 0, \tag{3}$$

*is verified, then the control $v(t) = -\langle i_\omega Bz(t), z(t)\rangle_{V',V}$ weakly stabilizes system (1) on $\omega$.*

***Proof.*** From hypothesis $(H_1)$, we have

$$\frac{d}{dt}\|\chi_\omega z(t)\|^2 \leq 2v(t)\langle i_\omega Bz(t), z(t)\rangle_{V',V}. \tag{4}$$

In order to make the energy non increasing, a natural choice for the control is

$$v(t) = -\langle i_\omega Bz(t), z(t)\rangle_{V',V}.$$

Since $A$ generates a semigroup of contractions, we have

$$\|z(t)\|^2 - \|z(0)\|^2 \leq -2\int_0^t \langle i_\omega Bz(s), z(s)\rangle_{V',V}\langle Bz(s), z(s)\rangle_{V',V} ds. \tag{5}$$

Due to $(H_3)$ and the fact that $B$ is positive, it follows that

$$\|z(t)\| \leq \|z_0\|. \tag{6}$$

From (2), (6), and Schwartz's inequality, we deduce

$$\|z(t) - S(t)z_0\| \leq \delta\|z_0\|\sqrt{T\lambda(0)}, \quad \forall t \in [0, T], \tag{7}$$

where $\delta = \|B\|_{\mathcal{L}(V,V')}$ and $\lambda(t) = \displaystyle\int_t^{t+T} |\langle i_\omega Bz(s), z(s)\rangle_{V',V}|^2 ds$.
For all $s \geq 0$, we have

$$\begin{aligned} \langle i_\omega BS(s)z_0, S(s)z_0\rangle_{V',V} &= -\langle i_\omega B(z(s) - S(s)z_0), S(s)z_0)\rangle_{V',V} \\ &- \langle i_\omega Bz(s), (z(s) - S(s)z_0)\rangle_{V',V} + \langle i_\omega Bz(s), z(s)\rangle_{V',V}. \end{aligned}$$

Since $\chi_\omega$ is continuous, there exists a constant $C > 0$ such that

$$|\langle i_\omega BS(s)z_0, S(s)z_0\rangle_{V',V}| \leq 2C\delta\|z_0\|\|z(s) - S(s)z_0\| + |\langle i_\omega Bz(s), z(s)\rangle_{V',V}|. \tag{8}$$

Now, let $\Gamma(t)z_0 := z(t)$ define a non-linear semigroup of contractions on $H$ and replacing $z_0$ by $\Gamma(t)z_0$ in (7) and (8), we have

$$\begin{aligned} |\langle i_\omega BS(s)\Gamma(t)z_0, S(s)\Gamma(t)z_0\rangle_{V',V}| &\leq 2C\delta^2\|z_0\|^2\sqrt{T\lambda(t)} \\ &+ |\langle i_\omega B\Gamma(s+t)z_0, \Gamma(s+t)z_0\rangle_{V',V}|, \quad \forall t, s \geq 0. \end{aligned} \tag{9}$$

Integrating (9) over the interval $[0,T]$ and using Schwartz's inequality, we arrive at

$$\int_0^T |\langle i_\omega BS(s)\Gamma(t)z_0, S(s)\Gamma(t)z_0\rangle_{V',V}|ds \leq M\sqrt{\lambda(t)}, \tag{10}$$

where $M = \left(2C\delta^2\|z_0\|^2T^{\frac{3}{2}} + \sqrt{T}\right)$ is a non-negative constant depending on $\|z_0\|$ and $T$. By virtue of (4), we have

$$\int_0^{+\infty} |\langle i_\omega B\Gamma(s)z_0, \Gamma(s)z_0\rangle_{V',V}|^2 ds < +\infty.$$

From the Cauchy criterion, we deduce that

$$\lambda(t) \to 0, \text{ as } t \to \infty. \tag{11}$$

To show that $\chi_\omega z(t) \rightharpoonup 0$, as $t \to +\infty$, let us consider a sequence $(t_n) \subset \mathbb{R}$ such that $t_n \to \infty$.
From (6), there exists a subsequence $(t_{\varphi(n)})$ of $(t_n)$ such that

$$\Gamma(t_{\varphi(n)})z_0 \rightharpoonup y \text{ in } V \text{ ,as } n \to \infty.$$

Using the continuity of $\chi_\omega$ and since $B$ is a compact operator from $V$ to $V'$, we have for all $t \geq 0$ that

$$S(t)\Gamma(t_{\varphi(n)})z_0 \rightharpoonup S(t)y \text{ in } V \text{ and } BS(t)\Gamma(t_{\varphi(n)})z_0 \rightharpoonup BS(t)y \text{ in } V', \text{ as } n \to \infty.$$

Then

$$\lim_{n\to\infty} \langle i_\omega BS(t)\Gamma(t_{\varphi(n)})z_0, S(t)\Gamma(t_{\varphi(n)}z_0)\rangle_{V',V} = \langle i_\omega BS(t)y, S(t)y\rangle_{V',V}.$$

By the dominated convergence theorem, we have

$$\lim_{n\to\infty} \int_0^T |\langle i_\omega BS(t)\Gamma(t_{\varphi(n)})z_0, S(t)\Gamma(t_{\varphi(n)}z_0)\rangle_{V',V}|dt = \int_0^T |\langle i_\omega BS(t)y, S(t)y)\rangle_{V',V}|dt.$$

From (10) and (11), it follows that

$$\int_0^T |\langle i_\omega BS(t)y, S(t)y)\rangle_{V',V}|dt = 0.$$

and then

$$\langle i_\omega BS(t)y, S(t)y)\rangle_{V',V} = 0, \ \forall t \in [0,T].$$

Using (3), we have

$$\chi_\omega\Gamma(t_{\varphi(n)})z_0 \rightharpoonup 0 \text{ as } n \to +\infty. \tag{12}$$

On the other hand, it is clear that (12) holds for each subsequence $(t_{\phi(n)})$ of $(t_n)$ such that $\chi_\omega\Gamma(t_{\phi(n)})z_0$ weakly converges in $L^2(\omega)$. This shows that for all $\varphi \in L^2(\omega)$, $\langle\chi_\omega\Gamma(t_n)z_0, \varphi\rangle \rightharpoonup 0$, as $n \to +\infty$. Hence $\chi_\omega\Gamma(t)z_0 \rightharpoonup 0$, as $t \to +\infty$, which completes the proof. □

**Remark 2.2** In the case $\omega = \Omega$, we retrieve the result established in [2] concerning the weak stabilisation of system (1) on the whole domain $\Omega$.

### 2.3 Regional strong stabilization

The following result gives sufficient conditions for strong stabilization of system (1) on $\omega$.

**Theorem 2.2** *Let $A$ generate a semigroup $S(t)$ of contractions on $H$, the assumptions $(H_1), (H_2), (H_3)$ hold, and assume that the condition*

$$\int_0^T |\langle i_\omega BS(t)y, S(t)y\rangle_{V',V}|dt \geq \mu \|\chi_\omega y\|^2_{L^2(\omega)}, \text{ for some } T, \mu > 0 \tag{13}$$

*is verified, then the control $v(t) = -\langle i_\omega Bz(t), z(t)\rangle_{V',V}$ strongly stabilizes system (1) on $\omega$ with the following decay estimate*

$$\|\chi_\omega z(t)\|_{L^2(\omega)} = O(t^{-1/2}), \quad as\ t \to +\infty.$$

***Proof.*** From (10) and (13), we deduce that

$$\beta\sqrt{\lambda(kT)} \geq \|\chi_\omega \Gamma(kT)z_0\|^2, \quad \forall k \geq 0, \tag{14}$$

where $\beta = \dfrac{1}{\mu}M$.
Integrating the following inequality

$$\frac{d}{dt}\|\chi_\omega \Gamma(t)z_0\|^2 \leq -2|\langle i_\omega B\Gamma(t)z_0, \Gamma(t)z_0\rangle_{V',V}|^2$$

from $kT$ to $(k+1)T$, $(k \in \mathbb{N})$, and using (14), we obtain

$$\|\chi_\omega \Gamma(kT)z_0\|^2 - \|\chi_\omega \Gamma(kT+T)z_0\|^2 \geq 2\lambda(kT), \quad \forall k \geq 0.$$

It follows that

$$\beta^2\|\chi_\omega \Gamma(kT+T)z_0\|^2 - \beta^2\|\chi_\omega \Gamma(kT)z_0\|^2 \leq -2\|\chi_\omega \Gamma(kT)z_0\|^4, \quad \forall k \geq 0. \tag{15}$$

Let us introduce the sequence $s_k = \|\chi_\omega \Gamma(kT)z_0\|^2, \ \forall k \geq 0$.
From (15), we deduce that

$$\frac{s_k - s_{k+1}}{s_k^2} \geq \frac{2}{\beta^2}, \quad \forall k \geq 0.$$

Since the sequence $(s_k)$ decreases, we get

$$\frac{s_k - s_{k+1}}{s_k . s_{k+1}} \geq \frac{2}{\beta^2}, \quad \forall k \geq 0,$$

so

$$\frac{1}{s_{k+1}} - \frac{1}{s_k} \geq \frac{2}{\beta^2}, \quad \forall k \geq 0.$$

We deduce that

$$s_k \leq \frac{s_0}{\frac{2s_0}{\beta^2}k + 1}, \quad \forall k \geq 0.$$

Finally, introducing the integer part $k = E(\frac{t}{T})$ and using the fact that $\|\chi_\omega \Gamma(t)z_0\|$ decreases, we deduce the estimate

$$\|\chi_\omega z(t)\| = O(t^{-1/2}), \quad \text{as } t \to +\infty.$$

□

## 3 Example and Simulations

Let us consider the system defined in $\Omega = ]0, +\infty[$ by the following equation

$$\begin{cases} \dfrac{\partial z(x,t)}{\partial t} = -\dfrac{\partial z(x,t)}{\partial x} + v(t)b(x)z(x,t), & \Omega \times ]0, +\infty[, \\ z(0,t) = 0, & ]0, +\infty[, \\ z(x,0) = z_0(x), & \Omega, \end{cases} \tag{16}$$

where $H = L^2(\Omega)$, $b(x) \geq 0$, a.e on $\Omega$, and $b(x) \geq c > 0$ a.e on $\omega$, $Az = -\dfrac{\partial z}{\partial x}$ with the domain $\mathcal{D}(A) = \{z \in H^1(\Omega) \mid z(0) = 0\}$, and consider the operator $B : \mathcal{D}(B)(\subset L^2(\Omega)) \to L^2(\Omega)$ given by $Bz = b(x)z$. The operator $B$ is unbounded on $L^2(\Omega)$. By the Sobolev embeddings $H^1(\Omega) \hookrightarrow L^2(\Omega)$, $B$ is bounded from $H^1(\Omega)$ to $L^2(\Omega)$. Hence, the space $V$ is given by $V = H^1(\Omega)$.

The operator $A$ generates the following semi-group of contractions

$$(S(t)z_0)(x) = \begin{cases} z_0(x-t), & if\ x \geq t, \\ 0, & if\ x < t. \end{cases}$$

Let $\omega = ]0, a[$, with $a > 0$ we have

$$\langle i_\omega Az, z \rangle \quad = \quad -\int_0^a z'(x)z(x)dx = -\frac{z^2(a)}{2} \leq 0,$$

so hypothesis $(H_1)$ holds. For $T = 1$, we have

$$\int_0^1 \langle i_\omega b(x)S(t)z_0, S(t)z_0 \rangle_{V',V} dt = \int_0^1 \int_0^{1-t} b(x)|z_0(x)|^2 dx dt \geq c\|\chi_\omega z_0\|^2.$$

Then, the control $v(t) = -\displaystyle\int_0^a b(x)|z(x,t)|^2 dx$ strongly stabilizes system (16) on $\omega$.

We consider system (16) with $b(x) = \dfrac{1}{\sqrt{x(x^2+1)}}$ and $z_0(x) = \sin(\pi x)$.

- For $\omega = ]0, 2[$, we have

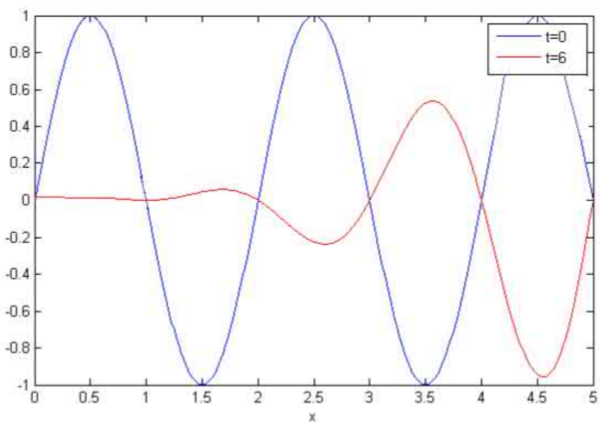


**Figure 1**: The stabilization on $\omega = ]0, 2[$.

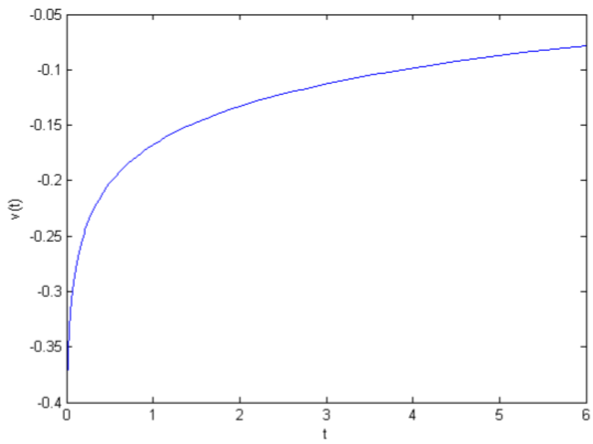


**Figure 2**: Control function.

Figure 1 shows that the system (16) is strongly stabilized on $\omega = ]0, 2[$.

• For $\omega = ]0, 3[$, we have

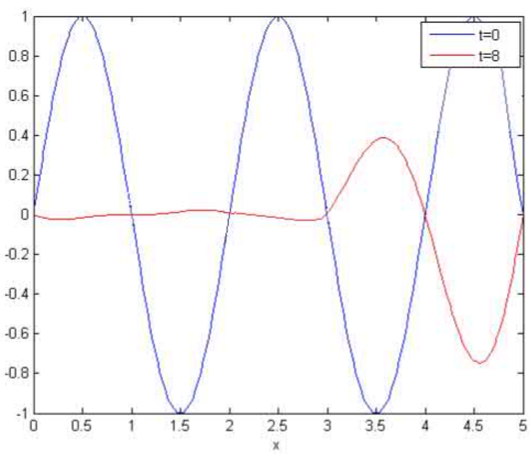


**Figure 3**: The stabilization on $\omega = ]0, 3[$.

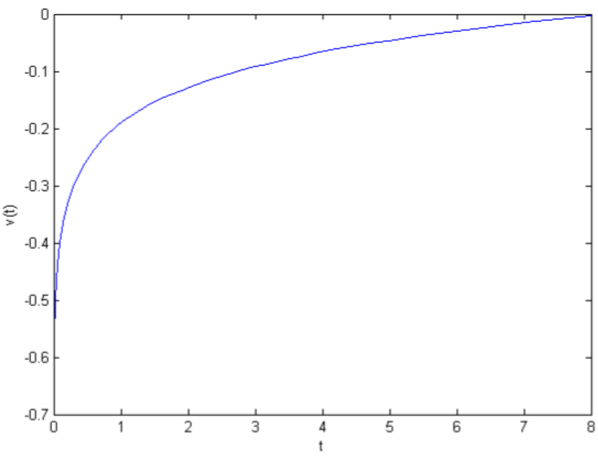


**Figure 4**: Control function.

Figure 3 shows that the system (16) is strongly stabilized on $\omega = ]0,3[$.

## 4 Conclusion

Regional stabilization of a class of infinite dimensional bilinear systems with unbounded control operator is considered. Under sufficient conditions, we give a control that ensures weak and strong regional stabilization. Questions are still open; this is, the case of boundary subregion.

## Acknowledgment

The work has been carried out with a grant from Hassan II Academy of Sciences and Technology.

## References

[1] Ball, J.M. and Slemrod, M. Feedback stabilization of distributed semilinear control systems. *Journal of Applied Mathematics and Optimization* **5** (1) (1979) 169–179.

[2] Berrahmoune, L. Stabilization of unbounded bilinear control systems in Hilbert space. *Journal of Mathematical Analysis and Applications* **372** (2) (2010) 645–655.

[3] Bounit, H. and Hammouri, H. Feedback stabilization for a class of distributed semilinear control systems. *Journal of Nonlinear Analysis* **37** (8) (1999) 953–969.

[4] Mohler, R.R. *Bilinear Control Processes.* Academic Press, New York, 1973.

[5] Ouzahra, M. Exponential and weak stabilization of constrained bilinear systems. *SIAM J. Control Optim* **48** (6) (2010) 3962-3974.

[6] Zerrik, E, Ouzahra, M. and Ztot, K. Regional stabilization for infinite bilinear systems. *IEE Proceeding of Control Theory and Applications* **151** (1) (2004) 109–116.

[7] Zerrik, E. and Ouzahra, M. Regional stabilization for infinite-dimensional systems. *International Journal of Control* **76** (1) (2003) 73–81.

Nonlinear Dynamics and Systems Theory, 18 (4) (2018) 426–427

## Contents of Volume 18, 2018

**Volume 18** **Number 1** **March 2018**



**Volume 18** **Number 2** **June 2018**

**Volume 18** **Number 3** **September 2018**



**Volume 18** **Number 4** **December 2018**

**BIRKHÄUSER**

**Book Series : Systems & Control: Foundations & Applications**

**Stability Theory for Dynamic Equations on Time Scales**

*Book Series: Systems & Control: Foundations & Applications: 223 pp., 2016*
*ISBN 978-3-319-42212-1, DOI 10.1007/978-3-319-42213-8,*

**A.A. Martynyuk**
*Institute of Mechanics, National Academy of Sciences of Ukraine, Kyiv, Ukraine*

This monograph is a first in the world to present three approaches for stability analysis of solutions of dynamic equations. The first approach is based on the application of dynamic integral inequalities and the fundamental matrix of solutions of linear approximation of dynamic equations. The second is based on the generalization of the direct Lyapunov method for equations on time scales, using scalar, vector and matrix-valued auxiliary functions. The third approach is the application of auxiliary functions (scalar, vector, or matrix-valued ones) in combination with differential dynamic inequalities. This is an alternative comparison method, developed for time continuous and time discrete systems. In recent decades, automatic control theory in the study of air- and spacecraft dynamics and in other areas of modern applied mathematics has encountered problems in the analysis of the behavior of solutions of time continuous-discrete linear and/or nonlinear equations of perturbed motion. In the book “Men of Mathematics,” 1937, E.T.Bell wrote: “A major task of mathematics today is to harmonize the continuous and the discrete, to include them in one comprehensive mathematics, and to eliminate obscurity from both.” Mathematical analysis on time scales accomplishes exactly this.

**CONTENTS** (5 chapters)



The Stability Theory for Dynamic Equations on Time Scales fulfills the reference needs of pure and applied mathematicians, applied physicist, industrial engineerings, operations researchers, and upper-level undegraduate and graduate students studying ordinary differential, difference and dynamic equations on time scales. This research has potential applications in such areas as theoretical and applied mechanics, neurodynamics, mathematical biology and finance among others.

---



---