Springer Optimization and Its Applications 111

Themistocles M. Rassias
Vijay Gupta *Editors*

Mathematical Analysis, Approximation Theory and Their Applications

Springer

Springer Optimization and Its Applications

VOLUME 111

Aims and Scope

Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics, and other sciences.

The series *Springer Optimization and Its Applications* publishes undergraduate and graduate textbooks, monographs and state-of-the-art expository work that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

More information about this series at http://www.springer.com/series/7393

Themistocles M. Rassias • Vijay Gupta
Editors

Mathematical Analysis, Approximation Theory and Their Applications

Springer

Editors
Themistocles M. Rassias
Department of Mathematics
National Technical Unversity of Athens
Athens, Greece

Vijay Gupta
Department of Mathematics
Netaji Subhas Institute of Technology
New Delhi, India

ISSN 1931-6828 ISSN 1931-6836 (electronic)
Springer Optimization and Its Applications
ISBN 978-3-319-31279-8 ISBN 978-3-319-31281-1 (eBook)
DOI 10.1007/978-3-319-31281-1

Library of Congress Control Number: 2016934411

Mathematics Subject Classification (2010): 15-XX, 26-XX, 30-XX, 33-XX, 39-XX, 41-XX, 43-XX, 44-XX, 46-XX, 47-XX, 65-X

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Mathematical Analysis, Approximation Theory, and Their Applications highlights the classical and new results in important areas of mathematical research. These include essential achievements in pure mathematical analysis and approximation theory, as well as some of their fascinating applications. The contributed papers have been written by experts from the international mathematical community. These papers deepen our understanding of some of the current research problems and theories.

The presentation of concepts and methods discussed here make it an invaluable source for a wide readership.

We would like to express our deepest thanks to all of the scientists who contributed to this book. We would also wish to acknowledge the superb assistance that the staff of Springer has provided for the publication of this book.

Athens, Greece — Themistocles M. Rassias
New Delhi, India — Vijay Gupta

Contents

Contributors

Ulrich Abel Technische Hochschule Mittelhessen, Fachbereich MND, Friedberg, Germany

P.N. Agrawal Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee, India

Douglas R. Anderson Department of Mathematics, Concordia College, Moorhead, MN, USA

V. Berinde Technical University Cluj-Napoca, Baia-Mare North University Center, Baia Mare, Romania

Daniel Cárdenas-Morales Department of Mathematics, University of Jaén, Jaén, Spain

O. Chau University of La Réunion, PIMENT EA4518, Saint-Denis, France

Monica-Gabriela Cojocaru Department of Mathematics & Statistics, University of Guelph, Guelph, ON, Canada

Stefan Czerwik Institute of Mathematics, Silesian University of Technology, Gliwice, Poland

Silvestru Sever Dragomir Mathematics, College of Engineering & Science, Victoria University, Melbourne, VIC, Australia

School of Computational & Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa

Elqorachi Elhoucien Faculty of Sciences, Department of Mathematics, Ibn Zohr University, Agadir, Morocco

Sorin G. Gal Department of Mathematics and Computer Science, University of Oradea, Oradea, Romania

Pedro Garrancho Department of Mathematics, University of Jaén, Jaén, Spain

Ioan Gavrea Department of Mathematics, Technical University of Cluj Napoca, Cluj-Napoca, Romania

Laura Găvruţa Department of Mathematics, Politehnica University of Timişoara, Timişoara, Romania

Paşc Găvruţa Department of Mathematics, Politehnica University of Timişoara, Timişoara, Romania

D. Goeleven University of La Réunion, PIMENT EA4518, Saint-Denis, France

N.K. Govil Department of Mathematics and Statistics, Auburn University, Auburn, AL, USA

Meenu Goyal Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee, India

Scott Greenhalgh Department of Mathematics & Statistics, Queen's University, Kingston, ON, Canada

Vijay Gupta Department of Mathematics, Netaji Subhas Institute of Technology, New Delhi, India

Mircea Ivan Department of Mathematics, Technical University of Cluj Napoca, Cluj-Napoca, Romania

Joaquín Jódar-Reyes Universidad de Jaén, Jaén, Spain

Yiannis Karagiorgos Department of Mathematics, National Technical University of Athens, Athens, Greece

Sammad Khalil Faculty of Sciences, Department of Mathematics, Ibn Tofail University, Kenitra, Morocco

Dae San Kim Department of Mathematics, Sogang University, Seoul, Republic of Korea

Taekyun Kim Department of Mathematics, Kwangwoon University, Seoul, Republic of Korea

Christos D. Kravvaritis Department of Mathematics, University of Athens, Athens, Greece

Krzysztof Król Education Department, Zabrze, Poland

José-Manuel Latorre-Palacios Universidad de Jaén, Jaén, Spain

Antonio-Jesús López-Moreno Universidad de Jaén, Jaén, Spain

Youssef Manar Superior School of Technology, Ibn Zohr University, Guelmim, Morocco

H.N. Mhaskar Department of Mathematics, California Institute of Technology, Pasadena, CA, USA

Institute of Mathematical Sciences, Claremont Graduate University, Claremont, CA, USA

Gradimir V. Milovanović The Serbian Academy of Sciences and Arts, Belgrade, Serbia

Eze R. Nwaeze Department of Mathematics and Statistics, Auburn University, Auburn, AL, USA

Sofiya Ostrovska Department of Mathematics, Atilim University, Ankara, Turkey

R. Oujja University of La Réunion, PIMENT EA4518, Saint-Denis, France

Ahmet Yaşar Özban Department of Mathematics, Atilim University, Ankara, Turkey

Basil K. Papadopoulos School of Engineering, Democritus University of Thrace, Xanthi, Greece

Kyriakos Papadopoulos College of Engineering, American University of the Middle East, Egaila, Kuwait

I.N. Parasidis Department of Electrical Engineering, TEI of Thessaly, Larissa, Greece

Dorothea Petraki Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece

A. Petruşel Department of Mathematics, Babeş-Bolyai University, Cluj-Napoca, Romania

E. Providas Department of Electrical Engineering, TEI of Thessaly, Larissa, Greece

I.A. Rus Department of Mathematics, Babeş-Bolyai University, Cluj-Napoca, Romania

Nikolaos Samaras Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece

M.A. Şerban Department of Mathematics, Babeş-Bolyai University, Cluj-Napoca, Romania

H.M. Srivastava Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada

Marija P. Stanić Faculty of Science, Department of Mathematics and Informatics, University of Kragujevac, Kragujevac, Serbia

Gancho Tachev Department of Mathematics, University of Architecture, Sofia, Bulgaria

Mihai Turinici “A. Myller” Mathematical Seminar, “A. I. Cuza” University, Iaşi, Romania

Christos Tzimopoulos School of Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

Rajendra G. Vyas Faculty of Science, Department of Mathematics, The M. S. University of Baroda, Vadodara, India

Nikos Yannakakis Department of Mathematics, National Technical University of Athens, Athens, Greece

K.V. Zhukovsky Faculty of Physics, Moscow State University, Moscow, Russian Federation

A Brief History of the Favard Operator and Its Variants

Ulrich Abel

Abstract In the year 1944 the well-known French mathematician Jean Favard (1902–1965) introduced a discretely defined operator which is a discrete analogue of the familiar Gauss–Weierstrass singular convolution integral. In the present chapter we sketch the history of this approximation operator during the past 70 years by presenting known results on the operator and its various extensions and variants. The first part after the introduction is dedicated to saturation of the classical Favard operator in weighted Banach spaces. Furthermore, we discuss the asymptotic behaviour of a slight generalization F_{n,σ_n} of the Favard operator and its Durrmeyer variant $\tilde{F}_{n,\sigma_n}$. In particular, the local rate of convergence when applied to locally smooth functions is considered. The main result of this part consists of the complete asymptotic expansions for the sequences $(F_{n,\sigma_n} f)(x)$ and $\left(\tilde{F}_{n,\sigma_n} f\right)(x)$ as n tends to infinity. Furthermore, these asymptotic expansions are valid also with respect to simultaneous approximation. A further part is devoted to the recent work of several Polish mathematicians on approximation in weighted function spaces. Finally, we define left quasi-interpolants for the Favard operator and its Durrmeyer variant in the sense of Sablonnière.

Keywords Approximation by positive operators • Rate of convergence • Degree of approximation.

Introduction

In the year 1944 J. Favard [7, pp. 229, 239] introduced the operator F_n defined by

$$(F_n f)(x) = \frac{1}{\sqrt{\pi n}} \sum_{\nu=-\infty}^{\infty} f\left(\frac{\nu}{n}\right) \exp\left(-n\left(\frac{\nu}{n} - x\right)^2\right) \tag{1}$$

U. Abel (✉)
Technische Hochschule Mittelhessen, Fachbereich MND, Wilhelm-Leuschner-Straße 13, 61169 Friedberg, Germany
e-mail: Ulrich.Abel@mnd.thm.de

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_1

which is a discrete analogue of the familiar Gauss–Weierstrass singular convolution integral

$$(W_n f)(x) = \sqrt{\frac{n}{\pi}} \int_{-\infty}^{\infty} f(t) \exp\left(-n(t-x)^2\right) dt$$

(see, e.g., the book [6, Eq. (3.1.32), p. 125], where the denotation $W(f; x; 1/(4n))$ is used) in order to approximate functions defined on the real axis. The operators can be applied to functions f defined on $\mathbb{R}$ satisfying the growth condition

$$f(t) = O\left(e^{Kt^2}\right) \qquad \text{as } |t| \to \infty, \tag{2}$$

for a constant $K > 0$. Favard proved that $(F_n f)(x)$ converges to $f(x)$ as $n \to \infty$ pointwise for every $x \in \mathbb{R}$, and even uniformly on any compact subinterval of $\mathbb{R}$, for all functions f continuous on $\mathbb{R}$ satisfying (2).

In the first part (section "Saturation in Weighted Banach Spaces") we present results by Becker et al. [5] on saturation in weighted Banach spaces consisting of functions of polynomial and exponential growth, respectively.

The second part (section "Complete Asymptotic Expansions") contains further developments with respect to results by Gawronski and Stadtmüller [8]. They considered in 1982, for a sequence of positive reals σ_n, the generalization

$$(F_{n,\sigma_n} f)(x) = \sum_{\nu=-\infty}^{\infty} p_{n,\nu,\sigma_n}(x) f\left(\frac{\nu}{n}\right), \tag{3}$$

where

$$p_{n,\nu,\sigma_n}(x) = \frac{1}{\sqrt{2\pi} n \sigma_n} \exp\left(-\frac{1}{2\sigma_n^2}\left(\frac{\nu}{n} - x\right)^2\right).$$

The particular case $\sigma_n^2 = 1/(2n)$ reduces to Favard's classical operator (1).

Under certain conditions on f and $(\sigma_n)_{n\in\mathbb{N}}$ the operators possess the basic property that $(F_{n,\sigma_n} f)(x)$ converges to $f(x)$ in each continuity point x of f. Among other results, Gawronski and Stadtmüller [8, Eq. (0.6)] established the Voronovskaja-type theorem

$$\lim_{n\to\infty} \frac{1}{\sigma_n^2}\left[(F_{n,\sigma_n} f)(x) - f(x)\right] = \frac{1}{2} f''(x) \tag{4}$$

uniformly on proper compact subsets of $[a, b]$, for $f \in C^2[a, b]$ $(a, b \in \mathbb{R})$ and $\sigma_n \to 0$ as $n \to \infty$, provided that certain conditions on the first three moments of F_{n,σ_n} are satisfied. Actually, Eq. (4) was proved for a truncated variant of (3) which possesses the same asymptotic properties as (3) [8, cf. Theorem 1 (iii) and

Remark (i), p. 393]. For a Voronovskaja-type theorem (cf. [21]) in the particular case $\sigma_n^2 = \gamma/(2n)$ see [5, Theorem 4.3]. Abel and Butzer extended Formula (4) by deriving a complete asymptotic expansion of the form

$$F_{n,\sigma_n} f \sim f + \sum_{k=0}^{\infty} c_k(f)\,\sigma_n^k \qquad (n \to \infty)$$

for f sufficiently smooth. The latter formula means that for all positive integers q there holds

$$F_{n,\sigma_n} f = f + \sum_{k=1}^{q} c_k(f)\,\sigma_n^k + o\left(\sigma_n^q\right) \qquad (n \to \infty)\,.$$

The coefficients c_k, which depend on f but are independent of n, are explicitly determined. It turns out that $c_k(f) = 0$, for all odd integers $k > 0$. Moreover, we present results on simultaneous approximation by the operators (3). A truncated version of them was defined by Gawronski and Stadtmüller [8, cf. Eq. (0.7)].

In 1999 Pych-Taberska and Nowak [18] defined the Kantorovich variant

$$\left(\hat{F}_{n,\sigma_n} f\right)(x) = n \sum_{\nu=-\infty}^{\infty} p_{n,\nu,\sigma_n}(x) \int_{\nu/n}^{(\nu+1)/n} p_{n,\nu,\sigma_n}(t) f(t)\,dt \tag{5}$$

and in 2001 a Durrmeyer variant [12–15]

$$\left(\tilde{F}_{n,\sigma_n} f\right)(x) = n \sum_{\nu=-\infty}^{\infty} p_{n,\nu,\sigma_n}(x) \int_{-\infty}^{\infty} p_{n,\nu,\sigma_n}(t) f(t)\,dt \tag{6}$$

of Favard operators. The third part (section "Approximation in Weighted Spaces") reports about recent results of several Polish mathematicians on approximation by those variants in weighted function spaces.

Also for the Durrmeyer variant (6) a complete asymptotic expansion

$$\tilde{F}_{n,\sigma_n} f \sim f + \sum_{k=0}^{\infty} \tilde{c}_k(f)\,\sigma_n^k \qquad (n \to \infty)\,,$$

for f sufficiently smooth will be presented. The coefficients $\tilde{c}_k$, which depend on f but are independent of n, are explicitly determined. As in the case of the Favard operators (3) we have $c_k(f) = 0$, for all odd integers $k > 0$. Moreover, results concerning simultaneous approximation by the operators (6) are included.

Finally, in section "Quasi-Interpolants" we define left quasi-interpolants (LQI) for the Favard operator and its Durrmeyer variant in the sense of Sablonnière and present their asymptotic expansions.

For more results on approximation by linear operators see, e.g., the recent textbook [9].

Saturation in Weighted Banach Spaces

In this section we sketch the main results by Becker et al. [5]. Actually, the authors considered the slightly more general operators F_{n,σ_n} when $\sigma_n^2 = \gamma / (2n)$ with a constant $\gamma > 0$. For the sake of simplicity we restrict our representation to the case $\gamma = 1$ which is Favard's classical operator (1).

Functions of Polynomial Growth

For $N \in \mathbb{N}$, define the weight function

$$w_N (x) = \left(1 + x^{2N}\right)^{-1} \tag{7}$$

and consider the Banach space

$$C_N (\mathbb{R}) := \left\{f \in C (\mathbb{R}) : f (x) = o \left(w_N^{-1} (x)\right) \text{ for } |x| \to +\infty\right\} \tag{8}$$

equipped with the norm

$$\|f\|_N := \|w_N f\|_\infty = \sup_{x\in\mathbb{R}} \left|\left(1 + x^{2N}\right)^{-1} f (x)\right| . \tag{9}$$

The Favard operators $F_n : C_N (\mathbb{R}) \to C_N (\mathbb{R})$ are bounded operators, more precisely there holds

Theorem 1 (Boundedness). *For each $N \in \mathbb{N}$, there exists a constant M_N such that*

$$\|F_n f\|_N \le \left(1 + \frac{M_N}{n}\right) \|f\|_N \qquad (f \in C_N (\mathbb{R})) .$$

Furthermore, the sequence $(F_n)_{n\in\mathbb{N}}$ constitutes an approximation process on each space $C_N (\mathbb{R})$.

Theorem 2 (Approximation). *For each $N \in \mathbb{N}$, there holds*

$$\lim_{n\to\infty} \|F_n f - f\|_N = 0 \qquad (f \in C_N (\mathbb{R})) .$$

If the Favard operators act on the subspace

$$D_N(\mathbb{R}) := \left\{f \in C_N(\mathbb{R}) : f', f'' \in C_N(\mathbb{R})\right\}$$

we have the following Voronovskaja-type formula.

Theorem 3 (Voronovskaja-Type Result). *For all $f \in D_N(\mathbb{R})$, there holds the asymptotic formula*

$$\lim_{n\to\infty} \left\| n\left(F_n f - f\right) - \frac{1}{4} f'' \right\|_N = 0.$$

A certain stability condition which follows by the boundedness and the Voronovskaja-type formula are the main assumptions of a general theorem by Trotter which implies the following saturation result.

Theorem 4 (Saturation). *Let $f \in C_N(\mathbb{R})$. Then the following three statements are equivalent:*

$$\|F_n f - f\|_N = O\left(n^{-1}\right) \qquad (n \to \infty),$$

$$\|f(x+h) - 2f(x) + f(x-h)\|_N = O\left(h^2\right) \qquad (h \to 0),$$

$$f \in \widetilde{D}_N(\mathbb{R}) \qquad (\textit{the completion relative to } C_N(\mathbb{R})).$$

An equivalence result for the Favard operator in polynomial weight spaces can be found in [4].

Functions of Exponential Growth

For real numbers $\beta > 0$, define the weight function

$$w_\beta(x) = e^{-\beta x^2} \tag{10}$$

and the norm

$$\|f\|_\beta := \|wf\|_\infty = \sup_{x\in\mathbb{R}} \left| e^{-\beta x^2} f(x) \right|.$$

We consider the space

$$X := \bigcap_{\beta>0} \left\{f \in C(\mathbb{R}) : \|f\|_\beta < +\infty\right\}.$$

Analogously to the preceding subsection there hold the following theorems.

Theorem 5 (Approximation). *For $\beta > 0$, there holds*

$$\lim_{n\to\infty} \|F_n f - f\|_\beta = 0 \qquad (f \in X) .$$

Theorem 6 (Voronovskaja-Type Result). *For all real numbers $\beta > 0$ and $f \in D' := \{f \in X : f', f'' \in X\}$, we have*

$$\lim_{n\to\infty} \left\| n\left(F_n f - f\right) - \frac{1}{4} f'' \right\|_\beta = 0.$$

Theorem 7 (Saturation). *Let $f \in X$. Then the following two statements*

$$\|F_n f - f\|_\beta = O\left(n^{-1}\right) \qquad (n \to \infty)$$

and

$$\|f(x+h) - 2f(x) + f(x-h)\|_\beta = O\left(h^2\right) \qquad (h \to 0)$$

are equivalent.

The paper [5] does not contain a result on the saturation class with regard to the space X.

Complete Asymptotic Expansions

Throughout this section we assume that

$$\sigma_n > 0, \quad \sigma_n \to 0, \quad \sigma_n^{-1} = O\left(n^{1-\eta}\right) \quad (n \to \infty) \tag{11}$$

with (an arbitrarily small) constant $\eta > 0$. Note that the latter condition implies that $n\sigma_n \to \infty$ as $n \to \infty$.

The following theorem (see [1–3]) presents the complete asymptotic expansion for the sequences $(F_{n,\sigma_n})(x)$ and $\left(\tilde{F}_{n,\sigma_n}\right)(x)$ as $n \to \infty$. For $r \in \mathbb{N}$ and $x \in \mathbb{R}$ let $W[r;x]$ be the class of functions on $\mathbb{R}$ satisfying growth condition (2), which admit a derivative of order r at the point x.

Theorem 8. *Let $q \in \mathbb{N}$ and $x \in \mathbb{R}$. Suppose that the real sequence (σ_n) satisfies the conditions (11). For each function $f \in W[2q;x]$, the generalized Favard operators (3) possess the complete asymptotic expansion*

$$(F_{n,\sigma_n} f)(x) = f(x) + \sum_{k=1}^{q} \frac{f^{(2k)}(x)}{(2k)!!} \sigma_n^{2k} + o\left(\sigma_n^{2q}\right) \tag{12}$$

and

$$\left(\tilde{F}_{n,\sigma_n}f\right)(x) = f(x) + \sum_{k=1}^{q} \frac{f^{(2k)}(x)}{k!}\sigma_n^{2k} + o\left(\sigma_n^{2q}\right) \tag{13}$$

as $n \to \infty$.

Here $m!!$ denote the double factorial numbers defined by $0!! = 1!! = 1$ and $m!! = m \cdot (m-2)!!$ for integers $m \geq 2$. Note that the asymptotic expansions contain only terms with even order derivatives of the function f.

The crucial tool for the proof of Theorem 8 is a classical transformation formula for the elliptic theta function (see, e.g., [6, p. 126]). An immediate consequence is the Voronosvkaja-type theorem of [8, Eq. (0.6), Theorem 1 (iii) and Remark (i)].

Corollary 1. *Let* $x \in \mathbb{R}$*. Suppose that the real sequence* (σ_n) *satisfies the conditions* (11)*. For each function* $f \in W[2;x]$*, we have the asymptotic relations*

$$\lim_{n\to\infty} \sigma_n^{-2} \left((F_{n,\sigma_n}f)(x) - f(x)\right) = \frac{1}{2} f''(x)$$

and

$$\lim_{n\to\infty} \sigma_n^{-2} \left(\left(\tilde{F}_{n,\sigma_n}f\right)(x) - f(x)\right) = f''(x).$$

Concerning simultaneous approximation, it turns out that the complete asymptotic expansion can be obtained by differentiating (12) term-by-term (see [1–3]).

Theorem 9. *Let* $\ell \in \mathbb{N}_0$*,* $q \in \mathbb{N}$ *and* $x \in \mathbb{R}$*. Suppose that the real sequence* (σ_n) *satisfies condition* (11)*. For each function* $f \in W[2(\ell+q);x]$*, the following complete asymptotic expansions are valid as* $n \to \infty$*:*

$$(F_{n,\sigma_n}f)^{(\ell)}(x) = f^{(\ell)}(x) + \sum_{k=1}^{q} \frac{f^{(2k+\ell)}(x)}{(2k)!!}\sigma_n^{2k} + o\left(\sigma_n^{2q}\right) \tag{14}$$

and

$$\left(\tilde{F}_{n,\sigma_n}f\right)^{(\ell)}(x) = f^{(\ell)}(x) + \sum_{k=1}^{q} \frac{f^{(2k+\ell)}(x)}{k!}\sigma_n^{2k} + o\left(\sigma_n^{2q}\right). \tag{15}$$

Remark 1. The latter formulas can be written in the equivalent form

$$\lim_{n\to\infty} \sigma_n^{-2q} \left((F_{n,\sigma_n}f)^{(\ell)}(x) - f^{(\ell)}(x) - \sum_{k=1}^{q} \frac{f^{(2k+\ell)}(x)}{(2k)!!}\sigma_n^{2k} \right) = 0,$$

$$\lim_{n\to\infty} \sigma_n^{-2q} \left(\left(\tilde{F}_{n,\sigma_n}f\right)^{(\ell)}(x) - f^{(\ell)}(x) - \sum_{k=1}^{q} \frac{f^{(2k+\ell)}(x)}{k!}\sigma_n^{2k} \right) = 0.$$

The proof of Eq. (14) in Theorem 9 essentially uses a representation of $(F_{n,\sigma_n}f)^{(\ell)}$ in terms of F_{n,σ_n} which follows by a certain identity of Hermite polynomials. A similar representation is used to derive Eq. (15).

Assuming smoothness of f on intervals $I = (a, b)$, $a, b \in \mathbb{R}$, it can be shown that the above expansions hold uniformly on compact subsets of I.

The proofs are based on a localization result which is interesting in itself. We quote only the result for the ordinary Favard operator (3).

Proposition 1. *Fix $x \in \mathbb{R}$ and let $\delta > 0$. Assume that the function $f : \mathbb{R} \to \mathbb{R}$ vanishes in $(x-\delta, x+\delta)$ and satisfies, for positive constants M_x, K_x, the growth condition*

$$|f(t)| \leq M_x e^{K_x (t-x)^2} \qquad (t \in \mathbb{R}). \tag{16}$$

Then, for positive $\sigma < 1/\sqrt{2K_x}$, there holds the estimate

$$|(F_{n,\sigma}f)(x)| \leq \sqrt{\frac{2}{\pi}} \frac{M_x \sigma/\delta}{1-2K_x\sigma^2} \exp\left(-\frac{1-2K_x\sigma^2}{2}\left(\frac{\delta}{\sigma}\right)^2\right).$$

Consequently, under the general assumption (11) there exists a positive constant A such that the sequence $((F_{n,\sigma_n}f)(x))$ can be estimated by

$$(F_{n,\sigma_n}f)(x) = o\left(\exp\left(-A\frac{\delta_n^2}{\sigma_n^2}\right)\right)$$

as $n \to \infty$.

Remark 2. Note that a function $f : \mathbb{R} \to \mathbb{R}$ satisfies condition (2) if and only if condition (16) is valid. The elementary inequality $(t-x)^2 \leq 2\left(t^2 + x^2\right)$ implies that

$$M_x e^{K_x(t-x)^2} \leq M e^{Kt^2} \qquad (t, x \in \mathbb{R})$$

with constants $M = M_x e^{2Kx^2}$ and $K = 2K_x$.

Gawronski and Stadtmüller [8, Eq. (0.7)] also considered a truncated version of F_{n,σ_n}, namely

$$\left(F^*_{n,\sigma_n,\delta_n}f\right)(x) = \frac{1}{\sqrt{2\pi}n\sigma_n} \sum_{\substack{\nu \\ |\nu/n-x|\leq\delta_n}} f\left(\frac{\nu}{n}\right) \exp\left(-\frac{1}{2\sigma_n^2}\left(\frac{\nu}{n}-x\right)^2\right) \tag{17}$$

for certain values $\delta_n > 0$. Note that in [8, Eq. (0.7)] $c_n = n\delta_n$. A direct consequence of Proposition (1) is the following:

Corollary 2. *Fix $x \in \mathbb{R}$ and let $\delta > 0$. If f satisfies condition (16) then, for each positive $\sigma < 1/\sqrt{2K_x}$, there holds the estimate*

$$\left|\left(F_{n,\sigma}f\right)(x) - \left(F^*_{n,\sigma,\delta}f\right)(x)\right| \leq \sqrt{\frac{2}{\pi}}\frac{M_x\sigma/\delta}{1-2K_x\sigma^2}\exp\left(-\frac{1-2K_x\sigma^2}{2}\left(\frac{\delta}{\sigma}\right)^2\right).$$

Consequently, under the general assumption (11) both sequences of operators (F_{n,σ_n}) and $\left(F^*_{n,\sigma_n,\delta_n}\right)$ are asymptotically equivalent if $\sigma_n = o(\delta_n)$ as $n \to \infty$. This means that under these conditions there is a positive constant A (depending on f and x) such that

$$\left(F_{n,\sigma_n}f\right)(x) - \left(F^*_{n,\sigma_n,\delta_n}f\right)(x) = o\left(\frac{\sigma_n}{\delta_n}\exp\left(-A\frac{\delta_n^2}{\sigma_n^2}\right)\right)$$

as $n \to \infty$. For the derivatives of all orders a similar estimate is valid.

A direct consequence is the fact that, under the condition

$$\log\frac{\delta_n}{\sigma_n} + A\frac{\delta_n^2}{\sigma_n^2} - 2q\left|\log\sigma_n\right| \to +\infty \qquad (n \to \infty),$$

the asymptotic expansions (12) as well as (14) are valid also for the sequences $\left(F^*_{n,\sigma_n,\delta_n}f\right)_{n\in\mathbb{N}}$ and $\left(\left(F^*_{n,\sigma_n,\delta_n}f\right)^{(\ell)}\right)_{n\in\mathbb{N}}$, respectively, of the truncated version (17) of the Favard operators.

Approximation in Weighted Spaces

Beginning with the year 1998 a couple of Polish mathematicians continued the study of approximation by Favard operators and its variants in weighted function spaces. We begin with polynomial weights (7), i.e. the spaces $D_N(\mathbb{R})$ equipped with the norm $\|f\|_N$ as defined in Eqs. (8) and (9), respectively.

In 1998 Pych-Taberska [17] proved the inequality

$$\omega_2\left(F_{n,\sigma}f, t\right)_N \leq M \cdot \left(\left(1+t_0^2\right)\omega_2\left(f,t\right)_N + t^2\left\|f\right\|_N\right), \qquad 0 < t < t_0,$$

where the second weighted modulus of smoothness of f is given by

$$\omega_2\left(f,t\right)_N := \sup_{0<h\leq t}\left\|\Delta_h^2 f\right\|_N.$$

In 1999 Pych-Taberska and Nowak [18] defined the Favard–Kantorovich operators (5) and estimated their rate of pointwise convergence at the Lebesgue or

Lebesgue–Denjoy points of f when applied to functions f which are locally integrable in the sense of Lebesgue or Denjoy–Perron. In 2001, the same authors defined the Favard–Durrmeyer operators (6).

In 2003 Nowak gave direct results for the Kantorovich and Durrmeyer variants of the generalized Favard operators in terms of the Ditzian–Totik modulus of smoothness.

In 2004 Rempulska and Tomczak [19] enhanced the rate of convergence by applying Favard operators to Taylor polynomials of smooth functions.

In 2006 Nowak obtained inverse approximation theorems for Kantorovich–Durrmeyer and Favard–Durrmeyer operators applied to functions f such that $w_N f \in L^p(\mathbb{R})$, where $1 \leq p \leq \infty$.

Now we consider the exponential weights (10). In 2007 and 2010 Nowak and Sikorska–Nowak considered the Kantorovich–Favard and the Durrmeyer–Favard operators for functions f such that $w_\beta f \in L^p(\mathbb{R})$, where $1 \leq p \leq \infty$ and w_β $(\beta > 0)$ is as defined in (10). In [15] both authors proved an inverse approximation theorem. In [16] they presented some direct approximation theorems.

Quasi-Interpolants

The results of the preceding section show that the optimal degree of approximation cannot be improved in general by higher smoothness properties of the function f. In order to obtain much faster convergence quasi-interpolants were considered. Let us briefly recall the definition of the quasi-interpolants in the sense of Sablonnière [20]. For another method to construct quasi-interpolants see [10, 11].

Denote Π_j the space of algebraic polynomials of order at most j $(j = 0, 1, 2, \ldots)$. If the operators $\mathscr{B}_n$ let invariant Π_j, for each j (the most approximation operators possess this property), i.e,

$$\mathscr{B}_n\left(\Pi_j\right) \subseteq \Pi_j \qquad (0 \leq j \leq n),$$

$\mathscr{B}_n : \Pi_n \to \Pi_n$ is an isomorphism which can be represented by linear differential operators

$$\mathscr{B}_n = \sum_{k=0}^{n} \beta_{n,k} D^k$$

with polynomial coefficients $\beta_{n,k}$ and $Df = f'$, $D^0 = \mathrm{id}$.

The inverse operator $\mathscr{B}_n^{-1} \equiv \mathscr{A}_n : \Pi_n \to \Pi_n$ satisfies

$$\mathscr{A}_n = \sum_{k=0}^{n} \alpha_{n,k} D^k$$

with polynomial coefficients $\alpha_{n,k}$. Sablonnière defined new families of intermediate operators obtained by composition of $\mathscr{B}_n$ and its truncated inverses

$$\mathscr{A}_n^{(r)} = \sum_{k=0}^{r} \alpha_{n,k} D^k.$$

In this way he obtained a family of LQI defined by

$$\mathscr{B}_n^{(r)} = \mathscr{A}_n^{(r)} \circ \mathscr{B}_n, \quad 0 \le r \le n,$$

and a family of right quasi-interpolants (RQI) defined by

$$\mathscr{B}_n^{[r]} = \mathscr{B}_n \circ \mathscr{A}_n^{(r)}, \quad 0 \le r \le n.$$

Obviously, there holds $\mathscr{B}_n^{(0)} = \mathscr{B}_n^{[0]} = \mathscr{B}_n$, and $\mathscr{B}_n^{(n)} = \mathscr{B}_n^{[n]} = I$ when acting on Π_n. In the following we consider only the family of LQI. The definition reveals that $\mathscr{B}_n^{(r)} f$ is a linear combination of derivatives of $\mathscr{B}_n f$. Furthermore, $\mathscr{B}_n^{(r)}$ $(0 \le r \le n)$ has the nice property to preserve polynomials of degree up to r, because, for $p \in \Pi_r$, we have

$$\mathscr{B}_n^{(r)} p = \left(\mathscr{A}_n^{(r)} \circ \mathscr{B}_n\right) p = \sum_{k=0}^{r} \alpha_{n,k} D^k \underbrace{(\mathscr{B}_n p)}_{\in \Pi_r} = \sum_{k=0}^{n} \alpha_{n,k} D^k (\mathscr{B}_n p)$$
$$= \left(\mathscr{A}_n^{-1} \circ \mathscr{B}_n\right) p = p.$$

For many concrete approximation operators we have $\mathscr{B}_n^{(r)} f - f = O\left(n^{-\lfloor r/2+1 \rfloor}\right)$ as $n \to \infty$.

Unfortunately, the Favard operator as well as its Durrmeyer variant doesn't let invariant the spaces Π_j, for $0 \le j \le n$. However, under appropriate assumptions on the sequence (σ_n) they do it asymptotically up to a remainder which decays exponentially fast as n tends to infinity. Writing $\simeq$ for this "asymptotic equality" we obtain, for fixed $n \in \mathbb{N}$,

$$F_{n,\sigma_n} p_k \simeq e_k$$
$$\text{with } p_k = k! \sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^j \frac{\sigma_n^{2j}}{2^j j!\,(k-2j)!} e_{k-2j},$$

where e_m denote the monomials $e_m(t) = t^m$ $(m = 0, 1, 2, \ldots)$. Hence, for the inverse,

$$(F_{n,\sigma_n})^{-1} e_k \simeq p_k = \sum_{j=0}^{\lfloor k/2 \rfloor} \underbrace{(-1)^j \frac{\sigma_n^{2j}}{2^j j!}}_{=\alpha_{n,2j}} D^{2j} e_k$$

Note that $\beta_{n,2k+1} = \alpha_{n,2k+1} = 0$ $(k = 0, 1, 2, \ldots)$ and that neither $\beta_{n,k}$ nor $\alpha_{n,k}$ depend on the variable x. The analogous results for the Favard–Durrmeyer operators are similar. Proceeding in this way we define the following operators:

Definition 1 (Favard Quasi-Interpolants). The LQI $F_{n,\sigma_n}^{(r)}$ and $\tilde{F}_{n,\sigma_n}^{(r)}$ $(r = 0, 1, 2, \ldots)$ of the Favard and Favard–Durrmeyer operators, respectively, are given by

$$F_{n,\sigma_n}^{(r)} = \sum_{k=0}^{r} \alpha_{n,k} D^k F_{n,\sigma_n} := \sum_{k=0}^{\lfloor r/2 \rfloor} (-1)^k \frac{\sigma_n^{2k}}{(2k)!!} D^{2k} F_{n,\sigma_n}$$

and

$$\tilde{F}_{n,\sigma_n}^{(r)} = \sum_{k=0}^{r} \tilde{\alpha}_{n,k} D_{n,\sigma_n}^k \tilde{F}_{n,\sigma_n} := \sum_{k=0}^{\lfloor r/2 \rfloor} (-1)^k \frac{\sigma_n^{2k}}{k!} D^{2k} \tilde{F}_{n,\sigma_n}.$$

Remark 3. Note that $F_{n,\sigma_n}^{(2r)} = F_{n,\sigma_n}^{(2r+1)}$ and $\tilde{F}_{n,\sigma_n}^{(2r)} = \tilde{F}_{n,\sigma_n}^{(2r+1)}$ $(r = 0, 1, 2, \ldots)$.

The local rate of convergence is given by the next theorem [1].

Theorem 10. *Let $\ell \in \mathbb{N}_0$, $q \in \mathbb{N}$ and $x \in \mathbb{R}$. Suppose that the real sequence (σ_n) satisfies condition (11). For each function $f \in W[2(\ell + q); x]$, the following complete asymptotic expansions are valid as $n \to \infty$:*

$$\left(F_{n,\sigma_n}^{(2r)} f\right)^{(\ell)}(x) \sim f^{(\ell)}(x) + (-1)^r \sum_{k=r+1}^{\infty} \binom{k-1}{r} \frac{f^{(2k+\ell)}(x)}{(2k)!!} \sigma_n^{2k}$$

and

$$\left(\tilde{F}_{n,\sigma_n}^{(2r)} f\right)^{(\ell)}(x) = f^{(\ell)}(x) + (-1)^r \sum_{k=1}^{q} \binom{k-1}{r} \frac{f^{(2k+\ell)}(x)}{k!} \sigma_n^{2k} + o\left(\sigma_n^{2q}\right).$$

Remark 4. An immediate consequence are the asymptotic relations

$$\left(F_{n,\sigma_n}^{(2r)} f\right)(x) - f(x) = O\left(\sigma_n^{2(r+1)}\right)$$

and

$$\left(\tilde{F}_{n,\sigma_n}^{(2r)} f\right)(x) - f(x) = O\left(\sigma_n^{2(r+1)}\right)$$

as $n \to \infty$.

References

1. Abel, U.: Asymptotic expansions for Favard operators and their left quasi-interpolants. Stud. Univ. Babeş-Bolyai Math. **56**, 199–206 (2011)
2. Abel, U., Butzer, P.L.: Complete asymptotic expansion for generalized favard operators. Constr. Approx. **35**, 73–88 (2012)
3. Abel, U., Butzer, P.L.: Old and new results on the Favard operator. In: Nikolov, G., Uluchev, R. (eds.) Constructive Theory of Functions, Sozopol 2010: In memory of Borislav Bojanov, pp. 1–8. Prof. Marin Drinov Academic Publishing House, Sofia (2012)
4. Becker, M.: Inverse theorems for Favard operators in polynomial weight spaces. Ann. Soc. Math. Pol. Ser. I Commentat. Math. **22**, 165–173 (1981)
5. Becker, M., Butzer, P.L., Nessel, R.J.: Saturation for Favard operators in weighted function spaces. Stud. Math. **59**, 139–153 (1976)
6. Butzer, P.L., Nessel, R.J.: Fourier Analysis and Approximation. One-Dimensional Theory, vol. 1. Birkhäuser, Basel/Stuttgart (1971)
7. Favard, J.: Sur les multiplicateurs d'interpolation. J. Math. Pures Appl. IX. **23**, 219–247 (1944)
8. Gawronski, W., Stadtmüller, U.: Approximation of continuous functions by generalized Favard operators. J. Approx. Theory **34**), 384–396 (1982)
9. Gupta, V., Agarwal, R.P.: Convergence Estimates in Approximation Theory. Springer, Cham (2014)
10. Kageyama, Y.: Generalization of the left Bernstein quasi-interpolants. J. Approx. Theory **94**, 306–329 (1998)
11. Kageyama, Y.: A new class of modified Bernstein operators. J. Approx. Theory **101**, 121–147 (1999)
12. Nowak, G.: Direct results for generalized Favard-Kantorovich and Favard-Durrmeyer operators in weighted function spaces. Demonstr. Math. **36**, 879–891 (2003)
13. Nowak, G.: Inverse results for generalized Favard-Kantorovich and Favard-Durrmeyer operators in weighted function spaces. Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) **9**, 183–195 (2006)
14. Nowak, G., Pych-Taberska, P.: Some properties of the generalized Favard-Durrmeyer operators. Funct. Approx. Comment. Math. **29**, 103–112 (2001)
15. Nowak, G., Sikorska-Nowak, A.: On the generalized Favard-Kantorovich and Favard-Durrmeyer operators in exponential function spaces. J. Inequal. Appl. **2007** (Article ID 75142), 12 (2007)
16. Nowak, G., Sikorska-Nowak, A.: Theorems for generalized Favard-Kantorovich and Favard-Durrmeyer operators in exponential function spaces. Math. Slovaca **60**, 265–278 (2010)
17. Pych-Taberska, P.: On the generalized Favard operators. Funct. Approx. Comment. Math. **26**, 265–273 (1998)
18. Pych-Taberska, P., Nowak, G.: Approximation properties of the generalized Favard-Kantorovich operators. Ann. Soc. Math. Pol. Ser. I Commentat. Math. **39**, 139–152 (1999)
19. Rempulska, L., Tomczak, K.: Approximation by modified Favard operators. Ann. Soc. Math. Pol. Ser. I Commentat. Math. **44**, 205–215 (2004)
20. Sablonnière, P.: Representation of quasi-interpolants as differential operators and applications. In: Müller, M.W., Buhmann, M., Mache, D.H. (eds.) New Developments in Approximation Theory. International Series of Numerical Mathematics, vol. 132, pp. 233–253. Birkhäuser, Basel (1998)
21. Voronovskaja, E.: Détermination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein. C. R. Acad. Sci. U.R.S.S. **1932**, 79–85 (1932)

Bivariate Extension of Linear Positive Operators

P.N. Agrawal and Meenu Goyal

Abstract The goal of this chapter is to present a survey of the literature on approximation of functions of two variables by linear positive operators. We study the approximation properties of these operators in the space of functions of two variables, continuous on a compact set. We also discuss the convergence of the operators in a weighted space of functions of two variables and find the rate of this convergence by means of modulus of continuity.

Keywords Rate of convergence • Simultaneous approximation • Bivariate modulus of continuity • Approximation by positive operators • Ssasz-Mirakjan operators • Baskakov operators • Asymptotic formula • Weighted approximation

Subject Classifications: 41A25, 41A28, 41A36

Bivariate Bernstein Operators

The well-known Bernstein polynomial of *mth* degree, corresponding to a bounded function $f(x)$ and the interval [0, 1] has the following expression:

$$B_m(f;x) = \sum_{i=0}^{m} \lambda_{m,i}(x) f\left(\frac{i}{m}\right),$$

where $\lambda_{m,i}(x) = \binom{m}{i} x^i (1-x)^{m-i}, x \in [0,1]$.

P.N. Agrawal • M. Goyal (✉)
Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee-247667, India
e-mail: pna_iitr@yahoo.co.in; meenu.goyaliitr@yahoo.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_2

The fundamental property of the polynomials B_m is that $B_m(f; x) \to f(x)$ as $m \to \infty$, uniformly on the interval [0, 1], if $f(x)$ is continuous on [0, 1]. We use the following approximation formula

$$f(x) \approx B_m(f; x). \tag{1}$$

Kingsley [30] introduced the Bernstein polynomials for functions of two variables of class $C^{(k)}(R)$, where $C^{(k)}(R)$ is the space of all continuous functions ϕ and having the partial derivatives $\frac{\partial^m}{\partial x^s \partial y^{m-s}} \in C(R)$, $s = 1, 2, ...m; m = 1, 2, ..., k$ and $R : 0 \leq x \leq 1, 0 \leq y \leq 1$. Let $\phi(x, y)$ be a continuous function in a closed region R. The Bernstein polynomials $B_{m,n}(x, y)$ associated with the function $\phi(x, y)$ are given by

$$B_{m,n}(x, y) = \sum_{p=0}^{n} \sum_{q=0}^{m} \phi\left(\frac{p}{n}, \frac{q}{m}\right) \lambda_{n,p}(x) \lambda_{m,q}(y). \tag{2}$$

Taking into account that

$$\sum_{p=0}^{n} \lambda_{n,p}(x) = 1 \qquad \text{and} \qquad \sum_{p=0}^{n} (nx - p)^2 \lambda_{n,p}(x) = nx(1 - x),$$

if we define for $k \geq 0, i = 0, 1, 2 \ldots k$,

$$A_{p,q}^{i,k-i} = \sum_{\alpha=0}^{i} \sum_{\beta=0}^{k-i} (-1)^{\alpha+\beta} \binom{i}{\alpha} \binom{k-i}{\beta} \phi\left(\frac{p + (i - \alpha)}{n}, \frac{q + (k - i - \alpha)}{m}\right),$$

then by mathematical induction, the following two lemmas hold:

Lemma 1 ([30]). *If* $0 \leq i \leq k$, $i \leq n$, $k \leq m$ *and* $x, y \in R$, *then the kth partial derivatives of Bernstein polynomials (2) are given by*

$$B_{m,n}^{(i,k-i)} = \frac{n!m!}{(n - i)!(m - k + i)!} \sum_{p=0}^{n-i} \sum_{q=0}^{m-k+i} A_{p,q}^{i,k-i} \lambda_{n-i,p} \lambda_{m-k+i,q}.$$

Lemma 2 ([30]). *If* $0 \leq i \leq k, 0 \leq p \leq n - i, 0 \leq q \leq m - k + i$, *and if* $\phi(x, y)$ *is of class* $C^{(k)}(R)$ *for x and y in R, then there exist two real numbers* $\xi = \xi(p), \gamma = \gamma(q)$ *such that* $0 < \xi < 1, 0 < \gamma < 1$ *and such that*

$$A_{p,q}^{i,k-i} = \frac{1}{n^i m^{k-i}} \phi^{(i,k-i)} \left(\frac{p + \xi i}{n}, \frac{q + \gamma(k - i)}{m}\right).$$

Lemma 3 ([30]). *For fixed x and y in R and for fixed positive integers M and N, let d be arbitrary positive number and let* $a(p, q)$ *be a quantity dependent upon p and q such that*

$$|a(p,q)| \le \pi_1 \qquad \text{for} \qquad \left|x - \frac{p}{N}\right| \le d \qquad \text{and} \qquad \left|y - \frac{q}{M}\right| \le d,$$
$$|a(p,q)| \le \pi_2 \qquad \text{for} \qquad \left|x - \frac{p}{N}\right| > d \qquad \text{and} \qquad \left|y - \frac{q}{M}\right| > d.$$

Furthermore, assume that it is possible to split off from $a(p,q)$ terms $a'(p)$ independent of q or terms $b'(q)$ independent of p, that is

$$a(p,q) = a''(p,q) + a'(p) = b''(p,q) + b'(q)$$

such that

$$|a''(p,q)| \le \pi_3 \qquad \textit{for} \qquad \left|x - \frac{p}{N}\right| \le d \qquad \textit{and} \qquad \left|y - \frac{q}{M}\right| > d,$$
$$|a'(p,q)| \le \pi_4 \qquad \textit{for} \qquad \left|x - \frac{p}{N}\right| \le d$$

$$|b''(p,q)| \le \pi_5 \qquad \textit{for} \qquad \left|x - \frac{p}{N}\right| > d \qquad \textit{and} \qquad \left|y - \frac{q}{M}\right| \le d,$$
$$|b'(q)| \le \pi_6 \qquad \textit{for} \qquad \left|y - \frac{q}{M}\right| \le d,$$

then,

$$\left|\sum_{p=0}^{N}\sum_{q=0}^{M} a(p,q)\lambda_{N,p}\lambda_{M,q}\right| \le \pi_1 + \pi_4 + \pi_6 + \frac{\pi_2(M+N)}{8MNd^2} + \frac{\pi_3}{4Md^2} + \frac{\pi_5}{4Nd^2}.$$

Theorem 1 ([30]). *If $\phi(x,y)$ is of class $C^{(k)}(R)$, then $\lim_{m,n\to\infty} B_{m,n}^{(s,k-s)} = \phi^{(s,k-s)}$ and the convergence is uniform in R.*

Butzer [14] also proved the above theorem in a direct manner. For this, he needed the following lemmas:

Lemma 4. *If $f(x,y)$ is bounded in the square R, then at every point of continuity (x,y) of f*

$$\lim_{n_1,n_2\to\infty} B^f_{n_1,n_2}(x,y) = f(x,y),$$

the result holding uniformly in x and y if $f(x,y)$ is continuous in R.

Lemma 5. *If all the partial derivatives of $f(x,y)$ of order $\le p$ exist and are continuous in R, then*

$$\frac{\partial^p}{\partial x^q \partial y^{p-q}} B^f_{n_1,n_2}(x,y) \to \frac{\partial^p}{\partial x^q \partial y^{p-q}} f(x,y),$$

uniformly in R as n_1, n_2 approach infinity in any manner whatever.

Lemma 6.

$$\frac{d^l}{du^l}(u^\nu(1-u)^{n-\nu}) = Q(u)u^{\nu-l}(1-u)^{n-\nu-l},$$

where $Q(u) = \sum_{i,j} n^i(\nu - nu)^j h^l_{i,j}(u),\ i,j \geq 0,\ 2i + j \leq l$ *and the* $h^l_{i,j}(u)$ *are polynomials in* u *independent of* ν *and* n.

Pop [38] obtained the rate of convergence in terms of the modulus of continuity and established the Voronovskaja type asymptotic theorem for the bivariate case of Bernstein polynomials:

Theorem 2. *Let* $f : [0,1] \times [0,1] \to \mathbb{R}$ *be a bivariate function.*

(*i*) *If* $(x,y) \in [0,1] \times [0,1]$ *and* f *admits partial derivatives of second order continuous in a neighborhood of the point* (x,y)*, then*

$$\lim_{m\to\infty} m(B_{m,m}(f;x,y) - f(x,y)) = \frac{x(1-x)}{2} f''_{xx}(x,y) + \frac{y(1-y)}{2} f''_{yy}(x,y). \quad (3)$$

If f *admits partial derivatives of second order continuous on* $[0,1] \times [0,1]$*, then the convergence given in (3) is uniform on* $[0,1] \times [0,1]$.

(*ii*) *If* f *is continuous on* $[0,1] \times [0,1]$*, then*

$$|B_{m,m}(f;x,y) - f(x,y)| \leq \frac{25}{16}\omega\left(f; \frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}}\right)$$

for any $(x,y) \in [0,1] \times [0,1]$*, any* $m \in \mathbb{N}$.

Stancu [42] defined another bivariate Bernstein polynomials on the triangle

$$\Delta := S = \{(x,y) : x + y \leq 1,\ 0 \leq x, y \leq 1\}.$$

For the functions $f : S \to \mathbb{R}$, he considered $G_n(f;x,y)$ as follows:

$$G_n(f;x,y) = \sum_{i=0}^{m}\sum_{j=0}^{m-i} \binom{m}{i}\binom{m-i}{j} x^i y^j (1-x-y)^{m-i-j} f\left(\frac{i}{m}, \frac{j}{m}\right),\ (x,y) \in S. \quad (4)$$

Let $\phi_1 = \phi_1(x)$ and $\phi_2 = \phi_2(x)$ be two polynomials which on the interval $[0,1]$ possess the following properties: $\phi_2 \geq \phi_1 \geq 0$. Assuming the function $f(x)$ is defined on the domain D, determined by the equations $y = \phi_1, y = \phi_2, x = 0$, $x = 1$, let us make the following change of the variable

$$y = (\phi_2 - \phi_1)t + \phi_1. \quad (5)$$

Using the formula (1), we get

$$f(x,y) = f(x, (\phi_2 - \phi_1)t + \phi_1) \approx \sum_{i=0}^{m} \lambda_{m,i}(x)\phi\left(\frac{i}{m}, t\right),$$

where $\phi\left(\frac{i}{m}, t\right) = f\left(\frac{i}{m}, \left(\phi_2\left(\frac{i}{m}\right) - \phi_1\left(\frac{i}{m}\right)\right)t + \phi_1\left(\frac{i}{m}\right)\right)$. Let us attach to the node $\frac{i}{m}$, the natural number n_i. By the formula (1), applied to the variable t we have

$$\phi\left(\frac{i}{m}, t\right) \approx \sum_{j=0}^{n_i} \lambda_{n_i,j}(t)\phi\left(\frac{i}{m}, \frac{j}{n_i}\right).$$

In this way, we obtain the approximation formula $f(x,y) \approx B(f;x,y)$, where

$$B(f;x,y) = \sum_{i=0}^{m}\sum_{j=0}^{n_i} \lambda_{m,i}(x)\lambda_{n_i,j}(t)\phi\left(\frac{i}{m}, \frac{j}{n_i}\right),$$

where t is given by (5). More explicitly,

$$\begin{aligned} B(f;x,y) &= \sum_{i=0}^{m}\sum_{j=0}^{n_i} \lambda_{m,i}(x)\lambda_{n_i,j}\left(\frac{y-\phi_1}{\phi_2-\phi_1}\right) \\ &\times f\left(\frac{i}{m}, \left(\phi_2\left(\frac{i}{m}\right) - \phi_1\left(\frac{i}{m}\right)\right)\frac{j}{n_i} + \phi_1\left(\frac{i}{m}\right)\right). \end{aligned}$$

For this function to represent a polynomial, we must conveniently particularize the polynomials ϕ_1 and ϕ_2. The following cases are remarkable:

(a) $\phi_1 = 0, \phi_2 = 1, n_i = n (i = \overline{0,m})$. In this case the domain D becomes the square R and we get the polynomial (2) of degree (m,n).
(b) $\phi_1 = 0, \phi_2 = 1-x, n_i = m-i (i = \overline{0,m})$. Now we obtain the polynomial (4) corresponding to the triangle Δ.
(c) $\phi_1 = 0, \phi_2 = x, n_i = m-i (i = \overline{0,m})$. The domain D transforms into the triangle $\overline{\Delta}: x \geq 0, y \geq 0, y - x \geq 0$ and we have

$$\overline{B}_m(f;x,y) = \sum_{i=0}^{m}\sum_{j=0}^{i} \overline{\lambda}_m^{i,j}(x,y) f\left(\frac{i}{m}, \frac{j}{m}\right), \tag{6}$$

where

$$\overline{\lambda}_m^{i,j}(x,y) = \binom{m}{i}\binom{i}{j}(1-x)^{m-i}y^j(x-y)^{i-j}.$$

(d) $\phi_1 = x, \phi_2 = 1, n_i = m - i (i = (0, m))$. The corresponding Bernstein polynomial is

$$\overline{\overline{B}}_m(f; x, y) = \sum_{i=0}^{m} \sum_{j=0}^{m-i} \overline{\overline{\lambda}}_m^{i,j}(x, y) f\left(\frac{i}{m}, \frac{i+j}{m}\right), \tag{7}$$

where

$$\overline{\overline{\lambda}}_m^{i,j}(x, y) = \binom{m}{i}\binom{m-i}{j} x^i (y-x)^j (1-x)^{m-i-j},$$

the domain D being reduced to the triangle $\overline{\overline{\Delta}}$ determined by the equations $y = x, y = 1, x = 0$.

(e) $\phi_1 = 1 - x, \phi_2 = 1, n_i = i (i = (0, m))$. In this case we have the triangle $\overline{\overline{\overline{\Delta}}}$ determined by $y = 1, y = 1 - x, x = 1$ and the Bernstein polynomial

$$\overline{\overline{\overline{B}}}_m(f; x, y) = \sum_{i=0}^{m} \sum_{j=0}^{i} \overline{\overline{\overline{\lambda}}}_m^{i,j}(x, y) f\left(\frac{i}{m}, 1 - \frac{i}{m} + \frac{j}{m}\right), \tag{8}$$

where

$$\overline{\overline{\overline{\lambda}}}_m^{i,j}(x, y) = \binom{m}{i}\binom{i}{j}(1-x)^{m-i}(x+y-1)^j(1-y)^{i-j}.$$

Remark 1. The polynomials (4), (6), (7), and (8) have the global degree equal to m. We can pass from one of these to another by a simple transformation. Thus, for instance, applying the transformation : $u = x, v = x + y$ to (4) gives us (7).

The complete modulus of continuity of a function $f(x, y)$ is defined by

$$\omega(\delta_1, \delta_2) = \sup |f(x'', y'') - f(x', y')|,$$

where $\delta_1 > 0, \delta_2 > 0$ are real numbers, whereas (x', y') and (x'', y'') are points of Δ such that $|x'' - x'| \leq \delta_1, |y'' - y'| \leq \delta_2$.
Alternately, the complete modulus of continuity of f which we denote by $\omega(f; \delta)$ is defined as

$$\omega(f; \delta) = \sup_{\sqrt{(x_1-x_2)^2+(y_1-y_2)^2} \leq \delta} |f(x_1, y_1) - f(x_2, y_2)|.$$

Taking into account that on Δ we have

$$\lambda_m(x,y) \geq 0, \quad \sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y) = 1,$$

$$|f(x'',y'') - f(x',y')| \leq \omega(|x''-x'|,|y''-y'|) \leq \omega(\delta_1,\delta_2)$$

and the inequality (see, e.g., [3])

$$\omega(\lambda_1\delta_1,\lambda_2\delta_2) \leq (\lambda_1+\lambda_2+1)\omega(\delta_1,\delta_2), \quad \lambda_1>0, \lambda_2>0,$$

Now, we give here the proof of convergence of the polynomials defined by (4) for a continuous function in the triangle Δ.

Theorem 3. *Let f be continuous on the Δ, then we have*

$$|f(x,y) - B_m(f;x,y)| \leq 2\omega\left(\frac{1}{\sqrt{m}},\frac{1}{\sqrt{m}}\right).$$

Proof. We can write successively

$$\begin{aligned}
|f(x,y) - B_m(f:x,y)| &\leq \sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y)\left|f(x,y) - f\left(\frac{i}{m},\frac{j}{m}\right)\right| \\
&\leq \sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y)\omega\left(\left|x-\frac{i}{m}\right|,\left|y-\frac{j}{m}\right|\right) \\
&\leq \sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y)\left(1+\frac{1}{\delta_1}\left|x-\frac{i}{m}\right|+\frac{1}{\delta_2}\left|y-\frac{j}{m}\right|\right)\omega(\delta_1,\delta_2) \\
&\leq \left(1+\frac{1}{2\sqrt{m}}\left(\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)\right)\omega(\delta_1,\delta_2),
\end{aligned}$$

since

$$\begin{aligned}
\sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y)\left|x-\frac{i}{m}\right| &\leq \left(\sum_{i=0}^{m}\sum_{j=0}^{m-i} \lambda_m^{i,j}(x,y)\left(x-\frac{i}{m}\right)^2\right)^{1/2} \\
&= \left(\sum_{i=0}^{m}\binom{m}{i}x^i(1-x)^{m-i}\left(x-\frac{i}{m}\right)^2\right)^{1/2} \\
&= \left(\frac{x(1-x)}{m}\right)^{1/2} \leq \frac{1}{2\sqrt{m}}, \text{ etc.}
\end{aligned}$$

By choosing $\delta_1, \delta_2 = \frac{1}{\sqrt{m}}$, we get the inequality

$$|f(x,y) - B_m(f;x,y)| \leq 2\omega\left(\frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}}\right).$$

Hence, uniformly on $\Delta : B_m(f;x,y) \to f(x,y)$, as $m \to \infty$.

For the polynomials (2), we have the following inequality of Popoviciu proved by Ipatov [26]

$$|f(x,y) - B_{m,n}(f;x,y)| \leq \frac{3}{2}\omega\left(\frac{1}{\sqrt{m}}, \frac{1}{\sqrt{n}}\right).$$

In 1989, Martinez [33] studied some approximation properties of two-dimensional Bernstein polynomials. He showed the convergence of the polynomials with integral coefficients $B_{m,n}^{(i,k-i),e}f$ to $f_{i,k-i}^{(k)}$ both in the uniform and L_p norms. Here the superscript "e" will denote a polynomial with integral coefficients in the above sense and $\|.\|_p$ will denote the $L_p[R]$ norm $(1 \leq p \leq \infty)$.
The partial moduli of continuity of f are defined as

$$\omega^{(1)}(f;\delta) = \omega(f;\delta,0) = \sup_{y} \sup_{|x_1-x_2|\leq\delta} |f(x_1,y) - f(x_2,y)|$$

$$\omega^{(2)}(f;\delta) = \omega(f;0,\delta) = \sup_{x} \sup_{|y_1-y_2|\leq\delta} |f(x,y_1) - f(x,y_2)|$$

and $\omega(f;\delta,\epsilon)$ is the complete modulus of continuity of f.

The complete modulus of continuity $\omega(f;\delta,\epsilon)$ of the function f is connected with its partial moduli of continuity $\omega(f;\delta,0)$ and $\omega(f;0,\epsilon)$ by the inequalities

$$\omega(f;\delta,\epsilon) \leq \omega(f;\delta,0) + \omega(f;0,\epsilon) \leq 2\omega(f;\delta,\epsilon).$$

First, we give some basic lemmas which are needed to prove the main theorem:

Lemma 7. *If f is continuous in R and $B_{m,n}(f;x)$ is the Bernstein polynomial of f, then*

$$|B_{m,n}(f;x,y) - f(x,y)| \leq \frac{3}{2}\left(\omega^{(1)}(f;n^{-1/2}) + \omega^{(2)}(f;m^{-1/2})\right) \leq 3\omega(f;n^{-1/2},m^{-1/2}),$$

where $\omega^{(i)}(f;\delta), i = 1,2$, are the partial moduli of continuity of f.

We now define the polynomial

$$\overline{B}_{m,n}^{(i,k-i)}f = \frac{(m-i)!}{m!}\frac{(n-k+i)!}{n!}m^i n^{k-i} B_{m,n}^{(i,k-i)}f$$

and let $\overline{B}_{m,n}^{(i,k-i),e}f$ be its corresponding polynomial with integral coefficients.

Lemma 8. *Let f be a function such that its partial derivatives of the first k orders exist and are continuous in R.*

(a) If $1 \le p < \infty$, *then*

$$\left\| f^{(k)}_{i,k-i} - \overline{B}^{(i,k-i),e}_{m,n} f \right\|_p \le A + B + O((mn)^{-1/p}),$$

where $A = 3\,\omega(f^{(k)}_{i,k-i}; (m-i)^{-1/2}, (n-k+i)^{-1/2})$, $B = \omega(f^{(k)}_{i,k-i}; i/m, (k-i)/n)$;

(b) If $p = \infty$ *and the numbers* $m^i n^{(k-i)} \Delta^i_{x,m^{-1}} \Delta^{k-i}_{y,n^{-1}} f(u_1, u_2)$ *are integers, where* u_1 *is either zero or* $(m-i)/m$ *and* u_2 *is either zero or* $(n-k+i)/n$, *then*

$$\left\| f^{(k)}_{i,k-i} - \overline{B}^{(i,k-i),e}_{m,n} f \right\|_\infty \le A + B + O((mn)^{-1})$$

where A and B are as in (a).

Theorem 4. *Let* $\max(q, r-q) > 1, n_1 = \min(m, n)$, *and let f be a function such that its partial derivatives of the first k orders exist and are continuous in R.*

(a) If $1 \le p < \infty$, *then*

$$\left\| f^{(k)}_{i,k-i} - \overline{B}^{(i,k-i),e}_{m,n} f \right\|_p \le A + B + O((n_1)^{-\lambda(p)}),$$

where A and B are as in Lemma 8 and

$$\lambda(p) = \begin{cases} 1 \quad , & \text{if } p = 1 \\ 2/p \quad , & \text{if } 2 \le p < \infty. \end{cases}$$

(b) If $p = \infty$ *and the numbers* $m!/(m-i)! \;\; n!/(n-k+i)! \;\; \Delta^i_{x,m^{-1}} \Delta^{k-i}_{y,n^{-1}} f(u_1, u_2)$ *are integers, where* u_1 *is either zero or* $(m-i)/m$, *and* u_2 *is either zero or* $(n-k+i)/n$, *then*

$$\left\| f^{(k)}_{i,k-i} - \overline{B}^{(i,k-i),e}_{m,n} f \right\|_\infty \le A + B + O((n_1)^{-1}),$$

where A and B are as in Lemma 8.

Bivariate Case of Bernstein–Schurer–Stancu Operators

Barbosu [11] introduced the Schurer–Stancu bivariate operators. The Bernstein–Schurer–Stancu operators $\tilde{S}^{\alpha,\beta}_{m,p} : C[0, 1+p] \to C[0, 1]$, defined for any $m \in \mathbb{N}$ and any $f \in C[0, 1+p]$ by

$$\tilde{S}_{m,p}^{\alpha,\beta}(f;x) = \sum_{k=0}^{m+p} \tilde{\lambda}_{m,k}(x) f\left(\frac{k+\alpha}{m+\beta}\right), \tag{9}$$

where $\tilde{\lambda}_{m,k}(x) = \left[{m+p \atop k}\right] x^k (1-x)^{m+p-k}$.

Note that for $\alpha = \beta = 0$, the operator (9) reduces to Schurer operators and for $p = 0$, (9) reduces to Stancu operator.

To construct the bivariate Schurer–Stancu type operators $\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2} : C([0,1+p] \times [0,1+q]) \to C([0,1] \times [0,1])$, where $p, q \geq 0$ are given integers and $0 \leq \alpha_1 \leq \beta_1, 0 \leq \alpha_2 \leq \beta_2$ are real parameters,

let

$$\tilde{S}_{m,p}^{\alpha_1,\beta_1} : C([0,1+p]) \to C([0,1]),$$

$$\tilde{S}_{n,q}^{\alpha_2,\beta_2} : C([0,1+q]) \to C([0,1])$$

be the univariate Schurer–Stancu operators defined by (9).

The parametric extensions (see [10]) of these operators are

$$x^{\tilde{S}_{m,p}^{\alpha_1,\beta_1}} \, y^{\tilde{S}_{n,q}^{\alpha_2,\beta_2}} : C([0,1+p] \times [0,1+q]) \to C([0,1] \times [0,1])$$

defined, respectively, by

$$(x^{\tilde{S}_{m,p}^{(\alpha_1,\beta_1)}} f)(x,y) : \sum_{k=0}^{m+p} \tilde{\lambda}_{m,k}(x) f((k+\alpha_1)/(m+\beta_1), y), \tag{10}$$

$$(y^{\tilde{S}_{n,q}^{(\alpha_2,\beta_2)}} f)(x,y) : \sum_{j=0}^{n+q} \tilde{\lambda}_{n,j}(y) f(x, (j+\alpha_2)/(n+\beta_2)). \tag{11}$$

Lemma 9. *The parametric extensions (10) and (11) are linear positive operators.*

Lemma 10. *The parametric extensions* $x^{\tilde{S}_{m,p}^{\alpha_1,\beta_1}} \, y^{\tilde{S}_{n,q}^{\alpha_2,\beta_2}}$ *of Schurer–Stancu operator commute on* $C([0,1+p] \times [0,1+q])$. *Their product is the bivariate Schurer–Stancu type operators*

$$\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2} : C([0,1+p] \times [0,1+q]) \to C([0,1] \times [0,1])$$

defined for any $f \in C([0,1+p] \times [0,1+q])$ *and any* $m, n \in \mathbb{N}$ *by*

$$(\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2} f)(x,y) = \sum_{k=0}^{m+p} \sum_{j=0}^{n+q} \tilde{\lambda}_{m,k}(x) \tilde{\lambda}_{n,j}(y) f\left(\frac{k+\alpha_1}{m+\beta_1}, \frac{j+\alpha_2}{n+\beta_2}\right). \tag{12}$$

Lemma 11. *The bivariate Schurer–Stancu type operator (12) is linear and positive.*

Approximations Properties of the Sequence $\{\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2} f\}$

First, we discuss the convergence of the sequence $\{\tilde{S}_{m,n,p,q}^{(\alpha_1,\beta_1,\alpha_2,\beta_2)} f\}_{m,n\in\mathbb{N}}$.
Let $e_{ij}(s,t) = s^i t^j$, $(i,j) \in \mathbb{N}\times\mathbb{N}$ denote the test functions.
We need the following Bohman–Korovkin Theorem for the approximation of bivariate functions (see [44]).

Theorem 5. *Let $(L_{m,n})_{m,n\in\mathbb{N}}$ be a sequence of linear positive operators applying the space $C([a,b]\times[c,d])$ into itself and having the properties*

(i) $L_{m,n}(1;x,y) = 1 + u_{m,n}(x,y)$;
(ii) $L_{m,n}((e_{10}-x)^2;x,y) = v_{m,n}(x,y)$;
(iii) $L_{m,n}((e_{01}-y)^2;x,y) = w_{m,n}(x,y)$;
(iv) $\lim\limits_{m,n\to\infty} u_{m,n}(x,y) = \lim\limits_{m,n\to\infty} v_{m,n}(x,y) = \lim\limits_{m,n\to\infty} w_{m,n}(x,y) = 0$, *uniformly on* $[a,b]\times[c,d]$.

Then the sequence $(L_{m,n}f)_{m,n\in\mathbb{N}}$ converges to f, uniformly on $[a,b]\times[c,d]$, for any $f \in C([a,b]\times[c,d])$.

Lemma 12. *The bivariate Schurer–Stancu type operators verify*

$$\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}(1;x,y) = 1;$$

$$\begin{aligned}\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}((e_{10}-x)^2;x,y) &= \frac{(p-\beta_1)^2}{(m+\beta_1)^2} + \frac{m+p}{(m+\beta_1)^2}x(1-x)\\ &\quad + \frac{2\alpha_1(mp-2m\beta_1-\beta_1^2)}{(m+\beta_2)^3}x + \frac{\alpha_1^2(3m+p)}{(m+\beta_1)^3};\end{aligned}$$

$$\begin{aligned}\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}((e_{01}-y)^2;x,y) &= \frac{(q-\beta_2)^2}{(n+\beta_2)^2} + \frac{n+q}{(n+\beta_2)^2}y(1-y)\\ &\quad + \frac{2\alpha_2(nq-2n\beta_2-\beta_2^2)}{(n+\beta_2)^3}y + \frac{\alpha_2^2(3n+q)}{(n+\beta_2)^3}.\end{aligned}$$

Theorem 6. *The sequence $(\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2} f)_{m,n\in\mathbb{N}}$ converges to f, uniformly on $[0,1]\times[0,1]$, for any $f \in C([0,1+p]\times[0,1+q])$.*

A quantitative estimation of approximation order of a continuous bivariate function using a sequence of linear positive bivariate operators is given in the following theorem, due to Shisha and Mond (see [44]).

Theorem 7. *Let $(L_{m,n})_{m,n\in\mathbb{N}}$ be a sequence of bivariate linear positive operators $L_{m,n} : C([a,b]\times[c,d]) \to C([a,b]\times[c,d])$, reproducing the constant functions. For any $f \in C([a,b]\times[c,d])$ and any $(x,y) \in [a,b]\times[c,d]$ and any $(\delta_1,\delta_2) \in [0,b-a]\times[0,d-c]$, the following*

$$|(L_{m,n}f)(x,y) - f(x,y)| \leq \left(1 + \delta_1^{-1}\sqrt{L_{m,n}((*-x)^2;x,y)} + \delta_2^{-1}\sqrt{L_{m,n}((*-y)^2;x,y)} + \delta^{-1}\delta_2^{-1}\sqrt{L_{m,n}((*-x)^2;x,y)L_{m,n}((*-y)^2;x,y)}\right)\omega(\delta_1,\delta_2)$$

holds.

Theorem 8. *For any $f \in C([0,1+p]\times[0,1+q])$, any $(x,y) \in [0,1]\times[0,1]$, the bivariate Schurer–Stancu type operators verify*

$$|(\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}f)(x,y) - f(x,y)| \leq 4\omega\left(\sqrt{\delta_{m,p,\alpha_1,\beta_1,x}}, \sqrt{\delta_{n,q,\alpha_2,\beta_2,y}}\right)$$

where

$$\delta_{m,p,\alpha_1,\beta_1,x} = \frac{(p-\beta_1)^2}{(m+\beta_1)^2} + \frac{m+p}{(m+\beta_1)^2}x(1-x) + \frac{2\alpha_1(mp-2m\beta_1-\beta_1^2)}{(m+\beta_2)^3}x + \frac{\alpha_1^2(3m+\beta_1)}{(m+\beta_1)^3}$$

and

$$\delta_{n,q,\alpha_2,\beta_2,y} = \frac{(q-\beta_2)^2}{(n+\beta_2)^2} + \frac{n+q}{(n+\beta_2)^2}y(1-y) + \frac{2\alpha_2(nq-2n\beta_2-\beta_2^2)}{(m+\beta_2)^2}y + \frac{\alpha_2^2(3n+\beta_2)}{(n+\beta_2)^3}.$$

The order of global approximation is expressed in:

Theorem 9. *For any $f \in C([0,1+p]\times[0,1+q])$, the operators $\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}$ satisfy*

$$|(\tilde{S}_{m,n,p,q}^{\alpha_1,\beta_1,\alpha_2,\beta_2}f)(x,y) - f(x,y)| \leq 4\omega(\sqrt{\delta_{m,p,\alpha_1,\beta_1}}, \sqrt{\delta_{n,q,\alpha_2,\beta_2}})$$

where

$$\delta_{m,p,\alpha_1,\beta_1} = \max_{x\in[0,1]} \delta_{m,p,\alpha_1,\beta_1,x}$$
$$\delta_{n,q,\alpha_2,\beta_2} = \max_{y\in[0,1]} \delta_{n,q,\alpha_2,\beta_2,y}$$

and $\delta_{m,p,\alpha_1,\beta_1,x}$, $\delta_{n,q,\alpha_2,\beta_2,y}$ are given, respectively, in Theorem 8, for all $(x,y) \in [0,1]\times[0,1]$.

Bivariate Case of Bernstein–Kantorovich Operators

Let $m \in \mathbb{N}$ be fixed. The Bernstein–Kantorovich operators [28] $K_m : L_1([0,1]) \to C[0,1]$, are defined by

$$K_m(f;x) = (m+1)\sum_{k=0}^{m}\binom{m}{k}x^k(1-x)^{m-k}\int_{\frac{k}{m+1}}^{\frac{k+1}{m+1}} f(t)dt.$$

The bivariate case of the Bernstein–Kantorovich was considered by Cabulea and Aldea [16] as follows:

$$K_{m,n}(f;x,y) = (m+1)(n+1)\sum_{k=0}^{n}\sum_{h=0}^{m}\binom{n}{k}\binom{m}{h}x^k(1-x)^{n-k}y^h(1-x)^{m-h}$$
$$\int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}}\int_{\frac{h}{m+1}}^{\frac{h+1}{m+1}} f(u,v)dudv,$$

where $f \in L_1([0,1]\times[0,1])$.

Lemma 13. *The bivariate operators of Bernstein–Kantorovich satisfy the relations:*

(i) $K_{m,n}(e_{00};x,y) = 1;$
(ii) $K_{m,n}(e_{10};x,y) = \frac{n}{n+1}x + \frac{1}{2(n+1)};$
(iii) $K_{m,n}(e_{01};x,y) = \frac{m}{m+1}y + \frac{1}{2(m+1)};$
(iv) $K_{m,n}(e_{11};x,y) = \left(\frac{n}{n+1}x + \frac{1}{2(n+1)}\right)\left(\frac{m}{m+1}y + \frac{1}{2(m+1)}\right);$
(v) $K_{m,n}(e_{20};x,y) = \frac{n(n-1)}{(n+1)^2}x^2 + \frac{2n}{(n+1)^2}x + \frac{1}{3(n+1)^2};$
(vi) $K_{m,n}(e_{02};x,y) = \frac{m(m-1)}{(m+1)^2}y^2 + \frac{2m}{(m+1)^2}y + \frac{1}{3(m+1)^2}.$

Theorem 10. *The bivariate operators of Kantorovich have the properties*

(i) $\lim_{m,n\to\infty} K_{m,n}f = f$ *uniformly on* $[0,1],\times[0,1],\ \forall f \in C([0,1]\times[0,1]);$
(ii) $\lim_{m,n\to\infty} K_{m,n}f = f,\ \forall f \in L_p([0,1]\times[0,1]), p \geq 1.$

The bivariate Stancu–Kantorovich operators [16] are defined by

$$K^{\alpha}_{m,n}(f;x,y) = (m+1)(n+1)\sum_{k=0}^{n}\sum_{h=0}^{m} w^{\alpha}_{n,k}(x)w^{\alpha}_{m,h}(y)\int_{\frac{k}{m+1}}^{\frac{k+1}{m+1}}\int_{\frac{h}{n+1}}^{\frac{h+1}{n+1}} f(u,v)dudv,$$

where

$$w^{\alpha}_{n,k}(x) = \binom{n}{k}\frac{x^{(k,-\alpha)}(1-x)^{(n-k,-\alpha)}}{1^{(n,-\alpha)}},\ x^{(k,-\alpha)} = x(x+\alpha)\dots(x+(k-1)\alpha)$$

and

$$w_{m,h}^{\alpha}(y) = \binom{m}{h} \frac{y^{(h,-\alpha)}(1-y)^{(m-h,-\alpha)}}{1^{(m,-\alpha)}}, \; y^{(h,-\alpha)} = y(y+\alpha)\ldots(y+(h-1)\alpha).$$

Lemma 14. *The bivariate operators of Kantorovich satisfy the relations:*

(i) $K_{m,n}^{\alpha}(e_{00};x,y) = 1;$
(ii) $K_{m,n}^{\alpha}(u-x;x,y) = \frac{1-2x}{2(n+1)};$
(iii) $K_{m,n}^{\alpha}(v-y;x,y) = \frac{1-2y}{2(m+1)};$
(iv) $K_{m,n}^{\alpha}((u-x)(v-y);x,y) = \frac{1-2x}{2(n+1)} \cdot \frac{1-2y}{2(m+1)};$
(v) $K_{m,n}^{\alpha}((u-x)^2;x,y) = x(1-x)\frac{n\frac{n\alpha+1}{\alpha+1}-1}{(n+1)^2} + \frac{1}{3(n+1)^2};$
(vi) $K_{m,n}^{\alpha}((v-y)^2;x,y) = y(1-y)\frac{m\frac{m\alpha+1}{\alpha+1}-1}{(m+1)^2} + \frac{1}{3(m+1)^2}.$

A result analogous to Theorem 10 can be proved similarly for the operators $K_{m,n}^{\alpha}$.

Pop [38] discussed the rate of approximation of f by $K_{m,n}(f;x,y)$ in terms of the modulus of continuity and proved a Voronovskaja type asymptotic theorem for the bivariate case.

Theorem 11. *Let $f : [0,1]\times[0,1] \to \mathbb{R}$ be a bivariate function.*

(*i*) *If $(x,y) \in [0,1]\times[0,1]$ and f admits partial derivatives of second order continuous in a neighborhood of the point (x,y), then*

$$\lim_{m\to\infty} m(K_{m,m}(f;x,y) - f(x,y)) = \frac{1-2x}{2} f'_x(x,y) + \frac{1-2y}{2} f'_y(x,y) + \frac{x(1-x)}{2} f''_{xx}(x,y) + \frac{y(1-y)}{2} f''_{yy}(x,y), \tag{13}$$

If f admits partial derivatives of second order continuous on $[0,1]\times[0,1]$, then the convergence given in (13) is uniform on $[0,1]\times[0,1]$.

(*ii*) *If f is continuous on $[0,1]\times[0,1]$, then*

$$|K_{m,m}(f;x,y) - f(x,y)| \le 4\omega\left(f; \frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}}\right),$$

for any $(x,y) \in [0,1]\times[0,1]$, any natural number $m, m \ge 3$.

Bivariate Operators of Kantorovich Type on the Triangle [39] Let the sets $\Delta_2(x,y) = \{(x,y) \in \mathbb{R}\times\mathbb{R} : x,y \ge 0, x+y \le 1\}$ and $\mathscr{F}(\Delta_2) = \{f : \Delta_2 \to \mathbb{R}\}$. For $m \in \mathbb{N}$, the bivariate operators of Kantorovich type are defined as

$$\mathscr{K}_m(f;x,y)=(m+1)^2\sum_{\substack{k,j=0\\k+j\le m}} p_{m,k,j}(x,y)\int_{\frac{k}{m+1}}^{\frac{k+1}{m+1}}\int_{\frac{j}{m+1}}^{\frac{j+1}{m+1}} f(s,t)dsdt \text{ for any } (x,y)\in\Delta_2,$$

where $p_{m,k,j}(x,y) = \frac{m!}{k!j!(m-k-j)!}x^k y^j (1-x-y)^{m-k-j}$, for any $k,j \ge 0, k+j \le m$.
Let the functions $e_{ij} : \Delta_2 \to \mathbb{R}, e_{ij}(x,y) = x^i y^j$ for any $(x,y) \in \Delta_2$, where $i,j \in \mathbb{N}_0$ (the set of nonnegative integers).

Lemma 15. *The operators $(\mathscr{K}_m)_{m\ge 1}$ verify for any $(x,y) \in \Delta_2$, the following equalities*

(1) $\mathscr{K}_m(e_{00};x,y) = 1;$
(2) $\mathscr{K}_m(e_{10};x,y) = \frac{2mx+1}{2(m+1)};$
(3) $\mathscr{K}_m(e_{01};x,y) = \frac{2my+1}{2(m+1)};$
(4) $\mathscr{K}_m((s-x)^2;x,y) = \frac{3(m-1)x(1-x)+1}{3(m+1)^2};$
(5) $\mathscr{K}_m((t-y)^2;x,y) = \frac{3(m-1)y(1-y)+1}{3(m+1)^2}.$

Lemma 16. *The operators $(\mathscr{K}_m)_{m\ge 1}$ verify for any $(x,y) \in \Delta_2$, the following estimations*

(1) $\mathscr{K}_m((s-x)^2;x,y) \le \frac{1}{m+1};$
(2) $\mathscr{K}_m((t-y)^2;x,y) \le \frac{1}{m+1};$
(3) $\mathscr{K}_m((s-x)^2(t-y)^2;x,y) \le \frac{1}{(m+1)^2};$

for any $m \in \mathbb{N}$.

Approximation and Convergence Theorems for the Bivariate Operators of Kantorovich Type

Theorem 12. *Let the function $f \in C([0,1]\times[0,1])$. Then, for any $(x,y) \in \Delta_2$, any $m \in \mathbb{N}, m \ge 4$, we have*

$$|\mathscr{K}_m(f;x,y) - f(x,y)| \le \left(1+\frac{1}{\delta_1\sqrt{m+1}}\right)\left(1+\frac{1}{\delta_2\sqrt{m+1}}\right)\omega(f;\delta_1,\delta_2)$$

for any $\delta_1,\delta_2 > 0$ and

$$|\mathscr{K}_m(f;x,y) - f(x,y)| \le 4\omega\left(f;\frac{1}{\sqrt{m+1}},\frac{1}{\sqrt{m+1}}\right).$$

Corollary 1. *If $f \in C([0,1]\times[0,1])$, then*

$$\lim_{m\to\infty}\mathscr{K}_m(f;x,y) = f(x,y),$$

uniformly on Δ_2.

Bivariate Rational Type Functions

Following [9], let $f(x, y)$ be a function of two variables, defined in $[0, \infty) \times [0, \infty)$. By Bernstein type rational functions of two variables corresponding to $f(x, y)$, we mean the following:

$$R_{m,n}(f; x, y) = \frac{1}{(1 + a_n x)^n} \frac{1}{(1 + c_m y)^m} \sum_{k=0}^{n} \sum_{j=0}^{m} P_{n,k}(x) P_{m,j}(y) f\left(\frac{k}{b_n}, \frac{j}{d_m}\right), (n = m = 1, 2, ...), \quad (14)$$

where a_n, b_n, c_m, d_m are suitably chosen real numbers, independent of x and y and $P_{n,k}(x) = \binom{n}{k}(a_n x)^k$ and therefore $P_{m,j}(x) = \binom{m}{j}(c_m y)^j$.

Convergence Theorem

Let $R_{m,n}(f; x, y)$ be the functions defined by (14) with $a_n = \frac{b_n}{n}, b_n = n^{2/3}, c_m = \frac{d_m}{m}, d_m = m^{2/3} (n = m = 1, 2, \ldots)$ and let $\omega_{2A}(\delta)$ be the modulus of continuity of the function $f(x, y)$ in $[0, 2A] \times [0, 2A]$, there holds the following theorem:

Theorem 13. *Let $f(x, y)$ be a continuous function defined in $[0, \infty) \times [0, \infty)$ such that $f(x, y) = O(e^{\alpha(x+y)}) (x \to \infty, y \to \infty)$ for some real number α. Then in any square $0 \le x \le A, 0 \le y \le A, (A \ge 0)$ the inequality*

$$|R_{m,n}(f; x, y) - f(x, y)| \le c_0 \omega_{2A}\left(\sqrt{\frac{1}{n^{2/3}} + \frac{1}{m^{2/3}}}\right) \quad (15)$$

is valid if n, m are sufficiently large, where c_0 is a constant depending on A and α only. The inequality (15) shows that $R_{m,n}(f; x, y) \to f(x, y)$ when $x \ge 0, y \ge 0$ if $m, n \to \infty$, and this convergence is uniform in every finite square.

Asymptotic Approximation

Theorem 14. *Let $f(t, \tau)$ be a function defined in $[0, \infty) \times [0, \infty)$ such that $f(t, \tau) = O(e^{\alpha(t+\tau)}) (t, \tau \to \infty, \alpha$ is a fixed real number), then at each point (x, y) in which f''_{xx}, f''_{xy}, and f''_{yy} exist finitely, we have*

$$R_{m,n}(f;x,y) = f(x,y) + a_n f'_x(x,y)\left(\frac{-x^2}{1+a_nx}\right) + c_n f'_y(x,y)\left(\frac{-y^2}{1+c_my}\right)$$
$$+a_n f''_{xx}(x,y)\left(\frac{a_nb_nx^4+\frac{x}{b_n}}{2b_n(1+a_nx)^2}\right)+a_nc_m f''_{xy}(x,y)\left(\frac{x^2y^2}{(1+a_nx)(1+c_my)}\right)$$
$$+c_m f''_{yy}(x,y)\left(\frac{c_md_my^4+\frac{y}{c_m}}{2d_m(1+c_my)^2}\right) + (a_n+c_m)\rho_{m,n},$$

where $\rho_{m,n} \to 0, a_n = \frac{b_n}{n} \to 0, c_m = \frac{d_m}{m} \to 0, \frac{n^{1/2}}{b_n} \to 0, \frac{m^{1/2}}{d_m} \to 0$*, as* $m, n \to \infty$.

Bivariate Operators of Durrmeyer Type

Let $m, n \in \mathbb{N}$. The bivariate operator of Durrmeyer type $M_{m,n} : L_1([0,1]\times[0,1]) \to C([0,1]\times[0,1])$ is defined for any function $f \in L_1([0,1]\times[0,1])$ and any $(x,y) \in [0,1]\times[0,1]$ by

$$M_{m,n}(f;x,y) = (m+1)(n+1)\sum_{k=0}^{m}\sum_{j=0}^{n}\lambda_{m,k}(x)\lambda_{n,j}(y)\int_0^1\int_0^1 \lambda_{m,k}(t)\lambda_{n,j}(s)f(t,s)dtds.$$

Theorem 15 ([38]). *Let* $f : [0,1]\times[0,1] \to \mathbb{R}$ *be a bivariate function.*

(i) *If* $(x,y) \in [0,1]\times[0,1]$ *and* f *admits partial derivatives of second order continuous in a neighborhood of the point* (x,y)*, then*

$$\lim_{m\to\infty} m(M_{m,m}(f;x,y) - f(x,y)) = (1-2x)f'_x(x,y) + (1-2y)f'_y(x,y) + x(1-x)f''_{xx}(x,y) + y(1-y)f''_{yy}(x,y). \quad (16)$$

If f *admits partial derivatives of second order continuous on* $[0,1]\times[0,1]$*, then the convergence given in (16) is uniform on* $[0,1]\times[0,1]$.

(ii) *If* f *is continuous on* $[0,1]\times[0,1]$*, then*

$$|M_{m,m}(f;x,y) - f(x,y)| \le \frac{25}{4}\omega\left(f;\frac{1}{\sqrt{m}},\frac{1}{\sqrt{m}}\right)$$

for any $(x,y) \in [0,1]\times[0,1]$*, any natural number* $m, m \ge 3$.

Bivariate Operators of Stancu Type

Following Pop [38], for the real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2$ with $\alpha_1 \geq 0$ and $\alpha_2 \geq 0, m_1, m_2, \mu^{\alpha_1,\beta_1}$ and μ^{α_2,β_2} are defined through

$$m_i = \begin{cases} \max\{1, -[\beta_i]\} & , \text{ if } \beta_i \in \mathbb{R} \setminus \mathbb{Z} \\ \max\{1, 1-\beta_i\} & , \text{ if } \beta_i \in \mathbb{Z}, \end{cases}$$

$$\gamma_{\beta_i} = m_i + \beta_i = \begin{cases} \max\{1+\beta_i, \{\beta_i\}\} & , \text{ if } \beta_i \in \mathbb{R} \setminus \mathbb{Z} \\ \max\{1+\beta_i, 1\} & , \text{ if } \beta_i \in \mathbb{Z}, \end{cases}$$

$$\mu^{\alpha_i,\beta_i} = \begin{cases} 1 & , \text{ if } \alpha_i \leq \beta_i \\ 1 + \frac{\alpha_i - \beta_i}{\gamma_{\beta_i}} & , \text{ if } \alpha_i > \beta_i, \end{cases}$$

where $i \in \{1, 2\}$.
Let the bivariate operators of Stancu type $P_{m,n}^{\alpha_1,\beta_1,\alpha_2,\beta_2} : C([0, \mu^{\alpha_1,\beta_1}] \times [0, \mu^{\alpha_2,\beta_2}]) \rightarrow C([0,1] \times [0,1])$ be defined for any function $f \in C([0, \mu^{\alpha_1,\beta_1}] \times [0, \mu^{\alpha_2,\beta_2}])$ by

$$P_{m,n}^{\alpha_1,\beta_1,\alpha_2,\beta_2}(f; x, y) = \sum_{k=0}^{m} \sum_{j=0}^{n} \lambda_{m,k}(x) \lambda_{n,j}(y) f\left(\frac{k+\alpha_1}{m+\beta_1}, \frac{j+\alpha_2}{n+\beta_2}\right)$$

for any $(x, y) \in [0, 1] \times [0, 1]$ and any natural numbers $m, n, m \geq m_1$, and $n \geq m_2$.

Theorem 16. *Let $f : C([0, \mu^{\alpha_1,\beta_1}] \times [0, \mu^{\alpha_2,\beta_2}]) \rightarrow \mathbb{R}$ be a bivariate function.*

(*i*) *If $(x, y) \in [0, 1] \times [0, 1]$ and f admits partial derivatives of second order continuous in a neighborhood of the point (x, y), then*

$$\lim_{m\to\infty} m(P_{m,m}^{\alpha_1,\beta_1,\alpha_2,\beta_2}(f; x, y) - f(x, y)) = (\alpha_1 - \beta_1 x) f'_x(x, y) + (\alpha_2 - \beta_2 y) f'_y(x, y) + \frac{x(1-x)}{2} f''_{xx}(x, y) + \frac{y(1-y)}{2} f''_{yy}(x, y). \tag{17}$$

If f admits partial derivatives of second order continuous on $[0, 1] \times [0, 1]$, then the convergence given in (17) is uniform on $[0, 1] \times [0, 1]$.

(*ii*) *If f is continuous on $[0, 1] \times [0, 1]$, then*

$$|P_{m,m}^{\alpha_1,\beta_1,\alpha_2,\beta_2}(f; x, y) - f(x, y)| \leq \frac{81}{16} \omega\left(f; \frac{1}{\sqrt{m}}, \frac{1}{\sqrt{m}}\right)$$

for any $(x, y) \in [0, 1] \times [0, 1]$, any natural number $m, m(0)$.

The Bernstein–Stancu polynomials [43] are defined by

$$B_{n,\alpha,\beta} = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k+\alpha}{n+\beta}\right) \tag{18}$$

Following [22], there are two type of extensions of the polynomial (18) in the case of two variables. The Bernstein type polynomials of nth degree with respect to x and mth degree with respect to y, corresponding to the square $S : \left[\frac{\alpha_1}{n+\beta_1}, \frac{n+\alpha_1}{n+\beta_1}\right] \times \left[\frac{\alpha_2}{n+\beta_2}, \frac{n+\alpha_2}{n+\beta_2}\right]$ is defined in the following form (see [42, 43]):

$$S_{n,m}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) = \sum_{i=0}^{n}\sum_{j=0}^{m} f\left(\frac{i+\gamma_1}{n+\sigma_1}, \frac{j+\gamma_2}{m+\sigma_2}\right) p_{n,i,\alpha_i,\beta_i}(x) q_{m,j,\alpha_2,\beta_2}(y),$$

where

$$p_{n,i,\alpha_i,\beta_i}(x) = \left(\frac{n+\beta}{n}\right)^n \binom{n}{i} (x - \frac{\alpha_1}{n+\beta_1})^i \left(\frac{n+\alpha_1}{n+\beta_2} - x\right)^{n-i},$$

$$q_{m,j,\alpha_2,\beta_2}(y) = \left(\frac{m+\beta}{m}\right)^m \binom{m}{j} \left(y - \frac{\alpha_2}{m+\beta_2}\right)^j \left(\frac{m+\alpha_2}{m+\beta_2} - y\right)^{m-j}$$

and $\alpha_k, \gamma_k, \beta_k, \sigma_k (k = 1, 2)$ are positive numbers providing $0 < \alpha_1 \le \gamma_1 \le \sigma_1 \le \beta_1$ and $\alpha_2, \gamma_2, \sigma_2, \beta_2$.

On the other hand, the Bernstein type polynomials on the triangle $\Delta := \{(x, y) : x + y \le \frac{n+2\alpha}{n+\beta}, x, y \ge \frac{\alpha}{n+\beta}\}$ are defined as

$$S_{n,n}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) = \sum_{k=0}^{n}\sum_{l=0}^{n-k} f\left(\frac{k+\alpha_1}{n+\beta_1}, \frac{l+\alpha_2}{n+\beta_2}\right) p_{n,\alpha,\beta}^{k,l}(x,y), \tag{19}$$

where

$$p_{n,\alpha,\beta}^{k,l}(x,y) = \left(\frac{n+\beta}{n}\right)^n \binom{n}{k}\binom{n-k}{l} \left(x - \frac{\alpha}{n+\beta}\right)^k \left(y - \frac{\alpha}{n+\beta}\right)^l \left(\frac{n+2\alpha}{n+\beta} - x - y\right)^{n-k-l}.$$

Theorem 17. *Let $f \in C(S)$. The following inequalities hold:*

(i) If $(\sigma_2 - \beta_2) \geq (\gamma_2 - \alpha_2)$ and $(\sigma_1 - \beta_1) \geq (\gamma_1 - \alpha_1)$

$$|S_{n,m}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) - f(x,y)| \leq \frac{3}{2}\left\{\omega^{(1)}\left(f;\frac{\sqrt{4(\sigma_1-\beta_1)^2(\frac{n+\alpha_1}{n+\beta_1})^2+n}}{(n+\sigma_1)}\right)\right.$$
$$\left.+\omega^{(2)}\left(f;\frac{\sqrt{4(\sigma_2-\beta_2)^2(\frac{m+\alpha_2}{m+\beta_2})^2+m}}{(m+\sigma_2)}\right)\right\};$$

and if $(\sigma_1 - \beta_1) < (\gamma_1 - \alpha_1)$ and $(\sigma_2 - \beta_2) < (\gamma_2 - \alpha_2)$

$$|S_{n,m}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) - f(x,y)| \leq \frac{3}{2}\left\{\omega^{(1)}\left(f;\frac{\sqrt{4(\gamma_1-\alpha_1)^2+n}}{(n+\sigma_1)}\right)\right.$$
$$\left.+\omega^{(2)}\left(f;\frac{\sqrt{4(\gamma_2-\alpha_2)^2+m}}{(m+\sigma_2)}\right)\right\}$$

(ii) If $(\sigma_2 - \beta_2) \geq (\gamma_2 - \alpha_2)$ and $(\sigma_1 - \beta_1) \geq (\gamma_1 - \alpha_1)$

$$|S_{n,m}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) - f(x,y)|$$
$$\leq \frac{3}{2}\left\{\omega\left(f;\sqrt{\frac{4(\sigma_1-\beta_1)^2(\frac{n+\alpha_1}{n+\beta_1})^2+n}{(n+\sigma_1)^2}+\frac{4(\sigma_2-\beta_2)^2(\frac{m+\alpha_2}{m+\beta_2})^2+m}{(m+\sigma_2)^2}}\right)\right\}.$$

Theorem 18. *Let $f \in C(\Delta)$, where Δ is defined in (19). Then the following inequalities hold:*

(i) If $(\beta - \beta_2) \geq (\alpha - \alpha_2)$ and $(\beta - \beta_1) \geq (\alpha - \alpha_1)$, then

$$|S_{n,n}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) - f(x,y)| \leq 2\omega\left(\frac{\sqrt{4(\beta-\beta_1)^2(\frac{n+\alpha}{n+\beta})^2+n}}{(n+\beta_1)},\right.$$
$$\left.\frac{\sqrt{4(\beta-\beta_2)^2(\frac{n+\alpha}{n+\beta})^2+n}}{(n+\beta_2)}\right);$$

(ii) If $(\beta - \beta_2) < (\alpha - \alpha_2)$ and $(\beta - \beta_1) < (\alpha - \alpha_1)$, then

$$|S_{n,n}^{\alpha_k,\gamma_k,\beta_k,\sigma_k}(f;x,y) - f(x,y)| \leq 2\omega\left(\frac{\sqrt{4(\alpha-\alpha_1)^2+n}}{(n+\beta_1)},\frac{\sqrt{4(\alpha-\alpha_2)^2+n}}{(n+\beta_2)}\right).$$

Bivariate Baskakov Operators

Let $\mathbb{N} := \{1, 2, \ldots\}, \mathbb{R}_+ := (0, \infty), \mathbb{R}_0 := \mathbb{R}_+ \cup \{0\}, \mathbb{R}^2_+ := \mathbb{R}_+ \times \mathbb{R}_+$ and $\mathbb{R}^2_0 := \mathbb{R}_0 \times \mathbb{R}_0$. Following [12], for a fixed $p \in \mathbb{N}_0$, we define the function w_p on $\mathbb{R}_0$ by

$$w_0(x) := 1, w_p(x) := (1 + x^p)^{-1} \qquad \text{if} \qquad p \geq 1.$$

Next, for fixed $p, q \in \mathbb{N}_0$, we define the weighted function $w_{p,q}$ on $\mathbb{R}^2_0$ by

$$w_{p,q}(x, y) := w_p(x)w_q(y),$$

and the weighted space $C_{p,q}$ of all real-valued continuous functions f on $\mathbb{R}^2_0$ for which $w_{p,q}f$ is uniformly continuous and bounded on $\mathbb{R}^2_0$ and the norm is given by the formula

$$||f||_{p,q} \equiv ||f(\cdot, \cdot)|| := \sup_{(x,y)\in\mathbb{R}^2_0} w_{p,q}(x, y)|f(x, y)|.$$

For $f \in C_{p,q}$, we define the modulus of continuity

$$\omega(f, C_{p,q}; t, s) := \sup_{0\leq h\leq t, 0\leq\delta\leq s} ||\Delta_{h,\delta} f(\cdot, \cdot)||_{p,q}, \qquad t, s \geq 0, \tag{20}$$

where $\Delta_{h,\delta} f(x, y) := f(x + h, y + \delta) - f(x, y)$ for $(x, y) \in \mathbb{R}^2_0$ and $h, \delta \in \mathbb{R}_0$.

For every $f \in C_{p,q}$, we have

$$\omega(f, C_{p,q}; t_1, s_1) \leq \omega(f, C_{p,q}; t_2, s_1) \leq \omega(f, C_{p,q}; t_2, s_2),$$

for $0 \leq t_1 < t_2$ and $0 \leq s_1 < s_2$, and

$$\lim_{t,s\to 0+} \omega(f, C_{p,q}; t, s) = 0. \tag{21}$$

Moreover, for fixed $m \in \mathbb{N}$ and $p, q \in \mathbb{N}_0$, let $C^m_{p,q}$ be the set of all functions $f \in C_{p,q}$ having the partial derivatives $\dfrac{\partial^k f}{\partial x^s \partial y^{k-s}} \in C_{p,q}, s = 1, 2, ...k, k = 1, 2, \ldots, m$.

In [25], Gurdek et al. considered the Baskakov operators

$$A_{m,n}(f; x, y) := \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} a_{m,j}(x)a_{n,k}(y)f\left(\frac{j}{m}, \frac{k}{n}\right),$$

where $a_{q,r}(z) := \binom{q-1+r}{r} z^r(1 + z)^{-q-r}$, and the Baskakov-Kantorovich operators

$$B_{m,n}(f; x, y) := \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} a_{m,j}(x)a_{n,k}(y)mn \int_{\frac{j}{m}}^{\frac{j+1}{m}} dt \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t, z)dz,$$

$(x, y) \in \mathbb{R}^2_0, m, n \in \mathbb{N}$ defined for functions $f \in C_{p,q}, p, q \in \mathbb{N}_0$.

Moreover, if $f \in C_{p,q}$ and if $f(x,y) = f_1(x)f_2(y)$ for all $(x,y) \in \mathbb{R}_0^2$, then

$$A_{m,n}(f(t,z);x,y) = A_m(f_1(t);x)A_n(f_2(z);y),\ B_{m,n}(f(t,z);x,y) = B_m(f_1(t);x)B_n(f_2(z);y),$$

for $(x,y) \in \mathbb{R}_0^2$, $m,n \in \mathbb{N}$. The authors [25] proved the following lemmas:

Lemma 17. *For all $n \in \mathbb{N}$ and $x \in \mathbb{R}_0$, we have*

$$A_n(t-x;x) = 0, \qquad B_n(t-x;x) = \frac{1}{2n},$$
$$A_n((t-x)^2;x) = \frac{\phi(x)}{n}, \qquad B_n((t-x)^2;x) = \frac{\phi(x)}{n} + \frac{1}{3n^2},$$

where $\phi(x) := x(1+x)$.

Lemma 18. *Let $L_n \in \{A_n, B_n\}$, i.e. $L_n = A_n$ for all $n \in \mathbb{N}$ or $L_n = B_n$ for all $n \in \mathbb{N}$. Then for every fixed $x_0 \in \mathbb{R}_0$ there exists a positive constant $M_1(x_0)$ such that for $n \in \mathbb{N}$,*

$$L_n((t-x)^4;x_0) \le M_1(x_0)n^{-2}.$$

Lemma 19. *Let $L_n \in \{A_n, B_n\}$. Then for every $p \in \mathbb{N}_0$ there exist positive constants $M_k(p)$, $k = 2,3$, such that*

$$w_p(x)L_n\left(\frac{1}{w_p(t)};x\right) \le M_2(p),$$
$$w_p(x)A_n\left(\frac{(t-x)^2}{w_p(t)};x\right) \le M_3(p)\frac{\phi(x)}{n},$$
$$w_p(x)B_n\left(\frac{(t-x)^2}{w_p(t)};x\right) \le M_3(p)\frac{\phi(x)+1}{n},$$

for all $n \in \mathbb{N}$ and $x \in \mathbb{R}_0$.

Lemma 20. *For every $p \in \mathbb{N}_0$ there exists a positive constant $M_4(p)$ such that*

$$w_p(x)\sum_{k=0}^{\infty}\left|\frac{d}{dx}a_{n,k}(x)\right|\left(w_p\left(\frac{k}{n}\right)\right)^{-1} \le M_4(p)n,$$
$$w_p(x)\sum_{k=0}^{\infty}\left|\frac{d}{dx}a_{n,k}(x)\right| n\int_{k/n}^{\frac{k+1}{n}}\frac{dt}{w_p(t)} \le M_4(p)n,$$

for all $n \in \mathbb{N}$ and $x \in \mathbb{R}_0$.

Lemma 21. *Let $L_{m,n} \in \{A_{m,n}, B_{m,n}\}$. Then for every $p,q \in \mathbb{N}_0$ there exist two positive constants $M_i(p,q)$, $i = 5,6$, such that*

$$\left\| L_{m,n}\left(\frac{1}{w_{p,q}(t,z)};\cdot,\cdot\right)\right\| \le M_5(p,q), \qquad m,n\in\mathbb{N},$$

and for every $f \in C_{p,q}$ *and for all* $m,n \in \mathbb{N}$

$$\|L_{m,n}(f;\cdot,\cdot)\|_{p,q} \le M_5(p,q)||f||_{p,q},$$

$$\left\|\frac{d}{dx}L_{m,n}(f;x,y)\right\|_{p,q} \le M_6(p,q)m||f||_{p,q},$$

$$\left\|\frac{d}{dy}L_{m,n}(f;x,y)\right\|_{p,q} \le M_6(p,q)n||f||_{p,q}.$$

Hence, $L_{m,n}$ *is linear positive operator from the space* $C_{p,q}$ *into* $C^1_{p,q}$.

Rate of Convergence

Theorem 19. *Suppose that* $f \in C^1_{p,q}$ *with some* $p,q \in \mathbb{N}_0$. *Then there exists a positive constant* $M_7(p,q)$ *such that for all* $(x,y) \in \mathbb{R}^2_+$ *and* $m,n \in \mathbb{N}$

$$w_{p,q}(x,y)|A_{m,n}(f;x,y)-f(x,y)| \le M_7(p,q)\left\{||f'_x||_{p,q}\sqrt{\frac{\phi(x)}{m}}+||f'_y||_{p,q}\sqrt{\frac{\phi(y)}{n}}\right\}$$

and

$$w_{p,q}(x,y)|B_{m,n}(f;x,y)-f(x,y)|\le M_7(p,q)\left\{||f'_x||_{p,q}\sqrt{\frac{\phi(x)+1}{m}}+||f'_y||_{p,q}\sqrt{\frac{\phi(y)+1}{n}}\right\}.$$

Theorem 20. *Suppose that* $f \in C_{p,q}$ *with some* $p,q \in \mathbb{N}_0$. *Then there exists a positive constant* $M_{11}(p,q)$ *such that for all* $(x,y) \in \mathbb{R}^2_+$ *and* $m,n \in \mathbb{N}$

$$w_{p,q}(x,y)|A_{m,n}(f;x,y)-f(x,y)| \le M_{11}(p,q)\omega\left(f, C_{p,q}; \sqrt{\frac{\phi(x)}{m}}, \sqrt{\frac{\phi(y)}{n}}\right)$$

and

$$w_{p,q}(x,y)|B_{m,n}(f;x,y)-f(x,y)| \le M_{11}(p,q)\omega\left(f, C_{p,q}; \sqrt{\frac{\phi(x)+1}{m}}, \sqrt{\frac{\phi(y)+1}{n}}\right)$$

for all $(x,y) \in \mathbb{R}^2_+$ *and* $m,n \in \mathbb{N}$.

Corollary 2. *Let* $L_{m,n} \in \{A_{m,n}, B_{m,n}\}$ *and let* $f \in C_{p,q}$ *with some* $p,q \in \mathbb{N}_0$. *Then for every* $(x,y) \in \mathbb{R}^2_+$

$$\lim_{m,n\to\infty} L_{m,n}(f;x,y) = f(x,y) \tag{22}$$

Moreover, the assertion (22) holds uniformly on every rectangle $0 \le x \le a, 0 \le y \le b$.

Voronovskaja Type Asymptotic Theorems

Theorem 21. *Suppose that* $f \in C^2_{p,q}$ *with some* $p, q \in \mathbb{N}_0$. *Then for every* $(x, y) \in \mathbb{R}^2_+$

$$\lim_{n\to\infty} n\{A_{n,n}(f; x, y) - f(x, y)\} = \frac{\phi(x)}{2} f''_{xx}(x, y) + \frac{\phi(y)}{2} f''_{yy}(x, y),$$

and

$$\lim_{n\to\infty} n\{B_{n,n}(f; x, y) - f(x, y)\} = \frac{1}{2}\{f'_x(x, y) + f'_y(x, y) + + \phi(x) f''_{xx}(x, y) + \phi(y) f''_{yy}(x, y)\}.$$

Theorem 22. *Let* $L_{m,n} \in \{A_{m,n}, B_{m,n}\}$ *and let* $f \in C^1_{p,q}$ *with some* $p, q \in \mathbb{N}_0$. *Then for every* $(x, y) \in \mathbb{R}^2_+$

$$\lim_{n\to\infty} \frac{\partial}{\partial x} L_{n,n}(f; x, y) = \frac{\partial f}{\partial x},$$
$$\lim_{n\to\infty} \frac{\partial}{\partial y} L_{n,n}(f; x, y) = \frac{\partial f}{\partial y}.$$

Generalized Baskakov Operators Mihesan [34] introduced an important generalization of the well-known Baskakov operators depending on a nonnegative constant a, independent of n as

$$B^a_n(f; x) = \sum_{k=0}^{\infty} W^a_{n,k}(x) f\left(\frac{k}{n}\right), \tag{23}$$

where

$$W^a_{n,k}(x) = e^{\frac{-ax}{1+x}} \frac{P_k(n, a)}{k!} \frac{x^k}{(1 + x)^{n-k}}, P_k(n, a) = \sum_{i=0}^{k} \binom{k}{i} (n)_i a^{k-i},$$

$$\text{with rising factorial given by } (n)_i = n(n + 1) \dots (n + i - 1), (n)_0 = 1. \tag{24}$$

Following [46, 47], in the space $C_{p,q}$, we define the operators for functions of two variables with nonnegative constants a, b, independent of n as

$$B^{a,b}_{m,n}(f; x, y) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} W^a_{m,j}(x) W^b_{n,k}(y) f\left(\frac{j}{m}, \frac{k}{n}\right), \qquad m, n \in \mathbb{N}, x, y \in \mathbb{R}^2_0, \tag{25}$$

Lemma 22. *For $f \in C_{p,q}$ and $B_{m,n}^{a,b}(f;x,y)$ as defined in (25), we have*

$$(B_{m,n}^{a,b}(f))'_x(x,y) = \frac{m}{x(1+x)}\{B_{m,n}^{a,b}(f(t,z)(t-x);x,y) - B_m^a(t-x;x)B_{m,n}^{a,b}(f(t,z);x,y)\}$$

and

$$(B_{m,n}^{a,b}(f))'_y(x,y) = \frac{n}{y(1+y)}\{B_{m,n}^{a,b}(f(t,z)(z-y);x,y) - B_n^b(z-y;y)B_{m,n}^{a,b}(f(t,z);x,y)\}.$$

Simultaneous Approximation

Theorem 23. *For $f \in C_{p,q}^1 a, b \geq 0$ and every $(x,y) \in \mathbb{R}_+^2$*

$$\lim_{n\to\infty}(B_{n,n}^{a,b}(f))'_x(x,y) = f'_x(x,y)$$

and

$$\lim_{n\to\infty}(B_{n,n}^{a,b}(f))'_y(x,y) = f'_y(x,y)$$

Voronovskaja Type Theorems

Theorem 24. *Let $f \in C_{p,q}^3$. Then, for every $(x,y) \in \mathbb{R}_+^2$, we have*

$$\begin{aligned} n\lim_{n\to\infty}\{(B_{n,n}^{a,b}(f))'_x(x,y) - f'_x(x,y)\} &= \frac{1+2x}{2}f''_{xx}(x,y) + \frac{y(1+y)}{2}f'''_{xyy}(x,y) \\ &+ \frac{x(1+x)}{2}f'''_{xxx}(x,y) + \frac{a}{(1+x)^2}f'_x(x,y) \\ &+ \frac{ax}{1+x}f''_{xx}(x,y) + \frac{ay}{1+y}f''_{xy}(x,y), \\ n\lim_{n\to\infty}\{(B_{n,n}^{a,b}(f))'_y(x,y) - f'_y(x,y)\} &= \frac{1+2y}{2}f''_{yy}(x,y) + \frac{x(1+x)}{2}f'''_{xxy}(x,y) \\ &+ \frac{y(1+y)}{2}f'''_{yyy}(x,y) + \frac{a}{(1+y)^2}f'_x(x,y) \\ &+ \frac{ay}{1+y}f''_{yy}(x,y) + \frac{ax}{1+x}f''_{xy}(x,y). \end{aligned}$$

Bivariate Case for Genaralized Baskakov–Kantorovich Operators Erencin [18] proposed a Durrmeyer type modification of the operators (23) by considering the weights of the beta basis function and established some direct results. Agrawal and Goyal [4] investigated the Kantorovich modification of the operators (23) and obtained some results concerning the rate of convergence in ordinary and simultaneous approximation. Subsequently, the same authors [24] considered the bivariate generalization of the generalized Baskakov–Kantorovich operators and discussed the rate of approximation.

In [24], for $f \in C_{(p,q)}$ we define a bivariate extension of the generalized Baskakov operators as follows:

$$G^a_{n_1,n_2}(f;x,y) = (n_1+1)(n_2+1)\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} W^a_{n_1,n_2,k_1,k_2}(x,y)$$
$$\int_{\frac{k_2}{n_2+1}}^{\frac{k_2+1}{n_2+1}}\int_{\frac{k_1}{n_1+1}}^{\frac{k_1+1}{n_1+1}} f(u,v)\,du\,dv, \qquad (26)$$

where

$$W^a_{n_1,n_2,k_1,k_2}(x,y) = \frac{x^{k_1}y^{k_2}p_{k_1}(n,a)p_{k_2}(n,a)e^{\frac{-ax}{1+x}}e^{\frac{-ay}{1+y}}}{k_1!k_2!(1+x)^{n_1+k_1}(1+y)^{n_2+k_2}}.$$

Auxiliary Results

Lemma 23. *Let $e_{i,j} : \mathbb{R}^2_0 \to \mathbb{R}^2_0, e_{i,j} = x^i y^j, 0 \le i,j \le 2$ be two-dimensional test functions. Then the bivariate operators defined in (26) satisfy the following results:*

(i) $G^a_{n_1,n_2}(e_{0,0};x,y) = 1;$

(ii) $G^a_{n_1,n_2}(e_{1,0};x,y) = \frac{1}{n_1+1}\left(n_1x + \frac{ax}{1+x} + \frac{1}{2}\right);$

(iii) $G^a_{n_1,n_2}(e_{0,1};x,y) = \frac{1}{n_2+1}\left(n_2y + \frac{ay}{1+y} + \frac{1}{2}\right);$

(iv) $G^a_{n_1,n_2}(e_{2,0};x,y) = \frac{1}{(n_1+1)^2}\left(n_1^2x^2 + n_1x^2 + 2n_1x + \frac{2an_1x^2}{1+x} + \frac{a^2x^2}{(1+x)^2} + \frac{2ax}{1+x} + \frac{1}{3}\right);$

(v) $G^a_{n_1,n_2}(e_{0,2};x,y) = \frac{1}{(n_2+1)^2}\left(n_2^2y^2 + n_2y^2 + 2n_2y + \frac{2an_2y^2}{1+y} + \frac{a^2y^2}{(1+y)^2} + \frac{2ay}{1+y} + \frac{1}{3}\right);$

(vi) $G^a_{n_1,n_2}(e_{3,0};x,y) = \frac{1}{(n_1+1)^3}\left(n_1^3x^3 + \frac{3n_1^2x^2}{2}(3+2x) + \frac{3n_1x^2}{2} + \frac{n_1x}{2}(4x^2 + 6x + 5) + \frac{3ax^3n_1^2}{1+x} + \frac{3an_1x^2}{1+x}\left\{(3+x) + \frac{ax}{1+x}\right\} + \frac{ax}{1+x}\left\{\frac{7}{2} + \frac{7}{2}\frac{ax}{(1+x)} + \frac{a^2x^2}{(1+x)^2}\right\} + \frac{1}{4}\right);$

(vii) $G^a_{n_1,n_2}(e_{0,3};x,y) = \frac{1}{(n_2+1)^3}\left(n_2^3y^3 + \frac{3n_2^2y^2}{2}(3+2y) + \frac{3n_2y^2}{2} + \frac{n_2y}{2}(4y^2 + 6y + 5) + \frac{3ay^3n_2^2}{1+y} + \frac{3an_2y^2}{1+y}\left\{(3+y) + \frac{ay}{1+y}\right\} + \frac{ay}{1+y}\left\{\frac{7}{2} + \frac{7}{2}\frac{ay}{(1+y)} + \frac{a^2y^2}{(1+y)^2}\right\} + \frac{1}{4}\right).$

Lemma 24. *For $n_1, n_2 \in \mathbb{N}$, we have*

(i) $G^a_{n_1,n_2}(u-x;x,y) = \dfrac{1}{n_1+1}\left(-x + \dfrac{ax}{1+x} + \dfrac{1}{2}\right)$;

(ii) $G^a_{n_1,n_2}(v-y;x,y) = \dfrac{1}{n_2+1}\left(-y + \dfrac{ay}{1+y} + \dfrac{1}{2}\right)$;

(iii) $G^a_{n_1,n_2}((u-x)^2;x,y) = \dfrac{1}{(n_1+1)^2}\left((n_1+1)x^2 + (n_1-1)x + \dfrac{a^2x^2}{(1+x)^2} + 2ax\left(\dfrac{1-x}{1+x}\right) + \dfrac{1}{3}\right)$;

(iv) $G^a_{n_1,n_2}((v-y)^2;x,y) = \dfrac{1}{(n_2+1)^2}\left((n_2+1)y^2 + (n_2-1)y + \dfrac{a^2y^2}{(1+y)^2} + 2ay\left(\dfrac{1-y}{1+y}\right) + \dfrac{1}{3}\right)$.

Remark 2. For every $x \in [0,\infty)$ and $n \in \mathbb{N}$, we have

$$G^a_{n_1,n_2}((u-x)^2;x,y) \le \frac{\{\delta^a_{n_1}(x)\}^2}{n_1+1},$$

where $\{\delta^a_{n_1}(x)\}^2 = \phi^2(x) + \frac{(1+a)^2}{n_1+1}$ and $\phi(x) = \sqrt{x(1+x)}$.

Lemma 25. *For every $p \in \mathbb{N}_0$ there exist positive constants $M_k(p), k = 1,2$ such that*

(i) $\omega_p(x)G^a_n\left(\dfrac{1}{\omega_p(t)};x\right) \le M_1(p)$,

(ii) $\omega_p(x)G^a_n\left(\dfrac{(t-x)^2}{\omega_p(t)};x\right) \le M_2(p)\dfrac{\{\delta^a_n(x)\}^2}{n+1}$,

for all $x \in \mathbb{R}_0 = \mathbb{R}_+ \cup \{0\}, \mathbb{R}_+ = (0,\infty)$ and $n \in \mathbb{N}$.

Lemma 26. *For every $p, q \in \mathbb{N}_0$ there exists positive constant $M_3(p,q)$, such that*

$$\| G^a_{n_1,n_2}(f;.,.) \|_{p,q} \le M_3(p,q) \| f \|_{p,q}$$

for every $f \in C_{p,q}$ and for all $n_1, n_2 \in \mathbb{N}$.

Local Approximation For $f \in C_B(\mathbb{R}^2_0)$ (the space of all bounded and uniformly continuous functions on I, let $C^2_B(\mathbb{R}^2_0) = \{f \in C_B(\mathbb{R}^2_0) : f^{(r,s)} \in C_B(\mathbb{R}^2_0), 1 \le r,s \le 2\}$, where $f^{(r,s)}$ is (r,s)th-order partial derivative with respect to x,y of f, equipped with the norm

$$||f||_{C^2_B(\mathbb{R}^2_0)} = ||f||_{C_B(\mathbb{R}^2_0)} + \sum_{i=1}^{2}\left\|\frac{\partial^i f}{\partial x^i}\right\| + \sum_{i=1}^{2}\left\|\frac{\partial^i f}{\partial y^i}\right\|.$$

The Peetre's K-functional of the function $f \in C_B(\mathbb{R}_0^2)$ is given by

$$\mathscr{K}(f;\delta) = \inf_{g \in C_B^2(\mathbb{R}_0^2)} \{||f-g||_{C_B(\mathbb{R}_0^2)} + \delta ||g||_{C_B^2(\mathbb{R}_0^2)}, \delta > 0\}.$$

It is also known that the following inequality

$$\mathscr{K}(f;\delta) \leq M_1\{\omega_2(f;\sqrt{\delta}) + \min(1,\delta)||f||_{C_B(\mathbb{R}_0^2)}\}$$

holds for all $\delta > 0$ ([15], page 192). The constant M_1 is independent of δ and f and $\omega_2(f;\sqrt{\delta})$ is the second order modulus of continuity which is defined in a similar manner as the second order modulus of continuity for one variable case:

$$\omega_2(f,\sqrt{\delta}) = \sup_{0<h\leq\sqrt{\delta}} \sup_{x,x+2h\in[0,\infty)} |f(x+2h) - 2f(x+h) + f(x)|.$$

Now, we find the order of approximation of the sequence $G^a_{n_1,n_2}(f;x,y)$ to the function $f(x;y) \in C_B^2(\mathbb{R}_0^2)$ by Peetre's K-functional.

Theorem 25. *For the function $f \in C_B(\mathbb{R}_0^2)$, the following inequality*

$$\begin{aligned}
&|G^a_{n_1,n_2}(f;x,y) - f(x,y)| \leq 4\mathscr{K}(f; M_{n_1,n_2}(x,y)) \\
&+\omega\left(f;\sqrt{\left(\frac{1}{n_1+1}\left(-x+\frac{ax}{1+x}+\frac{1}{2}\right)\right)^2+\left(\frac{1}{n_2+1}\left(-y+\frac{ay}{1+y}+\frac{1}{2}\right)\right)^2}\ \right) \\
&\leq M\left\{\omega_2\left(f;\sqrt{M_{n_1,n_2}(x,y)}\right) + min\{1, M_{n_1,n_2}(x,y)\}||f||_{C_B^2(I)}\right\} \\
&+\omega\left(f;\sqrt{\left(\frac{1}{n_1+1}\left(-x+\frac{ax}{1+x}+\frac{1}{2}\right)\right)^2+\left(\frac{1}{n_2+1}\left(-y+\frac{ay}{1+y}+\frac{1}{2}\right)\right)^2}\ \right)
\end{aligned}$$

holds, where M is a positive constant independent of f and
$M_{n_1,n_2}(x,y) = \dfrac{\{\delta^a_{n_1}(x)\}^2}{n_1+1} + \dfrac{\{\delta^a_{n_2}(y)\}^2}{n_2+1}.$

Rate of Convergence

Theorem 26. *Suppose that $f \in C^1_{p,q}$ with $p,q \in \mathbb{N}_0$ then there exists a positive constant $M_5(p,q)$ such that for all $(x,y) \in I$ and $n_1,n_2 \in \mathbb{N}$*

$$\omega_{p,q}(x,y)|G^a_{n_1,n_2}(f;x,y) - f(x,y)| \leq M_5(p,q)\left\{ \| f'_x \|_{p,q} \frac{\delta^a_{n_1}(x)}{\sqrt{n_1+1}} + \| f'_y \|_{p,q} \frac{\delta^a_{n_2}(y)}{\sqrt{n_2+1}} \right\}.$$

Theorem 27. *Suppose that $f \in C_{p,q}$ with some $p,q \in \mathbb{N}_0$. Then there exists a positive constant $M_9(p,q)$ such that*

$$\omega_{p,q}(x,y) \mid G^a_{n,n}(f(t,z);x,y) - f(x,y) \mid$$
$$\leq M_9(p,q)\omega\left(f; C_{p,q}; \frac{\delta^a_{n_1}(x)}{\sqrt{n_1+1}}, \frac{\delta^a_{n_2}(y)}{\sqrt{n_2+1}}\right),$$

for all $(x,y) \in \mathbb{R}^2_0$ *and* $n_1, n_2 \in \mathbb{N}$.

As a consequence of Theorem 27, we have

Theorem 28. *Let* $f \in C_{p,q}$ *with some* $p, q \in \mathbb{N}_0$. *Then for every* $(x,y) \in \mathbb{R}^2_0$,

$$\lim_{n_1,n_2 \to \infty} G^a_{n_1,n_2}(f;x,y) = f(x,y).$$

Voronovskaja Type Theorem

Theorem 29. *Let* $f \in C^2_{p,q}$. *Then for every* $(x,y) \in \mathbb{R}^2_0$,

$$\lim_{n\to\infty} n\{G^a_{n,n}(f;x,y) - f(x,y)\} = \left(-x + \frac{ax}{1+x} + \frac{1}{2}\right)f_x(x,y)$$
$$+\left(-y + \frac{ay}{1+y} + \frac{1}{2}\right)f_y(x,y)$$
$$+\frac{x}{2}(x+2)f_{xx}(x,y) + \frac{y}{2}(y+2)f_{yy}(x,y).$$

Simultaneous Approximation

Theorem 30. *Let* $f \in C^1_{p,q}$. *Then, for every* $(x,y) \in \mathbb{R}^2_+$ *we have*

$$\lim_{n\to\infty}\left(\frac{\partial}{\partial\omega}G^a_{n,n}(f;\omega,y)\right)_{\omega=x} = \frac{\partial f}{\partial x}(x,y),$$
$$\lim_{n\to\infty}\left(\frac{\partial}{\partial\nu}G^a_{n,n}(f;x,\nu)\right)_{\nu=y} = \frac{\partial f}{\partial y}(x,y).$$

Theorem 31. *Let* $f \in C^3_{p,q}$. *Then for every* $(x,y) \in \mathbb{R}^2_+$, *we have*

$$\lim_{n\to\infty} n\left\{\left(\frac{\partial}{\partial\omega}G^a_{n,n}(f;\omega,y)\right)_{\omega=x} - \frac{\partial f}{\partial x}(x,y)\right\} = \left(-1 + \frac{a}{(1+x)^2}\right)f_x(x,y)$$
$$+\left(1 + \frac{ax}{1+x}\right)f_{xx}(x,y)$$
$$+\left(-y + \frac{ay}{1+y} + \frac{1}{2}\right)f_{xy}(x,y)$$
$$+\frac{y}{2}(1+y)f_{xyy}(x,y)$$
$$+\frac{x}{2}(1+x)f_{xxx}(x,y)$$

and

$$\lim_{n\to\infty} n\left\{\left(\frac{\partial}{\partial \nu}G^{a}_{n,n}(f;x,\nu)\right)_{\nu=y} - \frac{\partial f}{\partial y}(x,y)\right\} = \left(-1+\frac{a}{(1+y)^2}\right)f_y(x,y)$$
$$+\left(1+\frac{ay}{1+y}\right)f_{yy}(x,y)$$
$$+\left(-x+\frac{ax}{1+x}+\frac{1}{2}\right)f_{xy}(x,y)$$
$$+\frac{x}{2}(1+x)f_{xxy}(x,y)$$
$$+\frac{y}{2}(1+y)f_{yyy}(x,y).$$

Bivariate Szász–Mirakyan Operators

For the functions $f:\mathbb{R}_0^2\to\mathbb{R}$, the Szász–Mirakyan operators [40] are defined as

$$S_{m,n}(f;x) := e^{-mx-ny}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(mx)^j(ny)^k}{j!k!}f\left(\frac{j}{m},\frac{k}{n}\right),$$

$(x,y)\in\mathbb{R}_0^2$ and $m,n\in\mathbb{N}$.

Bivariate Generalized Szász–Mirakyan Operators From [13] for $p>0$, let C_p denote the space of all functions $f(x)$ such that $e^{-px}f(x)$ is bounded and uniformly continuous on $\mathbb{R}_0$. For $f\in C_p$, Rempulska and Graczyk [41] introduced the generalized Szász–Mirakyan operators for $f\in C_p$ as follows:

$$S^{[r]}_{n,p}(f;x) = \frac{a}{A_r(nx)}\sum_{k=0}^{\infty}\frac{(nx)^{rk}}{(rk)!}f\left(\frac{rk}{n+p}\right),\quad x\in\mathbb{R}_0, n\in\mathbb{N}, \tag{27}$$

where $r\in\mathbb{N}$ is fixed and

$$A_r(t) := \sum_{k=0}^{\infty}\frac{t^{rk}}{(rk)!},\quad \text{for } t\in\mathbb{R}_0.$$

and studied the rate of convergence and the Voronovskaja type asymptotic theorem.

To discuss their results, first we give some definitions and auxiliary results. We denote by $E_{p,q}$, the exponential weighted space of functions of two variables.

For fixed $p\in\mathbb{N}_0$, we define the function υ_p on $\mathbb{R}_0$ by $\upsilon_p = e^{-px}$, for $p>0$ and $x\in\mathbb{R}_0$. Let $p,q>0$ be fixed numbers, let $\upsilon_{p,q}(x,y) = \upsilon_p(x)\upsilon_q(y) = e^{-px-qy}$ for $x,y\in\mathbb{R}_0^2$.

The space $E_{p,q}$ is the set of all functions $f : \mathbb{R}_0^2 \to \mathbb{R}$ such that $v_{p,q}f$ is uniformly continuous and bounded on $\mathbb{R}_0^2$ and the norm

$$\| f \|_{p,q} = \| f(.,.) \|_{p,q} := \sup_{(x,y)\in\mathbb{R}_0^2} v_{p,q}(x,y)|f(x,y)|.$$

Let $r, s \in \mathbb{N}$ be fixed and let A_r be given by (27). For functions $f \in E_{p,q}, p, q > 0$, we define the following generalized Szász–Mirakyan operators

$$S^{[r,s]}_{m,n;p,q}(f; x, y) = \frac{1}{A_r(mx)A_s(ny)} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(mx)^{rj}(ny)^{sk}}{(rj)!(sk)!} f\left(\frac{rj}{m+p}, \frac{sk}{n+q}\right), \tag{28}$$

$(x, y) \in \mathbb{R}_0^2$ and $m, n \in \mathbb{N}$. In particular, for $f_0(x, y) = 1$ we have

$$S^{[r,s]}_{m,n;p,q}(f_0; x, y) = 1, \text{ for } (x, y) \in \mathbb{R}_0^2, \text{ and } m, n \in \mathbb{N}.$$

If $r = s = 1$, then by (27) and (28),

$$S^{[1,1]}_{m,n;p,q}(f; x, y) = e^{-mx-ny} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(mx)^{j}(ny)^{k}}{(j)!(k)!} f\left(\frac{j}{m+p}, \frac{k}{n+q}\right).$$

Lemma 27. *For fixed $p, q > 0$ and $r, s \in \mathbb{N}$ there is $M_1(r, s) = M_0(r)M_0(s)$ where $M_0(t)$ is a positive constant depending only on t, for fixed $t \in \mathbb{N}$ such that for every $f \in E_{p,q}$,*

$$\| S^{[r,s]}_{m,n;p,q}(f) \|_{p,q} \le M_1(r, s) \| f \|_{p,q}, \ \textit{for } m, n \in \mathbb{N}.$$

Rate of Convergence

Theorem 33. *For fixed $p, q > 0$ and $r, s \in \mathbb{N}$ there exists $M_3 = M_3(p, q, r, s) =$ const. > 0 such that if $f \in E^1_{p,q}$, then*

$$v_{p,q}(x, y)|S^{[r,s]}_{m,n;p,q}f(x, y) - f(x, y)| \le M_3\left\{ \| f'_x \|_{p,q} \sqrt{\frac{x^2+x}{m+p}} + \| f'_y \|_{p,q} \sqrt{\frac{y^2+y}{n+q}} \right\},$$

for $(x, y) \in \mathbb{R}_0^2$ and $m, n \in \mathbb{N}$.

Theorem 32. *Let $p, q > 0$ and $r, s \in \mathbb{N}$ be fixed. Then there exists $M_2 = M_2(p, q, r, s) =$ const. > 0 such that if $f \in E_{p,q}$, then*

$$v_{p,q}(x, y)|S^{[r,s]}_{m,n;p,q}f(x, y) - f(x, y)| \le M_2\omega\left(f; E_{p,q}; \sqrt{\frac{x^2+x}{m+p}}, \sqrt{\frac{y^2+y}{n+q}}\right)$$

for $(x, y) \in \mathbb{R}_0^2$ *and* $m, n \in \mathbb{N}$, *where* $\omega(f; E_{p,q})$ *is the modulus of continuity of* f *defined by (20).*

From Theorem 32 and (21), we get the following:

Corollary 3. *For fixed* $p, q > 0$ *and* $r, s \in \mathbb{N}$ *and every* $f \in E_{p,q}$ *there holds*

$$\lim_{m,n\to\infty} S^{[r,s]}_{m,n;p,q} f(x, y) = f(x, y), \ \text{at every } (x, y) \in \mathbb{R}_0^2.$$

This convergence is uniform on every rectangle $[x_1, x_2] \times [y_1, y_2]$ *with* $x_1, y_1 \geq 0$.

Voronovskaja Type Theorem

Theorem 34. *Let* p, q, r, *and* s *satisfy the assumptions of Theorem 32. If* $f \in E^2_{p,q}$, *then*

$$\lim_{n\to\infty} n\left\{ S^{[r,s]}_{n,n;p,q} f(x, y) - f(x, y)\right\} = -pxf'_x(x, y) - qyf'_y(x, y) + \frac{x}{2} f''_{xx}(x, y) + \frac{y}{2} f''_{yy}(x, y)$$

for $(x, y) \in \mathbb{R}_0^2$.

Bivariate Case of BBH Operators

Adell et al. [2] exhibited a bivariate version of the BBH operators as follows: Let as set $\Delta := \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, xy < 1\}$ and define for $(x, y) \in \Delta, n \in \mathbb{N}$ and any real function f on Δ

$$L_n(f; x, y) = \sum_{k=0}^{n} \sum_{l=0}^{n-k} f\left(\frac{k}{n-k+1}, \frac{l}{n-l+1}\right) \binom{n}{k, l} \left(\frac{x}{1+x}\right)^k \left(\frac{y}{1+y}\right)^l \left(\frac{1-xy}{(1+x)(1+y)}\right)^{n-k-l}, \ n \in \mathbb{N}$$

with the multinomial coefficient $\binom{n}{k,l} = \frac{n!}{k!l!(n-k-l)!}$. Abel [1] obtained the complete asymptotic expansion for the operators L_n, as n tends to ∞.

Alternatively, the bivariate case for BBH operator [1] is defined by

$$L_n(f; x, y) = (1 + x + y)^{-n} \sum_{k=0}^{n} \sum_{j=0}^{n-k} f\left(\frac{k}{n-k-j+1}, \frac{j}{n-k-j+1}\right) \binom{n}{k, j} x^k y^j, \quad \text{for } f \in C(\mathbb{R}_0^2). \tag{29}$$

Khan [46] studied the convergence properties of these operators and a Voronovskaja type theorem. He proved the following results:

Theorem 35. *Let* $f \in C(\mathbb{R}_0^2)$*, where* $\mathbb{R}_0^2 = \{(x,y); x \geq 0, y \geq 0\}$*. Then*

$$|L_n(f;x,y) - f(x,y)| \leq \left(1 + \sqrt{(1+x+y)(9(x^2+y^2)+8(x+y))}\right)\omega\left(f, \frac{1}{\sqrt{n+1}}\right).$$

Theorem 36. *Let* $f \in C_B(\mathbb{R}_0^2)$*, and let* $N(x,y)$ *be the least integer* $\geq \frac{1}{2}\gamma(x,y) - 1$*, where* $\gamma(x,y) = (x+y)(1+x+y)\frac{b(x,y)+b(y,x)}{2} + \sqrt{b(x,y)b(y,x)}$*. Then for* $n \geq N(x,y)$ *we have*

$$L_n(f;x,y) - f(x,y) \leq 2M\left[\omega_2\left(f, \sqrt{\frac{\gamma(x,y)}{2(n+1)}}\right) + ||f||\frac{\gamma(x,y)}{2(n+1)}\right],$$

where M *is a constant.*

Let

$$\mathscr{B} := \{f \in C(\mathbb{R}_0^2), \text{ for each } \alpha > 0 \text{ and each } \beta > 0, f(x,y) = O(1)e^{\alpha x + \beta y}\}.$$

Theorem 37. $L_n(f;x,y) \to f(x,y)$*, as* $n \to \infty$*,* $\quad \forall f \in \mathscr{B}$ *and for each* $(x,y) \in \mathbb{R}_0^2$*.*

Theorem 38. *Let* $f \in \mathscr{B}$ *and suppose that the first two partial derivatives of* f *exist and are continuous in* $N := \{(\alpha,\beta) : |\alpha - x| < \delta_1, |\beta - y| < \delta_2, \delta_1, \delta_2 > 0\}$*. Let* $L_n(f,x,y)$ *be defined by (29). Then*

$$\lim_{n\to\infty}(L_n(f,x,y) - f(x,y)) = \frac{(1+x+y)}{2}\left\{x(1+x)f_{xx} + 2xyf_{xy} + y(1+y)f_{yy}\right\}.$$

Let $m, n \in \mathbb{N}$ and $C_B(\mathbb{R}_0^2)$ denote the space of all bounded and continuous functions on $\mathbb{R}_0^2$. For any function $f \in C_B(\mathbb{R}_0^2)$ and any $(x,y) \in \mathbb{R}_0^2$, the bivariate operator of Bleimann–Butzer–Hahn type $L_{m,n} : C_B(\mathbb{R}_0^2) \to C_B(\mathbb{R}_0^2)$ is defined by

$$L_{m,n}(f;x,y) = \frac{1}{(1+x)^m(1+y)^n}\sum_{k=0}^{m}\sum_{j=0}^{n}\binom{m}{k}\binom{n}{j}f\left(\frac{k}{m+1-k}, \frac{j}{n+1-j}\right).$$

Pop [38] proved that

Theorem 39. *Let* $f : [0,\infty) \times [0,\infty) \to \mathbb{R}$ *be a bivariate function.*

(i) If $(x,y) \in [0,\infty) \times [0,\infty)$ *and* f *admits partial derivatives of second order continuous in a neighborhood of the point* (x,y)*, then*

$$\lim_{m\to\infty} m(L_{m,m}(f;x,y) - f(x,y)) = \frac{x(1+x)^2}{2}f''_{xx}(x,y) + \frac{y(1+y)^2}{2}f''_{yy}(x,y).$$

(ii) If f is continuous on $[0,\infty)\times[0,\infty)$*, then*

$$|L_{m,m}(f;x,y)-f(x,y)|\leq(1+8b(1+b)^2+16b^2(1+b)^4)\omega\left(f;\frac{1}{\sqrt{m}},\frac{1}{\sqrt{m}}\right)$$

for any $(x,y)\in[0,b]\times[0,b]$*, any natural number* $m,\ m\geq 24(1+b)$.

Applications of q-Calculus In the last decade, one of the most important areas of research in approximation theory is the application of q-calculus. In the theory of approximation by linear positive operators, Lupas [31] was the first person who initiated the study in this direction by introducing a q-analogue of the classical Bernstein polynomials. In 1997, Phillips [37] gave another generalization of Bernstein polynomials based on q-integers which became more popular. After that many researchers published papers in this direction. Recently, Aral et al. [8] compiled the research work done on the q-analogues of different sequences and classes of operators and discussed their approximation properties. We mention below some of the notations and definitions. The details can be found in [19, 27]. Let parameter q be a positive real number and n a nonnegative integer. $[n]_q$ denotes a q-integer, defined by

$$[n]_q:=\begin{cases}\dfrac{1-q^n}{1-q}, & q\neq 1 \text{ for } n\in\mathbb{N} \text{ and } [0]_q=0,\\ n, & q=1,\end{cases}$$

The q-analogue of factorial n is defined by

$$[n]_q!=\begin{cases}[n]_q[n-1]_q\ldots[1]_q, & \mathrm{n}=1,2\ldots,\\ 1, & \mathrm{n}=0,\end{cases}$$

The q-binomial coefficient $\begin{bmatrix}n\\k\end{bmatrix}_q$ and the q-product $(a+b)_q^n$ are defined by

$$\begin{bmatrix}n\\k\end{bmatrix}_q=\frac{[n]_q!}{[n-k]_q![k]_q!}$$

and

$$(a+b)_q^n=\prod_{j=0}^{n-1}(a+q^jb)$$

respectively.

The q-Jackson integral and q-improper integral are given by

$$\int_a^b f(x)d_qx = (1-q)a\sum_{n=0}^{\infty} f(aq^n)q^n; a > 0$$

and

$$\int_a^{\infty/A} f(x)d_qx = (1-q)\sum_{n=-\infty}^{\infty} f\left(\frac{q^n}{A}\right)\frac{q^n}{A}; A > 0$$

respectively, where the sums are assumed to be absolutely convergent.

The Riemann type $q-$ integral introduced by Gauchman [47] and Marinkovic et al. [48] is defined as follows: $\int_a^b f(t)d_q^R t = (1-q)(b-a)\sum_{j=0}^{\infty} f(a+(b-a)q^j)$, where the real numbers a, b satisfy $0 < a < b$.

For $q \in (0,1)$ and any arbitrary real function $f : \mathbb{R} \to \mathbb{R}$, the q-derivative $D_qf(x)$ is defined as

$$D_qf(x) = \begin{cases} \dfrac{f(x) - f(qx)}{(1-q)x}, & x \neq 0, \\ \lim\limits_{x\to 0} D_qf(x), & x = 0, \end{cases}$$

The q-derivative of the product is given by the formula

$$D_q(f(x)g(x)) = f(qx)D_q(g(x)) + g(x)D_q(f(x)).$$

The q-analogue E_q^x of classical exponential function is defined as

$$E_q^x = \sum_{j=0}^{\infty} q^{j(j-1)/2}\frac{x^j}{[j]!}.$$

The q-analogue of beta function of second kind is defined by

$$B_q(t,s) = K(A,t)\int_0^{\infty/A} \frac{x^{t-1}}{(1+x)_q^{t+s}}d_qx,$$

where $K(x,t) = \dfrac{1}{x+1}x^t\left(1+\dfrac{1}{x}\right)_q^t(1+x)_q^{1-t}$, and $(a+b)_q^s = \Pi_{i=0}^{s-1}\left(a+q^ib\right)$, $s \in \mathbb{Z}^+$.

In particular, for any positive integers n, m, we have

$$K(x,n) = q^{\frac{n(n-1)}{2}}, \qquad K(x,0) = 1,$$

and

$$B_q(m,n) = \frac{\Gamma_q(m)\Gamma_q(n)}{\Gamma_q(m+n)}.$$

In what follows, we shall discuss the approximation properties of the bivariate case of the linear positive operators based on q-integers.

Bivariate Extension of q-Szász–Mirakjan–Kantorovich Operators

For $n \in \mathbb{N}, 0 < q_1, q_2 < 1$ and $0 \leq x < \frac{q_1}{1-q_1^n}, 0 \leq y < \frac{q_2}{1-q_2^n}$, the bivariate extension of q-Szász–Mirakjan–Kantorovich operators [35] is defined as follows :

$$K_n^{q_1,q_2}(f;x,y) = [n]_{q_1}[n]_{q_2}E_{q_1}\left(-[n]_{q_1}\frac{x}{q_1}\right)E_{q_2}\left(-[n]_{q_2}\frac{y}{q_2}\right)\sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\frac{([n]_{q_1}x)^k}{[k]_{q_1}q_1^k}\frac{([n]_{q_2}y)^l}{[l]_{q_2}q_2^l}$$
$$\int_{q_2[l]_{q_2}/[n]_{q_2}}^{[l+1]_{q_2}/[n]_{q_2}}\int_{q_1[k]_{q_1}/[n]_{q_1}}^{[k+1]_{q_1}/[n]_{q_1}} f(t,s)d_{q_1}^R t d_{q_2}^R s, \quad (30)$$

where

$$\int_c^d\int_a^b f(t,s)d_{q_1}^R t d_{q_2}^R s = (1-q_1)(1-q_2)(b-a)(c-d)$$
$$\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} f(a+(b-a)q_1^i, c+(c-d)q_2^j)q_1^i q_2^j. \quad (31)$$

Also, f is a q_R-integrable function, so the series in (31) converges.

First, we give the following Lemma:

Lemma 28. *Let $e_{ij} = x^i y^j, i,j \in \mathbb{N}_0 \times \mathbb{N}_0$ with $i+j \leq 2$ be the two-dimensional test functions, Then the following results hold for the operators given by (30):*

(i) $K_n^{q_1,q_2}(e_{00};x,y) = 1;$
(ii) $K_n^{q_1,q_2}(e_{10};x,y) = x + \frac{1}{[2]_{q_1}[n]_{q_1}};$
(iii) $K_n^{q_1,q_2}(e_{01};x,y) = y + \frac{1}{[2]_{q_2}[n]_{q_2}};$
(iv) $K_n^{q_1,q_2}(e_{20};x,y) = q_1x^2 + \left(q_1 + \frac{2}{[2]_{q_1}}\right)\frac{1}{[n]_{q_1}}x + \frac{1}{[3]_{q_1}[n]_{q_1}^2};$
(v) $K_n^{q_1,q_2}(e_{02};x,y) = q_2y^2 + \left(q_2 + \frac{2}{[2]_{q_2}}\right)\frac{1}{[n]_{q_2}}y + \frac{1}{[3]_{q_2}[n]_{q_2}^2}.$

Orkcu [33] established some $A-$ weighted statistical approximation properties of the operators given by (30). Further, the pointwise convergence result was obtained by means of modulus of continuity.
The Korovkin-type theorem for functions of two variables was proved by Volkov [45]. The theorem on weighted approximation for functions of several variables was proved by Gadjiev in [21].

Let B_ω be the space of real-valued functions defined on $\mathbb{R}^2$ and satisfying the bounded condition $|f(x,y)| \leq M_f \omega(x,y)$, where $\omega(x,y) \geq 1$ for all $(x,y) \in \mathbb{R}^2$ is called a weight function if it is continuous on $\mathbb{R}^2$ and $\lim\limits_{\sqrt{x^2+y^2}\to\infty} \omega(x,y) \to \infty$. We denote by C_ω the space of all continuous functions in the space B_ω with the norm $\| f \|_\omega = \sup_{(x,y)\in\mathbb{R}^2} \frac{|f(x,y)|}{\omega(x,y)}$.

Theorem 40 ([21]). *Let $\omega_1(x,y)$ and $\omega_2(x,y)$ be weight functions satisfying*

$$\lim_{\sqrt{x^2+y^2}\to\infty} \frac{\omega_1(x,y)}{\omega_2(x,y)} = 0.$$

Assume that T_n is a sequence of linear positive operators acting from C_{ω_1} to B_{ω_2}. Then, for $f \in C_{\omega_1}$,
$\lim\limits_{n\to\infty} \| T_n f - f \|_{\omega_2} = 0$ *if and only if* $\lim\limits_{n\to\infty} \| T_n F_v - F_v \|_{\omega_1} = 0$, $(v = 0,1,2,3)$,
where $F_0(x,y) = \frac{\omega_1(x)}{1+x^2+y^2}, F_1(x,y) = \frac{x\omega_1(x)}{1+x^2+y^2}, F_2(x,y) = \frac{y\omega_1(x)}{1+x^2+y^2}, F_3(x,y) = \frac{(x^2+y^2)\omega_1(x)}{1+x^2+y^2}$.

Let $A := (a_{nk})$ be an infinite summability matrix. For a given sequence $x := (x_k)$, the $A-$transform of x, denoted by $Ax := ((Ax)_n)$, is defined as $(Ax)_n := \sum\limits_{k=1}^{\infty} a_{nk}x_k$ provided the series converges for each n. A is said to be regular if $\lim\limits_n (Ax)_n = L$ whenever $\lim x = L$. Suppose that A is a nonnegative regular summability matrix. Then x is $A-$statistically convergent to L if for every $\epsilon > 0$, $\lim\limits_n \sum\limits_{k:|x_k-L|\geq\epsilon} a_{nk} = 0$, and we write $st_A - \lim x = L$. If $A = C_1$, the Cesaro matrix of order one, then A-statistical convergence reduces to the statistical convergence. Also, taking $A = I$, the identity matrix, $A-$statistical convergence coincides with the ordinary convergence. We consider $\omega_1(x,y) = 1 + x^2 + y^2$ and $\omega_2(x,y) = (1 + x^2 + y^2)^{1+\alpha}$ for $\alpha > 0$, $(x,y) \in \mathbb{R}_0^2$.
We obtain statistical approximation properties of the operator defined by (30) with the help of Korovkin-type Theorem 40. Let (q_1,n) and (q_2,n) be two sequences in the interval $(0,1)$ so that

$$st_A - \lim_n q_{1,n}^n = 1 \text{ and } st_A - \lim_n q_{2,n}^n = 1,$$
$$st_A - \lim_n \frac{1}{[n]_{q_{1,n}}} = 0 \text{ and } st_A - \lim_n \frac{1}{[n]_{q_{2,n}}} = 0. \tag{32}$$

Theorem 41. *Let $A = (a_{nk})$ be a nonnegative regular summability matrix, and let $(q_{1,n})$ and $(q_{2,n})$ be two sequences satisfying (32). Then, for any q_R-integrable function $f \in C_{\omega_1}(\mathbb{R}_0^2)$ and for $\alpha > 0$ we have*

$$st_A - \lim_n \| K_n^{q_{1,n},q_{2,n}} f - f \|_{\omega_2} = 0.$$

Let $C_B(\mathbb{R}_0^2)$ be the space of all bounded and uniformly continuous functions on $\mathbb{R}_0^2$. For $f \in C_B(\mathbb{R}_0^2)$, we have

$$|K_n^{q_1,q_2}(f;x,y) - f(x,y)| \leq 2\omega(\delta).$$

Orkcu [36] defined another extension of q-Szász–Mirakyan–Kantorovich operators in the bivariate case.

For $q_1, q_2 \in (0,1)$ and $m, n \in \mathbb{N} \times \mathbb{N}$, the new operators q-Szász–Mirakjan–Kantorovich operators for functions of two variables is defined as follows:

$$\begin{aligned} S_{m,n}^{q_1,q_2}(f;x,y) &= \sum_{l=0}^{\infty}\sum_{k=0}^{\infty} \frac{[m]_{q_1}^{k+1}x^k}{[k]_{q_1}!}\frac{[n]_{q_2}^{l+1}y^l}{[l]_{q_2}!} q_1^{k(k-1)-1} q_2^{l(l-1)-1} E_{q_1}(-[m]_{q_1}q_1^k x)E_{q_2}(-[n]_{q_2}q_2^l y) \\ &\times \int_{[l]_{q_2}/q_2^{k-1}[n]_{q_2}}^{[l+1]_{q_2}/q_2^{k-1}[n]_{q_2}} \int_{[k]_{q_1}/q_1^{k-1}[m]_{q_1}}^{[k+1]_{q_1}/q_1^{k-1}[m]_{q_1}} f(t,s) d_{q_1}^R t\, d_{q_2}^R s. \end{aligned} \tag{33}$$

Lemma 29. *Let $q_1, q_2 \in (0,1)$ and $m, n \in \mathbb{N} \times \mathbb{N}$. One has*

(i) $S_{m,n}^{q_1,q_2}(1;x,y) = 1$;
(ii) $S_{m,n}^{q_1,q_2}(e_{10};x,y) = x + \frac{q_1}{[2]_{q_1}[m]_{q_1}}$;
(iii) $S_{m,n}^{q_1,q_2}(e_{01};x,y) = y + \frac{q_2}{[2]_{q_2}[n]_{q_2}}$;
(iv) $S_{m,n}^{q_1,q_2}(e_{20};x,y) = x^2 + \left(1 + \frac{2q_1}{[2]_{q_1}}\right)\frac{1}{[m]_{q_1}}x + \frac{q_1^2}{[3]_{q_1}[m]_{q_1}^2}$;
(v) $S_{m,n}^{q_1,q_2}(e_{02};x,y) = y^2 + \left(1 + \frac{2q_2}{[2]_{q_2}}\right)\frac{1}{[n]_{q_2}}y + \frac{q_2^2}{[3]_{q_2}[n]_{q_2}^2}$.

Approximation Properties in Polynomial Weighted Spaces For bivariate operators $S_{m,n}^{q_1,q_2}$, let us recall the space $C_{p,q}$ as considered in section "Bivariate Baskakov operators":

For each $z \in \mathbb{R}_0$, define the function ϕ_z by $\phi_z^r(t) = (t-z)^r, t \in \mathbb{R}_0, r \in \mathbb{N}$. For the one-dimensional operator $L_s, s \in \mathbb{N}$, and for each $r \in \mathbb{N}$, a polynomial $\Gamma_r(t)$ exists such that

$$s^{r/2}(L_s\phi_z^r)(t) \leq \Gamma_r(t), \ deg\Gamma_r(t) \leq r. \tag{34}$$

Theorem 42 ([3]). *Let $(p,q) \in \mathbb{N}_0 \times \mathbb{N}_0$. Then for any $(m,n) \in \mathbb{N} \times \mathbb{N}$, the operator*

$$L_{m,n}(f;x,y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{m,i}(x)b_{n,j}(y)f(x_{m,i},y_{n,j}), \ (x,y) \in \mathbb{R}_0^2$$

verifies

$$\| L_{m,n}(\frac{1}{\omega_{p,q}};.) \|_{p,q} \leq c(p,q),$$

and hence $\| L_{m,n}(f;.) \|_{p,q} \leq c(p,q) \| f \|_{p,q}, \ f \in C_{p,q}$.

Theorem 43 ([3]). *Let* $(p, q) \in \mathbb{N}_0 \times \mathbb{N}_0$. *For any* $(m, n) \in \mathbb{N} \times \mathbb{N}$, *the operator* $L_{m,n}$ *verifies*

$$\omega_{p,q}(x, y)|(L_{m,n}f)(x, y) - f(x, y)| \leq c(p, q)\omega_f\left(\frac{\phi(x)}{\sqrt{m}}, \frac{\phi(y)}{\sqrt{n}}\right),$$

$(x, y) \in \mathbb{R}_0^2$, *where* ϕ *is given by*

$$\phi(t) = \sqrt{\Gamma_2(t)} + \sqrt{\Gamma_2(t) + \sum_{k=0}^{p}\binom{p}{k}\Gamma_{k+2}(t)},$$

with the polynomials $\Gamma_\nu(t), \nu = (2, p+2)$, *being indicated at (34) and* $c(p, q)$ *is a suitable constant.*

Convergence Theorem

Theorem 44 ([3]). *Let* $(p, q) \in \mathbb{N}_0 \times \mathbb{N}_0$ *and* $f \in C_{p,q}$. *For any* $(x, y) \in \mathbb{R}_0^2$, *there holds*

$$\lim_{m,n\to\infty}(L_{m,n}f)(x, y) = f(x, y), \; f \in C_{p,q}. \tag{35}$$

If K_1, K_2 *are compact intervals included in* $\mathbb{R}_0$, *then (35) holds uniformly on the domain* $K_1 \times K_2$.

Now, let $\rho(x, y) = (1 + x^2 + y^2)^{-1}$ and $C_2(\mathbb{R}_0^2)$ be the space of all functions $f(x, y)$ such that $\rho(x, y)f(x, y)$ is uniformly continuous and bounded on $\mathbb{R}_0^2$, endowed with the norm $\|f\|_2 = \sup_{(x,y)\in\mathbb{R}_0^2} |f(x, y)|$

Lemma 30. *The operator* $S_{m,n}^{q_1,q_2}, (m, n) \in \mathbb{N} \times \mathbb{N}, q_1, q_2 \in (0, 1)$ *given by (33) verifies*

$$\left\|S_{m,n}^{q_1,q_2}\left(\frac{1}{\rho(t, s)}; .\right)\right\|_2 \leq \frac{10}{3},$$

$$\left\|S_{m,n}^{q_1,q_2}(f; .)\right\|_2 \leq \frac{10}{3}\, \| f \|_2, \; f \in C(\mathbb{R}_0^2).$$

Theorem 45. *For any* $(m, n) \in \mathbb{N} \times \mathbb{N}$ *and* $f \in C_2(\mathbb{R}_0^2)$ *the operators* $S_{m,n}^{q_1,q_2}, q_1, q_2 \in (0, 1)$ *satisfy*

$$\rho(x, y)|(S_{m,n}^{q_1,q_2}f)(x, y) - f(x, y)| \leq \frac{10}{3}\omega_f\left(\frac{\phi(x, q_1)}{[m]_{q_1}}, \frac{\phi(y, q_2)}{[n]_{q_2}}\right),$$

where $(x, y) \in \mathbb{R}_0^2, \phi(x, q) = \frac{q}{[2]_q} + \left(1 + \frac{1}{q} + \frac{q}{[2]_q}\right)x^2 + \left(1 + \frac{3q}{[2]_q} + \frac{2q^2}{[3]_q}\right)x + \frac{q^3}{[4]_q}$.

We replace q_1 and q_2 in (33) by sequences $(q_{1,m}), (q_{2,n})$ so that

$$\lim_{m\to\infty} q_{1,m} = 1, \ \lim_{n\to\infty} q_{2,n} = 1,$$
$$\lim_{m\to\infty} \frac{1}{[n]_{q_{1,m}}} = 0, \ \lim_{n\to\infty} \frac{1}{[n]_{q_{2,n}}} = 0. \tag{36}$$

Let $C_2^2(\mathbb{R}_0^2)$ be the class of all functions f in which the second partial derivatives of f belong to $C_2(\mathbb{R}_0^2)$.

Theorem 46. *Let $(m, n) \in \mathbb{N} \times \mathbb{N}$ and let $(q_{1,m}), (q_{2,n})$ be sequences in the interval $(0, 1)$ satisfying (36). Let $f \in C_2^2(\mathbb{R}_0^2)$ q_R-integrable function. For any $(x, y) \in \mathbb{R}_0^2$, the pointwise convergence takes place*

$$\lim_{m,n\to\infty} (S_{m,n}^{q_{1,m},q_{2,n}} f) = f(x, y). \tag{37}$$

If K_1 and K_2 are compact intervals included in $\mathbb{R}_0$, then (37) holds uniformly on the domain $K_1 \times K_2$.

We present the Voronovskaya-type theorem.

Theorem 47. *Let $(q_{1,n}), (q_{2,n})$ be sequences in the interval $(0, 1)$ satisfying (36). Suppose that $f \in C_2^2(\mathbb{R}_0^2)$ and be a q_R-integrable function. Then, for every $(x, y) \in \mathbb{R}_0^2$, we have*

$$\lim_{n\to\infty} n\{(S_{n,n}^{q_{1,n},q_{2,n}} f) - f(x, y)\} = \frac{x}{2} f_{xx}(x, y) + \frac{y}{2} + \frac{1}{2} f_x(x, y) + \frac{1}{2} f_y(x, y).$$

Bivariate Extension for q-BBH Operators

Let $f : \mathbb{R}_0^2 \to R$ and $0 < q_m, q_n \leq 1$. Further, let $C_B(\mathbb{R}_0)$ be the space of all bounded and continuous functions on $\mathbb{R}_0$ and $H_\omega(\mathbb{R}_0)$, the subspace of $C_B(\mathbb{R}_0)$ consisting of all functions f satisfying $|f(t) - f(x)| \leq \omega(|(\frac{t}{1+t}) - (\frac{x}{1+x})|)$. Ersan [20] defined the bivariate extension of the q-Bleimann–Butzer and Hahn operators as follows:

$$L_{m,n}(f; q_m, q_n, x, y) = \frac{1}{l_{m,q_m}(x)} \frac{1}{l_{n,q_n}(y)} \times \sum_{k_1=0}^{m} \sum_{k_2=0}^{n} f\left(\frac{[k_1]}{[m-k_1+1]q_m^{k_1}}, \frac{[k_2]}{[n-k_2+1]q_n^{k_2}} \right)$$
$$\times q_m^{\frac{k_1(k_1-1)}{2}} q_n^{\frac{k_2(k_2-1)}{2}} \begin{bmatrix} m \\ k_1 \end{bmatrix}_{q_m} \begin{bmatrix} n \\ k_2 \end{bmatrix}_{q_n} x^{k_1} y^{k_2}. \tag{38}$$

Here $l_{m,q_m}(x) = \prod_{s=0}^{m-1} (1 + q_m^s x)$.

Choosing $q_m = q_n = 1$, BBH operators reduce to the classical bivariate BBH operators.

Lemma 31. *Let* $e_{ij} : \mathbb{R}_0^2 \to [0,1)$ *be the two-dimensional test function defined as* $e_{ij} = \left(\frac{x}{1+x}\right)^i \left(\frac{y}{1+y}\right)^j$. *Then we have the following items for the operator (38):*

(i) $L_{m,n}(e_{00}; q_m, q_n, x, y) = 1,$
(ii) $L_{m,n}(e_{10}; q_m, q_n, x, y) = \frac{[m]}{[m+1]}\frac{x}{1+x},$
(iii) $L_{m,n}(e_{01}; q_m, q_n, x, y) = \frac{[n]}{[n+1]}\frac{y}{1+y},$
(iv) $L_{m,n}(e_{20}; q_m, q_n, x, y) = \frac{[m][m-1]}{[m+1]^2}q_m^2 \times \frac{x^2}{(1+x)(1+q_m x)} + \frac{[m]}{[m+1]^2}\frac{x}{1+x},$
(v) $L_{m,n}(e_{02}; q_m, q_n, x, y) = \frac{[n][n-1]}{[n+1]^2}q_n^2 \times \frac{y^2}{(1+y)(1+q_n y)} + \frac{[n]}{[n+1]^2}\frac{y}{1+y}.$

Theorem 48 ([7]). *Let* A_n *be the sequence of linear positive operators acting from* $H_\omega(\mathbb{R}_0)$ *to* $C_B(\mathbb{R}_0)$ *satisfying*

$$\lim_{n\to\infty}\left\|A_n\left(\left(\frac{t}{1+t}\right)^\nu ; x\right) - \left(\frac{x}{1+x}\right)^\nu\right\|_{C_B} = 0, \qquad \nu = 0, 1, 2$$

then for any function $f \in H_\omega(\mathbb{R}_0)$

$$\lim_{n\to\infty}\|A_n(f) - f\|_{C_B} = 0$$

holds.

Let $H_\omega(\mathbb{R}_0^2)$, consisting of functions satisfying $|f(t,s) - f(x,y)| \le \omega(|(\frac{t}{1+t}, \frac{s}{1+s}) - (\frac{x}{1+x}, \frac{y}{1+y})|)$.

Theorem 49. *Let* $q = (q_m)$ *and* $q = q_n$ *satisfy* $0 < q_m \le 1, 0 < q_n \le 1$ *and let* $q_m \to 1$ *and* $q_n \to 1$ *for* $m, n \to \infty$. *If the sequence of linear positive operator* $A_{m,n}$:
$A_{m,n} : H_\omega(\mathbb{R}_0^2) \to C_B(\mathbb{R}_0^2)$ *satisfies the following conditions;*

(i) $\lim_{m,n\to\infty}\|A_{m,n}(e_{00}; q_m, q_n, x, y) - e_{00}\|_{C(\mathbb{R}_0^2)} = 0,$
(ii) $\lim_{m,n\to\infty}\|A_{m,n}(e_{10}; q_m, q_n, x, y) - e_{10}\|_{C(\mathbb{R}_0^2)} = 0,$
(iii) $\lim_{m,n\to\infty}\|A_{m,n}(e_{01}; q_m, q_n, x, y) - e_{01}\|_{C(\mathbb{R}_0^2)} = 0,$
(iv) $\lim_{m,n\to\infty}\|A_{m,n}(e_{20}; q_m, q_n, x, y) - e_{20}\|_{C(\mathbb{R}_0^2)} = 0,$

then for any function $f, f \in H_\omega(\mathbb{R}_0^2)$

$$\lim_{m,n\to\infty}\|A_{m,n}(f; q_m, q_n, x, y) - f\|_{C(\mathbb{R}_0^2)} = 0$$

holds. Here $e_{ij} : \mathbb{R}_0^2 \to [0,1], e_{ij} = \left(\frac{x}{1+x}\right)^i \left(\frac{y}{1+y}\right)^j$ *are two-dimensional test functions.*

Theorem 50. *Let $q = (q_m)$ and $q = q_n$ satisfy $0 < q_m \leq 1, 0 < q_n \leq 1$ and let $q_m \to 1$ and $q_n \to 1$ for $m, n \to \infty$. If the sequence of linear positive operator $L_{m,n} : H_\omega(\mathbb{R}_0^2) \to C_B(\mathbb{R}_0^2)$ satisfies conditions of Theorem 49 (i)–(iv), then $L_{m,n}$ converges uniformly to f in $\mathbb{R}_0^2$ for all $f \in H_\omega(\mathbb{R}_0^2)$. That is, $\forall f \in H_\omega(\mathbb{R}_0^2)$*

$$\lim_{m,n\to\infty} \|L_{m,n}(f; q_m, q_n, x, y) - f\|_{C(\mathbb{R}_0^2)} = 0.$$

Here $e_{ij} = \left(\frac{x}{1+x}\right)^i \left(\frac{y}{1+y}\right)^j$.

Theorem 51. *Let $q = (q_m)$ and $q = q_n$ satisfy $0 < q_m \leq 1, 0 < q_n \leq 1$ and let $q_m \to 1$ and $q_n \to 1$ for $m, n \to \infty$. So we have*

$$|L_{m,n}(f; q_m, q_n, x, y) - f(x, y)| \leq 4\tilde{\omega}(f; \delta_1(x), \delta_2(y))$$

$\forall f \in H_\omega(\mathbb{R}_0^2)$ and $x, y \geq 0$. Here

$$\delta_k(x) = bigg\{\frac{x^2}{(1+x)^2}\left(\frac{q_{n_k}[n_k]_{q_{n_k}}[n_k-1]_{q_{n_k}}(1+x)}{[n_k+1]_{q_{n_k}}^2(1+q_{n_k}x)} - \frac{2[n_k]_{q_{n_k}}}{[n_k+1]_{q_{n_k}}} + 1\right) + \frac{[n_k]_{q_{n_k}}x}{[n_k+1]_{q_{n_k}}^2(1+x)}\bigg\}^{\frac{1}{2}}, \quad k = 1, 2.$$

Bivariate Extension for *q*-MKZ Operators

Let $I^2 = [0, a] \times [0, a]$. Doğru and Gupta [17] considered a bivariate extension of q-MKZ operators

$$M_{m,n}(f; q_1; q_2; x, y) = u_{m,q_1}(x)u_{n,q_2}(y) \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} f\left(\frac{q_1^m}{k_1+m}, \frac{q_2^n}{k_2+n}\right)\begin{bmatrix} m+k \\ k_1 \end{bmatrix}\begin{bmatrix} n+k_2 \\ k_2 \end{bmatrix} x^{k_1}y^{k_2}, \quad (39)$$

where $u_{m,q}(x) = \prod_{s=0}^{n}(1 - xq^s)$.

Approximation Properties of the Bivariate *q*-MKZ Operators Let $C(I^2)$ be a Banach space with the norm $\| f \|_{C(I^2)} = \max_{(x,y)\in I^2} |f(x, y)|$. If $\lim_{m,n\to\infty} \| f_{m,n} - f \|_{C(I^2)} = 0$, then we say that the sequence $\{f_{m,n}\}$ converges uniformly to f.

Lemma 32. *Let $e_{ij} : I^2 \to I^2, e_{ij} = x^i y^j$, be the two-dimensional test functions. Then the following results hold for the operators (39):*

1. $M_{m,n}(e_{00};q_1;q_2;x,y) = 1$;
2. $M_{m,n}(e_{10};q_1;q_2;x,y) = q_1^m x$;
3. $M_{m,n}(e_{01};q_1;q_2;x,y) = q_1^n y$;
4. $q_1^{2m}x^2 \leq M_{m,n}(e_{20};q_1;q_2;x,y) \leq q_1^{2m+1}x^2 + \frac{q_1^{2m}}{[m]_{q_1}}$;
5. $q_2^{2n}y^2 \leq M_{m,n}(e_{02};q_1;q_2;x,y) \leq q_2^{2n+1}y^2 + \frac{q_2^{2n}}{[n]_{q_2}}$.

Let $(X_i)_{i=1,2}$ be compact subset of $\mathbb{R}$. For each $j = 1,2$ we denote by $pr_j : \prod_{i=1}^{2} X_i \to X_j$ the *jth* projection which is defined by $pr_j(x) := x_j$ for every $x = (x_i)_{i=1,2} \in \prod_{i=1}^{2} X_i$.

Theorem 52 ([6, 45]). *$\{1, pr_1, pr_2, pr_1^2 + pr_2^2\}$ is a Korovkin subset in $C([0,1] \times [0,1])$ for the identity operator.*

Replacing q in (39) by a sequence (q_n) in the interval $(0,1]$ so that

$$\lim_n q_n^n = 1 \text{ and } \lim_n \frac{1}{[n]_{q_n}} = 0. \tag{40}$$

In the light of above theorem, the following result holds.

Theorem 53. *If the sequences (q_1, n_1) and (q_2, n_2) satisfy conditions (40) in the interval $(0,1]$, then the sequence of operators (39) converges uniformly to $f(x,y)$ for any $f \in C(I^2)$.*

Corollary 4. *If the sequences (q_1, n_1) and (q_2, n_2) in $[0,1)$ satisfy conditions $q_i^{n_i} \to b_i$, $i = 1,2$ in the interval $(0,1]$, then*

$$\| M_{m,n}(f) - M_{\infty,\infty}(f) \| \to 0. \textit{ for every } f \in C(I^2).$$

Rates of Convergence of the Bivariate Operators

Theorem 54. *If the sequences (q_1, n_1) and (q_2, n_2) satisfy conditions (40) in the interval $(0,1]$, then*

$$\| M_{m,n}(f; q_{1,n_1}, q_{2,n_2}) - f \| \leq 4\omega(f, \delta_{1,n_1}, \delta_{2,n_2}), \quad \forall\ f \in C(I^2), 0 < a < 1,$$

where

$$\delta_{1,n_1} = \sqrt{(1 - q_{1,n_1}^{n_1})^2 a^2 + \frac{a q_{1,n_1}^{2n_1}}{[n_1]_{q_{1,n_1}}}} \text{ and } \delta_{2,n_2} = \sqrt{(1 - q_{2,n_2}^{n_2})^2 a^2 + \frac{a q_{2,n_2}^{2n_2}}{[n_2]_{q_{2,n_2}}}}.$$

Now, we introduce the following Lipschitz class in the bivariate case:

$$Lip_M(f; \alpha_1, \alpha_2) = \{f : |f(t,s) - f(x,y)| \leq M|t-x|^{\alpha_1}|s-y|^{\alpha_2}\},$$

where $0 \leq \alpha_1 \leq 1$ and $0 \leq \alpha_2 \leq 1$.

Theorem 55. *Let the sequences* (q_1, n_1) *and* (q_2, n_2) *satisfy conditions (40) in the interval* $(0, 1]$. *If* $f \in Lip_M(f; \alpha_1, \alpha_2)$*, then*

$$\| M_{m,n}(f; q_{1,n_1}, q_{2,n_2}) - f \| \leq M\delta_{1,n_1}^{\alpha_1}\delta_{2,n_2}^{\alpha_2},$$

where $\delta_{1,n_1}, \delta_{2,n_2}$ *are same as defined in Theorem 54.*

Bernstein–Schurer–Kantorovich Operators Based on q-Integers Let $I = [0, 1 + p]$, where $p \in \{0, 1, 2, \ldots\}$ and $J = [0, 1]$. For $I^2 = I \times I$, let $C(I^2)$ denote the space of all real-valued continuous functions on I^2 endowed with the norm $\|f\|_I = \sup_{(x,y)\in I^2} |f(x, y)|$. Analogously, for $J^2 = J \times J$, we denote by $\|f\|_J = \sup_{(x,y)\in J^2} |f(x, y)|$, the sup-norm on J^2, where $J = [0, 1]$.

For $f \in C(I^2)$ and $0 < q_{n_1}, q_{n_2} < 1$, Agrawal et al. [5] defined the bivariate case of Kantorovich type q-Bernstein–Schurer operators as follows:

$$K_{n_1,n_2,p}(f; q_{n_1}, q_{n_2}, x, y) = [n_1+1]_{q_{n_1}} [n_2+1]_{q_{n_2}} \sum_{k_1=0}^{n_1+p} \sum_{k_2=0}^{n_2+p} b_{n_1+p,n_2+p,k_1,k_2}^{q_{n_1},q_{n_2}}(x, y) q_{n_1}^{-k_1} q_{n_2}^{-k_2}$$
$$\times \int_{[k_2]_{q_{n_2}}/[n_2+1]_{q_{n_2}}}^{[k_2+1]_{q_{n_2}}/[n_2+1]_{q_{n_2}}} \int_{[k_1]_{q_{n_1}}/[n_1+1]_{q_{n_1}}}^{[k_1+1]_{q_{n_1}}/[n_1+1]_{q_{n_1}}} f(t, s) d_{q_{n_1}}^R t\, d_{q_{n_2}}^R s, \tag{41}$$

where $b_{n_1+p,n_2+p,k_1,k_2}^{q_{n_1},q_{n_2}}(x, y) = \binom{n_1 + p}{k_1}_{q_{n_1}} \binom{n_2 + p}{k_2}_{q_{n_2}} x^{k_1} y^{k_2} (1 - x)_{q_{n_1}}^{n_1+p-k_1} (1 - y)_{q_{n_2}}^{n_2+p-k_2}$, $(x, y) \in J^2$ and $\int_c^d \int_a^b f(t, s) d_{q_1}^R t d_{q_2}^R s$ is given by (31).

Lemma 33. *Let* $e_{ij} = x^i y^j$, $(x, y) \in I^2$, $(i, j) \in \mathbb{N}^0 \times \mathbb{N}^0$ *with* $i + j \leq 2$ *be the two-dimensional test functions. Then, the following equalities hold for the operators given by (41):*

(i) $K_{n_1,n_2,p}(e_{00}; q_{n_1}, q_{n_2}, x, y) = 1$;

(ii) $K_{n_1,n_2,p}(e_{10}; q_{n_1}, q_{n_2}, x, y) = \frac{[n_1+p]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}} \frac{2q_{n_1}}{[2]_{q_{n_1}}} x + \frac{1}{[2]_{q_{n_1}} [n_1+1]_{q_{n_1}}}$;

(iii) $K_{n_1,n_2,p}(e_{01}; q_{n_1}, q_{n_2}, x, y) = \frac{[n_2+p]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}} \frac{2q_{n_2}}{[2]_{q_{n_2}}} y + \frac{1}{[2]_{q_{n_2}} [n_2+1]_{q_{n_2}}}$;

(iv) $K_{n_1,n_2,p}(e_{20}; q_{n_1}, q_{n_2}, x, y) = \frac{1}{[n_1+1]_{q_{n_1}}^2 [3]_{q_{n_1}}} + \frac{q_{n_1}(3 + 5q_{n_1} + 4q_{n_1}^2)}{[2]_{q_{n_1}} [3]_{q_{n_1}}} \frac{[n_1+p]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}^2} x + \frac{q_{n_1}^2 (1+q_{n_1} + 4q_{n_1}^2)}{[2]_{q_{n_1}} [3]_{q_{n_1}}} \frac{[n_1+p]_{q_{n_1}} [n_1+p-1]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}^2} x^2$;

(v) $K_{n_1,n_2,p}(e_{02};q_{n_1},q_{n_2},x,y)=\frac{1}{[n_2+1]_{q_{n_2}}^2[3]_{q_{n_2}}}+\frac{q_{n_2}(3+5q_{n_2}+4q_{n_2}^2)}{[2]_{q_{n_2}}[3]_{q_{n_2}}}\frac{[n_2+p]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}^2}y+\frac{q_{n_2}^2(1+q_{n_2}+4q_{n_2}^2)}{[2]_{q_{n_2}}[3]_{q_{n_2}}}\frac{[n_2+p]_{q_{n_2}}[n_2+p-1]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}^2}y^2.$

Remark 3. From Lemma 33, we get

(i) $K_{n_1,n_2,p}((t-x);q_{n_1},q_{n_2},x,y)=\frac{1}{[2]_{q_{n_1}}[n_1+1]_{q_{n_1}}}+x\left(\frac{[n_1+p]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}}\frac{2q_{n_1}}{[2]_{q_{n_1}}}-1\right);$

(ii) $K_{n_1,n_2,p}((s-y);q_{n_1},q_{n_2},x,y)=\frac{1}{[2]_{q_{n_2}}[n_2+1]_{q_{n_2}}}+y\left(\frac{[n_2+p]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}}\frac{2q_{n_2}}{[2]_{q_{n_2}}}-1\right);$

(iii) $K_{n_1,n_2,p}((t-x)^2;q_{n_1},q_{n_2},x,y)=\left(\frac{q_{n_1}^2(1+q_{n_1}+4q_{n_1}^2)}{[2]_{q_{n_1}}[3]_{q_{n_1}}}\frac{[n_1+p]_{q_{n_1}}[n_1+p-1]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}^2}-\frac{[n_1+p]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}}\frac{4q_{n_1}}{[2]_{q_{n_1}}}+1\right)x^2+\left(\frac{q_{n_1}(3+5q_{n_1}+4q_{n_1}^2)}{[2]_{q_{n_1}}[3]_{q_{n_1}}}\frac{[n_1+p]_{q_{n_1}}}{[n_1+1]_{q_{n_1}}^2}-\frac{2}{[n_1+1]_{q_{n_1}}[2]_{q_{n_1}}}\right)x+\frac{1}{[n_1+1]_{q_{n_1}}^2[3]_{q_{n_1}}};$

(iv) $K_{n_1,n_2,p}((s-y)^2;q_{n_1},q_{n_2},x,y)=\left(\frac{q_{n_2}^2(1+q_{n_2}+4q_{n_2}^2)}{[2]_{q_{n_2}}[3]_{q_{n_2}}}\frac{[n_2+p]_{q_{n_2}}[n_2+p-1]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}^2}-\frac{[n_2+p]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}}\frac{4q_{n_2}}{[2]_{q_{n_2}}}+1\right)y^2+\left(\frac{q_{n_2}(3+5q_{n_2}+4q_{n_2}^2)}{[2]_{q_{n_2}}[3]_{q_{n_2}}}\frac{[n_2+p]_{q_{n_2}}}{[n_2+1]_{q_{n_2}}^2}-\frac{2}{[n_2+1]_{q_{n_2}}[2]_{q_{n_2}}}\right)y+\frac{1}{[n_2+1]_{q_{n_2}}^2[3]_{q_{n_2}}}.$

In what follows, let $\{q_{n_i}\}, i=1,2$ be sequences in $(0,1)$ such that $q_{n_i}\to 1$ and $\frac{1}{[n]_{q_i}}$ as $n_i\to\infty$.

Theorem 56. *For any $f\in C(I^2)$, we have*

$$\lim_{n_1,n_2\to\infty}\|K_{n_1,n_2,p}(f)-f\|_J=0.$$

Theorem 57. *Let $f\in C(I^2)$. Then for all $(x,y)\in J^2$, we have*

$$|K_{n_1,n_2,p}(f;q_{n_1},q_{n_2},x,y)-f(x,y)|\le 4\omega(f,\sqrt{\delta_{n_1}(x)},\sqrt{\delta_{n_2}(y)}),$$

where $\delta_{n_1}(x)=K_{n_1,p}((t-x)^2;q_{n_1},x)$ and $\delta_{n_2}(y)=K_{n_2,p}((s-y)^2;q_{n_2},y)$.

Theorem 58. *Let $f\in Lip_M(\alpha_1,\alpha_2)$. Then for all $(x,y)\in J^2$, we have*

$$|K_{n_1,n_2,p}(f;q_{n_1},q_{n_2},x,y)-f(x,y)|\le M\delta_{n_1}^{\alpha_1/2}(x)\delta_{n_2}^{\alpha_2/2}(y),$$

where $\delta_{n_1}(x)$ and $\delta_{n_2}(y)$ are defined as in Theorem 57.

Theorem 59. *Let $f \in C^1(I^2)$ and $(x, y) \in J^2$. Then, we have*

$$|K_{n_1,n_2,p}(f; q_{n_1}, q_{n_2}, x, y) - f(x, y)| \leq \|f'_x\|_I \sqrt{\delta_{n_1}(x)} + \|f'_y\|_I \sqrt{\delta_{n_2}(y)},$$

where $\delta_{n_1}(x)$ and $\delta_{n_2}(y)$ are defined as in Theorem 57.

Voronovskaja Type Theorem

Theorem 60. *Let $f \in C^2(I^2)$ and $\{q_n\}$ be a sequence in $(0, 1)$ such that $q_n \to 1$ and $q_n^n \to 0$ as $n \to \infty$. Then, we have*
$\lim_{n\to\infty} [n]_{q_n} (K_{n,n,p}(f; q_n, x, y) - f(x, y))$

$$= \frac{f'_x(x, y)}{2}(1 - x) + \frac{f'_y(x, y)}{2}(1 - y) + \frac{f''_{xx}(x, y)}{2}x(1 - x) + \frac{f''_{yy}(x, y)}{2}y(1 - y),$$

uniformly on J^2.

Acknowledgements The second author is thankful to the "Council of Scientific and Industrial Research" (Grant code: 09/143(0836)/2013-EMR-1) India for financial support to carry out the above research work.

References

1. Abel, U.: On the asymptotic approximation with bivariate operators of Bleimann, Butzer and Hahn. J. Approx. Theory **97**(1), 181–198 (1999)
2. Adell, J.A., de la Cal, J., Miguel, M.S.: On the property of monotonic convergence for multivariate Bernstein-type operators. J. Approx. Theory **80**(1), 132–137 (1995)
3. Agratini, O.: Bivariate positive operators in polynomial weighted spaces. Abstr. Appl. Anal. 2013, Art. ID 850760, 8 pp. (2013)
4. Agrawal, P.N., Goyal, M.: Generalized Baskakov Kantorovich operators FILOMAT. Accepted
5. Agrawal, P.N., Finta, Z., Sathish Kumar, A.: Bivariate q-Bernstein-Schurer-Kantorovich operators. Results Math. **67**(3–4), 365–380 (2015)
6. Altomare, F., Campiti, M.: Korovkin-Type Approximation Theory and Its Applications. Walter de Gruyter, Berlin (1994)
7. Andrews, G.E., Askey, R., Roy, R.: Special Functions. Cambridge University Press, Cambridge (1999)
8. Aral, A., Gupta, V., Agrawal, R.P.: Applications of q-Calculus in Operator Theory. Springer, New York (2013)
9. Atakut, Ç., İspir, N.: On Bernstein type rational functions of two variables. Math. Slovaca **54**(3), 291–301 (2004)
10. Bărbosu, D.: Aproximarea funcţiilor de mai multe variabile prin sume booleene de operatori liniari de tip interpolator. RISOPRINT, Cluj-Napoca (2002) (Romanian)
11. Bărbosu, D.: Bivariate operators of Schurer-Stancu type. An. Ştiinţ. Univ. Ovidius Constanţa Ser Mat. **11**(1), 1–8 (2003)
12. Baskakov, V.A.: An instance of a sequence of linear positive operators in the space of continuous functions. Dokl. Akad. Nauk. SSSR **113**, 249–251 (1957)

13. Becker, M., Kucharski, D., Nessel, R.J.: Global approximation theorems for the Szàsz-Mirakyan operators in exponential weight spaces. In: Linear Spaces and Approximation (Proceedings of the Conference, Oberwolfach, 1977), vol. 40, pp. 319–333. Birkhäuser, Basel (1978)
14. Butzer, P.L.: On two dimensional Bernstein polynomials. Can. J. Math. **5**, 107–113 (1953)
15. Butzer, P.L., Berens, H.: Semi-Groups of Operators and Approximation, xi+318 pp. Springer, New York (1967)
16. Căbulea, L.A., Aldea, M.: On some properties of Kantorovich bivariate operators. Proceedings of the International Conference on Theory and Applications of Mathematics and Informatics - ICTAMI. Alba Iulia (2003)
17. Doğru, O., Gupta, V.: Korovkin-type approximation properties of bivariate q-Meyer-König and Zeller operators. Calcolo **43**(1), 51–63 (2006)
18. Erencin, A.: Durrmeyer type modification of generalized Baskakov operators. Appl. Math. Comput. **218**(8), 4384–4390 (2011)
19. Ernst, T.: The history of q-calculus and a new method, U.U.D.M. Report 2000:16. Department of Mathematics, Uppsala University, Uppsala (2000)
20. Ersan, S.: Approximation properties of bivariate generalization of Bleimann, Butzer and Hahn operators Based on the q-integers. International Conference on Applied Mathematics, Cairo, Egypt, pp. 29–31 (2007)
21. Gadjiev, A.D.: Positive linear operators in weighted spaces of functions of several variables. Izv. Akad. Nauk Azerb. SSR, Fiz.-Tekhh. Mat. Nauk **1**(4), 32–37 (1980)
22. Gadjiev, A.D., Ghorbanalizadeh, A.M.: Approximation properties of a new type Bernstein-Stancu polynomials of one and two variables. Appl. Math. Comput. **216**(3), 890–901 (2010)
23. Gauchman, H.: Integral inequalities in q-calculus. Comput. Math. Appl. **47**(2–3), 281–300 (2004)
24. Goyal, M., Agrawal, P.N.: Bivariate generalized Baskakov-Kantorovich operators. (submitted)
25. Gurdek, M., Rempulska, L., Skorupka, M.: The Baskakov operators for functions of two variables. Collect. Math. **50**(3), 289–302 (1999)
26. Ipatov, A.F.: Estimation of error and order of approximation of functions of two variables by Bernstein polynomials, (in Russian). Uč Zap. Petrozavodsk Univ. **4**(4), 31–48 (1955)
27. Kac, V., Cheung, P.: Quantum Calculus, Universitext. Springer, New York (2002)
28. Kantorovich, L.V.: Sur certains développements suivant les polynômials de la formede S. Bernstein I, II. C. R. Acad. Sci. USSR **20**, 563–568, 595–600 (1930)
29. Khan, R.A.: A bivariate extension of Bleimann-Butzer-Hahn operator. Approx. Theory Appl. **18**(1), 90–100 (2002)
30. Kingsley, E.H.: Bernstein polynomials for functions of two variables of class $C^{(k)}$. Proc. Am. Math. Soc. **2**(1), 64–71 (1951)
31. Lupas, A.: A q-analogue of the Bernstein operator. In: Seminar on Numerical and Statistical Calculus, University of Cluj-Napoca, vol. 9, pp. 85–92 (1987)
32. Marinkovic, S., Rajkovic, P., Stankovic, M.: The inequalities for some types of q-integrals. Comput. Math. Appl. **56**(10), 2490–2498 (2008)
33. Martinez, F.L.: Some properties of two dimensional Bernstein polynomials. J. Approx. Theory **59**(3), 300–306 (1989)
34. Mihesan, V.: Uniform approximation with positive linear operators generated by generalized Baskakov method. Autom. Comput. Appl. Math. **7**(1), 34–37 (1998)
35. Örkcü, M.: Approximation properties of bivariate extension of q-Szász-Mirakjan-Kantorovich operators. J. Inequal. Appl. **324**, 1–10 (2013)
36. Örkcü, M.: q-Szász-Mirakjan-Kantorovich operators of Functions of two variables in polynomial weighted spaces. Hindawi Publ. Corp. Abstr. Appl. Anal. 2013 Art. ID 823803, 9 pp. (2013)
37. Phillips, G.M.: Bernstein polynomials based on the q-integers. The heritage of P. L. Chebyshev: a Festschrift in honour of the 70th-birthday of Prof. T. J. Rivlin.. Ann. Numer. Math. **4**, 511–518 (1997)

38. Pop, O.T.: The generalization of Voronovskaja's theorem for a class of bivariate operators. Stud. Univ. Babes-Bolyai Math. **53**(2), 85–107 (2008)
39. Pop, O.T., Fărcaş, M.D.: About the bivariate operators of Kantorovich type. Acta Math. Univ. Comen. **78**(1), 43–52 (2009)
40. Rempulska, L., Graczyk, S.: On certain class of Szàsz-Mirakyan operators in exponential weight spaces. Int. J. Pure Appl. Math. **60**(3), 259–267 (2010)
41. Rempulska, L., Graczyk, S.: On Generalized Szász-Mirakyan operators of functions of two variables. Math. Slovaca **62**(1), 87–98 (2012)
42. Stancu, D.D.: A method for obtaining polynomials of Bernstein type of two variables. Am. Math. Mon. **70**(3), 260–264 (1963)
43. Stancu, D.D.: Approximation of functions by a new class of linear polynomial operators. Rev. Roum. Math. Pures Appl. **13**(8), 1173–1194 (1968)
44. Stancu, F.: Aproximarea funcţiilor de două şi mai multe variabile cu ajutorul operatorilor liniari şi pozitivi. Ph.D. Thesis, Cluj-Napoca (1984) (Romanian)
45. Volkov, V.I.: On the convergence sequences of linear positive operators in the space of continuous functions of two variables. Dokl. Akad. Nauk SSSR **115**, 17–19 (1957) (Russian)
46. Wafi, A., Khatoon, S.: Approximation by generalized Baskakov operators for functions of one and two variables in exponential and polynomial weight spaces. Thai. J. Math. **2**, 53–66 (2004)
47. Wafi, A., Khatoon, S.: Convergence and Voronovskaja-type theorems for derivatives of generalized Baskakov operators. Cen. Eur. J. Math. **6**(2), 325–334 (2008)

Positive Green's Functions for Boundary Value Problems with Conformable Derivatives

Douglas R. Anderson

Abstract We use a newly introduced conformable derivative to formulate several boundary value problems with three or four conformable derivatives, including those with conjugate, right-focal, and Lidstone conditions. With the conformable differential equation and boundary conditions established, we find the corresponding Green's functions and prove their positivity under appropriate assumptions.

Keywords Positivity • Green's Function • Conformable Derivative

Introduction

The search for the existence of positive solutions and multiple positive solutions to nonlinear fractional boundary value problems has expanded greatly over the past decade; for some recent examples, please see [2, 5–11, 15–20]. In all of these works and the references cited therein, however, the definition of the fractional derivative used is either the Caputo or the Riemann–Liouville fractional derivative, involving an integral expression and the gamma function. Recently [1, 12, 14] a new local, limit-based definition of a conformable derivative has been formulated which is not a fractional derivative. In this paper, we use this conformable derivative of order α for $0 < \alpha \leq 1$ and $t \in [0, \infty)$ given by

$$D^{\alpha}f(t) := \lim_{\varepsilon \to 0} \frac{f(te^{\varepsilon t^{-\alpha}}) - f(t)}{\varepsilon}, \quad D^{\alpha}f(0) = \lim_{t \to 0+} D^{\alpha}f(t), \tag{1}$$

provided the limits exist; note that if f is fully differentiable at t, then

$$D^{\alpha}f(t) = t^{1-\alpha}f'(t), \tag{2}$$

where $f'(t) = \lim_{\varepsilon \to 0}[f(t+\varepsilon) - f(t)]/\varepsilon$. A function f is α-differentiable at a point $t \geq 0$ if the limit in (1) exists and is finite.

D.R. Anderson (✉)
Department of Mathematics, Concordia College, Moorhead, MN 56562, USA
e-mail: andersod@cord.edu

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_3

Using this new definition of a conformable derivative, we reformulate several common boundary value problems, including those with conjugate, right-focal, and Lidstone conditions. With the conformable differential equation and conformable boundary conditions established, we find the corresponding Green's functions and prove their positivity under appropriate assumptions. This work thus sets the stage for using fixed point theorems to prove the existence of positive and multiple positive solutions to nonlinear conformable problems based on the local conformable derivative (1) and these boundary value problems, as the kernel of the integral operator is often Green's function.

Recently, the conformable derivative (1) has been used successfully to explore two iterated conformable derivatives in comparison with a conformable derivative of order between 1 and 2 in [3]. There is also some potential application in quantum mechanics of this conformable derivative, as shown in [4]. In this work we will focus on boundary value problems with three or four iterated conformable derivatives.

Conformable Derivative Properties

Throughout this work we are considering the Katugampola derivative formulation of conformable calculus [12], specifically the operator D^α given in (1), where $0 < \alpha \leq 1$ and $t \geq 0$. This definition yields the following results [12, Theorem 2.3] for functions f and g that are α-differentiable at a point $t > 0$:

- $D^\alpha[af + bg] = aD^\alpha[f] + bD^\alpha[g]$ for all $a, b \in \mathbb{R}$;
- $D^\alpha t^n = nt^{n-\alpha}$ for all $n \in \mathbb{R}$;
- $D^\alpha c = 0$ for all constants $c \in \mathbb{R}$;
- $D^\alpha[fg] = fD^\alpha[g] + gD^\alpha[f]$;
- $D^\alpha[f/g] = \frac{gD^\alpha[f]-fD^\alpha[g]}{g^2}$;
- $D^\alpha[f \circ g](t) = f'(g(t))D^\alpha[g](t)$ for f differentiable at $g(t)$.

Three Iterated Conformable Derivatives

We begin the main discussion by considering the three iterated conformable derivative nonlinear right-focal problem

$$D^\gamma D^\beta D^\alpha x(t) = f(t, x(t)), \qquad 0 \leq t \leq 1, \tag{3}$$

$$x(0) = D^\alpha x(\tau) = D^\beta D^\alpha x(1) = 0, \tag{4}$$

where $\alpha, \beta, \gamma \in (0, 1]$, with boundary point τ satisfying $0 < \tau < 1$. One could impose further the conditions $\alpha + \beta \in (1, 2]$ and $\alpha + \beta + \gamma \in (2, 3]$ if one wishes to explore a conformable problem of order $(2, 3]$. Our approach to the existence of positive solutions would involve Green's function for this problem. Once the following three theorems are established, an interested reader could then apply a fixed point theorem to get positive solutions to (3), (4), although the details are omitted here.

Theorem 1 (Conformable Right-Focal Problem). *Let* $\alpha, \beta, \gamma \in (0, 1]$ *and* $0 < \tau < 1$. *The corresponding Green's function for the homogeneous problem*

$$D^{\gamma} D^{\beta} D^{\alpha} x(t) = 0$$

satisfying boundary conditions (4) *is given by*

$$G(t,s) = \begin{cases} s \in [0,\tau] & : \begin{cases} u(t,s) & : 0 \le t \le s \le 1 \\ x(0,s) & : 0 \le s \le t \le 1 \end{cases} \\ \\ s \in [\tau,1] & : \begin{cases} u(t,\tau) & : 0 \le t \le s \le 1 \\ u(t,\tau) + x(t,s) & : 0 \le s \le t \le 1 \end{cases} \end{cases} \tag{5}$$

where

$$u(t,s) = \frac{(\alpha+\beta)t^{\alpha}s^{\beta} - \alpha t^{\alpha+\beta}}{\alpha\beta(\alpha+\beta)} \tag{6}$$

and $x(\cdot,\cdot)$ *is the Cauchy function given by*

$$x(t,s) = \frac{\alpha t^{\alpha}\left(t^{\beta} - s^{\beta}\right) + \beta s^{\beta}\left(s^{\alpha} - t^{\alpha}\right)}{\alpha\beta(\alpha+\beta)}. \tag{7}$$

Proof. Let h be any continuous function. We will show that

$$x(t) = \int_0^1 G(t,s)h(s)s^{\gamma-1}ds,$$

for G given by (5), is a solution to the linear boundary value problem

$$D^{\gamma} D^{\beta} D^{\alpha} x(t) = h(t)$$

with boundary conditions (4).

First let $t \in [0, \tau]$. Then

$$x(t) = \int_0^t x(0,s)h(s)s^{\gamma-1}ds + \int_t^{\tau} u(t,s)h(s)s^{\gamma-1}ds + \int_{\tau}^1 u(t,\tau)h(s)s^{\gamma-1}ds; \tag{8}$$

clearly $x(0) = 0$ by (6). Differentiating (8), we have

$$\begin{aligned}
D^\alpha x(t) &= x(0,t)h(t)t^{\gamma-1}t^{1-\alpha} + \int_t^\tau \left(\frac{s^\beta - t^\beta}{\beta}\right) h(s)s^{\gamma-1}ds \\
&\quad -u(t,t)h(t)t^{\gamma-1}t^{1-\alpha} + \int_\tau^1 \left(\frac{\tau^\beta - t^\beta}{\beta}\right) h(s)s^{\gamma-1}ds \\
&= \frac{1}{\beta}\int_t^\tau \left(s^\beta - t^\beta\right) h(s)s^{\gamma-1}ds + \frac{1}{\beta}\int_\tau^1 \left(\tau^\beta - t^\beta\right) h(s)s^{\gamma-1}ds.
\end{aligned}$$

Clearly the second boundary condition $D^\alpha x(\tau) = 0$ is met. Differentiating again, we have

$$\begin{aligned}
D^\beta D^\alpha x(t) &= \frac{1}{\beta}\int_t^\tau \left(-\beta t^{\beta-1}t^{1-\beta}\right) h(s)s^{\gamma-1}ds + \frac{1}{\beta}\int_\tau^1 \left(-\beta t^{\beta-1}t^{1-\beta}\right) h(s)s^{\gamma-1}ds \\
&= \int_1^t h(s)s^{\gamma-1}ds.
\end{aligned}$$

It follows that $D^\beta D^\alpha x(1) = 0$ and $D^\gamma D^\beta D^\alpha x(t) = h(t)$, proving the claim in this case.

Next let $t \in [\tau, 1]$. Then

$$x(t) = \int_0^\tau x(0,s)h(s)s^{\gamma-1}ds + \int_\tau^t x(t,s)h(s)s^{\gamma-1}ds + u(t,\tau)\int_\tau^1 h(s)s^{\gamma-1}ds; \quad (9)$$

again $x(0) = 0$ by (6). Differentiating (9), we have

$$D^\alpha x(t) = \frac{1}{\beta}\int_\tau^t \left(t^\beta - s^\beta\right) h(s)s^{\gamma-1}ds + \left(\frac{\tau^\beta - t^\beta}{\beta}\right)\int_\tau^1 h(s)s^{\gamma-1}ds.$$

The second boundary condition $D^\alpha x(\tau) = 0$ is clearly met. Differentiating again, we have

$$D^\beta D^\alpha x(t) = \int_1^t h(s)s^{\gamma-1}ds;$$

as in the previous case, $D^\beta D^\alpha x(1) = 0$ and $D^\gamma D^\beta D^\alpha x(t) = h(t)$, proving the claim in this case as well. □

Theorem 2. *For $G(t,s)$ given in (5), we have that*

$$0 < G(t,s) \le G(\tau,s) \quad (10)$$

for $t \in (0,1]$ and $s \in (0,1]$, provided the condition $u(1,\tau) > 0$ is met for u in (6), that is to say the inequality

$$\tau > \left(\frac{\alpha}{\alpha+\beta}\right)^{1/\beta} \tag{11}$$

holds.

Proof. Referring to (5)–(7), we see that

$$u(s,s) = x(0,s), \quad x(0,\tau) = u(t,\tau) + x(t,\tau),$$

ensuring that $G(t,s)$ is a well-defined function; we also see that $G(t,s) = 0$ if $t = 0$ or $s = 0$. From (6) specifically, $u(0,s) = 0$ for all s and

$$\frac{d}{dt}u(t,s) = \frac{t^{\alpha-1}(s^\beta - t^\beta)}{\beta} \geq 0, \quad s \geq t.$$

Moreover, from (7) we have $x(s,s) = 0$ and

$$\frac{d}{dt}x(t,s) = \frac{t^{\alpha-1}(t^\beta - s^\beta)}{\beta} \geq 0, \quad t \geq s,$$

so that $G(t,s)$ is non-decreasing on $[0,\tau]$ and non-increasing on $[\tau,1]$. Thus (10) will hold if $G(1,\tau) > 0$ holds, which occurs if $u(1,\tau) > 0$. This is equivalent to (11), completing the proof. □

Theorem 3. *For all $t,s \in [0,1]$,*

$$g(t)G(\tau,s) \leq G(t,s) \leq G(\tau,s) \tag{12}$$

where

$$g(t) := \min\left\{\frac{t^\alpha}{\beta\tau^{\alpha+\beta}}\left[(\alpha+\beta)\tau^\beta - \alpha t^\beta\right], \frac{1-t}{1-\tau}\right\}. \tag{13}$$

Proof. From the preceding theorem, we have $G(t,s) \leq G(\tau,s)$ for all $t,s \in [0,1]$. For the lower bound, we proceed by cases on the branches of the Green's function (5), that is we use (6) and (7).

(*i*) $0 \leq t \leq s \leq \tau$: Here $G(t,s) = u(t,s)$, $G(\tau,s) = x(0,s) = \frac{1}{\alpha(\alpha+\beta)}s^{\alpha+\beta}$. For these t,s we have

$$\frac{u(t,s)}{x(0,s)} \geq \frac{u(t,\tau)}{x(0,\tau)} \geq \frac{t^\alpha}{\beta\tau^{\alpha+\beta}}\left[(\alpha+\beta)\tau^\beta - \alpha t^\beta\right]$$

which implies

$$G(t,s) \geq \frac{t^\alpha}{\beta\tau^{\alpha+\beta}}\left[(\alpha+\beta)\tau^\beta - \alpha t^\beta\right]G(\tau,s).$$

(ii) $0 \le t \le \tau \le s \le 1$: In this case $G(t,s) = u(t,\tau)$ and $G(\tau,s) = u(\tau,\tau)$, so again we have

$$G(t,s) \ge \frac{t^\alpha}{\beta\tau^{\alpha+\beta}}\left[(\alpha+\beta)\tau^\beta - \alpha t^\beta\right] G(\tau,s).$$

(iii) $0 \le s \le t \le \tau$ or $0 \le s \le \tau \le t \le 1$: Since $G(t,s) = G(\tau,s) = \frac{1}{\alpha(\alpha+\beta)}s^{\alpha+\beta}$, it follows that

$$G(t,s) = G(\tau,s).$$

(iv) $\tau \le t \le s \le 1$: As in (ii), $G(t,s) = u(t,\tau)$ and $G(\tau,s) = u(\tau,\tau)$. Define

$$\begin{aligned} w(t) &:= u(t,\tau) - \left(\frac{1-t}{1-\tau}\right) u(\tau,\tau) \qquad (14) \\ &= G(t,s) - \left(\frac{1-t}{1-\tau}\right) G(\tau,s). \end{aligned}$$

Now $w(\tau) = 0$, $w'(\tau) > 0$, and $w(1) = G(1,s) > 0$ by (11). Since w is concave down, $w(t) \ge 0$ on $[\tau, 1]$, hence

$$G(t,s) \ge \left(\frac{1-t}{1-\tau}\right) G(\tau,s).$$

(v) $\tau \le s \le t \le 1$: Note that $G(\tau,s) = u(\tau,\tau)$, while $G(t,s) = u(t,\tau) + x(t,s) \ge u(t,\tau)$; consequently, the employment of w as in (14) yields

$$G(t,s) \ge \left(\frac{1-t}{1-\tau}\right) G(\tau,s).$$

This completes the proof. □

Four Iterated Conformable Derivatives

In this section we consider four iterated conformable derivatives in the differential operator, with several types of boundary conditions. First, consider the nonlinear two-point cantilever beam eigenvalue problem

$$D^\delta D^\gamma D^\beta D^\alpha x(t) = \lambda a(t) f(x), \qquad 0 \le t \le 1, \qquad (15)$$

$$x(0) = D^\alpha x(0) = D^\beta D^\alpha x(1) = D^\gamma D^\beta D^\alpha x(1) = 0, \qquad (16)$$

where $\alpha, \beta, \gamma, \delta \in (0, 1]$, and with $\alpha + \beta \in (1, 2]$, $\alpha + \beta + \gamma \in (2, 3]$, and $\alpha + \beta + \gamma + \delta \in (3, 4]$ if one wishes to explore this problem as a conformable order between 3 and 4. The theme throughout this work has been to approach the existence of positive solutions to nonlinear equations such as (15) with boundary conditions (16) by involving Green's function for this problem. Please note: to obtain symmetry in Green's function below we must take $\gamma = \alpha$ in the conformable differential equation, but we will maintain the more general form (γ not necessarily equal to α) in the proofs to follow.

Theorem 4 (Conformable Cantilever Beam). *Let $\alpha, \beta, \gamma, \delta \in (0, 1]$. The corresponding Green's function for the homogeneous problem*

$$D^\delta D^\gamma D^\beta D^\alpha x(t) = 0 \tag{17}$$

satisfying boundary conditions (16) is given by

$$G(t,s) = \begin{cases} \dfrac{t^{\alpha+\beta}}{\gamma}\left[\dfrac{s^\gamma}{\beta(\alpha+\beta)} - \dfrac{t^\gamma}{(\beta+\gamma)(\alpha+\beta+\gamma)}\right] & : 0 \le t \le s \le 1, \\ \dfrac{s^{\beta+\gamma}}{\alpha}\left[\dfrac{t^\alpha}{\beta(\beta+\gamma)} - \dfrac{s^\alpha}{(\alpha+\beta)(\alpha+\beta+\gamma)}\right] & : 0 \le s \le t \le 1. \end{cases} \tag{18}$$

Proof. Let h be any continuous function. We will show that

$$x(t) = \int_0^1 G(t,s)h(s)s^{\delta-1}ds,$$

for G given by (18), is a solution to the linear boundary value problem

$$D^\delta D^\gamma D^\beta D^\alpha x(t) = h(t)$$

with boundary conditions (16).

For any $t \in [0, 1]$, using the branches of (18) we have

$$\begin{aligned} x(t) &= \int_0^t \frac{s^{\beta+\gamma}}{\alpha}\left[\frac{t^\alpha}{\beta(\beta+\gamma)} - \frac{s^\alpha}{(\alpha+\beta)(\alpha+\beta+\gamma)}\right]h(s)s^{\delta-1}ds \\ &\quad - \frac{t^{\alpha+\beta}}{\gamma}\int_1^t \left[\frac{s^\gamma}{\beta(\alpha+\beta)} - \frac{t^\gamma}{(\beta+\gamma)(\alpha+\beta+\gamma)}\right]h(s)s^{\delta-1}ds; \end{aligned}$$

clearly $x(0) = 0$. Taking the α-conformable derivative yields

$$\begin{aligned} D^\alpha x(t) &= \frac{1}{\beta(\beta+\gamma)}\int_0^t s^{\beta+\gamma+\delta-1}h(s)ds - \frac{t^\beta}{\beta\gamma}\int_1^t s^{\gamma+\delta-1}h(s)ds \\ &\quad + \frac{t^{\beta+\gamma}}{\gamma(\beta+\gamma)}\int_1^t s^{\delta-1}h(s)ds. \end{aligned}$$

It is easy to see that $D^\alpha x(0) = 0$. Taking the β-conformable derivative of the α-conformable derivative yields

$$D^\beta D^\alpha x(t) = \frac{1}{\gamma}\int_1^t (t^\gamma - s^\gamma)\, s^{\delta-1} h(s) ds,$$

so that $D^\beta D^\alpha x(1) = 0$. Next,

$$D^\gamma D^\beta D^\alpha x(t) = \int_1^t s^{\delta-1} h(s) ds,$$

from which we have $D^\gamma D^\beta D^\alpha x(1) = 0$, and

$$D^\delta D^\gamma D^\beta D^\alpha x(t) = h(t).$$

This finishes the proof. □

Theorem 5. *For all $t, s \in [0,1]$, the cantilever beam Green's function given by (18) satisfies*

$$0 \le G(t,s) \le G(1,s).$$

Proof. First observe from (18) that

$$\frac{d}{dt}G(t,s) = \begin{cases} \frac{t^{\alpha+\beta-1}}{\beta\gamma(\beta+\gamma)}\left[\gamma s^\gamma + \beta\left(s^\gamma - t^\gamma\right)\right] & : t \le s, \\ \frac{s^{\beta+\gamma} t^{\alpha-1}}{\beta(\beta+\gamma)} & : s \le t. \end{cases} \tag{19}$$

Now fix $s \in [0,1]$. By the first boundary condition $G(0,s) = 0$, and $\frac{d}{dt}G(t,s)$ as given above implies that $G(\cdot, s)$ is monotone increasing on $(0,1]$. In particular, $G(1,s) \ge G(t,s) \ge 0$ for all $t \in [0,1]$. □

The next Green's function we consider is for the conformable $(2,2)$-conjugate problem

$$D^\alpha D^\alpha D^\alpha D^\alpha x(t) = f(x), \qquad 0 \le t \le 1,$$
$$x(0) = 0 = D^\alpha x(0), \quad x(1) = 0 = D^\alpha x(1),$$

where $\alpha \in (0,1]$.

Theorem 6 (Conformable $(2,2)$-Conjugate). *Let $\alpha \in (0,1]$. The symmetric Green's function for the homogeneous problem*

$$D^\alpha D^\alpha D^\alpha D^\alpha x(t) = 0 \tag{20}$$

satisfying the conformable $(2,2)$*-conjugate boundary conditions*

$$x(0) = 0 = D^{\alpha}x(0), \quad x(1) = 0 = D^{\alpha}x(1) \tag{21}$$

is given by

$$G(t,s) = \begin{cases} u(t,s) & : 0 \le t \le s \le 1, \\ u(s,t) & : 0 \le s \le t \le 1, \end{cases} \tag{22}$$

where

$$u(t,s) = \frac{t^{2\alpha}\left(1-s^{\alpha}\right)^2}{6\alpha^3}\left[s^{\alpha} - t^{\alpha} + 2s^{\alpha}\left(1-t^{\alpha}\right)\right]. \tag{23}$$

Proof. In this proof we construct Green's function from scratch, modifying the classical approach, found, for example, in [13, Chap. 5]. Here Green's function takes the form of

$$G(t,s) = \begin{cases} u(t,s) & : t \le s, \\ u(t,s) + \frac{1}{3!}\left(\frac{t^{\alpha}-s^{\alpha}}{\alpha}\right)^3 & : s \le t, \end{cases}$$

where $u(t,s)$ satisfies (20) and the two boundary conditions in (21) set at $t = 0$. Thus

$$u(t,s) = a(s)t^{2\alpha} + b(s)t^{3\alpha};$$

the two boundary conditions at $t = 1$ are satisfied when

$$a(s) = \frac{s^{\alpha}\left(1-s^{\alpha}\right)^2}{2\alpha^3} \quad \text{and} \quad b(s) = \frac{\left(1-s^{\alpha}\right)^2\left(1+2s^{\alpha}\right)}{-6\alpha^3}.$$

Combining terms and simplifying leads to (22) and (23). □

Theorem 7. *For all* $t,s \in (0,1)$*, the* $(2,2)$*-conjugate Green's function given by* (22) *satisfies*

$$0 < G(t,s) \le \begin{cases} \frac{2}{3}\left(\frac{1-s^{\alpha}}{\alpha}\right)^3\left(\frac{s^{\alpha}}{3-2s^{\alpha}}\right)^2 & : 0 \le s \le 2^{-1/\alpha}, \\ \frac{2}{3}\left(\frac{s^{\alpha}}{\alpha}\right)^3\left(\frac{1-s^{\alpha}}{1+2s^{\alpha}}\right)^2 & : 2^{-1/\alpha} \le s \le 1. \end{cases}$$

Proof. From (23), it is evident that $u(0,s) = u(t,1) = 0$, and $u(t,s) \ge 0$ for $0 \le t \le s \le 1$. For fixed s, $u(t,s)$ has a local maximum when $t^{\alpha} = \frac{2s^{\alpha}}{1+2s^{\alpha}}$, and $u(s,t)$ has a local maximum when $t^{\alpha} = \frac{1}{3-2s^{\alpha}}$. After substitution and comparison, the result holds. □

Finally, we will end the present discussion by considering another common set of boundary conditions, namely the so-called Lidstone conditions. Unlike the argument used below, one could also approach this problem as the conjunction of two conjugate problems, whose Green's function is given in [3, (2.4)].

Theorem 8 (Conformable Lidstone). *Let $\alpha, \beta \in (0, 1]$. The symmetric Green's function for the homogeneous problem*

$$D^{\beta}D^{\alpha}D^{\beta}D^{\alpha}x(t) = 0 \tag{24}$$

satisfying the Lidstone-type boundary conditions

$$x(0) = 0 = D^{\beta}D^{\alpha}x(0), \quad x(1) = 0 = D^{\beta}D^{\alpha}x(1) \tag{25}$$

is given by

$$G(t,s) = \begin{cases} u(t,s) & : 0 \le t \le s \le 1, \\ u(s,t) & : 0 \le s \le t \le 1, \end{cases} \tag{26}$$

where

$$u(t,s) = \frac{t^{\alpha}}{\alpha\beta(\alpha+\beta)(2\alpha+\beta)} \left[2\alpha s^{\alpha}\left(1-s^{\beta}\right) - \beta\left(1-s^{\alpha}\right)\left(t^{\alpha+\beta}+s^{\alpha+\beta}\right)\right]. \tag{27}$$

Proof. Let $x(t,s)$ be the Cauchy function associated with (17), namely a function satisfying

$$x(s,s) = D^{\alpha}x(s,s) = D^{\beta}D^{\alpha}x(s,s) = 0, \quad D^{\gamma}D^{\beta}D^{\alpha}x(t,s) = 1.$$

(Note that we will use γ for now, and take $\gamma = \alpha$ for symmetry purposes at the conclusion.) Using (2) at each step, it is easy to verify that here the Cauchy function is given by

$$x(t,s) = \frac{1}{\gamma}\int_s^t \int_s^{\tau} \left(\xi^{\gamma} - s^{\gamma}\right)\xi^{\beta-1}\tau^{\alpha-1}d\xi d\tau, \tag{28}$$

and again Green's function takes the form of

$$G(t,s) = \begin{cases} u(t,s) & : t \le s, \\ u(t,s) + x(t,s) & : s \le t, \end{cases}$$

where $u(t,s)$ satisfies (17) and the two boundary conditions set at $t = 0$. Thus

$$u(t,s) = a(s) + b(s)t^{\alpha} + c(s)t^{\alpha+\beta} + d(s)t^{\alpha+\beta+\gamma};$$

the Lidstone boundary conditions force $a(s) = c(s) = 0$. The two boundary conditions at $t = 1$ are satisfied by $(u + x)$ for x given in (28). In particular, we use

$$D^{\beta}D^{\alpha}[u(t,s) + x(t,s)]|_{t\to 1} = 0$$

to solve for d, leading to

$$d(s) = \frac{s^{\gamma} - 1}{\gamma(\beta+\gamma)(\alpha+\beta+\gamma)},$$

and then $u(1,s) + x(1,s) = 0$ to solve for b, which yields

$$b(s) = \frac{s^{\alpha+\beta+\gamma}}{\alpha(\alpha+\beta)(\alpha+\beta+\gamma)} - \frac{s^{\beta+\gamma}}{\alpha\beta(\beta+\gamma)} + \frac{s^{\gamma}}{\gamma}\left(\frac{1}{\beta(\alpha+\beta)} - \frac{1}{(\beta+\gamma)(\alpha+\beta+\gamma)}\right).$$

Altogether we have

$$u(t,s) = s^{\gamma}t^{\alpha}\left(\frac{s^{\alpha+\beta}}{\alpha(\alpha+\beta)(\alpha+\beta+\gamma)} + \frac{\alpha+2\beta+\gamma}{\beta(\alpha+\beta)(\beta+\gamma)(\alpha+\beta+\gamma)} - \frac{s^{\beta}}{\alpha\beta(\beta+\gamma)}\right) + \frac{(s^{\gamma}-1)t^{\alpha+\beta+\gamma}}{\gamma(\beta+\gamma)(\alpha+\beta+\gamma)}.$$

To achieve symmetry and to satisfy (24), we take $\gamma = \alpha$ and arrive at (27). □

Theorem 9. *For all $t, s \in (0,1)$, the Lidstone Green's function given by (26) satisfies*

$$G(t,s) > 0.$$

Proof. Clearly $u(0,s) = u(t,1) = 0$ from (27), and

$$u(t,s) \geq \frac{2s^{\alpha}t^{\alpha}}{\alpha\beta(\alpha+\beta)(2\alpha+\beta)}\left[\alpha\left(1-s^{\beta}\right) - \beta\left(1-s^{\alpha}\right)s^{\beta}\right]$$

for $0 \leq t \leq s$. Define

$$k(s) := \alpha\left(1-s^{\beta}\right) - \beta\left(1-s^{\alpha}\right)s^{\beta} = \beta s^{\alpha+\beta} - (\alpha+\beta)s^{\beta} + \alpha.$$

Then $k(0) = \alpha > 0$, $k(1) = 0$, and

$$k'(s) = \beta(\alpha+\beta)s^{\beta-1}\left(s^{\alpha}-1\right) < 0.$$

Thus, k is strictly decreasing, so that $k(s) > 0$ for $s \in [0,1)$, forcing $u(t,s) \geq 0$ for $t, s \in [0,1]$. Therefore the result holds. □

References

1. Abu Hammad, M., Khalil, R.: Abel's formula and Wronskian for conformable fractional differential equations. Int. J. Differ. Equ. Appl. **13**(3), 177–183 (2014)
2. Ahmad, B., Nieto, J.J.: Existence results for a coupled system of nonlinear fractional differential equations with three-point boundary conditions. Comput. Math. Appl. **58**, 1838–1843 (2009)
3. Anderson, D.R., Avery, R.I.: Fractional-order boundary value problem with Sturm-Liouville boundary conditions. Electron. J. Differ. Equ. **2015**(29), 1–10 (2015)
4. Anderson, D.R., Ulness, D.J.: Properties of the Katugampola fractional derivative with potential application in quantum mechanics. J. Math. Phys. **56**, 063502 (2015). doi:10.1063/1.4922018
5. Bai, Z., Lü, H.: Positive solutions for boundary value problem of nonlinear fractional differential equation. J. Math. Anal. Appl. **311**(2), 495–505 (2005)
6. Cabada, A., Wang, G.: Positive solutions of nonlinear fractional differential equations with integral boundary value conditions. J. Math. Anal. Appl. **389**(1), 403–411 (2012)
7. Chai, G., Hu, S.: Existence of positive solutions for a fractional high-order three-point boundary value problem. Adv. Differ. Equ. **2014**, 90 (2014)
8. Chen, Y., Li, Y.: The existence of positive solutions for boundary value problem of nonlinear fractional differential equations. Abst. Appl. Anal. **2014**, 7 (2014). Article ID 681513
9. Chen, S., Liu, Y.: Solvability of boundary value problems for fractional order elastic beam equations. Adv. Differ. Equ. **2014**, 204 (2014)
10. Eloe, P.W., Neugebauer, J.T.: Existence and comparison of smallest eigenvalues for a fractional boundary-value problem. Electron. J. Differ. Equ. **2014**(43), 1–10 (2014)
11. Han, X., Gao, H.: Existence of positive solutions for eigenvalue problem of nonlinear fractional differential equations. Adv. Differ. Equ. **2012**, 66 (2012)
12. Katugampola, U.: A new fractional derivative with classical properties. arXiv:1410.6535v2
13. Kelley, W., Peterson, A.: The Theory of Differential Equations Classical and Qualitative. Pearson Prentice Hall, Upper Saddle River, NJ (2004)
14. Khalil, R., Al Horani, M., Yousef, A., Sababheh, M.: A new definition of fractional derivative. J. Comput. Appl. Math. **264**, 65–70 (2014)
15. Liu, X., Liu, Y.: Fractional differential equations with fractional non-separated boundary conditions. Electron. J. Differ. Equ. **2013**(25), 1–13 (2013)
16. Mahmudov, N.I., Unul, S.: Existence of solutions of $\alpha \in (2, 3]$ order fractional three-point boundary value problems with integral conditions. Abst. Appl. Anal. **2014**, 12 (2014). Article ID 198632
17. Sun, Y., Zhang, X.: Existence and nonexistence of positive solutions for fractional-order two-point boundary value problems. Adv. Differ. Equ. **2014**, 53 (2014)
18. Wu, J., Zhang, X.: Eigenvalue problem of nonlinear semipositone higher order fractional differential equations. Abst. Appl. Anal. **2012**, 14 (2012). Art. ID 740760
19. Zhang, X., Liu, L., Wiwatanapataphee, B., Wu, Y.: Positive solutions of eigenvalue problems for a class of fractional differential equations with derivatives. Abst. Appl. Anal. **2012**, 16 (2012). Art. ID 512127
20. Zhou, X., Wu, W.: Uniqueness and asymptotic behavior of positive solutions for a fractional-order integral boundary-value problem. Electron. J. Differ. Equ. **2013**(37), 1–10 (2013)

The Retraction-Displacement Condition in the Theory of Fixed Point Equation with a Convergent Iterative Algorithm

V. Berinde, A. Petruşel, I.A. Rus, and M.A. Şerban

Abstract Let (X, d) be a complete metric space and $f : X \to X$ be an operator with a nonempty fixed point set, i.e., $F_f := \{x \in X : x = f(x)\} \neq \emptyset$. We consider an iterative algorithm with the following properties:

(1) for each $x \in X$ there exists a convergent sequence $(x_n(x))$ such that $x_n(x) \to x^*(x) \in F_f$ as $n \to \infty$;
(2) if $x \in F_f$, then $x_n(x) = x$, for all $n \in \mathbb{N}$.

In this way, we get a retraction mapping $r : X \to F_f$, given by $r(x) = x^*(x)$. Notice that, in the case of Picard iteration, this retraction is the operator f^∞, see I.A. Rus (Picard operators and applications, Sci. Math. Jpn. 58(1):191–219, 2003). By definition, the operator f satisfies the retraction-displacement condition if there is an increasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ which is continuous at 0 and satisfies $\psi(0) = 0$, such that

$$d(x, r(x)) \leq \psi(d(x, f(x)), \text{ for all } x \in X.$$

In this paper, we study the fixed point equation $x = f(x)$ in terms of a retraction-displacement condition. Some examples, corresponding to Picard, Krasnoselskii, Mann and Halpern iterative algorithms, are given. Some new research directions and open questions are also presented.

Keywords Complete metric space • Fixed point • Weakly Picard operator • Comparison function • φ-Contraction • ψ-Picard operator • The retraction-displacement condition • Iterative algorithm • Krasnoselskii algorithm • Mann

---

V. Berinde
Technical University Cluj-Napoca, Baia-Mare North University Center, Victoriei Street, No. 76, Baia Mare, Romania
e-mail: vasile.berinde@math.utcluj.ro

A. Petruşel • I.A. Rus (✉) • M.A. Şerban
Department of Mathematics, Babeş-Bolyai University, Kogălniceanu Street, No. 1, Cluj-Napoca, Romania
e-mail: petrusel@math.ubbcluj.ro; iarus@math.ubbcluj.ro; mserban@math.ubbcluj.ro

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_4

algorithm • Halpern algorithm • Ulam-Hyers stability • Well-posedness of the fixed point equation • Ostrowski stability • Multi-valued operator

Introduction

In this paper we will consider the following two conditions involving a single-valued operator f from a metric space X to itself.

Definition 1. Let (X, d) be a metric space and $f : X \to X$ be an operator so that its fixed point set F_f is nonempty. Let $r : X \to F_f$ be a set retraction. Then, by definition, f satisfies the (ψ, r) retraction-displacement condition (ψ-condition in [11], (ψ, r)-operator in [41], ψ-weakly Picard operator in the case of Picard iterations in [37], the collage condition in [2]) if:

(i) $\psi : R_+ \to R_+$ is increasing, continuous at 0 and $\psi(0) = 0$;
(ii) $d(x, r(x)) \leq \psi(d(x, f(x))$, for every $x \in X$.

Remark 1. If $F_f = \{x^*\}$, then the (ψ, r) retraction-displacement condition takes the following form:

(i) $\psi : R_+ \to R_+$ is increasing, continuous at 0 and $\psi(0) = 0$;
(ii) $d(x, x^*) \leq \psi(d(x, f(x))$, for every $x \in X$.

We will call it the (x^*, ψ) retraction-displacement condition.

Definition 2. Let (X, d) be a metric space and $f : X \to X$ be an operator so that its fixed point set F_f is nonempty. Let

$$\begin{cases} x_{n+1} = f_n(x_n), n \in N \\ x_0 \in X \end{cases} \tag{1}$$

(where $f_n : X \to X$ is a sequence of single-valued operators, for $n \in N$) be an iterative algorithm such that

(i) $F_{f_n} = F_f$;
(ii) the sequence $(x_n)_{n \in N}$ converges to an element $x^*(x_0) \in F_f$ as $n \to +\infty$.

If we denote by $r : X \to F_f$ given by $r(x) := x^*(x)$ the retraction defined above, then, by definition, the above algorithm satisfies a retraction-displacement condition if the operator f satisfies a (ψ, r) retraction condition.

Notice that r is the limit operator of the iterative algorithm.

In this paper we study the fixed point equation $x = f(x)$ in terms of a retraction-displacement condition. Some examples, corresponding to Picard, Krasnoselskii, Mann and Halpern iterative algorithms, are given. Some new research directions are also presented.

Through the paper we will denote by R the set of real numbers, by N the set of natural numbers, by R_+ the set of positive numbers, by R_+^* the set of strict positive numbers and by $N^* := N\backslash\{0\}$. We also denote by R_+^m the space of all m-dimensional vectors with positive components.

Generalized Contractions: Examples

We will start this section by presenting some notions and results concerning generalized contractions, which will be used in our main section.

A first generalization of the Banach contraction principle involves the concept of comparison function.

Definition 3 ([36, 39]). A function $\varphi : R_+ \to R_+$ is called a comparison function if it satisfies:

$(i)_\varphi$ φ is increasing;
$(ii)_\varphi$ the sequence $(\varphi^n(t))_{n\in N}$ converges to 0 as $n \to \infty$, for all $t \in R_+$.
If the condition $(ii)_\varphi$ is replaced by the condition:
$(iii)_\varphi$ $\sum_{k=0}^{\infty} \varphi^k(t) < \infty$, for any $t > 0$,
then φ is called a strong comparison function.
Moreover, if the condition $(iii)_\varphi$ is replaced by the condition:
$(iv)_\varphi$ $t - \varphi(t) \to +\infty$, as $t \to +\infty$,
then φ is said to be a strict comparison function.

As a consequence of the above definition, we have the following lemmas.

Lemma 1 ([36, 39]). *If $\varphi : R_+ \to R_+$ is a comparison function, then the following hold:*

(i) $\varphi(t) < t$, *for any* $t > 0$;
(ii) $\varphi(0) = 0$;
(iii) φ *is continuous at* 0.

Lemma 2 ([3, 24, 39]). *If $\varphi : R_+ \to R_+$ is a strong comparison function, then the following hold:*

(i) φ *is a comparison function;*
(ii) the function $s : R_+ \to R_+$, *defined by*

$$s(t) = \sum_{k=0}^{\infty} \varphi^k(t), \ t \in R_+, \tag{2}$$

is increasing and continuous at 0;

(iii) there exist $k_0 \in N$, $a \in (0,1)$ *and a convergent series of nonnegative terms* $\sum_{k=1}^{\infty} v_k$ *such that*

$$\varphi^{k+1}(t) \leq a\varphi^k(t) + v_k, \text{ for } k \geq k_0 \text{ and any } t \in R_+.$$

Remark 2. Some authors use the notion of (c)-comparison function defined by the statements (i) and (iii) in Lemma 2. Actually, the concept of (c)-comparison function coincides with that of strong comparison function.

Example 1. (1) $\varphi : R_+ \to R_+, \varphi(t) = at$, where $a \in [0,1[$, is a strong comparison function and a strict comparison function. In this case, f is called a contraction with constant $a \in [0,1[$.

(2) $\varphi : R_+ \to R_+, \varphi(t) = \dfrac{t}{1+t}$ is a strict comparison function, but is not a strong comparison function.

(3) $\varphi : R_+ \to R_+$ defined by

$$\varphi(t) := \begin{cases} \frac{1}{2}t, & t \in [0,1] \\ t - \frac{1}{2}, & t > 1 \end{cases}$$

is a strong comparison function.

(4) $\varphi : R_+ \to R_+, \varphi(t) = at + \frac{1}{2}[t]$, where $a \in \left]0, \frac{1}{2}\right[$ is a strong comparison function.

For other considerations on comparison functions, see [3, 17, 18, 36, 39], and the references therein.

The Retraction-Displacement Condition in the Theory of Weakly Picard Operators

The first part of the following result is known as Matkowski's Theorem (see [20]), while the second part belongs to Rus [36].

Theorem 1. *Let* (X,d) *be a complete metric space and* $f : X \to X$ *be a* φ*-contraction, i.e.,* $\varphi : R_+ \to R_+$ *is a comparison function and*

$$d(f(x), f(y)) \leq \varphi(d(x,y)) \text{ for all } x, y \in X.$$

Then f *is a Picard operator, i.e.,* f *has a unique fixed point* $x^* \in X$ *and* $\lim_{n\to+\infty} f^n(x) = x^*$, *for all* $x \in X$. *Moreover, if* $\varphi : R_+ \to R_+$ *is a strict comparison function, then* f *is a* ψ*-Picard operator, i.e.,* f *is a Picard operator and*

$$d(x,x^*) \leq \psi_\varphi(d(x,f(x))),\ \textit{for all}\ x \in X,$$

where $\psi_\varphi : R_+ \to R_+$ *is given by* $\psi_\varphi(t) := \sup\{s | s - \varphi(s) \leq t\}$.

The second conclusion of the above theorem gives an answer to the following very general problem.

Problem 1. Let (X,d) be a metric space and $f : X \to X$. Which generalized contractions f are ψ-Picard operators? Which generalized contractions f satisfy a (ψ, r) retraction-displacement condition?

A general result concerning the above problem is the following.

Theorem 2. *Let* (X,d) *be a metric space,* $f : X \to X$ *be an operator,* $\varphi : R_+ \to R_+$ *a strict comparison function and* $\theta : R_+ \to R_+$ *be an increasing function, continuous at* 0 *with* $\theta(0) = 0$. *We suppose that:*

(i) $F_f = \{x^*\}$*;*
(ii) $d(f(x),x^*) \leq \varphi(d(x,x^*)) + \theta(d(x,f(x)))$*, for all* $x \in X$.

Then:

$$d(x,x^*) \leq \psi_\varphi\left(d\left(x,f(x)\right) + \theta\left(d(x,f(x))\right)\right),\ \textit{for all}\ x \in X, \tag{3}$$

i.e., f is a ψ*-Picard operator with* $\psi = \psi_\varphi \circ \left(1_{R_+} + \theta\right)$

Proof. We have

$$\begin{aligned} d(x,x^*) &\leq d\left(x,f(x)\right) + d\left(f(x),x^*\right) \leq \\ &\leq d\left(x,f(x)\right) + \varphi\left(d(x,x^*)\right) + \theta\left(d(x,f(x))\right). \end{aligned}$$

From the definition of ψ_φ we get the conclusion. We remark that the function $\psi = \psi_\varphi \circ \left(1_{R_+} + \theta\right)$ is increasing, continuous at 0 and $\psi(0) = 0$.

A class of ψ-Picard operators (with a particular $\psi(t) := \alpha t$, for some $\alpha \in [0,1[$ and with a particular $\theta(t) = Lt$, for some $L \geq 0$) is given by the following consequence.

Corollary 1. *Let* (X,d) *be a metric space and* $f : X \to X$ *be an operator. We suppose:*

(a) $F_f = \{x^*\}$*;*
(b) (see [32]) there exists $\alpha \in [0,1[$ *and* $L \geq 0$ *such that*

$$d(f(x),x^*) \leq \alpha d(x,x^*) + Ld(x,f(x)),\ \textit{for all}\ x \in X.$$

Then:

$$d(x,x^*) \leq \frac{1+L}{1-\alpha} d(x,f(x))),\ \textit{for all}\ x \in X.$$

We will present now some examples of generalized contractions which satisfy the assumptions (a) and (b) in the above theorems. For example, using the Hardy and Rogers type condition, we can prove the following result.

Theorem 3. *Let (X, d) be a complete metric space and $f : X \to X$ be an operator. We suppose there exist $a, b, c \in R_+$ with $a + 2b + 2c < 1$ such that, for all $x, y \in X$, we have*

$$d(f(x), f(y)) \leq ad(x, y) + b(d(x, f(x)) + d(y, f(y))) + c(d(x, f(y)) + d(y, f(x))).$$

Then:

(i) $F_f = \{x^*\}$*;*
(ii) $d(x, x^*) \leq \frac{1+b-c}{1-a-2c} d(x, f(x)))$, *for all* $x \in X$.

Proof. Let $x_0 \in X$ and $x_n := f^n(x_0)$, $n \in N^*$. Then, by Hardy–Rogers' fixed point theorem we get $F_f = \{x^*\}$. Now, we also have

$$d(f(x), x^*) = d(f(x), f(x^*)) \leq$$
$$ad(x, x^*) + b(d(x, f(x) + d(x^*, f(x^*)) + c(d(x, f(x^*)) + d(x^*, f(x))).$$

Thus

$$d(f(x), x^*) \leq \frac{a+c}{1-c} d(x, x^*) + \frac{b}{1-c} d(x, f(x), \text{ for all } x \in X.$$

The conclusion follows now from Theorem 2.

We will discuss now the case of so-called Suzuki type contractions.

Theorem 4. *Let (X, d) be a metric space, $\theta : R_+ \to R_+$ such that $\theta(0) = 0$, $\varphi : R_+ \to R_+$ be a strict comparison function and $f : X \to X$ be an operator. We suppose that*

(i) $F_f = \{x^*\}$*;*
(ii) $x, y \in X$, $\theta(d(f(x), x)) \leq d(x, y) \Longrightarrow d(f(x), f(y)) \leq \varphi(d(x, y))$.

Then:

$$d(x, x^*) \leq \psi_\varphi(d(x, f(x))), \text{ for all } x \in X. \tag{4}$$

Proof. We take in (ii) $x = x^*$ and we apply Theorem 2.

Remark 3. It is worth to notice that there exists operators which satisfy all the assumptions in Theorem 2, but which are not Picard operators. For example, $f : R \to R, f(x) := 2x$. In this case, $F_f := \{0\}$ and

$$|f(x)| \leq \frac{1}{2}|x| + 2|x - f(x)|, \text{ for all } x \in X,$$

but f is not a Picard operator.

In the next part of this section, the case of the cyclic φ-contractions is discussed.

One of the most important generalizations of the Banach Contraction Principle was given in 2003 by Kirk, Srinivasan and Veeramani, using the concept of cyclic operator. More precisely, they proved in [19] the following result.

Theorem 5 ([19, Theorem 2.4]). *Let $\{A_i\}_{i=1}^m$ be nonempty subsets of a complete metric space and suppose $f : \bigcup_{i=1}^m A_i \to \bigcup_{i=1}^m A_i$ satisfies the following conditions:*

(1) $f(A_i) \subseteq A_{i+1}$ for $1 \le i \le m$, where $A_{m+1} = A_1$;
(2) $d(f(x), f(y)) \le \varphi(d(x, y))$, for all $x \in A_i$, and all $y \in A_{i+1}$, for $1 \le i \le m$, where the mapping $\varphi : R_+ \to R_+$ is upper semi-continuous from the right and satisfies the condition $0 \le \varphi(t) < t$ for $t > 0$.

Then f has a unique fixed point.

An extension of this result was given by Păcurar and Rus in [24]. See also [38].

Theorem 6 ([24, Theorem 2.1]). *Let $\{A_i\}_{i=1}^m$ be nonempty subsets of a complete metric space and suppose $f : \bigcup_{i=1}^m A_i \to \bigcup_{i=1}^m A_i$ satisfies the following conditions:*

(1) $f(A_i) \subseteq A_{i+1}$ for $1 \le i \le m$, where $A_{m+1} = A_1$;
(2) there exists a comparison function $\varphi : R_+ \to R_+$ such that

$$d(f(x), f(y)) \le \varphi(d(x, y)), \text{ for all } x \in A_i \text{ and all } y \in A_{i+1},\ 1 \le i \le m.$$

Then:

(a) f has a unique fixed point $x^* \in \bigcap_{i=1}^m A_i$ and, for each $x \in A := \bigcup_{i=1}^m A_i$ the sequence $x_n := (f^n(x))_{n\in N}$ converges to x^* as $n \to +\infty$;
(b) the following estimates take place:

$$d(x_n, x^*) \le s(\varphi^n(d(x_0, f(x_0))) \text{ and } d(x_n, x^*) \le s(\varphi(d(x_n, f(x_n))), \text{ for all } n \in N^*;$$

(c) the following relation holds:

$$d(x, x^*) \le s(d(x, f(x))), \text{ for all } x \in A,$$

where $s : R_+ \to R_+$ is defined by $s(t) := \sum_{k=0}^{\infty} \varphi^k(t)$.

Our next result is an extension of the previous result.

Theorem 7. *Let $\{A_i\}_{i=1}^m$ be nonempty subsets of a complete metric space and suppose $f : \bigcup_{i=1}^m A_i \to \bigcup_{i=1}^m A_i$ satisfies the following conditions:*

(1) $f(A_i) \subseteq A_{i+1}$ for $1 \le i \le m$, where $A_{m+1} = A_1$;
(2) there exists a strict comparison function $\varphi : R_+ \to R_+$ such that

$$d(f(x), f(y)) \le \varphi(d(x, y)), \text{ for all } x \in A_i \text{ and all } y \in A_{i+1},\ 1 \le i \le m.$$

Then:

(a) f is a Picard operator, i.e., f has a unique fixed point and, for each $x \in A := \bigcup_{i=1}^m A_i$ the sequence $x_n := (f^n(x))_{n\in N}$ converges to x^* as $n \to +\infty$;
(b) $x^* \in \bigcap_{i=1}^m A_i$;
(c) the following relation holds:

$$d(x, x^*) \le \psi_\varphi(d(x, f(x))), \text{ for all } x \in A,$$

where $\psi_\varphi : R_+ \to R_+$ is given by $\psi_\varphi(t) := \sup\{s | s - \varphi(s) \le t\}$.

Proof. The conclusions (a) and (b) follow by Theorem 6. The conclusion (c) follows (since $x^* \in \bigcap_{i=1}^m A_i$) by the following relations:

$$d(x, x^*) \le d(x, f(x)) + d(f(x), f(x^*)) \le d(x, f(x)) + \varphi(d(x, x^*)).$$

Another general open problem is the following one.

Problem 2. Let (X, d) be a metric space and $f : X \to X$ be such that there exists $n_0 \in N^*$ such that f^{n_0} is a generalized contraction. Under which conditions f is a ψ-Picard operator (or a ψ-weakly Picard operator)? Under which conditions f satisfies the ψ-condition with respect to a set retraction r?

For the above problem, we have the following result (see [22, 35, 46, 52], etc.).

Theorem 8. *Let (X, d) be a complete metric space and $f : X \to X$ be such that there exists $n_0 \in N^*$ such that f^{n_0} is a contraction with constant $a \in [0, 1[$. We also suppose that f is L-Lipschitz with constant $L \ge 1$. Then, the following conclusions hold:*

(i) f is a Picard operator and $F_f = \{x^*\}$;
(ii) (a) if $L > 1$, then

$$d(x, x^*) \le \frac{L^{n_0} - 1}{(1 - a)(L - 1)} d(x, f(x)), \text{ for all } x \in X.$$

(b) if $L = 1$*, then*

$$d(x, x^*) \leq \frac{n_0}{1-a} d(x, f(x)), \text{ for all } x \in X.$$

Proof. We only need to prove (ii). Since f^{n_0} is a contraction with constant $a \in [0, 1[$, we get that f is a $\frac{1}{1-a}$-Picard operator. Thus, for all $x \in X$, we have

$$d(x, x^*) \leq \frac{1}{1-a} d(x, f^{n_0}(x)) \leq$$

$$\frac{1}{1-a}(d(x, f(x)) + d(f(x), f^2(x)) + d(f^{n_0-1}(x), f^{n_0}(x)))$$

$$\leq \frac{L^{n_0} - 1}{(1-a)(L-1)} d(x, f(x)).$$

In a similar way one obtain (b).

A third general problem is the following.

Problem 3. Let (X, d) be a metric space and $f : X \to X$ be such that there exists $n_0 \in N^*$ such that $f|_{f^{n_0}(X)} : f^{n_0}(X) \to f^{n_0}(X)$ is a generalized contraction. Under which conditions f is a ψ-Picard operator (or a ψ-weakly Picard operator)? Under which conditions f satisfies the ψ-condition with respect to a set retraction r?

A first answer to the above problem is the following theorem.

Theorem 9. *Let* (X, d) *be a complete metric space and* $f : X \to X$ *be such that there exists* $n_0 \in N^*$ *such that* $f|_{f^{n_0}(X)} : f^{n_0}(X) \to f^{n_0}(X)$ *is a contraction with constant* $a \in [0, 1[$. *Suppose that* f *is Lipschitz with constant* $L \geq 1$. *Then, the following conclusions hold:*

(i) f *is a Picard operator and* $F_f = \{x^*\}$*;*
(ii) (a) if $L > 1$*, then*

$$d(x, x^*) \leq \left(\frac{L^{n_0} - 1}{L-1} + \frac{L^{n_0}}{1-a}\right) \cdot d(x, f(x)), \text{ for all } x \in X.$$

(b) if $L = 1$*, then*

$$d(x, x^*) \leq \left(n_0 + \frac{1}{1-a}\right) \cdot d(x, f(x)), \text{ for all } x \in X.$$

Proof. (a) Since $f|_{f^{n_0}(X)}$ is a contraction with constant a, we obtain that $f|_{\overline{f^{n_0}(X)}}$ is a contraction with constant a, too. Thus, $F_f := \{x^*\}$. Moreover, since f^{n_0} is a $\frac{1}{1-a}$-Picard operator, for all $x \in X$, we have

$$d(f^{n_0}(x), x^*) \leq \frac{1}{1-a} d(f^{n_0}(x), f(f^{n_0}(x))) \leq \frac{L^{n_0}}{1-a} d(x, f(x)).$$

On the other hand, for all $x \in X$, we can write:

$$d(x, x^*) \leq d(x, f^{n_0}(x)) + d(f^{n_0}(x), x^*) \leq$$
$$(d(x, f(x)) + d(f(x), f^2(x)) + d(f^{n_0-1}(x), f^{n_0}(x))) + \frac{L^{n_0}}{1-a} d(x, f(x)) \leq \cdots \leq$$
$$\frac{L^{n_0}-1}{L-1} d(x, f(x)) + \frac{L^{n_0}}{1-a} d(x, f(x)).$$

Notice now that (b) follows by a similar approach.

Notice now that in Theorem 2 and Corollary 1 the uniqueness of the fixed point can be deduced by the imposed condition (b). In the absence of the uniqueness assumption for the fixed point, we can prove the following extension of Corollary 1.

Corollary 2. *Let (X, d) be a metric space and $f : X \to X$ be an operator. We suppose:*

(a) for each $x \in X$ there exists a sequence $(x_n(x))_{n \in N}$ and there exists $x^*(x) \in F_f$ with $\lim_{n \to +\infty} x_n(x) = x^*(x)$. In particular, if $x \in F_f$, then $x_n(x) = x$, for all $n \in N$. Thus, $r : X \to F_f \ \ x \mapsto x^*(x)$ is a set retraction.

(b) there exists $\alpha \in [0, 1[$ and $L \geq 0$ such that

$$d(f(x), r(x)) \leq \alpha d(x, r(x)) + L d(x, f(x)), \text{ for all } x \in X.$$

Then

$$d(x, r(x)) \leq \frac{1+L}{1-\alpha} d(x, f(x)), \text{ for all } x \in X.$$

Proof. Notice that, for every $x \in X$, we have

$$d(x, r(x)) \leq d(x, f(x)) + d(f(x), r(x)) \leq$$
$$\leq d(x, f(x)) + \alpha d(x, r(x)) + L d(x, f(x)).$$

As an illustrative example, we have the following result for graphic contractions.

Theorem 10. *Let (X, d) be a complete metric space and $f : X \to X$ be an operator. We suppose:*

(a) there exists $a \in [0, 1[$ such that

$$d(f(x), f^2(x)) \leq a d(x, f(x)), \text{ for all } x \in X.$$

(b) f has closed graph.
Then:

(i) for each $x \in X$, the sequence $x_n := f^n(x)$, $n \in N^*$ converges to an element $x^*(x) = f^\infty(x) \in F_f$;
(ii)

$$d(x, f^\infty(x)) \leq \frac{1}{1-a} d(x, f(x)) \text{ for all } x \in X.$$

Proof. (i) is the well-known Graphic Contraction Principle. For the sake of completeness, we recall here the proof. Let $x_0 \in X$ and $x_n := f^n(x_0)$, $n \in N$. Then, by (a), the sequence (x_n) is Cauchy and, by the completeness of the space (X, d), there exists $x^*(x_0) \in X$ such that $\lim_{n \to +\infty} x_n = x^*(x_0)$. By (b), we get that $x^*(x_0) \in F_f$. Now, for each $x \in X$, we also have

$$d(f(x), f^\infty(x)) \leq d(f(x), f^2(x)) + d(f^2(x), f^3(x)) + \cdots + d(f^n(x), f^\infty(x)) \leq$$
$$ad(x, f(x)) + a^2 d(x, f(x)) + \cdots + a^{n-1} d(x, f(x)) + d(f^n(x), f^\infty(x)) =$$
$$= \frac{a(1-a^{n-1})}{1-a} d(x, f(x)) + d(f^n(x), x^*(x)).$$

Letting $n \to +\infty$, we get that

$$d(f(x), f^\infty(x)) \leq \frac{a}{1-a} d(x, f(x)), \text{ for all } x \in X.$$

Now we can conclude

$$d(x, f^\infty(x)) \leq d(x, f(x)) + d(f(x), f^\infty(x)) \leq$$
$$\leq d(x, f(x)) + \frac{a}{1-a} d(x, f(x)) = \frac{1}{1-a} d(x, f(x)).$$

In the case of Caristi–Browder operators (see [18, 39]) we have a similar result.

Theorem 11. *Let (X, d) be a complete metric space, $f : X \to X$ be an operator and $\varphi : X \to R_+$ be a given function. We suppose:*

(a) $d(x, f(x)) \leq \varphi(x) - \varphi(f(x))$, for all $x \in X$;
(b) f has closed graph.
Then:

(i) $F_f \neq \emptyset$;
(ii) for each $x \in X$, the sequence $x_n := f^n(x)$, $n \in N^*$ converges to an element $f^\infty(x) \in F_f$;
(ii) if, additionally, there is $\alpha \in R_+^*$ such that $\varphi(x) \leq \alpha d(x, f(x))$, then

$$d(x,f^\infty(x)) \leq \alpha d(x,f(x)),\ \textit{for all } x \in X.$$

Proof. For (i) and (ii), let us consider $x \in X$. From (a) it follows

$$\sum_{k=0}^{n} d(f^k(x),f^{k+1}(x)) \leq \varphi(x) - \varphi(f^{n+1}(x)) \leq \varphi(x).$$

This implies that $(f^n(x))_{n\in N}$ is a convergent sequence. Let us denote by $f^\infty(x) \in X$ its limit. From (b) we have that $f^\infty(x) \in F_f$.

For (iii), notice that for each $x \in X$, we have

$$d(x,f^{n+1}(x)) \leq \sum_{k=0}^{n} d(f^k(x),f^{k+1}(x)) \leq \varphi(x) \leq \alpha d(x,f(x)).$$

Thus, $d(x,f^\infty(x)) \leq \alpha d(x,f(x))$, for all $x \in X$.

Remark 4. For other considerations on weakly Picard operator theory, see [6, 13, 27, 37, 40, 45], etc.

Remark 5. For generalized contractions conditions and related results, see [3, 17, 18, 23, 28, 30, 31, 34, 36, 39, 46, 52], etc.

The Retraction-Displacement Condition in the Case of Other Iterative Algorithms

Let $(X, +, R, \|\cdot\|)$ be a Banach space, $Y \subset X$ a nonempty convex subset, $f : Y \to Y$ an operator, $0 < \lambda < 1$ and $\Lambda := (\lambda_n)_{n\in N}$ with $0 < \lambda_n < 1$, $n \in N$.

Krasnoselskii Algorithm

By the Krasnoselskii perturbation of f we understand the operator $f_\lambda : Y \to Y$ defined by

$$f_\lambda(x) = (1-\lambda)x + \lambda f(x),\ x \in Y.$$

For this perturbation of f we have (see [3, 4, 10, 42, 50, 51], etc.):

Theorem 12. *Let f_λ be defined as above. Then:*

(i) $F_{f_\lambda} = F_f$. In general $F_{f_\lambda^n} \neq F_{f^n}$, $n \geq 2$.
(ii) If f is l-Lipschitz, then f_λ is l-Lipschitz.
(iii) If f is a φ-contraction, then f_λ is a φ_λ-contraction.

(iv) *If in addition Y is bounded and closed and f is nonexpansive, then f_λ is asymptotically regular.*

(v) *If f satisfies $a(r,\psi)$ retraction-displacement condition, then f_λ satisfies the (r,θ) retraction-displacement condition with $\theta(t)=\psi\left(\frac{1}{\lambda}t\right)$, $t\in R_+$.*

(vi) *If X is an ordered Banach space, then f increasing implies f_λ increasing.*

The following problem arises:

Problem 4. If f_λ is WPO, in which conditions on f, f_λ satisfies the (f_λ^∞,ψ) retraction-displacement condition?

Some results for this problem were given in section "The Retraction-Displacement Condition in the Case of Other Iterative Algorithms" when f is a generalized contraction (see (v) in Theorem 12).

Remark 6. For the condition in which f_λ is WPO see [3, 10, 18, 50, 51], etc. For example, the following result is well known, see [3].

Definition 4. Let H be a Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. An operator $f:H\to H$ is said to be a *generalized pseudo-contraction* if there exists a constant $r>0$ such that, for all x,y in H,

$$\|f(x)-f(y)\|^2\le r^2\|x-y\|^2+\|f(x)-f(y)-r(x-y)\|^2. \tag{5}$$

Remark 7. Condition 5 is equivalent to

$$\langle f(x)-f(y),\ x-y\rangle\ \le\ r\,\|x-y\|^2,\quad \text{for all}\quad x,y\in H, \tag{6}$$

or to

$$\langle (I-f)\,x-(I-f)\,y\rangle\ \ge\ (1-r)\ \|x-y\|^2. \tag{7}$$

Remark 8. Note that any Lipschitzian operator f, that is, any operator for which there exists $s>0$ such that

$$\|f(x)-f(y)\|\le s\cdot\|x-y\|,\quad x,y\in H, \tag{8}$$

is also a generalized pseudo-contractive operator, with $r=s$.

Theorem 13. *Let K be a nonempty closed convex subset of a real Hilbert space and let $f:K\to K$ be a generalized pseudocontractive and Lipschitzian operator with the corresponding constants r and s, respectively, such that*

$$0<r<1\quad \textit{and}\quad r\le s. \tag{9}$$

Then

(i) *f has a unique fixed point p;*

(ii) For each $x_0 \in K$, the Krasnoselskii iteration $\{x_n\}_{n=0}^{\infty}$, given by

$$x_{n+1} = (1-\lambda)x_n + \lambda f(x_n)\,, \quad n = 0, 1, 2, \ldots, \tag{10}$$

converges (strongly) to p, for all $\lambda \in (0, 1)$ satisfying

$$0 < \lambda < 2(1-r)/(1-2r+s^2). \tag{11}$$

(iii) The following retraction-displacement condition holds:

$$d(x, f_\lambda^\infty(x)) \le \frac{\lambda}{1-\theta} d(x, f(x)),\ \forall x \in K,$$

where

$$\theta = \left((1-\lambda)^2 + 2\lambda(1-\lambda)\, r + \lambda^2 s^2\right)^{1/2}. \tag{12}$$

Proof. Denote

$$f_\lambda(x) = (1-\lambda)x + \lambda \cdot f(x)\,, \quad x \in K, \tag{13}$$

for all $\lambda \in (0, 1)$.

Since f is generalized pseudo-contractive and Lipschitzian, we have

$$\|f_\lambda(x) - f_\lambda(y)\|^2 = \|\,(1-\lambda)\,x + \lambda f(x) - (1-\lambda)\,y - \lambda f(y)\|^2 = \tag{14}$$

$$\|\,(1-\lambda)(x-y) + \lambda(f(x) - f(y)\|^2 = \tag{15}$$

$$(1-\lambda)^2 \cdot \|\,x-y\|^2 + 2\lambda(1-\lambda)\cdot\langle f(x) - f(y),\, x-y\rangle + \lambda^2 \cdot \|f(x) - f(y)\|^2 \le \tag{16}$$

$$\left((1-\lambda)^2 + 2\lambda(1-\lambda)r + \lambda^2 s^2\right)\cdot\|\,x-y\|^2\,, \tag{17}$$

which yields

$$\|f_\lambda(x) - f_\lambda(y)\| \le \theta\cdot\|\,x-y\|\,, \quad \text{for all} \quad x, y \in K. \tag{18}$$

In view of condition (12), we get that $0 < \theta < 1$, so f_λ is a θ-contraction. The conclusion now follows by Theorem 3.6 in [3] and Theorem 12.

A more general result can be similarly proven.

Theorem 14. *Let K be a nonempty closed convex subset of a Banach space and let $f : K \to K$ be a mapping satisfying the following assumptions:*

(i) $F_f \neq \emptyset$;
(ii) The Krasnoselskii iteration $\{x_n\}_{n=0}^{\infty}$ converges to $x^*(x) \in F_f$, for any $x \in K$;

(iii) There exist $0 \leq \delta < 1$ *and a function* $\theta : R_+ \to R_+$*, continuous at* 0 *with* $\theta(0) = 0$*, such that*

$$\|f(x) - x^*\| \leq \delta\|x - x^*\| + \theta(\|x - f(x)\|),\ \forall x \in K, x^* \in F_f. \tag{19}$$

Then the following retraction-displacement condition holds:

$$\|x - f_\lambda^\infty(x)\| \leq \frac{1}{1-\delta}\left(\|x - f(x)\| + \theta(\|x - f(x)\|)\right),\ \forall x \in K.$$

Mann Algorithm

Let us consider the Mann algorithm corresponding to f and Λ (see [3, 10, 51]):

$$x_0 \in Y,\ x_{n+1}(x_0) = f_{\lambda_n}(x_n(x_0)),\ n \in N.$$

We suppose that this algorithm is convergent, i.e.,

$$\text{for all } x_0 \in Y,\ x_n(x_0) \longrightarrow x^*(x_0) \in F_f \text{ as } n \to \infty.$$

In this condition we define the operator

$$f_\Lambda^\infty : Y \to F_f,\ x \mapsto x^*(x)$$

operator which is a set retraction.

By definition a convergent Mann algorithm satisfies a retraction-displacement condition if

$$\|x - f_\Lambda^\infty(x)\| \leq \psi(\|x - f(x)\|),\ \forall\, x \in Y$$

with ψ as in Definition 1.

The basic problem is the following:

Problem 5. If the Mann algorithm is convergent, in which conditions on f and Λ it satisfies a retraction-displacement condition?

Remark 9. For the conditions in which the Mann algorithm is convergent see: [3, 10, 51], etc. For example, we have the following well-known result, see [3], Chap. 4.

Theorem 15. *Let E be an arbitrary Banach space, K a closed convex subset of E, and* $f : K \to K$ *a Zamfirescu operator. Let* $\{x_n\}_{n=0}^\infty$ *be the Mann iteration defined by* $x_0 \in K$ *and*

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n f(x_n),\ n = 1, 2, \ldots \tag{20}$$

with $\{\alpha_n\} \subset [0, 1]$ *satisfying*

$$(iv) \qquad \sum_{n=0}^{\infty} \alpha_n = \infty \,.$$

Then

(i) f *is a Picard operator with* $F_f = \{p\}$*;*
(ii) $\{x_n\}_{n=0}^{\infty}$ *converges strongly to the unique fixed point of* f*;*
(iii) f *satisfies the following retraction-displacement condition*

$$\|x - p\| \leq \frac{2\delta}{1-\delta} \|x - f(x)\|, \ \forall x \in K.$$

Proof. Remind that if f is a Zamfirescu mapping on K, then there exist the real numbers a, b, c satisfying $0 \leq a < 1,\ 0 \leq b < 0.5$ and $0 \leq c < 0.5$, such that, for each $x, y \in X$, at least one of the following is true:

(z_1) $\|f(x) - f(y)\| \leq a\ \|x - y\|$;
(z_2) $\|f(x) - f(y)\| \leq b[\|x - f(x)\| + \|y - f(y)\|]$;
(z_3) $\|f(x) - f(y)\| \leq c[\|x - f(y)\| + \|y - f(x)\|]$.

It is well known that f is a Picard operator (see, for example, Theorem 2.4 in [3]). By (z_1)–(z_3), we obtain that, for all $x, y \in K$, T satisfies

$$\|f(x) - f(y)\| \leq \delta\|x - y\| + 2\delta\|x - f(x)\| \tag{21}$$

where

$$\delta = \max\left\{a, \frac{b}{1-b}, \frac{c}{1-c}\right\} < 1. \tag{22}$$

Let $\{x_n\}_{n=0}^{\infty}$ be the Mann iteration given by 20, with $x_0 \in K$ arbitrary. Then

$$\begin{aligned} \|x_{n+1} - p\| &= \big\|(1-\alpha_n)x_n + \alpha_n f(x_n) - (1 - \alpha_n + \alpha_n)p\big\| = \\ &= \big\|(1-\alpha_n)(x_n - p) + \alpha_n(f(x_n) - p)\big\| \leq \\ &\leq (1-\alpha_n)\|x_n - p\| + \alpha_n\|f(x_n) - p\| \,. \end{aligned} \tag{23}$$

Take $x := p$ and $y := x_n$ in 21 to obtain

$$\|f(x_n) - p\| \leq \delta \cdot \|x_n - p\|,$$

which together with 23 yields

$$\|x_{n+1} - p\| \leq \big[1 - (1-\delta)\alpha_n\big]\|x_n - p\| \,, \quad n = 0, 1, 2, \ldots \,. \tag{24}$$

Inductively we get

$$\|x_{n+1} - p\| \leq \prod_{k=0}^{n} \left[1 - (1-\delta)\alpha_k\right] \cdot \|x_0 - p\| \,, \quad n = 0, 1, 2, \dots . \tag{25}$$

As $0 < \delta < 1$, $\alpha_k \in [0, 1]$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$, by a standard argument it results that

$$\lim_{n\to\infty} \prod_{k=0}^{n} \left[1 - (1-\delta)\alpha_k\right] = 0 \,,$$

which, together with the previous inequality, implies

$$\lim_{n\to\infty} \|x_{n+1} - p\| = 0 \,,$$

i.e., $\{x_n\}_{n=0}^{\infty}$ converges strongly to p. So, (i) and (ii) are proven. To prove (iii), we use the fact that

$$\|x - p\| \leq \|x - f(x)\| + \|f(x) - f(p)\|.$$

So, by inequality 21,

$$\|f(x) - f(p)\| \leq \delta\|x - p\| + 2\delta\|x - f(x)\|,$$

the desired estimate follows.

A more general result can be similarly proven.

Theorem 16. *Let K be a nonempty closed convex subset of a Banach space and let $f : K \to K$ be a mapping satisfying the following assumptions:*

(i) $F_f \neq \emptyset$*;*
(ii) The Mann iteration $\{x_n\}_{n=0}^{\infty}$ converges to $x^*(x) \in F_f$, for any $x \in K$;
(iii) There exist $0 \leq \delta < 1$ and a function $\theta : R_+ \to R_+$, continuous at 0 with $\theta(0) = 0$, such that

$$\|f(x) - x^*\| \leq \delta\|x - x^*\| + \theta(\|x - f(x)\|), \; \forall x \in K, x^* \in F_f. \tag{26}$$

Then, the following retraction-displacement condition holds:

$$\|x - f_\lambda^\infty(x)\| \leq \frac{1}{1-\delta} \left(\|x - f(x)\| + \theta(\|x - f(x)\|)\right) , \; \forall x \in K.$$

Remark 10. By Theorem 16, one can obtain a convergence theorem for Mann iteration by considering f an almost contraction with a unique fixed point, see, for example, [3].

Halpern Algorithm

Now we consider the Halpern algorithm (see [3, 10, 41, 50], etc.):

$$x_0 \in Y,\ x_{n+1}(x_0) = (1-\lambda_n)u + \lambda_n f(x_n(x_0)),\ n \in N,$$

where $u \in Y$ is a fixed anchor, see [47, 54, 55, 59] for more details.

We suppose that this algorithm is convergent, i.e.,

$$\text{for all } x_0 \in Y,\ x_n(x_0) \to x^*(x_0) \in F_f \text{ as } n \to \infty.$$

So, if we denote $\Lambda = \{\{\lambda_n\}_{n=0}^{\infty} : \lambda_n \in [0,1]\}$, then we have the set retraction, $f_\Lambda^\infty : Y \to F_f, f_\Lambda^\infty(x) = x^*(x)$.

By definition, a convergent Halpern algorithm satisfies a retraction-displacement condition if

$$\|x - f_\Lambda^\infty(x)\| \le \psi(\|x - f(x)\|),\ \forall\, x \in Y,$$

with ψ as in Definition 1.

Similarly to the case of the previous algorithms, we have

Problem 6. If the Halpern algorithm is convergent, under which condition on f and Λ it satisfies a retraction-displacement condition?

Remark 11. For some conditions under which the Halpern algorithm is convergent, see [1, 3, 33, 47, 50, 53–55, 59, 60], etc.

A general result similar to the ones established for Krasnoselskii and Mann algorithms can be easily proven for Halpern iteration, too.

Theorem 17. *Let K be a nonempty closed convex subset of a Banach space and let $f : K \to K$ be a mapping satisfying the following assumptions:*

(i) $F_f \neq \emptyset$*;*
(ii) The Halpern iteration $\{x_n(x)\}_{n=0}^{\infty}$ converges to $x^*(x) \in F_f$, for any $x \in K$;
(iii) There exist $0 \le \delta < 1$ and a function $\theta : R_+ \to R_+$, continuous at 0 with $\theta(0) = 0$, such that

$$\|f(x) - x^*\| \le \delta\|x - x^*\| + \theta(\|x - f(x)\|),\ \forall x \in K, x^* \in F_f. \tag{27}$$

Then the following retraction-displacement condition holds:

$$\|x - f_\Lambda^\infty(x)\| \le \frac{1}{1-\delta}\left(\|x - f(x)\| + \theta(\|x - f(x)\|)\right),\ \forall x \in K.$$

The next corollary provides an answer to Problem 6.

Corollary 3. *Let K be a nonempty closed convex subset of a Banach space and let $f : K \to K$ be a Zamfirescu mapping. Then*

(i) $F_f = \{x^*\}$;
(ii) The Halpern iteration $\{x_n\}_{n=0}^{\infty}$ *converges to* $x^* \in F_f$, *for any* $x_0 \in K$, *provided that* $\{\lambda_n\}_{n=0}^{\infty} \subset [0, 1]$ *satisfies the following condition:*

$$\lim_{n\to\infty} \lambda_n = 0. \tag{28}$$

(iii) The following retraction-displacement condition holds:

$$\|x - f_\Lambda^\infty(x)\| \le \frac{1+2\delta}{1-\delta}\|x - f(x)\|,\ \forall x \in K.$$

Proof. (i) This follows by Theorem 2.4 in [3].
(ii) Let $\{x_n\}_{n=0}^{\infty}$ be the Halpern iteration, defined by $x_0 \in K$, the fixed anchor $u \in K$ and the parameter sequence $\{\lambda_n\}_{n=0}^{\infty} \subset [0, 1]$ satisfying 28. Then we have

$$\|x_{n+1}-x^*\|=\|\lambda_n u+(1-\lambda_n)f(x_n)-x^*\|=\|\lambda_n u-\lambda_n x^*+(1-\lambda_n)(f(x_n)-f(x^*)\|$$
$$\le \lambda_n\|u - x^*\|+(1-\lambda_n)\|f(x_n)-f(x^*)\| \le \lambda_n\|u - x^*\| + (1-\lambda_n)\delta\|x_n - x^*\|$$
$$\le \delta\|x_n - x^*\| + \lambda_n\|u - x^*\|.$$

Thus

$$\|x_{n+1} - x^*\| \le \delta\|x_n - x^*\| + \lambda_n\|u - x^*\|,\ n \ge 0,$$

which, by applying Lemma 1.6 in [3], yields the conclusion.
(iii) Since f is a Zamfirescu mapping, see Theorem 15, the inequality 21 holds and so, by Theorem 16, we get the estimate

$$\|x - f_\Lambda^\infty(x)\| \le \frac{1+2\delta}{1-\delta}\|x - f(x)\|,\ \forall x \in K,$$

where δ is given by

$$\delta = \max\left\{a,\ \frac{b}{1-b},\ \frac{c}{1-c}\right\} < 1 \tag{29}$$

and a, b, c are the constants appearing in Zamfirescu's conditions $(z_1) - (z_3)$.

Retraction-Displacement Condition and the Condition (I), in the Case $F_f = \{x^*\}$

In this section we briefly discuss, in the context of fixed point iterative algorithms, the connection between the retraction-displacement condition considered in the present paper and the condition (I), the latter introduced by Senter and Dotson in [48] (see also [58]) for the case of single-valued mappings, and used by many authors mainly in the case of multi-valued operators, to study the convergence of Mann and Ishikawa iterations for nonexpansive type mappings, see [25, 48, 49, 56, 57] and the references therein.

Definition 5 ([48]). Let (X, d) be a metric space. A mapping $f : X \to X$ is said to satisfy condition (I) if there is a nondecreasing function $\theta : [0, +\infty) \to [0, +\infty)$ with $\theta(0) = 0$ and $\theta(r) > 0$ for all $r > 0$ such that

$$d(x, Tx) \geq \theta(d(x, F_f)), \forall x \in X,$$

where F_f denotes, as usually, the set of fixed points of f.

In the particular case announced in the title of this section, i.e., when $F_f = \{x^*\}$, it is easy to see that condition (I) requires in fact that

$$d(x, Tx) \geq \theta(d(x, x^*)), \forall x \in X.$$

Example 2. Let (X, d) be a metric space and $f : X \to X$ a λ-contraction $(0 \leq \lambda < 1)$. Then, f satisfies condition (I) with $\theta(r) = (1 - \lambda) \cdot r$, for all $r > 0$.

Example 3. Let (X, d) be a metric space and $f : X \to X$ a Kannan mapping, i.e., a mapping for which there exists $0 < \beta < 0.5$ such that

$$d(f(x), f(y)) \leq \beta[d(x, f(x)) + d(y, f(y))], \ \forall x, y \in X.$$

Then, f satisfies condition (I) with $\theta(r) = \dfrac{1 - 2\beta}{2\beta} r$, for all $r > 0$.

Example 4. Let (X, d) be a metric space and $f : X \to X$ a Zamfirescu mapping, see Theorem 16. Then, f satisfies condition (I) with $\theta(r) = \dfrac{1 - 2\delta}{2\delta} r$, for all $r > 0$, where

$$\delta = \max\left\{a, \frac{b}{1-b}, \frac{c}{1-c}\right\} < 1, \tag{30}$$

and a, b, c are the constants appearing in conditions (z_1)–(z_3).

Based on Examples 2–4, we can state the following generic result.

Proposition 1. *Let K be a nonempty closed convex subset of a Banach space and let $f : K \to K$ be a mapping satisfying the following assumptions:*

(i) $F_f = \{x^*\}$;
(ii) *A certain iterative algorithm* $f_P \equiv \{x_n(x_0)\}_{n=0}^{\infty}$ *converges to* x^*, *for any* $x_0 \in K$;
(iii) f *satisfies condition (I) with* θ *a bijection.*

Then the algorithm f_P satisfies the following retraction-displacement condition:

$$\|x - f_P^{\infty}(x)\| \leq \theta^{-1} \left(\|x - f(x)\|\right), \ \forall x \in K.$$

Remark 12. In the case of Krasnoselskii algorithm we have $P = \{\lambda\}$, $\lambda \in (0, 1)$, and hence $f_P \equiv f_\lambda$, while in the case of Mann iteration we have $P \equiv \Lambda = \{\lambda_n\}_{n=0}^{\infty}$, $\lambda_n \in (0, 1)$, and hence $f_P \equiv f_\Lambda$.

The Impact of Retraction-Displacement Condition on the Theory of Fixed Point Equations

Let (X, d) be a metric space and $f : X \to X$ an operator with $F_f \neq \emptyset$. We suppose that f satisfies a retraction-displacement condition as in Definition 1. In this section we consider the fixed point equation

$$x = f(x). \tag{31}$$

Data Dependence

Let us consider the fixed point equation (31) and let $g : X \to X$ be an operator such that $F_g \neq \emptyset$.

We have

Theorem 18. *We suppose that:*

(i) f *satisfies the* (r_1, ψ_1) *retraction-displacement condition;*
(ii) g *satisfies the* (r_2, ψ_2) *retraction-displacement condition;*
(iii) *there exists* $\eta > 0$ *such that*

$$d(f(x), g(x)) \leq \eta, \ \forall\, x \in X.$$

Then, $H_d(F_f, F_g) \leq \max(\psi_1(\eta), \psi_2(\eta))$.

Proof. Let $x^* \in F_f$. Then, $r_2(x^*) \in F_g$ and

$$d(x^*, r_2(x^*)) \leq \psi_2(d(x^*), g(x^*)) = \psi_2(d(f(x^*), g(x^*))) \leq \psi_2(\eta).$$

Let $y^* \in F_g$. Then $r_1(y^*) \in F_f$ and

$$d(y^*, r_1(y^*)) \leq \psi_1(d(y^*), f(y^*)) = \psi_1(d(g(y^*), f(y^*))) \leq \psi_1(\eta).$$

From a well-known property of the Pompeiu–Hausdorff functional (see, for example, [31], p. 76) it follows that

$$H_d(F_f, F_g) \leq \max(\psi_1(\eta), \psi_2(\eta)).$$

Theorem 19. *We suppose that*

(i) $F_f = \{x^*\}$ *and* f *satisfies a* (r, ψ) *retraction-displacement condition, where* $r(x) = x^*$, $\forall\, x \in X$;
(ii) there exists $\eta > 0$ *such that*

$$d(f(x), g(x)) \leq \eta, \ \forall\, x \in X.$$

Then, $d(y^*, x^*) \leq \psi(\eta)$, *for each* $y^* \in F_g$.

Proof. Let $y^* \in F_g$. Then

$$d(y^*, x^*) = d(y^*, r(y^*)) \leq \psi(d(y^*), f(y^*)) = \psi(d(g(y^*)), f(y^*)) \leq \psi(\eta).$$

Example 5. Let us consider the following functional integral equation

$$x(t) = \int_{t-\tau}^{t} f(s, x(s))ds, \text{ where } t \in R_+. \tag{32}$$

This equation is a mathematical model for epidemics and population growth (see, for example, [12, 14] and the references therein.

Let $0 < m < M$ and $) < \alpha < \beta$. We suppose:

(i) $f \in C(R \times [\alpha, \beta]);$
(ii) there exists $\omega > 0$ such that $f(t + \omega, u) = f(t, u)$, for all $t \in R$ and all $u \in [\alpha, \beta]);$
(iii) there exists $k > 0$ such that $k\tau < 1$ and $|f(t, u) - f(t, v)| \leq k|u - v|$, for all $t \in R$ and all $u, v \in [\alpha, \beta]);$
(iv) $m \leq f(t, u) \leq M$, for all $t \in R$ and all $u \in [\alpha, \beta]);$
(v) $\alpha \leq m\tau$ and $\beta \geq M\tau$.

If we define

$$X_\omega := \{x \in C(R, [\alpha, \beta]) \mid x(t + \omega) = x(t), \text{ for each } t \in R\}$$

endowed with the metric

$$d(x,y) := \max_{0\le t\le \omega} |x(t) - y(t)|,$$

then the operator A defined as

$$Ax(t) := \int_{t-\tau}^{t} f(s,x(s))ds, \text{ where } t \in R_+$$

has (using (i)–(iv)) the following properties:

(a) $A(X_\omega) \subset X_\omega$;
(b) A is a $k\tau$-contraction.

Then, by the Contraction Principle and Theorem 19, we obtain

Theorem 20. *Let us consider Eq. (32) and suppose that the assumptions (i)–(v) take place. Then:*

(1) Eq. (32) has in X_ω a unique solution x^*;
(2) $d(x,x^*) \le \frac{1}{1-k\tau} d(x,A(x))$, for all $x \in X_\omega$;
(3) let $g : R \times [\alpha,\beta] \to R$ be a function which satisfies the conditions (i),(ii),(iv) and (v) above. In addition, we suppose that there exists $\eta > 0$ such that

$$|f(t,u) - g(t,u)| \le \eta, \text{ for all } t \in R \text{ and } u \in [\alpha,\beta]);$$

If $y \in X_\omega$ is a solution of the integral equation

$$y(t) = \int_{t-\tau}^{t} g(s,y(s))ds, \text{ where } t \in R_+,$$

then

$$d(x,y) \le \frac{\tau\eta}{1-k\tau}.$$

Ulam Stability

We start our considerations with the following notions (see [44]).

Definition 6. By definition, the fixed point equation (31) is Ulam–Hyers stable if there exists a constant $c_f > 0$ such that: for each $\varepsilon > 0$ and each solution $y^* \in X$ of the inequation

$$d(y,f(y)) \le \varepsilon \tag{33}$$

there exists a solution x^* of Eq. (31) such that

$$d(y^*, x^*) \leq c_f \varepsilon.$$

Definition 7. By definition, Eq. (31) is generalized Ulam–Hyers stable if there exists $\theta : R_+ \to R_+$ increasing and continuous in 0 with $\theta(0) = 0$ such that: for each $\varepsilon > 0$ and for each solution y^* of (33) there exists a solution x^* of (31) such that

$$d(y^*, x^*) \leq \theta(\varepsilon).$$

We have

Theorem 21. *If f satisfies a (r, ψ) retraction-displacement condition, then Eq. (31) is generalized Ulam–Hyers stable.*

Proof. Let $y^* \in X$ be a solution of (33). Then $x^* = r(y^*) \in F_f$. Since f satisfies the (r, ψ) retraction-displacement condition we have

$$d(y^*, x^*) \leq \psi(d(y^*), f(y^*)) \leq \psi(\varepsilon).$$

Remark 13. If in the Theorem 21, the function $\psi(t) = c_f t$, $\forall t \in R_+$, then Eq. (31) is Ulam–Hyers stable.

Example 6. Let Ω be a bounded domain in R^m and let $X := C(\overline{\Omega}, R)$ endowed with the metric $d(x, y) := \max_{t \in \overline{\Omega}} |x(t) - y(t)|$.

We consider on X the following integral equation

$$x(t) = \int_{\Omega} K(t, s, x(s))ds + l(t), \; t \in \Omega. \tag{34}$$

With respect to the above equation, we suppose:

(i) $K \in C([\overline{\Omega} \times \overline{\Omega} \times R, R)$ and $l \in C(\overline{\Omega}, R)$;
(ii) there exists $k > 0$ such that

$$|K(t, s, u) - K(t, s, v)| \leq k|u - v|, \text{ for all } t, s \in \Omega \text{ and } u, v \in R;$$

(iii) $k \cdot mes(\Omega) < 1$.

Then, if we define $A : X \to X$ by

$$Ax(t) := \int_{\Omega} K(t, s, x(s))ds + l(t), \; t \in \Omega,$$

then, by (ii) and (iii), we obtain that A is a $k \cdot mes(\Omega)$-contraction. Applying the Contraction Principle and Theorem 21, we get

Theorem 22. *Consider Eq. (34) and suppose that the assumption (i)–(iv) take place. Then, Eq. (34) is Ulam–Hyers stable.*

Remark 14. For Ulam stability theory related to fixed point equations see [7, 21, 26, 44], etc.

Well-Posedness for Fixed Point Problems

Let (X, d) be a metric space and $f : X \to X$ be an operator such that its fixed point set $F_f = \{x^*\}$. Following F.S. De Blasi and J. Myjak (see [39] p. 42, see also [43]), we say, by definition, that the fixed point problem

$$x = f(x), x \in X$$

is well-posed if the following implication holds:

$$(x_n)_{n\in N} \subset X \text{ and } d(x_n, f(x_n)) \to 0 \text{ as } n \to +\infty \;\Rightarrow\; \lim_{n\to\infty} x_n = x^*.$$

In this setting, we have the following general result.

Theorem 23. *Let (X, d) be a metric space and $f : X \to X$ be an operator such that its fixed point set $F_f = \{x^*\}$. If the operator f satisfies an (r, ψ) retraction-displacement condition, then the fixed point problem for f is well-posed.*

Proof. Let $(x_n)_{n\in N} \subset X$ such that $d(x_n, f(x_n)) \to 0$ as $n \to +\infty$. Then, we have

$$d(x_n, x^*) \leq \psi(d(x_n, f(x_n)) \to 0 \text{ as } n \to \infty.$$

Ostrowski Stability

Let (X, d) be a metric space and $f : X \to X$ be an operator such that its fixed point set $F_f = \{x^*\}$. Let

$$x_{n+1} = f_n(x_n), n \in N$$

be an iterative algorithm with $f_n : X \to X$. By definition, this algorithm is said to be Ostrowski stable if the following implication holds:

$$(y_n)_{n\in N} \subset X \text{ and } d(y_{n+1}, f_n(y_n)) \to 0 \text{ as } n \to +\infty \;\Rightarrow\; \lim_{n\to\infty} y_n = x^*.$$

Some authors refer to the above property as the “limit shadowing property” (see [16, 23, 29, 41, 46], etc.).

The following open question seems to be a difficult one.

Open Question In which conditions a retraction-displacement condition on f implies that the fixed point problem for f is Ostrowski stable?

Remark 15. For a general fixed point iteration scheme and the concept of virtual stability of an iteration scheme see [9]. For other stability concepts see also [15].

Some New Research Directions

Examples in R^m

We will present in this section some examples related to the following problem:

Problem 7. Which operators $f : R^n \to R^n$ with $F_f \neq \emptyset$ and $r : R^n \to F_f$ a retraction satisfies the retraction-displacement condition

$$d\left(x, r\left(x\right)\right) \leq \psi\left(d\left(x, f\left(x\right)\right)\right), \quad \text{for all } x \in R^n \text{ ?}$$

Example 7. Let $f : R \to R$ such that $F_f = [a; b]$. Then f satisfies the retraction-displacement condition with $\psi(t) = ct$, $c > 0$ and $r : R \to F_f$

$$r\left(x\right) = \begin{cases} a, \ x < a \\ x, \ x \in [a; b] \\ b, \ x > b \end{cases}$$

if the following conditions are satisfied:

$$\frac{a}{c} + \frac{c-1}{c} x \leq f\left(x\right) \quad \text{or} \quad \frac{c+1}{c} x - \frac{a}{c} \geq f\left(x\right) \quad \text{for} \quad x < a,$$

and

$$\frac{b}{c} + \frac{c-1}{c} x \geq f\left(x\right) \quad \text{or} \quad \frac{c+1}{c} x - \frac{b}{c} \leq f\left(x\right) \quad \text{for} \quad x > b.$$

Example 8. Let $R > 0$ and

$$\overline{D_i} = \left\{(x, y) \in R^2 \mid x^2 + y^2 \leq R\right\},$$
$$D_e = \left\{(x, y) \in R^2 \mid x^2 + y^2 > R\right\}$$

Let $f : R^2 \to R^2$ be defined by

$$f(x,y)=\begin{cases}(x,y), & (x,y)\in\overline{D_i}\\ (x\cos\alpha-y\sin\alpha, x\sin\alpha+y\cos\alpha), & (x,y)\in D_e\end{cases}$$

with $\alpha\in]0;2\pi[$.

In this case we have that $F_f=\overline{D_i}$.

Let's consider

$$r(x,y)=\begin{cases}(x,y), & (x,y)\in\overline{D_i}\\ \left(R\frac{x}{\sqrt{x^2+y^2}}, R\frac{y}{\sqrt{x^2+y^2}}\right), & (x,y)\in D_e.\end{cases}$$

Then

$$\|(x,y)-r(x,y)\|\le c\cdot\|(x,y)-f(x,y)\|,\ (x,y)\in R^2,$$

with $c=\frac{1}{\sqrt{2(1-\cos\alpha)}}$, i.e., f satisfies the retraction-displacement condition with $\psi(t)=ct$.

Proof. For $(x,y)\in\overline{D_i}$ the retraction-displacement condition with $\psi(t)=ct$, $c>0$, and r is satisfied for any $c>0$.

If $(x,y)\in D_e$, then

$$\|(x,y)-r(x,y)\|=\|(x,y)\|-R$$

and

$$\|(x,y)-f(x,y)\|=\|(x,y)\|\cdot\sqrt{2(1-\cos\alpha)}.$$

The function $\varphi:[R;+\infty[\to R$ defined by

$$\varphi(t)=\frac{t-R}{t\sqrt{2(1-\cos\alpha)}}$$

is an increasing bounded function with

$$\varphi(t)<\frac{1}{\sqrt{2(1-\cos\alpha)}},\ t\in[R;+\infty[,$$

thus

$$\frac{\|(x,y)-r(x,y)\|}{\|(x,y)-f(x,y)\|}=\frac{\|(x,y)\|-R}{\|(x,y)\|\cdot\sqrt{2(1-\cos\alpha)}}<\frac{1}{\sqrt{2(1-\cos\alpha)}},\ \forall (x,y)\in D_e,$$

and we get the conclusion.

The Case of R_+^m-Metric Spaces

Another research direction is the study of the retraction-displacement condition in the case of generalized metric spaces (see [18, 23, 39, 44, 44, 45], etc.) For example, if we consider a vector-valued metric (i.e., $d(x, y) \in R_+^m$, then we can do the following commentaries:

(1) Let X be a nonempty set and let $d : X \times X \to R_+^m$ be a R_+^m-metric on X. Let $f : X \to X$ be an operator such that $F_f \neq \emptyset$ and let $r : X \to F_f$ be a set retraction. Then, by definition, f satisfies the (ψ, r) retraction-displacement condition if:

  (i) $\psi : R_+^m \to R_+^m$ is increasing, continuous at 0 with $\psi(0) = 0$;
  (ii) $d(x, r(x)) \leq \psi(d(x, f(x))$, for all $x \in X$.

(2) For the weakly Picard operator theory in R_+^m-metric spaces, see [36, 37, 39, 45], etc.
(3) For the Ulam stability in R_+^m-metric spaces, see [7, 26, 44], etc.

The Case of Nonself Operators

Let (X, d) be a metric space and Y be a nonempty subset of X. Then, by definition, the operator $f : Y \to X$ with $F_f \neq \emptyset$ satisfies the (ψ, r) retraction-displacement condition if:

(i) $\psi : R_+^m \to R_+^m$ is increasing, continuous at 0 with $\psi(0) = 0$;
(ii) $r : Y \to F_f$ is a set retraction;
(iii) $d(x, r(x)) \leq \psi(d(x, f(x)))$, for all $x \in Y$.

In this case, the problem is to study the fixed point equation $x = f(x)$ in terms of the (ψ, r) retraction-displacement condition.

Again some commentaries can be done:

(1) Let $\tilde{r} : X \to Y$ be a set retraction such that $F_f = F_{\tilde{r} \circ f}$. Then, the problem is to find a retraction $r : Y \to F_f$ as the limit operator of an iterative algorithm corresponding to the self-operator $\tilde{r} \circ f$.
(2) For some results concerning this problem, see [5, 11, 29, 43].

The Case of Multi-valued Operators

We will present first a concept of set-retraction related to multi-valued operators. Recall first that, if $T : X \to P(X)$ is a multi-valued operator, then we denote by

$$Graph(T) := \{(x, y) \in X \times X : \ y \in T(x)\}, \text{ the graphic of } T$$

and by

$$F_T := \{x \in X | \ x \in T(x)\}, \text{ the fixed point set of } T.$$

Definition 8. Let X be a nonempty set, $Y \in P(X)$ and $T : X \to P(X)$ be a multi-valued operator. An operator $r : Graph(T) \to Y$ is called a strong set-retraction of X onto Y if $r(x, x) = x$, for all $x \in Y$.

Then, we define the retraction-displacement condition for multi-valued operators as follows.

Definition 9. Let (X, d) be a metric space and let $T : X \to P(X)$ be a multi-valued operator such that its fixed point set F_T is nonempty. Then, by definition, T satisfies the (ψ, r) retraction-displacement condition if there exist $\psi : R_+ \to R_+$ and a strong set retraction $r : Graph(T) \to P(F_T)$ such that:

(i) $\psi : R_+ \to R_+$ is increasing, continuous at 0 with $\psi(0) = 0$;
(ii) $d(x, r(x, y)) \leq \psi(d(x, y))$, for all $(x, y) \in Graph(T)$.

In this case, the problem is to study the fixed point inclusion $x \in T(x)$ and the strict fixed point equation $\{x\} = T(x)$ in terms of the (ψ, r) retraction-displacement condition.

The case of nonself multi-valued operators can also be considered in a similar way.

Moreover, in particular, if $T : X \to P(X)$ is a multi-valued weakly Picard operator (i.e., for each $(x, y) \in Graph(T)$ there exists a sequence $(x_n)_{n\in N}$ such that:

(i) $x_0 = x, x_1 = y$;
(ii) $x_{n+1} \in T(x_n)$, for each $n \in N$;
(iii) $(x_n)_{n\in N}$ is convergent and its limit is a fixed point of T),
and we define the multi-valued operator $T^\infty : Graph(T) \to P(F_T)$ by the formula $T^\infty(x, y) := \{ z \in F_T \mid$ there exists a sequence $(x_n)_{n\in N}$ satisfying the assertions (i) and (ii) and convergent to z), then the strong set retraction r is any selection of T^∞ which satisfies the condition (ii) in Definition 9.

For the weakly Picard operator theory for multi-valued operators and related topics (data dependence, Ulam–Hyers stability, iterative algorithms) see [21, 26, 27, 29, 30, 36, 39, 44], etc. For retraction theory in the multi-valued operators context, see also [8].

References

1. Al-Rumaih, Z., Chebbi, S., Xu, H.-K.: A Halpern–Lions–Reich-like iterative method for nonexpansive mappings. Fixed Point Theory **14**, 289–300 (2013)
2. Barnsley, M.F., Ervin, V., Hardin, D., Lancaster, J.: Solution of an inverse problem for fractals and other sets. Proc. Nat. Acad. Sci. U. S. A. **83**, 1975–1976 (1986)
3. Berinde, V.: Iterative Approximation of Fixed Points. Springer, Berlin (2007)
4. Berinde, V.: Convergence theorems for fixed point iterative methods defined as admissible perturbation of a nonlinear operator. Carpathian J. Math. **29**, 9–18 (2013)
5. Berinde, V., Măruşter, Ş., Rus, I.A.: An abstract point of view on iterative approximation of fixed points of nonself operators. J. Nonlinear Convex Anal. **15**, 851–865 (2014)
6. Berinde, V., Păcurar, M., Rus, I.A.: From a Dieudonné theorem concerning the Cauchy problem to an open problem in the theory of weakly Picard operators. Carpathian J. Math. **30**, 283–292 (2014)
7. Bota-Boriceanu, M., Petruşel, A.: Ulam–Hyers stability for operatorial equations. Anal. Univ. Al.I. Cuza Iaşi **57**, 65–74 (2011)
8. Brown, R.F.: Retraction mapping principle in Nielsen fixed point theory. Pac. J. Math. **115**, 277–297(1984)
9. Chaoha, P., Chanthorm, P.: Fixed point sets through iteration schemes. J. Math. Anal. Appl. **386**, 273–277 (2012)
10. Chidume, C.E., Măruşter, S.: Iterative methods for the computation of fixed points of demicontractive mappings. J. Comput. Appl. Math. **234**, 861–882 (2010)
11. Chiş-Novac, A., Precup, R., Rus, I.A.: Data dependence of fixed points for nonself generalized contractions. Fixed Point Theory **10**, 73–87 (2009)
12. Cooke, K.L., Kaplan, J.L.: A periodicity threshold theorem for epidemics and population growth. Math. Biosci. **31**, 87–104 (1976)
13. Derafshpour, M., Rezapour, S.: Picard operators on ordered metric spaces. Fixed Point Theory **15**, 59–66 (2014)
14. Dobriţoiu, M., Rus, I.A., Şerban, M.A.: An integral equation arising from infectious diseases via Picard operators. Studia Univ. Babeş-Bolyai Math. **52**, 81–94 (2007)
15. Furi, M., Vignoli, A. A remark about some fixed point theorems. Boll. Un. Mat. Ital. **3**, 197–200 (1970)
16. Jachymski, J.: An extension of Ostrovski's theorem on the round-off stability of iterations. Aequa. Math. **53**, 242–253 (1977)
17. Jachymski, J., Józwik, I.: Nonlinear contractive conditions: a comparison and related problems. Banach Center Publ. **77**, 123–146 (2007)
18. Kirk, W.A., Sims, B. (eds.): Handbook of Fixed Point Theory. Kluwer Academic Publishers, Dordrecht (2001)
19. Kirk, W.A., Srinivasan, P.S., Veeramani, P.: Fixed points for mappings satisfying cyclical contractive conditions. Fixed Point Theory **4**, 79–89 (2003)
20. Matkowski, J.: Integrable solutions of functional equations. Dissertationes Math. (Rozprawy Mat.) **127**, 68 (1975)
21. Mleşniţe, O., Petruşel, A.: Existence and Ulam–Hyers stability results for multivalued coincidence problems. Filomat **26**, 965–976 (2012)
22. Mureşan, A.S.: The theory of some asymptotic fixed point theorems. Carpathian J. Math. **30**, 361–368 (2014)
23. Ortega, J.M., Rheinboldt, W.C.: Iterative Solutions of Nonlinear Equation in Several Variables.. Academic, New York (1970)
24. Păcurar, M., Rus, I.A.: Fixed point theory for cyclic φ-contractions. Nonlinear Anal. **72**, 1181–1187 (2010)
25. Panyanak, B.: Mann and Ishikawa iterative processes for multivalued mappings in Banach spaces. Comput. Math. Appl. **54**, 872–877 (2007)

26. Petru, T.P., Petruşel, A., Yao, J.-C.: Ulam–Hyers stability for operatorial equations and inclusions via nonself operators. Taiwan. J. Math. **15**, 2195–2212 (2011)
27. Petruşel, A.: Multivalued weakly Picard operators and applications. Sci. Math. Jpn. **59**, 169–202 (2004)
28. Petruşel, A.: Ćirić type fixed point theorems. Studia Univ. Babeş-Bolyai **59**, 233–245 (2014)
29. Petruşel, A., Rus, I.A.: An abstract point of view on iterative approximation schemes of fixed points for multivalued operators. J. Nonlinear Sci. Appl. **6**, 97–107 (2013)
30. Petruşel, A., Mleşniţe, O., Urs, C.: Vector-valued metrics in fixed point theory. Contemp. Math. **636**, 149–165 (2015). http://dx.doi.org/10.1090/conm/636/12734
31. Petruşel, A., Rus, I.A., Şerban, M.A.: The role of equivalent metrics in fixed point theory. Topol. Meth. Nonlinear Anal. **41**, 85–112 (2013)
32. Qing, Y., Rhoades, B.E.: T-stability of Picard iteration in metric space. Fixed Point Theory Appl. 4, Article ID 418971 (2008). doi:10.1155/2008/418971
33. Reich, S.: Strong convergence theorems for resolvents of accretive operators in Banach spaces. J. Math. Anal. Appl. **75**, 287–292 (1980)
34. Rhoades, B.E.: A comparison of various definitions of contractive mappings. Trans. Am. Math. Soc. **226**, 257–290 (1977)
35. Rus, I.A.: Maps with φ-contraction iterates. Studia Univ. Babeş-Bolyai Math. **25**(4), 47–51 (1980)
36. Rus, I.A.: Generalized Contractions and Applications. Cluj University Press, Cluj-Napoca (2001)
37. Rus, I.A.: Picard operators and applications. Sci. Math. Jpn. **58**, 191–219 (2003)
38. Rus, I.A. Cyclic representations and fixed points. Ann. T. Popoviciu Seminar Funct. Eq. Approx. Convexity **3**, 171–178 (2005)
39. Rus, I.A., Petruşel, A., Petruşel, G.: Fixed Point Theory. Cluj University Press, Cluj-Napoca, Romania (2008)
40. Rus, I.A.: Gronwall lemmas: ten open problems. Sci. Math. Jpn. **70**, 221–228 (2009)
41. Rus, I.A.: An abstract point of view on iterative approximation of fixed point equations. Fixed Point Theory **13**, 179–192 (2012)
42. Rus, I.A.: Properties of the solutions of those equations for which the Krasnoselskii iteration converges. Carpathian J. Math. **28**, 329–336 (2012)
43. Rus, I.A.: The generalized retraction methods in fixed point theory for nonself operators. Fixed Point Theory **15**, 559–578 (2014)
44. Rus, I.A.: Results and problems in Ulam stability of operatorial equations and inclusion. In: Rassias, T.M. (ed.) Handbook of Functional Equations: Stability Theory, pp. 323–352. Springer, Berlin (2014)
45. Rus, I.A., Petruşel, A., Şerban, M.A.: Weakly Picard operators: equivalent definitions, applications and open problems. Fixed Point Theory **7**, 3–22 (2006)
46. Rus, I.A., Şerban, M.-A.: Basic problems of the metric fixed point theory and the relevance of a metric fixed point theorem. Carpathian J. Math. **29**, 239–258 (2013)
47. Saejung, S.: Halpern's iteration in Banach spaces. Nonlinear Anal. **73**, 3431–3439 (2010)
48. Senter, H.F., Dotson, W.G.: Approximating fixed points of nonexpansive mappings. Proc. Am. Math. Soc. **44**, 375–380 (1974)
49. Shahzad, N., Zegeye, H.: On Mann and Ishikawa schemes for multivalued maps in Banach spaces. Nonlinear Anal. **71**, 838–844 (2009)
50. Sine, R.C. (ed.): Fixed Points and Nonexpansive Mappings. Contemporary Mathematics, Amer. Math. Soc., Providence vol. 18 (1983)
51. Singh, S.P., Watson, B.: On convergence results in fixed point theory. Rend. Sem. Mat. Univ. Politec. Torino **51**, 73–91 (1993)
52. Smart, D.R.: Fixed Point Theorems. Cambridge University Press, London (1974)
53. Suzuki, T.: Moudafi's viscosity approximations with Meir-Keller contractions. J. Math. Anal. Appl. **325**, 342–352 (2007)
54. Suzuki, T.: Reich's problem concerning Halpern's convergence. Arch. Math. (Basel) **92**, 602–613 (2009)

55. Song, Y.: A new sufficient condition for the strong convergence of Halpern type iterations. Appl. Math. Comput. **198**, 721–728 (2008)
56. Song, Y., Cho, Y.J.: Some notes on Ishikawa iteration for multi-valued mappings. Bull. Korean Math. Soc. **48**, 575–584 (2011)
57. Song, Y., Wang, H.: Convergence of iterative algorithms for multivalued mappings in Banach spaces. Nonlinear Anal. **70**, 1547–1556 (2009)
58. Thele, R.L.: Iterative technique for approximation of fixed points of certain nonlinear mappings in Banach spaces. Pac. J. Math. **53**, 259–266 (1974)
59. Wittman, R.: Approximation of fixed points of nonexpansive mappings. Arch. Math. (Basel) **58**, 486–492 (1992)
60. Xu, H.K.: Another control condition in an iterative method for nonexpansive mappings. Bull. Aust. Math. Soc. **65**, 109–113 (2002)

An Adaptive Finite Element Method for Solving a Free Boundary Problem with Periodic Boundary Conditions in Lubrication Theory

O. Chau, D. Goeleven, and R. Oujja

Abstract A duality approximation method combined with an adaptive finite element method is applied to solve a free boundary problem with periodic boundary conditions in lubrication theory.

Keywords Lubrication theory • Free boundary problem • Periodic conditions • Adaptive finite element method

Introduction

We consider a journal bearing consisting of two circular cylinders—the shaft and the bush—that are in relative motion with a constant angular velocity ω and have parallel axis (Fig. 1).

The narrow gap between the cylinders is occupied by a lubricant and the pressure p of this lubricant satisfies the Reynolds equation. Let r, θ, x be cylindrical coordinates with origin in the bottom of the cylinder and θ measured from the line of maximum clearance of the centers. As a consequence of the thin-film hypothesis, the pressure depends on the angular coordinate θ and the height of the cylinders x and does not depend on the normal coordinate r. Therefore, the mathematical model can be posed on the set $\Omega = [0, 2\pi] \times [0, 1]$ which represents the lateral area of the shaft.

Let r_s be the radius of the shaft and r_b the radius of the bush, e the distance between the axes of the cylinders, and $\eta = \frac{e}{r_s - r_b}$, $0 \leq \eta < 1$ the eccentricity ratio of the bearing. The thickness of the thin fluid film is represented by the function

$$h(\theta) = (r_s - r_b)(1 - \eta \cos(\theta - \alpha)) \tag{1}$$

O. Chau • D. Goeleven (✉) • R. Oujja
University of La Réunion, PIMENT EA4518, 97715 Saint-Denis Messag cedex
9 La Réunion, France
e-mail: oanh.chau@univ-reunion.fr; daniel.goeleven@univ-reunion.fr;
rachid.oujja@univ-reunion.fr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_5

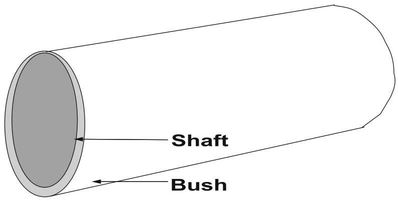


Fig. 1 Schematic representation of the bearing

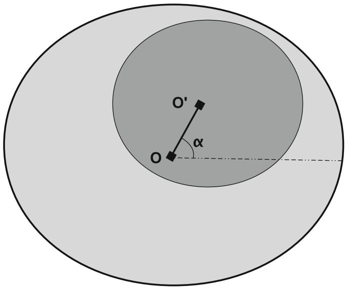


Fig. 2 A cross of the bearing

where α is the angle of vector $\overrightarrow{OO'}$ ($\overrightarrow{OO'} = (e\cos\alpha, e\sin\alpha)$) (Fig. 2).

When considering the non-coincidence of the axes of the cylinders (i.e., $\eta > 0$), overpressure areas and low pressure areas are produced, and it results in the appearance of bubbles phenomenon known as cavitation. Mathematical models of the phenomenon of cavitation consider that the lubricant pressure satisfies the Reynolds equation in the lubricated area Ω_+

$$-\frac{1}{r_b}\frac{\partial}{\partial\theta}\Big(\frac{h^3}{12\mu r_b}\frac{\partial p}{\partial\theta}\Big) - \frac{\partial}{\partial x}\Big(\frac{h^3}{12\mu}\frac{\partial p}{\partial x}\Big) = -\frac{r_b\omega}{2}\frac{dh}{d\theta}$$

where μ is a constant viscosity coefficient, while in the rest Ω_0 (cavitated zone) the pressure vanishes. By taking $r_b = 1$, $h := \frac{h}{r_s - rb}$ and $p := \frac{(r_s - r_b)^2 p}{6\mu\omega}$ we can write the Reynolds equation in the dimensionless form

$$\frac{\partial}{\partial\theta}\Big(h^3\frac{\partial p}{\partial\theta}\Big) + \frac{\partial}{\partial x}\Big(h^3\frac{\partial p}{\partial x}\Big) = \frac{dh}{d\theta}$$

where $h(\theta) = 1 - \eta\cos(\theta - \alpha)$.

The problem to solve is written in the following form (Fig. 3):

Find the pressure $p(X)$ ($X = (\theta, x)$) and the regions Ω_+ and Ω_0 such

$$\frac{\partial}{\partial\theta}\Big(h^3\frac{\partial p}{\partial\theta}\Big) + \frac{\partial}{\partial x}\Big(h^3\frac{\partial p}{\partial x}\Big) = \frac{dh}{d\theta}, \quad p > 0 \quad \text{in } \Omega_+. \tag{2}$$

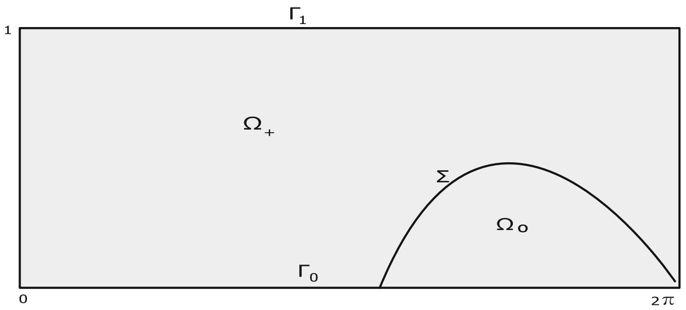


Fig. 3 The domain Ω

$$p = 0 \quad \text{in } \Omega_0. \tag{3}$$

$$p = \frac{\partial p}{\partial n} = 0 \quad \text{on } \Sigma = \Omega_0 \cap \Omega_+. \tag{4}$$

$$p(\theta, 0) = p(\theta, 1) = 0, \qquad 0 \leq \theta \leq 2\pi. \tag{5}$$

$$p(2\pi, x)) = p(0, x), \qquad 0 \leq x \leq 1. \tag{6}$$

The free boundary Σ is an additional unknown of the problem.
Equations (2)–(4) may be summarized as

$$p \geq 0 \quad \text{and} \quad p(div(h^3 \nabla p) - \frac{dh}{d\theta}) = 0 \quad \text{in} \quad \Omega.$$

Let us consider the set

$$K = \{\varphi \in H^1(\Omega),\ 2\pi - \text{periodic } \varphi(x, 0) = \varphi(x, 1) = 0 : \varphi \geq 0\}.$$

The above problem can be formulated as a variational inequality
find $p \in K$ such that

$$\int_\Omega h^3 \nabla p \cdot \nabla(\varphi - p)\, dX \geq \int_\Omega h \frac{\partial}{\partial \theta}(\varphi - p)\, dX \quad \forall \varphi \in K. \tag{7}$$

Existence and uniqueness results of (7) are known classical results. There exist a unique $p(\theta, x)$ satisfying (7) and a no empty cavity area $\Omega_0 \neq \emptyset$ (see [6, 7]).

Approximation Method

Numerical approach of inequality (7) can be placed in the frame of some approximation methods for variational inequalities based on classical results for monotone operators given in [3, 8]. For this purpose we give first the following formulation of this inequality,

find $p \in V$, such that

$$\int_{\Omega} h^3 \nabla p . \nabla(\varphi - p)\, dX + I_K(\varphi) - I_K(p) \geq \int_{\Omega} h \frac{\partial}{\partial\theta}(\varphi - p)\, dX, \forall \varphi \in V \quad (8)$$

where

$$V = \{\varphi \in H^1(\Omega), \varphi(.,0) = \varphi(.,1) = 0, \varphi \text{ is } 2\pi\text{- periodic}\}$$

and I_K is the indicatrix function of the non-empty closed and convex set K.

From the definition of the sub-differential ∂I_K we have

$$\beta \in \partial I_K(p) \iff I_K(\varphi) - I_K(p) \geq < \beta, \varphi - p >, \forall \varphi \in V.$$

It follows then from (8) that

$$\beta = Div(h^3 \nabla p) - \frac{\partial h}{\partial \theta} \in \partial I_K(p), \quad (9)$$

and we obtain this equivalent formulation of the problem,

find $p \in V$ such that

$$\int_{\Omega} h^3 \nabla p . \nabla \varphi\, dX + \int_{\Omega} \beta\varphi\, dX = -\int_{\Omega} \frac{dh}{d\theta}\varphi\, dX, \forall \varphi \in V, \quad (10)$$

$$\beta \in \partial I_K(p). \quad (11)$$

Now following [2] we introduce the multiplier $\gamma = \beta - \omega p$ with $\omega \geq 0$ and we get the formulation,

find $p \in V$ such that

$$\int_{\Omega} h^3 \nabla p . \nabla \varphi\, dX + \omega \int_{\Omega} p\varphi\, dX = -\int_{\Omega} \gamma\varphi\, dX - \int_{\Omega} \frac{\partial h}{\partial \theta}\varphi\, dX, \forall \varphi \in V \quad (12)$$

$$\gamma \in \partial I_K(p) - \omega p. \quad (13)$$

Let $\lambda > 0$, suppose that $\lambda\omega < 1$ and set $T = \partial I_K - \omega I$. We have

$$I + \lambda T = (1 - \lambda\omega)I + \lambda \partial I_K. \quad (14)$$

It can be proved that for all $f \in V$ there exists a unique $y \in V$ such that

$$f \in (I + \lambda T)(y).$$

The single-valued map

$$J_\lambda^T = (I + \lambda T)^{-1}$$

is the resolvent operator of T. And the map

$$T_\lambda = \frac{I - J_\lambda^T}{\lambda}$$

is the Moreau–Yosida approximation of T. The map T_λ is single-valued and $\frac{1}{\lambda}$-Lipschitz continuous. Moreover it satisfies the following property:

Lemma 1. *For all y and u in V, we have the equivalence property,*

$$u \in T(y) \iff u = T_\lambda(y + \lambda u). \tag{15}$$

Proof. Let $u = T_\lambda(x)$. Then

$$\begin{aligned} u = \frac{x - J_\lambda^T(x)}{\lambda} &\iff \lambda u = x - J_\lambda^T(x) \\ &\iff J_\lambda^T(x) = x - \lambda u \\ &\iff x \in (I + \lambda T)(x - \lambda u) = x - \lambda u + \lambda T(x - \lambda u) \\ &\iff \lambda u \in \lambda T(x - \lambda u) \\ &\iff u \in T(x - \lambda u) \end{aligned}$$

and by taking $x - \lambda u = y$ we get

$$u \in T(y) \iff u = T_\lambda(y + \lambda u)$$

□

Taking into account (15) in (12) and (13) we get the final formulation,
find $p \in V$ such that

$$\int_\Omega h^3 \nabla p . \nabla \varphi \, dX + \omega \int_\Omega p\varphi \, dX = -\int_\Omega \gamma\varphi \, dX - \int_\Omega \frac{dh}{d\theta}\varphi \, dX, \forall \varphi \in V \tag{16}$$

$$\gamma = T_\lambda(p + \lambda\gamma). \tag{17}$$

Iterative Algorithm

To compute the solution (p,γ) of (16) and (17) we apply the following iterative method:

(0) start with some arbitrary value of the multiplier γ_0.
(1) for γ_j known, compute p_j solution to

$$\int_\Omega h^3 \nabla p_j . \nabla \varphi \, dX + \omega \int_\Omega p_j \varphi \, dX = -\int_\Omega \gamma_j \varphi \, dX - \int_\Omega \frac{dh}{d\theta} \varphi \, dX, \forall \varphi \in V \quad (18)$$

(2) update multiplier γ_j as

$$\gamma_{j+1} = T_\lambda (p_j + \lambda \gamma_j). \quad (19)$$

(3) Go to (1) until stop criterion is reached.

Theorem 1. *For a parameter* $\lambda \geq \frac{1}{2\omega}$ *we have the convergence*

$$\lim_{j\to\infty} \|p_j - p\| = 0.$$

Proof. The mapping T_λ is $\frac{1}{\lambda}$-Lipschitz and thus

$$\begin{aligned} \|\gamma - \gamma_{j+1}\|^2 &= \|T_\lambda(p + \lambda\gamma) - T_\lambda(p_j + \lambda\gamma_j)\|^2 \\ &\leq \frac{1}{\lambda^2}\|(p + \lambda\gamma) - (p_j + \lambda\gamma_j)\|^2 \\ &= \frac{1}{\lambda^2}\|(p - p_j) + \lambda(\gamma - \gamma_j)\|^2 \\ &= \frac{1}{\lambda^2}\|p - p_j\|^2 + \frac{2}{\lambda}(p - p_j, \gamma - \gamma_j) + \|\gamma - \gamma_j\|^2. \end{aligned}$$

Therefore

$$\|\gamma - \gamma_j\|^2 - \|\gamma - \gamma_{j+1}\|^2 \geq -\frac{1}{\lambda^2}\|p - p_j\|^2 - \frac{2}{\lambda}(p - p_j, \gamma - \gamma_j) \quad (20)$$

From (12) and (18) we have

$$\int_\Omega h^3 \nabla (p - p_j) . \nabla \varphi \, dX + \omega \int_\Omega (p - p_j) \varphi \, dX = -\int_\Omega (\gamma - \gamma_j) \varphi \, dX, \quad \forall \varphi \in V.$$

And by taking $\varphi = p - p_j$ we obtain

$$\omega\|p-p_j\|^2 \leq \int_\Omega h^3|\nabla(p-p_j)|^2\,dX + \omega\int_\Omega (p-p_j)^2\,dX$$
$$= -\int_\Omega (\gamma-\gamma_j)(p-p_j)\,dX.$$

Now by substituting this inequality in (20) we obtain

$$\|\gamma-\gamma_j\|^2 - \|\gamma-\gamma_{j+1}\|^2 \geq -\frac{1}{\lambda^2}\|p-p_j\|^2 + \frac{2\omega}{\lambda}\|p-p_j\|^2$$
$$= \frac{1}{\lambda}(2\omega-\frac{1}{\lambda})\|p-p_j\|^2.$$

And if $\lambda \geq \frac{1}{2\omega}$, we get

$$\|\gamma-\gamma_j\|^2 - \|\gamma-\gamma_{j+1}\|^2 \geq 0.$$

The sequence $(\|\gamma-\gamma_j\|^2)_{j\geq 0}$ is then decreasing and positive, therefore

$$\lim_{j\to\infty}\|\gamma_j-\gamma\|^2 = 0$$

and finally

$$\lim_{j\to\infty}\|p_j-p\|^2 = 0$$

□

An Adaptive Finite Element Method

Our purpose in this section is to apply a P1-Galerkin finite element method to discretize equation (18). Let $\mathscr{T}$ be a regular triangulation of the domain into triangles and h an abstract discretization parameter. The discretization method consists in the construction of a finite-dimensional space

$$V_h := \{\psi \in \mathscr{C}(\Omega)\cap V : \forall T\in\mathscr{T} \quad \psi|_T \text{ is affine}\}.$$

The discrete solution $p_h \in V_h$ is defined as

$$\int_\Omega h^3 \nabla p_h.\nabla\varphi \, dX + \omega \int_\Omega p_h \varphi \, dX = -\int_\Omega \gamma_j \varphi \, dX - \int_\Omega \frac{dh}{d\theta}\varphi \, dX, \forall \varphi \in V_h \quad (21)$$

As mentioned before, high pressure variations occur in the bearing and cause the appearance of cavitated areas. The accuracy of discrete solution p_h depends then on the triangulation $\mathscr{T}$ in the sense that singularities and high variations of p_h have to be resolved by the triangulation. For this purpose we apply an adaptive method where triangulation $\mathscr{T}$ is improved automatically by use of a mesh-refinement where high variations occur. Moreover, the solution p_h can become smoother in some area of domain when iteration proceeds; we then remove certain elements from $\mathscr{T}$ and mesh coarsening is done.

More precisely to get a refined triangulation from the current triangulation, we first solve the equation to get the solution on the current triangulation. The error is estimated using the solution, and used to mark a set of triangles that are to be refined or coarsened. Triangles are refined or coarsened in such a way to keep regularity of the triangulations. This method is based on the following error estimator introduced by Babuska and Rheinboldt [1] and used in most works on convergence and optimality.

Theorem 3. *Given a triangulation $\mathscr{T}$ and let be p_h the solution of the discrete problem; there exists a constant $C > 0$ and an error estimator $\eta_h > 0$ depending on p_h such that*

$$\|p_j - p_h\|_{H^1(\Omega)} \leq C\eta_h.$$

Proof. Let $\varphi \in V$. We have with arbitrary $\varphi_h \in V_h$

$$\int_\Omega h^3 \nabla(p_j - p_h).\nabla\varphi \, dX + \omega \int_\Omega (p_j - p_h)\varphi \, dX =$$

$$\int_\Omega h^3 \nabla(p_j - p_h).\nabla(\varphi - \varphi_h) \, dX + \omega \int_\Omega (p_j - p_h)(\varphi - \varphi_h) \, dX$$

$$= -\int_\Omega \gamma_j(\varphi - \varphi_h) \, dX - \int_\Omega \frac{dh}{d\theta}(\varphi - \varphi_h) \, dX$$

$$- \int_\Omega h^3 \nabla p_h.\nabla(\varphi - \varphi_h) \, dX - \omega \int_\Omega p_h(\varphi - \varphi_h) \, dX$$

$$= \sum_{T\in\mathscr{T}} \Big[-\int_T \gamma_j(\varphi - \varphi_h) \, dX - \int_T \frac{dh}{d\theta}(\varphi - \varphi_h) \, dX + \int_T Div(h^3\nabla p_h)(\varphi - \varphi_h) \, dX$$

$$-\omega \int_T p_h(\varphi - \varphi_h) \, dX + \frac{1}{2}\int_{\partial T\setminus\partial\Omega} [\frac{\partial(h^3 p_h)}{\partial n}](\varphi - \varphi_h) dS\Big].$$

$$= \sum_{T\in\mathcal{T}} \Big[\int_T \Big(Div(h^3 \nabla p_h) - \omega p_h - \gamma_j - \frac{dh}{d\theta} \Big)(\varphi - \varphi_h)\, dX$$
$$+ \frac{1}{2} \int_{\partial T \setminus \partial\Omega} \Big[\frac{\partial(h^3 p_h)}{\partial n} \Big] (\varphi - \varphi_h) dS \Big].$$

We obtain two kinds of residuals, $Div(h^3\nabla p_h) - \omega p_h - \gamma_j - \frac{dh}{d\theta}$ is a point-wise residual and $\Big[\frac{\partial(h^3 p_h)}{\partial n}\Big]$ is a measure of regularity of the discrete solution.

By applying the Cauchy–Schwarz inequality we get

$$\int_\Omega h^3 \nabla(p_j - p_h).\nabla\varphi\, dX + \omega \int_\Omega (p_j - p_h)\varphi\, dX \le$$
$$\sum_{T\in\mathcal{T}} \Big[\|Div(h^3\nabla p_h) - \omega p_h - \gamma_j - \frac{\partial h}{\partial\theta}\|_T \|\varphi - \varphi_h\|_T + \|\frac{1}{2}\Big[\frac{\partial(h^3 p_h)}{\partial n}\Big]\|_{\partial T^*} \|\varphi - \varphi_h\|_{\partial T^*} \Big]$$

where $\partial T^* = \partial T \setminus \partial\Omega$. We now chose $\varphi_h = C_h\varphi$ with Clement interpolation operator $C_h : V \to V_h$ (see [4]) which verifies the interpolation estimate,

$$\|\varphi - C_h\varphi\|_T + d_T^{1/2} \|\varphi - C_h\varphi\|_{\partial T} \le C d_T \|\nabla\varphi\|_{\Omega_T}$$

with Ω_T denoting the set of neighboring elements of T and d_T its diameter.

It then follows

$$\int_\Omega h^3 \nabla(p_j - p_h).\nabla\varphi dX + \omega \int_\Omega (p_j - p_h)\varphi dX$$
$$\le \sum_{T\in\mathcal{T}} C\Big[d_T \|Div(h^3\nabla p_h) - \omega p_h - \gamma_j - \frac{\partial h}{\partial\theta}\|_T + \frac{d_T^{1/2}}{2} \|\Big[\frac{\partial(h^3 p_h)}{\partial n}\Big]\|_{\partial T^*} \Big] \|\nabla\varphi\|_{\Omega_T}$$
$$\le \sum_{T\in\mathcal{T}} 2C\Big[d_T^2 \|Div(h^3\nabla p_h) - \omega p_h - \gamma_j - \frac{\partial h}{\partial\theta}\|_T^2 + \frac{d_T}{4} \|\Big[\frac{\partial(h^3 p_h)}{\partial n}\Big]\|_{\partial T^*}^2 \Big]^{1/2} \|\nabla\varphi\|_{\Omega_T}$$
$$\le C\Big(\sum_{T\in\mathcal{T}} \eta_T^2 \Big)^{1/2} \Big(\sum_{T\in\mathcal{T}} \|\nabla\varphi\|_{\Omega_T}^2 \Big)^{1/2} \le C\eta_h \|\nabla\varphi\|.$$

where $\eta_T^2 = \Big[4d_T^2 \|Div(h^3\nabla p_h) - \omega p_h - \gamma_j - \frac{\partial h}{\partial\theta}\|_T^2 + d_T \|\Big[\frac{\partial(h^3 p_h)}{\partial n}\Big]\|_{\partial T^*}^2 \Big]$. We have used in the last step that the number of neighbors of any triangle T is bounded due to the uniform shape-regularity of the meshes. Taking $\varphi = p_j - p_h$ we obtain the global upper bounded

$$\min((1-\eta)^3, \omega) \|p_j - p_h\|_{H^1(\Omega)} \le C\eta_h,$$

with the residual-based error estimator $\eta_h = (\sum_{T\in\mathcal{T}} \eta_T^2)^{1/2}$ □

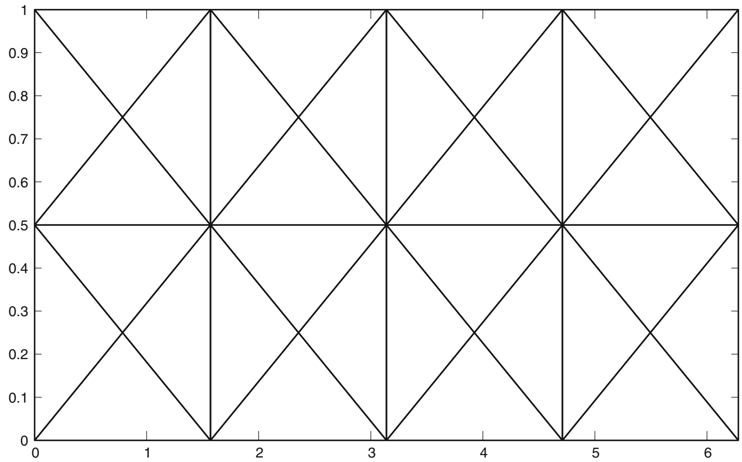


Fig. 4 Initial triangulation $\mathscr{T}_0$

Numerical Results

For numerical simulations we take $\omega = 0.1$ and $\lambda = 5$ so approximation condition (14) and theorem 2 convergence hypothesis are satisfied. We start algorithm (18) and (19) with $\gamma_0 = 0$ and take triangulation given in Fig. 4 as initial mesh for the adaptive finite element method in the first step.

At each step $j+1$ we take the final triangulation $\mathscr{T}_j$ obtained in the step j as initial triangulation for the adaptive method.

Given a triangulation $\mathscr{T}$, if the error estimator $\eta_h < \delta$, where δ is a fixed tolerance, the corresponding solution p_h is accepted and the adaptive method stopped. Otherwise, we apply the Dorfler criterion [5] to mark elements $T \in \mathscr{T}$ for refinement. This criterion seeks to determine the minimal set $\mathscr{M} \subset \mathscr{T}$ such that

$$\theta(\sum_{T\in\mathscr{T}} \eta_T^2) \leq \sum_{T\in\mathscr{M}} \eta_T^2,$$

for some parameter $\theta \in]0, 1[$. We fixed $\theta = 0.5$.

For coarsening we make elements $T \in \mathscr{T}$ such that $\eta_T^2 < \sigma \dfrac{\delta^2}{N}$; where $\sigma \in]0, 1[$, δ is a fixed tolerance and N_T the number of nodes of triangulation $\mathscr{T}$. We take here $\sigma = \delta = 1$. A stop criterion ϵ is fixed also for the algorithm (18) and (19). We summarize the method in the diagram shown below, and we give numerical simulations for $\alpha = \pi$ and $\alpha = 0$ in Figs. 5 and 6 respectively.

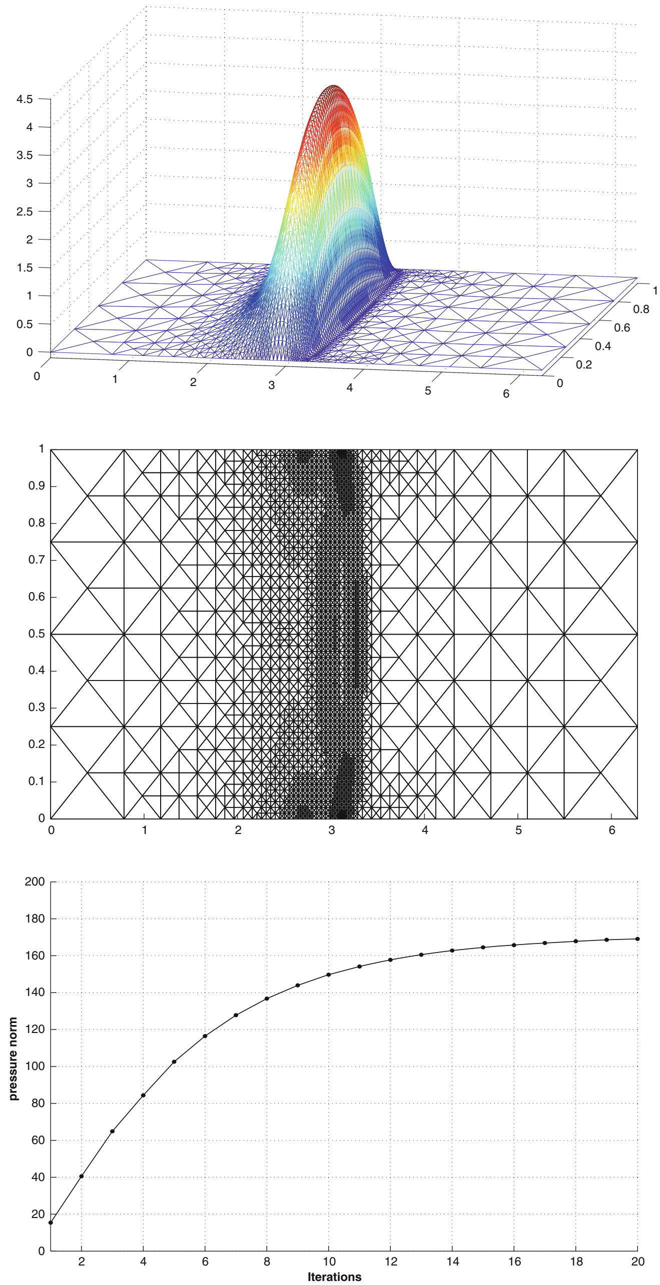


Fig. 5 Pressure p_h, the final mesh $\mathcal{T}_h$ and the pressure norm $\|p_j\|$ evolution, $(1 \leq j \leq 20)$, for $\alpha = \pi$

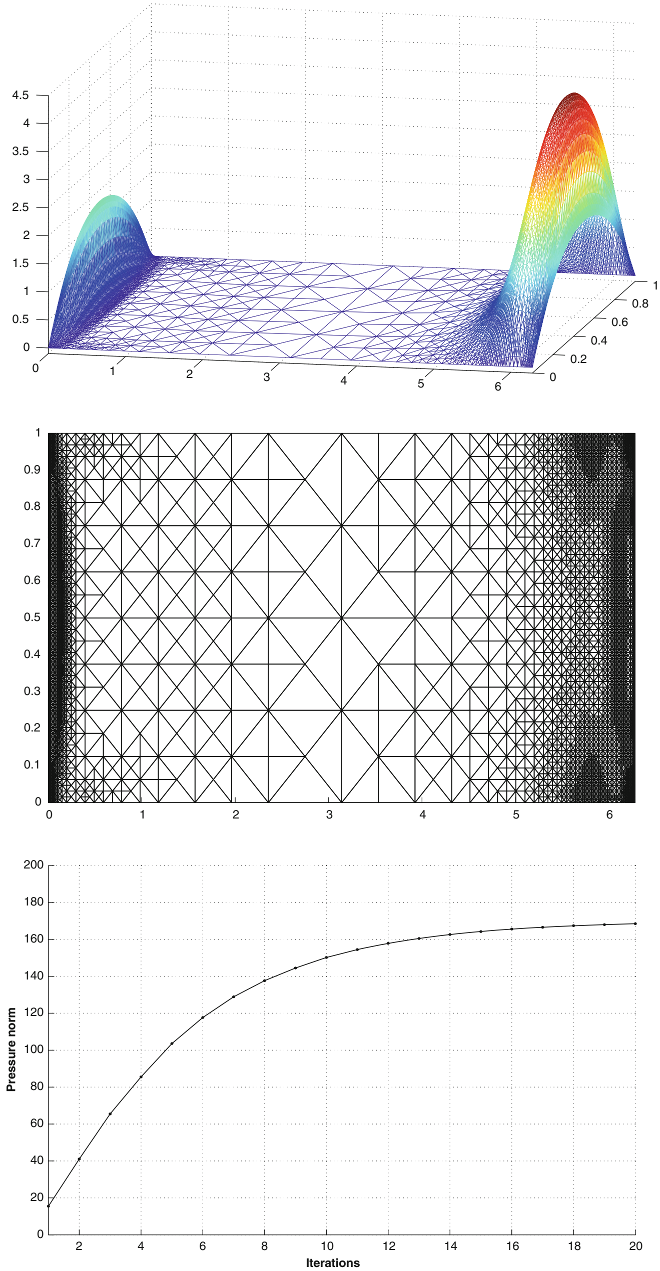
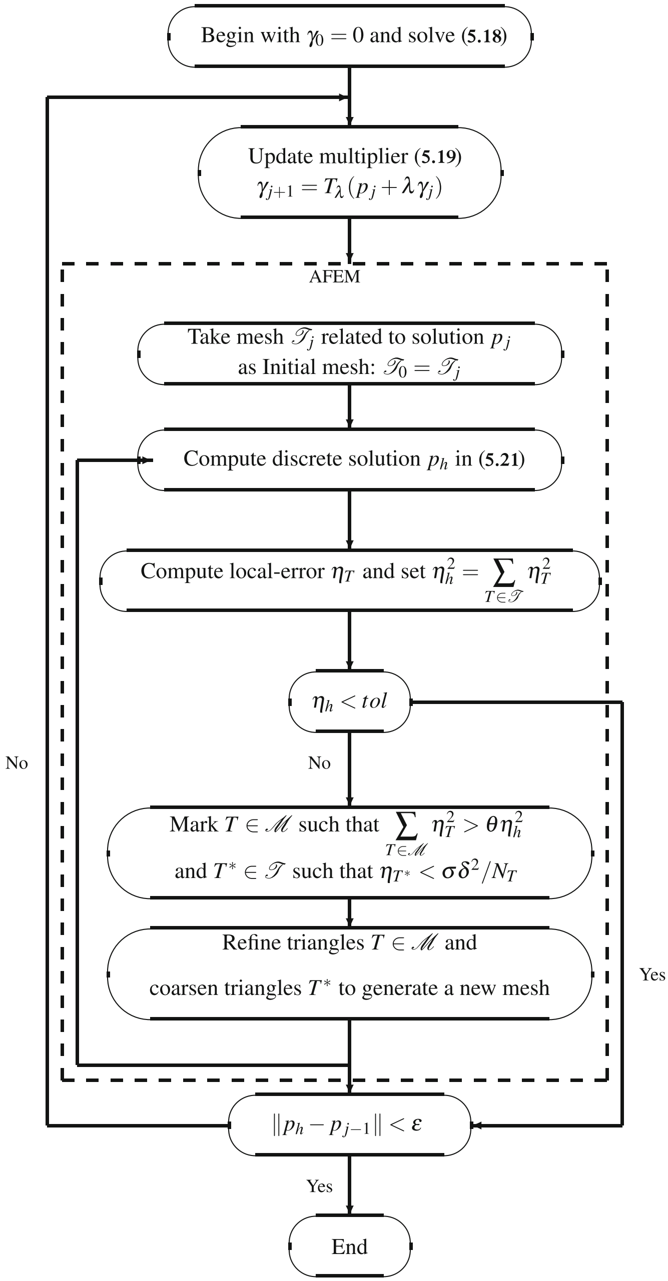


Fig. 6 Pressure p_h, the final mesh $\mathscr{T}_h$ and the pressure norm $\|p_j\|$ evolution $\|p_j\|$, $(1 \leq j \leq 20)$, for $\alpha = 0$


Begin with $\gamma_0 = 0$ and solve (5.18)
Update multiplier (5.19)
$\gamma_{j+1} = T_\lambda(p_j + \lambda\gamma_j)$
AFEM
Take mesh $\mathscr{T}_j$ related to solution p_j
as Initial mesh: $\mathscr{T}_0 = \mathscr{T}_j$
Compute discrete solution p_h in (5.21)
Compute local-error η_T and set $\eta_h^2 = \sum_{T\in\mathscr{T}} \eta_T^2$
$\eta_h < tol$
No
Mark $T \in \mathscr{M}$ such that $\sum_{T\in\mathscr{M}} \eta_T^2 > \theta\eta_h^2$
and $T^* \in \mathscr{T}$ such that $\eta_{T^*} < \sigma\delta^2/N_T$
Refine triangles $T \in \mathscr{M}$ and
coarsen triangles T^* to generate a new mesh
Yes
No
$\|p_h - p_{j-1}\| < \varepsilon$
Yes
End

References

1. Babuska, I., Rheinboldt, W.: Error estimates for adaptive finite element computations. SIAM J. Numer. Anal. **15**, 736–754 (1978)
2. Bermudez, A., Moreno, C.: Duality methods for solving variational inequalities. Comput. Methods Appl. Mech. Eng. **75**, 455–466 (1989)
3. Brezis, H.: Operateurs Maximaux Monotones et Semigroups de Contractions dans les Espaces de Hilbert. North-Holland, Amterdam (1973)
4. CLE MENT, P.: Approximation by finite element functions using local regularization. R.A.I.R.O. Anal. Numer. **9**, 77–84 (1975)
5. Dorfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
6. Ekeland, I., Teman, R.: Analyse Convexe et Problèmes Variationelles. Gautier-Villars, Paris (1984)
7. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic, New York (1980)
8. Pazy, A.: Semigroups of non-linear contractions in Hilbert Spaces In: Problems in Nonlinear Analysis. C.I.M.E., Ed. Cremonese, Roma (1971)

Evolution Solutions of Equilibrium Problems: A Computational Approach

Monica-Gabriela Cojocaru and Scott Greenhalgh

Abstract This paper proposes a computational method to describe evolution solutions of known classes of time-dependent equilibrium problems (such as time-dependent traffic network, market equilibrium or oligopoly problems, and dynamic noncooperative games). Equilibrium solutions for these classes have been studied extensively from both a theoretical (regularity, stability behaviour) and a computational point of view.

In this paper we highlight a method to further study the solution set of such problems from a dynamical systems perspective, namely we study their behaviour when they are not in an (market, traffic, financial, etc.) equilibrium state. To this end, we define what is meant by an evolution solution for a time-dependent equilibrium problem and we introduce a computational method for tracking and visualizing evolution solutions using a projected dynamical system defined on a carefully chosen L^2-space. We strengthen our results with various examples.

Keywords Projected dynamics • Time-dependent equilibrium problems • Dynamic adjustment • Applications

Introduction

Equilibrium problems have been studied over four decades and their formulation has become fairly complex. Most papers in the literature to date are concerned with the equilibrium states of these problems, where *equilibrium* is defined depending on the context of the problem: Wardrop [5, 15, 16, 27], Nash-Cournot [10, 11, 21, 27], market [4, 20, 27], physical/mathematical equilibrium [1, 2, 23, 25] and many others. The classes of equilibrium problems have been shown to be equivalent to different

M.-G. Cojocaru (✉)
Department of Mathematics & Statistics, University of Guelph, Guelph, ON, Canada N1G 2W1
e-mail: mcojocar@uoguelph.ca

S. Greenhalgh
Department of Mathematics & Statistics, Queen's University, Kingston, ON, Canada
e-mail: scottyg3@gmail.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_6

types of variational inequality problems both finite- and infinite-dimensional, as is the case in all cited papers above. Variational inequalities are well-studied mathematical constructs that allow substantial qualitative studies of equilibrium states of various problems. In their finite-dimensional form, they provide a (collection of) solution(s) to an equilibrium problem.

Time-dependent equilibrium problems (TDEP) (see, for instance, [1, 16, 19]) are those whose data depend on a parameter, usually taken to mean physical time. For instance, in a traffic network problem, the flows on routes will be time-dependent, and thus the demands on those routes can be thought of varying with time. In a game setting (see [10, 11]), strategies, as well as players' utilities or payoffs can be time-dependent. Infinite-dimensional variational inequalities provide a mathematical framework for studying equilibrium states of TDEP which are described as a succession of equilibrium states over a time interval of interest, generically denoted by $t \in [0, T]$.

Much studies have been devoted to identifying and computing—theoretically and algorithmically—the equilibrium states of such problems (see for instance [18, 22] and the references therein). We argue here that the continual drive for finding equilibria is an incomplete venue of investigation in the class of TDEP since such privileged states are not necessarily observed. For instance, in a classic traffic network equilibrium problem the non-equilibrium states are given by flow patterns on the network that do not satisfy the Wardrop-type conditions. To complete the investigation of TDEP, we propose to study their evolving states.

To this end, we associate to a TDEP a nonsmooth (projected) differential equation on a subset of the space $L^2([0, T], \mathbb{R}^k)$ whose stationary solutions also describe market equilibria, traffic equilibria, Nash/Cournot equilibria, etc. [8–11, 14]. Consequently, any non-stationary solution of the projected equation represents a time-dependent, evolving state of the underlying problem.

In this paper we show how non-equilibrium states of a TDEP can be computed with the help of the associated projected differential equation, and with the help of a selection of initial conditions for this equation. While it is known that, in theory, the projected equation on a specific type of Lebesgue integrable function space describes the dynamic evolution of an initial state of a problem towards later states at $t > 0$ (see [8, 12] where this was first introduced as double layer dynamics), there is no formalized approach to compute and visualize these evolutions in the literature to date.

The organization of this paper is as follows: We start by recalling several important concepts in section "PrDE." In section "TDEP," we briefly review (TDEPs) and their relation to projected differential equations. Section "Evolution Solutions of TDEP" outlines our computational method and section "Examples and Discussion" shows the method applied to several types of TDEP. We close with a few concluding remarks.

Preliminaries

PrDE

We start by recalling several concepts which are related to monotonicity. Let X be a Hilbert space of arbitrary dimension, let $\mathbb{K} \subset X$ be a non-empty, closed and convex set and let $F : \mathbb{K} \to X$ a mapping.

Definition 1. The problem of finding a point $x^* \in \mathbb{K}$ so that

$$\langle F(x^*), y - x^* \rangle \geq 0, \ \forall y \in \mathbb{K}$$

is called a **variational inequality problem**. The set of points $x^* \in \mathbb{K}$ satisfying the inequality above is the solution set of the variational inequality, denoted by $SOLVI(F, \mathbb{K})$.

Since their introduction in the 1960s [6, 26], VI problems have been extensively used in the study of Wardrop, Nash, Walras, Cournot and mathematical physics equilibrium problems. A classic result on the existence of solutions x^* as in Definition 1 is that of a compact set $\mathbb{K}$ and a continuous mapping F [25].

Definition 2. The **tangent cone** to a point $x \in \mathbb{K}$ is defined to be

$$T_{\mathbb{K}}(x) = \overline{\bigcup_{\delta > 0} \frac{1}{\delta} (\mathbb{K} - x)}.$$

In general a VI problem can be associated with the following differential equation (see [1, 12, 17, 27]):

Definition 3. Consider the VI problem of Definition 1. The discontinuous ordinary differential equation,

$$\frac{dx(\tau)}{d\tau} = P_{T_{\mathbb{K}}(x(\tau))}(-F(x(\tau))) := \Pi_{\mathbb{K}}(x(\tau), -F(x(\tau))), \quad x(0) \in \mathbb{K}, \tag{1}$$

where $P_{T_{\mathbb{K}}(z)}$, $z \in X$ is the closest element mapping from X to the set $T_{\mathbb{K}}(z) \in X$, is called a **projected differential equation** (PrDE).

It is known that Eq. (1) has unique solutions $x \in AC(\mathbb{R}_+, \mathbb{K})$ for each initial point $x(0) \in \mathbb{K}$ [1, 12, 13, 17]. Moreover, this equation has the property (see [1, 12, 13, 17] for proofs):

$$x^* \in SOL(VI(F, \mathbb{K})) \Leftrightarrow \Pi_{\mathbb{K}}(x^*, -F(x^*)) = 0,$$

That is, a solution of the VI problem is a stationary (or equilibrium) point of the PrDE and vice versa.

TDEP

A time-dependent equilibrium problem is an equilibrium problem whose data (variables, constraints) depend on time, t, taken to mean physical time. A state variable of such a problem is typically denoted by $x \in \mathbb{K}$, where $\mathbb{K}$ is traditionally thought of as a subset of $X := L^p([0,T],\mathbb{R}^q)$. As t varies over $[0,T]$, the data of the equilibrium problem vary with t and thus a state x is taken to belong to a constraint (feasible) set $\mathbb{K}$ given by

$$\mathbb{K} = \Big\{x \in L^p([0,T],\mathbb{R}^q) \mid \lambda(t) \le x(t) \le \mu(t),\ A(t)x(t) = \rho(t),\ \text{for a.a. } t \in [0,T]\Big\}, \tag{2}$$

where $\lambda, \mu \in L^p([0,T],\mathbb{R}^q)$, $A \in L^p([0,T],\mathbb{R}^{l\times q})$ and $\rho \in L^p([0,T],\mathbb{R}^l)$. Such a set $\mathbb{K}$ is closed, convex and bounded in X, provided that good conditions on λ, μ, A and ρ are considered.

There is a wide array of TDEPs (in the sense defined in this section) whose constraint sets amount to a subset of $L^p([0,T],\mathbb{R}^q)$ of the form (2) (models of time-dependent traffic network problems, spatial equilibrium problems, financial equilibrium problems, as well as oligopolies and dynamic games: [4, 10, 11, 16, 21] and the references therein).

It is also known that steady states $x^* \in \mathbb{K}$ of TDEP can be found by formulating the problem as an evolutionary variational inequality (EVI) (see [3, 16, 19]):

$$\text{find } x^* \in \mathbb{K} \text{ such that } \int_0^T \langle F(t,x^*(t)), v(t) - x^*(t)\rangle \ge 0,\ \ \forall v \in \mathbb{K}, \tag{3}$$

where $F : [0,T] \times \mathbb{K} \to (L^p([0,T],\mathbb{R}^q))^*$ is taken to be Lipschitz continuous with respect to the state variable belonging to $\mathbb{K}$; or equivalently as the variational problem:

$$\text{for a.a. } t \in [0,T], \text{find } x^*(t) \in \mathbb{K}(t) \text{ such that } \quad \langle F(t,x^*(t)), v(t) - x^*(t)\rangle \ge 0, \quad \forall v(t) \in \mathbb{K}(t). \tag{4}$$

The questions of existence, uniqueness and regularity of solutions to problems (3) and (4) are studied in detail in several works (see, for instance, [3, 4, 16, 19] and the references therein).

As in the case of a generic VI problem (see Definition 1 of section "Preliminaries"), we can associate a projected differential equation with an EVI problem (3) for $p = 2$:

$$\frac{dx(\cdot,\tau)}{d\tau} = P_{T_{\mathbb{K}}(x(\cdot,\tau))}(-F(x(\cdot,\tau))),\ \ x(\cdot,0) \in \mathbb{K} \subset L^2([0,T],\mathbb{R}^q), \tag{5}$$

where a solution of this equation is given by absolutely continuous functions $[0,\infty) \ni \tau \mapsto x(\cdot,\tau) \in \mathbb{K}$ (see []) with $\mathbb{K}$ of the form (2).

In [14] the association between a PrDE and an EVI as above is called **double layered dynamics** (DLD). Results on existence, uniqueness and stability of solutions to DLD (5) may be found in [8, 9, 14].

Evolution Solutions of TDEP

In this section we present a new framework for modeling evolution solutions of TDEPs. Unless otherwise mentioned, we consider a TDEP modeled on a constraint set as in (2) (with a possible modification as in Corollary 1 below).

In general, in the class of TDEP, much use is being made of existence results that ensure uniqueness of equilibrium states at almost all time moments $t \in [0, T]$. It is easy to see why, since most TDEP require a computational approach for finding approximates of their equilibrium states. In order to ensure that point-wise computed equilibria [typically using a formulation based on (4)] are meaningful to interpolate, uniqueness of point-wise states is required (see, for instance, [3, 14, 16, 27], etc.). Consequently, the current literature is concerned with imposing monotonicity type or pseudo-monotonicity conditions (strict or strong) specifically to ensure that a TDEP has a unique curve x^* of equilibrium states.

Given that we are not computing equilibrium states of TDEP we have no need for such restrictions, and in fact we highlight (as we did in previous works in related contexts [7, 24]) that dropping monotonicity-type conditions makes the dynamics of a problem much richer, but not complex enough to be intractable. Using the computational method which we present below in section "Computing Evolution Trajectories" completely removes the need to impose uniqueness of point-wise steady states and opens up the possibilities of finding perhaps several curves of steady states (i.e., non-unique points in the solution set of an EVI problem), periodic behaviour (evolution trapped in a periodic cycle, which has no counterpart in any VI or EVI model), or simply finding an estimate of the evolution of a particular initial state of TDEP into a later one. This is a completely novel approach as are the examples that we present to illustrate the method.

Evolution Solutions of TDEP

Definition 4. An evolution solution of a TDEP is $x(\cdot,\cdot) \in AC([0,\infty), \mathbb{K})$ such that $\frac{dx(\cdot,\tau)}{d\tau} = P_{T_{\mathbb{K}}(x(\cdot,\tau))}(-F(x(\cdot,\tau))) \neq 0$ for at least some time interval $[0,\tau]$, $\tau > 0$.

We remark that given an evolution solution x as in Definition 4, for each arbitrarily fixed $\tau \in [0,\infty]$, the point $x(\cdot,\tau) \in \mathbb{K}$. Thus, for each such τ, the mapping $t \mapsto x(t,\tau)$ describes the evolution of states for $t \in [0, T]$ at τ. We note here that Eq. (5) can be solved, as long as an initial distribution of flows $x(\cdot, 0)$ is given.

However, in applications, identifying an initial distribution of states of the problem over the interval $[0, T]$ cannot be realistically expected. Thus we show next that the amount of initial information regarding states of a modelled equilibrium problem can be significantly small, without imposing an obstacle to simulating and testing its evolution from these states into further ones.

Proposition 1. *Assuming the following properties on the functions λ, μ, A, ρ [introduced in (2)] are satisfied:*

1. *λ is convex and μ is concave in t;*
2. *A is a constant coefficient matrix, ρ is linear in t,*

then there exists an initial distribution of states $x(t,0) \in \mathbb{K}$ extrapolated from a finite number of discrete observation points $\{x(0,0), x(t_1,0), \dots, x(t_k = T, 0)\}$ over the interval of interest $[0,T]$ given by

$$x(t,0) := \begin{cases} \frac{t_1-t}{t_1}x(0,0) + (1-\frac{t_1-t}{t_1})x(t_1,0), & t \in [0,t_1] \\ \frac{t_2-t}{t_2-t_1}x(t_1,0) + (1-\frac{t_2-t}{t_2-t_1})x(t_2,0), & t \in (t_1,t_2] \\ \dots\dots.. & \\ \frac{T-t}{T-t_{k-1}}x(t_{k-1},0) + (1-\frac{T-t}{T-t_{k-1}})x(T,0), & t \in (t_{k-1},T] \end{cases}$$

with the property that the mapping $[0,T] \ni t \overset{x(\cdot,0)}{\mapsto} x(t,0)$ belongs to $\mathbb{K}$.

Proof. It is a quick check to see that by denoting $\gamma_i := \frac{t_i-t}{t_{i+1}-t_i} \in [0,1]$, for any $t \in (t_i, t_{i+1}]$, the function $x(t,0)$ is a convex combination of the end values $x(t_i,0)$, $x(t_{i+1},0)$.

Let us now fix $t \in [0,T]$ arbitrarily. Without loss of generality, we can consider $t \in (t_i, t_{i+1}]$ for some $i \in \{0, \dots, k\}$. We can now check that

$$x(t,0) \in \mathbb{K}(t) := \{w \in \mathbb{R}^q \mid \lambda(t) \le w \le \mu(t), A(t)w = \rho(t)\},$$

using the following facts: $x(t_i,0) \in \mathbb{K}(t_i)$, $x(t_{i+1},0) \in \mathbb{K}(t_{i+1})$ and λ, μ, A, ρ are as in the hypothesis.

We thus have

$$\lambda(t_i) \le x(t_i,0) \le \mu(t_i),\ A(t_i)x(t_i,0) = \rho(t_i) \text{ and}$$

$$\lambda(t_{i+1}) \le x(t_{i+1},0) \le \mu(t_{i+1}),\ A(t_{i+1})x(t_{i+1},0) = \rho(t_{i+1})$$

Since λ, μ are, respectively, convex and concave in t, then

$$\lambda(\gamma_i x(t_i,0) + (1-\gamma_i)x(t_{i+1},0)) \le x(t,0) \le \mu(\gamma_i x(t_i,0) + (1-\gamma_i)x(t_{i+1},0)),$$

which gives, given the expression of γ_i above, the following:

$$\lambda(t) \leq x(t,0) \leq \mu(t).$$

Similarly, assuming $A(t) = constant$ for any t, then $A(t_i) = A(t_{i+1}) =: A$ and so

$$\gamma_i A x(t_i, 0) + (1-\gamma_i) A x(t_{i+1}, 0) = \gamma_i \rho(t_i) + (1-\gamma_i)\rho(t_{i+1}) \implies Ax(t,0) = \rho(t),$$

by virtue of required linearity of ρ.

Corollary 1.

1) *Assuming the hypothesis of Proposition 1, but considering problems defined on a set $\mathbb{K}$ where equality constraints are replaced with inequality constraints (i.e., $A(t)x(t) \leq \rho(t)$) then the conclusion of Proposition 1 holds for functions ρ which are concave on* $[0, T]$.
2) *We can relax the linearity requirement of ρ in case of equality constraint by requiring a piecewise linear function, possibly with discontinuities at finite number of points.*

Lats but not least, it seems that requiring some form of linearity on the right-hand side of constraint of $\mathbb{K}$ might be a big restriction on the problem. However, in practice we might not have a priori knowledge of the shape or properties of this function, so, when it comes to estimating an initial condition $x(t,0)$ from a number of discrete points, we may benefit from using the information on discrete values of $\rho(t)$ at the observation times. Then we can also estimate a shape of $\rho(t)$ consistent with the values of $x(t,0)$. We illustrate this point in Examples 1, 3 in section "Examples and Discussion."

Computing Evolution Trajectories

We present next our computational approach for tracking the evolution of a TDEP, starting from an initial curve $x(t,0)$, given, for instance, by a procedure highlighted in Proposition 1.

Assumptions.

1) We consider the problem (5) where F is Lipschitz in x with constant b, for all $t \in [0, T]$.
2) Assume there exists $t \mapsto x(t,0)$ a function in $\mathbb{K}$, which is considered the initial condition of the problem (5).
3) Let $\Delta : \{0 \leq t_1 \leq t_2 \leq \ldots \leq T\}$ a division of $[0, T]$.

Step 1: Evaluate $x(t,0)$ at points $t_k \in \Delta$; Set the evolution time step for the projected equation to be $\tau \leq \frac{l}{m}$, where $l := \frac{L}{||F(x_0(\cdot,0))||_{L^2} + bL}$, for a given $L > 0$, the radius of a ball $B(x(\cdot,0), L) \in L^2([0,T], \mathbb{R}^q)$, and a given m to represent the number of steps covering the interval $[0, l]$.

Step 2: For each $t_k \in \Delta$, generate at every time step $\delta := n\tau$, $1 < n \leq m$, $n \in \mathbb{Z}_+$ a new set of points (of a new curve):

$$x(t_k, \delta) := P_{\mathbb{K}(t_k)}(x(t_k, \delta - \tau) - \tau F(x(t_k, \delta - \tau)).$$

Step 3(optional): Check if two consecutive sets of x values are close enough for a given tolerance TOL, i.e., check

$$||x(\cdot, \delta) - x(\cdot, \delta - \tau)||_\infty := \max_k \{|x(t_k, \delta) - x(t_k, \delta - \tau)|\} \leq TOL \tag{6}$$

Step 4: If inequality (6) holds true, then STOP. Otherwise, go back to **Step 2**.

Remark 1.

1. The step size τ is estimated according to the constructive proof from [12], adapted to the present context;
2. The computation of the projection of points in **Step 2** can be implemented in multiple ways, as is the minimization of distance between the points $x(t_k, \delta - \tau) - \tau F(x(t_k, \delta - \tau)$ and the set $\mathbb{K}(t_k)$; the modeler has a wide choice for computing these points.
3. Generally, **Step 3** is not necessary if all is needed is to track evolution for a finite number of steps τ.

   But **Step 3** may be of interest in some cases, since it detects potential equilibrium states arising from evolution from an initial condition in finite τ time (see conditions and examples of finite time evolution in [8, 9, 14]).

   In this paper, we use it to check the viability of the proposed computational method: we check if the method above is able to find equilibria of TDEP, in known examples, where these equilibria are point-wise unique, global, finite-time attractors (in τ), thus they should be detected from any initial conditions of (5). When Step 3 is implemented we set our $TOL = 10^{-5}$.

Examples and Discussion

Example 1. We first solve a test TDEP problem, with known dynamic behaviour, to show that the proposed method finds the known equilibrium solutions of this TDEP. Since the TDEP problem satisfies very strong conditions, the dynamic behaviour of its states is relatively simple, namely, we expect all initial states $x(t, 0) \in \mathbb{K}$ to evolve, over a finite amount of time steps τ, to the unique equilibrium curve $x^* \in \mathbb{K}$.

This example is presented in detail in [14]. Consider a transportation network with one OD pair, with two links; the flow on links is, respectively, $x_1(t), x_2(t)$, $t \in [120]$ so that

$$\mathbb{K} = \{x \in L^2([0, 120]), \mathbb{R}^2) \mid (0, 0) \leq x(t) \leq (100, 100),\ x_1(t) + x_2(t) = \rho(t) \text{ a.e.}\}$$

and the cost on the links is $F = (2x_1(t)+x_2(t)+1, x_1(t)+x_2(t)+2)$. It is shown that F is strongly monotone with degree $\alpha = 1$ and as such the EVI model (4) of this network admits a unique, globally, finite-time attracting curve of equilibria given by $x^*(t) = (1, \rho(t) - 1)$, with ρ a measurable function.

With this in mind, we apply our method above noticing that $\lambda(t) = (0,0)$, $\mu(t) = (100, 100)$. We assume the following observation points:

$$(x_1(t), x_2(t)) = \begin{cases} (50,20),\ t = 0 \\ (60,30),\ t = 20 \\ (35,40),\ t = 50 \\ (30,20),\ t = 70 \end{cases}$$

We assume now that the initial condition for problem (5) in this case is the piecewise continuous function:

$$(x_1(t), x_2(t)) = \begin{cases} (50,20),\ t \in [0,20] \\ (60,30),\ t \in (20,50] \\ (35,40),\ t \in (50,70] \\ (30,20),\ t \in (70,120] \end{cases}$$

Our Proposition 1 requires a linear function $\rho(t)$, or at least a piecewise one. We thus simply employ $\rho(t,0) = x_1(t,0) + x_2(t,0)$ to get values of $\{70, 90, 75, 50\}$ on the respective time subintervals. Now we note that $-F$ is Lipschitz with constant $5+\sqrt{3}$, and that $||-F(x(t,0))||_{L^2} \approx 2882$, thus taking $L = 200$ we obtain $l \approx 0.047$. Taking a number $m = 10$ of τ steps to cover $[0, l]$, we get $\tau \approx 0.0047$. Finally, we sampled the time interval [0, 120] with 13 points.

From the previous dynamic analysis of this problem in [14] we know that the equilibrium curve will be reached in a finite number n of steps τ. Our simulations below give $n \approx 200$ steps for the initial curves to reach the known equilibrium ones (Fig. 1):

Example 2. A related example to this one was first introduced in [24]. Here we adapt it to our context. Let us assume that a TDEP is given in a form of an EVI (4) where $F(t, x(t)) := (-x_1(t) + ax_2(t), -ax_1(t) - x_2(t))$, where $a \in [0, 1]$ and

$$\mathbb{K} := \{u \in L^2([0,10], \mathbb{R}^2) \mid -1 \leq x_i(t) \leq 1, \text{a.a. } t \in [0,10], i = \{1,2\}\}.$$

Here λ, μ are constant functions. The matrix $A(t)$ is the 0-matrix, so we do not have any equality constraints.

As stated above, we are interested in TDEP problems that may give rise to EVI problems whose F does not satisfy any monotonicity-type conditions. We can check right away that F above is not monotone, for any value of a (we get that $\langle F(x_1, x_2) - F(u, v), (x_1 - u, x_2 - v)\rangle = -(x_1 - u)^2 - (x_2 - v)^2$) and so all previous analyses based on variational inequalities theory are not of use here.

Let us now look at this TDEP as modelled by the projected equation (5). We have that $-F(x_1(t), x_2(t)) := (x_1(t) - ax_2(t), ax_1(t) + x_2(t))$ is linear in x, thus it is Lipschitz continuous; for instance, $b := \sqrt{1 + a^2} > 0$ is an L-constant for

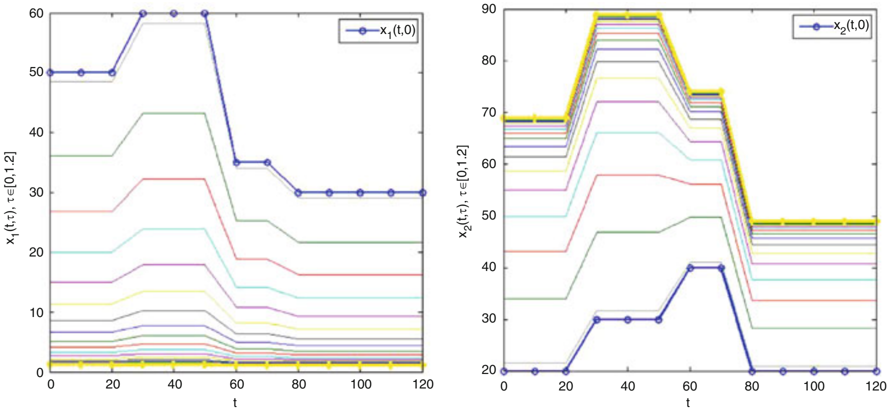


Fig. 1 We show here the evolution of the flows on this network, from initial piecewise values (*blue curve*) to the end equilibrium curve (*yellow*). *Left panel* represents $x_1(t,\tau)$ and *right panel* represents $x_2(t,\tau)$ over 200 τ steps. We see that $x_1^*(t) = 1$ constant, as predicted (*left*) where as $x_2^*(t) = \rho(t) - 1$

F. Assume now that we have a few observation data points for the states x_1, x_2 at $t \in \{0, 2, 4, 5, 7\}$ as

$$(x_1(t), x_2(t)) = \begin{cases} (0.5, 0.5), \; t = 0 \\ (0.8, 0), \; t = 2 \\ (0.9, -0.2), \; t = 4 \\ (0.4, 0.3), \; t = 5 \\ (-0.1, 0.3), \; t = 7 \end{cases}$$

These points indicate that we can consider as initial value for Eq. (5) the curve $x(t, 0)$ given by

$$(x_1(t,0), x_2(t,0)) = \begin{cases} (0.5 + 0.15t, 0.5 - 0.25t), \; t \in [0,2] \\ (0.7 + 0.05t, 0.2 - 0.1t), \; t \in (2,4] \\ (2.9 - 0.5t, -2.2 + 0.5t), \; t \in (4,5] \\ (1.65 - 0.25t, 1.8 - 0.3t), \; t \in (5,7] \\ (-0.1, 0.3), \; t \in (7,10] \end{cases}$$

This gives $|| - F(x(t,0))||_{L^2} \approx 5.43$ for $a = 1$ and 8.589 for $a = 2$.

We choose $L = 1$ and thus $l = \frac{L}{||-F(x(t,0))||_{L^2} + bL} = \frac{1}{4.75+\sqrt{1+a^2}}$. For values of $a = 1$, respectively $a = 2$ we get

$$l = 0.1461, \text{ respectively } l = 0.1304 \implies$$

for a given $m = 10$ we have

$$\tau = 0.0146, \text{respectively } \tau = 0.013.$$

We sampled the time interval [0, 10] with 21 points.

In Figs. 2 and 3 we present our simulations for $a = 1$, respectively, $a = 2$. Comparing the two rows of panels in Fig. 2, we see vastly different behaviour being displayed by the two curves $x_1(t, 0)$ and $x_2(t, 0)$ as τ increases. All four simulations have been run for 100 τ steps. Since we are unclear as to what type of dynamic behaviour we might expect, and exploiting the fact that this problem is 2-dimensional, we plotted in Fig. 3 the phase portraits of the evolutions, i.e. we eliminated t and plotted $x_1(\cdot, \tau)$ vs. $x_2(\cdot, \tau)$.

We note a much clearer picture emerging: when $a = 1$, the initial phase portrait at $\tau = 0$ represents a spiral curve ending at $(-1, 1)$ (Fig. 3, left panel, blue curve, filled in dots); it evolves into a different spiral curve ending at $(-1, -1)$ (Fig. 3, left panel, brown curve, marked with *).

However, for $a = 2$, we seem to evolve the initial curve into a periodic curve inside and around the boundary of $[-1, 1]$ (Fig. 3, right panel). This indicates that there will be periodic behaviour arising in the TDEP, specifically, that there exists $\tau_1 > \tau_2 > 0$ so that a curve $x(\cdot, \tau_2) = x(\cdot, \tau_1)$. So states in the TDEP may reoccur.

Example 3. Let us now assume that we observe at discrete moments a traffic network over a time interval [0, 90 min]. We consider that the user traffic equilibrium states are generically described as in [16], between 1 origin (*o1*) and 3 destinations (d_1, d_2, d_3). We assume the flows are functions of time $x_1(t)$, $x_2(t)$, $x_3(t)$, $t \in [0, 90\,\text{min}]$ (corresponding to pairs $(o1, d_1, o1, d_2$, respectively, $o1, d_3)$. It is known that there exist demand functions $\rho(t)$ and $\psi(t)$ on certain combination of routes, such as $x_1(t) + x_2(t) + x_3(t) = \rho(t)$ and $x_1(t) + 1/2x_2(t) = \psi(t)$. Finally, we reasonably assume that there is a lower and upper bound for flows on each route, thus $0 \le x_i(t) \le M$ for each $i \in \{1, 2, 3\}$.

We assume known that the cost of traveling on the three routes is given by the function $F : \mathbb{K} \to L^2([0, 90], \mathbb{R}^3)$, given by

$$F(x) = (3x_1 - x_2 + x_3,\ \frac{3}{2}x_1 - 4x_2 + x_3,\ 2x_1 - 7x_2 + x_3).$$

In general, such a traffic problem is modeled by an EVI problem (4) formulated on the constraint set:

$$\mathbb{K} = \{x \in L^2([0, 90], \mathbb{R}^3) \mid 0 \le x(t) \le M,\ x_1(t) + x_2(t) + x_3(t) = \rho(t),\ x_1(t) + \frac{1}{2}x_2(t) = \psi(t)\},$$

and depending on monotonicity properties of F above it would be studied from the perspective of its user equilibrium states only, given certain demand functions ρ, ψ. A simple computation (see Appendix) shows that F is monotone on $\mathbb{K}$ for any t, but it is not strictly monotone. Clearly from the Appendix, there are many points in each set $\mathbb{K}(t)$ where F is not strictly monotone. It is at these points that we cannot employ the theoretical results from EVI theory, instead we rely on the double layer

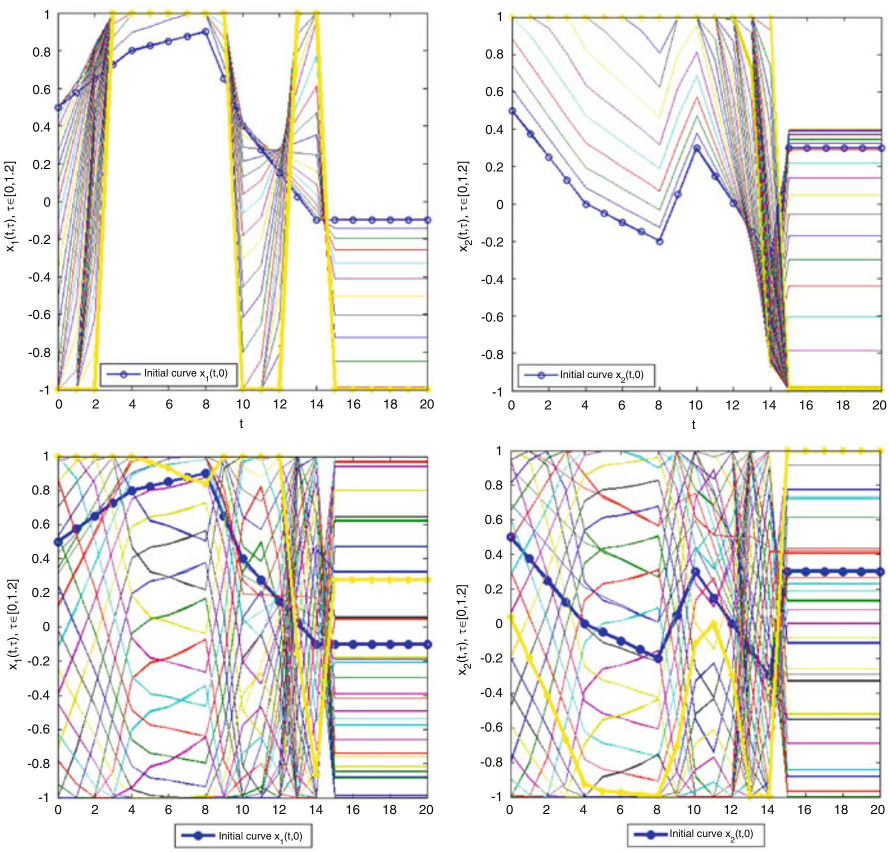


Fig. 2 The *upper two panels* display the evolution of the initial curves for $a = 1$; the *lower panels* display the evolution for $a = 2$

dynamics and evolution from initial states to study what happens to the traffic flows. We see that this example falls under the hypotheses of Proposition 1.

Moreover, given a set of discrete observation data points we extrapolate not only an initial curve of states for the traffic problem, but also demand functions ρ, ψ consistent with our observations. Unlike in the case of Example 1 above, we make here ρ, ψ linear and continuous.

We assume that we observe the traffic at 15-min intervals over a 90-min length of time, as follows:

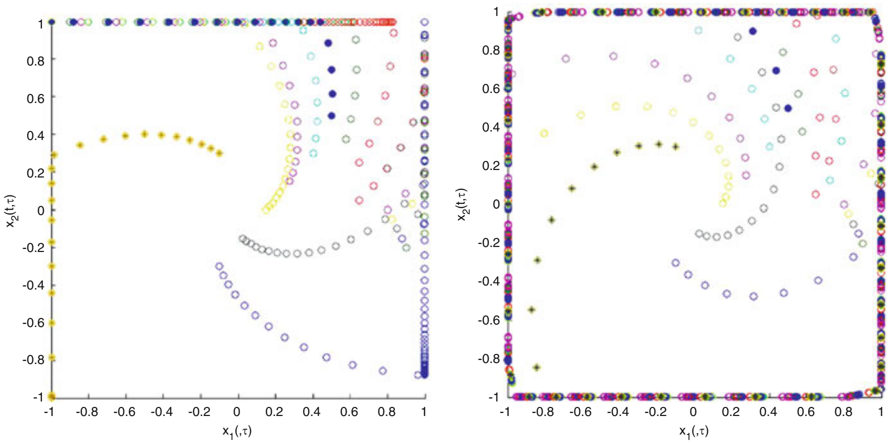


Fig. 3 Here we follow the evolution of curves in phase portrait depiction. *Left panel* has $a = 1$, *right panel* has $a = 2$

$$(x_1(t), x_2(t), x_3(t)) = \begin{cases} (10, 20, 25), \ t = 0 \\ (15, 20, 20), \ t = 15 \\ (20, 25, 25), \ t = 30 \\ (30, 10, 30), \ t = 45 \\ (30, 20, 30), \ t = 60 \\ (10, 20, 15), \ t = 75 \\ (10, 10, 10), \ t = 90 \end{cases}$$

Given the structure of the demands on certain route combinations, we then deduce

$\rho(t) = x_1(t) + x_2(t) + x_3(t))$	$\psi(t) = x_1(t) + \frac{1}{2}x_2(t)$
55, t=0	20, t=0
55, t=0	25, t=15
75, t=30	32.5 t=30
70, t=45	70, t=45
80, t=60	40, t=60
45, t=75	20, t=75
30, t=90	15, t=90

Then the initial curve of states for this transportation problem, as well as estimates for the demand functions consistent with observations, is as given by Proposition 1:

For this initial curve, we compute $\| - F(x(t, 0))\|_{L^2} \approx 2476$ and $b = 2$. For $L = 1000$ we thus get $l = \frac{L}{\|-F(x(t,0))\|_{L^2} + bL} \approx 0.22$; we thus choose $m = 10 \implies \tau :=$

$(x_1(t,0), x_2(t,0), x_3(t,0))$	$\rho(t)$	$\psi(t)$	Time
(10+1/3t, 20,25−t/3)	55	20+t/3	$t \in [0, 15]$
(10+t/3,15+t/3,15+t/3)	40+t	17.5+0.5t	$t \in (15, 30]$
(2t/3,55−t, 15+t/3)	70	0.16t+27.5	$t \in (30, 45]$
(30, −20+2t/3, 30)	40+2t/3	20+0.334t	$t \in (45, 60]$
(110−4t/3, 20,90−t)	220−7t/3	120−4t/3	$t \in (60, 75]$
(10, 70−2t/3,40−t/3)	120−t	45−0.334t	$t \in (75, 90]$

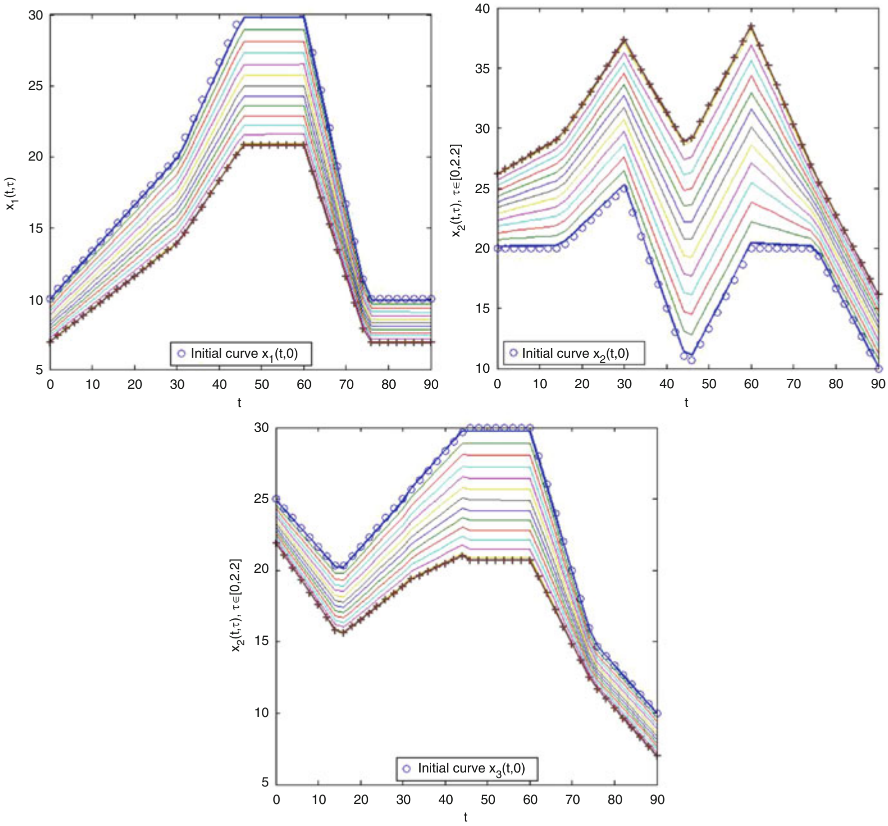


Fig. 4 We depict here the evolution of the initial states after 100 τ steps. Given that the demand functions are piecewise continuous, and given results in [3] we expect to see a piecewise continuous evolution

0.022. Here we sampled the time interval [0, 90] with 46 points. Our simulations are presented in Fig. 4 above. We see that in 100 τ steps we evolve the initial curves to later ones, depicted with a dagger sign in all panels of Fig. 4.

To see how different initial observed states behave under this dynamics, let us consider that ρ, ψ are as in the above table, and assume that we observe a constant

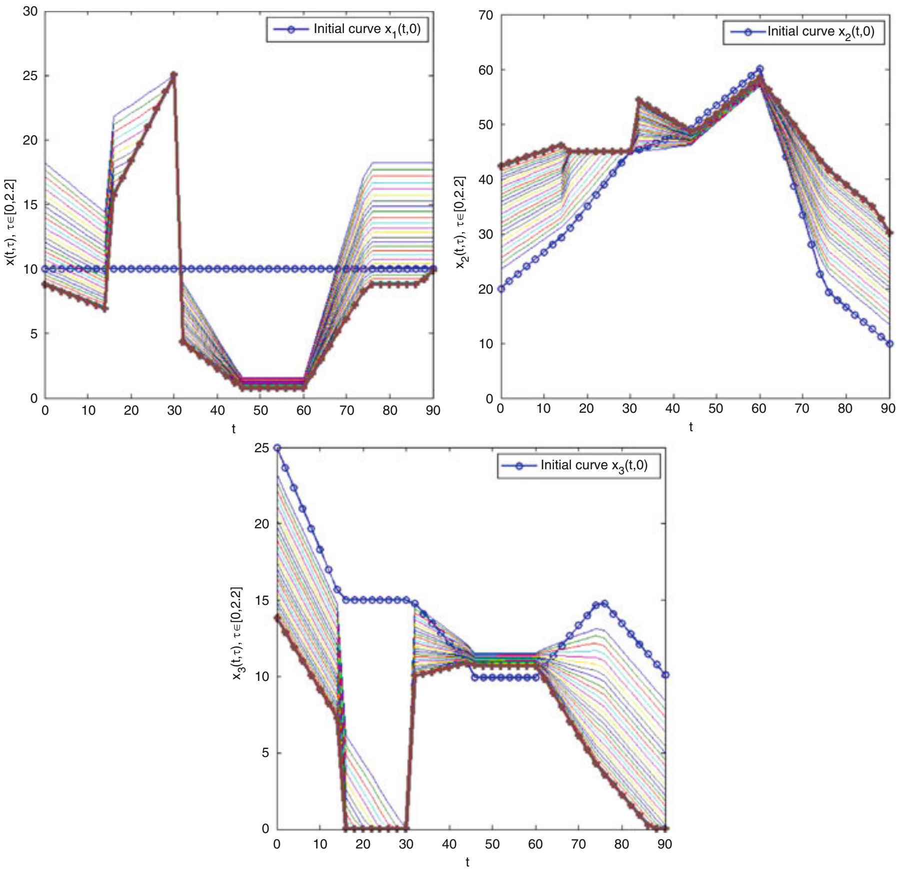


Fig. 5 We depict here the evolution of the initial states after 100 τ steps. The *dark brown curves* are the states of the problem the at end of simulated time. They are distinct than the ones obtained in Fig. 4, over the same length of simulated time

density flow $x_1(t, 0) = 10$, for any t. Then respecting the flow demands, we have $x_2(t, 0) = 2(\psi(t) - 10)$ and $x_3(t, 0) = \rho - (10 + x_2(t, 0))$. This initial condition gives the following evolution over 100 τ steps, shown in Fig. 5 above. We note that the flows have evolved in a distinct manner than for the previous initial condition of above table. This was expected, since we are not in possession of a global strict monotonicity condition for F over $\mathbb{K}$.

Conclusions and Future Work

We addressed here an open problem in the projected dynamics and EVI areas, namely that of giving a method to track evolution solutions of TDEP problems. This is by no means a complete answer to the question of how such evolution should be studied, however our main point here is that it should be studied. Our proposed method was mainly guided by reasonable assumptions on how much data one may be in possession of, when one studies a TDEP, and how one might use this information to get a picture of the changes in the states of a TDEP. The method is specifically concerned with cases of TDEP for which well-known VI theory fails to give a clear picture of the potential dynamics.

We included here examples (Examples 1 and 3) drawn from traffic network equilibrium problems, however, the method need not be restricted to this TDEP class. In Example 2, one can easily reinterpret the data as coming from a continuous time dynamic game, with $F := (-\nabla_{x_1} U_1, \nabla_{x_2} U_2)$ being a gradient of a quadratic payoff, such as $U_1 := x_1^2 - ax_1x_2 + t$, and $U_2 := ax_1x_2 - x_2^2/2 + t^2$, where each player wishes to maximize its U.

We hope that the reader will find it useful to apply and/or extend this method to specific classes of TDEP which fall under a generic formulation as variational inequality problems. One such application is currently concerning the evolution solutions of a dynamic casual encounters game. These are a sequence of repeated games between two potential sexual partners where one is $HIV+$ and the other is $HIV-$. It is sought to quantify the potential increase in infection in a homosexual population using a double layer dynamics model.

Appendix to Example 3

Recalling the mapping

$$F(x) = \begin{pmatrix} 3x_1 - x_2 + x_3 \\ \\ \frac{3}{2}x_1 - 4x_2 + x_3 \\ \\ 2x_1 - 7x_2 + x_3 \end{pmatrix},$$

with the constraint set

$$\mathbb{K}(t) = \{x \in L^2([0,90], \mathbb{R}^3) \mid 0 \le x(t), x_1(t) + x_2(t) + x_3(t) = \rho(t), x_1(t) + \frac{1}{2}x_2(t) = \psi(t)\},$$

We show below that F is monotone, but not strictly so. In what follows we make use of the following identities which can be derived from $\mathbb{K}$: for all $x, y \in \mathbb{K}$ we have that

$$x_3 - y_3 = -(x_1 - y_1) - (x_2 - y_2), \tag{7}$$

and

$$x_2 - y_2 = -2(x_1 - y_1). \tag{8}$$

We evaluate now $\langle F(y) - F(x), y - x\rangle = \sum_{i=1}^{3} F_i(x)(x_i - y_i)$ with (7), (8) yielding

$$2(x_1 - y_1)^2 \geq 0,\ \forall x \neq y \in \mathbb{K}(t).$$

References

1. Aubin, J.P., Cellina, A.: Differential Inclusions. Springer, Berlin (1984)
2. Baiocchi, C., Capelo, A.: Variational And Quasivariational Inequalities. Applicartions to Free Boundary Problems. Wiley, New York (1984)
3. Barbagallo, A., Cojocaru, M.: Continuity of solutions for parametric variational inequalities in Banach space. J. Math. Anal. Appl. **351**(2), 707–720 (2009)
4. Barbagallo, A., Cojocaru, M.-G.: Dynamic equilibrium formulation of the oligopolistic market problem. Math. Comput. Model. **49**, 966–976 (2009)
5. Beckmann, M.J., McGuire, C.B., Winsten, C.B.: Studies in the Economics of Transportation. Yale University Press, New Haven, CN (1956)
6. Brezis, H.: Inequations d'Evolution Abstraites. Comptes Rendus de l'Academie des Sciences, Paris (1967)
7. Cojocaru, M.-G.: Monotonicity and existence of periodic cycles for projected dynamical systems on Hilbert spaces. Proc. Am. Math. Soc. **134**, 793–804 (2006)
8. Cojocaru, M.-G.: Piecewise solutions of evolutionary variational inequalities. Consequences for the double-layer dynamics modelling of equilibrium problems. J. Inequal. Pure Appl. Math. **8**(2) (2007). Article 63
9. Cojocaru, M.-G.: Double-layer dynamics theory and human migration after catastrophic events. Nonlinear Analysis with applications in Economics, Energy and Transportation, Bergamo University Press, pp. 65–86 (2007)
10. Cojocaru, M.-G.: Dynamic equilibria of group vaccination strategies in a heterogeneous population. J. Glob. Optim. **40**, 51–63 (2008)
11. Cojocaru, M.-G., Greenhalgh, S.: Dynamic games and hybrid dynamical systems. Optim. Eng. Appl. Var. Inequal. Issue **13**(3), 505–517 (2011)
12. Cojocaru, M.-G., Jonker, L.B.: Existence of solutions to projected differential equations in Hilbert spaces. Proc. Am. Math. Soc. **132**(1), 183–193 (2003)
13. Cojocaru, M.-G., Pia, S.: Non-pivot and implicit projected dynamical systems on Hilbert spaces, J. Func. Spaces Appl. **2012**, (2012), 23 pp. (2011). Article ID 508570
14. Cojocaru, M.-G., Daniele, P., Nagurney, A.: Double-layered dynamics: a unified theory of projected dynamical systems and evolutionary variational inequalities. Eur. J. Oper. Res. **175**(6), 494–507 (2006)

15. Dafermos, S.: Traffic equilibrium and variational inequalities. Trans. Sci. **14**, 42–54 (1980)
16. Daniele, P.: Dynamic Networks and Evolutionary Variational Inequalities. Edward Elgar Publishing, Cheltenham (2006)
17. Dupuis, P., Nagurney, A.: Dynamical systems and variational inequalities. Ann. Oper. Res. **44**, 9–42 (1993)
18. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research and Financial Engineering. Springer, Berlin (2003)
19. Gwinner, J.: Time dependent variational inequalities - some recent trends. In: Daniele, P., Giannessi, F., Maugeri, A. (eds.) Equilibrium Problems and Variational Models, pp. 225–264. Kluwer Academic Publishers, Dordrecht (2003)
20. Harker, P.T.: A variational inequality approach for the determination of oligopolistic market equilibrium. Math. Program. **30**(1), 105–111 (1984)
21. Harker, P.T.: Generalized Nash games and quasivariational inequalities. Eur. J. Oper. Res. **54**, 81–94 (1991)
22. Harker, P., Pang, J.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
23. Isac, G.: Leray-Schauder Type Alternatives, Complementarity Problems and Variational Inequalities. Springer, Berlin (2006)
24. Johnston, M.D., Cojocaru, M.-G.: Equilibria and periodic solutions of projected dynamical systems on sets with corners. Aust. J. Math. Anal. Appl. **5**(2) (2008)
25. Kinderlehrer, D., Stampacchia, D.: An Introduction to Variational Inequalities and Their Application. Academic, New York (1980)
26. Lions, J.L., Stampacchia, G.: Variational inequalities. Commun. Pure Appl. Math. **22**, 493–519 (1967)
27. Nagurney, A., Zhang, D.: Projected Dynamical Systems and Variational Inequalities with Applications. Kluwer Academic Publishers, Boston (1996)

Cantor, Banach and Baire Theorems in Generalized Metric Spaces

Stefan Czerwik and Krzysztof Król

Abstract In the paper we present basic results for generalized metric spaces: Cantor, Banach and Baire theorems, well known and widely applied in metric spaces and mathematics.

Keywords Generalized metric space • Cantor • Banach • Baire theorems

Subject Classifications: 54E35, 46S20, 46A19

Introduction

At first we recall the idea of generalized metric space (shortly GMS) introduced by Luxemburg [5] (see also [2]). Let X be a set (nonempty). A function

$$d\colon X \times X \to [0, \infty]$$

is said to be a generalized metric on X, provided that for $x, y, z \in X$,

(i) $d(x, y) = 0$ if and only if $x = y$,
(ii) $d(x, y) = d(y, x)$,
(iii) $d(x, y) \le d(x, z) + d(z, y)$,

A pair (X, d) is called a generalized metric space.

S. Czerwik (✉)
Institute of Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland
e-mail: steczerw@gmail.com

K. Król
Education Department, Zabrze City Hall, Powstańców Śląskich 5-7, 41-800 Zabrze, Poland
e-mail: krolk@poczta.onet.pl

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_7

The specialized notation, which we use, is the same as in the relevant articles. Standard symbols: $\mathbb{N}$, $\mathbb{N}_0$ denote the set of all natural numbers and nonnegative integers, respectively. By $\overline{A}$ we denote the closure of a set A.

We shall need further on the following remark.

Remark 1. Let (X, d) be a gms and let $A \subset X$ be a set. If $x_n \in \overline{A}$, $n \in \mathbb{N}$ and $x_n \to x$ as $n \to \infty$, then $x \in \overline{A}$.

The simple proof is left to the reader.

Basic Results

At first we shall note the Cantor theorem for generalized metric spaces.

Theorem 1. *Let (X, d) be a generalized metric space. Assume that $A_n \subset X$, $n = 1, 2, \ldots$ are nonempty closed subsets such that $A_{n+1} \subset A_n$ for $n \in \mathbb{N}$ and*

$$\lim_{n\to\infty} \delta(A_n) = 0, \tag{1}$$

where

$$\delta(A_n) := \sup\{d(x, y) \colon x, y \in A_n\}. \tag{2}$$

Then there exists exactly one point $x \in X$ such that

$$\bigcap_{n=1}^{\infty} A_n = \{x\}. \tag{3}$$

Proof. Since, in view of Remark 1, the proof can be done similarly as for a metric space, we omit the details here. □

Now we shall apply the Cantor theorem to prove the Banach contraction principle for generalized metric spaces (for a different approach, see [2, 5]).

Theorem 2. *Let (X, d) be a complete generalized metric space and the function $T \colon X \to X$ satisfies the Lipschitz condition*

$$d\,(T(x), T(y)) \le \alpha d(x, y) \quad \textit{for all } x, y \in X, \tag{4}$$

where

$$0 \le \alpha < 1. \tag{5}$$

Assume that $x \in X$ is arbitrarily fixed. Then the following alternative holds: either

(A) *for every nonnegative integer* $n = 0, 1, 2, \ldots$*, one has*

$$d(T^n(x), T^{n+1}(x)) = \infty,$$

or

(B) *there exists a number* $k \in \mathbb{N}_0$ *such that*

$$d(T^k(x), T^{k+1}(x)) < \infty.$$

In the case (B):

(i) *the sequence* $\{T^n(x)\}$ *is convergent to a fixed point* $u \in X$ *of* T*, i.e.* $T(u) = u$*;*

(ii) u *is the unique fixed point of* T *in*

$$B := \{z \in X \colon d(T^k(x), z) < \infty\};$$

(iii) *for every* $z \in B$*,*

$$d(T^n(z), u) \to 0 \quad as \quad n \to \infty.$$

Proof. Put

$$\xi := d(T^k(x), T^{k+1}(x))$$

and

$$K = K(T^k(x), r) := \{z \in X \colon d(T^k(x), z) \le r\},$$

where $r \ge \frac{\xi}{1-\alpha}$.

Now we verify that $T(K) \subset K$. For if $y \in T(K)$, then $y = T(z)$ for some $z \in K$. One has

$$\begin{aligned} d(T^k(x), y) &\le d(T^k(x), T^{k+1}(x)) + d(T^{k+1}(x), T(z)) \\ &\le \xi + \alpha d(T^k(x), z) \le \xi + \alpha r \\ &\le r(1-\alpha) + \alpha r = r, \end{aligned}$$

whence $y \in K$ and consequently $T(K) \subset K$.

Clearly ($\overline{A}$ denotes the closure of A)

$$\overline{T(K)} \subset \overline{K} = K$$

and we get $\overline{T(K)} \subset K$. In the sequel

$$T^2(K) = T(T(K)) \subset T(K) \subset K$$

and also

$$\overline{T^2(K)} \subset \overline{T(K)}.$$

By the induction we get

$$\overline{T^{n+1}(K)} \subset \overline{T^n(K)} \quad \textit{for } n = 1, 2, \dots . \tag{6}$$

Define

$$\delta(T^n(K)) := \sup\{d(x, y) \colon x, y \in T^n(K)\}.$$

Obviously $\delta(K) \leq 2r$. We shall verify that

$$\delta(\overline{T(K)}) \leq 2\alpha r. \tag{7}$$

To this end let $x, y \in T(K)$, then $x = T(u),\ y = T(v),\ u, v \in K$ and

$$d(x, y) = d(T(u), T(v)) \leq \alpha d(u, v) \leq 2\alpha r,$$

whence since x and y are arbitrary from $T(K)$, we obtain $\delta(T(K)) \leq 2\alpha r$.

Now take $x, y \in \overline{T(K)}$, then there exist sequences $\{x_n\}, \{y_n\} \subset T(K)$, $x_n = T(u_n)$, $y_n = T(v_n),\ u_n, v_n \in K$, with $x_n \to x$ and $y_n \to y$ as $n \to \infty$. Therefore, for $\varepsilon > 0$ and $n \in \mathbb{N}$ sufficiently large

$$\begin{aligned} d(x, y) &\leq d(x, x_n) + d(x_n, y_n) + d(y_n, y) \\ &\leq \varepsilon + d(T(u_n), T(v_n)) \leq \varepsilon + \alpha d(u_n.v_n) \\ &\leq \varepsilon + 2\alpha r. \end{aligned}$$

Letting $\varepsilon \to 0$, we get $d(x, y) \leq 2\alpha r$ which implies

$$\delta(\overline{T(K)}) \leq 2\alpha r.$$

By the induction we get easily

$$\delta(\overline{T^s(K)}) \leq 2\alpha^s r, \quad s = 1, 2, \dots . \tag{8}$$

Taking into account (6) and (8), from the Cantor Theorem 1, there exists exactly one element $u \in X$ such that

$$\bigcap_{n=1}^{\infty} \overline{T^n(K)} = \{u\}. \tag{9}$$

Now we shall show that

$$T(\overline{A}) \subset \overline{T(A)}, \quad A \subset X. \tag{10}$$

In fact, if $x \in T(\overline{A})$, then $x = T(a)$ and $a \in \overline{A}$; so there exists a sequence $\{a_n\} \subset A$ such that $a_n \to a$ as $n \to \infty$. But $\{T(a_n)\} \subset T(A)$ and by the continuity of T, resulting from (4),

$$\lim_{n\to\infty} T(a_n) = T(a) = x \in \overline{T(A)},$$

which ends the proof of the condition (10).

Consequently, one has by (10)

$$T(u) = T\left(\bigcap_{n=1}^{\infty} \overline{T^n(K)}\right) \subset \bigcap_{n=1}^{\infty} T(\overline{T^n(K)})$$
$$\subset \bigcap_{n=1}^{\infty} \overline{T^{n+1}(K)} \subset \bigcap_{n=1}^{\infty} \overline{T^n(K)} = \{u\},$$

i.e. T(u)=u.

The condition (ii), in view of (4), is obvious.

Finally we shall prove the condition (iii) (so also (i)). Let $z \in B$, therefore for

$$r = d(T^k(x), z) < \infty,$$

$z \in K(T^k(x), r)$. But since $T^n(z) \in \overline{T^n(K)}$ for $n = 1, 2, \ldots$, we have

$$d(T^k(z), u) \le \delta(\overline{T^n(K)}) \le 2\alpha^n r, \quad n = 1, 2, \ldots,$$

and by (5)

$$d(T^k(z), u) \to 0 \quad as \quad n \to \infty,$$

which completes the proof of the theorem. □

Remark 2. If $z \in B$, then

$$d(z, u) \le \frac{1}{1-\alpha} d(z, T(z)). \tag{11}$$

In fact, one has

$$d(z, u) \le d(z, T(z)) + d(T(z), T(u)) \le d(z, T(z)) + \alpha d(z, u),$$

which implies the condition (11).

Remark 3. As a consequence of Theorem 2 we get the famous Banach contraction principle (see [1, 3]).

Now we shall present another basic theorem for complete generalized metric space, called the Baire theorem.

Theorem 3. *Any complete generalized metric space is a second category set.*

Proof. The proof of this statement can be done similarly as for a metric space. For details the reader is referred, e.g., to [4]. □

Let's note also the useful remark.

Remark 4. Any closed subset of a complete generalized metric space is also a complete generalized metric space.

References

1. Banach, S.: Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundam. Math. **3**, 133–181 (1922)
2. Diaz, J.B., Margolis, B.: A fixed point theorem of the alternative, for contractions on a generalized complete metric space. Bull. Am. Math. Soc. **74**, 305–309 (1968)
3. Dugundji, J., Granas, A.: Fixed Point Theory, pp. 5–209. Polish Scientific Publishers, Warsaw (1982)
4. Lusternik, L.A., Sobolew, W.L.: Elements of Functional Analysis, pp. 7–308. Polish Scientific Publishers, Warsaw (1959)
5. Luxemburg, W.A.J.: On the convergence of successive approximations in the theory of ordinary differential equations III. Nieuw. Arch. Wisk. **6**(3), 93–98 (1958)

A Survey of Perturbed Ostrowski Type Inequalities

Silvestru Sever Dragomir

Abstract In this paper we survey a number of recent perturbed versions of Ostrowski inequality that have been obtained by the author and provide their connections with numerous classical results of interest.

Keywords Ostrowski inequality • Lebesgue integral • Integral mean • Integral inequalities

Introduction

Ostrowski's Inequality

As revealed by a simple search in MathSciNet database with the key words "Ostrowski" and "inequality" in the title, an exponential evolution of research papers devoted to this result has been registered in the last decade. There are now at least 360 papers that can be found by performing the above search. Numerous extensions, generalizations in both the integral and discrete case have been discovered. More general versions for n-time differentiable functions, the corresponding versions on time scales, for vector valued functions or multiple integrals have been established as well. Numerous applications in Numerical Analysis, Approximation Theory, Probability Theory & Statistics, Information Theory, and other fields have also been given.

In this paper, after presenting some inequalities of Ostrowski type for absolutely continuous functions, we survey a number of recent perturbed versions of this inequality that have been obtained lately by the author and provide their connections

S.S. Dragomir (✉)
Mathematics, College of Engineering & Science, Victoria University, PO Box 14428, Melbourne, VIC 8001, Australia

School of Computer Science & Applied Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa
e-mail: sever.dragomir@vu.edu.au

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_8

with numerous classical results of interest. Complete proofs are provided and the necessary references where all these results have been obtained first are also given.

In 1938, Ostrowski [21] proved the following inequality concerning the distance between the integral mean $\frac{1}{b-a}\int_a^b f(t)\,dt$ and the value $f(x)$, $x \in [a,b]$ in the case of differentiable functions on an open interval:

Theorem 1 (Ostrowski, 1938 [21]). *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b) such that $f' : (a,b) \to \mathbb{R}$ is bounded on (a,b), i.e., $\|f'\|_\infty := \sup_{t\in(a,b)} |f'(t)| < \infty$. Then*

$$\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \le \left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a} \right)^2 \right] \|f'\|_\infty (b-a), \tag{1}$$

for all $x \in [a,b]$ and the constant $\frac{1}{4}$ is the best possible.

In [17], S.S. Dragomir and S. Wang, by the use of the *Montgomery integral identity* [20, p. 565],

$$f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt = \frac{1}{b-a}\int_a^b p(x,t) f'(t)\,dt, \quad x \in [a,b],$$

where $p : [a,b]^2 \to \mathbb{R}$ is given by

$$p(x,t) := \begin{cases} t-a \text{ if } t \in [a,x], \\ \\ t-b \text{ if } t \in (x,b], \end{cases}$$

gave a simple proof of Ostrowski's inequality and applied it for special means (identric mean, logarithmic mean, etc.) and to the problem of estimating the error bound in approximating the Riemann integral $\int_a^b f(t)\,dt$ by an arbitrary Riemann sum (see [17], Section Ostrowski for L_p-Norm). For related results see [14] and [19].

A Refinement for L_∞-Norm

The following result, which is an improvement on Ostrowski's inequality, holds.

Theorem 2 (Dragomir, 2002 [4]). *Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous function on $[a,b]$ whose derivative $f' \in L_\infty[a,b]$. Then*

$$\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{2}$$

$$\leq \frac{1}{2(b-a)}\left[\left\|f'\right\|_{[a,x],\infty}(x-a)^2+\left\|f'\right\|_{[x,b],\infty}(b-x)^2\right]$$

$$\leq \begin{cases} \|f'\|_{[a,b],\infty}\left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right](b-a)\,; \\ \\ \frac{1}{2}\left[\|f'\|^{\alpha}_{[a,x],\infty}+\|f'\|^{\alpha}_{[x,b],\infty}\right]^{\frac{1}{\alpha}}\left[\left(\frac{x-a}{b-a}\right)^{2\beta}+\left(\frac{b-x}{b-a}\right)^{2\beta}\right]^{\frac{1}{\beta}}(b-a)\,, \\ \text{where } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ \\ \frac{1}{2}\left[\|f'\|_{[a,x],\infty}+\|f'\|_{[x,b],\infty}\right]\left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]^2(b-a) \end{cases}$$

for all $x \in [a,b]$*, where* $\|\cdot\|_{[m,n],\infty}$ *denotes the usual norm on* $L_\infty[m,n]$*, i.e., we recall that*

$$\|g\|_{[m,n],\infty} = ess \sup_{t\in[m,n]} |g(t)| < \infty.$$

Proof. Using the integration by parts formula for absolutely continuous functions on $[a,b]$, we have

$$\int_a^x (t-a) f'(t)\,dt = (x-a) f(x) - \int_a^x f(t)\,dt \tag{3}$$

and

$$\int_x^b (t-b) f'(t)\,dt = (b-x) f(x) - \int_x^b f(t)\,dt, \tag{4}$$

for all $x \in [a,b]$.

Adding these two equalities, we obtain the Montgomery identity (see, for example, [20, p. 565]):

$$(b-a) f(x) - \int_a^b f(t)\,dt = \int_a^x (t-a) f'(t)\,dt + \int_x^b (t-b) f'(t)\,dt \tag{5}$$

for all $x \in [a,b]$.

Taking the modulus, we deduce

$$\left|(b-a) f(x) - \int_a^b f(t)\,dt\right| \leq \left|\int_a^x (t-a) f'(t)\,dt\right| + \left|\int_x^b (t-b) f'(t)\,dt\right| \tag{6}$$

$$\leq \int_a^x (t-a)\left|f'(t)\right|dt + \int_x^b (b-t)\left|f'(t)\right|dt$$
$$\leq \left\|f'\right\|_{[a,x],\infty} \int_a^x (t-a)\,dt + \left\|f'\right\|_{[x,b],\infty} \int_x^b (b-t)\,dt$$
$$= \frac{1}{2}\left[\left\|f'\right\|_{[a,x],\infty}(x-a)^2 + \left\|f'\right\|_{[x,b],\infty}(b-x)^2\right]$$

and the first inequality in (2) is proved.

Now, let us observe that

$$\left\|f'\right\|_{[a,x],\infty}(x-a)^2 + \left\|f'\right\|_{[x,b],\infty}(b-x)^2$$
$$\leq \max\left\{\left\|f'\right\|_{[a,x],\infty}, \left\|f'\right\|_{[x,b],\infty}\right\}\left[(x-a)^2 + (b-x)^2\right]$$
$$= \max\left\{\left\|f'\right\|_{[a,x],\infty}, \left\|f'\right\|_{[x,b],\infty}\right\}\left[\frac{1}{2}(b-a)^2 + 2\left(x - \frac{a+b}{2}\right)^2\right]$$
$$= (b-a)^2 \max\left\{\left\|f'\right\|_{[a,x],\infty}, \left\|f'\right\|_{[x,b],\infty}\right\}\left[\frac{1}{2} + 2\cdot\frac{\left(x-\frac{a+b}{2}\right)^2}{(b-a)^2}\right]$$
$$= (b-a)^2 \left\|f'\right\|_{[a,b],\infty}\left[\frac{1}{2} + 2\cdot\frac{\left(x-\frac{a+b}{2}\right)^2}{(b-a)^2}\right],$$

and the first part of the second inequality in (2) is proved.

For the second inequality, we employ the elementary inequality for real numbers which can be derived from *Hölder's discrete inequality*

$$0 \leq ms + nt \leq (m^\alpha + n^\alpha)^{\frac{1}{\alpha}} \times \left(s^\beta + t^\beta\right)^{\frac{1}{\beta}}, \tag{7}$$

provided that $m, s, n, t \geq 0$, $\alpha > 1$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Using (7), we obtain

$$\left\|f'\right\|_{[a,x],\infty}(x-a)^2 + \left\|f'\right\|_{[x,b],\infty}(b-x)^2$$
$$\leq \left(\left\|f'\right\|_{[a,x],\infty}^\alpha + \left\|f'\right\|_{[x,b],\infty}^\alpha\right)^{\frac{1}{\alpha}}\left[(x-a)^{2\beta} + (b-x)^{2\beta}\right]^{\frac{1}{\beta}}$$

and the second part of the second inequality in (2) is also obtained.

Finally, we observe that

$$\left\|f'\right\|_{[a,x],\infty}(x-a)^2 + \left\|f'\right\|_{[x,b],\infty}(b-x)^2$$
$$\leq \max\left\{(x-a)^2, (b-x)^2\right\}\left[\left\|f'\right\|_{[a,x],\infty} + \left\|f'\right\|_{[x,b],\infty}\right]$$

$$= \left[\frac{b-a}{2} + \left|x - \frac{a+b}{2}\right|\right]^2 \left[\|f'\|_{[a,x],\infty} + \|f'\|_{[x,b],\infty}\right]$$

and the last part of the second inequality in (2) is proved.

The following corollary is also natural.

Corollary 1. *Under the above assumptions, we have the midpoint inequality*

$$\left|f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{8}$$
$$\leq \frac{1}{8}(b-a)\left[\|f'\|_{\left[a,\frac{a+b}{2}\right],\infty} + \|f'\|_{\left[\frac{a+b}{2},b\right],\infty}\right]$$

$$\leq \begin{cases} \frac{1}{4}(b-a)\|f'\|_{[a,b],\infty}\,; \\[2ex] \frac{1}{2^{\frac{3\beta-1}{2}}}(b-a)\left[\|f'\|^{\alpha}_{\left[a,\frac{a+b}{2}\right],\infty} + \|f'\|^{\alpha}_{\left[\frac{a+b}{2},b\right],\infty}\right]^{\frac{1}{\alpha}}, \\[2ex] \text{where } \alpha > 1,\ \frac{1}{\alpha} + \frac{1}{\beta} = 1. \end{cases}$$

Ostrowski for L_1-Norm

L_1-Norm Inequality

In 1997, Dragomir and Wang proved the following Ostrowski type inequality [15].

Theorem 3. *Let $f : [a,b] \to \mathbb{R}$ be an absolutely continuous function on $[a,b]$. Then we have the inequality*

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \left[\frac{1}{2} + \frac{\left|x - \frac{a+b}{2}\right|}{b-a}\right]\|f'\|_{[a,b],1}\,, \tag{9}$$

for all $x \in [a,b]$, where $\|\cdot\|_1$ is the Lebesgue norm on $L_1[a,b]$, i.e., we recall it $\|g\|_{[a,b],1} := \int_a^b |g(t)|\,dt$. The constant $\frac{1}{2}$ is best possible.

Note that the fact that $\frac{1}{2}$ is the best constant for differentiable functions was proved in [22] and (9) can also be obtained from a more general result obtained by A. M. Fink in [18] choosing $n = 1$ and doing some appropriate computation. However the inequality (9) was not stated explicitly in [18].

A Refinement for L_1-Norm

The following result, which is an improvement on the inequality (9), holds.

Theorem 4 (Dragomir, 2002 [3]). *Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous function on $[a,b]$. Then*

$$\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{10}$$
$$\leq \frac{x-a}{b-a}\left\| f' \right\|_{[a,x],1} + \frac{b-x}{b-a}\left\| f' \right\|_{[x,b],1}$$
$$\leq \begin{cases} \frac{1}{2}\left[\|f'\|_{[a,b],1} + \left| \|f'\|_{[a,x],1} - \|f'\|_{[x,b],1} \right| \right] \\ \left[\left(\frac{x-a}{b-a}\right)^{\beta} + \left(\frac{b-x}{b-a}\right)^{\beta} \right]^{\frac{1}{\beta}} \left(\left| \|f'\|^{\alpha}_{[a,x],1} \right| + \|f'\|^{\alpha}_{[x,b],1} \right)^{\frac{1}{\alpha}} \\ \text{where } \alpha > 1 \text{ and } \frac{1}{\alpha} + \frac{1}{\beta} = 1, \\ \left[\frac{1}{2} + \frac{\left| x - \frac{a+b}{2} \right|}{b-a} \right] \|f'\|_{[a,b],1} \end{cases}$$

for all $x \in [a,b]$.

Proof. Start with the Montgomery identity for absolutely continuous functions proved in Theorem 2

$$(b-a)f(x) - \int_a^b f(t)\,dt = \int_a^x (t-a)f'(t)\,dt + \int_x^b (t-b)f'(t)\,dt \tag{11}$$

for all $x \in [a,b]$.

Taking the modulus, we deduce

$$\left| (b-a)f(x) - \int_a^b f(t)\,dt \right| \leq \left| \int_a^x (t-a)f'(t)\,dt \right| + \left| \int_x^b (t-b)f'(t)\,dt \right| \tag{12}$$
$$\leq \int_a^x (t-a)\left| f'(t) \right| dt + \int_x^b (b-t)\left| f'(t) \right| dt$$
$$\leq (x-a)\int_a^x \left| f'(t) \right| dt + (b-x)\int_x^b \left| f'(t) \right| dt$$
$$= (x-a)\left\| f' \right\|_{[a,x],1} + (b-x)\left\| f' \right\|_{[x,b],1}$$

and the first inequality in (10) is proved.

Now, let us observe that

$$
\begin{aligned}
&(x-a)\left\|f'\right\|_{[a,x],1}+(b-x)\left\|f'\right\|_{[x,b],1}\\
&\leq \max\left\{\left\|f'\right\|_{[a,x],1},\left\|f'\right\|_{[x,b],1}\right\}(b-a)\\
&=\frac{1}{2}\left[\left\|f'\right\|_{[a,b],1}+\left|\left\|f'\right\|_{[a,x],1}-\left\|f'\right\|_{[x,b],1}\right|\right](b-a)
\end{aligned}
$$

and the first part of the second inequality is proved.

For the second inequality, we employ the elementary inequality for real numbers which can be derived from Hölder's discrete inequality

$$0\leq ms+nt\leq\left(m^{\alpha}+n^{\alpha}\right)^{\frac{1}{\alpha}}\times\left(s^{\beta}+t^{\beta}\right)^{\frac{1}{\beta}}, \tag{13}$$

provided that $m,s,n,t\geq 0$, $\alpha>1$ and $\frac{1}{\alpha}+\frac{1}{\beta}=1$.

Using (13), we obtain

$$
\begin{aligned}
&(x-a)\left\|f'\right\|_{[a,x],1}+(b-x)\left\|f'\right\|_{[x,b],1}\\
&\leq\left(\left\|f'\right\|_{[a,x],1}^{\alpha}+\left\|f'\right\|_{[x,b],1}^{\alpha}\right)^{\frac{1}{\alpha}}\left[(x-a)^{\beta}+(b-x)^{\beta}\right]^{\frac{1}{\beta}}
\end{aligned}
$$

and the second part of the second inequality in (10) is also obtained.

Finally, we observe that

$$
\begin{aligned}
&(x-a)\left\|f'\right\|_{[a,x],1}+(b-x)\left\|f'\right\|_{[x,b],1}\\
&\leq\max\{x-a,b-x\}\left[\left\|f'\right\|_{[a,x],1}+\left\|f'\right\|_{[x,b],1}\right]\\
&=\left[\frac{b-a}{2}+\left|x-\frac{a+b}{2}\right|\right]\left\|f'\right\|_{[a,b],1}
\end{aligned}
$$

and the last part of the second inequality in (10) is proved.

The following corollary is also natural.

Corollary 2. *Under the above assumptions, we have*

$$\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\leq\frac{1}{2}\left\|f'\right\|_{[a,b],1}. \tag{14}$$

Another interesting result is the following one.

Corollary 3. *Under the above assumptions, and if there is an* $x_0 \in [a, b]$ *with*

$$\int_a^{x_0} \left|f'(t)\right| dt = \int_{x_0}^b \left|f'(t)\right| dt \tag{15}$$

then we have the inequality

$$\left|f(x_0) - \frac{1}{b-a}\int_a^b f(t)\, dt\right| \le \frac{1}{2}\left\|f'\right\|_{[a,b],1}. \tag{16}$$

Ostrowski for L_p-Norm

L_p-Norm Inequality

In 1998, Dragomir and Wang proved the following Ostrowski type inequality [16].

Theorem 5. *Let* $f : [a, b] \to \mathbb{R}$ *be an absolutely continuous function on* $[a, b]$. *If* $f' \in L_p[a, b]$, *then we have the inequality*

$$\begin{aligned} &\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\, dt\right| \\ &\le \frac{1}{(q+1)^{1/q}}\left[\left(\frac{x-a}{b-a}\right)^{q+1} + \left(\frac{b-x}{b-a}\right)^{q+1}\right]^{1/q} (b-a)^{1/q}\left\|f'\right\|_{[a,b],p}, \end{aligned} \tag{17}$$

for all $x \in [a, b]$, *where* $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ *and* $\|\cdot\|_{[a,b],p}$ *is the p-Lebesgue norm on* $L_p[a, b]$, *i.e., we recall it*

$$\|g\|_{[a,b],p} := \left(\int_a^b |g(t)|^p\, dt\right)^{1/p}.$$

From (17) we get the following midpoint inequality

$$\left|f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\, dt\right| \le \frac{1}{2(q+1)^{1/q}}(b-a)^{1/q}\left\|f'\right\|_{[a,b],p}, \tag{18}$$

and $\frac{1}{2}$ is a best possible constant.

Indeed, if we take $f : [a, b] \to \mathbb{R}$ with $f(t) = \left|t - \frac{a+b}{2}\right|$, then f is absolutely continuous $\int_a^b f(t)\, dt = \frac{(b-a)^2}{4}$, $\|f'\|_{[a,b],p} = (b-a)^{1/p}$ and if we assume that (18) holds with a constant $C > 0$ instead of $\frac{1}{2}$, then we get $\frac{1}{4}(b-a) \leq \frac{C}{(q+1)^{1/q}}(b-a)$ for any $q > 1$. Letting $q \to 1+$, we obtain $C \geq \frac{1}{2}$, which proves the sharpness of the constant.

In the following, we provide some refinements of (17) and (18).

A Refinement for L_p-Norm

The following result, which is an improvement on the inequality (17), holds.

Theorem 6 (Dragomir, 2013 [9]). *Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous function on $[a,b]$. If $f' \in L_p[a,b]$, then*

$$\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{19}$$
$$\leq \frac{1}{(q+1)^{1/q}}\left[\left(\frac{x-a}{b-a}\right)^{\frac{q+1}{q}} \|f'\|_{[a,x],p} + \left(\frac{b-x}{b-a}\right)^{\frac{q+1}{q}} \|f'\|_{[x,b],p}\right](b-a)^{1/q}$$
$$\leq \frac{1}{(q+1)^{1/q}}$$
$$\times \begin{cases} \frac{1}{2}\left[\|f'\|_{[a,x],p} + \|f'\|_{[x,b],p} + \left|\|f'\|_{[a,x],p} - \|f'\|_{[x,b],p}\right|\right] \\ \times \left[\left(\frac{x-a}{b-a}\right)^{\frac{q+1}{q}} + \left(\frac{b-x}{b-a}\right)^{\frac{q+1}{q}}\right](b-a)^{1/q} \\ \\ \left(\|f'\|^{\alpha}_{[a,x],p} + \|f'\|^{\alpha}_{[x,b],p}\right)^{\frac{1}{\alpha}}\left[\left(\frac{x-a}{b-a}\right)^{\frac{q+1}{q}\beta} + \left(\frac{b-x}{b-a}\right)^{\frac{q+1}{q}\beta}\right]^{\frac{1}{\beta}}(b-a)^{1/q} \\ \text{where } \alpha > 1 \text{ and } \frac{1}{\alpha} + \frac{1}{\beta} = 1, \\ \\ \left[\|f'\|_{[a,x],p} + \|f'\|_{[x,b],p}\right]\left[\frac{1}{2} + \left|\frac{x - \frac{a+b}{2}}{b-a}\right|\right]^{\frac{q+1}{q}}(b-a)^{1/q} \end{cases}$$

for all $x \in [a,b]$, where $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Start with the Montgomery identity for absolutely continuous functions

$$(b-a)f(x) - \int_a^b f(t)\,dt = \int_a^x (t-a)f'(t)\,dt + \int_x^b (t-b)f'(t)\,dt \tag{20}$$

for all $x \in [a,b]$.

Taking the modulus, we deduce

$$\left|(b-a)f(x)-\int_a^b f(t)\,dt\right| \le \left|\int_a^x (t-a)f'(t)\,dt\right| + \left|\int_x^b (t-b)f'(t)\,dt\right| \quad (21)$$
$$\le \int_a^x (t-a)\left|f'(t)\right| dt + \int_x^b (b-t)\left|f'(t)\right| dt.$$

Utilizing Hölder's integral inequality we have

$$\int_a^x (t-a)\left|f'(t)\right| dt + \int_x^b (b-t)\left|f'(t)\right| dt$$
$$\le \left(\int_a^x (t-a)^q\,dt\right)^{1/q}\left(\int_a^x \left|f'(t)\right|^p dt\right)^{1/p} + \left(\int_x^b (b-t)^q\,dt\right)^{1/q}\left(\int_x^b \left|f'(t)\right|^p dt\right)^{1/p}$$
$$= \frac{1}{(b-a)(q+1)^{1/q}}\left[(x-a)^{\frac{q+1}{q}}\left\|f'\right\|_{[a,x],p} + (b-x)^{\frac{q+1}{q}}\left\|f'\right\|_{[x,b],p}\right]$$

for all $x \in [a,b]$, and the first inequality in (19) is proved.

Now, let us observe that

$$(x-a)^{\frac{q+1}{q}}\left\|f'\right\|_{[a,x],p} + (b-x)^{\frac{q+1}{q}}\left\|f'\right\|_{[x,b],p}$$
$$\le \max\left\{\left\|f'\right\|_{[a,x],p}, \left\|f'\right\|_{[x,b],p}\right\}\left[(x-a)^{\frac{q+1}{q}} + (b-x)^{\frac{q+1}{q}}\right]$$
$$= \frac{1}{2}\left[\left\|f'\right\|_{[a,x],p} + \left\|f'\right\|_{[x,b],p} + \left|\left\|f'\right\|_{[a,x],p} - \left\|f'\right\|_{[x,b],p}\right|\right]$$
$$\times\left[(x-a)^{\frac{q+1}{q}} + (b-x)^{\frac{q+1}{q}}\right]$$

and the first part of the second inequality is proved.

For the second inequality, we employ the elementary inequality for real numbers which can be derived from Hölder's discrete inequality

$$0 \le ms + nt \le (m^\alpha + n^\alpha)^{\frac{1}{\alpha}} \times \left(s^\beta + t^\beta\right)^{\frac{1}{\beta}}, \quad (22)$$

provided that $m, s, n, t \ge 0$, $\alpha > 1$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Using (22), we obtain

$$(x-a)^{\frac{q+1}{q}}\left\|f'\right\|_{[a,x],p} + (b-x)^{\frac{q+1}{q}}\left\|f'\right\|_{[x,b],p}$$
$$\le \left(\left\|f'\right\|_{[a,x],p}^\alpha + \left\|f'\right\|_{[x,b],p}^\alpha\right)^{\frac{1}{\alpha}}\left[(x-a)^{\frac{q+1}{q}\beta} + (b-x)^{\frac{q+1}{q}\beta}\right]^{\frac{1}{\beta}}$$

and the second part of the second inequality in (19) is also obtained.

Finally, we observe that

$$\begin{aligned}&(x-a)^{\frac{q+1}{q}}\left\|f'\right\|_{[a,x],p}+(b-x)^{\frac{q+1}{q}}\left\|f'\right\|_{[x,b],p}\\&\le\max\left\{(x-a)^{\frac{q+1}{q}},(b-x)^{\frac{q+1}{q}}\right\}\left[\left\|f'\right\|_{[a,x],p}+\left\|f'\right\|_{[x,b],p}\right]\\&=\left[\frac{b-a}{2}+\left|x-\frac{a+b}{2}\right|\right]^{\frac{q+1}{q}}\left[\left\|f'\right\|_{[a,x],p}+\left\|f'\right\|_{[x,b],p}\right]\end{aligned}$$

and the last part of the second inequality in (19) is proved.

The following corollary is also natural.

Corollary 4. *Under the above assumptions, we have*

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\\&\le\frac{1}{2^{(q+1)/q}(q+1)^{1/q}}\left[\left\|f'\right\|_{\left[a,\frac{a+b}{2}\right],p}+\left\|f'\right\|_{\left[\frac{a+b}{2},b\right],p}\right](b-a)^{1/q}.\end{aligned}\tag{23}$$

Another interesting result is the following one.

Corollary 5. *Under the above assumptions, and if there is an* $x_0\in[a,b]$ *with*

$$\int_a^{x_0}\left|f'(t)\right|^p dt=\int_{x_0}^b\left|f'(t)\right|^p dt\tag{24}$$

then we have the inequality

$$\begin{aligned}&\left|f(x_0)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\\&\le\frac{1}{(q+1)^{1/q}}\left[\left(\frac{x_0-a}{b-a}\right)^{\frac{q+1}{q}}+\left(\frac{b-x_0}{b-a}\right)^{\frac{q+1}{q}}\right]\left\|f'\right\|_{[a,x_0],p}(b-a)^{1/q}.\end{aligned}\tag{25}$$

Remark 1. If we take in (19) $\alpha=p$ and $\beta=q$, where $p>1$, $\frac{1}{p}+\frac{1}{q}=1$, then we get the following refinement of (17)

$$\begin{aligned}&\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\\&\le\frac{1}{(q+1)^{1/q}}\left[\left(\frac{x-a}{b-a}\right)^{\frac{q+1}{q}}\left\|f'\right\|_{[a,x],p}+\left(\frac{b-x}{b-a}\right)^{\frac{q+1}{q}}\left\|f'\right\|_{[x,b],p}\right](b-a)^{1/q}\end{aligned}\tag{26}$$

$$\leq \frac{1}{(q+1)^{1/q}} \left[\left(\frac{x-a}{b-a} \right)^{q+1} + \left(\frac{b-x}{b-a} \right)^{q+1} \right]^{1/q} (b-a)^{1/q} \left\| f' \right\|_{[a,b],p},$$

for all $x \in [a, b]$.

This is true, since for $\alpha = p$ we have

$$\left(\left\| f' \right\|_{[a,x],p}^{\alpha} + \left\| f' \right\|_{[x,b],p}^{\alpha} \right)^{\frac{1}{\alpha}} = \left(\left\| f' \right\|_{[a,x],p}^{p} + \left\| f' \right\|_{[x,b],p}^{p} \right)^{\frac{1}{p}}$$
$$= \left(\int_a^x \left| f'(t) \right|^p dt + \int_x^b \left| f'(t) \right|^p \right)^{1/p} = \left\| f' \right\|_{[a,b],p}.$$

Ostrowski for Bounded Derivatives

We start with the following result.

Theorem 7 (Dragomir, 2003 [7]). *Let $f : [a, b] \to \mathbb{R}$ be an absolutely continuous function on $[a, b]$ and $x \in [a, b]$. Suppose that there exist the functions m_i, $M_i : [a, b] \to \mathbb{R}$ $\left(i = \overline{1, 2}\right)$ with the properties:*

$$m_1(x) \leq f'(t) \leq M_1(x) \text{ for a.e. } t \in [a, x] \tag{27}$$

and

$$m_2(x) \leq f'(t) \leq M_2(x) \text{ for a.e. } t \in (x, b]. \tag{28}$$

Then we have the inequalities:

$$\frac{1}{2(b-a)} \left[m_1(x)(x-a)^2 - M_2(x)(b-x)^2 \right] \tag{29}$$
$$\leq f(x) - \frac{1}{b-a} \int_a^b f(t)\, dt$$
$$\leq \frac{1}{2(b-a)} \left[M_1(x)(x-a)^2 - m_2(x)(b-x)^2 \right].$$

The constant $\frac{1}{2}$ is sharp on both sides.

Proof. Start with the Montgomery identity

$$f(x) = \frac{1}{b-a} \int_a^b f(t)\, dt + \frac{1}{b-a} \int_a^x (t-a) f'(t)\, dt + \frac{1}{b-a} \int_x^b (t-b) f'(t)\, dt \tag{30}$$

for any $x \in [a, b]$.

Using the assumption (27) and (28), we have

$$m_1(x)(t-a) \leq (t-a) f'(t) \leq M_1(x)(t-a) \text{ for a.e. } t \in [a, x] \tag{31}$$

and

$$M_2(x)(t-b) \leq f'(t)(t-b) \leq m_2(x)(t-b) \text{ for a.e. } t \in (x, b]. \tag{32}$$

Integrating (31) on $[a, x]$ and (32) on $[x, b]$ and summing the obtained inequalities, we have

$$\begin{aligned}
&\frac{1}{2} m_1(x)(x-a)^2 - \frac{1}{2} M_2(x)(b-x)^2 \\
&\leq \int_a^x (t-a) f'(t)\, dt + \int_x^b (t-b) f'(t)\, dt \\
&\leq \frac{1}{2} M_1(x)(x-a)^2 - \frac{1}{2} m_2(x)(b-x)^2.
\end{aligned}$$

Using the representation (30), we deduce (29).

Assume that the first inequality in (29) holds with a constant $c > 0$; that is

$$\frac{c}{b-a}\left[m_1(x)(x-a)^2 - M_2(x)(b-x)^2\right] \leq f(x) - \frac{1}{b-a}\int_a^b f(t)\, dt. \tag{33}$$

Consider the function $f : [a, b] \to \mathbb{R}, f(t) = M|t-x|, M > 0$. Then f is absolutely continuous and

$$f'(t) = \begin{cases} -M & \text{if } t \in [a, x] \\ M & \text{if } t \in (x, b]. \end{cases}$$

Thus, if we choose $m_1 = -M, m_2 = M$ in (33), we get

$$\begin{aligned}
-M\frac{c}{b-a}\left[(x-a)^2 + (b-x)^2\right] &\leq -\frac{M}{b-a}\int_a^b |t-x|\, dt \\
&= -\frac{M}{b-a}\left[\frac{(b-x)^2 + (x-a)^2}{2}\right]
\end{aligned}$$

for all $x \in [a, b]$, implying that $c \geq \frac{1}{2}$, that is, $\frac{1}{2}$ is the best constant in the first member of (29).

Using a similar process, we may prove that $\frac{1}{2}$ is the best constant in the third member of (29) and the theorem is completely proved.

Corollary 6. *If $f : [a,b] \to \mathbb{R}$ is absolutely continuous on $[a,b]$ and the derivative $f' : [a,b] \to \mathbb{R}$ is bounded above and below, that is,*

$$-\infty < m \leq f'(t) \leq M < \infty \text{ for a.e. } t \in [a,b], \tag{34}$$

then we have the inequality

$$\begin{aligned}\frac{1}{2(b-a)}\left[m(x-a)^2 - M(b-x)^2\right] &\leq f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \\ &\leq \frac{1}{2(b-a)}\left[M(x-a)^2 - m(b-x)^2\right]\end{aligned} \tag{35}$$

for all $x \in [a,b]$. The constant $\frac{1}{2}$ is the best in both inequalities.

Applying Taylor's formula

$$g(x) = g\left(\frac{a+b}{2}\right) + \left(x - \frac{a+b}{2}\right) g'\left(\frac{a+b}{2}\right) + \frac{1}{2}\left(x - \frac{a+b}{2}\right)^2 g''\left(\frac{a+b}{2}\right)$$

for $g(x) = M(x-a)^2 - m(b-x)^2$, we obtain

$$\begin{aligned} g(x) = &\frac{1}{4}(M-m)(b-a)^2 + 2\left(x - \frac{a+b}{2}\right)\left(\frac{M+m}{2}\right)(b-a) \\ &+ (M-m)\left(x - \frac{a+b}{2}\right)^2. \end{aligned}$$

The same formula applied for $h(x) = m(x-a)^2 - M(b-x)^2$ will reveal that

$$\begin{aligned} h(x) = &\frac{1}{4}(M-m)(b-a)^2 + 2\left(x - \frac{a+b}{2}\right)\left(\frac{M+m}{2}\right)(b-a) \\ &- (M-m)\left(x - \frac{a+b}{2}\right)^2. \end{aligned}$$

Consequently, we may rewrite Corollary 6 in the following equivalent manner:

Corollary 7. *With the assumptions on Corollary 6, we have*

$$\begin{aligned} &\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt - \left(x - \frac{a+b}{2}\right)\left(\frac{M+m}{2}\right)\right| \\ &\leq \frac{1}{2}(M-m)(b-a)\left[\left(\frac{x - \frac{a+b}{2}}{b-a}\right)^2 + \frac{1}{4}\right] \end{aligned} \tag{36}$$

for all $x \in [a,b]$.

Remark 2. If we assume that $\|f'\|_\infty := ess \sup_{t\in[a,b]} |f'(t)|$, then obviously we may choose in (35) $m = \|f'\|_\infty$ and $M = \|f'\|_\infty$, obtaining Ostrowski's inequality for absolutely continuous functions whose derivatives are essentially bounded:

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \le \frac{\|f'\|_\infty}{2(b-a)}\left[(x-a)^2 + (b-x)^2\right]$$
$$= \left[\frac{1}{4} + \left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right](b-a)\|f'\|_\infty,$$

for all $x \in [a,b]$. The constant $\frac{1}{4}$ here is best.

Remark 3. Ostrowski's inequality for absolutely continuous mappings in terms of $\|f'\|_\infty$ basically states that

$$-\frac{\|f'\|_\infty}{2(b-a)}\left[(x-a)^2 + (b-x)^2\right] \le f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{37}$$
$$\le \frac{\|f'\|_\infty}{2(b-a)}\left[(x-a)^2 + (b-x)^2\right]$$

for all $x \in [a,b]$.

Now, if we assume that (27) and (28) hold, then $-\|f'\|_\infty \le m_1(x), m_2(x)$ and $M_1(x), M_2(x) \le \|f'\|_\infty$, which implies

$$-\frac{\|f'\|_\infty}{2(b-a)}\left[(x-a)^2 + (b-x)^2\right] \tag{38}$$
$$\le \frac{1}{2(b-a)}\left[m_1(x)(x-a)^2 - M_2(x)(b-x)^2\right]$$
$$\le f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt$$
$$\le \frac{1}{2(b-a)}\left[M_1(x)(x-a)^2 - m_2(x)(b-x)^2\right]$$
$$\le \frac{\|f'\|_\infty}{2(b-a)}\left[(x-a)^2 + (b-x)^2\right].$$

Thus, the inequality (29) may also be regarded as a refinement of the classical Ostrowski result.

An important particular case is $x = \frac{a+b}{2}$ providing the following corollary.

Corollary 8. *Assume that the derivative* $f' : [a,b] \to \mathbb{R}$ *satisfy the conditions:*

$$-\infty < m_1 \leq f'(t) \leq M_1 < \infty \text{ for a.e. } t \in \left[a, \frac{a+b}{2}\right] \tag{39}$$

and

$$-\infty < m_2 \leq f'(t) \leq M_2 < \infty \text{ for a.e. } t \in \left(\frac{a+b}{2}, b\right]. \tag{40}$$

Then we have the inequalities

$$\begin{aligned}\frac{1}{8}(m_1 - M_2)(b-a) &\leq f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \\ &\leq \frac{1}{8}(M_1 - m_2)(b-a).\end{aligned} \tag{41}$$

The constant $\frac{1}{8}$ *is the best in both inequalities.*

Finally, if we know some global bounds for the derivative f' on $[a,b]$, then we may state the following corollary.

Corollary 9. *Under the assumptions of Corollary 6, we have the midpoint inequality:*

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \leq \frac{1}{8}(M-m)(b-a). \tag{42}$$

The constant $\frac{1}{8}$ *is best.*

Proof. The inequality is obvious by Corollary 6 putting $x = \frac{a+b}{2}$. We observe that the function $f : [a,b] \to \mathbb{R}, f(x) = k\left|x - \frac{a+b}{2}\right|, k > 0$ is absolutely continuous and $-k \leq f'(t) \leq k$ for all $t \in [a,b]$. Thus, we may choose $M = k, m = -k$ and as

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| = \frac{1}{8}(M-m)(b-a) = \frac{k(b-a)}{4},$$

we conclude that the constant $\frac{1}{8}$ is best in (42).

Functional Ostrowski Inequality for Convex Mappings

A Generalization of Ostrowski's Inequality

The following result holds:

Theorem 8 (Dragomir, 2013 [9]). *Let $f : [a,b] \to \mathbb{R}$ be absolutely continuous on $[a,b]$. If $\Phi : \mathbb{R} \to \mathbb{R}$ is convex (concave) on $\mathbb{R}$ then we have the inequalities*

$$\begin{aligned} &\Phi\left(f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right) \\ &\le (\ge) \frac{1}{b-a}\left[\int_a^x \Phi\left[(t-a)f'(t)\right]dt + \int_x^b \Phi\left[(t-b)f'(t)\right]dt\right] \end{aligned} \tag{43}$$

for any $x \in [a,b]$.

Proof. Utilizing the Montgomery identity

$$\begin{aligned} f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt &= \frac{1}{b-a}\left[\int_a^x (t-a)f'(t)\,dt + \int_x^b (t-b)f'(t)\,dt\right] \\ &= \frac{x-a}{b-a}\left(\frac{1}{x-a}\int_a^x (t-a)f'(t)\,dt\right) \\ &\quad + \frac{b-x}{b-a}\left(\frac{1}{b-x}\int_x^b (t-b)f'(t)\,dt\right), \end{aligned} \tag{44}$$

which holds for any $x \in (a,b)$ and the convexity of $\Phi : \mathbb{R} \to \mathbb{R}$, we have

$$\begin{aligned} &\Phi\left(f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right) \\ &\le \frac{x-a}{b-a}\Phi\left(\frac{1}{x-a}\int_a^x (t-a)f'(t)\,dt\right) \\ &\quad + \frac{b-x}{b-a}\Phi\left(\frac{1}{b-x}\int_x^b (t-b)f'(t)\,dt\right) \end{aligned} \tag{45}$$

for any $x \in (a,b)$, which is an inequality of interest in itself as well.

If we use Jensen's integral inequality

$$\Phi\left(\frac{1}{d-c}\int_c^d g(t)\,dt\right) \le \frac{1}{d-c}\int_c^d \Phi\left[g(t)\right]dt$$

we have

$$\Phi\left(\frac{1}{x-a}\int_a^x (t-a)f'(t)\,dt\right) \le \frac{1}{x-a}\int_a^x \Phi\left[(t-a)f'(t)\right]dt \tag{46}$$

and

$$\Phi\left(\frac{1}{b-x}\int_x^b (t-b)f'(t)\,dt\right) \le \frac{1}{b-x}\int_x^b \Phi\left[(t-b)f'(t)\right]dt \tag{47}$$

for any $x \in (a,b)$.

Making use of (45)–(47) we get the desired result (43) for the convex functions.

If $x = b$, then

$$f(b) - \frac{1}{b-a}\int_a^b f(t)\,dt = \frac{1}{b-a}\int_a^b (t-a)f'(t)\,dt$$

and by Jensen's inequality we get

$$\Phi\left(f(b) - \frac{1}{b-a}\int_a^b f(t)\,dt\right) \le \frac{1}{b-a}\int_a^b \Phi\left[(t-a)f'(t)\right]dt,$$

which proves the inequality (43) for $x = b$.

The same argument can be applied for $x = a$.

The case of concave functions goes likewise and the theorem is proved.

Corollary 10 (Dragomir, 2013 [9]). *With the assumptions of Theorem 8 we have*

$$\begin{aligned}\Phi(0) &\le (\ge) \frac{1}{b-a}\int_a^b \Phi\left(f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right)dx \\ &\le (\ge) \frac{1}{(b-a)^2}\left[\int_a^b (b-x)\,\Phi\left[(x-a)f'(x)\right]dx \right. \\ &\quad \left. + \int_a^b (x-a)\,\Phi\left[(x-b)f'(x)\right]dx\right].\end{aligned} \tag{48}$$

Proof. By Jensen's integral inequality we have

$$\begin{aligned}&\frac{1}{b-a}\int_a^b \Phi\left(f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right)dx \\ &\ge (\le)\, \Phi\left[\frac{1}{b-a}\int_a^b \left(f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right)dx\right] = \Phi(0),\end{aligned}$$

which proves the first inequality in (48).

Integrating the inequality (43) over x we have

$$\frac{1}{b-a}\int_a^b \Phi\left(f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right)dx \tag{49}$$
$$\leq (\geq)\ \frac{1}{(b-a)^2}\int_a^b\left[\int_a^x \Phi\left[(t-a)f'(t)\right]dt+\int_x^b \Phi\left[(t-b)f'(t)\right]dt\right]dx.$$

Integrating by parts we have

$$\begin{aligned}
&\int_a^b\left(\int_a^x \Phi\left[(t-a)f'(t)\right]dt\right)dx\\
&= x\int_a^x \Phi\left[(t-a)f'(t)\right]dt\Big|_a^b-\int_a^b xd\left(\int_a^x \Phi\left[(t-a)f'(t)\right]dt\right)\\
&= b\int_a^b \Phi\left[(t-a)f'(t)\right]dt-\int_a^b x\Phi\left[(x-a)f'(x)\right]dx\\
&= \int_a^b (b-x)\,\Phi\left[(x-a)f'(x)\right]dx
\end{aligned}$$

and

$$\begin{aligned}
&\int_a^b\left(\int_x^b \Phi\left[(t-b)f'(t)\right]dt\right)dx\\
&= x\left(\int_x^b \Phi\left[(t-b)f'(t)\right]dt\right)\Big|_a^b-\int_a^b xd\left(\int_x^b \Phi\left[(t-b)f'(t)\right]dt\right)\\
&= -a\left(\int_a^b \Phi\left[(t-b)f'(t)\right]dt\right)+\int_a^b x\Phi\left[(x-b)f'(x)\right]dx\\
&= \int_a^b (x-a)\,\Phi\left[(x-b)f'(x)\right]dx.
\end{aligned}$$

Utilizing the inequality (49) we deduce the desired inequality (48).

Remark 4. If we write the inequality (43) for the convex function $\Phi(x)=|x|^p$, $p\geq 1$, then we get the inequality

$$\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|^p \tag{50}$$
$$\leq \frac{1}{b-a}\left[\int_a^x (t-a)^p\left|f'(t)\right|^p dt+\int_x^b (b-t)^p\left|f'(t)\right|^p dt\right]$$

for $x\in[a,b]$.

Utilizing Hölder's inequality we have

$$B(x) := \int_a^x (t-a)^p \left|f'(t)\right|^p dt + \int_x^b (b-t)^p \left|f'(t)\right|^p dt \tag{51}$$

$$\leq \begin{cases} \frac{(x-a)^{p+1}}{p+1} \|f'\|^p_{[a,x],\infty} & \text{if } f' \in L_\infty[a,x]; \\ \frac{(x-a)^{p+1/\alpha}}{(p\alpha+1)^{1/\alpha}} \|f'\|^p_{[a,x],p\beta} & \text{if } f' \in L_{p\beta}[a,x], \\ & \alpha > 1, 1/\alpha + 1/\beta = 1; \\ (x-a)^p \|f'\|^p_{[a,x],p} \end{cases}$$

$$+ \begin{cases} \frac{(b-x)^{p+1}}{p+1} \|f'\|^p_{[x,b],\infty} & \text{if } f' \in L_\infty[x,b]; \\ \frac{(b-x)^{p+1/\alpha}}{(p\alpha+1)^{1/\alpha}} \|f'\|^p_{[x,b],p\beta} & \text{if } f' \in L_{p\beta}[x,b], \\ & \alpha > 1, 1/\alpha + 1/\beta = 1; \\ (b-x)^p \|f'\|^p_{[x,b],p} \end{cases}$$

for $x \in [a,b]$.

Utilizing the inequalities (50) and (51) we have for $x \in [a,b]$ that

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right|^p \tag{52}$$

$$\leq \frac{1}{(b-a)(p+1)} \left[(x-a)^{p+1} \left\|f'\right\|^p_{[a,x],\infty} + (b-x)^{p+1} \left\|f'\right\|^p_{[x,b],\infty}\right]$$

$$\leq \frac{1}{(p+1)} \left[\left(\frac{x-a}{b-a}\right)^{p+1} + \left(\frac{b-x}{b-a}\right)^{p+1}\right] (b-a)^p \left\|f'\right\|^p_{[a,b],\infty}$$

provided $f' \in L_\infty[a,b]$,

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right|^p \tag{53}$$

$$\leq \frac{1}{(b-a)(p\alpha+1)^{1/\alpha}} \left[(x-a)^{p+1/\alpha} \left\|f'\right\|^p_{[a,x],p\beta} + (b-x)^{p+1/\alpha} \left\|f'\right\|^p_{[x,b],p\beta}\right]$$

$$\leq \frac{1}{(p\alpha+1)^{1/\alpha}} \left[\left(\frac{x-a}{b-a}\right)^{p+1/\alpha} + \left(\frac{b-x}{b-a}\right)^{p+1/\alpha}\right] (b-a)^{p-1/\beta} \left\|f'\right\|^p_{[a,b],p\beta}$$

provided $f' \in L_{p\beta}[a,b]$, $\alpha > 1$, $1/\alpha + 1/\beta = 1$ and

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right|^p \tag{54}$$

$$\leq \frac{1}{b-a}\left[(x-a)^p \left\|f'\right\|_{[a,x],p}^p + (b-x)^p \left\|f'\right\|_{[x,b],p}^p\right]$$
$$\leq \max\left\{\left(\frac{x-a}{b-a}\right)^p, \left(\frac{b-x}{b-a}\right)^p\right\}(b-a)^{p-1}\left\|f'\right\|_{[a,b],p}^p$$
$$= \left\{\frac{1}{2} + \left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right\}^p (b-a)^{p-1}\left\|f'\right\|_{[a,b],p}^p$$

provided $f' \in L_p[a,b]$.

Remark 5. If we take $p = 1$ in the above inequalities (53)–(54), then we obtain

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{55}$$
$$\leq \frac{1}{2(b-a)}\left[(x-a)^2\left\|f'\right\|_{[a,x],\infty} + (b-x)^2\left\|f'\right\|_{[x,b],\infty}\right]$$
$$\leq \frac{1}{2}\left[\left(\frac{x-a}{b-a}\right)^2 + \left(\frac{b-x}{b-a}\right)^2\right](b-a)\left\|f'\right\|_{[a,b],\infty}$$
$$= \left[\frac{1}{4} + \left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right](b-a)\left\|f'\right\|_{[a,b],\infty}$$

for $x \in [a,b]$, provided $f' \in L_\infty[a,b]$,

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{56}$$
$$\leq \frac{1}{(b-a)(\alpha+1)^{1/\alpha}}\left[(x-a)^{1+1/\alpha}\left\|f'\right\|_{[a,x],\beta} + (b-x)^{1+1/\alpha}\left\|f'\right\|_{[x,b],\beta}\right]$$
$$\leq \frac{1}{(\alpha+1)^{1/\alpha}}\left[\left(\frac{x-a}{b-a}\right)^{1+1/\alpha} + \left(\frac{b-x}{b-a}\right)^{1+1/\alpha}\right](b-a)^{1/\alpha}\left\|f'\right\|_{[a,b],\beta}$$

for $x \in [a,b]$, provided $f' \in L_\beta[a,b]$, $\alpha > 1$, $1/\alpha + 1/\beta = 1$ and

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{57}$$
$$\leq \frac{1}{b-a}\left[(x-a)\left\|f'\right\|_{[a,x],1} + (b-x)\left\|f'\right\|_{[x,b],1}\right]$$
$$= \left\{\frac{1}{2} + \left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right\}\left\|f'\right\|_{[a,b],1}$$

for $x \in [a,b]$.

Applications for p-Norms

We have the following inequalities for Lebesgue norms of the deviation of a function from its integral mean:

Theorem 9 (Dragomir, 2013 [9]). *Let* $f : [a, b] \to \mathbb{R}$ *be absolutely continuous on* $[a, b]$.

(i) If $f' \in L_\infty [a, b]$, *then*

$$\left\| f - \frac{1}{b-a} \int_a^b f(t)\, dt \right\|_{[a,b],p} \leq \left[\frac{2}{(p+1)(p+2)} \right]^{1/p} (b-a)^{1+\frac{1}{p}} \left\| f' \right\|_{[a,b],\infty} . \tag{58}$$

(ii) If $f' \in L_{p\beta} [a, b]$, *with* $\alpha > 1, 1/\alpha + 1/\beta = 1$, *then*

$$\begin{aligned} & \left\| f - \frac{1}{b-a} \int_a^b f(t)\, dt \right\|_{[a,b],p} \\ & \leq \left[\frac{2}{(p\alpha+1)^{1/\alpha} (p + 1/\alpha + 1)} \right]^{1/p} \left\| f' \right\|_{[a,b],p\beta} (b-a)^{1+\frac{1}{\alpha p}} . \end{aligned} \tag{59}$$

(iii) We have

$$\left\| f - \frac{1}{b-a} \int_a^b f(t)\, dt \right\|_{[a,b],p} \leq \frac{1}{2} \left(\frac{2^{p+1}-1}{p+1} \right)^{1/p} (b-a) \left\| f' \right\|_{[a,b],p} . \tag{60}$$

Proof. Integrating on $[a, b]$ the inequality (52) we have

$$\begin{aligned} & \int_a^b \left| f(x) - \frac{1}{b-a} \int_a^b f(t)\, dt \right|^p dx \\ & \leq \frac{1}{(b-a)(p+1)} \left\| f' \right\|_{[a,b],\infty}^p \int_a^b \left[(x-a)^{p+1} + (b-x)^{p+1} \right] dx \\ & = \frac{1}{(b-a)(p+1)} \left\| f' \right\|_{[a,b],\infty}^p \left[\frac{2(b-a)^{p+2}}{p+2} \right] \\ & = \frac{2}{(p+1)(p+2)} \left\| f' \right\|_{[a,b],\infty}^p (b-a)^{p+1} \end{aligned} \tag{61}$$

which is equivalent with (58).

Integrating the inequality (53)

$$\int_a^b \left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right|^p dx \tag{62}$$
$$\leq \frac{1}{(b-a)(p\alpha+1)^{1/\alpha}} \left\| f' \right\|_{[a,b],p\beta}^p \int_a^b \left[(x-a)^{p+1/\alpha} + (b-x)^{p+1/\alpha} \right] dx$$
$$= \frac{1}{(b-a)(p\alpha+1)^{1/\alpha}} \left\| f' \right\|_{[a,b],p\beta}^p \left[\frac{2(b-a)^{p+1/\alpha+1}}{p+1/\alpha+1} \right]$$
$$= \frac{2}{(p\alpha+1)^{1/\alpha}(p+1/\alpha+1)} \left\| f' \right\|_{[a,b],p\beta}^p (b-a)^{p+1/\alpha}$$

which is equivalent with (59).

Integrating the inequality (54) we have

$$\int_a^b \left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right|^p dx \tag{63}$$
$$\leq \frac{1}{b-a} \left\| f' \right\|_{[a,b],p}^p \int_a^b \max\left\{ (x-a)^p, (b-x)^p \right\} dx.$$

Since

$$\int_a^b \max\left\{ (x-a)^p, (b-x)^p \right\} dx$$
$$= \int_a^{\frac{a+b}{2}} (b-x)^p\,dx + \int_{\frac{a+b}{2}}^b (x-a)^p\,dx$$
$$= -\frac{\left(\frac{b-a}{2}\right)^{p+1}}{p+1} + \frac{(b-a)^{p+1}}{p+1} + \frac{(b-a)^{p+1}}{p+1} - \frac{\left(\frac{b-a}{2}\right)^{p+1}}{p+1}$$
$$= \frac{1}{p+1}\left(\frac{2^{p+1}-1}{2^p} \right)(b-a)^{p+1}$$

then from (63) we get

$$\int_a^b \left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right|^p dx \leq \frac{1}{p+1}\left(\frac{2^{p+1}-1}{2^p} \right)(b-a)^p \left\| f' \right\|_{[a,b],p}^p ,$$

which is equivalent with (60).

Applications for the Exponential

If we write the inequality (43) for the convex function $\Phi(x) = \exp(x)$, then we get the inequality

$$\exp\left[f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right] \tag{64}$$
$$\le \frac{1}{b-a}\left[\int_a^x \exp\left[(t-a)f'(t)\right]dt + \int_x^b \exp\left[(t-b)f'(t)\right]dt\right]$$

for $x \in [a,b]$.

If we write the inequality (43) for the convex function $\Phi(x) = \cosh(x) := \frac{e^x+e^{-x}}{2}$, then we get the inequality

$$\cosh\left[f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right] \tag{65}$$
$$\le \frac{1}{b-a}\left[\int_a^x \cosh\left[(t-a)f'(t)\right]dt + \int_x^b \cosh\left[(t-b)f'(t)\right]dt\right]$$

for $x \in [a,b]$.

Utilizing the inequality (64) we have the following multiplicative version of Ostrowski's inequality:

Theorem 10 (Dragomir, 2013 [9]). *Let $f : [a,b] \to (0,\infty)$ be absolutely continuous on $[a,b]$. Then we have the inequalities*

$$\frac{f(x)}{\exp\left[\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right]} \tag{66}$$
$$\le \frac{1}{b-a}\left[\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt + \int_x^b \exp\left[(t-b)\frac{f'(t)}{f(t)}\right]dt\right]$$

for any $x \in [a,b]$ and

$$\frac{\int_a^b f(x)\,dx}{\exp\left[\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right]} \tag{67}$$
$$\le \frac{1}{b-a}\left[\int_a^b (b-x)\exp\left[(x-a)\frac{f'(x)}{f(x)}\right]dx\right.$$
$$\left.+\int_a^b (x-a)\exp\left[(x-b)\frac{f'(x)}{f(x)}\right]dx\right].$$

Proof. If we replace f by $\ln f$ in (64), we get

$$\exp\left[\ln f(x) - \frac{1}{b-a}\int_a^b \ln f(t)\,dt\right] \leq \frac{1}{b-a}\left[\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt + \int_x^b \exp\left[(t-b)\frac{f'(t)}{f(t)}\right]dt\right] \tag{68}$$

for any $x \in [a,b]$.

Since

$$\begin{aligned}
&\exp\left[\ln f(x) - \frac{1}{b-a}\int_a^b \ln f(t)\,dt\right] \\
&= \exp\left[\ln f(x) - \ln\left\{\exp\left(\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right)\right\}\right] \\
&= \exp\left[\ln\left(\frac{f(x)}{\exp\left(\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right)}\right)\right] \\
&= \frac{f(x)}{\exp\left(\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right)}
\end{aligned}$$

for any $x \in [a,b]$, then we get from (68) the desired inequality (66).

If we integrate the inequality (66), we get

$$\frac{\int_a^b f(x)\,dx}{\exp\left[\frac{1}{b-a}\int_a^b \ln f(t)\,dt\right]} \leq \frac{1}{b-a}\int_a^b \left[\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt + \int_x^b \exp\left[(t-b)\frac{f'(t)}{f(t)}\right]dt\right]dx. \tag{69}$$

Integrating by parts we have

$$\begin{aligned}
&\int_a^b \left(\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt\right)dx \\
&= x\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt\bigg|_a^b - \int_a^b x d\left(\int_a^x \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt\right) \\
&= b\int_a^b \exp\left[(t-a)\frac{f'(t)}{f(t)}\right]dt - \int_a^b x\exp\left[(x-a)\frac{f'(x)}{f(x)}\right]dx \\
&= \int_a^b (b-x)\exp\left[(x-a)\frac{f'(x)}{f(x)}\right]dx
\end{aligned}$$

and

$$
\begin{aligned}
&\int_a^b \left(\int_x^b \exp\left[(t-b) \frac{f'(t)}{f(t)} \right] dt \right) dx \\
&= x \int_x^b \exp\left[(t-b) \frac{f'(t)}{f(t)} \right] dt \Big|_a^b - \int_a^b x d \left(\int_x^b \exp\left[(t-b) \frac{f'(t)}{f(t)} \right] dt \right) \\
&= -a \int_a^b \exp\left[(t-b) \frac{f'(t)}{f(t)} \right] dt + \int_a^b x \exp\left[(x-b) \frac{f'(x)}{f(x)} \right] dx \\
&= \int_a^b (x-a) \exp\left[(x-b) \frac{f'(x)}{f(x)} \right] dx,
\end{aligned}
$$

then by (69) we deduce the desired inequality (67).

Applications for Midpoint-Inequalities

We have from the inequality (43) written for $-f$ the following result:

Proposition 1 (Dragomir, 2013 [9]). *Let $f : [a,b] \to \mathbb{R}$ be absolutely continuous on $[a,b]$. If $\Phi : \mathbb{R} \to \mathbb{R}$ is convex (concave) on $\mathbb{R}$, then from (43) we have the inequalities*

$$
\begin{aligned}
&\Phi \left(\frac{1}{b-a} \int_a^b f(t)\, dt - f\left(\frac{a+b}{2} \right) \right) \qquad (70) \\
&\le (\ge) \frac{1}{b-a} \left[\int_{\frac{a+b}{2}}^b \Phi \left[(b-t) f'(t) \right] dt + \int_a^{\frac{a+b}{2}} \Phi \left[(a-t) f'(t) \right] dt \right].
\end{aligned}
$$

If $f : [a,b] \to \mathbb{R}$ is convex on $[a,b]$, then by Hermite–Hadamard inequality we have

$$
\frac{1}{b-a} \int_a^b f(t)\, dt \ge f\left(\frac{a+b}{2} \right).
$$

We can state the following result in which the function Φ is assumed to be convex only on $[0,\infty)$ or $(0,\infty)$.

Proposition 2. *If $f : [a,b] \to \mathbb{R}$ is convex on $[a,b]$, monotonic nondecreasing on $\left[a, \frac{a+b}{2}\right]$ and monotonic nonincreasing $\left[a, \frac{a+b}{2}\right]$. If $\Phi : [0,\infty), (0,\infty) \to \mathbb{R}$ is convex (concave) on $[0,\infty)$ or $(0,\infty)$, then (70) holds true.*

If $f : [a,b] \to \mathbb{R}$ is strictly convex on $[a,b]$, monotonic nondecreasing on $\left[a, \frac{a+b}{2}\right]$ and monotonic nonincreasing $\left[a, \frac{a+b}{2}\right]$, then by taking $\Phi(x) = \ln x$, which is strictly concave on $(0, \infty)$, we get the logarithmic inequality

$$\ln\left(\frac{1}{b-a}\int_a^b f(t)\,dt - f\left(\frac{a+b}{2}\right)\right) \tag{71}$$
$$\geq \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \ln\left[(b-t)f'(t)\right]dt + \int_a^{\frac{a+b}{2}} \ln\left[(a-t)f'(t)\right]dt\right].$$

If $f : [a, b] \to \mathbb{R}$ is convex on $[a, b]$, monotonic nondecreasing on $\left[a, \frac{a+b}{2}\right]$ and monotonic nonincreasing $\left[a, \frac{a+b}{2}\right]$, then by taking $\Phi(x) = x^q$, with $q \in (0, 1)$ we also have

$$\left(\frac{1}{b-a}\int_a^b f(t)\,dt - f\left(\frac{a+b}{2}\right)\right)^q \tag{72}$$
$$\geq \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \left[(b-t)f'(t)\right]^q dt + \int_a^{\frac{a+b}{2}} \left[(a-t)f'(t)\right]^q dt\right].$$

If $\Phi : [0, \infty), (0, \infty) \to \mathbb{R}$ is convex (concave) on $[0, \infty)$ or $(0, \infty)$, and if we take $f(t) := \left|t - \frac{a+b}{2}\right|^p, p \geq 1$, then we get from (70)

$$\Phi\left(\frac{(b-a)^p}{2^p(p+1)}\right) \leq (\geq) \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \Phi\left[p(b-t)\left(t - \frac{a+b}{2}\right)^{p-1}\right]dt \right. \tag{73}$$
$$\left. + \int_a^{\frac{a+b}{2}} \Phi\left[(t-a)\left(\frac{a+b}{2} - t\right)^{p-1}\right]dt\right].$$

Assume that $\Phi : \mathbb{R} \to \mathbb{R}$ is convex (concave) on $\mathbb{R}$.

Now, if we take $f(t) = \frac{1}{t}$ in (70), where $t \in [a, b] \subset (0, \infty)$, then we have

$$\Phi\left(\frac{A(a,b) - L(a,b)}{A(a,b)\,L(a,b)}\right) \tag{74}$$
$$\leq (\geq) \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \Phi\left(\frac{t-b}{t^2}\right)dt + \int_a^{\frac{a+b}{2}} \Phi\left(\frac{t-a}{t^2}\right)dt\right].$$

If we take $f(t) = -\ln t$ in (70), where $t \in [a, b] \subset (0, \infty)$, then we have

$$\Phi\left(\ln\left(\frac{A(a,b)}{I(a,b)}\right)\right) \tag{75}$$
$$\leq (\geq) \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \Phi\left(\frac{t-b}{t}\right)dt + \int_a^{\frac{a+b}{2}} \Phi\left(\frac{t-a}{t}\right)dt\right].$$

If we take $f(t) = t^p$, $p \in \mathbb{R}\setminus\{0,-1\}$ in (70), where $t \in [a,b] \subset (0,\infty)$, then we have

$$\Phi\left(L_p^p(a,b) - A^p(a,b)\right) \tag{76}$$
$$\leq (\geq) \frac{1}{b-a}\left[\int_{\frac{a+b}{2}}^{b} \Phi\left[p(b-t)t^{p-1}\right]dt + \int_a^{\frac{a+b}{2}} \Phi\left[p(a-t)t^{p-1}\right]dt\right].$$

Perturbed Ostrowski Type Inequalities for Functions of Bounded Variation

Some Identities

We start with the following identity that will play an important role in the following:

Lemma 1 (Dragomir, 2013 [10]). *Let $f : [a,b] \to \mathbb{C}$ be a function of bounded variation on $[a,b]$ and $x \in [a,b]$. Then for any $\lambda_1(x)$ and $\lambda_2(x)$ complex numbers, we have*

$$f(x) + \frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x) - (x-a)^2\lambda_1(x)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{77}$$
$$= \frac{1}{b-a}\int_a^x (t-a)\,d[f(t) - \lambda_1(x)t] + \frac{1}{b-a}\int_x^b (t-b)\,d[f(t) - \lambda_2(x)t],$$

where the integrals in the right-hand side are taken in the Riemann–Stieltjes sense.

Proof. Utilizing the integration by parts formula in the Riemann–Stieltjes integral, we have

$$\int_a^x (t-a)\,d[f(t) - \lambda_1(x)t] \tag{78}$$
$$= (t-a)[f(t) - \lambda_1(x)t]\Big|_a^x - \int_a^x [f(t) - \lambda_1(x)t]\,dt$$
$$= (x-a)[f(x) - \lambda_1(x)x] - \int_a^x f(t)\,dt + \frac{1}{2}\lambda_1(x)\left(x^2 - a^2\right)$$
$$= (x-a)f(x) - \lambda_1(x)x(x-a) - \int_a^x f(t)\,dt + \frac{1}{2}\lambda_1(x)\left(x^2 - a^2\right)$$
$$= (x-a)f(x) - \int_a^x f(t)\,dt - \frac{1}{2}(x-a)^2\lambda_1(x)$$

and

$$\int_x^b (t-b)\, d\left[f(t) - \lambda_2(x)\, t\right] \tag{79}$$

$$= (t-b)\left[f(t) - \lambda_2(x)\, t\right]\Big|_x^b - \int_x^b \left[f(t) - \lambda_2(x)\, t\right] dt$$

$$= (b-x)\left[f(x) - \lambda_2(x)\, x\right] - \int_x^b f(t)\, dt + \frac{1}{2}\lambda_2(x)\left(b^2 - x^2\right)$$

$$= (b-x) f(x) - \int_x^b f(t)\, dt - (b-x)\,\lambda_2(x)\, x + \frac{1}{2}\lambda_2(x)\left(b^2 - x^2\right)$$

$$= (b-x) f(x) - \int_x^b f(t)\, dt + \frac{1}{2}(b-x)^2 \lambda_2(x).$$

By adding the equalities (78) and (79) and dividing by $b-a$ we get the desired representation (77).

Corollary 11. *With the assumption in Lemma 1, we have for any $\lambda(x) \in \mathbb{C}$ that*

$$f(x) + \left(\frac{a+b}{2} - x\right)\lambda(x) - \frac{1}{b-a}\int_a^b f(t)\, dt \tag{80}$$

$$= \frac{1}{b-a}\int_a^x (t-a)\, d\left[f(t) - \lambda(x)\, t\right] + \frac{1}{b-a}\int_x^b (t-b)\, d\left[f(t) - \lambda(x)\, t\right].$$

We have the following midpoint representation:

Corollary 12. *With the assumption in Lemma 1, we have for any $\lambda_1, \lambda_2 \in \mathbb{C}$ that*

$$f\left(\frac{a+b}{2}\right) + \frac{1}{8}(b-a)(\lambda_2 - \lambda_1) - \frac{1}{b-a}\int_a^b f(t)\, dt \tag{81}$$

$$= \frac{1}{b-a}\int_a^{\frac{a+b}{2}} (t-a)\, d\left[f(t) - \lambda_1 t\right]$$

$$+ \frac{1}{b-a}\int_{\frac{a+b}{2}}^b (t-b)\, d\left[f(t) - \lambda_2 t\right].$$

In particular, if $\lambda_1 = \lambda_2 = \lambda$, then we have the equality

$$f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\, dt \tag{82}$$

$$= \frac{1}{b-a}\int_a^{\frac{a+b}{2}} (t-a)\, d\left[f(t) - \lambda t\right] + \frac{1}{b-a}\int_{\frac{a+b}{2}}^b (t-b)\, d\left[f(t) - \lambda t\right].$$

Remark 6. If we take $\lambda(x) = 0$ in (80) we recapture the Montgomery type identity established in [2].

Inequalities for Functions of Bounded Variation

The following lemma will be used in the sequel and is of interest in itself as well [1, p. 177]. For a simple proof, see also [8].

Lemma 2. *Let $f, u : [a,b] \to \mathbb{C}$. If f is continuous on $[a,b]$ and u is of bounded variation on $[a,b]$, then the Riemann–Stieltjes integral $\int_a^b f(t)\,du(t)$ exists and*

$$\left|\int_a^b f(t)\,du(t)\right| \le \int_a^b |f(t)|\,d\left(\bigvee_a^t(u)\right) \le \max_{t\in[a,b]}|f(t)|\bigvee_a^b(u). \tag{83}$$

We denote by $\ell : [a,b] \to [a,b]$ the identity function, namely $\ell(t) = t$ for any $t \in [a,b]$.

We have the following result:

Theorem 11 (Dragomir, 2013 [10]). *Let $f : [a,b] \to \mathbb{C}$ be a function of bounded variation on $[a,b]$ and $x \in [a,b]$. Then for any $\lambda_1(x)$ and $\lambda_2(x)$ complex numbers, we have the inequality*

$$\left|f(x) + \frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x) - (x-a)^2\lambda_1(x)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{84}$$

$$\le \frac{1}{b-a}\left[\int_a^x\left(\bigvee_t^x(f-\lambda_1(x)\ell)\right)dt + \int_x^b\left(\bigvee_x^t(f-\lambda_2(x)\ell)\right)dt\right]$$

$$\le \frac{1}{b-a}\left[(x-a)\bigvee_a^x(f-\lambda_1(x)\ell) + (b-x)\bigvee_x^b(f-\lambda_2(x)\ell)\right]$$

$$\le \begin{cases} \max\left\{\bigvee_a^x(f-\lambda_1(x)\ell), \bigvee_x^b(f-\lambda_2(x)\ell)\right\} \\[2ex] \left[\frac{1}{2} + \left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]\left(\bigvee_a^x(f-\lambda_1(x)\ell) + \bigvee_x^b(f-\lambda_2(x)\ell)\right), \end{cases}$$

where $\bigvee_{c}^{d}(g)$ *denotes the total variation of* g *on the interval* $[c,d]$ *.*

Proof. Taking the modulus in (77) and using the property (83) we have

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x)-(x-a)^2\lambda_1(x)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{85}$$
$$\leq \frac{1}{b-a}\left|\int_a^x (t-a)\,d\left[f(t)-\lambda_1(x)\,t\right]\right|$$
$$+\frac{1}{b-a}\left|\int_x^b (t-b)\,d\left[f(t)-\lambda_2(x)\,t\right]\right|$$
$$\leq \frac{1}{b-a}\int_a^x (t-a)\,d\left(\bigvee_a^t(f-\lambda_1(x)\,\ell)\right)$$
$$+\frac{1}{b-a}\int_x^b (b-t)\,d\left(\bigvee_a^t(f-\lambda_2(x)\,\ell)\right).$$

Integrating by parts in the Riemann–Stieltjes integral, we have

$$\int_a^x (t-a)\,d\left(\bigvee_a^t(f-\lambda_1(x)\,\ell)\right)$$
$$=(t-a)\bigvee_a^t(f-\lambda_1(x)\,\ell)\Big|_a^x-\int_a^x\left(\bigvee_a^t(f-\lambda_1(x)\,\ell)\right)dt$$
$$=(x-a)\bigvee_a^x(f-\lambda_1(x)\,\ell)-\int_a^x\left(\bigvee_a^t(f-\lambda_1(x)\,\ell)\right)dt$$
$$=\int_a^x\left(\bigvee_t^x(f-\lambda_1(x)\,\ell)\right)dt$$

and

$$\int_x^b (b-t)\,d\left(\bigvee_a^t(f-\lambda_2(x)\,\ell)\right)$$
$$=(b-t)\bigvee_a^t(f-\lambda_2(x)\,\ell)\Big|_x^b+\int_x^b\left(\bigvee_a^t(f-\lambda_2(x)\,\ell)\right)dt$$
$$=\int_x^b\left(\bigvee_a^t(f-\lambda_2(x)\,\ell)\right)dt-(b-x)\bigvee_a^x(f-\lambda_2(x)\,\ell)$$

$$= \int_x^b \left(\bigvee_x^t (f - \lambda_2(x)\ell) \right) dt.$$

Using (85) we deduce the first inequality in (84).

We also have

$$\int_a^x \left(\bigvee_t^x (f - \lambda_1(x)\ell) \right) dt \leq (x-a) \bigvee_a^x (f - \lambda_1(x)\ell)$$

and

$$\int_x^b \left(\bigvee_x^t (f - \lambda_2(x)\ell) \right) dt \leq (b-x) \bigvee_x^b (f - \lambda_2(x)\ell),$$

which prove the second inequality in (84).

The last part is obvious.

The following result holds:

Corollary 13. *Let $f : [a,b] \to \mathbb{C}$ be a function of bounded variation on $[a,b]$ and $x \in [a,b]$. Then for any $\lambda(x)$ a complex number, we have the inequality*

$$\begin{aligned} &\left| f(x) + \left(\frac{a+b}{2} - x \right) \lambda(x) - \frac{1}{b-a} \int_a^b f(t)\,dt \right| && (86) \\ &\leq \frac{1}{b-a} \left[\int_a^x \left(\bigvee_t^x (f - \lambda(x)\ell) \right) dt + \int_x^b \left(\bigvee_x^t (f - \lambda(x)\ell) \right) dt \right] \\ &\leq \frac{1}{b-a} \left[(x-a) \bigvee_a^x (f - \lambda(x)\ell) + (b-x) \bigvee_x^b (f - \lambda(x)\ell) \right] \\ &\leq \begin{cases} \frac{1}{2} \bigvee_a^b (f - \lambda(x)\ell) + \frac{1}{2} \left| \bigvee_x^b (f - \lambda(x)\ell) - \bigvee_a^x (f - \lambda(x)\ell) \right| \\ \left[\frac{1}{2} + \left| \frac{x - \frac{a+b}{2}}{b-a} \right| \right] \bigvee_a^b (f - \lambda(x)\ell). \end{cases} \end{aligned}$$

Remark 7. Let $f : [a,b] \to \mathbb{C}$ be a function of bounded variation on $[a,b]$. Then for any $\lambda \in \mathbb{C}$ we have the inequalities

$$\left| f\left(\frac{a+b}{2} \right) - \frac{1}{b-a} \int_a^b f(t)\,dt \right| \qquad (87)$$

$$\leq \frac{1}{b-a}\left[\int_a^{\frac{a+b}{2}}\left(\bigvee_t^{\frac{a+b}{2}}(f-\lambda\ell)\right)dt+\int_{\frac{a+b}{2}}^b\left(\bigvee_{\frac{a+b}{2}}^t(f-\lambda\ell)\right)dt\right]$$
$$\leq \frac{1}{2}\bigvee_a^b(f-\lambda\ell).$$

This is equivalent to

$$\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{88}$$
$$\leq \frac{1}{b-a}\inf_{\lambda\in\mathbb{C}}\left[\int_a^{\frac{a+b}{2}}\left(\bigvee_t^{\frac{a+b}{2}}(f-\lambda\ell)\right)dt+\int_{\frac{a+b}{2}}^b\left(\bigvee_{\frac{a+b}{2}}^t(f-\lambda\ell)\right)dt\right]$$
$$\leq \frac{1}{2}\inf_{\lambda\in\mathbb{C}}\left[\bigvee_a^b(f-\lambda\ell)\right].$$

Inequalities for Lipshitzian Functions

We can state the following result:

Theorem 12 (Dragomir, 2013 [10]). *Let $f:[a,b]\to\mathbb{C}$ be a bounded function on $[a,b]$ and $x\in(a,b)$. If $\lambda_1(x)$ and $\lambda_2(x)$ are complex numbers and there exist the positive numbers $L_1(x)$ and $L_2(x)$ such that $f-\lambda_1(x)\ell$ is Lipschitzian with the constant $L_1(x)$ on the interval $[a,x]$ and $f-\lambda_2(x)\ell$ is Lipschitzian with the constant $L_2(x)$ on the interval $[x,b]$, then*

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x)-(x-a)^2\lambda_1(x)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{89}$$

$$\leq \frac{1}{2}\left[\left(\frac{x-a}{b-a}\right)^2 L_1(x)+\left(\frac{b-x}{b-a}\right)^2 L_2(x)\right](b-a)$$

$$\leq \begin{cases} \left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^{2}\right]\max\left\{L_{1}\left(x\right),L_{2}\left(x\right)\right\}\left(b-a\right), \\ \\ \frac{1}{2}\left[\left(\frac{x-a}{b-a}\right)^{2q}+\left(\frac{b-x}{b-a}\right)^{2q}\right]^{1/q}\left(L_{1}^{p}\left(x\right)+L_{2}^{p}\left(x\right)\right)^{1/p}\left(b-a\right), \\ p>1, \frac{1}{p}+\frac{1}{q}=1, \\ \\ \left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]\frac{L_{1}\left(x\right)+L_{2}\left(x\right)}{2}\left(b-a\right). \end{cases}$$

Proof. It is known that if $g : [c,d] \to \mathbb{C}$ is Riemann integrable and $u : [c,d] \to \mathbb{C}$ is Lipschitzian with the constant $L > 0$, then the Riemann–Stieltjes integral $\int_c^d g(t)\,du(t)$ exists and

$$\left|\int_{c}^{d} g\left(t\right)du\left(t\right)\right| \leq L\int_{c}^{d}\left|g\left(t\right)\right|dt. \tag{90}$$

Taking the modulus in (77) and using the property (90) we have

$$\begin{aligned} &\left|f\left(x\right)+\frac{1}{2\left(b-a\right)}\left[\left(b-x\right)^{2}\lambda_{2}\left(x\right)-\left(x-a\right)^{2}\lambda_{1}\left(x\right)\right]-\frac{1}{b-a}\int_{a}^{b}f\left(t\right)dt\right| \\ &\leq \frac{1}{b-a}\left|\int_{a}^{x}\left(t-a\right)d\left[f\left(t\right)-\lambda_{1}\left(x\right)t\right]\right| \\ &+\frac{1}{b-a}\left|\int_{x}^{b}\left(t-b\right)d\left[f\left(t\right)-\lambda_{2}\left(x\right)t\right]\right| \\ &\leq \frac{1}{b-a}\left[L_{1}\left(x\right)\int_{a}^{x}\left(t-a\right)dt+L_{2}\left(x\right)\int_{x}^{b}\left(b-t\right)dt\right] \\ &=\frac{L_{1}\left(x\right)\left(x-a\right)^{2}+L_{2}\left(x\right)\left(b-x\right)^{2}}{2\left(b-a\right)} \\ &=\frac{1}{2}\left[L_{1}\left(x\right)\left(\frac{x-a}{b-a}\right)^{2}+L_{2}\left(x\right)\left(\frac{b-x}{b-a}\right)^{2}\right]\left(b-a\right), \end{aligned}$$

and the first inequality in (89) is proved.

By Hölder's inequality we have

$$L_{1}\left(x\right)\left(\frac{x-a}{b-a}\right)^{2}+L_{2}\left(x\right)\left(\frac{b-x}{b-a}\right)^{2}$$

$$\leq \begin{cases} \left[\left(\frac{x-a}{b-a}\right)^2 + \left(\frac{b-x}{b-a}\right)^2\right] \max\{L_1(x), L_2(x)\} \\ \\ \left[\left(\frac{x-a}{b-a}\right)^{2q} + \left(\frac{b-x}{b-a}\right)^{2q}\right]^{1/q} \left(L_1^p(x) + L_2^p(x)\right)^{1/p}, \\ p > 1, \frac{1}{p} + \frac{1}{q} = 1 \\ \\ \max\left\{\left(\frac{x-a}{b-a}\right)^2, \left(\frac{b-x}{b-a}\right)^2\right\} [L_1(x) + L_2(x)], \end{cases}$$

which proves, upon simple calculations, the last part of the inequality (89).

Corollary 14. *Let $f : [a,b] \to \mathbb{C}$ be a bounded function on $[a,b]$ and $x \in (a,b)$. If $\lambda(x)$ is a complex number and there exist the positive number $L(x)$ such that $f - \lambda(x)\ell$ is Lipschitzian with the constant $L(x)$ on the interval $[a,b]$, then*

$$\left|f(x) + \left(\frac{a+b}{2} - x\right)\lambda(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{91}$$
$$\leq \left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a}\right)^2\right] L(x)(b-a).$$

Remark 8. If λ is a complex number and there exists the positive number L such that $f - \lambda\ell$ is Lipschitzian with the constant L on the interval $[a,b]$, then

$$\left|f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \frac{1}{4}L(b-a). \tag{92}$$

Inequalities for Monotonic Functions

Now, the case of monotonic integrators is as follows:

Theorem 13 (Dragomir, 2013 [10]). *Let $f : [a,b] \to \mathbb{R}$ be a bounded function on $[a,b]$ and $x \in (a,b)$. If $\lambda_1(x)$ and $\lambda_2(x)$ are real numbers such that $f - \lambda_1(x)\ell$ is monotonic nondecreasing on the interval $[a,x]$ and $f - \lambda_2(x)\ell$ is monotonic nondecreasing on the interval $[x,b]$, then*

$$\left|f(x) + \frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x) - (x-a)^2\lambda_1(x)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{93}$$

$$\leq \frac{1}{b-a}\left[(2x-a-b)f(x) + \int_a^b sgn(t-x)f(t)\,dt\right.$$

$$-\frac{1}{2}\left[\lambda_1(x)(x-a)^2+\lambda_2(x)(b-x)^2\right]\Bigg]$$

$$\le \frac{1}{b-a}\{(x-a)[f(x)-f(a)-\lambda_1(x)(x-a)]$$
$$+(b-x)[f(b)-f(x)-\lambda_2(x)(b-x)]\}$$

$$\le \begin{cases} \frac{1}{2}[f(b)-f(a)-\lambda_1(x)(x-a)-\lambda_2(x)(b-x)] \\ \quad +\left|f(x)-\frac{f(a)+f(b)}{2}-\frac{1}{2}\lambda_1(x)(x-a)+\frac{1}{2}\lambda_2(x)(b-x)\right|, \\ \\ \left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right] \\ \times[f(b)-f(a)-\lambda_1(x)(x-a)-\lambda_2(x)(b-x)]. \end{cases}$$

Proof. It is known that if $g:[c,d]\to\mathbb{C}$ is continuous and $u:[c,d]\to\mathbb{C}$ is monotonic nondecreasing, then the Riemann–Stieltjes integral $\int_c^d g(t)\,du(t)$ exists and

$$\left|\int_c^d g(t)\,du(t)\right| \le \int_c^d |g(t)|\,du(t). \tag{94}$$

Taking the modulus in (77) and using the property (94) we have

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x)-(x-a)^2\lambda_1(x)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{95}$$
$$\le \frac{1}{b-a}\left|\int_a^x (t-a)\,d[f(t)-\lambda_1(x)t]\right|$$
$$+\frac{1}{b-a}\left|\int_x^b (t-b)\,d[f(t)-\lambda_2(x)t]\right|$$
$$\le \frac{1}{b-a}\int_a^x (t-a)\,d[f(t)-\lambda_1(x)t]$$
$$+\frac{1}{b-a}\int_x^b (b-t)\,d[f(t)-\lambda_2(x)t].$$

Integrating by parts in the Riemann–Stieltjes integral we have

$$\int_a^x (t-a)\,d[f(t)-\lambda_1(x)t]$$
$$=(t-a)[f(t)-\lambda_1(x)t]\Big|_a^x-\int_a^x [f(t)-\lambda_1(x)t]\,dt$$

$$= (x-a)\left[f(x) - \lambda_1(x)x\right] - \int_a^x \left[f(t) - \lambda_1(x)t\right] dt$$

$$= (x-a)f(x) - \lambda_1(x)x(x-a) - \int_a^x f(t)\,dt + \lambda_1(x)\frac{x^2-a^2}{2}$$

$$= (x-a)f(x) - \int_a^x f(t)\,dt - \frac{1}{2}\lambda_1(x)(x-a)^2$$

and

$$\int_x^b (b-t)\,d\left[f(t) - \lambda_2(x)t\right]$$

$$= (b-t)\left[f(t) - \lambda_2(x)t\right]\Big|_x^b + \int_x^b \left[f(t) - \lambda_2(x)t\right] dt$$

$$= \int_x^b f(t)\,dt - \lambda_2(x)\int_x^b tdt - (b-x)\left[f(x) - \lambda_2(x)x\right]$$

$$= \int_x^b f(t)\,dt - \lambda_2(x)\frac{b^2-x^2}{2} - (b-x)f(x) + (b-x)\lambda_2(x)x$$

$$= \int_x^b f(t)\,dt - (b-x)f(x) - \frac{1}{2}\lambda_2(x)(b-x)^2.$$

If we add these equalities, we get

$$\int_a^x (t-a)\,d\left[f(t) - \lambda_1(x)t\right] + \int_x^b (b-t)\,d\left[f(t) - \lambda_2(x)t\right]$$

$$= (x-a)f(x) - \int_a^x f(t)\,dt - \frac{1}{2}\lambda_1(x)(x-a)^2$$

$$+ \int_x^b f(t)\,dt - (b-x)f(x) - \frac{1}{2}\lambda_2(x)(b-x)^2$$

$$= (2x-a-b)f(x) + \int_a^b sgn(t-x)f(t)\,dt$$

$$- \frac{1}{2}\left[\lambda_1(x)(x-a)^2 + \lambda_2(x)(b-x)^2\right]$$

and by (95) we get the first inequality in (93).

Now, since $f - \lambda_1(x)\,\ell$ is monotonic nondecreasing on the interval $[a, x]$, then

$$\begin{aligned}&\int_a^x (t-a)\,d\left[f(t)-\lambda_1(x)\,t\right]\\&\le (x-a)\left[f(x)-\lambda_1(x)\,x-f(a)+\lambda_1(x)\,a\right]\\&=(x-a)\left[f(x)-f(a)-\lambda_1(x)(x-a)\right]\end{aligned}$$

and, since $f-\lambda_2(x)\,\ell$ is monotonic nondecreasing on the interval $[x,b]$, then also

$$\begin{aligned}&\int_x^b (b-t)\,d\left[f(t)-\lambda_2(x)\,t\right]\\&\le (b-x)\left[f(b)-\lambda_2(x)\,b-f(x)+\lambda_2(x)\,x\right]\\&=(b-x)\left[f(b)-f(x)-\lambda_2(x)(b-x)\right].\end{aligned}$$

These prove the second inequality in (93).

The last part follows by the properties of maximum and the details are omitted.

Corollary 15. *Let* $f:[a,b]\to\mathbb{R}$ *be a bounded function on* $[a,b]$ *and* $x\in(a,b)$. *If* $\lambda(x)$ *is a real number such that* $f-\lambda(x)\,\ell$ *is monotonic nondecreasing on the interval* $[a,b]$, *then*

$$\left|f(x)+\left(\frac{a+b}{2}-x\right)\lambda(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{96}$$

$$\begin{aligned}&\le\frac{1}{b-a}\left[(2x-a-b)f(x)+\int_a^b sgn\,(t-x)f(t)\,dt\right.\\&\left.-\left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right](b-a)^2\lambda(x)\right]\end{aligned}$$

$$\begin{aligned}&\le\frac{1}{b-a}\left\{(x-a)\left[f(x)-f(a)-\lambda(x)(x-a)\right]\right.\\&\left.+(b-x)\left[f(b)-f(x)-\lambda(x)(b-x)\right]\right\}\end{aligned}$$

$$\leq \begin{cases} \frac{f(b)-f(a)}{2} - \frac{1}{2}\lambda(x)(b-a) \\ + \left| f(x) - \frac{f(a)+f(b)}{2} - \frac{1}{2}\lambda(x)(2x-a-b) \right|, \\ \left[\frac{1}{2} + \left| \frac{x-\frac{a+b}{2}}{b-a} \right| \right] \\ \times [f(b) - f(a) - \lambda(x)(b-a)]. \end{cases}$$

Remark 9. If λ is a real number such that $f - \lambda\ell$ is monotonic nondecreasing on the interval $[a,b]$, then

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{97}$$
$$\leq \frac{1}{b-a}\left[\int_a^b sgn\left(t - \frac{a+b}{2}\right) f(t)\,dt - \frac{1}{4}\lambda(b-a)^2 \right]$$
$$\leq \frac{1}{2}[f(b) - f(a) - \lambda(b-a)].$$

Some Perturbed Ostrowski Type Inequalities for Absolutely Continuous Functions

Some Identities

We start with the following identity that will play an important role in the following:

Lemma 3 (Dragomir, 2013 [11]). *Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous on $[a,b]$ and $x \in [a,b]$. Then for any $\lambda_1(x)$ and $\lambda_2(x)$ complex numbers, we have*

$$f(x) + \frac{1}{2(b-a)}\left[(b-x)^2\lambda_2(x) - (x-a)^2\lambda_1(x)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{98}$$
$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(t) - \lambda_1(x)\right]dt + \frac{1}{b-a}\int_x^b (t-b)\left[f'(t) - \lambda_2(x)\right]dt,$$

where the integrals in the right-hand side are taken in the Lebesgue sense.

Proof. Utilizing the integration by parts formula in the Lebesgue integral, we have

$$\int_a^x (t-a)\left[f'(t) - \lambda_1(x)\right]dt \tag{99}$$
$$= (t-a)[f(t) - \lambda_1(x)t]\big|_a^x - \int_a^x [f(t) - \lambda_1(x)t]\,dt$$

$$= (x-a)\left[f(x) - \lambda_1(x)x\right] - \int_a^x f(t)\,dt + \frac{1}{2}\lambda_1(x)\left(x^2 - a^2\right)$$

$$= (x-a)f(x) - \lambda_1(x)x(x-a) - \int_a^x f(t)\,dt + \frac{1}{2}\lambda_1(x)\left(x^2 - a^2\right)$$

$$= (x-a)f(x) - \int_a^x f(t)\,dt - \frac{1}{2}(x-a)^2\lambda_1(x)$$

and

$$\int_x^b (t-b)\left[f'(t) - \lambda_2(x)\right]dt \tag{100}$$

$$= (t-b)\left[f(t) - \lambda_2(x)t\right]\Big|_x^b - \int_x^b \left[f(t) - \lambda_2(x)t\right]dt$$

$$= (b-x)\left[f(x) - \lambda_2(x)x\right] - \int_x^b f(t)\,dt + \frac{1}{2}\lambda_2(x)\left(b^2 - x^2\right)$$

$$= (b-x)f(x) - \int_x^b f(t)\,dt - (b-x)\lambda_2(x)x + \frac{1}{2}\lambda_2(x)\left(b^2 - x^2\right)$$

$$= (b-x)f(x) - \int_x^b f(t)\,dt + \frac{1}{2}(b-x)^2\lambda_2(x).$$

If we add the identities (99) and (100) and divide by $b-a$, we deduce the desired identity (98).

Corollary 16. *With the assumption in Lemma 3, we have for any $\lambda(x) \in \mathbb{C}$ that*

$$f(x) + \left(\frac{a+b}{2} - x\right)\lambda(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{101}$$

$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(t) - \lambda(x)\right]dt + \frac{1}{b-a}\int_x^b (t-b)\left[f'(t) - \lambda(x)\right]dt.$$

Remark 10. If we take $\lambda(x) = 0$ in (101), then we get Montgomery's identity for absolutely continuous functions, i.e.

$$f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{102}$$

$$= \frac{1}{b-a}\int_a^x (t-a)f'(t)\,dt + \frac{1}{b-a}\int_x^b (t-b)f'(t)\,dt,$$

for $x \in [a,b]$.

We have the following midpoint representation:

Corollary 17. *With the assumption in Lemma 3, we have for any $\lambda_1, \lambda_2 \in \mathbb{C}$ that*

$$f\left(\frac{a+b}{2}\right)+\frac{1}{8}(b-a)(\lambda_2-\lambda_1)-\frac{1}{b-a}\int_a^b f(t)\,dt \tag{103}$$
$$=\frac{1}{b-a}\int_a^{\frac{a+b}{2}}(t-a)\left[f'(t)-\lambda_1\right]dt+\frac{1}{b-a}\int_{\frac{a+b}{2}}^b (t-b)\left[f'(t)-\lambda_2\right]dt.$$

In particular, if $\lambda_1=\lambda_2=\lambda$, then we have the equality

$$f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt \tag{104}$$
$$=\frac{1}{b-a}\int_a^{\frac{a+b}{2}}(t-a)\left[f'(t)-\lambda\right]dt+\frac{1}{b-a}\int_{\frac{a+b}{2}}^b (t-b)\left[f'(t)-\lambda\right]dt.$$

Remark 11. The identity (98) has many particular cases of interest.

If we assume that the derivatives $f'_+(a), f'_-(b)$ and $f'(x)$ exist and are finite, then by taking

$$\lambda_1(x)=\frac{f'_+(a)+f'(x)}{2} \text{ and } \lambda_2(x)=\frac{f'(x)+f'_-(b)}{2}$$

in (98) we get

$$f(x)+\frac{1}{2}\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt \tag{105}$$
$$+\frac{1}{4(b-a)}\left[(b-x)^2 f'_-(b)-(x-a)^2 f'_+(a)\right]$$
$$=\frac{1}{b-a}\int_a^x (t-a)\left[f'(t)-\frac{f'_+(a)+f'(x)}{2}\right]dt$$
$$+\frac{1}{b-a}\int_x^b (t-b)\left[f'(t)-\frac{f'(x)+f'_-(b)}{2}\right]dt.$$

In particular, we have

$$f\left(\frac{a+b}{2}\right)+\frac{1}{16}(b-a)\left[f'_-(b)-f'_+(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt \tag{106}$$
$$=\frac{1}{b-a}\int_a^{\frac{a+b}{2}}(t-a)\left[f'(t)-\frac{f'_+(a)+f'\left(\frac{a+b}{2}\right)}{2}\right]dt$$
$$+\frac{1}{b-a}\int_{\frac{a+b}{2}}^b (t-b)\left[f'(t)-\frac{f'\left(\frac{a+b}{2}\right)+f'_-(b)}{2}\right]dt.$$

Inequalities for Bounded Derivatives

Now, for $\gamma, \Gamma \in \mathbb{C}$ and $[a, b]$ an interval of real numbers, define the sets of complex-valued functions

$$\begin{aligned}&\bar{U}_{[a,b]}(\gamma,\Gamma)\\&:=\left\{f:[a,b]\to\mathbb{C}\,|\,\operatorname{Re}\left[(\Gamma-f(t))\left(\overline{f(t)}-\overline{\gamma}\right)\right]\geq 0 \text{ for almost every } t\in[a,b]\right\}\end{aligned}$$

and

$$\bar{\Delta}_{[a,b]}(\gamma,\Gamma):=\left\{f:[a,b]\to\mathbb{C}\,|\left|f(t)-\frac{\gamma+\Gamma}{2}\right|\leq\frac{1}{2}|\Gamma-\gamma| \text{ for a.e. } t\in[a,b]\right\}.$$

The following representation result may be stated.

Proposition 3 (Dragomir, 2013 [11]). *For any $\gamma, \Gamma \in \mathbb{C}$, $\gamma \neq \Gamma$, we have that $\bar{U}_{[a,b]}(\gamma,\Gamma)$ and $\bar{\Delta}_{[a,b]}(\gamma,\Gamma)$ are nonempty, convex, and closed sets and*

$$\bar{U}_{[a,b]}(\gamma,\Gamma)=\bar{\Delta}_{[a,b]}(\gamma,\Gamma). \tag{107}$$

Proof. We observe that for any $z \in \mathbb{C}$ we have the equivalence

$$\left|z-\frac{\gamma+\Gamma}{2}\right|\leq\frac{1}{2}|\Gamma-\gamma|$$

if and only if

$$\operatorname{Re}[(\Gamma-z)(\bar{z}-\bar{\gamma})]\geq 0.$$

This follows by the equality

$$\frac{1}{4}|\Gamma-\gamma|^2-\left|z-\frac{\gamma+\Gamma}{2}\right|^2=\operatorname{Re}[(\Gamma-z)(\bar{z}-\bar{\gamma})]$$

that holds for any $z \in \mathbb{C}$.

The equality (107) is thus a simple consequence of this fact.

On making use of the complex numbers field properties we can also state that

Corollary 18. *For any $\gamma, \Gamma \in \mathbb{C}$, $\gamma \neq \Gamma$, we have that*

$$\begin{aligned}\bar{U}_{[a,b]}(\gamma,\Gamma)=\{f:[a,b]\to\mathbb{C}\mid\ &(\operatorname{Re}\Gamma-\operatorname{Re}f(t))(\operatorname{Re}f(t)-\operatorname{Re}\gamma)\\&+(\operatorname{Im}\Gamma-\operatorname{Im}f(t))(\operatorname{Im}f(t)-\operatorname{Im}\gamma)\geq 0\, \textit{for a.e. } t\in[a,b]\}.\end{aligned} \tag{108}$$

Now, if we assume that $\mathrm{Re}(\Gamma) \geq \mathrm{Re}(\gamma)$ and $\mathrm{Im}(\Gamma) \geq \mathrm{Im}(\gamma)$, then we can define the following set of functions as well:

$$\bar{S}_{[a,b]}(\gamma,\Gamma) := \{f : [a,b] \to \mathbb{C} \mid \mathrm{Re}(\Gamma) \geq \mathrm{Re} f(t) \geq \mathrm{Re}(\gamma) \text{ and } \mathrm{Im}(\Gamma) \geq \mathrm{Im} f(t) \geq \mathrm{Im}(\gamma) \text{ for a.e. } t \in [a,b]\}. \tag{109}$$

One can easily observe that $\bar{S}_{[a,b]}(\gamma,\Gamma)$ is closed, convex and

$$\emptyset \neq \bar{S}_{[a,b]}(\gamma,\Gamma) \subseteq \bar{U}_{[a,b]}(\gamma,\Gamma). \tag{110}$$

Theorem 14 (Dragomir, 2013 [11]). *Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous on $[a,b]$ and $x \in (a,b)$. Suppose that $\gamma_i, \Gamma_i \in \mathbb{C}$ with $\gamma_i \neq \Gamma_i$, $i = 1,2$ and $f' \in \bar{U}_{[a,x]}(\gamma_1,\Gamma_1) \cap \bar{U}_{[x,b]}(\gamma_2,\Gamma_2)$, then we have*

$$\begin{aligned}
&\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right. \\
&\left. + \frac{1}{2(b-a)}\left[(b-x)^2\frac{\Gamma_2+\gamma_2}{2} - (x-a)^2\frac{\Gamma_1+\gamma_1}{2}\right]\right| \\
&\leq \frac{1}{4}\left[|\Gamma_1-\gamma_1|\left(\frac{x-a}{b-a}\right)^2 + |\Gamma_2-\gamma_2|\left(\frac{b-x}{b-a}\right)^2\right](b-a) \\
&\leq \frac{1}{4}(b-a) \\
&\times \begin{cases}
\left[\frac{1}{4} + \left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right] \max\{|\Gamma_1-\gamma_1|, |\Gamma_2-\gamma_2|\}. \\[2ex]
\left[\left(\frac{x-a}{b-a}\right)^{2p} + \left(\frac{b-x}{b-a}\right)^{2p}\right]^{1/p} \left[|\Gamma_1-\gamma_1|^q + |\Gamma_2-\gamma_2|^q\right]^{1/q} \\
p > 1, \frac{1}{p} + \frac{1}{q} = 1, \\[2ex]
\left[\frac{1}{2} + \left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right] \left[|\Gamma_1-\gamma_1| + |\Gamma_2-\gamma_2|\right].
\end{cases}
\end{aligned} \tag{111}$$

Proof. Since $f' \in \bar{U}_{[a,x]}(\gamma_1,\Gamma_1) \cap \bar{U}_{[x,b]}(\gamma_2,\Gamma_2)$, then by taking the modulus in (98) for $\lambda_1(x) = \frac{\Gamma_1+\gamma_1}{2}$ and $\lambda_2(x) = \frac{\Gamma_2+\gamma_2}{2}$ we get

$$\begin{aligned}
&\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right. \\
&\left. + \frac{1}{2(b-a)}\left[(b-x)^2\frac{\Gamma_2+\gamma_2}{2} - (x-a)^2\frac{\Gamma_1+\gamma_1}{2}\right]\right|
\end{aligned}$$

$$\leq \frac{1}{b-a}\left|\int_a^x (t-a)\left[f'(t)-\frac{\Gamma_1+\gamma_1}{2}\right]dt\right|$$
$$+\frac{1}{b-a}\left|\int_x^b (t-b)\left[f'(t)-\frac{\Gamma_2+\gamma_2}{2}\right]dt\right|$$
$$\leq \frac{1}{b-a}\int_a^x (t-a)\left|f'(t)-\frac{\Gamma_1+\gamma_1}{2}\right|dt$$
$$+\frac{1}{b-a}\int_x^b (t-b)\left|f'(t)-\frac{\Gamma_2+\gamma_2}{2}\right|dt$$
$$\leq \frac{1}{b-a}\frac{|\Gamma_1-\gamma_1|}{2}\int_a^x (t-a)\,dt+\frac{1}{b-a}\frac{|\Gamma_2-\gamma_2|}{2}\int_x^b (b-t)\,dt$$
$$=\frac{1}{4}\left[|\Gamma_1-\gamma_1|\left(\frac{x-a}{b-a}\right)^2+|\Gamma_2-\gamma_2|\left(\frac{b-x}{b-a}\right)^2\right](b-a)$$

and the first inequality in (111) is proved.

The last part follows by Hölder's inequality

$$mn+pq\leq (m^\alpha+p^\alpha)^{1/\alpha}\left(n^\beta+q^\beta\right)^{1/\beta},$$

where $m,n,p,q\geq 0$ and $\alpha>1$ with $\frac{1}{\alpha}+\frac{1}{\beta}=1$.

Corollary 19. *Let $f:[a,b]\to\mathbb{C}$ be an absolutely continuous on $[a,b]$ and $x\in(a,b)$. Suppose that $\gamma,\Gamma\in\mathbb{C}$ with $\gamma\neq\Gamma$, and $f'\in\bar{U}_{[a,b]}(\gamma,\Gamma)$, then we have*

$$\left|f(x)+\left(\frac{a+b}{2}-x\right)\frac{\Gamma+\gamma}{2}-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \quad (112)$$
$$\leq \frac{1}{2}|\Gamma-\gamma|\left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right](b-a).$$

In particular, we have

$$\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\leq\frac{1}{8}|\Gamma-\gamma|(b-a). \quad (113)$$

Remark 12. If the derivative $f':[a,b]\to\mathbb{R}$ is bounded above and below, that is, there exist the constants $M>m$ such that

$$-\infty<m\leq f'(t)\leq M<\infty \text{ for a.e. } t\in[a,b],$$

then we recapture from (112) the inequality [7]

$$\left| f(x) + \left(\frac{a+b}{2} - x\right)\frac{M+m}{2} - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \leq \frac{1}{2}(M-m)\left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a}\right)^2\right](b-a).$$

Remark 13. Let $f : [a,b] \to \mathbb{C}$ be an absolutely continuous on $[a,b]$. Suppose that $\gamma_i, \Gamma_i \in \mathbb{C}$ with $\gamma_i \neq \Gamma_i$, $i = 1,2$ and $f' \in \bar{U}_{\left[a,\frac{a+b}{2}\right]}(\gamma_1, \Gamma_1) \cap \bar{U}_{\left[\frac{a+b}{2},b\right]}(\gamma_2, \Gamma_2)$, then we have from (111) that

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt + \frac{1}{8}(b-a)\left(\frac{\Gamma_2+\gamma_2}{2} - \frac{\Gamma_1+\gamma_1}{2}\right) \right| \leq \frac{1}{16}\left[|\Gamma_1-\gamma_1| + |\Gamma_2-\gamma_2|\right](b-a). \tag{114}$$

Inequalities for Derivatives of Bounded Variation

Assume that the function $f : I \to \mathbb{C}$ is differentiable on the interior of I, denoted $\mathring{I}$, and $[a,b] \subset \mathring{I}$. Then, as in (105), we have the equality

$$\begin{aligned} & f(x) + \frac{1}{2}\left(\frac{a+b}{2} - x\right)f'(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \\ & + \frac{1}{4(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] \\ & = \frac{1}{b-a}\int_a^x (t-a)\left[f'(t) - \frac{f'(a)+f'(x)}{2}\right]dt \\ & + \frac{1}{b-a}\int_x^b (t-b)\left[f'(t) - \frac{f'(x)+f'(b)}{2}\right]dt, \end{aligned} \tag{115}$$

for any $x \in [a,b]$.

Theorem 15 (Dragomir, 2013 [11]). *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. If the derivative $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a,b]$, then*

$$
\begin{aligned}
&\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt+\frac{1}{2}\left(\frac{a+b}{2}-x\right)f'(x)\right. \\
&\left.+\frac{1}{4(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]\right|
\end{aligned}
\tag{116}
$$

$$
\begin{aligned}
&\leq \frac{1}{4}\left[\left(\frac{x-a}{b-a}\right)^2 \bigvee_a^x (f')+\left(\frac{b-x}{b-a}\right)^2 \bigvee_x^b (f')\right](b-a) \\
&\leq \frac{1}{4}(b-a) \\
&\times\begin{cases}
\left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right]\left[\frac{1}{2}\bigvee_a^b (f')+\frac{1}{2}\left|\bigvee_a^x (f')-\bigvee_x^b (f')\right|\right], \\
\left[\left(\frac{x-a}{b-a}\right)^{2p}+\left(\frac{b-x}{b-a}\right)^{2p}\right]^{1/p}\left[\left[\bigvee_a^x (f')\right]^q+\left[\bigvee_x^b (f')\right]^q\right]^{1/q} \\
p>1,\ \frac{1}{p}+\frac{1}{q}=1, \\
\left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]\bigvee_a^b (f'),
\end{cases}
\end{aligned}
$$

for any $x \in [a,b]$.

Proof. Taking the modulus in (115) we have

$$
\begin{aligned}
&\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt+\frac{1}{2}\left(\frac{a+b}{2}-x\right)f'(x)\right. \\
&\left.+\frac{1}{4(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]\right| \\
&\leq \frac{1}{b-a}\left|\int_a^x (t-a)\left[f'(t)-\frac{f'(a)+f'(x)}{2}\right]dt\right| \\
&+\frac{1}{b-a}\left|\int_x^b (t-b)\left[f'(t)-\frac{f'(x)+f'(b)}{2}\right]dt\right| \\
&\leq \frac{1}{b-a}\int_a^x (t-a)\left|f'(t)-\frac{f'(a)+f'(x)}{2}\right|dt \\
&+\frac{1}{b-a}\int_x^b (b-t)\left|f'(t)-\frac{f'(x)+f'(b)}{2}\right|dt.
\end{aligned}
\tag{117}
$$

Since $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a, x]$ and $[x, b]$, then

$$\begin{aligned}\left|f'(t) - \frac{f'(a) + f'(x)}{2}\right| &= \frac{|f'(t) - f'(a) + f'(t) - f'(x)|}{2} \\ &\leq \frac{1}{2}\left[\left|f'(t) - f'(a)\right| + \left|f'(x) - f'(t)\right|\right] \\ &\leq \frac{1}{2}\bigvee_a^x (f')\end{aligned}$$

for any $t \in [a, x]$ and, similarly,

$$\left|f'(t) - \frac{f'(x) + f'(b)}{2}\right| \leq \frac{1}{2}\bigvee_x^b (f')$$

for any $t \in [x, b]$.

Then

$$\begin{aligned}\int_a^x (t-a)\left|f'(t) - \frac{f'(a) + f'(x)}{2}\right| dt &\leq \frac{1}{2}\bigvee_a^x (f') \int_a^x (t-a)\, dt \\ &= \frac{1}{4}(x-a)^2 \bigvee_a^x (f')\end{aligned}$$

and

$$\begin{aligned}\int_x^b (b-t)\left|f'(t) - \frac{f'(x) + f'(b)}{2}\right| dt &\leq \frac{1}{2}\bigvee_x^b (f') \int_x^b (b-t)\, dt \\ &= \frac{1}{4}(b-x)^2 \bigvee_x^b (f')\end{aligned}$$

and by (117) we get the desired inequality (116).

The last part follows by Hölder's inequality

$$mn + pq \leq (m^\alpha + p^\alpha)^{1/\alpha}\left(n^\beta + q^\beta\right)^{1/\beta},$$

where $m, n, p, q \geq 0$ and $\alpha > 1$ with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Corollary 20. *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a, b] \subset \mathring{I}$. If the derivative $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a, b]$, then*

$$\left|f\left(\frac{a+b}{2}\right)-\frac{1}{b-a}\int_a^b f(t)\,dt+\frac{1}{16}(b-a)\left[f'(b)-f'(a)\right]\right| \tag{118}$$
$$\leq \frac{1}{16}(b-a)\bigvee_a^b(f').$$

Remark 14. If $p \in (a,b)$ is a median point in bounded variation for the derivative, i.e. $\bigvee_a^p(f') = \bigvee_p^b(f')$, then under the assumptions of Theorem 15, we have

$$\left|f(p)-\frac{1}{b-a}\int_a^b f(t)\,dt+\frac{1}{2}\left(\frac{a+b}{2}-p\right)f'(p)\right. \tag{119}$$
$$\left.+\frac{1}{4(b-a)}\left[(b-p)^2 f'(b)-(p-a)^2 f'(a)\right]\right|$$
$$\leq \frac{1}{8}(b-a)\left[\frac{1}{4}+\left(\frac{p-\frac{a+b}{2}}{b-a}\right)^2\right]\bigvee_a^b(f').$$

Inequalities for Lipschitzian Derivatives

We say that $v:[a,b]\to\mathbb{C}$ is *Lipschitzian* with the constant $L>0$, if it satisfies the condition

$$|v(t)-v(s)|\leq L|t-s| \text{ for any } t,s\in[a,b].$$

Theorem 16 (Dragomir, 2013 [11]). *Let $f: I\to\mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b]\subset\mathring{I}$. Let $x\in(a,b)$. If the derivative $f':\mathring{I}\to\mathbb{C}$ is Lipschitzian with the constant $K_1(x)$ on $[a,x]$ and constant $K_2(x)$ on $[x,b]$, then*

$$\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt+\frac{1}{2}\left(\frac{a+b}{2}-x\right)f'(x)\right. \tag{120}$$
$$\left.+\frac{1}{4(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]\right|$$

$$\leq \frac{1}{8}\left[\left(\frac{x-a}{b-a}\right)^3 K_1(x)+\left(\frac{b-x}{b-a}\right)^3 K_2(x)\right](b-a)^2$$
$$\leq \frac{1}{8}(b-a)^2$$

$$\times\begin{cases}\left[\left(\frac{x-a}{b-a}\right)^3+\left(\frac{b-x}{b-a}\right)^3\right]\max\left\{K_1(x),K_2(x)\right\},\\[2ex] \left[\left(\frac{x-a}{b-a}\right)^{2p}+\left(\frac{b-x}{b-a}\right)^{2p}\right]^{1/p}\left[K_1^q(x)+K_2^q(x)\right]^{1/q}\\ p>1,\frac{1}{p}+\frac{1}{q}=1,\\[2ex] \left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]^3\left[K_1(x)+K_2(x)\right].\end{cases}$$

Proof. Since $f' : \mathring{I} \to \mathbb{C}$ is Lipschitzian with the constant $K_1(x)$ on $[a,x]$ and constant $K_2(x)$ on $[x,b]$, then

$$\begin{aligned}\left|f'(t)-\frac{f'(a)+f'(x)}{2}\right| &= \frac{|f'(t)-f'(a)+f'(t)-f'(x)|}{2}\\ &\le \frac{1}{2}\left[\left|f'(t)-f'(a)\right|+\left|f'(x)-f'(t)\right|\right]\\ &\le \frac{1}{2}K_1(x)\left[|t-a|+|x-t|\right]\\ &= \frac{1}{2}K_1(x)(x-a)\end{aligned}$$

for any $t\in[a,x]$ and, similarly,

$$\begin{aligned}\left|f'(t)-\frac{f'(x)+f'(b)}{2}\right| &\le \frac{1}{2}K_2(x)\left[|t-x|+|b-t|\right]\\ &= \frac{1}{2}K_2(x)(b-x)\end{aligned}$$

for any $t\in[x,b]$.

Then

$$\begin{aligned}\int_a^x (t-a)\left|f'(t)-\frac{f'(a)+f'(x)}{2}\right|dt &\le \frac{1}{2}K_1(x)(x-a)\int_a^x (t-a)\,dt\\ &= \frac{1}{8}(x-a)^3 K_1(x)\end{aligned}$$

and

$$\begin{aligned}\int_x^b (b-t)\left|f'(t)-\frac{f'(x)+f'(b)}{2}\right|dt &\le \frac{1}{2}K_2(x)(b-x)\int_x^b (b-t)\,dt\\ &= \frac{1}{8}(b-x)^3 K_2(x).\end{aligned}$$

Making use of the inequality (117) we deduce the first bound in (120).

The second part is obvious.

Corollary 21. *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. If the derivative $f' : \mathring{I} \to \mathbb{C}$ is Lipschitzian with the constant K on $[a,b]$, then*

$$\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt + \frac{1}{2}\left(\frac{a+b}{2} - x\right) f'(x) \right. \tag{121}$$
$$\left. + \frac{1}{4(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right]\right|$$
$$\leq \frac{1}{8}\left[\left(\frac{x-a}{b-a}\right)^3 + \left(\frac{b-x}{b-a}\right)^3\right] K (b-a)^2$$

for any $x \in [a,b]$.

In particular, we have

$$\left| f\left(\frac{a+b}{2}\right) + \frac{1}{16}(b-a)\left[f'(b) - f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{122}$$
$$\leq \frac{1}{32} K (b-a)^2 .$$

Other Perturbed Ostrowski Type Inequalities for Absolutely Continuous Function

Inequalities for Derivatives of Bounded Variation

Assume that the function $f : I \to \mathbb{C}$ is differentiable on the interior of I, denoted $\mathring{I}$, and $[a,b] \subset \mathring{I}$. Then, we have the equality [12]

$$f(x) + \left(\frac{a+b}{2} - x\right) f'(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{123}$$
$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(t) - f'(x)\right] dt + \frac{1}{b-a}\int_x^b (t-b)\left[f'(t) - f'(x)\right] dt$$

for any $x \in [a,b]$.

We have the following result:

Theorem 17 (Dragomir, 2013 [12]). *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. If the derivative $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a,b]$, then*

$$\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{124}$$

$$\leq \frac{1}{b-a}\left[\int_a^x (t-a)\bigvee_t^x (f')\,dt+\int_x^b (b-t)\bigvee_x^t (f')\,dt\right]$$

$$\leq \frac{1}{2}(b-a)\left[\left(\frac{x-a}{b-a}\right)^2\bigvee_a^x (f')\,dt+\left(\frac{b-x}{b-a}\right)^2\bigvee_x^b (f')\right]$$

$$\leq \frac{1}{2}(b-a)$$

$$\times\begin{cases}\left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right]\left[\frac{1}{2}\bigvee_a^b (f')+\frac{1}{2}\left|\bigvee_a^x (f')-\bigvee_x^b (f')\right|\right],\\[2ex] \left[\left(\frac{x-a}{b-a}\right)^{2p}+\left(\frac{b-x}{b-a}\right)^{2p}\right]^{1/p}\left[\left[\bigvee_a^x (f')\right]^q+\left[\bigvee_x^b (f')\right]^q\right]^{1/q}\\ p>1,\ \frac{1}{p}+\frac{1}{q}=1,\\[2ex] \left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]\bigvee_a^b (f'),\end{cases}$$

for any $x\in[a,b]$.

Proof. Taking the modulus in (123) we have

$$\begin{aligned}&\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \\ &\leq \frac{1}{b-a}\left|\int_a^x (t-a)\left[f'(t)-f'(x)\right]dt\right| \\ &+\frac{1}{b-a}\left|\int_x^b (t-b)\left[f'(t)-f'(x)\right]dt\right| \\ &\leq \frac{1}{b-a}\int_a^x (t-a)\left|f'(t)-f'(x)\right|dt \\ &+\frac{1}{b-a}\int_x^b (b-t)\left|f'(t)-f'(x)\right|dt.\end{aligned} \tag{125}$$

Since the derivative $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a,x]$ and $[x,b]$, then

$$\left|f'(t) - f'(x)\right| \leq \bigvee_t^x (f') \text{ for } t \in [a,x]$$

and

$$\left|f'(t) - f'(x)\right| \leq \bigvee_x^t (f') \text{ for } t \in [x,b].$$

Therefore

$$\begin{aligned}\int_a^x (t-a)\left|f'(t) - f'(x)\right| dt &\leq \int_a^x (t-a) \bigvee_t^x (f')\, dt \\ &\leq \frac{1}{2}(x-a)^2 \bigvee_a^x (f')\, dt\end{aligned}$$

and

$$\begin{aligned}\int_x^b (b-t)\left|f'(t) - f'(x)\right| dt &\leq \int_x^b (b-t) \bigvee_x^t (f')\, dt \\ &\leq \frac{1}{2}(b-x)^2 \bigvee_x^b (f'),\end{aligned}$$

which, by (125) produce the first two inequalities in (124).

The last part follows by Hölder's inequality

$$mn + pq \leq (m^\alpha + p^\alpha)^{1/\alpha} \left(n^\beta + q^\beta\right)^{1/\beta},$$

where $m,n,p,q \geq 0$ and $\alpha > 1$ with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Corollary 22. *With the assumptions of Theorem 17, we have*

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\, dt\right| \qquad (126) \\ &\leq \frac{1}{b-a}\left[\int_a^{\frac{a+b}{2}} (t-a) \bigvee_t^{\frac{a+b}{2}} (f')\, dt + \int_{\frac{a+b}{2}}^b (b-t) \bigvee_{\frac{a+b}{2}}^t (f')\, dt\right] \\ &\leq \frac{1}{8}(b-a) \bigvee_a^b (f')\, dt.\end{aligned}$$

Remark 15. If $p \in (a,b)$ is a median point in bounded variation for the derivative, i.e. $\bigvee_a^p (f') = \bigvee_p^b (f')$, then under the assumptions of Theorem 17 we have

$$\left| f(p) + \left(\frac{a+b}{2} - p \right) f'(p) - \frac{1}{b-a} \int_a^b f(t)\,dt \right| \tag{127}$$
$$\leq \frac{1}{b-a} \left[\int_a^p (t-a) \bigvee_t^p (f')\,dt + \int_p^b (b-t) \bigvee_p^t (f')\,dt \right]$$
$$\leq \frac{1}{4}(b-a) \left[\frac{1}{4} + \left(\frac{p - \frac{a+b}{2}}{b-a} \right)^2 \right] \bigvee_a^b (f').$$

Inequalities for Lipschitzian Derivatives

We start with the following result.

Theorem 18 (Dragomir, 2013 [12]). *Let $f: I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. Let $x \in (a,b)$. If $\alpha_i > -1$ and $L_{\alpha_i} > 0$ with $i = 1,2$ are such that*

$$\left| f'(t) - f'(x) \right| \leq L_{\alpha_1} (x-t)^{\alpha_1} \text{ for any } t \in [a,x) \tag{128}$$

and

$$\left| f'(t) - f'(x) \right| \leq L_{\alpha_2} (t-x)^{\alpha_2} \text{ for any } t \in (x,b], \tag{129}$$

then we have

$$\left| f(x) + \left(\frac{a+b}{2} - x \right) f'(x) - \frac{1}{b-a} \int_a^b f(t)\,dt \right| \tag{130}$$
$$\leq \frac{1}{b-a} \left[\frac{L_{\alpha_1}}{(\alpha_1+1)(\alpha_1+2)} (x-a)^{\alpha_1+2} + \frac{L_{\alpha_2}}{(\alpha_2+1)(\alpha_2+2)} (b-x)^{\alpha_2+2} \right].$$

Proof. Taking the modulus in (123) we have

$$\left| f(x) + \left(\frac{a+b}{2} - x \right) f'(x) - \frac{1}{b-a} \int_a^b f(t)\,dt \right| \tag{131}$$
$$\leq \frac{1}{b-a} \int_a^x (t-a) \left| f'(t) - f'(x) \right| dt$$
$$+ \frac{1}{b-a} \int_x^b (b-t) \left| f'(t) - f'(x) \right| dt.$$

Using the properties (128) and (129) we have

$$\begin{aligned}\int_a^x (t-a)\left|f'(t)-f'(x)\right| dt &\le L_{\alpha_1}\int_a^x (t-a)(x-t)^{\alpha_1}\,dt\\ &= L_{\alpha_1}(x-a)^{\alpha_1+2}\int_0^1 u(1-u)^{\alpha_1}\,du\\ &= L_{\alpha_1}(x-a)^{\alpha_1+2}\int_0^1 u^{\alpha_1}(1-u)\,du\\ &= \frac{1}{(\alpha_1+1)(\alpha_1+2)}L_{\alpha_1}(x-a)^{\alpha_1+2}\end{aligned}$$

and

$$\begin{aligned}\int_x^b (b-t)\left|f'(t)-f'(x)\right| dt &\le L_{\alpha_2}\int_x^b (b-t)(t-x)^{\alpha_2}\,dt\\ &= \frac{1}{(\alpha_2+1)(\alpha_2+2)}L_{\alpha_2}(b-x)^{\alpha_2+2}.\end{aligned}$$

Utilizing (131) we get the desired result (130).

Corollary 23. *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. If the derivative is f' of r-H-Hölder type on $[a,b]$, i.e. we have the condition*

$$\left|f'(t)-f'(s)\right| \le H\,|t-s|^r$$

for any $t,s \in [a,b]$, where $r \in (0,1]$ and $H > 0$ are given, then

$$\begin{aligned}&\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \\ &\le \frac{H}{(r+1)(r+2)}\left[\left(\frac{x-a}{b-a}\right)^{r+2}+\left(\frac{b-x}{b-a}\right)^{r+2}\right](b-a)^{r+1},\end{aligned} \tag{132}$$

for any $x \in [a,b]$.

In particular, if f' is Lipschitzian with the constant $L > 0$, then

$$\begin{aligned}&\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \\ &\le \frac{1}{6}L\left[\left(\frac{x-a}{b-a}\right)^{3}+\left(\frac{b-x}{b-a}\right)^{3}\right](b-a)^{2},\end{aligned} \tag{133}$$

for any $x \in [a,b]$.

Inequalities for Differentiable Convex Functions

The case of convex functions is as follows:

Theorem 19 (Dragomir, 2013 [12]). *Let $f : I \to \mathbb{C}$ be a differentiable convex function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. Then for any $x \in [a,b]$ we have*

$$0 \leq \frac{1}{b-a}\int_a^b f(t)\,dt - f(x) - \left(\frac{a+b}{2}-x\right)f'(x) \leq \begin{cases} I_1(x) \\ \\ I_2(x) \\ \\ I_3(x) \end{cases} \qquad (134)$$

where

$$I_1(x) := \frac{(b-x)f(b)+(x-a)f(a)}{b-a} - f(x) - 2f'(x)\left(\frac{a+b}{2}-x\right),$$

$$I_2(x) := \frac{1}{2}\frac{f'(b)(b-x)^2 - f'(a)(x-a)^2}{b-a} - f'(x)\left(\frac{a+b}{2}-x\right)$$

and

$$I_3(x) := \frac{1}{2}\left[\frac{f(b)(b-x)+f(a)(x-a)}{b-a} - f(x)\right] - f'(x)\left(\frac{a+b}{2}-x\right).$$

Proof. We have the equality

$$\frac{1}{b-a}\int_a^b f(t)\,dt - f(x) - \left(\frac{a+b}{2}-x\right)f'(x) \qquad (135)$$
$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt + \frac{1}{b-a}\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt$$

for any $x \in [a,b]$.

Since f is a differentiable convex function on $\mathring{I}$, then f' is monotonic nondecreasing on $\mathring{I}$ and then

$$\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt \geq 0$$

and

$$\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt \geq 0,$$

which proves the first inequality in (134).

We have

$$\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt \le (x-a)\int_a^x \left[f'(x)-f'(t)\right]dt$$
$$= (x-a)\left[f'(x)(x-a)-f(x)+f(a)\right]$$

and

$$\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt \le (b-x)\int_x^b \left[f'(t)-f'(x)\right]dt$$
$$= (b-x)\left[f(b)-f(x)-f'(x)(b-x)\right].$$

Adding these inequalities we get

$$\begin{aligned}
&\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt + \int_x^b (b-t)\left[f'(t)-f'(x)\right]dt \\
&\le (x-a)\left[f'(x)(x-a)-f(x)+f(a)\right] \\
&+ (b-x)\left[f(b)-f(x)-f'(x)(b-x)\right] \\
&= (b-x)f(b)+(x-a)f(a)-(b-a)f(x) \\
&+ f'(x)\left[2x-(a+b)\right](b-a)
\end{aligned}$$

and by (135) we get the second inequality for $I_1(x)$.

We also have

$$\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt \le \int_a^x (t-a)\left[f'(x)-f'(a)\right]dt$$
$$= \frac{1}{2}\left[f'(x)-f'(a)\right](x-a)^2$$

and

$$\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt \le \int_x^b (b-t)\left[f'(b)-f'(x)\right]dt$$
$$= \frac{1}{2}\left[f'(b)-f'(x)\right](b-x)^2.$$

Adding these inequalities we get

$$\begin{aligned}
&\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt+\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt\\
&\le \frac{1}{2}\left[f'(x)-f'(a)\right](x-a)^2+\frac{1}{2}\left[f'(b)-f'(x)\right](b-x)^2\\
&=\frac{1}{2}\left[f'(b)(b-x)^2-f'(a)(x-a)^2+f'(x)(b-a)[2x-(a+b)]\right]
\end{aligned}$$

and by (135) we get the second inequality for $I_2(x)$.

Further, we use the Čebyšev inequality for asynchronous functions (functions of opposite monotonicity), namely

$$\frac{1}{d-c}\int_c^d g(t)\,h(t)\,dt\le\frac{1}{d-c}\int_c^d g(t)\,dt\cdot\frac{1}{d-c}\int_c^d h(t)\,dt.$$

Therefore

$$\begin{aligned}
&\frac{1}{x-a}\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt\\
&\le\frac{1}{x-a}\int_a^x (t-a)\,dt\cdot\frac{1}{x-a}\int_a^x\left[f'(x)-f'(t)\right]dt\\
&=\frac{(x-a)^2}{2(x-a)}\cdot\frac{f'(x)(x-a)-f(x)+f(a)}{x-a}\\
&=\frac{1}{2}\left[f'(x)(x-a)-f(x)+f(a)\right]
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{b-x}\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt\\
&\le\frac{1}{b-x}\int_x^b (b-t)\,dt\cdot\frac{1}{b-x}\int_x^b\left[f'(t)-f'(x)\right]dt\\
&=\frac{(b-x)^2}{2(b-x)}\cdot\frac{f(b)-f(x)-f'(x)(b-x)}{b-x}\\
&=\frac{1}{2}\left[f(b)-f(x)-f'(x)(b-x)\right].
\end{aligned}$$

Adding these inequalities, we have

$$\begin{aligned}
&\frac{1}{b-a}\int_a^x (t-a)\left[f'(x)-f'(t)\right]dt+\frac{1}{b-a}\int_x^b (b-t)\left[f'(t)-f'(x)\right]dt\\
&\le \frac{1}{2}\frac{[f'(x)(x-a)-f(x)+f(a)](x-a)}{b-a}\\
&+\frac{1}{2}\frac{[f(b)-f(x)-f'(x)(b-x)](b-x)}{b-a}\\
&=\frac{1}{2(b-a)}\left[\left[f'(x)(x-a)-f(x)+f(a)\right](x-a)\right]\\
&+\frac{1}{2(b-a)}\left[\left[f(b)-f(x)-f'(x)(b-x)\right](b-x)\right]\\
&=\frac{1}{2}\left[\frac{f(b)(b-x)+f(a)(x-a)}{b-a}-f(x)\right]+f'(x)\left(x-\frac{a+b}{2}\right)
\end{aligned}$$

which proves the inequality for $I_3(x)$.

Remark 16. From the first inequality in (134) we have

$$\frac{1}{b-a}\int_a^b f(t)\,dt \le \frac{(b-x)f(b)+(x-a)f(a)}{b-a}-f'(x)\left(\frac{a+b}{2}-x\right) \tag{136}$$

for any $x\in[a,b]$.

From the second inequality in (134) we have

$$\frac{1}{b-a}\int_a^b f(t)\,dt-f(x)\le \frac{1}{2}\cdot\frac{f'(b)(b-x)^2-f'(a)(x-a)^2}{b-a} \tag{137}$$

for any $x\in[a,b]$.

From the third inequality in (134) we have

$$\frac{1}{b-a}\int_a^b f(t)\,dt \le \frac{1}{2}\left[\frac{f(b)(b-x)+f(a)(x-a)}{b-a}+f(x)\right] \tag{138}$$

for any $x\in[a,b]$.

Inequalities for Absolutely Continuous Derivatives

The case of absolutely continuous derivatives is as follows:

Theorem 20 (Dragomir, 2013 [12]). *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. If the derivative f' is absolutely continuous on $[a,b]$, then for any $x \in [a,b]$*

$$\left| f(x) + \left(\frac{a+b}{2} - x\right) f'(x) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{139}$$

$$\leq \frac{1}{b-a} \times \begin{cases} \frac{1}{6}(x-a)^3 \|f''\|_{[a,x],\infty}, \\ \frac{q}{(q+1)(q+2)}(x-a)^{1/q+2}\|f''\|_{[a,x],p}, \\ \frac{1}{2}(x-a)^2\|f''\|_{[a,x],1}, \end{cases}$$

$$+ \frac{1}{b-a} \times \begin{cases} \frac{1}{6}(b-x)^3 \|f''\|_{[x,b],\infty}, \\ \frac{q}{(q+1)(q+2)}(b-x)^{1/q+2}\|f''\|_{[x,b],p}, \\ \frac{1}{2}(b-x)^2\|f''\|_{[x,b],1}, \end{cases}$$

where $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Taking the modulus in (123) we have

$$\begin{aligned} &\left| f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt + \left(\frac{a+b}{2} - x\right) f'(x) \right| \qquad (140) \\ &\leq \frac{1}{b-a}\int_a^x (t-a)\left|f'(t) - f'(x)\right| dt + \frac{1}{b-a}\int_x^b (b-t)\left|f'(t) - f'(x)\right| dt \\ &= \frac{1}{b-a}\int_a^x (t-a)\left|\int_x^t f''(s)\,ds\right| + \frac{1}{b-a}\int_x^b (b-t)\left|\int_x^t f''(s)\,ds\right| \\ &\leq \frac{1}{b-a}\int_a^x (t-a)\int_t^x \left|f''(s)\right| ds + \frac{1}{b-a}\int_x^b (b-t)\int_x^t \left|f''(s)\right| ds. \end{aligned}$$

Using Hölder's integral inequality we have for $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\int_a^x (t-a) \int_t^x \left| f''(s) \right| ds \le \begin{cases} \int_a^x (t-a)(x-t) \|f''\|_{[t,x],\infty} dt \\ \int_a^x (t-a)(x-t)^{1/q} \|f''\|_{[t,x],p} dt \\ \int_a^x (t-a) \|f''\|_{[t,x],1} dt \end{cases}$$

$$\le \begin{cases} \|f''\|_{[a,x],\infty} \int_a^x (t-a)(x-t)\, dt \\ \|f''\|_{[a,x],p} \int_a^x (t-a)(x-t)^{1/q} dt \\ \|f''\|_{[a,x],1} \int_a^x (t-a)\, dt \end{cases}$$

$$= \begin{cases} \frac{1}{6}(x-a)^3 \|f''\|_{[a,x],\infty} \\ \frac{q}{(q+1)(q+2)} (x-a)^{1/q+2} \|f''\|_{[a,x],p} \\ \frac{1}{2}(x-a)^2 \|f''\|_{[a,x],1} \end{cases}$$

and, similarly

$$\int_x^b (b-t) \int_x^t \left| f''(s) \right| ds \le \begin{cases} \frac{1}{6}(b-x)^3 \|f''\|_{[x,b],\infty} \\ \frac{q}{(q+1)(q+2)} (b-x)^{1/q+2} \|f''\|_{[x,b],p} \\ \frac{1}{2}(b-x)^2 \|f''\|_{[x,b],1} . \end{cases}$$

Utilizing the inequality (140) we get the desired result (139).

Remark 17. Since

$$\begin{aligned} &\frac{1}{6}(x-a)^3 \left\|f''\right\|_{[a,x],\infty} + \frac{1}{6}(b-x)^3 \left\|f''\right\|_{[x,b],\infty} \\ &\le \frac{1}{6}\left[(x-a)^3 + (b-x)^3\right] \max\left\{\left\|f''\right\|_{[a,x],\infty}, \left\|f''\right\|_{[x,b],\infty}\right\} \\ &= \frac{1}{6}(b-a)\left[(x-a)^2 - (x-a)(b-x) + (b-x)^2\right] \left\|f''\right\|_{[a,b],\infty}, \end{aligned}$$

then by (139) we get

$$\left| f(x) + \left(\frac{a+b}{2} - x\right) f'(x) - \frac{1}{b-a} \int_a^b f(t)\, dt \right| \tag{141}$$

$$\leq \frac{1}{6}\left[\left(\frac{x-a}{b-a}\right)^2-\left(\frac{x-a}{b-a}\right)\left(\frac{b-x}{b-a}\right)+\left(\frac{b-x}{b-a}\right)^2\right]$$
$$\times (b-a)^2 \left\|f''\right\|_{[a,b],\infty},$$

for any $x \in [a,b]$.

Since

$$(x-a)^{1/q+2}\left\|f''\right\|_{[a,x],p}+(b-x)^{1/q+2}\left\|f''\right\|_{[x,b],p}$$
$$\leq \left[(x-a)^{2q+1}+(b-x)^{2q+1}\right]^{1/q}\left[\left\|f''\right\|_{[a,x],p}^{p}+\left\|f''\right\|_{[x,b],p}\right]^{1/p}$$
$$=\left[(x-a)^{2q+1}+(b-x)^{2q+1}\right]^{1/q}\left\|f''\right\|_{[a,b],p},$$

then by (139) we get

$$\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{142}$$
$$\leq \frac{q}{(q+1)(q+2)}\left[\left(\frac{x-a}{b-a}\right)^{2q+1}+\left(\frac{b-x}{b-a}\right)^{2q+1}\right]^{1/q}$$
$$\times (b-a)^{1+1/q}\left\|f''\right\|_{[a,b],p},$$

for any $x \in [a,b]$.

Since

$$(x-a)^2\left\|f''\right\|_{[a,x],1}+(b-x)^2\left\|f''\right\|_{[x,b],1}$$
$$\leq \max\left\{(x-a)^2,(b-x)^2\right\}\left[\left\|f''\right\|_{[a,x],1}+\left\|f''\right\|_{[x,b],1}\right]$$
$$=\left[\frac{1}{2}(b-a)+\left|x-\frac{a+b}{2}\right|\right]^2\left\|f''\right\|_{[a,b],1},$$

then by (139) we get

$$\left|f(x)+\left(\frac{a+b}{2}-x\right)f'(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right|$$
$$\leq \frac{1}{2}\left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]^2(b-a)\left\|f''\right\|_{[a,b],1}$$

for any $x \in [a,b]$.

More Perturbed Ostrowski Type Inequalities for Absolutely Continuous Functions

Inequalities for Derivatives of Bounded Variation

Assume that the function $f : I \to \mathbb{C}$ is differentiable on the interior of I, denoted $\mathring{I}$, and $[a, b] \subset \mathring{I}$. Then, we have the equality

$$f(x) + \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{143}$$
$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(t) - f'(a)\right]dt + \frac{1}{b-a}\int_x^b (b-t)\left[f'(b) - f'(t)\right]dt,$$

for any $x \in [a, b]$.

In particular, for $x = \frac{a+b}{2}$, we have

$$f\left(\frac{a+b}{2}\right) + \frac{1}{8}(b-a)\left[f'(b) - f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{144}$$
$$= \frac{1}{b-a}\int_a^{\frac{a+b}{2}} (t-a)\left[f'(t) - f'(a)\right]dt$$
$$+ \frac{1}{b-a}\int_{\frac{a+b}{2}}^b (b-t)\left[f'(b) - f'(t)\right]dt.$$

Theorem 21 (Dragomir, 2013 [13]). *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a, b] \subset \mathring{I}$. If the derivative $f' : \mathring{I} \to \mathbb{C}$ is of bounded variation on $[a, b]$, then for any $x \in [a, b]$*

$$\left|f(x) + \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{145}$$
$$\leq \frac{1}{b-a}\left[\int_a^x (t-a)\bigvee_a^t (f')\,dt + \int_x^b (b-t)\bigvee_t^b (f')\,dt\right]$$

$$\leq \frac{1}{b-a}\begin{cases} \frac{1}{2}(x-a)^2 \bigvee_a^x (f'), \\ \frac{1}{(q+1)^{1/q}}(x-a)^{1+1/q}\left(\int_a^x \left(\bigvee_a^t (f')\right)^p dt\right)^{1/p}, \\ (x-a)\int_a^x \left(\bigvee_a^t (f')\right) dt \end{cases}$$

$$+\frac{1}{b-a}\begin{cases} \frac{1}{2}(b-x)^2 \bigvee_x^b (f'), \\ \frac{1}{(q+1)^{1/q}}(b-x)^{1+1/q}\left(\int_x^b \left(\bigvee_t^b (f')\right)^p dt\right)^{1/p}, \\ (b-x)\int_x^b \left(\bigvee_t^b (f')\right) dt. \end{cases}$$

where $p>1$, $\frac{1}{p}+\frac{1}{q}=1$.

Proof. Taking the modulus in (143) we have

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{146}$$
$$\leq \frac{1}{b-a}\int_a^x (t-a)\left|f'(t)-f'(a)\right| dt+\frac{1}{b-a}\int_x^b (b-t)\left|f'(b)-f'(t)\right| dt,$$

for any $x\in[a,b]$.

Since the derivative $f':\mathring{I}\to\mathbb{C}$ is of bounded variation on $[a,b]$, then

$$\left|f'(t)-f'(a)\right| \leq \bigvee_a^t (f') \text{ for any } t\in[a,x]$$

and

$$\left|f'(b)-f'(t)\right| \leq \bigvee_t^b (f') \text{ for any } t\in[x,b].$$

Therefore

$$\int_a^x (t-a)\left|f'(t)-f'(a)\right| dt \le \int_a^x (t-a)\bigvee_a^t (f')\, dt$$

and

$$\int_x^b (b-t)\left|f'(b)-f'(t)\right| dt \le \int_x^b (b-t)\bigvee_t^b (f')\, dt$$

for any $x \in [a,b]$.

Adding these two inequalities and dividing by $b-a$ we get the first inequality in (145).

Using Hölder's integral inequality we have

$$\int_a^x (t-a)\bigvee_a^t (f')\, dt \le \begin{cases} \bigvee_a^x (f') \int_a^x (t-a)\, dt, \\ \left(\int_a^x (t-a)^q\, dt\right)^{1/q} \left(\int_a^x \left(\bigvee_a^t (f')\right)^p dt\right)^{1/p}, \\ (x-a)\int_a^x \left(\bigvee_a^t (f')\right) dt, \end{cases}$$

$$= \begin{cases} \frac{1}{2}(x-a)^2 \bigvee_a^x (f'), \\ \frac{1}{(q+1)^{1/q}}(x-a)^{1+1/q} \left(\int_a^x \left(\bigvee_a^t (f')\right)^p dt\right)^{1/p}, \\ (x-a)\int_a^x \left(\bigvee_a^t (f')\right) dt \end{cases}$$

and

$$\int_x^b (b-t) \bigvee_t^b (f')\, dt \le \begin{cases} \frac{1}{2}(b-x)^2 \bigvee_x^b (f'), \\ \\ \frac{1}{(q+1)^{1/q}} (b-x)^{1+1/q} \left(\int_x^b \left(\bigvee_x^b (f') \right)^p dt \right)^{1/p}, \\ \\ (b-x) \int_x^b \left(\bigvee_x^b (f') \right) dt. \end{cases}$$

Remark 18. From the first branch in (145) we have the sequence of inequalities

$$\left| f(x) + \frac{1}{2(b-a)} \left[(b-x)^2 f'(b) - (x-a)^2 f'(a) \right] - \frac{1}{b-a} \int_a^b f(t)\, dt \right| \quad (147)$$

$$\le \frac{1}{b-a} \left[\int_a^x (t-a) \bigvee_a^t (f')\, dt + \int_x^b (b-t) \bigvee_t^b (f')\, dt \right]$$

$$\le \frac{1}{2}(b-a) \left[\left(\frac{x-a}{b-a} \right)^2 \bigvee_a^x (f') + \left(\frac{b-x}{b-a} \right)^2 \bigvee_x^b (f') \right]$$

$$\le \frac{1}{2}(b-a)$$

$$\times \begin{cases} \left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a} \right)^2 \right] \left[\frac{1}{2} \bigvee_a^b (f') + \frac{1}{2} \left| \bigvee_a^x (f') - \bigvee_x^b (f') \right| \right], \\ \\ \left[\left(\frac{x-a}{b-a} \right)^{2p} + \left(\frac{b-x}{b-a} \right)^{2p} \right]^{1/p} \left[\left[\bigvee_a^x (f') \right]^q + \left[\bigvee_x^b (f') \right]^q \right]^{1/q} \\ p > 1, \frac{1}{p} + \frac{1}{q} = 1, \\ \\ \left[\frac{1}{2} + \left| \frac{x - \frac{a+b}{2}}{b-a} \right| \right] \bigvee_a^b (f'), \end{cases}$$

for any $x \in [a, b]$.

From the second branch in (145) we have

$$\left| f(x) + \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{148}$$
$$\leq \frac{1}{b-a}\left[\int_a^x (t-a)\bigvee_a^t (f')\,dt + \int_x^b (b-t)\bigvee_t^b (f')\,dt\right]$$
$$\leq \frac{1}{(q+1)^{1/q}}\left\{\left(\frac{x-a}{b-a}\right)^{1+1/q}\left(\int_a^x \left(\bigvee_a^t (f')\right)^p dt\right)^{1/p}\right.$$
$$\left. + \left(\frac{b-x}{b-a}\right)^{1+1/q}\left(\int_x^b \left(\bigvee_t^b (f')\right)^p dt\right)^{1/p}\right\}(b-a)^{1/q}$$

$$\leq \frac{1}{(q+1)^{1/q}}\left[\left(\frac{x-a}{b-a}\right)^{q+1} + \left(\frac{b-x}{b-a}\right)^{q+1}\right]^{1/p}$$
$$\times\left[\int_a^x \left(\bigvee_a^t (f')\right)^p dt + \int_x^b \left(\bigvee_t^b (f')\right)^p dt\right]^{1/p}(b-a)^{1/q}$$
$$\leq \frac{1}{(q+1)^{1/q}}\left[\left(\frac{x-a}{b-a}\right)^{q+1} + \left(\frac{b-x}{b-a}\right)^{q+1}\right]^{1/p}$$
$$\times\left[(x-a)\left(\bigvee_a^x (f')\right)^p + (b-x)\left(\bigvee_x^b (f')\right)^p\right]^{1/p}(b-a)^{1/q}$$

for any $x \in [a,b]$ and $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$.

From the third branch in (145) we have

$$\left| f(x) + \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \tag{149}$$
$$\leq \frac{1}{b-a}\left[\int_a^x (t-a)\bigvee_a^t (f')\,dt + \int_x^b (b-t)\bigvee_t^b (f')\,dt\right]$$
$$\leq \left(\frac{x-a}{b-a}\right)\int_a^x \left(\bigvee_a^t (f')\right) dt + \left(\frac{b-x}{b-a}\right)\int_x^b \left(\bigvee_t^b (f')\right) dt$$

$$\leq \begin{cases} \left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]\left[\int_a^x \left(\bigvee_a^t (f')\right) dt + \int_x^b \left(\bigvee_t^b (f')\right) dt\right] \\ \left[\left(\frac{x-a}{b-a}\right)^q + \left(\frac{b-x}{b-a}\right)^q\right]^{1/q} \\ \times \left[\left[\int_a^x \left(\bigvee_a^t (f')\right) dt\right]^p + \left[\int_x^b \left(\bigvee_t^b (f')\right) dt\right]^p\right]^{1/p} \\ \max\left\{\int_a^x \left(\bigvee_a^t (f')\right) dt, \int_x^b \left(\bigvee_t^b (f')\right) dt\right\} \end{cases}$$

for any $x \in [a,b]$ and $p > 1$, $\frac{1}{p}+\frac{1}{q}=1$.

Remark 19. We observe that if we take $x = \frac{a+b}{2}$ in (147) then we get the perturbed midpoint inequality

$$\left|f\left(\frac{a+b}{2}\right)+\frac{1}{8}(b-a)\left[f'(b)-f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{150}$$
$$\leq \frac{1}{b-a}\left[\int_a^{\frac{a+b}{2}} (t-a)\bigvee_a^t (f')\,dt + \int_{\frac{a+b}{2}}^b (b-t)\bigvee_t^b (f')\,dt\right]$$
$$\leq \frac{1}{8}(b-a)\bigvee_a^b (f').$$

Inequalities for Lipschitzian Derivatives

We start with the following result.

Theorem 22 (Dragomir, 2013 [13]). *Let $f: I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. Let $x \in (a,b)$. If $\alpha_i > -1$ and $L_{\alpha_i} > 0$ with $i = 1,2$ are such that*

$$\left|f'(t)-f'(a)\right| \leq L_{\alpha_1}(t-a)^{\alpha_1} \text{ for any } t \in [a,x) \tag{151}$$

and

$$\left|f'(b)-f'(t)\right| \leq L_{\alpha_2}(b-t)^{\alpha_2} \text{ for any } t \in (x,b], \tag{152}$$

then we have

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^{2}f'(b)-(x-a)^{2}f'(a)\right]-\frac{1}{b-a}\int_{a}^{b}f(t)\,dt\right| \tag{153}$$
$$\leq\frac{1}{b-a}\left[\frac{L_{\alpha_1}}{\alpha_1+2}(x-a)^{\alpha_1+2}+\frac{L_{\alpha_2}}{\alpha_2+2}(b-x)^{\alpha_2+2}\right].$$

Proof. Using the conditions (151) and (152) we have

$$\begin{aligned}&\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^{2}f'(b)-(x-a)^{2}f'(a)\right]-\frac{1}{b-a}\int_{a}^{b}f(t)\,dt\right|\\&\leq\frac{1}{b-a}\int_{a}^{x}(t-a)\left|f'(t)-f'(a)\right|dt+\frac{1}{b-a}\int_{x}^{b}(b-t)\left|f'(b)-f'(t)\right|dt\\&\leq\frac{1}{b-a}L_{\alpha_1}\int_{a}^{x}(t-a)^{\alpha_1+1}dt+\frac{1}{b-a}L_{\alpha_2}\int_{x}^{b}(b-t)^{\alpha_2+1}dt\\&=\frac{1}{b-a}L_{\alpha_1}\frac{(x-a)^{\alpha_1+2}}{\alpha_1+2}+\frac{1}{b-a}L_{\alpha_2}\frac{(b-x)^{\alpha_2+2}}{\alpha_2+2}\\&=\frac{1}{b-a}\left[\frac{L_{\alpha_1}}{\alpha_1+2}(x-a)^{\alpha_1+2}+\frac{L_{\alpha_2}}{\alpha_2+2}(b-x)^{\alpha_2+2}\right]\end{aligned}$$

and the inequality (153) is obtained.

Corollary 24. *Let $f: I\to\mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b]\subset\mathring{I}$. If the derivative is f' of r-H-Hölder type on $[a,b]$, i.e. we have the condition*

$$\left|f'(t)-f'(s)\right|\leq H\left|t-s\right|^{r}$$

for any $t,s\in[a,b]$, where $r\in(0,1]$ and $H>0$ are given, then

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^{2}f'(b)-(x-a)^{2}f'(a)\right]-\frac{1}{b-a}\int_{a}^{b}f(t)\,dt\right| \tag{154}$$
$$\leq\frac{H}{r+2}\left[\left(\frac{x-a}{b-a}\right)^{r+2}+\left(\frac{b-x}{b-a}\right)^{r+2}\right](b-a)^{r+1},$$

for any $x\in[a,b]$.

In particular, if f' is Lipschitzian with the constant $L>0$, then

$$\left|f(x)+\frac{1}{2(b-a)}\left[(b-x)^{2}f'(b)-(x-a)^{2}f'(a)\right]-\frac{1}{b-a}\int_{a}^{b}f(t)\,dt\right| \tag{155}$$
$$\leq\frac{1}{3}L\left[\left(\frac{x-a}{b-a}\right)^{3}+\left(\frac{b-x}{b-a}\right)^{3}\right](b-a)^{2},$$

for any $x\in[a,b]$.

Remark 20. With the assumptions of Corollary 24 we have the midpoint inequality

$$\left|f\left(\frac{a+b}{2}\right)+\frac{1}{8}(b-a)\left[f'(b)-f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{156}$$
$$\leq \frac{H}{2^{r+1}(r+2)}(b-a)^{r+1}.$$

If if f' is Lipschitzian with the constant $L>0$, then

$$\left|f\left(\frac{a+b}{2}\right)+\frac{1}{8}(b-a)\left[f'(b)-f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \tag{157}$$
$$\leq \frac{1}{12}L(b-a)^2.$$

Inequalities for Differentiable Functions with the Property (S)

Let $f : I \to \mathbb{C}$ be a differentiable convex function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. Then f' is monotonic nondecreasing and by the equality (143) we have

$$f(x)+\frac{1}{2(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt \geq 0 \tag{158}$$

or, equivalently

$$\frac{1}{2(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right] \geq \frac{1}{b-a}\int_a^b f(t)\,dt - f(x) \tag{159}$$

for any $x \in [a,b]$.

We observe that the inequalities (158) and (159) remain valid for the larger class of differentiable functions f that satisfy the *property* (S) on the interval $[a,b]$, namely

$$f'(a) \leq f'(t) \leq f'(b) \tag{S}$$

for any $t \in [a,b]$.

Theorem 23 (Dragomir, 2013 [13]). *Let $f : I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$.*

(i) Let $x \in [a,b]$. If f satisfies the property (S) on the interval $[a,x]$ and $[x,b]$, then

$$f'(x)\left(\frac{a+b}{2}-x\right) \le \frac{1}{b-a}\int_a^b f(t)\,dt - f(x). \tag{160}$$

(ii) If f satisfies the property (S) on the interval $[a,b]$ *, then for any* $x \in [a,b]$

$$\frac{f(a)(x-a)+f(b)(b-x)}{b-a} - \frac{1}{b-a}\int_a^b f(t)\,dt \tag{161}$$
$$\le \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right].$$

Proof. (i) Since f satisfies the property (S) on the interval $[a,x]$ and $[x,b]$, then

$$f(x) + \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - \frac{1}{b-a}\int_a^b f(t)\,dt$$
$$= \frac{1}{b-a}\int_a^x (t-a)\left[f'(t)-f'(a)\right]dt + \frac{1}{b-a}\int_x^b (b-t)\left[f'(b)-f'(t)\right]dt$$
$$\le \frac{1}{b-a}\int_a^x (t-a)\left[f'(x)-f'(a)\right]dt + \frac{1}{b-a}\int_x^b (b-t)\left[f'(b)-f'(x)\right]dt$$
$$= \frac{f'(x)-f'(a)}{b-a}\int_a^x (t-a)\,dt + \frac{f'(b)-f'(x)}{b-a}\int_x^b (b-t)\,dt$$
$$= \frac{f'(x)-f'(a)}{b-a}\cdot\frac{(x-a)^2}{2} + \frac{f'(b)-f'(x)}{b-a}\cdot\frac{(b-x)^2}{2}$$
$$= \frac{1}{2(b-a)}\left[(b-x)^2 f'(b) - (x-a)^2 f'(a)\right] - f'(x)\left(\frac{a+b}{2}-x\right),$$

which proves the inequality (160).

(ii) If f satisfies the property (S) on the interval $[a,b]$, then for any $x \in [a,b]$

$$\frac{1}{b-a}\int_a^x (t-a)\left[f'(t)-f'(a)\right]dt + \frac{1}{b-a}\int_x^b (b-t)\left[f'(b)-f'(t)\right]dt$$
$$\le \frac{x-a}{b-a}\int_a^x \left[f'(t)-f'(a)\right]dt + \frac{b-x}{b-a}\int_x^b \left[f'(b)-f'(t)\right]dt$$
$$= \frac{1}{b-a}(x-a)\left[f(x)-f(a)-f'(a)(x-a)\right]$$
$$+ \frac{1}{b-a}(b-x)\left[f'(b)(b-x)-f(b)+f(x)\right]$$
$$= \frac{1}{b-a}\left[f(x)(x-a)-f(a)(x-a)-f'(a)(x-a)^2\right]$$

$$+\frac{1}{b-a}\left[f'(b)(b-x)^2-f(b)(b-x)+f(x)(b-x)\right]$$
$$=\frac{f'(b)(b-x)^2-f'(a)(x-a)^2}{b-a}+f(x)-\frac{f(a)(x-a)+f(b)(b-x)}{b-a},$$

which proves the inequality (161).

Remark 21. The inequality (160) was obtained for the case of convex functions in [5] while (161) was established for convex functions in [6] with different proofs.

Further, we use the Čebyšev inequality for synchronous functions (functions with same monotonicity), namely

$$\frac{1}{d-c}\int_c^d g(t)\,h(t)\,dt \geq \frac{1}{d-c}\int_c^d g(t)\,dt\cdot\frac{1}{d-c}\int_c^d h(t)\,dt. \tag{162}$$

Theorem 24 (Dragomir, 2013 [13]). *Let $f: I \to \mathbb{C}$ be a differentiable function on $\mathring{I}$ and $[a,b] \subset \mathring{I}$. Let $x \in [a,b]$. If f is convex on the interval $[a,x]$ and $[x,b]$, then*

$$\frac{1}{2}\left[f(x)+\frac{f(a)(x-a)+f(b)(b-x)}{b-a}\right] \geq \frac{1}{b-a}\int_a^b f(t)\,dt. \tag{163}$$

Proof. We have

$$f(x)+\frac{1}{2(b-a)}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]-\frac{1}{b-a}\int_a^b f(t)\,dt \tag{164}$$
$$=\frac{1}{b-a}\int_a^x (t-a)\left[f'(t)-f'(a)\right]dt+\frac{1}{b-a}\int_x^b (b-t)\left[f'(b)-f'(t)\right]dt$$

for any $x \in [a,b]$.

Since f' is monotonic nondecreasing on $[a,x]$, then by Čebyšev inequality (154) we have

$$\int_a^x (t-a)\left[f'(t)-f'(a)\right]dt \geq \frac{1}{x-a}\int_a^x (t-a)\,dt\cdot\int_a^x \left[f'(t)-f'(a)\right]dt$$
$$=\frac{1}{2}\left[f(x)(x-a)-f(a)(x-a)-f'(a)(x-a)^2\right]$$

and, by the same inequality,

$$\int_x^b (b-t)\left[f'(b)-f'(t)\right]dt \geq \frac{1}{b-x}\int_x^b (b-t)\,dt\cdot\int_x^b \left[f'(b)-f'(t)\right]dt$$
$$=\frac{1}{2}\left[f'(b)(b-x)^2-f(b)(b-x)+f(x)(b-x)\right].$$

If we add these two inequalities, then we get

$$\int_a^x (t-a)\left[f'(t)-f'(a)\right]dt+\int_x^b (b-t)\left[f'(b)-f'(t)\right]dt$$
$$\geq \frac{1}{2}\left[f(x)(x-a)-f(a)(x-a)-f'(a)(x-a)^2\right]$$
$$+\frac{1}{2}\left[f'(b)(b-x)^2-f(b)(b-x)+f(x)(b-x)\right]$$
$$=\frac{1}{2}\left[(b-x)^2 f'(b)-(x-a)^2 f'(a)\right]+\frac{1}{2}f(x)(b-a)$$
$$-\frac{1}{2}\left[f(a)(x-a)+f(b)(b-x)\right].$$

Dividing by $b-a$ and utilizing the equality (164) we deduce the inequality (163).

Remark 22. If the function is convex on the whole interval $[a,b]$, then the inequality (163) is true for any $x \in [a,b]$.

References

1. Apostol, T.M.: Mathematical Analysis, 2nd edn. Addison-Wesley, Reading (1975)
2. Dragomir, S.S.: On the Ostrowski's integral inequality for mappings with bounded variation and applications. Math. Inequal. Appl. **4**(1), 33–40 (2001). Preprint, RGMIA Res. Rep. Collect. **2**(1), Article 7 (1999). http://rgmia.org/papers/v2n1/v2n1-7.pdf
3. Dragomir, S.S.: A refinement of Ostrowski's inequality for absolutely continuous functions and applications. Acta Math. Vietnam. **27**(2), 203–217 (2002)
4. Dragomir, S.S.: A refinement of Ostrowski's inequality for absolutely continuous functions whose derivatives belong to L_∞ and applications. Lib. Math. **22**, 49–63 (2002)
5. Dragomir, S.S.: An inequality improving the first Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products. J. Inequal. Pure Appl. Math. **3**(2), Article 31, 8 pp. (2002)
6. Dragomir, S.S.: An inequality improving the second Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products. J. Inequal. Pure Appl. Math. **3**(3), Article 35, 8 pp. (2002)
7. Dragomir, S.S.: Improvements of Ostrowski and generalised trapezoid inequality in terms of the upper and lower bounds of the first derivative. Tamkang J. Math. **34**(3), 213–222 (2003)
8. Dragomir, S.S.: Refinements of the generalised trapezoid and Ostrowski inequalities for functions of bounded variation. Arch. Math. **91**(5), 450–460 (2008)
9. Dragomir, S.S.: A functional generalization of Ostrowski inequality via Montgomery identity. Preprint, Acta Math. Univ. Comenian. (2015), **84**(1), 63–78. RGMIA Res. Rep. Collect. **16**, Article 65, pp. 15 (2013). http://rgmia.org/papers/v16/v16a65.pdf
10. Dragomir, S.S.: Some perturbed Ostrowski type inequalities for functions of bounded variation. Preprint, Asian-Eur. J. Math. (2015), **8**(4), 1550069, pp. 14 RGMIA Res. Rep. Collect. **16**, Article 93 (2013). http://rgmia.org/papers/v16/v16a93.pdf
11. Dragomir, S.S.: Some perturbed Ostrowski type inequalities for absolutely continuous functions (I). Theory Appl. Math. Comput. Sci. (2015), **5**(2), 132–147. Preprint, RGMIA Res. Rep. Collect. **16**, Article 94 (2013). http://rgmia.org/papers/v16/v16a94.pdf

12. Dragomir, S.S.: Some perturbed Ostrowski type inequalities for absolutely continuous functions (II). Acta Univ. Apulensis Math. Inform. (43) (2015), 209–228. Preprint, RGMIA Res. Rep. Collect. **16**, Article 95 (2013)
13. Dragomir, S.S.: Some perturbed Ostrowski type inequalities for absolutely continuous functions (III). Transylv. J. Math. Mech. (2015), **7**(1), 31–43. Preprint RGMIA Res. Rep. Collect. **16**, Article 96 (2013). http://rgmia.org/papers/v16/v16a96.pdf
14. Dragomir, S.S., Rassias, Th.M. (eds.): Ostrowski Type Inequalities and Applications in Numerical Integration. Springer, Dordrecht (2002). Print ISBN 978-90-481-5990-1
15. Dragomir, S.S., Wang, S.: A new inequality of Ostrowski's type in L_1 norm and applications to some special means and to some numerical quadrature rules. Tamkang J. Math. **28**, 239–244 (1997)
16. Dragomir, S.S., Wang, S.: A new inequality of Ostrowski's type in L_p norm and applications to some special means and to some numerical quadrature rules. Indian J. Math. **40**(3), 299–304 (1998)
17. Dragomir, S.S., Wang, S.: Applications of Ostrowski's inequality to the estimation of error bounds for some special means and for some numerical quadrature rules. Appl. Math. Lett. **11**(1), 105–109 (1998)
18. Fink, A.M.: Bounds on the derivative of a function from its averages. Czechoslov. Math. J. **42**(117), 289–310 (1992)
19. Milovanović, G.V., Rassias, M.Th. (eds.): Analytic Number Theory, Approximation Theory, and Special Functions. Springer, New York (2014)
20. Mitrinović, D.S., Pečarić, J.E., Fink, A.M.: Inequalities for Functions and Their Integrals and Derivatives. Kluwer Academic, Dordrecht (1994)
21. Ostrowski, A.: Über die Absolutabweichung einer differentiierbaren Funktion von ihrem Integralmittelwert. Comment. Math. Helv. **10**, 226–227 (1938)
22. Peachey, T.C., Mcandrew, A., Dragomir, S.S.: The best constant in an inequality of Ostrowski type. Tamkang J. Math. **30**(3), 219–222 (1999)

Hyers–Ulam–Rassias Stability of the Generalized Wilson's Functional Equation

Elqorachi Elhoucien, Youssef Manar, and Sammad Khalil

Abstract In this chapter, we apply the fixed point theorem and the direct method to the proof of Hyers–Ulam–Rassias stability property for generalized Wilson's functional equation

$$\int_K \int_G f(xtk.y)dkd\mu(t) = f(x)g(y), \ x, y \in G,$$

where f, g are continuous complex valued functions on a locally compact group G, K is a compact subgroup of morphisms of G, dk is the normalized Haar measure on K and μ is a K-invariant complex measure with compact support.

Keywords Banach space • Complex measure • Fixed point • Hyers-Ulam-Rassias stability • Locally compact group • Wilson's functional equation

Introduction

In 1940, Ulam [33] gave a wide-ranging talk before the mathematics club of the University of Wisconsin in which he discussed a number of important unsolved problems. Among those was the question concerning the stability of a group homomorphisms:

> "Let G be a group and let (H, d) be a metric group. Given $\epsilon > 0$, does there exists a $\delta > 0$ such that if a function $f : G \to H$ satisfies the inequality $d(f(xy), f(x)f(y)) \leq \delta$ for all $x, y \in G$, then there exists a homomorphism $a : G \to H$ such that $d(f(x), a(x)) \leq \epsilon$ for all $x \in G$?"

E. Elhoucien (✉)
Faculty of Sciences, Department of Mathematics, Ibn Zohr University, Agadir, Morocco
e-mail: elqorachi@hotmail.com

Y. Manar
Superior School of Technology, Ibn Zohr University, Guelmim, Morocco
e-mail: manaryoussef1984@gmail.com

S. Khalil
Faculty of Sciences, Department of Mathematics, Ibn Tofail University, Kenitra, Morocco
e-mail: khalil.Sammad@lehman.cuny.edu

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_9

The case of approximately additive functions was solved by Hyers [17] under the condition that G and H are Banach spaces. Indeed the method used by Hyers is called the direct method. Rassias [26] attempted to weaken the condition for the bound of the norm of the Cauchy difference as follows: $\|f(x+y)-f(x)-f(y)\| \leq \epsilon$ $(\|x\|^p + \|y\|^p)$ and derived Hyer's theorem for the stability of the additive mapping as a special case. Thus in [26], a proof of the generalized Hyers–Ulam stability for the linear mapping between Banach spaces was obtained. A particular case of Rassias's theorem regarding the Hyers–Ulam stability of the additive mapping was proved by Aoki [1]. The stability concept that was introduced by Rassias's theorem provided a large influence to a number of mathematicians to develop the notion of what is known today as Hyers–Ulam–Rassias stability. Since then, the stability of several functional equations has been extensively investigated by several mathematicians. The terminology Hyers–Ulam–Rassias stability originates from those historical backgrounds. The terminology can also be applied to the case of other functional equations. For more detailed definitions of such terminologies, we refer to [14, 18–20, 22, 25, 27].

In 2003 Cadăriu and Radu [7] notice that a fixed point alternative method is very important for the solution of the Hyers–Ulam–Rassias stability problem. Subsequently, this method was applied to investigate the Hyers–Ulam–Rassias stability for Jensen's functional equation, as well as for the additive Cauchy's functional equation by considering a general control function $\varphi(x, y)$, with suitable properties. Using such an elegant idea, several authors applied the method to investigate the stability of some functional equations see, for example, [2, 3, 5, 9, 15, 16, 21, 23, 24].

In this chapter, we will apply the fixed point method and the direct method to prove the Hyers–Ulam–Rassias stability of generalized Wilson's functional equation

$$\int_K \int_G f(xtk.y)dkd\mu(t) = f(x)g(y), \; x, y \in G, \tag{1}$$

where K is a compact subgroup of morphisms of G, μ is a K-invariant measure with compact support and dk is the normalized Haar measure on K.

Solution for the particular cases of the functional equation (1) was investigated by several authors. An important example of (1) with $\mu = \delta_e$ (Dirac measure concentrated at the identity element of G) is the functional equation of a K-spherical function

$$\int_K f(xk.y)dk = f(x)f(y), \; x, y \in G. \tag{2}$$

This functional equation with $G = \mathbf{C}$, K finite ($K = \mathbf{Z}_q$), viz. $\sum_{k=0}^{q-1} f(x + w^k) = f(x)f(y)$, where $w = \exp(\frac{2\pi i}{q})$, occurs in Förg-Rob and Schwaiger [13].

The bounded solution of K-spherical function in Abelians group is obtained by Chojnacki [10] and later by Badora [4] while Stetkaer [29–32] studied some generalization of (1).

In [28] Shin'y described all continuous solution of (2) on Abelian groups.

Badora's equation:

$$\int_K \int_G f(xtk.y)dkd\mu(t) = f(x)f(y), \; x, y \in G, \tag{3}$$

is considered in [4]. Non-zero essentially bounded solutions of Eq. (3) are obtained, see [3, 4, 12].

The Hyers–Ulam stability of Eq. (1) was recently obtained by Bouikhalene and Elqorachi [6].

The results of the present chapter are organized as follows.

In section "Hyers–Ulam–Rassias Stability of Eq. (1): Fixed Point Method," we apply the fixed point method to prove the Hyers–Ulam–Rassias stability of the functional equation (1).

In section "Hyers–Ulam–Rassias Stability of Eq. (1): Direct Method," we apply the direct method to prove the Hyers–Ulam–Rassias stability of that equation.

Preliminaries

Throughout this chapter, let G denote a locally compact group, let K be a compact subgroup of morphisms of G. The action of $k \in K$ on $x \in G$ will be denoted by $k \cdot x$ and the normalized Haar measure on K by dk. Let $M(G)$ denote the topological dual of $C_0(G)$: the Banach space of continuous functions vanishing at infinity. For any $\mu \in M(G)$ with compact support and any continuous function f on G we put $< k \cdot \mu, f > = < \mu, k \cdot f >$, where $(k \cdot f)(x) = f(k^{-1} \cdot x)$ for all $x \in G$ and we say that μ is K-invariant if $k \cdot \mu = \mu$ for all $k \in K$. We say that f satisfy Kannappan type Condition if

$$\int_K \int_G f(ztxsy)d\mu(t)d\mu(s) = \int_K \int_G f(ztysx)d\mu(t)d\mu(s). \tag{4}$$

for all $x, y, z \in G$.

Let X be a set. A function $d : X \times X \to [0, \infty]$ is called a generalized metric on X if d satisfies the following properties.

(i) $d(x, y) = 0$ if and only if $x = y$;
(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$;
(iii) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Note that the only substantial difference on the generalized metric from the metric is that the range of generalized includes the infinity. We now recall one of the fundamental results of fixed point theory. For the proof, we refer to [11].

Theorem 1 ([11]). *Let* (X, d) *be complete generalized metric space. Assume that* $J : X \to X$ *is a strictly contractive operator with the Lipshitz constant* $0 < L < 1$.

If there exists a nonnegative integer n_0 *such that* $d(J^{n_0}x, J^{n_0+1}x) < +\infty$, *for some* $x \in X$, *then we have the following properties:*

1. *the sequence* $J^n x$ *converges to a fixed point* x^* *of* J;
2. x^* *is the unique fixed point of* J *in the set* $Y = \{y \in X : d(J^{n_0}x, y) < \infty\}$;
3. $d(y, x^*) \leq \frac{1}{1-L} d(y, Jy)$ *for all* $y \in Y$.

Hyers–Ulam–Rassias Stability of Eq. (1): Fixed Point Method

In this section, by using an idea of Cădariu and Radu (see [7, 8]), we prove the Hyers–Ulam–Rassias stability of generalized Wilson's functional equation (1).

For any complex valued function g on a group G we define the set $G_{g,\mu} = \{a \in G / |g(a)| > \|\mu\|\}$.

Theorem 2. *Let* $f, g : G \to \mathbf{C}$ *be the continuous functions satisfying the inequality*

$$\left| \int_K \int_G f(xtk \cdot y) dk d\mu(t) - g(y) f(x) \right| \leq \psi(x, y) \tag{5}$$

for all $x, y \in G$ *and for some continuous mapping* $\psi : G \to \mathbf{R}$. *If* $G_{\mu,g} \neq \emptyset$,

$$\left| \int_K \int_G \psi(xtk \cdot a, y) d|\nu|(t) dk \right| \leq \psi(x, y), \text{ for all } x, y \in G \text{ and for all } a \in G_{\mu,g} \tag{6}$$

where $\nu = \frac{\mu}{\|\mu\|}$ *and* f *satisfies the Kannappan type condition (4). Then there exists a unique function* $\mathscr{F} : G \to \mathbf{C}$ *such that*

$$\int_K \int_G \mathscr{F}(xtk \cdot y) dk d\mu(t) = g(y) \mathscr{F}(x) \tag{7}$$

for all $x, y \in G$,

$$\left(\int_K \int_G g(xtk \cdot y) dk d\mu(t) - g(x) g(y) \right) \mathscr{F}(z) = 0 \text{ for all } x, y, z \in G. \tag{8}$$

and

$$|f(x) - \mathscr{F}_\mu(x)| \leq \inf_{a \in S_{\mu,f}} \frac{\psi(a, x)}{|g(a)| - \|\mu\|} \text{ for all } x \in G. \tag{9}$$

Proof. Dividing the both sides of (5) by $\|\mu\|^2$ and replacing y by a, we get

$$\int_K \int_G F(xtk \cdot a) dk d\nu(t) - \alpha F(x) \leq \frac{\psi(x, a)}{\|\mu\|^2} \text{ for all } x \in G, \tag{10}$$

where $F = \frac{f}{\|\mu\|}$, $\nu = \frac{\mu}{|\mu\|}$ and $\alpha = \frac{g(a)}{\|\mu\|} > 1$. We consider the set $A = \{h : G \to \mathbf{C}\}$. First, we denote by d the generalized metric on A defined as

$$d(h,l) = \inf\{K \in [0,\infty] \; |h(x) - l(x)| \leq K\frac{\psi(x,a)}{\|\mu\|^2}, \; for \; all \; x \in G\}$$

Then, as in the proof of Cadariu and Radu [7, 8], we can show that (A, d) is complete.

Now, let us define an operator $J_a : A \to A$ by

$$(J_a h)(y) = \frac{1}{\alpha}\int_K \int_G h(ytk \cdot a)dkd\nu(t) \; for \; every \; y \in G.$$

Next, we are going to prove that J_a is a strictly contractive on A with the *Lipschitz constant L*. Given h and l in A and $K \geq 0$ an arbitrary constant with $d(g,h) \leq K$, that is,

$$|g(x) - h(x)| \leq K\frac{\psi(x,a)}{\|\mu\|^2} \tag{11}$$

for all $x \in G$ so, from inequality (5) we have

$$\begin{aligned} |(J_a l)(x) - (J_a h)(x)| &= \left|\frac{1}{\alpha}\int_K \int_G l(xtk \cdot a)d\nu(t)dk - \frac{1}{\alpha}\int_K \int_G h(xtk \cdot a)d\nu(t)dk\right| \\ &\leq \frac{1}{|\alpha|}\frac{K}{\|\mu\|^2}\int_K \int_G \psi(xtk.a, a)d|\nu|(t)dk \\ &\leq \frac{1}{|\alpha|}K\frac{\psi(x,a)}{\|\mu\|^2} \end{aligned}$$

for all $x \in G$. This means that $d(J_a(h), J_a(l)) \leq Ld(h,l)$ for any $h, l \in A$, where $L = \frac{\|\mu\|}{|g(a)|}$.

On the other hand inequality (10) can be written as follows, we get

$$|(J_a F)(x) - F(x)| \leq \frac{1}{|\alpha|}\frac{\psi(x,a)}{\|\mu\|^2}$$

for all $x \in G$, which implies that

$$d(J_a F, F) \leq \frac{1}{|\alpha|} < \infty. \tag{12}$$

Taking $n_0 = 0$ in Theorem 1 there exists a function $f_a : G \to \mathbf{C}$ such that

1. f_a is the unique fixed point of J_a in the set $A_\infty = \{h \in A/d(F,h) < \infty\}$;
2. $\lim_{n\longrightarrow\infty} d(J_a^n F, f_a) = 0$;

3. $d(F,f_a) \leq \frac{1}{1-L} d(J_a F, F) \leq \frac{\|\mu\|}{|g(a)|-\|\mu\|}$.

Moreover, since f satisfies the Kannappan type condition (4), we get by easy computation that $J_a^n F$ satisfies (4) and then the function f_a verify the Kannapaann type condition (4).

Now, we will prove that f_a is independent of the variable a. That is $f_a = f_b$ for all $a, b \in G_{\mu,g}$.

For every x in G and for all $n \in \mathbf{N}^*$, we define a sequence functions F_n by

$$F_1(x) = \int_K \int_G F(xtk \cdot a) dk d\nu(t), \quad F_{n+1}(x) = \int_K \int_G F_n(xtk \cdot a) dk d\nu(t),\ n \geq 1. \tag{13}$$

First step, we use induction on n to prove the following inequality

$$\|F_{n+1}(x) - \alpha F_n(x)\| \leq \frac{\psi(x,a)}{\|\mu\|^2} \tag{14}$$

for all $n \in \mathbf{N}$ and $x \in G$. It follows from (6) and (10) that

$$\begin{aligned}
&|F_2(x) - \alpha F_1(x)| \\
&= \left| \int_K \int_G F_1(xtk \cdot a) dk d\nu(t) - \alpha \int_K \int_G F(xtk \cdot a) dk d\nu(t) \right| \\
&= \left| \int_K \int_K \int_G \int_G F(xt_1k_1 \cdot at_2k_2 \cdot a) dk_1 dk_2 d\nu(t_1) d\nu(t_2) \right. \\
&\quad \left. -\alpha \int_K \int_G F(xt_1k_1 \cdot a) dk_1 d\nu(t_1) \right| \\
&\leq \int_K \int_G \left| \int_K \int_G F(xt_1k_1 \cdot at_2k_2 \cdot a) d\nu(t_2) dk_2 - \alpha F(xt_1k_1 \cdot a) \right| d|\nu(t_1)| dk_1 \\
&\leq \frac{1}{\|\mu\|^2} \int_K \int_G \psi(xt_1k_1 \cdot a, a) d|\nu(t_1)| dk_1 \\
&\leq \frac{\psi(x,a)}{\|\mu\|^2},
\end{aligned}$$

which proves that the assertion (14) is true for $n = 1$. Now, we assume that (14) is true for some $n \in \mathbf{N}$. By using (13) we have

$$\begin{aligned}
|F_{n+2}(x) - \alpha F_{n+1}(x)| &= \left| \int_K \int_G F_{n+1}(xtk \cdot a) dk d\nu(t) - \alpha \int_K \int_G F_n(xtk \cdot a) dk d\nu(t) \right| \\
&\leq \int_K \int_G |F_{n+1}(xtk \cdot a) - \alpha F_n(xtk \cdot a)| d|\nu(t)| dk
\end{aligned}$$

$$\leq \int_K \int_G \frac{\psi(x,a)}{\|\mu\|^2} dkd|\nu(t)|$$
$$\leq \frac{\psi(x,a)}{\|\mu\|^2},$$

which implies the validity of the inequality (14) for all $n \in \mathbf{N}$.

Second step, we improve that

$$|F_n(x) - \alpha^n F(x)| \leq \frac{\psi(x,a)}{\|\mu\|^2}(1 + |\alpha| + |\alpha|^2 + \ldots + |\alpha|^{n-1}). \tag{15}$$

For $n = 1$. From (10) we have

$$|F_1(x) - \alpha F(x)| = \left| \int_K \int_G F(xtk \cdot a) dkd\nu(t) - \alpha F(x) \right|$$
$$\leq \frac{\psi(x,a)}{\|\mu\|^2}$$

for all $x \in G$. Suppose that the inequality (15) holds for $n \in \mathbf{N}$. Using the triangle inequality and taking (15) into account we have

$$|F_{n+1}(x) - \alpha^{n+1} F(x)| \leq |F_{n+1}(x) - \alpha F_n(x)| + |\alpha||F_n(x) - \alpha^n F(x)|$$
$$\leq \frac{\psi(x,a)}{\|\mu\|^2} + |\alpha| \frac{\psi(x,a)}{\|\mu\|^2}(1 + |\alpha| + |\alpha|^2 + \ldots + |\alpha|^{n-1})$$
$$\leq \frac{\psi(x,a)}{\|\mu\|^2}(1 + |\alpha| + |\alpha|^2 + \ldots + |\alpha|^n).$$

This implies the inequality (15) for all $n \in \mathbf{N}$.

Now, we are ready to prove that $f_a = f_b$ for all $a, b \in G_{\mu,g}$. By using (15), it easy to check that, for all $a, b \in G_{\mu,g}$

$$|F_n(x) - \alpha^n F(x)| \leq \frac{\psi(x,a)}{\|\mu\|^2}(\frac{|\alpha|^n - 1}{|\alpha| - 1}) \tag{16}$$

and

$$|G_n(x) - \beta^n F(x)| \leq \psi(x,b)(\frac{|\beta|^n - 1}{|\beta| - 1}) \tag{17}$$

where G_n is defined by (13) by replacing the variable a by b.

Moreover, f satisfies the Kannappan type condition (4), so, we have

$$\int_K \int_G \int_K \int_G f(xtk \cdot ask' \cdot b) d\nu(t) d\nu(s) dk dk'$$
$$= \int_K \int_G \int_K \int_G f(xtk' \cdot bsk \cdot a) d\nu(t) d\nu(s) dk dk'$$

for all $x, a, b \in G$. On the other hand, we have

$$\left| \int_G \int_K \dots F(xt_1k_1 \cdot at_2k_2 \cdot a \dots t_nk_n \cdot a) dk_1 dk_2 \dots dk_n d\nu(t_1) d\nu(t_2) \dots d\nu(t_n) - \alpha^n F(x) \right|$$
$$\leq \psi(x, a)(\frac{|\alpha|^n - 1}{|\alpha| - 1})$$

and

$$\left| \int_G \int_K \dots F(xt_1k_1 \cdot bt_2k_2 \cdot b \dots t_nk_n \cdot b) dk_1 dk_2 \dots dk_n d\nu(t_1) d\nu(t_2) \dots d\nu(t_n) - \beta^n F(x) \right|$$
$$\leq \psi(x, a)(\frac{|\beta|^n - 1}{|\beta| - 1}),$$

where $\beta = \frac{g(b)}{\|\mu\|}$. So, we get

$$| \int_G \int_K \dots F(xs_1k_1' \cdot bs_2k_2' \cdot b \dots s_nk_n' \cdot bt_1k_1 \cdot at_2k_2 \cdot a \dots t_nk_n \cdot a)$$
$$dk_1' \dots dk_n' dk_1 \dots dk_n d\nu(s_1) \dots d\nu(s_n) d\nu(t_1) \dots d\nu(t_n)$$
$$-\alpha^n \int_G \int_K \dots F(xs_1k_1' \cdot bs_2k_2' \cdot b \dots s_nk_n' \cdot b) d\nu(s_1) \dots d\nu(s_n) dk_1' \dots dk_n' |$$
$$\leq (\frac{|\alpha|^n - 1}{|\alpha| - 1}) \left| \int_G \int_K \dots \psi(xs_1k_1' \cdot bs_2k_2' \cdot b \dots s_nk_n' \cdot b, a) d\nu(s_1) \dots d\nu(s_n) dk_1' \dots dk_n' \right|$$
$$\leq (\frac{|\alpha|^n - 1}{|\alpha| - 1}) \psi(x, a),$$

and

$$| \int_G \int_K \dots F(xs_1k_1' \cdot as_2k_2' \cdot a \dots s_nk_n' \cdot at_1k_1 \cdot bt_2k_2 \cdot b \dots t_nk_n \cdot b)$$
$$d\nu(s_1) \dots d\nu(s_n) dk_1' \dots dk_n' d\nu(t_1) \dots d\nu(t_n) dk_1 \dots dk_n$$
$$-\beta^n \int_G \int_K \dots F(xs_1k_1' \cdot as_2k_2' \cdot a \dots s_nk_n' \cdot a) d\nu(s_1) \dots d\nu(s_n) dk_1' \dots dk_n' |$$
$$\leq (\frac{|\beta|^n - 1}{|\beta| - 1}) \psi(x, b).$$

Since

$$\int_G \int_K \dots F(xs_1k_1^{'} \cdot as_2k_2^{'} \cdot a \dots s_nk_n^{'} \cdot at_1k_1 \cdot bt_2k_2 \cdot b \dots t_nk_n \cdot b)$$
$$d\nu(s_1)\dots d\nu(s_n)dk_1^{'}\dots dk_n^{'}d\nu(t_1)\dots d\nu(t_n)dk_1\dots dk_n$$
$$= \int_G \int_K \dots F(xt_1k_1 \cdot bt_2k_2 \cdot b \dots t_nk_n \cdot bs_1k_1^{'} \cdot as_2k_2^{'} \cdot a \dots s_nk_n^{'} \cdot a)$$
$$d\nu(t_1)\dots d\nu(t_n)dk_1\dots dk_nd\nu(s_1)\dots d\nu(s_n)dk_1^{'}\dots dk_n^{'},$$

and by the triangle inequality we obtain

$$\begin{aligned}&|\alpha^n \int_K \dots \int_K F(xs_1k_1^{'} \cdot bs_2k_2^{'} \cdot b \dots s_nk_n^{'} \cdot b)dk_1^{'}\dots dk_n^{'}d\nu(s_1)\dots d\nu(s_n) \\ &-\beta^n \int_K \dots \int_K F(xs_1k_1^{'} \cdot as_2k_2^{'} \cdot b \dots s_nk_n^{'} \cdot a)dk_1^{'}\dots dk_n^{'}d\nu(s_1)\dots d\nu(s_n)| \\ &\le \psi(x,a)(\frac{|\alpha|^n-1}{|\alpha|-1}) + \psi(x,b)(\frac{|\beta|^n-1}{|\beta|-1}).\end{aligned} \tag{18}$$

Dividing both sides of (18) by $|\alpha|^n|\beta|^n$ we get

$$\begin{aligned}&|\beta^{-n} \int_K \dots \int_K F(xs_1k_1^{'} \cdot bs_2k_2^{'} \cdot b \dots s_nk_n^{'} \cdot b)dk_1^{'}\dots dk_n^{'}d\nu(s_1)\dots d\nu(s_n) \\ &-\alpha^{-n} \int_K \dots \int_K F(xs_1k_1^{'} \cdot as_2k_2^{'} \cdot b \dots s_nk_n^{'} \cdot a)dk_1^{'}\dots dk_n^{'}d\nu(s_1)\dots d\nu(s_n)| \\ &\le \frac{\psi(x,a)}{|\beta|^n(|\alpha|-1)}(1-\frac{1}{|\alpha|^n}) + \frac{\psi(x,b)}{|\alpha|^n(|\beta|-1)}(1-\frac{1}{|\beta|^n}),\end{aligned} \tag{19}$$

and by letting $n \to \infty$ we obtain $f_a(x) = f_b(x)$ for all $x \in G$. Finally there exists a unique function $\frac{\mathscr{F}}{\|\mu\|}$ such that $\frac{\mathscr{F}}{\|\mu\|} = f_a$ for all $a \in G_{\mu,g}$. From (13) and $a \in G_{\mu,g}$ arbitrary we have

$$|F(x) - \frac{\mathscr{F}}{\|\mu\|}(x)| \le \inf_{a\in S_{\mu,g}} \frac{1}{\|\mu\|}\frac{\psi(x,a)}{|g(a)|-\|\mu\|}, \tag{20}$$

that is

$$|f(x) - \mathscr{F}(x)| \le \inf_{a\in S_{\mu,g}} \frac{\psi(x,a)}{|g(a)|-\|\mu\|} \tag{21}$$

for all $x \in G$.

Now, we will show that $\mathscr{F}$ is a solution of functional equation

$$\int_G \int_K \mathscr{F}(xtk \cdot y)dkd\mu(t) = g(y)\mathscr{F}(x) \tag{22}$$

for all $x, y \in G$. By using (10) and (6) we get

$$
\begin{aligned}
&|\int_G \dots \int_G \int_K \dots \int_K F(xt_1k_1 \cdot at_2k_2 \cdot a \dots t_nk_n \cdot atk.y)d\nu(t)d\nu(t_1)\dots d\nu(t_n)dkdk_1\dots dk_n \\
&-\frac{g(y)}{\|\mu\|}\int_G \dots \int_G \int_K \dots \int_K F(xt_1k_1.at_2k_2 \cdot a \dots t_nk_n \cdot a)d\nu(t_1)\dots d\nu(t_n)dk_1\dots dk_n| \\
&\leq \frac{1}{\|\mu\|^2}\int_G \dots \int_G \psi(xt_1k_1 \cdot at_2k_2 \cdot a \dots t_nk_n \cdot a, y)d\nu(t_1)\dots d\nu(t_n)dk_1\dots dk_n \\
&\leq \psi(x, y)
\end{aligned} \tag{23}
$$

for all $x, y \in S$. Dividing both sides of inequality (23) by α^n and letting $n \to \infty$ we deduce that $\mathscr{F}$ is a solution of the functional equation (22). Now, showing that the pair $(\mathscr{F}, g)$ satisfies (8). Since μ is k-invariant and $\mathscr{F}$ is a solution of Eq. (7) then, for all $x, y, a \in G$ we can write

$$
\int_K \int_G \int_K \int_G \mathscr{F}(ask' \cdot (xhk \cdot y))d\mu(h)dk'dk = \mathscr{F}(a)\int_K \int_G g(xhk \cdot y)d\mu(h)dk \tag{24}
$$

and on the other hand we rewrite by using the invariance of the Haar measure dk and the fact that $\mathscr{F}$ satisfies the Kannappan type condition (4) that

$$
\begin{aligned}
&\int_K \int_G \int_K \int_G \mathscr{F}(ask' \cdot (xhk \cdot y))d\mu(s)d\mu(h)dkdk' \\
&= \int_K \int_G \int_{K+} \int_G \mathscr{F}(ask' \cdot xhk'k \cdot y))d\mu(s)d\mu(h)dkdk' \\
&\quad + \int_K \int_G \int_{K-} \int_G \mathscr{F}(ask'k \cdot yk' \cdot x))d\mu(s)d\mu(h)dkdk' \\
&= \int_K \int_G \int_{K+} \int_G \mathscr{F}(ask \cdot xk \cdot x))d\mu(s)d\mu(h)dkdk' \\
&\quad + \int_K \int_G \int_{K-} \int_G \mathscr{F}(ask' \cdot xk \cdot y))d\mu(s)d\mu(h)dkdk' \\
&= \int_G \int_K \int_G \int_K \mathscr{F}(ask' \cdot xhk \cdot y)d\mu(s)d\mu(h)dkdk' \\
&= (\int_K \int_G \mathscr{F}(ask' \cdot x)d\mu(s)dk')g(y) \\
&= \mathscr{F}(a)g(x)g(y),
\end{aligned}
$$

which proves (8). This completes the proof.

Hyers–Ulam–Rassias Stability of Eq. (1): Direct Method

In this section we investigate the Hyers–Ulam–Rassias stability of Eq. (1) by using the direct method.

Theorem 3. *Let $f, g : G \to \mathbf{C}$ and $\varphi : G \to \mathbf{R}^+$ be a continuous functions satisfying*

$$\left| \int_K \int_G f(xtk \cdot y) dk d\mu(t) - g(y) f(x) \right| \leq \varphi(y) \tag{25}$$

for all $x, y \in G$ and such that f satisfies the Kannappan type Condition (4) and $G_{\mu,g} \neq \emptyset$. Then there exists a unique continuous function $\mathscr{F}_a : G \to \mathbf{C}$ such that

$$\int_K \int_G \mathscr{F}_a(xtk \cdot y) dk d\mu(t) = \mathscr{F}_a(x) g(y) \tag{26}$$

for all $x, y \in G$,

$$|f(x) - \mathscr{F}_a(x)| \leq \inf_{a \in G_{\mu,f}} \frac{1}{\|\mu\|} \frac{\varphi(a)}{|g(a)| - \|\mu\|} \tag{27}$$

for all $x \in G$ and

$$\left(\int_K \int_G g(xtk \cdot y) dk d\mu(t) - g(x) g(y) \right) \mathscr{F}_a(z) = 0 \tag{28}$$

for all $x, y, z \in G$.

Proof. Substituting y by $a \in S_{g,\mu}$ and dividing both sides of (25) by $\|\mu\|^2$, we get

$$\int_K \int_G F(xtk \cdot y) dk d\nu(t) - \alpha F(x) \leq \frac{\varphi(a)}{\|\mu\|^2},\ x, y \in G, \tag{29}$$

where $F = \frac{f}{\|\mu\|}$, $\nu = \frac{\mu}{|\mu\|}$ and $\alpha = \frac{g(a)}{\|\mu\|} > 1$.

For every x in G and for all $n \in \mathbf{N}^*$, we define a function sequence F_n by

$$F_1(x) = \int_K \int_G F(xtk \cdot a) dk d\nu(t), \quad F_{n+1}(x) = \int_K \int_G F_n(xtk \cdot a) dk d\nu(t),\ n \geq 1. \tag{30}$$

Now, by using the same computations as in the proof of Theorem 2 we prove by induction on n the following inequalities

$$|F_{n+1}(x) - \alpha F_n(x)| \leq \frac{\varphi(a)}{\|\mu\|^2}, \tag{31}$$

$$|F_n(x) - \alpha^n F(x)| \leq \frac{\varphi(a)}{\|\mu\|^2}(1 + |\alpha| + |\alpha|^2 + \ldots + |\alpha|^{n-1}) \tag{32}$$

and

$$|\alpha^{-(n+1)} F_{n+1}(x) - \alpha^{-n} F_n(x)| \leq \frac{1}{\alpha^{n+1}} \frac{\varphi(a)}{\|\mu\|^2}. \tag{33}$$

Define for all $x \in G$

$$\mathscr{F}_a(x) = \lim_{n \to \infty} \alpha^{-n} F_n(x).$$

Using (30) and letting $n \to \infty$ we get

$$\alpha \mathscr{F}_a(x) = \int_K \int_G \mathscr{F}_a(xtk.a) dk d\nu(t), \; x \in G, \tag{34}$$

and from (32) we get (27).

Proving now that $\mathscr{F}_a$ is a solution of the functional equation

$$\int_K \int_G \mathscr{F}_a(xtk \cdot y) dk d\nu(t) = \alpha \mathscr{F}_a(x) \frac{g(y)}{\|\mu\|}, \; x, y \in G. \tag{35}$$

First, make the induction assumption

$$\left| \alpha^{-n} \int_K \int_G F_n(xtk \cdot y) dk d\nu(t) - \alpha^{-n} F_n(x) \frac{g(y)}{\|\mu\|} \right| \leq \frac{\varphi(y)}{\|\mu\|^2} \frac{\|\mu\|^n}{\varphi(a)^n} \tag{36}$$

From (30) and inequality (25) and taking into account the condition (4), we obtain

$$\begin{aligned}
&|\alpha^{-1}| \left| \int_K \int_G F_1(xtk \cdot y) dk d\nu(t) - F_1(x) \frac{g(y)}{\|\mu\|} \right| \\
&= |\alpha^{-1}| \left| \int_K \int_K \int_G \int_G F(xtk \cdot ysk' \cdot a) dk dk' d\nu(t) d\nu(s) - \int_K \int_G F(xtk' \cdot a) dk' d\nu(t) \frac{g(y)}{\|\mu\|} \right| \\
&= |\alpha^{-1}| \left| \int_K \int_K \int_G \int_G F(xtk' \cdot ask \cdot y) dk dk' d\nu(t) d\nu(s) - \int_K \int_G F(xtk' \cdot a) dk' d\nu(t) \frac{g(y)}{\|\mu\|} \right| \\
&\leq \frac{\varphi(y)}{\|\mu\|^2} \frac{\|\mu\|}{\varphi(a)}
\end{aligned}$$

which proves (36) for $n = 1$. Now, We will show that the induction assumption (36) is true with n replaced by $n + 1$.

$$\left|\alpha^{-(n+1)}\int_K\int_G F_{n+1}(xtk\cdot y)dkd\nu(t) - \alpha^{-(n+1)}F_{n+1}(x)\frac{g(y)}{\|\mu\|}\right|$$

$$= |\alpha^{-1}|\left|\int_G\int_K\int_K\int_G \alpha^{-n}F_n(xtk\cdot ysk^{'}\cdot a)dkdk^{'}d\nu(t)d\nu(s)\right.$$

$$\left. - \int_K\int_G \alpha^{-n}F_n(xtk^{'}\cdot a)dk^{'}d\nu(t)\frac{g(y)}{\|\mu\|}\right|$$

$$= |\alpha^{-1}|\left|\int_G\int_K\int_K\int_G \alpha^{-n}F_n(xtk^{'}\cdot ask\cdot y)dkdk^{'}d\nu(t)d\nu(s)\right.$$

$$\left. - \int_K\int_G \alpha^{-n}F_n(xtk^{'}\cdot a)dk^{'}d\nu(t)\frac{g(y)}{\|\mu\|}\right|$$

$$\leq |\alpha^{-1}|\frac{\varphi(y)}{\|\mu\|^2}\left(\frac{\|\mu\|}{\varphi(a)}\right)^n = \frac{\varphi(y)}{\|\mu\|^2}\left(\frac{\|\mu\|}{\varphi(a)}\right)^{n+1},$$

this implies that inequality (36) holds for all n.

Assume now that there exists another solution of Eq. (35), namely $\mathscr{G}_a$ satisfying the inequality (27).

For all $x \in G$ and all $n \in \mathbf{N}$, we have

$$|\mathscr{F}_a(x) - \mathscr{G}_a(x)| = |\alpha|^{-1}$$

$$\left|\int_K\int_G \mathscr{F}_\mu(xt_1k_1\cdot a)dk_1d\nu(t_1) - \int_K\int_G \mathscr{G}_a(xt_1k_1\cdot a)dk_1d\nu(t_1)\right|$$

$$= |\alpha|^{-2}|\int_K\int_G\int_K\int_G \mathscr{F}_a(xt_1k_1\cdot at_2k_2\cdot a)dk_1dk_2d\nu(t_1)d\nu(t_2)$$

$$- \int_K\int_G\int_K\int_G \mathscr{G}_a(xt_1k_1\cdot at_2k_2\cdot a)dk_1dk_2d\nu(t_1)d\nu(t_2)|$$

$$= |\alpha|^{-n}|\int_K\cdots\int_K\int_G\cdots$$

$$\int_G \mathscr{F}_a(xt_1k_1\cdot at_2k_2\cdot a\ldots t_nk_n\cdot a)dk_1dk_2\ldots dk_nd\nu(t_1)d\nu(t_2)\ldots d\nu(t_n)$$

$$- \int_K\cdots\int_K\int_G\cdots$$

$$\int_G \mathscr{G}_\mu(xt_1k_1\cdot at_2k_2\cdot a\ldots t_nk_n\cdot a)dk_1dk_2\ldots dk_nd\nu(t_1)d\nu(t_2)\ldots d\nu(t_n)|$$

$$\leq |\alpha|^{-n}\frac{\varphi(a)}{\|\mu\|}\frac{2}{|g(a)| - \|\mu\|}.$$

Since $|\alpha| > 1$ so, by letting $n \to \infty$, we obtain $\mathscr{F}_a(x) = \mathscr{G}_a(x)$ for all $x \in G$. The rest of the proof is contained in the proof of Theorem 2. This completes the proof of theorem.

References

1. Aoki, T.: On the stability of the linear transformation in Banach spaces. J. Math. Soc. Jpn **2**, 64–66 (1950)
2. Akkouchi, M.: Stability of certain functional equations via fixed point of ĆIRIĆ. Filomat **25**(2), 121–127 (2011)
3. Akkouchi, M., Bouikhalene, B., Elqorachi, E.: Functional equations and μ-spherical functions. Georgian Math. J. **15**(1), 1–20 (2008)
4. Badora. R.: On a joint generalization of Cauchy's and d' Alembert functional equations. Aequationes Math. **43**(1), 72–89 (1992)
5. Baker, J.A.: The stability of certain functional equations. Proc. Am. Math. Soc. **112**(3), 729–732 (1991)
6. Bouikhalene, B., Elqoarchi, E.: Hyers-Ulam stability of spherical functions. Georgian Math. J. (2016). doi:10.1515/gmj-2015-0052
7. Cădariu, L., Radu, V.: Fixed points and the stability of Jensens functional equation. J. Inequal. Pure Appl. Math. **4**(1), Article 4, 15 pp. (2003)
8. Cadariu, L., Radu, V.: On the stability of the Cauchy functional equations: a fixed point approach. Grazer Math. Ber. **346**, 43–52 (2004)
9. Cadariu, L., Moslehian, M.S., Radu, V.: An application of Banach's fixed point theorem to the stability of a general functional equation. An. Univ. Vest Timis. Ser. Mat.-Inform. **47**(3), 21–26 (2009)
10. Chojnacki, W.: On some functional equation generalizing Chauchy's and d' Alembert's functional equations. Colloq. Math. **55**(1), 169–178 (1988)
11. Diaz, J.B., Margolis, B.: A fixed point theorem of the alternative, for contractions on a generalized complete metric space. Bull. Am. Math. Soc. **74**, 305–309 (1968)
12. Elqorachi, E., Akkouchi, M., Bakali, A., Bouikhalene, B,: Badora's equation on non abelian-locally compact group. Georgian Math. J. **11**(1), 449–466 (2004)
13. Forg-Rob, W., Shwaiger, J.: A generalization of the cosine equation to n summands. In: Selected Topics in Functional Equations and Iteration Theory (Graz, 1991). Grazer Mathematische Berichte, pp. 219–226. vol. 316. Karl-Franzens-University of Graz, Graz (1992)
14. Forti, G.L.: Hyers-Ulam stability of functional equations in several variables. Aequationes Math. **50**, 143–190 (1995)
15. Gavruta, P.: A generalization of the Hyers-Ulam-Rassias stability of approximately additive mappings. J. Math. Anal. Appl. **184**, 431–436 (1994)
16. Gavruta, P., Gavruta, L.: A new method for the generalized Hyers-Ulam-Rassias stability. Int. J. Nonlinear Anal. Appl. **1**(2), 11–18 (2010)
17. Hyers, D.H.: On the stability of the linear functional equation. Proc. Nat. Acad. Sci. U.S.A. **27**, 222–224 (1941)
18. Hyers, D.H., Isac, G.I., Rassias, T. M.: Stability of Functional Equations in Several Variables. Birkhauser, Basel (1998)
19. Hyers, D.H., Rassias, T.M.: Approximate homomorphisms. Aequationes Math. Soc. **44**, 125–153 (1992)
20. Jung, S.-M.: Hyers-Ulam-Rassias Stability of Functional Equations in Nonlinear Analysis. Springer, New York (2011)
21. Jung, S.-M., Lee, Z.-H.: A Fixed Point Approach to the Stability of Quadratic Functional Equation with Involution, Fixed Point Theory and Applications. **2008**, 11 pp. (2008). Article ID 732086

22. Kannappan, P.: Functional Equations and Inequalities with Applications. Springer, New York (2009)
23. Manar, Y., Elqoarchi, E., Bouikhalene, B.: Fixed points and Hyers-Ulam-Rassias stability of the quadratic and Jensen functional equations. Nonlinear Funct. Anal. Appl. **15**(4), 647–655 (2010)
24. Radu, V.: The fixed point alternative and the stability of functional equations. Fixed Point Theory **4**(1), 91–96 (2003)
25. Rassias, T.M.: On the stability of functional equations in Banach spaces. J. Math. Anal. Appl. **251**(1), 264–284 (2000)
26. Rassias, T.M.: On the stability of linear mapping in Banach spaces. Proc. Am. Math. Soc. **72**, 297–300 (1978)
27. Rassias, T.M., Brzde, k, J.: Functional Equations in Mathematical Analysis. Springer, New York (2011)
28. Shin'ya, H.: Spherical matrix functions and Banach representability for locally compact motion groups. Jpn. J. Math. (N.S.) **28**(2), 163–201 (2002)
29. Stetkær, H.: Functional equations and spherical functions. Aarhus Univ. (1994) (Preprint Series No.18)
30. Stetkær, H.: D'Alembert's equation and spherical functions. Aequations Math. **48**(2–3), 220–227 (1994)
31. Stetkær, H.: Wilson's functional equations on groups. Aequations Math. **49**(3), 252–275 (1995)
32. Stetkær, H.: Functional equations and matrix-valued spherical functions. Aequations Math. **69**(3), 271–292 (2005)
33. Ulam, S.M.: A Collection of Mathematical Problems. Interscience Publishers, New York (1961). Problems in Modern Mathematics. Wiley, New York (1964)

Approximation Under Exponential Growth Conditions by Szász and Baskakov Type Operators in the Complex Plane

Sorin G. Gal

Abstract In this chapter, firstly for $q > 1$ the exact order near to the best approximation, $\frac{1}{q^n}$, is obtained in approximation by complex q-Favard–Szász–Mirakjan, q-Szász–Kantorovich operators and q-Baskakov operators attached to functions of exponential growth conditions, which are entire functions or analytic functions defined only in compact disks (without to require to be defined on the whole axis $[0, +\infty)$). Quantitative Voronovskaja-type results of approximation order $\frac{1}{q^{2n}}$ are proved. For q-Szász–Kantorovich operators, the case $q = 1$ also is considered, when the exact order of approximation $\frac{1}{n}$ is obtained. Approximation results for a link operator between the Phillips and Favard–Szász–Mirakjan operators are also obtained. Then, by using a sequence $\frac{b_n}{a_n} := \lambda_n > 0, a_n, b_n > 0, n \in \mathbb{N}$ with the property that $\lambda_n \to 0$ as fast we want, we obtain the approximation order $O(\lambda_n)$ for the generalized Szász–Faber operators and the generalized Baskakov–Faber operators attached to analytic functions of exponential growth in a continuum $G \subset \mathbb{C}$. Several concrete examples of continuums G are given for which these generalized operators can explicitly be constructed. Finally, approximation results for complex Baskakov–Szász–Durrmeyer operators are presented.

Keywords Complex q-Favard–Szász–Mirakjan operator • Complex q-Szász–Kantorovich operator • Complex q-Baskakov operator • $q \geq 1$ • Complex Phillips operator • Link operator between Philips and Favard–Szász–Mirakjan operators • Entire functions • Analytic functions in compact disks • Voronovskaja-type result • Compact disk • Exponential growth conditions • Exact order of approximation • Compact disk • Continuum • Faber polynomials • Generalized Szász–Faber and Baskakov–Faber operators • Complex Baskakov–Szász–Durrmeyer operators

Mathematical subject classification: 30E10, 41A25

S.G. Gal (✉)
Department of Mathematics and Computer Science, University of Oradea, Str. Universitatii No. 1, 410087 Oradea, Romania
e-mail: galso@uoradea.ro

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_10

Introduction

The overconvergence phenomenon for positive and linear operators in the complex domain, that is the extension of approximation properties from real domain to complex domain is a topic of increasing interest in the recent years. In this context, the first qualitative kind results were obtained in the papers [4, 44, 45]. Then, in the books [8, 12] quantitative approximation results are presented for several type of approximation operators. For Szász–Mirakjan operator and its Stancu variant in complex domain, we refer the readers to [1, 3, 7, 24, 25, 30, 31, 40]. Also for complex Bernstein–Durrmeyer operators, several papers are available in the literature (see, e.g., [10, 11, 14, 15, 22, 32, 35, 37, 38]), for complex Szász–Durrmeyer operators see [16, 18], for complex Baskakov operators see [21, 39], while for complex q-Balász–Szabados operators see [26].

In this chapter we firstly deal with the approximation properties of the complex q-Favard–Szász–Mirakjan and q-Szász–Kantorovich operators, for $q \geq 1$, under exponential growth conditions on the approximated analytic functions. The results are new and appear for the first time here.

For that purpose, we need some concepts and notations in [27]. For $k \in \mathbb{N} \bigcup \{0\}$, the $q-$integer, $[k]_q$, and the $q-$factorial, $[k]_q!$, are defined by

$$[k]_q := \left\{ \begin{array}{c} \frac{1-q^k}{1-q} \quad \text{if } q > 0,\, q \neq 1, \\ k \ \text{ if } q = 1 \end{array} \right\}$$

and

$$[k]_q! := [1]_q\,[2]_q \ldots [k]_q \ \text{ for } k \in \mathbb{N} \text{ and } \ [0]! = 1,$$

respectively.

For $k \in \{0, 1, \ldots, n\}$, the q-binomial coefficient is given by

$$\begin{bmatrix} n \\ k \end{bmatrix} := \frac{[n]_q!}{[k]_q!\,[n-k]_q!}.$$

For $q > 1$, the q-derivative $D_q f(z)$ of f is defined by

$$D_q f(z) = \left\{ \begin{array}{ll} \frac{f(qz)-f(z)}{(q-1)z}, & z \neq 0, \\ f^{'}(0)\,, & z = 0 \end{array} \right\}.$$

If $|q| > 1$, or $0 < |q| < 1$ and $|z| < \frac{1}{1-q}$, then we define the q-exponential function $E_q\,(z)$ by

$$E_q\,(z) := \sum_{k=0}^{\infty} \frac{z^k}{[k]_q!}.$$

If $|q| > 1$, then it is known that $E_q(z)$ is an entire function and we have

$$E_q(z) = \prod_{j=0}^{\infty} \left(1 + (q-1)\frac{z}{q^{j+1}}\right), \quad |q| > 1.$$

For $f : [0, \infty) \to \mathbb{R}$, the classical Favard–Szász–Mirakjan operators are given by (see [5, 34, 42])

$$S_n(f; x) = e^{-nx} \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) \frac{n^k x^k}{k!}.$$

According to, e.g., [5], if f is of exponential growth, that is $|f(x)| \leq C \exp(Bx)$, for all $x \in [0, \infty)$, with $C, B > 0$, then $S_n(f; x)$ converges to $f(x)$. In addition, concerning quantitative estimates, in, e.g., [43] it is proved that we have $|S_n(f; x) - f(x)| \leq C/n$, for all $x \in [0, \infty)$, $n \in \mathbb{N}$, for some additional hypothesis on f.

If $q \geq 1$, then the complex $q-$Favard–Szász–Mirakjan operators were defined in [30] by

$$S_{n,q}(f; z) = \sum_{k=0}^{\infty} f\left(\frac{[k]_q}{[n]_q}\right) \frac{1}{q^{k(k-1)/2}} \frac{[n]_q^k z^k}{[k]_q!} E_q\left(-[n]_q q^{-k} z\right) \tag{1}$$

and the complex q- Szász–Kantorovich operators were defined in [23] by

$$K_{n,q}(f; z) = \sum_{j=0}^{\infty} E_q\left(-[n]_q q^{-j} z\right) \frac{\left([n]_q z\right)^j}{[j]_q!} \frac{1}{q^{\frac{j(j-1)}{2}}} \int_0^1 f\left(\frac{q[j]_q + t}{[n+1]}\right) dt. \tag{2}$$

Note that for $q = 1$ we have $S_{n,q}(f; z) = S_n(f; z)$.

Without to impose exponential growth conditions to the approximated functions in compact disks, the approximation properties of the complex q-Favard–Szász–Mirakjan operators were studied for $q = 1$ in [9], for $q > 1$ in [30], while the approximation properties of the complex q-Szász–Kantorovich operators were studied for $q = 1$ in [33] and for $q > 1$ in [23].

The absence of exponential growth conditions in compact disks in approximation by the complex Favard–Szász–Mirakjan-type operators, however, requires the boundedness on $[0, +\infty)$ and the analyticity in a disk of double ray of the approximated functions.

Under exponential growth conditions, the approximation properties of the complex q-Favard–Szász–Mirakjan operators were studied in [7] (see also the book [8], pp. 104–113), but only in the case of $q = 1$, while the study of approximation by the complex q-Szász–Kantorovich operators is missing for all $q \geq 1$.

The first main goal of the present work is to study under exponential growth conditions, approximation results for the complex q-Favard–Szász–Mirakjan-type operators in the case of $q > 1$ and for the complex q-Szász–Kantorovich-type operators in the cases of $q \geq 1$. In the case when $q > 1$, the near to the best approximation order, $1/q^n$, is obtained.

The formulas (1) and (2) evidently require that f be defined on $[0, +\infty)$ too, which seems to be not completely in accordance with the fact that the approximation results are obtained only inside the bounded disk of analyticity of f, $\mathbb{D}_R$. This disagreement can easily be avoided, by considering in the approximation of $f(z) = \sum_{k=0}^{\infty} c_k e_k(z)$, $e_k(z) = z^k$, $z \in \mathbb{D}_R$, the complex approximation operators

$$S_{n,q}^*(f;z) = \sum_{k=0}^{\infty} c_k S_{n,q}(e_k;z) \text{ and } K_{n,q}^*(f;z) = \sum_{k=0}^{\infty} c_k K_{n,q}(e_k;z), z \in \mathbb{D}_R, \tag{3}$$

with $S_{n,q}(e_k;z)$ given by (1) and $K_{n,q}(e_k;z)$ given by (2), cases when evidently that the values of f on $[0, +\infty)$ are not required.

To answer completely to this problem, we also point out that if f is entire function, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$ for all $z \in \mathbb{C}$, of some exponential growth, then the equality $L_n(f;z) = \sum_{k=0}^{\infty} c_k L_n(e_k;z)$, $z \in \mathbb{C}$ holds for $L_n = S_{n,q}$, $L_n = K_{n,q}$ and the quantitative estimates are valid (for L_n) in any disk $\mathbb{D}_r$, with $r \in [1, 1/A)$, $A \in (0, 1)$.

Approximation results for a link operator between the Phillips and Favard–Szász–Mirakjan operators are presented.

Also, we point out similar consideration on the q-Baskakov operators and on the Balázs–Szabados operators.

As a second aim of the present chapter, by using a sequence $\frac{b_n}{a_n} := \lambda_n > 0$, $a_n, b_n > 0$, $n \in \mathbb{N}$ with the property that $\lambda_n \to 0$ as fast we want, we obtain the approximation order $O(\lambda_n)$ for the generalized Szász–Faber operators and the generalized Baskakov–Faber operators attached to analytic functions of exponential growth in a continuum $G \subset \mathbb{C}$. Several concrete examples of continuums G are given for which these generalized operators can explicitly be constructed.

The plan of the chapter goes as follows. In section "Complex q-Favard–Szász–Mirakjan Operators, $q > 1$ Case" the case of the complex q-Szász–Kantorovich operators with $q = 1$ is studied. The case of the complex q-Favard–Szász–Mirakjan for $q > 1$ is presented in section "Complex q-Favard–Szász–Mirakjan Operators, $q > 1$ Case" and based on the results in section "Complex q-Favard–Szász–Mirakjan Operators, $q > 1$ Case," the case of the complex q-Szász–Kantorovich operators when $q > 1$ is considered by section "Complex q-Szász–Kantorovich Operators, $q > 1$ Case." In section "Link Operators Between Phillips and Szász–Mirakjan Operators," approximation results for a link operator between the Phillips and Szász–Mirakjan operators are presented. Section "Complex q-Baskakov Operators, $q \geq 1$ and Balázs–Szabados Operators" summarizes similar comments for the complex q-Baskakov operators, $q \geq 1$ and for the complex Balázs–Szabados operators. In sections "Generalized Szász–Faber Operators in Compact Sets" and "Generalized Baskakov–Faber Operators in Compact Sets," the cases of generalized Szász–Faber operators and of the generalized Baskakov–Faber operators attached to a continuum in $\mathbb{C}$ are considered, for which arbitrary orders of approximation, given from the beginning, are obtained. Section "Approximation by Baskakov–Szász–Durrmeyer Operators in Compact Disks" contains approximation results for the complex Baskakov–Szász–Durrmeyer operators.

Complex q-Favard–Szász–Kantorovich Operators, $q = 1$ Case

For $q = 1$, in this section we will consider

$$K_n(f; z) = e^{-nz} \sum_{k=0}^{\infty} \frac{(nz)^k}{k!} \int_0^1 f\left(\frac{k+t}{n}\right) dt.$$

For simplicity, we denote $S_{n,1}(f; z) = S_n(f; z)$.

The first result shows that for entire functions f, we have the equality $K_n(f; z) = \sum_{k=0}^{\infty} c_k K_n(e_k; z)$.

Lemma 1. *Let $f : \mathbb{C} \to \mathbb{C}$ be entire function, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$, for all $z \in \mathbb{C}$. Suppose that there exist $M > 0$ and $A \in (0, 1)$, with the property $|c_k| \leq M\frac{A^k}{k!}$, for all $k = 0, 1, \ldots,$ (which implies $|f(z)| \leq Me^{A|z|}$ for all $z \in \mathbb{C}$) and consider $1 \leq r < 1/A$.*

(i) Denoting $F(z) = \int_0^z f(t)dt$, F is entire function with $F(z) = \sum_{k=1}^{\infty} b_k z^k$, for all $z \in \mathbb{C}$, such that $|b_k| \leq \frac{M}{A} \cdot \frac{A^k}{k!}$, for all $k = 1, \ldots,$ (which implies $|F(z)| \leq \frac{M}{A} e^{A|z|}$ for all $z \in \mathbb{C}$).

(ii) We have $K_n(f; z) = \sum_{k=0}^{\infty} c_k K_n(e_k; z)$, $S_n(F; z) = \sum_{k=1}^{\infty} \frac{c_k}{c_{k-1}} S_n(e_k; z)$, for all $|z| \leq r$ and $K_n(f; z) = S'_n(F)(z)$, for all $z \in \mathbb{C}$ and $n \in \mathbb{N}$. Here $e_k(z) = z^k$.

Proof. (i) We get $F(z) = \int_0^z f(t)dt = \sum_{k=1}^{\infty} \frac{c_{k-1}}{k} z^k$ and denoting $b_k = \frac{c_{k-1}}{k}$ it follows $|b_k| \leq \frac{M}{k} \cdot \frac{A^{k-1}}{(k-1)!} = \frac{M}{A} \cdot \frac{A^k}{k!}$, for all $k \geq 1$.

(ii) Firstly, since by hypothesis we have $|f(t)| \leq Me^{At}$, for all $t \in [0, \infty)$, it follows

$$|K_n(f; z)| \leq M \cdot |e^{-nz}| \cdot \sum_{k=0}^{\infty} \frac{n^k|z|^k}{k!} \cdot \int_0^1 e^{A(k+t)/n} dt$$

$$= \frac{M}{A} \cdot (n+1)[e^{1/n} - 1] \cdot |e^{-nz}| \sum_{k=0}^{\infty} \frac{n^k|z|^k}{k!} \cdot e^{Ak/n} < +\infty,$$

since the last series is convergent by the ratio criterium.

Therefore, $K_n(f; z)$ is well defined for all $z \in \mathbb{C}$. Now, since the series $\sum_{k=0}^{\infty} c_k u^k$ is uniformly convergent for $u \in [j/n, (j+1)/n]$, we get

$$K_n(f;z) = e^{-nz}\sum_{j=0}^{\infty}\frac{(nz)^j}{j!}\cdot\int_0^1\left\{\sum_{k=0}^{\infty}c_k e_k[(j+t)/n]\right\}dt$$
$$= e^{-nz}\sum_{j=0}^{\infty}\frac{(nz)^j}{j!}\sum_{k=0}^{\infty}c_k\int_0^1 e_k[(j+t)/n]dt.$$

Now, if the two sums above would commute, then we would immediately get

$$K_n(f;z) = \sum_{k=0}^{\infty}c_k\cdot e^{-nz}\cdot\sum_{j=0}^{\infty}\frac{(nz)^j}{j!}\int_0^1 e_k[(j+t)/n]dt = \sum_{k=0}^{\infty}c_k\cdot K_n(e_k;z),$$

which would prove the lemma.

It is well known that a sufficient condition for the commutativity of the infinite sums sign, i.e. for $\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}a_{k,j} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}a_{k,j}$, is that $\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}|a_{k,j}| < +\infty$.

Applied to our case, this condition is

$$E := \frac{|e^{-nz}|}{e^{-n|z|}}\cdot\sum_{k=0}^{\infty}|c_k|\cdot e^{-n|z|}\cdot\sum_{j=0}^{\infty}\frac{(n|z|)^j}{j!}\cdot\int_0^1 e_k[(j+t)/n]dt < +\infty,$$

for all $|z| \le r$.

But, since by direct calculating it easily follows $K_n(f;z) = S_n'(F;z)$, we get

$$E = \frac{|e^{-nz}|}{e^{-n|z|}}\cdot\sum_{k=0}^{\infty}|c_k|\cdot K_n(e_k;|z|) \le M\frac{|e^{-nz}|}{e^{-n|z|}}\cdot\sum_{k=0}^{\infty}\frac{A^k}{k!}\cdot K_n(e_k;|z|)$$
$$= M\frac{|e^{-nz}|}{e^{-n|z|}}\cdot\sum_{k=0}^{\infty}\frac{A^k}{k!}\cdot S_n'(e_{k+1}/(k+1);|z|),$$

where $S_n(f;u) = e^{-nu}\cdot\sum_{j=0}^{\infty}\frac{(nu)^j}{j!}\cdot f\left(\frac{j}{n}\right)$, $u \ge 0$, is the classical Szász–Mirakjan operator.

Now, according to Theorem 2 in [28] (see also, e.g., [13], formula (1), replacing there b_n/a_n by $1/n$), we have the formula

$$S_n(f;u) = \sum_{j=0}^{\infty}[0,1/n,2/n,\ldots,j/n;f]u^j, u \ge 0,$$

where $[0,1/n,2/n,\ldots,j/n;f]$ denotes the divided difference of f on the knots $0,1/n,\ldots,j/n$. This implies

$$S_n(e_{k+1};u) = \sum_{j=0}^{k+1}[0,1/n,2/n,\ldots,j/n;e_{k+1}]u^j.$$

It follows

$$S'_n(e_{k+1};u) = \sum_{j=1}^{k+1} j \cdot [0, 1/n, 2/n, \dots, j/n; e_{k+1}]u^{j-1}$$
$$\leq (k+1) \sum_{j=1}^{k+1} [0, 1/n, 2/n, \dots, j/n; e_{k+1}]u^{j-1},$$

which for all $0 \leq u \leq r$ with $r > 1$ implies

$$|S'_n(e_{k+1};u)| \leq (k+1) \cdot r^k \cdot \sum_{j=0}^{k+1} [0, 1/n, 2/n, \dots, j/n; e_{k+1}].$$

Now, taking into account the first formula in the proof of Theorem 3.3 in [13] (written for $b_n/a_n = 1/n$), we get

$$\sum_{j=0}^{k+1} [0, 1/n, 2/n, \dots, j/n; e_{k+1}] \leq 2(k+2)!, \text{ for all } k \in \mathbb{N} \cup \{0\},$$

and therefore

$$|S'_n(e_{k+1};u)| \leq 2(k+1)(k+2)! \cdot r^k, \text{ for all } 0 \leq u \leq r.$$

Summarizing these estimates, for all $|z| \leq r$, it follows

$$E \leq M \frac{|e^{-nz}|}{e^{-n|z|}} \cdot \sum_{k=0}^{\infty} \frac{A^k}{k!} \cdot S'_n(e_{k+1}/(k+1))(|z|) \leq 2M \frac{|e^{-nz}|}{e^{-n|z|}} \cdot \sum_{k=0}^{\infty} \frac{A^k}{k!} \cdot (k+2)! \cdot r^k$$
$$= 2M \frac{|e^{-nz}|}{e^{-n|z|}} \cdot \sum_{k=0}^{\infty} (k+1)(k+2)(Ar)^k < +\infty.$$

This leads to $K_n(f;z) = \sum_{k=0}^{\infty} c_k K_n(e_k;z)$. Similarly we reason for $S_n(F;z)$. □

Remark 1. From the point of view of approximation theory, the hypothesis on f in Lemma 1 to be entire function can be weakened, by considering analytic functions only in a finite disk $\mathbb{D}_R$ and introducing in a natural way the approximation operator $K_n^*(f;z) = \sum_{k=0}^{\infty} c_k \cdot K_n(e_k;z)$. This definition has the advantage that omits the values of f on $[0, +\infty)$.

The first main result is the following.

Theorem 1. *Let* $1 < R < +\infty$ *and suppose that* $f : \mathbb{D}_R \to \mathbb{C}$ *is analytic in* $\mathbb{D}_R$*, i.e.* $f(z) = \sum_{k=0}^{\infty} c_k z^k$*, for all* $z \in \mathbb{D}_R$*, and that there exist* $M > 0$ *and* $A \in (1/R, 1)$*, with* *the property* $|c_k| \le M\frac{A^k}{k!}$*, for all* $k = 0, 1, \ldots,$ *(which implies* $|f(z)| \le Me^{A|z|}$ *for all* $z \in \mathbb{D}_R$*). Also, consider* $1 \le r < 1/A$.

The complex approximation operator

$$K_n^*(f; z) = \sum_{k=0}^{\infty} c_k \cdot K_n(e_k; z), z \in \overline{\mathbb{D}_r},$$

is well defined and if f *is not a polynomial of degree* ≤ 1*, then for all* $n \in \mathbb{N}$ *we have*

$$\|K_n^*(f) - f\|_r \sim \frac{1}{n},$$

where the constants in the equivalence depend only on f*,* r*,* A, M.

Proof. Firstly, the operator $K_n^*(f; z)$ is well defined for all $|z| \le r$ since we have

$$|K_n^*(f; z)| \le \sum_{k=0}^{\infty} |c_k| \cdot |K_n(e_k; z)| \le M \sum_{k=0}^{\infty} \frac{A^k}{k!} \|K_n(e_k)\|_r$$
$$\le 2M \sum_{k=0}^{\infty} \frac{A^k}{k!}(k+1)!r^k = 2M \sum_{k=0}^{\infty}(Ar)^k(k+1) < +\infty.$$

Above we used an analogue formula with that in Lemma 4 in [33],

$$K_n(e_k; z) = \frac{1}{n^k} \cdot \sum_{j=0}^{k} \binom{k}{j} \cdot \frac{n^j}{k-j+1} \cdot S_n(e_j; z)$$

and the inequality in [7], p. 1123, $\|S_n(e_j) - e_j\|_r \le \frac{(j+1)!}{2n} r^{j-1} \le (j+1)!r^j$, which implies $\|S_n(e_j)\|_r \le \|S_n(e_j) - e_j\|_r + r^j \le 2(j+1)!r^j$.

Finally, we can identically apply Corollary 2.2 in [7] (see also Theorem 1.8.3, p. 113 in [8]) and therefore, denoting $S_n^*(f; z) = \sum_{k=0}^{\infty} c_k \cdot S_n(e_k; z)$ and applying Lemma 1 (for the relationship between K_n^* and $[S_n^*]'$), we immediately get $\|K_n^*(f) - f\|_r = \|[S_n^*(F)]' - F'\|_r \sim \frac{1}{n}$. □

The following Voronovskaja's result holds.

Theorem 2. *Let* f, r, M, A *be satisfying the hypothesis in Theorem 1 and* $1 \le r < r_1 < 1/A$*. Then, for all* $|z| \le r$ *and* $n \in \mathbb{N}$*, we have*

$$|K_n^*(f)(z) - f(z) - \frac{z}{2n}f'(z)| \le \frac{3M|z|}{r_1^2 n^2} \sum_{k=0}^{\infty}(k+1)(r_1 A)^k \cdot \frac{r_1}{(r_1-1)^2}.$$

Proof. Denoting $S_n^*(f;z) = \sum_{k=0}^{\infty} c_k \cdot S_n(e;z)$ and replacing in the statement of Theorem 3.2 in [7] (or in the statement of Theorem 1.8.2, (i), p. 107 in [8]) f by F and r by r_1, we get

$$\left|S_n^*(F)(z) - F(z) - \frac{z}{2n}F''(z)\right| \leq \frac{3M|z|}{r_1^2 n^2}\sum_{k=0}^{\infty}(k+1)(r_1 A)^k,$$

for all $n \in \mathbb{N}$ and $|z| \leq r_1$.

Let Γ be the circle of radius $r_1 > r$ and center 0 and denote $E_n(F)(z) = S_n(F)(z) - F(z) - \frac{z}{2n}F''(z)$. Since for any $|z| \leq r$ and $v \in \Gamma$ we have $|v - z| \geq r_1 - r$, by the Cauchy formula, for all $n \in \mathbb{N}$ and $|z| \leq r$ we obtain

$$|E_n'(F)(z)| = \frac{1}{2\pi}\left|\int_{\Gamma} \frac{E_n(F)(v)}{(v-z)^2}dv\right| \leq \frac{3M|z|}{r^2 n^2}\sum_{k=0}^{\infty}(k+1)(r_1 A)^k \cdot \frac{1}{2\pi} \cdot \frac{2\pi r_1}{(r_1 - 1)^2},$$

which by Lemma 1 implies

$$|K_n^*(f)(z) - f(z) - \frac{z}{2n}f'(z)| \leq \frac{3M|z|}{r_1^2 n^2}\sum_{k=0}^{\infty}(k+1)(r_1 A)^k \cdot \frac{r_1}{(r_1 - 1)^2},$$

and proves the theorem. □

Remark 2. The conditions $r < 1/A$ and $A \in (1/R, 1)$ in the above results show that if we choose A as close we want to $1/R$, then r can be chosen as close we want to R, thus evidently improving the results in [33] where the necessary condition $r < R/2$ was required.

Remark 3. From Lemma 1, it is immediate that if f is supposed to be entire function, then the results in Theorems 1 and 2 hold in any disk $\mathbb{D}_r$ with $r \in [1, 1/A)$, $A \in (0, 1)$ and replacing $K_n^*(f)$ with $K_n(f)$.

Complex q-Favard–Szász–Mirakjan Operators, $q > 1$ Case

We begin with quantitative upper estimate in approximation by $S_{n,q}(f;z)$, with $q > 1$. But since the order of approximation is $O(1/q^n)$ and in approximation theory we are always looking for a better degree of approximation, without an essential loss in our considerations we can suppose that we are dealing with approximation by $S_{n,q}(f;z)$ with $q > 2$. This choice for q is made here simply because in this way we can use some relationships from previously published papers.

Theorem 3. *Let $f : \mathbb{D}_R \to \mathbb{C}$, $2 < q < R \leq +\infty$, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$, for all $z \in \mathbb{D}_R$. Suppose that there exist $M > 0$ and $A \in (1/R, 1/q)$, with the property $|c_k| \leq M\frac{A^k}{k!}$, for all $k = 0, 1, \ldots,$ (which implies $|f(z)| \leq Me^{A|z|}$ for all $z \in \mathbb{D}_R$). Also, consider $1 \leq r < 1/(Aq)$.*

(i) *If* $R = +\infty$, $(1/R = 0)$, *i.e.* f *is an entire function, then,* $S_{n,q}(f;z)$ *is entire function and for all* $|z| \le r$, $n \in \mathbb{N}$, *we have* $S_{n,q}(f;z) = \sum_{m=0}^{\infty} c_m S_{n,q}(e_m;z)$ *and*

$$|S_{n,q}(f;z) - f(z)| \le \frac{C_{r,M,A,q}}{[n]_q},$$

with $C_{r,M,A,q} = \frac{M}{2qr}\sum_{m=2}^{\infty}(m+1)(qrA)^m < \infty$.

(ii) *If* $R < +\infty$, *then the complex approximation operator*

$$S^*_{n,q}(f;z) = \sum_{k=0}^{\infty} c_k \cdot S_{n,q}(e_k;z), z \in \overline{\mathbb{D}_r},$$

is well defined and $|S^*_{n,q}(f;z) - f(z)| \le \frac{C_{r,M,A,q}}{[n]_q}$, *with* $C_{r,M,A,q}$ *as at the above point (i).*

Proof. (i) From the hypothesis on f (that is $|f(z)| \le Me^{A|z|}$, for all $z \in \mathbb{C}$), exactly as in the case of $q = 1$ (see, e.g., [4], pp. 1171–1172), it follows that $S_{n,q}(f;z)$ is analytic in $\mathbb{C}$, i. e. for all $z \in \mathbb{C}$, we have $|S_{n,q}(f;z)| < +\infty$.

Since we have

$$S_{n,q}(f;z) = \sum_{k=0}^{\infty}\left[\sum_{j=0}^{\infty} c_j \cdot e_j([k]_q/[n]_q)\right]\frac{1}{q^{k(k-1)/2}}\frac{[n]_q^k z^k}{[k]_q!}E_q\left(-[n]_q q^{-k} z\right),$$

if the above two infinite sums would commute, then we would obtain

$$S_{n,q}(f;z) = \sum_{j=0}^{\infty} c_j \sum_{k=0}^{\infty}\left[e_j([k]_q/[n]_q)\right]\frac{1}{q^{k(k-1)/2}}\frac{[n]_q^k z^k}{[k]_q!}E_q\left(-[n]_q q^{-k} z\right)$$
$$= \sum_{j=0}^{\infty} c_j \cdot S_{n,q}(e_j;z).$$

Reasoning exactly as in the proof of Lemma 1, (ii), it will be good enough if we prove that

$$E :=$$

$$= \frac{E_q(-[n]_q q^{-k} z)}{E_q(-[n]_q q^{-k}|z|} \cdot \sum_{j=0}^{\infty}|c_j| \cdot \sum_{k=0}^{\infty}\left[e_j([k]_q/[n]_q)\right]\frac{1}{q^{k(k-1)/2}}\frac{[n]_q^k |z|^k}{[k]_q!}E_q\left(-[n]_q q^{-k}|z|\right)$$
$$= \frac{E_q(-[n]_q q^{-k} z)}{E_q(-[n]_q q^{-k}|z|)} \cdot \sum_{j=0}^{\infty}|c_j| \cdot S_{n,q}(e_j;|z|) < +\infty.$$

But, from the below proved relationship in this proof, $\|S_{n,q}(e_m;z) - e_m\|_r \leq \frac{(m+1)!}{2[n]_q}(qr)^{m-1}$, for all $m \geq 2$, $n \geq 1$, we immediately get (taking into account that $q > 2$ too)

$$\|S_{n,q}(e_m)\|_r \leq \frac{(m+1)!}{2[n]_q}(qr)^{m-1} + r^m \leq \frac{(m+1)!}{2[n]_q}(qr)^m + (qr)^m$$
$$\leq \frac{(m+1)!}{2}(qr)^m + (qr)^m \leq (m+1)!(qr)^m.$$

Therefore, for all $|z|| \leq r$ we immediately get

$$E \leq M \cdot \frac{E_q(-[n]_q q^{-k}z)}{E_q(-[n]_q q^{-k}|z|} \cdot \sum_{j=0}^{\infty} \frac{A^j}{j!}(j+1)! \cdot (qr)^j$$
$$= M \cdot \frac{E_q(-[n]_q q^{-k}z)}{E_q(-[n]_q q^{-k}|z|} \cdot \sum_{j=0}^{\infty}(j+1) \cdot (qrA)^j < +\infty,$$

since $qrA < 1$.

According to the proof of Theorem 1, p. 1788 in [30], we have the recurrence formula $S_{n,q}(e_m;z) - z^m = \frac{z}{[n]_q}D_q[S_{n,q}(e_{m-1};z)] + z(S_{n,q}(e_{m-1};z) - z^{m-1})$, which immediately leads to

$$S_{n,q}(e_m;z) - z^m = \frac{z}{[n]_q}D_q[S_{n,q}(e_{m-1};z) - z^{m-1}] + z(S_{n,q}(e_{m-});z) - z^{m-1})$$
$$+ \frac{z^{m-1}}{[n]_q} \cdot \frac{q^{m-1}-1}{q-1}.$$

Then, by the proof of Theorem 1, p. 1789 in [30], for any polynomial of degree $\leq m$, we have

$$|D_q(P_m;z)| \leq \frac{m}{r}\|P_m\|_r.$$

Therefore, since $S_{n,q}(e_m;z)$ is a polynomial of degree $\leq m$, for all $|z| \leq r$, $m \geq 2$, we get

$$|S_{n,q}(e_m;z) - e_m(z)| \leq \frac{r}{[n]_q} \cdot \frac{m-1}{r}\|S_{n,q}(e_{m-1};z) - e_{m-1}\|_r$$
$$+ r\|S_{n,q}(e_{m-1};z) - e_{m-1}\|_r + \frac{(qr)^{m-1}}{(q-1)[n]_q}$$
$$\leq \left(r + \frac{m-1}{[n]_q}\right)\|S_{n,q}(e_{m-1};z) - e_{m-1}\|_r + \frac{(qr)^{m-1}(m-1)}{[n]_q}.$$

In what follows, by mathematical induction after $m \geq 2$, we prove the inequality

$$\|S_{n,q}(e_m) - e_m\|_r \leq \frac{(m+1)!}{2[n]_q}(qr)^{m-1},\ m \geq 2, n \geq 1.$$

Indeed, for $m = 2$, the left-hand side becomes (see, e.g., [30], Sect. 2, p. 1786) equal to $\frac{r}{[n]_q}$, while the right-hand side is $\frac{3r}{2[n]_q}$. Supposing now that it is valid for m, the above recurrence implies

$$\begin{aligned}\|S_{n,q}(e_{m+1}) - e_{m+1}\|_r &\leq \left(r + \frac{m}{[n]_q}\right) \cdot \frac{(m+1)!}{2[n]_q}(qr)^{m-1} + \frac{(qr)^m m}{[n]_q} \\ &\leq \left(qr + \frac{m}{[n]_q}\right) \cdot \frac{(m+1)!}{2[n]_q}(qr)^{m-1} + \frac{(qr)^m m}{[n]_q}.\end{aligned}$$

It remains to prove that

$$\left(qr + \frac{m}{[n]_q}\right) \cdot \frac{(m+1)!}{2[n]_q}(qr)^{m-1} + \frac{(qr)^m m}{[n]_q} \leq \frac{(m+2)!}{2[n]_q}(qr)^m,$$

which after simplification is equivalent to

$$\left(qr + \frac{m}{[n]_q}\right)[(m+1)!] + 2(qr)m \leq (m+2)!(qr).$$

Since for $q > 1$, $[n]_q \geq n$, it follows that it suffices if we prove the inequality

$$\left(qr + \frac{m}{n}\right)[(m+1)!] + 2(qr)m \leq (m+2)!(qr).$$

But this is exactly the inequality proved in [7], at p. 1123 (with $qr > 1$ instead of $r > 1$) (see also the book [8], p. 106).

Therefore, we obtain $S_{n,q}(f; z) = \sum_{m=0}^{\infty} c_m S_{n,q}(e_m; z)$, which from the hypothesis on c_m, implies for all $|z| \leq r$, $n \in \mathbb{N}$,

$$\begin{aligned}|S_{n,q}(f; z) - f(z)| &\leq \sum_{m=0}^{\infty} |c_m| \cdot \|S_{n,q}(e_m; z) - e_m(z)\|_r \\ &= \sum_{m=2}^{\infty} |c_m| \cdot \|S_{n,q}(e_m; z) - e_m(z)\|_r \leq M \sum_{k=2}^{\infty} \frac{A^m}{m!} \cdot \frac{(m+1)!}{2[n]_q}(qr)^{m-1} \\ &= \frac{1}{[n]_q} \cdot \frac{M}{2qr} \sum_{k=2}^{\infty}(m+1)(qrA)^m,\end{aligned}$$

where by the hypothesis $1 \leq r < \frac{1}{Aq}$ with $A \in (0, 1/q)$ implies $qrA < 1$ and $\sum_{k=2}^{\infty}(m+1)(qrA)^m < +\infty$.

(ii) The operator $S^*_{n,q}(f)(z)$ is well defined since for all $|z| \leq r$ we have

$$|S^*_{n,q}(f;z)| \leq \sum_{k=0}^{\infty} |c_k| \cdot |S_{n,q}(e_k;z)| \leq M \sum_{k=0}^{\infty} \frac{A^k}{k!} \cdot (k+1)!(qr)^k$$
$$= 2M \sum_{k=0}^{\infty}(k+1)(qrA)^k < +\infty.$$

The rest of the proof is identical with the proof of (i). □

The next result gives the Voronovskaja's result in compact disks.

Theorem 4. *Suppose that f, M, A, R, r, q, $S_{n,q}(f)$, $S^*_{n,q}(f)$ are as in the statement of Theorem 3. Then, for all $|z| \leq r$ and $n \in \mathbb{N}$, we have*

$$\left| S_{n,q}(f;z) - f(z) - \frac{1}{[n]_q} \cdot L_q(f;z) \right| \leq \frac{C_{r,M,A,q}}{[n]_q^2}, \text{ if } R = +\infty,$$
$$\left| S^*_{n,q}(f;z) - f(z) - \frac{1}{[n]_q} \cdot L_q(f;z) \right| \leq \frac{C_{r,M,A,q}}{[n]_q^2}, \text{ if } R < +\infty,$$

with $C_{r,M,A,q} = \frac{M}{2}\sum_{m=3}^{\infty}(m+1)(m+2)(qrA)^m < \infty$, $f(z) = \sum_{m=0}^{\infty} c_m z^m$ and

$$L_q(f;z) = \frac{D_q f(z) - f'(z)}{q-1} = \sum_{m=2}^{\infty} c_m([1]_q + \ldots + [m-1]_q)z^{m-1}.$$

Proof. Let $R = +\infty$. Denoting $E_{m,n}(z) = S_{n,q}(e_m;z) - e_m(z) - \frac{[1]_q + \ldots + [m-1]_q}{[n]_q} z^{m-1}$, by the proof of Theorem 2, p. 1789 in [30], for all $m \geq 2$, $z \in \mathbb{C}$ and $n \in \mathbb{N}$ we have the recurrence formula

$$E_{m,n}(z) = \frac{z}{[n]_q} D_q(S_{n,q}(e_{m-1};z) - e_{m-1}(z)) + zE_{m-1,n}(z).$$

Taking into account the estimate for $\|S_{n,q}(e_{m-1};z) - e_{m-1}(z)\|_r$ in the proof of the above Theorem 3, (i), for all $|z| \leq r$, $m \geq 2$, $n \in \mathbb{N}$, it follows

$$|E_{m,n}(z)| \leq \frac{z}{[n]_q} \|(S_{n,q}(e_{m-1}) - e_{m-1})'\|_r + |z| \cdot |E_{m-1,n}(z)|$$
$$\leq |z| \cdot |E_{m-1,n}(z)| + \frac{|z|}{[n]_q} \cdot \frac{m-1}{r} \|S_{n,q}(e_{m-1}) - e_{m-1}\|_r$$

$$\leq |z| \cdot |E_{m-1,n}(z)| + \frac{|z|(m-1)}{[n]_q} \cdot \frac{m!}{2[n]_q}(qr)^{m-1}$$

$$\leq |z| \cdot |E_{m-1,n}(z)| + \frac{|z|}{2[n]_q^2} \cdot [(m+1)!](qr)^{m-1}$$

$$\leq (qr) \cdot |E_{m-1,n}(z)| + \frac{1}{2[n]_q^2}](m+1)!(qr)^m.$$

Taking into account that $E_{2,n}(z) = 0$, step by step by mathematical induction we easily obtain, for all $m \geq 3$,

$$|E_{m,n}(z)| \leq \frac{(qr)^m}{2[n]_q^2}\left(\sum_{k=4}^{m+1} k!\right) \leq \frac{(qr)^m}{2[n]_q^2}(m+2)(m+1)! = \frac{m!(qr)^m(m+1)(m+2)}{2[n]_q^2}.$$

This implies, together with Theorem 3, (i), that for all $|z| \leq r$, $n \in \mathbb{N}$, we have

$$\left|S_{n,q}(f;z) - f(z) - \frac{1}{[n]_q} \cdot L_q(f;z)\right| \leq \sum_{m=1}^{\infty} |c_m| \cdot |E_{m,n}(z)|$$

$$= \sum_{m=3} |c_m| \cdot |S_{n,q}(e_m;z) - e_m(z) - \frac{[1]_q + \ldots + [m-1]_q}{[n]_q} z^{m-1}|$$

$$\leq \frac{M}{2[n]_q^2} \cdot \sum_{m=3}^{\infty} \frac{A^m}{m!} \cdot \frac{m!(qr)^m(m+1)(m+2)}{2[n]_q^2} = \frac{1}{[n]_q^2} \cdot \frac{M}{2} \sum_{m=3}^{\infty} (m+1)(m+2)(qrA)^m.$$

The proof in the case of $S^*_{n,q}$, i.e. when $R < +\infty$, follows identically the above lines in the case when $R = +\infty$. □

Corollary 1. *Suppose that f, M, A, R, r, q, $S_{n,q}(f)$, $S^*_{n,q}(f)$ are as in the statement of Theorem 3. If f is not a polynomial of degree ≤ 1, then*

$$\|S_{n,q}(f) - f\|_r \sim \frac{1}{[n]_q} \text{ if } R = +\infty, \; \|S^*_{n,q}(f) - f\|_r \sim \frac{1}{[n]_q} \text{ if } R < +\infty,$$

with the constants in the equivalences are independent of n.

Proof. Let $R = +\infty$. Since

$$S_{n,q}(f;z) - f(z) = \frac{1}{[n]_q}\left\{L_q(f;z) + [n]_q\left(S_{n,q}(f;z) - f(z) - \frac{1}{[n]_q} \cdot L_q(f;z)\right)\right\},$$

by the inequality $\|F + G\|_r \geq | \, \|F\|_r - \|G\|_r \, | \geq \|F\|_r - \|G\|_r$ we get

$$\|S_{n,q}(f;z) - f(z)\|_r \geq \frac{1}{[n]_q} \left\{ \|L_q(f)\|_r - [n]_q \|S_{n,q}(f) - f - \frac{1}{[n]_q} \cdot L_q(f)\|_r \right\}.$$

Since by the proof of Theorem 3 in [30] we have $\|L_q(f)\|_r > 0$ and taking into account that by Theorem 4 we get

$$[n]_q \left| S_{n,q}(f;z) - f(z) - \frac{1}{[n]_q} \cdot L_q(f;z) \right| \leq \frac{C_{r,M,A,q}}{[n]_q} \to 0 \text{ for } n \to \infty,$$

there exists n_1 (depending only on f and r) such that for all $n \geq n_1$ we have

$$\|L_q(f)\|_r - [n]_q \|S_{n,q}(f) - f - \frac{1}{[n]_q} \cdot L_q(f)\|_r \geq \frac{1}{2} \|L_q(f)\|_r.$$

This immediately implies that $\|S_{n,q}(f;z) - f(z)\|_r \geq \frac{1}{2[n]_q} \cdot \|L_q(f)\|_r$, for all $n \geq n_1$.

For $1 \leq n \leq n_1 - 1$ we have $\|S_{n,q}(f) - f\|_r = \frac{1}{[n]_q} \cdot M_{r,n}(f) > 0$, where $M_{r,n}(f) = [n]_q \|S_{n,q}(f) - f\|_r$, which finally implies

$$\|S_{n,q}(f;z) - f(z)\|_r \geq \frac{C_{r,q}(f)}{[n]_q}, \text{ for all } n \in \mathbb{N},$$

where $C_{r,q}(f) = \min \left\{ \frac{1}{2} \|L_q(f)\|_r, M_{r,1}(f), \ldots, M_{r,n_1-1}(f) \right\}$.

Combining this lower estimate with the upper estimate in Theorem 3, (i), the proof is finished.

The proof in the case of $S^*_{n,q}(f)$, i.e. when $R < +\infty$, is identical. □

Remark 4. If $R < +\infty$, then the conditions $r < \frac{1}{Aq}$ and $A \in (1/R, 1/q)$ in the above results show that if we choose A as close we want to $1/R$, then r can be chosen as close we want to $\frac{R}{q}$, thus evidently improving the results in [30] where the necessary condition $r < \frac{R}{2q}$ is required.

Remark 5. By the obvious inequalities $\frac{q-1}{q^n} \leq \frac{1}{[n]_q} \leq \frac{q}{q^n}$, for all $n \in \mathbb{N}$ and $q > 1$, it follows that the approximation order in Theorem 3 and Corollary 1 is geometrical, namely $\frac{1}{q^n}$.

Complex q-Szász–Kantorovich Operators, $q > 1$ Case

Firstly, we need the following auxiliary result.

Lemma 2. *Let $q > 2$, $1 \leq r$. For all $m, n \in \mathbb{N}$, we have*

$$\|K_{n,q}(e_m)\|_r \leq (m+1)! \, (qr)^m.$$

Proof. Since by Lemma 5 in [23], we have

$$K_{n,q}\left(e_m;z\right)=\frac{1}{[n+1]_q^m}\sum_{k=0}^{m}\binom{m}{k}\frac{q^k\,[n]_q^k}{(m-k+1)}S_{n,q}(e_k;z),$$

by the inequality $\|S_{n,q}(e_m)\|_r \le (m+1)!(qr)^m$ obtained in the proof of Theorem 3, (i), it follows, for all $|z|\le r$, $m\in\mathbb{N}\bigcup\{0\}$, $n\in\mathbb{N}$,

$$\left|K_{n,q}\left(e_m;z\right)\right|\le\frac{(m+1)!\,(qr)^m}{[n+1]_q^m}\sum_{j=0}^{m}\binom{m}{j}q^j\,[n]_q^j=(m+1)!\,(qr)^m\,,$$

which proves the lemma. □

The first main result is the following quantitative upper estimate.

Theorem 5. *Let $f:\mathbb{D}_R\to\mathbb{C}$, $2<q<R\le+\infty$, i.e. $f(z)=\sum_{k=0}^{\infty}c_kz^k$, for all $z\in\mathbb{D}_R$. Suppose that there exist $M>0$ and $A\in(1/R,1/q)$, with the property $|c_k|\le M\frac{A^k}{k!}$, for all $k=0,1,\dots$, (which implies $|f(z)|\le Me^{A|z|}$ for all $z\in\mathbb{D}_R$). Also, consider $1\le r<1/(Aq)$.*

(i_a) If $R=+\infty$, $(1/R=0)$, i.e. f is an entire function, then $K_{n,q}(f;z)$ is entire function and for all $|z|\le r$, $n\in\mathbb{N}$, we have $K_{n,q}(f;z)=\sum_{m=0}^{\infty}c_mK_{n,q}(e_m;z)$ and

$$|K_{n,q}(f;z)-f(z)|\le\frac{C_{r,M,A,q}}{[n]_q},$$

with $C_{r,M,A,q}=3M\sum_{m=1}^{\infty}m(m+1)(qrA)^m<\infty$.

(i_b) If $R<+\infty$, then the complex approximation operator

$$K^*_{n,q}(f;z)=\sum_{k=0}^{\infty}c_k\cdot K_{n,q}(e_k;z),\,z\in\overline{\mathbb{D}_r},$$

is well defined and $|K^*_{n,q}(f;z)-f(z)|\le\frac{C_{r,M,A,q}}{[n]_q}$, with $C_{r,M,A,q}$ as at the above point (i_a).

Proof. (i_a) Let $R=+\infty$. By the proof of Theorem 1 in [23], relationship (5), we have the relationship

$$K_{n,q}\left(e_m;z\right)-e_m\left(z\right)=\frac{z}{[n]}D_q\left(K_{n,q}\left(e_{m-1};z\right)\right)+z\left(K_{n,q}\left(e_{m-1};z\right)-e_{m-1}\left(z\right)\right)$$
$$+\frac{1}{[n+1]_q^m}\sum_{k=0}^{m}\binom{m}{k}\frac{q^k\,[n]_q^k}{m-k+1}\left\{1-\frac{k}{m}-\frac{k}{mq\,[n]_q}\right\}S_{n,q}\left(e_k;z\right),$$

where reasoning exactly as in the proof of Theorem 1 in [23] and taking into account the estimate for $\|S_{n,q}(e_k\|_r$ in the proof of Theorem 3, (i), for all $|z| \leq r$ we get

$$\left| \frac{1}{[n+1]_q^m} \sum_{k=0}^{m} \binom{m}{k} \frac{q^k [n]_q^k}{m-k+1} \left\{ 1 - \frac{k}{m} - \frac{k}{mq[n]_q} \right\} S_{n,q}(e_k; z) \right|$$
$$\leq \frac{2m+1}{[n+1]_q} \cdot (m+1)! (qr)^m .$$

Now, by using the Bernstein's inequality and reasoning exactly as in the proof of Theorem 1 in [23], we immediately obtain

$$|K_{n,q}(e_m; z) - e_m(z)| \leq r|K_{n,q}(e_{m-1}; z) - e_{m-1}(z)| + 3m \cdot \frac{(m+1)!(qr)^m}{[n]_q},$$

for $|z| \leq r$, $m, n \in \mathbb{N}$.

By writing the last inequality step by step for $m = 1, 2, \ldots$, we easily arrive at

$$|K_{n,q}(e_m; z) - e_m(z)| \leq \frac{3}{[n]_q}(qr)^m \sum_{k=1}^{m} k(k+1)! \leq \frac{3}{[n]_q} m(m+1)m!(qr)^m.$$

Since exactly in the proof of Theorem 3, (i), we can prove that $K_{n,q}(f; z) = \sum_{m=0}^{\infty} c_m K_{n,q}(e_m; z)$, for all $|z| \leq r$, it follows

$$\left|K_{n,q}(f; z) - f(z)\right| \leq \sum_{m=0}^{\infty} |c_m| \left|K_{n,q}(e_m; z) - e_m(z)\right| \leq \frac{3M}{[n]_q} \sum_{m=1}^{\infty} m(m+1)(qrA)^m,$$

which proves (i_a).

(i_b) Let $R < +\infty$. The operator $K^*_{n,q}(f; z)$ is well defined since by Lemma 2, for all $|z| \leq r$ we get

$$|K^*_{n,q}(f; z)| = \sum_{k=0}^{\infty} |c_k| \cdot |K_{n,q}(e_k; z)| \leq M \sum_{k=0}^{\infty} \frac{A^k}{k!} \cdot (k+1)!(qr)^k = M \cdot \sum_{k=0}^{\infty} (k+1)(qrA)^k.$$

The proof in the case of $K^*_{n,q}$, i.e. when $R < +\infty$, follows identically the above lines in the case (i_a) when $R = +\infty$. □

The next result gives the Voronovskaja's result in compact disks.

Theorem 6. *Suppose that f, M, R, r, A, q, $K_{n,q}(f)$, $K^*_{n,q}(f)$ are as in the statement of Theorem 5 and, in addition, that $2 < q < q^2 < R$. If $1 \leq r < \frac{1}{Aq^2}$, then for all $|z| \leq r$ and $n \in \mathbb{N}$, we have*

$$\left|K_{n,q}(f;z)-f(z)-\frac{1}{[n]_q}\cdot R_q(f;z)\right|\le\sum_{m=2}^{\infty}P_4(m)(q^2rA)^m\cdot\frac{C_{r,M,A,q}}{[n]_q^2},\ if\ R=+\infty,$$

$$\left|K^*_{n,q}(f;z)-f(z)-\frac{1}{[n]_q}\cdot R_q(f;z)\right|\le\sum_{m=2}^{\infty}P_4(m)(q^2rA)^m\cdot\frac{C_{r,M,A,q}}{[n]_q^2},\ if\ R<\infty,$$

with a constant $C_{r,M,A,q}>0$ independent of n, $P_4(m)$ a strictly positive polynomial of degree ≤ 4 in $m\in\mathbb{N}$, $R_q(f;z)=\sum_{m=1}^{\infty}c_m\cdot V_{m,q}(z)$, $V_{m,q}(z)=q([1]_q+\ldots+[m-1]_q)+\frac{mz^{m-1}}{2}(1-2z)$ and $\sum_{m=2}^{\infty}P_4(m)(q^2rA)^m<\infty$.

Proof. Let $R=+\infty$. Denote $E_{n,m}(z)=K_{n,q}(e_m;z)-e_m(z)-\frac{1}{[n]_q}V_{m,q}(z)$. By using the estimate for $|K_{n,q}(e_m;z)-e_m(z)|$ in the proof of Theorem 5, (i), the estimate for $|S_{n,q}(e_m;z)-e_m(z)|$ in the proof of Theorem 3, (i), and the estimate for $|K_{n,q}(e_m;z)|$ in Lemma 2, by following the lines in the proof of Theorem 2, (i) in [23], we easily arrive at an inequality of the form

$$|E_{n,m}(z)|\le r|E_{n-1,m}(z)|+\frac{P_3(m)}{[n]_q^2}(m+1)!\cdot(q^2r)^m,$$

where $P_3(m)$ is a strictly positive polynomial of degree 3 in $m\in\mathbb{N}$ ($P_3(m)$ could be explicitly obtained by calculation, but for simplicity we omit it here).

From this recurrence we get, step by step

$$|E_{n,m}(z)|\le\frac{C\cdot P_4(m)m!}{[n]_q^2}\cdot\sum_{m=2}^{\infty}(q^2rA)^m,$$

(here $P_4(m)$ is a polynomial of degree ≤ 4 in m and $C>0$ is a constant independent of n) which by $K_{n,q}(f;z)=\sum_{m=0}^{\infty}c_mK_{n,q}(e_m;z)$ immediately leads to

$$\left|K_{n,q}(f;z)-f(z)-\frac{1}{[n]_q}R_q(f;z)\right|\le\sum_{m=1}^{\infty}|c_m|\cdot|E_{n,m}(z)|\le\frac{C_{r,M,A,q}}{[n]_q^2}\sum_{m=2}^{\infty}P_4(m)(q^2rA)^m.$$

The proof in the case when $R<+\infty$ follows exactly the same lines as above. □

Corollary 2. *Suppose that f,R,r,M,A,q, $K_{n,q}(f)$, $K^*_n(f)$ are as in the statement of Theorem 5 and, in addition, that $2<q<q^2<R$. If $1\le r<\frac{1}{Aq^2}$ and f is not a constant function, then*

$$\|K_{n,q}(f)-f\|_r\sim\frac{1}{[n]_q}\ if\ R=\infty,\ \ \|K^*_{n,q}(f)-f\|_r\sim\frac{1}{[n]_q}\ if\ R<\infty.$$

Proof. Let $R=+\infty$. Reasoning exactly as in the proof of Theorem 2, (ii) in [23], by the above Theorem 6 we obtain that $\lim_{n\to\infty}[n]_q|K_{n,q}(f;z)-f(z)|=R_q(f;z)$.

Then, following word for word the proof of Theorem 3 in [23], we immediately arrive at the desired result.

The case when $R < +\infty$ is identical. □

Remark 6. The conditions $r < \frac{1}{Aq^2}$ and $A \in (1/R, 1)$ in the above results show that if we choose A as close we want to $1/R$, then r can be chosen as close we want to $\frac{R}{q^2}$, thus evidently improving the results in Theorem 3 in [23] where the necessary condition $r < \frac{R}{2q^2}$ is required.

Link Operators Between Phillips and Szász–Mirakjan Operators

Let $\alpha, \rho > 0$. The complex form of the link operator between the Phillips operator and the Szász–Mirakjan operator is formally defined by

$$P_\alpha^\rho(f, z) = \sum_{k=1}^{\infty} s_{\alpha,k}(z) \int_0^\infty \theta_{\alpha,k}^\rho(t) f(t) dt + e^{-\alpha z} f(0),\ z \in \mathbb{C},$$

where

$$s_{\alpha,k}(z) = e^{-\alpha z} \frac{(\alpha z)^k}{k!}, \theta_{\alpha,k}^\rho(t) = \frac{\alpha\rho}{\Gamma(k\rho)} e^{-\alpha\rho t} (\alpha\rho t)^{k\rho - 1}.$$

For $\rho = 1$, it is clear that it reduces to the Phillips operator [36] and for $\rho \to \infty$ one obtains the Szász–Mirakjan operator.

In the present section we present quantitative estimates in closed disks $\overline{\mathbb{D}_r} = \{z \in \mathbb{C}; |z| \leq r\}$, for the uniform convergence as $\alpha \to \infty$, of the complex operator $P_\alpha^\rho(f, z)$ attached to f entire function and of some exponential growth in $\mathbb{C}$.

Throughout the section we denote $\|f\|_r = \max\{|f(z)|; |z| \leq r\}$ and $e_k(z) = z^k$, $k \in \mathbb{N} \cup \{0\}$, $z \in \mathbb{C}$.

For $M > 0$ and $A > 0$, let us consider the following class of entire functions :

$$\mathscr{F}_{M,A} = \left\{ f : f(z) = \sum_{k=0}^{\infty} c_k z^k, z \in \mathbb{C}, |c_k| \leq M \frac{A^k}{k!},\ \text{for all } k \in \mathbb{N} \cup \{0\} \right\}.$$

Notice that for $f \in \mathscr{F}_{M,A}$, we have $|f(z)| \leq M e^{A|z|}$, for all $z \in \mathbb{C}$.

The following main results hold.

Theorem 7 (Upper Estimate, See [17]). *Let $\alpha, \rho > 0$, $E(\rho) := 1 + \frac{1}{\rho}$, $A \in (0, 1)$ and $f \in \mathscr{F}_{M,A}$. In this case, $P_\alpha^\rho(f)(z)$ is well defined for all $z \in \mathbb{C}$ and if $1 \leq r < \frac{1}{A}$, then for all $|z| \leq r$ and $\alpha \geq \max\{E(\rho), \frac{2A}{\rho}\}$, we have*

$$|P_\alpha^\rho(f)(z) - f(z)| \le \frac{C_{r,\rho,A,M}}{\alpha},$$

where $C_{r,\rho,A,M} = \frac{M \cdot E(\rho)}{r} \sum_{k=2}^{\infty} (k+1)(rA)^k < \infty$.

Theorem 8 (Voronovskaja Formula, See [17]). *Let* $\alpha, \rho > 0$, $E(\rho) := 1 + \frac{1}{\rho}$ *and* $f \in \mathscr{F}_{M,A}$ *with* $A \in \left(0, \frac{\rho}{1+\rho}\right)$.

If $1 \le r < r + 1/\rho < \frac{1}{A}$ *then for all* $|z| \le r$ *and* $\alpha, \rho > 0$ *with* $\alpha \ge \max\{2, E(\rho), \frac{2A}{\rho}\}$, *we have*

$$\left| P_\alpha^\rho(f)(z) - f(z) - \frac{(1+1/\rho)z}{2\alpha} f''(z) \right| \le \frac{C^{(1)}_{r,A,\rho,M}(f)}{\alpha^2},$$

where

$$C^{(1)}_{r,A,\rho,M}(f) = \frac{2M \cdot E^2(\rho)}{(r+1/\rho)^2} \cdot \sum_{k=2}^{\infty} (k+1)(k+2)(A(r+1/\rho))^k + \frac{6M \cdot E(\rho)}{r(1-Ar)^2 \cdot ln^2(1/Ar)}.$$

Theorem 9 (Exact Order, See [17]). *Under the hypothesis of Theorem 8, if* f *is not a polynomial of degree* ≤ 1 *then for all* $\alpha, \rho > 0$ *with* $\alpha \ge \max\{2, E(\rho), \frac{2A}{\rho}\}$, *we have*

$$\|P_\alpha^\rho(f) - f\|_r \sim \frac{C}{\alpha},$$

where the constant C *in the equivalence depends only on* f, r, *and* ρ.

If in our results we take $\rho = 1$, then we get the approximation order $\frac{1}{\alpha}$ for the complex Phillips operators and if we take $\rho \to \infty$, since $E(\rho) \to 1$ and $\frac{2A}{\rho} \to 0$, then we get for the Szász–Mirakjan operator the same order of approximation $\frac{1}{\alpha}$.

Complex q-Baskakov Operators, $q \ge 1$ and Balázs–Szabados Operators

For $0 \le \alpha \le \beta$ and $q \ge 1$, the complex q-Baskakov–Stancu operators

$$W_{n,q}^{\alpha,\beta}(f)(z) = \sum_{j=0}^{\infty} \frac{[n+j-1]_q!}{[n-1]_q!} q^{-j(j-1)/2}$$

$$\cdot [[\alpha]_q/([n]_q + [\beta]_q), \ldots, (q^{j-1}[\alpha] + [j]_q)/(q^{j-1}([n]_q + [\beta]_q)); f] \cdot \frac{z^j}{([n]_q + [\beta]_q)^j},$$

were introduced and studied in [39] for $q > 1$ and in [21] for $q = 1$, under the hypothesis that $f(z) = \sum_{k=0}^{\infty} c_k zk$ is analytic and of exponential growth in a disk $\mathbb{D}_R$ and with all derivatives bounded on $[0, \infty)$.

As in the previous sections, for the study of these operators, we can consider two approaches:

1. under the hypothesis that f is defined, analytic and of exponential growth only in the finite disk $\mathbb{D}_R$, one can consider the complex approximation operators

$$W_{n,q}^{(\alpha,\beta)*}(f)(z) = \sum_{j=0}^{\infty} c_k W_{n,q}^{\alpha,\beta}(e_k)(z),$$

recapturing for $W_{n,q}^{(\alpha,\beta)*}$ the order of approximation $1/[n]_q$, if $q > 1$.

2. under the stronger hypothesis that f is entire and of exponential growth in $\mathbb{C}$, one can recapture the same orders of approximation for $W_{n,q}^{\alpha,\beta}(f)$.

Also, the complex q-Balász–Szabados operators $R_{n,q}(f)(z)$, $0 < q \leq 1$, were introduced and studied for $q \geq 1$ in [26] and for $q = 1$ in [8], pp. 139–148. Defining

$$R_{n,q}^{*}(f)(z) = \sum_{j=0}^{\infty} c_k R_{n,q}(e_k)(z),$$

similar approaches can be applied in this case too.

We omit here the details.

Remark 7. Since from the exponential growth condition on f imposed to the coefficients c_k in the finite disk $\mathbb{D}_R$, we easily get that $|\sum_{k=0}^{\infty} c_k z^k| \leq \sum_{k=0}^{\infty} |c_k| \cdot r^k \leq M \sum_{k=0}^{\infty} \frac{(Ar)k}{k!} = Me^{Ar} < +\infty$, for all $r \geq 1$, it follows that in fact f can be analytically extended in the whole complex plane. Therefore, under the exponential growth conditions in the results of all sections "Complex q-Favard–Szász–Kantorovich Operators, $q = 1$ Case," "Complex q-Favard–Szász–Mirakjan Operators, $q > 1$ Case," "Complex q-Szász–Kantorovich Operators, $q > 1$ Case," and "Link Operators Between Phillips and Szász–Mirakjan Operators," the operators denoted with $*$ are in fact the restrictions of the initial operators to the finite disks $\mathbb{D}_R$.

Generalized Szász–Faber Operators in Compact Sets

Let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, be two sequences with the properties that $\frac{b_n}{a_n} \leq 1$, for all $n \in \mathbb{N}$ and $\frac{b_n}{a_n} \searrow 0$ as fast as we want. In [3], a remarkable generalization of the Favard–Szász–Mirakjan operators defined by

$$S_n(f;a_n;b_n)(z) = e^{-\frac{a_n}{b_n}z}\sum_{j=0}^{\infty} f\left(\frac{jb_n}{a_n}\right)\cdot\frac{(a_n z)^j}{j!b_n^j},\ z\in\mathbb{C}, n\in\mathbb{N}$$

was introduced, by proving that it approximates analytic functions f of some exponential growth in compact disks, with the approximation order $\frac{b_n}{a_n}$.

Also, according to Theorem 1 in [3], one can write

$$S_n(f;a_n;b_n)(z) = \sum_{j=0}^{\infty}\left[0, \frac{b_n}{a_n}, \frac{2b_n}{a_n}, \ldots, \frac{jb_n}{a_n}; f\right] z^j, \tag{4}$$

where $\left[0, \frac{b_n}{a_n}, \frac{2b_n}{a_n}, \ldots, \frac{jb_n}{a_n}; f\right]$ denotes the divided difference of f on the knots $0, \ldots, \frac{jb_n}{a_n}$.

Suggested by the Open Problem 1.11.8, pp. 152–153 in [8] and by Cetin and Ispir [3], the aim of this section is to generalize this result to the approximation by the so-called Favard–Szász–Mirakjan–Faber operators, attached to analytic functions of exponential growth in a continuum in $\mathbb{C}$. An upper estimate of order $\frac{b_n}{a_n}$ in approximation by these operators is obtained.

In what follows, let us briefly recall some basic concepts on Faber polynomials and Faber expansions, which will be useful in the next section.

Throughout the paper, $G \subset \mathbb{C}$ will be considered a compact set such that $\tilde{\mathbb{C}} \setminus G$ is connected. Let $A(G)$ be the Banach space of all functions that are continuous on G and analytic in the interior of G endowed with the uniform norm $\|f\|_G = \sup\{|f(z)|; z \in G\}$. If we denote $\mathbb{D}_r = \{z \in \mathbb{C}; |z| < r\}$, then according to the Riemann Mapping Theorem, a unique conformal mapping Ψ of $\tilde{\mathbb{C}} \setminus \overline{\mathbb{D}}_1$ onto $\tilde{\mathbb{C}} \setminus G$ exists so that $\Psi(\infty) = \infty$ and $\Psi'(\infty) > 0$. The n-th *Faber polynomial* $F_n(z)$ attached to G may be defined by $\frac{\Psi'(w)}{\Psi(w)-z} = \sum_{n=0}^{\infty}\frac{F_n(z)}{w^{n+1}},\ z \in G, |w| > 1$. Then $F_n(z)$ is a polynomial of exact degree n.

If $f \in A(G)$, then

$$a_n(f) = \frac{1}{2\pi i}\int_{|u|=1}\frac{f(\Psi(u))}{u^{n+1}}du = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\Psi(e^{it}))e^{-int}dt, n \in \mathbb{N}\cup\{0\} \tag{5}$$

are called the Faber coefficients of f and $\sum_{n=0}^{\infty} a_n(f)F_n(z)$ is called the Faber expansion (series) attached to f on G. The Faber series represents a natural generalization of the Taylor series, when the unit disk is replaced by an arbitrary simply connected domain bounded by a “nice” curve.

For further properties of Faber polynomials and Faber expansions, see, e.g., Gaier [6] and Suetin [41].

Now, let G be a continuum (that is a connected compact subset of $\mathbb{C}$) and suppose that f is analytic on G, that is, there exists $R > 1$ such that f is analytic in G_R, given by $f(z) = \sum_{k=0}^{\infty} a_k(f)F_k(z),\ z \in G_R$. Here recall that G_R denotes the interior of the closed level curve Γ_R given by $\Gamma_R = \{\Psi(w); |w| = R\}$ ($G \subset \overline{G_r}$ for all $1 < r < R$).

Suggested by the Open Problem 1.11.8, pp. 152–153 in [8] and by (4), the Favard–Szász–Mirakjan–Faber operators attached to G and f might be defined by the formula

$$M_n(f;a_n,b_n,G;z)=\sum_{j=0}^{\infty}\left[0,\frac{b_n}{a_n},\frac{2b_n}{a_n},\ldots,\frac{jb_n}{a_n};F(f)\right]\cdot F_j(z),\ z\in G,\ n\in\mathbb{N},\quad (6)$$

where $F_j(z)$ is the Faber polynomial of degree j attached to G in the Faber's expansion of f and $F(f)$ is defined by

$$F(f)(w)=\frac{1}{2\pi i}\int_{|u|=1}\frac{f(\Psi(u))}{u-w}du=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f(\Psi(e^{it}))e^{it}}{e^{it}-w}dt, w\in\mathbb{D}_1\bigcup\mathbb{C}\setminus\mathbb{D}_1. \quad (7)$$

Here $\left[0,\frac{b_n}{a_n},\frac{2b_n}{a_n},\ldots,\frac{jb_n}{a_n};F(f)\right]$ denotes the divided difference of $F(f)$ on the knots $0,\frac{b_n}{a_n},\ldots,\frac{jb_n}{a_n}$ and we suppose that there exist the constants $B,C>0$ such that $|F(f)(x)|\le Ce^{Bx}$, for all $x\in[0,+\infty)$.

Under the above hypothesis, $M_n(f;a_n,b_n,G;z)$ is well defined for all $n\in\mathbb{N}$ and $z\in G_r$. Indeed, by the formula

$$\begin{aligned}M_n(f;a_n,b_n,G;z)&=\sum_{j=0}^{\infty}\frac{\Delta^j_{b_n/a_n}}{j!(b_n/a_n)^j}\cdot F_j(z)\\&=\sum_{j=0}^{\infty}\frac{1}{j!}\cdot\left(\frac{a_n}{b_n}\right)^j\cdot F_j(z)\cdot\sum_{p=0}^{j}\binom{j}{p}F(f)[(j-p)b_n/a_n]\end{aligned}$$

and by the well-known estimates $|F_j(z)|\le C(r)r^j$, valid for all $z\in\overline{G_r}$, $j\ge 0$ (see, e.g., inequality (8), p. 43 in Suetin [41]), it is easily seen that

$$\begin{aligned}|M_n(f;a_n,b_n,G;z)|&\le C\cdot C(r)\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{a_n}{b_n}\right)^j\cdot\sum_{p=0}^{j}\binom{j}{p}\cdot e^{B(j-p)b_n/a_n}\cdot r^j\\&=C\cdot C(r)\sum_{j=0}^{\infty}\frac{1}{j!}\cdot\left(\frac{a_n}{b_n}\right)^j\cdot\left(r+e^{Bb_n/a_n}\right)^j\\&=C\cdot C(r)\cdot e^{(r+e^{Bb_n/a_n})\cdot a_n/b_n}<+\infty.\end{aligned}$$

Notice that according to Plemejl–Sokhoski formula, if, for example, $f\circ\Psi$ is α-Hölder function on $\{u\in\mathbb{C};|u|=1\}$ with $\alpha\in(0,1)$, then it is known that $F(f)$ given by (7) is continuous and bounded on $[0,\infty)$, which obviously implies that $F(f)$ is of exponential growth on $[0,\infty)$.

Since from the proofs of all quantitative results in approximation by complex operators it follows that a necessary key result is $M_n(f; a_n, b_n, G; z) = \sum_{k=0}^{\infty} a_k(f) \cdot M_n(F_k; a_n, b_n, G; z)$, instead of defining $M_n(f; a_n, b_n, G; z)$ as in (6), for simplicity throughout the paper we will consider the following.

Definition 1 ([13]). Let G be a continuum (that is a connected compact subset of $\mathbb{C}$) and suppose that f is analytic on G, that is, there exists $R > 1$ such that $f(z) = \sum_{k=0}^{\infty} a_k(f)F_k(z)$ for all $z \in G_R$.

The Favard–Szász–Mirakjan–Faber operators attached to G and f will be defined by

$$M_n(f; a_n, b_n, G; z) = \sum_{k=0}^{\infty} a_k(f) \cdot M_n(F_k; a_n, b_n, G; z),$$

(supposing that the above series is convergent)), where making use of

$$F(F_k)(w) = \frac{1}{2\pi i} \int_{|u|=1} \frac{F_k(\Psi(u))}{u - w} du = w^k := e_k(w), \text{ for all } |w| < 1,$$

(see [6], p. 48, first relation before relation (6.17)), it formally becomes

$$M_n(f; a_n, b_n, G; z) = \sum_{k=0}^{\infty} a_k(f) \cdot \sum_{j=0}^{k} \left[0, \frac{b_n}{a_n}, \ldots, \frac{jb_n}{a_n}; e_k\right] \cdot F_j(z). \tag{8}$$

Remark 8. Formula (8) presents the advantage that the pretty strong hypothesis on $F(f)$ in formula (6) can be replaced by direct simple requirements on f, as we will see in the next Theorem 7.

Remark 9. For $G = \overline{\mathbb{D}}_1$, since $F_j(z) = z^j$ and $F(f)(z) = f(z)$, by Theorem 1 in [3], the above Favard–Szász–Mirakjan–Faber operators reduce to the complex Favard–Szász–Mirakjan operators introduced and studied in [3].

For the proof of the main result, we need two lemmas.

Lemma 3 ([13]). *Let $a_n, b_n > 0$, $n \in \mathbb{N}$, with $1 \geq \frac{b_n}{a_n} \searrow 0$. For all $k, n \in \mathbb{N}$ with $k \leq [a_n/b_n]$ (here $[a]$ denotes the integer part of a) we have the inequality*

$$E_{k,n} := \sum_{j=0}^{k-1} \left[0, \frac{b_n}{a_n}, \ldots, \frac{jb_n}{a_n}; e_k\right] \leq \frac{b_n}{a_n} \cdot \frac{(k+1)!}{2}.$$

Lemma 4 ([13]). *Let $a_n, b_n > 0$, $n \in \mathbb{N}$, with $1 \geq \frac{b_n}{a_n} \searrow 0$. Denoting $m(n) = [a_n/b_n]$, for all $n \in \mathbb{N}$ and $k \geq m(n) + 1$ we have*

$$\sum_{j=0}^{k}\left[0,\frac{b_n}{a_n},\ldots,\frac{jb_n}{a_n};e_k\right]\leq (k+1)!.$$

Note that this lemma is an immediate consequence of the Lemma 6 in the next section "Generalized Baskakov–Faber Operators in Compact Sets."

We are now in a position to state the main result of this section.

Theorem 10 ([13]). *Let f be analytic on the continuum G, that is, there exists $R > 1$ such that f is analytic in G_R, given by $f(z) = \sum_{k=0}^{\infty} a_k(f)F_k(z)$, $z \in G_R$. Also, suppose that there exist $M > 0$ and $A \in \left(\frac{1}{R}, 1\right)$, with $|a_k(f)| \leq M\frac{A^k}{k!}$, for all $k = 0, 1, \ldots,$ (which implies $|f(z)| \leq C(r)Me^{Ar}$ for all $z \in G_r$, $1 < r < R$). Here G_R denotes the interior of the closed level curve Γ_R given by $\Gamma_R = \{\Psi(w); |w| = R\}$ and $G \subset \overline{G_r}$ for all $1 < r < R$.*

Let $1 < r < \frac{1}{A}$ be arbitrary fixed. Then, there exist an index $n_0 \in \mathbb{N}$ and a constant $C(r,f) > 0$ depending on r and f only, such that for all $z \in \overline{G_r}$ and $n \geq n_0$ we have

$$|M_n(f;a_n,b_n,G;z) - f(z)| \leq C(r,f)\cdot\frac{b_n}{a_n}.$$

Remark 10. If G is a disk centered at the origin and of radius R, then M_n in Theorem 10 becomes the operator considered in [3] and we recapture the upper estimate obtained there.

Remark 11. Theorem 10 holds under the more general hypothesis $|a_k(f)| \leq P_m(k)\cdot\frac{A^k}{k!}$, for all $k \geq 0$, where P_m is an algebraic polynomial of degree m, satisfying $P_m(k) > 0$ for all $k \geq 0$. Indeed, since $\sum_{k=0}^{\infty}(k+1)P_m(k)(Ar)^k < +\infty$ and $\sum_{k=0}^{\infty} P_m(k+1)\cdot\frac{(Ar)^k}{k!} < +\infty$, this is immediate from the proof of Theorem 10.

Remark 12. There are many concrete examples for G when the conformal mapping Ψ and the Faber polynomials associated with G, and consequently when the Favard–Szász–Mirakjan–Faber operators too, can explicitly be written (for details, see, e.g., [12], pp. 81–83):

(i) G is the continuum bounded by the m-cusped hypocycloid H_m ($m = 2, 3, \ldots,$), given by the parametric equation $z = e^{i\theta} + \frac{1}{m-1}e^{-(m-1)i\theta}$, $\theta \in [0, 2\pi)$, case when $\Psi(w) = w + \frac{1}{(m-1)w^{m-1}}$ and the Faber polynomials can explicitly be calculated ;

(ii) G is the regular m-star ($m = 2, 3, \ldots,$) given by

$$S_m = \{x\omega^k; 0 \leq x \leq 4^{1/m}, k = 0, 1, \ldots, m-1,\ \omega^m = 1\},$$

case when $\Psi(w) = w\left(1 + \frac{1}{w^m}\right)^{2/m}$ and the Faber polynomials can explicitly be calculated;

(iii) G is the m-leafed symmetric lemniscate, $m = 2, 3, \ldots$, with its boundary given by

$$L_m = \{z \in \mathbb{C}; |z^m - 1| = 1\},$$

case when $\Psi(w) = w\left(1 + \frac{1}{w^m}\right)^{1/m}$ and the Faber polynomials can explicitly be calculated;

(iv) G is the semidisk $SD = \{z \in \mathbb{C}; |z| \leq 1$ and $|Arg(z)| \leq \pi/2\}$, case when $\Psi(w) = \frac{2(w^3-1)+3(w^2-w)+2(w^2+w+1)^{3/2}}{w(w+1)\sqrt{3}}$ and the attached Faber polynomials can explicitly be calculated ;

(v) G is a circular lune or G is an annulus sector, cases when again the conformal mapping Ψ and the Faber polynomials can explicitly be calculated.

Generalized Baskakov–Faber Operators in Compact Sets

For x real and ≥ 0, the original formula of the classical now Baskakov operator is given by (see [2])

$$Z_n(f)(x) = (1+x)^{-n} \sum_{k=0}^{\infty} \binom{n+k-1}{k} \left(\frac{x}{1+x}\right)^k f(k/n)$$

and many approximation results of this operators were published.

According to [29], Theorem 2, under the same hypothesis on f that $Z_n(f)(x)$ is well defined and denoting by $[0, 1/n, \ldots, j/n; f]$ the divided difference of f on the knots $0, \ldots, j/n$, for $x \geq 0$ we can write

$$Z_n(f)(x) = W_n(f)(x) = \sum_{j=0}^{\infty} \left(1 + \frac{1}{n}\right) \cdot \ldots \cdot \left(1 + \frac{j-1}{n}\right) \cdot [0, 1/n, \ldots, j/n; f] x^j, x \geq 0, \tag{9}$$

(where for $j = 0$ and $j = 1$ we take $(1 + 1/n) \cdot \ldots \cdot (1 + (j-1)/n) = 1$). But as it was remarked in [8], p. 124, if $|x| < 1$ is not positive then $W_n(f)(x)$ and $Z_n(f)(x)$ do not necessarily coincide and because of this reason in Sect. 1.8 of the book [8], pp. 124–138, they were studied separately, under different hypothesis on f and $z \in \mathbb{C}$.

For analytic functions satisfying some exponential-type growth condition, quantitative estimates of order $O\left(\frac{1}{n}\right)$ in approximation by $W_n(f)(z)$ in compact disks with center at origin were obtained in [8], Sect. 1.9, pp. 124–138. For $f(z) = \sum_{k=0}^{\infty} a_k z^k$, all the quantitative results are based on the formula $W_n(f)(z) = \sum_{k=0}^{\infty} a_k \cdot W_n(e_k)(z)$, with $e_k(z) = z^k$, i.e. by using (9) too,

$$W_n(f)(z) = \sum_{k=0}^{\infty} a_k \cdot \sum_{j=0}^{k} \left(1 + \frac{1}{n}\right) \cdot \ldots \cdot \left(1 + \frac{j-1}{n}\right) \cdot [0, 1/n, \ldots, j/n; e_k] z^j. \tag{10}$$

By using a sequence of real positive numbers, $(\lambda_n)_{n\in\mathbb{N}}$, with the properties that $\lambda_n \to 0$ as fast we want, suggested by the formula (10) too, the aim of this section is to generalize the approximation by the operators $W_n(f)(z)$, to the approximation by the so-called by us generalized Baskakov–Faber operators attached to analytic functions of some exponential growth in a continuum in $\mathbb{C}$, obtaining the approximation order $O(\lambda_n)$.

Since $\lambda_n \to 0$, obviously that without to lose the generality, everywhere in the paper we may suppose that $0 < \lambda_n \le \frac{1}{2}$, for all $n \in \mathbb{N}$.

In what follows, let us consider the basic concepts on Faber polynomials and Faber expansions presented in the previous section "Complex q-Baskakov Operators, $q \ge 1$ and Balázs–Szabados Operators."

Suggested by the formula (10), we can introduce the following.

Definition 2 ([20]). The generalized Baskakov–Faber operators attached to G and f is (formally) defined by $W_n(f;\lambda_n,G;z) = \sum_{k=0}^{\infty} a_k(f)\cdot W_n(e_k;\lambda_n,G;z)$, i.e.,

$$W_n(f;\lambda_n,G;z)=\sum_{k=0}^{\infty} a_k(f)\cdot\sum_{j=0}^{k}(1+\lambda_n)\cdot\ldots\cdot(1+(j-1)\lambda_n)\cdot[0,\lambda_n,\ldots,j\lambda_n;e_k]\cdot F_j(z), \tag{11}$$

where for $j=0$ and $j=1$, by convention we take $(1+\lambda_n)\cdot\ldots\cdot(1+(j-1)\lambda_n)=1$.

Remark 13. For $\lambda_n = 1/n$, $n \in \mathbb{N}$ and $G = \overline{\mathbb{D}}_1$, since $F_j(z) = z^j$, the above generalized Baskakov–Faber operators reduce to the complex Baskakov operators of the form given by (10), introduced and studied in [8], Sect. 1.9.

For the proof of the main result, we need two lemmas, as follows.

Lemma 5 ([20]). *Let $0 < \lambda_n \le \frac{1}{2} < 1$, $n \in \mathbb{N}$, be with $\lambda_n \to 0$. For all $k, n \in \mathbb{N}$ with $k \le [1/\lambda_n]$ (here $[a]$ denotes the integer part of a) we have the inequality*

$$E_{k,n} := \sum_{j=0}^{k-1}(1+\lambda_n)\cdot\ldots\cdot(1+(j-1)\lambda_n)\cdot[0,\lambda_n,\ldots,j\lambda_n;e_k] \le \lambda_n\cdot(k+3)!.$$

Here, by convention, for $j=0$ and $j=1$ we take $(1+\lambda_n)\cdot\ldots\cdot(1+(j-1)\lambda_n)=1$.

Lemma 6 ([20]). *Let $0 < \lambda_n \le \frac{1}{2}$, $n \in \mathbb{N}$, be with $\lambda_n \to 0$. For all $k \ge 0$ and $n \in \mathbb{N}$, we have*

$$G_{k,n} := \sum_{j=0}^{k}(1+\lambda_n)\cdot\ldots\cdot(1+(j-1)\lambda_n)\cdot[0,\lambda_n,\ldots,j\lambda_n;e_k] \le (k+1)!.$$

The main result of this section is the following.

Theorem 11 ([20]). *Let f be analytic on the continuum G, that is there exists $R > 1$ such that f is analytic in G_R, given by $f(z) = \sum_{k=0}^{\infty} a_k(f)F_k(z)$, $z \in G_R$. Also,* *suppose that there exist* $M > 0$ *and* $A \in \left(\frac{1}{R}, 1\right)$*, with* $|a_k(f)| \leq M\frac{A^k}{k!}$*, for all* $k = 0, 1, \ldots,$ *(which implies* $|f(z)| \leq C(r)Me^{Ar}$ *for all* $z \in G_r$*,* $1 < r < R$*).*

Let $1 < r < \frac{1}{A}$ *be arbitrary fixed. Then, there exist an index* $n_0 \in \mathbb{N}$ *and a constant* $C(r,f) > 0$ *depending on r and f only, such that for all* $z \in \overline{G_r}$ *and* $n \geq n_0$ *we have*

$$|W_n(f; \lambda_n, G; z) - f(z)| \leq C(r,f) \cdot \lambda_n.$$

Remark 14. Theorem 11 generalizes Theorem 1.9.1, p. 126 in [8], in two senses: firstly, it is extended from compact disks with center at origin to compact sets and secondly, the order of approximation $O\left(\frac{1}{n}\right)$ is essentially improved to the order $O(\lambda_n)$, with $\lambda_n \to 0$ as fast we want.

Remark 15. It is clear that Theorem 11 holds under the more general hypothesis $|a_k(f)| \leq P_m(k) \cdot \frac{A^k}{k!}$, for all $k \geq 0$, where P_m is an algebraic polynomial of degree m with $P_m(k) > 0$ for all $k \geq 0$.

Remark 16. There are many concrete examples for G when the conformal mapping Ψ and the Faber polynomials associated with G, and consequently when the Baskakov–Faber operators too, can explicitly be written, see the Remark 12, (i)–(v), from the end of the previous section "Generalized Szász–Faber Operators in Compact Sets."

Approximation by Baskakov–Szász–Durrmeyer Operators in Compact Disks

In the present section, we study the rate of approximation of analytic functions in a disk $\mathbb{D}_R = \{z \in \mathbb{C}; |z| < R\}$, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$, of exponential growth, and the Voronovskaja type result, for a natural derivation from the complex operator $L_n(f)(z)$ formally defined as operator of complex variable by

$$L_n(f)(z) := n\sum_{v=1}^{\infty} b_{n,v}(z)\int_0^{\infty} s_{n,v-1}(t)f(t)dt + (1+z)^{-n}f(0), z \in \mathbb{C}, \quad (12)$$

where

$$b_{n,k}(z) = \binom{n+k-1}{k}\frac{z^k}{(1+z)^{n+k}}, s_{n,k}(t) = e^{-nt}\frac{(nt)^k}{k!}.$$

An important relationship used for the quantitative results in approximation of an analytic function f by the complex operator $L_n(f)$ would be $L_n(f) = \sum_{k=0}^{\infty} c_k L_n(e_k)$, but which requires some additional hypothesis on f (because the definition of $L_n(f)(z)$ involves the values of f on $[0, +\infty)$ too) and implies restrictions on the domain of convergence. This situation can naturally be avoided, by defining directly the approximation complex operator

$$L_n^*(f)(z) = \sum_{k=0}^{\infty} c_k \cdot L_n(e_k)(z),$$

whose definition evidently omits the values of f outside of its disk of analyticity.

In this section we deal with the approximation properties of the complex operator $L_n^*(f)(z)$.

It is worth noting here that if instead of the above defined $L_n(f)(z)$ we consider any other Szász-type or Baskakov-type complex operator, then for $L_n^*(f)(z)$ defined as above, all the quantitative estimates in, e.g., [1, 3, 9, 13, 16, 24, 25, 30] hold true identically, without to need the additional hypothesis on the values of f on $[0,\infty)$ imposed there.

Our first main result is the following theorem for upper bound.

Theorem 12 ([19]). *For $f : \mathbb{D}_R \to \mathbb{C}$, $1 < R < +\infty$, analytic on $\mathbb{D}_R$, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$, for all $z \in \mathbb{D}_R$, suppose that there exist $M > 0$ and $A \in (\frac{1}{R}, 1)$, with the property that $|c_k| \le M\frac{A^k}{k!}$, for all $k = 0, 1, \ldots,$ (which implies $|f(z)| \le Me^{A|z|}$ for all $z \in \mathbb{D}_R$).*

(i) If $1 \le r < \frac{1}{A}$, then for all $|z| \le r$ and $n \in \mathbb{N}$ with $n > r+2$, $L_n^*(f)(z)$ is well defined and we have

$$|L_n^*(f)(z) - f(z)| \le \frac{C_{r,A,M}}{n},$$

where $C_{r,A,M} = \frac{M(r+2)}{r} \cdot \sum_{k=2}^{\infty}(k+1)\cdot(rA)^k < \infty$;

(ii) If $1 \le r < r_1 < \frac{1}{A}$, then for all $|z| \le r$ and $n, p \in \mathbb{N}$ with $n > r+2$, we have

$$|[L_n^*(f)]^{(p)}(z) - f^{(p)}(z)| \le \frac{p! r_1 C_{r_1,A,M}}{n(r_1 - r)^{p+1}},$$

where $C_{r_1,A,M}$ is given as at the above point (i).

The following Voronovskaja type result holds.

Theorem 13 ([19]). *For $f : \mathbb{D}_R \to \mathbb{C}$, $2 < R < +\infty$, analytic on $\mathbb{D}_R$, i.e. $f(z) = \sum_{k=0}^{\infty} c_k z^k$, for all $z \in \mathbb{D}_R$, suppose that there exist $M > 0$ and $A \in (\frac{1}{R}, 1)$, with the property that $|c_k| \le M\frac{A^k}{k!}$, for all $k = 0, 1, \ldots,$.*

If $1 \le r < r+1 < \frac{1}{A}$, then for all $|z| \le r$ and $n \in \mathbb{N}$ with $n > r+2$, we have

$$\left| L_n^*(f)(z) - f(z) - \frac{z(z+2)}{2n} f''(z) \right| \le \frac{C_{r,A,M}(f)}{n^2},$$

where $C_{r,A,M}(f) = M\sum_{k=2}^{\infty}\frac{k-1}{k!}[A(r+1)]^k B_{k,r} + \frac{4M(r+2)}{r}\cdot\frac{1}{ln^2(1/\rho)}\cdot\left(\frac{1}{(1-Ar)^2}+\frac{4}{1-Ar}\right) < \infty$ *and*

$$B_{k,r} = (k-1)^2(k-2)r^2 + 2(k-1)(k-2)(2k-3)(r+1) + (r+1)(r+2)\cdot(k+1)!.$$

The following exact order of approximation can be obtained.

Theorem 14 ([19]). *Suppose that the hypothesis in the previous theorem hold.*

(i) If f is not a polynomial of degree ≤ 1, then for all $n > r+2$ we have

$$\|L_n^*(f) - f\|_r \sim \frac{1}{n}.$$

where the constants in the equivalence depend only on f, R, A, and r.

(ii) If $1 \leq r < r_1 < r_1 + 1 < 1/A$ and f is not a polynomial of degree $\leq p-1$, $(p \geq 1)$, then

$$\|[L_n^*(f)]^{(p)} - f^{(p)}\|_r \sim \frac{1}{n}, \text{ for all } n > r+2,$$

where the constants in the equivalence depend only on f, R, A, r, r_1, and p.

References

1. Aydin, D.: On complex q-Szász-Mirakjan operators. Commun. Fac. Sci. Univ. Ankara. Sér. A1 Math. Stat. **61**(2), 51–66 (2012)
2. Baskakov, V.A.: An example of a sequence of linear positive operators in the space of continuous functions (Russian). Dokl. Akad. Nauk SSSR. **113**, 249–251 (1957)
3. Cetin, N., Ispir, N.: Approximation by complex modified Szász-Mirakjan operators. Stud. Sci. Math. Hung. **50**(3), 355–372 (2013)
4. Dressel, F.G., Gergen, J.J., Purcell, W.H.: Convergence of extended Bernstein polynomials in the complex plane. Pac. J. Math. **13**(4), 1171–1180 (1963)
5. Favard, J.: Sur les multiplicateurs d'interpolation. J. Math. Pures Appl. **23**, 219–247 (1944)
6. Gaier, D.: Lectures on Complex Approximation. Birkhauser, Boston (1987)
7. Gal, S.G.: Approximation and geometric properties of complex Favard-Szász-Mirakjan operators in compact disks. Comput. Math. Appl. **56**, 1121–1127 (2008)
8. Gal, S.G.: Approximation by Complex Bernstein and Convolution Type Operators. World Scientific, Singapore/Hong Kong/London/NJ (2009)
9. Gal, S.G.: Approximation of analytic functions without exponential growth conditions by complex Favard-Szász-Mirakjan operators. Rend. Circ. Mat. Palermo. **59**, 367–376 (2010)
10. Gal, S.G.: Approximation by complex Bernstein-Durrmeyer polynomials with Jacobi weights in compact disks. Mathematica Balkanica (N.S.) **24**(1–2), 103–119 (2010)
11. Gal, S.G.: Approximation by complex genuine Durrmeyer type polynomials in compact disks. Appl. Math. Comput. **217**, 1913–1920 (2010)
12. Gal, S.G.: Overconvergence in Complex Approximation. Springer, New York (2013)
13. Gal, S.G.: Approximation of analytic functions by generalized Favard-Szász-Mirakjan-Faber operators in compact sets. Compl. Anal. Oper. Theory. **9**(5), 975–985 (2015)

14. Gal, S.G., Gupta, V.: Approximation by a Durrmeyer-type operator in compact disks. Annali dell'Universita di Ferrara **57**(2), 261–274 (2011)
15. Gal, S.G., Gupta, V.: Quantitative estimates for a new complex Durrmeyer operator in compact disks. Appl. Math. Comput. **218**(6), 2944–2951 (2011)
16. Gal, S.G., Gupta, V.: Approximation by complex Szász-Durrmeyer operators in compact disks. Acta Math. Sci. Ser. B. **34**(4), 1157–1165 (2014)
17. Gal, S.G., Gupta, V.: Approximation by the complex form of a link operator between the Phillips and the Szász-Mirakjan operators. Results Math. **67**(3-4), 381–393 (2015)
18. Gal, S.G., Gupta, V.: Approximation by complex Durrmeyer type operators in compact disks. In: Pardalos, P.M., Rassias, T.M. (eds.) Mathematics Without Boundaries: Surveys in Interdisciplinary Research, pp. 263–284. Springer, New York/Heidelberg/Dordrecht/London (2014)
19. Gal, S.G., Gupta, V.: Approximation by complex Baskakov-Szász-Durrmeyer operators in compact disks. Anal. Theory Appl. **31**(2), 207–220 (2015)
20. Gal, S.G., Opris, B.D.: Approximation of analytic functions with an arbitrary order by generalized Baskakov-Faber operators in compact sets. Compl. Anal. Oper. Theory. **10**(2), 369–377 (2016)
21. Gal, S.G., Gupta, V., Verma, D.K., Agrawal, P.N.: Approximation by complex Baskakov-Stancu operators in compact disks. Rendiconti Circ. Mat. Palermo **61**(2), 153–165 (2012)
22. Gal, S.G., Gupta, V., Mahmudov, N.I.: Approximation by a complex q-Durrmeyer type operator. Ann. Univ. Ferrara **58**(1), 65–87 (2012)
23. Gal, S.G., Mahmudov, N.I., Kara, M.: Approximation by complex q-Szász-Kantorovich operators in compact disks. $q > 1$. Compl. Anal. oper. Theory **7**(6), 1853–1867 (2013)
24. Gupta, V., Verma, D.K.: Approximation by complex Favard-Szász-Mirakjan-Stancu operators in compact disks. Math. Sci. **6**(25) (2012). doi:10.1186/2251-7456-6-25
25. Ispir, N.: Approximation by modified complex Szász-Mirakjan operators. Azerb. J. Math. **3**(2), 95–107 (2013)
26. Ispir, N., Özkan, E.Y.: Approximation properties of complex q-Balázs-Szabados operators in compact disks. J. Inequal. Appl. **2013**(361), 12pp. (2013)
27. Kac, V., Cheung, P.: Quantum Calculus. Springer, New York (2002)
28. Lupaş, A.: Some properties of the linear positive operators. I. Mathematica (Cluj) **9**(32), 77–83 (1967)
29. Lupaş, A.: Some properties of the linear positive operators, II. Mathematica (Cluj) **9**(32), 295–298 (1967)
30. Mahmudov, N.I.: Approximation properties of complex q-Szász-Mirakjan operators in compact disks. Comput. Math. Appl. **60**(6), 1784–1791 (2010)
31. Mahmudov, N.I.: Convergence properties and iterations for q-Stancu polynomials in compact disks. Comput. Math. Appl. **59**(12), 3763–3769 (2010)
32. Mahmudov, N.I.: Approximation by Bernstein-Durrmeyer-type operators in compact disks. Appl. Math. Lett. **24**(7), 1231–1238 (2011)
33. Mahmudov, N.I., Kara, M.: Approximation theorems for complex Szász-Kantorovich operators. J. Comput. Anal. Appl. **15**, 32–38 (2013)
34. Mirakjan, G.M.: Approximation des fonctiones continues au moyen de polynomes de la forme $e^{-nx}\sum_{k=0}^{n} C_{k,n}x^k$ (French). Dokl. Akad. Nauk SSSR **31**, 201–205 (1941)
35. Ostrovska, S.: q-Bernstein polynomials and their iterates. J. Approx. Theory **123**(2), 232–255 (2003)
36. Phillips, R.S.: An inversion formula and semi-groups of linear operators. Ann. Math. **59**, 325–356 (1954)
37. Ren, M.-Y., Zeng, X.-M.: Approximation by complex q-Bernstein-Schurer operators in compact disks. Georgian Math. J. **20**(2), 377–395 (2013)
38. Ren, M.-Y., Zeng, X.-M., Zeng, L.: Approximation by complex Durrmeyer-Stancu type operators in compact disks. J. Inequal. Appl. **2013**(442) (2013), p. 13
39. Soylemez, D., Aydin, D.: Approximation by a complex q-Baskakov-Stancu operator in compact disks. J. Inequal. Appl. **2014**(249) (2014), p. 15

40. Sucu, S., Ibikli, E.: Approximation by Jakimovski-Leviatan type operators on a complex domain. Complex Anal. Oper. Theory **8**(1), 177–188 (2014)
41. Suetin, P.K.: Series of Faber Polynomials. Gordon and Breach, Amsterdam (1998)
42. Szász, O.: Generalizations of S. Bernstein's polynomial to the infinite interval. J. Res. Natl. Bur. Stand. **45**, 239–245 (1950)
43. Totik, V.: Uniform approximation by positive operators on infinite intervals. Anal. Math. **10**, 163–182 (1984)
44. Wood, B.: On a generalized Bernstein polynomial of Jakimovski and Leviatan. Math. Zeitschr. **106**, 170–174 (1968)
45. Wright, E.M.: The Bernstein approximation polynomials in the complex plane. J. Lond. Math. Soc. **5**, 265–269 (1930)

On the Asymptotic Behavior of Sequences of Positive Linear Approximation Operators

Ioan Gavrea and Mircea Ivan

Abstract We provide an analysis of the rate of convergence of positive linear approximation operators defined on $C[0,1]$. We obtain a sufficient condition for a sequence of positive linear approximation operators to possess a Mamedov-type property and give an application to the Durrmeyer approximation process.

Keywords Asymptotic expansion • Bernstein operators • Central moments • Durrmeyer operstors • Least concave majorant • Mamedov property • Positive linear operators • Rate of convergence • Voronovsakja type formulas

Introduction

A main problem in approximation theory is to estimate the rate of convergence for sequences of positive linear approximation operators. Voronovsakja type formulas are one of the most important tools for studying their asymptotic behavior.

Generalizations of the classical Voronovskaja's theorem [16] related to the Bernstein polynomials were intensively studied by Bernstein [1], Mamedov [11], Sikkema and van der Meer [13], Gonska [8], Telyakovskii [15], Tachev [14], Gavrea and Ivan [5–7], and in many subsequent papers that deal with this topic. A key tool in this analysis is a limit involving the central moments of the Bernstein operators. G. Tachev conjectured that this limit is zero [14, p. 1183]. Recently, the authors [5] provided a positive answer to this question.

The main objective of this paper is to give an analysis, as general as possible, of the rate of convergence of positive linear approximation operators defined on $C[0,1]$.

I. Gavrea • M. Ivan (✉)
Department of Mathematics, Technical University of Cluj Napoca, Str. Memorandumului nr. 28, Cluj-Napoca, 400114, Romania
e-mail: ioan.gavrea@math.utcluj.ro; mircea.ivan@math.utcluj.ro

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_11

Throughout the paper we will use the following notations:

$C^k[0,1]$, the space of all functions $f\colon [0,1] \to \mathbb{R}$, possessing a continuous derivative of order k, $k \in \mathbb{N}_0$, endowed with the sup norm $\|f\| := \sup_{x\in[0,1]} |f(x)|$ ($C[0,1] := C^0[0,1]$);

$L_n\colon C[0,1] \to C[0,1]$, $n = 0, 1, \ldots$, a sequence of *positive linear approximation operators preserving constants*;

$R(L_n, f, q, x) :=\quad L_n(f;x) - \sum_{i=0}^{q} L_n((\cdot - x)^i; x)\frac{f^{(i)}(x)}{i!}$;

$\tilde{\omega}(f;\delta)$, the *least concave majorant* of $f\colon [0,1] \to \mathbb{R}$ given by

$$\tilde{\omega}(f;\delta) = \begin{cases} \sup\limits_{0\le x\le\delta<y\le 1} \dfrac{(\delta - x)\omega(f;y) + (y-\delta)\omega(f;x)}{y-x}, & 0 < \delta \le 1 \\ \omega(f;1), & \delta > 1 \end{cases}$$

(see, e.g., [4, 10, 12] and [3, Chaps. 2 & 6] for details).

$a_n = o(b_n)$, Landau's little-o notation (there exists a sequence 0_n converging to zero such that $a_n = 0_n\, b_n$).

Let $f\colon [0,1] \to \mathbb{R}$. For any non-negative integer n, the classical Bernstein operator $B_n\colon C[0,1] \to C[0,1]$ is defined by

$$B_n(f;x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right), \qquad x \in [0,1].$$

The classical Voronovskaja's theorem on the asymptotic behavior of Bernstein polynomials can be stated as follows.

Theorem 1 (Voronovskaja [16]). *If f is bounded on $[0,1]$ and, for some $x \in [0,1]$, f is differentiable in a neighborhood of x and possesses a second derivative $f''(x)$, then*

$$\lim_{n\to\infty} n\,(B_n(f;x) - f(x)) = \frac{x(1-x)}{2} f''(x).$$

If $f \in C^2[0,1]$, the convergence is uniform.

In 1932, S. N. Bernstein gave the following generalization of Voronovskaja's result (Theorem 1).

Theorem 2 (Bernstein [1]). *If $q \in \mathbb{N}$ is even and $f \in C^q[0,1]$, then,*

$$\lim_{n\to\infty} n^{q/2} \left(B_n(f;x) - \sum_{r=0}^{q} B_n((\cdot - x)^r; x) \frac{f^{(r)}(x)}{r!} \right) = 0,$$

uniformly on $[0,1]$.

In 1962, R.G. Mamedov obtained the following generalization of Bernstein's result.

Theorem 3 (Mamedov [11]). *If $q \in \mathbb{N}$ is even, $f \in C^q[0,1]$, and $L_n \colon C[0,1] \to C[0,1]$ is a sequence of positive linear operators preserving constants such that*

$$\lim_{n\to\infty} \frac{L_n((\cdot - x)^{q+2j}; x)}{L_n((\cdot - x)^q; x)} = 0 \tag{1}$$

for some $x \in [0,1]$ and for at least one integer $j > 0$, then

$$\lim_{n\to\infty} \frac{L_n(f;x) - \sum\limits_{r=0}^{q} L_n((\cdot - x)^r; x)\frac{f^{(r)}(x)}{r!}}{L_n((\cdot - x)^q; x)} = 0.$$

Let $WC^q[0,1]$, $q \geq 0$ denote the set of all functions in $\mathbb{R}^{[0,1]}$ possessing a piecewise continuous derivative of order q. A complete asymptotic expansion in quantitative form was given by Sikkema and van der Meer [13].

Theorem 4 (Sikkema and van der Meer [13]). *Let $(L_n)_{n\in\mathbb{N}}$ be a sequence of positive linear operators $L_n \colon WC^q[0,1] \to C[0,1]$ satisfying $L_n(e_0) = 1$. Then for all $f \in WC^q[0,1]$, $x \in [0,1]$, $n \in \mathbb{N}$ and $\delta > 0$, one has*

$$\left| L_n(f;x) - \sum_{r=0}^{q} L_n((\cdot - x)^r; x)\frac{f^{(r)}(x)}{r!} \right| \leq C_{n,q}(x,\delta)\; \omega(f^{(q)}; \delta),$$

where

$$C_{n,q}(x,\delta) = \delta^q L_n\left(s_{q,\mu}\left(\frac{\cdot - x}{\delta}\right); x\right),$$

$$\mu = \begin{cases} \frac{1}{2}, & \text{if} L_n\left((\cdot - x)^q; x\right) \geq 0, \\ -\frac{1}{2}, & \text{if} L_n\left((\cdot - x)^q; x\right) < 0, \end{cases}$$

$$s_{q,\mu}(u) = \frac{1}{q!}\left(\frac{1}{2}|u|^q + \mu\, u^q\right) + \frac{1}{(q+1)!}\left(b_{q+1}(|u|) - b_{q+1}(|u| - \lfloor |u| \rfloor)\right).$$

Here b_{q+1} is the Bernoulli polynomial of degree $q+1$, $\lfloor\ \rfloor$ denotes the integer part function, and $\omega(g;\delta)$ is the classical first order modulus of continuity of $g \in C[0,1]$. Moreover the functions $C_{n,q}(x,\delta)$ are the best possible.

In 2007, H. Gonska provided a quantitative improvement of Mamedov's Theorem 3:

Theorem 5 (Gonska [8, Theorem 3.2.]). *Let $q \in \mathbb{N}$, $f \in C^q[0,1]$ and $L \colon C[0,1] \to C[0,1]$ be a positive linear operator. Then*

$$\left| L(f;x) - \sum_{r=0}^{q} L((\cdot - x)^r;x)\frac{f^{(r)}(x)}{r!} \right| \tag{2}$$
$$\leq \frac{L(|\cdot - x|^q;x)}{q!}\,\tilde{\omega}\left(f^{(q)}; \frac{1}{q+1}\frac{L\left(|\cdot - x|^{q+1};x\right)}{L(|\cdot - x|^q;x)}\right),$$

where $\tilde{\omega}(f;\delta)$ is the least concave majorant of f.

In 2012, by using the K-functionals technique, Tachev [14, Theorem 2] proved that the classical result of Bernstein (Theorem 2) remains valid for odd values of q. In this context we mention the following conjecture.

Conjecture 1 (Tachev [14, p. 1183]). For $q \in \mathbb{N}$ and any $x \in (0,1)$, the following statement holds true.

$$\lim_{n\to\infty} \frac{B_n\left((\cdot - x)^{q+1};x\right)}{B_n(|\cdot - x|^q;x)} = 0.$$

Recently we have given [5, Theorem 13] an affirmative answer to Conjecture 1. By using our result it follows that both the classical result of Bernstein (Theorem 2), and its extension to odd values of q [14, Theorem 2] are immediate consequences of Theorem 5.

The Mamedov Property

Let $x \in [0,1]$ and $q \in \mathbb{N}$. As a necessity revealed by Conjecture 1, and in view of condition (1), we propose the following definition.

Definition 1. We will refer to a sequence of operators (L_n) satisfying

$$\lim_{n\to\infty} \frac{L_n\left(|\cdot - x|^{q+1};x\right)}{L_n(|\cdot - x|^q;x)} = 0, \tag{3}$$

as possessing *the Mamedov property.*

Remark 1. In connection with the Mamedov property we would like to emphasize the following points:

- If (L_n) is a sequence of positive linear operators preserving constants, then the sequence $\left(\frac{L_n\left(|\cdot - x|^{i+1};x\right)}{L_n(|\cdot - x|^i;x)}\right)_{i\in\mathbb{N}}$ is nondecreasing.
- If a sequence of positive linear operators preserving constants (L_n) possesses the Mamedov property for a certain integer q, then:
  - the same is true for all integers $i = 0, 1, \ldots, q$;
  - the sequence $\left(L_n\left(|\cdot - x|^i;x\right)\right)_{i=0}^{q+1}$ constitutes an asymptotic scale as $n \to \infty$.

Indeed, if $L_n\left(|\cdot - x|^q; x\right) \neq 0$ then $L_n\left(|\cdot - x|^i; x\right) \neq 0$ for all $i \in \mathbb{N}$ (cf. Lemma 1). Moreover, taking into account the fact that the sequence $\dfrac{L_n\left(|\cdot - x|^{i+1}; x\right)}{L_n\left(|\cdot - x|^i; x\right)}$, $i \in \mathbb{N}$, is nondecreasing (cf. the Cauchy–Schwarz inequality for positive linear functionals), we deduce that

$$\lim_{n\to\infty} \frac{L_n\left(|\cdot - x|^{i+1}; x\right)}{L_n\left(|\cdot - x|^i; x\right)} = 0, \qquad i = 0, 1, \ldots, q.$$

To check whether a sequence of operators possesses the Mamedov property is not an easy task, even if we restrict ourselves to classical approximation operators. Note that the particular case of the Bernstein operators has only been solved recently. As far as we know, the problem in its general form has not been studied in the literature. The Mamedov condition (3) will be a key tool in our analysis.

Some Operators Failing to Satisfy the Mamedov Property

There was a *strong belief that positive linear approximation operators possess the Mamedov property.* This was probably induced by the case of the Bernstein operators.

It is in this regard that the following remarks are particularly significant. They warn us that, in general, sequences of positive linear approximation operators fail to meet the Mamedov conditions (3).

Remark 2 (Gavrea and Ivan [7, Theorem 4]). There exist positive linear approximation operators preserving constants, $L_n\colon C[0,1] \to C[0,1]$, such that

$$\frac{L_n(|\cdot - x|^{q+1}; x)}{L_n(|\cdot - x|^q; x)} \not\to 0, \qquad \text{as} \quad n \to \infty, \qquad x \in [0,1],$$

for any $q > 0$.

An example is the sequence of operators

Proof (cf. [7]).

$$L_n f = \frac{f(0) + f(1) + nf}{n+2}, \qquad n \in \mathbb{N}.$$

$$\lim_{n\to\infty} \frac{L_n(|\cdot - x|^{q+1}; x)}{L_n(|\cdot - x|^q; x)} = \frac{x^{q+1} + (1-x)^{q+1}}{x^q + (1-x)^q} \neq 0, \qquad x \in [0,1], \quad q > 0.$$

□

Remark 3. For any $q > 0$, there exists a sequence of positive linear approximation operators preserving constants, $L_n\colon C[0,1] \to C[0,1]$, such that

$$\lim_{n\to\infty} \frac{L_n(|x-\cdot|^{q+1};x)}{L_n(|x-\cdot|^{q};x)} = 0, \qquad \text{and} \qquad \lim_{n\to\infty} \frac{L_n(|x-\cdot|^{q+2};x)}{L_n(|x-\cdot|^{q+1};x)} \neq 0.$$

Proof. Indeed, let $q > 0$ and define the positive linear approximation operators preserving constants $L_n\colon C[0,1] \to C[0,1]$,

$$L_n f = f(0)\, n^{-1} + \left(1 - n^{-1}\right) f\left(\left(1 - n^{-\frac{1}{q+1}}\right)\cdot\right), \qquad n = 1, 2, \ldots$$

For $f = |\cdot - x|^a$, we obtain

$$L_n(f;x) = \frac{x^a}{n} + \left(1 - \frac{1}{n}\right) n^{-\frac{a}{1+q}}\, x^a,$$

and

$$\frac{L_n(|\cdot - x|^{a+1};x)}{L_n(|\cdot - x|^{a};x)} = \frac{n^{\frac{a}{1+q}} + n^{\frac{q}{1+q}} - n^{-\frac{1}{1+q}}}{n + n^{\frac{a}{1+q}} - 1}\, x, \qquad x \in (0,1].$$

Therefore, we obtain

$$\lim_{n\to\infty} \frac{L_n(|\cdot - x|^{q+1};x)}{L_n(|\cdot - x|^{q};x)} = \lim_{n\to\infty} \frac{2n^{\frac{q}{1+q}} - n^{-\frac{1}{1+q}}}{n + n^{\frac{q}{1+q}} - 1}\, x = 0$$

$$\lim_{n\to\infty} \frac{L_n(|\cdot - x|^{q+2};x)}{L_n(|\cdot - x|^{q+1};x)} = \lim_{n\to\infty} \frac{n + n^{\frac{q}{1+q}} - n^{-\frac{1}{1+q}}}{2n - 1}\, x = \frac{x}{2}.$$

□

Some Properties Related to the Rate of Convergence of Positive Linear Operators

In this section we provide some auxiliary results.

Lemma 1 (Gavrea and Ivan [7, Lemma 5]). *Let $L\colon C[0,1] \to C[0,1]$ be a positive linear operator preserving constants and $x \in [0,1]$. If there exists a function $\psi \in C[0,1]$ such that*

(a) $\psi(x) = 0$ and $\psi(t) > 0$ for all $t \in [0,1] \setminus \{x\}$ (in particular, e.g., $\psi = |\cdot - x|^q$),
(b) $L\psi = 0$,

then x is an interpolation point for L, i.e., $L(f,x) = f(x)$, $\forall f \in C[0,1]$. In particular, if there exists $q > 0$ such that $L(|\cdot - x|^q, x) = 0$, then $L(|\cdot - x|^p, x) = 0$ for all $p > 0$.

In connection with the interpolation, we denote by S_L, or simply by S when there is no danger of ambiguity, the set of all non-interpolation points of the operator L. Similarly, $x \in S_{(L_n)}$ means that x is not an interpolation point for any operator L_n, $n \in \mathbb{N}$.

Since at any interpolation point of L "the approximation is exact," in virtue of Lemma 1 it remains to study the Mamedov property of the operator L only on the set S_L.

Proposition 1. *Let $x \in S_{(L_n)}$. The necessary and sufficient condition such that*

$$\lim_{n\to\infty} \frac{L_n(f;x) - \sum_{i=1}^{q} \frac{f^{(i)}(x)}{i!} L_n((\cdot - x)^i; x)}{L_n(|\cdot - x|^q; x)} = 0, \tag{4}$$

for all $f \in C^q[0,1]$, is that the Mamedov condition

$$\lim_{n\to\infty} \frac{L_n(|\cdot - x|^{q+1}; x)}{L_n(|\cdot - x|^q; x)} = 0, \tag{5}$$

be satisfied.

Proof. Suppose that (4) is satisfied for all $f \in C^q[0,1]$. Then, by taking $f_x = |\cdot - x|^{q+1}$ in (4), we obtain

$$L_n(f_x;x) - \sum_{i=0}^{q} \frac{f_x^{(i)}(x)}{i!} L_n((\cdot - x)^i; x) = L_n(|\cdot - x|^{q+1}; x),$$

i.e., (5) is true. Conversely, if (5) holds true, then by using (2) we deduce that (4) also holds true. □

Remark 4. We mention that the constant $\frac{1}{q!}$ in (2) is the best possible.

Proof. Suppose that there exists $\alpha \in (0,1)$ such that

$$\left| L_n(f;x) - \sum_{i=0}^{q} \frac{f^{(i)}(x)}{i!} L_n((\cdot - x)^i; x) \right| \tag{6}$$
$$\le \alpha \frac{L_n(|\cdot - x|^q; x)}{q!} \tilde{\omega}\left(f^{(q)}; \frac{L_n(|\cdot - x|^{q+1}; x)}{(q+1)\, L_n(|\cdot - x|^q; x)} \right),$$

for all $f \in C^q[0,1]$. For $f = |\cdot - x|^{q+1}$, Eq. (6) becomes

$$L_n(|\cdot - x|^{q+1}; x) \le \alpha\, L_n(|\cdot - x|^q; x)(q+1)\, \tilde{\omega}\left(|\cdot - x|; \frac{L_n(|\cdot - x|^{q+1}; x)}{(q+1)\, L_n(|\cdot - x|^q; x)} \right), \tag{7}$$

Since the Lipschitz constant of the function $|\cdot - x|$ is 1, we obtain

$$\tilde{\omega}\left(|\cdot - x|; \frac{L_n(|\cdot - x|^{q+1}; x)}{(q+1)\, L_n(|\cdot - x|^q; x)}\right) \le \frac{L_n(|\cdot - x|^{q+1}; x)}{(q+1)\, L_n(|\cdot - x|^q)}. \tag{8}$$

From (7) and (8) we obtain

$$L_n(|\cdot - x|^{q+1}; x) \le \alpha\, L_n(|\cdot - x|^{q+1}; x),$$

which is false (cf. $x \in S_{(L_n)}$). □

A Sufficient Condition for the Mamedov Property

A general Bernstein–Voronovskaja property is provided in the following theorem:

Theorem 6 (Gavrea and Ivan [7, Theorem 10]). *If* $L_n : C[0,1] \to C[0,1]$ *is a sequence of positive linear approximation operators,* $f \in C^q[0,1]$ *and* $x \in [0,1]$*, then*

$$R(L_n, f, q, x) = o\left(\sqrt{L_n((\cdot - x)^{2q}; x)}\right), \quad \text{as } n \to \infty.$$

In order to obtain a better order of approximation, e.g., $o\,(L_n(|\cdot - x|^q; x))$, in view of (2), we must restrict ourselves to the class of positive linear approximation operators possessing the Mamedov property.

The following theorem provides a sufficient condition for a sequence $(L_n)_{n\in\mathbb{N}}$ of positive linear approximation operators to possess the Mamedov property.

Theorem 7. *Let* $x \in S_{(L_n)}$ *and* $q > 0$*. If there exists a constant* $K > 0$ *such that*

$$\frac{L_n(|\cdot - x|^{2q}; x)}{L_n^2(|\cdot - x|^q; x)} \le K, \qquad \forall n \in \mathbb{N},$$

then the Mamedov condition

$$\lim_{n\to\infty} \frac{L_n(|\cdot - x|^{q+1}; x)}{L_n(|\cdot - x|^q; x)} = 0,$$

is satisfied.

Proof. Let us suppose that the Mamedov condition is not satisfied. It follows that there exist $\lambda > 0$ and a sequence of integers $(n_k)_{k\ge 0}$ such that

$$\frac{L_{n_k}(|\cdot - x|^{q+1}; x)}{L_{n_k}(|\cdot - x|^q; x)} > \lambda, \qquad \forall\, k \ge 0.$$

For any fixed n_k, an application of the Cauchy–Schwarz inequality for positive linear functionals proves that the sequence

$$\left(\frac{L_{n_k}(|\cdot - x|^{q+1}; x)}{L_{n_k}(|\cdot - x|^{q}; x)}\right)_{q\geq 1}$$

is non-decreasing. It follows that

$$\frac{L_{n_k}(|\cdot - x|^{q+i}; x)}{L_{n_k}(|\cdot - x|^{q+i-1}; x)} > \lambda, \qquad i = 1, \ldots, q. \tag{9}$$

Taking the product of both sides of (9) over $1 \leq i \leq q$, we obtain

$$\frac{L_{n_k}(|\cdot - x|^{2q}; x)}{L_{n_k}(|\cdot - x|^{q}; x)} > \lambda^q,$$

hence

$$0 < \frac{1}{K} \leq \frac{L^2_{n_k}(|\cdot - x|^{q}; x)}{L_{n_k}(|\cdot - x|^{2q}; x)} < \left(\frac{1}{\lambda}\right)^q L_{n_k}(|\cdot - x|^{q}; x) \overset{k\to\infty}{\longrightarrow} 0,$$

and this contradiction completes the proof. □

The Rate of Convergence of the Durrmeyer Operators

In this section we give an application of Theorem 7 to a classical approximation process, namely the Durrmeyer one. We prove that the Durrmeyer operators satisfy the Mamedov property, and implicitly find an optimal estimate for the remainder term in the approximation process.

Let $f\colon [0,1] \to \mathbb{R}$ be integrable. The Durrmeyer operators are defined by (see, e.g., Derriennic [2])

$$D_n(f; x) = (n+1) \sum_{k=0}^{n} p_{n,k}(x) \int_0^1 p_{n,k}(t) f(t)\, \mathrm{d}t, \qquad n \in \mathbb{N}.$$

For fixed $n \in \mathbb{N}^*$, we use the notation

$$M_r(x) = D_n((\cdot - x)^r; x), \qquad x \in [0,1], \quad r \in \mathbb{N}^*.$$

The recurrence formula given in the theorem below will be essential to obtain estimates for the central moments of the Durrmeyer operators.

Theorem 8. *For any integer $r \geq 1$, the following identity is satisfied.*

$$\begin{aligned} & M_{r+1}(x) \qquad (10) \\ &= \frac{1}{n+r+2}\Big((r+1)(1-2x)M_r(x) + x(1-x)M_r'(x) + 2\,r\,x(1-x)M_{r-1}(x)\Big) \end{aligned}$$

Proof. We have

$$\begin{aligned} & x(1-x)M_r'(x) \qquad (11) \\ &= (n+1)\sum_{k=0}^{n}(k-nx)p_{n,k}(x)\int_0^1 p_{n,k}(t)(t-x)^r\,\mathrm{d}t - rx(1-x)M_{r-1}. \end{aligned}$$

Note that

$$\begin{aligned} & (k-nx)\,p_{n,k}(x)\int_0^1 p_{n,k}(t)(t-x)^r\,\mathrm{d}t \\ &= p_{n,k}(x)\int_0^1 (k-nt)p_{n,k}(t)(t-x)^r\,\mathrm{d}t + n\int_0^1 p_{n,k}(t)(t-x)^{r+1}\,\mathrm{d}t, \end{aligned}$$

for $k = 0,\ldots,n$. We have

$$\int_0^1 (k-nt)p_{n,k}(t)(t-x)^r\,\mathrm{d}t = \int_0^1 (p_{n,k}(t))'\,t(1-t)\,(t-x)^r\,\mathrm{d}t.$$

Integrating by parts, we get

$$\begin{aligned} & \int_0^1 (k-nt)\,p_{n,k}(t)(t-x)^r\,\mathrm{d}t \qquad (12) \\ &= -\int_0^1 p_{n,k}(t)\,(t-x)^{r-1}\big((1-2t)(t-x) + rt(1-t)\big)\,\mathrm{d}t \\ &= -\int_0^1 p_{n,k}(t)\,(t-x)^{r-1}\Big(rx(1-x) + (r+1)(1-2x)(t-x) - (r+2)(t-x)^2\Big)\,\mathrm{d}t \\ &\quad + n\int_0^1 p_{n,k}(t)(t-x)^{r+1}\,\mathrm{d}t. \end{aligned}$$

From (11) and (12) we obtain (10) and the proof is complete. ◻

Corollary 1. *The following statements hold true.*

(a) $M_{2s}(x) = \dfrac{\sum_{i=1}^{s}\lambda_i\,n^i X^i + \sum_{i=0}^{s}\beta_i X^i}{(n+2)(n+3)\cdots(n+2s+1)}$
where λ_i, β_i are independent of x and n, and $\beta_0 = (2s)!$

(b) $M_{2s+1}(x) = (1-2x)\dfrac{\sum_{i=1}^{s} a_i n^i X^i + \sum_{i=0}^{s} b_i X^i}{(n+2)(n+3)\cdots(n+2s+2)}$
where a_i, b_i are independent of x and n, and $b_0 = (2s+1)!$

Proof. We prove the statements using mathematical induction on r, the equalities

$$M_1(x) = \frac{1-2x}{n+2}, \qquad M_2(x) = \frac{2nX + 2 - 6X}{(n+2)(n+3)},$$

and the recurrence formula (10). □

Corollary 2. *The following relations are satisfied.*

(a) *For $x \in [0,1]$, $r > 0$ and $nX \geq 1$, the following inequality is satisfied*

$$D_n(|\cdot - x|^r; x) \leq C(r)\left(\frac{X}{n}\right)^{r/2} + \frac{B(r)}{n^r},$$

where $C(r)$ and $B(r)$ are independent of x and n.

(b) *For $x \in [0,1]$, $r > 0$ and $nX < 1$, we have*

$$D_n(|\cdot - x|^r; x) \leq \frac{D(r)}{n^r},$$

where $D(r)$ is independent of x and n.

Proof. We start with the proof of statement (a). We first consider the case $r = 2m > 0$. By Corollary 1(a)

$$\begin{aligned} D_n(|\cdot - x|^{2m}; x) &\leq \sum_{i=1}^{m} \frac{|\lambda_i| X^i}{n^{2m-i}} + \sum_{i=1}^{m} \frac{|\beta_i| X^i}{n^{2m}} \\ &= \frac{1}{n^m} \sum_{i=1}^{m} \frac{|\lambda_i| X^m}{(nX)^{m-i}} + \sum_{i=1}^{m} \frac{|\beta_i| X^i}{n^{2m}} \end{aligned}$$

With $C(m) = \sum_{i=1}^{m} |\lambda_i|$, $B(m) = \max\limits_{x \in [0,1]} \sum_{i=1}^{m} |\beta_i| X^i$, we obtain

$$D_n(|\cdot - x|^{2m}; x) \leq C(m)\left(\frac{X}{m}\right)^m + \frac{B(m)}{n^{2m}}.$$

This proves statement (a) for even r.

Let $2m$ be the least even number greater than r. Since the function $t \mapsto t^{\frac{r}{2m}}$ is concave on $(0,\infty)$, by using the Jessen inequality for functionals [9], we obtain

$$\begin{aligned} D_n(|\cdot - x|^r; x) &= \sum_{k=0}^{n} p_{n,k}(x)(n+1)\int_0^1 p_{n,k}(t)\,|t-x|^r\,\mathrm{d}t \\ &= \sum_{k=0}^{n} p_{n,k}(x)(n+1)\int_0^1 p_{n,k}(t)\,\left(|t-x|^{2m}\right)^{\frac{r}{2m}}\,\mathrm{d}t \\ &\le \left(\sum_{k=0}^{n} p_{n,k}(x)(n+1)\int_0^1 p_{n,k}(t)\,|t-x|^{2m}\,\mathrm{d}t\right)^{\frac{r}{2m}}. \end{aligned}$$

Since $0 < \frac{r}{2m} < 1$, we obtain

$$\begin{aligned} D_n(|\cdot - x|^r; x) &\le \left((n+1)\sum_{k=0}^{n} p_{n,k}(x)\int_0^1 p_{n,k}(t)\,|t-x|^{2m}\,\mathrm{d}t\right)^{\frac{r}{2m}} \\ &\le \left(C(2m)\left(\frac{X}{n}\right)^m + \frac{B(2m)}{n^{2m}}\right)^{\frac{r}{2m}} \\ &\le (C(2m))^{\frac{r}{2m}}\left(\frac{X}{n}\right)^{\frac{r}{2}} + \frac{(B(2m))^{\frac{r}{2m}}}{n^r}. \end{aligned}$$

and the proof of statement (a) is complete.

Let us now proceed with the proof of statement (b). For $r = 2m > 0$, from Corollary 1, statement (a), we obtain

$$D_n(|\cdot - x|^{2m}; x) \le \frac{\sum_{i=1}^{m}(|\lambda_i| + |\beta_i|)}{(n+2)\cdots(n+2m+1)} \le \frac{D(2m)}{n^{2m}}$$

where $D(2m) = \sum_{i=1}^{m}(|\lambda_i| + |\beta_i|)$.

Let $r > 0$ and let $2m$ be the least even number greater than r. The proof repeats the steps of proving the statement (a). □

The next theorem shows us that the sequence of the Durrmeyer operators satisfies the Mamedov property.

Theorem 9. *For any integer $q > 0$, the following limit holds uniformly in x,*

$$\lim_{n\to\infty} \frac{D_n(|\cdot - x|^{q+1}; x)}{D_n(|\cdot - x|^q; x)} = 0.$$

Proof. By Theorem 7, it is sufficient to prove that there exists a number $K > 0$ such that

$$\frac{D_n(|\cdot - x|^{2q};x)}{D_n^2(|\cdot - x|^q;x)} \leq K, \quad \forall x \in [0,1], \quad \forall n \in \mathbb{N}.$$

Consider the case $nX \geq 1$. We have

$$D_n(|\cdot - x|^{2q};x) \leq C(2q)\frac{X^q}{n^q} + \frac{B(2q)}{n^{2q}} \leq C(2q)\frac{X^q}{n^q} + \frac{B(2q)X^q}{n^q}$$

By Jessen inequality for positive linear functionals [9], we obtain

$$D_n(|\cdot - x|^q;x) \geq (D_n((\cdot - x)^2;x))^{q/2}. \tag{13}$$

By using the equality

$$D_n((\cdot - x)^2;x) = \frac{X(2n-6)+2}{(n+2)(n+3)},$$

for $n \geq 4$, we obtain

$$D_n((\cdot - x)^2;x) \geq \frac{X(2n-6)}{(n+2)(n+3)} \tag{14}$$

From (13) and (14) we obtain

$$D_n(|\cdot - x|^q;x) \geq \frac{X^{q/2}(2n-6)^{q/2}}{(n+2)^{q/2}(n+3)^{q/2}} \tag{15}$$

From (13) and (15) we obtain

$$\frac{D_n(|\cdot - x|^{2q};x)}{D_n^2(|\cdot - x|^q;x)} \leq A(2q)\left(\frac{(n+2)(n+3)}{(2n-6)n}\right)^q \leq A(2q)\left(\frac{21}{4}\right)^q \tag{16}$$

In the case $nX < 1$, we have

$$D_n(|\cdot - x|^{2q};x) \leq \frac{D(2q)}{n^{2q}}, \tag{17}$$

and

$$D_n((\cdot - x)^2;x) \geq \frac{2}{(n+2)(n+3)}, \tag{18}$$

From (17) and (18) we get

$$\frac{D_n((\cdot - x)^{2q};x)}{D_n^2(|\cdot - x|^q;x)} \leq D(2q)\left(\frac{(n+2)(n+3)}{2n^2}\right)^q \leq D(2q)\left(\frac{21}{19}\right)^q, \tag{19}$$

Equations (16) and (19) yield

$$\frac{D_n((\cdot - x)^{2q}; x)}{D_n^2(|\cdot - x|^q; x)} \le K = \max\left\{A(2q)\left(\frac{21}{4}\right)^q, D(2q)\left(\frac{21}{19}\right)^q\right\},$$

and the proof is complete. □

Since the Durrmeyer operators possess the Mamedov property, the (D_n) remainder estimate deduced from Eq. (2) becomes effective.

References

1. Bernstein, S.: Complément à l'article de E. Voronovskaya "Détermination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein". C. R. Acad. Sci. URSS **1932**, 86–92 (1932)
2. Derriennic, M.M.: Sur l'approximation de fonctions intégrables sur [0, 1] par des polynômes de Bernstein modifiés. J. Approx. Theory **31**(4), 325–343 (1981)
3. DeVore, R.A., Lorentz, G.G.: Constructive approximation. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 303. Springer, Berlin (1993)
4. Gal, S.G.: Properties of the modulus of continuity for monotonous convex functions and applications. Int. J. Math. Math. Sci. **18**(3), 443–446 (1995)
5. Gavrea, I., Ivan, M.: An answer to a conjecture on Bernstein operators. J. Math. Anal. Appl. **390**(1), 86–92 (2012)
6. Gavrea, I., Ivan, M.: A simple proof of a generalized Bernstein Voronovskaja-type theorem. Ann. Tiberiu Popoviciu Semin. Funct. Equ. Approx. Convexity **12**, 33–36 (2014)
7. Gavrea, I., Ivan, M.: The Bernstein Voronovskaja-type theorem for positive linear approximation operators. J. Approx. Theory **192**, 291–296 (2015)
8. Gonska, H.: On the degree of approximation in Voronovskaja's theorem. Stud. Univ. Babeş-Bolyai Math. **52**(3), 103–115 (2007)
9. Jessen, B.: Bemærkninger om konvekse Funktioner og Uligheder imellem Middelværdier. I. Mat. Tidsskrift **B**, 17–28 (1931)
10. Kornejchuk, N.: The best uniform approximation on certain classes of continuous functions. Sov. Math. Dokl. **2**, 1254–1257 (1961)
11. Mamedov, R.G.: The asymptotic value of the approximation of multiply differentiable functions by positive linear operators. Dokl. Akad. Nauk SSSR **146**, 1013–1016 (1962)
12. Mitjagin, B.S., Semenov, E.M.: Absence of interpolation of linear operators in spaces of smooth functions. Izv. Akad. Nauk SSSR Ser. Mat. **41**(6), 1289–1328, 1447 (1977)
13. Sikkema, P.C., van der Meer, P.J.C.: The exact degree of local approximation by linear positive operators involving the modulus of continuity of the pth derivative. Nederl. Akad. Wetensch. Indag. Math. **41**(1), 63–76 (1979)
14. Tachev, G.T.: The complete asymptotic expansion for Bernstein operators. J. Math. Anal. Appl. **385**(2), 1179–1183 (2012)
15. Telyakovskii, S.A.: On the rate of approximation of functions by the Bernstein polynomials. Proc. Steklov Inst. Math. **264**(suppl. 1), S177–S184 (2009)
16. Voronovskaja, E.: Détermination de la forme asyptotique de l'approximation des fonctions par les polynômes de M. Bernstein. C. R. Acad. Sci. URSS **1932**, 79–85 (1932)

Approximation of Functions by Additive and by Quadratic Mappings

Laura Găvruţa and Paşc Găvruţa

Abstract In this chapter, we characterize the functions with values in a Banach space which can be approximated by additive mappings, with a given error. Also, we give a characterization of functions with values in a Banach space which can be approximated by a quadratic mapping, with a given error.

Keywords Hyers-Ulam-Rassias stability • Additive mapping • Quadratic mapping

Introduction

The study of stability problems for various functional equations originated from a question posed by Ulam [17] in 1940, at the University of Wisconsin. He presented there a number of unsolved problems and among them was the following question concerning the stability of group homomorphisms.

Given G_1 a group, G_2 a metric group and $\varepsilon > 0$, find $\delta > 0$ such that, if

$$f : G_1 \to G_2$$

satisfies

$$d(f(xy), f(x)f(y)) \leq \delta, \quad \text{for all } x, y \in G_1$$

then there exists a homomorphism $g : G_1 \to G_2$ such that

$$d(f(x), g(x)) \leq \varepsilon, \quad \text{for all } x \in G_1.$$

The first affirmative answer to this question was the one provided by Hyers [7], who solved the problem for Banach spaces, with

L. Găvruţa (✉) • P. Găvruţa
Department of Mathematics, Politehnica University of Timişoara, Piaţa Victoriei no. 2, 300006 Timişoara, Romania
e-mail: gavruta_laura@yahoo.com; pgavruta@yahoo.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_12

$$\delta = \varepsilon \text{ and } g(x) = \lim_{n\to\infty} \frac{f(2^n x)}{2^n}.$$

Theorem 1 (Hyers [7]). *Let $f : E_1 \to E_2$ (E_1, E_2 are Banach spaces) be a function such that*

$$\|f(x+y) - f(x) - f(y)\| \le \delta$$

for some $\delta > 0$ and for all $x, y \in E_1$. Then the limit

$$T(x) = \lim_{n\to\infty} \frac{f(2^n x)}{2^n}$$

exists for each $x \in E_1$ and $T : E_1 \to E_2$ is the unique additive mapping such that

$$\|f(x) - T(x)\| \le \delta, \quad \text{for every } x \in E_1.$$

Moreover, if $f(tx)$ is continuous in t for each fixed $x \in E_1$, then the function T is linear.

Another important result was obtained by Rassias [13] for approximately additive mappings, by using the so-called *the direct method*.

Theorem 2 (Rassias [13]). *Let $f : E_1 \to E_2$ be a function between Banach spaces, such that $f(tx)$ is continuous in t for each fixed x. If f satisfies the functional inequality*

$$\|f(x+y) - f(x) - f(y)\| \le \theta(\|x\|^p + \|y\|^p)$$

for some $\theta \ge 0$, $0 \le p < 1$ and for all $x, y \in E_1$, then there exists a unique linear mapping $T : E_1 \to E_2$ such that

$$\|f(x) - T(x)\| \le \frac{2\theta}{2 - 2^p}\|x\|^p, \quad \text{for each } x \in E_1.$$

A further generalization was obtained by Găvruţa [4], by replacing the Cauchy difference by a control mapping φ and also introduced the concept of generalized Hyers–Ulam–Rassias stability in the spirit of Th.M. Rassias' approach.

Theorem 3 (Găvruţa [4]). *Let be $\varphi : S \times S \to [0, \infty)$ (where S is an abelian semigroup) a mapping such that*

$$\Phi(x, y) := \sum_{n=0}^{\infty} \frac{\varphi(2^n x, 2^n y)}{2^{n+1}} < \infty$$

and let $f : S \to X$ be such that

$$\|f(x+y)-f(x)-f(y)\| \leq \varphi(x,y), \quad \text{for all } x,y \in S. \tag{1}$$

Then there exists a unique mapping $T : S \to X$ such that

$$T(x+y) = T(x) + T(y), \quad \text{for all } x,y \in S$$

and $\|f(x) - T(x)\| \leq \Phi(x,x) = \Phi(x), \quad$ for all $x \in S$.

Moreover, if S is a Banach space and $f(tx)$ is continuous in t for each fixed $x \in S$, then the function T is linear.

The algebraic part of the Theorem 3 was generalized in the following form

Theorem 4 ([5]). *Let $(G, +)$ be an abelian group, k an integer, $k \geq 2$, X a Banach space, $\varphi : G \times G \to [0, \infty)$ a mapping such that*

$$\Phi_k(x,y) = \sum_{n=0}^{\infty} \frac{1}{k^{n+1}} \varphi(k^n x, k^n y) < \infty$$

and $f : G \to X$ a mapping with the property

$$\|f(x+y)-f(x)-f(y)\| \leq \varphi(x,y), \quad (\forall)x,y \in G.$$

Then there exists a unique additive mapping $T : G \to X$ such that

$$\|f(x) - T(x)\| \leq \sum_{n=0}^{\infty} \Phi_k(x, mx), \quad (\forall)x \in G.$$

In the paper [6] we give a solution to the Ulam stability problem for continuous Parseval frames in Hilbert spaces. We prove that if F is a nearly Parseval frame then there exist a Parseval frame near F. Also, we give generalizations of this result. The theory of frames is closely related to some notions of the quantum information theory. In [15], Scott introduced a class of informationally complete positive-operator-valued measures which are, in analogy with a tight frame, as close at possible to orthonormal bases for the space of quantum states. For beautiful presentations of the connections between frames and POVM, see [14, 18].

In our opinion, it is possible that the results or ideas of this paper can be applied in quantum information theory.

For recent results on the *Hyers–Ulam–Rassias stability*, see also [3, 8–12].

Our paper is organized at follows. In section "Approximation of Functions by Additive Mappings" we characterize the functions with values in a Banach space, which can be approximated by additive mappings with a given error. In section "Approximation of Functions by Quadratic Mappings," we consider functions with values in a Banach space and we approximate these functions by a quadratic ones, with a given error.

Approximation of Functions by Additive Mappings

We consider S to be an abelian semigroup, X to be a Banach space and the following given functions:

$$f : S \to X \text{ and } \Phi : S \to \mathbb{R}_+.$$

Definition 1. We say that f is Φ-approximable by an additive map if there exists $T : S \to X$ additive such that

$$\|f(x) - T(x)\| \leq \Phi(x), \quad x \in S. \tag{2}$$

We say that T is the additive Φ-approximation of f.

Problem 1. Give conditions on f such that f to be Φ-approximable by an additive map.

We solve this problem by posing minimal conditions on Φ. We denote by

$$\mathscr{A} = \left\{ \Phi : S \to \mathbb{R}_+ : \lim_{n\to\infty} \frac{\Phi(2^n x)}{2^n} = 0 \right\}.$$

Theorem 5. *Let be $\Phi \in \mathscr{A}$. Then f is Φ-approximable by an additive map if and only if*

$$\lim_{n\to\infty} \frac{\|f(2^n x + 2^n y) - f(2^n x) - f(2^n y)\|}{2^n} = 0, \quad (\forall)\, x, y \in S$$

and there exists $\Psi \in \mathscr{A}$ such that

$$\|f(2^n x) - 2^n f(x)\| \leq \Psi(2^n x) + 2^n \Phi(x), \quad x \in S.$$

In this case, the additive Φ-approximation of f is unique and is given by

$$T(x) = \lim_{n\to\infty} \frac{f(2^n x)}{2^n}.$$

Proof. First, we assume that f is Φ-approximable by an additive map, i.e. condition (2) holds.

Hence $\|f(x+y) - T(x+y)\| \leq \Phi(x+y), \quad (\forall)\, x, y \in S.$

Now, for $x, y \in S$, we have

$$\begin{aligned}\|f(x+y) - f(x) - f(y)\| &\leq \|f(x+y) - T(x+y)\| + \|T(x) - f(x)\| + \|T(y) - f(y)\| \\ &\leq \Phi(x+y) + \Phi(x) + \Phi(y).\end{aligned}$$

It follows

$$\frac{\|f(2^n(x+y))-f(2^nx)-f(2^ny)\|}{2^n} \le \frac{\Phi(2^n(x+y))+\Phi(2^n(x))+\Phi(2^ny)}{2^n}$$

By letting n go to infinity in the above inequality, we obtain the desired condition, i.e.

$$\lim_{n\to\infty}\frac{\|f(2^nx+2^ny)-f(2^nx)-f(2^ny)\|}{2^n}=0.$$

In (2) we put instead of $x \mapsto 2^nx$ and then we multiply (2) by 2^n. So, we get

$$\begin{aligned}\|f(2^nx)-2^nf(x)\| &\le \|f(2^nx)-T(2^nx)\|+\|2^nT(x)-2^nf(x)\|\\ &\le \Phi(2^nx)+2^n\Phi(x)\end{aligned}$$

Conversely, using hypothesis, we have

$$\left\|\frac{f(2^nx)}{2^n}-f(x)\right\| \le \frac{\Psi(2^nx)}{2^n}+\Phi(x), \text{for } x\in S. \tag{3}$$

But $\left\{\frac{f(2^nx)}{2^n}\right\}$ is a Cauchy sequence. Indeed, if we put instead of $x \mapsto 2^mx$ in (3), we obtain

$$\left\|\frac{f(2^n2^mx)}{2^n}-f(2^mx)\right\| \le \frac{\Psi(2^n2^mx)}{2^n}+\Phi(2^mx)$$

and by dividing the above inequality by 2^m we get

$$\left\|\frac{f(2^{n+m}x)}{2^{n+m}}-\frac{f(2^mx)}{2^m}\right\| \le \frac{\Psi(2^{n+m}x)}{2^{n+m}}+\frac{\Phi(2^mx)}{2^m}$$

hence

$$\left\|\frac{f(2^{n+m}x)}{2^{n+m}}-\frac{f(2^mx)}{2^m}\right\| \to 0,\ m,n\to\infty.$$

Since X is a Banach space it follows that the limit $T(x)=\lim_{n\to\infty}\frac{f(2^nx)}{2^n}$ exists.

In (3), we make $n\to\infty$ and we obtain relation (2).

Now we prove that T is additive. We know, from hypothesis, that

$$\lim_{n\to\infty}\left\|\frac{f(2^nx+2^ny)}{2^n}-\frac{f(2^nx)}{2^n}-\frac{f(2^ny)}{2^n}\right\|=0$$

hence $\|T(x+y)-T(x)-T(y)\|=0$ and from here we get that T is additive.

Now we show that T is unique. We suppose that T satisfies relation (2) and exists T' which satisfies

$$\|T'(x) - f(x)\| \leq \Phi(x)$$

In (2), we take instead of $x \mapsto 2^n x$ and we get

$$\|2^n T(x) - f(2^n x)\| \leq \Phi(2^n x)$$

and from here $\|T(x) - \frac{f(2^n x)}{2^n}\| \leq \frac{\Phi(2^n x)}{2^n}$.

We use the same computation for T' and we obtain

$$\|T(x) - T'(x)\| \leq 2\frac{\Phi(2^n x)}{2^n}.$$

But $\lim_{n\to\infty} \frac{\Phi(2^n x)}{2^n} = 0$ so $T(x) = T'(x)$.

Remark 1. The algebraic part of Theorem 3 follows from the above Theorem.

The first condition from Theorem 5 it follows from hypothesis. We prove that the second condition holds. Indeed, if in (1) we take $x = y$, we obtain

$$\|f(2x) - 2f(x)\| \leq \varphi(x, x)$$

and instead of $x \mapsto 2x$ in the above inequality, we get

$$\|f(2^2 x) - 2f(2x)\| \leq \varphi(2x, 2x).$$

So

$$\begin{aligned}\|f(2^2 x) - 2^2 f(x)\| &\leq \|f(2^2 x - 2f(2x))\| + \|2f(2x) - 2^2 f(x)\| \\ &\leq \varphi(2x, 2x) + 2\varphi(x, x) \\ &= 2^2\left[\frac{\varphi(x, x)}{2} + \frac{\varphi(2x, 2x)}{2^2}\right] \\ &\leq 2^2 \Phi(x, x)\end{aligned}$$

We suppose that $\|f(2^n x) - 2^n f(x)\| \leq 2^n \Phi(x, x)$ and by taking $x \mapsto 2x$ we get

$$\|f(2^{n+1} x) - 2^n f(2x)\| \leq 2^n \Phi(2x, 2x)$$

Hence

$$\begin{aligned}\|f(2^{n+1}x) - 2^{n+1}f(x)\| &\leq 2^n \Phi(2x, 2x) + 2^n \varphi(x, x) \\ &= 2^{n+1}\left[\frac{\varphi(x, x)}{2} + \frac{\Phi(2x, 2x)}{2}\right] \\ &= 2^{n+1}\Phi(x, x).\end{aligned}$$

And we apply Theorem 5.

Approximation of Functions by Quadratic Mappings

We extend now our previous result for quadratic functional equations. First, we recall some known facts about this type of functional equations. For a detailed survey on the fast developing of the field of Hyers–Ulam stability, see the book of Jung [8].

The functional equation

$$f(x + y) + f(x - y) = 2f(x) + 2f(y)$$

is called a *quadratic functional equation*. Every solution of the quadratic functional equation is said to be a quadratic mapping. The Hyers–Ulam stability for quadratic functional equation was proved by Skof [16], for mappings acting between a normed space and a Banach space. Cholewa [2] showed that Skof's Theorem remains true when the normed space is replaced with an abelian group.

Theorem 6 (Skof [16]). *Let $(G, +)$ be an abelian group and let E be a Banach space. If a function $f : G \to E$ satisfies the inequality*

$$\|f(x + y) + f(x - y) - 2f(x) - 2f(y)\| \leq \delta$$

for some $\delta \geq 0$ and for all $x, y \in G$, then there exists a unique quadratic function $Q : G \to E$ such that

$$\|f(x) - Q(x)\| \leq (1/2)\delta,$$

for any $x \in G$.

Let $(G, +)$ be an abelian group and X a Banach space.

Definition 2. We say that f is Φ-approximable by a quadratic map if there exists $Q : G \to X$ quadratic mapping such that

$$\|f(x) - Q(x)\| \le \Phi(x), \quad x \in G. \tag{4}$$

We say that Q is the quadratic Φ approximation of f.

Problem 2. Give conditions on f such that f to be Φ-approximable by a quadratic map.

We denote by

$$\mathscr{Q} = \left\{ \Phi : G \to \mathbb{R}_+ : \lim_{n\to\infty} \frac{\Phi(2^n x)}{4^n} = 0 \right\}.$$

The set $\mathscr{Q}$ is the analogous of the set $\mathscr{A}$ from the case of approximation by additive mappings.

In this case, we have the following characterization of functions which can be approximated by quadratic ones.

Theorem 7. *Let be $Q \in \mathscr{Q}$. Then f is Φ-approximable by a quadratic map if and only if the following two conditions hold*

(i) $\lim_{n\to\infty} \frac{\|f(2^n x + 2^n y) + f(2^n x - 2^n y) - 2f(2^n x) - 2f(2^n y)\|}{4^n} = 0, \ (\forall)\, x, y \in G$

(ii) there exists $\Psi \in \mathscr{Q}$ such that

$$\|f(2^n x) - 4^n f(x)\| \le \Psi(2^n x) + 4^n \Phi(x), x \in G.$$

In this case, the quadratic Φ–approximation of f is unique and is given by

$$Q(x) = \lim_{n\to\infty} \frac{f(2^n x)}{4^n}$$

Proof. First, we assume that f is Φ-approximable by a quadratic map, i.e. the condition (4) holds. Let be $x, y \in G$

We have

$$\|f(x+y) - Q(x+y)\| \le \Phi(x+y)$$

$$\|f(x-y) - Q(x-y)\| \le \Phi(x-y)$$

It follows

$$\begin{aligned} &\|f(x+y) + f(x-y) - 2f(x) - 2f(y)\| \le \\ &\quad \le \|f(x+y) - Q(x+y)\| + \|f(x-y) - Q(x-y)\| \\ &\quad\quad + \|2Q(x) - 2f(x)\| + \|2Q(y) - 2f(y)\| \\ &\quad \le \Phi(x+y) + \Phi(x-y) + 2\Phi(x) + 2\Phi(y) \end{aligned}$$

hence

$$\frac{\|f(2^n(x+y)) + f(2^n(x-y)) - 2f(2^n x) - 2f(2^n y)\|}{4^n} \leq$$
$$\leq \frac{\Phi(2^n(x+y)) + \Phi(2^n(x-y)) + 2\Phi(2^n x) + 2\Phi(2^n y)}{4^n}$$

By letting n go to infinity in the above inequality, we obtain condition (i).

In (4) we put instead of $x \mapsto 2^n x$ to get

$$\|f(2^n x) - Q(2^n x)\| \leq \Phi(2^n x)$$

and we multiply equation (4) by 4^n to get

$$\|4^n f(x) - 4^n Q(x)\| \leq 4^n \Phi(x).$$

So

$$\begin{aligned}\|f(2^n x) - 4^n f(x)\| &= \|f(2^n x) - Q(2^n x) + 4^n Q(x) - 4^n f(x)\| \\ &\leq \|f(2^n x) - Q(2^n x)\| + \|4^n Q(x) - 4^n f(x)\| \\ &\leq \Phi(2^n x) + 4^n \Phi(x)\end{aligned}$$

hence (ii) holds with $\Psi \in \mathcal{Q}$.

Conversely, we suppose that (i) and (ii) holds. From (ii) it follows, for $x \in G$,

$$\left\| \frac{f(2^n x)}{4^n} - f(x) \right\| \leq \frac{\Psi(2^n x)}{4^n} + \Phi(x). \tag{5}$$

But $\left\{ \dfrac{f(2^n x)}{4^n} \right\}$ is a Cauchy sequence. Indeed, by putting instead of $x \mapsto 2^m x$, we get

$$\left\| \frac{f(2^n \cdot 2^m x)}{4^n} - f(2^m x) \right\| \leq \frac{\Psi(2^n \cdot 2^m x)}{4^n} + \Phi(2^m x)$$

and by dividing the above inequality by 4^m, we obtain

$$\left\| \frac{f(2^{n+m} x)}{4^{n+m}} - \frac{f(2^m x)}{4^m} \right\| \leq \frac{\Psi(2^n \cdot 2^m x)}{4^{n+m}} + \frac{\Phi(2^m x)}{4^m}$$

hence

$$\left\| \frac{f(2^{n+m} x)}{4^{n+m}} - \frac{f(2^m x)}{4^m} \right\| \to 0, \ m, n \to \infty$$

Since X is a Banach space it follows that there exists the limit $Q(x) := \lim_{n\to\infty} \frac{f(2^n x)}{4^n}$ and by letting $n \to \infty$ in relation (5) we obtain condition (4).

We prove now that Q is a quadratic mapping. From hypothesis, we know that

$$\lim_{n\to\infty} \left\| \frac{f(2^n(x+y))}{4^n} + \frac{f(2^n(x-y))}{4^n} - \frac{2f(2^n x)}{4^n} - \frac{2f(2^n y)}{4^n} \right\| = 0$$

hence $\|Q(x+y) + Q(x-y) - 2Q(x) - 2Q(y)\| = 0$ so Q is a quadratic mapping.

Now we show that Q is unique. We suppose that Q satisfies

$$\|Q(x) - f(x)\| \le \Phi(x)$$

and there exists Q' which satisfies $\|Q'(x) - f(x)\| \le \Phi(x)$. We apply the norm inequality to get $\|Q(x) - Q'(x)\| \le 2\Phi(x)$. But Q and Q' are quadratic mappings and by putting instead of $x \mapsto 2^n x$ we get

$$\|4^n Q(x) - 4^n Q'(x)\| \le 2\Phi(2^n x).$$

Dividing the above inequality by 4^n we obtain

$$\|Q(x) - Q'(x)\| \le 2 \cdot \frac{\Phi(2^n x)}{4^n}$$

But $\lim_{n\to\infty} \frac{\Phi(2^n x)}{4^n} = 0$ so $Q(x) = Q'(x)$.

From Theorem 7 we have immediately the following result.

Corollary 1 (Borelli and Forti [1]). *Let $(G, +)$ be an abelian group, X a Banach space and $f : G \to X$ a function with $f(0) = 0$ and fulfilling*

$$\|f(x+y) + f(x-y) - 2f(x) - 2f(y)\| \le \varphi(x, y), \tag{6}$$

for all $x, y \in G$. Assume that the series

$$\sum_{n=0}^{\infty} 2^{-2(n+1)} \varphi(2^n x, 2^n x)$$

converges for every x and call $\Phi(x)$ its sum. If for every x, y we have

$$\lim_{n\to\infty} 2^{-2(n+1)} \varphi(2^n x, 2^n y) = 0,$$

then there exists a unique quadratic function $Q : G \to X$ such that

$$\|f(x) - Q(x)\| \le \Phi(x), \; x \in X.$$

Proof. In (6) we take $x = y$:

$$\|f(2x) - 4f(x)\| \leq \varphi(x, x)$$

and in the above inequality $x \mapsto 2x$:

$$\|f(2^2x) - 4f(2x)\| \leq \varphi(2x, 2x).$$

Hence

$$\begin{aligned}\|f(2^2x) - 4^2f(x)\| &\leq \varphi(2x, 2x) + 4\varphi(x, x)\\ &\leq 4^2\Phi(x, x)\end{aligned}$$

Using the induction, we have

$$\|f(2^nx) - 4^nf(x)\| \leq 4^n\Phi(x, x).$$

Hence the condition (*ii*) in Theorem 7 holds. Condition (*i*) holds from hypothesis.

Acknowledgements The work of P. Găvruţa was partially supported by a grant of Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-JRP-RO-FR-2011-2-11-RO-FR/01.03.2013.

References

1. Borelli, C., Forti, G.L.: On a general Hyers-Ulam stability result. Int. J. Math. Math. Sci. **18**, 229–236 (1995)
2. Cholewa, P.W.: Remarks on the stability of functional equations. Aequationes Math. **27**, 76–86 (1984)
3. Czerwik, S.: Stability of Functional Equations of Ulam-Hyers-Rassias Type. Hadronic Press, Palm Harbor (2003)
4. Găvruţa, P.: A generalization of the Hyers-Ulam-Rassias stability of approximately additive mappings. J. Math. Anal. Appl. **184**, 431–436 (1994)
5. Găvruţa, P., Hossu, M., Popescu, D., Căprău, C.: On the stability of mappings and an answer to a problem of Th. M. Rassias. Ann. Math. Blaise Pascal **2**(2), 55–60 (1995)
6. Găvruţa, L., Găvruţa, P.: Ulam stability problem for frames. In: Rassias, Th.M., Brzdek, J. (Eds.) Functional Equations in Mathematical Analysis. Springer Optimization and Its Applications, pp. 139–152. Springer, New York (2012)
7. Hyers, D.H.: On the stability of the linear functional equation. Proc. Natl. Acad. Soc. USA **27**, 222–224 (1941)
8. Jung, S.-M.: Hyers-Ulam-Rassias Stability of Functional Equations in Nonlinear Analysis. Springer, New York (2011)
9. Jung, S.-M., Rassias, M.Th.: A linear functional equation of third order associated to the Fibonacci numbers. Abstr. Appl. Anal. 2014, Article ID 137468, 7 pp (2014)
10. Jung, S.-M., Popa, D., Rassias, M.Th.: On the stability of the linear functional equation in a single variable on complete metric groups. J. Glob. Optim. **59**, 165–171 (2014)

11. Lee, Y.-H., Jung, S.-M., Rassias, M.Th.: On an n-dimensional mixed type additive and quadratic functional equation. Appl. Math. Comput. **228**, 13–16 (2014)
12. Mortici, C., Rassias, M. Th., Jung, S.-M.: On the stability of a functional equation associated with the Fibonacci numbers. Abst. Appl. Anal. 2014, Article ID 546046, 6 pp (2014)
13. Rassias, Th.M: On the stability of the linear mapping in Banach spaces. Proc. Am. Math. Soc. **72**, 297–300 (1978)
14. Ruskai, M.B.: Some connections between frames,mutually unbiased bases, and POVM's in quantum information theory. Acta Appl. Math. **108**(3), 709–719 (2009)
15. Scott, A.J.: Tight informationally complete quantum measurements. J. Phys. A **39**, 13507–13530 (2006)
16. Skof, F.: Proprietá locali e approssimazione di operatori. Rend. Sem. Mat. Fis. Milano **53**, 113–129 (1983)
17. Ulam, S.M.: A Collection of Mathematical Problems. Interscience Publishers, New York (1960)
18. Wolf, M. M.: Quantum Channels & Operations: Guided Tour. Published electronically at http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf (2012)

Bernstein Type Inequalities Concerning Growth of Polynomials

N.K. Govil and Eze R. Nwaeze

Abstract Let $p(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots + a_n z^n$ be a polynomial of degree n, where the coefficients a_j, for $0 \le j \le n$, may be complex, and $p(z) \ne 0$ for $|z| < 1$. Then

$$M(p,R) \le \left(\frac{R^n + 1}{2}\right) ||p||, \quad \text{for } R \ge 1, \tag{1}$$

and

$$M(p,r) \ge \left(\frac{r+1}{2}\right)^n ||p||, \quad \text{for } 0 < r \le 1, \tag{2}$$

where $M(p,R) := \max_{|z|=R\ge 1} |p(z)|$, $M(p,r) := \max_{|z|=r\le 1} |p(z)|$, and $||p|| := \max_{|z|=1} |p(z)|$. Inequality (1) is due to Ankeny and Rivlin (Pac. J. Math. **5**, 849–852, 1955), whereas Inequality (2) is due to Rivlin (Am. Math. Mon. **67**, 251–253, 1960). These inequalities, which due to their applications are of great importance, have been the starting point of a considerable literature in Approximation Theory, and in this paper we study some of the developments that have taken place around these inequalities. The paper is expository in nature and would provide results dealing with extensions, generalizations and refinements of these inequalities starting from the beginning of this subject to some of the recent ones.

Keywords Maximum modulus • Complex polynomials • Restricted zeros • Inequalities • Growth

---

N.K. Govil (✉) • E.R. Nwaeze
Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849, USA
e-mail: govilnk@auburn.edu; ern0002@auburn.edu

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_13

Introduction

Several years after chemist Mendeleev invented the periodic table of elements he made a study of the specific gravity of a solution as a function of the percentage of the dissolved substance, and for this he needed an answer to the following question.

Question If $p(x)$ is a quadratic polynomial with real coefficients and $|p(x)| \leq 1$ on $-1 \leq x \leq 1$, then how large can $|p'(x)|$ be on $-1 \leq x \leq 1$?

To see how an answer to the above question of Mendeleev helped him in the solution of the problem in Chemistry he was interested in, we refer to the paper of Boas [6].

Note that, even though Mendeleev was a chemist, he was able to show that $|p'(x)| \leq 4$ for $-1 \leq x \leq 1$. This estimate is best possible in the sense that there is a quadratic polynomial $p(x) = 1 - 2x^2$ for which $|p(x)| \leq 1$ on $[-1, 1]$ but $|p'(\pm 1)| = 4$. In the general case when $p(x)$ is a polynomial of degree n with real coefficients the problem was solved by Markov [26], who proved the following result which is known as Markov's Theorem.

Theorem 1.1. *Let* $p(x) = \sum_{j=0}^{n} a_j x^j$ *be an algebraic polynomial of degree n such that* $|p(x)| \leq 1$ *for* $x \in [-1, 1]$. *Then*

$$|p'(x)| \leq n^2, \qquad x \in [-1, 1] \tag{3}$$

The inequality is sharp. Equality holds only if $p(x) = \alpha T_n(x)$, *where* α *is a complex number such that* $|\alpha| = 1$, *and*

$$T_n(x) = \cos(n \cos^{-1} x) = 2^{n-1} \prod_{j=1}^{n} \left[x - \cos((j - \frac{1}{2})\pi/n)\right]$$

is the nth degree Tchebycheff polynomial of the first kind. It can be easily verified that $|T_n(x)| \leq 1$ *for* $x \in [-1, 1]$ *and* $|T_n'(1)| = n^2$.

It would be natural to go on and ask for an upper bound for $|p^{(k)}(x)|$ where $1 \leq k \leq n$. Iterating Markov's Theorem yields $|p^{(k)}(x)| \leq n^{2k} L$ if $|p(x)| \leq L$. However, this inequality is not sharp; the best possible inequality was found by Markov's brother, Markov [27], who proved the following.

Theorem 1.2. *Let* $p(x) = \sum_{j=0}^{n} a_j x^j$ *be an algebraic polynomial of degree n with real coefficients such that* $|p(x)| \leq 1$ *for* $x \in [-1, 1]$. *Then*

$$|p^{(k)}(x)| \leq \frac{(n^2 - 1^2)(n^2 - 2^2) \cdots (n^2 - (k-1)^2)}{1 \cdot 3 \cdots (2k-1)}, \qquad x \in [-1, 1]. \tag{4}$$

The inequality is sharp, and the equality holds again only for $p(x) = T_n(x)$*, where* $T_n(x) = \cos(n \cos^{-1} x)$ *is the Chebyschev polynomial of degree n.*

Several years later, around 1926, Serge Bernstein needed the analogue of the above result Theorem 1.1 of A. A. Markov for polynomials in the complex domain and proved the following, which in the literature is known as Bernstein's Inequality.

Theorem 1.3. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a complex polynomial of degree at most n. Then*

$$\max_{|z|=1} |p'(z)| \le n \max_{|z|=1} |p(z)|. \tag{5}$$

The inequality is best possible and equality holds only for polynomials of the form $p(z) = \alpha z^n$, $\alpha \neq 0$ *being a complex number.*

The above theorem is, in fact, a special case of a more general result due to Riesz [35] for trigonometric polynomials.

For the sake of brevity, throughout in this paper, we shall be using the following notations.

Definition 1. Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a complex polynomial of degree at most n. We will denote

$$M(p, r) := \max_{|z|=r} |p(z)|, \ r > 0,$$

$$||p|| := \max_{|z|=1} |p(z)|,$$

and

$$D(0, K) := \{z : |z| < K\}, \ K > 0.$$

In 1945, S. Bernstein initiated and observed the following result, which in fact is a simple consequence of the maximum modulus principle (see [29] or [35, volume 1, p. 137]). This inequality is also known as the Bernstein's Inequality.

Theorem 1.4. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n. Then for* $R \ge 1$*,*

$$M(p, R) \le R^n ||p||. \tag{6}$$

Equality holds for $p(z) = \alpha z^n$*,* α *being a complex number.*

If one applies the above inequality to the polynomial $P(z) = z^n p(1/z)$ and use maximum modulus principle, one easily gets

Theorem 1.5. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n. Then for* $0 < r \leq 1$,

$$M(p,r) \geq r^n ||p||. \tag{7}$$

Equality holds for $p(z) = \alpha z^n$, α *being a complex number.*

The above result is due to Varga [38] who attributes it to E. H. Zarantonello.

By use of the transformation $P(z) = z^n p(1/z)$ and the maximum modulus principle it is not difficult to see that Theorems 1.4 and 1.5 can be obtained from each other. The fact that Theorem 1.3 can be obtained from Theorem 1.4 was proved by Bernstein himself. However, it was not known if Theorem 1.4 can also be obtained from Theorem 1.3, and this has been shown by Govil et al. [22]. Thus all the above three Theorems 1.3, 1.4 and 1.5 are equivalent in the sense that anyone can be obtained from any of the others.

For the sharpening of Theorems 1.3, 1.4 and 1.5 we refer the reader to the paper of Frappier et al. [12] (also, see Sharma and Singh [37]).

For polynomial of degree n not vanishing in the interior of the unit circle, Ankeny and Rivlin [1] proved the following result.

Theorem 1.6. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $R \geq 1$,

$$M(p,R) \leq \left(\frac{R^n + 1}{2}\right) ||p||. \tag{8}$$

Here equality holds for $p(z) = \dfrac{\alpha + \beta z^n}{2}$, *where* $|\alpha| = |\beta| = 1$.

For some generalizations of inequalities (6) and (8), see Govil et al. [23].

The analogue of Inequality (7) for polynomials not vanishing in the interior of a unit circle was proved later in 1960 by Rivlin [36], who in fact proved

Theorem 1.7. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $0 < r \leq 1$,

$$M(p,r) \geq \left(\frac{r+1}{2}\right)^n ||p||, \tag{9}$$

and equality holding for $p(z) = \left(\dfrac{\alpha + \beta z}{2}\right)^n$, *where* $|\alpha| = |\beta| = 1$.

The above results, Theorems 1.4–1.5 which are known as Bernstein inequalities concerning growth of polynomials, and Theorems 1.6–1.7 have been the starting point of a considerable literature in Approximation Theory. Several books and research monographs have been written on this subject of inequalities (see, for example, Govil and Mohapatra [20], Milovanović et al. [28], Pinkus [30], and Rahman and Schmeisser [34]) and in this chapter we study some of the developments that have, over a period, taken place around these inequalities.

The present chapter is expository in nature and consists of four sections. Having introduced Bernstein type inequalities concerning the growth of polynomials, Theorems 1.4–1.7 in section "Introduction," we in section "Results Concerning Generalizations, Extensions and Refinements of Theorem 1.6" study inequalities related to the inequality in Theorem 1.6 while section "Results Concerning Generalizations, Extensions and Refinements of Theorem 1.7" will consist of inequalities related to Theorem 1.7. Lastly, the section "Polynomials Having all the Zeros on $S(0, K)$" deals with inequalities concerning the growth of polynomials having all their zeros on a circle.

Results Concerning Generalizations, Extensions and Refinements of Theorem 1.6

We begin with Theorem 1.6 stated in section "Introduction," and which is due to Ankeny and Rivlin [1].

Theorem 2.1. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0, 1)$. *Then for* $R \geq 1$,

$$M(p, R) \leq \left(\frac{R^n + 1}{2}\right)||p||. \tag{10}$$

Equality holding for $p(z) = \dfrac{\alpha + \beta z^n}{2}$, *where* $|\alpha| = |\beta| = 1$.

We present here the brief outlines of the proof of Theorem 2.1 as given by Ankeny and Rivlin in [1], which makes use of Erdös-Lax Theorem. As is well known, the Erdös-Lax Theorem which is stated below as Lemma 2.2, was conjectured by Erdös and proved by Lax [25]. It may be remarked that a simpler proof of Lemma 2.2 was provided by Aziz and Mohammad [3].

Lemma 2.2 (Lax [25]). *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0, 1)$. *Then*

$$M(p', 1) \leq \frac{n}{2}||p||. \tag{11}$$

Proof (Proof of Theorem 2.1). Let us assume that $p(z)$ does not have the form $\dfrac{\alpha + \beta z^n}{2}$. In view of Lemma 2.2

$$|p'(e^{i\theta})| \le \frac{n}{2}||p||, \qquad 0 \le \theta < 2\pi, \tag{12}$$

from which we may deduce that

$$|p'(re^{i\theta})| < \frac{n}{2}r^{n-1}||p||, \qquad 0 \le \theta < 2\pi, \quad r > 1, \tag{13}$$

by applying Theorem 1.4 to the polynomial $p'(z)/(n/2)$ and observing that we have the strict inequality in (13) because $p(z)$ does not have the form $\dfrac{\alpha+\beta z^n}{2}$. But for each $\theta, 0 \le \theta < 2\pi$, we have

$$p(Re^{i\theta}) - p(e^{i\theta}) = \int_1^R e^{i\theta} p'(re^{i\theta})dr.$$

Hence

$$\left|p(Re^{i\theta}) - p(e^{i\theta})\right| \le \int_1^R |p'(re^{i\theta})|dr < \frac{n}{2}||p|| \int_1^R r^{n-1}dr = \frac{||p||}{2}(R^n - 1),$$

and

$$\left|p(Re^{i\theta})\right| < \frac{||p||}{2}(R^n - 1) + \left|p(e^{i\theta})\right| \le \frac{||p||}{2}(1 + R^n).$$

Finally, if $p(z) = \dfrac{\alpha+\beta z^n}{2}$, $|\alpha| = |\beta| = 1$, then clearly

$$M(p, R) = \frac{1+R^n}{2}, \; R > 1,$$

and the proof of Theorem 2.1 is thus complete.

It may be remarked that later a simpler proof of Theorem 2.1 which does not make use of Erdös-Lax Theorem was given by Dewan [7].

Remark 1. The converse of Theorem 2.1 is false as the simple example $p(z) = (z+\frac{1}{2})(z+3)$ shows. However, the following result in the converse direction, which is also due to Ankeny and Rivlin [1], is valid.

Theorem 2.3. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n such that* $p(1) = 1$ *and*

$$M(p, R) \le \left(\frac{R^n + 1}{2}\right)||p||$$

for $0 < R - 1 < \delta$*, where* δ *is any positive number. Then* $p(z)$ *does not have all its zeros within the unit circle.*

In 1989, Govil [17] observed that since the equality in (10) holds only for polynomials $p(z) = \alpha + \beta z^n,\ |\alpha| = |\beta|$, which satisfy

$$\left|\text{coefficient of } z^n\right| = \frac{1}{2}||p||, \tag{14}$$

it should be possible to improve upon the bound in (10) for polynomials not satisfying (14), and therefore in this connection he proved the following refinement of (10).

Theorem 2.4. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$*. Then for* $R \geq 1$*,*

$$M(p,R) \leq \left(\frac{R^n+1}{2}\right)||p|| - \frac{n(||p||^2 - 4|a_n|^2)}{2||p||}\left\{\frac{(R-1)||p||}{||p|| + 2|a_n|} - \ln\left[1 + \frac{(R-1)||p||}{||p|| + 2|a_n|}\right]\right\} \tag{15}$$

Equality holding for $p(z) = \alpha + \beta z^n$*, where* $|\alpha| = |\beta|$*.*

In 1998, Dewan and Bhat [9] sharpened the above Theorem 2.4 as follows:

Theorem 2.5. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$*. Then for* $R \geq 1$*,*

$$\begin{aligned} M(p,R) \leq {} & \left(\frac{R^n+1}{2}\right)||p|| - \left(\frac{R^n-1}{2}\right)m \\ & - \frac{n}{2}\left[\frac{(||p||-m)^2 - 4|a_n|^2}{(||p||-m)}\right]\left\{\frac{(R-1)(||p||-m)}{(||p||-m)+2|a_n|}\right. \\ & \left. - \ln\left[1 + \frac{(R-1)(||p||-m)}{(||p||-m)+2|a_n|}\right]\right\}, \end{aligned} \tag{16}$$

where $m = \min_{|z|=1}|p(z)|$*. Here again, equality holds for* $p(z) = \alpha + \beta z^n$*, where* $|\alpha| = |\beta|$*.*

In 2001, Govil and Nyuydinkong [21] generalized Theorem 2.5, where they considered polynomials not vanishing in $D(0,K),\ K \geq 1$. More specifically, they proved

Theorem 2.6. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,K),\ K \geq 1$. *Then for* $R \geq 1$,

$$M(p,R) \leq \Big(\frac{R^n + K}{1+K}\Big)||p|| - \Big(\frac{R^n - 1}{1+K}\Big)m - \frac{n}{1+K}\left[\frac{(||p||-m)^2 - (1+K)^2|a_n|^2}{(||p||-m)}\right]$$
$$\times\left\{\frac{(R-1)(||p||-m)}{(||p||-m)+(1+K)|a_n|} - \ln\left[1 + \frac{(R-1)(||p||-m)}{(||p||-m)+(1+K)|a_n|}\right]\right\}, \tag{17}$$

where $m = \min_{|z|=K} |p(z)|$.

Following immediately, Gardner et al. [14] generalized Theorem 2.6 by considering polynomials of the form $a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \leq t \leq n$, and for this, they proved the following:

Theorem 2.7. *Let* $p(z) = a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \leq t \leq n$, *be a polynomial of degree* n *and* $p(z) \neq 0$ *in* $D(0,K),\ K \geq 1$. *Then for* $R \geq 1$,

$$M(p,R) \leq \Big(\frac{R^n + K^t}{1+K^t}\Big)||p|| - \Big(\frac{R^n - 1}{1+K^t}\Big)m - \frac{n}{1+K^t}\left[\frac{(||p||-m)^2 - (1+K^t)^2|a_n|^2}{(||p||-m)}\right]$$
$$\times\left\{\frac{(R-1)(||p||-m)}{(||p||-m)+(1+K^t)|a_n|} - \ln\left[1 + \frac{(R-1)(||p||-m)}{(||p||-m)+(1+K^t)|a_n|}\right]\right\}, \tag{18}$$

where $m = \min_{|z|=K} |p(z)|$.

Clearly, for $t = 1$, Theorem 2.7 gives Theorem 2.6, which for $K = 1$ reduces to Theorem 2.5.

In 2005, Gardner et al. [15] proved the following generalization and sharpening of Theorem 2.4.

Theorem 2.8. *Let* $p(z) = a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \leq t \leq n$, *be a polynomial of degree* n *and* $p(z) \neq 0$ *in* $D(0,K),\ K \geq 1$. *Then for* $R \geq 1$,

$$M(p,R) \leq \left(\frac{R^n + s_0}{1+s_0}\right)||p|| - \left(\frac{R^n - 1}{1+s_0}\right)m - \frac{n}{1+s_0}\left[\frac{(||p||-m)^2 - (1+s_0)^2|a_n|^2}{(||p||-m)}\right]$$
$$\times\left\{\frac{(R-1)(||p||-m)}{(||p||-m)+(1+s_0)|a_n|} - \ln\left[1 + \frac{(R-1)(||p||-m)}{(||p||-m)+(1+s_0)|a_n|}\right]\right\}, \tag{19}$$

where $m = \min_{|z|=K} |p(z)|$, *and*

$$s_0 = K^{t+1}\frac{\frac{t}{n}\cdot\frac{|a_t|}{|a_0|-m}K^{t-1}+1}{\frac{t}{n}\cdot\frac{|a_t|}{|a_0|-m}K^{t+1}+1}.$$

Dividing both sides of (19) by R^n, and letting $R \to \infty$, one gets

Corollary 2.9. *Let* $p(z) = a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \leq t \leq n$, *be a polynomial of degree* n *and* $p(z) \neq 0$ *in* $D(0,K),\ K \geq 1$. *Then*

$$|a_n| \leq \frac{1}{1+s_0}\big(||p||-m\big), \tag{20}$$

where $m = \min_{|z|=K} |p(z)|$.

In case one does not have knowledge of $m = \min_{|z|=K} |p(z)|$, one could use the following result due to Gardner et al. [15] which does not depend on m.

Theorem 2.10. *Let* $p(z) = a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \leq t \leq n$, *be a polynomial of degree* n *and* $p(z) \neq 0$ *in* $D(0,K),\ K \geq 1$. *Then for* $R \geq 1$,

$$M(p,R) \leq \left(\frac{R^n + s_1}{1+s_1}\right)||p|| - \frac{n}{1+s_1}\left[\frac{||p||^2 - (1+s_1)^2|a_n|^2}{||p||}\right]$$
$$\times\left\{\frac{(R-1)||p||}{||p||+(1+s_1)|a_n|} - \ln\left[1 + \frac{(R-1)||p||}{||p||+(1+s_1)|a_n|}\right]\right\}, \tag{21}$$

where

$$s_1 = K^{t+1}\frac{\frac{t}{n}\cdot\frac{|a_t|}{|a_0|}K^{t-1}+1}{\frac{t}{n}\cdot\frac{|a_t|}{|a_0|}K^{t+1}+1}.$$

If, in the above theorem, one divides both sides of (21) by R^n and let $R \to \infty$, one obtains the following

Corollary 2.11. *Let* $p(z) = a_0 + \sum_{j=t}^{n} a_j z^j,\ 1 \le t \le n$, *be a polynomial of degree* n *and* $p(z) \neq 0$ *in* $D(0, K),\ K \geq 1$. *Then*

$$|a_n| \leq \frac{1}{1+s_1}||p||. \tag{22}$$

Both Corollaries 2.9 and 2.11 generalize and sharpen the well-known inequality, obtainable by an application of Visser's Inequality [39], that if $p(z) = \sum_{j=0}^{n} a_j z^j$ is a polynomial of degree n and $p(z) \neq 0$ in $D(0, 1)$ then $|a_n| \leq \frac{1}{2}||p||$.

We present some of the examples Gardner et al. [15] gave to illustrate the quality of Theorems 2.7, 2.8 and 2.10.

Example 1. Let $p(z) = 1000 + z^2 + z^3 + z^4$. Clearly, here $t = 2$ and $n = 4$, and one can take $K = 5.4$, since numerically $p \neq 0$ for $|z| < 5.4483$. For this polynomial, the bound for $M(p, 2)$ by Theorem 2.7 comes out to be 1447.503, and by Theorem 2.8, it comes out to be 1101.84, which is a significant improvement over the bound obtained from Theorem 2.7. Numerically, for this polynomial $M(p, 2) \approx 1028$, which is quite close to the bound 1101.84, that is obtainable by Theorem 2.10. The bound for $M(p, 2)$ obtained by Theorem 2.10 is 1105.05, which is also quite close to the actual bound ≈ 1028. However, in this case Theorem 2.8 gives the best bound.

Example 2. Let $p(z) = 1000 + z^2 - z^3 - z^4$. Here also, $t = 2$ and $n = 4$. Again, numerically $p(z) \neq 0$ for $|z| < 5.43003$, and thus take $K = 5.4$. If $R = 3$, then for this polynomial the bound for $M(p, 3)$ obtained by Theorem 2.7 comes out to be 3479.408, while by Theorem 2.10 it comes out to be 1545.3, and by Theorem 2.8 it comes out to be 1534.5, a considerable improvement. Thus again the bounds obtained from Theorem 2.8 and Theorem 2.10 are considerably smaller than the bound obtained from Theorem 2.7, and the bound 1534.5 obtained by Theorem 2.8 is much closer to the actual bound $M(p, 3) \approx 1100.6$ than the bound 3479.408, obtained from Theorem 2.7.

In 1981, Aziz and Mohammad [4] sharpened Theorem 2.1 by proving the following.

Theorem 2.12. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n and* $p(z) \neq 0$ *in* $D(0, K), K \geq 1$. *Then for* $R \geq 1$,

$$M(p, R) \leq \frac{(R^n + 1)(R + K)^n}{(R + K)^n + (1 + RK)^n} ||p|| \tag{23}$$

Theorem 2.12 is a generalization of Theorem 2.1 in a compact form but unfortunately with the exception of $n = 1$, the Inequality (23) does not appear to be sharp for $K > 1$. However, a precise estimate is given by the following theorem, which is also due to Aziz and Mohammad [4].

Theorem 2.13. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n and* $p(z) \neq 0$ *in* $D(0, K), K \geq 1$. *Then for* $R \geq 1$,

$$M(p, R) \leq \frac{(R + K)^n}{(1 + K)^n} ||p|| \text{ for } 1 \leq R \leq K^2 \tag{24}$$

and

$$M(p, R) \leq \frac{R^n + K^n}{1 + K^n} ||p|| \text{ for } R \geq K^2 \tag{25}$$

The result is best possible with equality in (25) for $p(z) = (z^n + K^n)/(1 + K^n)$ *and in (24) for* $p(z) = (z + K)^n/(1 + K)^n$.

The following theorem which is due to Govil et al. [22] sharpens Inequality (24) in the above Theorem 2.13.

Theorem 2.14. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0, K),\ K \geq 1$. *Then for* $1 \leq R \leq K^2$,

$$M(p, R) \leq \left(\frac{K^2 + 2R|\lambda|K + R^2}{K^2 + 2|\lambda|K + 1} \right)^{n/2} M(p, 1), \tag{26}$$

where $\lambda = \lambda(K) := \dfrac{Ka_1}{na_0}$.

The fact that the Inequality (26) sharpens Inequality (24) follows because if $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ for $|z| < K$, where $K \geq 1$, then $|\lambda| = |Ka_1/na_0| \leq 1$, and therefore

$$\left(\frac{K^2 + 2R|\lambda|K + R^2}{K^2 + 2|\lambda|K + 1}\right) \le \left(\frac{K^2 + 2RK + R^2}{K^2 + 2K + 1}\right) = \left(\frac{R+K}{K+1}\right)^2,$$

from which the conclusion follows.

In the case $R \ge K^2$, Govil et al. [22] proved

Theorem 2.15. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,K),\ K > 1$. *Then for* $R \ge K^2$,

$$M(p,R) \le \frac{R^n}{K^n}\left(\frac{K^n}{K^n+1}\right)^{(R-K^2)/(R+K^2)} M(p,1). \tag{27}$$

About 6 years after Aziz and Mohammad [4] proved Theorem 2.13, Aziz [2] obtained more results in this direction, some of which are presented below.

Theorem 2.16. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n and* $p(z) \neq 0$ *in* $D(0,K), K \ge 1$. *Then for* $R \ge 1$,

$$M(p,R) \le \frac{R^n+1}{2}||p|| - \frac{R^n-1}{2}\min_{|z|=1}|p(z)|. \tag{28}$$

The result is best possible and equality in (28) holds for the polynomial $p(z) = \alpha z^n + \beta K^n,\ |\alpha| = |\beta| = 1,\ K \ge 1$.

As an application of Theorem 2.16, Aziz [2] established

Theorem 2.17. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n and* $p(z) \neq 0$ *in* $D(0,K), K \le 1$. *Then for* $0 \le r \le K$,

$$(1+r^n)M(p,r) - (1-r^n)\min_{|z|=1}|p(z)| \ge 2r^n||p||. \tag{29}$$

The result is best possible and equality in (29) holds for the polynomial $p(z) = \alpha z^n + \beta K^n$, *where* $|\alpha| = |\beta| = 1$, *and* $K \le 1$.

In 2002, Aziz and Zargar [5] proved the following refinement of Theorem 2.12 which includes Theorem 2.16 as a special case.

Theorem 2.18. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n and* $p(z) \neq 0$ *in* $D(0,K), K \ge 1$. *Then for* $R \ge 1$,

$$M(p,R) \le F(R)\left\{(R^n+1)||p|| - \left[R^n - \left(\frac{1+RK}{R+K}\right)^n\right]m\right\}, \tag{30}$$

where $m := \min_{|z|=K} |p(z)|$ *and* $F(R) := \dfrac{(R+K)^n}{(R+K)^n + (1+RK)^n}$.

For $K = 1$, the above theorem reduces to Theorem 2.16. In the same paper, they [5] proved the following which is a refinement of (24).

Theorem 2.19. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree* n *and* $p(z) \ne 0$ *in* $D(0,K), K \ge 1$. *Then for* $1 \le R \le K^2$,

$$M(p,R) \le \left(\frac{R+K}{1+K}\right)^n ||p|| - \left\{\left(\frac{R+K}{1+K}\right)^n - 1\right\}m, \tag{31}$$

where $m := \min_{|z|=K} |p(z)|$.

Theorem 2.18, in some special cases, can provide much better information than Theorem 2.12 regarding $M(p,R)$, $R > 1$, and thus they [5] illustrate with the help of the following examples.

Example 3. Let $p(z) = (z^2+9)(z-19)$. Then $p(z)$ is a polynomial of degree 3 which does not vanish in $D(0,t)$, where $t \in (0,3]$. Clearly

$$|p(z)| \ge \left(9 - |z|^2\right)\left(19 - |z|\right)$$

which in particular gives $M(p,2) \ge 85$ and $||p|| = 200$. Using Theorem 2.12 with $K = t = 3,\ R = 2$, it follows that

$$M(p,2) \le 480.8, \tag{32}$$

whereas using Theorem 2.18 with $K = 2$, and $R = 2$, we get

$$M(p,2) \le 435.5, \tag{33}$$

which is much better than (32).

Example 4. Let $p(z) = z^3 + 3^3$. Then $p(z)$ does not vanish in $D(0,t)$, where $t \in (0,3]$. Evidently,

$$M(p,2) \ge 19 \quad \text{and} \quad M(p,1) = 28. \tag{34}$$

Using Theorem 2.12 with $K = t = 3,\ R = 2$, it follows that

$$M(p,2) \le 67.4. \tag{35}$$

Using Theorem 2.18 with $K = t = 2,\ R = 2$, gives

$$M(p,2) \leq 46.5, \tag{36}$$

which is much better than (35).

For more results in this direction we refer the reader to [18, 19, 33, 37]. We wrap up this part by presenting the following results where the maximum modulus is taken on an ellipse rather than on a circle.

Theorem 2.20 (Duffin and Schaeffer [10]). *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n such that* $\max_{-1\leq x\leq 1} |p(x)| \leq 1$ *and* $p(z)$ *is real for real* z,*. Then for* $R > 1$,

$$\max_{z\in\mathscr{E}_{\mathscr{R}}} |p(z)| \leq \frac{R^n + R^{-n}}{2},$$

where $\mathscr{E}_{\mathscr{R}} := \left\{ z = x + iy : \frac{x^2}{\left(\frac{R+R^{-1}}{2}\right)^2} + \frac{y^2}{\left(\frac{R-R^{-1}}{2}\right)^2} = 1 \right\}$

In the following result the hypothesis on the polynomial $p(z)$ that $p(z)$ is real for real z has been dropped.

Theorem 2.21 (Frappier and Rahman [11]). *If* $p(z) = \sum_{j=0}^{n} a_j z^j$ *is a polynomial of degree n such that* $\max_{-1\leq x\leq 1} |p(x)| \leq 1$, *then for* $R > 1$, *we have*

$$\max_{z\in\mathscr{E}_{\mathscr{R}}} |p(z)| \leq \frac{R^n}{2} + \frac{5+\sqrt{17}}{4} R^{n-2},$$

where $\mathscr{E}_{\mathscr{R}}$ *is the same as above.*

Results Concerning Generalizations, Extensions and Refinements of Theorem 1.7

So far we have been dealing with improvements and generalizations of Inequality (10), and now we turn our attention to Inequality (9), given in Theorem 1.7. In this regard, Govil [16] generalized this Theorem 1.7 by proving

Theorem 3.1. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $0 < r \leq \rho \leq 1$,

$$M(p,r) \geq \left(\frac{1+r}{1+\rho}\right)^n M(p,\rho). \tag{37}$$

The result is best possible and equality holds for the polynomial $p(z) = \left(\frac{1+z}{1+\rho}\right)^n$.

If polynomial $p(z)$ has all its zeros on $|z| = 1$, the polynomial $q(z) = z^n p(\frac{1}{z})$ also has its zeros on $|z| = 1$. Further, if $1 \leq \rho \leq r$, then $\frac{1}{r} \leq \frac{1}{\rho} \leq 1$, and when (37) is applied to $q(z)$, it yields

$$M\Big(q, \frac{1}{r}\Big) \geq \left(\frac{1+\frac{1}{r}}{1+\frac{1}{\rho}}\right)^n M\Big(q, \frac{1}{\rho}\Big),$$

which is equivalent to (37).

The above explanation thus leads to the following corollary.

Corollary 2.2. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n having all its zeros on the unit circle. Then for* $0 < r \leq \rho \leq 1$, *and for* $1 \leq \rho \leq r$,

$$M(p,r) \geq \left(\frac{1+r}{1+\rho}\right)^n M(p,\rho). \tag{38}$$

The result is best possible and equality holds for the polynomial $p(z) = (1+z)^n$.

If in Theorem 3.1 one also assumes that $p'(0) = 0$, the bound in (37) can be considerably improved. Govil [16] in the same paper obtained the following in this direction.

Theorem 3.2. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Let* $p'(0) = 0$. *Then for* $0 < r \leq \rho \leq 1$,

$$M(p,r) \geq \left(\frac{1+r}{1+\rho}\right)^n \left\{\frac{1}{1 - \frac{(1-\rho)(\rho-r)n}{4}\left(\frac{1+r}{1+\rho}\right)^{n-1}}\right\} M(p,\rho). \tag{39}$$

Theorem 3.1 is best possible, however, if $0 < r < \rho < 1$, then for any polynomial $p(z)$ having no zeros in $D(0,1)$, and $p'(0) = 0$, the bound obtained by Theorem 3.2 can be considerably sharper than the bound obtained by Theorem 3.1. Govil [16] illustrated this by means of the following examples.

Example 5. Let $p(z) = 1 + z^3$, $\rho = 0.5$, $r = 0.1$. Theorem 3.1 gives $M(p,r) \geq (0.3943704)M(p,\rho)$, while by Theorem 3.2, $M(p,r) \geq (0.4289743)M(p,\rho)$.

Example 6. Let $p(z) = 1+z^7$, $\rho = 0.168$, $r = 0.022$. Theorem 3.1 gives $M(p,r) \geq (0.3926959)M(p,\rho)$, while by Theorem 3.2, $M(p,r) \geq (0.4341115)M(p,\rho)$.

In 1992, Qazi [32] extended Theorem 3.1 to polynomials with gaps. Specifically he proved

Theorem 3.4. *Let* $1 \leq m \leq n$ *and* $p(z) = a_0 + \sum_{j=m}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $0 < r < \rho \leq 1$,

$$M(p,r) \geq \left(\frac{1+r^m}{1+\rho^m}\right)^{n/m} M(p,\rho); \tag{40}$$

more precisely

$$M(p,r) \geq \exp\left(-n\int_r^\rho \frac{t^m + (m/n)|a_m/a_0|t^{m-1}}{t^{m+1} + (m/n)|a_m/a_0|(t^m+t)+1}dt\right)M(p,\rho). \tag{41}$$

If $m = 1$, Theorem 3.4 gives

Corollary 2.5. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $0 < r < \rho \leq 1$,

$$M(p,r) \geq \left(\frac{1+2|\lambda|r+r^2}{1+2|\lambda|\rho+\rho^2}\right)^{n/2} M(p,\rho), \tag{42}$$

where $\lambda := \dfrac{a_1}{na_0}$.

Proof. It is easy to see that

$$\exp\left(-n\int_r^\rho \frac{t+(1/n)|a_1/a_0|}{t^2+(2/n)|a_1/a_0|t+1}dt\right) = \left(\frac{1+2|a_1/na_0|r+r^2}{1+2|a_1/na_0|\rho+\rho^2}\right)^{n/2},$$

and on applying Inequality (41), the corollary follows.

For $\rho = 1$, the above corollary reduces to

Corollary 2.6. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0,1)$. *Then for* $0 < r < 1$,

$$M(p,r) \geq \left(\frac{1+2|\lambda|r+r^2}{2+2|\lambda|}\right)^{n/2} M(p,1), \tag{43}$$

where $\lambda := \dfrac{a_1}{na_0}$.

In 2003, Govil et al. [22] extended Corollary 2.6 to polynomials not vanishing in $D(0, K),\ K \geq 1$. They, in fact, proved

Theorem 3.7. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0, K),\ K \geq 1$. *Then for* $0 < r < 1$,

$$M(p, r) \geq \left(\frac{K^2 + 2r|\lambda|K + r^2}{K^2 + 2|\lambda|K + 1} \right)^{n/2} M(p, 1), \tag{44}$$

where $\lambda = \lambda(K) := \dfrac{Ka_1}{na_0}$.

In the case where n is even, (44) becomes an equality for polynomials of the form $c(K^2 + 2Kze^{i\beta}\cos\alpha + z^2 e^{2i\beta})^{n/2},\ c \in \mathbf{C},\ c \neq 0,\ \alpha \in \mathbf{R},\ \beta \in \mathbf{R}$.

For any n, Inequality (44) may be replaced by

$$M(p, r) \geq \left(\frac{r + K}{K + 1} \right)^{n} M(p, 1), \quad 0 < r < 1, \tag{45}$$

where the bound is attained if $p(z) = c(ze^{i\beta} + K)^n,\ c \in \mathbf{C},\ c \neq 0,\ \beta \in \mathbf{R}$. It may be noted that even (45) is a generalization of (9).

Assuming that $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ in $D(0, K),\ K \leq 1$, Govil et al. [22] proved the following result.

Theorem 3.8. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j \neq 0$ *in* $D(0, K),\ K \leq 1$. *Then for* $0 < r \leq K^2$,

$$M(p, r) \geq \left(\frac{K^2 + 2r|\lambda|K + r^2}{K^2 + 2|\lambda|K + 1} \right)^{n/2} M(p, 1), \tag{46}$$

where $\lambda = \lambda(K) := \dfrac{Ka_1}{na_0}$.

In the case where n is even, (46) becomes an equality for polynomials of the form $c(K^2 + 2Kze^{i\beta}\cos\alpha + z^2 e^{2i\beta})^{n/2},\ c \in \mathbf{C},\ c \neq 0,\ \alpha \in \mathbf{R},\ \beta \in \mathbf{R}$.

For any n, Inequality (46) may be replaced by

$$M(p, r) \geq \left(\frac{r + K}{K + 1} \right)^{n} M(p, 1), \quad 0 < r \leq K^2, \tag{47}$$

where the bound is attained if $p(z) = c(ze^{i\beta} + K)^n,\ c \in \mathbf{C},\ c \neq 0,\ \beta \in \mathbf{R}$.

Inequality (47) extends and refines a result of Jain [24, Inequality (1.4)], who had obtained it under the assumption that all the zeros of p lie on the circle $S(0,K) := \{z : |z| = K\},\ K > 0$.

Polynomials Having all the Zeros on $S(0,K)$

While trying to obtain inequality analogous to (10) for polynomials not vanishing in $D(0,K), K \leq 1$, Dewan and Ahuja [8] were able to prove this only for polynomials having all the zeros on the circle $S(0,K) := \{z : |z| = K\},\ 0 < K \leq 1$.

Theorem 4.1. *Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n having all its zeros on $S(0,K),\ K \leq 1$. Then for $R \geq 1$ and for every positive integer s,*

$$\{M(p,R)\}^s \leq \left[\frac{K^{n-1}(1+K) + (R^{ns}-1)}{K^{n-1}+K^n}\right]\{M(p,1)\}^s. \tag{48}$$

For $s = 1$, the Theorem 4.1 yields

Corollary 4.2. *Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n having all its zeros on $S(0,K),\ K \leq 1$. Then for $R \geq 1$,*

$$M(p,R) \leq \left[\frac{K^{n-1}(1+K) + (R^{n}-1)}{K^{n-1}+K^n}\right]M(p,1). \tag{49}$$

The following corollary immediately follows from Inequality (49), if one takes $K = 1$.

Corollary 4.3. *Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n having all its zeros on $S(0,1)$. Then for $R \geq 1$,*

$$M(p,R) \leq \left(\frac{R^n+1}{2}\right)M(p,1). \tag{50}$$

In the same paper, Dewan and Ahuja [8] also proved

Theorem 4.4. *Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n having all its zeros on $S(0,K),\ K \leq 1$. Then for $R \geq 1$ and for every positive integer s,*

$$\{M(p,R)\}^s \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^{ns}-1)\}+|a_{n-1}|\{2K^n+R^{ns}-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right] \{M(p,1)\}^s. \tag{51}$$

The following example illustrates that in some case the bound obtained in Theorem 4.4 is considerably better than the bound obtained in Theorem 4.1.

Example 7. Let $p(z) = z^4 - \frac{1}{50}z^2 + \frac{1}{100^2}$ and $K = \frac{1}{10}$, $R = 1.5$ and $s = 2$. Using Theorem 4.1, one gets

$$\{M(p,R)\}^s \leq 22390.91477\{M(p,1)\}^s,$$

while Theorem 4.4 gives

$$\{M(p,R)\}^s \leq 2439.505569\{M(p,1)\}^s,$$

showing that the bound obtained by Theorem 4.4 can be considerably sharper than what one gets from Theorem 4.1.

For $s = 1$ in Theorem 4.4 yields

Corollary 4.5. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree* n *having all its zeros on* $S(0,K)$, $K \leq 1$. *Then for* $R \geq 1$,

$$M(p,R) \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^n-1)\}+|a_{n-1}|\{2K^n+R^n-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right]M(p,1). \tag{52}$$

By restricting ourselves to polynomials of degree $n \geq 2$, Pukhta [30] obtained an improvement and generalization of Theorem 4.1 and Theorem 4.4. More precisely, he proved

Theorem 4.6. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree* n *having all its zeros on* $S(0,K)$, $K \leq 1$. *Then for* $R \geq 1$ *and for every positive integer* s,

$$\{M(p,R)\}^s \leq \left[\frac{K^{n-1}(1+K)+(R^{ns}-1)}{K^{n-1}+K^n}\right]\{M(p,1)\}^s - s|a_1|\left[\frac{R^{ns}-1}{ns} - \frac{R^{ns-2}-1}{ns-2}\right]\{M(p,1)\}^{s-1}, \tag{53}$$

for $n > 2$*; and*

$$\{M(p,R)\}^s \le \left[\frac{K^{n-1}(1+K)+(R^{ns}-1)}{K^{n-1}+K^n}\right]\{M(p,1)\}^s \tag{54}$$
$$-s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-1}-1}{ns-1}\right]\{M(p,1)\}^{s-1},$$

for $n = 2$.

Setting $s = 1$ in Theorem 4.6 reduces to

Corollary 4.7. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n having all its zeros on* $S(0,K)$, $K \le 1$. *Then for* $R \ge 1$,

$$M(p,R) \le \left[\frac{K^{n-1}(1+K)+(R^{n}-1)}{K^{n-1}+K^n}\right]M(p,1) \tag{55}$$
$$-|a_1|\left[\frac{R^{n}-1}{n}-\frac{R^{n-2}-1}{n-2}\right],$$

for $n > 2$*; and*

$$M(p,R) \le \left[\frac{K^{n-1}(1+K)+(R^{n}-1)}{K^{n-1}+K^n}\right]M(p,1) \tag{56}$$
$$-|a_1|\left[\frac{R^{n}-1}{n}-\frac{R^{n-1}-1}{n-1}\right],$$

for $n = 2$.

The following result which deals for polynomials having all the zeros on the circle $S(0,K), K \le 1$ is also due to Pukhta [30].

Theorem 4.8. *Let* $p(z) = \sum_{j=0}^{n} a_j z^j$ *be a polynomial of degree n having all its zeros on* $S(0,K)$, $K \le 1$. *Then for* $R \ge 1$ *and for every positive integer s,*

$$\{M(p,R)\}^s \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^{ns}-1)\}+|a_{n-1}|\{2K^n+R^{ns}-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right]$$
$$\{M(p,1)\}^s - s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-2}-1}{ns-2}\right]\{M(p,1)\}^{s-1}, \quad \textit{for } n>2$$

and

$$\{M(p,R)\}^s \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^{ns}-1)\}+|a_{n-1}|\{2K^n+R^{ns}-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right]$$
$$\{M(p,1)\}^s - s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-1}-1}{ns-1}\right]\{M(p,1)\}^{s-1}, \quad \textit{for } n=2.$$

Choosing $s=1$ in Theorem 4.8 gives

Corollary 4.9. *Let $p(z)=\sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n having all its zeros on $S(0,K)$, $K\leq 1$. Then for $R\geq 1$,*

$$M(p,R) \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^{n}-1)\}+|a_{n-1}|\{2K^n+R^{n}-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right]M(p,1)$$
$$-|a_1|\left[\frac{R^{n}-1}{n}-\frac{R^{n-2}-1}{n-2}\right], \quad \textit{for } n>2$$

and

$$M(p,R) \leq \frac{1}{K^n}\left[\frac{n|a_n|\{K^n(1+K^2)+K^2(R^{n}-1)\}+|a_{n-1}|\{2K^n+R^{n}-1\}}{2|a_{n-1}|+n|a_n|(1+K^2)}\right]M(p,1)$$
$$-|a_1|\left[\frac{R^{n}-1}{n}-\frac{R^{n-1}-1}{n-1}\right], \quad \textit{for } n=2.$$

We close this section by stating the following two recent results due to Pukhta [31] which are for polynomials with gaps, and therefore generalize the above results.

Theorem 4.10. *Let $p(z)=a_n z^n+\sum_{j=\mu}^{n} a_{n-j} z^{n-j}$, $1\leq\mu<n$ be a polynomial of degree n having all its zeros on $S(0,K)$, $K\leq 1$. Then for $R\geq 1$ and for every positive integer s,*

$$\{M(p,R)\}^s \le \left[\frac{K^{n-2\mu+1}(1+K^{\mu})+(R^{ns}-1)}{K^{n-2\mu+1}+K^{n-\mu+1}}\right]\{M(p,1)\}^s$$
$$-s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-2}-1}{ns-2}\right]\{M(p,1)\}^{s-1},$$

for $n>2$; *and*

$$\{M(p,R)\}^s \le \left[\frac{K^{n-2\mu+1}(1+K^{\mu})+(R^{ns}-1)}{K^{n-2\mu+1}+K^{n-\mu+1}}\right]\{M(p,1)\}^s$$
$$-s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-1}-1}{ns-1}\right]\{M(p,1)\}^{s-1},$$

for $n=2$.

Note that for $\mu = 1$, the above Theorem 4.10 reduces to Theorem 4.6.

Theorem 4.11. *Let* $p(z) = a_n z^n + \sum_{j=\mu}^{n} a_{n-j} z^{n-j}$, $1 \le \mu < n$ *be a polynomial of degree* n *having all its zeros on* $S(0,K)$, $K \le 1$. *Then for* $R \ge 1$ *and for every positive integer* s,

$$\{M(p,R)\}^s \le \frac{G(s,k,\mu)}{K^q}\{M(p,1)\}^s$$
$$-s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-2}-1}{ns-2}\right]\{M(p,1)\}^{s-1}, \quad n>2$$

and

$$\{M(p,R)\}^s \le \frac{G(s,k,\mu)}{K^q}\{M(p,1)\}^s$$
$$-s|a_1|\left[\frac{R^{ns}-1}{ns}-\frac{R^{ns-1}-1}{ns-1}\right]\{M(p,1)\}^{s-1}, \quad n=2,$$

where

$$G(s,k,\mu) = \frac{n|a_n|\{K^q(K^{\mu-1}+K^{2\mu})+K^{2\mu}(R^{ns}-1)\}+|a_{n-\mu}|\{\mu(K^n+K^q+K^{\mu-1}(R^{ns}-1))\}}{\mu|a_{n-\mu}|(K^{\mu-1}+1)+n|a_n|(K^{\mu-1}+K^{2\mu})},$$
$$q = n-\mu+1.$$

References

1. Ankeny, N.C., Rivlin, T.J.: On a theorem of S. Bernstein. Pac. J. Math. **5**, 849–852 (1955)
2. Aziz, A.: Growth of polynomials whose zeros are within or outside a circle. Bull. Aust. Math. Soc. **35**, 247–256 (1987)
3. Aziz, A., Mohammad, Q.G.: Simple proof of a theorem of Erdös and Lax. Proc. Am. Math. Soc. **80**, 119–122 (1980)
4. Aziz, A., Mohammad, G.: Growth of polynomials with zeros outside a circle. Proc. Am. Math. Soc. **81**, 549–553 (1981)
5. Aziz, A., Zargar, B.A.: Growth of maximum modulus of polynomials with prescribed zeros. Glas. Mat. **37**, 73–81 (2002)
6. Boas, R.P.: Inequalities for the derivatives of polynomials. Math. Mag. **42**, 165–174 (1969)
7. Dewan, K.K.: Another proof of a theorem of Ankeny and Rivlin. Glas. Math. **18**, 291–293 (1983)
8. Dewan, K.K., Ahuja, A.: Growth of polynomials with prescribed zeros. J. Math. Inequal. **5**, 355–361 (2011)
9. Dewan, K.K., Bhat, A.A.: On the maximum modulus of polynomials not vanishing inside the unit circle. J. Interdisciplinary Math. **1**, 129–140 (1998)
10. Duffin, R.J., Schaeffer, C.: Some properties of functions of exponential type. Bull. Am. Math. Soc. **44**, 236–240 (1938)
11. Frappier, C., Rahman, Q.I.: On an Inequality of S. Bernstein. Can. J. Math. **34**(4), 932–944 (1982)
12. Frappier, C., Rahman, Q.I., Ruscheweyh, S.: New inequalities for polynomials. Trans. Am. Math. Soc. **288**, 69–99 (1985)
13. Frappier, C., Rahman, Q. I., Ruscheweyh, S. : New Inequalities for Polynomials. Trans. Am. Math. Soc. **288**, 69–99 (1985)
14. Gardner, R.B., Govil, N.K., Weems, A.: Some results concerning rate of growth of polynomials. East J. Approx. **10**, 301–312 (2004)
15. Gardner, R.B., Govil, N.K., Musukula, S.R.: Rate of growth of polynomials not vanishing inside a circle. J. Inequal. Pure Appl. Math. **6**(2), article 53 (2005)
16. Govil, N.K.: On the maximum modulus of polynomials. J. Math. Anal. Appl. **112**, 253–258 (1985)
17. Govil, N.K.: On the maximum modulus of polynomials not vanishing inside the unit circle. Approx. Theory Appl. **5**, 79–82 (1989)
18. Govil, N.K.: On a theorem of Ankeny and Rivlin concerning maximum modulus of polynomials. Complex Variables **40**, 249–259 (2000)
19. Govil, N.K.: On Growth of polynomials. J. Inequal. Appl. **7**(5), 623–631 (2002)
20. Govil, N.K., Mohapatra, R.N. : Markov and Bernstein type inequalities for polynomials. J. Inequal. Appl. **3**, 349–387 (1999)
21. Govil, N.K., Nyuydinkong, G.: On maximum modulus of polynomials not vanishing inside a circle. J. Interdisciplinary Math. **4**, 93–100 (2001)
22. Govil, N.K., Qazi, M.A., Rahman, Q.I.: Inequalities describing the growth of polynomials not vanishing in a disk of prescribed radius. Math. Inequal. Appl. **6**(3), 453–467 (2003)
23. Govil, N.K., Liman, A., Shah, W.M.: Some inequalities concerning derivative and maximum modulus of polynomials. Aust. J. Math. Anal. Appl. **8**, 1–8 (2011)
24. Jain, V.K.: Certain interesting implications of T. J. Rivlins's result on maximum modulus of a polynomial. Glas. Math. **33**, 33–36 (1998)
25. Lax, P.D.: Proof of a conjecture of P. Erdös on the derivatives of a polynomial. Bull. Am. Math. Soc. **50**, 509–513 (1944)
26. Markov, A.A.: On a problem of D. I. Mendeleev (Russian), Zapiski Imp. Akad. Nauk **62**, 1–24 (1889)
27. Markov, V.A.: Über Polynome die in einem gegebenen Intervalle möglichst wenig von null abweichen, Math. Ann. **77**, 213–258 (1916)

28. Milovanović, G.V., Mitrinović, D.S., Rassias, T.M.: Topics in Polynomials: Extremal Problems, Inequalities, Zeros. World Scientific Publishing Co. Pte. Ltd., Singapore (1994)
29. Nehari, Z.: Conformal Mapping. McGraw Hill, New York (1952)
30. Pukhta, M.S.: Rate of growth of polynomials with zeros on the unit Disc. Prog. Appl. Math. **6**(2), 50–58 (2013)
31. Pukhta, M.S.: Growth of the maximum modulus of polynomials with prescribed zeros. J. Class. Anal. **5**(2), 107–113 (2014)
32. Qazi, M.A.: On the maximum modulus of polynomials. Proc. Am. Math. Soc. **115**(2), 337–343 (1992)
33. Rahman, Q.I., Schmeisser, G.: Les inégalités de Markoff et de Bernstein. Les Presses de l' Université de Montréal, Québec (1983)
34. Rahman, Q.I., Schmeisser, G.: Analytic Theory of Polynomials. Oxford University Press, New York (2002)
35. Riesz, M.: Über einen Satz des Herrn Serge Bernstein. Acta. Math. **40**, 337–347 (1916)
36. Rivlin, T.J.: On the maximum modulus of polynomials. Am. Math. Mon. **67**, 251–253 (1960)
37. Sharma, A., Singh, V.: Some Bernstein type inequalities for polynomials. Analysis **5**, 321–341 (1985)
38. Varga, R.S.: A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials. J. Soc. Ind. Appl. Math. **5**, 39–46 (1957)
39. Visser, C.: A simple proof of certain inequalities concerning polynomials. Nederl. Akad. Wetensch. Proc. **47**, 276–281 (1945)

Approximation for Generalization of Baskakov–Durrmeyer Operators

Vijay Gupta

Abstract In the present article, we study certain approximation properties of the modified form of generalized Baskakov operators introduced by Erencin (Appl. Math. Comput. 218(3):4384–4390, 2011). We estimate a recurrence relation for the moments of their Durrmeyer type modification. First we estimate rate of convergence for functions having derivatives of bounded variation. Next, we discuss some direct results in simultaneous approximation by these operators, e.g. pointwise convergence theorem, Voronovskaja-type theorem and an estimate of error in terms of the modulus of continuity.

Keywords Baskakov operators • Simultaneous approximation • Rate of convergence • Modulus of continuity • Bounded variation

Introduction

The well-known Baskakov operators for $f \in C[0, \infty)$ are defined as

$$\begin{aligned} B_n(f;x) &= \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{x^k}{(1+x)^{n+k}} f\left(\frac{k}{n}\right) \\ &= \sum_{k=0}^{\infty} \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!} \frac{x^k}{(1+x)^{n+k}} f\left(\frac{k}{n}\right). \end{aligned} \tag{1}$$

Several variants of these operators have been discussed, Gupta [6] proposed the Bézier variant of the Baskakov operators defined by (1). In [13], Mihesan proposed a generalization of the operators (1). He started by considering the following identity

V. Gupta (✉)
Department of Mathematics, Netaji Subhas Institute of Technology, Sector 3, Dwarka, New Delhi 110078, India
e-mail: vijaygupta2001@hotmail.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_14

$$e^{at}.\frac{1}{(1-t)^n} = \left\{\sum_{k=0}^{\infty}\frac{(at)^k}{k!}\right\}\left\{\sum_{k=0}^{\infty}\binom{n+k-1}{k}t^k\right\}$$
$$= \sum_{k=0}^{\infty}\sum_{i=0}^{k}\binom{n+i-1}{i}t^i\frac{(at)^{k-i}}{(k-i)!}$$

Replacing $t = x/(1+x)$, we get

$$e^{ax/(1+x)}.(1+x)^n = \sum_{k=0}^{\infty}\sum_{i=0}^{k}\binom{k}{i}\frac{(n)_i a^{k-i}}{k!}\left(\frac{x}{1+x}\right)^k$$

introduced the following generalized Baskakov operators with a non-negative constant $a \geq 0$ independent of n :

$$B_n^a(f;x) = \sum_{k=0}^{\infty} v_{n,k}^a(x) f\left(\frac{k}{n}\right), \tag{2}$$

where

$$v_{n,k}^a(x) = e^{-ax/(1+x)}\frac{P_k(n,a)}{k!}\frac{x^k}{(1+x)^{n+k}}$$

such that $\sum_{k=0}^{\infty} v_{n,k}^a(x) = 1$ and $P_k(n,a) = \sum_{i=0}^{k}\binom{k}{i}(n)_i a^{k-i}$. In case $a = 0$, we get at once the Baskakov operators (1).

Also, another generalization of the Baskakov operators defined by Chen [2] was considered in the following way:

$$B_n^{\beta}(f;x) = \sum_{k=0}^{\infty}\frac{n(n+\beta)(n+2\beta)\cdots(n+(k-1)\beta)}{k!}\frac{x^k}{(1+\beta x)^{n/\beta+k}} f\left(\frac{k}{n}\right). \tag{3}$$

Combining (2) and (3), we propose the following more general form for $a \geq 0$ and $\beta > 0$, which include these two generalizations of the Baskakov operators as special cases:

$$B_n^{a,\beta}(f;x) = \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) f\left(\frac{k}{n}\right), \tag{4}$$

where

$$v_{n,k}^{a,\beta}(x) = e^{-a\beta x/(1+\beta x)} \frac{\sum_{i=0}^{k} \binom{k}{i} \prod_{j=0}^{i-1} (\frac{n}{\beta} + j).a^{k-i}}{k!} \frac{(\beta x)^k}{(1+\beta x)^{n/\beta+k}}$$

such that $\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) = 1$. The operators (4) as such are not possible to approximate Lebesgue integrable functions on the interval $[0,\infty)$, to approximate integrable functions, we introduce the Durrmeyer type modification by considering the parameters $a \geq 0$ and $\beta > 0$, in the following form:

$$L_n^{a,\beta}(f;x) = \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t) f(t) dt, \tag{5}$$

where $v_{n,k}^{a,\beta}(x)$ is as defined in (4) and the Beta basis function is defined as

$$b_{n,k}(t) = \frac{1}{B(k+1,n)} \frac{t^k}{(1+t)^{n+k+1}}$$

the beta function is given by

$$B(m,n) = \int_0^{\infty} \frac{t^{m-1}}{(1+t)^{m+n}} dt = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}, \quad m,n > 0.$$

As a special case, for $a = 0$ and $\beta = 1$ these operators become the well-known Baskakov–Beta operators introduced by the author Gupta [7] (see also [3, 4]). We may point out that one may consider the Stancu type generalization by considering two more parameters as done in [15], here we consider the form (5).

Let L denote the class of all Lebesgue measurable functions f on $[0,\infty)$ as

$$\mathrm{L} = \left\{ f : \int_0^{\infty} \frac{|f(t)|}{(1+t)^m} dt < \infty, \text{for some positive integer } m \right\}.$$

Note that the class L is bigger than the class of Lebesgue integrable functions on the interval $[0,\infty)$.

For different problems related to the present article, we refer the reader to [8, 9]. The aim of the present paper is to discuss some approximation properties of the operators $L_n^{a,\beta}$. We estimate rate of convergence for functions having derivatives of bounded variation. We also discuss simultaneous approximation properties, which include point-wise convergence, asymptotic formula and error estimations.

Basic Lemmas

In this section, we present some basic lemmas, which are necessary to prove main results.

Lemma 1. *For every $x \in (0,\infty)$, we have*

$$x(1+\beta x)^2\left\{\frac{d}{dx}v_{n,k}^{a,\beta}(x)\right\} = \left\{(k-nx)(1+\beta x)-a\beta x\right\}v_{n,k}^{a,\beta}(x).$$

Proof. Differentiating $v_{n,k}^{a,\beta}(x)$ with respect to x and multiplying by $x(1+\beta x)^2$ on both sides, the result easily follows, hence the details are omitted.

Lemma 2. *For $m \in \mathrm{N}^0$, if the mth order moment of the generalized Baskakov operators B_n^a is defined as*

$$\eta_{n,m}^{a,\beta}(x) = B_n^{a,\beta}((t-x)^m;x) = \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\left(\frac{k}{n}-x\right)^m.$$

Then, for $m \geq 1$, we have the following recurrence relation

$$n(1+\beta x)\eta_{n,m+1}^{a,\beta}(x) = x(1+\beta x)^2(\eta_{n,m}^{a,\beta}(x))' + a\beta x\eta_{n,m}^{a,\beta}(x) + mx(1+\beta x)^2\eta_{n,m-1}^{a,\beta}(x). \tag{6}$$

In particular

$$\eta_{n,0}^{a,\beta}(x) = 1, \quad \eta_{n,1}^{a,\beta}(x) = \frac{a\beta x}{n(1+\beta x)}.$$

Consequently,

(i) *$\eta_{n,m}^{a,\beta}(x)$ is a rational function of x depending on the parameters a, β;*
(ii) *for each $x \in (0,\infty)$ and $m \in \mathrm{N}^0$, $\eta_{n,m}^{a,\beta}(x) = O(n^{-[(m+1)/2]})$, where $[s]$ denotes the integer part of s.*

Proof. By definition of $\eta_{n,m}^{a,\beta}(x)$ we have

$$x(1+\beta x)^2[\eta_{n,m}^{a,\beta}(x)]' = x(1+\beta x)^2\left[\sum_{k=0}^{\infty}[v_{n,k}^{a,\beta}(x)]'\left(\frac{k}{n}-x\right)^m \right.$$
$$\left. -m\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\left(\frac{k}{n}-x\right)^{m-1}\right]$$

$$= \sum_{k=0}^{\infty}[(k-nx)(1+\beta x)-a\beta x]v_{n,k}^{a,\beta}(x)\left(\frac{k}{n}-x\right)^m$$
$$-mx(1+\beta x)^2\eta_{n,m-1}^{a,\beta}(x)$$
$$= n(1+\beta x)\eta_{n,m+1}^{a,\beta}(x)-a\beta x\eta_{n,m}^{a,\beta}(x)-mx(1+\beta x)^2\eta_{n,m-1}^{a,\beta}(x).$$

This completes the proof of recurrence relation. The consequences (i) and (ii) follow from (6) by using induction on m.

We establish below a Lorentz-type lemma for the derivatives of the kernel $v_{n,k}^{a,\beta}(x)$ of the generalized Baskakov operators (4).

Lemma 3. *For each $x \in (0,\infty)$ and $r \in \mathrm{N}^0$, there exist polynomials $q_{i,j,r}(x)$ in x independent of n and k such that*

$$\frac{d^r}{dx^r}v_{n,k}^{a,\beta}(x) = v_{n,k}^{a,\beta}(x)\sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i(k-nx)^j\frac{q_{i,j,r}(x)}{(x(1+\beta x)^2)^r}.$$

The proof of the above lemma follows along the lines of Gupta and Agarwal [8], we omit the details.

Lemma 4. *The* mth *order* $(m \in \mathrm{N}^0 := \mathrm{N}\cup\{0\})$ *moment for the operators (5) is defined as*

$$T_{n,m}^{a,\beta}(x) := L_n^{a,\beta}(t^m;x) = \sum_{k=0}^{\infty}v_{n,k}^{a,\beta}(x)\int_0^{\infty}b_{n,k}(t)t^m dt, \quad n>m.$$

Then, $T_{n,0}^{a,\beta}(x)=1$ and there holds the following recurrence relation for $n>m+1$:

$$(n-m-1)T_{n,m+1}^{a,\beta}(x) = x(1+\beta x)(T_{n,m}^{a,\beta}(x))'$$
$$+\left((m+1)+nx+\frac{a\beta x}{(1+\beta x)}\right)T_{n,m}^{a,\beta}(x). \qquad (7)$$

Proof. For $x=0$, the relation (7) is easily verified. For $x \in (0,\infty)$, we proceed as follows:

From Lemma 1, we may write

$$(T_{n,m}^{a,\beta}(x))' = \sum_{k=0}^{\infty}\frac{d}{dx}(v_{n,k}^{a,\beta}(x))\int_0^{\infty}b_{n,k}(t)t^m dt$$
$$= \sum_{k=0}^{\infty}\frac{((k-nx)(1+\beta x)-a\beta x)v_{n,k}^{a,\beta}(x)}{x(1+\beta x)^2}\int_0^{\infty}b_{n,k}(t)t^m dt$$

$$= \frac{1}{x(1+\beta x)} \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} (k-nx) b_{n,k}(t) t^m dt$$

$$- \frac{a\beta}{(1+\beta x)^2} \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t) t^m dt$$

$$= I_1 - \frac{a\beta}{(1+\beta x)^2} T_{n,m}^{a,\beta}(x). \tag{8}$$

We may write I_1 as

$$I_1 = \frac{1}{x(1+\beta x)} \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t)[k-(n+1)t+(n+1)t-nx] t^m dt$$

$$= \frac{1}{x(1+\beta x)} \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t)(k-(n+1)t) t^m dt + \frac{(n+1)}{x(1+\beta x)} T_{n,m+1}^{a,\beta}(x)$$

$$- \frac{nx}{x(1+\beta x)} T_{n,m}^{a,\beta}(x)$$

$$= I_2 + \frac{(n+1)}{x(1+\beta x)} T_{n,m+1}^{a,\beta}(x) - \frac{nx}{x(1+\beta x)} T_{n,m}^{a,\beta}(x). \tag{9}$$

Now, we consider I_2. Making use of the identity

$$t(1+t)\frac{d}{dt}\Big(b_{n,k}(t)\Big) = (k-(n+1)t) b_{n,k}(t), t \in (0,\infty)$$

and then integrating by parts, we obtain for $n > m+1$ the following:

$$I_2 = -\frac{1}{x(1+\beta x)} \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t)\Big((m+1)t^m + (m+2)t^{m+1}\Big) dt$$

$$= \frac{1}{x(1+\beta x)} \Big(-(m+1) T_{n,m}^{a,\beta}(x) - (m+2) T_{n,m+1}^{a,\beta}(x) \Big). \tag{10}$$

From Eqs. (8)–(10), we get the required recurrence relation.

Corollary 1. *For the function $T_{n,m}^{a,\beta}(x)$, we have*

1. For $n > 1$, we have

$$L_n^{a,\beta}(t;x) = \frac{1}{(n-1)} \Big(nx + \frac{a\beta x}{(1+\beta x)} + 1 \Big).$$

2. *For $n > 2$, we have*

$$L_n^{a,\beta}(t^2;x) = \frac{1}{(n-1)(n-2)}\Bigg((n^2+n)x^2 + 4nx + \frac{a^2x^2}{(1+\beta x)^2} + \frac{2nax^2}{(1+\beta x)} + \frac{4ax}{(1+\beta x)} + 2\Bigg).$$

3. *For each $x \in (0,\infty)$ and $m \in \mathrm{N}$,*

$$T_{n,m}^{a,\beta}(x) = a_m(n)x^m + n^{-1}(p_m(x,a,\beta) + o(1)),$$

where

$$a_m(n) = \frac{\prod_{j=0}^{m-1}(n+j)}{\prod_{j=1}^{m}(n-j)}$$

and $p_m(x,a,\beta)$ is a rational function of x depending on the parameters a, β and m.

Lemma 5. *For the* mth *order central moment $\mu_{n,m}^{a,\beta}(x), n > m$ defined as*

$$\mu_{n,m}^{a,\beta}(x) := L_n^{a,\beta}((t-x)^m;x) = \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\int_0^{\infty} b_{n,k}(t)(t-x)^m dt, \quad n > m$$

we have

$$\mu_{n,0}^{a,\beta}(x) = 1,\ \mu_{n,1}^{a,\beta}(x) = \frac{(1+x)}{(n-1)} + \frac{a\beta x}{(1+\beta x)(n-1)}$$

and the following recurrence relation for all $n \geq m+1$:

$$(n-m-1)\mu_{n,m+1}^{a,\beta}(x) = x\Big[(1+\beta x)(\mu_{n,m}^{a,\beta}(x))' + m(2+(1+\beta)x)\mu_{n,m-1}^{a,\beta}(x)\Big] + \Big[(1+2x)(m+1) - \frac{x(1+\beta x - a\beta)}{(1+\beta x)}\Big]\mu_{n,m}^{a,\beta}(x).$$

Proof. Using Lemma 2, we have

$$x(1+\beta x)^2(\mu_{n,m}^{a,\beta}(x))'$$
$$= \sum_{k=0}^{\infty}[(k-nx)(1+\beta x) - a\beta x]v_{n,k}^{a,\beta}(x)\int_0^{\infty} b_{n,k}(t)(t-x)^m dt$$
$$-mx(1+\beta x)^2\mu_{n,m-1}^{a,\beta}(x)$$

$$= (1+\beta x)\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\int_0^{\infty}[(k-(n+1)t+(n+1)(t-x)+x]b_{n,k}(t)(t-x)^m dt$$

$$-a\beta x\mu_{n,m}^{a,\beta}(x) - mx(1+\beta x)^2\mu_{n,m-1}^{a,\beta}(x)$$

$$= (1+\beta x)\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\int_0^{\infty}[k-(n+1)t]b_{n,k}(t)(t-x)^m dt$$

$$+ (n+1)(1+\beta x)\mu_{n,m+1}^{a,\beta}(x)+(1+\beta x-a\beta)x\mu_{n,m}^{a,\beta}(x) - mx(1+\beta x)^2\mu_{n,m-1}^{a,\beta}(x)$$

$$= (1+\beta x)\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)\int_0^{\infty} t(1+t)b'_{n,k}(t)(t-x)^m dt$$

$$+ (n+1)(1+\beta x)\mu_{n,m+1}^{a,\beta}(x)+(1+\beta x-a\beta)x\mu_{n,m}^{a,\beta}(x) - mx(1+\beta x)^2\mu_{n,m-1}^{a,\beta}(x).$$

Using the identity

$$t(1+t) = (t-x)^2 + (1+2x)(t-x) + x(1+x),$$

and integrating by parts, we have

$$x(1+\beta x)^2(\mu_{n,m}^{a,\beta}(x))'$$

$$= -(1+\beta x)\left[(m+2)\mu_{n,m+1}^{a,\beta}(x)+(1+2x)(m+1)\mu_{n,m}^{a,\beta}(x)+mx(1+x)\mu_{n,m-1}^{a,\beta}(x)\right]$$

$$+ (n+1)(1+\beta x)\mu_{n,m+1}^{a,\beta}(x)+(1+\beta x-a\beta)x\mu_{n,m}^{a,\beta}(x) - mx(1+\beta x)^2\mu_{n,m-1}^{a,\beta}(x).$$

This completes the proof of recurrence relation.

Corollary 2. *For the function $\mu_{n,m}^{a,\beta}(x)$, we have*

(i) $\mu_{n,m}^{a,\beta}$ *is a rational function of x and n depending on the parameter a, β and m;*
(ii) for every $x \in (0,\infty)$, $\mu_{n,m}^{a,\beta}(x) = O(n^{-[(m+1)/2]})$, where $[\alpha]$ denotes the integer part of α.

Let $\psi_x(t) = (t-x)$. For $r \in N^0$ and $x \in (0,\infty)$, applying Schwarz inequality and Corollary 2 (ii), we have

$$L_n^{a,\beta}(|\psi_x^r(t)|, x) \le \sqrt{L_n^{a,\beta}(\psi_x^{2r}(t); x)} = O\left(n^{-r/2}\right).$$

From Lemma 5, for $n > 2$, we obtain

$$\mu_{n,2}^{a,\beta}(x) = \frac{1}{(n-1)(n-2)}\Bigg[[(1+\beta)n+2]x^2 + (2n+4)x$$
$$+\frac{a^2\beta^2x^2}{(1+\beta x)^2} + \frac{4a\beta x^2}{1+\beta x} + \frac{3a\beta x}{1+\beta x} + 2\Bigg]. \qquad (11)$$

Hence, for $\lambda > 1$ and n sufficiently large, by using Schwarz inequality and (11), we have

$$L_n^{a,\beta}(|\psi_x(t)|;x) \leq \sqrt{\frac{\lambda x[2+(1+\beta)x]}{n}}. \tag{12}$$

To study the approximation of functions having a derivative of bounded variation, let us rewrite the operators (5) as

$$L_n^{a,\beta}(f;x) = \int_0^\infty K_n^{a,\beta}(x;t)f(t)dt,$$

where $K_n^{a,\beta}(x,t)$ is the kernel function given by

$$K_n^{a,\beta}(x;t) = \sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)b_{n,k}(t).$$

Lemma 6. *For a fixed $x \in (0,\infty)$ and n sufficiently large, we have*

$$\lambda_{n,a,\beta}(x;y) := \int_0^y K_n^{a,\beta}(x;t)dt \leq \frac{\lambda x[2+(1+\beta)x]}{n(x-y)^2}, \quad 0 \leq y < x, \tag{13}$$

$$1-\lambda_{n,a,\beta}(x;z) := \int_z^\infty K_n^{a,\beta}(x;t)dt \leq \frac{\lambda x[2+(1+\beta)x]}{n(z-x)^2}, \quad x < z < \infty. \tag{14}$$

The proof of the above lemma easily follows and so the details are omitted.

Rate of Approximation

In this section, we shall estimate the rate of convergence for the generalized Baskakov–Durrmeyer operators $L_n^{a,\beta}$ for functions with derivatives of bounded variation. In the recent years, several researchers have made significant contributions in this direction. We refer the reader to some of the related papers (cf. [1, 11, 12, 14] etc).

Let $f \in DBV_\gamma[0,\infty)$, $\gamma \geq 0$ be the class of all functions defined on $[0,\infty)$, having a derivative of bounded variation on every finite subinterval of $[0,\infty)$ and $|f(t)| \leq Mt^\gamma$, $\forall \ t > 0$.

It turns out that for $f \in DBV_\gamma[0,\infty)$, we may write

$$f(x) = \int_0^x g(t)dt + f(0),$$

where $g(t)$ is a function of bounded variation on each finite subinterval of $[0,\infty)$.

Theorem 1. *Let $f \in DBV_\gamma[0,\infty)$, $\gamma \geq 0$. Then, for every $x \in (0,\infty)$, $\lambda > 1$, $r \in \mathbb{N}$ $(2r \geq \gamma)$ and for n sufficiently large, we have*

$$\left| L_n^{a,\beta}(f;x) - f(x) - \frac{f'(x+) - f'(x-)}{2}\sqrt{\frac{\lambda x[2+(1+\beta)x]}{n}} \right|$$
$$\leq \left| \frac{f'(x+) + f'(x-)}{2} \right| \left(\frac{1+x}{n-1} + \frac{a\beta x}{(1+\beta x)(n-1)} \right)$$
$$+ \frac{x}{\sqrt{n}} \bigvee_{x-\frac{x}{\sqrt{n}}}^{x+\frac{x}{\sqrt{n}}} (f'_x) + \frac{\lambda[2+(1+\beta)x]}{n} \sum_{k=1}^{[\sqrt{n}]} \bigvee_{x-\frac{x}{k}}^{x+\frac{x}{k}} (f'_x)$$
$$+ \frac{C_1(x,r,a,\beta)}{n^r} + |f(x)| \frac{\mu_{n,2}^a(x)}{x^2} + |f'(x+)| \sqrt{\frac{\lambda x[2+(1+\beta)x]}{n}}$$
$$+ \frac{\lambda x[2+(1+\beta)x]}{n} |f(2x) - f(x) - xf'(x+)|,$$

where

$$f'_x(t) = \begin{cases} f'(t) - f'(x+), & x < t < \infty \\ 0, & t = x \\ f'(t) - f'(x-), & 0 \leq t < x \end{cases} \tag{15}$$

and $\vee_c^d(f'_x)$ is the total variation of f'_x on $[c,d] \subset (0,\infty)$.

Proof. It follows from (15) that

$$f'(t) = \frac{1}{2}(f'(x+) + f'(x-)) + f'_x(t) + \frac{f'(x+) - f'(x-)}{2} sgn(t-x)$$
$$+ \delta_x(t)\left(f'(t) - \frac{1}{2}(f'(x+) + f'(x-)) \right),$$

where

$$\delta_x(t) = \begin{cases} 1, & x = t \\ 0, & x \neq t \end{cases}$$

By simple computation, we have

$$|L_n^{a,\beta}(f;x) - f(x)| \leq \left| \frac{f'(x+) + f'(x-)}{2} \right| |L_n^{a,\beta}(t-x,x)|$$
$$+ \left| \frac{f'(x+) - f'(x-)}{2} \right| L_n^{a,\beta}(|t-x|,x)$$
$$+ |E_1(n,x,a,\beta)| + |E_2(n,x,a,\beta)|, \tag{16}$$

where

$$E_1(n,x,a,\beta) = \int_0^x \left(\int_t^x f_x'(u)du\right) K_n^{a,\beta}(x,t)dt$$

and

$$E_2(n,x,a,\beta) = \int_x^\infty \left(\int_x^t f_x'(u)du\right) K_n^{a,\beta}(x,t)dt.$$

First, we obtain an estimate of $E_1(n,x,a,\beta)$. From the definition of $\lambda_{n,a}(x,t)$ given in (13), we may write

$$E_1(n,x,a,\beta) = \int_0^x \left(\int_t^x f_x'(u)du\right) \frac{\partial}{\partial t}\lambda_{n,a,\beta}(x,t)dt.$$

Applying the integration by parts, we get

$$\begin{aligned}|E_1(n,x,a,\beta)| &\le \int_0^x |f_x'(t)|\lambda_{n,a,\beta}(x,t)dt \\ &\le \int_0^{x-\frac{x}{\sqrt{n}}} |f_x'(t)|\lambda_{n,a,\beta}(x,t)dt + \int_{x-\frac{x}{\sqrt{n}}}^x |f_x'(t)|\lambda_{n,a,\beta}(x,t)dt \\ &= J_1 + J_2, \ \text{say.} \qquad (17)\end{aligned}$$

Since $f_x'(x) = 0$ and $\lambda_{n,a,\beta}(x,t) \le 1$, we have

$$\begin{aligned}J_2 &= \int_{x-\frac{x}{\sqrt{n}}}^x |f_x'(t) - f_x'(x)|\lambda_{n,a,\beta}(x,t)dt \\ &\le \int_{x-\frac{x}{\sqrt{n}}}^x \bigvee_t^x (f_x')dt \\ &\le \bigvee_{x-\frac{x}{\sqrt{n}}}^x (f_x') \int_{x-\frac{x}{\sqrt{n}}}^x dt = \frac{x}{\sqrt{n}} \bigvee_{x-\frac{x}{\sqrt{n}}}^x (f_x'). \qquad (18)\end{aligned}$$

Next, we estimate J_1. By applying Lemma 6 and putting $t = x - \frac{x}{u}$, we have

$$\begin{aligned}J_1 &\le \frac{\lambda x[2+(1+\beta)x]}{n} \int_0^{x-\frac{x}{\sqrt{n}}} |f_x'(t) - f_x'(x)| \frac{dt}{(x-t)^2} \\ &\le \frac{\lambda x[2+(1+\beta)x]}{n} \int_x^{x-\frac{x}{\sqrt{n}}} \bigvee_t^x (f_x') \frac{dt}{(x-t)^2}\end{aligned}$$

$$= \frac{\lambda[2+(1+\beta)x]}{n}\int_1^{\sqrt{n}} \bigvee_{x-\frac{x}{u}}^{x} (f_x')du$$

$$\leq \frac{\lambda[2+(1+\beta)x]}{n}\sum_{k=1}^{[\sqrt{n}]} \bigvee_{x-\frac{x}{u}}^{x} (f_x'). \tag{19}$$

Substituting the values of J_1 and J_2 in (17), we obtain

$$|E_1(n,x,a,\beta)| \leq \frac{x}{\sqrt{n}} \bigvee_{x-\frac{x}{\sqrt{n}}}^{x} (f_x') + \frac{\lambda[2+(1+\beta)x]}{n}\sum_{k=1}^{[\sqrt{n}]} \bigvee_{x-\frac{x}{u}}^{x} (f_x'). \tag{20}$$

Now, we find an estimate of $E_2(n,x,a,\beta)$.

$$\begin{aligned}
E_2(n,x,a,\beta) &\leq \left| \int_{2x}^{\infty} \left(\int_x^t f_x'(u)du \right) K_n^{a,\beta}(x,t)dt \right| \\
&\quad + \left| \int_{x}^{2x} \left(\int_x^t f_x'(u)du \right) \frac{\partial}{\partial t}(1-\lambda_{n,a,\beta}(x,t))dt \right| \\
&\leq \left| \int_{2x}^{\infty} [f(t)-f(x)]K_n^{a,\beta}(x,t)dt \right| + |f'(x+)| \left| \int_{2x}^{\infty} (t-x)K_n^{a,\beta}(x,t)dt \right| \\
&\quad + \left| \int_{x}^{2x} f_x'(u)du \right| |1-\lambda_{n,a,\beta}(x,2x)| + \left| \int_{x}^{2x} f_x'(t)dt(1-\lambda_{n,a,\beta}(x,t))dt \right|.
\end{aligned}$$

We note that there exists an integer r $(2r \geq \gamma)$, such that $f(t) = O(t^{2r})$, for every $t > 0$. Proceeding in a manner similar to the treatment of $E_1(n,x,a,\beta)$ in (17), on using (12) and Lemma 6, we have

$$\begin{aligned}
|E_2(n,x,a,\beta)| &\leq M\int_{2x}^{\infty} t^{2r}K_n^{a,\beta}(x,t)dt + |f(x)|\int_{2x}^{\infty} K_n^{a,\beta}(x,t)dt \\
&\quad + |f'(x+)|\sqrt{\frac{\lambda x[2+(1+\beta)x]}{n}} \\
&\quad + \frac{\lambda[2+(1+\beta)x]}{n}|f(2x)-f(x)-xf'(x+)| + \frac{x}{\sqrt{n}} \bigvee_{x}^{x+\frac{x}{\sqrt{n}}} (f_x') \\
&\quad + \frac{\lambda[2+(1+\beta)x]}{n}\sum_{k=1}^{[\sqrt{n}]} \bigvee_{x}^{x+\frac{x}{k}} (f_x').
\end{aligned} \tag{21}$$

Since $t \leq 2(t-x)$ and $x \leq t-x$ when $t \geq 2x$, in view of Hölder's inequality and Corollary 2 (ii), we obtain

$$\begin{aligned}
&\int_{2x}^{\infty} t^{2r} K_n^{a,\beta}(x,t)dt + |f(x)| \int_{2x}^{\infty} K_n^{a,\beta}(x,t)dt \\
&\leq 2^{2r} \int_{2x}^{\infty} (t-x)^{2r} K_n^{a,\beta}(x,t)dt + \frac{|f(x)|}{x^2} \int_{2x}^{\infty} (t-x)^2 K_n^{a,\beta}(x,t)dt \\
&\leq 2^{2r} \int_{2x}^{\infty} (t-x)^{2r} K_n^{a,\beta}(x,t)dt + \frac{|f(x)|}{x^2} \mu_{n,2}^{a,\beta}(x) \\
&\leq \frac{C_1(x,r,a,\beta)}{n^r} + |f(x)| \frac{\mu_{n,2}^{a,\beta}(x)}{x^2}. \qquad (22)
\end{aligned}$$

Finally, combining (35) and (20)–(22), we get the required result. Hence, the proof is completed.

Simultaneous Approximation

In the following theorem, we show that the derivative $\left(\frac{d^r}{dw^r} L_n^{a,\beta}(f;w)\right)_{w=x}$ converges to $f^{(r)}(x)$.

Theorem 2. *Let $f \in \mathrm{L}$ be bounded on every finite sub-interval of $[0,\infty)$ and $f(t) = O(t^\alpha)$ as $t \to \infty$ for some $\alpha > 0$. If $f^{(r)}$ exists at a point $x \in (0,\infty)$, then we have*

$$\lim_{n\to\infty} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f;w)\right)_{w=x} = f^{(r)}(x). \qquad (23)$$

Proof. By our hypothesis, we have

$$f(t) = \sum_{i=0}^{r} \frac{f^{(i)}(x)}{i!}(t-x)^i + \psi(t,x)(t-x)^r, \quad t \in [0,\infty), \qquad (24)$$

where the function $\psi(t,x) \to 0$ as $t \to x$ and $\psi(t,x) = O((t-x)^\gamma)$ as $t \to \infty$ for some $\gamma > 0$.

From Eq. (24), we can write

$$\begin{aligned}
\left(\frac{d^r}{dw^r} L_n^{a,\beta}(f(t);w)\right)_{w=x} &= \sum_{i=0}^{r} \frac{f^{(i)}(x)}{i!} \left(\frac{d^r}{dw^r} L_n^{a,\beta}((t-x)^i;w)\right)_{w=x} \\
&\quad + \left(\frac{d^r}{dw^r} L_n^{a,\beta}(\psi(t,x)(t-x)^r;w)\right)_{w=x} \\
&:= I_1 + I_2. \qquad (25)
\end{aligned}$$

First, we estimate I_1.

$$\begin{aligned} I_1 &= \sum_{i=0}^{r} \frac{f^{(i)}(x)}{i!}\left[\frac{d^r}{dw^r}\left(\sum_{v=0}^{i}\binom{i}{v}(-x)^{i-v}L_n^{a,\beta}(t^v;w)\right)_{w=x}\right] \\ &= \sum_{i=0}^{r} \frac{f^{(i)}(x)}{i!}\sum_{v=0}^{i}\binom{i}{v}(-x)^{i-v}\left(\frac{d^r}{dw^r}L_n^{a,\beta}(t^v;w)\right)_{w=x} \\ &= \sum_{i=0}^{r-1} \frac{f^{(i)}(x)}{i!}\sum_{v=0}^{i}\binom{i}{v}(-x)^{i-v}\left(\frac{d^r}{dw^r}L_n^{a,\beta}(t^v;w)\right)_{w=x} \\ &\quad + \frac{f^{(r)}(x)}{r!}\sum_{v=0}^{r}\binom{r}{v}(-x)^{r-v}\left(\frac{d^r}{dw^r}L_n^{a,\beta}(t^v;w)\right)_{w=x} \\ &:= I_3 + I_4. \end{aligned} \tag{26}$$

Now, we may write

$$\begin{aligned} I_4 &= \frac{f^{(r)}(x)}{r!}\sum_{v=0}^{r-1}\binom{r}{v}(-x)^{r-v}\left(\frac{d^r}{dw^r}L_n^{a,\beta}(t^v;w)\right)_{w=x} + \frac{f^{(r)}(x)}{r!}\left(\frac{d^r}{dw^r}L_n^{a,\beta}(t^r;w)\right)_{w=x} \\ &:= I_5 + I_6. \end{aligned} \tag{27}$$

Making use of Corollary 1 (iii), we obtain

$I_6 = f^{(r)}(x)a_r(n) + O\left(\frac{1}{n}\right)$, and the terms $I_3 = O\left(\frac{1}{n}\right)$ and $I_5 = O\left(\frac{1}{n}\right)$.

Combining (25)–(27), for each $x \in (0,\infty)$ we have $I_1 \to f^{(r)}(x)$ as $n \to \infty$.

In view of Lemma 1, we have

$$|I_2| \leq \sum_{k=0}^{\infty}\sum_{\substack{2i+j\leq r\\ i,j\geq 0}} n^i|k-nx|^j\frac{|q_{i,j,r}(x)|}{(x(1+\beta x))^r}v_{n,k}^{a,\beta}(x)\int_0^\infty b_{n,k}(t)|\psi(t,x)||(t-x)|^r dt\Bigg)_{w=x}. \tag{28}$$

Since $\psi(t,x) \to 0$ as $t \to x$, for a given $\epsilon > 0$ there exists a $\delta > 0$ such that $|\psi(t,x)| < \epsilon$ whenever $|t-x| < \delta$. For $|t-x| \geq \delta$, we can find some $\gamma > 0$ such that $|\psi(t,x)| \leq M|t-x|^\gamma$, for some $M > 0$. Thus, from Eq. (28) we may write

$$
\begin{aligned}
|I_2| \le & \sum_{k=0}^{\infty} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i |k-nx|^j \frac{|q_{i,j,r}(x)|}{(x(1+\beta x))^r} v_{n,k}^{a,\beta}(x) \Big(\epsilon \int_{|t-x|<\delta} b_{n,k}(t)|t-x|^r dt \\
& + M \int_{|t-x|\ge\delta} b_{n,k}(t)|t-x|^{r+\gamma} dt \Big) \\
:= & J_1 + J_2, \quad \text{say.}
\end{aligned} \tag{29}
$$

Let $K = \sup\limits_{\substack{2i+j\le r\\ i,j\ge 0}} \frac{|q_{i,j,r}(x)|}{(x(1+\beta x))^r}$.

Using Schwarz inequality, Lemma 2 and Corollary 2 (ii) we have

$$
\begin{aligned}
J_1 &\le \epsilon \; K \sum_{k=0}^{\infty} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i |k-nx|^j v_{n,k}^{a,\beta}(x) \Big(\int_0^{\infty} b_{n,k}(t)(t-x)^{2r} dt \Big)^{1/2} \\
&\le \epsilon \; K \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i \Big(\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)(k-nx)^{2j} \Big)^{1/2} \Big(\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t)(t-x)^{2r} dt \Big)^{1/2} \\
&= \epsilon \; K \sum_{\substack{2i+j\le r\\ i,j\ge 0}} O(n^{\frac{2i+j}{2}}) O(n^{-r/2}) \\
&= \epsilon \; .O(1).
\end{aligned} \tag{30}
$$

Since $\epsilon > 0$ is arbitrary, $J_1 \to 0$ as $n \to \infty$.

Let $s(\in \mathrm{N}) > r+\gamma$. Proceeding in a manner similar to the estimate of J_1, we obtain

$$
\begin{aligned}
J_2 &\le \frac{M\,K}{\delta^{s-r-\gamma}} \sum_{k=0}^{\infty} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i |k-nx|^j v_{n,k}^{a,\beta}(x) \int_{|t-x|\ge\delta} b_{n,k}(t)|t-x|^s dt \\
&\le \frac{M_1}{\delta^{s-r-\gamma}} \sum_{k=0}^{\infty} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i |k-nx|^j v_{n,k}^{a,\beta}(x) \Big(\int_0^{\infty} b_{n,k}(t)(t-x)^{2s} dt \Big)^{1/2}, \quad \text{where } M_1 = M\,K \\
&\le \frac{M_1}{\delta^{s-r-\gamma}} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i \Big(\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x)(k-nx)^{2j} \Big)^{1/2} \Big(\sum_{k=0}^{\infty} v_{n,k}^{a,\beta}(x) \int_0^{\infty} b_{n,k}(t)(t-x)^{2s} dt \Big)^{1/2} \\
&= \frac{M_1}{\delta^{s-r-\gamma}} \sum_{\substack{2i+j\le r\\ i,j\ge 0}} n^i O(n^{j/2}) O(n^{-s/2}) \\
&= \frac{M_1}{\delta^{s-r-\gamma}} O(n^{(r-s)/2}),
\end{aligned} \tag{31}
$$

which implies that $J_2 \to 0$ as $n \to \infty$.

Combining the estimates of J_1 and J_2, $I_2 \to 0$ as $n \to \infty$.

Thus, from the estimates of I_1 and I_2, the required result follows. This completes the proof.

Next, we establish a Voronovskaja type asymptotic formula in simultaneous approximation.

Theorem 3. *Let $f \in \mathrm{L}$ be bounded on every finite sub-interval of $[0, \infty)$ and $f(t) = O(t^{\alpha})$ as $t \to \infty$ for some $\alpha > 0$. If f admits a derivative of order $(r+2)$ at a fixed point $x \in (0, \infty)$, then we have*

$$\lim_{n\to\infty} n\left(\left(\frac{d^r}{dw^r} L_n^{a,\beta}(f; w)\right)_{w=x} - f^{(r)}(x)\right) = \sum_{v=1}^{r+2} Q(v, r, a, \beta, x) f^{(v)}(x),$$

where $Q(v, r, a, \beta, x)$ are certain rational functions of x depending on the parameter a.

Proof. From the Taylor's theorem, we may write

$$f(t) = \sum_{v=0}^{r+2} \frac{f^{(v)}(x)}{v!}(t-x)^v + \psi(t,x)(t-x)^{r+2}, \quad t \in [0, \infty), \tag{32}$$

where the function $\psi(t,x) \to 0$ as $t \to x$ and $\psi(t,x) = O((t-x)^{\gamma})$ as $t \to \infty$ for some $\gamma > 0$. From Eq. (32), we obtain

$$\begin{aligned}
\left(\frac{d^r}{dw^r} L_n^{a,\beta}(f(t); w)\right)_{w=x} &= \sum_{v=0}^{r+2} \frac{f^{(v)}(x)}{v!}\left(\frac{d^r}{dw^r} L_n^{a,\beta}((t-x)^v; w)\right)_{w=x} \\
&\quad + \left(\frac{d^r}{dw^r} L_n^{a,\beta}(\psi(t,x)(t-x)^{r+2}; w)\right)_{w=x} \\
&= \sum_{v=0}^{r+2} \frac{f^{(v)}(x)}{v!} \sum_{j=0}^{v} \binom{v}{j} (-x)^{v-j} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(t^j; w)\right)_{w=x} \\
&\quad + \left(\frac{d^r}{dw^r} L_n^{a,\beta}(\psi(t,x)(t-x)^{r+2}; w)\right)_{w=x} \\
&:= I_1 + I_2.
\end{aligned} \tag{33}$$

Proceeding along the lines of the estimate of I_2 of Theorem 2, it follows that for each $x \in (0, \infty)$

$$\lim_{n\to\infty} n\left(\frac{d}{dw}(L_n^{a,\beta}(\psi(t,x)(t-x)^{r+2}; w)\right)_{w=x} = 0.$$

Now, we estimate I_1.

$$\begin{aligned} I_1 &= \sum_{v=0}^{r-1} \frac{f^{(v)}(x)}{v!} \sum_{j=0}^{v} \binom{v}{j} (-x)^{v-j} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(t^j; w) \right)_{w=x} \\ &+ \frac{f^{(r)}(x)}{r!} \sum_{j=0}^{r} \binom{r}{j} (-x)^{r-j} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(t^j; w) \right)_{w=x} \\ &+ \frac{f^{(r+1)}(x)}{(r+1)!} \sum_{j=0}^{r+1} \binom{r+1}{j} (-x)^{r+1-j} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(t^j; w) \right)_{w=x} \\ &+ \frac{f^{(r+2)}(x)}{(r+2)!} \sum_{j=0}^{r+2} \binom{r+2}{j} (-x)^{r+2-j} \left(\frac{d^r}{dw^r} L_n^{a,\beta}(t^j; w) \right)_{w=x}. \end{aligned} \tag{34}$$

In view of Corollary 1 (iii), we have

$$\begin{aligned} I_1 &= \sum_{v=1}^{r-1} \frac{f^{(v)}(x)}{v!} O\left(\frac{1}{n}\right) + \frac{f^{(r)}(x)}{r!} \left(a_r(n) r! + O\left(\frac{1}{n}\right) \right) \\ &+ \frac{f^{(r+1)}(x)}{(r+1)!} \left((r+1) r! (-x) a_r(n) + a_{r+1}(n)(r+1)! x + O\left(\frac{1}{n}\right) \right) \\ &+ \frac{f^{(r+2)}(x)}{(r+2)!} \left((r+2)! \frac{x^2}{2} a_r(n) - (r+2)! x^2 a_{r+1}(n) + (r+2)! \frac{x^2}{2} a_{r+2}(n) + O\left(\frac{1}{n}\right) \right) \\ &= \sum_{v=1}^{r-1} f^{(v)}(x) O\left(\frac{1}{n}\right) + f^{(r)}(x) \left(a_r(n) + O\left(\frac{1}{n}\right) \right) \\ &+ f^{(r+1)}(x) \left((a_{r+1}(n) - a_r(n)) x + O\left(\frac{1}{n}\right) \right) \\ &+ \frac{f^{(r+1)}(x) x^2}{2} \left(a_r(n) - 2a_{r+1}(n) + a_{r+2}(n) + O\left(\frac{1}{n}\right) \right), \end{aligned} \tag{35}$$

where the coefficients of n^{-1} in $O\left(\frac{1}{n}\right)$ terms are certain rational functions of x depending on the parameter a.

Since $a_r(n) = \frac{n(n+1)(n+2)...(n+r-1)}{(n-1)(n-2)(n-3)...(n-r)}$, which implies that $a_{r+1}(n) - a_r(n) = a_r(n)\left(\frac{2r+1}{n-r-1}\right)$ and $a_r(n) - 2a_{r+1}(n) + a_{r+2}(n) = 2a_r(n)\left(\frac{n+(2r+1)^2}{(n-r-1)(n-r-2)}\right)$, it follows that

$$I_1 = f^{(r)}(x) + n^{-1}\left(\sum_{v=1}^{r+2} Q(v, r, a, \beta, x) f^{(v)}(x) + o(1) \right).$$

Combining the estimates of I_1 and I_2, we get the required result.

Corollary 3. *From the above lemma, we have for $r = 0$*

$$\lim_{n\to\infty} n(L_n^{a,\beta}(f;x) - f(x)) = \left(1 + x + \frac{a\beta x}{1+\beta x}\right) f'(x) + x\left[1 + x\frac{(1+\beta)}{2}\right] f''(x);$$

Next, we give an estimate of the degree of the approximation by $\left(\frac{d^r}{dw^r} L_n^{a,\beta}(.;w)\right)_{w=x}$ for smooth functions.

Theorem 4. *Let $r \le p \le r+2$ and $f \in \mathrm{L}$ be bounded on every finite sub-interval of $[0,\infty)$. Let $f(t) = O(t^\alpha)$ as $t \to \infty$ for some $\alpha > 0$. If $f^{(p)}$ exists and is continuous on $(a-\eta, b+\eta) \subset (0,\infty)$, $\eta > 0$, then for sufficiently large n*

$$\left\| \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f;w)\right)_{w=x} - f^{(r)}(x) \right\| \le C_1 n^{-1} \left(\sum_{i=1}^{p} \|f^{(i)}\|\right)$$
$$+ C_2 n^{-(p-r)/2} \omega_{f^{(p)}}(n^{-1/2}) + O(n^{-2}),$$

where C_1 and C_2 are both independent of f and n, $\omega_{f^{(p)}}(\delta)$ is the modulus of continuity of $f^{(p)}$ on $(a-\eta, b+\eta)$ and $\|.\|$ denotes the sup-norm on $[a,b]$.

Proof. By our hypothesis

$$f(t) = \sum_{i=0}^{p} \frac{f^{(i)}(x)}{i!}(t-x)^i + \frac{f^{(p)}(\xi) - f^{(p)}(x)}{p!}(t-x)^p \chi(t) + h(t,x)(1-\chi(t)), \tag{36}$$

where ξ lies between t and x and $\chi(t)$ is the characteristic function of $(a-\eta, b+\eta)$.

For $t \in (a-\eta, b+\eta)$ and $x \in [a,b]$, we have

$$f(t) = \sum_{i=0}^{p} \frac{f^{(i)}(x)}{i!}(t-x)^i + \frac{f^{(p)}(\xi) - f^{(p)}(x)}{p!}(t-x)^p.$$

For $t \in [0,\infty) \setminus (a-\eta, b+\eta)$ and $x \in [a,b]$, we define

$$h(t,x) = f(t) - \sum_{i=0}^{p} \frac{f^{(i)}(x)}{i!}(t-x)^i,$$

which implies that we can choose a $\delta > 0$ in such a way that $|t-x| \ge \delta$ for all $x \in [a,b]$ and thus there exists $\gamma > 0$ such that $|h(t,x)| \le M|t-x|^\gamma$, for some constant M.

Operating on the equality (36) by $\left(\frac{d^r}{dw^r}L_n^{a,\beta}(.;w)\right)_{w=x}$ and breaking the right-hand side into three parts I_1, I_2 and I_3, say, corresponding to the three terms on the right-hand side of (36), as in the estimate of I_1 of Theorem 3, we obtain

$$\|I_1\| \leq C_1 n^{-1}\left(\sum_{i=1}^{p}\|f^{(i)}\|\right) + O(n^{-2}),$$

uniformly in $x \in [a, b]$.

Applying Schwarz inequality, Lemma 2 and Corollary 2 (ii), it follows that for $s = 0, 1, 2, \ldots$

$$\sum_{k=0}^{\infty}|k-nx|^j v_{n,k}^{a,\beta}(x)\int_0^{\infty} b_{n,k}(t)|t-x|^s dt = O(n^{(j-s)/2}), \tag{37}$$

uniformly in $x \in [a, b]$.

Now, in order to estimate I_2, using $|f^{(p)}(\xi) - f^{(p)}(x)| \leq \left(1 + \frac{|t-x|}{\delta}\right)\omega_{f^{(p)}}(\delta)$, $\delta > 0$ and (37), we get $\|I_2\| \leq C_2 n^{-(p-r)/2}\omega_{f^{(p)}}(n^{-1/2})$, on choosing $\delta = n^{-1/2}$.

Finally, applying Cauchy–Schwarz inequality, Lemma 2 and Corollary 2 (ii) we obtain $\|I_3\| = O(n^{-s})$, for any $s > 0$, uniformly in $x \in [a, b]$.

Combining the estimates of I_1-I_3, the required result follows.

For sufficiently small $\eta > 0$, the Steklov mean $f_{\eta,2}$ of 2nd order corresponding to $f \in C_\gamma[0, \infty)$ and $t \in I_i = [a_i, b_i]$, $i = 1, 2$ is defined as follows:

$$f_{\eta,2}(t) = \eta^{-2}\int_{-\eta/2}^{\eta/2}\int_{-\eta/2}^{\eta/2}(f(t) - \Delta_h^2 f(t))dt_1 dt_2$$

where $h = \frac{t_1+t_2}{2}$ and Δ_h^2 is the second order forward difference operator with step length h. The following properties are satisfied (see [8, 10] and the references therein):

1. $f_{\eta,2}$ has continuous derivatives up to order 2 over I_1;
2. $\| f_{\eta,2}^{(r)} \|_{C(I_2)} \leq C\eta^{-r}\omega_r(f, \eta, I_2), r = 1, 2;$
3. $\| f - f_{\eta,2} \|_{C(I_2)} \leq C\omega_2(f, \eta, I_1);$
4. $\| f_{\eta,2} \|_{C(I_2)} \leq C \| f \|_{C(I_1)} \leq C||f||_\gamma,$

where C is a constant not necessarily the same at each occurrence and is independent of f and η.

Lemma 7 ([5]). *Let $f \in C(I)$. Then,*

$$\| f_{\eta,2k}^{(i)} \|_{C(I)} \leq C_i\{\| f_{\eta,2} \|_{C(I)} + \| f_{\eta,2}^{(2k)} \|_{C(I)}\},\ i = 1, 2, \ldots, 2k-1,$$

where C_i's are certain constants independent of f.

Theorem 5. *Let* $f \in C_\gamma[0,\infty)$ *for some* $\gamma > 0$ *and* $0 < a < a_1 < b_1 < b < \infty$. *Then for* α *sufficiently large, we have*

$$\left\| \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f;w) \right)_{w=x} - f^{(r)} \right\|_{C(I_1)} \leq K_1 \omega_2(f^{(r)}, \alpha^{-1/2}, I) + K_2 \alpha^{-1} \|f\|_\gamma,$$

where $K_1 = K_1(r)$ *and* $K_2 = K_2(r,f)$.

Proof. We can write

$$\begin{aligned} \left\| \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f;w) \right)_{w=x} - f^{(r)} \right\|_{C(I_1)} &\leq \left\| \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f - f_{\eta,2};w) \right)_{w=x} \right\|_{C(I_1)} \\ &\quad + \left\| \left(\frac{d^r}{dw^r} L_n^{a,\beta}(f_{\eta,2};w) \right)_{w=x} - f_{\eta,2}^{(r)} \right\|_{C(I_1)} \\ &\quad + \|f^{(r)} - f_{\eta,2}^{(r)}\|_{C(I_1)} \\ &= M_1 + M_2 + M_3. \end{aligned}$$

Since $f_{\eta,2}^{(r)} = (f^{(r)})_{\eta,2}$, hence by property (iii) of the Steklov mean, we get

$$M_3 \leq K_1 \omega_2(f^{(r)}, \eta, I).$$

Next, applying Theorem 3 and Lemma 7, we obtain

$$M_2 \leq K_2 \alpha^{-1} \sum_{i=r}^{r+2} \| f_{\eta,2}^{(i)} \|_{C(I_1)} \leq K_3 \alpha^{-1} \{ \| f_{\eta,2} \|_{C(I_1)} + \| f_{\eta,2}^{(r+2)} \|_{C(I_1)} \}.$$

By using properties (ii) and (iv) of Steklov mean, we get

$$M_2 \leq K_4 \alpha^{-1} \{ \| f \|_\gamma + \eta^{-2} \omega_2(f^{(r)}), \eta, I) \}.$$

Let a^* and b^* be such that $0 < a < a^* < a_1 < b_1 < b^* < b < \infty$ and I^* denote the interval $[a^*, b^*]$.

Now, we estimate M_1. Let $f - f_{\eta,2} \equiv F$. By our hypothesis we can write

$$F(t) = \sum_{m=0}^{r} \frac{F^{(m)}(x)}{m!}(t-x)^m + \frac{F^{(r)}(\xi) + F^{(r)}(x)}{r!}(t-x)^r \chi(t) + \theta(t,x)(1 - \chi(t)), \tag{38}$$

where ξ lies between t and x, and χ denotes the characteristic function of the interval I^*. For $t \in I^*$ and $x \in I_1$, we get

$$F(t) = \sum_{m=0}^{r} \frac{F^{(m)}(x)}{m!}(t-x)^m + \frac{F^{(r)}(\xi) + F^{(r)}(x)}{r!}(t-x)^r,$$

and for $t \in [0,\infty) \setminus I^*, x \in I_1$ we set

$$\theta(t,x) = F(t) - \sum_{m=0}^{r} \frac{F^{(m)}(x)}{m!}(t-x)^m.$$

Operating $D^r L_n^{a,\beta}$ on both sides of (38), we get three terms J_1, J_2, and J_3, corresponding to three terms in right-hand side of (38). Using Theorem 3, we get

$$|J_1| \leq \| f^{(r)} - f_{\eta,2}^{(r)} \|_{C(I_1)} .$$

Next, by using Theorem 3, we obtain

$$\begin{aligned} |J_2| &\leq \frac{2 \parallel F^{(r)} \parallel}{r!} L_\alpha^{\rho(r)}((t-x)^r \chi(t), x) \\ &\leq K_5 \parallel f^{(r)} - f_{\eta,2}^{(r)} \parallel_{C(I_*)} . \end{aligned}$$

Lastly, we easily have

$$|J_3| = \frac{d^r}{dx^r} L_n^{a,\beta}(1 - \chi(t)\theta(t,x), x) = O(\alpha^{-s}), \text{ for any } s > 0.$$

Combining $J_1 - J_3$, and from property (*iii*) of the Steklov mean, we obtain

$$M_1 \leq K_6 \parallel f^{(r)} - f_{\eta,2}^{(r)} \parallel_{C(I_*)} \leq K_6\, \omega_2(f^{(r)}, \eta, I).$$

Finally choosing $\eta = \alpha^{-1/2}$, the required result follows at once.

References

1. Acar, T., Gupta, V., Aral, A.: Rate of convergence for generalized Szász operators. Bull. Math. Sci. **1**(1), 99–113 (2011)
2. Chen, W.Z.: Approximation Theory of Operators. Xiamen University Publishing House, Xiamen (1989)
3. Erencin, A.: Durrmeyer type modification of generalized Baskakov operators. Appl. Math. Comput. **218**(3), 4384–4390 (2011)
4. Erencin, A., Bascanbaz-Tunca, G.: Approximation properties of a class of linear positive operators in weighted spaces. C. R. Acad. Bulgare Sci. **63**(10), 1397–1404 (2010)
5. Goldberg, S., Meir, V.: Minimum moduli of ordinary differential operators. Proc. Lond. Math. Soc. **23**(3), 1–15 (1971)
6. Gupta, V.: An estimate on the convergence of Baskakov–Bézier operators. J. Math. Anal. Appl. **312**(1), 280–288 (2005)
7. Gupta, V.: A note on modified Baskakov type operators. Approx. Theory Appl. **10**(3), 74–78 (1994)

8. Gupta, V., Agarwal, R.P.: Convergence Estimates in Approximation Theory. Springer, New York (2014)
9. Gupta, V., Srivastava, G.S.: Simultaneous approximation by Baskakov-Szász type operators. Bull. Math. de la Soc. Sci. de Roumanie (N. S.) **37** (85)(3–4), 73–85 (1993)
10. Hewitt, E., Stromberg, K.: Real and Abstract Analysis. McGraw Hill, New York (1956)
11. Ispir, N.: Rate of convergence of generalized rational type Baskakov operators. Math. Comput. Model. **46**, 625–631 (2007)
12. Karsli, H.: Rate of convergence of new gamma type operators for functions with derivatives of bounded variation. Math. Comput. Model. **45**, 617–624 (2007)
13. Mihesan, V.: Uniform approximation with positive linear operators generated by generalized Baskakov method. Autom. Comput. Appl. Math. **7**(1), 34–37 (1998)
14. Özarslan, M.A., Duman, O., Kannağlu, C.: Rates of convergence of certain King-type operators for functions with derivative of bounded variation. Math. Comput. Model. **52**, 334–345 (2010)
15. Verma, D.K., Gupta, V., Agrawal, P.N.: Some approximation properties of Baskakov-Durrmeyer-Stancu operators. Appl. Math. Comput. **218**(11), 6549–6556 (2012)

A Tour on $p(x)$-Laplacian Problems When $p = \infty$

Yiannis Karagiorgos and Nikos Yannakakis

Abstract Most of the times, in problems where the $p(x)$-Laplacian is involved, the variable exponent $p(\cdot)$ is assumed to be bounded. The main reason for this is to be able to apply standard variational methods. The aim of this paper is to present the work that has been done so far, in problems where the variable exponent $p(\cdot)$ equals infinity in some part of the domain. In this case the infinity Laplace operator arises naturally and the notion of weak solution does not apply in the part where $p(\cdot)$ becomes infinite. Thus the notion of viscosity solution enters into the picture. We study both the Dirichlet and the Neumann case.

Keywords $p(x)$-Laplacian • Viscosity solution • Infinity harmonic function • Dirichlet problem • Neumann problem • Lipschitz constant

Introduction

In this paper we study problems concerning the equation

$$\Delta_{p(x)} u = 0, \quad \text{in } \Omega, \tag{1}$$

where $\Omega \subset \mathbb{R}^N$ is a bounded smooth domain,

$$\Delta_{p(x)} u := \operatorname{div}(|\nabla u|^{p(x)-2} \nabla u),$$

is the $p(x)$-Laplace operator and the variable exponent $p(\cdot)$ satisfies

$$p|_D = \infty, \tag{2}$$

in a subdomain D of Ω.

Y. Karagiorgos • N. Yannakakis (✉)
Department of Mathematics, National Technical University of Athens,
Zografou Campus, Athens 15780, Greece
e-mail: ykarag@math.ntua.gr; nyian@math.ntua.gr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_15

Manfredi et al. in [20] considered condition (2) for the first time for the Dirichlet problem with Lipschitz boundary conditions. Working in the same direction the authors in [17] studied the corresponding Neumann problem. Our purpose here is to briefly present these two cases. We will also mention some further results concerning the Dirichlet case, studied in [28].

Recently, the case of $p(x) \to \infty$ has been studied in several problems where the $p(x)$-Laplacian is involved. See, for instance, [21] or [24] and [25]. On the other hand, when p is constant the case of $p \to \infty$ in problems with the p-Laplacian was first studied in [5], in which the physical motivation was given as well. On both cases the notion of infinity Laplacian arises naturally as the limit case.

Partial Differential Equations involving the $p(x)$-Laplacian appear in a variety of applications. In [7] the authors proposed a framework for image restoration based on a variable exponent Laplacian. This was the starting point for the research on the connection between PDEs with variable exponents and image processing. Recently there has been quite a rapid progress in this direction.[1] Other applications that use variable exponent type Laplacians are elasticity theory and the modelling of electrorheological fluids (see [26]).

Under these preliminaries, the structure of our paper is as follows. In section "Preliminaries," we state some general definitions and propositions concerning the tools to be needed for the main results. Section "The Dirichlet Case" is devoted to the study of Dirichlet problem and a discussion of the results in [20]. Furthermore, we briefly present some interesting results of [28] concerning the emptiness and non-emptiness of the functional set S introduced there. Finally, in section "The Neumann Case" we study the corresponding Neumann problem [17].

Preliminaries

In this section we summarize some basic properties of the variable exponent Lebesgue and Sobolev spaces. For details the interested reader is referred to [13, 19] and [10].

Let $\Omega \subset \mathbb{R}^N$ and denote by $L^0(\Omega)$ the space of real valued measurable functions in Ω. If $p\colon \Omega \to [1,\infty]$ is a measurable function, we define a semimodular $\varrho_{L^{p(\cdot)}}$ on $L^0(\Omega)$ by setting

$$\varrho_{L^{p(\cdot)}}(u) := \int_{\Omega} |u(x)|^{p(x)} dx\,;$$

where we use the convention $t^{\infty} = \infty\chi_{(1,\infty]}(t)$. We define the variable exponent Lebesgue space as

[1] The reader can visit the website http://www.helsinki.fi/~pharjule/varsob/index.shtml for further details.

$$L^{p(\cdot)}(\Omega) = \left\{ u \in L^0(\Omega) : \varrho_{L^{p(\cdot)}}(\lambda u) < \infty, \text{ for some } \lambda > 0 \right\}$$

equipped with the norm

$$\|u\|_{L^{p(\cdot)}(\Omega)} := \|u\|_{p(\cdot)} = \inf \left\{ \lambda > 0 : \varrho_{L^{p(\cdot)}}\left(\frac{u}{\lambda}\right) \leq 1 \right\}.$$

The variable exponent Sobolev space is defined by

$$W^{1,p(\cdot)}(\Omega) = \left\{ u \in L^{p(\cdot)}(\Omega) : \nabla u \in L^{p(\cdot)}(\Omega, \mathbb{R}^N) \right\}$$

with norm

$$\|u\|_{W^{1,p(\cdot)}(\Omega)} = \|u\|_{L^{p(\cdot)}(\Omega)} + \|\nabla u\|_{L^{p(\cdot)}(\Omega,\mathbb{R}^N)}.$$

Remark 1. For the definition of the variable exponent Lebesgue space, $L^{p(\cdot)}$, we follow [10, Chap. 3] using a semimodular instead of the modular given in [19]. Nevertheless the two definitions are equivalent.

The spaces $L^{p(\cdot)}(\Omega)$, $W^{1,p(\cdot)}(\Omega)$ are Banach spaces and if

$$1 < p_- := \operatorname{ess}\inf_{x\in\Omega} p(x) \leq p^+ := \operatorname{ess}\sup_{x\in\Omega} p(x) < \infty,$$

they are also separable and reflexive.

We also define the variable exponent *Sobolev space with zero boundary values*, $W_0^{1,p(\cdot)}(\Omega)$ as the closure of the set of compactly supported $W^{1,p(\cdot)}(\Omega)$-functions (see [10, Definition 8.1.10]). When p is bounded and smooth functions are dense in $W^{1,p}(\Omega)$, we could equivalently define $W_0^{1,p(\cdot)}(\Omega)$ as the closure of $C_0^\infty(\Omega)$.

When p is constant, it is well known that smooth functions are dense in $W^{1,p}(\Omega)$. This is no longer true when we deal with variable exponent spaces, (see [12, 27] and [10]). In fact, we have to consider additional conditions for the variable exponent. The most natural one is the so-called *log-Hölder continuity*: there exists $C > 0$, such that

$$|p(x) - p(y)| \leq \frac{C}{\log(e + \frac{1}{|x-y|})}, \quad \text{for } x, y \in \Omega.$$

Log-Hölder continuity is also a sufficient condition for the variable exponent version of Poincaré's inequality, i.e.,

$$\|u\|_{L^{p(\cdot)}(\Omega)} \leq C(N, p)\,\text{diam}(\Omega)\|\nabla u\|_{L^{p(\cdot)}(\Omega,\mathbb{R}^N)}, \quad \text{for all } u \in W_0^{1,p(\cdot)}(\Omega).$$

On the next Proposition we state another sufficient condition for the Poincaré inequality, the so-called *jump condition* (see [15] and [10, Theorem 8.2.18]).

Proposition 1. *Let $\Omega \subset \mathbb{R}^N$ and $p\colon \Omega \to [1,\infty]$ a measurable function. If there exists $\delta > 0$ such that for every $x \in \Omega$ either*

$$p^+_{B(x,\delta)\cap\Omega} \le \frac{Np^-_{B(x,\delta)\cap\Omega}}{N - p^-_{B(x,\delta)\cap\Omega}} \quad or \quad p^-_{B(x,\delta)\cap\Omega} \ge N \tag{3}$$

then the Poincaré inequality holds in $W_0^{1,p(\cdot)}(\Omega)$.

Remark 2. The Poincaré inequality plays a crucial role in the proof of Lemma 1 in obtaining the coercivity of the functional I_k.

Proposition 2. *Let $p\colon \Omega \to \mathbb{R}$ be a measurable function. The dual space of $L^{p(\cdot)}(\Omega)$ is the space $L^{q(\cdot)}(\Omega)$, where $\frac{1}{p(\cdot)} + \frac{1}{q(\cdot)} = 1$ in Ω. Also, the variable exponent version of Hölder's inequality holds, namely*

$$\int_\Omega |u(x)v(x)|\, dx \le 2\|u\|_{p(\cdot)}\|v\|_{q(\cdot)}, \quad \text{forall } u \in L^{p(\cdot)}(\Omega), v \in L^{q(\cdot)}(\Omega).$$

The next proposition is very useful, since it allows us to compare the norm $\|u\|_{p(\cdot)}$ with the integral $\int_\Omega |u(x)|^{p(x)}dx$.

Proposition 3. *Let $u \in L^{p(\cdot)}(\Omega)$.*

(i) If $\|u\|_{p(\cdot)} > 1$, then

$$\|u\|^{p-}_{p(\cdot)} \le \int_\Omega |u(x)|^{p(x)}dx \le \|u\|^{p^+}_{p(\cdot)}.$$

(ii) If $\|u\|_{p(\cdot)} < 1$, then

$$\|u\|^{p^+}_{p(\cdot)} \le \int_\Omega |u(x)|^{p(x)}dx \le \|u\|^{p-}_{p(\cdot)}.$$

(iii) $\|u\|_{p(\cdot)} = 1 \Leftrightarrow \int_\Omega |u(x)|^{p(x)}dx = 1$.

Let $p(\cdot)$ be a variable exponent in Ω which satisfies the following hypothesis

$$p|_D = \infty,$$

where D is a subdomain of Ω. Moreover, $p \in C^1(\Omega \setminus \overline{D})$ with

$$p^+ := \sup_{x\in\Omega\setminus\overline{D}} p(x) < \infty \tag{4}$$

and

$$p_- := \inf_{x \in \Omega} p(x) > N. \tag{5}$$

For $k \in \mathbb{N}$, consider $p_k(\cdot)$ such that

$$p_k(x) = \min\{p(x), k\}. \tag{6}$$

Then $p_k(x) \to p(x)$, as $k \to \infty$, while for $k > p^+$ we have that

$$p_k(x) = \begin{cases} p(x), & x \in \overline{\Omega} \setminus D \\ k, & x \in D. \end{cases} \tag{7}$$

Remark 3. The sequence $(p_k(\cdot))$ has been introduced in [20]. The reason is the following: let u_k be the unique solution of the corresponding $p_k(\cdot)$-problem. If we assume for a moment that the limit $\lim_{k\to\infty} u_k$ exists, then it is a candidate solution for the original $p(\cdot)$-problem.

Proposition 4. *The space $C^\infty(\overline{\Omega})$ is dense in $W^{1,p_k(\cdot)}(\Omega)$.*

Proof. This is straightforward, if we use Theorem 9.3.5 of [10, p. 298] with $\Omega_1 = \Omega \backslash \overline{D}$ and $\Omega_2 = D$, where each of $\Omega_i, i = 1, 2$ has Lipschitz boundary. □

Remark 4. The above Proposition is instrumental for both the Dirichlet and the Neumann case. For the Dirichlet case it provides us with the density of the set $C_0^\infty(\Omega)$ in $W_0^{1,p_k(\cdot)}(\Omega)$ which is necessary for the weak formulation of problem (P_k). Also it provides us with the direct method of calculus of variations which we use in Lemmas 1 and 3. Finally it is needed at Lemmas 2 and 4.

The next proposition is needed for the Neumann case.

Proposition 5. *There exists $C > 0$ such that the following Poincaré type inequality holds*

$$\|u\|_{1,p_k(\cdot)} \leq C\|\nabla u\|_{L^{p_k(\cdot)}}, \quad \text{for all } u \in W^{1,p_k(\cdot)}(\Omega) \text{ s.t} \int_\Omega u = 0. \tag{8}$$

Proof. Apply Theorem 8.2.17 in [10, p. 256] with $D_1 = \Omega \backslash D, D_2 = D$. Then,

$$(p_k|_{D_1})_- := \inf_{x \in D_1} p(x) \geq p_- > N$$

and

$$(p_k|_{D_2})_- = k > N.$$

□

Proposition 6. *Let p be a variable exponent such that $p_- > N$. Then the following holds*

(i)

$$W^{1,p(\cdot)}(\Omega) \hookrightarrow W^{1,p_-}(\Omega) \hookrightarrow C(\overline{\Omega}). \tag{9}$$

(ii) If $q \in C(\partial\Omega)$, *the embedding*

$$W^{1,p(\cdot)}(\Omega) \hookrightarrow L^{q(\cdot)}(\partial\Omega), \tag{10}$$

is compact and continuous.

For reference, see [10, 13, 19] for (i) and [30, Proposition 2.6] for (ii).

Remark 5. From (i) of Proposition 6 we know that we always deal with continuous functions. Also in the Dirichlet case we ensure that we can check the boundary data pointwise. Note that (ii) is needed for the Neumann case in Proposition 10, with

$$(p_k)_- > N \quad \text{and} \quad p_k|_{\partial\Omega} = p \in C(\partial\Omega).$$

The notion of viscosity solution is of great importance throughout this paper, since the notion of weak solution does not apply in some cases. Here, we give the classical definition. Later in the following sections we give the ones that fit to each problem. For reference see [8].

Let $F\colon \overline{\Omega}\times\mathbb{R}^N\times\mathbb{S}^N \to \mathbb{R}$, where $\mathbb{S}^N$ is the space of the N-dimensional symmetric matrices. F is said to be degenerate elliptic, if for each $X, Y \in \mathbb{S}^N$ with $X \geq Y$ (i.e. $\langle X\xi, \xi\rangle \geq \langle Y\xi, \xi\rangle$, for $\xi \in \mathbb{R}^N$), then

$$F(x, \xi, X) \leq F(x, \xi, Y).$$

We are ready to give the definition of a viscosity solution based on pointwise evaluation of the operator F. Recall that by D^2u we denote the Hessian matrix of u.

Definition 1. Let $F\colon \Omega \times \mathbb{R}^N \times \mathbb{S}^N \to \mathbb{R}$ be a degenerate elliptic operator and $u\colon \Omega \to \mathbb{R}$.

(i) Let u be a lower semicontinuous function in Ω. We say that u is a viscosity supersolution of $F(x, \nabla u, D^2u) = 0$ in Ω, if for every $\varphi \in C^2(\Omega)$ such that $u - \varphi$ attains its strict minimum at $x_0 \in \Omega$ with $u(x_0) = \varphi(x_0)$, we have

$$F(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \geq 0.$$

(ii) Let u be an upper semicontinuous function in Ω. We say that u is a viscosity subsolution of $F(x, \nabla u, D^2u) = 0$ in Ω, if for every $\varphi \in C^2(\Omega)$ such that $u-\varphi$ attains its strict maximum at $x_0 \in \Omega$ with $u(x_0) = \varphi(x_0)$, we have

$$F(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \leq 0.$$

(iii) We say that $u \in C(\Omega)$ is a viscosity solution of $F(x, \nabla u, D^2u) = 0$ in Ω, if it is both a viscosity supesolution and subsolution.

Next we turn our attention to the infinity Laplace operator

$$\Delta_\infty u := \sum_{i,j=1}^{N} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \frac{\partial^2 u}{\partial x_i x_j};$$

which appears naturally in the part of the domain where the variable exponent $p(\cdot)$ equals infinity. It was discovered by G. Aronsson (see [1, 2]) when he studied the Lipschitz extension problem as a limit procedure (as $p \to \infty$) for the variational problem of minimizing the functional

$$I_p(u) = \int_\Omega |\nabla u(x)|^p dx.$$

Aronsson showed that there is a connection between optimal Lipschitz extensions and the solutions of the equation

$$-\Delta_\infty u = 0$$

in the classical sense. In particular he proved that if Ω is a bounded domain of $\mathbb{R}^N$ and $g \in C(\partial\Omega)$, then a classical solution u of the Dirichlet problem

$$\begin{cases} -\Delta_\infty u = 0, & \text{in } \Omega \\ \quad u = g, & \text{in } \partial\Omega \end{cases} \tag{11}$$

is also an *absolutely minimizing Lipschitz extension* (AMLE for short). The latter means, that for every bounded subset V of Ω and for each $v \in C(\overline{V})$, with $u = v$ in ∂V, one has

$$\|\nabla u\|_{L^\infty(V)} \leq \|\nabla v\|_{L^\infty(V)}. \tag{12}$$

When the viscosity theory appeared, Jensen in [16] proved that problem (11) has a unique viscosity solution $u \in C(\overline{\Omega})$, which is also an absolutely minimizing Lipschitz extension. Crandall, Evans and Gariepy (see [9] or the survey paper [3]) used the so-called *comparison with cones principle* and proved the above connection in a direct way (without using the approximation procedure).

Definition 2. We say that $u \in C(\Omega)$ is infinity harmonic (in the viscosity sense) in Ω, if it is a viscosity solution of the equation

$$-\Delta_\infty u = 0, \quad \text{in } \Omega.$$

Smooth solutions of the equation $\Delta_\infty u = 0$ have some nice geometric properties concerning their gradient which are very useful in image processing (see [6]). Infinity harmonic functions appear also in optimal transportation (see [11, 14]), and more recently in tug of war games (see [23]).

The Dirichlet Case

In this section we focus on the Dirichlet case. In particular we present the results of [20, 28]. Consider the following problem

$$\begin{cases} -\Delta_{p(x)}u(x) = 0, & x \in \Omega \\ \qquad u(x) = f(x), & x \in \partial\Omega \end{cases} \tag{P$_1$}$$

where $\Omega \subset \mathbb{R}^N$ is a bounded smooth domain and $N \geq 2$.

$$\Delta_{p(x)}u := \operatorname{div}(|\nabla u|^{p(x)-2}\nabla u)$$

is the $p(x)$-Laplacian operator which is the variable exponent version of the p-Laplacian. Also, the function $f : \partial\Omega \to \mathbb{R}$ appearing in the boundary condition is Lipschitz continuous on $\partial\Omega$. The variable exponent p satisfies the following hypothesis

$$p|_D = \infty$$

where D is a convex subdomain of Ω with C^1 boundary. Also, $p \in C^1(\Omega \setminus \overline{D})$ with

$$p^+ := \sup_{x \in \Omega \setminus \overline{D}} p(x) < \infty \tag{13}$$

and

$$p_- := \inf_{x \in \Omega} p(x) > N \tag{14}$$

Remark 6. For a C^2 function u, if we compute the $p(x)$-Laplacian we get

$$\begin{aligned} -\Delta_{p(x)}u(x) =& -|\nabla u(x)|^{p(x)-2}\Delta u(x) \\ &= -(p(x)-2)|\nabla u(x)|^{p(x)-4}\Delta_\infty u(x) \\ &\quad - |\nabla u(x)|^{p(x)-2}\ln(|\nabla u(x)|)\nabla u(x) \cdot \nabla p(x). \end{aligned}$$

Due to the appearance of the term $\nabla p(x)$ we need $p \in C^1(\Omega \setminus \overline{D})$ to ensure that the map $x \mapsto -\Delta_{p(x)}u(x)$ is continuous.

Remark 7. The convexity of D is needed so that the set of Lipschitz functions on D and $W^{1,\infty}(D)$ coincide and the Lipschitz constant is the L^∞- norm of their gradient.

Remark 8. Condition (14) guarantees that the embedding

$$W^{1,p(\cdot)}(\Omega) \hookrightarrow W^{1,p_-}(\Omega) \hookrightarrow C(\overline{\Omega})$$

holds (see (9)) and so we are considering the boundary condition in the classical sense.

To find out what a solution of (P_1) might be, the authors in [20] considered a sequence of bounded variable exponents p_k such that

$$p_k(x) = \min\{p(x), k\}. \tag{15}$$

Then $p_k(x) \to p(x)$ as $k \to \infty$, while for $k > p^+$ we have that

$$p_k(x) = \begin{cases} p(x), & x \in \overline{\Omega} \setminus D \\ k, & x \in D. \end{cases} \tag{16}$$

Remark 9. Note that for $k > p^+$, the boundary of the set $\{x : p(x) > k\}$ coincides with the boundary of D and so is independent of k. Due to this fact we have no problems when passing to the limit as $k \to \infty$.

If we replace p with p_k in problem (P_1) we have the intermediate boundary value problems,

$$\begin{cases} -\Delta_{p_k(x)} u(x) = 0, & x \in \Omega \\ \qquad u(x) = f(x), & x \in \partial\Omega. \end{cases} \tag{P_k}$$

Using standard variational methods, one can obtain the existence of a unique weak solution u_k for problem (P_k). In this case, $\lim_k u_k$ is a natural candidate as a solution to the problem (P_1). Nevertheless the existence of the above limit is not a trivial issue. The following questions arise naturally :

- When does this limit exist?
- What is the equation that this limit should satisfy?

The set

$$S = \left\{ u \in W^{1,p-}(\Omega) : u|_{\Omega \setminus \overline{D}} \in W^{1,p(x)}(\Omega \setminus \overline{D}), \|\nabla u\|_{L^\infty(D)} \leq 1 \text{ and } u|_{\partial\Omega} = f \right\}$$

will play crucial role for the first question, while the infinity Laplace operator

$$\Delta_\infty u := \sum_{i,j=1}^{N} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \frac{\partial^2 u}{\partial x_i x_j}.$$

for the second one.

We proceed with two subsections. On the first one we give a brief presentation of the results of [20] and on the second one we discuss the results of [28].

Main Results

Our aim here is to state the main results of [20], and try to point out the basic tools and ideas that are necessary. (For a nice analytic presentation of [20], see also [18, Sect. 4].) We also give a sketch for some proofs. As we discussed in the previous section a natural candidate for a solution to problem (P_1) is the limit $\lim_k u_k$, where u_k is the unique solution to problem (P_k). Before we study the existence of the above limit, we have to deal with some independent results concerning the intermediate problems (P_k). The following Lemmas are in this direction. Recall that a function $u \in W^{1,p_k(\cdot)}(\Omega)$ is a weak solution to problem (P_k) if $u|_{\partial\Omega} = f$ and

$$\int_\Omega |\nabla u|^{p_k(x)-2}\nabla u \cdot \nabla v dx = 0, \quad \text{for all} \quad v \in C_0^\infty(\Omega)\,. \tag{17}$$

Lemma 1. *There exists a unique weak solution u_k to problem (P_k), which is the unique minimizer of the functional*

$$I_k(u) = \int_\Omega \frac{|\nabla u|^{p_k(x)}}{p_k(x)}\,dx$$

in the set

$$S_k = \left\{ u \in W^{1,p_k(\cdot)}(\Omega) : u|_{\partial\Omega} = f \right\}.$$

Proof. [Sketch] The basic tool to apply the direct method of calculus of variations for functional I_k is Poincaré's inequality and the density of $C_0^\infty(\Omega)$ functions in $W_0^{1,p_k(\cdot)}(\Omega)$ (see Remarks 2 and 4). For $k > p^+$ the variable exponent $p_k(\cdot)$ satisfies

$$p_k(\cdot) \geq (p_k)_- \geq p_- > N \tag{18}$$

and the jump condition (1) is true. Thus, the Poincaré inequality holds. Also from Proposition 4 smooth functions are dense $W^{1,p_k(\cdot)}(\Omega)$. Hence $C_0^\infty(\Omega)$ is dense in $W_0^{1,p_k(\cdot)}(\Omega)$. Finally we can check the boundary condition pointwise since from (18) we have that

$$W^{1,p_k(\cdot)}(\Omega) \hookrightarrow W^{1,p_-}(\Omega) \hookrightarrow C(\overline{\Omega}). \tag{19}$$

□

The next Lemma is very useful since it provides us with a problem that has the same weak solutions as problem (P_k) but which allows us to take separate cases.

Lemma 2. *The following problem*

$$\begin{cases} -\Delta_{p_k(x)} u(x) = 0, & x \in \Omega \setminus \overline{D} \\ -\Delta_k u(x) = 0, & x \in D \\ |\nabla u(x)|^{k-2} \frac{\partial u}{\partial \nu}(x) = |\nabla u(x)|^{p(x)-2} \frac{\partial u}{\partial \nu}(x), & x \in \partial D \cap \Omega \\ u(x) = f(x), & x \in \partial \Omega. \end{cases} \tag{20}$$

has the same weak solutions as problem (P_k)*, where* ν *is the exterior unit normal to* ∂D *in* Ω.

Remark 10. To conclude that the weak formulation of (20) is the same with the one of problem (P_k) we multiply by a test function $v \in C_0^{\infty}(\Omega)$ and use integration by parts and the Gauss–Green theorem.

Next we state the definition of viscosity solutions that correspond to problems like (20) and (25) which involve a transmission condition. For reference see [8].

Definition 3. Let $F_1, F_2 \colon \overline{\Omega} \times \mathbb{R}^N \times \mathbb{S}^N \to \mathbb{R}$, $B \colon \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be degenerate elliptic operators and $u \colon \overline{\Omega} \to \mathbb{R}$. Consider the following problem

$$F_1(x, \nabla u, D^2 u) = 0, \qquad \text{in} \quad \Omega \setminus \overline{D} \tag{21}$$

$$F_2(x, \nabla u, D^2 u) = 0, \qquad \text{in} \quad D \tag{22}$$

with transmission condition

$$B(x, u, \nabla u) = 0, \qquad \text{on} \quad \partial D \cap \Omega \tag{23}$$

and boundary condition

$$u = f, \qquad \text{on} \quad \partial \Omega. \tag{24}$$

(i) Let u be a lower semicontinuous function in $\overline{\Omega}$. We say that u is a viscosity supersolution of problem (21)–(24) if for every $\varphi \in C^2(\overline{\Omega})$ such that $u - \varphi$ attains its strict minimum at $x_0 \in \overline{\Omega}$ with $u(x_0) = \varphi(x_0)$, we have

- if $x_0 \in \Omega \setminus \overline{D}$, then $F_1(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \geq 0$.
- If $x_0 \in D$, then $F_2(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \geq 0$.
- If $x_0 \in \partial D \cap \Omega$, then

$$\max\{F_1(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)), F_2(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)), \\ B(x_0, \varphi(x_0), \nabla\varphi(x_0))\} \geq 0.$$

(ii) Let u be an upper semicontinuous function in $\overline{\Omega}$. We say that u is a viscosity subsolution of problem (21)–(24), if for every $\varphi \in C^2(\overline{\Omega})$ such that $u - \varphi$ attains its strict maximum at $x_0 \in \overline{\Omega}$ with $u(x_0) = \varphi(x_0)$, we have

- if $x_0 \in \Omega \backslash \overline{D}$, then $F_1(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \leq 0$.
- If $x_0 \in D$, then $F_2(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)) \leq 0$.
- If $x_0 \in \partial D \cap \Omega$, then

$$\max\{F_1(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)), F_2(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)), \\ B(x_0, \nabla\varphi(x_0), D^2\varphi(x_0)\} \leq 0.$$

(iii) Finally, u is called a viscosity solution of problem (21)–(24), if it is both a viscosity subsolution and a viscosity supersolution.

The next proposition is of great importance when passing to the limit.

Proposition 7. *Let u be a continuous weak solution of (P_k). Then u is a solution of (P_k) in the viscosity sense.*

Remark 11. For the proof we have to verify Definition 3 with

$$F_1(x, \nabla\varphi(x), D^2\varphi(x)) = -\Delta_{p(x)}\varphi(x), \quad F_2(x, \nabla\varphi(x), D^2\varphi(x)) = -\Delta_k\varphi(x)$$

and

$$B(x, \varphi(x), \nabla\varphi(x)) = |\nabla\varphi(x)|^{k-2}\frac{\partial\varphi}{\partial\nu}(x) - |\nabla\varphi(x)|^{p(x)-2}\frac{\partial\varphi}{\partial\nu}(x),$$

where $\varphi \in C^2(\overline{\Omega})$. Recall also that for $x \in \Omega \setminus D$ the map $x \mapsto -\Delta_{p(x)}\varphi(x)$ is continuous and this is also needed for the proof.

Proposition 8. *Assume that S is non-empty and let u_k be the minimizer of I_k in S_k. If $v \in S$, then*

$$I_k(u_k) = \int_\Omega \frac{|\nabla u_k|^{p_k(x)}}{p_k(x)}\,dx \leq \int_D \frac{|\nabla v|^k}{k}\,dx + \int_{\Omega\backslash\overline{D}} \frac{|\nabla v|^{p(x)}}{p(x)}\,dx\,.$$

Hence the sequence $(I_k(u_k))_k$ is uniformly bounded and $(u_k)_k$ is uniformly bounded in $W^{1,p-}(\Omega)$ and equicontinuous in $C(\overline{\Omega})$.

Remark 12. The crucial step for the proof of Proposition 8 is to compare the norm $\|\nabla u_k\|_{L^p(\Omega,\mathbb{R}^N)}$ with $I_k(u_k)$ and give a uniform bound. Then the result follows using Poincaré's inequality for the norm $\|u_k - f\|_{W_0^{1,p-}(\Omega)}$.

Remark 13. The previous proposition is instrumental for the limit case, since it allows the use of Arzellà–Ascoli theorem which will play a crucial role in later results.

Now we have all the ingredients we need to study the limit case. We first treat the case $S = \emptyset$.

Theorem 1. *Assume that $\partial\Omega \cap \overline{D} \neq \emptyset$ and the Lipschitz constant of $f_{\partial\Omega\cap\overline{D}}$ is strictly greater than 1. Then*

$$\lim_{n\to\infty} \inf(I_k(u_k))^{\frac{1}{k}} > 1.$$

Hence, $I_k(u_k) \to \infty$ and the natural energy associated with u_k is unbounded.

Proof. [Sketch] Arguing by contradiction, suppose that

$$\lim_{n\to\infty} \inf(I_k(u_k))^{\frac{1}{k}} \leq 1.$$

Then one can construct a function $u \in W^{1,\infty}(D)$ which is a Lipschitz extension of $f_{\partial\Omega\cap\overline{D}}$ to D, such that $\|\nabla u\|_{L^\infty(D)} \leq 1$. This contradicts the hypotheses, since a Lipschitz extension cannot decrease the Lipschitz constant. □

Remark 14. In the case of $\partial\Omega \cap \overline{D} \neq \emptyset$, the condition that the Lipschitz constant of $f_{\partial\Omega\cap\overline{D}}$ is strictly greater than 1 (i.e, $Lip(f, \partial\Omega \cap \overline{D}) > 1$) is sufficient for the emptiness of the set S. Indeed, any Lipschitz extension u of $f|_{\partial\Omega\cap\overline{D}}$ to D satisfies $\|\nabla u\|_{L^\infty(D)} > 1$.

Next, we focus on the case $S \neq \emptyset$. In this case the next theorem guarantees that the uniform limit, $u_\infty := \lim_k u_k$ exists and is some kind of solution to the original problem (P_1). A characterization of general conditions that guarantee the non-emptiness of S remains an open problem. We state some trivial sufficient conditions for the non-emptiness of S.

- If $Lip(f, \partial\Omega) \leq 1$, then we obtain an element in S by taking the McShane–Whitney extension (see [22, 29]) of f in $\overline{\Omega}$. Hence $S \neq \emptyset$.
- If $\partial\Omega \cap \overline{D} = \emptyset$ (i.e, D is compactly supported in Ω), then $S \neq \emptyset$. Indeed we can obtain an element of S by assuming a constant function u in D which satisfies $u|_{\partial\Omega} = f$ and extending it (as a Lipschitz function) to Ω.

So far we have one sufficient condition for the emptiness of S (see Remark 14) and two for the non-emptiness of S. The tricky case is when $\partial\Omega \cap \overline{D} \neq \emptyset$ and the Lipschitz constant of f satisfies

$$Lip(f, \partial\Omega \cap \overline{D}) \leq 1, \quad \text{and} \quad Lip(f, \partial\Omega) > 1.$$

The authors in [20] constructed a counterexample showing that the above condition is necessary but not sufficient for the non-emptiness of S. The authors in [28] studied this case exhaustively where they tried to identify conditions that guarantee the emptiness (and the non-emptiness) of S. We study this case in the next section.

We return now on the main purpose of [20] which is to figure out what a solution of (P_1) might be and what is the problem (as a limit case) that it solves.

For the next theorem, consider the functional $I\colon S \to \mathbb{R}$, such that

$$I(v) := \int_{\Omega\setminus\overline{D}} \frac{|\nabla v|^{p(x)}}{p(x)}\,dx, \quad \text{for each } v \in S \subset S_k.$$

Theorem 2. *Let u_k be the unique minimizer of I_k in S_k. Then there exists a function $u_\infty \in S$, such that u_∞ minimizes I in S and is also infinity harmonic in D.*

Proof. [Sketch] From Proposition 8 there exists a function $u_\infty \in C(\overline{\Omega})$ such that along subsequences we have

$$u_k \to u_\infty\,, \quad \text{uniformlyin} \quad \overline{\Omega};$$

$$u_k \overset{w}{\to} u_\infty\,, \quad \text{in} \quad W^{1,p-}(\Omega);$$

and

$$u_k \overset{w}{\to} u_\infty\,, \quad \text{in} \quad W^{1,p(\cdot)}(\Omega\setminus\overline{D}).$$

Also u_∞ is a minimizer of I in S. To prove that u_∞ is infinity harmonic in D we use the fact that u_k is k-harmonic in D and $u_k \to u_\infty$ uniformly in $\overline{\Omega}$ (see [5, 16, Proposition 2.2] or Theorem 2.8 in [18, p. 17]). □

Remark 15. The uniform convergence of (u_k) to u_∞ allows us to find a sequence (x_k) in $\Omega\setminus\overline{D}$ such that $x_k \to x_0$ and $u_k - \varphi$ attains its strict minimum at x_k (see, for instance, [5, Proposition 2.2] or Theorem 2.8 in [18, p. 17]). This is necessary when one wants to prove that a uniform limit of k-harmonic functions is ∞-harmonic. This fact plays also a key role in the proof of Theorem 3.

Remark 16. The minimizer u_∞ is unique in the following sense. Any other minimizer of I in S that is infinity harmonic in D coincides with u_∞. This is a straightforward result due to the strict convexity of I in S and the uniqueness of the Dirichlet problem for the infinity Laplacian due to R. Jensen (see [16]).

In the following theorem we state the problem of which u_∞ is a viscosity solution.

Theorem 3. *Let (u_k) be the sequence of solutions of problems (P_k) and u_∞ the uniform limit of a subsequence of (u_k). Then u_∞ is a viscosity solution of the following problem*

$$\begin{cases} -\Delta_{p(x)}u(x) = 0, & x \in \Omega\setminus\overline{D} \\ -\Delta_\infty u(x) = 0, & x \in D \\ sgn(|\nabla u(x)| - 1)sgn(\frac{\partial u}{\partial \nu}(x)) = 0, & x \in \partial D \cap \partial\Omega \\ u(x) = f(x), & x \in \partial\Omega \end{cases} \tag{25}$$

Proof. [Sketch] For the proof one has to verify Definition 3. The uniform convergence of (u_k) to u_∞ plays again a crucial role (see Remark 15). Also Lemma 2 and Proposition 7 together with the fact that for $\varphi \in C^2(\overline{\Omega})$, the mapping $x \mapsto -\Delta_{p(x)}\varphi(x)$ is continuous in $\Omega \setminus D$ allow us to pass to the limit safely. Note that the boundary condition is trivially satisfied for u_∞, since $u_k|_{\partial\Omega} = f|_{\partial\Omega}$ for each k. □

Remark 17. For the one-dimensional case, see [20, Sect. 5] and [18, Sect. 5].

More on the Set S

In this section we discuss further the results of [28]. As mentioned above we are interested in conditions that guarantee us that $S \neq \emptyset$. Recall that so far we have the following sufficient conditions for non-emptiness and emptiness of S, respectively.

- If $Lip(f, \partial\Omega) \leq 1$, then $S \neq \emptyset$.
- If $\partial D \cap \partial\Omega = \emptyset$, then $S \neq \emptyset$.
- If $\partial D \cap \partial\Omega \neq \emptyset$ and $Lip(f, \partial D \cap \partial\Omega) > 1$, then $S = \emptyset$,

while the condition

- $\partial D \cap \partial\Omega \neq \emptyset,\ Lip(f, \partial D \cap \partial\Omega) \leq 1$ *and* $Lip(f, \partial\Omega) > 1$,

is necessary but not sufficient for the emptiness of S. In this direction, the authors in [28] tried to identify conditions for the emptiness and non-emptiness of S. They also gave several concrete examples that are quite illustrative (see [28, Sect. 5]). To avoid technicalities we are just going through the main results. To simplify the discussion we give the following notation (see [28]). Define the sets,

$$\Omega_* = \Omega \setminus \overline{D}, \quad Q = \partial D \cap \partial\Omega, \quad Q_* = Q \cap \partial\Omega_*.$$

Also define the following subset of S.

$$S_L = \{f : \overline{\Omega} \to \mathbb{R} :\ f \text{ isLipschitz}\}.$$

- **Non-emptiness of** S ([28, Sect. 3]).

  The problem arises when we are close to the points of Q_*. The following results hold.

  1. If $\partial\Omega$, ∂D are C^1-smooth and there is an open neighborhood W of Q such that

  $$Lip(f, W \cap \partial\Omega) \leq 1,$$

  then $S_L \neq \emptyset$. Thus S is non-empty ([28, Theorem 3.2 (i)]).

2. If we drop the C^1-smoothness condition for D and Ω, we have to put some additional hypotheses to control the distance between the points of Q_* (roughly speaking), for concluding that $S_L \neq \emptyset$ ([28, Theorem 3.1]).
3. Now consider that $\partial\Omega$ and ∂D are C^1-smooth but we cannot find a neighborhood of Q where the Lipschitz constant of f is less than or equal to 1. Assume further that f can be written as a sum of two components with the one satisfying condition 1. Then if we put some extra hypotheses to handle things near Q_*, we have again that $S_L \neq \emptyset$ ([28, Theorem 3.2 (ii)]).
4. Assume again that $\partial\Omega$ and ∂D are C^1-smooth and f does not satisfy conditions 1 and 3. In this case we have to construct an element of S step by step, since we might have that $S_L = \emptyset$. The key step here is to construct a function f' such that for each $z \in Q_*$ and Z_z a sufficiently small neighborhood of z we have that $f' \in W^{1,p(x)}(Z_z \cap \Omega_*)$. (Note that if $f \in S_L$, then immediately $f \in W^{1,(p(x))}(\Omega \setminus \overline{D})$.) For this purpose we have to put some extra conditions concerning the variable exponent $p(\cdot)$, the dimension N and the set Q_* to gain the non-emptiness of S ([28, Theorem 3.2 (iii)]).

Remark 18. Recall (see the discussion about S in section "Main Results") that finding a necessary and sufficient condition for the non-emptiness of S is an open problem.

- **Emptiness of** S ([28, Sect. 4]).
  Here, we briefly mention some sufficient conditions for the emptiness of S.

1. If there are points in some neighborhood of Q that keep the Lipschitz constant of f away from 1 (that is roughly speaking), then $S_L = \emptyset$ ([28, Theorem 4.1 (i)]).
2. If there are some points $z \in Q_*$ where the variable exponent $p(\cdot)$ and the dimension N satisfy certain inequalities, then $S = \emptyset$. The reason is that if we assume that there exists an element $f \in S$, then at such a point z one can show that $f \notin W^{1,p(x)}(Z_z \cap \Omega_*)$, where Z_z is again a neighborhood of z ([28, Theorem 4.1 (ii), (iii)]).

The Neumann Case

Now we are turning our attention to the Neumann case studied in [17]. Consider the following problem

$$\begin{cases} -\Delta_{p(x)}u(x) = 0, & x \in \Omega \\ |\nabla u(x)|^{p(x)-2}\frac{\partial u}{\partial \nu}(x) = g(x), & x \in \partial\Omega \end{cases} \tag{P$_2$}$$

where $\Omega \subset \mathbb{R}^N$ is a bounded smooth domain, with $N \geq 2$, and $\frac{\partial}{\partial \nu}$ is the outer normal derivative. Also, $g \in C(\partial\Omega)$ and satisfies

$$\int_{\partial\Omega} g = 0.$$

Note that this latter condition is necessary, since otherwise problem (P_2) has no solution. The setting is similar to the one of problem (P_1). In particular the variable exponent p satisfies the following hypothesis

$$p|_D = \infty$$

where D is a compactly supported subdomain of Ω, with Lipschitz boundary.

Moreover, $p \in C^1(\overline{\Omega} \setminus \overline{D})$ with

$$p^+ := \sup_{x\in\overline{\Omega}\setminus\overline{D}} p(x) < \infty \tag{26}$$

and

$$p_- := \inf_{x\in\overline{\Omega}} p(x) > N \tag{27}$$

Remark 19. Note that p is defined also in $\partial\Omega$ due to the boundary condition, while here we demand D to be compactly supported to Ω. The reason for this is to have a Lemma similar to Lemma 2.

Remark 20. In [20] the set D is assumed to be convex with smooth boundary. The main reason for this is that the set of Lipcshitz function on D and $W^{1,\infty}(D)$ coincide. In this case we only assume that D has Lipschitz boundary which we need to have the density of smooth functions in $W^{1,p_k(\cdot)}(\Omega)$ by Proposition 4.

We consider again the sequence of variable exponents (p_k) introduced in section "The Dirichlet Case" (see (15)). If we replace p with p_k, we have the intermediate boundary value problems corresponding to problem (P_2).

$$\begin{cases} -\Delta_{p_k(x)}u(x) = 0, & x \in \Omega \\ |\nabla u(x)|^{p_k(x)-2}\frac{\partial u}{\partial \nu}(x) = g(x), & x \in \partial\Omega \end{cases} \tag{P'_k}$$

We follow the same strategy that is used in [20]. Using standard methods we prove the existence of a unique weak solution u_k, for problem (P'_k), that is also a viscosity solution. From the Arzellà–Ascoli theorem, we then show that the uniform limit of (u_k) exists. We call this uniform limit u_∞ and show that it satisfies a variational characterization in the set

$$S' = \left\{u \in W^{1,p_-}(\Omega) : u|_{\Omega\setminus\overline{D}} \in W^{1,p(x)}(\Omega \setminus \overline{D}), \|\nabla u\|_{L^\infty(D)} \leq 1 \text{ and } \int_\Omega u = 0\right\} \tag{28}$$

and that it is infinity harmonic in D.

Remark 21. The condition $\int_\Omega u = 0$ in the definition of S' plays a crucial role in the proof of the existence and uniqueness of the solutions u_k (since it provides us with Proposition 6) and also in their uniform boundedness.

Remark 22. Note that in the Neumann case we do not have to worry about the emptiness of S', since the zero function is always an element of S'.

We state now the main results of [17]. As we did in the Dirichlet case, we give first the results corresponding to the intermediate problem (P'_k).

Definition 4. Let $u \in W^{1,p_k(\cdot)}(\Omega)$. We say that u is a weak solution of problem (P'_k) if

$$\int_\Omega |\nabla u|^{p_k(x)-2}\nabla u \cdot \nabla v dx = \int_{\partial\Omega} gv dS, \quad \text{forall } v \in W^{1,p_k(\cdot)}(\Omega)\,. \tag{29}$$

Lemma 3. *There exists a unique weak solution u_k to problem (P'_k), which is the unique minimizer of the functional*

$$I'_k(u) = \int_\Omega \frac{|\nabla u|^{p_k(x)}}{p_k(x)}\,dx - \int_{\partial\Omega} gu\,dS$$

in the set

$$S'_k = \left\{u \in W^{1,p_k(\cdot)}(\Omega) : \int_\Omega u = 0\right\}.$$

The next Lemma is similar to Lemma 2.

Lemma 4. *The following problem*

$$\begin{cases} -\Delta_{p_k(x)}u(x) = 0, & x \in \Omega\backslash\overline{D} \\ -\Delta_k u(x) = 0, & x \in D \\ |\nabla u(x)|^{k-2}\frac{\partial u}{\partial\nu}(x) = |\nabla u(x)|^{p(x)-2}\frac{\partial u}{\partial\nu}(x), & x \in \partial D \\ |\nabla u(x)|^{p(x)-2}\frac{\partial u}{\partial\nu}(x) = g(x), & x \in \partial\Omega\,. \end{cases} \tag{30}$$

has the same weak solutions as problem (P'_k).

Remark 23. Unlike the Dirichlet case, for the validity of the above Lemma we want D to be compactly supported in Ω.

Next we follow the ideas of [4] to give the definition of a viscosity solution for a problem like (30). A similar definition corresponds to problem (31), which follows later.

Recall that D has only Lipschitz boundary. Thus, the normal vector field $\nu(\cdot)$ is not uniquely defined. For this reason to each point $x \in \partial D$ we define the set of outward unit normals $N(x)$, as the collection of all vectors ν for which we can find a sequence (x_k) in ∂D, such that $x_k \to x$ and for each k there exists a unique outward unit normal vector ν_k on ∂D at x_k, such that $\nu_k \to \nu$. Note that since D has Lipschitz boundary, $N(x)$ is nonempty.

Definition 5. (i) Let u be a lower semicontinuous function in $\overline{\Omega}$. We say that u is a viscosity supersolution of problem (30), if for every $\varphi \in C^2(\overline{\Omega})$ such that $u - \varphi$ attains its strict minimum at $x_0 \in \overline{\Omega}$ with $u(x_0) = \varphi(x_0)$, we have

- if $x_0 \in \Omega \setminus \overline{D}$, then $-\Delta_{p(x_0)}\varphi(x_0) \geq 0$.
- If $x_0 \in D$, then $-\Delta_k \varphi(x_0) \geq 0$.
- If $x_0 \in \partial D$, then

$$\max\{-\Delta_{p(x_0)}\varphi(x_0), -\Delta_k \varphi(x_0), \sup_{\nu \in N(x_0)} \{(|\nabla\varphi(x_0)|^{k-2} - |\nabla\varphi(x_0)|^{p(x_0)-2})\nabla\varphi(x_0) \cdot \nu\}\} \geq 0\,.$$

- If $x_0 \in \partial\Omega$, then

$$\max\{|\nabla\varphi(x_0)|^{p(x_0)-2}\frac{\partial\varphi}{\partial\nu}(x_0) - g(x_0), -\Delta_{p(x_0)}\varphi(x_0)\} \geq 0.$$

(ii) Let u be an upper semicontinuous function in $\overline{\Omega}$. We say that u is a viscosity subsolution of problem (30), if for every $\varphi \in C^2(\overline{\Omega})$ such that $u - \varphi$ attains its strict maximum at $x_0 \in \overline{\Omega}$ with $u(x_0) = \varphi(x_0)$, we have

- if $x_0 \in \Omega \setminus \overline{D}$, then $-\Delta_{p(x_0)}\varphi(x_0) \leq 0$.
- If $x_0 \in D$, then $-\Delta_k \varphi(x_0) \leq 0$.
- If $x_0 \in \partial D$, then

$$\min\{-\Delta_{p(x_0)}\varphi(x_0), -\Delta_k \varphi(x_0), \inf_{\nu \in N(x_0)} \{(|\nabla\varphi(x_0)|^{k-2} - |\nabla\varphi(x_0)|^{p(x_0)-2})\nabla\varphi(x_0) \cdot \nu\}\} \leq 0.$$

- If $x_0 \in \partial\Omega$, then

$$\min\{|\nabla\varphi(x_0)|^{p(x_0)-2}\frac{\partial\varphi}{\partial\nu}(x_0) - g(x_0), -\Delta_{p(x_0)}\varphi(x_0)\} \leq 0.$$

(iii) Finally, u is called a viscosity solution of problem (30), if it is both a viscosity subsolution and a viscosity supersolution.

Proposition 9. *Let u be a continuous weak solution of (P_k'). Then u is a solution of (30) in the viscosity sense.*

Proof. [Sketch] Depending on the location of x_0, one has to verify Definition 5. When $x_0 \in \Omega \setminus \overline{D}$ or $x_0 \in D$ the proof is similar to the one that corresponds to the Dirichlet case. When $x_0 \in \partial D$ we follow the ideas of [4]. In [20], the case of $x_0 \in \partial\Omega$ is trivially verified because of the Dirichlet boundary condition. Here, the boundary condition in the viscosity sense is not immediately satisfied and one has to use the continuity of the boundary data g. □

Next we state a Proposition similar to Proposition 8 that allows us to characterize the candidate solution of (P_2) as a uniform limit.

Proposition 10. *Let u_k be a weak solution of problem (P_2). Then the sequence (u_k) is equicontinuous and uniformly bounded.*

Remark 24. Note that in the above Proposition, in contrast to Proposition 8 there can be no comparison of the p_- norm of $|\nabla u_k|$ with $I_k(u_k)$, due to existence of the term $\int_{\partial\Omega} g u_k dS$, in the definition of I_k. To overcome this difficulty we have to use the estimates given in Proposition 3. Also the hypothesis $\int_\Omega u_k = 0$ allows us to use the Poincaré–Wirtinger inequality to conclude the proof.

Remember that (see 28)

$$S' = \left\{ u \in W^{1,p_-}(\Omega) : u|_{\Omega\setminus\overline{D}} \in W^{1,p(\cdot)}(\Omega\setminus\overline{D}),\ \|\nabla u\|_{L^\infty(D)} \le 1 \text{ and } \int_\Omega u = 0 \right\}.$$

Let $I' : S' \to \mathbb{R}$ be such that, if $v \in S \subset S_k$, we have

$$I'(v) := \int\limits_{\Omega\setminus\overline{D}} \frac{|\nabla v|^{p(x)}}{p(x)}\, dx - \int\limits_{\partial\Omega} g v\, dS$$

The next theorem is the analogue of Theorem 2 for the Neumann case. We give a variational characterization of the limit function u_∞ in $\Omega \setminus \overline{D}$, where

$$p^+ := \sup_{\overline{\Omega}\setminus\overline{D}} p(x) < \infty$$

and next we prove that u_∞ is infinity harmonic in D in the viscosity sense.

Theorem 4. *Let u_k be the unique minimizer of I_k in S_k. Then there exists a function $u_\infty \in S$, such that u_∞ minimizes I' in S' and is also infinity harmonic in D.*

Remark 25. The proof is similar to the one of Theorem 2. Also, the minimizer of I' in S' is unique. Indeed let u_1, u_2 be minimizers of I' in S' that are also infinity harmonic in D. Then, there exists some $C \in \mathbb{R}$ such that

$$u_2 = u_1 + C \quad \text{in} \quad \overline{\Omega}\setminus D,$$

which by the uniqueness of the Dirichlet problem for the infinity Laplacian gives us that $u_2 = u_1 + C$ also in D. Hence $u_2 = u_1 + C$ in Ω which implies uniqueness by the fact that the mean value of both u_1, u_2 is zero.

Next we state the problem of which u_∞ is a viscosity solution. This arises naturally from Lemma 4 and Proposition 9 as a limit case.

Theorem 5. *Let (u_k) be the sequence of solutions of problems (P'_k) and u_∞ the uniform limit of a subsequence of (u_k). Then u_∞ is a viscosity solution of the following problem*

$$\begin{cases} -\Delta_{p(x)}u(x) = 0, & x \in \Omega\backslash\overline{D} \\ -\Delta_\infty u(x) = 0, & x \in D \\ sgn(|\nabla u(x)| - 1)sgn(\frac{\partial u}{\partial \nu}(x)) = 0, & x \in \partial D \\ |\nabla u(x)|^{p(x)-2}\frac{\partial u}{\partial \nu}(x) = g(x), & x \in \partial\Omega \end{cases} \tag{31}$$

Proof. [Sketch] To show that u_∞ is a viscosity solution of (31), we have to verify a definition similar to Definition 5. When $x_0 \in \Omega \setminus \overline{D}$ or $x_0 \in D$ the proof is similar to the one of Theorem 3. When $x_0 \in \partial D$ we prove that

$$\max\{-\Delta_{p(x_0)}\varphi(x_0), -\Delta_\infty\varphi(x_0), \sup_{\nu\in N(x_0)}\{\text{sgn}(|\nabla\varphi(x_0)| - 1)\text{sgn}(\nabla\varphi(x_0)\cdot\nu)\}\} \geq 0,$$

to show that u_∞ is a viscosity supersolution. To verify the boundary condition in the viscosity sense we use again the continuity of the boundary data g. □

Conclusions

As mentioned at the beginning of this paper, the problems where the variable exponent assumed to be infinite are rare in the literature. Nevertheless, these problems are quite interesting, since they combine the concepts of variable exponent theory and the infinity Laplacian. The latter means that one should try different approaches than the usual variational one, such as viscosity theory, absolutely minimal Lipschitz extensions, etc.

This paper is an attempt to briefly present the work has been done so far, and hopefully give a motivation for further research.

Acknowledgement The authors would like to thank the anonymous referee for his useful remarks and suggestions.

References

1. Aronsson, G.: Minimization problems for the functional $\sup_x F(x,f(x),f'(x))$. Ark. Mat. **6**, 33–53 (1965)
2. Aronsson, G.: Extensions of functions satisfying Lipschitz conditions. Ark. Mat. **6**, 551–561 (1967)
3. Aronsson, G., Crandall, M.G., Juutinen, P.: A tour of the theory of absolutely minimizing functions. Bull. Am. Math. Soc. **41**, 439–505 (2004)
4. Barles, G., Georgelin, C.A., Jakobsen, E.R.: On Neumann and oblique derivatives boundary conditions for nonlocal elliptic equations. J. Differ. Equ. **256**(4), 1368–1394 (2014)
5. Bhattacharya, T., DiBenedetto, E., Maneuverer, J.: Limits as $p \to \infty$ of $\Delta_p u_p = f$ and related extremal problems. Rend. Sem. Mat. Univ. Politec. Torino **1989**, 15–68 (1991)
6. Caselles, V., Morel, J.M., Sbert, C.: An axiomatic approach to image interpolation. IEEE Trans. Image Process. **7**, 376–386 (1998)
7. Chen, Y.M., Levine, S., Rao, M.: Variable exponent, linear growth functionals in image restoration. SIAM J. Appl. Math. **66**, 1383–1406 (2006)
8. Crandall, M.G., Ishii, H., Lions, P.L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. **27**, 1–67 (1992)
9. Crandall, M.G., Evans, L.C., Gariepy, R.: Optimal Lipschitz extensions and the infinity Laplacian. Calc. Var. Partial Differ. Equ. **13**(2), 123–139 (2001)
10. Diening, L., Harjuleto, P., Hästö, P., Ruzisca, M.: Lebesgue and Sobolev Spaces with Variable Exponents. Lecture notes in mathematics, vol. 2017, Springer, Heidelberg, (2011). Available at http://www.helsinki.fi/~pharjule/varsob/publications.shtml
11. Evans, L.C., Gangbo, W.: Differential equations methods for the Monge–Kantorovich mass transfer problem. Mem. Am. Math. Soc. **137**(653), viii+66 (1999)
12. Edmunds, D.E., Rakosník, J.: Density of smooth functions in $W^{k,p(x)}(\Omega)$. Proc. R. Soc. Lond. Ser. A **437**, 229–236 (1992)
13. Fan, X., Zhao, D.: On the spaces $L^{p(x)}(\Omega)$ and $W^{m,p(x)}(\Omega)$. J. Math. Anal. Appl. **263**, 424–446 (2001)
14. García-Azorero, J., Manfredi, J.J., Peral, I., Rossi, J.D.: The Neumann problem for the ∞-Laplacian and the Monge–Kantorovich mass transfer problem. Nonlinear Anal. **66**, 349–366 (2007)
15. Harjuleto, P., H*ästö*, P., Koskenoja, M., Varonen, S.: The Dirichlet energy integral and variable exponent Sobolev spaces with zero boundary values. Potential Anal. **25**(3), 205–222 (2006)
16. Jensen, R.: Uniqueness of Lipschitz extensions: minimizing the sup norm of the gradient. Ark. Ration. Mech. Anal. **123**, 51–74 (1993)
17. Karagiorgos, Y., Yannakakis, N.: A Neumann problem involving the p(x)-Laplacian with $p = \infty$ in a subdomain. In: Advances in Calculus of Variations, pp. 1864–8266. De Gruyter, Berlin (Online) (2014). doi:10.1515/acv-2014-0003
18. Keisala, J.: Existence and uniqueness of $p(x)$-harmonic functions for bounded and unbounded $p(x)$. Licentiate thesis, Rep. Univ. Jyv*ä*skyl*ä* Dept. Math. Stat. 130 (2011). Available at http://www.helsinki.fi/~pharjule/varsob/publications.shtml
19. Kováčik, O., Rakosník, J.: On spaces $L^{p(x)}$ and $W^{1,p(x)}$. Czech. Math. J. **41**(116), 592–618 (1991)
20. Manfredi, J.J., Rossi, J.D., Urbano, J.M.: $p(x)$-harmonic functions with unbounded exponent in a subdomain. Ann. l'Inst. Henri Poincarè C. Anal. Non Linéaire **26** (6), 2581–2595 (2009)
21. Manfredi, J.J., Rossi, J.D., Urbano, J.M.: Limits as $p(x) \to \infty$ of $p(x)$-harmonic functions. Nonlinear Anal. **72**, 309–315 (2010)
22. McShane, E.J.: Extension of range of functions. Bull. Am. Math. Soc. **40**(12), 837–842 (1934)
23. Peres, Y., Schramm, O., Sheffield, S., Wilson, D.B.: Tug-of-war and the infinity Laplacian. J. Am. Math. Soc. **22**, 167–210 (2009)
24. Pérez-Llanos, M., Rossi, J.D.: The behaviour of the $p(x)$-Laplacian eigenvalue problems $p(x) \to \infty$. J. Math. Anal. Appl. **363**, 502–511 (2010)

25. Pérez-Llanos, M., Rossi, J.D.: Limits as $p(x) \to \infty$ of $p(x)$-harmonic functions with non-homogeneous Neumann boundary conditions. Contem. Math. AMS **540**, 187–202 (2011)
26. Ruzicka, M. : Electrorheological Fluids: Modelling and Mathematical Theory. Springer, Berlin (2000)
27. Samko, S.: Denseness of $C_0^\infty(\mathbb{R}^n)$ in the generalized Sobolev spaces $W^{m,p(x)}(\mathbb{R}^n)$. In: Direct and Inverse Problems of Mathematical Physics. International Society for Analysis, Applications and Computation, Newark, DE, 1997, vol. 5, pp. 333–342. Kluwer Academic, Dordrecht (2000)
28. Urbano, J.M., Vorotnikov, D.: On the well-posedness of a two-phase minimization problem. J. Math. Anal. Appl. **378**, 159–168 (2011)
29. Whitney, H.: Analytic extensions of differentiable functions defined in closed sets. Trans. Am. Math. Soc. **36**(1), 63–89 (1934)
30. Yao, J.: Solutions for Neumann boundary value problems involving $p(x)$-Laplace operators. Nonlinear Anal. **68**, 1271–1283 (2008)

An Umbral Calculus Approach to Bernoulli–Padé Polynomials

Dae San Kim and Taekyun Kim

Abstract In this paper, we consider Bernoulli–Padé polynomials of fixed order whose generating function is based on the Padé approximant of the exponential function. We derive, by using umbral calculus techniques, several recurrence relations for these polynomials and investigate connections between our polynomials and several known families of polynomials.

Keywords Bernoulli–Padé polynomials • Umbral calculus

Introduction

Let $\mathbb{C}$ be the complex field and let $\mathscr{F}$ be the set of all formal power series in the variable t:

$$\mathscr{F} = \left\{ f(t) = \sum_{k=0}^{\infty} a_k \frac{t^k}{k!} \,\middle|\, a_k \in \mathbb{C} \right\}. \tag{1}$$

Let $\mathbb{P} = \mathbb{C}[x]$ and let $\mathbb{P}^*$ be the vector space of all linear functionals on $\mathbb{P}$. We denote by $\langle L \mid p(x) \rangle$ the action of the linear functional L on the polynomial $p(x)$, and recall that the vector space operations on $\mathbb{P}^*$ are defined by $\langle L + M \mid p(x) \rangle = \langle L \mid p(x) \rangle + \langle M \mid p(x) \rangle$, and $\langle cL \mid p(x) \rangle = c \langle L \mid p(x) \rangle$, where c is a complex constant in $\mathbb{C}$. Let $f(t) \in \mathscr{F}$. Then we define the linear functional on $\mathbb{P}$ by setting

$$\langle f(t) | x^n \rangle = a_n, \quad (n \geq 0), \quad (\text{see [11–13]}). \tag{2}$$

D.S. Kim
Department of Mathematics, Sogang University, Seoul 121-742, Republic of Korea
e-mail: dskim@sogang.ac.kr

T. Kim (✉)
Department of Mathematics, Kwangwoon University, Seoul 139-701, Republic of Korea
e-mail: tkkim@kw.ac.kr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_16

By (1) and (2), we get

$$\left\langle t^{k}\middle| x^{n}\right\rangle = n!\delta_{n,k}, \quad (n,k \geq 0), \quad (\text{see [10, 13]}), \tag{3}$$

where $\delta_{n,k}$ is the Kronecker symbol.

For $f_L(t) = \sum_{k=0}^{\infty} \frac{\langle L|x^k\rangle}{k!} t^k$, we have $\langle f_L(t)| x^n\rangle = \langle L \mid x^n\rangle$. That is, $f_L(t) = L$. So, the map $L \mapsto f_L(t)$ is vector space isomorphism from $\mathbb{P}^*$ onto $\mathscr{F}$.

Henceforth, $\mathscr{F}$ denotes both the algebra of formal power series in t and the vector space of all linear functionals on $\mathbb{P}$, and so an element $f(t)$ of $\mathscr{F}$ will be thought of as both a formal power series and a linear functional. $\mathscr{F}$ is called the umbral algebra, the study of which umbral calculus is. The order of $f(t)$ $(\in \mathscr{F})$ is the smallest integer k such that the coefficient of t^k does not vanish and is denoted by $o(f(t))$. If $o(f(t)) = 1$, then $f(t)$ is called a delta series; if $o(g(t)) = 0$, then $g(t)$ is called an invertible series. For $f(t), g(t) \in \mathscr{F}$ with $o(f(t)) = 1$ and $o(g(t)) = 0$, there exists a unique sequence $s_n(x)$ $(\deg s_n(x) = n)$ such that $\left\langle g(t) f(t)^k \middle| s_n(x)\right\rangle = n!\delta_{n,k}$, for $n, k \geq 0$. Such a sequence $s_n(x)$ is called the Sheffer sequence for $(g(t), f(t))$, which is denoted by $s_n(x) \sim (g(t), f(t))$.

The sequence $s_n(x)$ is Sheffer for $(g(t), f(t))$ if and only if

$$\frac{1}{g\left(\bar{f}(t)\right)} e^{y\bar{f}(t)} = \sum_{k=0}^{\infty} \frac{s_k(y)}{k!} t^k, \quad (y \in \mathbb{C}), \quad (\text{see [3, 9, 13]}), \tag{4}$$

where $\bar{f}(t)$ is the compositional inverse of $f(t)$, with $\bar{f}(f(t)) = f\left(\bar{f}(t)\right) = t$.

Let $f(t), g(t) \in \mathscr{F}$, and $p(x) \in \mathbb{C}[x]$. Then we have

$$\langle f(t) g(t)| p(x)\rangle = \langle f(t)| g(t) p(x)\rangle = \langle g(t)| f(t) p(x)\rangle, \tag{5}$$

and

$$f(t) = \sum_{k=0}^{\infty} \left\langle f(t)| x^k\right\rangle \frac{t^k}{k!}, \quad p(x) = \sum_{k=0}^{\infty} \left\langle t^k \mid p(x)\right\rangle \frac{x^k}{k!}, \quad (\text{see [13]}). \tag{6}$$

By (6), we easily get

$$t^k p(x) = p^{(k)}(x) = \frac{d^k p(x)}{dx^k}, \quad e^{yt} p(x) = p(x+y), \tag{7}$$

and

$$\left\langle e^{yt}\middle| p(x)\right\rangle = p(y).$$

The following Eqs. (8)–(10) are equivalent to the fact that $s_n(x)$ is Sheffer for $(g(t), f(t))$, for some invertible $g(t)$ (see [13]):

$$f(t)\, s_n(x) = n s_{n-1}(x), \quad (n \geq 0), \tag{8}$$

$$s_n(x+y) = \sum_{j=0}^{n} \binom{n}{j} s_j(x)\, p_{n-j}(y), \tag{9}$$

with $p_n(x) = g(t)\, s_n(x)$,

$$s_n(x) = \sum_{j=0}^{n} \frac{1}{j!} \left\langle g\left(\bar{f}(t)\right)^{-1} \bar{f}(t)^j \middle| x^n \right\rangle x^j. \tag{10}$$

Note that

$$s_{n+1}(x) = \left(x - \frac{g'(t)}{g(t)}\right) \frac{1}{f'(t)} s_n(x), \tag{11}$$

$$\langle f(t) | xp(x) \rangle = \langle \partial_t f(t) | p(x) \rangle, \tag{12}$$

$$\frac{d}{dx} s_n(x) = \sum_{l=0}^{n-1} \binom{n}{l} \left\langle \bar{f}(t) \middle| x^{n-l} \right\rangle s_l(x), \quad (n \geq 1), \tag{13}$$

where $s_n(x) \sim (g(t), f(t))$.

For $s_n(x) \sim (g(t), f(t))$, and $r_n(x) \sim (h(t), l(t))$, we have

$$s_n(x) = \sum_{m=0}^{n} C_{n,m} r_m(x), \quad (n \geq 0), \tag{14}$$

where

$$C_{n,m} = \frac{1}{m!} \left\langle \frac{h\left(\bar{f}(t)\right)}{g\left(\bar{f}(t)\right)} l\left(\bar{f}(t)\right)^m \middle| x^n \right\rangle, \quad \text{(see [9, 13])}. \tag{15}$$

The Padé approximant, introduced by Georg Frobenius and developed by Henri Padé, is a rational function of given order which best approximates a given function (c.f. [1, 2, 4, 7]). It gives better approximation of the function than just truncating its Taylor series. For this reason, it is widely used in electronic computer computations, has connection with certain type of continued fractions, and is used in Diophantine approximation and transcendental number theory. In this paper, we will consider Bernoulli–Padé polynomials of fixed order whose generating function is based on the Padé approximant of the exponential function. These Bernoulli–Padé polynomials include the ordinary Bernoulli polynomials and several other types of polynomials such as van der Pol polynomials as special cases.

An Umbral Calculus Approach to Bernoulli–Padé Polynomials

Let r, s be given nonnegative integers. Then the n-th Bernoulli–Padé polynomials $B_n^{(r,s)}(x)$ of order (r, s) are defined by

$$\frac{(-1)^s \, r!s!t^{r+s+1}e^{xt}}{(r+s)!\,(r+s+1)!\left[Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right]} = \sum_{n=0}^{\infty} B_n^{(r,s)}(x)\,\frac{t^n}{n!}, \tag{16}$$

where

$$P^{(r,s)}(t) = \sum_{j=0}^{r} \binom{r}{j} \frac{(r+s-j)!}{(r+s)!} t^j, \tag{17}$$

$$Q^{(r,s)}(t) = \sum_{j=0}^{s} \binom{s}{j} \frac{(r+s-j)!}{(r+s)!} (-t)^j .$$

When $x = 0$, $B_n^{(r,s)} = B_n^{(r,s)}(0)$ are called the n-th Bernoulli–Padé numbers of order (r, s).

Here, we note that $B_n^{(0,0)}(x) = B_n(x)\,, B_n^{(0,0)} = B_n(0) = B_n$, where $B_n(x)$ are Bernoulli polynomials given by the generating function

$$\frac{t}{e^t - 1}e^{xt} = \sum_{n=0}^{\infty} B_n(x)\,\frac{t^n}{n!}, \quad (\text{see } [1-9, 11-13])\,.$$

Actually,

$$[r/s] = \frac{P^{(r,s)}(t)}{Q^{(r,s)}(t)}, \quad (r, s = 0, 1, 2, \ldots) \tag{18}$$

are the Padé approximants to e^t, so that

$$e^t - \frac{P^{(r,s)}(t)}{Q^{(r,s)}(t)} \equiv 0 \pmod{t^{r+s+1}}. \tag{19}$$

More precisely, one can show that

$$e^t - \frac{P^{(r,s)}(t)}{Q^{(r,s)}(t)} = \frac{(-1)^s \, r!s!}{(r+s)!\,(r+s+1)!} t^{r+s+1} + \text{higher order terms}, \tag{20}$$

and hence $B_0^{(r,s)} = 1$.

The Bernoulli–Padé polynomials and numbers were introduced in [2]. These polynomials and numbers were further generalized, with certain exceptions, to the case of arbitrary complex parameters in [1]. Special cases of Bernoulli–Padé numbers $A_{k,n}$ and $V_{k,n}$ were studied earlier by Howard in [4, 7]. They are given by

$$\frac{\frac{1}{k!}t^k}{e^t - \sum_{j=0}^{k-1} \frac{t^j}{j!}} = \sum_{n=0}^{\infty} A_{k,n} \frac{t^n}{n!} = \sum_{n=0}^{\infty} B_n^{(k-1,0)} \frac{t^n}{n!}, \quad (k \geq 1), \tag{21}$$

and

$$\frac{\frac{k!}{(2k+1)!}t^{2k+1}}{t^k y_k\left(-\frac{2}{t}\right) e^t - (-t)^k y_k\left(\frac{2}{t}\right)} = \sum_{n=0}^{\infty} V_{k,n} \frac{t^n}{n!} = \sum_{n=0}^{\infty} B_n^{(k,k)} \frac{t^n}{n!}, \tag{22}$$

where $y_k(x)$ is the Bessel polynomial of degree k defined by

$$y_k(x) = \sum_{j=0}^{k} \frac{(k+j)!}{j!\,(k-j)!} \left(\frac{x}{2}\right)^j. \tag{23}$$

For $k = 1$ in (22), we have

$$\frac{\frac{1}{6}t^3}{(t-2)\,e^t + (t+2)} = \sum_{n=0}^{\infty} V_{1,n} \frac{t^n}{n!} = \sum_{n=0}^{\infty} V_n \frac{t^n}{n!}, \tag{24}$$

and V_n are called van der Pol numbers, which were investigated in [5, 6, 8]. From (16), we note that $B_n^{(r,s)}(x)$ is the Appell sequence for

$$\frac{(r+s)!\,(r+s+1)!\left[Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right]}{(-1)^s\, r!s!t^{r+s+1}},$$

so that

$$B_n^{(r,s)}(x) \sim \left(\frac{(r+s)!\,(r+s+1)!\left[Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right]}{(-1)^s\, r!s!t^{r+s+1}}, t\right). \tag{25}$$

Sometimes, for simplicity, we will put $d_{r,s} = \frac{(-1)^s r!s!}{(r+s)!(r+s+1)!}$.

As $B_n^{(r,s)}(x)$ is an Appell sequence, we have

$$tB_n^{(r,s)}(x) = \frac{d}{dx}B_n^{(r,s)}(x) = nB_{n-1}^{(r,s)}(x), \tag{26}$$

$$B_n^{(r,s)}(x+y) = \sum_{j=0}^{n} \binom{n}{j} B_j^{(r,s)}(x)\, y^{n-j}, \tag{27}$$

and

$$B_n^{(r,s)}(x) = \sum_{j=0}^{n} \binom{n}{j} B_j^{(r,s)} x^{n-j}. \tag{28}$$

In addition, we get

$$B_n^{(r,s)}(1-x) = (-1)^n B_n^{(s,r)}(x). \tag{29}$$

Here, we note that (29) follows from (16) by using the fact $Q^{(s,r)}(-t) = p^{(r,s)}(t)$. The above (26)–(29) were stated as Proposition 2.2 in [2].

For nonnegative integer r, s with $r + s + 1 \leq n$, we have

$$Q^{(r,s)}(D) B_n^{(r,s)}(x+1) - P^{(r,s)}(D) B_n^{(r,s)}(x) = (-1)^s \frac{\binom{n}{r+s+1}}{\binom{r+s}{r}} x^{n-(r+s+1)}, \tag{30}$$

where $D = \frac{d}{dx}$ (see [2, Proposition 2.3]).

Proof (Proof of (30)).

Taking the advantage of umbral calculus, we can give very short proof for this. Since $B_n^{(r,s)}(x)$ is an Appell sequence from (16), we have

$$B_n^{(r,s)}(x) = \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} x^n. \tag{31}$$

Multiplying $Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)$ on both sides gives

$$\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right) B_n^{(r,s)}(x) = d_{r,s} t^{r+s+1} x^n. \tag{32}$$

Thus, by (32), we get

$$\begin{aligned} & Q^{(r,s)}(D) B_n^{(r,s)}(x+1) - P^{(r,s)}(D) B_n^{(r,s)}(x) \\ =& d_{r,s} (n)_{r+s+1} x^{n-(r+s+1)} \\ =& \frac{(-1)^s \binom{n}{r+s+1}}{\binom{r+s}{r}} x^{n-(r+s+1)}. \end{aligned} \tag{33}$$

For $r + s + 1 \leq n$, we have

$$\begin{aligned} & \sum_{j=0}^{s} (-1)^j \frac{\binom{s}{j}\binom{n}{j}}{\binom{r+s}{j}} B_{n-j}^{(r,s)}(x+1) - \sum_{j=0}^{r} \frac{\binom{r}{j}\binom{n}{j}}{\binom{r+s}{j}} B_{n-j}^{(r,s)}(x) \\ &= (-1)^s \frac{\binom{n}{r+s+1}}{\binom{r+s}{r}} x^{n-(r+s+1)}, \quad \text{(see [2, Corollary 2.1])}. \end{aligned} \tag{34}$$

Proof (Proof of (34)*).* This is immediate from (30) by using the definition of $Q^{(r,s)}(t)$ and $P^{(r,s)}(t)$ in (17).

From (11) and (25), we have

$$B_{n+1}^{(r,s)}(x) = xB_n^{(r,s)}(x) - \frac{g'(t)}{g(t)} B_n^{(r,s)}(x), \tag{35}$$

where

$$g(t) = \frac{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)}{d_{r,s} t^{r+s+1}}.$$

Now, we observe that

$$\begin{aligned}
\frac{g'(t)}{g(t)} &= (\log(g(t)))' \qquad (36)\\
&= \left(\log\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right) - \log d_{r,s} - (r+s+1)\log t\right)' \\
&= \frac{\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} - \frac{r+s+1}{t}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{g'(t)}{g(t)} B_n^{(r,s)}(x) \qquad (37)\\
=&\left(\frac{\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} - \frac{r+s+1}{t}\right) B_n^{(r,s)}(x) \\
=&t^{r+s+1}\left(\frac{\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} - \frac{r+s+1}{t}\right) \\
&\times \frac{B_{n+r+s+1}^{(r,s)}(x)}{(n+1)(n+2)\cdots(n+r+s+1)} \\
=&\frac{n!}{(n+r+s+1)!}\left(\frac{1}{d_{r,s}}\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right)' \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)}\right. \\
&\left. - (r+s+1)\, t^{r+s}\right) B_{n+r+s+1}^{(r,s)}(x) \\
=&\frac{n!}{(n+r+s+1)!}\frac{1}{d_{r,s}}\left(Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)\right)' \\
&\times \sum_{m=0}^{n+r+s+1}\binom{n+r+s+1}{m} B_{n+r+s+1-m}^{(r,s)} \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} x^m
\end{aligned}$$

$$
\begin{aligned}
&-\frac{r+s+1}{n+1}B_{n+1}^{(r,s)}(x) \\
&=(-1)^{s}\frac{\binom{r+s}{r}}{\binom{n+r+s+1}{n}}\left(Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)\right)' \\
&\times\sum_{m=0}^{n+r+s+1}\binom{n+r+s+1}{m}B_{n+r+s+1-m}^{(r,s)}B_{m}^{(r,s)}(x)-\frac{r+s+1}{n+1}B_{n+1}^{(r,s)}(x).
\end{aligned}
$$

Before proceeding, we note that

$$
\left(Q^{(r,s)}(t)\right)'=-\frac{s}{r+s}Q^{(r,s-1)}(t), \tag{38}
$$

and

$$
\left(P^{(r,s)}(t)\right)'=\frac{r}{r+s}P^{(r-1,s)}(t). \tag{39}
$$

Thus, by (38) and (39), we get

$$
\begin{aligned}
&\left(Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)\right)' \qquad (40)\\
&=\left(Q^{(r,s)}(t)\right)'e^{t}+Q^{(r,s)}(t)e^{t}-\left(P^{(r,s)}(t)\right)' \\
&=-\frac{s}{r+s}Q^{(r,s-1)}(t)e^{t}+Q^{(r,s)}(t)e^{t}-\frac{r}{r+s}P^{(r-1,s)}(t) \\
&=-\frac{s}{r+s}\sum_{j=0}^{s-1}\binom{s-1}{j}\frac{(r+s-1-j)!}{(r+s-1)!}(-1)^{j}e^{t}t^{j} \\
&+\sum_{j=0}^{s}\binom{s}{j}\frac{(r+s-j)!}{(r+s)!}(-1)^{j}e^{t}t^{j} \\
&-\frac{r}{r+s}\sum_{j=0}^{r-1}\binom{r-1}{j}\frac{(r+s-1-j)!}{(r+s-1)!}t^{j} \\
&=\frac{e^{t}}{s!}\binom{r+s}{r}^{-1}(-1)^{s}t^{s}+\left(\sum_{j=0}^{s-1}\frac{1}{j!}\binom{r+s-j}{r}(-1)^{j}e^{t}t^{j}\right)\binom{r+s}{r}^{-1} \\
&-\binom{r+s}{r}^{-1}\sum_{j=0}^{s-1}\frac{1}{j!}\binom{r+s-1-j}{r}(-1)^{j}e^{t}t^{j} \\
&-\binom{r+s}{r}^{-1}\sum_{j=0}^{r-1}\frac{1}{j!}\binom{r+s-1-j}{s}t^{j}
\end{aligned}
$$

$$=\frac{1}{s!}\binom{r+s}{r}^{-1}(-1)^{s}e^{t}t^{s}+\binom{r+s}{r}^{-1}\sum_{j=0}^{s-1}\frac{1}{j!}\binom{r+s-1-j}{r-1}(-1)^{j}e^{t}t^{j}$$
$$-\binom{r+s}{r}^{-1}\sum_{j=0}^{r-1}\frac{1}{j!}\binom{r+s-1-j}{s}t^{j}.$$

From (40), we have

$$(-1)^{s}\frac{\binom{r+s}{r}}{\binom{n+r+s+1}{n}}\left(Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)\right)' \tag{41}$$
$$\times\sum_{m=0}^{n+r+s+1}\binom{n+r+s+1}{m}B_{n+r+s+1-m}^{(r,s)}B_{m}^{(r,s)}(x)$$
$$=(-1)^{s}\binom{n+r+s+1}{n}^{-1}\left\{(-1)^{s}\sum_{m=s}^{n+r+s+1}\binom{n+r+s+1}{m}\binom{m}{s}\right.$$
$$\times B_{n+r+s+1-m}^{(r,s)}B_{m-s}^{(r,s)}(x+1)$$
$$+\sum_{j=0}^{s-1}\sum_{m=j}^{n+r+s+1}(-1)^{j}\binom{r+s-1-j}{r-1}\binom{n+r+s+1}{m}\binom{m}{j}$$
$$\times B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x+1)$$
$$\left.-\sum_{j=0}^{r-1}\sum_{m=j}^{n+r+s+1}\binom{r+s-1-j}{s}\binom{n+r+s+1}{m}\binom{m}{j}B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x)\right\}.$$

By (35), (37), and (41), we get

$$B_{n+1}^{(r,s)}(x) \tag{42}$$
$$=xB_{n}^{(r,s)}(x)-(-1)^{s}\binom{n+r+s+1}{n}^{-1}$$
$$\times\left\{(-1)^{s}\sum_{m=s}^{n+r+s+1}\binom{n+r+s+1}{m}\binom{m}{s}B_{n+r+s+1-m}^{(r,s)}B_{m-s}^{(r,s)}(x+1)\right.$$
$$+\sum_{j=0}^{s-1}\sum_{m=j}^{n+r+s+1}(-1)^{j}\binom{r+s-1-j}{r-1}\binom{n+r+s+1}{m}\binom{m}{j}$$
$$\times B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x+1)$$

$$-\sum_{j=0}^{r-1}\sum_{m=j}^{n+r+s+1}\binom{r+s-1-j}{s}\binom{n+r+s+1}{m}\binom{m}{j}B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x)\Bigg\}$$

$$+\frac{r+s+1}{n+1}B_{n+1}^{(r,s)}(x).$$

Thus, by (42), we get the following equation.
For $s, r \geq 1$, $n \geq 0$, we have

$$\begin{aligned}
&\frac{n-r-s}{n+1}B_{n+1}^{(r,s)}(x)-xB_{n}^{(r,s)}(x) \qquad (43)\\
=&(-1)^{s-1}\binom{n+r+s+1}{n}^{-1}\Bigg\{(-1)^{s}\sum_{m=s}^{n+r+s+1}\binom{n+r+s+1}{m}\binom{m}{s}\\
&\times B_{n+r+s+1-m}^{(r,s)}B_{m-s}^{(r,s)}(x+1)\\
&+\sum_{j=0}^{s-1}\sum_{m=j}^{n+r+s+1}(-1)^{j}\binom{r+s-1-j}{r-1}\binom{n+r+s+1}{m}\binom{m}{j}\\
&\times B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x+1)\\
&-\sum_{j=0}^{r-1}\sum_{m=j}^{n+r+s+1}\binom{r+s-1-j}{s}\binom{n+r+s+1}{m}\binom{m}{j}B_{n+r+s+1-m}^{(r,s)}B_{m-j}^{(r,s)}(x)\Bigg\}.
\end{aligned}$$

Let $n \geq 1$. Then, by (3), we get

$$\begin{aligned}
&B_{n}^{(r,s)}(y) \qquad (44)\\
=&\left\langle \sum_{i=0}^{\infty}B_{i}^{(r,s)}(y)\frac{t^{i}}{i!}\middle| x^{n}\right\rangle\\
=&\left\langle \frac{d_{r,s}t^{r+s+1}e^{yt}}{Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)}\middle| x^{n}\right\rangle\\
=&\left\langle \partial_t\left(\frac{d_{r,s}t^{r+s+1}e^{yt}}{Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)}\right)\middle| x^{n-1}\right\rangle\\
=&\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)}\partial_t\left(e^{yt}\right)\middle| x^{n-1}\right\rangle\\
&+\left\langle \partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)}\right)e^{yt}\middle| x^{n-1}\right\rangle\\
=&yB_{n-1}^{(r,s)}(y)+\left\langle t\partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^{t}-P^{(r,s)}(t)}\right)e^{yt}\middle| \frac{x^{n}}{n}\right\rangle.
\end{aligned}$$

Note that the order of $t\partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t-P^{(r,s)}(t)}\right)$ is at least one.
We observe that

$$\begin{aligned} & t\partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\right) \\ &=(r+s+1)\,\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}-\left(\frac{t\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\right) \\ &\quad\times\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\right). \end{aligned} \tag{45}$$

Thus, by (45), we get

$$\begin{aligned} & \left\langle t\partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\right)e^{yt}\middle|\,\frac{x^n}{n}\right\rangle \\ &=\frac{r+s+1}{n}\left\langle\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}e^{yt}\middle|\,x^n\right\rangle \\ &\quad-\frac{1}{n}\left\langle\frac{t\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}e^{yt}\middle|\,\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}x^n\right\rangle \\ &=\frac{r+s+1}{n}B_n^{(r,s)}(y)-\frac{1}{n}\left\langle\frac{t\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}e^{yt}\middle|\,B_n^{(r,s)}(x)\right\rangle. \end{aligned} \tag{46}$$

Note that

$$\begin{aligned} & -\frac{1}{n}\left\langle\frac{t\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}e^{yt}\middle|\,B_n^{(r,s)}(x)\right\rangle \\ &=-\frac{1}{n}\left\langle\frac{t^{r+s+1}\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}e^{yt}\middle|\,\frac{B_{n+r+s}^{(r,s)}(x)}{(n+1)\cdots(n+r+s)}\right\rangle \\ &=-\frac{(n-1)!}{d_{r,s}(n+r+s)!} \\ &\quad\times\left\langle\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'e^{yt}\middle|\,\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}B_{n+r+s}^{(r,s)}(x)\right\rangle \\ &=-\frac{(n-1)!}{d_{r,s}(n+r+s)!}\sum_{m=0}^{n+r+s}\binom{n+r+s}{m}B_{n+r+s-m}^{(r,s)} \\ &\quad\times\left\langle\left(Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)\right)'e^{yt}\middle|\,B_m^{(r,s)}(x)\right\rangle. \end{aligned} \tag{47}$$

From (40), we have

$$\left(Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)\right)'e^{yt} \tag{48}$$
$$=\frac{1}{s!}\binom{r+s}{r}^{-1}(-1)^{s}e^{(y+1)t}t^{s}+\binom{r+s}{r}^{-1}$$
$$\times\sum_{j=0}^{s-1}\frac{1}{j!}\binom{r+s-1-j}{r-1}(-1)^{j}e^{(y+1)t}t^{j}$$
$$-\binom{r+s}{r}^{-1}\sum_{j=0}^{r-1}\frac{1}{j!}\binom{r+s-1-j}{s}e^{yt}t^{j}.$$

By (47) and (48), we get

$$-\frac{1}{n}\left\langle\frac{t\left(Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}e^{yt}\,\middle|\,B_{n}^{(r,s)}(x)\right\rangle \tag{49}$$
$$=\frac{(-1)^{s-1}}{\binom{n+r+s}{n-1}}\sum_{m=0}^{n+r+s}\binom{n+r+s}{m}B_{n+r+s-m}^{(r,s)}\left\{\frac{1}{s!}(-1)^{s}\left\langle e^{(y+1)t}t^{s}\,\middle|\,B_{m}^{(r,s)}(x)\right\rangle\right.$$
$$+\sum_{j=0}^{s-1}\frac{1}{j!}\binom{r+s-1-j}{r-1}(-1)^{j}\left\langle e^{(y+1)t}t^{j}\,\middle|\,B_{m}^{(r,s)}(x)\right\rangle$$
$$\left.-\sum_{j=0}^{r-1}\frac{1}{j!}\binom{r+s-1-j}{s}\left\langle e^{yt}t^{j}\,\middle|\,B_{m}^{(r,s)}(x)\right\rangle\right\}$$
$$=\frac{(-1)^{s-1}}{\binom{n+r+s}{n-1}}\left\{(-1)^{s}\sum_{m=s}^{n+r+s}\binom{n+r+s}{m}\binom{m}{s}B_{n+r+s-m}^{(r,s)}B_{m-s}^{(r,s)}(y+1)\right.$$
$$+\sum_{j=0}^{s-1}\sum_{m=j}^{n+r+s}(-1)^{j}\binom{r+s-1-j}{r-1}\binom{n+r+s}{m}\binom{m}{j}$$
$$\times B_{n+r+s-m}^{(r,s)}B_{m-j}^{(r,s)}(y+1)-\sum_{j=0}^{r-1}\sum_{m=j}^{n+r+s}\binom{r+s-1-j}{s}\binom{n+r+s}{m}\binom{m}{j}$$
$$\left.\times B_{n+r+s-m}^{(r,s)}B_{m-j}^{(r,s)}(y)\right\}.$$

For $r, s \geq 1$, $n \geq 1$, by (44), (46), and (49), we obtain the following equation:

$$\begin{aligned}
&\frac{n-r-s-1}{n} B_n^{(r,s)}(x) - x B_{n-1}^{(r,s)}(x) \qquad (50)\\
&= (-1)^{s-1} \binom{n+r+s}{n-1}^{-1} \Bigg\{ (-1)^s \sum_{m=s}^{n+r+s} \binom{n+r+s}{m} \binom{m}{s} B_{n+r+s-m}^{(r,s)} B_{m-s}^{(r,s)}(x+1) \\
&+ \sum_{j=0}^{s-1} \sum_{m=j}^{n+r+s} (-1)^j \binom{r+s-1-j}{r-1} \binom{n+r+s}{m} \binom{m}{j} B_{n+r+s-m}^{(r,s)} B_{m-j}^{(r,s)}(x+1) \\
&- \sum_{j=0}^{r-1} \sum_{m=j}^{n+r+s} \binom{r+s-1-j}{s} \binom{n+r+s}{m} \binom{m}{j} B_{n+r+s-m}^{(r,s)} B_{m-j}^{(r,s)}(x) \Bigg\}.
\end{aligned}$$

We see that this is the same as the result in (43) upon replacing n by $n+1$ and hence that it gives another proof for (43).

We will derive an identity by computing the following expression in two different ways. Here, we assume $m \leq r+s+1$.

$$\left\langle \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} t^m \,\middle|\, x^n \right\rangle. \qquad (51)$$

On one hand, it is

$$\begin{aligned}
&\left\langle \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} t^m \,\middle|\, x^n \right\rangle \qquad (52)\\
&= \left\langle \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} \,\middle|\, t^m x^n \right\rangle \\
&= (n)_m \left\langle \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} \,\middle|\, x^{n-m} \right\rangle \\
&= (n)_m \left\langle \sum_{i=0}^{\infty} B_i^{(r,s)} \frac{t^i}{i!} \,\middle|\, x^{n-m} \right\rangle \\
&= (n)_m B_{n-m}^{(r,s)}.
\end{aligned}$$

On the other hand, it is

$$\begin{aligned}
&\left\langle \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} t^m \,\middle|\, x^n \right\rangle \qquad (53)\\
&= \left\langle \partial_t \left(\frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\, e^t - P^{(r,s)}(t)} t^m \right) \,\middle|\, x^{n-1} \right\rangle
\end{aligned}$$

$$=\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\partial_t\left(t^m\right)\middle|\, x^{n-1}\right\rangle$$
$$+\left\langle \left(\partial_t \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\right)t^m\middle|\, x^{n-1}\right\rangle.$$

The first term of (53) is

$$\begin{aligned}
&\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\partial_t\left(t^m\right)\middle|\, x^{n-1}\right\rangle \\
=&m\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}t^{m-1}\middle|\, x^{n-1}\right\rangle \\
=&m\,(n-1)_{m-1}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\middle|\, x^{n-m}\right\rangle \\
=&m\,(n-1)_{m-1}\left\langle \sum_{i=0}^{\infty}B_i^{(r,s)}\frac{t^i}{i!}\middle|\, x^{n-m}\right\rangle \\
=&m\,(n-1)_{m-1}\,B_{n-m}^{(r,s)}.
\end{aligned} \tag{54}$$

We recall that

$$\begin{aligned}
&\partial_t\left(\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\right) \\
=&(r+s+1)\,\frac{d_{r,s}t^{r+s}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)} \\
&-\frac{\left(Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}.
\end{aligned}$$

Now, the second term of (53) is

$$\begin{aligned}
&\left\langle \left(\partial_t \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}\right)t^m\middle|\, x^{n-1}\right\rangle \\
=&(r+s+1)\left\langle t^{m-1}\middle|\, \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}x^{n-1}\right\rangle \\
&-\left\langle \frac{\left(Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}t^m\middle|\, \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^{t}-P^{(r,s)}(t)}x^{n-1}\right\rangle
\end{aligned} \tag{55}$$

$$= (r+s+1)\left\langle t^{m-1} \,\middle|\, B_{n-1}^{(r,s)}(x)\right\rangle - \left\langle \frac{\left(Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)} t^m \,\middle|\, B_{n-1}^{(r,s)}(x)\right\rangle$$

$$= (r+s+1)\,(n-1)_{m-1}\,B_{n-m}^{(r,s)} - \left\langle \frac{\left(Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)} t^m \,\middle|\, B_{n-1}^{(r,s)}(x)\right\rangle.$$

We observe that

$$\left\langle \frac{\left(Q^{(r,s)} e^t - P^{(r,s)}(t)\right)'}{Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)} t^m \,\middle|\, B_{n-1}^{(r,s)}(x)\right\rangle \tag{56}$$

$$= \frac{1}{d_{r,s}} \left\langle \left(Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right)' \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)} \,\middle|\, \frac{B_{n+r+s-m}^{(r,s)}(x)}{n\,(n+1)\cdots(n+r+s-m)}\right\rangle$$

$$= (-1)^s\,(r+s+1)!\frac{(n-1)!}{(n+r+s-m)!}\binom{r+s}{r}$$

$$\times \left\langle \left(Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right)' \,\middle|\, \frac{d_{r,s} t^{r+s+1}}{Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)} B_{n+r+s-m}^{(r,s)}(x)\right\rangle$$

$$= (-1)^s\,(r+s+1)!\frac{(n-1)!}{(n+r+s-m)!}\binom{r+s}{r}\sum_{l=0}^{n+r+s-m}\binom{n+r+s-m}{l} B_{n+r+s-m-l}^{(r,s)}$$

$$\times \left\langle \left(Q^{(r,s)}(t)\,e^t - P^{(r,s)}(t)\right)' \,\middle|\, B_l^{(r,s)}(x)\right\rangle$$

$$= (-1)^s\,(r+s+1)!\frac{(n-1)!}{(n+r+s-m)!}$$

$$\times \sum_{l=0}^{n+r+s-m}\binom{n+r+s-m}{l} B_{n+r+s-m-l}^{(r,s)} \left\{ \frac{1}{s!}(-1)^s \left\langle e^t t^s \,\middle|\, B_l^{(r,s)}(x)\right\rangle \right.$$

$$+ \sum_{j=0}^{s-1} \frac{1}{j!}\binom{r+s-1-j}{r-1}(-1)^j \left\langle e^t t^j \,\middle|\, B_l^{(r,s)}(x)\right\rangle$$

$$\left. - \sum_{j=0}^{r-1} \frac{1}{j!}\binom{r+s-1-j}{s} \left\langle t^j \,\middle|\, B_l^{(r,s)}(x)\right\rangle \right\}$$

$$= (-1)^s\,(r+s+1)!\frac{(n-1)!}{(n+r+s-m)!}$$

$$\times \left\{ (-1)^s \sum_{l=s}^{n+r+s-m} \binom{n+r+s-m}{l} \right.$$
$$\times \binom{l}{s} B_{n+r+s-m-l}^{(r,s)} B_{l-s}^{(r,s)}(1)$$
$$+ \sum_{j=0}^{s-1} \sum_{l=j}^{n+r+s-m} (-1)^j \binom{r+s-1-j}{r-1}$$
$$\times \binom{n+r+s-m}{l} \binom{l}{j} B_{n+r+s-m-l}^{(r,s)} B_{l-j}^{(r,s)}(1)$$
$$\left. - \sum_{j=0}^{r-1} \sum_{l=j}^{n+r+s-m} \binom{r+s-1-j}{s} \binom{n+r+s-m}{l} \binom{l}{j} B_{n+r+s-m-l}^{(r,s)} B_{l-j}^{(r,s)} \right\}.$$

For $1 \le m \le \min\{n, r+s+1\}$, by (52)–(56), we get

$$(n)_m B_{n-m}^{(r,s)} \tag{57}$$
$$= m(n-1)_{m-1} B_{n-m}^{(r,s)} + (r+s+1)(n-1)_{m-1} B_{n-m}^{(r,s)}$$
$$+ (-1)^{s-1} (r+s+1)! \frac{(n-1)!}{(n+r+s-m)!}$$
$$\times \left\{ (-1)^s \sum_{l=s}^{n+r+s-m} \binom{n+r+s-m}{l} \binom{l}{s} B_{n+r+s-m-l}^{(r,s)} B_{l-s}^{(r,s)}(1) \right.$$
$$+ \sum_{j=0}^{s-1} \sum_{l=j}^{n+r+s-m} (-1)^j \binom{r+s-1-j}{r-1}$$
$$\times \binom{n+r+s-m}{l} \binom{l}{j} B_{n+r+s-m-l}^{(r,s)} B_{l-j}^{(r,s)}(1)$$
$$\left. - \sum_{j=0}^{r-1} \sum_{l=j}^{n+r+s-m} \binom{r+s-1-j}{s} \binom{n+r+s-m}{l} \binom{l}{j} B_{n+r+s-m-l}^{(r,s)} B_{l-j}^{(r,s)} \right\}.$$

Dividing both sides by $(n-1)_{m-1}$, we obtain: for $1 \le m \le \min\{n, r+s+1\}$,

$$\{n-m-(r+s+1)\} B_{n-m}^{(r,s)} \tag{58}$$
$$= (-1)^{s-1} (r+s+1) \binom{n+r+s-m}{r+s}^{-1}$$

$$\times\left\{(-1)^s\sum_{l=s}^{n+r+s-m}\binom{n+r+s-m}{l}\binom{l}{s}B_{n+r+s-m-l}^{(r,s)}B_{l-s}^{(r,s)}(1)\right.$$
$$+\sum_{j=0}^{s-1}\sum_{l=j}^{n+r+s-m}(-1)^j\binom{r+s-1-j}{r-1}\binom{n+r+s-m}{l}\binom{l}{j}$$
$$\times B_{n+r+s-m-l}^{(r,s)}B_{l-j}^{(r,s)}(1)$$
$$\left.-\sum_{j=0}^{r-1}\sum_{l=j}^{n+r+s-m}\binom{r+s-1-j}{s}\binom{n+r+s-m}{l}\binom{l}{j}B_{n+r+s-m-l}^{(r,s)}B_{l-j}^{(r,s)}\right\}.$$

Let us consider the following two Sheffer sequences:

$$B_n^{(r,s)}(x)\sim\left(\frac{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}{d_{r,s}t^{r+s+1}},t\right),\quad (x)_n\sim\left(1,e^t-1\right). \tag{59}$$

From (14), (15), and (59), we have

$$B_n^{(r,s)}(x)=\sum_{m=0}^{n}C_{n,m}\,(x)_m\,, \tag{60}$$

where

$$\begin{aligned}C_{n,m}&=\left\langle\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\middle|\frac{1}{m!}\left(e^t-1\right)^m x^n\right\rangle\\&=\left\langle\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\middle|\sum_{l=m}^{\infty}S_2(l,m)\,\frac{t^l}{l!}x^n\right\rangle\\&=\sum_{l=m}^{n}\binom{n}{l}S_2(l,m)\left\langle\frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\middle|x^{n-l}\right\rangle\\&=\sum_{l=m}^{n}\binom{n}{l}S_2(l,m)\,B_{n-l}^{(r,s)}.\end{aligned} \tag{61}$$

Therefore, by (60) and (61), we obtain the following equation:

$$B_n^{(r,s)}(x)=\sum_{m=0}^{n}\sum_{l=m}^{n}\binom{n}{l}S_2(l,m)\,B_{n-l}^{(r,s)}(x)_m\,,\quad (n\geq 0)\,. \tag{62}$$

For $B_n^{(r,s)}(x) \sim \left(\frac{Q^{(r,s)}(t)e^t-P^{(r,s)}(t)}{d_{r,s}t^{r+s+1}}, t\right)$, $x^{(n)} = x(x+1)\cdots(x+n-1) \sim (1, 1-e^{-t})$, we have

$$B_n^{(r,s)}(x) = \sum_{m=0}^{n} C_{n,m} x^{(m)}, \quad (n \geq 0), \tag{63}$$

where

$$\begin{aligned} C_{n,m} &= \frac{1}{m!}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}e^t - P^{(r,s)}(t)} \left(1-e^{-t}\right)^m \right| x^n \right\rangle \\ &= \left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t - P^{(r,s)}(t)} e^{-mt} \right| \frac{1}{m!}\left(e^t-1\right)^m x^n \right\rangle \\ &= \left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t - P^{(r,s)}(t)} e^{-mt} \right| \sum_{l=m}^{\infty} S_2(l,m)\frac{t^l}{l!} x^n \right\rangle \\ &= \sum_{l=m}^{n} \binom{n}{l} S_2(l,m) \left\langle e^{-mt} \right| \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t - P^{(r,s)}(t)} x^{n-l} \right\rangle \\ &= \sum_{l=m}^{n} \binom{n}{l} S_2(l,m) \left\langle e^{-mt} \right| B_{n-l}^{(r,s)}(x) \right\rangle \\ &= \sum_{l=m}^{n} \binom{n}{l} S_2(l,m) B_{n-l}^{(r,s)}(-m). \end{aligned} \tag{64}$$

Therefore, by (63) and (64), we obtain the following equation:

$$B_n^{(r,s)}(x) = \sum_{m=0}^{n}\sum_{l=m}^{n} \binom{n}{l} S_2(l,m) B_{n-l}^{(r,s)}(-m) x^{(m)}. \tag{65}$$

For $B_n^{(r,s)}(x) \sim \left(\frac{Q^{(r,s)}(t)e^t-P^{(r,s)}(t)}{d_{r,s}t^{r+s+1}}, t\right)$, $B_n^{(s)}(x) \sim \left(\left(\frac{e^t-1}{t}\right)^s, t\right)$, we have

$$B_n^{(r,s)}(x) = \sum_{m=0}^{n} C_{n,m} B_m^{(s)}(x), \tag{66}$$

where

$$\begin{aligned} C_{n,m} &= \frac{1}{m!}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t - P^{(r,s)}(t)} \left(\frac{e^t-1}{t}\right)^s t^m \right| x^n \right\rangle \\ &= \frac{1}{m!}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)e^t - P^{(r,s)}(t)} \left(\frac{e^t-1}{t}\right)^s \right| t^m x^n \right\rangle \end{aligned} \tag{67}$$

$$=\binom{n}{m}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)} \middle| \left(\frac{e^t-1}{t}\right)^s x^{n-m}\right\rangle$$

$$=\binom{n}{m}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)} \middle| \sum_{l=0}^{\infty}\frac{s!}{(l+s)!}S_2(l+s,s)\,t^l x^{n-m}\right\rangle$$

$$=\binom{n}{m}\sum_{l=0}^{n-m}\frac{s!}{(l+s)!}S_2(l+s,s)\,(n-m)_l$$

$$=\times\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)} \middle| x^{n-m-l}\right\rangle$$

$$=\binom{n}{m}\sum_{l=0}^{n-m}\frac{s!}{(l+s)!}S_2(l+s,s)\,(n-m)_l\,B_{n-m-l}^{(r,s)}.$$

Therefore, by (66) and (67), we obtain the following equation.

$$B_n^{(r,s)}(x)=\sum_{m=0}^{n}\binom{n}{m}\sum_{l=0}^{n-m}\frac{s!}{(l+s)!}S_2(l+s,s)\,(n-m)_l\,B_{n-m-l}^{(r,s)}B_m^{(s)}(x)\,. \tag{68}$$

It is well known that Frobenius–Euler polynomials of order $s\ (\in \mathbb{N})$ are defined by the generating function to be

$$\left(\frac{1-\lambda}{e^t-\lambda}\right)^s e^{xt}=\sum_{n=0}^{\infty}H_n^{(s)}(x\mid\lambda)\,\frac{t^n}{n!},\quad (\lambda\in\mathbb{C}\text{ with }\lambda\neq 1)\,. \tag{69}$$

From (69), we have

$$H_n^{(s)}(x\mid\lambda)\sim\left(\left(\frac{e^t-\lambda}{1-\lambda}\right)^s,t\right). \tag{70}$$

By (14), (15), (25), and (70), we have

$$B_n^{(r,s)}(x)=\sum_{m=0}^{n}C_{n,m}H_m^{(s)}(x\mid\lambda)\,, \tag{71}$$

where

$$C_{n,m}=\frac{1}{m!}\left\langle \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}\left(\frac{e^t-\lambda}{1-\lambda}\right)^s t^m \middle| x^n\right\rangle \tag{72}$$

$$=\frac{\binom{n}{m}}{(1-\lambda)^s}\left\langle (e^t-\lambda)^s \middle| \frac{d_{r,s}t^{r+s+1}}{Q^{(r,s)}(t)\,e^t-P^{(r,s)}(t)}x^{n-m}\right\rangle$$

$$= \frac{\binom{n}{m}}{(1-\lambda)^s} \left\langle \sum_{j=0}^{s} \binom{s}{j} (-\lambda)^{s-j} e^{jt} \middle| B_{n-m}^{(r,s)}(x) \right\rangle$$

$$= \frac{\binom{n}{m}}{(1-\lambda)^s} \sum_{j=0}^{s} \binom{s}{j} (-\lambda)^{s-j} B_{n-m}^{(r,s)}(j) .$$

Therefore, by (71) and (72), we obtain the following equation:

$$B_n^{(r,s)}(x) = (1-\lambda)^{-s} \sum_{m=0}^{n} \binom{n}{m} \sum_{j=0}^{s} \binom{s}{j} (-\lambda)^{s-j} B_{n-m}^{(r,s)}(j) H_m^{(s)}(x \mid \lambda) ,$$

where $n \geq 0$.

References

1. Dilcher, K.: Bernoulli numbers and confluent hypergeometric functions. In: Number Theory for the Millennium, I (Urbana, IL, 2000), pp. 343–363. A K Peters, Natick, MA (2002)
2. Dilcher, K., Malloch, L.: Arithmetic properties of Bernoulli–Padé numbers and polynomials. J Number Theory **92**(2), 330–347 (2002) doi:10.1006/jnth.2001.2696. http://dx.doi.org/10.1006/jnth.2001.2696
3. Fang, Q, Wang, T.: Umbral calculus and invariant sequences. Ars Combin. **101**, 257–264 (2011)
4. Howard, F.T.: Some sequences of rational numbers related to the exponential function. Duke Math. J. **34**, 701–716 (1967)
5. Howard, F.T.: The van der Pol numbers and a related sequence of rational numbers. Math. Nachr. **42**, 89–102 (1969)
6. Howard, F.T.: Properties of the van der Pol numbers and polynomials. J. Reine Angew. Math. **260**, 35–46 (1973)
7. Howard, F.T.: Polynomials related to the Bessel functions. Trans. Am. Math. Soc. **210**, 233–248 (1975)
8. Howard, F.T.: Properties of the van der Pol numbers and polynomails, J. Reine Angew. Math. **260**, 35–46 (1973)
9. Kim, D.S., Kim, T.: q-Bernoulli polynomials and q-umbral calculus. Sci. China Math. **57**(9), 1867–1874 (2014). doi:10.1007/s11425-014-4821-3. URL http://dx.doi.org/10.1007/s11425-014-4821-3
10. Kim, D.S., Kim, T., Dolgy, D.V., Komatsu, T.: Barnes-type degenerate Bernoulli polynomials. Adv. Stud. Contemp. Math. **25**, 121–146 (2015)
11. Kim, T.: Identities involving Laguerre polynomials derived from umbral calculus. Russ. J. Math. Phys. **21**(1), 36–45 (2014). doi:10.1134/S1061920814010038. URL http://dx.doi.org/10.1134/S1061920814010038
12. Kim, T., Mansour, T.: Umbral calculus associated with Frobenius-type Eulerian polynomials. Russ. J. Math. Phys. **21**(4), 484–493 (2014). doi:10.1134/S1061920814040062. URL http://dx.doi.org/10.1134/S1061920814040062
13. Roman, S.: The umbral calculus. In: Pure and Applied Mathematics, vol. 111. Academic/[Harcourt Brace Jovanovich Publishers], London/New York (1984)

Hadamard Matrices: Insights into Their Growth Factor and Determinant Computations

Christos D. Kravvaritis

Abstract In this expository paper we survey the most important progress in the growth problem for Hadamard matrices. The history of the problem is presented, the importance of determinant calculations is highlighted, and the relevant open problems are discussed. Emphasis is laid on the contribution of determinant manipulations to the study of the growth factor for Hadamard matrices after application of Gaussian Elimination with complete pivoting on them, which is an important scientific field in Numerical Analysis.

Keywords Hadamard matrices • Gaussian Elimination • Complete pivoting • Growth factor • Determinant calculus

Introduction

Hadamard Matrices

It is interesting to study the pivot patterns of Hadamard matrices, as they are the only known matrices that attain growth factor equal to their order. Besides, they are an attractive tool for several applications [18, 25, 29, 47, 55, 59, 64].

For instance, in Coding Theory (error correction coding) Hadamard matrices are used for generating the error correcting Hadamard codes, which have good coding properties. Hence, data sequences are transmitted more assuredly over a noisy channel and high error correction rate is provided. A famous application of the Hadamard code was in the NASA Mariner spacecraft missions in 1969 and 1972 to Mars, for correcting the picture transmission errors and for digitalizing and transmitting photos of Mars back to Earth, where the messages were fairly weak. Information transmission during recent flybys of the outer planets in the solar system is based on Hadamard matrices, too. Hadamard matrices are associated with symmetric balanced incomplete block designs (SBIBDs) in the way that a

C.D. Kravvaritis (✉)
Department of Mathematics, University of Athens, Panepistimiopolis 15784, Athens, Greece
e-mail: ckrav@math.uoa.gr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_17

Hadamard matrix of order $4t$ exists if and only if there is an SBIBD with parameters $(4t-1, 2t-1, t-1)$. Moreover, Hadamard matrices are used in Cryptography for guaranteeing the encryption, concealment, and secure transmission of data over a nonsecure channel. Hadamard matrices are related to bent functions, which have the highest possible non-linearity. Thus they effectively disguise and confuse characteristics of data sequences. Applications of Hadamard matrices in digital Signal/Image Processing, modern telecommunications, optical multiplexing, and design/analysis of statistics can be found as well.

Definition 1. A *Hadamard matrix* $H \equiv H_n$ of order n (denoted by H_n) has entries ± 1 and satisfies $HH^T = H^T H = nI_n$.

Example 1.

$$H_1 = (1)\,,\; H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},\; H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

It follows from the above definition that every two distinct rows and columns of a Hadamard matrix are mutually orthogonal. Hadamard matrices are defined only for $n = 1, 2$ and $n \equiv 0 \bmod 4$. Hadamard matrices can be considered as a special case of weighing matrices. A $(0, 1, -1)$ matrix $W = W(n, n-k)$, $k = 0, 1, 2, \ldots$, of order n satisfying $WW^T = W^T W = (n-k)I_n$ is called a *weighing matrix of order n and weight $n-k$* or simply a *weighing matrix*. A $W(n,n)$, $n \equiv 0 \pmod 4$, is a Hadamard matrix.

In 1893 Hadamard [22] specified quadratic matrices of orders 12 and 20 with entries ± 1 having all their rows and columns mutually orthogonal. These matrices satisfied the following famous Hadamard inequality.

Theorem 1 (Hadamard's Inequality [22], [8, p. 496]). *For any matrix $A = (a_{ij})_{1 \le i,j \le n}$ with entries on the unit disc*

$$|\det A| \le \prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}} = \prod_{j=1}^{n} \|a_j\|_2 \le n^{\frac{n}{2}},$$

where a_j denotes the jth column of A. The equalities hold if and only if $|a_{ij}| = 1$ for all i, j and the rows of A are mutually orthogonal.

However, these matrices for every order being a power of 2 have been first discovered in 1867 by Sylvester [52]. The following well-known result describes the possible order n of a Hadamard matrix.

Theorem 2 ([22]). *If H is an $n \times n$ Hadamard matrix and $n > 2$, then n is a multiple of 4.*

Theorem 2 is usually proved by describing the sign patterns in three rows of a Hadamard matrix like

$$\begin{array}{cccc} \overbrace{+\cdots+}^{x} & \overbrace{+\cdots+}^{y} & \overbrace{+\cdots+}^{z} & \overbrace{+\cdots+}^{w} \\ +\cdots+ & +\cdots+ & -\cdots- & -\cdots- \\ +\cdots+ & -\cdots- & +\cdots+ & -\cdots- \end{array},$$

where x, y, z, w denote the number of columns of the respective form. Another elegant proof for Theorem 2 is given in [8], which is facilitated by the next result.

Proposition 1 ([8, 38]). *Let B be an $n \times n$ matrix with elements ± 1. Then*

(i) $\det B$ *is an integer and* 2^{n-1} *divides* $\det B$*;*
(ii) *when $n \leq 6$, the only possible values for* $\det B$ *are given in Table 1, and they do all occur.*

Proposition 1 leads to the following open conjecture.

Conjecture 1.

Determinants of ± 1 matrices
The determinant of an $n \times n$ matrix with elements ± 1 is

$$0 \text{ or } p, \text{ for } p = 2^{n-1},\ 2 \cdot 2^{n-1},\ 3 \cdot 2^{n-1}, \ldots, s \cdot 2^{n-1},$$

where

$$s \cdot 2^{n-1} = \max\{\det(A) | A \in \mathbb{R}^{n \times n}, \text{with entries } \pm 1\}$$

and the value 0 is excluded from the case $n = 1$.

One more recent proof for Theorem 2 can be found in [4]. Theorem 2 does not assure the existence of Hadamard matrices for every n being a multiple of 4. There exists the following relevant conjecture.

Table 1 Possible determinant values for $n \times n$ ± 1 matrices

n	det
1	1
2	0,2
3	0,4
4	0,8,16
5	0,16,32,48
6	0,32,64,96,128,160
7	0,64,128,192,256,320,384,448,512,576

Conjecture 2.

Existence of Hadamard matrices
There exists a Hadamard matrix of order n for every n multiple of 4.

This conjecture seems very realistic but is not proved yet. The smallest order, for which a Hadamard matrix has not been found, is 668 [29]. Recently a Hadamard matrix of order 428 was found [32]. Other orders smaller than 1000, for which Hadamard matrices have not been found yet, are 716, 764, 892.

Finally, we introduce an important notion for the study of the problem, the *Hadamard equivalence*.

Definition 2. Two matrices are said to be *Hadamard equivalent* or *H-equivalent* if one can be obtained from the other by a sequence of the operations:

1. interchange any pairs of rows and/or columns;
2. multiply any rows and/or columns through by -1.

Notations

Throughout this paper we assume, without loss of generality, that the entries of the first row and column of a Hadamard matrix are always $+1$ (normalized form), because this can be achieved easily by multiplying columns and/or rows with -1 and leaves unaffected the properties of the matrix. The entries of an $(1,-1)$ matrix will be denoted by $(+,-)$. I_n and J_n stand for the identity matrix of order n and the matrix with ones of order n, respectively. We write $A(j)$ for the absolute value of the determinant of the $j \times j$ principal submatrix in the upper left corner of the matrix A, i.e. the $j \times j$ leading principal minor. Similarly, $A[j]$ denotes the magnitude of the determinant of the lower right $j \times j$ principal submatrix of A.

We write $J_{b_1,b_2,\cdots,b_z}$ for the all ones matrix with diagonal blocks of sizes $b_1 \times b_1,\ b_2 \times b_2\ \cdots\ b_z \times b_z$, and $a_{ij}J_{b_1,b_2,\cdots,b_z}$ for the matrix, for which the elements of the block with corners

$(i + b_1 + b_2 + \cdots + b_{j-1}, i + b_1 + b_2 + \cdots + b_{i-1})$,

$(i + b_1 + b_2 + \cdots + b_{j-1}, b_1 + b_2 + \cdots + b_i)$,

$(b_1 + b_2 + \cdots + b_j, i + b_1 + b_2 + \cdots + b_{i-1})$,

$(b_1+b_2+\cdots+b_j, b_1+b_2+\cdots+b_i)$ are the integers a_{ij}. We write $(k_i-a_{ii})I_{b_1,b_2,\cdots,b_z}$ for the matrix direct sum

$(k_1 - a_{11})I_{b_1} + (k_2 - a_{22})I_{b_2} + \cdots + (k_z - a_{zz})I_{b_z}$.

Minors of Hadamard Matrices

Sharpe's Results

The first known effort for calculating minors of Hadamard matrices is estimated to be accomplished in 1907 by Sharpe [49]. The essence of his results is summarized in the next Theorem.

Theorem 3. *Let H be a Hadamard matrix of order n. The cofactor of any element of H is* $\pm n^{\frac{n}{2}-1}$*, the sign being the same as the sign of that element. The second minors of H are* $2n^{\frac{n}{2}-2}$ *or 0, according as the complementary minor is 2 or 0. The third minors of H are* $4n^{\frac{n}{2}-3}$ *or 0, according as the complementary minor is 4 or 0.*

Theorem 3 actually gives all possible values of $n-j$, $j = 1, 2, 3$, minors of Hadamard matrices according to the determinant of the respective excluded $j \times j$ matrix. Sharpe's idea, which leads to short, elegant proofs, is based on considering a special arrangement of the entries in a Hadamard matrix. It takes appropriately into account the definition $HH^T = nI_n$ by observing that when comparing any pair of rows or columns, there is always an equal number of changes and permanences of sign amongst the corresponding elements. But unfortunately it doesn't seem to be applicable for calculating minors of orders $n-j$ for $j > 3$.

A New Technique

A new idea for computing $n-j$ minors was proposed in 2001 [34]. First of all, the notion of a matrix denoted as U_j, containing all possible columns of a normalized Hadamard matrix clustered together for some number of rows j, was introduced as follows.

Let $\mathbf{y}_{\beta+1}^T$ be the vectors containing the binary representation of each integer $\beta + 2^{j-1}$ for $\beta = 0, \ldots, 2^{j-1} - 1$. Replace all zero entries of $\mathbf{y}_{\beta+1}^T$ by -1 and define the $j \times 1$ vectors $\mathbf{u}_k = \mathbf{y}_{2^{j-1}-k+1}$, $k = 1, \ldots, 2^{j-1}$. U_j shall denote all the matrices with j rows and the appropriate number of columns, in which $\mathbf{u}_k$ occurs u_k times. In other words, U_j is the matrix containing all possible 2^{j-1} columns of size j with elements ± 1 starting with $+1$. So,

$$U_j = \begin{array}{ccccc} \overbrace{+\ldots+}^{u_1} & \overbrace{+\ldots+}^{u_2} & \ldots & \overbrace{+\ldots+}^{u_{2^{j-1}-1}} & \overbrace{+\ldots+}^{u_{2^{j-1}}} \\ +\ldots+ & +\ldots+ & \ldots & -\ldots- & -\ldots- \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ +\ldots+ & +\ldots+ & \ldots & -\ldots- & -\ldots- \\ +\ldots+ & -\ldots- & \ldots & +\ldots+ & -\ldots- \end{array} = \begin{array}{ccccc} u_1 & u_2 & \ldots & u_{2^{j-1}-1} & u_{2^{j-1}} \\ + & + & \ldots & + & + \\ + & + & \ldots & - & - \\ \vdots & \vdots & & \vdots & \vdots \\ + & + & \ldots & - & - \\ + & - & \ldots & + & - \end{array}.$$

Example 2. $U_3 = \begin{matrix} u_1 & u_2 & u_3 & u_4 \\ + & + & + & + \\ + & + & - & - \\ + & - & + & - \end{matrix}$, $U_4 = \begin{matrix} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 & u_8 \\ + & + & + & + & + & + & + & + \\ + & + & + & + & - & - & - & - \\ + & + & - & - & + & + & - & - \\ + & - & + & - & + & - & + & - \end{matrix}$

The matrix U_j is important in this study because it depicts a general form for the first j rows of a normalized Hadamard matrix.

With respect to this definition, every Hadamard matrix H can be written in the form

$$H = \begin{bmatrix} M & U_j \\ & C \end{bmatrix}, \qquad (1)$$

where M and C are of orders j and $n-j$, respectively. The various values of $\det C$ are the required $n-j$ minors. It is important to emphasize that this method calculated all possible $n-j$ minors. The selected rows, which are written as the first j rows of H, do not necessarily appear there but they can be located everywhere in the matrix. They can be made to appear as first rows with appropriate row and/or column interchanges, which leave unaffected the properties of the matrix. They are written at the top only for the sake of better presentation and without any loss of generality. This idea serves also the idea of standardizing a technique. The fact that *all possible upper left* $j \times j$ *corners* are examined guarantees that with this technique *all possible* $(n-j) \times (n-j)$ *minors of* A are calculated and no appearing values are missed out.

Next, the matrix CC^T is formed and, taking into account the definition of Hadamard matrices $HH^T = nI_n$, it can be written as a block matrix of the form

$$CC^T = (n-j)I_{u_1,u_2,\ldots,u_z} + a_{ik}J_{u_1,u_2,\ldots,u_z}, \quad z = 2^{j-1}, \qquad (2)$$

where $(a_{ik}) = (-\mathbf{u}_i \cdot \mathbf{u}_k)$, with $\cdot$ the inner product. The required minor $\det C$ can be computed using the following Theorem.

Theorem 4 (Determinant Simplification Theorem [35]). *Let* $A = (k_i - a_{ii})I_{b_1,b_2,\cdots,b_z} + a_{ij}J_{b_1,b_2,\cdots,b_z}$, $i,j = 1,\ldots,z$. *Then*

$$\det A = \prod_{i=1}^{z}(k_i - a_{ii})^{b_i-1} \det D,$$

where

$$D = \begin{bmatrix} k_1 + (b_1-1)a_{11} & b_2a_{12} & b_3a_{13} \cdots & b_za_{1z} \\ b_1a_{21} & k_2 + (b_2-1)a_{22} & b_3a_{23} \cdots & b_za_{2z} \\ \vdots & \vdots & \vdots & \vdots \\ b_1a_{z1} & b_2a_{z2} & b_3a_{z2} \cdots & k_z + (b_z-1)a_{zz} \end{bmatrix}.$$

Hence,

$$\det C = (n^{n-2^{j-1}-j} \det D)^{1/2}, \tag{3}$$

where

$$D = \begin{bmatrix} n-ju_1 & u_2a_{12} & u_3a_{13} & \cdots & u_za_{1z} \\ u_1a_{21} & n-ju_2 & u_3a_{23} & \cdots & u_za_{2z} \\ \vdots & \vdots & \vdots & & \vdots \\ u_1a_{z1} & u_2a_{z2} & u_3a_{z2} & \cdots & n-ju_z \end{bmatrix}. \tag{4}$$

In order to find only the feasible values of u_i, the properties of a Hadamard matrix led to the formulation of the following constraints for the number of columns of U_j

$$0 \le t - u_1 - u_2 - \ldots - u_{2^{j-3}} \le j,$$
$$0 \le t - u_{2^{j-3}+1} - \ldots - u_{2^{j-2}} \le j,$$
$$0 \le t - u_{2^{j-2}+1} - \ldots - u_{2^{j-2}+2^{j-3}} \le j,$$
$$0 \le t - u_{2^{j-2}+2^{j-3}+1} - \ldots - u_{2^{j-3}1} \le j.$$

This problem has been treated from an optimization point of view and the appropriate methods have been applied for deriving the required results. This technique confirmed Sharpe's already known results and extended them to the $n-4$ case, as given in the next Theorem.

Theorem 5. *The $n-4$ minors of a Hadamard matrix of order n are 0, $8n^{\frac{n}{2}-4}$ or $16n^{\frac{n}{2}-4}$.*

The idea used for proving Theorem 5 could be theoretically generalized to an algorithm MINORS1 for computing all possible $n-j$ minors of Hadamard matrices. But it was impossible to implement this idea in practice for $j > 4$ due to the demanding involved complexity. This issue is mainly connected with the huge amount of all possible $j \times j$ matrices M, which can exist in a Hadamard matrix, especially while j grows, and also with the difficulty to handle the values of u_i effectively.

A Further Development

A few years later in 2003 a new algorithm MINORS2 for computing $n-j$ minors of Hadamard matrices was proposed [37]. This idea is similar to the one used in [34] but there are some essential differences, which made it possible to obtain new results.

This paper suggests the notion of IP equivalence as a criterion for reducing the total number of matrices M, when the form (1) is considered. Two matrices are called *IP equivalent* (inner product equivalent) if they have the same IP profile vector. For a matrix M of order j the IP profile vector is the vector $(m_{12}, m_{13}, \ldots, m_{1j}, m_{23}, m_{24}, \ldots, m_{2j}, \ldots, m_{j-1,j})$ formed by all possible inner products m_{kl} for every pair of rows k and l of M. The usefulness of this criterion will be explained later in this subsection.

In order to skip more unimportant matrices M, the notion of Hadamard submatrix was also introduced. The matrix $N := [M \quad U_j]$ is formed and the linear system, occurring from the order of N and the orthogonality of any three of its rows (which constitute the matrix N_1), with unknowns the number of columns of U_j is set up. If the solution of this system has nonnegative integer components and the solution satisfies the following Lemma 1, then N_1 is called a *Hadamard submatrix*.

Lemma 1. *For the first j rows, $j \geq 3$, of a normalized Hadamard matrix H of order n, $n > 3$, and for all the 2^{j-1} possible columns $\mathbf{u}_1, \ldots, \mathbf{u}_{2^{j-1}}$ of U_j, it holds*

$$0 \leq u_i \leq \frac{n}{4}, \textit{for } i = 1, \ldots, 2^{j-1}.$$

This Lemma is useful for finding only feasible values of u_i, and moreover for limiting any calculations, since the numbers of columns in U_j are proven to be subject to the constraints given in Lemma 1. The notion of Hadamard submatrix is useful since it identifies matrices, which are possible to exist embedded inside a Hadamard matrix, and skips the rest.

The proposed algorithm for computing minors has the following functionality. First, all IP inequivalent ± 1 matrices M of order j are generated. From these matrices only the possible Hadamard submatrices are kept. Then, for each possible Hadamard submatrix, a linear system with unknowns the numbers of columns of U_j is formed, taking into account the dimension and the orthogonality of every two rows of a Hadamard matrix. The determinant is evaluated with the help of formulas (3) and (4).

The concept of IP equivalence is useful in the above procedure for constructing *only the different* linear systems that can arise by considering the first j rows of a Hadamard matrix, i.e. only IP inequivalent matrices give different linear systems. The IP equivalent ones yield same systems and don't have to be taken into account for the rest of the algorithm.

The new technique verifies the already known results up to $j = 4$. The technique seems to be applicable for some more values of j, but the result won't be of general nature, i.e. it can't be proved for every n but only for specific values of n fixed. This happens because in such cases the number of unknowns in the linear systems is greater than the number of equations, so there exist parameters in the solutions. The only known possibility so far for estimating the parameters, which actually correspond to some columns of U_j, is Lemma 1. The upper bound given there is n-dependent, so the same will hold for the resulting values of minors.

Algorithm MINORS2 has been tested for $n = 12, 16$ and for $j = 1, 2, \ldots, 7$ and the results are presented in Table 2. These results lead to formulating the following conjecture for minors of Hadamard matrices.

Table 2 All possible values of minors of orders $n-1,\ldots,n-7$ for Hadamard matrices of order $n = 12, 16$, where $D = n^{n/2}$

Order	Values of minors
$n-1$	$(1/n)D$
$n-2$	0, $(2/n^2)D$
$n-3$	0, $(4/n^3)D$
$n-4$	0, $(q/n^4)D,\ q = 8, 16$
$n-5$	0, $(q/n^5)D,\ q = 16, 32, 48$
$n-6$	0, $(q/n^6)D,\ q = 32, 64, 96, 128, 160$
$n-7$	0, $(q/n^7)D,\ q = 64, 128, 192, 256, 320, 384, 448, 512, 576$

Conjecture 3.

$(n-j) \times (n-j)$ **minors of Hadamard matrices**
All possible $(n-j) \times (n-j), j \geq 1$, minors of Hadamard matrices are

$$0 \text{ or } p \cdot n^{(n/2)-j},\ \text{ for } p = 2^{j-1}, 2 \cdot 2^{j-1}, 3 \cdot 2^{j-1}, \ldots, s \cdot 2^{j-1},$$

where

$$s \cdot 2^{j-1} = \max\{\det A | A \in \mathbb{R}^{j \times j}, \text{with entries } \pm 1\}$$

and the value 0 is excluded from the case $j = 1$.

Alternatively, the obtained results can be summarized with the formula

$$M_{n-j} = 0 \text{ or } p \cdot n^{(n/2)-j},\ j = 0, 1, 2, \ldots,$$

where for the evaluation of the coefficient p the following procedure is adopted:
$p := 2^{j-1}$
$m := \max\{\det(A) | A \in \mathbb{R}^{j \times j}, \text{with elements} \pm 1\}$
$k := 1$
repeat
  $p = k \cdot p$
  $k = k + 1$
until
  $p = m.$

The maximum determinant values for ± 1 $n \times n$ matrices are given in Table 1.

The Latest Algorithm

The state-of-the-art algorithm MINORS3 for computing $n-j$ minors of Hadamard matrices was introduced in [40]. The algorithm employs the same main idea as [37], but it handles in a totally different and more effective way the classification of matrices M and the final determinant.

The idea is again based on the partitioning (1). The possible existing matrices M in the upper left $j \times j$ corner are identified with the help of algorithm EXIST. This algorithm specifies whether a set of $j \times j$ matrices M_i can always exist or not among all possible columns that can appear in the first j rows of a Hadamard matrix of order n, or equivalently in its upper left corner (this can be done with H-equivalent column interchanges). The main idea of the algorithm is to form the linear system that results (1) from summing all possible columns of length j with elements ± 1 starting with $+1$ (namely, the total number of columns of the matrix U_j must be equal to the order n) and (2) from the property that every two distinct rows of a Hadamard matrix are orthogonal. The system is solved (the parameters, if there are any, are bounded with the help of Lemma 1) and it is examined whether the columns, which constitute the matrices under investigation M_i, always appear among the solutions. If they appear, it is implied that the matrices M_i always exist within the first j rows of a Hadamard matrix of order n, otherwise not.

EXIST is actually a generalization of the criterion "Hadamard submatrix," which is implemented in a symbolical computer algebra setting that guarantees furthermore its precision. Algorithm EXIST is based formally on the special structure and properties of Hadamard matrices, in contrast with the more intuitive and less stringent concepts of "IP equivalent matrices" and "Hadamard submatrix." However, the performance of EXIST is about the same as the combination of these two criteria, i.e. it doesn't achieve any new remarkable reduction in the total amount of possible matrices under examination M. From now on, the intention is to write down for every possible upper left $j \times j$ corner M the appearing values for the determinant of C, which is the required minor. For this purpose, the following preliminary results are used.

Lemma 2. *Let $A = (k-\lambda)I_v + \lambda J_v$, where $k,\ \lambda$ are integers. Then,*

$$\det A = [k + (v-1)\lambda](k-\lambda)^{v-1} \tag{5}$$

and for $k \neq \lambda, -(v-1)\lambda$, A is nonsingular with

$$A^{-1} = \frac{1}{k^2 + (v-2)k\lambda - (v-1)\lambda^2}\{[k + (v-1)\lambda]I_v - \lambda J_v\}. \tag{6}$$

The first part of Lemma 2 is straightforward to show, while the second part is a special case of the Sherman–Morrison formula [7, p. 239], which computes the inverse of a rank-one-correction of a nonsingular matrix B as

$$(B - uv^T)^{-1} = B^{-1} + \frac{B^{-1}uv^TB^{-1}}{1 - v^TB^{-1}u},$$

where u, v are vectors and $v^TB^{-1}u \neq 1$. Indeed, (6) occurs for $B = (\kappa - \lambda)I_v$ and $u = -\lambda[1\ 1\ \dots\ 1]^T$ and $v = [1\ 1\ \dots\ 1]^T$.

Lemma 3 (Schur Determinant Formula [30, p. 21]). *Let* $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$, B_1 *nonsingular. Then*

$$\det B = \det B_1 \cdot \det(B_4 - B_3B_1^{-1}B_2). \tag{7}$$

In the sequel, a system of $1 + \binom{j}{2}$ equations with 2^{j-1} unknowns the numbers of columns of U_j is solved for every M; the system results from the order of H and from the orthogonality of its first j rows. The matrix C^TC is formed taking into account $H^TH = nI_n$ and the result is written in block form as

$$C^TC \equiv \begin{bmatrix} E_1 & F_1 \\ F_1^T & G_1 \end{bmatrix},$$

where $E_1 = nI_{u_1} - jJ_{u_1}$. Initially, the matrix

$$G_1 - F_1^TE_1^{-1}F_1 \equiv \begin{bmatrix} E_2 & F_2 \\ F_2^T & G_2 \end{bmatrix}$$

is computed, where E_2 is of order $u_2 \times u_2$. Then, according to (7),

$$\begin{aligned} \det C^TC &= \det E_1 \cdot \det(G_1 - F_1^TE_1^{-1}F_1) \\ &= \det E_1 \cdot \det E_2 \cdot \det(G_2 - F_2^TE_2^{-1}F_2). \end{aligned}$$

From (5) $\det E_1$ and $\det E_2$ are calculated and the algorithm proceeds with calculating $G_2 - F_2^TE_2^{-1}F_2$ with the help of (6). So the aim is to derive $\det C^TC$ by consecutive applications of formula (7), with help of (5) and (6).

Due to the clustering of same columns in U_j, C^TC and all the intermediate matrices $G_k - F_k^TE_k^{-1}F_k$, $k = 1, \dots, 2^{j-1} - 1$, will appear in block form with diagonal blocks of known orders $u_i \times u_i$, $i = k+1, \dots, 2^{j-1}$, and due to the orthogonality of H, all blocks will be of the form $(a - b)I + bJ$, for various a, b.

For $k = 2, \dots, 2^{j-1} - 1$ the sequence of matrices

$$G_k - F_k^TE_k^{-1}F_k \equiv \begin{bmatrix} E_{k+1} & F_{k+1} \\ F_{k+1}^T & G_{k+1} \end{bmatrix}$$

is computed, where each E_k is of order $u_k \times u_k$. The determinants of E_k are stored. The last matrix of the sequence

$$G_{2^{j-1}-1} - F_{2^{j-1}-1}^T E_{2^{j-1}-1}^{-1} F_{2^{j-1}-1} = E_{2^{j-1}}$$

consists of one block of order $u_{2^{j-1}} \times u_{2^{j-1}}$ and its determinant is evaluated directly with (5). The required absolute value of the determinant is computed from the formulas

$$\det D^T D := \prod_{i=1}^{2^{j-1}} \det E_i, \; |\det D| = \sqrt{\det D^T D}.$$

Although algorithm MINORS3 might seem extremely complicated at a first glance, it is designed in such a way that the special structure and properties of every appearing matrix are taken into account. So, all necessary matrix multiplications and inversions are not performed explicitly but in an efficient manner and the total computational cost remains at relatively low levels. The computational cost is $O(2^{3(j-1)})$ operations for matrix multiplications and inversions. This amount is significantly less than the number of flops required, if the calculations are performed explicitly, even if the fastest possible algorithm for matrix multiplication and inversion is used (see [26]). It is interesting to mention that some matrix multiplications in MINORS3 are carried out with the lowest possible cost, which is n^2 operations for $n \times n$ matrices.

Regardless of all this effort and although the minimum possible cost for matrix multiplication is reached, the total computational cost of the algorithm remains at high levels, because of the exhaustive (complete) searches performed for determining the possible matrices M and the feasible values of parameters according to Lemma 1. Unfortunately these searches cannot be avoided because if some cases are excluded, some values of minors might be skipped, which could probably appear, and in this way the result is false, since not all possible values will be calculated. Algorithm MINORS3 verified the results of Table 2 and extended them only for the case $n = 20$.

Gaussian Elimination and the Growth Problem

Consider a linear system of the form $A \cdot x = b$, where $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is nonsingular. Gaussian elimination (GE) [7, 19, 26, 27, 57] is the simplest way to solve such a system by hand, and also the standard method for solving it on computers. GE without pivoting fails if any of the pivots is zero and it behaves worse if any pivot becomes close to zero. In this case the method can be carried out to completion but it is totally unstable and the obtained results may be totally wrong, as it is already demonstrated in a famous example by Forsythe and Moler [15].

Therefore, a search for the element with maximum absolute value is performed. If the search is done in the respective column, then we have GE with partial pivoting, else if it is done in the respective lower right submatrix we have GE with complete pivoting. Let $A^{(k)} = [a_{ij}^{(k)}]$ denote the matrix obtained after the first k pivoting operations, so $A^{(n-1)}$ is the final upper triangular matrix. A diagonal entry of that final matrix will be called a *pivot*.

Traditionally, backward error analysis for GE is expressed in terms of the *growth factor*

$$g(n, A) = \frac{\max_{i,j,k} |a_{ij}^{(k)}|}{\max_{i,j} |a_{ij}|},$$

which involves all the elements $a_{ij}^{(k)}$, $k = 1, 2, \ldots, n$ that occur during the elimination. The growth factor actually measures how large the entries become during the process of elimination. The following classic theorem illustrates the accuracy of the computed solution.

Theorem 6 (Wilkinson [26, p. 165]). *Let $A \in \mathbb{R}^{n\times n}$ and suppose GE with partial pivoting produces a computed solution $\hat{x}$ to $A \cdot x = b$. Then there exists a matrix ΔA and a constant c_{3n} such that*

$$(A + \Delta A)\hat{x} = b, \quad \|\Delta A\|_\infty \leq c_{3n} n^2 g(n, A) \|A\|_\infty.$$

It is clear that the stability of GE depends on the growth factor. If $g(n, A)$ is of order 1, not much growth has taken place, and the elimination process is stable. If $g(n, A)$ is bigger than this, we must expect instability. If GE can be unstable, why is it so famous and so popular? The answer seems to be that although some matrices cause instability, these represent such an extraordinarily small proportion of the set of all matrices that they "never" arise in practice simply for statistical reasons. This explanation gives rise to a statistical approach to the growth factor and motivates the study of its behavior for random matrices [9, 58]. In practice the growth factor is usually of order 10 and therefore most numerical analysts regard the occurrence of serious element growth in GE with partial pivoting as highly unlikely in practice. So, the method can be used with confidence and constitutes a stable algorithm in practice [56].

The determination of $g(n, A)$ in general remains a mystery. Regarding the possible magnitude of the growth factor for partial pivoting, it is easy to show that $g(n, A) \leq 2^{n-1}$. Wilkinson in [61, p. 289] and [62, p. 212] has reported of special form matrices attaining this upper bound. Higham and Higham characterize all matrices attaining this upper bound in the following theorem.

Theorem 7 ([28]). *All real $n \times n$ matrices A, for which $g(n, A) = 2^{n-1}$ for partial pivoting, are of the form*

$$A = DM \begin{bmatrix} T & \vdots & \theta d \\ 0 & \vdots & \end{bmatrix},$$

Table 3 Values of $f(n)$ for Wilkinson's bound

n	10	20	50	100	200	1000
$f(n)$	19	67	530	3300	26,000	7,900,000

where $D = diag(\pm 1)$*,* M *is unit lower triangular with* $m_{ij} = -1$ *for* $i > j$*,* T *is an arbitrary nonsingular upper triangular matrix of order* $n-1$*,* $d = (1,2,4,\ldots,2^{n-1})^T$*, and* θ *is a scalar such that* $\theta := |a_{1n}| = \max_{i,j}|a_{ij}|$.

Although the growth factor can be as large as 2^{n-1} for an $n \times n$ matrix, the occurrence of a growth factor even as large as n is rare. However, later than the early 1990s some applications with large growth factors have been published [14, 16, 63]. Efforts have been made to explain this phenomenon with some success [28, 57, 58], yet the matter is far from completely understood.

For complete pivoting, Wilkinson has showed in [61, pp. 282–285] that

$$g(n,A) \leq [n\,2\,3^{1/2}\ldots n^{1/(n-1)}]^{1/2} \equiv f(n) \sim cn^{1/2}n^{\frac{1}{4}\log n}$$

and that this bound is not attainable. The bound is a much more slowly growing function than 2^{n-1}, but it can still be quite large, cf. Table 3.

Matrices with the property that no exchanges are actually needed during GE with complete pivoting are called *completely pivoted (CP)* or *feasible*. Equivalently, a matrix is CP if the rows and columns have been permuted so that GE without pivoting satisfies the requirements for complete pivoting, hence the maximum elements on each elimination step appear on the diagonal position. For a CP matrix A we have

$$g(n,A) = \frac{\max\{p_1,p_2,\ldots,p_n\}}{|a_{11}|}, \tag{8}$$

where $p_1,p_2,\ldots,p_n$ are the pivots of A. The study of the values appearing for $g(n,A)$ and the specification of pivot patterns are referred to as *the growth problem*.

Cryer [6] defined

$$g(n) = \sup\{\, g(n,A) \mid A \in \mathbb{R}^{n\times n}, \mathrm{CP}\}.$$

The function $g(n)$ plays a role in the analysis of roundoff errors in GE ([15, p. 103] and [62, p. 213]). Wilkinson in [62, p. 213] noted that there were no known examples of matrices for which $g(n,A) > n$. The problem of determining $g(n)$ for various values of n is called the *growth problem*. The following results are known:

- $g(2) = 2$ (trivial)
- $g(3) = 2\frac{1}{4}$ [5, 6, 8, 54]
- $g(4) = 4$ [5, 6, 54]
- $g(5) < 5.005$ [5]

History of the Problem

In 1964 Tornheim [54] (see also [6]) proved that $g(n,H) \geq n$ for a CP $n \times n$ Hadamard matrix H. In 1968 Cryer [6] formulated the following confecture.

Conjecture 4.

The Complete Pivoting Conjecture for Gaussian Elimination
$g(n,A) \leq n$, with equality iff A is a Hadamard matrix".

20 years later, in 1988 Day and Peterson [8] provided useful insight into the pivot structures of Hadamard matrices belonging to the Sylvester class [29], the matrices of which are constructed according to the scheme

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}.$$

Actually, they have shown the following result.

Proposition 2 ([8]). *Suppose A is $n \times n$ and CP and let $\otimes^k A$ denote the Kronecker product of k copies of A. Then $A \otimes H_2$ is CP and its pivots are the Kronecker product of the pivots of A with those of H_2, that is, the entries of*

$$[a_{11}^{(0)} \; \ldots \; a_{nn}^{(n-1)}] \otimes [1\ 2].$$

Therefore $A \otimes H_n$ is CP for all n and the pivots are the Kronecker products of the pivots of A with $\otimes^n [1\ 2] = [1\ 2\ 2\ 4\ 2\ 4\ 4\ 8\ \ldots]$. If A is replaced by H_2, we obtain the pivot structure of Hadamard matrices of orders $2n$ that belong to the Sylvester equivalence class, and the next corollary holds.

Corollary 1. *For any Hadamard matrix H of order n with growth factor $g(n,H)$ and maximum pivot p_n, there exists a Hadamard matrix of order $2^t n$ with growth factor $2^t g(n,A)$ and maximum pivot $2^t p_n$.*

Corollary 1 shows in fact Cryer's conjecture for Hadamard matrices belonging in the Sylvester equivalence class. Day and Peterson provided also some experimental results for pivot patterns of 16×16 Hadamard matrices and conjectured that the fourth pivot from the end is $n/4$, which was shown to be false in 1998 by Edelman and Friedman [11], who found the first H_{16} with fourth pivot from the end $n/2$. In 1991 Gould reported on matrices that exhibited, in presence of roundoff error, growth larger than n [10, 20], cf. Table 4. These matrices were created by simulating the process of GE as an appropriate optimization problem. Thus the first part of Cryer's conjecture was shown to be false. The second part of the conjecture concerning the growth factor of Hadamard matrices still remains an open problem.

Over all these years small progress has been made towards the Hadamard part of Cryer's conjecture. In 1995 Edelman and Mascarenhas [12] proved that the growth factor of H_{12} is 12 by demonstrating its unique pivot pattern. In this work the

Table 4 Gould's results on growth factors greater than n

n	Growth
13	13.02
18	20.45
20	24.25
25	32.99

difficulties of such problems in general have been stated for first time, and also more specifically:

"it appears very difficult to prove that $g(16, H_{16}) = 16$".

It is known that if GE with complete pivoting is applied to Hadamard matrices of order 16, over 30 different pivot patterns are attained, by contrast with Hadamard matrices of orders less than or equal to 12, which yield only one pivot pattern [12]. Hadamard matrices of order 16 can be classified with respect to the H-equivalence in five classes of equivalence I,. . . ,V, see [23, 59]. Classes IV and V are one another's transpose, and so they are identical for GE with complete pivoting, since a matrix is CP iff its transpose is CP, in which case both give the same pivot pattern.

Extensive experiments revealed 34 possible pivot patterns for Hadamard matrices of order 16, though not all patterns appeared for each equivalence class. A great difficulty arises at the study of this problem because H-equivalence operations do not preserve pivots, i.e. H-equivalent matrices do not have necessarily the same pivot pattern and the pivot pattern is not invariant under H-equivalence. Therefore it is not sufficient to specify the pivot pattern of a representative matrix of an equivalence class, but there must be available information from all matrices. This property is an obstacle for studying theoretically the pivots of H_{16}, which are classified in five equivalence classes [59] and had given until that time more than 30 pivot sequences. So, for the case of proving the pivot structures of H_{16}, a naive computer exhaustive search creating all possible H_{16} by performing all possible H-equivalence operations would require $(16!)^2(2^{16})^2 \approx 10^{36}$ trials. Taking into account that GE should be applied additionally to these matrices, then it becomes clear that the total amount of trials and computations cannot be carried out within a realistic, sensible time framework, even by very powerful computers. In 2007 [40] all 34 possible pivot patterns of H_{16} were demonstrated theoretically and the complete pivoting conjecture for H_{16} was proved.

The Importance of Determinant Calculations

In the following we will describe the principal technique used in the majority of the works for calculating pivots of Hadamard matrices, which is based on determinant evaluations. The idea takes advantage of the following lemma, which constitutes a powerful tool for this research, since it offers a possibility for calculating pivots in connection with minors.

Lemma 4 ([6, 8, 12, 17, 42]). *Let A be a CP matrix.*

(i) *The magnitude of the pivots appearing after application of GE operations on A is given by*

$$p_j = \frac{A(j)}{A(j-1)}, \quad j = 1, 2, \ldots, n, \quad A(0) = 1. \tag{9}$$

(ii) *The maximum $j \times j$ leading principal minor of A, when the first $j-1$ rows and columns are fixed, is $A(j)$.*

So, it is obvious that the calculation of minors is important in order to study pivot structures, and moreover the growth problem for CP Hadamard matrices. Since Cryer's conjecture is connected with CP Hadamard matrices, Eq. (9) can be exploited for calculating pivots for them but it is important to emphasize that the results are valid for every H_{16} if GE with complete pivoting is applied. Indeed, if a matrix isn't initially CP, the row and column operations of GE with complete pivoting bring it always in CP form, hence we can deal without loss of generality only with CP Hadamard matrices.

The second part of Lemma 4 assures that the maximum $j \times j$ minor appears in the upper left $j \times j$ corner of A. So, if the existence of a matrix with maximal determinant is proved for a CP Hadamard matrix, we can indeed assume that it always appears in the upper left corner.

Computation of $(n-j) \times (n-j)$ Minors

Pivots from the end of the pivot structure of Hadamard matrices are calculated in the bibliography with relation (9). In order to take advantage of (9) for this purpose, one has to calculate minors of orders $(n-j) \times (n-j)$, $j = 1, 2, \ldots$ for Hadamard matrices of orders n.

Day and Peterson's Results

After many years of unremarkable progress towards the growth factor of Hadamard matrices, Day and Peterson [8] presented an interesting variety of results regarding the growth factor generally. The portion relevant to the growth problem for Hadamard matrices is summarized in the next proposition.

Proposition 3 ([8]). *Let H be an $n \times n$ CP Hadamard matrix and reduce H by GE. Let $D = n^{n/2}$. Then the possible values for $H(k)$ and $|h_{kk}^{(k-1)}|$ when $k \geq n-6$ are*

k	$H(k)$	$\lvert h_{kk}^{(k-1)}\rvert$
n	D	n
$n-1$	$(1/n)D$	$n/2$
$n-2$	$(2/n^2)D$	$n/2$
$n-3$	$(4/n^3)D$	$n/2\ \ n/4$
$n-4$	$(q/n^4)D,\ q=8,16$	$np,\ p=1,1/2,1/4,1/3,1/6$
$n-5$	$(q/n^5)D,\ q=16,32,48$	$np,\ p=3/2,1,3/4,1/2,3/8,1/3,$
		$3/10,1/4,1/5,1/8,1/10$
$n-6$	$(q/n^6)D,\ q=32,64,96,128,160$	

The values of $H(k)$ in Proposition 3 are proved by exploiting the next proposition, which establishes an explicit connection between $k\times k$ upper left and $(n-k)\times(n-k)$ lower right principal minors of Hadamard matrices.

Proposition 4 ([8]). *If H is a Hadamard matrix of order n, then*

$$n^{n/2}H(k) = n^k H[n-k]. \tag{10}$$

The values of $|h_{kk}^{(k-1)}|$ in Proposition 3 are proved by substituting appropriately the values of $H(k)$ in (9).

The results presented in section "Minors of Hadamard Matrices" concerning the evaluation of minors of orders $(n-j)\times(n-j), j=1,2,\ldots$ form a recent progress to this issue which can be adopted for the computation of pivots from the end.

Computation of $j\times j$ Minors

In this section we will illustrate briefly the two principal techniques that have been used in the works [12] and [40] for calculating $j\times j$ minors of CP Hadamard matrices.These calculations can lead us to the specification of pivots from the beginning. The results $H(1)=1$, $H(2)=2$ and $H(3)=4$ for a CP Hadamard matrix are trivial.

Lemma 5 ([8]). *If H is a CP Hadamard matrix, then* $H(4)=16$*, i.e. the* 4×4 *principal submatrix of H is a Hadamard matrix of order 4.*

The Technique Used for H_{12}

For estimating $H(5)$ for a CP H_{12}, the authors of [12] utilize the interpretation of a Hadamard matrix as a symmetric block design (cf. section "The Pivot Pattern of H_{12}") in appropriate combination with the property of Lemma 4(ii).

Lemma 6 ([12]). *If H is a CP Hadamard matrix of order 12, then* $H(5)=48$.

The Technique Used for H_{16}

The main idea presented in [40] for deriving values of $j \times j$ minors is to specify all possible matrices that can always exist embedded inside a CP Hadamard matrix. Hence, if a matrix is proved to exist as a possible submatrix of a CP Hadamard matrix, then one can claim that its determinant is a possible minor for the Hadamard matrix. The first question towards applying this idea is, which matrices should be examined. An obvious answer would be to construct all possible $j \times j$ matrices with entries ± 1 and first row and column with $+1$'s and to use them as input for an appropriately designed algorithm. This idea would lead to a forbidding from a computational point of view strategy, because, e.g., for $j = 7$ one should construct $2^{36} \approx 10^{10}$ matrices. Furthermore, any information obtained from the order n of the Hadamard matrix wouldn't be taken into account under this perspective and it would be also hard to apply the CP property as a selection criterion.

The proposed remedy is a stepwise extension procedure. It starts with the 2×2 matrix

$$\begin{bmatrix} + & + \\ + & - \end{bmatrix}$$

(which is actually H_2). An important assumption is that H_2 exists embedded inside any H_n. This follows straightforwardly from the orthogonality of the first two rows, which implies the existence of some entry -1 in the second row, and so this entry can be brought by means of a column permutation at position $(2, 2)$. In the sequel, the following procedure was adopted for producing and classifying the matrices of higher orders.

Extension Procedure

For $j = 3, 4, \ldots$

1. Extension of the $(j-1) \times (j-1)$ existing matrices to all possible normalized ± 1 $j \times j$ matrices;
2. From all possible extensions only the CP ones are kept;
3. The CP extensions are separated into classes according to the appearing values of determinants;
4. The set $\mathcal{M}_j$ with the matrices that can always exist among the first j rows of H_{16} is kept. $\mathcal{M}_j$ is created by examining the extensions with the maximum determinant. If they always exist, their set is denoted as $\mathcal{M}_j$, otherwise we proceed with the second maximum appearing determinant value. If the matrices with the two largest determinant values always exist, their set is denoted as $\mathcal{M}_j$, otherwise we proceed with the third maximum appearing value, and so on.

The second crucial question is, how to examine whether a matrix exists embedded inside a CP H_n. The answer is partially given by the previously described algorithm EXIST, which verifies the occurrence or not of a matrix inside H_n but doesn't consider at all the CP property of H_n. This is actually done with the help of Lemma 4 (ii), which guarantees that the minor with maximum magnitude appears in the upper left corner of a CP matrix. So, if a matrix is proved to exist inside H_n and at the same time this matrix attains the maximal value of determinant with respect to the range of determinant values of all possible existing matrices of the same order, then it can be deduced that this matrix can always appear (with H-equivalence operations, if necessary) in the upper left corner of a CP Hadamard matrix. This strategy is incorporated as Step 4 in the above described extension procedure.

For instance, according to this procedure, the matrix

$$A \equiv H_4 = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix},$$

which is proved to exist in every H_n for $n > 4$, is extended to all possible 5×5 matrices of the form

$$\begin{bmatrix} + & + & + & + & + \\ + & - & + & - & * \\ + & + & - & - & * \\ + & - & - & + & * \\ + & * & * & * & * \end{bmatrix},$$

where the elements $*$ can be ± 1. From these $2^7 = 128$ possible 5×5 matrices only the 52 CP matrices with determinants 48 and 32 forming $\mathscr{M}_5$, and tested with algorithm EXIST can always exist among all possible columns that can appear in the first five rows, and furthermore in the upper left corner of a CP H_{16}. Since the 52 matrices $\mathscr{M}_5$ with determinants 48 and 32 always exist among the rows of H_{16}, from Lemma 4(ii) we derive that

$$H_{16}(5) = 32 \text{ or } 48.$$

Following a similar manner the following results have been also derived.

- $H_{16}(6) = 128$ or 160;
- $H_{16}(7) = 256$ or 384 or 512 or 576;
- $H_{16}(8) = 1024$ or 1536 or 2048 or 2304 or 2560 or 3072 or 4096.

Computation of Pivot Patterns

The first four pivots of a CP Hadamard matrix were determined by Day and Peterson [8] as given in the next result.

Proposition 5 ([8]). *If H_n is a CP Hadamard matrix, $n \geq 4$, then the first four pivots are 1, 2, 2, and 4, respectively. For $n > 4$, the fifth pivot is 2 or 3.*

When performing GE with complete pivoting on a Hadamard matrix, the final pivot always has magnitude n [6, Theorem 2.4]. This result was also verified by the technique introduced in [34] and extended as follows.

Proposition 6 ([34]). *If H_n is a CP Hadamard matrix, $n \geq 4$, then the last four pivots are (in backward order) $n, \frac{n}{2}, \frac{n}{2}$ and $\frac{n}{4}$ or $\frac{n}{2}$.*

After having derived the results of section "A New Technique," this proposition was easily proved by substituting appropriately the values of minors in the relative quotients obtained by Eq. (9). This result verifies theoretically the occurrence of the two possible values as fourth pivot from the end.

The Pivot Pattern of H_{12}

Husain [31] showed that there is only one Hadamard matrix of order 12 up to H-equivalence. Edelman and Mascarenhas [12] take the first step towards proving the Hadamard part of Cryer's conjecture by settling a conjecture by Day and Peterson [8, p. 508] that only one set of pivot magnitudes is possible for a Hadamard matrix of order 12. Their proof is carried out in a very elegant and compact manner by utilizing the notion of symmetric block designs to reduce the complexity that would be encountered in enumerating cases. Lemma 5 is taken also into account. So, they take for granted that H_4 appears in the upper left corner of a H_{12} and show that some 5×5 submatrix of H_{12} with determinant 48 includes H_4. Hence, one can conclude that $H_{12}(5) = 48$. Indeed, if there is a submatrix M_5 with determinant 48, then Lemma 4(ii) in combination with Table 2 tell that $48 = |\det M_5| \leq H_{12}(5) \leq 48$, implying $H_{12}(5) = 48$.

A *symmetric block design* with parameters v, k, and λ is a collection of v objects and v blocks such that

- every block contains k objects;
- every object appears in k blocks;
- every pair of blocks has λ objects in common;
- every pair of objects can be found in λ blocks.

A Hadamard matrix of order $n = 4t$ is equivalent to a symmetric block design with parameters $v = 4t-1, k = 2t-1$ and $\lambda = t-1$ [24]. Hence, H_{12} can be interpreted as an arrangement of 11 objects into 11 blocks containing five objects such that each object appears in exactly five blocks, every pair of distinct objects appears together exactly twice and every pair of distinct blocks has exactly two elements in common.

The fifth pivot of a CP H_{12} is calculated by substituting the results of Lemmas 5 and 6 in (9) as

$$p_5 = \frac{H_{12}(5)}{H_{12}(4)} = 3.$$

For deriving the rest of the pivots of H_{12}, the authors of [12] take advantage of Proposition 4 and of the following corollary.

Corollary 2 ([12, Corollary 2.1]). *If H is a Hadamard matrix of order n, then the kth pivot from the end is*

$$p_{n+1-k} = \frac{nH[k-1]}{H[k]}.$$

Substituting $H_{12}(5) = 48$ in (10) gives $H[7] = 576$. The remaining values of minors are obtained by taking into account that each $n \times n$ ± 1 matrix attaining the maximum determinant contains an $(n-1) \times (n-1)$ submatrix with maximum determinant. So, one can conclude $H[5] = 48$, $H[4] = 16$, $H[3] = 4$, $H[2] = 2$ and $H[1] = 1$. The last seven pivots follow from Corollary 2. Hence, the unique pivot pattern of a CP H_{12} is

$$[1,\ 2,\ 2,\ 4,\ 3,\ 10/3,\ 18/5,\ 4,\ 3,\ 6,\ 6,\ 12].$$

The Pivot Patterns of H_{16}

The main idea for trying initially to reduce the forbidding cost for finding the pivots of H_{16} described in section "History of the Problem" was to split this task into two independent subproblems: calculation of pivots from the beginning and from the end of the pivot pattern. So, the results of sections "The Technique Used for H_{16}" and "The Latest Algorithm" can be utilized separately. A benefit from the discrimination of the two tasks is that one deals in this manner only with relatively small matrices not exceeding the order 8. On the contrary, if one would like to use only the technique of section "The Latest Algorithm" or "The Technique Used for H_{16}" for the computation of the whole pivot pattern, then it would be mandatory to deal with matrices of order up to 12, what would complicate furthermore the computations.

Hence, the pivots p_5, p_6, p_7, and p_8 were calculated by taking advantage of the strategy of section "The Technique Used for H_{16}" and the pivots p_{10}, p_{11}, and p_{12} with help of the results of section "The Latest Algorithm." In both cases, the appropriate values of minors were substituted *carefully* in formula (9) for deriving

Table 5 Pivots of H_{16}

Pivot	Values
1	1
2	2
3	2
4	4
5	2,3
6	$4, \frac{8}{3}, \frac{10}{3}$
7	$2, 4, \frac{8}{10/3}, \frac{16}{5}, \frac{18}{5}$
8	$4, \frac{9}{2}, 5, 6, 8$
9	$2, \frac{8}{3}, \frac{16}{5}, 4, \frac{9}{2}, \frac{16}{3}$
10	$4, \frac{16}{8/5}, 5, 8$
11	$4, \frac{16}{10/3}, 6, 8$
12	$8, \frac{16}{3}$
13	4,8
14	8
15	8
16	16

magnitudes of pivots. By saying carefully it is meant that one should always take the origin of a minor into account, i.e. to check whether a minor of order j is really obtained from a minor of order $j-1$ and not substitute blindly all values of minors of orders j and $j-1$ in (9). For example, all possible minor quotients for forming p_{12} would be $p_{12} = \frac{H_{16}(12)}{H_{16}(11)} =$

$$\frac{16 \cdot 16^4}{32 \cdot 16^3} \text{ or } \frac{16 \cdot 16^4}{48 \cdot 16^3} \text{ or } \frac{8 \cdot 16^4}{16 \cdot 16^3} \text{ or } \frac{16 \cdot 16^4}{16 \cdot 16^3} \text{ or } \frac{8 \cdot 16^4}{32 \cdot 16^3} \text{ or } \frac{8 \cdot 16^4}{48 \cdot 16^3},$$

by taking into account the values $n-4$ and $n-5$ from Table 2. The three last quotients are rejected as possible values for p_{12} by using algorithm EXIST appropriately and determining that the respective matrices cannot exist always in a H_{16}.

The ninth pivot is calculated afterwards separately by using the property that the product of the pivots equals the determinant of the matrix. The resulting pivot values for H_{16} are summarized in Table 5. All 34 possible different pivot patterns of H_{16} are given in [40] and the classification of the pivot patterns in the five equivalence classes is done *experimentally only* in [33].

Further Progress and a Good Pivots Property

Here we present the latest results achieved on this topic. Additionally, we highlight the significance of the newly introduction of a new notion of good pivots for the analysis of the problem and study its properties, contribution, and a further outlook.

State-of-the-Art Results

Moreover, further progress on the complete pivoting conjecture for Hadamard matrices and a new good pivots property are established in [41]. A fundamental result of this study is presented in Proposition 7.

Proposition 7 ([41]). *For every pivot p_k, $k = 2, \ldots, n$, of a CP Hadamard matrix H of order n it holds $p_k > 1$. Furthermore, the leading principal minors form an increasing sequence, namely $H(k) > H(k-1)$, $k = 1, \ldots, n$.*

The result is proved by taking advantage of the special block form of the submatrices of H after the first k elimination steps, Hadamard's inequality successively for them and relation (9). The outcomes of Proposition 7 can be utilized appropriately for demonstrating explicitly some possible values of their theoretically admissible minors, resp., pivots, leading to the formulation of Corollary 3.

Corollary 3. *Let H be a CP Hadamard matrix of order n. The possible values of the leading principal minors of orders 5, 6, 7, and 8 of H are given in Table 6.*

Applying Eq. (9) one can proceed to the specification of the *theoretically admissible* minors and pivot values according to Proposition 7. It is not explicitly proved that all of them appear actually. By means of the aforementioned theoretical tools and argumentation, one can compute explicitly the possible values that the fifth pivot of a CP Hadamard matrix can attain.

Corollary 4. *The fifth pivot of a CP Hadamard matrix H of order $n \geq 8$ is* 2 *or* 3*, after application of GE on H.*

Table 6 All possible values $H(5)$, $H(6)$, $H(7)$, and $H(8)$ for any CP Hadamard matrix H

$H(5)$	32, 48
$H(6)$	64, 96, 128, 160
$H(7)$	128, 192, 256, 320, 384, 448, 512, 576
$H(8)$	256, 384, 512, 640, 768, 896, 1024, 1152, 1280, 1408,
	1536, 1664, 1792, 1920, 2048, 2176, 2304, 2560, 3072, 4096

The following Lemma certifies that some specific values cannot be attained by the sixth and seventh pivots. As a consequence they can be excluded from the theoretical considerations.

Lemma 7. *Let H be a CP Hadamard matrix H of order $n \geq 8$. After application of GE with complete pivoting on H, the sixth pivot p_6 and the seventh pivot p_7 cannot take the value 5 and 9, respectively.*

An essential ingredient of the novel strategy for the computation of growth factors for certain matrices is not necessarily the explicit specification of these matrices, as it was done so far. It is sufficient to derive inequalities for their pivots and to utilize them for calculating finally the growth factor simply with the aid of its definition. For instance, we illustrate here the result concerning the possible values of the sixth and seventh pivots appearing after application of GE on a CP Hadamard matrix.

Corollary 5. *The following bounds hold for the sixth and seventh pivots, resp., after application of GE on a CP Hadamard matrix H*

$$4/3 \leq p_6 \leq 4 \tag{11}$$

$$6/5 \leq p_7 \leq 8 \tag{12}$$

Moreover, careful substitution of minor values in (9) and taking into account Proposition 7, Corollary 4 and the possible values of ± 1 determinants of order 6 and 7 presented in [8], one obtains the possible values for the sixth and seventh pivots given in Table 7.

With the previously illustrated theoretical tools one can directly compute the growth factor of a Hadamard matrix of order 12, offering a new proof for the issue, which is simpler and more practical than the ones presented in [12] and [42]. These techniques computed explicitly the pivot pattern of a Hadamard matrix of order 12 taking advantage of several strategies, in contrast to the latest approach with the aforementioned theoretical results that doesn't require analytical computation of pivot patterns. The new proof is brief, demonstrates that all pivots are less or equal to 12, and deploys also the definition (8) of the growth factor for CP matrices.

Corollary 6. *The growth factor of a CP Hadamard matrix of order 12 is 12, when GE is applied.*

Table 7 All possible values of the sixth and seventh, resp., sixth last and seventh last pivots for any CP Hadamard matrix

p_6	4/3, 2, 8/3, 3, 10/3, 4
p_7	6/5, 4/3, 8/5, 2, 12/5, 5/2, 8/3, 14/5, 3,
	16/5, 10/3, 7/2, 18/5, 4, 9/2, 14/3, 5, 16/3, 6, 7, 8

Good Pivots

We present a new notion of *good pivot patterns* (or simply *good pivots*) that was introduced in [41], in addition, e.g., to the one elaborated in [44].

Definition 3. A pivot pattern $[p_1, p_2, \dots, p_n]$ appearing after application of GE on a matrix of order n is called *good*, if its pivots satisfy

$$p_i p_{n-i+1} = n, \quad i = 1, \dots, n.$$

According to this definition, good pivot patterns have evidently the form

$$\left[p_1, p_2, \dots, p_{\frac{n}{2}}, \frac{n}{p_{\frac{n}{2}}}, \dots, \frac{n}{p_2}, \frac{n}{p_1} \right],$$

i.e. their pivots appear with a specific sense of symmetry in the pattern. Hadamard matrices are the only matrices known so far that lead to good pivot patterns. An interesting observation is that numerical experiments suggest the rare appearance of Hadamard matrices with this good pivots property. It is also difficult to construct Hadamard matrices artificially so that they acquire this feature. It is meaningful to study good pivot patterns obtained after GE with complete pivoting, since this particular elimination strategy is important in view of Cryer's growth conjecture.

The following Proposition 8 constitutes a criterion for the specification of matrices possessing good pivot patterns.

Proposition 8 (The Equality of Quotients of Minors Property [41]). *Let H be a CP Hadamard matrix of order n with good pivot pattern. Then*

$$\frac{H(k)}{H(k-1)} = \frac{H[k]}{H[k-1]}, \tag{13}$$

where $H[k]$ denotes the magnitude of the lower right $k \times k$ principal minor of H.

Utilizing the previous results in this section one can demonstrate the following fundamental Theorem 8. It reveals the usefulness and significance of the appearance of good pivot patterns with respect to the computation of the growth factor of a CP Hadamard matrix.

Theorem 8 ([41]). *If the pivots $[p_1, p_2, \dots, p_n]$ of a CP Hadamard matrix H of order n are good, then $g(n, H) = n$.*

In the sequel we present the latest results concerning pivot patterns of any CP Hadamard matrix of order up to 20. They lead to posing the good pivots conjecture. It concerns the specification of a Hadamard matrix in each equivalence class satisfying the equality of quotients of minors criterion, which lead to growth equal to their dimension.

Table 8 On the application of GE with complete pivoting for Hadamard matrices of orders n up to 20 (order, number of H-equivalence classes, number of pivot patterns, number of good pivot patterns)

n	H-equivalence classes	Pivot patterns	Good pivots patterns
8	1	1	1
12	1	1	1
16	5	34	11
20	3	≥ 1128	≥ 47

Results concerning the application of GE with complete pivoting to Hadamard matrices of order up to 20 are presented in Table 8. The results for orders 8, 12, and 16 are presented in [8, 12, 40], respectively. The results for $n = 20$ correspond to extensive numerical experiments on a sufficiently large number of matrices carried out for the needs of [41].

It is interesting to highlight that the 11 good pivot sequences appearing for H_{16} are split in all five H-equivalence classes existing when $n = 16$. Indeed, it is straightforward to infer the exact distribution in the classes in view of all resulting possible pivot patterns for H_{16} in [40]. Analogous experimentation can be performed numerically for H_{20}.

Hence, we notice that for Hadamard matrices of orders up to 20 there exists in each equivalence class a representative Hadamard matrix possessing the good pivots property. This m matrix fulfills the equality of quotients of minors criterion. With the aid of this specific result, one can detect the representative Hadamard matrix of a H-equivalence class with the good pivot pattern.

All the above results and remarks lead to stating the following conjecture.

Conjecture 5.

The good pivots conjecture

Every CP Hadamard matrix H of order n can be transformed with H-equivalence operations to a form $\widetilde{H}$, satisfying the equality of quotients of minors property. The growth of $\widetilde{H}$ is equal to n.

The good pivots conjecture is confirmed for orders up to 20. For instance, the H_{20} has operated as follows. A computer search algorithm generates H-equivalent H_{20} matrices from the three possible equivalence classes. It has shown that some pivot patterns, which appear for H_{20} belonging to the three equivalence classes, are good. We notice that Hadamard matrices of order 20 belong to three H-equivalence classes, cf., e.g., [46, 50]. For example, the following pivot pattern appears for Hadamard matrices of order 20 from the three H-equivalence classes.

$$[1, 2, 2, 4, 3, 10/3, 18/5, 4, 40/9, 24/5, 20/(24/5), 9/2, 5, 20/(18/5),$$
$$6, 20/3, 5, 10, 10, 20]$$

Thus, one can conclude that there can be found always a H_{20} being H-equivalent to a H_{20} with good pivots. Concluding, one can find three nonequivalent Hadamard matrices of order 20 in [46, 50]. The matrix in [46] from the first equivalence class gives good pivots.

Infinite Classes of Hadamard Matrices with Good Pivots

Here we specify a possible construction of an infinite family of Hadamard matrices with good pivots by means of the ordinary Kronecker product, cf, e.g., [43, p. 598]. The class can be generated starting from a CP Hadamard matrix with good pivots and applying the following Theorem 9 consecutively.

Theorem 9 ([41]). *Suppose the pivots of a CP matrix A_n of order n are good. Then the pivots of $A_n \otimes H_2$ are also good.*

The proof takes also advantage of the fact that if the pivots $[p_1, p_2, \ldots, p_n]$ of a CP matrix A_n of order n are good, then the pivots of $A_n \otimes H_2$ are known [8, Proposition 5.12] to be

$$[p_1, p_2, \ldots, p_n] \otimes [1,2] = [p_1 \otimes [1,2], p_2 \otimes [1,2], \ldots, p_n \otimes [1,2]].$$

Calculation of the last expression discloses that the pivots of $A_n \otimes H_2$ are also good, with the respective pivots product constant being $2n$.

Relevant Open Problems

Through the description of the various techniques in sections "A New Technique," "A Further Development," and "The Latest Algorithm" it becomes evident that the huge amount of matrices M, which can exist in the upper left $j \times j$ corner of a Hadamard matrix according to scheme (1), plays a crucial role in making algorithms MINORS1, 2, and 3 impossible to implement in practice within a realistic period of time, especially for $j \geq 7$. The criteria developed so far for reducing the total amount of these matrices, based either on more intuitive ideas ("IP equivalent matrices," "Hadamard submatrix") or on more formal aspects (algorithm EXIST), didn't succeed in a significant reduction of the total amount of matrices to the "really useful" matrices M. By saying "really useful" it is meant the minimum necessary amount of matrices M, which would give all possible different values of the $n-j$ minors. Hence there emerges the need to establish a theoretically correct and practically feasible criterion for eliminating the unnecessary for the algorithms matrices M, which at the same time doesn't skip any appearing values of minors. The same framework can be regarded also for reducing the amount of matrices in the sets $\mathscr{M}_j$, but in this case one should explore the mechanisms of Gaussian Elimination in order to find out which $j \times j$ matrices yield ultimately same pivot patterns, and therefore are useless with respect to this study.

The study of the conjecture concerning $n-j$ minors of Hadamard matrices might furthermore lead to useful results concerning the possible values of determinants of ± 1 matrices, which are not exactly specified even for relatively small orders ($n = 8$). The relevant known results are given in Table 1. Comparing Conjectures 1 and 3 (or, equivalently Tables 1 and 2), one sees that there seems to be a connection between the possible determinant values of $j \times j$ ± 1 matrices and the magnitudes of minors of Hadamard matrices of respective orders $n-j$. Namely, the former values appear as coefficients for the latter values before an n- and j-term. This observation gives rise to the intuition that these two problems could be somehow connected. So, any progress for the $n-j$ minors of Hadamard matrices for $j > 7$ could yield useful results concerning the possible determinants of ± 1 $j \times j$ matrices for $j > 7$, and vice versa. Additionally, if one manages to implement a version of algorithm MINORS (or another technique with the same purpose) up to the $n-15$ case, then it would be possible to obtain an estimation for the maximum determinant of a 15×15 matrix with entries ± 1, which is an unsolved problem so far. Related problems about determinants of $(0, \pm 1)$ matrices can be found in [3, 13, 22, 48].

Another open problem is to classify theoretically the appearing pivot patterns of H_{16} or of other Hadamard matrices of higher orders in the appropriate equivalence classes. It is known [59, p. 421] that H_{16} matrices can be classified with respect to the H-equivalence in five classes I,. . . ,V. This research would require more information on the properties of the classes, see [23, 59]. For example, the following reasoning could be a motivation for such an effort. We observe [33] that the value $10/3$ for the sixth pivot does not appear in the I-Class of equivalence. Looking closely at the possible relative quotients we see that this value corresponds to $H_{16}(6) = 160$. Hence, in order to prove the nonoccurrence of $p_6 = 10/3$ in I-Class, it is equivalent to show that a 6×6 submatrix with determinant 160 cannot exist inside a matrix of this class. The 6×6 ± 1 matrix attaining the maximum determinant value 160 is called *the D-optimal design of order 6* (D_6). Or, whenever $p_5 = 3$, equivalently this means that the D-optimal design of order 5 exists embedded in these H_{16}.

The methods presented in this work can be used as the basis for calculating the pivot pattern of Hadamard matrices of higher orders, such as H_{20}, H_{24} etc. The complexity of such problems points out the need for developing algorithms that can implement very effectively the ideas introduced in this work, or other, more elaborate ideas. For instance, a question towards this direction is, is there a reliable criterion for reducing the total amount of matrices M used as input for algorithm Minors, so that no appearing values are skipped? The reduction of the amount of matrices occurring with the extension procedure of section "The Technique Used for H_{16}" is a matter of concern, too. Another observation, which could lead to a computational improvement, is the fact that all acceptable solutions of the linear systems of algorithm EXIST, $k \geq 5$, are always obtained for the parameter values $0, \ldots, n/8$. Hence, there arises the obvious question whether there is a more precise upper bound for the possible columns in the first rows of a Hadamard matrix than the one given in Lemma 1. Furthermore, a parallel implementation, which would limit significantly the computational time needed by the algorithms, could be a matter under investigation.

It is also interesting to mention that the values of the pivots depend on the choice of the maximum element at each elimination step, when there are at least two equal maximum entries in the respective submatrix. More pivoting strategies leading to several choices of the maximum entry are described in [14, 26]. For example, it is known that the D-optimal design of order 5 can exist embedded inside a H_{16}, leading to the value of the fifth pivot $p_5 = 3$. We attempt to find the pivot pattern of a H_{16}, which is equivalent to the one constructed with the command hadamard(16) in Matlab. If we use a straightforward selection of the maximum element, i.e. to choose the *first* maximum entry of every respective lower right matrix as pivot, we get $p_5 = 2$. But if we select as pivot the *last* maximum entry we might be led to $p_5 = 3$. Hence, it is challenging to investigate furthermore this phenomenon and to determine the choices of maximum entries that lead to specific pivot values.

Finally, it is a mystery why the value 8 as fourth pivot from the end appears only for matrices from the Hadamard–Sylvester class (called also I-Class of equivalence), and specifically only in one pivot pattern. It is interesting to study this issue and to find out whether the construction properties of this class can explain this exceptional case. This issue is discussed in detail in [36], where it is proved that the construction of the Hadamard–Sylvester class gives an infinite family of Hadamard matrices with fourth last pivot $n/2$. It is also conjectured that this value can appear only for Hadamard matrices belonging to the Sylvester equivalence class.

In conclusion, the study of the growth factor is a intriguing and important issue in Numerical Analysis because it characterizes the roundoff estimates of Gaussian Elimination and captures its stability properties. Particularly, its investigation for Hadamard matrices led to the formulation of the open Complete Pivoting conjecture for GE and to the publication of several relevant articles. It would be also interesting to study the growth factors for other matrix factorizations, e.g. as it was done recently in [60] for the modified Gram–Schmidt algorithm by deriving upper bounds for growth factors arising at the solution of least squares problems.

Research on the values of minors of Hadamard matrices is ongoing, cf., e.g., [53] and the references therein. This work provides a distribution for any minors of orders j of Hadamard matrices with reference to minors of orders $n-j$, up to a factor $n^{j-\frac{n}{2}}$. This issue is closely related to the specification of the possible determinant values of matrices with entries ± 1. A survey on this problem and current updates and results can be found in [45], cf. also [8, 38]. Further progress on this issue can help in the study of Cryer's growth conjecture for Hadamard matrices, since Eq. (9) provides a powerful tool for computing pivots in terms of leading principal minors.

Further ongoing research is focused on how the criterion of Proposition 8 could be devised effectively for specifying good pivots for Hadamard matrices of general order. Such a study would contribute significantly to proving Cryer's growth conjecture. The specification of the H-equivalence classes giving good pivots remains an open problem. The identification of good pivot patterns for specific constructions except for Sylvester–Kronecker (e.g., Paley, Williamson, etc. [29, 51]) is under consideration, too. Also the role of the maximum element selection strategy in pivoting procedures is investigated. It is important to clarify which maximum selection approaches leads to good pivot patterns.

Other recent worthwhile scientific results and applications concerning Hadamard matrices, which are given from a Numerical Analysis viewpoint and without necessarily any reference to their minors, are presented in [1, 2]. The significance of developing efficient algorithms for the computation of principal minors of a matrix is highlighted, e.g., in [21]. Similar techniques to the ones presented here for Hadamard matrices can be adopted also for other classes of orthogonal designs as well, cf., e.g., [38, 39].

Acknowledgements The work of C. D. Kravvaritis was supported by Alexander von Humboldt Foundation, Bonn, Germany.

References

1. Bekas, C., Kokiopoulou, E., Saad, Y.: An estimator for the diagonal of a matrix. Appl. Numer. Math. **57**, 1214–1229 (2007)
2. Bella, T., Olshevsky, V., Sakhnovich, L.: Ranks of Hadamard Matrices and equivalence of Sylvester Hadamard and pseudo-noise matrices. Oper. Theory Adv. Appl. **1**, 1–13 (2007)
3. Brenner, J., Cummings, L.: The Hadamard maximum determinant problem. Am. Math. Mon. **79**, 626–630 (1972)
4. Christian, J.D., Shader, B.L.: Nonexitence results for Hadamard-like matrices. Electron. J. Comb. **11**, #N1 (2004)
5. Cohen, A.M.: A note on pivot size in Gaussian Elimination. Linear Algebra Appl. **8**, 361–368 (1974)
6. Cryer, C.W.: Pivot size in Gaussian Elimination. Numer. Math. **12**, 335–345 (1968)
7. Datta, B.N.: Numerical Linear Algebra and Applications, 2nd. edn. SIAM, Philadelphia (2010)
8. Day,J., Peterson,B.: Growth in Gaussian Elimination. Am. Math. Mon. **95**, 489–513 (1988)
9. Driscoll,T.A., Maki, K.L.: Searching for rare growth factors using multicanonical Monte Carlo methods. SIAM Rev. **49**, 673–692 (2007)
10. Edelman, A.: The complete pivoting conjecture for Gaussian Elimination is false. Math. J. **2**, 58–61 (1992)
11. Edelman, A., Friedman, D.: A counterexample to a Hadamard matrix pivot conjecture. Linear Multilinear Algebra **44**, 53–56 (1998)
12. Edelman, A., Mascarenhas, W.: On the complete pivoting conjecture for a Hadamard matrix of order 12. Linear Multilinear Algebra **38**, 181–187 (1995)
13. Ehlich, H.: Determinanten Abschätzungen für binäre Matrizen. Math. Z. **83**, 123–132 (1964)
14. Favati, P., Leoncini, M., Martiinez, A.: On the robustness of Gaussian Elimination with partial pivoting. BIT Numer. Math. **40**, 62–73 (2000)
15. Forsythe, G.E., Moler, C.B.: Computer Solution of Linear Algebraic Systems. Prentice Hall, Englewood Cliffs (1967)
16. Foster, L.V.: Gaussian Elimination with partial pivoting can fail in practice. SIAM J. Matrix Anal. Appl. **15**, 1354–1362 (1994)
17. Gantmacher, F.R.: The Theory of Matrices, vol. 1. Chelsea, New York (1959)
18. Geramita, A.V., Seberry, J.: Orthogonal Designs: Quadratic Forms and Hadamard Matrices. Marcel Dekker, New York (1979)
19. Golub, G.H., Van Loan, C.E.: Matrix Computations, 3rd edn. John Hopkins University Press, Baltimore (1996)

20. Gould, N.: On growth in Gaussian Elimination with pivoting. SIAM J. Matrix Anal. Appl. **12**, 354–361 (1991)
21. Griffin, K., Tsatsomeros, M.J. : Principal minors, Part I: a method for computing all the principal minors of a matrix. Linear Algebra Appl. **419**, 107–124 (2006)
22. Hadamard, J.: Résolution d'une question relative aux déterminants. Bull. Sci. Math. **17**, 240–246 (1893)
23. Hall, M.: Hadamard matrices of order 16. Jet Propulsion Laboratory, Technical Report 36–10, **1**, 21–26 (1961)
24. Hall, M.: Combinatorial Theory. Wiley, New York (1986)
25. Harwit, M., Sloane, N.J.A.: Hadamard Transform Optics. Academic, New York (1979)
26. Higham, N.J.: Accuracy and Stability of Numerical Algorithms. SIAM, Philadelphia (2002)
27. Higham, N.J.: Gaussian Elimination. WIREs Comput. Stat. **3**, 230–238 (2011)
28. Higham, N.J., Higham, D.J. : Large growth factors in Gaussian Elimination with pivoting. SIAM J. Matrix Anal. Appl. **10**, 155–164 (1989)
29. Horadam, K.J.: Hadamard Matrices and Their Appplications. Princeton University Press, Princeton (2007)
30. Horn,R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
31. Husain, Q.M.: On the totality of the solutions for the symmetrical incomplete block designs:$\lambda = 2$, $k = 5$ or 6. Sankhyā **7**, 204–208 (1945)
32. Kharaghani, H., Tayfeh-Rezaie, B.: A Hadamard matrix of order 428. J. Comb. Des. **13**, 435–440 (2005)
33. Koukouvinos, C., Lappas, E., Mitrouli, M., Seberry, J.: On the complete pivoting conjecture for Hadamard matrices of small orders. J. Res. Pract. Inf. Techol. **33**, 298–309 (2001)
34. Koukouvinos, C., Mitrouli, M., Seberry, J.: An algorithm to find formulae and values of minors of Hadamard matrices. Linear Algebra Appl. **330**, 129–147 (2001)
35. Koukouvinos, C., Mitrouli, M., Seberry, J.: Values of minors of an infinite family of D-optimal designs and their application to the growth problem. SIAM J. Matrix Anal. Appl. **23**, 1–14 (2001)
36. Koukouvinos, C., Mitrouli, M., Seberry, J.: An infinite family of Hadamard matrices with fourth last pivot n/2. Linear Multilinear Algebra **50**, 167–173 (2002)
37. Koukouvinos, C., Mitrouli, M., Seberry, J.: An algorithm to find formulae and values of minors of Hadamard matrices: II. Linear Algebra Appl. **371**, 111–124 (2003)
38. Kravvaritis, C., Mitrouli, M.: Evaluation of Minors associated to weighing matrices. Linear Algebra Appl. **426**, 774–809 (2007)
39. Kravvaritis, C., Mitrouli, M.: A technique for computing minors of binary Hadamard matrices and application to the growth problem. Electron. Trans. Numer. Anal. **31**, 49–67 (2008)
40. Kravvaritis, C., Mitrouli, M.: The growth factor of a Hadamard matrix of order 16 is 16. Numer. Linear Algebra Appl. **16**, 715–743 (2009)
41. Kravvaritis, C., Mitrouli, M.: On the complete pivoting conjecture for Hadamard matrices: further progress and a good pivots property. Numer. Algorithms **62**, 571–582 (2013)
42. Kravvaritis, C., Mitrouli, M., Seberry, J.: On the pivot structure for the weighing matrix $W(12, 11)$. Linear Multilinear Algebra **55**, 471–490 (2007)
43. Meyer, C.D.: Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia (2000)
44. Olschowka, M., Neumaier, A.: A new pivoting strategy for Gaussian Elimination. Linear Algebra Appl. **240**, 131–151 (1996)
45. Orrick, W. P., Solomon, B.: Spectrum of the determinant function. Indiana University. http://www.indiana.edu/~maxdet/spectrum.html (2012). Cited 19 May 2015
46. Seberry, J.: Hadamard matrices: some constructions for order 20. University of Wollongong. http://www.uow.edu.au/$\sim$jennie/hadamard.html (2001). Cited 19 May 2015
47. Seberry, J., Yamada, M.: Hadamard matrices, sequences and block designs. In: Stinson, D.J., Dinitz, J.H. (eds.) Contemporary Design Theory: A Collection of Surveys, pp. 431–560. Wiley, New York (1992)
48. Seberry, J., Xia, T., Koukouvinos, C., Mitrouli, M.: The maximal determinant and subdeterminants of ± 1 matrices. Linear Algebra Appl. **373**, 297–310 (2003)

49. Sharpe, F.R.: The maximum value of a determinant. Bull. Am. Math. Soc. **14**, 121–123 (1907)
50. Sloane, N.J.A.: A library of Hadamard matrices. http://neilsloane.com/hadamard/ (1999). Cited 19 May 2015
51. Stinson, D.R.: Combinatorial Designs: Constructions and Analysis. Springer, New York (2004)
52. Sylvester, J.J.: Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. Philos. Mag. **34**, 461–475 (1867)
53. Szöllősi, F.: Exotic complex Hadamard matrices and their equivalence. Cryptogr. Commun. **2**, 187–198 (2010)
54. Tornheim, L.: Pivot size in Gauss reduction. Tech. Report, Calif. Res. Corp., Richmond (1964)
55. Trappe, W., Washington, L.: Introduction to Cryptography with Coding Theory, 2nd edn. Pearson Education, Upper Saddle River (2006)
56. Trefethen, L.N.: Three mysteries of Gaussian elimination. ACM SIGNUM Newsl. **20**, 2–5 (1985)
57. Trefethen, L.N., Bau III, D.: Numerical Linear Algebra. SIAM, Philadelphia (1997)
58. Trefethen, L.N., Schreiber, R.S.: Average-case stability of Gaussian elimination. SIAM J. Matrix Anal. Appl. **11**, 335–360 (1990)
59. Wallis, W.D., Street, A.P., Wallis, J.S.: Combinatorics: Room Squares, Sum-Free Sets, Hadamard Matrices. Lectures Notes in Mathematics, vol. 292. Springer, New York (1972)
60. Wei, M., Liu, Q.: On growth factors of the modified Gram-Schmidt algorithm. Numer. Linear Algebra Appl. **15**, 621–626 (2008)
61. Wilkinson, J.H.: Error analysis of direct methods of matrix inversion. J. Assoc. Comput. Mach. **8**, 281–330 (1961)
62. Wilkinson, J.H.: The Algebraic Eigenvalue Problem. Oxford University Press, London (1965)
63. Wright, S.J.: A collection of problems for which Gaussian elimination with partial pivoting is unstable. SIAM J. Sci. Comput. **14**, 231–238 (1993)
64. Yarlagadda, R.K., Hershey, J.E.: Hadamard Matrix Analysis and Synthesis: With Applications to Communications and Signal/Image Processing. Kluwer, Boston (1997)

Localized Summability Kernels for Jacobi Expansions

H.N. Mhaskar

Abstract While the direct and converse theorems of approximation theory enable us to characterize the smoothness of a function $f : [-1, 1] \to \mathbb{R}$ in terms of its degree of polynomial approximation, they do not account for local smoothness. The use of localized summability kernels leads to a wavelet-like representation, using the Fourier–Jacobi coefficients of f, so as to characterize the smoothness of f in a neighborhood of each point in terms of the behavior of the terms of this representation. In this paper, we study the localization properties of a class of kernels, which have explicit forms in the "space domain," and establish explicit bounds on the Lebesgue constants on the summability kernels corresponding to some of these.

Keywords Jacobi expansions • Localized kernels • Jacobi translates • Product formulas

Introduction

Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function, and $n \geq 1$. In this paper, we denote the class of all algebraic polynomials of degree $< n$ by Π_n. A very simple and oft-used way to approximate f from Π_n is to minimize

$$\int_{-1}^{1} |f(y) - P(y)|^2 dy \tag{1}$$

over all polynomials $P \in \Pi_n$. Of course, it is well known that a solution to this problem can be written explicitly in terms of the Legendre expansion of f. Denoting by P_k the Legendre polynomial of degree k defined by

H.N. Mhaskar (✉)
Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA

Institute of Mathematical Sciences, Claremont Graduate University, Claremont, CA 91711, USA
e-mail: hrushikesh.mhaskar@cgu.edu

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_18

$$P_k(x) = \frac{(-1)^k}{2^k k!}\frac{d^k}{dx^k}(1-x^2)^k, \qquad x \in [-1,1],$$

and

$$\hat{f}(k) = \sqrt{k+1/2}\int_{-1}^{1} f(y)P_k(y)dy, \qquad k = 0, 1, \cdots,$$

the polynomial that minimizes the square integral error in (1) is given by

$$s_n(f)(x) = \sum_{k=0}^{n-1} \sqrt{k+1/2}\hat{f}(k)P_k(x).$$

(For clarity of exposition, the notation used in the introduction will be slightly different from the one used in the rest of the paper.) We will refer to $s_n(f)$ as the projection of f on Π_n.

Although the projection is the best approximation to f in the sense of L^2 approximation, the projections do not converge to f uniformly unless f is sufficiently smooth. Moreover, even if f is analytic everywhere on $[-1,1]$ except at a small number of points, the behavior of these projections is affected adversely at all points of $[-1,1]$ (cf. for example, the left figure in Fig. 1.)

In the theory of spectral methods for numerical solutions of differential equations, a standard strategy therefore is to first estimate the location of such "singularities," and device different methods for approximation of f on intervals where it is smooth, and where it is not [4, 13, 14]. This task, however, is difficult, because the starting point of spectral methods is the data $\{\hat{f}(k)\}_{k=0}^{\infty}$ (actually, a finite set of the coefficients), and this sequence by itself does not reveal the locations of the singularities. For this reason, polynomial approximation is often abandoned in practice in favor of spline or wavelet approximations.

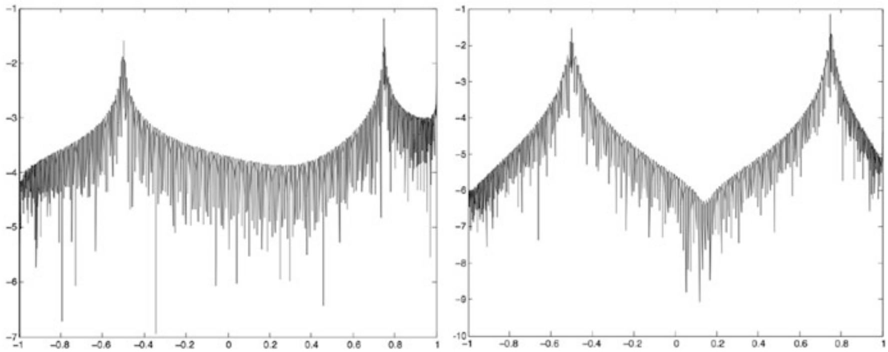


Fig. 1 On the *left*, the plot of $\log_{10}|f(x) - s_n(f)(x)|$, on the *right*, the plot of $\log_{10}|f(x) - \sigma_n(f)(x)|$, $n = 256$, where f is defined in (18.4)

Clearly, it is desirable to devise methods based on the information $\{\hat{f}(k)\}_{k=0}^{n-1}$ for a finite integer $n \geq 1$ that would detect the location of singularities accurately, as well as yield pointwise approximation automatically adjusting the accuracy to the optimal one depending on the local smoothness of the function f. This would avoid the use of elaborate separate methods for approximating f on different intervals.

We have carried out such an analysis in a series of papers [2, 7, 8, 10]. In particular, we have shown that a filtered projection of the form

$$\sigma_n(f)(x) = \sum_{k=0}^{n-1} h(k/n)\sqrt{k+1/2}\hat{f}(k)P_k(x)$$

has the necessary localization properties, if h is sufficiently smooth, and satisfies some other technical conditions. From the definition of $\hat{f}(k)$, it is clear that

$$\sigma_n(f)(x) = \int_{-1}^{1} f(y)\Phi_n(x,y)dy, \tag{2}$$

where

$$\Phi_n(x,y) = \sum_{k=0}^{n-1} h(k/n)(k+1/2)P_k(x)P_k(y).$$

In [7, 8], we have proved in particular that for any integer $S \geq 2$, h can be chosen to ensure that

$$|\Phi_n(\cos\theta, \cos\phi)| \leq c\frac{n^2}{\max(1, (n|\theta-\phi|)^S)}, \qquad n = 1, 2, \cdots, \ \theta, \phi \in [0, \pi], \tag{3}$$

where $c > 0$ is a constant depending on S alone.

We discuss an example to illustrate the effect of this localization. Let

$$f(x) = |x+1/2|^{1/2} + |x-3/4|^{1/3}, \qquad x \in [-1, 1]. \tag{4}$$

Clearly, f is analytic on $[-1, 1]$ except at $-1/2, 3/4$. We choose

$$h(t) := \begin{cases} 1, & \text{if } 0 \leq t \leq 1/2, \\ \exp\left(-\dfrac{\exp(2/(1-2t))}{1-t}\right), & \text{if } 1/2 < t < 1, \\ 0, & \text{if } t \geq 1. \end{cases}$$

It is seen readily from Fig. 1 that because of the singularities at $-1/2$ and $3/4$, the error in approximation by the projection $s_n(f)$ exceeds 10^{-4} also at "far away" points such as $\pm 1, 0.2$. The corresponding errors by the filtered projection $\sigma_n(f)$ are at least 100 times smaller.

In [8], we have obtained the necessary quadrature formulas to enable us to approximate the integral expression (2) for $\sigma_n(f)$ based on "scattered data" on f; i.e., values of f at points which are not necessarily the Gauss–Jacobi quadrature nodes, while making sure that the degree of approximation estimates are the same as those for the integral operator based on $\{\hat{f}(k)\}$. The filtered projection method with this modification was used recently for numerical differentiation, and applied to the problem of short-term prediction of blood glucose levels based on continuous glucose monitoring devices [9]. The result was substantially better than the state-of-the-art technology.

For such applications, it is desirable to use a kernel which has a closed form formula. The purpose of this paper is to propose such a kernel, and study its necessary localization properties, analogous to (3) for the kernels Φ_n. A version of this kernel was given in [3] in the case of spherical polynomials. We will do this in the more general setting of Jacobi polynomials. In section "Main Results," we introduce these kernels, and state their localization properties. Some ideas on the computational aspects of evaluating the kernels are discussed in section "Comments on Implementation." The proofs of the results in section "Main Results" are given in section "Proofs."

Main Results

For $\alpha, \beta > -1$, we define the Jacobi weight (with parameters α, β) by

$$w_{\alpha,\beta}(y) = (1-y)^{\alpha}(1+y)^{\beta}, \qquad y \in (-1,1). \tag{5}$$

For integer $k \geq 0$, the (orthonormalized) Jacobi polynomial $p_k^{(\alpha,\beta)}$ is a polynomial of degree k, with positive leading coefficient, such that the sequence $\{p_k^{(\alpha,\beta)}\}_{k=0}^{\infty}$ satisfies the orthogonality condition

$$\int_{-1}^{1} p_k^{(\alpha,\beta)}(y) p_j^{(\alpha,\beta)}(y) w_{\alpha,\beta}(y) dy = \delta_{k,j}, \qquad k,j = 0,1,\cdots. \tag{6}$$

Clearly,

$$p_k^{(\beta,\alpha)}(y) = (-1)^k p_k^{(\alpha,\beta)}(-y), \qquad k = 0,1,\cdots, \ y \in [-1,1].$$

Therefore, it is customary to state and prove theorems in the theory of Jacobi polynomials assuming that $\alpha \geq \beta$, and we will make this assumption throughout this paper.

If $f : [-1,1] \to \mathbb{R}$, and $w_{\alpha,\beta} f$ is Lebesgue integrable on $[-1,1]$, then we define

$$\hat{f}(k) = \hat{f}(\alpha,\beta;k) = \int_{-1}^{1} f(y) p_k^{(\alpha,\beta)}(y) w_{\alpha,\beta}(y) dy, \qquad k = 0,1,\cdots, \tag{7}$$

and define the Fourier–Jacobi projection operator by

$$s_n(f)(x) = s_n(\alpha,\beta;f)(x) = \sum_{k=0}^{n-1} \hat{f}(k) p_k^{(\alpha,\beta)}(x), \qquad n = 1, 2, \cdots. \tag{8}$$

Clearly,

$$s_n(f)(x) = \int_{-1}^{1} f(y) K_n(x,y) w_{\alpha,\beta}(y) dy, \qquad n = 1, 2, \cdots, \ x \in [-1,1], \tag{9}$$

where the *Christoffel–Darboux* kernel K_n is defined by

$$K_n(x,y) = K_n^{(\alpha,\beta)}(x,y) = \sum_{k=0}^{n-1} p_k^{(\alpha,\beta)}(x) p_k^{(\alpha,\beta)}(y), \qquad n = 1, 2, \cdots, \ x, y \in [-1,1]. \tag{10}$$

In the theory of Fourier–Jacobi expansions, the Christoffel–Darboux kernel plays the role of the Dirichlet kernel in the theory of classical trigonometric Fourier series.

An important fact in the latter is that the partial sum of the trigonometric Fourier series can be expressed as a convolution with the Dirichlet kernel, so that in the proofs of several theorems, one may prove statements only at the point 0 rather than arbitrary points on $[-\pi,\pi]$. We now describe a convolution structure for Jacobi polynomials, where the role of 0 will be played by 1. Koornwinder has shown in [6] that for $\alpha \geq \beta \geq -1/2$, $k = 0, 1, \cdots$, the Jacobi polynomials $p_k^{(\alpha,\beta)}$ satisfy the following product formula: for $x, y \in [-1,1]$,

$$p_k^{(\alpha,\beta)}(x)\, p_k^{(\alpha,\beta)}(y) = p_k^{(\alpha,\beta)}(1) \int_0^{\pi} \int_0^{1} p_k^{(\alpha,\beta)}(Z(x,y;r,\psi))\, dm_{\alpha,\beta}(r,\psi), \tag{11}$$

where

$$Z(x,y;r,\psi) = \frac{1}{2}(1+x)(1+y) + \frac{1}{2}(1-x)(1-y)\, r^2 + \sqrt{1-x^2}\sqrt{1-y^2}\, r\cos\psi \, -1 \tag{12}$$

and

$$dm_{\alpha,\beta}(r,\psi) = \begin{cases} \frac{2\,\Gamma(\alpha+1)}{\sqrt{\pi}\,\Gamma(\alpha-\beta)\,\Gamma(\beta+1/2)} (1-r^2)^{\alpha-\beta-1}\, r^{2\beta+1} \sin^{2\beta}\psi\, dr\, d\psi, \\ \qquad \text{if } \alpha > \beta > -1/2, \\ \frac{2\,\Gamma(\alpha+1)}{\sqrt{\pi}\,\Gamma(\alpha+1/2)} (1-r^2)^{\alpha-1/2}\, dr\, \frac{1}{2}\{\delta(0)+\delta(\pi)\}, \\ \qquad \text{if } \alpha > \beta = -1/2, \\ \frac{\Gamma(\alpha+1)}{\sqrt{\pi}\,\Gamma(\alpha+1/2)} \delta(1-r)\, \sin^{2\beta}\psi\, d\psi\,, \\ \qquad \text{if } \alpha = \beta > -1/2, \\ \delta(1-r)\, \frac{1}{2}\{\delta(0)+\delta(\pi)\}, \\ \qquad \text{if } \alpha = \beta = -1/2, \end{cases} \tag{13}$$

where δ denotes the Dirac delta measure supported at 0. The constants are chosen such that

$$\int_0^\pi \int_0^1 dm_{\alpha,\beta}(r,\psi) = 1, \tag{14}$$

i.e., $m_{\alpha,\beta}$ is a probability measure.

The *Jacboi translation* $\mathscr{T}_y$ of a function f by $y \in [-1,1]$ evaluated at $x \in [-1,1]$ is defined (whenever the integral expression below is well defined) by

$$\mathscr{T}_y f(x) = \int_0^\pi \int_0^1 f(Z(x,y;r,\psi))\, dm_{\alpha,\beta}(r,\psi). \tag{15}$$

In view of (11),

$$\mathscr{T}_y(p_k^{(\alpha,\beta)}(1)\,p_k^{(\alpha,\beta)})(x) = p_k^{(\alpha,\beta)}(x)\,p_k^{(\alpha,\beta)}(y) = \mathscr{T}_x(p_k^{(\alpha,\beta)}(1)\,p_k^{(\alpha,\beta)})(y), \tag{16}$$

and we may rewrite (9) in form

$$s_n(f)(x) = \int_{-1}^1 f(y)\mathscr{T}_x K_n(y,1) w_{\alpha,\beta}(y)dy, \qquad n = 1,2,\cdots,\ x \in [-1,1], \tag{17}$$

in complete analogy with the classical expression for the trigonometric Fourier projection in terms of convolution with the Dirichlet kernel. It is convenient to overload the notation and denote

$$K_n(y) = K_n(y,1), \qquad y \in [-1,1]. \tag{18}$$

With this preparation, we are now ready to define our kernel and state our main theorem. Let $S \geq 2$ be an integer. For integer $n = 1,2,\cdots$ and $y \in [-1,1]$, we define

$$\mathfrak{F}_{S,n}(y) := \mathfrak{F}_{\alpha,\beta;S,n}(y) := K_{S(n-1)+1}(y)\left(\frac{K_n(y)}{K_n(1)}\right)^{S-1}, \tag{19}$$

Clearly, $\mathfrak{F}_{S,n} \in \Pi_{(2S-1)(n-1)+1}$. Corresponding to this kernel, we define

$$\mathscr{V}_{S,n}(f)(x) = \int_{-1}^1 f(y)\mathscr{T}_x \mathfrak{F}_{S,n}(y) w_{\alpha,\beta}(y)dy, \qquad n = 1,2,\cdots,\ x \in [-1,1]. \tag{20}$$

In order to state our theorem, we need some further notation. For a Lebesgue measurable function $f : [-1,1] \to \mathbb{R}$, we define

$$\|f\|_p = \|f\|_{\alpha,\beta;p} = \begin{cases} \left(\int_{-1}^1 |f(y)|^p w_{\alpha,\beta}(y)dy\right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \text{ess sup}_{y\in[-1,1]}|f(y)|, & \text{if } p = \infty. \end{cases}$$

The space of all functions f for which $\|f\|_p < \infty$ will be denoted by $L^p = L^p_{\alpha,\beta}$, with the usual convention that functions which are equal almost everywhere are considered to be equal as members of L^p. The notation $X^p = X^p_{\alpha,\beta}$ will denote L^p if $1 \le p < \infty$, and the space of all continuous functions on $[-1, 1]$, endowed with the uniform norm, if $p = \infty$. Thus, X^p is the L^p–closure of $\bigcup_{n\ge1} \Pi_n$. For $f \in X^p$, we define the degree of approximation of f by Π_n by

$$E_{n,p}(f) = E_{n,p}(\alpha, \beta; f) = \min_{P\in\Pi_n} \|f - P\|_p. \tag{21}$$

In the sequel, it is convenient to write

$$\varpi_{\alpha,\beta}(y) = (1-y)^{\alpha/2+1/4}(1+y)^{\beta/2+1/4}, \qquad y \in [-1, 1]. \tag{22}$$

Also, we make the following convention regarding constants. In the sequel, $c, c_1, \cdots$ denote positive constants depending only on α, β, S, and other such obvious parameters fixed in the discussion. Their values may be different at different occurrences, even with a single formula. The notation $A \sim B$ means $c_1 A \le B \le c_2 A$.

Our main theorem is the following.

Theorem 1. *Let $\alpha \ge \beta \ge -1/2$, $x = \cos\theta$, $y = \cos\phi$, $\theta, \phi \in [0, \pi]$, $S \ge 2$, $n \ge 2$ be integers, and $1 \le p \le \infty$.*

(a) *For $P \in \Pi_n$, $\mathscr{V}_n(P) = P$.*
(b) *We have*

$$\int_{-1}^{1} |\mathscr{T}_x \mathfrak{F}_{S,n}(y)| w_{\alpha,\beta}(y) dy \le S^{\alpha+1}. \tag{23}$$

Hence, for $f \in L^p$,

$$\|\mathscr{V}_{S,n}(f)\|_p \le S^{\alpha+1}\|f\|_p, \tag{24}$$

and

$$E_{(2S-1)(n-1)+1,p}(f) \le \|f - \mathscr{V}_{S,n}(f)\|_p \le (1 + S^{\alpha+1})E_{n,p}(f). \tag{25}$$

(c) *We have*

$$|\mathscr{T}_x \mathfrak{F}_{S,n}(y)| \le c \frac{n^{2\alpha+2}}{\max(1, (n|\theta - \phi|)^{(\alpha-\beta+1)(S-1)})}. \tag{26}$$

and

$$\varpi(x)\varpi(y)|\mathscr{T}_x \mathfrak{F}_{S,n}(y)| \le c \frac{n^{\alpha+3/2}}{\max(1, (n|\theta - \phi|)^{(\alpha-\beta+1)(S-1)+1})}. \tag{27}$$

To illustrate the localization of the kernels $\mathfrak{F}_{S,n}$, we observe the following corollary, which shows that if S is sufficiently large, then the maximum absolute value of $|\mathscr{V}_{S,n}(f)(x)|$ on a neighborhood of a point x_0 is bounded by the maximum absolute value of $|f(x)|$ on a slightly larger neighborhood plus a term that tends to 0 as $n \to \infty$. For $x_0 = \cos\theta_0 \in [-1,1]$, and $r > 0$, we denote

$$\mathbb{B}(x_0, r) = \{\cos\theta : |\theta - \theta_0| \le r\}.$$

Corollary 1. *Let* $\alpha \ge \beta \ge -1/2, f \in X^\infty$, $x_0 \in [-1,1]$, $\delta > 0$. *Then*

$$\max_{x \in \mathbb{B}(x_0,\delta)} |\mathscr{V}_{S,n}(f)(x)| \le c(x_0,\delta)\left\{\max_{x\in\mathbb{B}(x_0,3\delta/2)} |f(x)| + \frac{\|f\|_\infty}{n^{(\alpha-\beta+1)(S-1)-2\alpha-2}}\right\}. \tag{28}$$

Comments on Implementation

Since it is difficult to evaluate the Jacobi translates directly from the definition (15), we make some comments on computing the kernel $\mathfrak{F}_{S,n}$ in this section. The idea is to express the kernel first as an expansion in Jacobi polynomials using (41) below, and then use (16).

In view of (41),

$$\mathfrak{F}_{S,n}(y) = \frac{2(\alpha+1)}{2S(n-1)+\alpha+\beta+2} \frac{p_{2S(n-1)}^{(\alpha+1,\beta)}(1)}{\left(p_{n-1}^{(\alpha+1,\beta)}(1)\right)^{S-1}} p_{2S(n-1)}^{(\alpha+1,\beta)}(y)\left(p_{n-1}^{(\alpha+1,\beta)}(y)\right)^{S-1}. \tag{29}$$

Since the fraction expression above can be computed using (40), we will describe the implementation of $p_{n-1}^{(\alpha+1,\beta)}(y)p_{n-1}^{(\alpha+1,\beta)}(y)^{S-1}$. More generally, we will describe an expansion of $p_m^{(\alpha+1,\beta)}(y)p_\ell^{(\alpha+1,\beta)}(y)$ for integers $\ell, m = 0, 1, \cdots$. It is well known [1, Theorem 5.1] that under our assumption that $\alpha \ge \beta \ge -1/2$,

$$p_m^{(\alpha+1,\beta)}(y)p_\ell^{(\alpha+1,\beta)}(y) = \sum_{k=|\ell-m|}^{\ell+m} C(k,\ell,m)p_k^{(\alpha+1,\beta)}(y), \qquad \ell, m \ge 0. \tag{30}$$

where *linearization coefficients* $C(k,\ell,m)$ are positive numbers. Setting $C(k,\ell,m) = 0$ when any of the variables k, ℓ, m are out of their indicated ranges, one can obtain a recurrence formula for these coefficients (cf. [5]).

To describe this recurrence, we first observe the recurrence formula for the Jacobi polynomials themselves:

$$yp_k^{(\alpha+1,\beta)}(y) = \rho_k p_{k+1}^{(\alpha+1,\beta)}(y) + d_k p_k^{(\alpha+1,\beta)}(y) + \rho_{k-1}p_{k-1}^{(\alpha+1,\beta)}(y), \tag{31}$$

with the initial terms $p_{-1}^{(\alpha+1,\beta)}(y) = 0$, and

$$p_0^{(\alpha+1,\beta)}(y) = \sqrt{\frac{\Gamma(\alpha+\beta+3)}{2^{\alpha+\beta+2}\Gamma(\alpha+2)\Gamma(\beta+1)}}, \tag{32}$$

where

$$\rho_0 := \frac{2}{\alpha+\beta+3}\sqrt{\frac{(\alpha+2)(\beta+1)}{\alpha+\beta+4}},\ d_0 := \frac{2(\beta-\alpha-1)}{2\alpha+2\beta+5}, \tag{33}$$

and for $k = 1, 2, \cdots$,

$$\begin{aligned} \rho_k &:= 2\sqrt{\frac{(k+1)(k+\alpha+2)(k+\beta+1)(k+\alpha+\beta+2)}{(2k+\alpha+\beta+2)(2k+\alpha+\beta+3)^2(2k+\alpha+\beta+4)}}, \\ d_k &:= \frac{\beta^2-(\alpha+1)^2}{(2k+\alpha+\beta+1)(2k+\alpha+\beta+2)}. \end{aligned} \tag{34}$$

The recurrence (31) leads to

$$\begin{aligned} & y p_m^{(\alpha+1,\beta)}(y) p_\ell^{(\alpha+1,\beta)}(y) \\ &= \rho_m p_{m+1}^{(\alpha+1,\beta)}(y) p_\ell^{(\alpha+1,\beta)}(y) + d_m p_m^{(\alpha+1,\beta)}(y) p_\ell^{(\alpha+1,\beta)}(y) + \\ &\qquad \rho_{m-1} p_{m-1}^{(\alpha+1,\beta)}(y) p_\ell^{(\alpha+1,\beta)}(y) \\ &= \sum_k \{\rho_m C(k,\ell,m+1) + d_m C(k,\ell,m) + \rho_{m-1} C(k,\ell,m-1)\}\, p_k^{(\alpha+1,\beta)}(y). \end{aligned} \tag{35}$$

On the other hand, (30) and (31) imply that

$$\begin{aligned} & y p_m^{(\alpha+1,\beta)}(y) p_\ell^{(\alpha+1,\beta)}(y) \\ &= \sum_k C(k,\ell,m)\, y p_k^{(\alpha+1,\beta)}(y) \\ &= \sum_k C(k,\ell,m) \left\{\rho_k p_{k+1}^{(\alpha+1,\beta)}(y) + d_k p_k^{(\alpha+1,\beta)}(y) + \rho_{k-1} p_{k-1}^{(\alpha+1,\beta)}(y)\right\} \\ &= \sum_k \{\rho_{k-1} C(k-1,\ell,m) + d_k C(k,\ell,m) + \rho_{k+1} C(k+1,\ell,m)\}\, p_k^{(\alpha+1,\beta)}(y). \end{aligned} \tag{36}$$

Comparing (36) and (35), we get a "five star" recurrence formula

$$\begin{aligned}
&\rho_m C(k,\ell,m+1)\\
&= \rho_{k-1}C(k-1,\ell,m) + d_k C(k,\ell,m) + \rho_{k+1}C(k+1,\ell,m)\\
&\qquad -d_m C(k,\ell,m) - \rho_{m-1}C(k,\ell,m-1). \qquad (37)
\end{aligned}$$

To apply this formula, we note that for every $k,\ell = 0,1,\cdots,$

$$C(k,\ell,-1) = 0, \quad C(k,\ell,0) = \delta_{k,\ell}. \qquad (38)$$

Assuming that we know $C(k,\ell,j)$ for all $k,\ell = 0,1,\cdots,$ and $-1 \le j \le m-1$, we can use (37) recursively to obtain $C(k,\ell,m)$ in (30).

Now, to compute $\mathfrak{F}_{S,n}(y)$, we first use (30) repeatedly to obtain

$$\mathfrak{F}_{S,n}(y) = \sum_k a_k(S,n,\alpha,\beta)p_k^{(\alpha+1,\beta)}(y),$$

and hence, using (41),

$$\mathfrak{F}_{S,n}(y) = \sum_k A_k(S,n,\alpha,\beta)p_k^{(\alpha,\beta)}(1)p_k^{(\alpha,\beta)}(y).$$

The required Jacobi translate is then given by

$$\mathfrak{F}_{S,n}(y) = \sum_k A_k(S,n,\alpha,\beta)p_k^{(\alpha,\beta)}(x)p_k^{(\alpha,\beta)}(y). \qquad (39)$$

Proofs

The proof of Theorem 1 requires some preparation. We begin by listing some well-known results about Jacobi polynomials.

Lemma 1. *Let* $\alpha \ge \beta \ge -1/2$*,* $y \in [-1,1]$*,* $n \ge 1$ *be an integer.*

(a)

$$\begin{aligned}
&|p_n^{(\alpha,\beta)}(y)| \le p_n^{(\alpha,\beta)}(1)\\
&\quad = \left\{\frac{2n+\alpha+\beta+1}{2^{\alpha+\beta+1}}\frac{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}\right\}^{1/2}\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\Gamma(n+1)}\\
&\quad \sim n^{\alpha+1/2}. \qquad (40)
\end{aligned}$$

(b)

$$K_n(y) = \frac{2(\alpha+1)}{2n+\alpha+\beta} p_{n-1}^{(\alpha+1,\beta)}(1) p_{n-1}^{(\alpha+1,\beta)}(y). \tag{41}$$

(c)

$$\begin{aligned} K_n(1) &= 2^{-\alpha-\beta-1} \frac{\Gamma(n+\alpha+\beta+1)\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\Gamma(\alpha+2)\Gamma(n)\Gamma(n+\beta)} \\ &= \frac{n^{2\alpha+2}}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\alpha+2)}(1+o(1/n)). \end{aligned} \tag{42}$$

(d)

$$|p_n^{(\alpha,\beta)}(y)| \le c \begin{cases} \min((1-y)^{-\alpha/2-1/4}, n^{\alpha+1/2}), & \text{if } 0 \le y \le 1, \\ \min((1+y)^{-\beta/2-1/4}, n^{\beta+1/2}), & \text{if } -1 \le y < 0, \end{cases} \tag{43}$$

In particular,

$$\max_{y\in[-1,1]} |\varpi(y) p_n^{(\alpha,\beta)}(y)| \le c, \tag{44}$$

and

$$(\sqrt{1-y} + 1/n)^{2\alpha+1}(\sqrt{1+y} + 1/n)^{2\beta+1} K_n(y,y) \sim n, \tag{45}$$

where the constants are independent of y*.*

Proof. Part (a) is [12, Estimate (7.32.2), Eq. (4.1.1)]. Part (b) follows from [12, Eqs. (4.5.3), (4.3.4)] by writing $n-1$ in place of n to account for our different notation. Part (c) is [12, Eq. (4.5.8)], again with the same replacement. The estimates (43) and (44) in part (d) can be deduced easily from [12, Theorem 7.32.2], and the estimate (45) is given in [11, Lemma 5, p. 108]. □

Next, we state a few important properties of the Jacobi translation, summarized from [1].

Proposition 1. *Let* $1 \le p \le \infty$*,* $f \in X^p$*,* $x, y \in [-1,1]$*. Then each of the following statements holds.*

(a) $\mathscr{T}_y f(x) = \mathscr{T}_x f(y)$.
(b) $\|\mathscr{T}_x f\|_p \le \|f\|_p$*, with equality if* $p = 1$.
(c)

$$\int_{-1}^{1} \mathscr{T}_x f(y) w_{\alpha,\beta}(y) dx = \int_{-1}^{1} f(y) w_{\alpha,\beta}(y) dy.$$

We need some more refined estimates than those given in Proposition 1. We find it easier to prove the statements than to find a reference.

Theorem 2. *Let* $g : [0,2] \to [0,\infty)$ *be non-increasing on* $[0,2]$, $P \in \Pi_n$, $f \in X^\infty$, $x = \cos\theta$, $y = \cos\phi$, $\theta, \phi \in [0,\pi]$.

(a) *If* $|f(z)| \le g(1-z)$ *for all* $z \in [-1,1]$,

$$|\mathscr{T}_y f(x)| \le g(1-\cos(\theta-\phi)).$$

(b) *If*

$$\varpi(z)|f(z)| \le g(1-z), \qquad z \in [-1,1], \tag{46}$$

then

$$\varpi(x)\varpi(y)|\mathscr{T}_y f(x)| \le cg(1-\cos(\theta-\phi)). \tag{47}$$

In order to prove Theorem 2, we first prove a lemma.

Lemma 2. *Let* $A \ge B \ge 0$, $\theta, \phi \in [0,\pi]$, $x = \cos\theta$, $y = \cos\phi$, $r \in [0,1]$, $\psi \in [0,\pi]$, *and* $Z = Z(x,y;r,\psi)$ *be as defined in (12). We have*

$$(1-Z)^A(1+Z)^B \ge c(1-x)^A(1-y)^A(1+x)^B(1+y)^B(1-r^2)^{A-B}r^B\sin^{2B}(\psi). \tag{48}$$

Proof. We observe that

$$\begin{aligned}1-Z &= 1-xy-\sqrt{1-x^2}\sqrt{1-y^2}-\frac{1}{2}(1-x)(1-y)(r^2-1)\\ &\qquad -\sqrt{1-x^2}\sqrt{1-y^2}(r\cos\psi-1)\\ &= 1-\cos(\theta-\phi)+\frac{1}{2}(1-x)(1-y)(1-r^2)\\ &\qquad +\sqrt{1-x^2}\sqrt{1-y^2}(1-r\cos\psi)\\ &\ge \frac{1}{2}(1-x)(1-y)(1-r^2), \qquad \psi \in [0,\pi].\end{aligned} \tag{49}$$

Let $\psi \in [0,\pi/2)$, so that $\cos\psi > 0$. Then the definition (12) of Z and the arithmetic–geometric mean inequality show that

$$\begin{aligned}1+Z &= \frac{1}{2}(1+x)(1+y)+\frac{1}{2}(1-x)(1-y)\,r^2+\sqrt{1-x^2}\sqrt{1-y^2}\,r\cos\psi\\ &\ge \frac{1}{2}(1+x)(1+y)+\frac{1}{2}(1-x)(1-y)\,r^2 \ge \sqrt{1-x^2}\sqrt{1-y^2}r.\end{aligned} \tag{50}$$

Further, [cf. (49)]

$$\begin{aligned}1-Z &= 1-\cos(\theta-\phi)+\frac{1}{2}(1-x)(1-y)(1-r^2)\\ &\quad +\sqrt{1-x^2}\sqrt{1-y^2}(1-r\cos\psi)\\ &\geq \sqrt{1-x^2}\sqrt{1-y^2}(1-r\cos\psi)\geq\sqrt{1-x^2}\sqrt{1-y^2}(1-\cos\psi)\\ &= 2\sqrt{1-x^2}\sqrt{1-y^2}\sin^2(\psi/2). \end{aligned}\tag{51}$$

Since $\sin^2(\psi/2)\geq(1/4)\sin^2\psi$, (50) and (51) yield

$$1-Z^2\geq(1/2)(1-x^2)(1-y^2)r\sin^2\psi,\qquad \psi\in[0,\pi/2).\tag{52}$$

Since

$$(1-Z)^A(1+Z)^B=(1-Z)^{A-B}(1-Z^2)^B,\tag{53}$$

we obtain (48) from (49) and (52) in the case when $\psi\in[0,\pi/2]$.

Next, let $\psi\in[\pi/2,\pi]$, so that $\cos\psi\leq 0$. In view of the arithmetic–geometric mean inequality,

$$\left|\sqrt{1-x^2}\sqrt{1-y^2}\,r\cos\psi\right|\leq\frac{1}{2}(1+x)(1+y)\cos^2\psi+\frac{1}{2}(1-x)(1-y)\,r^2.$$

Hence, the first equation in (50) shows that

$$\begin{aligned}1+Z &\geq \frac{1}{2}(1+x)(1+y)\sin^2\psi+\frac{1}{2}(1+x)(1+y)\cos^2\psi+\frac{1}{2}(1-x)(1-y)\,r^2\\ &\quad +\sqrt{1-x^2}\sqrt{1-y^2}\,r\cos\psi\\ &\geq \frac{1}{2}(1+x)(1+y)\sin^2\psi. \end{aligned}\tag{54}$$

Similarly, using the arithmetic–geometric mean inequality we obtain

$$\begin{aligned}1+Z &\geq \frac{1}{2}(1-x)(1-y)\,r^2\sin^2\psi+\frac{1}{2}(1-x)(1-y)\,r^2\cos^2\psi+\frac{1}{2}(1+x)(1+y)\\ &\quad +\sqrt{1-x^2}\sqrt{1-y^2}\,r\cos\psi\\ &\geq \frac{1}{2}(1-x)(1-y)r^2\sin^2\psi. \end{aligned}\tag{55}$$

Using (54), (55), and the arithmetic–geometric mean inequality once more,

$$\begin{aligned}1+Z &\geq \frac{1}{4}(1+x)(1+y)\sin^2\psi+\frac{1}{4}(1-x)(1-y)r^2\sin^2\psi\\ &\geq (1/2)\sqrt{1-x^2}\sqrt{1-y^2}r\sin^2\psi. \end{aligned}\tag{56}$$

Since $\cos\psi \le 0$, (51) shows that

$$1 - Z \ge \sqrt{1-x^2}\sqrt{1-y^2},$$

we conclude from (56) that

$$1 - Z^2 \ge (1/2)(1-x^2)(1-y^2)r\sin^2\psi, \qquad \psi \in [\pi/2, \pi]. \tag{57}$$

The estimate (48) follows from (49) and (57) in the case when $\psi \in [\pi/2, \pi]$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Part (a) was proved in [2, Lemma 3.1]. To prove part (b), we need only to consider the cases when $\alpha > \beta > -1/2$, $\alpha = \beta > -1/2$, and $\alpha > \beta = -1/2$. We will use the notation $Z = Z(x, y; r, \psi)$, $A = \alpha/2 + 1/4$, $B = \beta/2 + 1/4$, so that $A - B = (\alpha - \beta)/2$. The conditions of Lemma 2 are satisfied with these choices. We note that (49) shows that

$$1 - Z(x, y; r, \psi) \ge 1 - \cos(\theta - \phi). \tag{58}$$

Case $\alpha > \beta > -1/2$.
The definition (13) shows that

$$\begin{aligned}
&\int_0^\pi \int_0^1 (1-r^2)^{-(A-B)} r^{-B} \sin^{-2B}\psi\, dm_{\alpha,\beta}(r,\psi)\\
&= c\int_0^\pi \int_0^1 (1-r^2)^{\alpha/2-\beta/2-1} r^{(3/2)(\beta+1/2)} \sin^{\beta-1/2}\psi\, dr d\psi \le c.
\end{aligned} \tag{59}$$

Hence, using (11), (46), (48), and (58) we conclude that

$$\begin{aligned}
&(1-x)^A(1-y)^A(1+x)^B(1+y)^B|\mathscr{T}_y f(x)|\\
&\le c\int_0^\pi \int_0^1 (1-Z)^A(1+Z)^B|f(Z)|(1-r^2)^{-(A-B)} r^{-B}\sin^{-2B}\psi\, dm_{\alpha,\beta}(r,\psi)\\
&\le c\int_0^\pi \int_0^1 g(1-Z)(1-r^2)^{-(A-B)} r^{-B}\sin^{-2B}\psi\, dm_{\alpha,\beta}(r,\psi)\\
&\le cg(1-\cos(\theta-\phi))\int_0^\pi \int_0^1 (1-r^2)^{-(A-B)} r^{-B}\sin^{-2B}\psi\, dm_{\alpha,\beta}(r,\psi)\\
&\le cg(1-\cos(\theta-\phi)).
\end{aligned}$$

This proves (47) in the case $\alpha > \beta > -1/2$.
Case $\alpha = \beta > -1/2$
In this case, $A = B$, and the measure $dm_{\alpha,\alpha}(r,\psi)$ is supported only on $\{1\}\times[0,\pi]$. In view of (13),

$$\int_0^\pi \sin^{-2B}\psi\, dm_{\alpha,\alpha}(1,\psi) \le c\int_0^\pi \sin^{\beta-1/2} d\psi \le c. \tag{60}$$

For $r = 1, A = B$, (48) takes the form

$$(1-Z^2)^A \ge c(1-x)^A(1-y)^A(1+x)^B(1+y)^B \sin^{2B}(\psi).$$

Hence, (11), (46), (58), and (60) imply as before that

$$\begin{aligned}(1-x^2)^A(1-y^2)^A|\mathscr{T}_y f(x)| &\le cg(1-\cos(\theta-\phi))\int_0^\pi \sin^{-2B}\psi\, dm_{\alpha,\alpha}(1,\psi)\\ &\le cg(1-\cos(\theta-\phi)).\end{aligned}$$

This proves (47) in the case $\alpha = \beta > -1/2$.

Case $\alpha > \beta = -1/2$

In this case, $B = 0$, and (48) takes the form

$$(1-x)^A(1-y)^A \le c(1-Z(x,y;r,\psi))^A(1-r^2)^{1/4-\alpha/2}, \qquad r\in[0,1],\ \psi\in[0,\pi].$$

Therefore, using (15), (13), (46), and (58) we conclude as before that

$$\begin{aligned}(1-x)^A(1-y)^B|\mathscr{T}_y f(x)| &\le cg(1-\cos(\theta-\phi))\int_0^1 (1-r^2)^{\alpha/2-1/4}dr\\ &\le cg(1-\cos(\theta-\phi)).\end{aligned}$$

This proves (47) also in this final case. □

We are now in a position to prove Theorem 1.

Proof of Theorem 1. Let $P \in \Pi_n$. Then $K_n^{S-1}P \in \Pi_{S(n-1)+1}$, and hence, the reproducing property of $K_{S(n-1)+1}$ shows that

$$\begin{aligned}\int_{-1}^1 \mathfrak{F}_{S,n}(y)P(y)w_{\alpha,\beta}(y)dy &= \int_{-1}^1 K_{S(n-1)+1}(y)\,(K_n(y)/K_n(1))^{S-1}P(y)w_{\alpha,\beta}(y)dy\\ &= (K_n(1)/K_n(1))^{S-1}P(1) = P(1).\end{aligned} \tag{61}$$

If $P(y) = \sum_{k=0}^{n-1} a_k p_k^{(\alpha,\beta)}(1)p_k^{(\alpha,\beta)}(y)$, then (16) shows that

$$\mathscr{T}_x P(y) = \sum_{k=0}^{n-1} a_k p_k^{(\alpha,\beta)}(x)p_k^{(\alpha,\beta)}(y),$$

so that $\mathscr{T}_x P(1) = P(x)$, and part (a) follows by using (61) with $\mathscr{T}_x P$ in place of P.

It follows from (41) and (40) that $|K_n(y)| \le K_n(1)$. Hence, using Schwarz inequality, the reproducing property of the kernels K_n, and (41) again, we obtain

$$\left\{\int_{-1}^{1} |\mathfrak{F}_{S,n}(y)| w_{\alpha,\beta}(y)dy\right\}^2 \le \left\{\int_{-1}^{1} |K_{S(n-1)+1}(y)|\,|K_n(y)/K_n(1)|\,w_{\alpha,\beta}(y)dy\right\}^2$$
$$\le \left\{\int_{-1}^{1} |K_{S(n-1)+1}(y)|^2 w_{\alpha,\beta}(y)dy\right\}\left\{\int_{-1}^{1} |K_n(y)/K_n(1)|^2\, w_{\alpha,\beta}(y)dy\right\}$$
$$= \frac{K_{S(n-1)+1}(1)}{K_n(1)} \le S^{2\alpha+2}.$$

In view of Proposition 1(b), this leads to (23).

The definition (20) now leads easily to (24) when $p = \infty$ and $p = 1$. The general case follows using the Riesz–Thorin interpolation theorem [15, Theorem XII.1.11]. The first estimate in (25) is clear since $\mathscr{V}_{S,n}(f) \in \Pi_{(2S-1)(n-1)+1}$. To prove the second inequality in (25), let $P \in \Pi_n$ be arbitrary. Then part (a) and (24) imply that

$$\|f - \mathscr{V}_{S,n}(f)\|_p = \|f - P - \mathscr{V}_{S,n}(f-P)\|_p \le \|f-P\|_p + \|\mathscr{V}_{S,n}(f-P)\|_p$$
$$\le (1 + S^{\alpha+1})\|f-P\|_p.$$

Since $P \in \Pi_n$ was arbitrary, this leads to the second inequality in (25), and completes the proof of part (b).

In order to prove part (c), we first estimate $K_n(y)/K_n(1)$. Using (41), (40), and (43), we obtain

$$|K_n(y)| \le cn^{\alpha+1/2}\begin{cases}(1-y)^{-\alpha/2-3/4}, & \text{if } 0 \le y < 1,\\ n^{\beta+1/2}, & \text{if } -1 \le y < 0.\end{cases} \tag{62}$$

Hence, (42) implies that

$$\left|\frac{K_n(y)}{K_n(1)}\right| \le cn^{-\alpha-3/2}\begin{cases}(1-y)^{-\alpha/2-3/4}, & \text{if } 0 \le y < 1,\\ n^{\beta+1/2}, & \text{if } -1 \le y < 0.\end{cases}$$
$$= c\begin{cases}(n^2(1-y))^{-\alpha/2-3/4}, & \text{if } 0 \le y < 1,\\ n^{\beta-\alpha-1}, & \text{if } -1 \le y < 0.\end{cases} \tag{63}$$

Since $\beta \ge -1/2$, $\alpha - \beta + 1 \le \alpha + 3/2$. Also, if $-1 \le y < 0$, then $1 < 1 - y \le 2$. Hence, (63) shows that for all $y \in [-1, 1]$ and $n \ge 1$,

$$\left|\frac{K_n(y)}{K_n(1)}\right| \le \frac{c}{(n^2(1-y))^{(\alpha-\beta+1)/2}}. \tag{64}$$

Next, using (41) and (40), we obtain

$$|K_{S(n-1)+1}(y)| \le cn^{\alpha+1/2}|p_{S(n-1)}^{(\alpha+1,\beta)}(y)| \le cn^{2\alpha+2}. \tag{65}$$

Together with (64), this implies that

$$|\mathfrak{F}_{S,n}(y)| \le cn^{2\alpha+2-(\alpha-\beta+1)(S-1)}(1-y)^{-(a-\beta+1)(S-1)/2}.$$

In view of Theorem 2(a), this leads to (26).

In order to prove (27), we use (41), (40), and (43) to obtain

$$\varpi(y)|K_{(2S-1)(n-1)+1}(y)| \le cn^{\alpha+1/2}\varpi(y)|p_{(2S-1)(n-1)}^{(\alpha+1,\beta)}(y)| \le c\frac{n^{\alpha+3/2}}{n(1-y)^{-1/2}}.$$

In view of (63), this yields

$$\varpi(y)|\mathfrak{F}_{S,n}(y)| \le cn^{\alpha+3/2-(\alpha-\beta+1)(S-1)-1}(1-y)^{-(\alpha-\beta+1)(S-1)/2-1/2}.$$

An application of Theorem 2(b) now leads to (27). □

Proof of Corollary 1. Let ψ be an infinitely differentiable function such that $\psi(x) = 1$ if $x \in \mathbb{B}(x_0, 3\delta/2)$, and $\psi(x) = 0$ if $x \in [-1,1] \setminus \mathbb{B}(x_0, 2\delta)$. If $x = \cos\theta \in \mathbb{B}(x_0, \delta/2)$, and $y = \cos\phi \in [-1,1] \setminus \mathbb{B}(x_0, 3\delta/2)$, then

$$|\phi - \theta| \ge |\phi - \theta_0| - |\theta - \theta_0| \ge \delta/2.$$

Therefore, (26) shows that

$$|\mathscr{T}_x\mathfrak{F}_{S,n}(y)| \le c(\delta)n^{2\alpha+2-(\alpha-\beta+1)(S-1)}.$$

Consequently,

$$\begin{aligned}
&|\mathscr{V}_{S,n}((1-\psi)f)(x)| \\
&= \left|\int_{-1}^{1}(1-\psi(y))f(y)\mathscr{T}_x\mathfrak{F}_{S,n}(y)w_{\alpha,\beta}(y)dy\right| \\
&\le \int_{y\in\mathbb{B}(x,\delta/2)}|f(y)||\mathscr{T}_x\mathfrak{F}_{S,n}(y)|w_{\alpha,\beta}(y)dy \\
&\le c(\delta)n^{2\alpha+2-(\alpha-\beta+1)(S-1)}\|f\|_\infty.
\end{aligned} \tag{66}$$

Further, in view of (24),

$$|\mathscr{V}_{S,n}(\psi f)(x)| \le c\|\psi f\|_\infty \le c \max_{x\in\mathbb{B}(x_0,3\delta/2)}|f(x)|, \tag{67}$$

Since

$$\mathscr{V}_{S,n}(f)(x) = \mathscr{V}_{S,n}(\psi f)(x) + \mathscr{V}_{S,n}((1-\psi)f)(x),$$

the estimates (66), (66) lead to (28). □

Acknowledgements The research of the author is supported in part by Grant W911NF-15-1-0385 from the U.S. Army Research Office. The author thanks Dr. Frank Filbir for many helpful discussions.

References

1. Askey, R.: Orthogonal Polynomials and Special Functions. SIAM, Philadelphia (1975)
2. Filbir, F., Mhaskar, H.N., Prestin, J.: On a filter for exponentially localized kernels based on jacobi polynomials. J. Approx. Theory **160**(1), 256–280 (2009)
3. Filbir, F., Themistoclakis, W.: On the construction of de la vallée poussin means for orthogonal polynomials using convolution structures. J. Comput. Anal. Appl. **6**, 297–312 (2004)
4. Gelb, A., Tadmor, E.: Enhanced spectral viscosity approximations for conservation laws. Appl. Numer. Math. **33**(1), 3–21 (2000)
5. Hylleraas, E.: Linearization of products of jacobi polynomials. Math. Scand. **10**, 189–200 (1962)
6. Koornwinder, T.: Jacobi polynomials, ii. an analytic proof of the product formula. SIAM J. Math. Anal. **5**(1), 125–137 (1974)
7. Mhaskar, H.N.: Polynomial operators and local smoothness classes on the unit interval. J. Approx. Theory **131**(2), 243–267 (2004)
8. Mhaskar, H.N.: Polynomial operators and local smoothness classes on the unit interval, ii. Jaén J. Approx. **1**(1), 1–25 (2009)
9. Mhaskar, H.N., Naumova, V., Pereverzyev, S.V.: Filtered legendre expansion method for numerical differentiation at the boundary point with application to blood glucose predictions. Appl. Math. Comput. **224**, 835–847 (2013)
10. Mhaskar, H.N., Prestin, J.: Polynomial operators for spectral approximation of piecewise analytic functions. Appl. Comput. Harmon. Anal. **26**(1), 121–142 (2009)
11. Nevai, P.G.: Orthogonal Polynomials, vol. 23. American Mathematical Society, Providence (1979)
12. Szegö, G.: Orthogonal Polynomials, vol. 23. Colloquium Publications/American Mathematical Society, Providence (1975)
13. Tadmor, E.: Convergence of spectral methods for nonlinear conservation laws. SIAM J. Numer. Anal. **26**(1), 30–44 (1989)
14. Tadmor, E., Tanner, J.: Adaptive filters for piecewise smooth spectral data. IMA J. Numer. Anal. **25**(4), 635–647 (2005)
15. Zygmund, A.: Trigonometric Series, vol. 1. Cambridge University Press, Cambridge (2002)

Quadrature Rules with Multiple Nodes

Gradimir V. Milovanović and Marija P. Stanić

Abstract In this paper a brief historical survey of the development of quadrature rules with multiple nodes and the maximal algebraic degree of exactness is given. The natural generalization of such rules are quadrature rules with multiple nodes and the maximal degree of exactness in some functional spaces that are different from the space of algebraic polynomial. For that purpose we present a generalized quadrature rules considered by Ghizzeti and Ossicini (Quadrature Formulae, Academie, Berlin, 1970) and apply their ideas in order to obtain quadrature rules with multiple nodes and the maximal trigonometric degree of exactness. Such quadrature rules are characterized by the so-called s- and σ-orthogonal trigonometric polynomials. Numerical method for constructing such quadrature rules is given, as well as a numerical example to illustrate the obtained theoretical results.

Keywords Multiple nodes • Quadrature rules • Degree of exactness • s-orthogonal polynomials • σ-orthogonal polynomials

Introduction and Preliminaries

The most significant discovery in the field of numerical integration in the nineteenth century was made by Carl Friedrich Gauß in 1814. His method [14] dramatically improved the earlier Newton's method. Gauss' method was enriched by significant contributions of Jacobi [21] and Christoffel [5]. More than hundred years after Gauß published his famous method there appeared the idea of numerical integration involving multiple nodes. Taking any system of n distinct points $\{\tau_1, \ldots, \tau_n\}$ and

G.V. Milovanović (✉)
The Serbian Academy of Sciences and Arts, Kneza Mihaila 35, 11000 Belgrade, Serbia
e-mail: gvm@mi.sanu.ac.rs

M.P. Stanić
Faculty of Science, Department of Mathematics and Informatics, University of Kragujevac, Radoja Domanovića 12, 34000 Kragujevac, Serbia
e-mail: stanicm@kg.ac.rs

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_19

n nonnegative integers $m_1, \ldots, m_n$, and starting from the Hermite interpolation formula, Chakalov [2] in 1948 obtained the quadrature formula

$$\int_{-1}^{1} f(t)\,\mathrm{d}t = \sum_{\nu=1}^{n} \big[A_{0,\nu} f(\tau_\nu) + A_{1,\nu} f'(\tau_\nu) + \cdots + A_{m_\nu-1,\nu} f^{(m_\nu-1)}(\tau_\nu)\big], \tag{1}$$

which is exact for all polynomials of degree at most $m_1 + \cdots + m_n - 1$. He gave a method for computing the coefficients $A_{i,\nu}$ in (1). Such coefficients are given by $A_{i,\nu} = \int_{-1}^{1} \ell_{i,\nu}(t)\,\mathrm{d}t$, $\nu = 1, \ldots, n$, $i = 0, 1, \ldots, m_\nu - 1$, where $\ell_{i,\nu}(t)$ are the fundamental functions of Hermite interpolation.

Taking $m_1 = \cdots = m_n = k$ in (1), Turán [48] in 1950 studied numerical quadratures of the form

$$\int_{-1}^{1} f(t)\,\mathrm{d}t = \sum_{i=0}^{k-1} \sum_{\nu=1}^{n} A_{i,\nu} f^{(i)}(\tau_\nu) + R_{n,k}(f). \tag{2}$$

Obviously, for any given nodes $-1 \le \tau_1 \le \cdots \le \tau_n \le 1$ the formula (2) can be made exact for $f \in \mathcal{P}_{kn-1}$ ($\mathcal{P}_m$, $m \in \mathbb{N}_0$, is the set of all algebraic polynomials of degree at most m). However, for $k = 1$ the formula (2), i.e.,

$$\int_{-1}^{1} f(t)\,\mathrm{d}t = \sum_{\nu=1}^{n} A_{0,\nu} f(\tau_\nu) + R_{n,1}(f)$$

is exact for all polynomials of degree at most $2n - 1$ if the nodes τ_ν are the zeros of the Legendre polynomial P_n (it is the well-known Gauss–Legendre quadrature rule). Because of that it is quite natural to consider whether nodes τ_ν can be chosen in such a way that the quadrature formula (2) will be exact for all algebraic polynomials of degree less than or equal to $(k+1)n - 1$. Turán [48] showed that the answer is negative for $k = 2$, while it is positive for $k = 3$. He proved that the nodes τ_ν, $\nu = 1, 2, \ldots, n$, should be chosen as the zeros of the monic polynomial $\pi_n^*(t) = t^n + \cdots$ which minimizes the integral $\int_{-1}^{1} [\pi_n(t)]^4\,\mathrm{d}t$, where $\pi_n(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. In the general case, the answer is negative for even, and positive for odd k, when τ_ν, $\nu = 1, 2, \ldots, n$, must be the zeros of the monic polynomial π_n^* minimizing $\int_{-1}^{1} [\pi_n(t)]^{k+1}\,\mathrm{d}t$. Specially, for $k = 1$, π_n^* is the monic Legendre polynomial $\widehat{P}_n$.

Let us now assume that $k = 2s + 1$, $s \ge 0$. Instead of (2), it is also interesting to consider a more general *Gauss–Turán type* quadrature formula

$$\int_{\mathbb{R}} f(t)\,\mathrm{d}\lambda(t) = \sum_{i=0}^{2s} \sum_{\nu=1}^{n} A_{i,\nu} f^{(i)}(\tau_\nu) + R_{n,2s}(f), \tag{3}$$

where $\mathrm{d}\lambda(t)$ is a given nonnegative measure on the real line $\mathbb{R}$, with compact or unbounded support, for which all moments $\mu_k = \int_{\mathbb{R}} t^k\,\mathrm{d}\lambda(t)$, $k = 0, 1, \ldots$, exist and are finite, and $\mu_0 > 0$. It is known that formula (3) is exact for all polynomials of degree at most $2(s+1)n-1$, i.e., $R_{n,2s}(f) = 0$ for $f \in \mathcal{P}_{2(s+1)n-1}$. The nodes τ_ν, $\nu = 1, \dots, n$, in (3) are the zeros of the monic polynomial $\pi_{n,s}(t)$, which minimizes the integral

$$F(a_0, a_1, \dots, a_{n-1}) = \int_{\mathbb{R}} [\pi_n(t)]^{2s+2} \, \mathrm{d}\lambda(t),$$

where $\pi_n(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. This minimization leads to the conditions

$$\frac{1}{2s+2} \cdot \frac{\partial F}{\partial a_k} = \int_{\mathbb{R}} [\pi_n(t)]^{2s+1} t^k \, \mathrm{d}\lambda(t) = 0, \quad k = 0, 1, \dots, n-1. \tag{4}$$

These polynomials $\pi_n = \pi_{n,s}$ are known as *s-orthogonal* (or *s-self associated*) *polynomials* on $\mathbb{R}$ with respect to the measure $\mathrm{d}\lambda(t)$ (for more details, see [15, 35–37]. For $s = 0$ they reduce to the standard orthogonal polynomials and (3) becomes the well-known Gauss–Christoffel formula.

Using some facts about monosplines, Micchelli [22] investigated the sign of the Cotes coefficients $A_{i,\nu}$ in the Turán quadrature formula.

For the numerical methods for the construction of s-orthogonal polynomials as well as for the construction of Gauss–Turán-type quadrature formulae, we refer readers to the articles [17, 23, 24].

A next natural generalization of the Turán quadrature formula (3) is generalization to rules having nodes with different multiplicities. Such rule for $\mathrm{d}\lambda(t) = \mathrm{d}t$ on (a, b) was derived independently by Chakalov [3, 4] and Popoviciu [38]. Important theoretical progress on this subject was made by Stancu [43, 44] (see also [46]).

Let us assume that the nodes $\tau_1 < \tau_2 < \cdots < \tau_n$, have multiplicities $m_1, m_2, \dots, m_n$, respectively (a permutation of the multiplicities $m_1, m_2, \dots, m_n$, with the nodes held fixed, in general yields to a new quadrature rule).

It can be shown that the quadrature formula (1) is exact for all polynomials of degree less than $2\sum_{\nu=1}^{n}[(m_\nu + 1)/2]$. Thus, the multiplicities m_ν that are even do not contribute toward an increase in the degree of exactness, so that it is reasonable to assume that all m_ν are odd integers, i.e., that $m_\nu = 2s_\nu + 1$, $\nu = 1, 2, \dots, n$. Therefore, for a given sequence of nonnegative integers $\sigma = (s_1, s_2, \dots, s_n)$ the corresponding quadrature formula

$$\int_{\mathbb{R}} f(t) \, \mathrm{d}\lambda(t) = \sum_{\nu=1}^{n} \sum_{i=0}^{2s_\nu} A_{i,\nu} f^{(i)}(\tau_\nu) + R(f)$$

has the maximal degree of exactness $d_{\max} = 2\sum_{\nu=1}^{n} s_\nu + 2n - 1$ *if and only if*

$$\int_{\mathbb{R}} \prod_{\nu=1}^{n} (t-\tau_\nu)^{2s_\nu+1} t^k \, \mathrm{d}\lambda(t) = 0, \quad k = 0, 1, \ldots, n-1. \tag{5}$$

The previous orthogonality conditions correspond to (4) and they could be obtained by the minimization of the integral

$$\int_{\mathbb{R}} \prod_{\nu=1}^{n} (t-\tau_\nu)^{2s_\nu+2} \, \mathrm{d}\lambda(t).$$

The existence of such quadrature rules was proved by Chakalov [4], Popoviciu [38], Morelli and Verna [33], and existence and uniqueness by Ghizzetti and Ossicini [19].

The conditions (5) define a sequence of polynomials $\{\pi_{n,\sigma}\}_{n\in\mathbb{N}_0}$,

$$\pi_{n,\sigma}(t) = \prod_{\nu=1}^{n} (t-\tau_\nu^{(n,\sigma)}), \quad \tau_1^{(n,\sigma)} < \tau_2^{(n,\sigma)} < \cdots < \tau_n^{(n,\sigma)},$$

such that

$$\int_{\mathbb{R}} \pi_{k,\sigma}(t) \prod_{\nu=1}^{n} (t-\tau_\nu^{(n,\sigma)})^{2s_\nu+1} \, \mathrm{d}\lambda(t) = 0, \quad k = 0, 1, \ldots, n-1.$$

These polynomials $\pi_{k,\sigma}$ are called *σ-orthogonal polynomials*, and they correspond to the sequence $\sigma = (s_1, s_2, \ldots)$. Specially, for $s = s_1 = s_2 = \cdots$, the σ-orthogonal polynomials reduce to the s-orthogonal polynomials.

At the end of this section we mention a general problem investigated by Stancu [43, 44, 46]. Namely, let $\eta_1, \ldots, \eta_m$ $(\eta_1 < \cdots < \eta_m)$ be given *fixed* (or *prescribed*) nodes, with multiplicities $\ell_1, \ldots, \ell_m$, respectively, and $\tau_1, \ldots, \tau_n$ $(\tau_1 < \cdots < \tau_n)$ be *free* nodes, with given multiplicities $m_1, \ldots, m_n$, respectively. Stancu considered interpolatory quadrature formulae of a general form

$$I(f) = \int_{\mathbb{R}} f(t) \, \mathrm{d}\lambda(t) \cong \sum_{\nu=1}^{n} \sum_{i=0}^{m_\nu-1} A_{i,\nu} f^{(i)}(\tau_\nu) + \sum_{\nu=1}^{m} \sum_{i=0}^{\ell_\nu-1} B_{i,\nu} f^{(i)}(\eta_\nu), \tag{6}$$

with an algebraic degree of exactness at least $M + L - 1$, where $M = \sum_{\nu=1}^{n} m_\nu$ and $L = \sum_{\nu=1}^{m} \ell_\nu$.

Using free and fixed nodes we can introduce two polynomials

$$Q_M(t) := \prod_{\nu=1}^{n} (t-\tau_\nu)^{m_\nu} \quad \text{and} \quad q_L(t) := \prod_{\nu=1}^{m} (t-\eta_\nu)^{\ell_\nu}.$$

Choosing the free nodes to increase the degree of exactness leads to the so-called Gaussian type of quadratures. If the free (or *Gaussian*) nodes $\tau_1, \ldots, \tau_n$ are such that the quadrature rule (6) is exact for each $f \in \mathcal{P}_{M+L+n-1}$, then we call it the *Gauss–Stancu formula* (see [26]). Stancu [45] proved that $\tau_1, \ldots, \tau_n$ *are the Gaussian nodes if and only if*

$$\int_{\mathbb{R}} t^k Q_M(t) q_L(t)\, \mathrm{d}\lambda(t) = 0, \quad k = 0, 1, \ldots, n-1. \tag{7}$$

Under some restrictions of node polynomials $q_L(t)$ and $Q_M(t)$ on the support interval of the measure $\mathrm{d}\lambda(t)$ we can give sufficient conditions for the existence of Gaussian nodes (cf. Stancu [20, 45]). For example, *if the multiplicities of the Gaussian nodes are odd, e.g., $m_\nu = 2s_\nu + 1$, $\nu = 1, \ldots, n$, and if the polynomial with fixed nodes $q_L(t)$ does not change its sign in the support interval of the measure $\mathrm{d}\lambda(t)$, then, in this interval, there exist real distinct nodes τ_ν, $\nu = 1, \ldots, n$.* This condition for the polynomial $q_L(t)$ means that the multiplicities of the internal fixed nodes must be even. Defining a new (nonnegative) measure $\mathrm{d}\hat{\lambda}(t) := |q_L(t)|\, \mathrm{d}\lambda(t)$, the "orthogonality conditions" (7) can be expressed in a simpler form

$$\int_{\mathbb{R}} t^k Q_M(t)\, \mathrm{d}\hat{\lambda}(t) = 0, \quad k = 0, 1, \ldots, n-1.$$

This means that the general quadrature problem (6), under these conditions, can be reduced to a problem with only Gaussian nodes, but with respect to another modified measure. Computational methods for this purpose are based on Christoffel's theorem and described in detail in [16] (see also [20]).

For more details about the concept of power orthogonality and the corresponding quadrature with multiple nodes and the maximal algebraic degree of exactness and about the numerical methods for their construction we refer readers to the article [25] and the references therein, as well as to a nice book by Shi [41].

The next natural generalization represents quadrature rules with multiple nodes and the maximal degree of exactness in some functional spaces which are different from the space of algebraic polynomials.

A Generalized Gaussian Problem

For a finite interval $[a, b]$ and a positive integer N, Ghizzeti and Osicini in [18] considered quadrature rules for the computation of an integral of the type $\int_a^b f(x)\, w(x)\, \mathrm{d}x$, where the weight function w and function f satisfy the following conditions: $w(x) \in L[a, b]$, $f(x) \in AC^{N-1}[a, b]$ ($AC^m[a, b]$, $m \in \mathbb{N}_0$, is the class of functions f whose mth derivative is absolutely continuous in $[a, b]$).

For a fixed number $m \geq 1$ of points $a \leq x_1 < x_2 < \cdots < x_m \leq b$ and a fixed linear differential operator E of order N of the form

$$E = \sum_{k=0}^{N} a_k(x) \frac{\mathrm{d}^{N-k}}{\mathrm{d}x^{N-k}},$$

where $a_1(x) = 1$, $a_k(x) \in AC^{N-k-1}[a, b]$, $k = 1, 2, \ldots, N-1$, and $a_N(x) \in L(a, b)$, they considered (see [18, pp. 41–43]) a generalized Gaussian problem, i.e., they studied the quadrature rule

$$\int_a^b f(x)\, w(x)\, \mathrm{d}x = \sum_{\nu=1}^{m} \sum_{j=0}^{N-1} A_{j,\nu}, f^{(j)}(x_\nu) + R(f)$$

such that $E(f) = 0$ implies $R(f) = 0$.

According to [18, p. 28] quadrature rule which is exact for all algebraic polynomials of degree less than or equal $\nu - 1$ is relative to the differential operator $E = \frac{\mathrm{d}^\nu}{\mathrm{d}x^\nu}$, and quadrature rule which is exact for all trigonometric polynomials of degree less than or equal to ν is relative to the differential operator (of order $2\nu + 1$): $E = \frac{\mathrm{d}}{\mathrm{d}x} \prod_{k=1}^{\nu} \left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + k^2 \right)$.

They considered whether there can exist a rule of the form

$$\int_a^b f(x) w(x)\, \mathrm{d}x = \sum_{\nu=1}^{m} \sum_{j=0}^{N-p_\nu-1} A_{j,\nu} f^{(j)}(x_\nu) + R(f), \tag{8}$$

with fixed integers p_ν, $0 \leq p_\nu \leq N-1$, $\nu = 1, 2, \ldots, m$, such that at least one of the integers p_ν is greater than or equal to 1, satisfying that $E(f) = 0$ implies $R(f) = 0$, too. The answer is given in the following theorem (see [18, p. 45] for the proof).

Theorem 1. *For the given nodes $x_1, x_2, \ldots, x_m$, the linear differential operator E of order N and the nonnegative integers $p_1, p_2, \ldots, p_m$, $0 \leq p_\nu \leq N-1$, $\nu = 1, 2, \ldots, m$, such that there exists $\nu \in \{1, 2, \ldots, m\}$ for which $p_\nu \geq 1$, consider the following homogenous boundary differential problem*

$$E(f) = 0, \quad f^{(j)}(x_\nu) = 0, \quad j = 0, 1, \ldots, N - p_\nu - 1, \quad \nu = 1, 2, \ldots, m. \tag{9}$$

If this problem has no non-trivial solutions (whence $N \leq mN - \sum_{\nu=1}^{m} p_\nu$) it is possible to write a quadrature rule of the type (8) with $mN - \sum_{\nu=1}^{m} p_\nu - N$ parameters chosen arbitrarily. If, on the other hand, the problem (9) has q linearly independent solutions U_r, $r = 1, 2, \ldots, q$, with $N - mN + \sum_{\nu=1}^{m} p_\nu \leq q \leq p_\nu$ for all $\nu = 1, 2, \ldots, m$, then the formula (8) may apply only if the following q conditions

$$\int_a^b U_r(x)\, w(x)\, \mathrm{d}x = 0, \quad r = 1, 2, \ldots, q,$$

are satisfied; if so, $mN - \sum_{\nu=1}^{m} p_\nu - N + q$ *parameters in the formula* (8) *can be chosen arbitrary.*

Since the quadrature rules with multiple nodes and the maximal algebraic degree of exactness are widespread in literature, in what follows we restrict our attention to the quadrature rules with multiple nodes and the maximal trigonometric degree of exactness.

Quadrature Rules with Multiple Nodes and the Maximal Trigonometric Degree of Exactness

For a nonnegative integer n and for $\gamma \in \{0, 1/2\}$ by $\mathcal{T}_n^\gamma$ we denote the linear span of the set $\{\cos(k+\gamma)x, \sin(k+\gamma)x : k = 0, 1, \ldots, n\}$. It is obvious that $\mathcal{T}_n^0 = \mathcal{T}_n$ is the linear space of all trigonometric polynomials of degree less than or equal to n, $\mathcal{T}_n^{1/2}$ is the linear space of all trigonometric polynomials of semi-integer degree less than or equal to $n + 1/2$ and $\dim(\mathcal{T}_n^\gamma) = 2(n+\gamma) + 1$. By $\mathcal{T}^\gamma = \bigcup_{n\in\mathbb{N}_0} \mathcal{T}_n^\gamma$ we denote the set of all trigonometric polynomials (for $\gamma = 0$) and the set of all trigonometric polynomials of semi-integer degree (for $\gamma = 1/2$). For $\gamma = 0$ we simple write $\mathcal{T}$ instead of $\mathcal{T}^0$. Finally, by $\widetilde{\mathcal{T}}_n$ we denote the linear space $\mathcal{T}_n \ominus \operatorname{span}\{\sin nx\}$ or $\mathcal{T}_n \ominus \operatorname{span}\{\cos nx\}$.

We use the following notation: $\hat{\gamma} = 1 - 2\gamma$, $\gamma \in \{0, 1/2\}$.

In what follows we simple say that trigonometric polynomial has degree $k + \gamma$, $k \in \mathbb{N}_0$, $\gamma \in \{0, 1/2\}$, which always means that for $\gamma = 0$ it is trigonometric polynomial of precise degree k, while for $\gamma = 1/2$ it is trigonometric polynomial of precise semi-integer degree $k + 1/2$.

Obviously, every trigonometric polynomial of degree $n + \gamma$ can be represented in the form

$$T_n^\gamma(x) = \sum_{k=0}^{n} \left(c_k \cos(k+\gamma)x + d_k \sin(k+\gamma)x\right), \quad c_n^2 + d_n^2 \neq 0, \tag{10}$$

where $c_k, d_k \in \mathbb{R}$, $k = 0, 1, \ldots, n$. For $\gamma = 0$ we always set $d_0 = 0$. Coefficients c_n and d_n are called *the leading coefficients*.

Let us suppose that w is a weight function, integrable and nonnegative on the interval $[-\pi, \pi)$, vanishing there only on a set of a measure zero.

The quadrature rules with simple nodes and the maximal trigonometric degree of exactness were considered by many authors. For a brief historical survey of available approaches for the construction of such quadrature rules we refer readers to the article [30] and the references therein. Following Turetzkii's approach (see [49]), which is a simulation of the development of Gaussian quadrature rules for algebraic polynomials, quadrature rules with the maximal trigonometric degree of exactness with an odd number of nodes ($\gamma = 1/2$) and with an even number of nodes ($\gamma = 0$) were considered in [30] (see also [9]) and in [47], respectively.

It is well known that in a case of quadrature rule with maximal algebraic degree of exactness the nodes are the zeros of the corresponding orthogonal algebraic polynomial. In a case of quadrature with an odd maximal trigonometric degree of exactness the nodes are the zeros of the corresponding orthogonal trigonometric polynomials, but for an even maximal trigonometric degree of exactness the nodes are not zeros of trigonometric polynomial, but zeros of the corresponding orthogonal trigonometric polynomial of semi-integer degree. Similarly, when the quadrature rules with multiple nodes and the maximal algebraic degree of exactness are in question it is necessary to consider power orthogonality in the space of algebraic polynomials. But, if we want to consider quadrature rules with multiple nodes and the maximal trigonometric degree of exactness, we have to consider power orthogonality in the space of trigonometric polynomials or in the space of trigonometric polynomials of semi-integer degree, which depends on the number of nodes.

In mentioned papers [30, 47] the existence and the uniqueness of orthogonal trigonometric polynomials of integer or of semi-integer degree were proved under the assumption that two of the coefficients in expansion (10) are given in advance (it is usual to choose the leading coefficients in advance). Also, if we directly compute the zeros of orthogonal trigonometric polynomial (of integer or semi-integer degree), we can fix one of them in advance since for $2(n+\gamma)$ zeros we have $2(n+\gamma)-1$ orthogonality conditions.

Let n be positive integer, $\gamma \in \{0, 1/2\}$, s_ν nonnegative integers for $\nu = \hat{\gamma}, \hat{\gamma}+1, \ldots, 2n$, and $\sigma = (s_{\hat{\gamma}}, s_{\hat{\gamma}+1}, \ldots, s_{2n})$. For the given weight function w we considered a quadrature rule of the following type

$$\int_{-\pi}^{\pi} f(x)\, w(x)\, \mathrm{d}x = \sum_{\nu=\hat{\gamma}}^{2n} \sum_{j=0}^{2s_\nu} A_{j,\nu} f^{(j)}(x_\nu) + R_n(f). \tag{11}$$

which have maximal trigonometric degree of exactness, i.e., for which $R_n(f) = 0$ for all $f \in \mathcal{T}_{N_1}$, where $N_1 = \sum_{\nu=\hat{\gamma}}^{2n}(s_\nu + 1) - 1$. In this case boundary differential problem (9) has the following form

$$E(f) = 0, \quad f^{(j)}(x_\nu) = 0, \quad j = 0, 1, \ldots, 2s_\nu, \quad \nu = \hat{\gamma}, \hat{\gamma}+1, \ldots, 2n, \tag{12}$$

where E is the following differential operator of order $N = 2N_1 + 1$:

$$E = \frac{\mathrm{d}}{\mathrm{d}x} \prod_{k=1}^{N_1} \left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + k^2 \right).$$

The boundary problem (12) has the following $2n - \hat{\gamma}$ linear independent nontrivial solutions

$$U_\ell(x) = \prod_{\nu=\hat{\gamma}}^{2n} \left(\sin \frac{x - x_\nu}{2}\right)^{2s_\nu+1} \cos(\ell+\gamma)x, \quad \ell = 0, 1, \ldots, n-1,$$

$$V_\ell(x) = \prod_{\nu=\hat{\gamma}}^{2n} \left(\sin \frac{x - x_\nu}{2}\right)^{2s_\nu+1} \sin(\ell+\gamma)x, \quad \ell = \hat{\gamma}, \hat{\gamma}+1, \ldots, n-1.$$

According to Theorem 1, the nodes $x_{\hat{\gamma}}, x_{\hat{\gamma}+1}, \ldots, x_{2n}$ of the quadrature rule (11) satisfy conditions

$$\int_{-\pi}^{\pi} \prod_{\nu=\hat{\gamma}}^{2n} \left(\sin \frac{x - x_\nu}{2}\right)^{2s_\nu+1} \cos(\ell+\gamma)x\, w(x)\,\mathrm{d}x = 0, \quad \ell = 0, 1, \ldots, n-1,$$

$$\int_{-\pi}^{\pi} \prod_{\nu=\hat{\gamma}}^{2n} \left(\sin \frac{x - x_\nu}{2}\right)^{2s_\nu+1} \sin(\ell+\gamma)x\, w(x)\,\mathrm{d}x = 0, \quad \ell = \hat{\gamma}, \hat{\gamma}+1, \ldots, n-1,$$

i.e.,

$$\int_{-\pi}^{\pi} \prod_{\nu=\hat{\gamma}}^{2n} \left(\sin \frac{x - x_\nu}{2}\right)^{2s_\nu+1} t_{n-1}^{\gamma}(x)\, w(x)\,\mathrm{d}x = 0, \quad \text{for all } t_{n-1}^{\gamma} \in \mathfrak{T}_{n-1}^{\gamma}. \tag{13}$$

Trigonometric polynomial $T_{\sigma,n}^{\gamma} = \prod_{\nu=\hat{\gamma}}^{2n} \sin(x - x_\nu)/2 \in \mathfrak{T}_n^{\gamma}$ which satisfies orthogonality conditions (13) is called *σ-orthogonal trigonometric polynomial* of degree $n+\gamma$ with respect to weight function w on $[-\pi, \pi)$. Therefore, the nodes of quadrature rule (11) with the maximal trigonometric degree of exactness are the zeros of the corresponding σ-orthogonal trigonometric polynomial.

Specially, for $s_{\hat{\gamma}} = s_{\hat{\gamma}+1} = \cdots = s_{2n} = s$, σ-orthogonality conditions (13) reduce to

$$\int_{-\pi}^{\pi} \left(\prod_{\nu=\hat{\gamma}}^{2n} \sin \frac{x - x_\nu}{2}\right)^{2s+1} t_{n-1}^{\gamma}(x)\, w(x)\,\mathrm{d}x = 0, \quad t_{n-1}^{\gamma} \in \mathfrak{T}_{n-1}^{\gamma}. \tag{14}$$

Trigonometric polynomial $T_{s,n}^{\gamma} = \prod_{\nu=\hat{\gamma}}^{2n} \sin(x - x_\nu)/2 \in \mathfrak{T}_n^{\gamma}$ which satisfies (14) is called *s-orthogonal trigonometric polynomial* of degree $n+\gamma$ with respect to w on $[-\pi, \pi)$.

For $\gamma = 1/2$, s-orthogonal trigonometric polynomials of semi-integer degree were defined and considered in [29], while σ-orthogonal trigonometric polynomials of semi-integer degree were defined and their main properties were proved in [31]. For $\gamma = 0$, the both s- and σ-orthogonal trigonometric polynomials were defined and studied in [47].

It is obvious that s-orthogonal trigonometric polynomials can be considered as a special case of σ-orthogonal trigonometric polynomials. However, we first prove the existence and uniqueness of s-orthogonal trigonometric polynomials because it can be done in a quite simple way, by using the well-known facts about the best approximation given in the Remark 1 below (see [10, pp. 58–60]). The proofs of the existence and uniqueness of σ-orthogonal trigonometric polynomials are more complicated. We prove them by using theory of implicitly defined orthogonality (see [13, Sect. 5.3] for implicitly defined orthogonal algebraic polynomials).

Remark 1. Let X be a Banach space and Y be a closed linear subspace of X. For each $f \in X$, $\inf_{g\in Y} \|f - g\|$ is the error of approximation of f by elements from Y. If there exists some $g = g_0 \in Y$ for which that infimum is attained, then g_0 is called the best approximation to f from Y. For each finite dimensional subspace X_n of X and each $f \in X$, there exists the best approximation to f from X_n. In addition, if X is a strictly convex space, then each $f \in X$ has at most one element of the best approximation in each closed linear subspace $Y \subset X$.

Analogously as it was proved in [30, 47] for orthogonal trigonometric polynomials, one can prove the existence and uniqueness of s- and σ-orthogonal trigonometric polynomials fixing in advance two of their coefficients or one of their zeros. Before we prove the existence and uniqueness of s- and σ-orthogonal trigonometric polynomials, we prove that all of their zeros in $[-\pi, \pi)$ are simple. Of course, it is enough to prove that for σ-orthogonal trigonometric polynomials.

Theorem 2. *The σ-orthogonal trigonometric polynomial $T_{\sigma,n}^{\gamma} \in \mathfrak{T}_n^{\gamma}$ with respect to the weight function $w(x)$ on $[-\pi, \pi)$ has in $[-\pi, \pi)$ exactly $2(n + \gamma)$ distinct simple zeros.*

Proof. It is easy to see that the trigonometric polynomial $T_{\sigma,n}^{\gamma}(x)$ must have at least one zero of odd multiplicity in $[-\pi, \pi)$, because if we assume the contrary, then for $n \in \mathbb{N}$ we obtain that

$$\int_{-\pi}^{\pi} \prod_{\nu=\hat{\gamma}}^{2n} \Big(\sin \frac{x - x_\nu}{2}\Big)^{2s_\nu+1} \cos(0 + \gamma)x\, w(x)\, \mathrm{d}x = 0,$$

which is impossible, because the integrand does not change its sign on $[-\pi, \pi)$. We know (see [6, 7]) that $T_{\sigma,n}^{\gamma}(x)$ must change its sign an even number of times for $\gamma = 0$ and an odd number of times for $\gamma = 1/2$.

First, we suppose that $\gamma = 0$ and that the number of zeros of odd multiplicities of $T_{\sigma,n}(x)$ on $[-\pi, \pi)$ is $2m$, $m < n$. We denote these zeros by $y_1, y_2, \ldots, y_{2m}$, and set $t(x) = \prod_{k=1}^{2m} \sin((x - y_k)/2)$. Since $t \in \mathfrak{T}_m$, $m < n$, we have

$$\int_{-\pi}^{\pi} \prod_{\nu=1}^{2n} \Big(\sin \frac{x - x_\nu}{2}\Big)^{2s_\nu+1} t(x)\, w(x)\, \mathrm{d}x = 0,$$

which also gives a contradiction, since the integrand does not change its sign on $[-\pi, \pi)$. Thus, $T_{\sigma,n}$ must have exactly $2n$ distinct simple zeros on $[-\pi, \pi)$.

Similarly, for $\gamma = 1/2$ by $y_1, y_2, \ldots, y_{2m+1}$, $m < n$, we denote the zeros of odd multiplicities of $T_{\sigma,n}^{1/2}(x)$ on $[-\pi, \pi)$ and set $\hat{t}(x) = \prod_{k=1}^{2m+1} \sin((x - y_k)/2)$. Since $\hat{t} \in \mathcal{T}_m^{1/2}$, $m < n$, we have

$$\int_{-\pi}^{\pi} \prod_{\nu=0}^{2n} \Big(\sin \frac{x - x_\nu}{2}\Big)^{2s_\nu+1}(x)\hat{t}(x)\, w(x)\,\mathrm{d}x = 0,$$

which again gives a contradiction, since the integrand does not change its sign on $[-\pi, \pi)$, which means that $T_{\sigma,n}^{1/2}(x)$ must have exactly $2n + 1$ distinct simple zeros on $[-\pi, \pi)$. □

s-Orthogonal Trigonometric Polynomials

Theorem 3. *Trigonometric polynomial $T_{s,n}^{\gamma}(x) \in \mathcal{T}_n^{\gamma}$, $\gamma \in \{0, 1/2\}$, with given leading coefficients, which is s-orthogonal on $[-\pi, \pi)$ with respect to a given weight function w is determined uniquely.*

Proof. Let us set $X = L^{2s+2}[-\pi, \pi]$,

$$u = w(x)^{1/(2s+2)}(c_{n,\gamma} \cos(n + \gamma)x + d_{n,\gamma} \sin(n + \gamma)x) \in L^{2s+2}[-\pi, \pi],$$

and fix the following $2n - \hat{\gamma}$ linearly independent elements in $L^{2s+2}[-\pi, \pi]$:

$$u_j = w(x)^{1/(2s+2)} \cos(j + \gamma)x, \quad j = 0, 1, \ldots, n - 1,$$
$$v_k = w(x)^{1/(2s+2)} \sin(k + \gamma)x, \quad k = \hat{\gamma}, \hat{\gamma} + 1, \ldots, n - 1.$$

Then $Y = \operatorname{span}\{u_j, v_k : j = 0, 1, \ldots, n - 1, k = \hat{\gamma}, \hat{\gamma} + 1, \ldots, n - 1\}$ is a finite dimensional subspace of X and, according to Remark 1, for each element from X there exists the best approximation from Y, i.e., there exist $2n - \hat{\gamma}$ constants $\alpha_{j,\gamma}$, $j = 0, 1, \ldots, n - 1$, and $\beta_{k,\gamma}$, $k = \hat{\gamma}, \hat{\gamma} + 1, \ldots, n - 1$, such that the error

$$\Big\| u - \Big(\sum_{k=0}^{n-1} \alpha_{k,\gamma} u_k + \sum_{k=\hat{\gamma}}^{n-1} \beta_{k,\gamma} v_k\Big)\Big\| = \Big(\int_{-\pi}^{\pi} \Big(c_{n,\gamma} \cos(n + \gamma)x + d_{n,\gamma} \sin(n + \gamma)x$$
$$- \Big(\sum_{k=0}^{n-1} \alpha_{k,\gamma} \cos(k + \gamma)x + \sum_{k=\hat{\gamma}}^{n-1} \beta_{k,\gamma} \sin(k + \gamma)x\Big)\Big)^{2s+2} w(x)\,\mathrm{d}x\Big)^{1/(2s+2)},$$

is minimal, i.e., for every n and for every choice of the leading coefficients $c_{n,\gamma}$, $d_{n,\gamma}$, $c_{n,\gamma}^2 + d_{n,\gamma}^2 \neq 0$, there exists a trigonometric polynomial of degree $n + \gamma$

$$T_{s,n}^{\gamma}(x) = c_{n,\gamma}\cos(n+\gamma)x + d_{n,\gamma}\sin(n+\gamma)x \\ -\left(\sum_{k=0}^{n-1}\alpha_{k,\gamma}\cos(k+\gamma)x + \sum_{k=\hat{\gamma}}^{n-1}\beta_{k,\gamma}\sin(k+\gamma)x\right),$$

such that

$$\int_{-\pi}^{\pi}(T_{s,n}^{\gamma}(x))^{2s+2}\,w(x)\,\mathrm{d}x$$

is minimal. Since the space $L^{2s+2}[-\pi,\pi]$ is strictly convex, according to Remark 1, the problem of the best approximation has the unique solution, i.e., the trigonometric polynomial $T_{s,n}^{\gamma}$ is unique.

Therefore, for each of the following $2n-\hat{\gamma}$ functions

$$F_k^C(\lambda) = \int_{-\pi}^{\pi}\big(T_{s,n}^{\gamma}(x) + \lambda\cos(k+\gamma)x\big)^{2s+2}\,w(x)\,\mathrm{d}x, \quad k = 0,1,\ldots,n-1,$$

$$F_k^S(\lambda) = \int_{-\pi}^{\pi}\big(T_{s,n}^{\gamma}(x) + \lambda\sin(k+\gamma)x\big)^{2s+2}\,w(x)\,\mathrm{d}x, \quad k = \hat{\gamma},\hat{\gamma}+1,\ldots,n-1,$$

its derivative must be equal to zero for $\lambda = 0$. Hence,

$$\int_{-\pi}^{\pi}\big(T_{s,n}^{\gamma}(x)\big)^{2s+1}\cos(k+\gamma)x\,w(x)\,\mathrm{d}x = 0, \quad k = 0,1,\ldots,n-1,$$

$$\int_{-\pi}^{\pi}\big(T_{s,n}^{\gamma}(x)\big)^{2s+1}\sin(k+\gamma)x\,w(x)\,\mathrm{d}x = 0, \quad k = \hat{\gamma},\hat{\gamma}+1,\ldots,n-1,$$

which means that the polynomial $T_{s,n}^{\gamma}(x)$ satisfies (14). □

σ-Orthogonal Trigonometric Polynomials

In order to prove the existence and uniqueness of σ-orthogonal trigonometric polynomials we fix in advance the zero $x_{\hat{\gamma}} = -\pi$ and consider the σ-orthogonality conditions (13) as the system with unknown zeros.

Theorem 4. *Let w be the weight function on $[-\pi,\pi)$ and let p be a nonnegative continuous function, vanishing only on a set of a measure zero. Then there exists a trigonometric polynomial T_n^{γ}, $\gamma \in \{0,1/2\}$, of degree $n+\gamma$, orthogonal on $[-\pi,\pi)$ to every trigonometric polynomial of degree less than or equal to $n-1+\gamma$ with respect to the weight function $p\big(T_n^{\gamma}(x)\big)w(x)$.*

Proof. By $\widehat{\mathfrak{T}}_n^\gamma$ we denote the set of all trigonometric polynomials of degree $n+\gamma$ which have $2(n+\gamma)$ real distinct zeros $-\pi = x_{\hat\gamma} < x_{\hat\gamma+1} < \cdots < x_{2n} < \pi$ and denote

$$S_{2n-\hat\gamma} = \big\{\mathbf{x} = (x_{\hat\gamma+1}, x_{\hat\gamma+2}, \ldots, x_{2n}) \in \mathbb{R}^{2n-\hat\gamma} : -\pi < x_{\hat\gamma+1} < x_{\hat\gamma+2} < \cdots < x_{2n} < \pi\big\}.$$

For a given function p and an arbitrary $Q_n^\gamma \in \widehat{\mathfrak{T}}_n^\gamma$, we introduce the inner product as follows:

$$\big(f, g\big)_{Q_n^\gamma} = \int_{-\pi}^{\pi} f(x)g(x)\, p(Q_n^\gamma(x))w(x)\,\mathrm{d}x.$$

It is obvious that there is one to one correspondence between the sets $\widehat{\mathfrak{T}}_n^\gamma$ and $S_{2n-\hat\gamma}$. Indeed, for every element $\mathbf{x} = (x_{\hat\gamma+1}, x_{\hat\gamma+2}, \ldots, x_{2n}) \in S_{2n-\hat\gamma}$ and for

$$Q_n^\gamma(x) = \cos\frac{x}{2}\prod_{\nu=\hat\gamma+1}^{2n}\sin\frac{x-x_\nu}{2} \tag{15}$$

there exists a unique system of orthogonal trigonometric polynomials $U_k^\gamma \in \widehat{\mathfrak{T}}_k^\gamma$, $k = 0, 1, \ldots, n$, such that $\big(U_n^\gamma, \cos(k+\gamma)x\big)_{Q_n^\gamma} = \big(U_n^\gamma, \sin(j+\gamma)x\big)_{Q_n^\gamma} = 0$, for all $k = 0, 1, \ldots, n-1, j = \hat\gamma, \hat\gamma+1, \ldots, n-1$, and $\big(U_n^\gamma, U_n^\gamma\big)_{Q_n^\gamma} \neq 0$.

In such a way, we introduce a mapping $F_n : S_{2n-\hat\gamma} \to S_{2n-\hat\gamma}$, defined in the following way: for any $\mathbf{x} = (x_{\hat\gamma+1}, x_{\hat\gamma+2}, \ldots, x_{2n}) \in S_{2n-\hat\gamma}$ we have $F_n(\mathbf{x}) = \mathbf{y}$, where $\mathbf{y} = (y_{\hat\gamma+1}, y_{\hat\gamma+2}, \ldots, y_{2n}) \in S_{2n-\hat\gamma}$ is such that $-\pi, y_{\hat\gamma+1}, y_{\hat\gamma+2}, \ldots, y_{2n}$ are the zeros of the orthogonal trigonometric polynomial of degree $n+\gamma$ with respect to the weight function $p(Q_n^\gamma(x))w(x)$, where $Q_n^\gamma(x)$ is given by (15). It is easy to see that the function $p\big(\cos(x/2)\prod_{\nu=\hat\gamma+1}^{2n}\sin((x-x_\nu)/2)\big)w(x)$ is also an admissible weight function for arbitrary $\mathbf{x} = (x_{\hat\gamma+1}, x_{\hat\gamma+2}, \ldots, x_{2n}) \in \overline{S}_{2n-\hat\gamma}\backslash S_{2n-\hat\gamma}$.

Now, we prove that F_n is continuous mapping on $\overline{S}_{2n-\hat\gamma}$. Let $\mathbf{x} \in \overline{S}_{2n-\hat\gamma}$ be an arbitrary point, $\{\mathbf{x}^{(m)}\}$, $m \in \mathbb{N}$, a convergent sequence of points from $S_{2n-\hat\gamma}$, which converges to $\mathbf{x}$, $\mathbf{y} = F_n(\mathbf{x})$, and $\mathbf{y}^{(m)} = F_n(\mathbf{x}^{(m)})$, $m \in \mathbb{N}$. Let $\mathbf{y}^* \in \overline{S}_{2n-\hat\gamma}$ be an arbitrary limit point of the sequence $\{\mathbf{y}^{(m)}\}$ when $m \to \infty$. Thus,

$$\int_{-\pi}^{\pi}\cos\frac{x}{2}\prod_{\nu=\hat\gamma+1}^{2n}\sin\frac{x-y_\nu^{(m)}}{2}\cos(k+\gamma)x$$
$$\times p\left(\cos\frac{x}{2}\prod_{\nu=\hat\gamma+1}^{2n}\sin\frac{x-x_\nu^{(m)}}{2}\right)w(x)\,\mathrm{d}x = 0, \quad k = 0, 1, \ldots, n-1,$$

$$\int_{-\pi}^{\pi}\cos\frac{x}{2}\prod_{\nu=\hat\gamma+1}^{2n}\sin\frac{x-y_\nu^{(m)}}{2}\sin(j+\gamma)x$$

$$\times p\left(\cos\frac{x}{2}\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-x_\nu^{(m)}}{2}\right)w(x)\,\mathrm{d}x=0,\quad j=\hat{\gamma},\hat{\gamma}+1,2,\ldots,n-1.$$

Applying Lebesgue Theorem of dominant convergence (see [32, p. 83]), when $m\to+\infty$ we obtain

$$\int_{-\pi}^{\pi}\cos\frac{x}{2}\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-y_\nu^*}{2}\cos(k+\gamma)x$$
$$p\left(\cos\frac{x}{2}\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-x_\nu}{2}\right)w(x)\,\mathrm{d}x=0,\quad k=0,1,\ldots,n-1,$$

$$\int_{-\pi}^{\pi}\cos\frac{x}{2}\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-y_\nu^*}{2}\sin(j+\gamma)x$$
$$\times p\left(\cos\frac{x}{2}\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-x_\nu}{2}\right)w(x)\,\mathrm{d}x=0,\quad j=\hat{\gamma},\hat{\gamma}+1,\ldots,n-1,$$

i.e., $\cos(x/2)\prod_{\nu=\hat{\gamma}+1}^{2n}\sin((x-y_\nu^*)/2)$ is the trigonometric polynomial of degree $n+\gamma$ which is orthogonal to all trigonometric polynomials from $\mathcal{T}_{n-1}^{\gamma}$ with respect to the weight function $p\big(\cos(x/2)\prod_{\nu=\hat{\gamma}+1}^{2n}\sin((x-x_\nu)/2)\big)w(x)$ on $[-\pi,\pi)$. Such trigonometric polynomial has $2(n+\gamma)$ distinct simple zeros in $[-\pi,\pi)$ (see [30, 47]). Therefore, $\mathbf{y}^*\in S_{2n-\hat{\gamma}}$ and $\mathbf{y}^*=F_n(\mathbf{x})$. Since $\mathbf{y}=F_n(\mathbf{x})$, because of the uniqueness we have $\mathbf{y}^*=\mathbf{y}$, i.e., the mapping F_n is continuous on $\overline{S}_{2n-\hat{\gamma}}$.

Finally, we prove that the mapping F_n has a fixed point. The mapping F_n is continuous on the bounded, convex and closed set $\overline{S}_{2n-\hat{\gamma}}\subset\mathbb{R}^{2n-\hat{\gamma}}$. Applying the Brouwer fixed point theorem (see [34]) we conclude that there exists a fixed point of F_n. Since $F_n(\mathbf{x})\in S_{2n-\hat{\gamma}}$ for all $\mathbf{x}\in\overline{S}_{2n-\hat{\gamma}}\backslash S_{2n-\hat{\gamma}}$, the fixed point of F_n belongs to $S_{2n-\hat{\gamma}}$.

If we denote the fixed point of F_n by $\mathbf{x}=(x_{\hat{\gamma}+1},x_{\hat{\gamma}+2},\ldots,x_{2n})$, then for $T_n(x)=\cos(x/2)\prod_{\nu=\hat{\gamma}+1}^{2n}\sin((x-x_\nu)/2)$ we get

$$\int_{-\pi}^{\pi}T_n^{\gamma}(x)\cos(k+\gamma)x\,p\big(T_n^{\gamma}(x)\big)w(x)\,\mathrm{d}x=0,\quad k=0,1,\ldots,n-1,$$
$$\int_{-\pi}^{\pi}T_n^{\gamma}(x)\sin(j+\gamma)x\,p\big(T_n^{\gamma}(x)\big)w(x)\,\mathrm{d}x=0,\quad j=\hat{\gamma},\hat{\gamma}+1,\ldots,n-1.\quad\square$$

Construction of Quadrature Rules with Multiple Nodes and the Maximal Trigonometric Degree of Exactness

The construction of quadrature rule with multiple nodes and the maximal trigonometric degree of exactness is based on the application of some properties of the topological degree of mapping (see [39]). For that purposes we use a certain modification of ideas of Bojanov [1], Shi [40], Shi and Xu [42] (for algebraic σ-orthogonal polynomials) and Dryanov [11]. Using a unique notation for $\gamma \in \{0, 1/2\}$ and omitting some details we present here results obtained in [31] and [47] for $\gamma = 1/2$ and $\gamma = 0$, respectively.

For $a \in [0, 1]$, $n \in \mathbb{N}$, and

$$F(\mathbf{x}, a) = \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \prod_{\nu=\hat{\gamma}+1}^{2n} \left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+1} \times \operatorname{sgn}\left(\prod_{\nu=\hat{\gamma}+1}^{2n} \sin\frac{x-x_\nu}{2}\right) t_{n-1}^{\gamma}(x)\, w(x)\, \mathrm{d}x, \quad t_{n-1}^{\gamma} \in \mathcal{T}_{n-1}^{\gamma},$$

we consider the following problem

$$F(\mathbf{x}, a) = 0 \quad \text{for all } t_{n-1}^{\gamma} \in \mathcal{T}_{n-1}^{\gamma}, \tag{16}$$

with unknowns $x_{\hat{\gamma}+1}, x_{\hat{\gamma}+2}, \ldots, x_{2n}$.

The σ-orthogonality conditions (13) (with $x_1 = -\pi$) are equivalent to the problem (16) for $a = 1$, which means that the nodes of the quadrature rule (11) can be obtained as a solution of the problem (16) for $a = 1$.

For $a = 0$ and $a = 1$ the problem (16) has the unique solution in the simplex $S_{2n-\hat{\gamma}}$. According to Theorem 4 we know that the problem (16) has solutions in the simplex $S_{2n-\hat{\gamma}}$ for every $a \in [0, 1]$. We will prove the uniqueness of the solution $\mathbf{x} \in S_{2n-\hat{\gamma}}$ of the problem (16) for all $a \in (0, 1)$.

Let us denote

$$W(\mathbf{x}, a, x) = \prod_{\nu=\hat{\gamma}+1}^{2n} \left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+2}$$

and

$$\phi_k(\mathbf{x}, a) = \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \frac{W(\mathbf{x}, a, x)}{\sin\dfrac{x-x_k}{2}}\, w(x)\, \mathrm{d}x, \tag{17}$$

for $k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n$.

Applying the same arguments as in [11, Lemma 3.2], the following auxiliary result can be easily proved.

Lemma 1. *There exists $\varepsilon > 0$ such that for every $a \in [0,1]$ the solutions $\mathbf{x}$ of the problem (16) belong to the simplex*

$$\overline{S}_{\varepsilon,\gamma} = \left\{\mathbf{y} : \varepsilon \le y_{\hat{\gamma}+1} + \pi, \varepsilon \le y_{\hat{\gamma}+2} - y_{\hat{\gamma}+1}, \dots, \varepsilon \le y_{2n} - y_{2n-1}, \varepsilon \le \pi - y_{2n}\right\}.$$

Also, the following result holds (see [31, Lemma 2.3] and [47, Lemma 4.4] for the proof).

Lemma 2. *The problem (16) and the following problem*

$$\phi_k(\mathbf{x}, a) = 0, \quad k = \hat{\gamma} + 1, \hat{\gamma} + 2, \dots, 2n, \tag{18}$$

where $\phi_k(\mathbf{x}, a)$ are given by (17), are equivalent in the simplex $S_{2n-\hat{\gamma}}$.

It is important to emphasize that the problems (16) and (18) are equivalent in the simplex $\overline{S}_{\varepsilon,\gamma}$, for some $\varepsilon > 0$, as well as in the simplex $\overline{S}_{\varepsilon_1,\gamma}$, for all $0 < \varepsilon_1 < \varepsilon$, but they are not equivalent in $\overline{S}_{2n-\hat{\gamma}}$.

For the proof of the main result we need the following lemma (for the proof see [31, Lemma 2.4]).

Lemma 3. *Let $p_{\xi,\eta}(x)$ be a continuous function on $[-\pi, \pi]$, which depends continuously on parameters $\xi, \eta \in [c, d]$, i.e., if (ξ_m, η_m) approaches (ξ_0, η_0) then the sequences $p_{\xi_m,\eta_m}(x)$ tend to $p_{\xi_0,\eta_0}(x)$ for every fixed x. If the solution $\mathbf{x}(\xi, \eta)$ of the problem (18) with the weight function $p_{\xi,\eta}(x)w(x)$ is always unique for every $(\xi, \eta) \in [c, d]^2$, then the solution $\mathbf{x}(\xi, \eta)$ depends continuously on $(\xi, \eta) \in [c, d]^2$.*

Now, we are ready to prove the main result.

Theorem 5. *The problem (18) has a unique solution in the simplex $S_{2n-\hat{\gamma}}$ for all $a \in [0, 1]$.*

Proof. We call the problem (18) as $(a; s_{\hat{\gamma}+1}, s_{\hat{\gamma}+2}, \dots, s_{2n}; w)$ problem and prove this theorem by mathematical induction on n.

The uniqueness for $n = 0$ is trivial.

As an induction hypothesis, we suppose that the $(a; s_{\hat{\gamma}+1}, s_{\hat{\gamma}+2}, \dots, s_{2n-2}; w)$ problem has a unique solution for every $a \in [0, 1]$ and for every weight function w, integrable and nonnegative on the interval $[-\pi, \pi)$, vanishing there only on a set of a measure zero.

For $(\xi, \eta) \in [-\pi, \pi]^2$ we define the weight functions

$$p_{\xi,\eta}(x) = \left|\sin\frac{x-\xi}{2}\right|^{2as_{2n-1}+2} \left|\sin\frac{x-\eta}{2}\right|^{2as_{2n}+2} w(x).$$

According to the induction hypothesis, the $(a; s_{\hat{\gamma}+1}, s_{\hat{\gamma}+2}, \dots, s_{2n-2}; p_{\xi,\eta})$ problem has a unique solution $(x_{\hat{\gamma}+1}(\xi, \eta), x_{\hat{\gamma}+2}(\xi, \eta), \dots, x_{2n-2}(\xi, \eta))$ for every $a \in [0, 1]$, such that $-\pi < x_{\hat{\gamma}+1}(\xi,\eta) < \cdots < x_{2n-2}(\xi,\eta) < \pi$, and $x_\nu(\xi,\eta)$ for all $\nu = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n-2$ depends continuously on $(\xi,\eta) \in [-\pi,\pi]^2$ (see Lemma 3).

We will prove that the solution of the $(a; s_{\hat{\gamma}+1}, s_{\hat{\gamma}+2}, \ldots, s_{2n}; w)$ problem is unique for every $a \in [0,1]$.

Let us denote $x_{2n-1}(\xi,\eta) = \xi$, $x_{2n}(\xi,\eta) = \eta$,

$$W(\mathbf{x}(\xi,\eta), a, x) = \prod_{\nu=\hat{\gamma}+1}^{2n} \left| \sin \frac{x - x_\nu(\xi,\eta)}{2} \right|^{2as_\nu+2},$$

and

$$\phi_k(\mathbf{x}(\xi,\eta), a) = \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \frac{W(\mathbf{x}(\xi,\eta), a, x)}{\sin \dfrac{x - x_k(\xi,\eta)}{2}}\, w(x)\, \mathrm{d}x, \quad k = \hat{\gamma}+1, \ldots, 2n.$$

The induction hypothesis gives us that

$$\phi_k(\mathbf{x}(\xi,\eta), a) = 0, \quad k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n-2, \tag{19}$$

for $(\xi,\eta) \in D$, where $D = \{(\xi,\eta) : x_{2n-2}(\xi,\eta) < \xi < \eta < \pi\}$.

Let us now consider the following problem in D with unknown $\mathbf{t} = (\xi,\eta)$:

$$\varphi_1(\mathbf{t}, a) = \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \frac{W(\mathbf{x}(\xi,\eta), a, x)}{\sin \dfrac{x-\xi}{2}}\, w(x)\, \mathrm{d}x = 0, \tag{20}$$

$$\varphi_2(\mathbf{t}, a) = \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \frac{W(\mathbf{x}(\xi,\eta), a, x)}{\sin \dfrac{x-\eta}{2}}\, w(x)\, \mathrm{d}x = 0.$$

If $(\xi,\eta) \in D$ is a solution of the problem (20), then, according to (19), $\mathbf{x}(\xi,\eta)$ is a solution of the $(a; s_{\hat{\gamma}+1}, \ldots, s_{2n}; w)$ problem in the simplex $S_{2n-\hat{\gamma}}$. On the contrary, if $\mathbf{x} = (x_{\hat{\gamma}+1}, x_{\hat{\gamma}+2}, \ldots, x_{2n})$ is a solution of the $(a; s_{\hat{\gamma}+1}, \ldots, s_{2n}; w)$ problem in the simplex $S_{2n-\hat{\gamma}}$, then (x_{2n-1}, x_{2n}) is a solution of the problem (20). Applying Lemma 1 we conclude that every solution of the problem (20) belongs to $\overline{D}_\varepsilon = \{(\xi,\eta) : \varepsilon \le \xi - x_{2n-2}(\xi,\eta), \varepsilon \le \eta - \xi, \varepsilon \le \pi - \eta\}$, for some $\varepsilon > 0$.

Let $\mathbf{x}$ be a solution of the problem $(a; s_{\hat{\gamma}+1}, s_{\hat{\gamma}+2}, \ldots, s_{2n}; w)$ in $\overline{S}_{\varepsilon,\gamma}$. Then, differentiating φ_1 with respect to the x_k, for all $k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n-2, 2n$, we have

$$\frac{\partial \varphi_1}{\partial x_k} = -\,(as_k+1) \int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1} \prod_{\nu=\hat{\gamma}+1}^{2n} \left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+1}$$
$$\times \operatorname{sgn}\left(\prod_{\nu=\hat{\gamma}+1}^{2n} \sin\frac{x-x_\nu}{2}\right) \prod_{\substack{\nu=\hat{\gamma}+1 \\ \nu\ne k,\, \nu\ne 2n-1}}^{2n} \sin\frac{x-x_\nu}{2} \cos\frac{x-x_k}{2}\, w(x)\, \mathrm{d}x.$$

Since

$$\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq k,\,\nu\neq 2n-1}}^{2n} \sin\frac{x-x_\nu}{2}\cos\frac{x-x_k}{2} \in \mathcal{T}_{n-1}^{\gamma},$$

applying Lemma 2, we conclude that $\frac{\partial\varphi_1}{\partial x_k} = 0$, for all $k = \hat{\gamma}+1,\ldots,2n-2,2n$. Further, applying elementary trigonometric transformations we get the following identity

$$\cos\frac{x-y}{2} = \cos\frac{x}{2}\cos\frac{y}{2} + \sin\frac{x-y}{2}\cos\frac{y}{2}\sin\frac{y}{2} + \cos\frac{x-y}{2}\sin^2\frac{y}{2},$$

and for

$$I := \frac{\partial\varphi_1}{\partial x_{2n-1}} = -\frac{2as_{2n-1}+1}{2}\int_{-\pi}^{\pi}\left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1}\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq 2n-1}}^{2n}\left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+2}$$
$$\times\left|\sin\frac{x-x_{2n-1}}{2}\right|^{2as_{2n-1}}\cos\frac{x-x_{2n-1}}{2}\,w(x)\,\mathrm{d}x,$$

we obtain

$$\cos\frac{x_{2n-1}}{2} I = -\frac{2as_{2n-1}+1}{2}\left(I_1 + \sin\frac{x_{2n-1}}{2} I_2\right), \tag{21}$$

where

$$I_1 = \int_{-\pi}^{\pi}\left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1}\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq 2n-1}}^{2n}\left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+2}\left|\sin\frac{x-x_{2n-1}}{2}\right|^{2as_{2n-1}}\cos\frac{x}{2}\,w(x)\,\mathrm{d}x,$$

$$I_2 = \int_{-\pi}^{\pi}\left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1}\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq 2n-1}}^{2n}\left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+2}$$
$$\times\left|\sin\frac{x-x_{2n-1}}{2}\right|^{2as_{2n-1}}\sin\frac{x-x_{2n-1}}{2}\,w(x)\,\mathrm{d}x.$$

Obviously, $I_1 > 0$ (the integrand does not change its sign on $[-\pi,\pi)$) and $I_2 = 0$ (because of (20)). Since $-\pi < x_{2n-1} < \pi$, from (21) we get

$$\operatorname{sgn}\left(\frac{\partial\varphi_1}{\partial x_{2n-1}}\right) = -1.$$

Analogously,

$$\frac{\partial \varphi_2}{\partial x_k} = 0, \quad k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n-1, \quad \operatorname{sgn}\left(\frac{\partial \varphi_2}{\partial x_{2n}}\right) = -1.$$

Thus, at any solution $\mathbf{t} = (\xi, \eta)$ of the problem (20) from $\overline{D}_\varepsilon$ we have

$$\frac{\partial \varphi_1}{\partial \xi} = \sum_{\substack{k=\hat{\gamma}+1 \\ k \neq 2n-1}}^{2n} \frac{\partial \varphi_1}{\partial x_k} \cdot \frac{\partial x_k}{\partial \xi} + \frac{\partial \varphi_1}{\partial x_{2n-1}} = \frac{\partial \varphi_1}{\partial x_{2n-1}}, \quad \frac{\partial \varphi_1}{\partial \eta} = 0,$$

$$\frac{\partial \varphi_2}{\partial \eta} = \sum_{k=\hat{\gamma}+1}^{2n-1} \frac{\partial \varphi_2}{\partial x_k} \cdot \frac{\partial x_k}{\partial \eta} + \frac{\partial \varphi_2}{\partial x_{2n}} = \frac{\partial \varphi_2}{\partial x_{2n}}, \quad \frac{\partial \varphi_2}{\partial \xi} = 0.$$

Hence, the determinant of the corresponding Jacobian $J(\mathbf{t})$ is positive.

We are going to prove that the problem (20) has a unique solution in $\overline{D}_{\varepsilon/2}$. We prove it by using the topological degree of a mapping. Let us define the mapping $\boldsymbol{\varphi}(\mathbf{t}, a) : \overline{D}_{\varepsilon/2} \times [0,1] \to \mathbb{R}^2$, by $\boldsymbol{\varphi}(\mathbf{t}, a) = (\varphi_1(\mathbf{t}, a), \varphi_2(\mathbf{t}, a))$, $\mathbf{t} \in \mathbb{R}^2$, $a \in [0,1]$. If $\mathbf{x}(\xi, \eta)$ is a solution of the $(a; s_{\hat{\gamma}+1}, \ldots, s_{2n}; w)$ problem in $\overline{S}_{\varepsilon,\gamma}$, then the solutions of the problem $\boldsymbol{\varphi}(\mathbf{t}, a) = (0,0)$ in $\overline{D}_{\varepsilon/2}$ belong to $\overline{D}_\varepsilon$. The problem $\boldsymbol{\varphi}(\mathbf{t}, a) = (0,0)$ has the unique solution in $\overline{D}_{\varepsilon/2}$ for $a = 0$. It is obvious that $\boldsymbol{\varphi}(\,\cdot\,, 0)$ is differentiable on $D_{\varepsilon/2}$, the mapping $\boldsymbol{\varphi}(\mathbf{t}, a)$ is continuous in $\overline{D}_{\varepsilon/2} \times [0,1]$, and $\boldsymbol{\varphi}(\mathbf{t}, a) \neq (0,0)$ for all $\mathbf{t} \in \partial D_{\varepsilon/2}$ and $a \in [0,1]$. Then $\deg(\boldsymbol{\varphi}(\,\cdot\,, a), D_{\varepsilon/2}, (0,0)) = \operatorname{sgn}(\det(J(\mathbf{t}))) = 1$, for all $a \in [0,1]$, which means that $\deg(\boldsymbol{\varphi}(\,\cdot\,, a), D_{\varepsilon/2}, (0,0))$ is a constant independent of a. Therefore, the problem $\boldsymbol{\varphi}(\mathbf{t}, a) = (0,0)$ has the unique solution in $D_{\varepsilon/2}$ for all $a \in [0,1]$. Hence, the $(a; s_{\hat{\gamma}+1}, \ldots, s_{2n}; w)$ problem has a unique solution in $S_{2n-\hat{\gamma}}$ which belongs to $\overline{S}_{\varepsilon,\gamma}$. □

Theorem 6. *The solution* $\mathbf{x} = \mathbf{x}(a)$ *of the problem* (18) *depends continuously on* $a \in [0,1]$.

Proof. Let $\{a_m\}$, $a_m \in [0,1]$, $m \in \mathbb{N}$, be a convergent sequence, which converges to $a^* \in [0,1]$. Then for every a_m, $m \in \mathbb{N}$, there exists the unique solution $\mathbf{x}(a_m) = (x_{\hat{\gamma}+1}(a_m), \ldots, x_{2n}(a_m))$ of the system (18) with $a = a_m$. The unique solution of system (18) for $a = a^*$ is $\mathbf{x}(a^*) = (x_{\hat{\gamma}+1}(a^*), \ldots, x_{2n}(a^*))$. Let $\mathbf{x}^* = (x^*_{\hat{\gamma}+1}, \ldots, x^*_{2n})$ be an arbitrary limit point of the sequence $\mathbf{x}(a_m)$ when $a_m \to a^*$. According to Theorem 5, for each $m \in \mathbb{N}$ we have

$$\int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2a_m s_{\hat{\gamma}}+1} \frac{W(\mathbf{x}(a_m), a_m, x)}{\sin\dfrac{x - x_k(a_m)}{2}}\, w(x)\, \mathrm{d}x = 0, \quad k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n.$$

When $a_m \to a^*$ the above equations lead to

$$\int_{-\pi}^{\pi} \left(\cos\frac{x}{2}\right)^{2a^* s_{\hat{\gamma}}+1} \frac{W(\mathbf{x}^*, a^*, x)}{\sin\dfrac{x - x_k^*}{2}}\, w(x)\, \mathrm{d}x = 0, \quad k = \hat{\gamma}+1, \hat{\gamma}+2, \ldots, 2n.$$

According to Theorem 5 we get that $\mathbf{x}^* = \mathbf{x}(a^*)$, i.e., $\lim\limits_{a_m \to a^*} \mathbf{x}(a_m) = \mathbf{x}(a^*)$. □

Numerical Method for Construction of Quadrature Rules with Multiple Nodes and the Maximal Trigonometric Degree of Exactness

Based on the previously given theoretical results we present the numerical method for construction of quadrature rules of the form (11), for which $R_n(f) = 0$ for all $f \in \mathcal{T}_{N_1}$, where $N_1 = \sum_{\nu=\hat{\gamma}}^{2n}(s_\nu + 1) - 1$. The first step is a construction of nodes, and the second one is a construction of weights knowing nodes.

Construction of Nodes

We chose $x_{\hat{\gamma}} = -\pi$ and obtain the nodes $x_{\hat{\gamma}+1}, x_{\hat{\gamma}+2}, \cdots, x_{2n}$ by solving system of nonlinear equations (18) for $a = 1$, applying Newton–Kantorovič method. Obtained theoretical results suggest that for fixed n the system (18) can be solved progressively, for an increasing sequence of values for a, up to the target value $a = 1$. The solution for some $a^{(i)}$ can be used as the initial iteration in Newton–Kantorovič method for calculating the solution for $a^{(i+1)}$, $a^{(i)} < a^{(i+1)} \leq 1$. If for some chosen $a^{(i+1)}$ Newton–Kantorovič method does not converge, we decrease $a^{(i+1)}$ such that it becomes convergent, which is always possible according to Theorem 6. Thus, we can set $a^{(i+1)} = 1$ in each step, and, if that iterative process is not convergent, we set $a^{(i+1)} := (a^{(i+1)} + a^{(i)})/2$ until it becomes convergent. As the initial iteration for the first iterative process, for some $a^{(1)} > 0$, we choose the zeros of the corresponding orthogonal trigonometric polynomial of degree $n + \gamma$, i.e., the solution of system (18) for $a = 0$.

Let us introduce the following matrix notation

$$\mathbf{x} = \left[x_{\hat{\gamma}+1}\ x_{\hat{\gamma}+2} \cdots x_{2n}\right]^T,$$

$$\mathbf{x}^{(m)} = \left[x_{\hat{\gamma}+1}^{(m)}\ x_{\hat{\gamma}+2}^{(m)} \cdots x_{2n}^{(m)}\right]^T, \quad m = 0, 1, \ldots,$$

$$\boldsymbol{\phi}(\mathbf{x}) = \left[\phi_{\hat{\gamma}+1}(\mathbf{x})\ \phi_{\hat{\gamma}+2}(\mathbf{x}) \cdots \phi_{2n}(\mathbf{x})\right]^T.$$

Jacobian of $\boldsymbol{\phi}(\mathbf{x})$,

$$\mathbf{W} = \mathbf{W}(\mathbf{x}) = [w_{i,j}]_{(2n-\hat{\gamma})\times(2n-\hat{\gamma})} = \left[\frac{\partial \phi_{i+\hat{\gamma}}}{\partial x_{j+\hat{\gamma}}}\right]_{(2n-\hat{\gamma})\times(2n-\hat{\gamma})},$$

has the following entries

$$\frac{\partial \phi_i}{\partial x_j} = -(1+as_j)\int_{-\pi}^{\pi}\left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1}\prod_{\nu=\hat{\gamma}+1}^{2n}\left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+1}\operatorname{sgn}\left(\prod_{\nu=\hat{\gamma}+1}^{2n}\sin\frac{x-x_\nu}{2}\right)$$
$$\times\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq i,\nu\neq j}}^{2n}\sin\frac{x-x_\nu}{2}\cos\frac{x-x_j}{2}\,w(x)\,\mathrm{d}x,\ \ i\neq j,\ \ i,j=\hat{\gamma}+1,\hat{\gamma}+2,\ldots,2n,$$

$$\frac{\partial \phi_i}{\partial x_i} = -\frac{1+2as_i}{2}\int_{-\pi}^{\pi}\left(\cos\frac{x}{2}\right)^{2as_{\hat{\gamma}}+1}\prod_{\substack{\nu=\hat{\gamma}+1\\ \nu\neq i}}^{2n}\left|\sin\frac{x-x_\nu}{2}\right|^{2as_\nu+2}$$
$$\times\left|\sin\frac{x-x_i}{2}\right|^{2as_i}\cos\frac{x-x_i}{2}\,w(x)\,\mathrm{d}x,\quad i=\hat{\gamma}+1,\hat{\gamma}+2,\ldots,2n.$$

All of the above integrals can be computed by using a Gaussian type quadrature rule for trigonometric polynomials (see [30, 47])

$$\int_{-\pi}^{\pi} f(x)\,w(x)\,\mathrm{d}x = \sum_{\nu=\hat{\gamma}}^{2N} A_\nu f(x_\nu) + R_N(f),$$

with $2N \geq \sum_{\nu=\hat{\gamma}}^{2n} s_\nu + 2n$.

The Newton–Kantorovič method for calculating the zeros of the σ-orthogonal trigonometric polynomial $T_{\sigma,n}^{\gamma}$ is given as follows

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} - \mathbf{W}^{-1}(\mathbf{x}^{(m)})\boldsymbol{\phi}(\mathbf{x}^{(m)}),\quad m=0,1,\ldots,$$

and for sufficiently good chosen initial approximation $\mathbf{x}^{(0)}$, it has the quadratic convergence.

Construction of Weights

Knowing the nodes of quadrature rule (11), it is possible to calculate the corresponding weights. The weights can be calculating by using the Hermite trigonometric interpolation polynomial (see [11, 12]). Since the construction of the Hermite interpolation trigonometric polynomial is more difficult than in the algebraic case, we use an adaptation on the method which was first given in [17] for construction of Gauss–Turán quadrature rules (for algebraic polynomials) and then generalized for Chakalov–Popoviciu's type quadrature rules (also for algebraic polynomials) in [27, 28]. For $\gamma = 1/2$ and $\gamma = 0$ the corresponding method for construction of weights of quadrature rule (11) is presented in [31] and [47], respectively.

Our method is based on the facts that quadrature rule (11) is of interpolatory type and that it is exact for all trigonometric polynomials of degree less than or equal to $\sum_{\nu=\hat{\gamma}}^{2n}(s_\nu+1)-1=\sum_{\nu=\hat{\gamma}}^{2n}s_\nu+2(n+\gamma)-1$. Thus, the weights can be calculated requiring that quadrature rule (11) integrates exactly all trigonometric polynomials of degree less than or equal to $\sum_{\nu=\hat{\gamma}}^{2n}s_\nu+(n+\gamma)-1$. Additionally, in the case $\gamma=0$ (i.e., $\hat{\gamma}=1$) we require that (11) integrates exactly $\cos\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$, when $\sum_{\nu=1}^{2n}(2s_\nu+1)x_\nu=\ell\pi$ for an odd integer ℓ, or $\sin\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$ when $\sum_{\nu=1}^{2n}(2s_\nu+1)x_\nu=\ell\pi$ for an even integer ℓ, while if $\sum_{\nu=1}^{2n}(2s_\nu+1)x_\nu\neq\ell\pi$, $\ell\in\mathbb{Z}$, one can choose to require exactness for $\cos\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$ or $\sin\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$ arbitrary (see [6, 7, 47]). In such a way we obtain a system of linear equations for the unknown weights. That system can be solved by decomposing into a set of $2(n+\gamma)$ upper triangular systems. For $\gamma=0$ we explain that decomposing method in the case when $\sum_{\nu=1}^{2n}(2s_\nu+1)x_\nu=\ell\pi$, $\ell\in\mathbb{Z}$.

Let us denote

$$\Omega_\nu(x)=\prod_{\substack{i=\hat{\gamma}\\ i\neq\nu}}^{2n}\left(\sin\frac{x-x_i}{2}\right)^{2s_i+1},\quad \nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n,$$

$$u_{k,\nu}(x)=\left(\sin\frac{x-x_\nu}{2}\right)^{k}\Omega_\nu(x),\qquad \nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n,\ k=0,1,\ldots,2s_\nu,$$

and

$$t_{k,\nu}(x)=\begin{cases}u_{k,\nu}(x)\cos\dfrac{x-x_\nu}{2}, & k\text{-even},\ \gamma=0,\ \text{or}\ k\text{-odd},\ \gamma=1/2,\\ u_{k,\nu}(x), & k\text{-odd},\ \gamma=0,\ \text{or}\ k\text{-even},\ \gamma=1/2,\end{cases}\tag{22}$$

for all $\nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n$, $k=0,1,\ldots,2s_\nu$. It is easy to see that $t_{k,\nu}$ is a trigonometric polynomial of degree less than or equal to $\sum_{\nu=\hat{\gamma}}^{2n}s_\nu+n$, which for $\gamma=0$ has the leading term $\cos\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$ or $\sin\big(\sum_{\nu=1}^{2n}s_\nu+n\big)x$, but not the both of them. Quadrature rule (11) is exact for all such trigonometric polynomials, i.e., $R_n(t_{k,\nu})=0$ for all $\nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n$, $k=0,1,\ldots,2s_\nu$.

By using the Leibniz differentiation formula it is easy to check that $t_{k,\nu}^{(j)}(x_i)=0$, $j=0,1,\ldots,2s_\nu$, for all $i\neq\nu$. Hence, for all $k=0,1,\ldots,2s_\nu$ we have

$$\mu_{k,\nu}=\int_{-\pi}^{\pi}t_{k,\nu}(x)\,w(x)\,\mathrm{d}x=\sum_{j=0}^{2s_\nu}A_{j,\nu}t_{k,\nu}^{(j)}(x_\nu),\qquad \nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n.\tag{23}$$

In such a way, we obtain $2(n+\gamma)$ independent systems for calculating the weights $A_{j,\nu}$, $j=0,1,\ldots,2s_\nu$, $\nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n$. Here, we need to calculate the derivatives $t_{k,\nu}^{(j)}(x_\nu)$, $k=0,1,\ldots,2s_\nu$, $j=0,1,\ldots,2s_\nu$, for each $\nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n$. For that purpose we use the following result (see [31] for details).

Lemma 4. *For the trigonometric polynomials $t_{k,\nu}$, given by (22) we have*

$$t_{k,\nu}^{(i)}(x_\nu)=0,\quad i<k;\quad t_{k,\nu}^{(k)}(x_\nu)=\frac{k!}{2^k}\Omega_\nu(x_\nu),$$

and for $i>k$

$$t_{k,\nu}^{(i)}(x_\nu)=\begin{cases} v_{k,\nu}^{(i)}(x_\nu), & k\text{-even}, \gamma=0, \text{ or } k\text{-odd}, \gamma=1/2,\\ u_{k,\nu}^{(i)}(x_\nu), & k\text{-odd}, \gamma=0, \text{ or } k\text{-even}, \gamma=1/2,\end{cases}$$

where

$$v_{k,\nu}^{(i)}(x_\nu)=\sum_{m=0}^{[i/2]}\binom{i}{2m}\frac{(-1)^m}{2^{2m}}u_{k,\nu}^{(i-2m)}(x_\nu),$$

and the sequence $u_{k,\nu}^{(i)}(x_\nu)$, $k\in\mathbb{N}_0$, $i\in\mathbb{N}_0$, is the solution of the difference equation

$$f_{k,\nu}^{(i)}=\sum_{m=[(i-k)/2]}^{[(i-1)/2]}\binom{i}{2m+1}\frac{(-1)^m}{2^{2m+1}}f_{k-1,\nu}^{(i-2m-1)}(x_\nu),\ \nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n,\ k\in\mathbb{N},\ i\in\mathbb{N}_0,$$

with the initial conditions $f_{0,\nu}^{(i)}=\Omega_\nu^{(i)}(x_\nu)$.

Finally, the remaining problem of calculating $\Omega_\nu^{(i)}(x_\nu)$, $\nu=\hat{\gamma},\hat{\gamma}+1,\ldots,2n$, is solved in [31, Lemma 3.3].

Lemma 5. *Let us denote*

$$\Omega_{\nu,k}(x)=\prod_{\substack{i\le k\\ i\ne\nu}}\left(\sin\frac{x-x_i}{2}\right)^{2s_i+1}.$$

Then $\Omega_{\nu,\nu-1}(x)=\Omega_{\nu,\nu}(x)$ and $\Omega_\nu(x)=\Omega_{\nu,2n}(x)$. The sequence $\Omega_{\nu,k,\ell}^{(i)}(x)$, $\ell,i\in\mathbb{N}_0$, where

$$\Omega_{\nu,k,\ell}(x)=\left(\sin\frac{x-x_{k+1}}{2}\right)^\ell\Omega_{\nu,k}(x),\quad k=\hat{\gamma},\hat{\gamma}+1,\ldots,2n-1,\ k\ne\nu-1,$$

is the solution of the following difference equation

$$f_{\nu,k,\ell}^{(i)}(x)=\sum_{m=0}^{i}\binom{i}{m}\frac{1}{2^m}\sin\left(\frac{x-x_{k+1}}{2}+\frac{m\pi}{2}\right)f_{\nu,k,\ell-1}^{(i-m)}(x),\quad \ell\in\mathbb{N}_0,$$

with the initial conditions $f^{(i)}_{\nu,k,0} = \Omega^{(i)}_{\nu,k}$, $i \in \mathbb{N}_0$. *The following equalities*

$$\begin{aligned} \Omega^{(i)}_{\nu,k,2s_{k+1}+1}(x) &= \Omega^{(i)}_{\nu,k+1,0}(x), \quad k \neq \nu - 1,\\ \Omega^{(i)}_{\nu}(x) &= \Omega^{(i)}_{\nu,2n}(x) = \Omega^{(i)}_{\nu,2n-1,2s_{2n}+1}(x), \quad \nu \neq 2n,\\ \Omega^{(i)}_{2n}(x) &= \Omega^{(i)}_{2n,2n-1}(x) = \Omega^{(i)}_{2n,2n-2,2s_{2n-1}+1}(x) \end{aligned}$$

hold.

According to Lemma 4, the systems (23) are the following upper triangular systems

$$\begin{bmatrix} t_{0,\nu}(x_\nu) & t'_{0,\nu}(x_\nu) & \cdots & t^{(2s_\nu)}_{0,\nu}(x_\nu) \\ & t'_{1,\nu}(x_\nu) & \cdots & t^{(2s_\nu)}_{1,\nu}(x_\nu) \\ & & \ddots & \vdots \\ & & & t^{(2s_\nu)}_{2s_\nu,\nu}(x_\nu) \end{bmatrix} \cdot \begin{bmatrix} A_{0,\nu} \\ A_{1,\nu} \\ \vdots \\ A_{2s_\nu,\nu} \end{bmatrix} = \begin{bmatrix} \mu_{0,\nu} \\ \mu_{1,\nu} \\ \vdots \\ \mu_{2s_\nu,\nu} \end{bmatrix}, \quad \nu = \hat{\gamma}, \hat{\gamma}+1, \ldots, 2n.$$

Remark 2. For the case $\gamma = 0$, if $\sum_{\nu=1}^{2n}(2s_\nu + 1)x_\nu \neq \ell\pi$, $\ell \in \mathbb{Z}$, then we choose $t_{k,\nu}$, $k = 0, 1, \ldots, 2s_\nu$, $\nu = 1, 2, \ldots, 2n$, in one of the following ways

$$t_{k,\nu}(x) = u_{k,\nu}(x) \cos \frac{x + \sum\limits_{\substack{i=1\\ i\neq\nu}}^{2n}(2s_i + 1)x_i + 2s_\nu x_\nu}{2}$$

or

$$t_{k,\nu}(x) = u_{k,\nu}(x) \sin \frac{x + \sum\limits_{\substack{i=1\\ i\neq\nu}}^{2n}(2s_i + 1)x_i + 2s_\nu x_\nu}{2},$$

which provides that

$$t_{k,\nu} \in \mathcal{T}_{\sum_{\nu=1}^{2n} s_\nu + n} \ominus \operatorname{span}\left\{\cos\left(\sum_{\nu=1}^{2n} s_\nu + n\right)x\right\}$$

or

$$t_{k,\nu} \in \mathcal{T}_{\sum_{\nu=1}^{2n} s_\nu + n} \ominus \operatorname{span}\left\{\sin\left(\sum_{\nu=1}^{2n} s_\nu + n\right)x\right\},$$

$k = 0, 1, \ldots, 2s_\nu$, $\nu = 1, 2, \ldots, 2n$, respectively.

Numerical Example

In this section we present one numerical example as illustration of the obtained theoretical results. Numerical results are obtained using our proposed progressive approach. For all computations we use MATHEMATICA and the software package *OrthogonalPolynomials* explained in [8] in double precision arithmetic.

Example 1. We construct quadrature rule (11) for $\gamma = 0$, $n = 3$, $\sigma = (3, 3, 3, 4, 4, 4)$, with respect to the weight function

$$w(x) = 1 + \cos 2x, \quad x \in [-\pi, \pi).$$

For calculation of nodes we use proposed progressive approach with only two steps, i.e., two iterative processes: for $a = 1/2$ (with 9 iterations) and for $a = 1$ (with 8 iterations). The nodes x_ν, $\nu = 1, 2, \ldots, 6$, are the following:

$$-3.141592653589793, \; -2.264556388673865, \; -1.179320242581565,$$
$$-0.1955027724705077, \; 0.8612188670819011, \; 2.178685249095223.$$

The corresponding weight coefficients $A_{j,\nu}$, $j = 0, 1, \ldots, 2s_\nu$, $\nu = 1, 2, \ldots, 6$, are given in Table 1 (numbers in parentheses denote decimal exponents).

Table 1 Weight coefficients $A_{j,\nu}, j = 0, 1, \ldots, 2s_\nu$, $\nu = 1, 2, \ldots, 2n$, for $w(x) = 1 + \cos 2x$, $x \in [-\pi, \pi)$, $n = 3$ and $\sigma = (3, 3, 3, 4, 4, 4)$

j	$A_{j,1}$	$A_{j,2}$	$A_{j,3}$
0	1.676594372192496	7.305332409592605(−1)	3.487838013502532(−1)
1	−3.606295932640399(−3)	−8.705060748567067(−2)	6.802235520408829(−2)
2	4.450266475268382(−2)	1.929684165460602(−2)	1.120920249826093(−2)
3	−6.813768854008439(−2)	−9.825070962211830(−4)	8.081866652580276(−4)
4	2.575834718221877(−4)	1.050037121874394(−4)	6.594073161399015(−5)
5	−2.025745530496365(−7)	−2.102329131299247(−6)	1.808851807084914(−6)
6	3.780621226622374(−7)	1.413143818423462(−7)	9.154154335491590(−8)
j	$A_{j,4}$	$A_{j,5}$	$A_{j,6}$
0	1.880646586862865	9.170128528217915(−1)	7.296144529929201(−1)
1	6.328640684710125(−2)	−1.557489854117211(−1)	1.399304206896031(−1)
2	7.316493158593520(−2)	3.683058320865013(−2)	2.916174720939193(−2)
3	1.252054136886738(−3)	−2.974661584717919(−3)	2.588432069654582(−3)
4	7.029326017474412(−4)	3.381006758789474(−4)	2.606855435340879(−4)
5	6.219410600844099(−6)	−1.428080829445919(−5)	1.209621855269665(−5)
6	2.280678345497009(−6)	1.008832884300956(−6)	7.540115736887533(−7)
7	8.397304312672816(−9)	−1.865855244589502(−8)	1.546523556082379(−8)
8	2.276058254314503(−9)	8.981189180630658(−10)	6.501031023479647(−10)

Acknowledgements The authors were supported in part by the Serbian Ministry of Education, Science and Technological Development (grant number #174015).

References

1. Bojanov, B.D.: Oscillating polynomials of least L_1-norm. In: Hämmerlin, G. (ed.) Numerical Integration. ISNM, vol. 57, pp. 25–33. Birkhäuser, Basel (1982)
2. Chakalov, L.: Über eine allgemeine Quadraturformel. C.R. Acad. Bulg. Sci. **1**, 9–12 (1948)
3. Chakalov, L.: Formules générales de quadrature mécanique du type de Gauss. Colloq. Math. **5**, 69–73 (1957)
4. Chakalov, L.: General quadrature formulae of Gaussian type. Bulgar. Akad. Nauk Izv. Mat. Inst. **1**, 67–84 (1954) (Bulgarian) [English transl. East J. Approx. **1**, 261–276 (1995)]
5. Christoffel, E.B.: Über die Gaußische Quadratur und eine Verallgemeinerung derselben. J. Reine Angew. Math. **55**, 61–82 (1858) [Also in Ges. Math. Abhandlungen I, pp. 42–50.]
6. Cruz–Barroso, R., Darius, L., Gonzáles–Vera, P., Njåstad, O.: Quadrature rules for periodic integrands. Bi–orthogonality and para–orthogonality. Ann. Math. et Informancae. **32**, 5–44 (2005)
7. Cruz-Barroso, R., Gonzáles–Vera, P., Njåstad, O.: On bi-orthogonal systems of trigonometric functions and quadrature formulas for periodic integrands. Numer. Algor. **44**(4), 309–333 (2007)
8. Cvetković, A.S., Milovanović, G.V.: The mathematica package "OrthogonalPolynomials". Facta Univ. Ser. Math. Inform. **19**, 17–36 (2004)
9. Cvetković, A.S., Stanić, M.P.: Trigonometric orthogonal systems. In: Gautschi, W., Mastroianni, G., Rassias, Th.M. (eds.) Approximation and Computation – In Honor of Gradimir V. Milovanović. Springer Optimization and Its Applications, vol. 42, pp. 103–116. Springer, Berlin–Heidelberg–New York (2011)
10. DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Springer, Berlin, Heildeberg (1993)
11. Dryanov, D.P.: Quadrature formulae with free nodes for periodic functions. Numer. Math. **67**, 441–464 (1994)
12. Du, J., Han, H., Jin, G.: On trigonometric and paratrigonometric Hermite interpolation. J. Approx. Theory **131**, 74–99 (2004)
13. Engels, H.: Numerical Quadrature and Qubature. Academic, London (1980)
14. Gauss, C.F.: Methodus nova integralium valores per approximationem inveniendi. Commentationes Societatis Regiae Scientarium Göttingensis Recentiores **3**, 39–76 (1814) [Also in Werke III, pp. 163–196.]
15. Gautschi, W.: A survey of Gauss–Christoffel quadrature formulae. In: Butzer, P.L., Fehér, F., Christoffel, E.B. (eds.) The Influence of His Work on Mathematics and the Physical Sciences, pp. 72–147. Birkhäuser, Basel (1981)
16. Gautschi, W.: Orthogonal polynomials: applications and computation. Acta Numer. **5**, 45–119 (1996)
17. Gautschi, W., Milovanović, G.V.: S-orthogonality and construction of Gauss–Turán-type quadrature formulae. J. Comput. Appl. Math. **86**, 205–218 (1997)
18. Ghizzeti, A., Ossicini, A.: Quadrature Formulae. Academie, Berlin (1970)
19. Ghizzetti, A., Ossicini, A.: Sull' esistenza e unicità delle formule di quadratura gaussiane. Rend. Mat. **8**(6), 1–15 (1975)
20. Golub, G.H., Kautsky, J.: Calculation of Gauss quadratures with multiple free and fixed knots. Numer. Math. **41**, 147–163 (1983)
21. Jacobi, C.G.J.: Ueber Gaußs neue Methode, die Werthe der Integrale näherungsweise zu finden. J. Reine Angew. Math. **1**, 301–308 (1826)

22. Micchelli, C.A.: The fundamental theorem of algebra for monosplines with multiplicities. In: Butzer, P., Kahane, J.P., Nagy, B.Sz. (eds.) Linear Operators and Approximation. ISNM, vol. 20, pp. 372–379. Birkhäuser, Basel (1972)
23. Milovanović, G.V.: Construction of *s*–orthogonal polynomials and Turán quadrature formulae. In: Milovanović, G.V. (ed.) Numerical Methods and Approximation Theory III (Niš, 1987), pp. 311–328. University of Niš, Niš (1988)
24. Milovanović, G.V.: *S*–orthogonality and generalized Turán quadratures: construction and applications. In: Stancu, D.D., Coman, Ch., Breckner, W.W., Blaga, P. (Eds.), Approximation and Optimization (Cluj-Napoca, 1996). vol. I, pp. 91–106, Transilvania Press, Cluj-Napoca (1997)
25. Milovanović, G.V.: Quadrature with multiple nodes, power orthogonality, and moment–preserving spline approximation. J. Comput. Appl. Math. **127**, 267–286 (2001). (Quadrature and orthogonal polynomials. In: Gautschi, W., Marcellan, F., Reichel, L. (eds.) Numerical analysis 2000, vol. V)
26. Milovanović, G.V.: Numerical quadratures and orthogonal polynomials. Stud. Univ. Babeş-Bolyai, Math. **56**, 449–464 (2011)
27. Milovanović, G.V., Spalević, M.M.: Construction of Chakalov–Popoviciu's type quadrature formulae. Rend. Circ. Mat. Palermo, Serie II, Suppl. **52**, 625–636 (1998)
28. Milovanović, G.V., Spalević, M.M., Cvetković, A.S.: Calculation of Gaussian type quadratures with multiple nodes. Math. Comput. Model. **39**, 325–347 (2004)
29. Milovanović, G.V., Cvetković, A.S., Stanić, M.P.: Trigonometric orthogonal systems and quadrature formulae with maximal trigonometric degree of exactness. Numerical Methods and Applications, vol. 2006. Lecture Notes in Computer Science, vol. 4310, pp. 402–409 (2007)
30. Milovanović, G.V., Cvetković, A.S., Stanić, M.P.: Trigonometric orthogonal systems and quadrature formulae. Comput. Math. Appl. **56**(11), 2915–2931 (2008)
31. Milovanović, G.V., Cvetković, A.S., Stanić, M.P.: Quadrature formulae with multiple nodes and a maximal trigonometric degree of exactness. Numer. Math. **112**, 425–448 (2009)
32. Mirković, B.: Theory of Measures and Integrals. Naučna knjiga, Beograd, 1990 (in Serbian)
33. Morelli, A., Verna, I.: Formula di quadratura in cui compaiono i valori della funzione e delle derivate con ordine massimo variabile da nodo a nodo. Rend. Circ. Mat. Palermo **18**(2), 91–98 (1969)
34. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables. In: Classics in Applied Mathematics, vol. 30. SIAM, Philadelphia (2000)
35. Ossicini, A: Costruzione di formule di quadratura di tipo Gaussiano. Ann. Mat. Pura Appl. **72**(4), 213–237 (1966)
36. Ossicini, A., Rosati, F.: Funzioni caratteristiche nelle formule di quadratura gaussiane con nodi multipli. Boll. Un. Mat. Ital. **11**(4), 224–237 (1975)
37. Ossicini, A., Rosati, F.: Sulla convergenza dei funzionali ipergaussiani. Rend. Mat. **11**(6), 97–108 (1978)
38. Popoviciu, T. : Sur une généralisation de la formule d'intégration numérique de Gauss. Acad. R. P. Romîne Fil. Iaşi Stud. Cerc. Şti. **6**, 29–57 (1955) (Romanian)
39. Schwartz, J.T.: Nonlinear Functional Analysis. Gordon and Breach, New York (1969)
40. Shi, Y.G.: A kind of extremal problem of integration on an arbitrary measure. Acta Sci. Math. (Szeged) **65**, 567–575 (1999)
41. Shi, Y.G.: Power Orthogonal Polynomials. Nova Science Publishers, New York (2006)
42. Shi, Y.G., Xu, G.: Construction of σ-orthogonal polynomials and Gaussian quadrature formulas. Adv. Comput. Math. **27**(1), 79–94 (2007)
43. Stancu, D.D.: On a class of orthogonal polynomials and on some general quadrature formulas with minimum number of terms. Bull. Math. Soc. Sci. Math. Phys. R. P. Romîne (N.S) **1**(49), 479–498 (1957)
44. Stancu, D.D.: On certain general numerical integration formulas. Acad. R. P. Romîne. Stud. Cerc. Mat. **9**, 209–216 (1958) (Romanian)
45. Stancu, D.D.: Sur quelques formules générales de quadrature du type Gauss-Christoffel. Mathematica (Cluj) **1**(24), 167–182 (1959)

46. Stroud, A.H., Stancu, D.D.: Quadrature formulas with multiple Gaussian nodes. J. SIAM Numer. Anal. Ser. B **2**, 129–143 (1965)
47. Tomović, T.V., Stanić, M.P.: Quadrature rules with an even number of multiple nodes and a maximal trigonometric degree of exactness. FILOMAT **29**(10), 2239–2255 (2015)
48. Turán, P.: On the theory of the mechanical quadrature. Acta Sci. Math. Szeged **12**, 30–37 (1950)
49. Turetzkii, A.H.: On quadrature rule that are exact for trigonometric polynomials. East J. Approx. 11, 337–359 (2005). (Translation in English from Uchenye Zapiski, Vypusk 1(149), Seria Math. Theory of Functions, Collection of papers, Izdatel'stvo Belgosuniversiteta imeni V.I. Lenina, Minsk (1959) 31–54)

$\mathscr{A}$-Summability of Sequences of Linear Conservative Operators

Daniel Cárdenas-Morales and Pedro Garrancho

Abstract This work deals with the approximation of functions by sequences of linear operators. Here the classical convergence is replaced by matrix summability. Beyond the usual positivity of the operators involved in the approximation processes, more general conservative approximation properties are considered. Quantitative results, as well as results on asymptotic formulae and saturation are stated. It is the intention of the authors to show the way in which some concepts of generalized convergence entered Korovkin-type approximation theory. This is a survey work that gathers and orders the results stated by the authors and other researchers within the aforesaid subject.

Keywords Linear operator • Matrix summability method • Conservative approximation • Asymptotic formula • Saturation

Introduction

Given a sequence of infinite regular matrices, denoted by $\mathscr{A}$, our concern with this work is to show the way in which $\mathscr{A}$-summability theory entered the so-called Korovkin-type approximation theory, and then show as well how this matter was investigated through the classical topics that one encounters on studying approximation processes by sequences of linear operators, as those the aforesaid theory deals with. Special emphasis is made on the shape approximation properties of the operators. This is a survey work that gathers and orders some results stated mainly by the authors.

In the rest of this section we first review some classical results on (conservative) approximation by linear operators, and then present the notion of $\mathscr{A}$-summability and how it was introduced in approximation theory. In the way, we anticipate the contents of the other three sections of this work, devoted respectively to

D. Cárdenas-Morales (✉) • P. Garrancho
Department of Mathematics, University of Jaén, Campus Las Lagunillas s/n 23071 Jaén, Spain
e-mail: cardenas@ujaen.es; pgarran@ujaen.es

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_20

quantitative results, asymptotic formulae and saturation results, via $\mathscr{A}$-summability, and considering conservative approximation processes.

Basics on Korovkin-Type (Conservative) Approximation

Korovkin-type theory treats the problem of approximating a function f by using sequences of functions L_nf defined from a sequence L_n of linear operators. Let us recall that given an operator L, $Lf = L(f)$ denotes the image of f by L; an operator L between two function spaces is linear if given any two functions f and g in the domain, then $\alpha f + \beta g$ belongs to the domain as well for any $\alpha, \beta \in \mathbb{R}$, and $L(\alpha f + \beta g) = \alpha L(f) + \beta L(g)$; the operator L is said to be positive if $f \geq 0$ implies $Lf \geq 0$.

The foundations of this branch of approximation theory were laid in the 1950s by Popoviciu, Bohmann and finally by Korovkin who gave the name to the theory probably after the publication of his book [16]. The theory found much further development since then (see [3]).

The most basic statement reads as follows. Here and in the sequel $\|\cdot\|$ denotes the uniform norm, $e_i(t) = t^i$ and as usual $C^k(X)$ stands for the space of all k-times continuously differentiable functions defined on the set X; $C^0(X) = C(X)$.

Theorem 1. *Let X be a compact real interval and let K_n be a sequence of positive linear operators defined on $C(X)$. If $\|K_ne_i - e_i\| \to 0$ for all $i \in \{0, 1, 2\}$, then*

$$\|K_nf - f\| \to 0 \text{ for each function } f \in C(X).$$

The list of variations and extensions of this result is endless. According to our aim for this work, we shall focus our attention on the more general setting where the element to be approximated is Lf instead of f, L being a linear operator, and where the hypotheses of positivity of the operators is extended and replaced by a conservative property related to a cone a functions, that is to say, a set of functions closed under multiplication by non-negative real numbers. To fix ideas the reader may think of L as D^k, i.e. the classical kth differential operator of order k, and think of the conservative property as if it were k-convexity. Recall that an operator is said to be k-convex if it maps k-convex functions onto k-convex functions, and recall also that a function f, k-times differentiable on certain interval, is said to be k-convex on that interval if D^kf is non-negative there. This situation naturally takes us to the setting of the so-called simultaneous approximation. In this point, we just refer the reader to the paper [22] for a general Korovkin-type result on conservative approximation.

Once an approximation process is proved to be convergent, the next step consists of studying the rate of convergence, leading this way to quantitative versions of the first results. Here the basic important tool is represented by the modulus of continuity of a function, defined for $f \in C(X)$ by

$$\omega(f,\delta) = sup\{|f(x) - f(y)| : x, y \in X, |x-y| < \delta\}.$$

The most known quantitative version of Theorem 1 was stated in 1968 by Shisha and Mond [23] and reads as follows. Here and throughout the paper we use the notation $e_i^x(t) = (t-x)^i$.

Theorem 2. *Let X be a compact real interval and let K_n be a sequence of positive linear operators defined on $C(X)$. Assume that the function $K_n e_0$ is bounded for all $n \in \mathbb{N}$. Then for each function $f \in C(X)$, and all $n \in \mathbb{N}$,*

$$\|f - K_n f\| \leq \|f\| \cdot \|K_n e_0 - e_0\| + \|K_n e_0 + e_0\| \cdot \omega(f, \mu_n),$$

where $\mu_n^2 = \sup_{x \in X}\{K_n e_2^x(x)\}$.

In [7] the reader can consult a generalization of this last theorem, stated under the more general setting we shall be concerned with, roughly described above.

Just to mention in advance one of our aims, we note that the results in [7, 22], extensions of the pioneers statements by Korovkin, Shisha, and Mond, remain valid when considering notions of convergence more general than the usual, and maintaining similar shape preserving assumptions on the operators. We could introduce right now these notions, however we now try to present further aims that turn to be extensions of the results that deal with the next natural step within approximation by linear operators, namely the study of the goodness of the estimates that yield the quantitative results. Obviously we are referring to the topic of saturation and to the natural previous step represented by the establishment of asymptotic formulae.

One of the seminal results in this context was given in 1932 by Voronovskaya [26] for the classical Bernstein operator,

$$B_n f(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k},\ f \in C[0,1],\ x \in [0,1],$$

namely,

Theorem 3. *If f is a function bounded on $[0,1]$, differentiable in a neighbourhood of $x \in [0,1]$, such that $f''(x)$ exists, then*

$$\lim_{n \to +\infty} n(B_n f(x) - f(x)) = \frac{x(1-x)}{2} f''(x).$$

If $f \in C^2[0,1]$, then the convergence is uniform.

The result assures that $B_n f(x) - f(x)$ is of order not better that $1/n$ if $f''(x) \neq 0$. A proof of the following extension can be consulted in [9]:

Theorem 4. *Let $L_n : C^k[a,b] \longrightarrow C^k[c,d]$, with $[c,d] \subset [a,b]$, be a sequence of linear k-convex operators and let $x \in [c,d]$. Assume that $D^k L_n e_{k+4}^x(x) = o(\lambda_n^{-1})$ and* *that there exists a sequence of positive real numbers* $\lambda_n \to +\infty$ *and a function* $p = p(t)$, *k times differentiable and positive on* (a,b), *such that for each* $i \in \{0,1,\ldots,k+2\}$,

$$\lim_{n\to+\infty} \lambda_n (D^k L_n e_i^x(x) - D^k e_i^x(x)) = D^k(pD^2 e_i^x)(x).$$

Then for each $f \in C^k[a,b]$, $k+2$ *times differentiable in a neighbourhood of* x,

$$\lim_{n\to+\infty} \lambda_n (D^k L_n f(x) - D^k f(x)) = D^k(pD^2 f)(x).$$

As regards saturation, we refer the reader to a result of Lorentz [18] for the Bernstein operators, and to an extension of it stated in 1972 by Lorentz and Schumaker [19]. Then, following the same scheme as above, we find a paper of the authors [6], where the results are moved into the setting of simultaneous approximation. For illustrative purposes we merely detail an application to the Bernstein operators.

Theorem 5. *Let* $0 < a < b < 1$. *Then for* $k \in \mathbb{N}$ *and* $f \in C^k[0,1]$,

$$2n\left|D^k B_n f(x) - D^k f(x)\right| \le M\frac{1}{(1-x)^{k-1}} + o(1),\ x \in (a,b),$$

if and only if

$$(e_1 - e_2)^k\left(\frac{1}{e_{k-1}}D^{k+1}f + \frac{k-1}{e_k}D^k f\right) \in Lip_M 1 \ en\ (a,b).$$

$\mathscr{A}$-Summability and Korovin-Type Approximation

In 1948 Lorentz [17] introduced and characterized the notion of almost convergence. Here, for the sake of clarity, we use that characterization as a definition.

Definition 1. A sequence of real numbers x_j is said to be almost convergent to ℓ, written in short as $x_j \rightharpoonup \ell$, if

$$\lim_{k\to+\infty} \frac{1}{k}\sum_{j=n}^{n+k-1} x_j = \ell \qquad \textit{uniformly in } n.$$

Twenty years later this notion was studied in approximation theory by King and Swetits [14]. They handled sequences of the type

$$L_i f(x) = \sum_{j=1}^{\infty} a_{ij}(x) f(x_{ij}),\ i = 1, 2, \ldots$$

related to interesting summability matrices $(a_{ij}(x))$, $i, j = 1, 2, \ldots$, and came up with certain connections between the convergence of the sequences and the regularity of those infinite matrices. We recall now these concepts that we shall use later in a different context.

Definition 2. Let $A = (a_{kj})$ be an infinite matrix with real entries, and let x_j be a sequence. We define the A-transform of x_j to be the sequence whose elements are given by

$$\sum_{j=1}^{\infty} a_{kj} x_j, \quad k = 1, 2, 3, \ldots$$

whenever the series converge for all k. If this last sequence converges to a limit ℓ, then we say that x_j is A-summable to ℓ.

As for the aforementioned regularity we quote a definition and ask the reader to recall the popular characterization theorem of Silverman–Toeplitz.

Definition 3. An infinite matrix A is said to be regular if the convergence of a sequence x towards its limit ℓ implies the convergence to the same limit of its A-transform.

King and Swetits proved the following result analogous to the classical Korovkin's one:

Theorem 6. *Let K_j be a sequence of positive linear operators defined on $C[a, b]$. Then $K_j f \rightharpoonup f \ \ \forall f \in C[a, b]$ if and only if $K_j e_i \rightharpoonup e_i \ \forall i \in \{0, 1, 2\}$.*

Later on, Mohapatra [20] studied the error of approximation by stating a result similar to Theorem 2. He introduced some notation and used a norm as follows: with a sequence of linear operators $K_j : C(X) \to C(X)$, a function $f \in C(X)$, and a non-negative integer k, one associates the sequence of functions $t_k f$, whose terms are given by

$$t_k^n f = \frac{1}{k} \sum_{j=n}^{n+k-1} K_j f,\ n = 1, 2, \ldots$$

Its norm, natural extension of the uniform one, is defined as

$$\|t_k f\| = \sup_n \sup_{x \in X} |t_k^n f(x)|.$$

With this notation, the sequence $K_j f$ is almost convergent to $g \in C(X)$, uniformly in X, if and only if $\|t_k f - g\|$ tends to 0 as k tends to infinity.

Under the setting of simultaneous approximation, the almost convergence was first studied in [12] from a qualitative point of view. The corresponding quantitative version will be studied in the next section, under the more general framework introduced in 1973 by Bell [4] via the following definition:

Definition 4. Let $A^{(n)}$, $n = 1, 2, \ldots$, be a sequence of infinite matrices with $A^{(n)} = (a_{k,j}^{(n)})$ and $a_{k,j}^{(n)} \geq 0$ for $n, k, j = 1, 2, \ldots$. The sequence x_j is said to be $A^{(n)}$-summable to ℓ if

$$\lim_{k\to\infty} \sum_{j=1}^{\infty} a_{k,j}^{(n)} x_j = \ell$$

uniformly in n, provided the series are convergent for all k and all n.

This concept unifies convergence and almost convergence. Indeed, if $a_{k,j}^{(n)} = 1/k$ for $n \leq j < k + n$ and $a_{k,j}^{(n)} = 0$ otherwise, then $A^{(n)}$-summability coincides with almost convergence. On the other hand, if $a_{k,j}^{(n)} = \delta_{kj}$, δ_{kj} being the Kronecker delta, then $A^{(n)}$-summability amounts to convergence. Moreover, if $A^{(n)} = A$ for all n and certain matrix A, then $A^{(n)}$-summability refers to classical summability matrix, and if in addition A is lower triangular and $a_{kj} = 1/k$, then $A^{(n)}$-summability reduces to the well-known Cesàro or $(C, 1)$ summability.

Swetits [25] proved a quantitative Korovkin-type result for sequences of positive linear operators under $A^{(n)}$-summability. He considered $L_j : C[a,b] \to C[a,b]$ and $A^{(n)}$, and defined for $f \in C[a,b]$ and $k, n = 1, 2, \ldots,$

$$A_k^{(n)}(f,x) = \sum_{j=1}^{\infty} a_{kj}^{(n)} L_j f(x), \quad \|A_k f\| = \sup_n \sup_{x\in[a,b]} |A_k^{(n)}(f,x)|.$$

Then he proved the following result, extension of Theorem 2 and the corresponding one by Mohapatra:

Theorem 7. *Under the previous conditions, assume that $\|A_k e_0\| < \infty$ and let $\mu_k^2 = \|A_k e_2^x\|$. Then for each $f \in C[a,b]$ and $k = 1, 2, \ldots$,*

$$\|f - A_k f\| \leq \|f\| \cdot \|A_k e_0 - e_0\| + \|A_k e_0 + e_0\| \cdot \omega(f, \mu_k).$$

The following section is devoted to present an extension of this last result to the setting of conservative approximation. It turns to be a quantitative version via $A^{(n)}$-summability of the general result proved in [12]. The last two sections deal with asymptotic formulae and saturation results. These results complete the study of $A^{(n)}$-summability in conservative approximation theory. They were published in the last few years in [1, 2, 11].

Quantitative Results

Let $X = [a, b]$ or $X = \mathbb{T}$ (understood to be $\mathbb{R}$ with the identifications of points modulo 2π), both endowed with the usual metric that we denote by d, i.e. $d(x, y) = |x-y|$ or $d(x, y) = \min\{|x-y|, 2\pi - |x-y|\}$, respectively. For $f \in C(X)$, the modulus of continuity is defined accordingly.

Let $\mathscr{A}$ be a sequence of infinite matrices $A^{(n)} = (a_{k,j}^{(n)})$, $n = 1, 2, \ldots$ with $a_{k,j}^{(n)} \geq 0$ for $n, k, j = 1, 2, \ldots$. Let $\mathscr{L}$ be a sequence of linear operators $L_j : C(X) \to C(X)$, $j = 1, 2, \ldots$, and for each $f \in C(X)$, let us consider the double sequence (whenever the series converge)

$$\mathscr{A}_{\mathscr{L}}^{k,n} f := \sum_{j=1}^{\infty} a_{kj}^{(n)} L_j f, \quad k, n = 1, 2, \ldots.$$

Our attention is now paid in the approximation process of $\mathscr{A}_{\mathscr{L}}^{k,n} f$ towards f as $k \to \infty$, uniformly in X and uniformly for $n = 1, 2, \ldots$. The convergence amounts to the $\mathscr{A}$-summability of $L_j f$ towards f uniformly in X. Equivalently, if we call $\mathscr{A}_{\mathscr{L}}^{k} f$ the sequence whose elements are given by $\mathscr{A}_{\mathscr{L}}^{k,n} f$, $n = 1, 2, \ldots$, then we can consider the norm

$$\|\mathscr{A}_{\mathscr{L}}^{k} f\| = \sup_{n} \sup_{x \in X} \left| \mathscr{A}_{\mathscr{L}}^{k,n} f(x) \right|.$$

This way, $L_j f$ is $\mathscr{A}$-summable to f uniformly in X, if and only if

$$\sup_{n} \sup_{x \in X} |\mathscr{A}_{\mathscr{L}}^{k,n} f(x) - f(x)| \to 0 \text{ cuando } k \to \infty,$$

which amounts to the fact that

$$\|\mathscr{A}_{\mathscr{L}}^{k} f - f\| \to 0 \text{ as } k \to \infty.$$

Now we state a quantitative result on general conservative approximation via $\mathscr{A}$-summability. With this aim, we consider the following ingredients: a subspace $W \subset C(X)$, a cone $C \subset W$, a linear operator $T : W \longrightarrow C(X)$ and, from it, the set $P = \{f \in W : Tf(x) \geq 0 \ \forall x \in X\}$; and finally a sequence of linear operators $L_j : W \longrightarrow W$ denoted by $\mathscr{L}$.

Obviously, our objective is to study the $\mathscr{A}$-summability of $TL_j f$ towards Tf ($TL_j f$ represents $(T \circ L_j)(f) = T(L_j(f))$). We shall assume the following conservative property related to the cone C: $L_j(P \cap C) \subset P$, $j = 1, 2, \ldots$. Specifically we shall estimate the quantity

$$\|\mathscr{A}_{T \circ \mathscr{L}}^{k} f - Tf\| = \sup_{n} \sup_{x \in X} |\mathscr{A}_{T \circ \mathscr{L}}^{k,n} f(x) - Tf(x)|,$$

where

$$\mathscr{A}^{k,n}_{T\circ\mathscr{L}}f = \sum_{j=1}^{\infty} a_{kj}^{(n)} TL_j f.$$

Theorem 8 ([1]). *Under the conditions above, let* $u \in W$ *such that* $Tu = e_0$*, and for each* $x \in X$*, let* $\varphi_x \in C$ *such that, for certain constants* $M > 0$ *and* $\lambda \geq 1$*,* $T\varphi_x(x) = 0$ *for all* $x \in X$ *and* $T\varphi_x(t) \geq Md(t,x)^{\lambda}$ *for all* $t, x \in X$*. Let us also assume that for each* $f \in W$ *there exists* $\alpha_f \geq 0$ *such that for* $\alpha \geq \alpha_f$*,* $\gamma \in [-1, 3]$ *and* $x \in X$*, we have that* $\pm f + \alpha\varphi_x + \gamma \|Tf\| u \in C$*. Finally, for each* $k \in \mathbb{N}$*, let us suppose that* $\left\|\mathscr{A}^{k}_{T\circ\mathscr{L}}u\right\| < \infty$ *and* $\mu_k^{\lambda} := \sup_n \sup_{x\in X} \mathscr{A}^{k,n}_{T\circ\mathscr{L}}\varphi_x(x) < \infty$*. Then, for* $f \in W$ *and* $k \in \mathbb{N}$

$$\left\|\mathscr{A}^{k}_{T\circ\mathscr{L}}f - Tf\right\| \leq \|Tf\| \left\|\mathscr{A}^{k}_{T\circ\mathscr{L}}u - e_0\right\| + \left\|\mathscr{A}^{k}_{T\circ\mathscr{L}}u\right\| \omega(Tf, \mu_k)$$
$$+max\left\{\frac{\omega(Tf,\mu_k)}{M}, \mu_k^{\lambda}\alpha_f\right\}.$$

Next we show some applications of the result. However, we firstly try to make it easier its understanding. With this purpose, let $W = C^m(X)$ and $T = D^m$. Thus $P \subset C^m(X)$ is the set of all m-convex functions and we are studying the $\mathscr{A}$-summability of $D^m L_j f$ towards $D^m f$, by estimating the quantity

$$\|\mathscr{A}^{k}_{D^m\circ\mathscr{L}}f - D^m f\| = \sup_n \sup_x \left|\sum_{j=1}^{\infty} a_{kj}^{(n)} D^m L_j f(x) - D^m f(x)\right|.$$

Notice that if the series above is finite, then $\mathscr{A}^{k}_{D^m\circ\mathscr{L}}f = D^m\left(\mathscr{A}^{k}_{\mathscr{L}}f\right)$, which takes us, if we were restricted to the usual convergence, to the topic of simultaneous approximation. Some papers in this line are [7, 8, 21, 22].

On the other hand, assume that for each $x \in X$, $\varphi_x = \sum_{i=1}^{s}\alpha_i(x)f_i$ for certain functions $f_i \in W$ and α_i defined on X. Then, from the hypotheses of the theorem,

$$\mathscr{A}^{k,n}_{T\circ\mathscr{L}}\varphi_x(x) = \mathscr{A}^{k,n}_{T\circ\mathscr{L}}\varphi_x(x) - T\varphi_x(x) = \sum_{j=1}^{\infty} a_{kj}^{(n)} TL_j\varphi_x(x) - T\varphi_x(x)$$

$$= \sum_{i=1}^{s}\alpha_i(x)\left(\sum_{j=1}^{\infty} a_{kj}^{(n)} TL_j f_i(x) - Tf_i(x)\right) = \sum_{i=1}^{s}\alpha_i(x)\left(\mathscr{A}^{k,n}_{T\circ\mathscr{L}}f_i(x) - Tf_i(x)\right).$$

Thus, if for $i = 1, \ldots, s$, the functions α_i are bounded on X and the functions $TL_j f_i$ are $\mathscr{A}$-summable to Tf_i uniformly in X, then

$$\mu_k^\lambda \leq \sum_{i=1}^{s} \|\alpha_i\| \left\| \mathscr{A}^k_{T\circ\mathscr{L}} f_i - Tf_i \right\|, \qquad \mu_k \longrightarrow 0.$$

If in addition $TL_j u$ is $\mathscr{A}$-summable to Tu uniformly in X, then for each $f \in W$, $TL_j f$ is $\mathscr{A}$-summable to Tf and we have an estimate of the order of $\mathscr{A}$-summability in terms of the corresponding orders of $TL_j f_i$ and $TL_j u$.

Applications

From Theorem 8 we first recover classical results on convergence and almost convergence. Then we consider more general domains and more general conservative properties. Finally, we deal with some particular well-known sequences of operators.

Particular Cases and Extensions

Let $X = [a, b]$, $W = C = C(X)$, $T = \mathbb{I}$ (the identity operator) and $a_{kj}^{(n)} = \delta_{kj}$. Let us apply Theorem 8 with $u = e_0$ and $\varphi_x(t) = (t - x)^2$. It appears the classical result of Shisha and Mond [23]. Analogously, if we let $X = \mathbb{T}$, $W = C = C(X)$, $T = \mathbb{I}$ and $a_{kj}^{(n)} = \delta_{kj}$, and then apply the theorem with $u = e_0$ and $\varphi_x(t) = \sin^2(\frac{t-x}{2})$, it appears the other result of Shisha and Mond [24] in the trigonometric case.

If, under the previous framework, we consider $a_{kj}^{(n)} = 1/k$ for $n \leq j < k + n$, and $a_{kj}^{(n)} = 0$ otherwise, then we recover the result of Mohapatra [20].

On the other hand, it is important to notice that, with obvious modifications, Theorem 8 remains valid if X is replaced by any compact, convex subset of $\mathbb{R}^m$. Under the usual convergence, this case was studied in [7].

On Almost Convex Operators

Now we deal with almost convexity of order $m - 1$, for certain $m \in \mathbb{N}$. This shape preserving property was first considered in [15]: the operator L_j is said to be almost convex of order $m - 1$ if there exists a subset $\Omega \subset \{0, 1, \ldots, m - 1\}$ such that

$$L_j \left(\bigcap_{i \in \Omega \cup \{m\}} H_X^i \right) \subset H_X^m, \tag{1}$$

where $H_X^i := \{f \in C^i(X) : D^i f(x) \geq 0 \ \forall x \in X\}$.

Next result follows directly from Theorem 8.

Corollary 1. *Let $m \in \mathbb{N}$ and let $L_j : C^m[0,1] \to C^m[0,1]$ be a sequence of linear operators fulfilling (1) for $X = [0,1]$ and certain subset $\Omega \subset \{0,1,\ldots,m-1\}$. Let $\varphi_x \in \langle e_{\min\Omega},\ldots,e_{m+2}\rangle$ such that $D^m\varphi_x(t) = (t-x)^2$ and $D^i\varphi_x(0) = 1 + \|D^{i+1}\varphi_x\|$ for $i = m-1, m-2, \ldots, \min\Omega$, and finally let*

$$\mu_k^2 = \sup_{x\in X}\sup_{n\in\mathbb{N}} \mathscr{A}^{k,n}_{D^m\circ\mathscr{L}}\varphi_x(x).$$

Then, for each $f \in C^m[0,1]$ and $k \in \mathbb{N}$,

$$\left\|\mathscr{A}^k_{D^m\circ\mathscr{L}}f - D^mf\right\| \le \frac{\omega(D^mf,\mu_k)}{m!}\left\|\mathscr{A}^k_{D^m\circ\mathscr{L}}e_m\right\| + \frac{1}{m!}\|D^mf\|\|\mathscr{A}^k_{D^m\circ\mathscr{L}}e_m - D^me_m\|$$
$$+ \max\left\{\omega(D^mf,\mu_k), \mu_k^2\max_{i\in\Omega}\left\{\|D^if\| + \frac{\|D^mf\|}{(m-i)!}\right\}\right\}.$$

We apply now the previous corollary to the well-known sequence of Meyer–König and Zeller operators, defined by

$$M_nf(x) = (1-x)^{n+1}\sum_{p=0}^{\infty} f\left(\frac{p}{p+n}\right)\binom{n+p}{p}x^p,\ \ x\in[0,1),\quad M_nf(1) = f(1).$$

Some shape preserving properties of these operators are known:

$$M_n(H^i_{[0,1]}) \subset H^i_{[0,1]},\ i = 0,1,2,\quad M_n(H^3_{[0,1]}) \not\subset H^3_{[0,1]},$$
$$M_n(H^2_{[0,1]}\cap H^3_{[0,1]}) \subset H^2_{[0,1]}\cap H^3_{[0,1]},$$
$$M_n(H^2_{[0,1]}\cap H^3_{[0,1]}\cap H^4_{[0,1]}) \subset H^2_{[0,1]}\cap H^3_{[0,1]}\cap H^4_{[0,1]}$$

and

$$M_n(H^2_{[0,1]}\cap H^3_{[0,1]}\cap H^4_{[0,1]}\cap H^5_{[0,1]}) \subset H^2_{[0,1]}\cap H^3_{[0,1]}\cap H^4_{[0,1]}\cap H^5_{[0,1]}.$$

Equivalently, these operators are positive, 1-convex and 2-convex. They are not 3-convex but almost convex of order k, for $k = 3,4,5$ with $\Omega = \{2,\ldots,k-1\}$.

Knoop and Pottinger [15] obtained estimates for $||D^1M_nf - D^1f||$ and $||D^2M_nf - D^2f||$. Under the usual convergence, Corollary 1 applies for $m = 3,4,5$. But also, given any sequence of matrices $\mathscr{A}$ under the conditions of Theorem 8, we can consider a sequence of positive real numbers β_n that is $\mathscr{A}$-summable to 0 and not necessarily convergent in the classical sense, and consider from it

$$\widehat{\mathscr{M}_n}f = (1+\beta_n)\mathscr{M}_nf.$$

The new sequence $\widehat{\mathscr{M}_n}$ inherits from M_n the aforesaid shape preserving properties, so Corollary 1 applies.

A Trigonometric Case

For $i \in \mathbb{N}_0 \cup \{+\infty\}$ and $\phi \in C(\mathbb{T})$, let

$$T_i\phi(t) = \frac{a_0}{2} + \sum_{v=1}^{i} a_v \cos(vt) + b_v \sin(vt),$$

a_v, b_v being the Fourier coefficients of $\phi \in C(\mathbb{T})$ and let $H^i_{\mathbb{T}} := \{\phi \in C(\mathbb{T}) : T_i\phi \geq 0\}$. We fix $m \in \mathbb{N}$ and consider $X = \mathbb{T}$, $W = C(\mathbb{T})$, $T = T_m$ and $C = \bigcap_{i\in\Omega} H^i_{\mathbb{T}}$, where Ω is a subset of $\{0, 1, \ldots, m-1\}$, and assume that the operators L_j satisfy that $L_j(P \cap C) \subset P$. Under these conditions the following result can be proved with the aid of Theorem 8.

Corollary 2. *Let* $h = \max \Omega$, $\varphi_x(t) = \sum_{v=1}^{h+1}(1 - \cos(v(t-x)))$ *and*

$$\mu_k^2 = \sup_{x\in X} \sup_{n\in\mathbb{N}} \mathscr{A}^{k,n}_{T_m\circ\mathscr{L}} \varphi_x(x).$$

Then for each $f \in W$ *and* $k \in \mathbb{N}$

$$\begin{aligned}\|\mathscr{A}^k_{T_m\circ\mathscr{L}} f - T_m f\| \leq{}& w(T_m f, \mu_k) \left\|\mathscr{A}^k_{T_m\circ\mathscr{L}} e_0\right\| + \|T_m f\| \left\|\mathscr{A}^k_{T_m\circ\mathscr{L}} e_0 - e_0\right\| \\ &+ \max\left\{\frac{\pi^2\omega(T_m f, \mu_k)}{2}, \mu_k^2 \max_{i\in\Omega}\{\|T_i f\| + \|T_m f\|\}\right\}.\end{aligned}$$

Asymptotic Formulae

In this section we pass on to study the next natural topic, once qualitative and quantitative results have been studied. To have a quick idea of what we are referring to, we recall the following expression for the mth derivative of the classical Bernstein operators: given f, m-times continuously differentiable on $[0, 1]$ and $m + 2$-times differentiable at $x \in (0, 1)$,

$$\lim_{n\to\infty} n(D^m B_n f(x) - D^m f(x)) = D^m(pD^2 f)(x),$$

where $p = p(t) = t(1-t)/2$. And also take a look at this other that follows, related to the almost convergence of $D^m B_k f(x)$ towards $D^m f(x)$:

$$\lim_{k\to\infty} \frac{k}{\log k} \left(\sum_{j=n}^{n+k-1} \frac{1}{k} D^m B_j f(x) - D^m f(x) \right) = D^m(pD^2 f)(x)$$

uniformly en n.

This last formula is illustrative but of low interest, as the usual convergence of $D^m B_k f(x)$ towards $D^m f(x)$ implies its almost convergence. Nevertheless it shows the type of result that we are about to state in this section. Here we do not consider a framework as general as the one considered in the previous section in Theorem 8. We consider the space $C[a,b]$, we do not consider any cones and restrict our attention to the case $T = D^m$, that is to say, we assume that the operators are m-convex. Moreover, in the applications, instead of considering general $\mathscr{A}$-summability, we focus on almost convergence.

Let $\mathscr{L}$ be a sequence of linear operators $L_j : C^m[a,b] \to C^m[a,b]$ and let $\mathscr{A} := \{A^{(n)}\} = \{a_{kj}^{(n)}\}$. Let us assume that the operators L_j are m-convex. For $f \in C^m[a,b]$ and $x \in [a,b]$ we consider the double sequence (whenever it makes sense)

$$\mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} f(x) := \sum_{j=1}^{\infty} a_{kj}^{(n)} D^m L_j f(x), \quad k, n = 1, 2, \ldots$$

Our concern now is to study the $\mathscr{A}$-summability of $D^m L_j f(x)$ towards $D^m f(x)$, specifically, the asymptotic behaviour of the expression

$$\mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} f(x) - D^m f(x).$$

In this respect, under the conditions above, we present the following result published in [11]. It can be considered a sort of Korovkin-type result for asymptotic formulae.

Theorem 9. *Let $x \in (a,b)$ such that the series $\mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} e_i^x(x)$ is convergent for $i = 0, 1, \ldots, m+4$ and for $k, n = 1, 2, \ldots$.*

Assume that:

(a) There exists a sequence of positive real numbers $\lambda_k \to +\infty$ and a function $p = p(t)$, m-times differentiable and strictly positive on (a,b), such that for each $i \in \{0, 1, \ldots, m+2\}$,

$$\lim_{k \to +\infty} \lambda_k \left(\mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} e_i^x(x) - D^m e_i^x(x) \right) = D^m \left(p D^2 e_i^x \right)(x)$$

uniformly for $n \in \mathbb{N}$,

(b)

$$\lim_{k \to +\infty} \lambda_k \mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} e_{m+4}^x(x) = 0$$

uniformly for $n \in \mathbb{N}$.

Then for each $f \in C^m[a,b]$, $m+2$-times differentiable in a neighbourhood of the point x,

$$\lim_{k \to +\infty} \lambda_k \left(\mathscr{A}_{D^m \circ \mathscr{L}}^{k,n} f(x) - D^m f(x) \right) = D^m \left(p D^2 f \right)(x)$$

uniformly for $n \in \mathbb{N}$.

Applications

First of all we notice that [10, Lemma 5.2] and Theorem 4 proved in [9] are particular cases of Theorem 9. On the other hand, we roughly pointed out at the beginning of the section the type of formulae that can be obtained for the derivatives of the Bernstein polynomials via almost convergence. Now we pay attention to some generalizations.

Modified Bernstein Operators and Almost Convergence

We make use of generalized Lototsky matrices (see [13, 14]). Given $f \in C[0,1]$ we consider the sequence of operators L_j defined by

$$L_j f(x) = \sum_{v=0}^{j} b_{jv}(x) f(v/j), \; x \in [0,1], \tag{2}$$

where the infinite matrix $B = (b_{jv}(x))$ is defined as follows:

$$b_{00} = 1, \quad b_{0v} = 0, \; v > 0,$$
$$\phi_j(x) \prod_{i=1}^{j} (h_i(x)y + 1 - h_i(x)) = \sum_{v=0}^{j} b_{jv}(x) y^v,$$

$\phi_j(x)$ being a sequence of functions defined on $[0,1]$ with $0 \leq \phi_j(x)$ for each $j = 1, 2, \ldots$ and $x \in [0,1]$, and $h_i(x)$ being a sequence of functions defined on $[0,1]$ with $0 \leq h_i(x) \leq 1$ for each $i = 1, 2, \ldots$ and $x \in [0,1]$.

Notice that if $\phi_j = e_0$ for all j and $h_i = e_1$ for all i, then the operators given in (2) coincide with the classical Bernstein ones.

In [14] the authors stated a qualitative result on the almost convergence of the sequence given in (2). Specifically, it was proved that if the $(C,1)$-transform of h_i converges to e_1 uniformly in $[0,1]$, and if $\phi_j(x)$ is almost convergent to 1 uniformly in $[0,1]$, then for each $f \in C[0,1]$, the sequence $L_j f(x)$ is almost convergent to $f(x)$ uniformly in $[0,1]$.

For the sake of simplicity, we consider the particular case where $h_i = e_1$ for all i and $\phi_j(x)$ is almost convergent to 1, non-necessarily convergent. Thus, the operators L_j turn to be a sort of Bernstein type operators. We can consider as example the sequence of functions $\phi_j(x) = 1 + (-1)^j x/2$, which does not converge but almost converges to 1. In this way, if we rename as L_j^{ϕ} the operators of this particular sequence, then we have that $L_j^{\phi} f(x)$ does not converge to $f(x)$, but does almost converge, that is to say

$$\lim_{k\to\infty}\sum_{j=n}^{n+k-1}\frac{1}{k}L_j^{\phi}f(x)=f(x)\ \text{ uniformemente para } n\in\mathbb{N}.$$

For sequences of this type the corollary below follows as a direct application of Theorem 9.

Corollary 3. *Let $f \in C[0,1]$ twice differentiable in a neighbourhood of a point $x \in (0,1)$, then, uniformly for $n \in \mathbb{N}$,*

$$\lim_{k\to+\infty}\frac{k}{\log k}\left(\sum_{j=n}^{n+k-1}\frac{1}{k}L_j^{\phi}f(x)-f(x)\right)=\frac{x(1-x)}{2}D^2f(x).$$

Saturation Results

In this section we review the topic of saturation under the same general framework that was considered in the previous section.

Our concern now is to study the saturation in the $\mathscr{A}$-summability processes of $D^mL_jf(x)$ towards $D^mf(x)$. To do this we start from the following three assumptions:

(P1) for each $f \in C^m[a,b]$ and $x \in [a,b]$, $D^mL_jf(x)$ is $\mathscr{A}$-summable to $D^mf(x)$, or equivalently

$$\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}}f(x):=\sum_{j=1}^{\infty}a_{kj}^{(n)}D^mL_jf(x)$$

converges to D^mf as k tends to infinity, uniformly in n,

(P2) each L_j is m-convex.
As a direct consequence of this property, for $f, g \in C^m[a,b]$,

$$D^mf(t)\le D^mg(t)\ \forall t\in[a,b]\ \Rightarrow \mathscr{A}^{k,n}_{D^m\circ\mathscr{L}}f(t)\le \mathscr{A}^{k,n}_{D^m\circ\mathscr{L}}g(t)\ \forall t\in[a,b].$$

(P3) there exists a sequence of real positive numbers $\lambda_k \to +\infty$ and three strictly positive functions w_0, w_1, w_2 defined on (a,b) with $w_i \in C^{2-i}(a,b)$ such that for $f \in C^m[a,b]$, $m+2$-times differentiable in some neighbourhood of a point $x \in (a,b)$,

$$\lim_{k\to+\infty}\lambda_k\left(\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}}f(x)-D^mf(x)\right)=\frac{1}{w_2}D^1\left(\frac{1}{w_1}D^1\left(\frac{1}{w_0}D^mf\right)\right)(x)$$

uniformly in n.

The formula given in (P3) gives information about the order of convergence of $\mathscr{A}^{k,n}_{D^m \circ \mathscr{L}} f(x)$ towards $D^m f(x)$; it assures that it is not better than λ_k^{-1} if that limit is different from zero. Thus, λ_k^{-1} is called optimal order of convergence and those functions which attained it form the so-called saturation class.

Our aim with this part of the work is to obtain that saturation class for the A-summability processes we are dealing with. We follow the line of work of [6], based upon the pioneers results by Lorentz and Schumaker [19]. We include some applications in the end.

Firstly, we discuss the reason why we use in (P3) the expression

$$\frac{1}{w_2} D^1 \left(\frac{1}{w_1} D^1 \left(\frac{1}{w_0} D^m f \right) \right) (x),$$

instead of the one that we considered in the previous section, i.e.

$$D^m \left(pD^2 f + qD^1 f \right) (x).$$

The reason is that the results look nicer and the saturation classes are clearly expressed from the functions w_i.

Secondly, if consider a subinterval $J \subset [a, b]$ and fix a point $c \in J$, it is well known that the functions

$$\begin{aligned} u_0(t) &= w_0(t), \\ u_1(t) &= w_0(t) \int_c^t w_1(s) ds, \\ u_2(t) &= w_0(t) \int_c^t w_1(t_1) \int_c^{t_1} w_2(t_2) dt_2 dt_1 \end{aligned} \tag{3}$$

form (in J) an extended complete Tchebychev system $\mathscr{T} = \{u_0, u_1, u_2\}$. On the other hand, $\{u_0, u_1\}$ is a fundamental system of solutions of the second order differential equation (with unknown v) that appears in (P3)

$$\mathscr{D}v := w_2^{-1} D^1 \left(w_1^{-1} D^1 \left(w_0^{-1} v \right) \right) = 0. \tag{4}$$

In addition, $\mathscr{D}u_2 = e_0$.

From the machinery introduced by Bonsall [5] in 1950, these considerations allow to talk about sub-($\mathscr{D}$) functions in J. This notion, that we recall now, will play an important role in the sequel. It amounts to certain generalized convexity, as we try to state with a proposition.

Definition 5. A function f defined on J is said to be sub-($\mathscr{D}$) in J if for each $t, t_1, t_2 \in J$ such that $t_1 < t < t_2$, one has that

$$f(t) \leq F(f, t_1, t_2),$$

where $F(f, t_1, t_2)$ is the unique solution of (4) which takes the values $f(t_1)$ and $f(t_2)$, respectively, at t_1 and t_2.

Proposition 1. *Let f be continuous on J. The following statements are equivalent:*

(i) *f is sub-($\mathscr{D}$) in J,*
(ii) *f is convex on J with respect to $\{u_0, u_1\}$,*
(iii) *for any $t_1, t, t_2 \in J$ with $t_1 < t < t_2$*

$$\begin{vmatrix} u_0(t_1) & u_1(t_1) & f(t_1) \\ u_0(t) & u_1(t) & f(t) \\ u_0(t_2) & u_1(t_2) & f(t_2) \end{vmatrix} \geq 0.$$

Moreover, in connection with the Tchebychev system, we present the operator

$$\triangle_{\mathscr{T}} f = \frac{1}{w_1} D^1 \left(\frac{1}{w_0} f \right),$$

and then the class $Lip_M^{\mathscr{T}} 1$, $M \geq 0$, both introduced and studied in [19], formed by those functions f, such that $\triangle_{\mathscr{T}} f$ is absolutely continuous and fulfill

$$|\triangle_{\mathscr{T}} f(t_2) - \triangle_{\mathscr{T}} f(t_1)| \leq M \int_{t_1}^{t_2} w_2(s) ds, \tag{5}$$

which amounts to the fact that f' is absolutely continuous and fulfill almost everywhere the inequality $|\mathscr{D} f| \leq M$.

It is important to notice that if $w_2 = e_0$, then the fact that f belongs to the class $Lip_M^{\mathscr{T}} 1$ is equivalent to $\triangle_{\mathscr{T}} f$ belongs to the classical Lipschitz class $Lip_M 1$.

Now we present a lemma proved in [19], which relates convexity and Liptschitz conditions.

Lemma 1. *A function $f \in C(J)$ belongs to $Lip_M^{\mathscr{T}} 1$ if and only if the functions $(-f + Mu_2)$ and $(f + Mu_2)$ are both convex with respect to $\{u_0, u_1\}$.*

As a last ingredient to achieve the desired result, we state some notation to compare rates of convergence. If $\alpha_k^{(n)}$ is a double sequence of real numbers such that $\lim_{k \to +\infty} \alpha_k^{(n)} = 0$ uniformly in $n \in \mathbb{N}$, and β_k is another sequence with $\lim_{k \to +\infty} \beta_k = 0$, then we shall write

$$\alpha_k^{(n)} = o^{(n)}(\beta_k)$$

to indicate that

$$\lim_{k \to +\infty} \frac{\alpha_k^{(n)}}{\beta_k} = 0 \text{ uniformly for } n \in \mathbb{N}.$$

Now we write one after the other the results that takes us to the main statement of this section.

Lemma 2. *Let J be an open subinterval of $[a,b]$. Let $g,h \in C(J)$ and $t_0,t_1,t_2 \in J$ such that $t_0 \in (t_1,t_2)$, $g(t_1) = g(t_2) = 0$ and $g(t_0) > 0$. Then there exist a real number $\epsilon < 0$, a solution of the differential equation (4) in J, say z, and a point $x \in (t_1,t_2)$ such that*

$$\begin{aligned}\epsilon h(t) + z(t) \geq g(t) \quad \forall t \in [t_1,t_2],\\ \epsilon h(x) + z(x) = g(x).\end{aligned}$$

Lemma 3. *Let $f \in C^m[a,b]$. If $D^m f$ is a solution of the differential equation (4) in some neighbourhood of $x \in (a,b)$, then*

$$\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) - D^m f(x) = o^{(n)}(\lambda_k^{-1}).$$

Lemma 4. *Let $f,g \in C^m[a,b]$ and let $x \in (a,b)$. Assume that there exists a neighbourhood N_x of x where $D^m f \leq D^m g$. Then*

$$\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) \leq \mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} g(x) + o^{(n)}(\lambda_k^{-1}).$$

Proposition 2. *Let $f \in C^m[a,b]$, then*

(a) $D^m f$ is convex with respect to $\{u_0,u_1\}$ in (a,b) if and only if for each $x \in (a,b)$

$$\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) \geq D^m f(x) + o^{(n)}(\lambda_k^{-1}).$$

(b) If $\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) \geq \mathscr{A}^{k+1,n}_{D^m\circ\mathscr{L}} f(x) + o^{(n)}(\lambda_k^{-1})$ for all $x \in (a,b)$, then $D^m f$ is convex with respect to $\{u_0,u_1\}$ in (a,b).

Proposition 3. *Let M be a positive constant and let $f,w \in C^m[a,b]$. Then the following statements are equivalent:*

(i) $MD^m w + D^m f$ and $MD^m w - D^m f$ are convex with respect to $\{u_0,u_1\}$ in (a,b),
(ii) for each $x \in (a,b)$,

$$\left|\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) - D^m f(x)\right| \leq M\left(\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} w(x) - D^m w(x)\right) + o^{(n)}(\lambda_k^{-1}).$$

We find different saturation results by taking different choices of the function w. We show two. In the first one, the saturation class is expressed in terms of the classical Lipschitz spaces, whereas in the second it appears the differential operator which appears in the corresponding asymptotic formula.

Theorem 10. *Let $f \in C^m[a,b]$. Then*

$$\lambda_k \left|\mathscr{A}^{k,n}_{D^m\circ\mathscr{L}} f(x) - D^m f(x)\right| \leq \frac{M}{w_2(x)} + o^{(n)}(1), \quad x \in (a,b)$$

if and only if, in (a, b),

$$\triangle_{\mathscr{T}} D^m f = w_1^{-1} D^1 (w_0^{-1}(D^m f) \in Lip_M 1.$$

Theorem 11. *Let* $f \in C^m[a, b]$. *Then*

$$\lambda_k \left| \mathscr{A}^{k,n}_{D^m \circ \mathscr{L}} f(x) - D^m f(x) \right| \le M + o^{(n)}(1), \quad x \in (a, b)$$

if and only if,

$$\left| \frac{1}{w_2} D^1 \left(\frac{1}{w_1} D^1 \left(\frac{1}{w_0} D^m f \right)\right)\right| \le M \ \ a.e. \ on \ (a, b).$$

Applications

Finally, we illustrate the use of the previous results. One has to make use of the corresponding asymptotic formulae that can be obtained from the previous section. We restrict our attention to Bernstein type operators and we do not go beyond almost convergence, as a particular case of $\mathscr{A}$-summability.

Corollary 4. *Let* $M > 0$ *and* $f \in C^3[0, 1]$, *then*

$$\frac{k}{\log k} \left| \sum_{j=n}^{n+k-1} \frac{1}{k} D^3 B_j f(x) - D^3 f(x) \right| \le M \frac{1}{2(1-x)^2} + o^{(n)}(1), \ x \in (0, 1)$$

if and only if

$$(e_1 - e_2)^3 \left(\frac{1}{e_2} D^4 f + \frac{2}{e_3} D^3 f \right) \in Lip_M 1 \ in \ (0, 1).$$

Corollary 5. *Let* $M > 0$ *and* $f \in C^3[0, 1]$, *then*

$$\frac{k}{\log k} \left| \sum_{j=n}^{n+k-1} \frac{1}{k} D^3 B_j f(x) - D^3 f(x) \right| \le M + o^{(n)}(1), \ x \in (a, b)$$

if and only if

$$|D^3((e_1 - e_2) D^2 f)| \le M \ \ a.e. \ on \ (0, 1).$$

Now we consider the sequence of linear operators L_j given in section "Modified Bernstein Operators and Almost Convergence."

Corollary 6. *Let* $M > 0$ *and* $f \in C[0,1]$*, then*

$$\frac{k}{\log k}\left|\sum_{j=n}^{n+k-1}\frac{1}{k}L_jf(x)-f(x)\right| \leq M\frac{x(1-x)}{2}+o^{(n)}(1),\ x\in(0,1)$$

if and only if

$$D^1 f \in Lip_M 1 \textit{ in } (0,1).$$

Corollary 7. *Let* $M > 0$ *and* $f \in C[0,1]$*, then*

$$\frac{k}{\log k}\left|\sum_{j=n}^{n+k-1}\frac{1}{k}L_jf(x)-f(x)\right| \leq M+o^{(n)}(1),\ x\in(0,1)$$

if and only if

$$|((e_1-e_2)D^2 f)| \leq M \ \textit{ a.e. on } (0,1).$$

Acknowledgement This work is partially supported by Junta de Andalucía, Spain (Research group FQM-0178).

References

1. Aguilera, F., Cárdenas-Morales, D., Garrancho, P., Hernández, J.M.: Quantitative results in conservative approximation via summability. Autom. Comput. Appl. Math. **17**(2), 201–208 (2008)
2. Aguilera, F., Cárdenas-Morales, D., Garrancho, P.: Optimal simultaneous approximation via A-summability. Abstr. Appl. Anal. Vol. 2013, 12p. (2013). Article ID 824058
3. Altomare, F., Campiti, M.: Koronkin-type Approximation Theory and its Applications. De Gruyter Studies in Mathematics, vol. 17. Walter de Gruyter, Berlin-New York (1994)
4. Bell, H.T.: Order summability and almost convergence. Proc. Am. Math. Soc. **38**, 548–552 (1973)
5. Bonsall, F.F.: The characterization of generalized convex functions. Q. J. Math. Oxford Ser. (2) **1**, 100–111 (1950)
6. Cárdenas-Morales, D., Garrancho, P., Local saturation of conservative operators. Acta Math. Hung. **100**(1–2), 83–95 (2003)
7. Cárdenas-Morales, D., Muñoz-Delgado, F.J.: A general uniform estimate of Korovkin-Type for conservative operators. Commun. Appl. Anal. (5) **4**, 445–454 (2001)
8. Cárdenas-Morales, D., Muñoz-Delgado, F.J.: A Korovkin-type result in C^k, an application to the M_n operators. Approx. Theory Appl. (N.S.) **17**(3), 1–13 (2001)
9. Cárdenas-Morales, D., Garrancho, P., Muñoz-Delgado, F.J.: A result of asymptotic formulae for linear k-convex operators. Int. J. Differ. Equ. Appl. **2**(3), 335–347 (2001)
10. DeVore, R.A.: The Approximation of Continuous Functions by Positive Linear Operators. Lecture Notes in Mathematics, vol. 293. Springer, Berlin-Heidelberg-New York (1972)

11. Garrancho, P., Cárdenas-Morales, D., Aguilera, F.: On asymptotic formulae via summability. Math. Comput. Simul. **81**, 2174–2180 (2011)
12. Hernández Guerra, J.M., Cárdenas-Morales, D.: Qualitative Korovkin-type results on almost convergence. Monogr. Semin. Mat. García Galdeano Universidad de Zaragoza **27**, 331–336 (2003)
13. King, J.P.: The Lototsky transform and Bernstein polynomials. Can. J. Math. **18**, 89–91 (1966)
14. King, J.P., Swetits, J.J.: Positive linear operators and summability. Aust. J. Math. **11**, 281–291 (1970)
15. Knoop, H.B., Pottinger, P.: Eim Satz von Korovkin-Typ für C^k-Räume. Math. Z. **148**, 123–132 (1976)
16. Korovkin, P.P.: Linear Operators and Approximation Theory. Hindustan Publishing, Delhi (1960)
17. Lorentz, G.G.: A contribution to the theory of divergent sequences. Acta Math. **80**, 167–190 (1948)
18. Lorentz, G.G.: Inequalities and saturation classes of Bernstein polynomials. In: On Approximation Theory. Proceedings of Conference Oberwolfach, 1963, pp. 200–207. Birkhäuser, Basel and Stuttgart (1964)
19. Lorentz, G.G., Schumaker, L.L.: Saturation of positive operators. J. Approx. Theory **5**, 413–424 (1972)
20. Mohapatra, R.N.: Quantitative results on almost convergence of sequence of positive linear operators. J. Approx. Theory **20**, 239–250 (1977)
21. Muñoz-Delgado, F.J., Cárdenas-Morales, D.: Almost convexity and quantitative Korovkin type results. Appl. Math. Lett. **11**(4), 105–108 (1998)
22. Muñoz-Delgado, F.J., Ramírez-González, V., Cárdenas-Morales, D.: Qualitative Korovkin-type results on conservative approximation. J. Approx. Theory **94**, 144–159 (1998)
23. Shisha, O., Mond, B.: The degree of convergence of sequences of linear positive operators. Proc. Natl. Acad. Sci. USA **60**, 1196–1200 (1968)
24. Shisha, O., Mond, B.: The degree of convergence of sequences of linear positive operators. J. Approx. Theory **1**, 335–339 (1968)
25. Swetits, J.J.: On summability and positive linear operators. J. Approx. Theory **25** 186–188 (1979)
26. Voronovskaya, E.: Détermination de la forme asymptotique d'approximation des fonctions par les polynômes de S. Bernstein. Dokl. Akad. Nauk. USSR A 79–85 (1932)

Simultaneous Weighted Approximation with Multivariate Baskakov–Schurer Operators

Antonio-Jesús López-Moreno, Joaquín Jódar-Reyes, and José-Manuel Latorre-Palacios

Abstract We study the properties of weighted simultaneous approximation of multivariate Baskakov–Schurer operators. We obtain quantitative estimates with explicit constants of the weighted approximation error for the partial derivatives. Moreover, we analyze the behavior of the operators with respect to weighted Lipschitz functions. For this purpose, we first compute the best constants, $M \in \mathbb{R}$, in the inequalities of the type $A_{n,p}\left((1 + |t|)^r\right) \leq M(1 + |t|)^r$.

Keywords Weighted simultaneous approximation • Lipschitz functions • Moduli of continuity • Multivariate operators

Introduction

Along the last years we find in the literature many papers devoted to the analysis of weighted approximation properties of different operators by means of several weighted modifications of classical moduli of continuity. We compute in this paper certain sharp constants for the multivariate version of the Baskakov–Schurer operators [11] that in combination with some extensions of the ideas in [9, 10] allow us to establish several inequalities for the weighted approximation error with explicit constants and for the Lipschitz constants preservation. In this section we start with the notation and main definitions of the paper.

Basic Notation

Given $m \in \mathbb{N} = \{1, 2, \ldots\}$ and $\alpha \in \mathbb{R}^m$, the i-th, $i = 1, \ldots, m$, component of the vector α is denoted by α_i, so that $\alpha = (\alpha_1, \ldots, \alpha_m)$. For $\beta \in \mathbb{R}^m$ and $a \in \mathbb{R}$, as usual, we will write, whenever it makes sense, $|\alpha| = \alpha_1 + \cdots + \alpha_m$, $\alpha! = \alpha_1! \cdots \alpha_m!$, $\alpha^\beta = \alpha_1^{\beta_1} \cdots \alpha_m^{\beta_m}$, $\alpha^a = (\alpha_1^a, \ldots, \alpha_m^a)$, $a^\alpha = a^{|\alpha|}$, $a\alpha = (a\alpha_1, \ldots, a\alpha_m)$,

A.-J. López-Moreno (✉) • J. Jódar-Reyes • J.-M. Latorre-Palacios
Universidad de Jaén, Campus Las Lagunillas, 23071 Jaén, Spain
e-mail: ajlopez@ujaen.es

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_21

$\alpha\beta = (\alpha_1\beta_1, \ldots, \alpha_m\beta_m)$, $\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}$ and also $\binom{a}{\beta} = \frac{a!}{\beta!(a-|\beta|)!}$. Besides, we will make use of the following notations for sums with several indexes,

$$\sum_{i=\alpha}^{\beta} A_i = \sum_{i\in\mathbb{Z}^m, \alpha\le i\le\beta} A_i \quad \text{and} \quad \sum_{i=\alpha}^{\infty} A_i = \sum_{i\in\mathbb{Z}^m, \alpha\le i} A_i,$$

where $\alpha \le \beta$ stands for $\alpha_i \le \beta_i$ for all $i = 1, \ldots, m$. Finally, we define $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^m$ and $e_i = (0, \ldots, 0, \overset{i)}{1}, 0, \ldots, 0) \in \mathbb{R}^m$, $i = 1, \ldots, m$.

In order to avoid confusions we will write the boldface symbol $\mathbf{|}\cdot\mathbf{|}$ for the absolute value of a number while $|\cdot|$ stands for the sum of the components of a vector.

We will denote by $t : \mathbb{R}^m \ni z \mapsto t(z) = z \in \mathbb{R}^m$ the identity map in $\mathbb{R}^m$ and by $t_i : \mathbb{R}^m \ni z = (z_1, \ldots, z_m) \mapsto t_i(z) = z_i \in \mathbb{R}$ the ith projection. The restrictions of t and t_i to any subset of $\mathbb{R}^m$ will also be denoted by t and t_i, respectively.

Given $\alpha \in \mathbb{N}_0^m$ (where $\mathbb{N}_0 = \{0, 1, 2\ldots\}$), we denote by D^α the partial derivative operator α_1 times with respect to the first variable, …, α_m times with respect to the m-th variable. For $U \subseteq \mathbb{R}^m$, $\mathbf{C}(U)$ is the set of continuous functions on U.

Finally, given $i, j \in \mathbb{N}_0$, the Stirling numbers, σ_i^j, of the second kind are the unique coefficients that make true the identity $x^i = \sum_{j=0}^{i} \sigma_i^j x^{\underline{j}}$, where $x^{\underline{j}} = x(x-1)\ldots(x-j+1)$ is the falling factorial, being $x^{\underline{0}} = 1$. If we take $\alpha, \beta \in \mathbb{N}_0^m$ we will extend the definition of the second kind Stirling numbers by $\sigma_\alpha^\beta = \sigma_{\alpha_1}^{\beta_1} \cdots \sigma_{\alpha_m}^{\beta_m}$.

The Baskakov–Schurer Operators

From now on, let us fix $m \in \mathbb{N}$ and take $H = (\mathbb{R}_0^+)^m$. Given $p \in \mathbb{N}_0$ and $n \in \mathbb{N}$ we define the multivariate Baskakov–Schurer operators, $A_{n,p}$, as follows: for $f : H \to \mathbb{R}$ and $x \in H$,

$$A_{n,p}f(x) = (1 + |x|)^{-(n+p)} \sum_{k=\mathbf{0}}^{\infty} f\left(\frac{k}{n}\right) \binom{n+p+|k|-1}{k} \left(\frac{x}{1+|x|}\right)^k .$$

When $m = 1$, we obtain the original univariate Baskakov–Schurer operators [14, 15] defined on $[0, \infty)$. Notice that this multivariate extension is not trivial in the sense that it is not a tensor product of the operators on $[0, \infty)$. The univariate version was introduced by Schurer as an extension of the original operators defined by Baskakov [3] (case $p = 0$) to approximate continuous functions on the semi-axis $[0, \infty)$. Multivariate Baskakov type operators are studied in several papers [8, 13, 16]. In particular, the authors have computed their complete asymptotic expansion in [12] and several properties relative to the approximation of functions of exponential growth in [11].

It is known that the operators $A_{n,p}$ represent an approximation method for functions of polynomial growth on H and their partial derivatives. The aim of this paper is to study this simultaneous approximation process. We do so from two points of view. We first compute quantitative estimates for the approximation error of f by $A_{n,p}f$ in section "Weighted Approximation Error for $A_{n,p}f$" and for the corresponding partial derivatives in section "Simultaneous Approximation." Second, we analyze the conservation of Lipschitz properties by $A_{n,p}$ and its partial derivatives in the last section.

Since H is a noncompact set, if we want to deal with functions of polynomial growth we have to use the weighted norms and moduli of continuity that we consider in section "Weighted Approximation Error for $A_{n,p}f$."

The starting point of the estimates that we are going to obtain is the calculation of the best constants, $M \in \mathbb{R}^+$, in the inequality

$$A_{n,p}\left((1+|t|)^r\right) \leq M(1+|t|)^r.$$

Such best constants are strongly related to the eigenvalues and eigenvectors of the operators. Section "Best Constants for Ψ^r" is devoted to this last problem.

In the sequel we denote $\Psi = 1 + |t|$ for short and throughout the paper we consider p to be a fixed number of $\mathbb{N}_0$.

Best Constants for Ψ^r

In [12] it is proved that for $\nu \in \mathbb{N}_0^m$,

$$A_{n,p}(t^\nu) = \sum_{i=\mathbf{0}}^{\nu} \frac{(n+p+|i|-1)^{\underline{|i|}}}{n^\nu} \sigma_\nu^i t^i. \tag{1}$$

Let $\mathbb{P}_r$ denote the space of polynomials of total degree $r \in \mathbb{N}$. From (1) it is immediate to conclude that the matrix of the linear operator $A_{n,p} : \mathbb{P}_r \to \mathbb{P}_r$ with respect to the basis $\{t^\nu : \nu \in \mathbb{N}_0^m, |\nu| \leq r\}$ is diagonal whenever we order its elements in an increasing way depending on the sum $|\nu|$. Then, its matrix has the block decomposition,

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_1 I_{d_1} & R_{1,2} & R_{1,3} & \cdots & R_{1,s} \\ 0 & 0 & \lambda_2 I_{d_2} & R_{2,3} & \cdots & R_{2,s} \\ 0 & 0 & 0 & \lambda_3 I_{d_3} & \cdots & R_{3,s} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_r I_{d_r} \end{pmatrix},$$

where $\lambda_i = \frac{(n+p+i-1)^{\underline{i}}}{n^i}$, $R_{i,j}$ are certain matrices, I_d stands for the identity matrix of order d, and $d_i = \binom{m+i}{i} - \binom{m+i-1}{i-1}$ for $i = 1, \ldots, r$. It is obvious that λ_i is an eigenvalue of $A_{n,p}$ with d_i independent eigenvectors of the form

$$p_{i,\nu} = t^\nu + \text{ lower powers of } t,$$

for each $\nu \in \mathbb{N}_0^m$ with $|\nu| = i$. The explicit computation of such eigenvectors is not easy and can only be carried out by means of recurrence relations as it is done for the Bernstein operators in [6, 7].

In order to obtain bounds for the approximation error we would like to have simple functions with a behavior close to that of the eigenvectors (once we know that we cannot compute them easily). For our purpose it would be useful to find a polynomial function $\phi_r \geq 0$ of total degree r such that $A_{n,p}(\phi_r) \leq \lambda_r \phi_r$. We solve this problem by proving that

$$A_{n,p}\left(\Psi^r\right) \leq \lambda_r \Psi^r.$$

The initial remarks about the eigenstructure of $A_{n,p}$ stress the fact that the eigenvalue λ_r is the best constant that makes true this inequality.

One could think of simpler functions, ϕ_r, as $|t|^r$, $1 + |t|^r$ or t^ν with $|\nu| = r$ for this purpose. However, a simple computation shows that they are not suitable (take, for instance, the case $m = 2$ and use (1)).

First of all, notice that from the definition of binomial numbers with vector arguments given above, we can write a couple of multivariate versions of Newton's binomial formula: given $\alpha, \beta \in \mathbb{N}_0^m$, $r \in \mathbb{N}_0$ and $x, y \in \mathbb{R}^m$,

$$(x+y)^\alpha = \sum_{i=\mathbf{0}}^{\alpha} \binom{\alpha}{i} x^i y^{\alpha - i}, \quad |x|^r = \sum_{i \in \mathbb{N}_0^m, |i| = r} \binom{r}{i} x^i. \tag{2}$$

Moreover, some well-known properties of binomial and Stirling numbers (see [1, Sect. 24.1.4.II.A]) can be extended to the multivariate notation as follows:

$$\sigma_{\alpha+e_j}^{\beta} = \beta_j \sigma_\alpha^\beta + \sigma_\alpha^{\beta - e_j}, j = 1, \ldots, m \quad \text{and} \quad \binom{r+1}{\alpha} = \binom{r}{\alpha} + \sum_{\substack{j=1 \\ \alpha_j \neq 0}}^{m} \binom{r}{\alpha - e_j}. \tag{3}$$

We start proving a couple of technical lemmas.

Lemma 1. *Given $r \in \mathbb{N}$, $N \in \mathbb{R}^+$ and $i \in \mathbb{N}_0^m$ such that $1 \leq |i| \leq r$, let us denote*

$$G(r,i,N) = \sum_{\substack{\nu \in \mathbb{N}_0^m \\ i \leq \nu, |\nu| \leq r}} \binom{r}{\nu} \frac{\sigma_\nu^i}{N^\nu}.$$

Then,

$$G(r,i,N) \leq \frac{(N+r-1)^{\underline{r-|i|}}}{N^r} \binom{r}{i}. \tag{4}$$

To prove the preceding lemma, we need the following sub-lemma:

Lemma 2. *Given $r \in \mathbb{N}$ and $i \in \mathbb{N}_0^m$ such that $1 < |i| \leq r$, it is satisfied that*

$$G(r+1,i,N) = \left(1+\frac{|i|}{N}\right) G(r,i,N) + \frac{1}{N}\sum_{\substack{j=1\\ i_j\neq 0}}^{m} G(r,i-e_j,N).$$

Proof (of Lemma 2). Applying the definition of $G(r+1,i,N)$ and the second identity in (3), we have

$$G(r+1,i,N) = \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|\leq r+1}} \binom{r+1}{\nu}\frac{\sigma_\nu^i}{N^\nu} = \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|\leq r+1}} \left[\binom{r}{\nu} + \sum_{\substack{j=1\\ \nu_j\neq 0}}^{m}\binom{r}{\nu-e_j}\right]\frac{\sigma_\nu^i}{N^\nu}.$$

Taking into account that $\binom{r}{\nu} = 0$ when $|\nu| = r+1$,

$$\begin{aligned} G(r+1,i,N) &= G(r,i,N) + \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|\leq r+1}} \sum_{\substack{j=1\\ \nu_j\neq 0}}^{m}\binom{r}{\nu-e_j}\frac{\sigma_\nu^i}{N^\nu} \\ &= G(r,i,N) + \sum_{\substack{j=1\\ i_j\neq 0}}^{m} \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i-e_j\leq\nu,|\nu|\leq r}} \binom{r}{\nu}\frac{\sigma_{\nu+e_j}^i}{N^{\nu+e_j}} \end{aligned}$$

and from the first identity in (3),

$$\begin{aligned} G(r+1,i,N) &= G(r,i,N) + \sum_{j=1}^{m}\frac{i_j}{N} \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|\leq r}} \binom{r}{\nu}\frac{\sigma_\nu^i}{N^\nu} + \frac{1}{N}\sum_{\substack{j=1\\ i_j\neq 0}}^{m} \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i-e_j\leq\nu,|\nu|\leq r}} \binom{r}{\nu}\frac{\sigma_\nu^{i-e_j}}{N^\nu} \\ &= G(r,i,N) + \frac{|i|}{N}G(r,i,N) + \frac{1}{N}\sum_{\substack{j=1\\ i_j\neq 0}}^{m} G(r,i-e_j,N). \end{aligned}$$

Proof (of Lemma 1). The result is clear for $r = |i|$. It is also true for $|i| = 1$ since then $i = e_j$ for some $j \in \{1,\ldots,m\}$ and

$$\begin{aligned} G(r,e_j,N) &= \sum_{\substack{\nu\in\mathbb{N}_0^m\\ e_j\leq\nu,|\nu|\leq r}} \binom{r}{\nu}\frac{\sigma_\nu^{e_j}}{N^\nu} = \sum_{s=1}^{r}\binom{r}{se_j}\frac{\sigma_{se_j}^{e_j}}{N^{se_j}} = \sum_{s=1}^{r}\binom{r}{s}\frac{1}{N^s} \\ &= \left(1+\frac{1}{N}\right)^r - 1 \leq r\frac{(N+r-1)^{\underline{r-1}}}{N^r} = \frac{(N+r-1)^{\underline{r-|e_j|}}}{N^r}\binom{r}{e_j}, \end{aligned}$$

where the last inequality can be easily proved by induction on r.

We are ready to prove (4) again by induction on r. For $r = 1$ the result is immediate. Let us suppose that it is true for $r \in \mathbb{N}$ and $i \in \mathbb{N}_0^m$ with $1 \leq |i| \leq r$ and prove it for $r+1$ and $i \in \mathbb{N}_0^m$ with $1 < |i| < r+1$ (for $|i| = 1, r+1$ we have already the result as we indicated before):

$$\begin{aligned} G(r+1,i,N) &= \left(1+\frac{|i|}{N}\right)G(r,i,N) + \frac{1}{N}\sum_{\substack{j=1\\ i_j\neq 0}}^{m} G(r,i-e_j,N) \\ &\leq \left(1+\frac{|i|}{N}\right)\frac{(N+r-1)^{\underline{r-|i|}}}{N^r}\binom{r}{i} + \frac{1}{N}\sum_{\substack{j=1\\ i_j\neq 0}}^{m}\frac{(N+r-1)^{\underline{r-|i-e_j|}}}{N^r}\binom{r}{i-e_j} \\ &\leq \frac{(N+r)^{\underline{r+1-|i|}}}{N^{r+1}}\left[\binom{r}{i} + \sum_{\substack{j=1\\ i_j\neq 0}}^{m}\binom{r}{i-e_j}\right] = \frac{(N+r)^{\underline{r+1-|i|}}}{N^{r+1}}\binom{r+1}{i} \end{aligned}$$

and the induction process is finished.

Let us establish now the announced optimal inequality for functions of the kind $(K+|t|)^r$.

Theorem 1. *Given $n \in \mathbb{N}$, $r \in \mathbb{N}_0$ and $K \in [1,\infty)$,*

$$A_{n,p}\left((K+|t|)^r\right) \leq \frac{(n+p+r-1)^{\underline{r}}}{n^r}(K+|t|)^r.$$

Proof. From (2) and (1) we have

$$\begin{aligned} A_{n,p}\left((K+|t|)^r\right) &= \sum_{s=0}^{r}\binom{r}{s}K^{r-s}A_{n,p}\left(|t|^s\right) = \sum_{s=0}^{r}\binom{r}{s}K^{r-s}\sum_{\substack{\nu\in\mathbb{N}_0^m\\ |\nu|=s}}\binom{s}{\nu}A_{n,p}(t^\nu) \\ &= \sum_{s=0}^{r}\binom{r}{s}K^{r-s}\sum_{\substack{\nu\in\mathbb{N}_0^m\\ |\nu|=s}}\binom{s}{\nu}\sum_{i=\mathbf{0}}^{\nu}\frac{(n+p+|i|-1)^{\underline{|i|}}}{n^\nu}\sigma_\nu^i t^i \\ &= \sum_{s=0}^{r}\binom{r}{s}K^{r-s}\sum_{\substack{i\in\mathbb{N}_0^m\\ |i|\leq s}}\sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|=s}}\binom{s}{\nu}\frac{(n+p+|i|-1)^{\underline{|i|}}}{n^\nu}\sigma_\nu^i t^i \\ &= \sum_{\substack{i\in\mathbb{N}_0^m\\ |i|\leq r}}(n+p+|i|-1)^{\underline{|i|}}t^i\sum_{s=|i|}^{r}K^{r-s}\sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\leq\nu,|\nu|=s}}\binom{r}{s}\binom{s}{\nu}\frac{\sigma_\nu^i}{n^\nu}. \end{aligned}$$

If $|\nu| = s$, then $\binom{r}{s}\binom{s}{\nu} = \binom{r}{|\nu|}\binom{|\nu|}{\nu} = \binom{r}{\nu}$, so that

$$\begin{aligned}
A_{n,p}\left((K+|t|)^r\right) &= \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} (n+p+|i|-1)^{\underline{|i|}} t^i \sum_{s=|i|}^{r} K^{r-s} \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\le\nu,|\nu|=s}} \binom{r}{\nu}\frac{\sigma_\nu^i}{n^\nu} \\
&= \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} (n+p+|i|-1)^{\underline{|i|}} t^i K^r \sum_{\substack{\nu\in\mathbb{N}_0^m\\ i\le\nu,|\nu|\le r}} \binom{r}{\nu}\frac{\sigma_\nu^i}{(Kn)^\nu} \\
&= \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} (n+p+|i|-1)^{\underline{|i|}} t^i K^r G(r,i,Kn) \\
&\le \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} (n+p+|i|-1)^{\underline{|i|}} t^i K^r \frac{(Kn+r-1)^{\underline{r-|i|}}}{(Kn)^r}\binom{r}{i} \\
&\le \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} (n+p+|i|-1)^{\underline{|i|}} t^i K^{r-|i|} \frac{(n+p+r-1)^{\underline{r-|i|}}}{n^r}\binom{r}{i} \\
&= \sum_{\substack{i\in\mathbb{N}_0^m\\|i|\le r}} t^i \frac{(n+p+r-1)^{\underline{r}}}{n^r} K^{r-|i|}\binom{r}{i} = \frac{(n+p+r-1)^{\underline{r}}}{n^r}\sum_{s=0}^{r} K^{r-s} \sum_{\substack{i\in\mathbb{N}_0^m\\|i|=s}} t^i \binom{r}{i} \\
&= \frac{(n+p+r-1)^{\underline{r}}}{n^r}\sum_{s=0}^{r}\binom{r}{s} K^{r-s} \sum_{\substack{i\in\mathbb{N}_0^m\\|i|=s}} t^i \binom{s}{i} \\
&= \frac{(n+p+r-1)^{\underline{r}}}{n^r}\sum_{s=0}^{r}\binom{r}{s} K^{r-s}\,|t|^s = \frac{(n+p+r-1)^{\underline{r}}}{n^r}(K+|t|)^r.
\end{aligned}$$

Weighted Approximation Error for $A_{n,p}f$

Since H is a noncompact set, we can find unbounded continuous functions on H for which the classical sup norm and moduli of continuity give no information (see for instance [2]). For this reason we need to make approximation with weights. The computations of the preceding section allow us to think of weight functions of the type Ψ^α. We are going to consider spaces of functions, norms and moduli of continuity relatives to these polynomial weights. We will find estimates of the weighted approximation error between $A_{n,p}f$ and f in terms of such weighted moduli of continuity.

Given $q \in \mathbb{N}$, $\|\cdot\|_q$ is the classical sub-q norm on $\mathbb{R}^m$. Let us also write $H_S = S^{m-1} \cap H$, where

$$S^{m-1} = \{x \in \mathbb{R}^m : \|x\|_2 = 1\},$$

and set $\mathbf{B}(x,\delta)$ the Euclidean ball centered at $x \in \mathbb{R}^m$ with radius $\delta \in \mathbb{R}^+$. The weighted moduli of continuity that we introduce below make necessary to consider functions, $f : H \to \mathbb{R}$, with a good behavior as x tends to infinity in the following sense:

Definition 1. Given $f : H \to \mathbb{R}$ we will say that

$$\lim f = \phi,$$

for a function $\phi : H_S \to \mathbb{R}$, whenever

$$\forall \varepsilon \in \mathbb{R}^+,\ \exists\, \delta \in \mathbb{R}^+ / x \in H - \mathbf{B}(\mathbf{0},\delta) \Rightarrow \left| f(x) - \phi\left(\frac{x}{\|x\|_2}\right) \right| \leq \varepsilon.$$

Now, for any $\alpha \in \mathbb{N}_0$ we will consider the classes of functions

$$B_{(\alpha)} = \{f \in \mathbf{C}(H) : \lim \frac{f}{\Psi^\alpha} = \phi \in \mathbf{C}(H_S)\},$$

and

$$\tilde{B}_{(\alpha)} = \{f : H \to \mathbb{R} : \exists K \in \mathbb{R}^+ / \forall x \in H, \frac{|f(x)|}{\Psi(x)^\alpha} \leq K\}$$

both of them equipped with the norm $\|f\|_{(\alpha)} = \sup_{x\in H} \frac{|f(x)|}{\Psi(x)^\alpha}$. It is immediate that $B_{(\alpha)} \subseteq \tilde{B}_{(\alpha)}$. The usage of the following modulus of continuity is suitable for the functions of $B_{(\alpha)}$: given $f \in B_{(\alpha)}$ and $\delta \in \mathbb{R}^+$,

$$\Omega_\alpha(f,\delta) = \sup_{\substack{h\in\mathbb{R}^m, \|h\|_2\leq\delta \\ x, x+h\in H}} \frac{|f(x+h) - f(x)|}{(1 + |x| + \|h\|_1)^\alpha}.$$

First of all, it is clear that

$$f \in \tilde{B}_{(\alpha)} \Rightarrow \Omega_\alpha(f,\delta) \leq 2\|f\|_{(\alpha)} < \infty.$$

That is to say, for those functions these moduli are well defined. Furthermore, Ω_α satisfies standard inequalities of the classical moduli of continuity and meet good properties with respect to the class $B_{(\alpha)}$. Let us see it in the following two lemmas.

Lemma 3. *Given* $f \in B_{(\alpha)}$, $\mu \in \mathbb{N}$ *and* $\delta \in \mathbb{R}^+$,

$$\Omega_\alpha(f, \mu\delta) \le \mu \Omega_\alpha(f, \delta).$$

Proof. We have $\Omega_\alpha(f, \mu\delta) = \sup_{\substack{h \in \mathbb{R}^m, \|h\|_2 \le \delta \\ x, x+\mu h \in H}} \underbrace{\frac{|f(x+\mu h) - f(x)|}{(1 + |x| + \|\mu h\|_1)^\alpha}}_{(*)}$ and

$$\begin{aligned}
(*) &\le \sum_{i=0}^{\mu-1} \frac{|f(x+(i+1)h) - f(x+ih)|}{(1+|x+ih| + \|h\|_1)^\alpha} \frac{(1+|x+ih| + \|h\|_1)^\alpha}{(1+|x| + \mu\|h\|_1)^\alpha} \\
&\le \sum_{i=0}^{\mu-1} \frac{|f(x+(i+1)h) - f(x+ih)|}{(1+|x+ih| + \|h\|_1)^\alpha} \frac{(1+|x| + (i+1)\|h\|_1)^\alpha}{(1+|x| + \mu\|h\|_1)^\alpha} \\
&\le \mu \Omega_\alpha(f, \delta).
\end{aligned}$$

Given $U \subseteq \mathbb{R}^m$ we call ω_U the classical modulus of continuity on U. That is to say, for $f : U \to \mathbb{R}$ and $\delta \in \mathbb{R}^+$,

$$\omega_U(f, \delta) = \sup_{\substack{h \in \mathbb{R}^m, \|h\|_2 \le \delta \\ x, x+h \in U}} |f(x+h) - f(x)|.$$

We also denote $\|f\|_U = \sup_{x \in U} |f(x)|$, the sup norm on U. It is immediate that for a compact set $U, f \in \mathbf{C}(U)$ implies $\lim_{\delta \to 0} \omega_U(f, \delta) = 0$ and $\|f\|_U < \infty$. We employ this all to prove our second lemma.

Lemma 4. *Given* $f \in B_{(\alpha)}$, $\lim_{\delta \to 0} \Omega_\alpha(f, \delta) = 0$.

Proof. Since $f \in B_{(\alpha)}$, $\lim \frac{f}{\Psi^\alpha} = \phi \in \mathbf{C}(H_S)$. Then, for any $\varepsilon \in \mathbb{R}^+$ we can find $\gamma \in [1, \infty)$ such that $x \in H - \mathbf{B}(\mathbf{0}, \gamma)$ implies $\left| \frac{f(x)}{(1+|x|)^\alpha} - \phi\left(\frac{x}{\|x\|_2}\right) \right| \le \varepsilon$.

Take any fixed number $\gamma_0 \ge \gamma + \delta$. We have

$$\Omega_\alpha(f, \delta) \le$$
$$\max\left\{ \underbrace{\sup_{\substack{h \in \mathbb{R}^m, \|h\|_2 \le \delta \\ x, x+h \in \mathbf{B}(\mathbf{0}, \gamma_0) \cap H}} \frac{|f(x+h) - f(x)|}{(1+|x| + \|h\|_1)^\alpha}}_{\le \omega_{\overline{\mathbf{B}(\mathbf{0},\gamma_0)} \cap H}(f, \delta)}, \sup_{\substack{h \in \mathbb{R}^m, \|h\|_2 \le \delta \\ x, x+h \in H - \mathbf{B}(\mathbf{0}, \gamma)}} \underbrace{\frac{|f(x+h) - f(x)|}{(1+|x| + \|h\|_1)^\alpha}}_{(*)} \right\},$$

and

$$\begin{aligned}
(*) = \Bigg| &\frac{f(x+h) - f(x)}{(1+|x| + \|h\|_1)^\alpha} - \phi\left(\frac{x}{\|x\|_2}\right) + \phi\left(\frac{x}{\|x\|_2}\right) \\
&+ \phi\left(\frac{x+h}{\|x+h\|_2}\right) - \phi\left(\frac{x+h}{\|x+h\|_2}\right) \Bigg|
\end{aligned}$$

$$\leq \left|\frac{f(x)}{(1+|x|+\|h\|_1)^\alpha} - \phi\left(\frac{x}{\|x\|_2}\right)\right| + \left|\frac{f(x+h)}{(1+|x|+\|h\|_1)^\alpha} - \phi\left(\frac{x+h}{\|x+h\|_2}\right)\right|$$

$$+ \left|\phi\left(\frac{x+h}{\|x+h\|_2}\right) - \phi\left(\frac{x}{\|x\|_2}\right)\right|$$

$$\leq \left|\frac{f(x)}{(1+|x|)^\alpha} - \phi\left(\frac{x}{\|x\|_2}\right)\right| + \left|\frac{f(x)}{(1+|x|)^\alpha}\right|\left(1 - \frac{(1+|x|)^\alpha}{(1+|x|+\|h\|_1)^\alpha}\right)$$

$$+ \left|\frac{f(x+h)}{(1+|x+h|)^\alpha} - \phi\left(\frac{x+h}{\|x+h\|_2}\right)\right|$$

$$+ \left|\frac{f(x+h)}{(1+|x+h|)^\alpha}\right|\left(1 - \frac{(1+|x+h|)^\alpha}{(1+|x|+\|h\|_1)^\alpha}\right) + \omega_{H_S}(\phi, \|h\|_2)$$

$$\leq 2\varepsilon + \|f\|_{(\alpha)}\left(1 - \frac{(1+|x|+\|h\|_1 - \|h\|_1)^\alpha}{(1+|x|+\|h\|_1)^\alpha}\right)$$

$$+ \|f\|_{(\alpha)}\left(1 - \frac{(1+|x|+\|h\|_1 + |h| - \|h\|_1)^\alpha}{(1+|x|+\|h\|_1)^\alpha}\right) + \omega_{H_S}(\phi, \|h\|_2)$$

$$= 2\varepsilon + \frac{\|f\|_{(\alpha)}}{(1+|x|+\|h\|_1)^\alpha}\sum_{j=1}^{\alpha}\binom{\alpha}{j}(1+|x|+\|h\|_1)^{\alpha-j}(-1)^{j+1}\|h\|_1^j$$

$$- \frac{\|f\|_{(\alpha)}}{(1+|x|+\|h\|_1)^\alpha}\sum_{j=1}^{\alpha}\binom{\alpha}{j}(1+|x|+\|h\|_1)^{\alpha-j}(|h| - \|h\|_1)^j$$

$$+ \omega_{H_S}(\phi, \|h\|_2)$$

$$\leq 2\varepsilon + \|f\|_{(\alpha)}\sum_{j=1}^{\alpha}\binom{\alpha}{j}\|h\|_1^j + \|f\|_{(\alpha)}\sum_{j=1}^{\alpha}\binom{\alpha}{j}\,\big||h| - \|h\|_1\big|^j + \omega_{H_S}(\phi, \|h\|_2)$$

$$\leq 2\varepsilon + (1+2^\alpha)\|f\|_{(\alpha)}\|h\|_1\sum_{j=1}^{\alpha}\binom{\alpha}{j}\|h\|_1^{j-1} + \omega_{H_S}(\phi, \|h\|_2)$$

$$\leq 2\varepsilon + (1+2^\alpha)m^{\frac{\alpha}{2}}\|f\|_{(\alpha)}\delta\sum_{j=1}^{\alpha}\binom{\alpha}{j}\delta^{j-1} + \omega_{H_S}(\phi, \delta).$$

Hence, we finally obtain that

$$\Omega_\alpha(f,\delta) \leq$$

$$\max\left\{\omega_{\overline{\mathbf{B}}(\mathbf{0},\gamma_0)\cap H}(f,\delta),\, 2\varepsilon + (1+2^\alpha)m^{\frac{\alpha}{2}}\|f\|_{(\alpha)}\delta\sum_{j=1}^{\alpha}\binom{\alpha}{j}\delta^{j-1} + \omega_{H_S}(\phi,\delta)\right\}.$$

Since $\overline{\mathbf{B}}(\mathbf{0}, \gamma_0) \cap H$ and H_S are compact sets and both f and ϕ are continuous functions, we have that $\lim_{\delta\to 0} \omega_{\overline{\mathbf{B}}(\mathbf{0},\gamma_0)\cap H}(f, \delta) = \lim_{\delta\to 0} \omega_{H_S}(\phi, \delta) = 0$ and then $\lim_{\delta\to 0} \Omega_\alpha(f, \delta) \leq 2\varepsilon$. As ε has been chosen arbitrarily this ends the proof.

Once we have a weighted modulus of continuity suitable for functions of polynomial growth on H we are going to use it to estimate the approximation error $\|A_{n,p}f - f\|_{(\alpha)}$. For this purpose we use a multivariate extension of the techniques that we find in [9].

Theorem 2. *Let $f \in B_{(\alpha)}$ and $\alpha > 1$,*

$$\|A_{n,p}f - f\|_{(\alpha+1)} \leq 3^\alpha \left(1 + \frac{(p + 2\alpha)^{\underline{\alpha}} - 1}{n}\right)\left(1 + \frac{p + p^2}{n}\right)^{\frac{1}{2}} \Omega_\alpha(f, \frac{1}{\sqrt{n}}).$$

Proof. Given $z, x \in H$, from the definition of Ω_α we have

$$|f(z) - f(x)| \leq (1 + |x| + \|z - x\|_1)^\alpha \Omega_\alpha(f, \|z - x\|_2).$$

From Lemma 3 it is easy to check that for any $\lambda, \delta \in \mathbb{R}^+$, $\Omega_\alpha(f, \lambda\delta) \leq (1 + \lambda)$ $\Omega_\alpha(f, \delta)$. Take $\lambda = \frac{\|z-x\|_2}{\delta}$ to obtain

$$\begin{aligned}|f(z) - f(x)| &\leq (1 + |x| + \|z - x\|_1)^\alpha (1 + \frac{\|z - x\|_2}{\delta})\Omega_\alpha(f, \delta)\\ &\leq (1 + 2\,|x| + |z|)^\alpha (1 + \frac{\|z - x\|_2}{\delta})\Omega_\alpha(f, \delta).\end{aligned}$$

If we apply $A_{n,p}$ and evaluate at x,

$$|A_{n,p}f(x) - f(x)| \leq A_{n,p}\left((1 + 2\,|x| + |t|)^\alpha (1 + \frac{\|t - x\|_2}{\delta})\right)(x)\Omega_\alpha(f, \delta).$$

By means of a Cauchy-Schwartz type inequality we have

$$\begin{aligned}&A_{n,p}\left((1 + 2\,|x| + |t|)^\alpha (1 + \frac{\|t - x\|_2}{\delta})\right)(x) \leq\\ &\leq A_{n,p}\left((1 + 2\,|x| + |t|)^\alpha\right)(x) + A_{n,p}^{\frac{1}{2}}\left((1 + 2\,|x| + |t|)^{2\alpha}\right)(x)A_{n,p}^{\frac{1}{2}}\left(\frac{\|t - x\|_2^2}{\delta^2}\right)(x)\\ &\leq \frac{(n + p + \alpha - 1)^{\underline{\alpha}}}{n^\alpha}(1 + 3\,|x|)^\alpha\\ &\quad + \left(\frac{(n + p + 2\alpha - 1)^{\underline{2\alpha}}}{n^{2\alpha}}\right)^{\frac{1}{2}} (1 + 3\,|x|)^\alpha \left(1 + \frac{p + p^2}{n}\right)^{\frac{1}{2}} (1 + |x|)\end{aligned}$$

$$\leq 3^{\alpha}\frac{(n+p+2\alpha-1)^{\underline{\alpha}}}{n^{\alpha}}(1+|x|)^{\alpha+1}\left(1+\frac{p+p^2}{n}\right)^{\frac{1}{2}}$$

$$\leq 3^{\alpha}\left(1+\frac{(p+2\alpha)^{\underline{\alpha}}-1}{n}\right)(1+|x|)^{\alpha+1}\left(1+\frac{p+p^2}{n}\right)^{\frac{1}{2}},$$

where we have taken $\delta = \frac{1}{\sqrt{n}}$ and we have used that $(1+3|x|)^{\alpha} + (1+|x|)(1+3|x|)^{\alpha} \leq 3^{\alpha}(1+|x|)^{\alpha+1}$ and also that

$$A_{n,p}\left(\|t-x\|_2^2\right)(x) \leq \left(\frac{1}{n}+\frac{p+p^2}{n^2}\right)(1+|x|)^2$$

which can be deduced from (1).

The inequality of the preceding result is also valid for $\alpha = 0$ and $\alpha = 1$ if we write 2 and $\frac{25}{8}$, respectively, in place of 3^{α}.

Simultaneous Approximation

In this section we study the error of simultaneous approximation. In other words we compute bounds for $\|D^{\gamma}A_{n,p}f - D^{\gamma}f\|_{(\alpha)}$ with $\gamma \in \mathbb{N}_0^m$. Our computations are based on the results of the preceding section and on the formula proved for the partial derivatives of $A_{n,p}$ in [12]: given $\gamma \in \mathbb{N}_0^m$,

$$D^{\gamma}A_{n,p}f = \frac{1}{\Psi^{\gamma}}A_{n,p}\left((n\Psi+p+|\gamma|-1)^{\underline{|\gamma|}}\Delta^{\gamma}_{\frac{1}{n}}f\right), \tag{5}$$

where for a function $g : H \to \mathbb{R}$, $x \in H$ and $h \in \mathbb{R}^+$, $\Delta_h^{\gamma}g(x)$ is the forward difference of g of order γ at x recursively defined by $\Delta_h^{\mathbf{0}}g(x) = g(x)$ and $\Delta_h^{\gamma+e_i}g(x) = \Delta_h^{\gamma}g(x+he_i) - \Delta_h^{\gamma}g(x)$. Recall that the mean value theorem implies that if $D^{\gamma}g$ is continuous then $\Delta_h^{\gamma}g(x) = h^{\gamma}D^{\gamma}g(\xi_x)$ for a ξ_x with $x \leq \xi_x \leq x + h\gamma$.

By means of (5) and Theorem 1 we can transfer Theorem 2 to the partial derivatives.

Fix $\gamma \in \mathbb{N}_0^m$ and $\alpha \in \mathbb{N}_0$ along this section.

Theorem 3. *Let $f : H \to \mathbb{R}$ be such that $D^{\gamma}f \in B_{(\alpha)}$. Then,*

$$\left|D^{\gamma}A_{n,p}f - D^{\gamma}f\right| \leq E_nC_n\Omega_{\alpha}(D^{\gamma}f, \frac{\|\gamma\|_2}{n})(1+\frac{|\gamma|}{n}+|t|)^{\alpha}$$
$$+E_n\left(C_n - \frac{(n+k)^{\gamma}}{n^{\gamma}}\right)\|D^{\gamma}f\|_{(\alpha)}\Psi^{\alpha} + \frac{1}{\Psi^{\gamma}}\left|A_{n,p}\left(\Psi^{\gamma}D^{\gamma}f\right) - \Psi^{\gamma}D^{\gamma}f\right|,$$

where $C_n : H \to \mathbb{R}$ *is given by* $C_n = \dfrac{((n+k)\Psi + k)^{\underline{|\gamma|}}}{n^{\gamma}\Psi^{\gamma}}$, $E_n = \dfrac{(n+\alpha+k)^{\underline{\alpha}}}{n^{\alpha}}$ *and* $k = p + |\gamma| - 1$.

Proof. From (5) we have

$$D^{\gamma}A_{n,p}f - D^{\gamma}f = \frac{1}{\Psi^{\gamma}}A_{n,p}\left((n\Psi + k)^{\underline{|\gamma|}}\left(\Delta^{\gamma}_{\frac{1}{n}}f - \frac{1}{n^{\gamma}}D^{\gamma}f\right)\right) + \frac{1}{\Psi^{\gamma}}\left(A_{n,p}\left(\frac{(n\Psi + k)^{\underline{|\gamma|}}}{n^{\gamma}}D^{\gamma}f\right) - \Psi^{\gamma}D^{\gamma}f\right). \tag{6}$$

On the other hand, for any $z \in H$, $\Delta^{\gamma}_{\frac{1}{n}}f(z) = \frac{1}{n^{\gamma}}D^{\gamma}f(\xi_z)$ for certain $\xi_z \in H$ with $z \le \xi_z \le z + \frac{\gamma}{n}$ and then,

$$\begin{aligned}\left|\Delta^{\gamma}_{\frac{1}{n}}f(z) - \frac{1}{n^{\gamma}}D^{\gamma}f(z)\right| &= \left|\frac{1}{n^{\gamma}}\left(D^{\gamma}f(\xi_z) - D^{\gamma}f(z)\right)\right| \\ &\le (1 + \|\xi_z - z\|_1 + |z|)^{\alpha}\frac{1}{n^{\gamma}}\Omega_{\alpha}(D^{\gamma}f, \|\frac{\gamma}{n}\|_2) \\ &\le \left(1 + \frac{|\gamma|}{n} + |z|\right)^{\alpha}\frac{1}{n^{\gamma}}\Omega_{\alpha}(D^{\gamma}f, \frac{\|\gamma\|_2}{n}).\end{aligned}$$

In this case,

$$\begin{aligned}&\left|A_{n,p}\left((n\Psi + k)^{\underline{|\gamma|}}\left(\Delta^{\gamma}_{\frac{1}{n}}f - \frac{1}{n^{\gamma}}D^{\gamma}f\right)\right)\right| \\ &\le \frac{1}{n^{\gamma}}\Omega_{\alpha}(D^{\gamma}f, \frac{\|\gamma\|_2}{n})A_{n,p}\left((n\Psi + k)^{\underline{|\gamma|}}(1 + \frac{|\gamma|}{n} + |t|)^{\alpha}\right).\end{aligned} \tag{7}$$

For any $N \in \mathbb{R}$, it is immediate that $(N+k)^{\underline{|\gamma|}} = \sum_{r=0}^{|\gamma|}\rho_r N^r$ for certain constants $\rho_0, \rho_1, \dots, \rho_{|\gamma|} \in \mathbb{N}_0$, being $\rho_{|\gamma|} = 1$. Then,

$$\begin{aligned}&A_{n,p}\left((n\Psi + k)^{\underline{|\gamma|}}(1 + \frac{|\gamma|}{n} + |t|)^{\alpha}\right) = \sum_{r=0}^{|\gamma|}\rho_r n^r A_{n,p}\left(\Psi^r(1 + \frac{|\gamma|}{n} + |t|)^{\alpha}\right) \\ &= \sum_{r=0}^{|\gamma|}\rho_r n^r \sum_{j=0}^{\alpha}\binom{\alpha}{j}\frac{|\gamma|^{\alpha-j}}{n^{\alpha-j}}A_{n,p}\left(\Psi^{r+j}\right) \\ &\le \sum_{r=0}^{|\gamma|}\rho_r n^r \sum_{j=0}^{\alpha}\binom{\alpha}{j}\frac{|\gamma|^{\alpha-j}}{n^{\alpha-j}}\frac{(n+p+r+j-1)^{\underline{r+j}}}{n^{r+j}}\Psi^{r+j}\end{aligned}$$

$$= \sum_{r=0}^{|\gamma|} \rho_r n^r \sum_{j=0}^{\alpha} \binom{\alpha}{j} \frac{|\gamma|^{\alpha-j}}{n^{\alpha-j}} \frac{(n+p+r+j-1)^{\underline{j}}}{n^j} \frac{(n+p+r-1)^{\underline{r}}}{n^r} \Psi^{r+j}$$

$$\leq E_n \sum_{r=0}^{|\gamma|} \rho_r (n+p+r-1)^{\underline{r}} \Psi^r \sum_{j=0}^{\alpha} \binom{\alpha}{j} \frac{|\gamma|^{\alpha-j}}{n^{\alpha-j}} \Psi^j$$

$$\leq E_n (1 + \frac{|\gamma|}{n} + |t|)^{\alpha} \sum_{r=0}^{|\gamma|} \rho_r \left[(n+k)\Psi\right]^r$$

$$\leq E_n (1 + \frac{|\gamma|}{n} + |t|)^{\alpha} \left((n+k)\Psi + k\right)^{\underline{|\gamma|}} . \tag{8}$$

Furthermore, if we denote $F_n = \left| A_{n,p} \left(\Psi^{\gamma} D^{\gamma} f\right) - \Psi^{\gamma} D^{\gamma} f \right|$,

$$\left| A_{n,p} \left(\frac{(n\Psi + k)^{\underline{|\gamma|}}}{n^{\gamma}} D^{\gamma} f \right) - \Psi^{\gamma} D^{\gamma} f \right| \leq \left| A_{n,p} \left(\sum_{r=0}^{|\gamma|} \rho_r n^{r-|\gamma|} \Psi^r D^{\gamma} f \right) - \Psi^{\gamma} D^{\gamma} f \right|$$

$$\leq \left| A_{n,p} \left(\sum_{r=0}^{|\gamma|-1} \rho_r n^{r-|\gamma|} \Psi^r D^{\gamma} f \right) \right| + F_n$$

$$\leq \|D^{\gamma} f\|_{(\alpha)} \left| A_{n,p} \left(\sum_{r=0}^{|\gamma|-1} \rho_r n^{r-|\gamma|} \Psi^{r+\alpha} \right) \right| + F_n$$

$$\leq \|D^{\gamma} f\|_{(\alpha)} \sum_{r=0}^{|\gamma|-1} \rho_r n^{-(\alpha+|\gamma|)} (n+p+r+\alpha-1)^{\underline{r+\alpha}} \Psi^{r+\alpha} + F_n$$

$$\leq \|D^{\gamma} f\|_{(\alpha)} E_n \Psi^{\alpha} \sum_{r=0}^{|\gamma|-1} \rho_r n^{-|\gamma|} (n+k)^r \Psi^r + F_n$$

$$\leq \|D^{\gamma} f\|_{(\alpha)} E_n \Psi^{\alpha} \left[\sum_{r=0}^{|\gamma|} \frac{\rho_r}{n^{\gamma}} (n+k)^r \Psi^r - \frac{(n+k)^{|\gamma|}}{n^{\gamma}} \Psi^{\gamma} \right] + F_n$$

$$= \|D^{\gamma} f\|_{(\alpha)} E_n \Psi^{\alpha} \left[\frac{\left[(n+k)\Psi + k\right]^{\underline{|\gamma|}}}{n^{\gamma}} - \frac{(n+k)^{|\gamma|}}{n^{\gamma}} \Psi^{\gamma} \right] + F_n. \tag{9}$$

If we use (7)–(9) in (6) we finish the proof.

Notice that C_n is a decreasing function of $|t|$ so $C_n \leq C_n(\mathbf{0})$. Then, it suffices to employ Theorem 2 with the preceding one to obtain the following quantitative estimates:

Theorem 4. *Suppose that* $1 \le |\gamma| \le \alpha$ *and take* $f : H \to \mathbb{R}$ *such that* $D^{\gamma} f \in B_{(\alpha-|\gamma|)}$*, the following holds true:*

$$
\begin{aligned}
|D^{\gamma} A_{n,p} f - D^{\gamma} f| &\le E_n C_n(\mathbf{0}) \Omega_{\alpha}(D^{\gamma} f, \frac{\|\gamma\|_2}{n})(1 + \frac{|\gamma|}{n} + |t|)^{\alpha} \\
&+ E_n \left(C_n(\mathbf{0}) - \frac{(n+k)^{\gamma}}{n^{\gamma}} \right) \|D^{\gamma} f\|_{(\alpha)} \Psi^{\alpha} \\
&+ 3^{\alpha} \Psi^{\alpha-|\gamma|+1} \left(1 + \frac{(p+2\alpha)^{\underline{\alpha}} - 1}{n} \right) \left(1 + \frac{p+p^2}{n} \right)^{\frac{1}{2}} \Omega_{\alpha}(\Psi^{\gamma} D^{\gamma} f, \frac{1}{\sqrt{n}})
\end{aligned}
$$

and

$$
\begin{aligned}
\|D^{\gamma} A_{n,p} f - D^{\gamma} f\|_{(\alpha)} &\le \left(1 + \frac{|\gamma|}{n} \right)^{\alpha} E_n C_n(\mathbf{0}) \Omega_{\alpha}(D^{\gamma} f, \frac{\|\gamma\|_2}{n}) \\
&+ E_n \left(C_n(\mathbf{0}) - \frac{(n+k)^{\gamma}}{n^{\gamma}} \right) \|D^{\gamma} f\|_{(\alpha)} \\
&+ 3^{\alpha} \left(1 + \frac{(p+2\alpha)^{\underline{\alpha}} - 1}{n} \right) \left(1 + \frac{p+p^2}{n} \right)^{\frac{1}{2}} \Omega_{\alpha} \left(\Psi^{\gamma} D^{\gamma} f, \frac{1}{\sqrt{n}} \right).
\end{aligned}
$$

Remark 1. The coefficients C_n that appear in Theorems 3 and 4 are functions defined on H but they are actually uniformly bounded on that domain. It is not difficult to obtain inequalities that describe more precisely their behavior.

Take $R \ge 1$. It is clear that

$$
((n+k)R + k)^{\underline{|\gamma|}} - (n+k)^{\gamma} R^{\gamma} = \sum_{i,j=0}^{|\gamma|-1} B_{i,j} n^i R^j
$$

for certain constants $B_{i,j} \in \mathbb{N}_0$. Taking $n = R = 1$, the former identity gives $(1 + 2k)^{\underline{|\gamma|}} - (1+k)^{|\gamma|} = \sum_{i,j=0}^{|\gamma|-1} B_{i,j}$. Moreover, $B_{|\gamma|-1,|\gamma|-1} = \frac{1}{2}(p+k)\,|\gamma|$ and $B_{|\gamma|-1,j} = 0$ for $j = 0, \ldots, |\gamma| - 2$. Since $R, n \ge 1$, we have

$$
\frac{((n+k)R+k)^{\underline{|\gamma|}} - (n+k)^{\gamma} R^{\gamma}}{n^{\gamma} R^{\gamma}} \le \frac{(p+k)\,|\gamma|}{2nR} + \frac{(1+2k)^{\underline{|\gamma|}} - (k+1)^{\gamma} - \frac{(p+k)|\gamma|}{2}}{n^2 R}.
$$

Hence, taking $R = \Psi$,

$$
0 \le C_n - \frac{(n+k)^{\gamma}}{n^{\gamma}} \le \frac{(p+k)\,|\gamma|}{2n\Psi} + \frac{(1+2k)^{\underline{|\gamma|}} - (k+1)^{\gamma} - \frac{(p+k)|\gamma|}{2}}{n^2 \Psi}.
$$

On the other hand, it is also easy to check that

$$
\frac{(n+k)^{\gamma}}{n^{\gamma}} \le 1 + \frac{|\gamma|\,k}{n} + \frac{(k+1)^{\gamma} - 1 - |\gamma|\,k}{n^2}
$$

and these two last inequalities yield

$$C_n \le 1 + \frac{|\gamma|\,k}{n} + \frac{(k+1)^{\gamma} - 1 - |\gamma|\,k}{n^2} + \frac{(p+k)\,|\gamma|}{2n\Psi} + \frac{(1+2k)\frac{|\gamma|}{} - (k+1)^{\gamma} - \frac{(p+k)|\gamma|}{2}}{n^2\Psi}.$$

Finally, the same techniques lead us to $E_n \le 1 + \frac{\alpha k + \frac{\alpha(\alpha+1)}{2}}{n} + \frac{(1+\alpha+k)\frac{\alpha}{} - 1 - \alpha k - \frac{\alpha(\alpha+1)}{2}}{n^2}$.

Preservation of Lipschitz Constants

It is a well-known fact that most of the classical sequences of linear positive operators present Lipschitz constants preservation properties. One of the first papers devoted to this question is the one by Brown et al. [4] where, for the Bernstein operators on $[0,1]$, B_n, it is proved that given a Lipschitz function $f : [0,1] \to \mathbb{R}$ we have that $B_n f$ is also Lipschitz with the same Lipschitz constant as f. From that moment on, many other authors have investigated these kind of preserving properties for different sequences of operators and different Lipschitz type constants both in the univariate [10] and multivariate [5, 17] case.

In this last section we consider the problem of the preservation of Lipschitz constants by the Baskakov–Schurer operators again for weighted approximation.

For this purpose let us take the following definition for the weighted Lipschitz constants: given $\alpha \in \mathbb{N}_0$, $\mathbf{0} < s \in \mathbb{N}_0^m$, $\mu \le |s|$ and $\delta \in \mathbb{R}^+$,

$$\theta^{(s)}_{\alpha,\mu}(f,\delta) = \sup_{\substack{h \in \mathbb{R}^+,\ 0<h\le\delta \\ x \in H}} \frac{\left|\Delta_h^s f(x)\right|}{h^{\mu}(1+h+|x|)^{\alpha}}$$

and

$$\theta^{(s)}_{\alpha,\mu}(f) = \sup_{\delta\in\mathbb{R}^+} \theta^{(s)}_{\alpha,\mu}(f,\delta).$$

We will say that a function f is Lipschitz of order (s,μ) whenever $\theta^{(s)}_{\alpha,\mu}(f) < \infty$ for some α.

In our last result we are going to establish relations between the Lipschitz constant for $D^{\beta} f$ and $D^{\beta} A_{n,p} f$, $\beta \in \mathbb{N}_0^m$. An immediate consequence of these relationship is the fact that $D^{\beta} f$ is Lipschitz of order $(s,|s|)$ implies that $D^{\beta} A_{n,p} f$ is so.

Theorem 5. *Given $\beta \in \mathbb{N}_0^m$ and $\alpha \in \mathbb{N}_0$, if $\mu \leq |s|$ then*

$$\theta^{(s)}_{\alpha,\mu}(D^\beta A_{n,p}f, \frac{1}{n}) \leq (F_n)^\alpha \left(1 + \frac{|\beta|}{n+1}\right)^\alpha \tilde{C}_n(\mathbf{0})\tilde{E}_n \theta^{(s)}_{\alpha,\mu}(D^\beta f, \frac{1}{n})$$

and if $\mu = |s|$ then

$$\theta^{(s)}_{\alpha,|s|}(D^\beta A_{n,p}f) \leq (G_n)^\alpha \left(1 + \frac{|\beta|}{n+1}\right)^\alpha \tilde{C}_n(\mathbf{0})\tilde{E}_n \theta^{(s)}_{\alpha,|s|}(D^\beta f, \frac{1}{n}),$$

where $\tilde{C}_n$ and $\tilde{E}_n$ are the coefficients C_n, E_n of Theorem 3 obtained for $\gamma = s + \beta$, $F_n = G_n = 1 + \frac{1}{n}$, if $|s| = 1$, and $F_n = 1 + \frac{|s|}{n+1}$ and $G_n = |s|$, if $|s| > 1$.

Proof. If we use (5), we know that

$$\frac{\left|\Delta_h^s D^\beta A_{n,p}f(x)\right|}{h^\mu(1+h+|x|)^\alpha} = \frac{h^s\left|D^{s+\beta}A_{n,p}f(\xi_x)\right|}{h^\mu(1+h+|x|)^\alpha}$$

$$= \frac{h^{|s|-\mu}n^{-\mu}}{(1+h+|x|)^\alpha}\left|\frac{A_{n,p}\left((n\Psi+\tilde{k})^{\underline{|s+\beta|}}\frac{\Delta^{s+\beta}_{\frac{1}{n}}f}{n^{-\mu}(1+\frac{1}{n}+|t|)^\alpha}(1+\frac{1}{n}+|t|)^\alpha\right)(\xi_x)}{\Psi(\xi_x)^{s+\beta}}\right|,$$

where $\tilde{k} = p + |s+\beta| - 1$ and $x \leq \xi_x \leq x + hs$. On the other hand, for any $z \in H$ there exists $z \leq \xi_z \leq z + \frac{\beta}{n}$ such that

$$\Delta^{s+\beta}_{\frac{1}{n}}f(z) = \Delta^\beta_{\frac{1}{n}}\Delta^s_{\frac{1}{n}}f(z) = n^{-\beta}D^\beta\Delta^s_{\frac{1}{n}}f(\xi_z) = n^{-\beta}\Delta^s_{\frac{1}{n}}D^\beta f(\xi_z)$$

and hence

$$\frac{\left|\Delta^{s+\beta}_{\frac{1}{n}}f(z)\right|}{n^{-\mu}(1+\frac{1}{n}+|z|)^\alpha} = n^{-\beta}\frac{\left|\Delta^s_{\frac{1}{n}}D^\beta f(\xi_z)\right|}{n^{-\mu}(1+\frac{1}{n}+|z|)^\alpha} \leq n^{-\beta}\theta^{(s)}_{\alpha,\mu}(D^\beta f, \frac{1}{n})\frac{(1+\frac{1}{n}+|\xi_z|)^\alpha}{(1+\frac{1}{n}+|z|)^\alpha}$$

$$\leq n^{-\beta}\theta^{(s)}_{\alpha,\mu}(D^\beta f, \frac{1}{n})\frac{(1+\frac{1}{n}+|z|+\frac{|\beta|}{n})^\alpha}{(1+\frac{1}{n}+|z|)^\alpha} \leq n^{-\beta}\left(1+\frac{|\beta|}{n+1}\right)^\alpha \theta^{(s)}_{\alpha,\mu}(D^\beta f, \frac{1}{n}).$$

Therefore,

$$\frac{\left|\Delta_h^s D^\beta A_{n,p}f(x)\right|}{h^\mu(1+h+|x|)^\alpha} \leq \frac{h^{|s|-\mu}n^{|s|-\mu}}{(1+h+|x|)^\alpha}\left(1+\frac{|\beta|}{n+1}\right)^\alpha \theta^{(s)}_{\alpha,\mu}(D^\beta f, \frac{1}{n})$$
$$\times\frac{A_{n,p}\left((n\Psi+\tilde{k})^{\underline{|s+\beta|}}(1+\frac{1}{n}+|t|)^\alpha\right)(\xi_x)}{n^{s+\beta}(1+|\xi_x|)^{s+\beta}}$$

$$\leq h^{|s|-\mu} n^{|s|-\mu} \left(1 + \frac{|\beta|}{n+1}\right)^{\alpha} \tilde{C}_n(\xi_x) \tilde{E}_n \theta^{(s)}_{\alpha,\mu}(D^{\beta} f, \frac{1}{n}) \frac{(1 + \frac{1}{n} + |\xi_x|)^{\alpha}}{(1 + h + |x|)^{\alpha}}$$

$$\leq \begin{cases} h^{|s|-\mu} n^{|s|-\mu} (F_n)^{\alpha} \left(1 + \frac{|\beta|}{n+1}\right)^{\alpha} \tilde{C}_n(\mathbf{0}) \tilde{E}_n \theta^{(s)}_{\alpha,\mu}(D^{\beta} f, \frac{1}{n}), & \text{if } h \leq \frac{1}{n}, \\ h^{|s|-\mu} n^{|s|-\mu} (G_n)^{\alpha} \left(1 + \frac{|\beta|}{n+1}\right)^{\alpha} \tilde{C}_n(\mathbf{0}) \tilde{E}_n \theta^{(s)}_{\alpha,\mu}(D^{\beta} f, \frac{1}{n}), & \text{in all cases,} \end{cases}$$

where we have obtained the third step by using the same arguments as in (8) and the last one taking into account that $\xi_x \leq x + hs$. In the last expression of the chain of inequalities we can cancel out h and n if $\mu = |s|$ or when $\mu < |s|$ and $h \leq \frac{1}{n}$.

References

1. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover Publications, New York (1972)
2. Achieser, N.I.: Theory of Approximation. Frederik Ungar Publishing, New York (1956)
3. Baskakov, V.A.: An example of a sequence of linear positive operators in the space of continuous functions. Dokl. Akad. Nauk SSSR **113**, 249–251 (1957)
4. Brown, B.M., Elliott, D., Paget, D.F.: Lipschitz constants for the Bernstein polynomials of a Lipschitz continuous function. J. Approx. Theory **49**, 196–199 (1987)
5. Chen, F.: The best Lipschitz constants of Bernstein polynomials and Bezier nets over a given triangle. Approx. Theory Appl. **11**(2), 1–8 (1995)
6. Cooper, S., Waldron, S.: The eigenstructure of the Bernstein operator. J. Approx. Theory **150**(1), 133–165 (2000)
7. Cooper, S., Waldron, S.: The diagonalisation of the multivariate Bernstein operator. J. Approx. Theory **117**, 103–131 (2002)
8. de la Cal, J. , Valle, A.M.: Best constants for tensor products of Bernstein, Szasz and Baskakov operators. Bull. Aust. Math. Soc. 62 **2**, 211–220 (2000)
9. López-Moreno, A.J.: Weighted simultaneous approximation with Baskakov type operators. Acta Math. Hung. **102**(1–2), 143–151 (2004)
10. López-Moreno, A.J.: Preservation of Lipschitz constants by Bernstein type operators. Acta Math. Hung. **106**(4), 343–354 (2005)
11. López-Moreno, A.J., Jódar, J., Muñoz-Delgado, F.J.: Exponential type moments and localization results for multivariate Baskakov–Schurer operators. Int. J. Differ. Equ. Appl. **6**(1), 53–67 (2002)
12. López-Moreno, A.J., Martínez-Moreno, J., Muñoz-Delgado, F.-J.: Asymptotic expressions for multivariate operators. In: Chui, C.K., Schumaker, L.L., Stöeckler, J. (eds.) Approximation Theory X: Abstract and Classical Analysis, pp. 287–307. Vanderbilt University Press, Nashville, TN (2002)
13. Rempulska, L., Skorupka, M.: The Baskakov operators for functions of two variables. Collect. Math. **50**(3), 289–302 (1999)
14. Schurer, F.: Linear positive operators in approximation theory. Math. Inst. Techn. Univ. Delft, report (1962)
15. Schurer, F.: On linear positive operators in approximation theory. Dissertation, University of Delft (1965)
16. Stancu, F.: On the approximation of functions of one and two variables by means of the operators of Baskakov. Stud. Cerc. Mat. **22**, 531–542 (1970)
17. Yu Yu, F.: Lipschitz constants for the Bernstein polynomials defined over a triangle. J. Math. Res. Exposition **19** 1, 105–108 (1990)

Approximation of Discontinuous Functions by q-Bernstein Polynomials

Sofia Ostrovska and Ahmet Yaşar Özban

In the memory of Professor Hüseyin Şirin Hüseyin

Abstract This chapter presents an overview of the results related to the q-Bernstein polynomials with $q > 1$ attached to discontinuous functions on $[0, 1]$. It is emphasized that the singularities of such functions located on the set

$$\mathbb{J}_q := \{0\} \cup \{q^{-l}\}_{l=0}^{\infty}, \quad q > 1,$$

are definitive for the investigation of the convergence properties of their q-Bernstein polynomials.

Keywords q-Bernstein polynomial • Discontinuous function • Time scale • Convergence

Introduction

The Bernstein polynomials first appeared in [3] published by S.N. Bernstein in 1912. There, it was proved that, given $f \in C[0, 1]$, polynomials

$$B_n(f; x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}, \; n = 1, 2, \ldots, \qquad (1)$$

converge to f *uniformly* on $[0, 1]$. Nowadays, polynomials (1) are called the *Bernstein polynomials* and their remarkable properties have made them an area of intensive research. A vast number of studies have been conducted not only on the behavior, but also on various applications of the Bernstein polynomials see, for example, [9] and [19]. In [20], V.S. Videnskii analyzed ideas and problems, leading to the discovery of polynomials (1).

S. Ostrovska (✉) • A.Y. Özban
Department of Mathematics, Atilim University, TR-06836 İncek, Ankara, Turkey
e-mail: sofia.ostrovska@atilim.edu.tr; ahmet.ozban@atilim.edu.tr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_22

Originally, the Bernstein polynomials were used only for the continuous functions and, later on, L.V. Kantorovich pioneered the application of these polynomials to wider classes of functions. In [8], he introduced modified Bernstein polynomials, known today as the *Kantorovich polynomials*, to approximate integrable functions on $[0, 1]$ and, subsequently, G.G. Lorentz employed the latter polynomials to approximate functions $f \in L^p[0, 1]$ in the L_p-metric. Historical notes on the subject can be found in [21]. Afterwards, Lorentz considered approximation of unbounded functions by the Bernstein polynomials, stating '*Remarkable phenomena can occur for unbounded functions*' (cf. [9, Chap. 1, Sect. 1.9]), which has been confirmed by profound theorems proved by Chlodovsky, Herzog, Hill, and Lorentz himself. Recently, Weba [22] obtained new results on the approximation of unbounded functions by the Bernstein polynomials.

On the other hand, due to an intensive development in the q-calculus, various q-versions of the classical Bernstein polynomials have appeared. Regrettably, the first one proposed by Lupaş in 1987 [10] remained in the shadow caused by the very limited availability of his article published in regional conference proceedings. Today, though, this situation has changed—see, for example, [1]—along with improvements in the communication technology. The most popular q-generalization which appeared 10 years after Lupaş's paper [10] belongs to G.M. Philips [17], who constructed new polynomials known today as the *q-Bernstein polynomials*. The bibliography on the q-Bernstein polynomials comprises about 200 papers and the research in the area is still going on.

Initially, the q-Bernstein polynomials had been considered exclusively for continuous functions, where both the similarities and distinctions with the classical Bernstein polynomials had been established. Recently, it has been discovered that the investigation of the q-Bernstein polynomials attached to discontinuous functions is also a fruitful field of research because, in this direction, new phenomena have been revealed. More precisely, the behavior of these polynomials depends on whether singularity belongs to a certain sequence, in which case the type of singularity affects the approximation properties.

In this work, an overview of the results related to the q-Bernstein polynomials for discontinuous functions is presented.

Preliminaries

Let $\mathbb{N}_0$ and $\mathbb{N}$ denote the set of nonnegative integers and positive integers, respectively, $0 < q \in \mathbb{R}$, and let $f : [0, 1] \to \mathbb{C}$. For any positive integer n and $k \in \mathbb{N}_0$ with $0 \le k \le n$, let $[k]_q$, $[k]_q!$, and $\begin{bmatrix} n \\ k \end{bmatrix}_q$ be the *q-integer, q-factorial, and q-binomial coefficients* defined by

$$[k]_q := 1 + q + \dots + q^{k-1} \ \ (k \in \mathbb{N}), \ \ [0]_q := 0,$$

$$[k]_q! := [1]_q[2]_q \dots [k]_q \ (k \in \mathbb{N}), \ \ [0]_q! := 1,$$

and

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{[n]_q!}{[k]_q![n-k]_q!},$$

respectively ([2], Chap. 10).

Definition 1 ([17]). The *q-Bernstein polynomials* are defined by:

$$B_{n,q}(f;x) := \sum_{k=0}^{n} f\left(\frac{[k]_q}{[n]_q}\right) p_{nk}(q;x), \quad n \in \mathbb{N},$$

where

$$p_{nk}(q;x) := \begin{bmatrix} n \\ k \end{bmatrix}_q x^k \prod_{s=0}^{n-k-1} (1 - q^s x), \quad k = 0, 1, \ldots n, \tag{2}$$

For $q = 1, B_{n,q}(f;x)$ is the classical Bernstein polynomial $B_n(f;x)$ and, conventionally, the term *q-Bernstein polynomial* implies that $q \neq 1$. It is a noteworthy fact that both the Bernstein and q-Bernstein polynomials are closely related to probability theory; more specifically, to the binomial and the q-binomial distributions. See [3, 5–7].

Throughout the chapter, we take $q > 1$. Although, in general, the discussion on the approximation properties of the q-Bernstein polynomials deals with the interval [0, 1], the set $\mathbb{J}_q$ defined by

$$\mathbb{J}_q := \{0\} \cup \{q^{-l}\}_{l=0}^{\infty},$$

plays an important role in our studies. Let us point out that $\mathbb{J}_q$ is a *time scale* and that the authors will continue to use this term. The definition of a time scale and an introduction to the calculus on time scales can be found, for example, in [4, Chap. 1], where the notation $(1/q)^{\mathbb{N}_0}$ is employed for this set.

In this chapter, the results are presented on the q-Bernstein polynomials for the functions of the class $\mathscr{F}_\alpha$ consisting of all functions continuous on $[0, 1] \setminus \{\alpha\}$ which possess an analytic continuation into disc $\mathscr{D}_\alpha = \{z \in \mathbb{C} : |z| < \alpha\}$. All theoretical outcomes are illustrated by numerical examples, which have been obtained in a MAPLE 8 environment using 1000 digits in precision and displayed in Figs. 1, 2, 3, 4, 5, 6, 7, 8, and 9. In each of these figures, part **a** demonstrates general behaviour of the functions on the entire interval [0, 1], while the other parts emphasize specific details on smaller subintervals. In all the numerical computations, we take $q = 2$, and all of the figures exhibit the well-known *end-point interpolation property* of the q-Bernstein polynomials, namely:

$$B_{n,q}(f;0) = f(0),\ B_{n,q}(f;1) = f(1) \ \text{ for all } \ n \in \mathbb{N} \ \text{ and all } \ q > 0$$

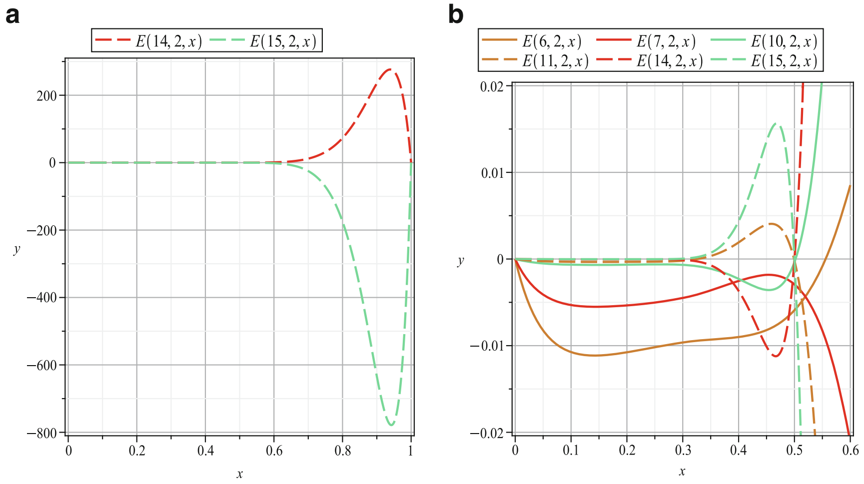


Fig. 1 Graphs of the error functions $E(n, 2, x)$, for $n = 6, 7, 10, 11, 14$, and 15

The Role of Time Scale $\mathbb{J}_q$

The investigation of convergence of the q-Bernstein polynomials reveals that the set $\mathbb{J}_q$ plays a key role in this circle of problems.

Proposition 1. *The set $\mathbb{J}_q$ is a 'minimal' set, where the sequence $\{B_{n,q}(f;x)\}$ converges for all $f \in C[0,1]$, in the sense that:*

(i) $\{B_{n,q}(f;x)\} \to f(x)$ *when* $x \in \mathbb{J}_q$, $f \in C[0,1]$. *Moreover, the convergence on $\mathbb{J}_q$ is uniform;*
(ii) There exists $f \in C[0,1]$, such that $x \notin \mathbb{J}_q \Rightarrow \left|B_{n,q}(f;x)\right| \to \infty$ as $n \to \infty$.

If $f \in \mathscr{F}_\alpha$, the set $\mathbb{J}_q$ is still essential when the convergence of $\{B_{n,q}(f;.)\}$ is concerned. To begin with, one has to distinguish whether $\alpha \in \mathbb{J}_q$ or $\alpha \notin \mathbb{J}_q$. The results of the next section demonstrate that, when $\alpha \notin \mathbb{J}_q$, the approximation results are quite similar to those for $f \in C[0,1]$ and also to the well-known convergence theorem for the Maclaurin polynomials. In contrast, the case $\alpha \in \mathbb{J}_q$ reveals a new phenomenon, thoroughly considered in section "Singularity at $q^{-m}, m \in \mathbb{N}$." It should be mentioned that the convergence theorems of the forthcoming sections are proved with the help of Vitali's Theorem [18, Chap. 5, Theorem 5.21], where the convergence on $\mathbb{J}_q$ is a crucial issue.

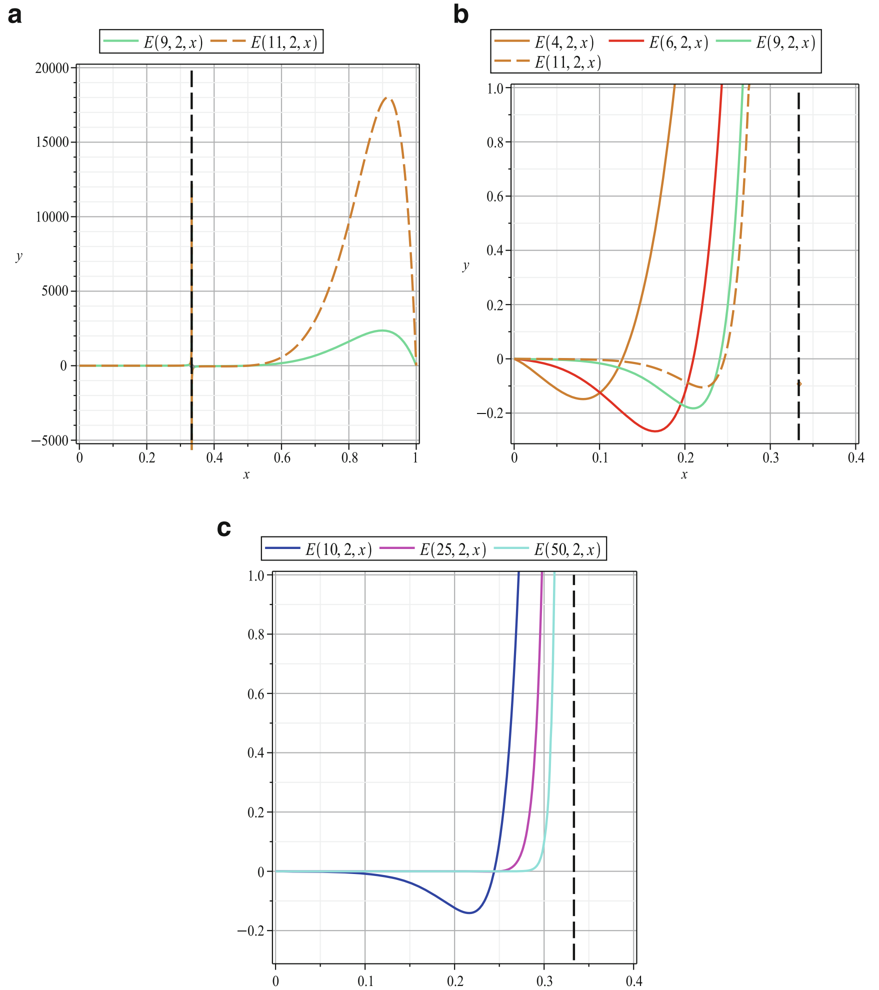


Fig. 2 Graphs of the error functions $E(n, 2, x)$, for $n = 4, 6, 9, 10, 11, 25$, and 50

Singularity Outside of the Time Scale $\mathbb{J}_q$

As mentioned already, if $\alpha \notin \mathbb{J}_q$, then the situation in terms of convergence for $\{B_{n,q}(f; .)\}$ resembles the one for $f \in C[0, 1]$. In particular, statement (i) of Proposition 1 remains true (see, Lemma 2 of [14]). To trace the resemblance, let us refer to Theorem 2.3 of [11] concerning the continuous case and its counterpart Theorem 1 in [16] concerning $\mathscr{F}_\alpha$. The respective findings are summarized as follows.

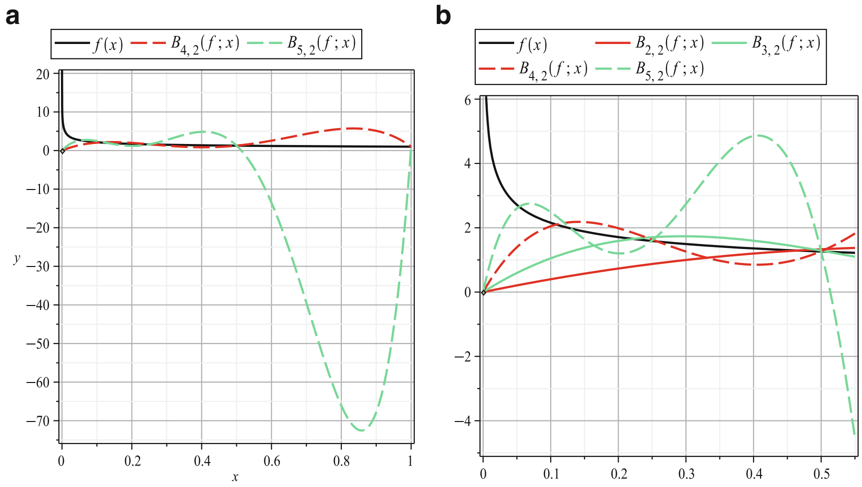


Fig. 3 Graphs of the function $y = f(x)$ and its q-Bernstein polynomial $B_{n,2}(f;x)$ for $n = 2, 3, 4$, and 5

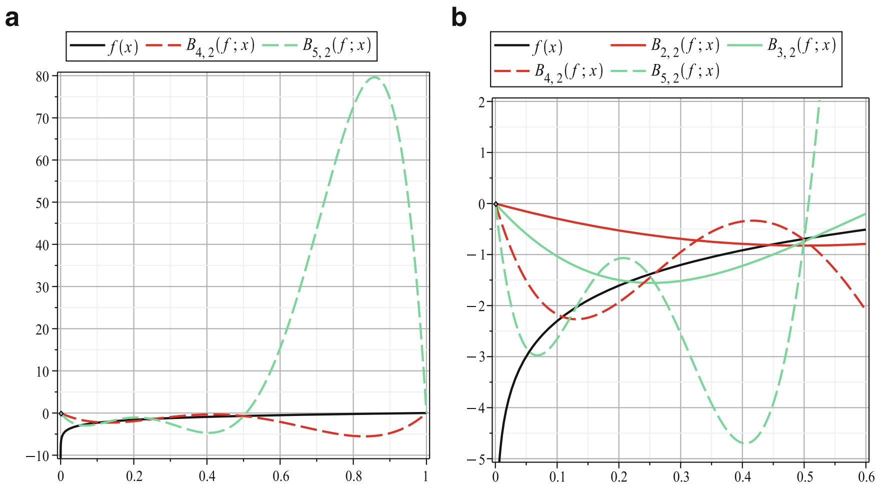


Fig. 4 Graphs of the function $y = f(x)$ and its q-Bernstein polynomial $B_{n,2}(f;x)$ for $n = 2, 3, 4$, and 5

Theorem 1. *(i) Let $f \in C[0,1]$ admit an analytic continuation into a disc D_a, $a > 0$. Then, f is uniformly approximated by its q-Bernstein polynomials on any compact subset of $(-a, a)$.*

(ii) Let $\alpha \in [0,1] \setminus \mathbb{J}_q$ and $f \in \mathscr{F}_\alpha$. Then, f is uniformly approximated by its q-Bernstein polynomials on any compact subset of $(-\alpha, \alpha)$.

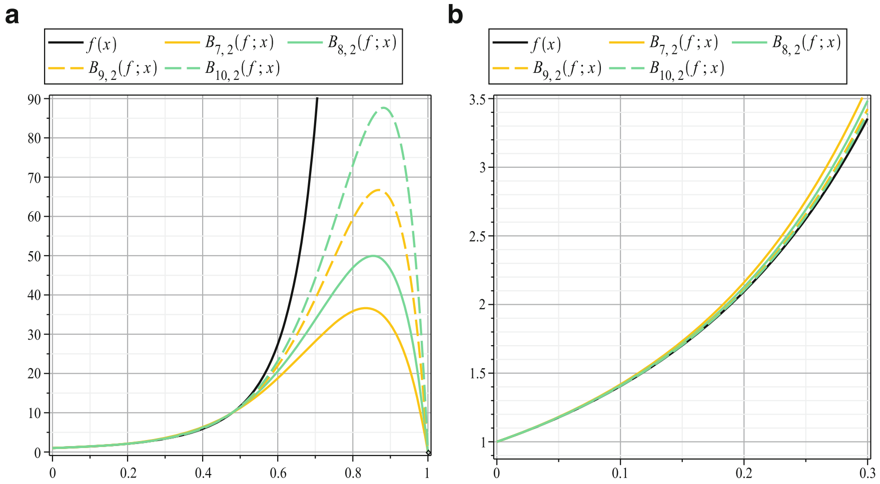


Fig. 5 Graphs of the function $y = f(x)$ and its q-Bernstein polynomial $B_{n,2}(f;x)$ for $n = 7, 8, 9$, and 10

These statements are illustrated with the help of the examples below.

Example 1. Let $f(x) = \frac{\sin(2x)}{x+1/3}$.

Figure 1a and b displays the graphs of the *error functions* $E(n, q, x) := B_{n,q}(f;x) - f(x)$ for $n = 6, 7, 10, 11, 14$, and 15. Figure 1a demonstrates the general behavior of $E(14, 2, x)$ and $E(15, 2, x)$ on $[0, 1]$, while related details are provided in Fig. 1b. It is observed that, as n increases, the values of $E(n, q, x)$ approach 0 if $x \in [0, 1/3)$ or $x = 0.5 \in \mathbb{J}_q$, and grow in magnitude otherwise. It can be proved that although f is continuous on $[0, 1]$, it is not approximated by its q-Bernstein polynomials on this interval.

Example 2. Let $f(x) = \frac{\sin(2x)}{x-1/3}$ if $x \neq 1/3$ and $f(1/3) = 0$.

Figure 2a–c shows the graphs of the error functions $E(n, 2, x)$ when $n = 4, 6, 9, 10, 11, 25$, and 50. Despite the fact that the graphs are very different from those in Fig. 1, the qualitative behavior is rather similar. Namely, as n increases, the values of $E(n, 2, x)$ approach 0 on $[0, 1/3)$ and at $x = 0.5 \in \mathbb{J}_q$ while, outside of those points, the values of the error functions grow in magnitude. However, to observe these tendencies, one needs to use polynomials of higher degrees because the convergence near the pole $x = 1/3$ is slow. In Fig. 2c the graphs of the error functions when the degrees of polynomials equal 10, 25, and 50 are shown.

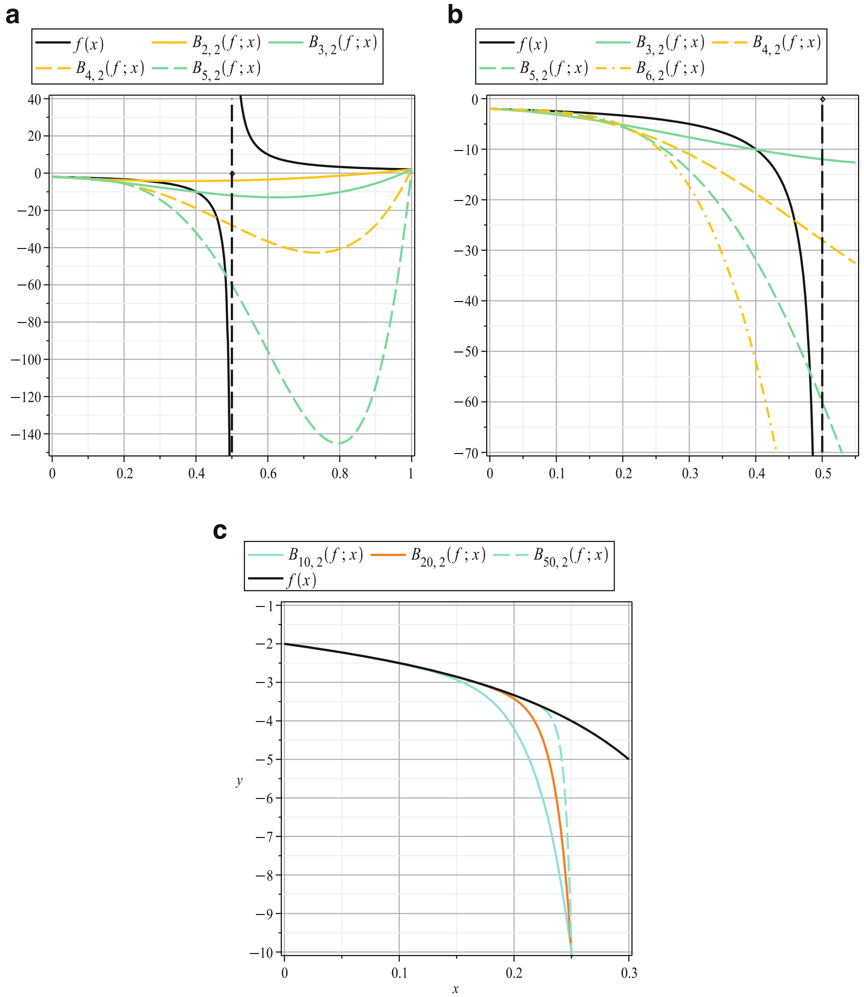


Fig. 6 Graphs of the function $y = f(x)$ and its q-Bernstein polynomials $B_{n,2}(f; x)$ for $n = 2, 3, 4, 5$ and 6

Singularity on the Time Scale $\mathbb{J}_q$

On the whole, if $\alpha \notin \mathbb{J}_q$, then the results cannot be viewed as significantly different from those corresponding to $f \in C[0, 1]$ and, as such, we focus on $\alpha \in \mathbb{J}_q$. As for the time scale itself, the points are distinguished regarding their characteristic properties as follows:

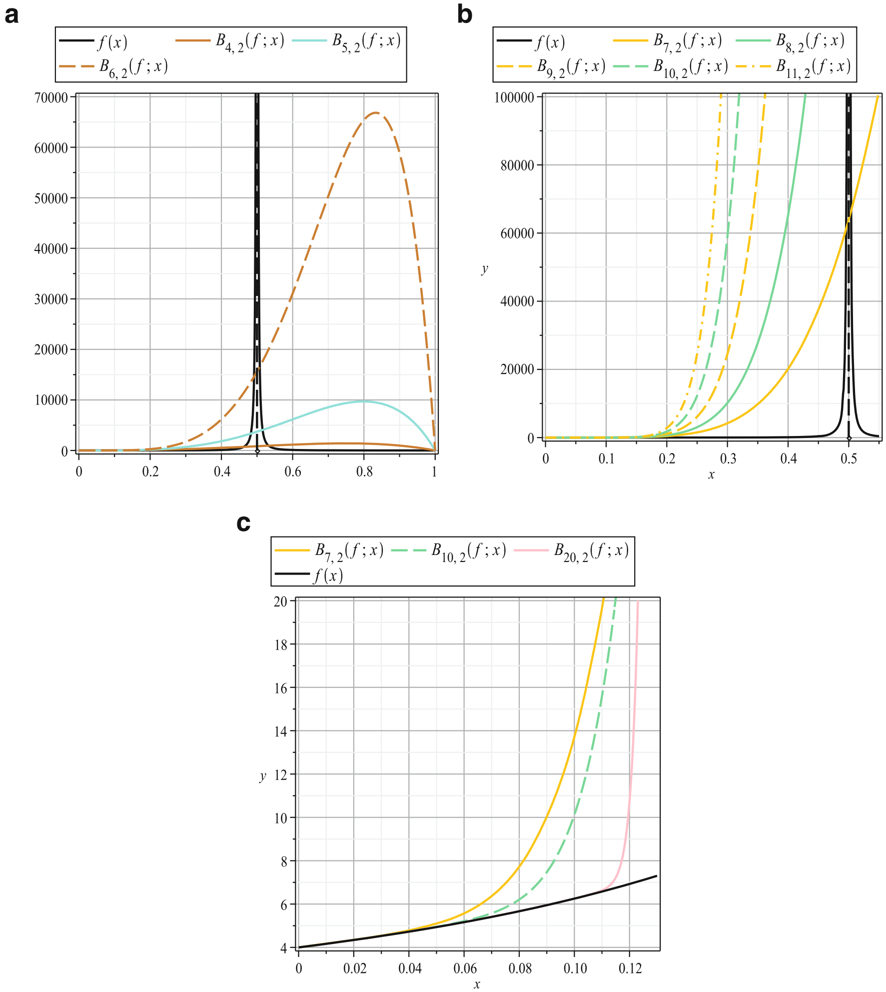


Fig. 7 Graphs of the function $y = f(x)$ and its q-Bernstein polynomials $B_{n,2}(f; x)$ for $n = 4, 5, 6, 7, 8, 9, 10$ and 11

- $x = 0$ is an accumulation point for the set $\mathbb{J}_q$;
- $x = 1$ is one of the nodes for all $n \in \mathbb{N}$ and also an isolated point for the set of nodes;
- $x = q^{-m}, m \in \mathbb{N}$, is an accumulation point for the nodes $\dfrac{[n-m]_q}{[n]_q}$.

In the same order, the cases when α is one of those will be discussed separately in the sequel.

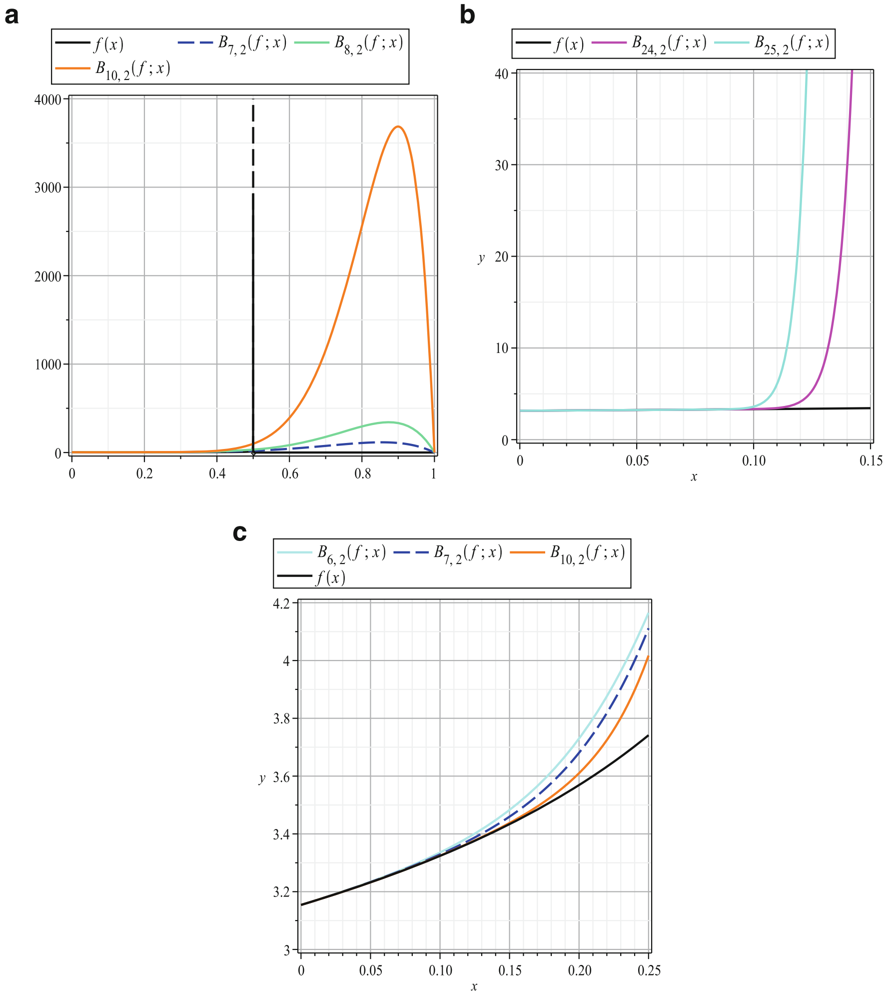


Fig. 8 Graphs of the function $y = f(x)$ and its q-Bernstein polynomials $B_{n,2}(f;x)$ for $n = 6, 7, 8, 10, 24$ and 25

Singularity at 0

Here, the time scale is still a minimal set of convergence since, by Corollary 4 of [14], the sequence $\{B_{n,q}(f;x)\} \to f(x)$ for $x \in \mathbb{J}_q$, $f \in \mathscr{F}_0$. Meanwhile, the following results show that the convergence at all other points may fail. The next theorem deals with the functions possessing either the power or the logarithmic singularities at 0. For their proofs, we refer to [12, 14].

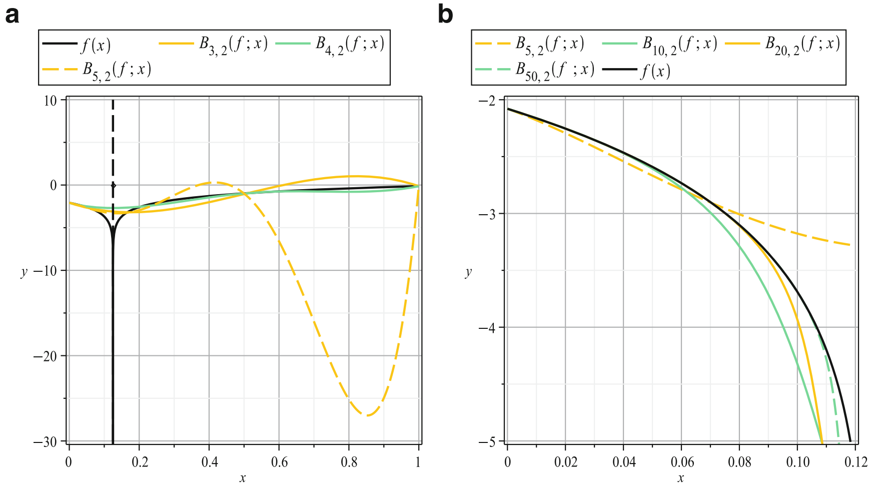


Fig. 9 Graphs of the function $y = f(x)$ and its q-Bernstein polynomial $B_{n,2}(f;x)$ for $n = 3, 4, 5, 10, 20$, and 50

Theorem 2. *(i) Let* $f \in \mathscr{F}_0$, *so that* $\lim_{x\to 0+} x^{\gamma} f(x) = K \neq 0$ *for some* $\gamma > 0$ *and* $f(0) = A \in \mathbb{R}$. *Then, for* $q \geq 2$,

$$|B_{n,q}(f;x)| \to \infty \quad as \quad n \to \infty \quad \forall \quad x \in \mathbb{R} \setminus \mathbb{J}_q.$$

(ii) Let $f \in \mathscr{F}_0$, *so that* $\lim_{x\to 0+} x^{-m} f(x) = K \neq 0$ *for some* $m \in \mathbb{N}$ *and* $f(0) = A \in \mathbb{R}$. *Then,*

$$|B_{n,q}(f;x)| \to \infty \quad as \quad n \to \infty \quad \forall \quad x \in \mathbb{R} \setminus \mathbb{J}_q.$$

Example 3. Let $f(x) = \frac{1}{\sqrt[3]{x}}, x \neq 0$ and $f(0) = 0$.

Figure 3a and b demonstrates the graphs of the function f and its q-Bernstein polynomials of degree $2, 3, 4$, and 5. Both the convergence to f on the time scale and the divergence elsewhere are noticeable.

Example 4. Let $f(x) = \ln x, x \neq 0$ and $f(0) = 0$.

The graphs of f and its q-Bernstein polynomials $B_{n,q}(f;x)$ when $n = 2, 3, 4$ and 5 are shown in Fig. 4a and b. The general behavior of the polynomials regarding the convergence is similar to the preceding example.

Singularity at 1

It can be readily seen that since $x = 1$ is an isolated node for all q-Bernstein polynomials, the singularity at 1 does not affect the expression for the $\{B_{n,q}(f;x)\}$. As a result, the sequence $\{B_{n,q}(f;x)\}$ is pointwise convergent to $f(x)$ on the entire interval [0, 1], while the convergence on $\mathbb{J}_q$ is uniform. More precisely, the following statement holds.

Theorem 3. *Let $f \in \mathscr{F}_1$. Then,*

$$B_{n,q}(f;x) \to f(x) \ \textit{uniformly on any compact subset of} \ (-1,1)$$

and, consequently, $B_{n,q}(f;x) \to f(x)$ at every $x \in [0,1]$.

The assertion has been previously obtained only for rational functions on [0, 1]—see [13] and [15]—but it is not difficult to extend it for the entire class $\mathscr{F}_1$.

Example 5. $f(x) = \frac{1}{(\sin(\frac{\pi}{2}x)-1)^2}, x \neq 1$ and $f(1) = 0$.

The graphs of f and $B_{n,q}(f;x)$ are given in Fig. 5a and b. Since f is unbounded on [0, 1], its uniform approximation on the entire interval is not possible. The figure illustrates the pointwise approximation on [0, 1]; however, one can observe that the approximation at the points close to 1 is rather slow.

Singularity at q^{-m}, $m \in \mathbb{N}$

This case is the most interesting one because a new remarkable phenomenon occurs at this stage: *the interval of convergence is governed by the type of singularity at* $\boldsymbol{\alpha}$. A comprehensive analysis of this situation has been performed in [16], while certain preliminary efforts were made in [13, 15].

Further, the following subsets of $\mathscr{F}_\alpha$—depending on the type of singularity—are taken into account:

- $\mathscr{A} = \{f \in \mathscr{F}_\alpha$: there exists $\gamma > 0$, such that $\lim_{x\to\alpha^-} f(x)(\alpha - x)^\gamma = K \in \mathbb{R} \setminus \{0\}\}$.
- $\mathscr{B} = \{f \in \mathscr{F}_\alpha : \lim_{x\to\alpha^-} f(x)(\alpha - x)^\gamma = \infty$ for all $\gamma > 0\}$.
- $\mathscr{C} = \{f \in \mathscr{F}_\alpha : \lim_{x\to\alpha^-} f(x)(\alpha - x)^\gamma = 0$ for all $\gamma > 0\}$.

Given this classification, the sets of convergence for the q-Bernstein polynomials are identified.

The first result is related to class $\mathscr{A}$. It turns out that γ, the parameter describing the singularity of f at α, is crucial for finding the interval of convergence for polynomials $B_{n,q}(f;\cdot)$.

Theorem 4. *Let $f \in \mathscr{A}$ and $\alpha = q^{-m}, m \in \mathbb{N}$. Then,*

(i) $B_{n,q}(x; q^{-l}) \to f(q^{-l})$, $l = 0, 1, \ldots, m-1$ *and* $l > m + \gamma$;
(ii) $f(x)$ *is uniformly approximated by* $B_{n,q}(f; x)$ *on any compact subset of* $(-\alpha q^{-\gamma}, \alpha q^{-\gamma})$;
(iii) if $|x| > \alpha q^{-\gamma}$ *and* $x \notin \mathbb{J}_q$, *then* $|B_{n,q}(f; x)| \to \infty$ *as* $n \to \infty$.

It appears that (i) is sharp in the sense that, for the other points of the time scale, this approximation does not occur.

Example 6. Let $f(x) = \frac{1}{(x-\frac{1}{2})}, x \neq \frac{1}{2}$ and $f(\frac{1}{2}) = 0$.

The graphs of f and the associated q-Bernstein polynomials for $n = 2, 3, 4, 5$, and 6 are given in Fig. 6a–c. We have $m = 1, \alpha = \frac{1}{2}$ and $\gamma = 1$. While the uniform approximation on any compact subset of $(0, \frac{1}{4})$ is easily viewed in Fig. 6b and c, the divergence for $x > \frac{1}{4}$ when $x \notin \mathbb{J}_q$ can also be detected in Fig. 6 as a whole.

Example 7. Let $f(x) = \frac{1}{(x-\frac{1}{2})^2}, x \neq \frac{1}{2}$ and $f(\frac{1}{2}) = 0$.

Figure 7a–c supplies the graphs of the function f and the corresponding q-Bernstein polynomials for $n = 4, 5, 6, 7, 8, 9, 10$, and 11. Here, $m = 1, \alpha = \frac{1}{2}$ but, in distinction from the previous example, $\gamma = 2$. As a result, the interval on which the convergence occurs has been halved compared with Example 6. Apart from that, the situation in terms of approximation is rather similar, and only the number of time scale points at which $B_{n,2}(f; x)$ converge to $f(x)$ is reduced by 1 point.

On a different note, for functions belonging to $\mathscr{B}$, the q-Bernstein polynomials have very poor convergence properties. This fact is expressed by:

Theorem 5. *Let $f \in \mathscr{B}$. Then,* $B_{n,q}(f; x) \to f(x)$ *if* $x \in \{1, q^{-1}, \ldots, q^{-(m-1)}, 0\}$ *and* $|B_{n,q}(f; x)| \to \infty$, *otherwise.*

To summarize, the set of convergence for $\{B_{n,q}(f; x)\}$ is reduced just to a finite number of points in $\mathbb{J}_q$.

Example 8. Let $f(x) = e^{(x-\frac{1}{2})^{-\frac{1}{5}}}, x \neq \frac{1}{2}$ and $f(\frac{1}{2}) = 0$.

The graphs of f and $B_{n,q}(f; x)$ are given in Fig. 8a–c, which pictures the lack of convergence on $(0, 1)$ for the q-Bernstein polynomials of f. The convergence at $x = 0$ and $x = 1$ originates from the end-point interpolation property.

Finally, let $f \in \mathscr{C}$, that is, f possesses a relatively 'mild' singularity at α. Then, it has been established that the approximation by the q-Bernstein polynomials occurs on the same interval as by the Maclaurin polynomials.

Theorem 6. *Let $f \in \mathscr{C}$. Then,* $B_{n,q}(f; x) \to f(x)$ *uniformly on any compact subset of* $(-\alpha, \alpha)$.

The statement of Theorem 6 contains no information on the convergence of $\{B_{n,q}(f; \cdot)\}$ outside of $(-\alpha, \alpha)$. This is because, in general, the functions of class $\mathscr{C}$ may have a wider interval of the approximation than $(-\alpha, \alpha)$. Therefore, it is beneficial to demonstrate the sharpness of Theorem 6. That is, to show that there exists $f \in \mathscr{C}$, such that its q-Bernstein polynomials diverge for all $|x| > \alpha,\ x \notin \mathbb{J}_q$. This sharpness is achieved by the following result proved in [16].

Theorem 7. *Let $f \in \mathscr{C}$ and $f(x) = g(x)\ln|x-\alpha|$ for $x \in [0, \alpha)$, where the function g admits an analytic continuation from $[0, \alpha)$ into $\mathscr{D}_\beta$ with $\beta > \alpha$. Then, $|B_{n,q}(f; x)| \to \infty$ for $|x| > \alpha$ with $x \notin \mathbb{J}_q$.*

Example 9. Let $f(x) = \ln\left|x - \frac{1}{8}\right|, x \neq \frac{1}{8}$ and $f(\frac{1}{8}) = 0$.

The graphs of f and corresponding q-Bernstein polynomials for $n = 3, 4, 5, 10, 20,$ and 50 are shown in Fig. 9a and b. Here, the q-Bernstein polynomials uniformly approximate f on all compact subsets of $(-\alpha, \alpha)$ and, hence, on the compact subsets of $[0, \alpha) = [0, \frac{1}{8})$, as can be observed in the figure. In addition, outside of the interval $(0, \frac{1}{8})$, the convergence to f appears only on the time scale.

Acknowledgement We would like to express our sincere appreciation to Mr. P. Danesh from Atilim University Academic Writing and Advisory Center for his valuable help in the preparation of the manuscript.

References

1. Agratini, O.: On certain q-analogues of the Bernstein operators. Carpathian J. Math. **20**(1), 1–6 (2008)
2. Andrews, G.E., Askey, R., Roy, R.: Special Functions. Cambridge University Press, Cambridge (1999)
3. Bernstein, S.N.: Démonstration du théorème de Weierstrass fondée sur la calcul des probabilités. Commun. Soc. Math. Charkow Série 2 **13**, 1–2 (1912)
4. Bohner, M., Guseinov, G., Peterson, A.: Introduction to the time scales calculus. In: Advances in Dynamic Equations on Time Scales, pp. 1–15. Birkhäuser, Boston, MA (2003)
5. Charalambides, C.A.: The q-Bernstein basis as a q-binomial distribution. J. Stat. Plan. Infer. **140**(8), 2184–2190 (2010)
6. Il'inskii, A.: A probabilistic approach to q-polynomial coefficients, Euler and Stirling numbers I. Mat. Fiz. Anal. Geom. **11**(4), 434–448 (2004)
7. Jing, S.: The q-deformed binomial distribution and its asymptotic behaviour. J. Phys. A: Math. Gen. **27**, 493–499 (1994)
8. Kantorovich, L.V.: La representation explicite d'une fonction mesurable arbitraire dans la forme de la limite d'une suite de polynomes. Math. Sbornik **41**, 508–510 (1934)
9. Lorentz, G.G.: Bernstein Polynomials. Chelsea, New York (1986)
10. Lupaş, A.: A q-analogue of the Bernstein operator. University of Cluj-Napoca, Seminar on numerical and statistical calculus, University of Babeş-Bolyai, Cluj-Napoca, vol. 87–89, pp. 85–92 (1987)
11. Ostrovska, S.: On the approximation of analytic functions by the q-Bernstein polynomials in the case $q > 1$. Electron. Trans. Numer. Anal. **37**, 105–112 (2010)
12. Ostrovska, S.: On the q-Bernstein polynomials of the logarithmic function in the case $q > 1$. Math. Slovaca (in press)
13. Ostrovska, S., Özban, A.Y.: The q-Bernstein polynomials of the Cauchy kernel with a pole on $[0, 1]$. Appl. Math. Comp. **220**, 735–747 (2013)

14. Ostrovska, S., Özban, A.Y.: On the q-Bernstein polynomials of unbounded functions with $q > 1$. Abstr. Appl. Anal. **2013**, 8 (2013). Article ID 349156
15. Ostrovska, S., Özban, A.Y.: On the q-Bernstein polynomials of rational functions with real poles. J. Math. Anal. Appl. **413**, 547–556 (2014)
16. Ostrovska, S., Özban, A.Y., Turan, M.: How do singularities of functions affect the convergence of q-Bernstein polynomials? J. Math. Inequal. **9**(1), 121–136 (2015)
17. Phillips, G.M.: Bernstein polynomials based on the q-integers. Ann. Numer. Math. **4**, 511–518 (1997)
18. Titchmarsh, E.C.: The Theory of Functions. Oxford University Press, Oxford (1986)
19. Videnskii, V.S.: Bernstein Polynomials. Leningrad State Pedagogical University, Leningrad (1990, Russian)
20. Videnskii, V.S.: On the centenary of the discovery of Bernstein polynomials. Some current problems in modern mathematics and education in mathematics, pp. 5–11, Izdat. RGPU im. A.I. Gertsena, St. Petersburg (2013, Russian)
21. Videnskii, V.S.: Papers of L.V. Kantorovich on Bernstein polynomials. Vestnik St. Petersburg Univ. Math. **46**(2), 85–88 (2013)
22. Weba, M.: On the approximation of unbounded functions by Bernstein polynomials. East J. Approx. **9**(3), 351–356 (2003)

Nests, and Their Role in the Orderability Problem

Kyriakos Papadopoulos

Abstract This chapter is divided into two parts. The first part is a survey of some recent results on nests and the orderability problem. The second part consists of results, partial results and open questions, all viewed in the light of nests. From connected LOTS, to products of LOTS and function spaces, up to the order relation in the Fermat Real Line.

Keywords Nests • Orderability problem • LOTS • Products of LOTS • Function spaces • Fermat reals

Introduction

> ... confusion connotes something which possesses no ***order***,
> the individual parts of which are so strangely admixed and interwined,
> that it is impossible to detect where each element actually belongs...

(Extract from The Musical Dialogue, by *Nikolaus Harnoncourt*, Amadeus Press, 1997.)

What is an orderability theorem? In particular in S. Purisch's account of results on orderability and suborderability (see [12]), one can read the formulation and development of several orderability theorems, starting from the beginning of the twentieth century and reaching our days. By an orderability theorem, in topology, we mean the following. Let $(X, \mathscr{T})$ be a topological space. Under what conditions does there exist an order relation $<$ on X such that the topology $\mathscr{T}_<$ induced by the order $<$ is equal to $\mathscr{T}$? As we can see, this problem is very fundamental as it is of the same weight as the metrizability problem, for example (let X be a topological space: is there a metric d, on X, such that the metric topology generated by this metric to be equal to the original topology of X?).

K. Papadopoulos (✉)
College of Engineering and Technology, American University of the Middle East, Egaila, Kuwait
e-mail: kyriakos.papadopoulos1981@gmail.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_23

Some History

> **Order** is a concept as old as the idea of number and much of early mathematics was devoted to constructing and studying various subsets of the real line. (Steve Purisch [12])

The great German mathematician Georg Cantor (1845–1918) is credited to be one of the inventors of set theory. This fact makes him automatically one of the inventors of order-theory as well, as he is the one who first introduced the class of cardinals and the class of ordinal numbers, two classes of rich order-theoretic properties. Cantor was not only interested in defining classes of ordered sets, and studying their arithmetic; he also produced major results while examining order-isomorphisms, that is, bijective order-preserving mappings between sets whose inverses are also order-preserving. S. Purisch gives a complete list of these historic papers written by Cantor, in his article "A History of Results on Orderability and Suborderability" [12].

Together with set theory, the field of topology met a rapid rising in the early twentieth century and new problems, combining both fields, appeared. A topologist's temptation is always to examine what sort of topology can be introduced in a given set. So, a very early question was what is the relationship between the natural topology of a set and the topology which is induced by an ordering in this set; this question led to the formulation of the *orderability problem*.

According to Purisch, one of the earliest orderability theorems was introduced by O. Veblen and N.J. Lennes, who were both students of the American mathematician E.H. Moore (1862–1932), and who attended his geometry seminar. This theorem stated that *every metric continuum, with exactly two non-cut points, is homeomorphic to the unit interval*. For the statement of the theorem, Veblen combined the notions of ordered set and topology, for defining a simple arc. Lennes used up-to-date machinery to prove Veblen's statement, a proof that was published in 1911.

In the meanwhile, some of the greatest mathematicians of the first half of the twentieth century, like the French mathematicians R. Baire, M. Fréchet, the Dutch mathematician L.E.J. Brouwer, the Jewish-German mathematician F. Hausdorff, the Polish mathematicians S. Mazurkiewicz, W. Sierpínski, the Russian mathematicians P. Alexandroff and P. Urysohn and others, were devoted to constructing various subsets of the real line. In particular, Baire used ideas of the Yugoslavian mathematician D. Kurepa and of the Dutch mathematician A.F. Monna, on non-Archimedean spaces, in order to characterize the set of irrational numbers. The British mathematician, A.J. Ward, found a topological characterization of the real line (1936), stating that *the real line is homeomorphic to a separable, connected and locally connected metric space* X*, such that* $X - \{p\}$ *consists of exactly two components, for every* $p \in X$.

A more general result (1920), by Mazurkiewicz and Sierpínski, stated that *compact, countable metric spaces are homeomorphic to well-ordered sets*; this is one of the first, if not the first, topological characterizations of abstract ordered sets.

Having in mind that a special version of the orderability problem was solved in the beginning of the 1970s (J. van Dalen and E. Wattel), its formulation started from the beginning of the 1940s. In particular, the Polish-American mathematician S. Eilenberg, gave in 1941 the following result: *a connected space, X, is weakly orderable, if and only if* $X \times X$ *minus the diagonal is not connected*. This condition is also necessary and sufficient for a connected, locally connected space to be orderable.

The American mathematician E. Michael extended this work and showed, in 1951, that *a connected Hausdorff space X is a weakly orderable space, if and only if X admits a continuous selection*.

It took two more decades, for a complete topological characterization of GO-spaces and LOTS to appear. In 1972 J. de Groot and P.S. Schnare showed [2] that *a compact* T_1 *space X is LOTS, if and only if there exists an open subbase* $\mathscr{S}$ *of X which is the union of two nests, such that every cover of the space, by elements of* $\mathscr{S}$*, has a two element subcover*. **J. van Dalen and E. Wattel** used the characterization of de Groot and Schnare as a basis for their construction [20], which led to a solution of the orderability problem via nests. We revisited van Dalen and Wattel's characterization in [8], and we introduced a simpler proof of their main characterization theorem.

In [16] we analyzed further the order-theoretic results from [10] and gave a list of open problems.

The study of ordered spaces did not finish with the solution to the orderability problem that was proposed by van Dalen and Wattel. On the contrary, many interesting and important results have appeared since then. We will now refer to those results which have motivated our own research in particular.

In 1986, G.M. Reed published an article with title "The Intersection Topology w.r.t. the Real Line and the Countable Ordinals" [13]. The author constructed there a class which was shown to be a surprisingly useful tool in the study of abstract spaces. We know that if $\mathscr{T}_1, \mathscr{T}_2$ are topologies on a set X, then the *intersection topology*, with respect to $\mathscr{T}_1$ and $\mathscr{T}_2$, is the topology $\mathscr{T}$ on X such that the set $\{U_1 \cap U_2 : U_1 \in \mathscr{T}_1 \text{ and } U_2 \in \mathscr{T}_2\}$ forms a base for $(X, \mathscr{T})$. Reed introduced the class $\mathscr{C}$, where $(X, \mathscr{T}) \in \mathscr{C}$ if and only if $X = \{x_\alpha : \alpha < \omega_1\} \subset \mathbb{R}$, where $\mathscr{T}_1 = \mathscr{T}_{\mathbb{R}}$ and $\mathscr{T}_2 = \mathscr{T}_{\omega_1}$ and $\mathscr{T}$ is the intersection of $\mathscr{T}_{\mathbb{R}}$ (the subspace real line topology on X) and $\mathscr{T}_{\omega_1}$ (the order topology on X, of type ω_1). In particular, Reed showed that if $(X, \mathscr{T}) \in \mathscr{C}$, then X has rich topological, but not very rich order-theoretic properties. In particular, X is a completely regular, submetrizable, pseudo-normal, collectionwise Hausdorff, countably metacompact, first countable, locally countable space, with a base of countable order, that is neither subparacompact, metalindelöf, cometrizable nor locally compact. That an $(X, \mathscr{T}) \in \mathscr{C}$ does not necessarily have rich order-theoretic properties comes from the fact that there exists, in ZFC, an $(X, \mathscr{T}) \in \mathscr{C}$ which is not normal.

Eric K. van Douwen characterized in 1993 [15] the noncompact spaces, whose every noncompact image is orderable, as the noncompact continuous images of ω_1. Van Douwen refers to a closed non-compact set as *cub* (corresponding to closed unbounded sets in ordinals—also met as club in the literature), and he calls *bear* a space which is noncompact and has no disjoint cubs. Here we state his result that has motivated our research on ordinals (see [8]):

For a noncompact space X, the following are equivalent:

1. X is a continuous image of ω_1.
2. Every noncompact continuous image of X is orderable.
3. X is scattered first countable orderable bear.
4. X is locally countable orderable bear.
5. X has a compatible linear order, all initial closed segments of which are compact and countable.

A Survey of Recent Results on Nests

Characterizations of LOTS

As we also mentioned in section "Some History," van Dalen and Watten used nests in order to give a solution to the orderability problem, and in [8] we gave a more order- and set-theoretic dimension to this characterization. In particular, we did not declare our space being T_1, but its topology generated by a (so-called) T_1-separating subbase.

Definition 1. Let X be a set. We say that a collection of subsets $\mathscr{S}$ of X:

1. T_0-*separates* X, if and only if for all $x, y \in X$, such that $x \neq y$, there exists $S \in \mathscr{S}$, such that $x \in S$ and $y \notin S$ or $y \in S$ and $x \notin S$,
2. T_1-*separates* X, if and only if for all $x, y \in X$, such that $x \neq y$, there exist $S, T \in \mathscr{S}$, such that $x \in S$ and $y \notin S$ and also $y \in T$ and $x \notin T$ and

One can easily see that a space is T_0 (resp. T_1) if and only if its topology is generated by a T_0- (resp. T_1-) separating subbase, but the statement of Definition 1 is not valid for the T_2 separation axiom, if one defines a T_2-separating subbase in an analogous way.

Definition 2. Let X be a set and let $\mathscr{L} \subset \mathscr{P}(X)$. We define an order $\lhd_{\mathscr{L}}$ on X by declaring that $x \lhd_{\mathscr{L}} y$, if and only if there exists some $L \in \mathscr{L}$, such that $x \in L$ and $y \notin L$.

In [8] we showed the close link between nests and linear orders in Theorem 1 that follows below.

Theorem 1. *Let X be a set and let $\mathscr{L} \subset \mathscr{P}(X)$. Then, the following hold:*

1. *If $\mathscr{L}$ is a nest, then $\lhd_{\mathscr{L}}$ is a transitive relation.*
2. *$\mathscr{L}$ is a nest, if and only if for every $x, y \in X$, either $x = y$ or $x \not\lhd_{\mathscr{L}} y$ or $y \not\lhd_{\mathscr{L}} x$.*

3. *$\mathscr{L}$ is T_0-separating, if and only if for every $x, y \in X$, either $x = y$ or $x \lhd_{\mathscr{L}} y$ or $y \lhd_{\mathscr{L}} x$.*
4. *$\mathscr{L}$ is a T_0-separating nest, if and only if $\lhd_{\mathscr{L}}$ is a linear order.*

We still needed some more tools, in order to restate van Dalen and Wattel's characterization theorem in a more elementary way. Theorem 2 shows the connection between a subbase which generates a GO-topology and two T_0-separating nests with reverse orders, whose union T_1-separates the space.

Theorem 2. *Let X be a set. Suppose that $\mathscr{L}$ and $\mathscr{R}$ are two nests on X. Then, $\mathscr{L} \cup \mathscr{R}$ is T_1-separating, if and only if $\mathscr{L}$ and $\mathscr{R}$ are both T_0-separating and $\lhd_{\mathscr{L}} = \rhd_{\mathscr{R}}$.*

A key tool, for van Dalen and Wattel's solution of the Orderability Problem, was the notion of interlocking.

Definition 3. Let X be a set and let $\mathscr{S} \subset \mathscr{P}(X)$. We say that $\mathscr{S}$ is interlocking, if and only if for each $T \in \mathscr{S}$, such that:

$$T = \bigcap \{S : T \subset S, S \in \mathscr{S} - \{T\}\}$$

we have that:

$$T = \bigcup \{S : S \subset T, S \in \mathscr{S} - \{T\}\}.$$

By Lemma 1 that follows, we clarified the relationship between an interlocking nest and the properties of its induced order.

Lemma 1. *Let X be a set and let $\mathscr{L}$ be a T_0-separating nest on X. Then, the following hold for $L \in \mathscr{L}$:*

1. *$L = \bigcap \{M \in \mathscr{L} : L \subsetneq M\}$, if and only if $X - L$ has no $\lhd_{\mathscr{L}}$-minimal element.*
2. *$L = \bigcup \{M \in \mathscr{L} : M \subsetneq L\}$, if and only if L has no $\lhd_{\mathscr{L}}$-maximal element.*

It is immediate, from Definition 3, that a collection $\mathscr{L}$ is interlocking, if and only if for all $L \in \mathscr{L}$, either $L = \bigcup \{N \in \mathscr{L} : N \subsetneq L\}$ or $L \neq \bigcap \{N \in \mathscr{L} : L \subsetneq N\}$. So, we observed that Theorem 1 and Lemma 1 therefore imply the following.

Theorem 3. *Let X be a set and let $\mathscr{L}$ be a T_0-separating nest on X. The following are equivalent:*

1. *$\mathscr{L}$ is interlocking;*
2. *for each $L \in \mathscr{L}$, if L has a $\lhd_{\mathscr{L}}$-maximal element, then $X - L$ has a $\lhd_{\mathscr{L}}$-minimal element;*
3. *for all $L \in \mathscr{L}$, either L has no $\lhd_{\mathscr{L}}$-maximal element or $X - L$ has a $\lhd_{\mathscr{L}}$-minimal element.*

So, Theorem 3 is a specific version of the notion interlocking in the case of a linearly ordered topological space, and this gave us enough tools to prove the following alteration of van Dalen and Wattel's Theorem:

Theorem 4 (van Dalen & Wattel). *Let $(X, \mathscr{T})$ be a topological space. Then:*

1. *If $\mathscr{L}$ and $\mathscr{R}$ are two nests of open sets, whose union is T_1-separating, then every $\lhd_{\mathscr{L}}$-order open set is open, in X.*
2. *X is a GO space, if and only if there are two nests, $\mathscr{L}$ and $\mathscr{R}$, of open sets, whose union is T_1-separating and forms a subbase for $\mathscr{T}$.*
3. *X is a LOTS, if and only if there are two interlocking nests $\mathscr{L}$ and $\mathscr{R}$, of open sets, whose union is T_1-separating and forms a subbase for $\mathscr{T}$.*

Characterizations of Ordinals

Ordinals, like LOTS and GO-spaces, are fundamental building blocks for set-theoretic and topological examples. In [8] we used properties of nests in order to characterize ordinals topologically. To achieve this, we considered our spaces to be "scattered by a nest".

Definition 4. A topological space X is *scattered*, if every non-empty subset $A \subset X$ has an isolated point, i.e. for every non-empty $A \subset X$, there exists $a \in A$ and U open in X, such that $U \cap A = \{a\}$.

Therefore, a space X is *scattered*, if for every non-empty $A \subset X$, there exists U open in X, such that $|U \cap A| = 1$.

Definition 5. Let $\mathscr{S}$ be a family of subsets of a set X. We say that X is *scattered* by $\mathscr{S}$, if and only if for every $A \subset X$, there exists $S \in \mathscr{S}$, such that $|A \cap S| = 1$.

Theorem 5. *Let X be a set and let $\mathscr{L}$ be a nest on X. Then, the following are equivalent:*

1. *$\mathscr{L}$ scatters X.*
2. *$\lhd_{\mathscr{L}}$ is a well-order on X.*
3. *$\mathscr{L}$ is T_0-separating and well ordered by $\subset$.*
4. *$\mathscr{L}$ is T_0-separating and, for every non-empty subset A of X, there is an $a \in A$, such that for any $x \in A$ and any $L \in \mathscr{L}$, if $x \in L$, then $a \in L$.*

Theorem 6. *Let X be a space. The following are equivalent:*

1. *X is homeomorphic to an ordinal.*
2. *X has two interlocking nests $\mathscr{L}$ and $\mathscr{R}$, of open sets, whose union is a T_1-separating subbase, such that $\mathscr{L}$ scatters X.*
3. *X has two interlocking nests $\mathscr{L}$ and $\mathscr{R}$, of open sets, whose union is a T_1-separating subbase, one of which is well-ordered by $\subset$ or $\supset$.*

4. *X is scattered by a nest* $\mathscr{L}$*, of clopen sets, such that:*
   a. $L \neq \bigcup\{M : M \subsetneq L\}$*, for any* $L \in \mathscr{L}$ *and*
   b. $\{L - M : L, M \in \mathscr{L}\}$ *is a base for X.*
5. *X is scattered by a nest of compact clopen sets.*

Corollary 1 that follows leaded us to a characterization of the ordinal ω_1, with clear links to the well-known Pressing (or Pushing) Down Lemma in Set Theory.

Corollary 1. *X is homeomorphic to a cardinal, if and only if X is scattered by a nest* $\mathscr{L}$*, of compact clopen sets, such that* $|L| < |X|$*, for each* $L \in \mathscr{L}$*.*

In particular, X is homeomorphic to ω_1*, if and only if X is uncountable and is scattered by a nest of compact, clopen, countable sets.*

A Generalization of the Orderability Problem

In [11], we restated Theorem 4 via the interval topology, in the corollary that follows.

Corollary 2. *A topological space* $(X, \mathscr{T})$ *is:*

1. *a LOTS, iff there exists a nest* $\mathscr{L}$ *on X, such that* $\mathscr{L}$ *is* T_0*-separating and interlocking and also* $\mathscr{T} = \mathscr{T}_{in}^{\trianglelefteq_{\mathscr{L}}}$.
2. *a GO-space, iff there exists a nest* $\mathscr{L}$ *on X, such that* $\mathscr{L}$ *is* T_0*-separating and also* $\mathscr{T} = \mathscr{T}_{in}^{\trianglelefteq_{\mathscr{L}}}$.

An answer to the following question will give a weaker orderability theorem.

Question Let X be a set equipped with a transitive relation $<$ and the interval topology $\mathscr{T}_{in}^{\leq}$, defined via $\leq$, where $\leq$ is $<$ plus reflexivity. Under which necessary and sufficient conditions will $\mathscr{T}_{<}$ be equal to $\mathscr{T}_{in}^{\leq}$?

Some New Thoughts

Connectedness and Orderability

In this section we give a characterization of interlockingness via connectedness. This will give a condition for a connected space to be LOTS.

Definition 6. A partial order $<$, on a set X, is said to be dense if, for all x and y in X for which $x < y$, there exists some z in X, such that $x < z < y$.

So, given Definition 6, the next lemma follows naturally.

Lemma 2. *Let X be a set and let $\mathscr{L}$ be a nest on X. Then, the order $\lhd_{\mathscr{L}}$ is dense in X, if and only if for every $x, y \in X$, $x \neq y$, there exist $L, M \in \mathscr{L}$, $L \subsetneq M$, such that $x \in L$ and $y \notin M$ or $y \in L$ and $x \notin M$.*

Proposition 1. *Let X be a set and let $\mathscr{L}, \mathscr{R}$ be two nests of open sets on X, such that $\mathscr{L} \cup \mathscr{R}$ creates a T_1-separating subbase for a topology on X. If X is connected, with respect to the topology that is induced by the union of $\mathscr{L}$ and $\mathscr{R}$, then $\lhd_{\mathscr{L}}$ is dense in X.*

Proof. Suppose $\lhd_{\mathscr{L}}$ is not dense. Then, there exist $x, y \in X$, such that $(x, y) = \emptyset$. So, there exists $L \in \mathscr{L}$, such that $x \in L$ and $y \notin L$ and there also exists $R \in \mathscr{R}$, such that $x \notin R$ and $y \in R$ and also $L \cap R = \emptyset$ and $L \cup R = X$. So, X is not connected.

In Theorem 3 we described interlocking nests, in terms of maximal and minimal elements. Here we use this result, in order to give a characterization of connected spaces via nests.

Theorem 7. *Let X be a set and let $\mathscr{L}, \mathscr{R}$ be two nests of open sets on X, such that $\mathscr{L} \cup \mathscr{R}$ creates a T_1-separating subbase for a topology on X. If X is connected, with respect to the topology with subbase $\mathscr{L} \cup \mathscr{R}$, then $\mathscr{L}$ and $\mathscr{R}$ are interlocking nests.*

Proof. If $\mathscr{L}$ is not interlocking then, according to Theorem 3, there exists $L \in \mathscr{L}$, such that $L = (-\infty, x]$, but $X - L$ has no minimal element. The set L is open, as a subbasic element for the topology that is generated by $\mathscr{L} \cup \mathscr{R}$. So, for every $z \in X - L$, there exists z', such that $x \lhd_{\mathscr{L}} z' \lhd_{\mathscr{L}} z$. But, there exists $R_z \in \mathscr{R}$, such that $z' \notin R_z$ and $z \in R_z$. So, $X - L = \bigcup_{z \notin L} R_z$, i.e. $R_z \cap L = \emptyset$. Thus, $X - L$ is open and L is open, hence X is not connected. In a similar way, $\mathscr{R}$ is interlocking, too.

Theorem 7 permits us now to view LOTS, in the light of connectedness.

Corollary 3. *Let X be a set and let $\mathscr{L}, \mathscr{R}$ be two nests of open sets on X, such that $\mathscr{L} \cup \mathscr{R}$ creates a T_1-separating subbase for a topology on X. If X is connected with respect to the topology with subbase $\mathscr{L} \cup \mathscr{R}$, then X is a LOTS.*

Proof. The proof follows immediately from the statements of Theorem 4 and Theorem 7.

Powers of LOTS

Let I be a set of indices. Let X be a LOTS and let π its i-th canonical projection. Here we examine properties of powers of LOTS, linking X with X^I via projections.

Proposition 2. *Let X be a LOTS and let $\mathscr{L}_{X^I}$ be a nest on X^I. Then, $\pi_i(\mathscr{L}_{X^I}) = \{\pi_i(L) : L \in \mathscr{L}_{X^I}\}$ will be a nest on X, for every $i \in I$.*

Proof. Let $\pi_i(L_1), \pi_i(L_2) \in \pi_i(\mathscr{L}_{X^I})$, where $L_1, L_2 \in \mathscr{L}_{X^I}$. Then, $L_1 \subset L_2$ or $L_2 \subset L_1$, which implies that $\pi_i(L_1) \subset \pi_i(L_2)$ or $\pi_i(L_2) \subset \pi_i(L_1)$, proving that $\pi_i(\mathscr{L}_{X^I})$ is a nest, too.

Proposition 3. *Let $\mathscr{L}_X$ be a nest on X. Then, $\pi_i^{-1}(\mathscr{L}_X) = \{\pi_i^{-1}(L) : L \in \mathscr{L}_X\}$ will be a nest on X^I, for every $i \in I$.*

Proof. Let $\pi_i^{-1}(L_1), \pi_i^{-1}(L_2) \in \pi_i^{-1}(\mathscr{L}_X)$. Since $\mathscr{L}_X$ is a nest, then either $L_1 \subset L_2$ or $L_2 \subset L_1$. If $L_1 \subset L_2$, then $\pi_i^{-1}(L_1) \subset \pi_i^{-1}(L_2)$, and if $L_2 \subset L_1$, then $\pi_i^{-1}(L_2) \subset \pi_i^{-1}(L_1)$. Thus, $\pi_i^{-1}(\mathscr{L}_X)$ will be a nest, too.

Definition 7. Let X be a set, and let $\mathscr{L}_{X^I}$ be a nest on X^I, satisfying the condition that if $(x_i)_{i\in I}, (y_i)_{i\in I} \in X^I$, such that $x_j \neq y_j$, $j \in I$, then there exists $L \in \mathscr{L}_{X^I}$, such that $(x_i)_{i\in I} \in L$ and $(y_i)_{i\in I} \notin L$ or $(y_i)_{i\in I} \in L$ and $(x_i)_{i\in I} \notin L$. Then, we say that the nest $\mathscr{L}_{X^I}$ is *weakly T_0-separating*, with respect to the j-th variable.

Definition 8. Let X be a set and let $\mathscr{L}_{X^I}$ be a nest on X^I. Let also $(x_i)_{i\in I}$, $(y_i)_{i\in I}$, be such that $x_j \neq y_j$, for a fixed $j \in I$. Then, we define $(x_i)_{i\in I} \lhd_{\mathscr{L}_{X^I}} (y_i)_{i\in I}$, if there exists a set $L \in \mathscr{L}_{X^I}$, such that $(x_i)_{i\in I} \in L$ and $(y_i)_{i\in I} \notin L$.

Theorem 8. *If $\mathscr{L}_{X^I}$ is a weakly T_0-separating nest on X^I, with respect to the j-th variable, such that it satisfies the condition that if $(x_i)_{i\in I} \notin L \in \mathscr{L}_{X^I}$, then $x_j \notin \pi_j(L)$, then $\pi_j(\mathscr{L}_{X^I}) = \{\pi_j(L) : L \in \mathscr{L}_{X^I}\}$ is a T_0-separating nest on X.*

Proof. Proposition 2 gives that $\pi_j(\mathscr{L}_{X^I})$ is a nest.

For proving that $\pi_j(\mathscr{L}_{X^I})$ is T_0-separating, let $x_1, x_2 \in X$, such that $x_1 \neq x_2$. Then, we form $(y_i)_{i\in I}$, $(z_i)_{i\in I}$, so that we place x_1 in the j-th position of $(y_i)_{i\in I}$ and x_2 in the j-th position of $(z_i)_{i\in I}$. The rest y_i and z_i are considered arbitrary.

Since $\mathscr{L}_{X^I}$ is T_0-separating, with respect to the j-th variable, then there exists $L \in \mathscr{L}_{X^I}$, such that $(y_i)_{i\in I} \in L$ and $(z_i)_{i\in I} \notin L$ or $(z_i)_{i\in I} \in L$ and $(y_i)_{i\in I} \notin L$.

So, $\pi_j((y_i)_{i\in I}) = x_1 \in \pi_j(L)$ and $\pi_j((z_i)_{i\in I}) = x_2 \notin \pi_j(L)$ or $\pi_j((z_i)_{i\in I}) = x_2 \in \pi_j(L)$ and $\pi_j((y_i)_{i\in I}) = x_1 \notin \pi_j(L)$, which proves that $\pi_j(\mathscr{L}_{X^I})$ is T_0-separating.

Remark 1. Let $\mathscr{L}_{X^I}$ be a weakly T_0-separating nest in X^I. Then, if $(y_i)_{i\in I}$, $(z_i)_{i\in I}$ have in the j-th position the elements y_j and z_j, respectively, then $(y_i)_{i\in I} \lhd_{\mathscr{L}_{X^I}} (z_i)_{i\in I}$ implies that $y_j \lhd_{\pi_j(\mathscr{L}_{X^I})} z_j$.

Definition 9. Let X be a set and let $\mathscr{L}_{X^I}, \mathscr{R}_{X^I}$ be nests on X^I. Then, $\mathscr{L}_{X^I} \cup \mathscr{R}_{X^I}$ will be called *weakly T_1-separating*, with respect to the j-th variable, if and only if for every $(x_i)_{i\in I}, (y_i)_{i\in I} \in X^I$, such that $x_j \neq y_j$, there exist $L \in \mathscr{L}_{X^I}$ and $R \in \mathscr{R}_{X^I}$, such that $(x_i)_{i\in I} \in L$ and $(y_i)_{i\in I} \notin L$ and also $(y_i)_{i\in I} \in R$ and $(x_i)_{i\in I} \notin R$.

In this case, it is easy to see that $(x_i)_{i\in I} \lhd_{\mathscr{L}_{X^I}} (y_i)_{i\in I}$, if and only if $(y_i)_{i\in I} \lhd_{\mathscr{R}_{X^I}} (x_i)_{i\in I}$.

Proposition 4. *Let X be a set and let also $\mathscr{L}_X$ and $\mathscr{R}_X$ be two nests on X, such that $\mathscr{L}_X \cup \mathscr{R}_X$ is T_1-separating in X. Then, $\pi_j^{-1}(\mathscr{L}_X) \cup \pi_j^{-1}(\mathscr{R}_X)$ is weakly T_1-separating in $X \times X$, with respect to the j-th variable.*

Proof. Let $(x_i)_{i\in I}, (y_i)_{i\in I} \in X^I$, such that $x_j \neq y_j$. Then, there exist $L \in \mathscr{L}_X$ and $R \in \mathscr{R}_X$, such that $x_j \in L$ and $y_j \notin L$ and also $y_j \in R$ and $x_j \notin R$, which implies that $(x_i)_{i\in I} \in \pi_j^{-1}(L)$, $(y_i)_{i\in I} \notin \pi_j^{-1}(R)$, and also $(y_i)_{i\in I} \in \pi_j^{-1}(R)$ and $(x_i)_{i\in I} \notin \pi_j^{-1}(R)$. Thus, $\pi_j^{-1}(\mathscr{L}_X) \cup \pi_j^{-1}(\mathscr{R}_X)$ is weakly T_1-separating, with respect to the j-th variable.

Proposition 5. *Let X be a set and let $\mathscr{L}_{X^I}$ and $\mathscr{R}_{X^I}$ be two nests in X^I, such that $\mathscr{L}_{X^I} \cup \mathscr{R}_{X^I}$ is weakly T_1-separating in X^I, with respect to the j-th variable. Let also $\mathscr{L}_{X^I}$ and $\mathscr{R}_{X^I}$ satisfy the condition that if $(x_i)_{i\in I} \notin L \in \mathscr{L}_{X^I}$, then $\pi_j((x_i)_{i\in I}) = x_j \notin \pi_j(\mathscr{L}_{X^I})$, and if $(x_i)_{i\in I} \notin R \in \mathscr{R}_{X^I}$, then $\pi_j((x_i)_{i\in I}) \notin \pi_j(R)$. Then, $\pi_j(\mathscr{L}_{X^I}) \cup \pi_j(\mathscr{R}_{X^I})$ is T_1- separating in X.*

Proof. Let $x_1 \neq x_2$. Then, $(y_i)_{i\in I} \neq (z_i)_{i\in I}$, where $y_j = x_1$, $z_j = x_2$, and the rest y_i and z_i are arbitrary. Since $\mathscr{L}_{X^I} \cup \mathscr{R}_{X^I}$ is weakly T_1-separating, there exist $L \in \mathscr{L}_{X^I}$, $R \in \mathscr{R}_{X^I}$, such that $(y_i)_{i\in I} \in L$ and $(z_i)_{i\in I} \notin L$ and also $(z_i)_{i\in I} \in R$ and $(y_i)_{i\in I} \notin R$, which implies that $x_1 \in \pi_j(L)$ and $x_2 \notin \pi_j(R)$ and also $x_2 \in \pi_j(R)$ and $x_1 \notin \pi_j(R)$.

Theorem 9. *Let X be a set and let $\mathscr{L}_X$ be an interlocking nest in X. Then, $\mathscr{M} = \{\pi_j^{-1}(L) : L \in \mathscr{L}_X\}$ will be an interlocking nest in X^I.*

Proof. Suppose $M \in \mathscr{M}$ be such that $M = \bigcap\{M' \in \mathscr{M} : M' \supsetneq M\}$. By the definition of $\mathscr{M}$, there exists $L \in \mathscr{L}$ such that: $M = \pi_j^{-1}(L) = \Pi_i\{Y_i : Y_i = X, \text{if} i \neq j \text{and} Y_i = L \text{if} i = j\}$. Making a similar substitution for all $M' \in \mathscr{M}$, we deduce that: $\Pi_i\{Y_i : Y_i = X, \text{if} i \neq j \text{and} Y_i = L, \text{if} i = j\} = \bigcap \Pi_i\{Z_i : Z_i = X, \text{if} i \neq j \text{and} Z_i = L', \textit{if}\ \text{i} = \text{j}, L' \in \mathscr{L}, L' \supset L\} = \bigcap\{W_i : W_i = X, \text{if} i \neq j \text{and} W_i = \bigcap\{L' \in \mathscr{L} : L' \supsetneq L\}, \text{if} i = j\}$. So, $L = \bigcap\{L' \in \mathscr{L} : L' \supsetneq L\}$, which implies that $L = \bigcup\{L' \in \mathscr{L} : L' \subsetneq L\}$. Hence, $M = \Pi_i\{Y_i : Y_i = X, \text{if} i \neq j \text{and} Y_i = L, \text{if} i = j\} = \Pi_i\{\Theta_i : \Theta_i = X, \text{if}, i \neq j \text{and} \Theta_i = \bigcup\{L' \in L : L' \subsetneq L\}, \text{if} i = j\}$. So, $M = \bigcup\{M' \in \mathscr{M} : M' \subsetneq M\}$, which proves that $\mathscr{M}$ is interlocking.

Lemma 3. *Let X be a set and let $\mathscr{L}_{X^I}$ be a collection of subsets of X^I. If the following condition holds: [if $(x_i)_{i\in I} \notin L$, $L \in \mathscr{L}_{X^I}$, then $x_j \notin \pi_j(L)$], then $\pi_j(L) \supset \bigcap\{\pi_j(L') : \pi_j(L') \supsetneq \pi_j(L)\}$ implies that $L \supset \bigcap\{L' : L' \supsetneq L\}$.*

Proof. If $\bigcap\{L' : L' \supsetneq L\} \subsetneq L$, then there exists $(x_i)_{i\in I}$, such that $(x_i)_{i\in I} \in \bigcap\{L' : L' \supsetneq L\}$ and $(x_i)_{i\in I} \notin L$. So, $(x_i)_{i\in I} \in L'$, for every $L' \supsetneq L$, and $\pi_j((x_i)_{i\in I}) = x_j \notin \pi_j(L)$. Thus, $x_j \in \pi_j(L')$, for all $\pi_j(L') \supsetneq \pi_j(L)$ and $x_j \notin \pi_j(L)$, which contradicts the statement of the Lemma 3.

Theorem 10. *Let $\mathscr{L}_{X^I}$ be an interlocking nest in X^I. Then, $\pi_j(\mathscr{L}_{X^I})$, $j \in I$, is an interlocking nest in X, if the condition in Lemma 3 holds.*

Proof. We have that $\pi_j(\mathscr{L}_{X^I}) = \{\pi_j(L) : L \in \mathscr{L}_{X^I}\}$. Let

$$\pi_j(L) = \bigcap\{\pi_j(L')\} : \pi_j(L') \supsetneq \pi_j(L). \tag{1}$$

We will prove that

$$\pi_j(L) = \bigcup\{\pi_j(L') : \pi_j(L') \subsetneq \pi_j(L)\} \tag{2}$$

or, equivalently, we will prove that:

$$\pi_j(L) \subset \bigcup\{\pi_j(L') : \pi_j(L') \subsetneq \pi_j(L)\}$$

Since (1) is satisfied, we have that $\pi_j(L) \supset \bigcap\{\pi_j(L') : \pi_j(L') \supsetneq \pi_j(L)\}$ which implies, by Lemma 3, that $L \supset \bigcap\{L' : L' \supsetneq L\}$. But since it is always true that $L \subset \bigcap\{L' : L' \supsetneq L\}$, we have that $L = \bigcap\{L' : L' \supsetneq L\}$, and since $\mathscr{L}_{X^I}$ is interlocking, we have that $\pi_j(L) \subset \bigcup\{\pi_j(L') : L' \subsetneq L\} \subset \bigcup\{\pi_j(L') : \pi_j(L') \subsetneq \pi_j(L)\}$, which completes the proof.

Theorem 11. *Let X be a topological space and let $\mathscr{L}_{X^I}$, $\mathscr{R}_{X^I}$ be two interlocking, weakly T_0-separating nests in X^I, such that their union, $\mathscr{L}_{X^I} \cup \mathscr{R}_{X^I}$ is weakly T_1-separating, with respect to the j-th variable. Let also for $L \in \mathscr{L}_{X^I}$ and $R \in \mathscr{R}_{X^I}$, $\mathscr{L}_{X^I}$ and $\mathscr{R}_{X^I}$ satisfy the following two conditions:*

1. *If $x_j \notin L$, then $x_j \notin \pi_j(L)$.*
2. *If $x_j \notin R$, then $x_j \notin \pi_j(R)$*

Then, $\pi_j(\mathscr{L}_{X^I})$ and $\pi_j(\mathscr{R}_{X^I})$ are interlocking, T_0-separated nests of open sets, in X, such that their union, $\pi_j(\mathscr{L}_{X^I}) \cup \pi_j(\mathscr{R}_{X^I})$ is T_1-separating (thus, the topology of X will coincide with the order topology).

Proof. We have already shown that the canonical projection of a weakly T_0-separating nest is a T_0-separating nest, that the projection of a weakly T_1-separating union of two nests of open sets is T_1-separating, and also that interlockingness is preserved in a nest, if we project it via canonical projection. The only thing that remains to complete the proof is to remark that π_j is an open mapping, so for each L open in X^I, $\pi_j(L)$ and $\pi_j(R)$ are open sets in X, and this completes the proof.

Corollary 4. *Let X be a topological space and let $\mathscr{L}$ and $\mathscr{R}$ be two T_0-separating, interlocking nests of open sets, in X, such that $\mathscr{L} \cup \mathscr{R}$ is T_1-separating. Then, $\mathscr{S} = \{\bigcap_{j_k \in J_k} \pi_{j_k}^{-1}(L \cap R)\}$, $J_k \subset I$, $L \in \mathscr{L}, R \in \mathscr{R}\}$ will be a base for a topology in X^I.*

LOTS and Function Spaces

Let X and Y be two sets and let $\mathscr{F}(X,Y) = \{f : f \mathrm{isafunction}, f : X \to Y\}$. Then, it is known that $\mathscr{F}(X,Y) = \Pi_{x \in X} Y_X$, where $Y_X = Y$, for all $x \in X$.

Theorem 12. *Let X and Y be two sets, and let $\mathscr{F}(X,Y)$ be the function space, that consists of all functions from X to Y. Let also $\mathscr{L}$ be a nest on Y. Then, for each $x \in X$, the set $\mathscr{L}^x_{\mathscr{F}(X,Y)} = \{(x,L) : L \in \mathscr{L}\}$, where $(x,L) = \{f \in \mathscr{F}(X,Y) : f(x) \in L\}$, will be a nest on $\mathscr{F}(X,Y)$.*

Proof. We remark that $\mathscr{L}^x_{\mathscr{F}(X,Y)} = \{\pi_x^{-1}(L) : L \in \mathscr{L}\}$ is a nest, and this proves the assertion of the theorem.

Remark 2. Let $\mathscr{F}(X,Y)$ be a function space and let also $\mathscr{L}_Y$ and $\mathscr{R}_Y$ be two nests on Y, such that $\mathscr{L}_Y \cup \mathscr{R}_Y$ is T_1-separating. Let also $x \in X$ be a point in X. Then, $\{(x,L) : L \in \mathscr{L}_Y\} \cup \{(x,R) : R \in \mathscr{R}_Y\}$ is weakly T_1-separating, with respect to x. This means that if $f, g \in \mathscr{F}(X,Y)$, such that $f(x) \neq g(x)$, then there exist L, R in

$\mathscr{L}_Y, \mathscr{R}_Y$, respectively, such that $f \in (x, L)$ and $g \notin (x, L)$ and also $g \in (x, R)$ and $f \notin (x, R)$, which is an immediate consequence of Proposition 4.

Last but not least, the union $\bigcup_{x \in X} \mathscr{L}_{\mathscr{F}(X,Y)} \cup \bigcup_{x \in X} \mathscr{R}_{\mathscr{F}(x,y)}$ is a subbase for the point-open topology.

Corollary 5. *Let X and Y be two sets and let also $\mathscr{L}$ be a nest on Y. Then, for each $x \in X$, all the nests of the form $\mathscr{L}^x = \{(x, L) : L \in \mathscr{L}\}$ are interlocking.*

Nests and the Ring ${}^\bullet\mathbb{R}$ of Fermat Reals

A Short Introduction

The idea of the ring of Fermat Reals ${}^\bullet\mathbb{R}$ has come as a possible alternative to Synthetic Differential Geometry (see, e.g., [4–7]) and its main aim is the development of a new foundation of smooth differential geometry for finite and infinite-dimensional spaces. In addition, ${}^\bullet\mathbb{R}$ could play a role of a potential alternative in some certain problems in the field ${}^\star\mathbb{R}$ in Nonstandard Analysis (NSA), because the applications of NSA in differential geometry are very few. One of the "weak" points of ${}^\bullet\mathbb{R}$ at the moment is the lack of a natural topology, carrying the strong topological properties of the line.

P. Giordano and M. Kunzinger have recently done brave steps towards the topologization of the ring ${}^\bullet\mathbb{R}$ of Fermat Reals. In particular, they have constructed two topologies: the Fermat topology and the omega topology (see [7]). The Fermat topology is generated by a complete pseudo-metric and is linked to the differentiation of non-standard smooth functions. The omega topology is generated by a complete metric and is linked to the differentiation of smooth functions on infinitesimals. Although both topologies are very useful in developing infinitesimal instruments for smooth differential geometry, none of these two topologies aims to characterize the Fermat real line from an order-theoretic perspective. In fact, neither makes the space T_1, while an appropriate order-topology would equip the Fermat Real Line with the structure of a monotonically normal space, at least. The possibility to define a linear order relation on ${}^\bullet\mathbb{R}$, so that it can be viewed as a LOTS (linearly ordered topological space) can be considered important, because ${}^\bullet\mathbb{R}$ is an alternative mathematical model of the real line, having some features with respect to applications in smooth differential geometry and mathematical physics. It is therefore natural to ask whether for ${}^\bullet\mathbb{R}$ peculiar characteristics of $\mathbb{R}$ hold or not.

In this section we will focus in the order relation which is stated in [4], and we will interpreted through nests.

As we shall see in Definition 11, the idea of the formation of ${}^\bullet\mathbb{R}$ starts with an equivalence relation in the little-oh polynomials, where ${}^\bullet\mathbb{R}$ is the quotient space under this relation. This treatment permits us to view these little-oh polynomials as numbers.

Definitions

Definition 10. A little-oh polynomial x_t (or $x(t)$) is an ordinary set-theoretical function, defined as follows:

1. $x : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and
2. $x_t = r + \sum_{i=1}^{k} \alpha_i t^{a_i} + o(t)$, as $t \to 0^+$, for suitable $k \in \mathbb{N}$, $r, \alpha_1, \cdots, \alpha_k \in \mathbb{R}$ and $a_1, \cdots, a_k \in \mathbb{R}_{\geq 0}$.

The set of all little-oh polynomials is denoted by the symbol $\mathbb{R}_o[t]$. So, $x \in \mathbb{R}_o(t)$, if and only if x is a polynomial function with real coefficients, of a real variable $t \geq 0$, with generic positive powers of t and up to a little-oh function $o(t)$, as $t \to 0^+$.

Definition 11. Let $x, y \in \mathbb{R}_o[t]$. We declare $x \sim y$ (and we say $x = y$ in $^\bullet\mathbb{R}$), if and only if $x(t) = y(t) + o(t)$, as $t \to 0^+$.

The relation $\sim$ in Definition 11 is an equivalence relation and $^\bullet\mathbb{R} := \mathbb{R}_o[t]/\sim$. A first attempt to define an order in $^\bullet\mathbb{R}$ has come from Giordano.

Definition 12 (Giordano). Let $x, y \in {}^\bullet\mathbb{R}$. We declare $x \leq y$, if and only if there exists $z \in {}^\bullet\mathbb{R}$, such that $z = 0$ in $^\bullet\mathbb{R}$ (i.e. $\lim_{t\to 0^+} z_t/t = 0$) and for every $t \geq 0$ sufficiently small, $x_t \leq y_t + z_t$.

For simplicity, one does not use equivalence relation but works with an equivalent language of representatives. If one chooses to use the notations of [4], one has to note that Definition 12 does not depend on representatives.

As the author describes in [4], the order relation in NSA admits all formal properties among all the theories of (actual) infinitesimals, but there is no good dialectic of these properties with their informal interpretation. In particular, the order in $^\star\mathbb{R}$ inherits by transfer all the first order properties but, on the other hand, in the quotient field $^\star\mathbb{R}$ it is difficult to interpret these properties of the order relation as intuitive properties of the corresponding representatives. For example, a geometrical interpretation like that of $^\bullet\mathbb{R}$ seems not possible for $^\star\mathbb{R}$. Definition 12 provides a clear geometrical representation of the ring $^\bullet\mathbb{R}$ (see, for instance, Sect. 4.4 of [4]).

The Fermat Topology and the Omega-Topology on $^\bullet\mathbb{R}$

A subset $A \subset {}^\bullet\mathbb{R}^n$ is open in the Fermat topology, if it can be written as $A = \bigcup\{{}^\bullet U \subset A : U \text{isopeninthenaturaltopologyin} \mathbb{R}^n\}$. Giordano and Kunzinger describe this topology as the best possible one for sets having a “sufficient amount of standard points”, for example $^\bullet U$. They add that this connection between the Fermat topology and standard reals can be glimpsed by saying that the monad $\mu(r) := \{x \in {}^\bullet\mathbb{R} : \text{standardpartof} x = r\}$ of a real $r \in \mathbb{R}$ is the set of all points which are limits of sequences with respect to the Fermat topology. However it is obvious that in sets of infinitesimals there is a need for constructing a (pseudo-)metric generating a finer topology that the authors call the omega-topology (see [7]). Since neither the Fermat nor the omega-topology is Hausdorff when restricted to ${}^\bullet\mathbb{R}$ and since each of them describes sets having a "sufficient amount" of standard points or infinitesimals, respectively, there is a need for defining a natural topology on ${}^\bullet\mathbb{R}$ describing sufficiently all Fermat reals and carrying the best possible properties.

Interlocking Nests on ${}^\bullet\mathbb{R}$

Theorem 13. *The pair* $({}^\bullet\mathbb{R}, <_F)$*, where* $<_F$ *is defined as follows:*

$$x <_F y \Leftrightarrow \begin{cases} \exists\,\{k \in {}^\bullet\mathbb{R} : k \leq l\},\ \text{some } l \in {}^\bullet\mathbb{R},\ \text{such that } x \in \{k \in {}^\bullet\mathbb{R} : k \leq l\} \not\ni y, l \in {}^\bullet\mathbb{R} \\ \textit{or} \\ x = \max\{k \in {}^\bullet\mathbb{R} : k \leq l\},\ \text{some } l \in {}^\bullet\mathbb{R} \text{ and } \exists h \in {}^\bullet\mathbb{R} : h > 0,\ y = x + h \\ \textit{or} \\ y = \min\{k \in {}^\bullet\mathbb{R} : l \leq k\},\ \text{some } l \in {}^\bullet\mathbb{R} \text{ and } \exists h \in {}^\bullet\mathbb{R} : h > 0,\ x = y - h \end{cases}$$

where x, y *are distinct Fermat reals, is a linearly ordered set.*

Proof. The order of Definition 12 gives two nests, namely the nest $\mathscr{L}$, which consists of all sets $L = \{k \in {}^\bullet\mathbb{R} : k \leq l\}$, some $l \in {}^\bullet\mathbb{R}$ and the nest $\mathscr{R}$, which consists of all sets $R = \{k \in {}^\bullet\mathbb{R} : l \leq k\}$, some $l \in {}^\bullet\mathbb{R}$. In addition, we have that $\trianglelefteq_{\mathscr{L}} = \trianglerighteq_{\mathscr{R}} = \leq$.

We remark that, for any $L \in \mathscr{L}$ (respectively for any $R \in \mathscr{R}$), L (resp. R) has a $\trianglelefteq_{\mathscr{L}}$-maximal element (resp. $\trianglelefteq_{\mathscr{R}}$-maximal element for R), such that $X - L$ has no $\trianglelefteq_{\mathscr{L}}$-minimal element (resp. $X - R$ has no $\trianglelefteq_{\mathscr{R}}$-minimal element). So, neither $\mathscr{L}$ nor $\mathscr{R}$ are interlocking.

Now, for all $L = \{k \in {}^\bullet\mathbb{R} : k \leq l\} \in \mathscr{L}$, some $l \in {}^\bullet\mathbb{R}$, let x_L denote the $\trianglelefteq_{\mathscr{L}}$-maximal element of L and for all $R = \{k \in {}^\bullet\mathbb{R} : l \leq k\} \in \mathscr{R}$, some $l \in {}^\bullet\mathbb{R}$ let y_R denote the $\trianglelefteq_{\mathscr{L}}$-minimal element of R.

Furthermore, for each $L \in \mathscr{L}$ choose $x_L^+ \in {}^\bullet\mathbb{R}$ and for each $R \in \mathscr{R}$ choose $y_R^- \in {}^\bullet\mathbb{R}$, where x_L^+ and y_R^- are distinct points in ${}^\bullet\mathbb{R}$, and define a map $p : {}^\bullet\mathbb{R} \to {}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\})$, as follows:

$$p(x) = \begin{cases} x, & \text{if } x \in {}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\}) \\ x_L, & \text{if } x = x_L^+ \\ y_R, & \text{if } x = y_R^- \end{cases}$$

Now, define an order $<_F$ on ${}^\bullet\mathbb{R}$, so that:

$$x <_F y \Leftrightarrow \begin{cases} p(x) \lhd_{\mathscr{L}} p(y) \\ or \\ x = x_L \text{ and } y = x_L^+ \\ or \\ x = y_R^- \text{ and } y = y_R \end{cases}$$

Obviously, $<_F$ is a linear order and the restriction of $<_F$ to ${}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\})$ equals $\trianglelefteq_{\mathscr{L}}$, the order in Definition 12. In addition, we can set $x_L^+ = x_L + h$, where h is not zero in ${}^\bullet\mathbb{R}$ and $h > 0$, that is, $\lim_{t\to 0^+} h_t/t \neq 0$ and, respectively, we set $x_R^- = x_R - h$, and this completes the proof.

Theorem 14. ${}^\bullet\mathbb{R}$ *equipped with the order topology from* $<_F$ *is a LOTS.*

Proof. We will now show that the topology $\mathscr{T}$ on ${}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\})$ coincides with the subspace topology on ${}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\})$ that is inherited from the $<_F$-order topology on ${}^\bullet\mathbb{R}$.

But, since $\mathscr{L} \cup \mathscr{R}$ forms a subbasis for $\mathscr{T}$, that consists of two nests, every set in $\mathscr{T}$ can be written as a union of sets of the form $L \cap R$, where $L \in \mathscr{L}$ and $R \in \mathscr{R}$. It suffices therefore to show that every $L \in \mathscr{L}$ and $R \in \mathscr{R}$ can be written as the intersection of an order-open set with ${}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\})$. But this is always true, since if $L \in \mathscr{L}$, with $\trianglelefteq_{\mathscr{L}}$-maximal element x_L, then $L = {}^\bullet\mathbb{R} - (\{x_L^+ : L \in \mathscr{L}\} \cup \{y_R^- : R \in \mathscr{R}\}) \cap \{x \in {}^\bullet\mathbb{R} : x <_F x_L^+\}$.

The argument for $R \in \mathscr{R}$ is similar, and this completes the proof.

Remarks

1. The order topology $\mathscr{T}_{<_F}$ equals the topology $\mathscr{T}_{\mathscr{L}_{<_F} \cup \mathscr{R}_{<_F}}$, where $\mathscr{L}_{<_F} = \{k \in {}^\bullet\mathbb{R} : k <_F l\}$, some $l \in {}^\bullet\mathbb{R}$ and $\mathscr{R}_{<_F} = \{k \in {}^\bullet\mathbb{R} : l <_F k\}$, some $l \in {}^\bullet\mathbb{R}$. This is because $\mathscr{L}_{<_F} \cup \mathscr{R}_{<_F}$ T_1-separates ${}^\bullet\mathbb{R}$ and both $\mathscr{L}_{<_F}$ and $\mathscr{R}_{<_F}$ are interlocking nests. So, unlike the GO-space topology $\mathscr{T}_{\leq}$ on ${}^\bullet\mathbb{R}$, where $\mathscr{T}_{\leq} \subset \mathscr{T}_{\mathscr{L} \cup \mathscr{R}}$, $<_F$ provides a natural extension of the natural linear order of the set of real numbers to the Fermat real line and the order topology from $<_F$ can be completely described via the nests $\mathscr{L}_{<_F}$ and $\mathscr{R}_{<_F}$.
2. Viewing the Fermat real line as a LOTS and working with nests $\mathscr{L}_{<_F}$ and $\mathscr{R}_{<_F}$, one can now define the product topology for ${}^\bullet\mathbb{R}^n$, some positive integer n, or even more generally for $\Pi_{i\in I}{}^\bullet\mathbb{R}_i$, some arbitrary indexing set I, in the usual way via the subbasis $\pi_{j_0}^{-1}(A_{j_0}) = \Pi_{i\in I}\{{}^\bullet\mathbb{R}_i : i \neq j_0\} \times A_{j_0}$, where A_{j_0} is an open subset in the coordinate space ${}^\bullet\mathbb{R}_{j_0}$ in the order topology $\mathscr{T}_{<_F}$ and $\pi_i : \Pi_{i\in I}{}^\bullet\mathbb{R}_i \to {}^\bullet\mathbb{R}_i$ the projection.
3. In this way one can define continuity for any function f from a topological space Y into the product space $\Pi_{i\in I}{}^\bullet\mathbb{R}_i$ via the continuity of $\pi_i \circ f : Y \to {}^\bullet\mathbb{R}_i$.

4. The neight of ${}^{\bullet}\mathbb{R}$ is 2 and the neight of ${}^{\bullet}\mathbb{R}^n = n + 1$ (see [1]). Using the product topology, as stated in Remark 2, we use four nests in order to define -for example- the topology in ${}^{\bullet}\mathbb{R}^2$, but since the neight of ${}^{\bullet}\mathbb{R}^2$ is 3, one can define a topology using three nests exclusively.

Questions

1. As a LOTS, $({}^{\bullet}\mathbb{R}, <_F)$ has rich topological properties. It is, for example, a monotone normal space. It would be interesting though to have an extensive study on the metrizability of this space. It is known that in a GO-space the terms metrizable, developable, semistratifiable, etc. are equivalent (see [3] and [9]). The real line (i.e. the set of all standard reals, from the point of view of ${}^{\bullet}\mathbb{R}$) is a developable LOTS and this is equivalent to say that it is also a metrizable LOTS. Is $({}^{\bullet}\mathbb{R}, \mathscr{T}_{<_F})$ developable?
2. Which of the subspaces of $({}^{\bullet}\mathbb{R}, \mathscr{T}_{<_F})$ are developable?

Since any sequence $x_1, x_2, \cdots$ of points in $\Pi_{i\in I}{}^{\bullet}\mathbb{R}_i$ will converge to a point $x \in \Pi_{i\in I}{}^{\bullet}\mathbb{R}_i$, iff for every projection $\pi_i : \Pi_{i\in I}{}^{\bullet}\mathbb{R}_i \to {}^{\bullet}\mathbb{R}_i$ the sequence $\pi_i(x_1), \pi_i(x_2), \cdots$ converges to $\pi_i(x)$ in the coordinate space ${}^{\bullet}\mathbb{R}_i$, any answer to the above questions will be fundamental towards our understanding of convergence in the ring of Fermat Reals.

References

1. Brian, W.R.: Neight: the nested weight of a topological space. Topol. Proc. **47**, 279–296 (2016)
2. de Groot, J., Schnare, P.S.: A topological characterization of products of compact totally ordered spaces. Gen. Topol. Appl. **2**, 67–73 (1972)
3. Faber, M.J.: Metrizability in Generalized Ordered Spaces. Matematisch Centrum, Amsterdam (1974)
4. Giordano, P.: Fermat Reals: Nilpotent Infinitesimals and Infinite Dimensional Spaces. Prebook http://arxiv.org/abs/0907.1872
5. Giordano, P.: The ring of fermat reals. Adv. Math. **225**, 2050–2075 (2010)
6. Giordano, P.: Fermat reals: infinitesimals without logic. Miskolc Math. Notes **14**(3), 65–80 (2013)
7. Giordano, P., Kunzinger, M.: Topological and algebraic structures on the ring of fermat reals. Isr. J. Math. **193**, 459–505 (2013)
8. Good, C., Papadopoulos, K.: A topological characterization of ordinals: van Dalen and Wattel revisited. Topol. Appl. **159**, 1565–1572 (2012)
9. Hart, K.P., Nagata, J.-I., Vaughan, J.E.: Encyclopedia of General Topology. Elsevier Science and Technology Books (2014) ISBN: 978-0-444-50355-8
10. Papadopoulos, K: On properties of nests: some answers and questions. Quest. Answ. Gen. Topol. **33**, 71–91 (2015)
11. Papadopoulos, K.: On the orderability problem and the interval topology. In: Rassias, T., Toth, L. (eds.) Topics in Mathematical Analysis and Applications. Optimization and Its Applications Springer Series. Springer, New York (2014)

12. Purisch, S.: A History of Results on Orderability and Suborderability. Handbook of the History of General Topology, vol. 2 (San Antonio, TX, 1993), volume 2 of Hist. Topol., pp. 689–702. Kluwer Academic Publications, Dordrecht (1998)
13. Reed, G.M.: The intersection topology w.r.t. the real line and the countable ordinals. Trans. Am. Math. Soc. **247**(2), 509–520 (1986)
14. van Dalen, J., Wattel, E.: A topological characterization of ordered spaces. Gen. Topol. Appl. **3**, 347–354 (1973)
15. van Douwen, E.K.: On Mike's misnamed intersection topologies. Topol. Appl. **51**(2), 197–201 (1993)

Resolvent Operators for Some Classes of Integro-Differential Equations

I.N. Parasidis and E. Providas

Abstract Explicit representations are constructed for the resolvents of the operators of the form $B = \widehat{A} + Q_1$ and $\mathbf{B} = \widehat{A}^2 + Q_2$, where $\widehat{A}$ and $\widehat{A}^2$ are linear closed operators with known resolvents and Q_1 and Q_2 are perturbation operators embedding inner products of $\widehat{A}$ and $\widehat{A}^2$ as they appear in integro-differential equations and other applications.

Keywords Resolvent Operator • Integro-Differential Equations • Boundary value Problems • Exact Solution

Introduction

The present article is concerned with the study of generalized boundary value problems containing differential or integro-differential operators by means of the resolvent operator. Specifically, we derive explicit representations for the resolvent operators for three classes of problems.

The first problem involves a linear operator defined in the complex Hilbert space H of the form

$$B = \widehat{A} + Q_1 = \widehat{A} - \sum_{i=1}^{m} g_i \langle \widehat{A}\,\cdot, \phi_i \rangle_H, \quad D(B) = D(\widehat{A}), \tag{1}$$

where $\widehat{A}$ is a linear closed, not necessarily bounded, operator, $g_1, g_2, \ldots, g_m$ are linearly independent elements of H, $\phi_1, \phi_2, \ldots, \phi_m \in H$ and $\langle \cdot, \cdot \rangle_H$ denotes the inner product in H. Note that the operator B can be viewed as a perturbation of the operator $\widehat{A}$ by the operator Q_1. Moreover, the operators $\widehat{A}$, B are extensions of

I.N. Parasidis
Department of Electrical Engineering, TEI of Thessaly, 41110 Larissa, Greece
e-mail: paras@teilar.gr

E. Providas (✉)
Department of Mechanical Engineering, TEI of Thessaly, 41110 Larissa, Greece
e-mail: providas@teilar.gr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_24

a minimal operator $A_0 \subset \widehat{A}$ with $D(A_0) = D(\widehat{A}) \cap \ker Q_1$. We prove that when the resolvent set $\rho(\widehat{A})$ and the resolvent operator $R_\lambda(\widehat{A}) = (\widehat{A} - \lambda I)^{-1}$ of $\widehat{A}$ are known then we can find the resolvent set $\rho(B) \cap \rho(\widehat{A})$ and the resolvent operator $R_\lambda(B) = (B - \lambda I)^{-1}$ of B in closed form.

The second problem encompasses an operator of the kind

$$\mathbf{B} = \widehat{A}^2 + Q_2 = \widehat{A}^2 - \sum_{i=1}^{m} s_i \langle \widehat{A}\cdot, \phi_i \rangle_H - \sum_{i=1}^{m} g_i \langle \widehat{A}^2\cdot, \phi_i \rangle_H, \quad D(\mathbf{B}) = D(\widehat{A}^2), \tag{2}$$

where in addition $s_i \in H,\ i = 1, \ldots, m$. We show that the resolvent set $\rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$ and the resolvent operator $R_\lambda(\mathbf{B}) = (\mathbf{B} - \lambda I)^{-1}$ of the operator $\mathbf{B}$ can be evaluated when their counterparts $\rho(\widehat{A}^2)$ and $R_\lambda(\widehat{A}^2) = (\widehat{A}^2 - \lambda I)^{-1}$ of the simpler operator $\widehat{A}^2$ are known.

A special form of the second problem with interest is obtained when we take $g_i \in D(\widehat{A})$ and $s_i = Bg_i,\ i = 1, \ldots, m$. In this case $\mathbf{B} = B^2$, i.e. $\mathbf{B}$ becomes a quadratic operator. The explicit formula for the resolvent operator $R_{\lambda^2}(\mathbf{B})$ for $\lambda^2 \in \rho(\mathbf{B})$ is constructed from the resolvent operators $R_\lambda(\widehat{A})$ and $R_{-\lambda}(\widehat{A})$ for $\pm\lambda \in \rho(\widehat{A})$.

Resolvent operators are associated with the spectral theory and their origin goes as back as to the early days of functional analysis, see, e.g., [13] and [10]. When the resolvent $R_\lambda(B)$ of an operator B exists and is provided in an analytic form, it is valuable for the study of the operator B itself and the solution of the problems $(B - \lambda I)x = f,\ f \in H$ and $Bx = f\ (\lambda = 0)$. Perturbation theory for linear operators was first introduced by Rayleigh and Schrödinger [17] and founded later by Kato [10]. Since then it occupies an important place in theoretical physics, mechanics and applied mathematics. Extension theory was initiated by von Neumann [19] and developed further by [12, 23] and [1], commonly known as Birman–Kreĭn–Vishik theory, as well as [3, 5, 11, 14, 21, 22] and many others. Integro-differential equations appear in the mathematical modeling in biology, engineering, telecommunications and economics. Of interest here are the following works. In [4, 6, 7, 15, 16] and [8], resolvent methods have been employed to study a class of integro-differential equations, occurring in heat conduction and viscoelasticity, where $\widehat{A}$ in (1) is a specific operator and Q_1 is a Volterra operator. By the same means certain integro-differential equations, arising in quantum-mechanical scattering theory, where $\widehat{A}$ is a special operator and Q_1 is a one-dimensional perturbation ($m = 1$), have also been investigated, see, e.g., [2]. The resolvent and the spectrum of perturbed operators of the type (1) where $\widehat{A}$ is a symmetric operator and $\langle \widehat{A}\cdot, \phi_i \rangle_H = a_i \langle \cdot, g_i \rangle_H,\ a_i \in \mathbf{R}$ have been studied by [9] and the references therein. Finally, a Fredholm type boundary integral equation in elasticity with $\langle \widehat{A}\cdot, \phi_i \rangle_H = p_i(\cdot)$ has been considered in [18].

In the rest we make use of the following notation. Namely, $F = (\phi_1, \ldots, \phi_m)$, $G = (g_1, \ldots, g_m)$ and $AF = (A\phi_1, \ldots, A\phi_m)$ are vectors of H^m. We write F^t and $\langle Ax, F^t \rangle_{H^m}$ for the column vectors $\mathrm{col}(\phi_1, \ldots, \phi_m)$ and $\mathrm{col}(\langle Ax, \phi_1 \rangle_H, \ldots, \langle Ax, \phi_m \rangle_H)$, respectively. We denote by $\overline{M}$ $(resp. M^t)$ the conjugate (resp. transpose) matrix of M and by $\langle G^t, F \rangle_{H^m}$ the $m \times m$ matrix whose

i, j-th entry is the inner product $\langle g_i, \phi_j \rangle_H$. Notice that $\langle G^t, F \rangle_{H^m}$ defines the matrix inner product and has the properties:

$$\langle G^t, FC \rangle_{H^m} = \langle G^t, F \rangle_{H^m} \overline{C}, \quad \langle G^t, F \rangle_{H^m} = \overline{\langle F^t, G \rangle}^{\,t}_{H^m}, \tag{3}$$

where C is an $m \times k$ constant complex matrix. We denote by I_m the $m \times m$ identity matrix and by 0_m the $m \times m$ zero matrix. It is understood that $D(A)$ and $R(A)$ stand for the domain and the range of A, respectively.

The paper is organized as follows. In sections "Resolvent of Extensions of a Minimal Operator," "Resolvent of Extensions of the Square of a Minimal Operator," and "Resolvent of Quadratic Operators" we develop the theory for acquiring analytic formulas for the resolvent operators corresponding to each of the three classes of problems presented. In section "Resolvent of Extensions of a Minimal Operator" we apply the theory to three Fredholm type, generalized integro-differential boundary value problems to demonstrate the power of the theory developed.

Resolvent of Extensions of a Minimal Operator

In this section we consider the operator B in Eq. (1) and determine the necessary and sufficient conditions for the existence of the resolvent $R_\lambda(B)$ and we find it in an explicit form provided $R_\lambda(\widehat{A})$ is known. In [22] the perturbed operator B has been studied as the extension of the minimal operator A_0, i.e.

$$A_0 x = \widehat{A}x \;\text{ for }\; x \in D(A_0) = \{x \in D(\widehat{A}) : \; \langle \widehat{A}x, F^t \rangle_{H^m} = \mathbf{0}^t\}$$

We begin by giving the definition of the resolvent. Let $A : H \to H$ be a linear operator. We say that $\lambda \in \mathbf{C}$ belongs to the resolvent set $\rho(A)$ of A if there exists the operator $R_\lambda = R_\lambda(A) = (A - \lambda I)^{-1}$ which is bounded and $D(R_\lambda)$ is dense in H ($\overline{D(R_\lambda)} = H$). The operator $R_\lambda(A)$ is called the resolvent of A. We recall from [13, Lemma 7.2-3] the following result.

Lemma 1. *Let X be a complex Banach space, $A : X \to X$ a linear operator, and $\lambda \in \rho(A)$. Assume that (a) A is closed or (b) A is bounded. Then the resolvent operator $R_\lambda(A)$ of A is defined on the whole space X and is bounded.*

The proposition below is well known and is utilized here several times.

Proposition 1. *If in a Hilbert space H, $\widehat{A} : H \to H$ is a closed linear operator and $\lambda \in \rho(\widehat{A})$, then*

$$\widehat{A}R_\lambda(\widehat{A}) = I + \lambda R_\lambda(\widehat{A}) \quad \text{and} \quad \widehat{A}R_\lambda(\widehat{A}) = R_\lambda(\widehat{A})\widehat{A}. \tag{4}$$

We now prove the key theorem for obtaining the resolvent of the operator B.

Theorem 1. *Let H be a complex Hilbert space, $\widehat{A} : H \to H$ a linear closed operator, $\lambda \in \rho(\widehat{A})$ and $B : H \to H$ the operator defined by*

$$Bx = \widehat{A}x - G\langle \widehat{A}x, F^t \rangle_{H^m}, \quad D(B) = D(\widehat{A}), \tag{5}$$

where the vectors $G = (g_1, \ldots, g_m)$, with $g_1, \ldots, g_m$ being linearly independent elements, $F^t = \mathrm{col}(\phi_1, \ldots, \phi_m) \in H^m$ and $x \in D(B)$. Let the operator

$$B_\lambda x = (B - \lambda I)x = f, \quad D(B_\lambda) = D(\widehat{A}), \tag{6}$$

where I is the identity operator on $D(B)$ and $f \in H$. Then:

(i) $\lambda \in \rho(B)$ *if and only if*

$$\det L_\lambda = \det \left[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G \rangle_{H^m} \right] \neq 0. \tag{7}$$

(ii) $\rho(B) \cap \rho(\widehat{A}) = \{\lambda \in \rho(\widehat{A}) : \det L_\lambda \neq 0\}$.

(iii) *For $\lambda \in \rho(B) \cap \rho(\widehat{A})$ the resolvent operator $R_\lambda(B)$ is defined on the whole space H, is bounded and has the representation*

$$R_\lambda(B)f = R_\lambda(\widehat{A})f + R_\lambda(\widehat{A})G\overline{L_\lambda^{-1}}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t \rangle_{H^m}. \tag{8}$$

Proof. (i)–(iii) Suppose that $\det L_\lambda \neq 0$. Since $D(\widehat{A}) = D(\widehat{A} - \lambda I) = D(B - \lambda I)$ then from (6) for $x \in D(\widehat{A})$ we have

$$(\widehat{A} - \lambda I)x - G\langle \widehat{A}x, F^t \rangle_{H^m} = f, \quad \forall f \in H. \tag{9}$$

Since $\widehat{A}$ is closed, the resolvent operator $R_\lambda(\widehat{A})$ is defined on the whole space H and is bounded as it is implied by Lemma 1. By applying first the operator $R_\lambda(\widehat{A})$ and then the operator $\widehat{A}$ on (9) we get

$$x - R_\lambda(\widehat{A})G\langle \widehat{A}x, F^t \rangle_{H^m} = R_\lambda(\widehat{A})f, \tag{10}$$

$$\widehat{A}x - \widehat{A}R_\lambda(\widehat{A})G\overline{\langle F^t, \widehat{A}x \rangle}_{H^m} = \widehat{A}R_\lambda(\widehat{A})f. \tag{11}$$

By taking the inner products of both sides of (11) with the components of the vector F^t and by observing that the inner product is conjugate linear in the second factor, see Eq. (3), we obtain

$$\langle F^t, \widehat{A}x \rangle_{H^m} - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G \rangle_{H^m} \langle F^t, \widehat{A}x \rangle_{H^m} = \langle F^t, \widehat{A}R_\lambda(\widehat{A})f \rangle_{H^m},$$

$$\left[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G \rangle_{H^m} \right] \langle F^t, \widehat{A}x \rangle_{H^m} = \langle F^t, \widehat{A}R_\lambda(\widehat{A})f \rangle_{H^m},$$

$$L_\lambda \langle F^t, \widehat{A}x \rangle_{H^m} = \langle F^t, \widehat{A}R_\lambda(\widehat{A})f \rangle_{H^m}. \tag{12}$$

Hence, since $\det L_\lambda \neq 0$,

$$\langle F^t, \widehat{A}x\rangle_{H^m} = L_\lambda^{-1}\langle F^t, \widehat{A}R_\lambda(\widehat{A})f\rangle_{H^m},$$
$$\langle \widehat{A}x, F^t\rangle_{H^m} = \overline{L_\lambda^{-1}}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m}. \tag{13}$$

By substituting (13) and $x = (B - \lambda I)^{-1} f$ into (10), we have

$$x = R_\lambda(\widehat{A})f + R_\lambda(\widehat{A})G\overline{L_\lambda^{-1}}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m}, \tag{14}$$

from where Eq. (8) follows. From Proposition 1 we have

$$\widehat{A}R_\lambda(\widehat{A})f = (I + \lambda R_\lambda(\widehat{A}))f$$

and since $R_\lambda(\widehat{A})$ is defined on the whole space H and is bounded, it follows that the resolvent operator $R_\lambda(B)$ is also defined on the whole space H and is bounded. Consequently, $\lambda \in \rho(B)$.

Conversely, let $\lambda \in \rho(B) \cap \rho(\widehat{A})$. We will show that $\det L_\lambda \neq 0$. Assume that $\det L_\lambda = 0$, then $\det \overline{L_\lambda} = 0$. Hence, exists a nonzero vector $\mathbf{a}^t = \mathrm{col}(a_1, \ldots, a_m) \in \mathbf{C}^m$ such that $\overline{L_\lambda}\mathbf{a}^t = \mathbf{0}^t$. We consider the element $x_0 = R_\lambda(\widehat{A})G\mathbf{a}^t \in D(\widehat{A})$. Since the components of the vector G are linearly independent elements we have $G\mathbf{a}^t \neq 0$ and therefore $x_0 \neq 0$. By substituting in (6) we get

$$\begin{aligned}
B_\lambda x_0 &= (\widehat{A} - \lambda I)x_0 - G\overline{\langle F^t, \widehat{A}x_0\rangle}_{H^m} \\
&= (\widehat{A} - \lambda I)R_\lambda(\widehat{A})G\mathbf{a}^t - G\overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A})G\mathbf{a}^t\rangle}_{H^m} \\
&= G\mathbf{a}^t - G\overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle}_{H^m}\mathbf{a}^t \\
&= G\left[I_m - \overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle}_{H^m}\right]\mathbf{a}^t \\
&= G\overline{L_\lambda}\mathbf{a}^t = G\mathbf{0}^t = 0.
\end{aligned} \tag{15}$$

This means that $x_0 \in \ker B_\lambda$ and so $\lambda \notin \rho(B)$, which contradicts the assumption that $\lambda \in \rho(B) \cap \rho(\widehat{A})$. Thus, $\det L_\lambda \neq 0$ and $\rho(B) \cap \rho(\widehat{A}) = \{\lambda \in \rho(\widehat{A}) : \det L_\lambda \neq 0\}$. □

Remark 1. The linear independence of the components of the vector G is required to prove the necessary condition of (i). Hence, if the components of the vector G are not linearly independent, then holds only: $\lambda \in \rho(B)$ if $\det L_\lambda \neq 0$.

Resolvent of Extensions of the Square of a Minimal Operator

In this section we extend the results of the previous section for the case of the operator $\mathbf{B}$ defined in Eq. (2). This operator has been studied in [22] as an extension of the minimal operator A_0^2, namely

$$A_0^2 x = \widehat{A}^2 x \text{ for } x \in D(A_0^2) = \{x \in D(\widehat{A}^2) : \langle \widehat{A}x, F^t \rangle_{H^m} = \langle \widehat{A}^2 x, F^t \rangle_{H^m} = \mathbf{0}^t\}.$$

We find here the resolvent set $\rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$ and a resolvent $R_\lambda(\mathbf{B})$ when the $\rho(\widehat{A}^2)$ and $R_\lambda(\widehat{A}^2)$ are known.

First, we recall the following proposition from [20, pp. 39]

Proposition 2. *If A is a closed operator on a complex Banach space and its resolvent set $\rho(A)$ is not empty, then A^n is closed too for each $n \in \mathbf{N}$.*

We now prove the theorem for the resolvent of the operator $\mathbf{B}$.

Theorem 2. *Let H be a complex Hilbert space, $\widehat{A} : H \to H$ a linear closed operator, $\rho(\widehat{A})$ is not empty, $\lambda \in \rho(\widehat{A}^2)$ and $\mathbf{B} : H \to H$ the operator defined by*

$$\mathbf{B}x = \widehat{A}^2 x - S\langle \widehat{A}x, F^t \rangle_{H^m} - G\langle \widehat{A}^2 x, F^t \rangle_{H^m}, \quad D(\mathbf{B}) = D(\widehat{A}^2), \tag{16}$$

where $F^t = \text{col}(\phi_1, \ldots, \phi_m)$, $S = (s_1, \ldots, s_m)$ and $G = (g_1, \ldots, g_m) \in H^m$, with the components of the vector (S, G) being linearly independent elements of H, and $x \in D(\mathbf{B})$. Let the operator

$$\mathbf{B}_\lambda x = (\mathbf{B} - \lambda I)x = f, \quad D(\mathbf{B}_\lambda) = D(\widehat{A}^2), \tag{17}$$

where I is the identity operator on $D(\mathbf{B})$ and $f \in H$. Then:

(i) $\lambda \in \rho(\mathbf{B})$ *if and only if*

$$\det W_\lambda = \det \begin{pmatrix} I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)S \rangle_{H^m} & -\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)G \rangle_{H^m} \\ -\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)S \rangle_{H^m} & I_m - \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)G \rangle_{H^m} \end{pmatrix} \neq 0. \tag{18}$$

(ii) $\rho(\mathbf{B}) \cap \rho(\widehat{A}^2) = \{\lambda \in \rho(\widehat{A}^2) : \det W_\lambda \neq 0\}$.

(iii) For $\lambda \in \rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$ the resolvent operator $R_\lambda(\mathbf{B})$ is defined on the whole space H, is bounded and has the representation

$$\begin{aligned} R_\lambda(\mathbf{B})f = R_\lambda(\widehat{A}^2)f &+ R_\lambda(\widehat{A}^2)(S\,\overline{W_{\lambda 11}^{-1}} + G\,\overline{W_{\lambda 21}^{-1}})\langle \widehat{A}R_\lambda(\widehat{A}^2)f, F^t \rangle_{H^m} \\ &+ R_\lambda(\widehat{A}^2)(S\,\overline{W_{\lambda 12}^{-1}} + G\,\overline{W_{\lambda 22}^{-1}})\langle \widehat{A}^2 R_\lambda(\widehat{A}^2)f, F^t \rangle_{H^m}, \end{aligned} \tag{19}$$

where $W_{\lambda ij}^{-1}$, $i, j = 1, 2$ are the $m \times m$ submatrices of the partition of the $2m \times 2m$ inverse matrix W_λ^{-1}.

Proof. (i)–(iii)Let us assume that (18) is true, specifically $\det W_\lambda \neq 0$. Since $D(\widehat{A}^2) = D(\widehat{A}^2 - \lambda I) = D(\mathbf{B} - \lambda I)$ then from (17) and for $x \in D(\widehat{A}^2)$ we have

$$(\widehat{A}^2 - \lambda I)x - S\langle \widehat{A}x, F^t\rangle_{H^m} - G\langle \widehat{A}^2 x, F^t\rangle_{H^m} = f. \tag{20}$$

Proposition 2 states that the operator $\widehat{A}^2$ is closed and because $\lambda \in \rho(\widehat{A}^2)$, it follows from Lemma 1 that the resolvent operator $R_\lambda(\widehat{A}^2)$ is defined and is bounded on the whole space H. By applying the resolvent operator $R_\lambda(\widehat{A}^2)$ on (20), we get

$$x - R_\lambda(\widehat{A}^2)\left[S\langle \widehat{A}x, F^t\rangle_{H^m} + G\langle \widehat{A}^2 x, F^t\rangle_{H^m}\right] = R_\lambda(\widehat{A}^2)f. \tag{21}$$

By employing the operators $\widehat{A}$ and $\widehat{A}^2$, we have

$$\widehat{A}x - \widehat{A}R_\lambda(\widehat{A}^2)\left[S\langle \widehat{A}x, F^t\rangle_{H^m} + G\langle \widehat{A}^2 x, F^t\rangle_{H^m}\right] = \widehat{A}R_\lambda(\widehat{A}^2)f, \tag{22}$$

$$\widehat{A}^2 x - \widehat{A}^2 R_\lambda(\widehat{A}^2)\left[S\langle \widehat{A}x, F^t\rangle_{H^m} + G\langle \widehat{A}^2 x, F^t\rangle_{H^m}\right] = \widehat{A}^2 R_\lambda(\widehat{A}^2)f. \tag{23}$$

Taking the inner products of (22) and (23) with the components of the vector F^t, we obtain the system

$$\langle F^t, \widehat{A}x\rangle_{H^m} - \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)\left[S\overline{\langle F^t, \widehat{A}x\rangle}_{H^m} + G\overline{\langle F^t, \widehat{A}^2 x\rangle}_{H^m}\right]\rangle_{H^m} = \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)f\rangle_{H^m}, \tag{24}$$

$$\langle F^t, \widehat{A}^2 x\rangle_{H^m} - \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)\left[S\overline{\langle F^t, \widehat{A}x\rangle}_{H^m} + G\overline{\langle F^t, \widehat{A}^2 x\rangle}_{H^m}\right]\rangle_{H^m} = \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)f\rangle_{H^m}. \tag{25}$$

Exploiting the conjugate linear property of the inner product as in (3), we get

$$\left[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)S\rangle_{H^m}\right]\langle F^t, \widehat{A}x\rangle_{H^m} - \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)G\rangle_{H^m}\langle F^t, \widehat{A}^2 x\rangle_{H^m} = \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)f\rangle_{H^m}, \tag{26}$$

$$-\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)S\rangle_{H^m}\langle F^t, \widehat{A}x\rangle_{H^m} + \left[I_m - \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)G\rangle_{H^m}\right]\langle F^t, \widehat{A}^2 x\rangle_{H^m} = \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)f\rangle_{H^m}. \tag{27}$$

Writing Eqs. (26) and (27) in a matrix form by using the matrix W_λ in (18) and inverting, since $\det W_\lambda \neq 0$, we have

$$\begin{pmatrix} \langle F^t, \widehat{A}x\rangle_{H^m} \\ \langle F^t, \widehat{A}^2 x\rangle_{H^m} \end{pmatrix} = \begin{pmatrix} W^{-1}_{\lambda 11} & W^{-1}_{\lambda 12} \\ W^{-1}_{\lambda 21} & W^{-1}_{\lambda 22} \end{pmatrix} \begin{pmatrix} \langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)f\rangle_{H^m} \\ \langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)f\rangle_{H^m} \end{pmatrix}, \tag{28}$$

or, by taking the conjugates,

$$\begin{pmatrix} \langle \widehat{A}x, F^t\rangle_{H^m} \\ \langle \widehat{A}^2 x, F^t\rangle_{H^m} \end{pmatrix} = \overline{\begin{pmatrix} W^{-1}_{\lambda 11} & W^{-1}_{\lambda 12} \\ W^{-1}_{\lambda 21} & W^{-1}_{\lambda 22} \end{pmatrix}} \begin{pmatrix} \langle \widehat{A}R_\lambda(\widehat{A}^2)f, F^t\rangle_{H^m} \\ \langle \widehat{A}^2 R_\lambda(\widehat{A}^2)f, F^t\rangle_{H^m} \end{pmatrix}. \tag{29}$$

Substituting (29) into (21) and using $x = (\mathbf{B}-\lambda I)^{-1}f$ the resolvent operator $R_\lambda(\mathbf{B})$ is obtained as in Eq. (19). Since $\widehat{A}^2 R_\lambda(\widehat{A}^2)f = (I + \lambda R_\lambda(\widehat{A}^2))f$ by Proposition 1 and $R_\lambda(\widehat{A}^2)$ is defined on the whole space H and is bounded, it follows that $R_\lambda(\mathbf{B})$ is also defined on the whole space H and is bounded. Hence $\lambda \in \rho(\mathbf{B})$.

Conversely, let $\lambda \in \rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$. We will show that $\det W_\lambda \neq 0$. Suppose $\det W_\lambda = 0$, then $\det \overline{W}_\lambda = 0$. Hence, there exists a nonzero vector $\mathbf{a}^t = \operatorname{col}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{col}(a_1, \ldots, a_m, a_{m+1}, \ldots, a_{2m}) \in \mathbf{C}^{2m}$ such that $\overline{W}_\lambda \mathbf{a}^t = \mathbf{0}^t$. We consider the element $x_0 = R_\lambda(\widehat{A}^2)(S\mathbf{a}_1^t + G\mathbf{a}_2^t) \in D(\widehat{A}^2)$. The components of the vector (S, G) are linearly independent and therefore $S\mathbf{a}_1^t + G\mathbf{a}_2^t \neq 0$ and hence $x_0 \neq 0$. Substituting into (17), we have

$$\begin{aligned}
\mathbf{B}_\lambda x_0 &= (\widehat{A}^2 - \lambda I)x_0 - S\overline{\langle F^t, \widehat{A}x_0\rangle}_{H^m} - G\overline{\langle F^t, \widehat{A}^2 x_0\rangle}_{H^m} \\
&= (\widehat{A}^2 - \lambda I)[R_\lambda(\widehat{A}^2)(S\mathbf{a}_1^t + G\mathbf{a}_2^t)] - S\overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)(S\mathbf{a}_1^t + G\mathbf{a}_2^t)\rangle}_{H^m} \\
&\quad - G\overline{\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)(S\mathbf{a}_1^t + G\mathbf{a}_2^t)\rangle}_{H^m} \\
&= S\mathbf{a}_1^t + G\mathbf{a}_2^t - S\overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)S\rangle}_{H^m}\mathbf{a}_1^t - S\overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)G\rangle}_{H^m}\mathbf{a}_2^t \\
&\quad - G\overline{\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)S\rangle}_{H^m}\mathbf{a}_1^t - G\overline{\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)G\rangle}_{H^m}\mathbf{a}_2^t \\
&= S\left[\left(I_m - \overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)S\rangle}_{H^m}\right)\mathbf{a}_1^t - \overline{\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)G\rangle}_{H^m}\mathbf{a}_2^t\right] \\
&\quad + G\left[\left(I_m - \overline{\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)G\rangle}_{H^m}\right)\mathbf{a}_2^t - \overline{\langle F^t, \widehat{A}^2 R_\lambda(\widehat{A}^2)S\rangle}_{H^m}\mathbf{a}_1^t\right] \\
&= (S, G)\overline{W}_\lambda \mathbf{a}^t = (S, G)\mathbf{0}^t = 0.
\end{aligned} \tag{30}$$

Consequently, $x_0 \in \ker \mathbf{B}_\lambda$ and so $\lambda \notin \rho(\mathbf{B})$, which is not true since $\lambda \in \rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$. Therefore $\det W_\lambda \neq 0$ and $\rho(\mathbf{B}) \cap \rho(\widehat{A}^2) = \{\lambda \in \rho(\widehat{A}^2) : \det W_\lambda \neq 0\}$. □

Remark 2. The linear independence of the components of the vector (S, G) is required to prove the necessary condition of (i). So, if the components of the vector (S, G) are not linearly independent, then holds only: $\lambda \in \rho(\mathbf{B}))$ if $\det W_\lambda \neq 0$.

Resolvent of Quadratic Operators

In reference [22, Theorem 4.6] it is stated that for the operator $\mathbf{B}$ defined in (16) holds $\mathbf{B} = B^2$, where B is as in (5), if and only if

$$G \in D(\widehat{A}) \text{ and } S = BG = \widehat{A}G - G\overline{\langle F^t, \widehat{A}G\rangle}_{H^m}. \tag{31}$$

This is important because it provides the means to construct the resolvent $R_{\lambda^2}(B^2)$ of the quadratic operator B^2 from the resolvents $R_\lambda(\widehat{A})$ and $R_{-\lambda}(\widehat{A})$ as it is shown below.

Before we articulate the main theorem we have to prove first the next lemma.

Lemma 2. *If $\widehat{A} : H \to H$ is a closed linear operator in a complex Hilbert space H and the resolvent set $\rho(\widehat{A})$ is nonempty, then:*

(i) For $\pm\lambda \in \rho(\widehat{A})$ the resolvent operators $R_\lambda(\widehat{A})$ and $R_{-\lambda}(\widehat{A})$ are bounded and are defined on the whole H and commute.
(ii) The resolvent set $\rho(\widehat{A}^2) = \{\lambda^2 \in \mathbf{C} : \pm\lambda \in \rho(\widehat{A})\}$, i.e. $\lambda^2 \in R_{\lambda^2}(\widehat{A}^2)$ if and only if $\pm\lambda \in \rho(\widehat{A})$. The resolvent operator $R_{\lambda^2}(\widehat{A}^2)$ is bounded and is defined on the whole H. Moreover,

$$R_{\lambda^2}(\widehat{A}^2) = R_\lambda(\widehat{A})R_{-\lambda}(\widehat{A}), \tag{32}$$

$$\widehat{A}R_{\lambda^2}(\widehat{A}^2) = R_{-\lambda}(\widehat{A}) + \lambda R_{\lambda^2}(\widehat{A}^2), \tag{33}$$

$$\widehat{A}R_{\lambda^2}(\widehat{A}^2) = R_\lambda(\widehat{A}) - \lambda R_{\lambda^2}(\widehat{A}^2). \tag{34}$$

(iii) If in addition $\lambda \neq 0$, then

$$R_{\lambda^2}(\widehat{A}^2) = \frac{1}{2\lambda}[R_\lambda(\widehat{A}) - R_{-\lambda}(\widehat{A})], \tag{35}$$

$$\widehat{A}R_{\lambda^2}(\widehat{A}^2) = \frac{1}{2}\left[R_\lambda(\widehat{A}) + R_{-\lambda}(\widehat{A})\right], \tag{36}$$

$$\widehat{A}^2R_{\lambda^2}(\widehat{A}^2) = I + \frac{\lambda}{2}\left[R_\lambda(\widehat{A}) - R_{-\lambda}(\widehat{A})\right]. \tag{37}$$

Proof. (i) Since $\widehat{A}$ is a closed linear operator and $\pm\lambda \in \rho(\widehat{A})$ then the resolvent operators $R_{-\lambda}(\widehat{A})$ and $R_\lambda(\widehat{A})$ are bounded and are defined on the whole H by Lemma 1. The commuting property $R_\lambda(\widehat{A})R_{-\lambda}(\widehat{A}) = R_{-\lambda}(\widehat{A})R_\lambda(\widehat{A})$ follows from the resolvent equation

$$R_{\lambda_1}(\widehat{A}) - R_{\lambda_2}(\widehat{A}) = (\lambda_1 - \lambda_2)R_{\lambda_1}(\widehat{A})R_{\lambda_2}(\widehat{A}) \tag{38}$$

which holds for every closed operator $\widehat{A}$ in a complex Banach space and for every $\lambda_1, \lambda_2 \in \rho(\widehat{A})$, see, e.g., [10].

(ii) Note that for each $\lambda \in \mathbf{C}$ and $D(\widehat{A} \pm \lambda I) = D(\widehat{A})$, $D(\widehat{A}^2 \pm \lambda^2 I) = D(\widehat{A}^2)$ we have

$$\begin{aligned}&\widehat{A}^2 - \lambda^2 I = (\widehat{A} + \lambda I)(\widehat{A} - \lambda I),\\&(\widehat{A}^2 - \lambda^2 I)x = (\widehat{A} + \lambda I)(\widehat{A} - \lambda I)x = f, \quad x \in D(\widehat{A}^2), \quad f \in H.\end{aligned} \tag{39}$$

Let $\pm\lambda \in \rho(\widehat{A})$. Then because of case (i) we get

$$\begin{aligned}&R_{-\lambda}(\widehat{A})(\widehat{A} + \lambda I)(\widehat{A} - \lambda I)x = R_{-\lambda}(\widehat{A})f,\\&R_{\lambda}(\widehat{A})(\widehat{A} - \lambda I)x = R_{\lambda}(\widehat{A})R_{-\lambda}(\widehat{A})f,\\&x = R_{\lambda}(\widehat{A})R_{-\lambda}(\widehat{A})f.\end{aligned} \tag{40}$$

It follows from (39) that

$$x = (\widehat{A}^2 - \lambda^2 I)^{-1} f = R_{\lambda}(\widehat{A})R_{-\lambda}(\widehat{A})f \tag{41}$$

and hence $R_{\lambda^2}(\widehat{A}^2)$ is bounded and is defined on the whole H. Thus, $\lambda^2 \in \rho(\widehat{A}^2)$.

Conversely, by Proposition 2 the operator $\widehat{A}^2$ is closed. Let $\lambda^2 \in \rho(\widehat{A}^2)$. Then by Lemma 1, the resolvent operator $R_{\lambda^2}(\widehat{A}^2)$ is bounded and defined on the whole H. From (39) and since $\ker(\widehat{A}^2 - \lambda^2 I) = \{0\}$ and the operators $(\widehat{A} - \lambda I)$, $(\widehat{A} + \lambda I)$ commute, it follows that $\ker(\widehat{A} \pm \lambda I) = \{0\}$ and $(\widehat{A}^2 - \lambda^2 I)^{-1} = (\widehat{A} - \lambda I)^{-1}(\widehat{A} + \lambda I)^{-1}$. Because $R(\widehat{A}^2 - \lambda^2 I) = H$ it is implied that $R(\widehat{A} \pm \lambda I) = H$. Since $\widehat{A}$ is closed then $\widehat{A} \pm \lambda I$ and $(\widehat{A} \pm \lambda I)^{-1}$ are closed too. By the Closed Graph theorem the operators $(\widehat{A} \pm \lambda I)^{-1}$ are bounded. Hence $\pm\lambda \in \rho(\widehat{A})$.

Furthermore, the identity (32) follows from Eq. (41) while the identities (33) and (34) are easily proved, viz.

$$\begin{aligned}&\widehat{A}R_{\lambda^2}(\widehat{A}^2) = \widehat{A}R_{\lambda}(\widehat{A})R_{-\lambda}(\widehat{A}) = [I + \lambda R_{\lambda}(\widehat{A})]R_{-\lambda}(\widehat{A}) = R_{-\lambda}(\widehat{A}) + \lambda R_{\lambda^2}(\widehat{A}^2),\\&\widehat{A}R_{\lambda^2}(\widehat{A}^2) = \widehat{A}R_{-\lambda}(\widehat{A})R_{\lambda}(\widehat{A}) = [I - \lambda R_{-\lambda}(\widehat{A})]R_{\lambda}(\widehat{A}) = R_{\lambda}(\widehat{A}) - \lambda R_{\lambda^2}(\widehat{A}^2).\end{aligned}$$

(iii) From (41) and (38) we get

$$R_{\lambda^2}(\widehat{A}^2) = R_{\lambda}(\widehat{A})R_{-\lambda}(\widehat{A}) = \frac{1}{2\lambda}[R_{\lambda}(\widehat{A}) - R_{-\lambda}(\widehat{A})].$$

By acting $\widehat{A}$ on the last equation we obtain

$$\begin{aligned}\widehat{A}R_{\lambda^2}(\widehat{A}^2) &= \frac{1}{2\lambda}[\widehat{A}R_{\lambda}(\widehat{A}) - \widehat{A}R_{-\lambda}(\widehat{A})]\\&= \frac{1}{2\lambda}\{I + \lambda R_{\lambda}(\widehat{A}) - [I + (-\lambda)R_{-\lambda}(\widehat{A})]\} = \frac{1}{2}\left[R_{\lambda}(\widehat{A}) + R_{-\lambda}(\widehat{A})\right].\end{aligned}$$

Finally, operating $\widehat{A}$ on this equation we get (37). This completes the proof. □

Now we present the main theorem for the resolvent of the quadratic operator $\mathbf{B}$.

Theorem 3. *Let H be a complex Hilbert space, $\widehat{A} : H \to H$ a linear closed operator, $\pm\lambda \in \rho(\widehat{A})$ and $\mathbf{B}$ the operator defined by (16). Suppose the vectors G, S satisfy (31) and the components of the vector (S, G) are linearly independent elements. Let the operator*

$$\mathbf{B}_{\lambda^2}x = (\mathbf{B} - \lambda^2 I)x = \widehat{A}^2 x - S\langle \widehat{A}x, F^t\rangle_{H^m} - G\langle \widehat{A}^2 x, F^t\rangle_{H^m} - \lambda^2 x = f,$$
$$D(\mathbf{B}_{\lambda^2}) = D(\widehat{A}^2), \tag{42}$$

where $f \in H$. Then:

(i) $\lambda^2 \in \rho(\mathbf{B})$ *if and only if*

$$\det L_\lambda = \det[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle_{H^m}] \neq 0, \tag{43}$$

$$\det L_{-\lambda} = \det[I_m - \langle F^t, \widehat{A}R_{-\lambda}(\widehat{A})G\rangle_{H^m}] \neq 0. \tag{44}$$

(ii) $\rho(\mathbf{B}) \cap \rho(\widehat{A}^2) = \{\lambda^2 \in \rho(\widehat{A}^2) : \det L_\lambda \neq 0, \ \det L_{-\lambda} \neq 0\}$.

(iii) *If* $\lambda \neq 0$, $\det L_\lambda \neq 0$ *and* $\det L_{-\lambda} \neq 0$, *then there exists the resolvent operator* $R_{\lambda^2}(\mathbf{B})$ *which is defined on the whole space H, is bounded and is given by*

$$R_{\lambda^2}(\mathbf{B})f = \frac{1}{2\lambda}\Big[R_\lambda(\widehat{A})f - R_{-\lambda}(\widehat{A})f + R_\lambda(\widehat{A})G\overline{L_\lambda^{-1}}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m}$$
$$- R_{-\lambda}(\widehat{A})G\overline{L_{-\lambda}^{-1}}\langle \widehat{A}R_{-\lambda}(\widehat{A})f, F^t\rangle_{H^m}\Big]. \tag{45}$$

Proof. Note that here, for brevity and for ease of presentation, we denote the inner products without their index $_{H^m}$, e.g. $\langle F^t, G\rangle_{H^m} = \langle F^t, G\rangle$. Also, we use some others shorthands which are explained as they appear.

(i) and (ii) Since $\widehat{A}$ is a closed, so is $\widehat{A}^2$ by Proposition 2. By hypothesis the vectors S and G satisfy (31) and hence by [22, Theorem 4.6] the operator $\mathbf{B} = B^2$, where B as in (5). Theorem 2 affirms that $\lambda^2 \in \rho(\mathbf{B})$ if and only if $\det W_{\lambda^2} \neq 0$. By introducing the notations $T_1 = \widehat{A}R_{\lambda^2}(\widehat{A}^2)$ and $T_2 = \widehat{A}^2 R_{\lambda^2}(\widehat{A}^2)$ for convenience, the $\det W_{\lambda^2}$ in Eq. (18) is written as follows:

$$\det W_{\lambda^2} = |W_{\lambda^2}| = \begin{vmatrix} I_m - \langle F^t, T_1 S\rangle & -\langle F^t, T_1 G\rangle \\ -\langle F^t, T_2 S\rangle & I_m - \langle F^t, T_2 G\rangle \end{vmatrix}. \tag{46}$$

Substituting $S = \widehat{A}G - G\overline{\langle F^t, \widehat{A}G\rangle}$ and utilizing property (3), we have

$$|W_{\lambda^2}| = \begin{vmatrix} I_m - \langle F^t, T_1\Big[\widehat{A}G - G\overline{\langle F^t, \widehat{A}G\rangle}\Big]\rangle & -\langle F^t, T_1 G\rangle \\ -\langle F^t, T_2\Big[\widehat{A}G - G\overline{\langle F^t, \widehat{A}G\rangle}\Big]\rangle & I_m - \langle F^t, T_2 G\rangle \end{vmatrix}$$

$$= \begin{vmatrix} I_m - \langle F^t, T_1\widehat{A}G\rangle + \langle F^t, T_1G\rangle\langle F^t, \widehat{A}G\rangle & -\langle F^t, T_1G\rangle \\ -\langle F^t, T_2\widehat{A}G\rangle + \langle F^t, T_2G\rangle\langle F^t, \widehat{A}G\rangle & I_m - \langle F^t, T_2G\rangle \end{vmatrix}. \tag{47}$$

Multiplying from the right the elements of the second column by $\langle F^t, \widehat{A}G\rangle$ and adding to the matching elements of the first column, and replacing $T_1 = \widehat{A}R_{\lambda^2}(\widehat{A^2})$ and $T_2 = \widehat{A}^2R_{\lambda^2}(\widehat{A^2})$, we obtain

$$|W_{\lambda^2}| = \begin{vmatrix} I_m - \langle F^t, \widehat{A}R_{\lambda^2}(\widehat{A^2})\widehat{A}G\rangle & -\langle F^t, \widehat{A}R_{\lambda^2}(\widehat{A^2})G\rangle \\ \langle F^t, \widehat{A}G\rangle - \langle F^t, \widehat{A}^2R_{\lambda^2}(\widehat{A^2})\widehat{A}G\rangle & I_m - \langle F^t, \widehat{A}^2R_{\lambda^2}(\widehat{A^2})G\rangle \end{vmatrix}. \tag{48}$$

Now, by using Proposition 1, Eqs. (33), (34) and the commuting property $R_\lambda(\widehat{A})\widehat{A}G = \widehat{A}R_\lambda(\widehat{A})G$, and denoting $R_{\pm\lambda} = R_{\pm\lambda}(\widehat{A})$ and $P = R_\lambda R_{-\lambda}$ for ease of presentation, we get

$$\begin{aligned} |W_{\lambda^2}| &= \begin{vmatrix} I_m - \langle F^t, (I + \lambda^2P)G\rangle & -\langle F^t, (R_{-\lambda} + \lambda P)G\rangle \\ \langle F^t, \widehat{A}G\rangle - \langle F^t, \widehat{A}(I + \lambda^2P)G\rangle & I_m - \langle F^t, (I + \lambda^2P)G\rangle \end{vmatrix} \\ &= \begin{vmatrix} I_m - \langle F^t, G\rangle - \overline{\lambda^2}\langle F^t, PG\rangle & -\langle F^t, R_{-\lambda}G\rangle - \overline{\lambda}\langle F^t, PG\rangle \\ -\overline{\lambda^2}\langle F^t, R_{-\lambda}G\rangle - \overline{\lambda^3}\langle F^t, PG\rangle & I_m - \langle F^t, G\rangle - \overline{\lambda^2}\langle F^t, PG\rangle \end{vmatrix}. \end{aligned} \tag{49}$$

Multiplying the elements of the first row by $-\overline{\lambda}$ and adding to the corresponding elements of the second row, we obtain

$$|W_{\lambda^2}| = \begin{vmatrix} I_m - \langle F^t, G\rangle - \overline{\lambda^2}\langle F^t, PG\rangle & -\langle F^t, R_{-\lambda}G\rangle - \overline{\lambda}\langle F^t, PG\rangle \\ -\overline{\lambda}I + \overline{\lambda}\langle F^t, G\rangle - \overline{\lambda^2}\langle F^t, R_{-\lambda}G\rangle & I_m - \langle F^t, G\rangle + \overline{\lambda}\langle F^t, R_{-\lambda}G\rangle \end{vmatrix}. \tag{50}$$

Furthermore, multiplying the elements of the second column by $\overline{\lambda}$ and adding to the matching elements of the first column, and replacing $P = R_\lambda R_{-\lambda}$, we get

$$|W_{\lambda^2}| = \begin{vmatrix} I_m - \langle F^t, (I + \lambda R_{-\lambda} + 2\lambda^2R_\lambda R_{-\lambda})G\rangle & -\langle F^t, (I + \lambda R_\lambda)R_{-\lambda}G\rangle \\ 0_m & I_m - \langle F^t, (I - \lambda R_{-\lambda})G\rangle \end{vmatrix}. \tag{51}$$

Using Proposition 1 and (32)–(34) we get $I - \lambda R_{-\lambda} = \widehat{A}R_{-\lambda}$ and

$$\begin{aligned} I + \lambda R_{-\lambda} + 2\lambda^2R_\lambda R_{-\lambda} &= I + \lambda(\widehat{A}R_\lambda R_{-\lambda} - \lambda R_\lambda R_{-\lambda}) + 2\lambda^2R_\lambda R_{-\lambda} \\ &= I + \lambda\widehat{A}R_\lambda R_{-\lambda} + \lambda^2R_\lambda R_{-\lambda} \\ &= I + \lambda(R_\lambda - \lambda R_{-\lambda}R_\lambda) + \lambda^2R_\lambda R_{-\lambda} \\ &= I + \lambda R_\lambda = \widehat{A}R_\lambda. \end{aligned} \tag{52}$$

Substituting these results in (51), we acquire

$$\det W_{\lambda^2} = \det\left[I_m - \langle F^t, \widehat{A}R_\lambda G\rangle\right] \cdot \det\left[I_m - \langle F^t, \widehat{A}R_{-\lambda} G\rangle\right], \tag{53}$$

Hence, $\lambda^2 \in \rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$ if and only if $\det L_{\pm\lambda} \neq 0$.

(iii) It has been proven that $\widehat{A}^2$ is a closed operator and $\lambda^2 \in \rho(\mathbf{B}) \cap \rho(\widehat{A}^2)$. Then $\pm\lambda \in \rho(\widehat{A})$ by Lemma 2 and also $\pm\lambda \in \rho(B)$ by Theorem 1. From Lemma 2 for $\lambda \neq 0$ we find

$$R_{\lambda^2}(\mathbf{B}) = \frac{1}{2\lambda}[R_\lambda(B) - R_{-\lambda}(B)], \tag{54}$$

where the resolvent operator $R_\lambda(B)$ is set out in (8) and $R_{-\lambda}(B)$ is the same as in (8) except that λ is replaced by $-\lambda$. Substituting these formulas into (54) we obtain (45). The resolvent operators $R_{\pm\lambda}(B)$ are bounded and are defined on the whole H and so is $R_{\lambda^2}(\mathbf{B})$ by (54). This completes the proof.

Applications

In this section we apply the theory presented in the previous sections to boundary value problems involving integro-differential equations of the Fredholm type. In particular, we find the resolvent sets and provide closed form representations for the resolvent operators. Some auxiliary results needed are quoted in Appendix for ease of reference. By $H^1(0,1)$ (resp. $H^2(0,1)$) is denoted the Sobolev space of all complex functions of $L_2(0,1)$ which have generalized derivatives up to the first (resp. second) order that are Lebesgue integrable.

First Order Integro-differential Equation

Consider the following integro-differential boundary value problem

$$\begin{aligned} &iu'(t) - ie^{i\pi t}\int_0^1 xu'(x)dx - \lambda u(t) = f(t), \\ &u(0) + u(1) = 0, \quad u(t) \in H^1(0,1). \end{aligned} \tag{55}$$

In Problem 1 in the Appendix it is quoted that the operator $\widehat{A} : L_2(0,1) \to L_2(0,1)$ defined by

$$\widehat{A}u = iu' = f, \quad D(\widehat{A}) = \{u(t) \in H^1(0,1) : u(0) + u(1) = 0\}, \tag{56}$$

is a linear closed operator and that $\rho(\widehat{A}) = \{\lambda \in \mathbf{C} : \lambda \neq (2k+1)\pi,\ k \in \mathbf{Z}\}$, while for $\lambda \in \rho(\widehat{A})$ the resolvent operator $R_\lambda(\widehat{A})$ is defined on the whole space $L_2(0,1)$, is bounded and is given explicitly by the formula (98). Let $B : L_2(0,1) \to L_2(0,1)$ be the operator

$$\begin{aligned} Bu(t) &= iu'(t) - ie^{i\pi t}\int_0^1 xu'(x)dx \\ &= \widehat{A}u(t) - G\langle \widehat{A}u(t), F^t\rangle, \quad D(B) = D(\widehat{A}), \end{aligned} \tag{57}$$

where $G = e^{i\pi t}$ and $F = t$. We express the boundary value problem (55) in the operator form

$$B_\lambda u(t) = (B - \lambda I)u(t) = f(t), \quad D(B_\lambda) = D(\widehat{A}). \tag{58}$$

In applying Theorem 1 we have first to compute the determinant

$$\det L_\lambda = \det\Big[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle_{H^m}\Big] \neq 0. \tag{59}$$

Acting $R_\lambda(\widehat{A})$ on G, operating by $\widehat{A}$ and taking the inner products, we find

$$\begin{aligned} &R_\lambda(\widehat{A})G = ie^{-i\lambda t}\left[(e^{i\lambda}+1)^{-1}\int_0^1 e^{i\pi x}e^{i\lambda x}dx - \int_0^t e^{i\pi x}e^{i\lambda x}dx\right] = -\frac{e^{i\pi t}}{\pi+\lambda}, \\ &\widehat{A}R_\lambda(\widehat{A})G = \frac{\pi}{\pi+\lambda}e^{i\pi t}, \\ &\langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle_{H^m} = \frac{\pi}{\pi+\overline{\lambda}}\int_0^1 xe^{-i\pi x}dx = -\frac{2+i\pi}{(\pi+\overline{\lambda})\pi}. \end{aligned} \tag{60}$$

Substituting Eq. (60) into (59) we have

$$\det L_\lambda = \det\left[1 + \frac{2+i\pi}{(\pi+\overline{\lambda})\pi}\right] = \frac{\pi^2+\overline{\lambda}\pi+i\pi+2}{(\pi+\overline{\lambda})\pi} \neq 0, \tag{61}$$

which implies that $\det L_\lambda \neq 0$ if and only if $\lambda \neq -(\pi^2+2)/\pi + i$. Then from Theorem 1 it follows that $\lambda \in \rho(B)$ if $\lambda \neq (2k+1)\pi,\ k \in \mathbf{Z}$ and $\lambda \neq -(\pi^2+2)/\pi + i$. Moreover, the resolvent operator $R_\lambda(B)$ is bounded and defined in all $L_2(0,1)$ and is given by

$$R_\lambda(B)f = R_\lambda(\widehat{A})f + R_\lambda(\widehat{A})G\overline{L_\lambda^{-1}}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m}. \tag{62}$$

Applying $\widehat{A}$ on Eq. (98) and then forming the inner products we get

$$\widehat{A}R_\lambda(\widehat{A})f = e^{-i\lambda t} i\lambda \left[(e^{i\lambda}+1)^{-1} \int_0^1 f(x)e^{i\lambda x}dx - \int_0^t f(x)e^{i\lambda x}dx \right] + f(t),$$

$$\begin{aligned}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m} &= \int_0^1 \Big\{ e^{-i\lambda t} i\lambda \Big[(e^{i\lambda}+1)^{-1} \int_0^1 f(x)e^{i\lambda x}dx \\ &\quad - \int_0^t f(x)e^{i\lambda x}dx \Big] + f(t) \Big\}\, tdt \\ &= i\lambda (e^{i\lambda}+1)^{-1} \int_0^1 e^{-i\lambda t}tdt \int_0^1 f(x)e^{i\lambda x}dx \\ &\quad - i\lambda \int_0^1 e^{-i\lambda t} \left[\int_0^t f(x)e^{i\lambda x}dx \right] tdt + \int_0^1 tf(t)dt. \qquad (63)\end{aligned}$$

Exploiting the Fubini theorem we find

$$\begin{aligned}\langle \widehat{A}R_\lambda(\widehat{A})f, F^t\rangle_{H^m} &= i\lambda(e^{i\lambda}+1)^{-1} \int_0^1 f(x)e^{i\lambda x}dx \,\frac{1}{\lambda^2}\left[e^{-i\lambda}(i\lambda+1)-1 \right] \\ &\quad - i\lambda \int_0^1 f(x)e^{i\lambda x}dx \int_x^1 e^{-i\lambda t}tdt + \int_0^1 f(t)tdt \\ &= \frac{i\left[e^{-i\lambda}(i\lambda+1)-1\right]}{\lambda(e^{i\lambda}+1)} \int_0^1 f(x)e^{i\lambda x}dx \\ &\quad - \frac{i(1+i\lambda)}{\lambda} \int_0^1 f(x)e^{i\lambda(x-1)}dx \\ &\quad + \int_0^1 \frac{i(i\lambda x+1)}{\lambda} f(x)dx + \int_0^1 f(x)xdx \\ &= \frac{i\left[e^{-i\lambda}(i\lambda+1)-1\right]}{\lambda(e^{i\lambda}+1)} \int_0^1 f(x)e^{i\lambda x}dx \\ &\quad - \frac{i(1+i\lambda)}{\lambda} \int_0^1 f(x)e^{i\lambda(x-1)}dx + \frac{i}{\lambda}\int_0^1 f(x)dx \\ &= \frac{i\left[e^{-i\lambda}(i\lambda+1)-1\right]}{\lambda(e^{i\lambda}+1)} \int_0^1 e^{i\lambda x}f(x)dx \\ &\quad - i\int_0^1 \frac{(1+i\lambda)e^{i\lambda(x-1)}-1}{\lambda} f(x)dx. \qquad (64)\end{aligned}$$

Finally, by substituting (98), (60), (64) and the inverse $\overline{L_\lambda^{-1}}$ from (61) into (62), we get

$$\begin{aligned} R_\lambda(B)f(t) &= R_\lambda(\widehat{A})f(t) \\ &\quad - \frac{i\pi e^{i\pi t}}{\pi^2+\lambda\pi - i\pi + 2} \Big\{ \frac{e^{-i\lambda}(i\lambda+1)-1}{\lambda(e^{i\lambda}+1)} \int_0^1 f(x)e^{i\lambda x}dx \end{aligned}$$

$$
\begin{aligned}
&\quad - \int_0^1 \frac{(1+i\lambda)e^{i\lambda(x-1)} - 1}{\lambda} f(x)dx \Big\} \\
&= R_\lambda(\widehat{A})f(t) \\
&\quad - \frac{i\pi e^{i\pi t}(e^{i\lambda}+1)^{-1}}{\pi^2 + (\lambda - i)\pi + 2} \int_0^1 \frac{e^{i\lambda} + 1 - (2+i\lambda)e^{i\lambda x}}{\lambda} f(x)dx\,. \qquad (65)
\end{aligned}
$$

Second Order Integro-differential Equation

Consider the following boundary value problem involving a second order Fredholm integro-differential equation

$$
\begin{aligned}
&u''(t) - t\int_0^1 u'(x)\cos 2\pi x dx - \int_0^1 u''(x)\cos 2\pi x dx - \lambda u(t) = f(t), \\
&u(0) = u(1), \quad u'(0) = u'(1), \quad u(t) \in H^2(0,1). \qquad (66)
\end{aligned}
$$

We define the operators $\widehat{A},\ \widehat{A}^2 : L_2(0,1) \to L_2(0,1)$ as follows:

$$
\widehat{A}u = u'(t), \quad D(\widehat{A}) = \{u(t) \in H^1(0,1) : u(0) = u(1)\}, \qquad (67)
$$

$$
\widehat{A}^2 u = u''(t), \quad D(\widehat{A}^2) = \{u(t) \in H^2(0,1) : u(0) = u(1),\, u'(0) = u'(1)\}. \qquad (68)
$$

In Problem 2 in the Appendix it is given that the operator $\widehat{A}^2$ is closed and that the resolvent set is $\rho(\widehat{A}^2) = \{\lambda \in \mathbf{C} : \lambda \neq -4k^2\pi^2,\ k \in \mathbf{Z}\}$, whereas for $\lambda \in \rho(\widehat{A}^2)$ the resolvent operator $R_\lambda(\widehat{A}^2)$ is defined on the whole space $L_2(0,1)$, is bounded and is expressed analytically in (105). Additionally, we define the operator $\mathbf{B} : L_2(0,1) \to L_2(0,1)$ as

$$
\begin{aligned}
\mathbf{B}u(t) &= u''(t) - t\int_0^1 u'(x)\cos 2\pi x dx - \int_0^1 u''(x)\cos 2\pi x dx \\
&= \widehat{A}^2 x - S\langle \widehat{A}x, F^t\rangle_{H^m} - G\langle \widehat{A}^2 x, F^t\rangle_{H^m}, \quad D(\mathbf{B}) = D(\widehat{A}^2), \qquad (69)
\end{aligned}
$$

where $S = t$, $G = 1$ and $F = \cos 2\pi t$. We reformulate the integro-differential equation (66) as

$$
\mathbf{B}_\lambda u(t) = (\mathbf{B} - \lambda I)u(t) = f(t), \quad D(\mathbf{B}_\lambda) = D(\widehat{A}^2). \qquad (70)
$$

According to Theorem 2, any λ from $\rho(\widehat{A}^2)$ belongs to $\rho(\mathbf{B})$ if and only if Eq. (18) is satisfied, i.e. $\det W_\lambda \neq 0$. By acting $R_\lambda(\widehat{A}^2)$ from Eq. (105) on S, applying $\widehat{A}$ and $\widehat{A}^2$ and taking the inner products, we have

$$R_\lambda(\widehat{A}^2)S = \frac{1}{2\lambda}\left[\frac{e^{t\sqrt{\lambda}}}{e^{\sqrt{\lambda}}-1} + \frac{e^{(1-t)\sqrt{\lambda}}}{1-e^{\sqrt{\lambda}}} - 2t\right], \tag{71}$$

$$\widehat{A}R_\lambda(\widehat{A}^2)S = \frac{1}{2\sqrt{\lambda}}\left[\frac{e^{t\sqrt{\lambda}}}{e^{\sqrt{\lambda}}-1} + \frac{e^{(1-t)\sqrt{\lambda}}}{e^{\sqrt{\lambda}}-1} - \frac{2}{\sqrt{\lambda}}\right], \tag{72}$$

$$\widehat{A}^2R_\lambda(\widehat{A}^2)S = \frac{e^{t\sqrt{\lambda}}}{2(e^{\sqrt{\lambda}}-1)} + \frac{e^{(1-t)\sqrt{\lambda}}}{2(1-e^{\sqrt{\lambda}})}, \tag{73}$$

$$\begin{aligned}\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)S\rangle_H &= \frac{1}{2\sqrt{\overline{\lambda}}}\int_0^1\left[\frac{e^{t\sqrt{\overline{\lambda}}}}{e^{\sqrt{\overline{\lambda}}}-1} + \frac{e^{(1-t)\sqrt{\overline{\lambda}}}}{e^{\sqrt{\overline{\lambda}}}-1} - \frac{2}{\sqrt{\overline{\lambda}}}\right]\cos 2\pi t dt \\ &= \frac{1}{\overline{\lambda}+4\pi^2},\end{aligned} \tag{74}$$

$$\langle F^t, \widehat{A}^2R_\lambda(\widehat{A}^2)S\rangle_H = 0. \tag{75}$$

Imitating the same procedure for G, we get

$$R_\lambda(\widehat{A}^2)G = -\frac{1}{\lambda}, \tag{76}$$

$$\widehat{A}R_\lambda(\widehat{A}^2)G = \widehat{A}^2R_\lambda(\widehat{A}^2)G = 0, \tag{77}$$

$$\langle F^t, \widehat{A}R_\lambda(\widehat{A}^2)G\rangle_H = \langle F^t, \widehat{A}^2R_\lambda(\widehat{A}^2)G\rangle_H = 0. \tag{78}$$

Substituting Eqs. (74), (75) and (78) into (18), we obtain

$$\det\, W_\lambda = \det\begin{pmatrix} 1-\frac{1}{(\overline{\lambda}+4\pi^2)} & 0 \\ 0 & 1\end{pmatrix} = \frac{\overline{\lambda}+4\pi^2-1}{\overline{\lambda}+4\pi^2} \neq 0. \tag{79}$$

Hence $\lambda \in \rho(\mathbf{B})$ if $\lambda \neq -4k^2\pi^2,\ k \in \mathbf{Z}$ and $\lambda \neq 1-4\pi^2$. Moreover, Theorem 2 provides the resolvent operator $R_\lambda(\mathbf{B})$ as in formula (19).

The inverse matrix W_λ^{-1} is easily computed as

$$W_\lambda^{-1} = \begin{pmatrix} W_{\lambda 11}^{-1} & W_{\lambda 12}^{-1} \\ W_{\lambda 21}^{-1} & W_{\lambda 22}^{-1}\end{pmatrix} = \begin{pmatrix} \frac{\bar{\lambda}+4\pi^2}{\bar{\lambda}+4\pi^2-1} & 0 \\ 0 & 1\end{pmatrix}. \tag{80}$$

Using Eqs. (71), (76) and (80) as well as the fact that the operator $R_\lambda(\widehat{A}^2)$ is linear, we find

$$R_\lambda(\widehat{A}^2)(S\overline{W_{\lambda 11}^{-1}} + G\overline{W_{\lambda 21}^{-1}}) = \frac{\lambda+4\pi^2}{2\lambda(\lambda+4\pi^2-1)}\left[\frac{e^{t\sqrt{\lambda}}}{e^{\sqrt{\lambda}}-1} + \frac{e^{(1-t)\sqrt{\lambda}}}{1-e^{\sqrt{\lambda}}} - 2t\right], \tag{81}$$

$$R_\lambda(\widehat{A}^2)(S\overline{W^{-1}_{\lambda 12}} + G\overline{W^{-1}_{\lambda 22}}) = -\frac{1}{\lambda}. \tag{82}$$

Acting $\widehat{A}$ on (105), taking the inner products and utilizing the Fubini theorem, we obtain

$$\widehat{A}R_\lambda(\widehat{A}^2)f(t) = \frac{1}{2(1-e^{\sqrt{\lambda}})}\int_0^1 \left[e^{\sqrt{\lambda}(t-x+1)} - e^{-\sqrt{\lambda}(t-x)}\right]f(x)dx$$
$$+\frac{1}{2}\int_0^t \left[e^{\sqrt{\lambda}(t-x)} + e^{-\sqrt{\lambda}(t-x)}\right]f(x)dx,$$

$$\begin{aligned}
&\langle \widehat{A}R_\lambda(\widehat{A}^2)f(t), F^t\rangle_H \\
&= \frac{1}{2(1-e^{\sqrt{\lambda}})}\int_0^1 \cos 2\pi t dt \int_0^1 \left[e^{\sqrt{\lambda}(t-x+1)} - e^{-\sqrt{\lambda}(t-x)}\right]f(x)dx \\
&+\frac{1}{2}\int_0^1 \cos 2\pi t dt \int_0^t \left[e^{\sqrt{\lambda}(t-x)} + e^{-\sqrt{\lambda}(t-x)}\right]f(x)dx \\
&= \frac{1}{2(1-e^{\sqrt{\lambda}})}\int_0^1 f(x)dx \int_0^1 \cos 2\pi t \left[e^{\sqrt{\lambda}(t-x+1)} - e^{-\sqrt{\lambda}(t-x)}\right]dt + \\
&+\frac{1}{2}\int_0^1 f(x)dx \int_x^1 \cos 2\pi t \left[e^{\sqrt{\lambda}(t-x)} + e^{-\sqrt{\lambda}(t-x)}\right]dt \\
&= \frac{1}{2(1-e^{\sqrt{\lambda}})}\int_0^1 \frac{\sqrt{\lambda}\left(e^{\sqrt{\lambda}(x-1)} - e^{\sqrt{\lambda}x} + e^{\sqrt{\lambda}(2-x)} - e^{\sqrt{\lambda}(1-x)}\right)}{\lambda+4\pi^2} f(x)dx + \\
&+\frac{1}{2}\int_0^1 \frac{-\sqrt{\lambda}e^{\sqrt{\lambda}(x-1)} + \sqrt{\lambda}e^{\sqrt{\lambda}(1-x)} - 4\pi \sin 2\pi x}{\lambda+4\pi^2} f(x)dx \\
&= \frac{-2\pi}{\lambda+4\pi^2}\int_0^1 \sin 2\pi x f(x)dx.
\end{aligned} \tag{83}$$

Working in the same way with $\widehat{A}^2$, we get

$$\begin{aligned}
&\widehat{A}^2R_\lambda(\widehat{A}^2)f(t) = f(t) + \lambda R_\lambda(\widehat{A}^2)f(t), \\
&\langle \widehat{A}^2R_\lambda(\widehat{A}^2)f(t), F^t\rangle_H = \langle f(t) + \lambda R_\lambda(\widehat{A}^2)f(t), F^t\rangle_H \\
&= \langle f(t), F^t\rangle_H + \lambda\langle R_\lambda(\widehat{A}^2)f(t), F^t\rangle_H \\
&= \langle f(t), F^t\rangle_H \\
&\quad + \frac{\sqrt{\lambda}}{2(1-e^{\sqrt{\lambda}})}\int_0^1 \cos 2\pi t dt \int_0^1 \left[e^{\sqrt{\lambda}(t-x+1)} + e^{\sqrt{\lambda}(x-t)}\right]f(x)dx
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\sqrt{\lambda}}{2}\int_0^1 \cos 2\pi t dt \int_0^t \left[e^{\sqrt{\lambda}(t-x)} - e^{\sqrt{\lambda}(x-t)}\right] f(x)dx. \\
&= \langle f(t), F^t\rangle_H \\
&+\frac{\sqrt{\lambda}}{2(1-e^{\sqrt{\lambda}})}\int_0^1 f(x)dx \int_0^1 \left[e^{\sqrt{\lambda}(t-x+1)} + e^{\sqrt{\lambda}(x-t)}\right]\cos 2\pi t dt \\
&+\frac{\sqrt{\lambda}}{2}\int_0^1 f(x)dx \int_x^1 \left[e^{\sqrt{\lambda}(t-x)} - e^{\sqrt{\lambda}(x-t)}\right]\cos 2\pi t dt \\
&= \langle f(t), F^t\rangle_H \\
&+\frac{\lambda}{2(1-e^{\sqrt{\lambda}})(\lambda+4\pi^2)}\int_0^1 \left[e^{\sqrt{\lambda}x} - e^{\sqrt{\lambda}(x-1)} + e^{\sqrt{\lambda}(2-x)} - e^{\sqrt{\lambda}(1-x)}\right] f(x)dx \\
&+\frac{\lambda}{2(\lambda+4\pi^2)}\int_0^1 \left[e^{\sqrt{\lambda}(x-1)} + e^{\sqrt{\lambda}(1-x)} - 2\cos 2\pi x\right] f(x)dx \\
&= \int_0^1 \cos 2\pi x f(x)dx - \frac{\lambda}{\lambda+4\pi^2}\int_0^1 \cos 2\pi x f(x)dx \\
&= \frac{4\pi^2}{\lambda+4\pi^2}\int_0^1 \cos 2\pi x f(x)dx.
\end{aligned} \tag{84}
$$

Lastly, by substituting (105) and (81)–(84) into (19) we obtain the resolvent operator $R_\lambda(\mathbf{B})$ in the following closed form

$$
\begin{aligned}
R_\lambda(\mathbf{B})f &= R_\lambda(\widehat{A}^2)f(t) \\
&\quad - \frac{\pi(e^{\sqrt{\lambda}t} - e^{\sqrt{\lambda}(1-t)} + 2t(1-e^{\sqrt{\lambda}})}{\lambda(\lambda+4\pi^2-1)(e^{\sqrt{\lambda}}-1)}\int_0^1 \sin 2\pi x f(x)dx \\
&\quad - \frac{4\pi^2}{\lambda(\lambda+4\pi^2)}\int_0^1 \cos 2\pi x f(x)dx.
\end{aligned} \tag{85}
$$

Integro-Differential Equation with a Quadratic Operator

Consider the following boundary value problem which can be expressed in terms of a quadratic operator

$$
\begin{aligned}
u(t) &= u''(t) - \pi(2\cos 2\pi t - \sin 2\pi t)\int_0^1 u'(x)\cos 2\pi x dx \\
&\quad - \sin 2\pi t\int_0^1 u''(x)\cos 2\pi x dx - \lambda^2 u(t) = f(t), \\
u(0) &= u(1), \quad u'(0) = u'(1), \quad u(t) \in H^2(0,1).
\end{aligned} \tag{86}
$$

Let $\widehat{A},\ \widehat{A}^2 : L_2(0,1) \to L_2(0,1)$ be the operators

$$\widehat{A}u = u'(t),\ \ D(\widehat{A}) = \{u(t) \in H^1(0,1) : u(0) = u(1)\}, \tag{87}$$

$$\widehat{A}^2 u = u''(t),\ \ D(\widehat{A}^2) = \{u(t) \in H^2(0,1) : u(0) = u(1),\ u'(0) = u'(1)\}. \tag{88}$$

Note that, from Problem 2 in the Appendix, $\lambda \in \rho(\widehat{A})$ if and only if $\lambda \neq 2k\pi i,\ k \in \mathbf{Z}$. Then $\lambda^2 \in \rho(\widehat{A}^2)$ iff $\pm\lambda \in \rho(\widehat{A})$ by Lemma 2. Consequently, $\lambda^2 \in \rho(\widehat{A}^2)$ if and only if $\lambda \neq \pm 2k\pi i,\ k \in \mathbf{Z}$. In addition, we define the operator $\mathbf{B} : L_2(0,1) \to L_2(0,1)$ as follows:

$$\begin{aligned}
\mathbf{B}u(t) &= u''(t) - \pi(2\cos 2\pi t - \sin 2\pi t)\int_0^1 u'(x)\cos 2\pi x dx \\
&\quad - \sin 2\pi t \int_0^1 u''(x)\cos 2\pi x dx \\
&= \widehat{A}^2 u(t) - S\langle \widehat{A}u(t), F^t\rangle_{H^m} - G\langle \widehat{A}^2 u(t), F^t\rangle_{H^m}, \quad D(\mathbf{B}) = D(\widehat{A}^2),
\end{aligned} \tag{89}$$

where $S = \pi(2\cos 2\pi t - \sin 2\pi t)$, $G = \sin 2\pi t$ and $F = \cos 2\pi t$. We observe that the components of the vector (S, G) are linearly independent and most important S and G satisfy (31). Therefore the operator $\mathbf{B}$ is quadratic and hence we apply Theorem 3. Accordingly, we rewrite the integro-differential equation (86) in the operator form

$$\mathbf{B}_{\lambda^2} u(t) = (\mathbf{B} - \lambda^2 I)u(t) = f(t), \quad D(\mathbf{B}_{\lambda^2}) = D(\widehat{A}^2). \tag{90}$$

Theorem 3 claims that the resolvent operator $R_{\lambda^2}(\mathbf{B})$ exists if and only if Eqs. (43) and (44) are satisfied, namely $\det L_\lambda \neq 0$ and $\det L_{-\lambda} \neq 0$. In computing these determinants we need the resolvent operators $R_\lambda(\widehat{A})$ and $R_{-\lambda}(\widehat{A})$ which are set out in (103) and (104) in the Appendix. By applying $R_\lambda(\widehat{A})$ on G, employing $\widehat{A}$ and taking the inner products, we obtain

$$\begin{aligned}
R_\lambda(\widehat{A})G &= \frac{1}{1-e^\lambda}\int_0^1 e^{\lambda(t-x+1)}\sin 2\pi x dx + \int_0^t e^{\lambda(t-x)}\sin 2\pi x dx \\
&= \frac{-\lambda \sin 2\pi t - 2\pi\cos 2\pi t}{\lambda^2 + 4\pi^2}, \\
\widehat{A}R_\lambda(\widehat{A})G &= \frac{-2\pi\lambda\cos 2\pi t + 4\pi^2 \sin 2\pi t}{\lambda^2 + 4\pi^2}, \\
\langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle &= \frac{-\pi\overline{\lambda}}{\overline{\lambda}^2 + 4\pi^2}, \\
L_\lambda &= \left[I_m - \langle F^t, \widehat{A}R_\lambda(\widehat{A})G\rangle\right] = \frac{\overline{\lambda}^2 + 4\pi^2 + \pi\overline{\lambda}}{\overline{\lambda}^2 + 4\pi^2}.
\end{aligned} \tag{91}$$

Repeating the same sequence of operations for $R_{-\lambda}(\widehat{A})$, we have

$$R_{-\lambda}(\widehat{A})G = \frac{1}{1-e^{-\lambda}}\int_0^1 e^{-\lambda(t-x+1)}\sin 2\pi x dx + \int_0^t e^{-\lambda(t-x)}\sin 2\pi x dx,$$

$$\widehat{A}R_{-\lambda}(\widehat{A})G = \frac{2\pi\lambda\cos 2\pi t + 4\pi^2\sin 2\pi t}{\lambda^2+4\pi^2},$$

$$\langle F^t, \widehat{A}R_{-\lambda}(\widehat{A})G\rangle = \frac{\pi\overline{\lambda}}{\overline{\lambda}^2+4\pi^2},$$

$$L_{-\lambda} = \left[I_m - \langle F^t, \widehat{A}R_{-\lambda}(\widehat{A})G\rangle\right] = \frac{\overline{\lambda}^2+4\pi^2-\pi\overline{\lambda}}{\overline{\lambda}^2+4\pi^2}. \tag{92}$$

It is evident that $\det L_\lambda \neq 0$ if and only if $\lambda \neq \pi(-1\pm i\sqrt{15})/2$ and $\det L_{-\lambda} \neq 0$ if and only if $\lambda \neq \pi(1\pm i\sqrt{15})/2$. From Theorem 3, $\lambda^2 \in \rho(\mathbf{B})$ if $\lambda \neq \pm 2k\pi i,\ k\in\mathbf{Z}$ and $\lambda \neq \pi(\pm 1 \pm i\sqrt{15})/2\}$ and the resolvent operator $R_{\lambda^2}(\mathbf{B})$ exists and has the representation as in (45).

By acting $\widehat{A}$ on Eq. (103), making use of Proposition 1 and applying the Fubini theorem, we get

$$\begin{aligned}
&\langle \widehat{A}R_\lambda(\widehat{A})f(t), F^t\rangle = \langle f + \lambda R_\lambda(\widehat{A})f(t), F^t\rangle \\
&\quad = \int_0^1 \cos 2\pi x f(x)dx \\
&\qquad + \int_0^1 \cos 2\pi t\left[\frac{\lambda}{1-e^\lambda}\int_0^1 e^{\lambda(t-x+1)}f(x)dx + \lambda\int_0^t e^{\lambda(t-x)}f(x)dx\right]dt \\
&\quad = \int_0^1 \cos 2\pi x f(x)dx + \frac{\lambda}{1-e^\lambda}\int_0^1 \cos 2\pi t e^{\lambda t}dt\int_0^1 e^{\lambda(-x+1)}f(x)dx \\
&\qquad + \lambda\int_0^1 e^{-\lambda x}f(x)dx\int_x^1 e^{\lambda t}\cos 2\pi t dt \\
&\quad = \int_0^1 \cos 2\pi x f(x)dx - \frac{\lambda^2}{\lambda^2+4\pi^2}\int_0^1 e^{\lambda(1-x)}f(x)dx \\
&\qquad + \frac{\lambda}{\lambda^2+4\pi^2}\int_0^1 [\lambda e^{\lambda(1-x)} - (\lambda\cos 2\pi x + 2\pi\sin 2\pi x)]f(x)dx \\
&\quad = -\frac{2\pi\lambda}{\lambda^2+4\pi^2}\int_0^1 \sin 2\pi x f(x)dx + \frac{4\pi^2}{\lambda^2+4\pi^2}\int_0^1 \cos 2\pi x f(x)dx \\
&\quad = \frac{2\pi}{\lambda^2+4\pi^2}\int_0^1 (2\pi\cos 2\pi x - \lambda\sin 2\pi x)f(x)dx.
\end{aligned} \tag{93}$$

By operating alike on (104), we acquire

$$\langle \widehat{A}R_{-\lambda}(\widehat{A})f(t), F^t\rangle = \frac{2\pi}{\lambda^2+4\pi^2}\int_0^1 (2\pi\cos 2\pi x + \lambda\sin 2\pi x)f(x)dx. \tag{94}$$

Substituting (103), (104), $\overline{L^{-1}_{\pm\lambda}}$ from (91) and (92), (93) and (94) into (45), we get

$$R_{\lambda^2}(\mathbf{B})f = \frac{1}{2\lambda}\left\{\frac{1}{1-e^{\lambda}}\int_0^1 \left[e^{\lambda(t-x+1)} + e^{-\lambda(t-x)}\right]f(x)dx \right.$$
$$+\int_0^t \left[e^{\lambda(t-x)} - e^{-\lambda(t-x)}\right]f(x)dx$$
$$-\frac{2\pi(\lambda\sin 2\pi t + 2\pi\cos 2\pi t)}{(\lambda^2+4\pi^2)(\lambda^2+4\pi^2+\pi\lambda)}\int_0^1 (2\pi\cos 2\pi x - \lambda\sin 2\pi x)f(x)dx$$
$$\left.-\frac{2\pi(\lambda\sin 2\pi t - 2\pi\cos 2\pi t)}{(\lambda^2+4\pi^2)(\lambda^2+4\pi^2-\pi\lambda)}\int_0^1 (2\pi\cos 2\pi x + \lambda\sin 2\pi x)f(x)dx\right\}. \tag{95}$$

The resolvent operator $R_{\lambda^2}(\mathbf{B})$ for every $\lambda^2 \in \rho(\mathbf{B})$ is defined on the whole space $L_2(0,1)$ and is bounded.

Appendix

Problem 1. Let the operator $\widehat{A} : L_2(0,1) \to L_2(0,1)$ be defined by

$$\widehat{A}u = iu' = f, \quad D(\widehat{A}) = \{u(t) \in H^1(0,1) : u(0) + u(1) = 0\} \tag{96}$$

Then $\hat{A}$ is closed and:

(i) $\lambda \in \rho(\widehat{A})$ if and only if $\lambda \neq (2k+1)\pi, \quad k \in \mathbf{Z}$, i.e.

$$\rho(\widehat{A}) = \{\lambda \in \mathbf{C} : \lambda \neq (2k+1)\pi, \quad k \in \mathbf{Z}\}. \tag{97}$$

(ii) For $\lambda \in \rho(\widehat{A})$ the resolvent operator $R_\lambda(\widehat{A})$ is bounded and defined on the whole space $L_2(0,1)$ by the formula

$$R_\lambda(\widehat{A})f(t) = i\int_0^1 e^{i\lambda(x-t)}\left[(e^{i\lambda}+1)^{-1} - \eta(t-x)\right]f(x)dx, \tag{98}$$

where

$$\eta(t-x) = \begin{cases} 1, & x \le t \\ 0, & x > t \end{cases} \text{ is the Heaviside's function.}$$

Problem 2. Let the operator $\widehat{A} : L_2(0,1) \to L_2(0,1)$ be defined by

$$\widehat{A}u = u' = f, \quad D(\widehat{A}) = \{u(t) \in H^1(0,1) : \ u(0) = u(1)\}. \tag{99}$$

Then the quadratic operator $\widehat{A}^2 : L_2(0,1) \to L_2(0,1)$ is closed and defined by

$$\widehat{A}^2 u = u'' = f, \quad D(\widehat{A}^2) = \{u(t) \in H^2(0,1) : u(0) = u(1), \, u'(0) = u'(1)\}, \tag{100}$$

the resolvent sets of $\widehat{A}$ and $\widehat{A}^2$ are

$$\rho(\widehat{A}) = \ \{\lambda \in \mathbf{C} : \lambda \neq 2k\pi i, \ k \in \mathbf{Z}\}, \tag{101}$$

$$\rho(\widehat{A}^2) = \{\lambda \in \mathbf{C} : \lambda \neq -4k^2\pi^2, \, k \in \mathbf{Z}\} \tag{102}$$

and the resolvent operators $R_{\pm\lambda}(\widehat{A}), \, R_\lambda(\widehat{A}^2)$ are bounded and defined on the whole space $L_2(0,1)$ by

$$R_\lambda(\widehat{A})f(t) = \frac{1}{1-e^\lambda}\int_0^1 e^{\lambda(t-x+1)}f(x)dx + \int_0^t e^{\lambda(t-x)}f(x)dx \tag{103}$$

$$R_{-\lambda}(\widehat{A})f(t) = \frac{1}{1-e^{-\lambda}}\int_0^1 e^{-\lambda(t-x+1)}f(x)dx + \int_0^t e^{-\lambda(t-x)}f(x)dx \tag{104}$$

$$\begin{aligned} R_\lambda(\widehat{A}^2)f(t) = {} & \frac{1}{2\sqrt{\lambda}(1-e^{\sqrt{\lambda}})}\int_0^1 \left[e^{\sqrt{\lambda}(t-x+1)} + e^{-\sqrt{\lambda}(t-x)}\right] f(x)dx \\ & + \frac{1}{2\sqrt{\lambda}}\int_0^t \left[e^{\sqrt{\lambda}(t-x)} - e^{-\sqrt{\lambda}(t-x)}\right] f(x)dx. \end{aligned} \tag{105}$$

References

1. Birman, M.S.: On the theory of self-adjoint extensions of positive definite operators. Mat. Sb. **38**, 431–450 (1956) (Russian)
2. Catchpole, E.A.: An integro-differential operator. J. London Math. Soc. **6**, 513–523 (1973)
3. Derkach, V.A., Malamud, M.M.: Generalized resolvents and the boundary value problems for Hermitian operators with gaps. J. Funct. Anal. **95**, 1–95 (1991)
4. Desch, W., Grimmer, R.C., Schappacher, W.: Some considerations for linear integrodifferential equations. J. Math. Anal. Appl. **104**, 219–234 (1984)
5. Gorbachuk, V.I., Gorbachuk, M.L.: Boundary Value Problems for Operator Differential Equations. Kluwer, Dordrecht (1991)
6. Grimmer, R.C.: Resolvent operators for integral equations in a Banach space. Trans. Am. Math. Soc. **273**, 333–349 (1982)
7. Grimmer, R.C., Pritchard, A.J.: Analytic resolvent operators for integral equations in a Banach space. J. Differ. Equ. **50**, 234–259 (1983)
8. He, M.: A class of integrodifferential equations and applications. Discret. Contin. Dyn. S. (2005). doi: 10.3934/proc.2005.2005.386

9. Hebbeche, A.: Generalized resolvents and spectrum for a certain class of perturbed symmetric operators. J. Appl. Math. Math. **1**, 81–92 (2005)
10. Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin (1980)
11. Kočubeĭ, A.N.: Extensions of symmetric operators and symmetric binary relations. Math. Notes. **17**, 25–28 (1975)
12. Kreĭn, M.G.: Theory of self-adjoint extensions of symmetric semi-bounded operators and applications. I. Mat. Sb. **20**, 431–495 (1947) (Russian)
13. Kreyszig, E.: Introductory Functional Analysis with Applications. Wiley, New York (1978)
14. Kurasov, P., Kuroda, S.T.: Krein's resolvent formula and perturbation theory. J. Operator Theory **51**, 321–334 (2004)
15. Lin, Y., Liu, J.H.: Semilinear integrodifferential equations with nonlocal Cauchy problem. Nonlinear Analysis THA **26**, 1023–1033 (1996)
16. Liu, J.H.: Resolvent operators and weak solutions of integrodifferential equations. Differ. Integral Equ. **7**, 523–534 (1994)
17. McCartin, B.J.: Rayleigh-Schrödinger Perturbation Theory: Pseudoinverse Formulation. Hikari Ltd, Bulgaria (2009)
18. Mikhailov, S.E.: Finite-dimensional perturbations of linear operators and some applications to boundary integral equations. Eng. Anal. Boundary Elem. **23**, 805–813 (1999)
19. von Neumann J.: Allgemeine eigenwerttheorie Hermitescher funktionaloperatoren. Math. Ann. **102**, 49–131 (1929)
20. Otelbaev, M.O., Mynbaev, K.T.: Weighting Functional Spaces and Differential Operator Spectrum. Nauka, Moscow (1988)
21. Parasidis, I.N., Tsekrekos, P.C.: Correct self-adjoint and positive extensions of nondensely defined minimal symmetric operators. Abstr. Appl. Anal. **7**, 767–790 (2005)
22. Parasidis, I.N., Tsekrekos, P.C.: Some quadratic correct extensions of minimal operators in Banach spaces. Oper. Matrices **4**, 225–243 (2010)
23. Vishik, M.I.: On general boundary value problems for elliptic differential equations. Tr. Moskv. Mat. Obšv. **1**, 187–246 (1952), translated in AMS Transl. **24**, 107–172 (1963)

Component Matrices of a Square Matrix and Their Properties

Dorothea Petraki and Nikolaos Samaras

Abstract The definition of the component matrices of a square matrix A is well-known [Lancaster]. This paper is concerned with all the basic properties of component matrices of a square matrix A, where $A \in M_{\nu\times\nu}(K), K = \mathbb{R}$, or $K = \mathbb{C}$. This is very useful for the studies of the spectral resolution of a matrix function $f(A)$, the convergence of sequences and series of matrices and also the convergence of matrix functions. It is also useful to solve differential equations and control system problems.

Keywords Component matrices • Jordan canonical form • Projection matrices • Resolvent of a matrix • Hermite interpolation

AMS Classification: 11Cxx, 15Axx, 65H04

Introduction

A brief history of matrix functions, as referred by Nicholas J. Higham, follows.

The British mathematician Arthur Cayley (1821–1895) is the first who began the study of functions of matrices. The French mathematician Edmond Laguerre (1834–1886) defined the exponential of a matrix A via its power series. The interpolating polynomial definition of a matrix function $f(A)$ was stated by the British mathematician James Joseph Sylvester (1814–1897) for matrices with distinct eigenvalues. In 1884 Artur Buchheim gave a derivation of Sylvester polynomial interpolation formulae for matrix function $f(A)$ and then generalized it to multiple eigenvalues using Hermite interpolation. The Austrian mathematician Emil Weyr (1848–1894) defined $f(A)$ using a power series of function f and studied the convergence of the defined series. The French mathematician Augustin-Louis Cauchy (1789–1857) defined $f(A)$ as the sum of the residues of $(\lambda I - A)^{-1}f(\lambda)$ at the eigenvalues of

D. Petraki (✉) • N. Samaras
Department of Applied Informatics, University of Macedonia, 156 Egnatia Str., 54006 Thessaloniki, Greece
e-mail: dpetraki@uom.gr; samaras@uom.gr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_25

A. It is well known that a large number of famous mathematicians such as A. Poincare, Giorgi, Cipolla, etc. also worked in the area of matrix functions. Another definition of $f(A)$ is given by the formulae $f(A) = \sum_{i=1}^{s} \sum_{j=0}^{d_i-1} f^{(j)}(\lambda_i)Z_{ij}$, where $\lambda_1, \lambda_2, \cdots, \lambda_s$ are the different eigenvalues and Z_{ij} are the component matrices of the matrix A.

Robert F. Rinehart in his 1955 paper "The Equivalence of Definitions of a Matrix Function" proved that all the above definitions are equivalents.

In this paper we present the definition of component matrices of a matrix A. Then we give a large number of their properties [4, 6, 8, 11].

Component Matrices

Let $A \in \mathbb{C}^{nxn}$ and $\lambda_1, \lambda_2, \cdots, \lambda_s$ be are different by two eigenvalues of matrix A. Let $\alpha_1, \alpha_2, \cdots, \alpha_s$ and $\gamma_1, \gamma_2, \cdots, \gamma_s$ be, respectively, the algebraic and geometric multiplicities of the above eigenvalues. Then the minimal polynomial of the matrix A is $q(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\cdots(x-\lambda_s)^{d_s}$, where $d_i = \alpha_i - \gamma_i + 1, i = 1, 2, \cdots, s$ are the indices of the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_s$ [9, 10, 12, 13].

Definition of the Component Matrices

Case 1.

The matrix $A \in \mathbb{C}^{nxn}$ is simple. Then $\alpha_i = \gamma_i$, so $d_i = 1$, for $i = 1, 2, \cdots, s$. The Jordan Form of the matrix A is $J_A = diag[\lambda_1 I_{\alpha_1}, \cdots, \lambda_{i-1} I_{\alpha_{i-1}} \lambda_i, I_{\alpha_i}, \lambda_{i+1} I_{\alpha_{i+1}}, \cdots, \lambda_s I_{\alpha_s}]$. Let φ_{i0}, $i = 1, 2, \cdots, s$, be the interpolation polynomials with $\varphi_{i0}(\lambda_k) = \delta_{ik}$, for $i, k = 1, 2, \cdots, s$ and δ_{ik} is the Kronecker Delta. Then it is

$A = P \cdot J_A \cdot P^{-1} \Rightarrow$ ($P \in \mathbb{C}^{nxn}$ is the transition matrix of A)
$\phi_{i0}(A) = P \cdot \phi_{i0}(J_A).\ P^{-1} \Rightarrow$
$\phi_{i0}(A) = P \cdot diag[\phi_{i0}(\lambda_1 I_{\alpha_1}), \cdots, \phi_{i0}(\lambda_{i-1} I_{\alpha_{i-1}}), \phi_{i0}(\lambda_i I_{\alpha_i}), \phi_{i0}(\lambda_{i+1} I_{\alpha_{i+1}}), \cdots, \phi_{i0}(\lambda_s I_{\alpha_s})] P^{-1} \Rightarrow$
$\phi_{i0}(A) = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s})] P^{-1}, i = 1, 2, ..., s$.
Let $Z_{i0} = \varphi_{i0}(A)$, then
$Z_{i0} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_{s-1}}, O_{\alpha_s})] P^{-1}, i = 1, 2, \cdots, s.$

Case 2.

The matrix $A \in \mathbb{C}^{nxn}$ is derogatory. Then there exists $d_i > 1, i = 1, 2, \cdots, s$.
The Jordan Form of matrix A is not diagonal and has the form

$J_A = diag[J_1, \cdots, J_{i-1}, J_i, J_{i+1}, \cdots, J_s]$, where $J_i = diag[M_{d_i}, \lambda_i I_{\gamma_i - 1}]$

$$\text{with } M_{d_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & \lambda i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}, \text{ matrix } d_i \times d_i,$$

for $i, k = 1, 2, \cdots, d_i - 1$ [1–3, 5, 7–9].

Let φ_{ij}, $i = 1, 2, \cdots, s$, and $j = 0, 1, \cdots, d_i - 1$ be the interpolation polynomials with $\varphi_{ij}^{(r)}(\lambda_k) = \delta_{ik}\delta_{jr}$, for $i, k = 1, 2, \cdots, s$ and $j, r = 0, 1, 2, \cdots, d_i - 1$.

For $i = 1, 2, \cdots, s$ it is $\varphi_{i0}(J_i) = I_{\alpha_i}$ and $\varphi_{i0}(J_k) = O_{\alpha_k}$ for $i \neq k$.

Applying the polynomial function φ_{i0} on to the matrix A, we have

$A = P \cdot J_A \cdot P^{-1} \Rightarrow$ ($P \epsilon \mathbb{C}^{nxn}$ is the transition matrix of A)
$\phi_{i0}(A) = P \cdot \phi_{i0} \cdot (J_A)P^{-1} \Rightarrow$
$\phi_{i0}(A) = P \cdot diag[\phi_{i0}(J_1), \cdots, \phi_{i0}(J_{i-1}), \phi_{i0}(J_i), \phi_{i0}(J_{i+1}), \cdots, \phi_{i0}(J_s)] \cdot P^{-1} \Rightarrow$
$\phi_{i0}(A) = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s})] \cdot P^{-1}$, $i = 1, 2, \cdots, s$.
Let $Z_{i0} = \varphi_{i0}(A)$, then
$Z_{i0\cdot} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_{s-1}}, O_{\alpha_s})]P^{-1}$, $i = 1, 2, \cdots, s$.
For $i = 1, 2, \cdots, s$, and $j = 1, \cdots, d_i - 1$ it is
$\varphi_{ij}(\lambda_i I_{\gamma_i - 1}) = diag[\varphi_{ij}(\lambda_i), \varphi_{ij}(\lambda_i), \cdots, \varphi_{ij}(\lambda_i)] \Rightarrow$
$\varphi_{ij}(\lambda_i I_{\gamma_i - 1}) = diag[\delta_{j0}, \delta_{j0}, \cdots, \delta_{j0}] \Rightarrow$
$\varphi_{ij}(\lambda_i I_{\gamma_i - 1}) = diag[0, 0, \cdots, 0] = O_{\gamma_i - 1}$.

Also it is

$$\varphi_{ij}(M_{d_i}) = \begin{bmatrix} \varphi_{ij}(\lambda_i) & \frac{\varphi_{ij}^{(1)}(\lambda_i)}{1!} & \frac{\varphi_{ij}^{(2)}(\lambda_i)}{2!} & \cdots & \frac{\varphi_{ij}^{(d_i-2)}(\lambda_i)}{(d_i-2)!} & \frac{\varphi_{ij}^{(d_i-1)}(\lambda_i)}{(d_i-1)!} \\ 0 & \varphi_{ij}(\lambda_i) & \frac{\varphi_{ij}^{(1)}(\lambda_i)}{1!} & \cdots & \frac{\varphi_{ij}^{(d_i-3)}(\lambda_i)}{(d_i-3)!} & \frac{\varphi_{ij}^{(d_i-2)}(\lambda_i)}{(d_i-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \varphi_{ij}(\lambda_i) & \frac{\varphi_{ij}^{(1)}(\lambda_i)}{1!} \\ 0 & 0 & 0 & \cdots & 0 & \varphi_{ij}(\lambda_i) \end{bmatrix} \Rightarrow$$

$$\varphi_{ij}(M_{d_i}) = \begin{bmatrix} \delta_{j0} & \frac{\delta_{j1}}{1!} & \frac{\delta_{j2}}{2!} & \cdots & \frac{\delta_{j,d_i-2}}{(d_i-2)!} & \frac{\delta_{j,d_i-1}}{(d_i-1)!} \\ 0 & \delta_{j0} & \frac{\delta_{j1}}{1!} & \cdots & \frac{\delta_{j,d_i-3}}{(d_i-3)!} & \frac{\delta_{j,d_i-2}}{(d_i-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \delta_{j0} & \frac{\delta_{j1}}{1!} \\ 0 & 0 & 0 & \cdots & 0 & \delta_{j0} \end{bmatrix} \Rightarrow$$

$$\varphi_{ij}(M_{d_i}) = \begin{bmatrix} 0 & 0 & \cdots & \frac{1}{j!} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & \frac{1}{j!} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \frac{1}{j!} & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{1}{j!} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} = [b_{mr}],$$

With $b_{m,j+m} = \frac{1}{j!}$, for $m = 1, 2, \cdots, d_i - j$ and $b_{mr} = 0$ for the other cases.
Let N_{d_i} be the $d_i \times d_i$ matrix with

$$N_{d_i} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0\cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, i = 1, 2, \cdots, s,$$

then $\varphi_{ij}(M_{d_i}) = \frac{1}{j!}N_{d_i}^j$ and finally $\varphi_{ij}(J_i) = diag[\frac{1}{j!}N_{d_i}^j, O_{\gamma_i - 1}] = \frac{1}{j!}diag[N_{d_i}^j, O_{\gamma_i - 1}]$.
It is obvious that $\varphi_{ij}(J_k) = O_{\alpha_k}$, for $i \neq k$.
Applying the polynomial function φ_{ij} on the matrix A, we have

$$A = P \cdot J_A \cdot P^{-1} \Rightarrow$$

$$\phi_{ij}(A) = P \cdot \phi_{ij}(J_A) \cdot P^{-1} \Rightarrow$$

$$\phi_{ij}(A) = P \cdot diag[\phi_{ij}(J_1), \cdots, \phi_{ij}(J_{i-1}), \phi_{ij}(J_i), \phi_{ij}(J_{i+1}), \cdots, \phi_{ij}(J_s)] \cdot P^{-1} \Rightarrow$$

$$\phi_{ij}(A) = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, \frac{1}{j!}N_{d_i}^j, O_{\gamma_i - 1}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1} \Rightarrow$$

$$\phi_{ij}(A) = \frac{1}{j!}P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, N_{d_i}^j, O_{\gamma_i - 1}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}.$$

Let $Z_{ij} = \varphi_{ij}(A)$, for $i = 1, 2, \cdots, s, j = 1, 2, \cdots, d_i - 1$, then

$$\phi_{ij}(A) = \frac{1}{j!}P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, N_{d_i}^j, O_{\gamma_i - 1}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}.$$

Illustrative Examples

Example

We will calculate the component matrices of the square matrix $A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 0 & -1 & 0 & 5 \end{bmatrix}$.

The characteristic polynomial of matrix A is
$p(x) = (x+1)(x-1)(x-2)(x-5)$.

The eigenvalues of the matrix A are $\lambda_1 = 5, \lambda_2 = 2, \lambda_3 = -1, \lambda_4 = 1$ with algebraic multiplicities $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, and indices
$d_1 = d_2 = d_3 = d_4 = 1$, respectively.

So the minimal polynomial of the matrix A is
$q(x) = (x+1)(x-1)(x-2)(x-5)$.

(i) The interpolating polynomial $\varphi_{10}(x)$ with $\varphi_{10}(\lambda_1) = 1, \varphi_{10}(\lambda_i) = 0$, for $i = 2, 3, 4$ is $\varphi_{10}(x) = \frac{x^3}{72} - \frac{x^2}{36} - \frac{x}{72} + \frac{1}{36}$.

So the corresponding component matrix is

$$Z_{10} = \varphi_{10}(A) = \begin{bmatrix} 0 & -\frac{1}{24} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{24} & 0 & \frac{1}{4} \\ 0 & -\frac{1}{6} & 0 & 1 \end{bmatrix}.$$

(ii) The interpolating polynomial $\varphi_{20}(x)$ with $\varphi_{20}(\lambda_2) = 1, \varphi_{20}(\lambda_i) = 0$, for $i = 1, 3, 4$ is $\varphi_{20}(x) = -\frac{x^3}{9} + \frac{5x^2}{9} + \frac{x}{9} - \frac{5}{9}$.

So the corresponding component matrix is

$$Z_{20} = \varphi_{20}(A) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & \frac{1}{3} & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(iii) The interpolating polynomial $\varphi_{30}(x)$ with $\varphi_{30}(\lambda_3) = 1, \varphi_{30}(\lambda_i) = 0$, for $i = 1, 2, 4$ is $\varphi_{30}(x) = -\frac{x^3}{36} + \frac{2x^2}{9} - \frac{17x}{36} + \frac{5}{18}$.

So the corresponding component matrix is

$$Z_{30} = \varphi_{30}(A) = \begin{bmatrix} 0 & -\frac{1}{12} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & \frac{1}{12} & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 \end{bmatrix}.$$

(iv) The interpolating polynomial $\varphi_{40}(x)$ with $\varphi_{40}(\lambda_4) = 1, \varphi_{40}(\lambda_i) = 0$, for $i = 1, 2, 3$ is $\varphi_{40}(x) = -\frac{x^3}{8} - \frac{3x^2}{4} + \frac{3x}{8} + \frac{5}{4}$.

So the corresponding component matrix is

$$Z_{40} = \varphi_{40}(A) = \begin{bmatrix} 1 & \frac{1}{8} & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 \\ -3 & -\frac{3}{8} & 0 & \frac{3}{4} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example

We will calculate the component matrices of the square matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \imath & -1 & -\imath \\ 1 & -1 & 1 & -1 \\ 1 & -\imath & -1 & \imath \end{bmatrix}.$$

The characteristic polynomial of matrix A is

$$p(x) = (x-2)^2(x-2\imath)(x+2).$$

The eigenvalues of the matrix A are $\lambda_1 = -2, \lambda_2 = 2\imath, \lambda_3 = 2$, with algebraic multiplicities $\alpha_1 = \alpha_2 = 1, \alpha_3 = 2$, and indices $d_1 = d_2 = d_3 = 1$, respectively.

So the minimal polynomial of the matrix A is $q(x) = (x-2)(x-2\imath)(x+2)$.

(i) The interpolating polynomial $\varphi_{10}(x)$ with $\varphi_{10}(\lambda_1) = 1, \varphi_{10}(\lambda_i) = 0$, for $i = 2, 3$ is

$$\varphi_{10}(x) = (\frac{1}{16} - \frac{\imath}{16})x^2 - \frac{x}{4} + (\frac{1}{4} + \frac{\imath}{4}).$$

So the corresponding component matrix is

$$Z_{10} = \varphi_{10}(A) = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

(ii) The interpolating polynomial $\varphi_{20}(x)$ with $\varphi_{20}(\lambda_2) = 1, \varphi_{20}(\lambda_i) = 0$, for $i = 1, 3$ is $\varphi_{20}(x) = -\frac{x^2}{8} + \frac{1}{2}$.

So the corresponding component matrix is

$$Z_{20} = \varphi_{20}(A) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

(*iii*) The interpolating polynomial $\varphi_{30}(x)$ with $\varphi_{30}(\lambda_3) = 1, \varphi_{30}(\lambda_i) = 0$, for $i = 1, 2$ is $\varphi_{30}(x) = (\frac{1}{16} + \frac{\imath}{16})x^2 + \frac{x}{4} + (\frac{1}{4} - \frac{\imath}{4})$.
So the corresponding component matrix is

$$Z_{30} = \varphi_{30}(A) = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Example

We will calculate the component matrices of the square matrix

$$A = \begin{bmatrix} 0 & 4 & 1 & -2 \\ -1 & 4 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ -1 & 3 & 0 & 0 \end{bmatrix}.$$

The characteristic polynomial of matrix A is $p(x) = (x-2)(x-1)^3$.

The eigenvalues of the matrix A are $\lambda_1 = 2, \lambda_2 = 1$ with algebraic multiplicities $\alpha_1 = 1, \alpha_2 = 3$, and indices $d_1 = 1, d_2 = 2$ respectively.

So the minimal polynomial of the matrix A is $q(x) = (x-2)(x-1)^2$.

(*i*) The interpolating polynomial $\varphi_{10}(x)$ with $\varphi_{10}(\lambda_1) = 1, \varphi_{10}(\lambda_2) = 0$ and $\varphi_{10}^{(1)}(\lambda_2) = 0$ is $\varphi_{10}(x) = x^2 - 2x + 1$.
So the corresponding component matrix is

$$Z_{10} = \varphi_{10}(A) = \begin{bmatrix} -1 & 2 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 \end{bmatrix}.$$

(*ii*) The interpolating polynomial $\varphi_{20}(x)$ with $\varphi_{20}(\lambda_2) = 1, \varphi_{20}^{(1)}(\lambda_2) = 0$ is $\varphi_{20}(x) = -x^2 + 2x$.
So the first corresponding component matrix is the matrix:

$$Z_{20} = \varphi_{20}(A) = \begin{bmatrix} 2 & -2 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -2 & 1 & 1 \end{bmatrix}.$$

The interpolating polynomial $\varphi_{21}(x)$ with $\varphi_{21}(\lambda_1) = 0, \varphi_{21}(\lambda_2) = 0$ and $\varphi_{21}^{(1)}(\lambda_2) = 1$ is $\varphi_{21}(x) = -x^2 + 3x - 2$.

So the second corresponding component is the matrix:

$$Z_{21} = \varphi_{21}(A) = \begin{bmatrix} 0 & 2 & 2 & -2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}.$$

Basic Properties

In this section we describe the basic properties related to the component Z_{ij}, $i = 1, 2, \cdots, s, j = 0, 1, 2, \cdots, d_i - 1$ of a square matrix A [8].

Property 1. $Z_{i0}Z_{ij} = Z_{ij}$, for $i = 1, 2, \cdots, s, j = 1, 2, \cdots, d_i - 1$.

Proof. The proof is obvious from the definition of the matrices Z_{i0}, Z_{ij}.

Property 2. $Z_{i0}Z_{jk} = O_n$, for every $i \neq j$.

Proof. The proof is obvious from the definition of the matrices Z_{i0}, Z_{jk}.

Property 3. It is $\sum_{i=1}^{s} Z_{i0} = I_n$.

Proof. It is $Z_{i0}=P \cdot diag[O_{\alpha_1}, \cdots, I_{\alpha_i}, \cdots, O_{\alpha_s}] \cdot P^{-1}$, $i=1, 2, \cdots, s$ so $\sum_{i=1}^{s} Z_{i0}=P \cdot \sum_{i=1}^{s} diag[O_{\alpha_1}, \cdots, I_{\alpha_i}, \cdots, O_{\alpha_s}] \cdot P^{-1} = P \cdot I_n \cdot P^{-1} = I_n$.

Property 4. It is $Z_{i0}Z_{j0} = \delta_{ij}Z_{i0}$, for every $i, j = 1, 2, \cdots, s$.

Proof. It is $Z_{i0} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}$ and
$Z_{j0} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{j-1}}, I_{\alpha_j}, O_{\alpha_{j+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}$.
So, if $i \neq j$, then
$Z_{i0}Z_{j0} = P \cdot diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, O_{\alpha_i}, \cdots, O_{\alpha_j}, \cdots, O_{\alpha_s}] \cdot P^{-1} = O_n$.
If $i = j$, then
$Z_{i0}Z_{i0} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1} = Z_{i0}$.

Property 5. (i) It is $Z_{i0}^r = Z_{i0}$, for $r = 1, 2, 3, \cdots$.
(ii) The matrix Z_{i0} is a projection matrix.
(iii) $Range(Z_{i0}) = Kernel(I_n - Z_{i0})$ and $Kernel(Z_{i0}) = Range(I_n - Z_{i0})$.
(iv) $Range(Z_{i0}) = Im(Z_{i0})$ and $Range(I_n - Z_{i0}) = Im(I_n - Z_{i0})$.
(v) $K^n = Range(Z_{i0}) \oplus Kernel(Z_{i0})$, for $i = 1, 2, \cdots, s$.

Proof. (i) It is

$$Z_{i0} = P \cdot diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_j}, \cdots, O_{\alpha_s}] \cdot P^{-1} \Rightarrow$$

$$Z_{i0}^r = P \cdot diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, I_{\alpha_i}^r, O_{\alpha_{i+1}}, \cdots, O_{\alpha_j}, \cdots, O_{\alpha_s}] \cdot P^{-1} \Rightarrow$$

$$Z_{i0}^r = P \cdot diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_j}, \cdots, O_{\alpha_s}] \cdot P^{-1} Z_{i0}.$$

(ii) From the relation $Z_{i0}^2 = Z_{i0}$, we conclude that the matrix Z_{i0} is a projection matrix.

(iii) It is $Range(Z_{i0}) = \{x \in K^n : Z_{i0}x = x\}$, and
$Kernel(I_n - Z_{i0}) = x \in K^n : (I_n - Z_{i0})x = O$ so,
$Range(Z_{i0}) = Kernel(I_n - Z_{i0})$.
It is also $Kernel(Z_{i0}) = \{x \in K^n : Z_{i0}x = O\}$ and $Range(I_n - Z_{i0}) = \{x \in K^n : (I_n - Z_{i0})x = x\}$ so $Kernel(Z_{i0}) = Range(I_n - Z_{i0})$.

(iv) It is obvious that $Range(Z_{i0}) \subseteq Im(Z_{i0})$. Let $y \in Im(Z_{i0})$, then there exist $x \in K^n$: $y = Z_{i0}x \Rightarrow Z_{i0}y = Z_{i0}^2x = Z_{i0}x = y \Rightarrow Z_{i0}y = y \Rightarrow y \in Range(Z_{i0}) \Rightarrow Im(Z_{i0}) \subseteq Range(Z_{i0})$, so finally $Range(Z_{i0}) = Im(Z_{i0})$. The matrix $I_n - Z_{i0}$ is also a projection matrix so $Range(I_n - Z_{i0}) = Im(I_n - Z_{i0})$.

(v) It is obvious that $Range(Z_{i0}) \cap KernelZ_{i0}) = \{O\}$.
Let $x \in K^n$, then $x = (I_n - Z_{i0})x + Z_{i0}x$, with
$Z_{i0}x \in Range(Z_{i0})$ and $(I_n - Z_{i0})x \in Kernel(Z_{i0})$, so
$K^n = Range(Z_{i0}) \oplus Kernel(Z_{i0})$.

Property 6. $Z_{i,j+1} = \frac{1}{j+1}(A - \lambda_i I_n)Z_{ij}$, $i = 1, 2, \cdots, s, j = 1, 2, \cdots, d_i - 1$.

Proof. It is $J_A - \lambda_i I_n = P^{-1}(A - \lambda_i I_n)P \Rightarrow J_A - \lambda_i I_n = diag[J_{d_1} - \lambda_i I_{d_1}, \cdots, N_{d_i}, O_{\gamma_i - 1}, \cdots, J_{d_s} - \lambda_i I_s] \Rightarrow P^{-1}(A - \lambda_i I_n)P = diag[J_{d_1} - \lambda_i I_{d_1}, \cdots, N_{d_i}, O_{\gamma_i-1}, \cdots, J_{d_s} - \lambda_i I_s] \Rightarrow P^{-1}(A - \lambda_i I_n)Z_{ij}P = diag[J_{d_1} - \lambda_i I_{d_1}, \cdots, N_{d_i}, O_{\gamma_i-1}, \cdots, J_{d_s} - \lambda_i I_s]P^{-1}Z_{ij}P \Rightarrow P^{-1}(A - \lambda_i I_n)Z_{ij}P = \frac{1}{j!}diag[O_{a_1}, \cdots, N_{d_i}^{j+1}, O_{\gamma_i-1}, \cdots, O_{a_s}] \Rightarrow$ (because $P^{-1}Z_{ij}P = \frac{1}{j!}diag[O_{a_1}, \cdots, O_{a_s-1}, N_{d_i}^j, O_{\gamma_i-1}, O_{a_{i+1}} \cdots, O_{a_s}]$) $(A - \lambda_i I_n)Z_{ij} = \frac{1}{j!}P \cdot diag[O_{a_1}, \cdots, N_{d_i}^{j+1}, O_{\gamma_i-1}, \cdots, O_{a_s}] \cdot P^{-1} \Rightarrow (A - \lambda_i I_n)Z_{ij} = (j+1)Z_{i,j+1} \Rightarrow Z_{i,j+1} = \frac{1}{j+1}(A - \lambda_i I_n)Z_{ij}$.

Property 7. $Z_{ij} = \frac{1}{j!}(A - \lambda_i I_n)^j Z_{i0}$, $i = 1, 2, \cdots, s, j = 1, 2, \cdots, d_i - 1$.

Proof. It is obvious conclusion from Property 6.

Property 8. $Z_{ij}Z_{ik} = \binom{j+k}{j}Z_{i,j+k}$.

Proof. It is $Z_{ij}Z_{ik} = \frac{1}{j!}(A - \lambda_i I_n)^j Z_{i0}\frac{1}{k!}(A - \lambda_i I_n)^k Z_{i0} \Rightarrow$
$Z_{ij}Z_{ik} = \frac{(j+k)!}{j!k!}\frac{1}{(j+k)!}(A - \lambda_i I_n)^{j+k}Z_{i0} \Rightarrow$
$Z_{ij}Z_{ik} = \binom{j+k}{j}Z_{i,j+k}$.

Corollary

$$Z_{ij_1}Z_{ij_2}\cdots Z_{ij_k} = \binom{j_1 + j_2 + \cdots + j_k}{j_1, j_2, \cdots, j_k}Z_{i,j_1+j_2+\cdots+j_k}.$$

Property 9. $Z_{ij}^r = \frac{(j \cdot r)!}{(j!)^r}Z_{i,j \cdot r}$, for $j, r = 1, 2, 3, \cdots, d_i - 1$ and $Z_{ij}^r = O_n$ for $j \cdot r = d_i, d_i + 1, \cdots$.

Proof. For $j, r = 1, 2, 3, \cdots, d_i - 1$, it is $Z_{ij} = \frac{1}{j!}(A - \lambda_i I_n)^j Z_{i0}$ so
$Z_{ij}^r = \frac{1}{(j!)^r}(A - \lambda_i I_n)^{j \cdot r}Z_{i0} \Rightarrow Z_{ij}^r = \frac{(j \cdot r)!}{(j)!} \cdot \frac{1}{(j \cdot r)!}(A - \lambda_i I_n)^{j \cdot r}Z_{i0} \Rightarrow$
$Z_{ij}^r = \frac{(j \cdot r)!}{(j!)^r}Z_{i,j \cdot r}$.
For $j \cdot r = d_i, d_i + 1, \cdots$ it is $Z_{i,j \cdot r} = O_n$ so and $Z_{ij}^r = O_n$.

Property 10. $Z_{ij}^r \neq Z_{ij}$, for $r = 2, 3, \cdots, d_i - 1$.

Proof. It is $Z_{ij} = \frac{1}{j!} P diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, O_{\alpha_{i-1}}, N_{d_i}^j, O_{\gamma_i - 1}, O_{\alpha_{i+1}}, \cdots,$
$O_{\alpha_{s-1}}, O_{\alpha_s}] \cdot P^{-1} \Rightarrow$
$Z_{ij}^r = (\frac{1}{j!})^r P diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, O_{\alpha_{i-1}}, (N_{d_i}^j)^r, O_{\gamma_i - 1}, O_{\alpha_{i+1}}, \cdots,$
$O_{\alpha_{s-1}}, O_{\alpha_s}] \cdot P^{-1} \neq Z_{ij}$.

Property 11. The m component matrices Z_{ij} are linear independent.

Proof.
Case 1. The matrix A is simple.
In this case the components of the matrix A are defined by
$Z_{i0} = P \cdot diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_{s-1}}, O_{\alpha_s}] \cdot P^{-1}$,
$i = 1, 2, \cdots, s$.
So, $\sum_{i=1}^s \lambda_i Z_{i0} = O_n \Leftrightarrow diag[\lambda_1 I_{a_1}, \cdots, \lambda_i I_{a_i}, \cdots, \lambda_s I_{a_s}] = O_n \Leftrightarrow$
$\lambda i = 0$, for every $i = 1, 2, \cdots, s$.

Proof. **Case 2.** The matrix A is derogatory.
It is, $\sum_{i=1}^s \sum_{j=0}^{d_i-1} \lambda_{ij} Z_{ij} = O_n \Leftrightarrow$
$\sum_{i=1}^s (\lambda_{i0} Z_{i0} + \lambda_{i1} Z_{i1} + \cdots + \lambda_{id_i-1} Z_{id_i-1}) = O_n \Leftrightarrow$
$\sum_{i=1}^s diag[O_{\alpha_1}, \cdots, \lambda_{i0} I_{\alpha_i} + \frac{\lambda_{i1}}{1!} N_{d_i}^1 + \cdots + \frac{\lambda_{i,d_i-1}}{(d_i-1)!} N_{d_i}^{d_i-1}, \cdots, O_{\alpha_s}] = O_n \Leftrightarrow$
$\lambda_{i0} I_{\alpha_i} + \frac{\lambda_{i1}}{1!} N_{d_i}^1 + \cdots + \frac{\lambda_{i,d_i-1}}{(d_i-1)!} N_{d_i}^{d_i-1} = O_n$, for every $i = 1, 2, \cdots, s \Leftrightarrow$
$\lambda_{ij} = 0$, for every $i = 1, 2, \cdots, s$ and $j = 0, 1, 2, \cdots, d_i - 1$.

Property 12. If the matrix A is simple, then

(*i*) $A = \sum_{i=1}^s \lambda_i Z_{i0}$.
(*ii*) $\lambda I_n - A = \sum_{i=1}^s (\lambda - \lambda_i) Z_{i0}$.
(*iii*) If $\lambda \in \mathbb{C} - \sigma(A)$, then the matrix $\lambda I_n - A$ is invertible with $(\lambda I_n - A)^{-1} = \sum_{i=1}^s \frac{1}{(\lambda - \lambda_i)} Z_{i0}$ where $\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda s\}$ is the spectrum of matrix A.

Proof. (*i*) It is $\sum_{i=1}^s \lambda_i Z_{i0} = P \cdot diag[\lambda_1 I_{\alpha_1}, \lambda_2 I_{\alpha_2}, \cdots, \lambda_i I_{\alpha_i}, \cdots, \lambda_s I_{\alpha_s}] \cdot P^{-1} = P \cdot J_A \cdot P^{-1} = A$.
(*ii*) Also it is $\lambda I_n - A = \lambda \sum_{i=1}^s Z_{i0} - \sum_{i=1}^s \lambda_i Z_{i0} = \sum_{i=1}^s (\lambda - \lambda_i) Z_{i0}$.
(*iii*) Finally $(\lambda I_n - A)(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}) = (\sum_{j=1}^s (\lambda - \lambda_j) Z_{j0})(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{j0}) \Rightarrow$
$(\lambda I_n - A)(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}) = (\sum_{i,j=1}^s (\lambda - \lambda_j) \frac{1}{\lambda - \lambda_i} Z_{j0} Z_{i0} \Rightarrow$
$(\lambda I_n - A)(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}) = (\sum_{i,j=1}^s (\lambda - \lambda_j) \frac{1}{\lambda - \lambda_i} \delta_{ij} Z_{i0} \Rightarrow$
$(\lambda I_n - A)(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}) = (\sum_{i=1}^s (\lambda - \lambda_i) \frac{1}{\lambda - \lambda_i} Z_{i0} \Rightarrow$
$(\lambda I_n - A)(\sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}) = \sum_{i=1}^s Z_{i0} = I_n$, so
$(\lambda I_n - A)^{-1} = \sum_{i=1}^s \frac{1}{\lambda - \lambda_i} Z_{i0}$.

Property 13. If the matrix A is derogatory, then

(*i*) $A = \sum_{i=1}^s \lambda_i Z_{i0} + \sum_{i=1}^s Z_{i1}$.
(*ii*) $\lambda I_n - A = \sum_{i=1}^s (\lambda - \lambda_i) Z_{i0} - \sum_{i=1}^s Z_{i1}$.
(*iii*) If $\lambda \in \mathbb{C} - \sigma(A)$ then the matrix $\lambda I_n - A$ is invertible with $(\lambda I_n - A)^{-1} = \sum_{i=1}^s (\sum_{j=0}^{d_i-1} \frac{j!}{(\lambda - \lambda i)^{j+1}} Z_{ij})$.

Proof. (*i*) It is $\sum_{i=1}^{s}\lambda_i Z_{i0}+\sum_{i=1}^{s}=P\cdot diag[\lambda_1 I_{d_1}+N_{d_i},\lambda_1 I_{\gamma_1-1},\cdots,\lambda_i I_{d_i}+N_{d_i},\lambda_i I_{\gamma_i-1},\cdots,\lambda_s I_{d_s}+N_{d_s},\lambda_s I_{\gamma_s-1}]\cdot P^{-1}=P\cdot J_A\cdot P^{-1}=A$.

(*ii*) Also it is $\lambda I_n-A=\lambda\sum_{i=1}^{s}Z_{i0}-\sum_{i=1}^{s}\lambda_i Z_{i0}-\sum_{i=1}^{s}Z_{i1}=\sum_{i=1}^{s}(\lambda-\lambda_i)Z_{i0}-\sum_{i=1}^{s}Z_{i1}$.

(*iii*) Finally if $B=(\lambda I_n-A)\sum_{j=1}^{s}(\sum_{k=0}^{d_j-1}\frac{k!}{(\lambda-\lambda_j)^{k+1}}Z_{jk})$, then it is $B=(\sum_{i=1}^{s}(\lambda-\lambda_i)Z_{i0}-\sum_{i=1}^{s}Z_{i1})(\sum_{j=1}^{s}(\sum_{k=0}^{d_j-1}\frac{k!}{(\lambda-\lambda_j)^{k+1}}Z_{jk}))\Rightarrow$

$B=(\sum_{i=1}^{s}(\lambda-\lambda_i)Z_{i0})(\sum_{j=1}^{s}(\sum_{k=0}^{d_j-1}\frac{k!}{(\lambda-\lambda_j)^{k+1}}Z_{jk}))-(\sum_{i=1}^{s}Z_{i1})(\sum_{j=1}^{s}(\sum_{k=0}^{d_j-1}\frac{k!}{(\lambda-\lambda_j)^{k+1}}Z_{jk}))\Rightarrow$

$B=\sum_{i=1}^{s}(\sum_{j=0}^{d_i-1}\frac{j!}{(\lambda-\lambda_i)^j}Z_{ij})-\sum_{i=1}^{s}(\sum_{j=0}^{d_i-1}\frac{j!}{(\lambda-\lambda_j)^{j+1}}Z_{i1}Z_{ij})\Rightarrow$

$B=\sum_{i=1}^{s}(\sum_{j=0}^{d_i-1}\frac{j!}{(\lambda-\lambda_i)^j}Z_{ij})-\sum_{i=1}^{s}(\sum_{j=0}^{d_i-1}\frac{(j+1)!}{(\lambda-\lambda_j)^{j+1}}Z_{i,j+1})\Rightarrow$

$B=\sum_{i=1}^{s}Z_{i0}+\sum_{i=1}^{s}(\sum_{j=1}^{d_i-1}\frac{j!}{(\lambda-\lambda_i)^j}Z_{ij})-\sum_{i=1}^{s}(\sum_{j=1}^{d_i}\frac{j!}{(\lambda-\lambda_j)^j}Z_{ij})\Rightarrow$

$B=\sum_{i=1}^{s}Z_{i0}+\sum_{i=1}^{s}(\sum_{j=1}^{d_i-1}\frac{j!}{(\lambda-\lambda_i)^j}Z_{ij})-\sum_{i=1}^{s}(\sum_{j=1}^{d_{i-1}}\frac{j!}{(\lambda-\lambda_j)^j}Z_{ij})\Rightarrow$ (because $Z_{i,d_i}=O_n$)

$B=\sum_{i=1}^{s}Z_{i0}=I_n$, so

$(\lambda I_n-A)^{-1}=\sum_{i=1}^{s}(\sum_{j=0}^{d_i-1}\frac{j!}{(\lambda-\lambda_i)^{j+1}}Z_{ij})$.

Property 14. If the matrix A is invertible with spectrum $\sigma(A)=\{\lambda_1,\lambda_2,\cdots,\lambda_s\}$ then $A^{-1}=\sum_{i=1}^{s}\sum_{j=0}^{d_i-1}\frac{(-1)^j j!}{\lambda_i^{j+1}}Z_{ij}$.

Notations *Let* $y=\sum_{j=1}^{n}y_j v_j \epsilon K^n$, $Y_{\alpha_1}=[y_1,y_2,\cdots,y_{\alpha_1}]^T$, $Y_{\alpha_2}=[y_{\alpha_1+1},y_{\alpha_1+2},\cdots,y_{\alpha_1+\alpha_2}]^T,\cdots,Y_{\alpha_s}=[y_{\alpha_1+\alpha_2+\cdots\alpha_{s-1}+1},y_{\alpha_1+\alpha_2+\cdots\alpha_{s-1}+2},\cdots,y_{\alpha_1+\alpha_2+\cdots\alpha_{s-1}+\alpha_s}]^T$ *then* $Y=[y_1,y_2,\cdots,y_n]^T=[Y_{\alpha_1},Y_{\alpha_1},\cdots,Y_{\alpha_s}]^T$

Lemma 1. *Let* $L_1=Im(diag[O_{\alpha_1},O_{\alpha_2},\cdots,O_{\alpha_i-1},I_{\alpha_i},O_{\alpha_i+1},\cdots,O_{\alpha_s-1},O_{\alpha_s}]\cdot P^{-1})$ *then* $L_1=\{[O_{1\times\alpha_1},O_{1\times\alpha_2},\cdots,O_{1\times\alpha_{i-1}},Y_{\alpha_i},O_{1\times\alpha_{i+1}},\cdots,O_{1\times\alpha_{s-1}},O_{1\times\alpha_s}]^T/Y_{a_i}\epsilon K^{\alpha_i}\}$.

Proof. Let be

$$P^{-1}=\begin{bmatrix}W_{\alpha_1\alpha_1}&\cdots&W_{\alpha_1\alpha_i}&\cdots&W_{\alpha_1\alpha_s}\\\cdots&\cdots&\cdots&\cdots&\cdots\\W_{\alpha_i\alpha_1}&\cdots&W_{\alpha_i\alpha_i}&\cdots&W_{\alpha_i\alpha_s}\\\cdots&\cdots&\cdots&\cdots&\cdots\\W_{\alpha_s\alpha_1}&\cdots&W_{\alpha_s\alpha_i}&\cdots&W_{\alpha_s\alpha_s}\end{bmatrix},$$

then $W_{\alpha_i\alpha_j}$ are $\alpha_i\times\alpha_j$ sub-matrices.

Then it is $diag[O_{\alpha_1},O_{\alpha_2},\cdots,O_{\alpha_{i-1}},I_{\alpha_i},O_{\alpha_{i+1}},\cdots,O_{\alpha_{s-1}},O_{\alpha_s}]\cdot P^{-1}=diag[O_{\alpha_1},O_{\alpha_2},\cdots,O_{\alpha_{i-1}},W_{\alpha_i\alpha_i},O_{\alpha_{i+1}},\cdots,O_{\alpha_{s-1}},O_{\alpha_s}]$.

So $L_1=\{diag[O_{\alpha_1},\cdots,O_{\alpha_{i-1}},W_{\alpha_i\alpha_i},\cdots,O_{\alpha_s}]x\ \epsilon\ K^n\}\ \Rightarrow.\ L_1=\{[O_{1\times\alpha_1},\cdots,O_{1\times\alpha_{i-1}},W_{\alpha_i\alpha_i}x_{\alpha_i},O_{1\times\alpha_{i+1}},\cdots,O_{1\times\alpha_s}]^T/x_{\alpha_i}\ \epsilon\ K^{\alpha_i}\}\ \Rightarrow\ L_1=\{[O_{1\times\alpha_1},\cdots,O_{1\times\alpha_{i-1}},Y_{\alpha_i},O_{1\times\alpha_{i+1}},\cdots,O_{1\times\alpha_s}]^T/Y_{\alpha_i}\ \epsilon\ K^{\alpha_i}\}$.

Lemma 2. *Let* $L_2=Ker(J_A-\lambda_i I_n)^{d_i}$, *then* $L_2=\{[O_{1\times\alpha_1},\cdots,O_{1\times\alpha_{i-1}},Y_{\alpha_i},O_{1\times\alpha_{i+1}},\cdots,O_{1\times\alpha_s}]^T/Y_{a_i}\epsilon K^{\alpha_i}\}$.

Proof. Let $y \in L_2$, then $(J_A - \lambda_i I_n)^{d_i} y = 0$ and

$(J_A - \lambda_i I_n)^{d_i} = diag[(\lambda_1 - \lambda_i)^{d_i} I_{\alpha_1}, \cdots, (\lambda_{\alpha_{i-1}} - \lambda_i)^{d_i} I_{\alpha_{i-1}}, O_{\alpha_i}, (\lambda_{\alpha_{i+1}} - \lambda_i)^{d_i} I_{\alpha_{i+1}}, \cdots, (\lambda_s - \lambda_i)^{d_i} I_{\alpha_s}]$.

So, $diag[(\lambda_1 - \lambda_i)^{d_i} I_{\alpha_1}, \cdots, (\lambda_{\alpha_{i-1}} - \lambda_i)^{d_i} I_{\alpha_{i-1}}, O_{\alpha_i}, (\lambda_{\alpha_{i+1}} - \lambda_i)^{d_i} I_{\alpha_{i+1}}, \cdots, (\lambda_s - \lambda_i)^{d_i} I_{\alpha_s}] y = 0 \Leftrightarrow$

$diag[(\lambda_1 - \lambda_i)^{d_i} I_{\alpha_1}, \cdots, (\lambda_{\alpha_{i-1}} - \lambda_i)^{d_i} I_{\alpha_{i-1}}, O_{\alpha_i}, (\lambda_{\alpha_{i+1}} - \lambda_i)^{d_i} I_{\alpha_{i+1}}, \cdots, (\lambda_s - \lambda_i)^{d_i} I_{\alpha_s}][Y_{\alpha_1} \cdots Y_{\alpha_i} \cdots Y_{\alpha_s}]^T = O \Leftrightarrow$

$Y_{\alpha_j} = 0, \forall j \neq i \Rightarrow$

$y = [O_{1\times\alpha_1}, \cdots, O_{1\times\alpha_{i-1}}, Y_{\alpha_i}, \cdots, O_{1\times\alpha_{i+1}}, \cdots, O_{1\times\alpha_s}]^T$ so $L_2 = Ker(J_A - \lambda_i I_n)^{d_i} = \{[O_{1\times\alpha_1}, \cdots, O_{1\times\alpha_{i-1}}, Y_{\alpha_i}, \cdots O_{1\times\alpha_{i+1}}, \cdots, O_{1\times\alpha_s}]^T / Y_{\alpha_i} \in K^{\alpha_i}\}$.

Corollary $Im(diag[O_{\alpha_1}, O_{\alpha_2}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_{s-1}}, O_{\alpha_s}] \cdot P^{-1}) = Ker(J_A - \lambda_i I_n)^{d_i}$.

Property 15. $Range(Z_{i0}) = Im(Z_{i0}) = Kernel[(A - \lambda_i I_n)^{d_i}], i = 1, 2, \cdots, s$.

Proof. It is $x \in Im(Z_{i0}) \Leftrightarrow x = Z_{i0} x \Leftrightarrow$

$x = P \cdot diag[O_{\alpha_1}, \cdots, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1} x \Leftrightarrow$

$P^{-1} x = (diag[O_{\alpha_1}, \cdots, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}) x \Leftrightarrow$

$P^{-1} x \in Im(diag[O_{\alpha_1}, \cdots, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}) \Leftrightarrow$

$P^{-1} x \in Kernel[(J_A - \lambda_i I_n)^{d_i}] \Leftrightarrow (J_A - \lambda_i I_n)^{d_i} \cdot P^{-1} x = O \Leftrightarrow$

$(A - \lambda_i I_n)^{d_i} x = O \Leftrightarrow x \in Kernel[(A - \lambda_i I_n)^{d_i}]$.

So $Im(Z_{i0}) = Kernel[(A - \lambda_i I_n)^{d_i}]$, is $i = 1, 2, \cdots, s$.

Property 16. $Kernel(Z_{i0}) = Im(Z_{i0}) \oplus \cdots \oplus Im(Z_{i-1,0}) \oplus Im(Z_{i+1,0}) \oplus \cdots \oplus Im(Z_{s0})$, $i = 1, 2, \cdots, s$.

Proof. It is $Kernel(Z_{i0}) = Im(I_n - Z_{i0}) =$

$Im(Z_{i1} + \cdots + Z_{i-1,0} + Z_{i+1,0} + \cdots + Z_{s0} = Im(Z_{10}) \oplus \cdots \oplus Im(Z_{i-1,0}) \oplus Im(Z_{i+1,0}) \oplus \cdots \oplus Im(Z_{s0})$,

because $Im(Z_{j0}) \cap Im(Z_{k0}) = \{O\}$, for $j \neq k$.

Property 17. $(A - \lambda_i I_n)^{d_i}) Z_{i0} = O_n, i = 1, 2, \cdots, s$.

Proof. It is a direct result from the relation

$Z_{i0} = P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}$.

Property 18. $Im(Z_{i0}) \supset Im(Z_{i1}) \supset \cdots \supset Im(Z_{i,d_i-1}), i = 1, 2, \cdots, s$.

Proof. It is $x \in Im(Z_{ij}) \Leftrightarrow \exists y \in K^n : x = Z_{ij} \Leftrightarrow x = \frac{1}{j}(A - \lambda_i I_n) Z_{i,j-1} y$.

Let $y_1 = \frac{1}{j}(A - \lambda_i I_n) y$, then

$x = Z_{i,j-1} y_1 \Rightarrow x \in Im(Z_{i,j-1}) \Rightarrow Im(Z_{ij}) \subseteq Im(Z_{i,j-1})$.

Property 19. If the matrix $A \in \mathbb{R}^{n\times n}$ is Symmetric then the matrix Z_{i0} is an orthogonal projector onto the eigenspace of λ_i, $i = 1, 2, \cdots, s$.

Proof. Let $A = P \cdot J_A \cdot P^{-1}$, where J_A is the Jordan form and P is the transition matrix of the matrix A. The matrix A is Symmetric, so the matrix J_A is Symmetric and the matrix P is Orthogonal. Hence the matrix Z_{i0} is Symmetric with $Z_{i0}^2 = Z_{i0}$. So the matrix Z_{i0} is an orthogonal projector onto the eigenspace of λ_i, for $i = 1, 2, \cdots, s$.

Property 20. If the matrix $A \in \mathbb{C}^{n\times n}$ is Hermitian then the matrix Z_{i0} is an orthogonal projector onto the eigenspace of λ_i,$i = 1, 2, \cdots, s$.

Proof. Let $A = P \cdot J_A \cdot P^{-1}$, where J_A is the Jordan form and P is the transition matrix of the matrix A. The matrix A is Hermitian so, the matrix J_A is Hermitian and the matrix P is Unitary. Hence the matrix Z_{i0} is Hermitian with $Z_{i0}^2 = Z_{i0}$. So the matrix Z_{i0} is an orthogonal projector onto the eigenspace of λ_i, for $i = 1, 2, \cdots, s$.

Property 21. $Kernel(Z_{i0}) \bigcap Kernel(Z_{j0}) \supset \{0\}$.

Proof. It is $x \in Kernel(Z_{i0}) \Leftrightarrow Z_{i0}x = 0 \Leftrightarrow$
$P \cdot diag[O_{\alpha_1}, \cdots, O_{\alpha_{i-1}}, I_{\alpha_i}, O_{\alpha_{i+1}}, \cdots, O_{\alpha_s}] \cdot P^{-1}x = O \Leftrightarrow X_{\alpha_i} = O$.
So $x \in KernelZ_(i0) \Leftrightarrow X_{\alpha_i}{=}O \Leftrightarrow X = \left[X_{\alpha_1}\ X_{\alpha_2} \cdots X_{\alpha_{i-1}}\ O_{\alpha_i}\ X_{\alpha_{i+1}} \cdots X_{\alpha_{s-1}}\ X_{\alpha_s}\right]^T$.

It is $x \in Kernel(Z_{j0}) \Leftrightarrow X_{\alpha_j}{=}O \Leftrightarrow X{=}\left[X_{\alpha_1}\ X_{\alpha_2} \cdots X_{\alpha_{j-1}}\ O_{\alpha_j}\ X_{\alpha_{j+1}} \cdots X_{\alpha_{s-1}}\ X_{\alpha_s}\right]^T$.

Finally for $i < j$ it is
$Kernel(Z_{i0}) \bigcap Kernel(Z_{j0}){=}\{\left[X_{\alpha_1} \cdots X_{\alpha_{i-1}}\ O_{\alpha_i} \cdots X_{\alpha_{j-1}}\ O_{\alpha_j} \cdots X_{\alpha_s}\right]^T / X_{\alpha_k} \in K^{\alpha_k}\}$.

Property 22. $Kernel(Z_{i0}) + Kernel(Z_{j0}) = K^n, \forall i \neq j$.

Proof. It is $dimKernel(Z_{i0}) = n - \alpha_i$, $dimKernel(Z_{i0}) = n - \alpha_j$ and
$dim[Kernel(Z_{i0}) \bigcap Kernel(Z_{j0})] = n - \alpha_i - \alpha_j$, so
$dim[Kernel(Z_{i0}) + Kernel(Z_{j0})] = n \Rightarrow$
$Kernel(Z_{i0}) + Kernel(Z_{j0})] = K^n, \forall i \neq j$.

Conclusions

In this paper the definitions of the component matrices are presented and a main number of properties are proved. Furthermore some illustrative examples are given.

The results obtained from the work are the necessary and sufficient tools to study the sequences and series of matrices.

Acknowledgements This research has been co-financed by the European Union (European Social Fund-EFS) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)—Research Funding Program: ARCHIMEDES III.

Investing in knowledge society through the European Social Fund.

Archimedes III: Funding of Research Groups in TEI of W. Macedonia, project with the title "Applications of dynamics systems", (MIS 383583).

References

1. Bavely, C.A., Stewart, G.W.: An algorithm for computing reducing subspaces by block diagonalization. SIAM J. Numer. Anal. **16**(2), 359–367 (1979)
2. Bhatia, R.: Matrix Analysis. Springer, New York (1997)

3. Belman, R.: Introduction to Matrix Analysis, 2nd edn. McGraw-Hill, New York (1970)
4. Cullen, C.G.: Matrices and Linear Transformations, 2nd edn. Addison-Wesley, Reading (1972)
5. Demmel, J.W.: Applied Numerical Linear Algebra. Society for Industrial and Applied Mathematics, Philadelphia (1997)
6. Dennis, J.E., Traub, J.F., Weber, R.P.: The algebraic theory of matrix polynomials. SIAM J. Numer. Anal. **13**(6), 831–845 (1976)
7. Dieudonne, J.: Foundations of Modern Algebra. Academic, New York (1960)
8. Higham, N.J.: Functions of Matrices, Theory and Computation. Society of Industrial and Applied Mathematics, Philadelphia (2008)
9. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
10. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
11. Lal Mehta, M.: Matrix Theory: Selected Topics and Useful Results, 2nd edn. Hindustan Publishing Company, Delhi (1989)
12. Lancaster, P., Tismenetsky, M.: Theory of Matrices, 2nd edn. Academic, London (1995)
13. Pullman, N.J.: Matrix Theory and Its Applications. Marcel Dekker, New York (1976)

Solutions of Some Types of Differential Equations and of Their Associated Physical Problems by Means of Inverse Differential Operators

H.M. Srivastava and K.V. Zhukovsky

Abstract We present an operational method, involving an inverse derivative operator, in order to obtain solutions for differential equations, which describe a broad range of physical problems. Inverse differential operators are proposed to solve a variety of differential equations. Integral transforms and the operational exponent are used to obtain the solutions. Generalized families of orthogonal polynomials and special functions are also employed with recourse to their operational definitions. Examples of solutions of physical problems, related to the mass, the heat and other processes of propagation are demonstrated by the developed operational technique. In particular, the evolutional type problems, the generalizations of the Black–Scholes, of the heat, of the Fokker–Plank and of the telegraph equations are considered as well as equations, involving the Laguerre derivative operator.

Keywords Inverse operator • Inverse derivative • Exponential operator • Differential equation • Hermite and Laguerre polynomials • Special functions

Introduction

Differential equations play very important rôle in mathematics and their rôle in the description of physical processes cannot be overestimated. Thus, obtaining the solutions for differential equations is of paramount importance. Few types

H.M. Srivastava (✉)
Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3R4
e-mail: harimsri@math.uvic.ca

K.V. Zhukovsky
Faculty of Physics, Moscow State University, Leninskie Gory, Moscow 119991, Russian Federation
e-mail: zhukovsk@physics.msu.ru

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_26

of differential equations allow explicit and straightforward analytical solutions. Lately, development of computer methods facilitated equations solving. However, the understanding of the obtained solutions and their application to description of physical processes in some cases is still challenging. Despite the revolutionary breakthrough in computer methods and techniques in the twenty-first century, analytical studies remain required because, in general, they allow transparent meaning of the solutions obtained. Indeed, expansion in series of orthogonal polynomials [1] is useful for solving many physical problems (see, for example, [17] and [18]). Hermite, Laguerre and other polynomial families are usually defined by series or sums; for the purpose of differential equations solution and in the context of the operational approach, they are best defined by operational relations. These polynomials possess generalized forms with many variables and indices (see [6] and [8]), which arise naturally in studies of physical problems, related to propagation and radiation of accelerated charges, heat and mass transfer (see, for example, [5, 10] and [28, 29, 31]). In what follows, we develop an analytical method to obtain the solutions for various types of differential equations on the base of the operational identities, employing special functions, expansions in the series of Hermite and Laguerre polynomials and their modified forms (see [1] and [15]). The above-mentioned mathematical instruments are very useful for solution of a remarkably broad range of physical problems. Recent developments in technique and technology in synchrotron radiation (SR) and undulator radiation (UR) induced particular interest to the analysis of the radiation from charged particle beams and their propagation in accelerators and in insertion devices. Free electron lasers (FEL) now only open new frontiers for researchers, but they also require high-quality radiation sources and high precision undulators. This calls for modern mathematical instruments, suitable for the analysis of the undergoing physical processes and for the adequate description of the performance of the devices. Analytical solutions have this value since they usually possess more transparent physical meaning than numerical solutions and allow deep understanding of the processes. When it comes to a numerical analysis, there are also practical and theoretical reasons for examining the process of inverting differential operators. Indeed, the inverse or integral form of a differential equation displays explicitly the input–output relationship of the system. Moreover, integral operators are computationally and theoretically less troublesome than differential operators; for example, differentiation emphasizes data errors, whereas integration averages them. Thus, clearly, it may be advantageous to apply computational procedures to differential systems, based upon the inverse or integral description of the system.

The evident concept of an inverse function is a function that undoes another function: if an input x into the function f produces an output y, then putting y into the inverse function g produces the output x, and vice versa. That is,

$$f(x) = y \qquad \text{and} \qquad g(y) = x \quad \text{or} \quad g\big(f(x)\big) = x.$$

If a function f has an inverse f^{-1}, it is invertible and the inverse function is then uniquely determined by f. We can develop similar approach with regard to differential operators. In what follows, we will further develop this technique and explore its relation with extended forms of orthogonal polynomials, producing useful relations for solutions of a variety of differential equations, by means of inverse derivatives. The relevant physical problems will be considered. Recent studies of the UR with recourse to the generalized Bessel [9] and Airy type functions [12, 22, 30, 32] allowed for the analytical description of the influence of additional non-periodic magnetic components in undulators on the UR properties. For these studies, the expansion in series of Hermite and Laguerre polynomials [1] were employed, including those with many indices and variables (see [6] and [8]), which were earlier used with relations to other physical problems (see [17] and [18]). In this framework the operational definitions for the polynomials are useful (see [15]). In what follows, we shall develop the operational approach to obtain analytical solutions for a broad class of differential equations, including evolution type equations, generalized forms of heat and mass transfer and Black–Scholes type equations, involving also the Laguerre derivative operator.

Let us denote a common differential operator $D = \frac{d}{dx}$. The inverse derivative of a function $f(x)$ is another function $F(x)$:

$$D^{-1}\{f(x)\} = F(x), \tag{1}$$

whose derivative is given by

$$F'(x) = f(x).$$

Naturally, we expect the anti-derivative or the inverse derivative D^{-1} as the inverse operation of differentiation to be an integral operator. Evidently, the generalized form of the inverse derivative of $f(x)$ with respect to x is given by

$$\int f(x) = F(x) + C,$$

where C is a constant of integration. The action of the inverse derivative operator D_x^{-n} of the nth order, which is given by

$$D_x^{-n}f(x) = \frac{1}{(n-1)!}\int_0^x (x-\xi)^{n-1}f(\xi)d\xi \qquad (n \in \mathbb{N} := \{1, 2, 3, \ldots\}), \tag{2}$$

can be complemented with the definition for its 0th order action as follows:

$$D_x^0 f(x) = f(x). \tag{3}$$

Hence, we can write

$$D_x^{-n}\, 1 = \frac{x^n}{n!} \qquad (n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}). \tag{4}$$

For the following general form of a differential equation:

$$\psi(D)F(x) = f(x), \tag{5}$$

where $\psi(D)$ is a differential operator, the inverse differential operator $1/\psi(D)$ or $\big(\psi(D)\big)^{-1}$ is such that

$$\psi(D)\big(\psi(D)\big)^{-1}\{f(x)\} = f(x). \tag{6}$$

By using (5), we obtain the following particular integral:

$$F(x) = \big(\psi(D)\big)^{-1}\{f(x)\}. \tag{7}$$

The inverse differential operator $(\psi(D))^{-1}$ is evidently linear, that is,

$$\frac{1}{\psi(D)}\{af(x) + bg(x)\} = a\frac{1}{\psi(D)}\{f(x)\} + b\frac{1}{\psi(D)}\{g(x)\}, \tag{8}$$

where a and b are constants, and $f(x)$ and $g(x)$ are some suitable functions of x.

Let us consider an elementary example of the following simple equation:

$$\psi(D)F(x) = \mathrm{e}^{\alpha s}, \tag{9}$$

where $\psi(D)$ consists of derivatives of various orders. The action of $\psi(D)$ on $\exp(\alpha x)$ results in the following equation:

$$\psi(D)\,\mathrm{e}^{\alpha x} = (D^2 + c_1 D + c_0)\,\mathrm{e}^{\alpha x} = \psi(\alpha)\,\mathrm{e}^{\alpha x}.$$

By applying the inverse operator $\big(\hat{\psi}(D)\big)^{-1}$ to both sides, we obtain

$$\mathrm{e}^{ax} = \psi(\alpha)\frac{1}{\psi(D)}\{\mathrm{e}^{\alpha x}\}$$

or

$$\frac{\mathrm{e}^{ax}}{\psi(\alpha)} = \frac{1}{\psi(D)}\{\mathrm{e}^{\alpha x}\}.$$

We thus conclude that Eq. (9) possesses the following particular integral:

$$F(x) = \big(\psi(D)\big)^{-1}\{\mathrm{e}^{\alpha x}\} = \frac{\mathrm{e}^{\alpha x}}{\psi(\alpha)}. \tag{10}$$

It can be easily shown by means of the inverse derivative operator that (10), being the solution of Eq. (9), is true also for the operator $\psi(D)$ with higher than the second-order derivatives. Moreover, it is easy to prove the following identity:

$$\left(\psi(D)\right)^{-1}\{e^{\alpha x}f(x)\} = e^{\alpha x}\left(\psi(D+\alpha)\right)^{-1}f(x). \tag{11}$$

Moreover, the action of the inverse operator $\left(\psi(D+\alpha)\right)^{-1}$ on a function $f(x)$, which can be expressed via the inverse differential operator $\left(\psi(D)\right)^{-1}$, reads as follows:

$$F(x) = \left(\psi(D+\alpha)\right)^{-1}\{f(x)\} = e^{-\alpha x}\left(\psi(D)\right)^{-1}\{e^{\alpha x}f(x)\}. \tag{12}$$

Inverse Differential and Exponential Operators for Solutions of Some Differential Equations

Let us consider the following equation:

$$\left(\beta^2 - (D+\alpha)^2\right)^{\nu}\{F(x)\} = f(x) \qquad \left(D+\alpha =: \tilde{D}\right), \tag{13}$$

where we denote by $\tilde{D}$ the operator $D+\alpha$ and the parameters α, β and ν are constants. In order to find the particular integral $F(x)$ given by

$$F(x) = \left(\beta^2 - \tilde{D}^2\right)^{-\nu}\{f(x)\}, \tag{14}$$

we shall make use of the following well-known operational identity (see [15] and [24]), frequently used in fractional derivative calculus:

$$\hat{q}^{-\nu} = \frac{1}{\Gamma(\nu)}\int_0^{\infty}\exp(-\hat{q}t)t^{\nu-1}\,dt \qquad \left(\min\{\Re(q),\Re(\nu)\} > 0\right), \tag{15}$$

which reads for the operator $\hat{q} = \beta^2 - \tilde{D}^2$:

$$\left(\beta^2 - \tilde{D}^2\right)^{-\nu}\{f(x)\} = \frac{1}{\Gamma(\nu)}\int_0^{\infty}\exp(-\beta^2 t)t^{\nu-1}\exp(t\tilde{D}^2)f(x)\,dt. \tag{16}$$

There are several ways to proceed with the solution. One of them consists in the following. Note that differential operator $-t\tilde{D}^2$ in the exponential reduces to the first order derivative with the help of the following integral presentation for the exponential of a square of an operator $\hat{p}$ [7]:

$$\exp\left(\hat{p}^2\right) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left(-\xi^2 + 2\xi\hat{p}\right)d\xi, \tag{17}$$

where $\hat{p} = \sqrt{t}\tilde{D}$ in our case. Thus, the above formula reads as follows:

$$\exp(t\tilde{D}^2)f(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-\xi^2 + 2\xi\sqrt{t}\tilde{D})f(x)d\xi. \tag{18}$$

Now, if we take into account the action of the operator of translation $\exp(\eta\tilde{D})$ for $\tilde{D} = D + \alpha$:

$$\exp[\eta(D+\alpha)]f(x) = \exp(\eta\alpha)f(x+\eta), \tag{19}$$

the above-sketched operational procedure yields the following expression for the particular integral (14):

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)} \int_0^{\infty} t^{\nu-1} \exp\left[(\alpha^2 - \beta^2)t\right] \cdot \int_{-\infty}^{\infty} \exp\left[-(\xi - \sqrt{t}\alpha)^2\right] f(x + 2\xi\sqrt{t})\, d\xi\, dt. \tag{20}$$

Upon performing the change of variables, given by

$$\eta = x + 2\xi\sqrt{t} \qquad \text{and} \qquad t = \tau^2, \tag{21}$$

we finally obtain the following solution of Eq. (13):

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)} \int_0^{\infty} \tau^{2(\nu-1)} \exp\left[-(\beta\tau)^2\right] \cdot \int_{-\infty}^{\infty} \exp\left[-\left(\frac{\eta - x}{2\tau}\right)^2 + \alpha(\eta - x)\right] f(\eta)\, d\eta\, d\tau. \tag{22}$$

Evidently, for $\alpha = 0$, the equation reduces to

$$\left(\beta^2 - D^2\right)^{\nu} F(x) = f(x), \tag{23}$$

and its solution becomes the particular case of (16) with the substitution $\tilde{D} \to D$:

$$F(x) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \exp(-\beta^2 t)t^{\nu-1}\, \hat{S}f(x)\, dt, \tag{24}$$

where the differential operator $\hat{S}$ given by

$$\hat{S} = \exp(tD_x^2) \equiv \exp(tD^2) \tag{25}$$

was thoroughly explored by Srivastava and Manocha in [24]. In particular, for

$$f(x) = \exp(-x^2),$$

we can make use of the Gleisher operational rule [24]:

$$\hat{S}f(x) = \exp\left(y\frac{\partial^2}{\partial x^2}\right)\exp(-x^2) = \frac{1}{\sqrt{1+4y}}\exp\left(-\frac{x^2}{1+4y}\right) \tag{26}$$

to obtain the following particular solution:

$$F(x) = \frac{1}{\Gamma(\nu)}\int_0^\infty \frac{\mathrm{e}^{-\beta^2 t}\ t^{\nu-1}}{\sqrt{1+4t}}\exp\left(-\frac{x^2}{1+4t}\right)\ dt. \tag{27}$$

The other approach to Eq. (13) consists in combining the exponential operator technique, the inverse derivative formalism and the Gauss transform. Indeed, when solving equations with $D+\alpha$, we can write the particular integral based upon the operational rule (12), where

$$\psi^{-1}(D) = \left(\beta^2 - D^2\right)^{-\nu}. \tag{28}$$

Proceeding with account for (23), (24) and (25), we compute the result of the action of the operator $\exp(\partial_x^2)$ on $\exp(\alpha x)g(x)$ with the help of the following chain rule:

$$\exp(y\partial_x^2)\exp(\alpha x)g\,(x) = \exp(\alpha\, x)\exp(\alpha^2 y\,)\exp(2\alpha y\partial_x)\exp(y\partial_x^2)g\,(x)\,, \tag{29}$$

where y and α are the parameters. It eventually yields the following solution for Eq. (13):

$$F(x) = \frac{1}{\Gamma\,(\nu)}\int_0^\infty t^{\nu-1}\exp\left\{-(\beta^2-\alpha^2)t\right\}\hat{\Theta}\hat{S}f(x)dt, \tag{30}$$

where $\hat{\Theta}$ is the well-known operator of translation

$$\hat{\Theta} = \exp(2\alpha t D_x) \equiv \exp\left(2\alpha t\frac{\partial}{\partial x}\right) \quad \text{and} \quad \hat{\Theta}f(x) = f(x+2\alpha t) \tag{31}$$

and operator $\hat{S}$ is encountered in problems, related to heat propagation and defined in (25). Its action can be written in integral form by means of common Gauss transforms:

$$Fi(x,t) \equiv \hat{S}f(x) = \frac{1}{2\sqrt{\pi t}}\int_{-\infty}^\infty \exp\left(-\frac{(x-\xi)^2}{4t}\right)f(\xi)\ d\xi. \tag{32}$$

Thus, we conclude that the integrand of the solution (30) of the Eq. (13), apart from the phase and the factor, responsible for the equation dimension, ν, is a result of consequent action of operators of heat propagation $\hat{S}$ and operator of translation $\hat{\Theta}$ on the function $f(x)$:

$$F(x,t) \equiv \hat{\Theta}\hat{S}f(x) = Fi(x+2\alpha t, t)$$
$$= \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x+2\alpha t-\xi)^2}{4t}\right) f(\xi)\, d\xi. \tag{33}$$

With these notations, we can write the solution as follows:

$$F(x) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} t^{\nu-1} \exp\left[-(\beta^2-\alpha^2)t\right]\ F(x,t)\, dt. \tag{34}$$

The solution is best illustrated by the example of a Gaussian function $f(x)$:

$$f(x) = \exp(-x^2). \tag{35}$$

With the help of the above described operational procedure and on account of Eq. (26) we easily obtain the following solution of Eq. (13) for $f(x)$ given by a Gaussian:

$$F(x) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \frac{t^{\nu-1}\exp\left[-(\beta^2-\alpha^2)t\right]}{\sqrt{1+4t}} \exp\left(-\frac{(x+2\alpha t)^2}{1+4t}\right) dt. \tag{36}$$

So far we have demonstrated on simple examples how the usage of inverse derivative together with operational formalism and, in particular, with exponential operator technique provides elegant and easy way to find solutions in some classes of differential equations. In what follows we will apply the concept of inverse differential operator to find solutions of more sophisticated problems, expressed by differential equations.

Inverse Differential Operators and Orthogonal Polynomials

Despite the traditional presentation of many polynomial families as the expansion in series, they are worth being viewed from operational point of view too. Particularly interesting appears their relation with the exponential operators of derivatives and inverse derivatives and also, special functions. Recently, Hermite, Laguerre and other polynomial families were reconsidered by means of the operational technique (see [6, 8] and [11]). Hermite polynomials of two variables are explicitly given by the following operational rule (see [6]) and the series expansion (see [16]):

$$H_n^{(m)}(x,y) = \exp\left(y\frac{\partial^m}{\partial x^m}\right)x^n = n! \sum_{r=0}^{[n/m]} \frac{x^{n-mr}y^r}{(n-mr)!r!}. \tag{37}$$

We note that

$$H_n^{(1)}(x,y) = (x+y)^n \qquad \text{and} \qquad H_n^{(2)}(x,y) = H_n(x,y)\,, \tag{38}$$

where $H_n(x,y)$ are more commonly known Hermite polynomials of two variables defined by

$$H_n(x,y) = \exp\left(y\frac{\partial^2}{\partial x^2}\right)x^n \quad \text{and} \quad H_n(x,y) = n!\sum_{r=0}^{[n/2]}\frac{x^{n-2r}y^r}{(n-2r)!r!} \tag{39}$$

with the following generating exponent:

$$\exp\left(xt+yt^2\right) = \sum_{n=0}^{\infty}\frac{t^n}{n!}H_n(x,y). \tag{40}$$

They can be reduced to the well-known forms of Hermite polynomials of a single variable:

$$H_n(x,y) = (-i)^n y^{n/2} H_n\left(\frac{ix}{2\sqrt{y}}\right) = i^n(2y)^{n/2}He_n\left(\frac{x}{i\sqrt{2y}}\right). \tag{41}$$

We note also the following useful and easy to prove relation (see [16]) for Hermite polynomials:

$$z^n H_n(x,y) = H_n(xz, yz^2). \tag{42}$$

Laguerre polynomials of two variables can be given by an operational relation (see [6]) or a sum as follows:

$$L_n(x,y) = \exp\left(-y\frac{\partial}{\partial x}x\frac{\partial}{\partial x}\right)\frac{(-x)^n}{n!} = n!\sum_{r=0}^{n}\frac{(-1)^r y^{n-r}x^r}{(n-r)!\,(r!)^2}. \tag{43}$$

They also reduce to polynomials of a single variable (see [24]) as follows:

$$L_n(x,y) = y^n L_n(x/y) \qquad \text{and} \qquad L_n(x) = y^{-n}L_n(xy,y) = L_n(x,1). \tag{44}$$

The introduction of the second variable in Hermite and Laguerre polynomials allows us to consider them as solutions of partial differential equations with proper initial conditions:

$$\partial_y L_n(x,y) = -(\partial_x x\,\partial_x)L_n(x,y) \qquad \text{with} \qquad L_n(x,0) = (-x)^n n! \tag{45}$$

for the Laguerre polynomials $L_n(x,y)$ and

$$\partial_y H_n(x,y) = \partial_x^2 H_n(x,y) \qquad \text{with} \qquad H_n(x,0) = x^n \tag{46}$$

for the Hermite polynomials $H_n\,(x,y)$. It is important that the following differential operators:

$$_L D_x = \frac{\partial}{\partial x} x \frac{\partial}{\partial x} = -\hat{P} \qquad \text{and} \qquad \hat{M} = y - D_x^{-1}, \tag{47}$$

where

$$D_x^{-1}\{f(x)\} = \int_0^x f(\xi)d\xi$$

are non-commutative:

$$[_L D_x, D_x^{-1}] = -1 \tag{48}$$

with the standard definition for the commutator

$$[A,B] = AB - BA,$$

and the following operational relation between them (see [11]):

$$_L D_x = \frac{\partial}{\partial x} x \frac{\partial}{\partial x} = \frac{\partial}{\partial D_x^{-1}}. \tag{49}$$

With the help of this relation, we can extend our approach on differential equations, including operator $\partial_x x \partial_x$, sometimes called Laguerre derivative $_L D_x$. Then, from the definition (47), we immediately conclude for Laguerre polynomials $L_n(x,y)$, defined in (43) that (in terms of the inverse derivative operator) they are expressed as follows:

$$L_n(x,y) = n! \sum_{k=0}^{n} \frac{(-x)^k y^{n-k}}{(n-k)!(k!)^2} = (y - D_x^{-1})^n\{1\}. \tag{50}$$

Moreover, from (43) and (50), we conclude that the following operational identity is true for the Laguerre polynomials:

$$\exp\left(\alpha \frac{\partial}{\partial D_x^{-1}}\right) L_n(x,y) = L_n(x, y-\alpha). \tag{51}$$

Extensive study of the relations between various polynomial families, inverse derivative operator, exponential operator and realization of operators $\hat{M}$ and $\hat{P}$, which stand, respectively, for multiplicative operator and differential operator for quasi-monomial polynomials can be found in [7] and [13]. Various polynomial families, such as Hermite, Laguerre, Legendre, Shaffer and hybrid polynomials were discussed there in the context of umbral calculus and of a more general family of Appell polynomials, which they belong to. Such consideration was possible in the framework of operational approach, where inverse derivative plays important role as an instrument for the study of relevant polynomial families, their features and properties.

Operational Approach and Orthogonal Polynomials for the Solution of Some Differential Equations

Some Solutions for Laguerre Type Ordinary Differential Equations

Let us consider the following equation with the Laguerre derivative:

$$\left(\beta - \frac{\partial}{\partial x} x \frac{\partial}{\partial x}\right)^{\nu} \{F(x)\} = f(x). \tag{52}$$

From the operational point of view, its solution reads as follows:

$$\begin{aligned} F(x) &= \left(\beta - \frac{\partial}{\partial x} x \frac{\partial}{\partial x}\right)^{-\nu} \{f(x)\} \\ &= \frac{1}{\Gamma(\nu)} \int_0^\infty \exp(-\beta t) t^{\nu-1} \exp\left(t_L D_x\right) f(x)\, dt. \end{aligned} \tag{53}$$

We note that, upon the change of variable $t \to e^t$, the last integral can be transformed into the following one:

$$F(x) = \left(\beta - \frac{\partial}{\partial x} x \frac{\partial}{\partial x}\right)^{-\nu} \{f(x)\} = \frac{1}{\Gamma(\nu)} \int_{-\infty}^\infty e^{t\nu}\, e^{-\beta\, e^t}\, e^{e^t\, {}_L D_x} f(x) dt. \tag{54}$$

In the particular case when $\beta = 1$ and $\nu = 1$, the above integral presentation reduces to the Laplace transforms for the differential operator ${}_L D_x$, involved in (53), which is identical (except for the change ${}_L D_x \leftrightarrow D_x$) to the well-known Laplace transforms for the operator D_x:

$$\frac{1}{1 - \hat{D}_x} = \int_0^\infty \exp[-s(1 - \hat{D}_x)]\, ds. \tag{55}$$

By choosing $f(x) = x^n$ in Eq. (53), we obtain the following result:

$$(\beta - {}_L D_x)^{-\nu}\{x^n\} = \frac{n!}{\Gamma(\nu)} \int_0^\infty \exp(-\beta t) t^{n+\nu-1}\, L_n\left(\frac{x}{t}\right) dt. \tag{56}$$

Moreover, we can write a solution of Eq. (52) for an arbitrary function $f(x)$, if its expansion into series of the simple Laguerre polynomials $L_n(x)$ exists. Then, provided that

$$f(x) = \sum_{n=0}^{\infty} c_n L_n(x), \tag{57}$$

and taking into account (51), the solution (53) of Eq. (52) can also be written as the following integral of series of Laguerre polynomials:

$$F(x) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \exp(-\beta t) t^{\nu-1} \sum_{n=0}^{\infty} c_n L_n(x, 1-t) dt. \tag{58}$$

For the exponential function $f(x) = \exp(-\gamma x)$, we can employ the generalized form of the Glaisher operational rule (see [5]):

$$\exp(-t_L D_x) \cdot \exp(-\gamma x) = \frac{1}{1-\gamma t} \exp\left(-\frac{\gamma x}{1-\gamma t}\right), \tag{59}$$

which immediately yields the following result:

$$\begin{aligned} &\left(\beta - \frac{\partial}{\partial x} x \frac{\partial}{\partial x}\right)^{-\nu} \exp(-\gamma x) \\ &\quad = \frac{1}{\Gamma(\nu} \int_0^{\infty} \exp(-\beta t) t^{\nu-1} \frac{1}{1+\gamma t} \exp\left(-\frac{\gamma x}{1+\gamma t}\right) dt. \end{aligned} \tag{60}$$

Another interesting case arises in the case of the Laguerre derivative ${}_L D_x$ instead of the common derivative ∂_x in Eq. (13). Let us choose the following initial condition function:

$$f(x) = W_0(-x^2, 2), \tag{61}$$

where

$$W_n(x, m) = \sum_{s=0}^{\infty} \frac{x^s}{s!(ms+n)!}, \tag{62}$$

where $W_n(x, m)$ is the particular case of the Bessel–Write function [11]. Then, following (15), we can write the solution given by

$$\begin{aligned} &\left(\beta^2 - \left(\frac{\partial}{\partial x} x \frac{\partial}{\partial x}\right)^2\right)^{-\nu} W_0(-x^2, 2) \\ &\quad = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \exp(-\beta^2 t) t^{\nu-1} \exp(t_L D_x^2) f(x) dt. \end{aligned} \tag{63}$$

Eventually, basing upon the operational definition of Laguerre polynomials (43) and exploiting the other generalized form of the Gleisher operational rule [15] in the form:

$$\exp(t_L D_x^2)\, W_0(-x^2,2) = \frac{1}{\sqrt{1+4t}} W_0\left(-\frac{1}{1+4t},2\right), \tag{64}$$

we obtain

$$\left(\beta^2 - \left(\frac{\partial}{\partial x}x\frac{\partial}{\partial x}\right)^2\right)^{-\nu} W_0(-x^2,2) \\ = \frac{1}{\Gamma(\nu)}\int_0^\infty \exp(-\beta^2 t)t^{\nu-1}\frac{1}{\sqrt{1+4t}} W_0\left(-\frac{1}{1+4t},2\right)dt. \tag{65}$$

Moreover, according to the developed procedure, we can write the solutions for other types of equations different from those already specified above. For example, we can take advantage of the following generalization of the Laguerre polynomials $L_n^{(\alpha)}(x,y)$:

$$L_n^{(\alpha)}(x,y) = \exp\left(-y\breve{D}_x\right)\left\{\frac{(-x)^n}{n!}\right\}, \tag{66}$$

where operator $\breve{D}_x$ is defined as follows:

$$\breve{D}_x = x\partial_x^2 + (\alpha+1)\partial_x. \tag{67}$$

Inverse operator technique easily allows us to write the solution of the equation

$$\left(x\,\partial_x^2 + (\alpha+1)\,\partial_x\right)F(x) = f(x). \tag{68}$$

Indeed, by following the operational rule (15), we get

$$\breve{D}_x^{-\nu} f(x) = \frac{1}{\Gamma(\nu)}\int_0^\infty \exp(-\beta t)t^{\nu-1}\exp\left(t\breve{D}_x\right)f(x)dt \tag{69}$$

and for the initial condition function $f(x) = x^n$, we then write

$$\breve{D}_x^{-\nu} x^n = \frac{(-1)^n n!}{\Gamma(\nu)}\int_0^\infty \exp(-\beta t)t^{\nu-1}L_n^{(\alpha)}(x,-t)dt. \tag{70}$$

Similarly to (56), by taking advantage of the generalized form of the Gleisher operational rule from [11], we find for the operator $\breve{D}_x$ and

$$f(x) = \exp(-\gamma x)$$

that

$$\breve{D}_x^{-\nu}\exp(-\gamma x) = \frac{1}{\Gamma(\nu)}\int_0^\infty \exp(-\beta t)t^{\nu-1}\frac{1}{(1+\gamma t)^{\alpha+1}}\exp\left(-\frac{\gamma x}{1+\gamma t}\right)dt. \quad (71)$$

Orthogonal Polynomials and Special Functions for the Convolution Forms of Solutions for Ordinary Differential Equations

In what follows, we will focus on the technique of inverse operator, applied for the derivation of a general convolution form of solution for Eq. (13). First of all, let us note that for the particular case of the initial function $f(x) = x^k$ we can make use of the operational rule (12) and of the identity:

$$\exp(yD_x^2)x^k\,\mathrm{e}^{\alpha x} = \mathrm{e}^{(\alpha x+\alpha^2 y)}\,H_k(x+2\alpha y, y), \quad (72)$$

which arises from the operational relation:

$$\exp\left(y\frac{\partial^m}{\partial x^m}\right)f(x) = f\left(x + my\frac{\partial^{m-1}}{\partial x^{m-1}}\right)\{1\} \quad (73)$$

and from generating function (40). Then, with account for (72), we write the particular integral (14) for the function

$$f(x) = x^k$$

as follows:

$$\begin{aligned} F(x) &= (\beta^2 - (D_x+\alpha)^2)^{-\nu}x^k \\ &= \frac{1}{\Gamma(\nu)}\int_0^\infty \mathrm{e}^{-t(\beta^2-\alpha^2)}\,t^{\nu-1}H_k(x+2\alpha t, t)dt. \end{aligned} \quad (74)$$

The above expression with the shifted argument of the Hermite polynomial can be derived directly from the general form of the solution (30) and the operational definition of the Hermite polynomials (39). Particular solutions for (13) with $f(x) = x^k$ and $\alpha,\ \beta = 0$, obviously follow from (74).

Now, let us consider the most general case even without specifying the type of the function f in the r.h.s. of (13) and the values of ν and α. We can still disentangle two integrals in (22) by involving Hermite polynomials of two variables (40) as follows:

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)} \sum_{n=0}^{\infty} \int_0^{\infty} \tau^{2(\nu-1)} \exp(-\beta^2\tau^2) H_n \left(\alpha, -\frac{1}{4\tau^2}\right) d\tau \cdot \frac{1}{n!} \int_{-\infty}^{\infty} (\eta - x)^n f(\eta) d\eta. \tag{75}$$

Thus, we obtain the solution of Eq. (13) as the series of the convolution given by

$$\phi(x) = \Phi(x) * f(\eta)$$

with the kernel

$$\Phi(x) = x^n$$

and the coefficient, given by the Hermite polynomials, as follows:

$$F(x) = \sum_{n=0}^{\infty} \phi(x)\ C(\nu, \alpha, \beta), \tag{76}$$

where

$$\phi(x) = \int_{-\infty}^{\infty} \Phi(x - \eta) f(\eta) d\eta \equiv \Phi(x) * f(\eta),$$
$$\Phi\ (x - \eta) = (\eta - x)^n,$$
$$\Phi(x) = (-x)^n$$

and

$$C(\nu, \alpha, \beta) = \frac{1}{\sqrt{\pi}\Gamma(\nu)} \int_0^{\infty} \tau^{2(\nu-1)} \exp(-\beta^2\tau^2) \frac{1}{n!} H_n(\alpha, -\frac{1}{4\tau^2}) d\tau.$$

Moreover, it follows directly from (22) and from the above convolution with account for the generating function (40), that the solution of (13) is given by the integral of the following convolution with the kernel, equal to the Gauss frequency function $\Omega(x, \tau)$:

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)} \int_0^{\infty} \tau^{2(\nu-1)}\, e^{-\beta^2\tau^2}\, \varphi(x, \tau) d\tau, \tag{77}$$

where

$$\varphi\ (x, \tau) = \int_{-\infty}^{\infty} \Omega\ (x - \eta, \tau) f(\eta) d\eta \equiv \Omega\ (x, \tau) * f\ (\eta)\,,$$
$$\Omega\ (x - \eta, \tau) = \exp\left(\alpha\ (\eta - x) - \frac{(\eta - x)^2}{4\tau^2}\right)$$

and

$$\Omega\,(x,\tau) = \exp\left[-\alpha x - \left(\frac{x}{2\tau}\right)^2\right].$$

Accounting for the integral given by

$$\int_0^\infty \tau^{2(\nu-1)}\,\mathrm{e}^{-(\beta\tau)^2-\frac{(x-\eta)^2}{4\tau^2}}\,d\tau = \left(\frac{|x-\nu|}{2\beta}\right)^{\nu-1/2} K_{\nu-\frac{1}{2}}(\beta\,|x-\eta|),$$

where $K_n\,(x)$ is the modified Bessel (or the Macdonald) function, we conclude that

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)}\int_{-\infty}^\infty \left(\frac{|x-\nu|}{2\beta}\right)^{\nu-1/2} K_{\nu-\frac{1}{2}}\,(\beta\,|x-\eta|)\,\mathrm{e}^{-\alpha(x-\eta)}f(\eta)d\eta. \quad (78)$$

Thus, following the general approach of [19], we have obtained the solution

$$F(x) = \left(\beta^2-(D+\alpha)^2\right)^{-\nu} f(x)$$

for the Eq. (13) in the form of the following convolution:

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma\,(\nu)}\int_{-\infty}^\infty \chi\,(x-\eta)f(\eta)d\eta, \quad (79)$$

where

$$\chi\,(x-\eta) = \left(\frac{|x-\nu|}{2\beta}\right)^{\nu-\frac{1}{2}} K_{\nu-\frac{1}{2}}\,(\beta\,|x-\eta|)\,\mathrm{e}^{-\alpha(x-\eta)}$$

with the kernel χ otherwise written as follows:

$$F(x) = \frac{1}{\sqrt{\pi}\Gamma(\nu)}\chi * f \qquad \text{and} \qquad \chi = \left(\frac{|x|}{2\beta}\right)^{\nu-1/2} K_{\nu-\frac{1}{2}}\,(\beta\,|x|)\,\mathrm{e}^{-\alpha x}. \quad (80)$$

Solutions of Some Partial Differential Equations by the Operational Technique

The method of the inverse differential operators has multiple applications for solving mathematical problems, describing wide range of physical processes, such as the heat transfer, the diffusion, wave propagation, etc. Some of the examples of solution of the heat equation, of the diffusion equation and of their modified forms, the Laguerre heat equation and others were considered by the inverse derivative method in [5, 15, 21]. In what follows solutions for more complicated generalized forms of the aforementioned equations will be explored, as well as some second order over the time variable partial differential equations will be touched upon.

It is worth mentioning that, despite the fact that relation (12) seems trivial to all appearance, it is very useful for solution of a broad family of differential equations by operational method. Indeed, for the differential equation:

$$\psi(D_x + \alpha)F(x,t) = f(x,t),$$

we can rewrite Eq. (12) in the following form:

$$\mathrm{e}^{\alpha x} F(x,t) = \psi^{-1}(D_x)\, \mathrm{e}^{\alpha x} f(x,t)$$

and, for example, for the evolutional type equations, where

$$f(x,t) = \partial_t F(x,t),$$

we obtain

$$\psi(D_x)\, \mathrm{e}^{\alpha x} F(x,t) = \partial_t\, \mathrm{e}^{\alpha x} F(x,t).$$

By writing

$$\mathrm{e}^{\alpha x} F(x,t) = G(x,t),$$

we have the following equation:

$$\psi(D_x)G(x,t) = \partial_t G(x,t)$$

with $\psi(D_x)$ and with the initial condition given by

$$g(x) = G(x,0) = \mathrm{e}^{\alpha x} F(x,0) = \mathrm{e}^{\alpha x} f(x).$$

Thus, in order to obtain the required solution:

$$F(x,t) = \mathrm{e}^{-\alpha x} G(x,t)$$

of the following equation:

$$\psi(D_x + \alpha)F(x,t) = f(x,t)$$

with the initial condition:

$$F(x,0) = f(x).$$

we end up with the necessity to solve the equation with $\psi(D_x)$ for the function $G(x, t)$ with the initial condition:

$$g(x) = \mathrm{e}^{\alpha x} f(x).$$

We note that the above-discussed method is applicable not only to the evolutional type equations with ∂_t in the right-hand side, but also to other operators $\hat{D}(t)$, acting over the time variable. Indeed, if $G(x, t)$ is the solution of the equation:

$$\psi(D_x)G(x, t) = \hat{D}(t)G(x, t)$$

with

$$g(x) = G(x, 0) = \mathrm{e}^{\alpha x} F(x, 0) = \mathrm{e}^{\alpha x} f(x),$$

then, following the above-described scheme, it is easy to demonstrate that

$$F(x, t) = \mathrm{e}^{-\alpha x} G(x, t)$$

is the solution of the equation

$$\psi(D_x + \alpha)F(x, t) = \hat{D}(t)F(x, t)$$

with

$$F(x, 0) = f(x).$$

Evidently, in the case of the second-order differential operator $\hat{D}(t)$, the second boundary or initial condition has to be chosen for the differential equation for $F(x, t)$ and, accordingly, for $G(x, t)$. In what follows, we shall apply the above-discussed method to several examples of equations, common in physics and other applications.

Black–Scholes Type Equations

In order to demonstrate the solution of differential equations by the operational method, we first consider the following differential equation, which is a generalized form of a Black–Scholes equation, frequently used in financial models:

$$\frac{1}{\rho}\frac{\partial}{\partial t}F(x, t) = \left(x^2\frac{\partial^2}{\partial x^2} + 2\alpha x^2\frac{\partial}{\partial x} + \lambda x\frac{\partial}{\partial x} + (\alpha x)^2 - \mu\right)F(x, t) \tag{81}$$

with

$$f(x) = F(x, 0),$$

where α, ρ, λ and μ are the constant coefficients and $f(x)$ is the initial condition function. Then the apparently complicated equation (81) reduces to the following form:

$$l\frac{1}{\rho}\frac{\partial}{\partial t}G(x,t) = x^2\frac{\partial^2}{\partial x^2}G(x,t) + \lambda x\frac{\partial}{\partial x}G(x,t) - \mu G(x,t) \qquad (82)$$

with

$$g(x) = G(x,0),$$

by the substitution given by

$$\partial_x \to \partial_x + \alpha.$$

Therefore, according to (12) and to the discussion at the beginning of this section, the solution of Eq. (81) will be found, if we obtain the solution of the Black–Scholes equation (82) for $G(x,t)$ with the initial condition function $g(x)$ given by

$$g(x) = G(x,0) = \mathrm{e}^{\alpha x}F(x,0).$$

Then, the solution of (81) reads as follows:

$$F(x,t) = \mathrm{e}^{-\alpha x}G(x,t) \qquad \text{with} \qquad g(x) = G(x,0) = \mathrm{e}^{\alpha x}F(x,0). \qquad (83)$$

The Eq. (82) can be easily solved with the help of the operational approach if we distinguish the perfect square of the operator $x\partial_x$ (see [10]):

$$G(x,t) = \frac{\exp(-\rho\varepsilon t)}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left(-\sigma^2 + \sigma\gamma\frac{\lambda}{2\rho}\right)g\left(x\,\mathrm{e}^{\sigma\gamma}\right)\,d\sigma, \qquad (84)$$

where

$$\gamma = \gamma(t) = 2\sqrt{\rho t} \qquad \text{and} \qquad \varepsilon = \mu + \left(\frac{\lambda}{2}\right)^2.$$

Let us choose the following initial condition:

$$f(x) = \mathrm{e}^{-\alpha x}x^n$$

for the Black–Scholes type equation (81), that is,

$$g(x) = x^n.$$

Then the solution (84) has the simple form (see [10]):

$$G(x,t) = x^n \exp[\rho t(n^2 + \lambda n - \mu)]$$

and Eq. (81) has the following solution:

$$F(x,t) = \mathrm{e}^{-\alpha x} x^n \exp[\rho t(n^2 + \lambda n - \mu)]. \tag{85}$$

Let us consider another generalization of the Black–Scholes type differential equation with the Laguerre derivative $\partial_x x \partial_x$ [see Eqs. (43) and (47)] instead of $x\partial_x$:

$$\frac{1}{\rho}\frac{\partial}{\partial t}A(x,t) = (\partial_x x \partial_x)^2 A(x,t) + \lambda(\partial_x x \partial_x)A(x,t) - \mu A(x,t) \tag{86}$$

with

$$g(x) = A(x,0),$$

where ρ, λ and μ are the constant coefficients and $g(x)$ is the initial condition function. The Eq. (86) generalizes and unifies equations of Laguerre diffusion of matter and of heat, considered in [8] and [10]. This equation can also be solved by the operational method developed above. Indeed, by distinguishing the perfect square of the Laguerre derivative

$${}_L D_x = \partial_x x \partial_x$$

in (86), the solution evidently reads in the form of the exponential:

$$A\,(x,t) = \exp\left[\rho t \left\{\left({}_L D_x + \frac{\lambda}{2}\right)^2 - \varepsilon\right\}\right] g(x),$$

where

$$\varepsilon = \mu + \left(\frac{\lambda}{2}\right)^2.$$

We now apply the operational identity (17) to $\exp(a_L D_x)$ to obtain the following solution for $A(x,t)$:

$$A\,(x,t) = \frac{\exp(-\varepsilon\alpha^2)}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-\sigma^2 - \sigma\alpha\lambda - 2\sigma\alpha_L D_x\right) g(x)\, d\sigma, \tag{87}$$

where

$$\alpha = \alpha(t) = \sqrt{\rho t}.$$

For a given function $g(x)$, we still have to find the result of the action of the operational exponent:

$$\exp(-a_L D_x) g(x)$$

and to take the integral

$$\int (\cdots) d\sigma.$$

Let us choose, for example, the initial condition function given by

$$g(x) = \frac{(-x)^n}{n!}.$$

Then, clearly, we can make use of the operational definition of the Laguerre polynomials (43) and obtain the following integral form for $A(x,t)$

$$A(x,t) = \frac{\exp\left(-\frac{\varepsilon\alpha^2}{4}\right)}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-\sigma^2 - \sigma\alpha\lambda\right) L_n(x, 2\sigma\alpha)\, d\sigma. \tag{88}$$

Further integration over $d\sigma$ yields the solution of the Black–Scholes like equation with the Laguerre derivative in (86) and with the initial condition given by

$$g(x) = \frac{(-x)^n}{n!}$$

in the following form:

$$A(x,t) = \frac{\exp(-\alpha^2\mu)}{\sqrt{\pi}} n! \sum_{r=0}^{n} \frac{(-x)^r (2\alpha)^{n-r}}{(n-r)!(r!)^2} \mathscr{I}, \tag{89}$$

where

$$\mathscr{I} := \frac{\alpha\lambda}{2}\left(e^{i(n-r)\pi} - 1\right) \Gamma\left(1 + \frac{n-r}{2}\right) {}_1F_1\left[\frac{1-(n-r)}{2}; \frac{3}{2}; -\left(\frac{\alpha\lambda}{2}\right)^2\right] + \frac{1}{2}\left(e^{i(n-r)\pi} + 1\right) \Gamma\left(\frac{1+n-r}{2}\right) {}_1F_1\left[-\frac{n-r}{2}; \frac{1}{2}; -\left(\frac{\alpha\lambda}{2}\right)^2\right]. \tag{90}$$

Here Γ denotes the (Euler's) gamma function and ${}_1F_1$ is the familiar confluent hypergeometric function. Evidently, if the initial condition function can be expanded in the power series of x^n, then the respective solution represents series of the already obtained solution (89). Moreover, if the expansion in series of the Laguerre polynomials for the initial condition function given by

$$g(x) = \sum_{n=0}^{\infty} a_n L_n(x)$$

exists, then we can exploit the relationships (51) and (44) to obtain the solution in the following form:

$$A\,(x,t) = \frac{\exp\left(-\frac{\varepsilon\alpha^2}{4}\right)}{\sqrt{\pi}} \sum_{n=0}^{\infty} a_n \int_{-\infty}^{\infty} \exp\left(-\sigma^2 - \sigma\alpha\lambda\right) L_n(x, 2\sigma\alpha + 1)\, d\sigma. \quad (91)$$

In the most general case, the solution $A\,(x,t)$ can be obtained through the following procedure: We employ the operational definitions (47) and the definition of the inverse derivative given by Eq. (2):

$$D_x^{-1} f(x) = \int_0^x f(\xi)\, d\xi$$

in order to write $A(x,t)$ in the form:

$$A\,(x,t) = \frac{\exp\left(-\varepsilon\alpha^2\right)}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-\sigma^2 - \sigma\alpha\lambda\right) \cdot \exp\left(-2\sigma\alpha \frac{\partial}{\partial D_x^{-1}}\right) \varphi(D_x^{-1})\, d\sigma, \quad (92)$$

where

$$\varphi(D_x^{-1})\{1\} = g(x)$$

and the explicit form of the function φ is given by the following integral:

$$\varphi(x) = \int_0^{\infty} \exp(-\kappa) g(x\kappa) d\kappa,$$

provided that this integral converges. We note that $\exp(-t_L D_x) g(x)$ is the solution of the Laguerre diffusion equation (see [10]):

$$\partial_t f(x,t) = -_L D_x f(x,t)$$

with the initial condition given by

$$f(x,0) = g(x).$$

Therefore, the result of the action of the exponential operator in (92) is, in fact, given by

$$f(x,t) = \exp\left(-t\,\frac{\partial}{\partial D_x^{-1}}\right) g(x) = \varphi\big(D_x^{-1} - t\big)\{1\},$$

which is the solution of the above-mentioned Laguerre diffusion equation. Then the required solution of Eq. (86) takes the following form:

$$A\,(x,t) = \frac{\exp(-\varepsilon\alpha^2)}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-\sigma^2 - \sigma\alpha\lambda\right) g(x,t) d\sigma, \tag{93}$$

where

$$g(x,t) = \varphi(D_x^{-1} - 2\sigma\alpha)\,1 = \exp\left(-2\sigma\alpha\frac{\partial}{\partial D_x^{-1}}\right)\varphi(D_x^{-1})\,1. \tag{94}$$

Consider the following initial condition:

$$g(x) = W_0(-x^2, 2),$$

where

$$W_n(x,m) = \sum_{r=0}^{\infty} \frac{x^r}{r!(mr+n)!} \qquad (m \in \mathbb{N};\ n \in \mathbb{N}_0)$$

is the particular case of the Bessel-Wright function (see [24]). The corresponding image function is

$$\varphi(x) = \exp(-x^2).$$

With account for (17) and (94) we obtain (see also [5])

$$g(x,t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-\xi^2 + 4i\sigma\alpha\xi\right) C_0(2i\xi x) d\xi, \tag{95}$$

where $C_n(x)$ given by

$$C_n(x) := \sum_{r=0}^{\infty} \frac{(-x)^r}{r!(r+n)!} \qquad (n \in \mathbb{N}_0)$$

is the Bessel-Tricomi function [27], which is related to the Bessel-Wright function:

$$C_n(x) = W_n(-x, 1)$$

and also to the commonly known cylindrical Bessel functions:

$$C_n(x) = x^{-n/2} J_n(2\sqrt{x}).$$

Thus, the solution (86) with the initial condition:

$$g(x) = W_0(-x^2, 2)$$

is explicitly determined by the formulas (93) and (95).

Heat Diffusion Type Equations

Let us first consider the following form of the Schrödinger equation:

$$i\,\partial_\tau \Psi(x,\tau) = -\partial_x^2 \Psi(x,\tau) + b\,x\,\Psi(x,\tau). \tag{96}$$

It can be solved by means of the operational method just as we did for Eq. (13) with the help of the heat-diffusion operator (25) and of the translation operator (31). Disentanglement of the operators in the exponential according to the rule (29) leads to the solution of the Schrödinger equation in the form of the sequence of the transformations of the initial function $f(x)$ under the action of the operators $\hat{\hat{S}}$ and $\hat{\hat{\Theta}}$:

$$\Psi(x,t) = \exp[-i\,\Phi(x,\tau;\,b)]\,\hat{\hat{\Theta}}\hat{\hat{S}} f(x), \tag{97}$$

where

$$\hat{\hat{S}} = \exp\left(i\,\tau\,\partial_x^2\right),$$

$$\hat{\hat{S}} f(x) = f(x, i\tau),$$

$$\hat{\hat{\Theta}} = \exp(b\,\tau^2 \partial_x)$$

and

$$\hat{\hat{\Theta}} f(x,\tau) = f(x + b\tau^2, \tau)$$

and Φ is the phase. The integral presentation of the solution reads as follows:

$$\Psi(x,t) = \exp(-i\,\Phi(x,\ \tau; b)) \frac{1}{2\sqrt{i\,\pi\,\tau}} \int_{-\infty}^{\infty} \exp\left(-\frac{\left(x + b\,\tau^2 - \xi\right)^2}{4\,i\,\tau}\right) f(\xi) d\xi. \tag{98}$$

We note that, in the solution (98), as well as in the formula (33), we did not make any suppositions about the initial function $f(x)$ and we obtained the solution as sequential action of the heat diffusion operator $\hat{S}$ and the translation operator $\hat{\Theta}$ on $f(x)$ (see also [22]). In this context $\hat{S}$ can be viewed as the evolution operator of a free particle. The solution given by (see [21])

$$\Psi(x,\tau) = \exp[-i\Phi(x,\tau;b)]Ai(x + b\,\tau^2, i\tau)$$

of the Schrödinger equation (96) with the Airy initial condition (see [26]):

$$f(x) = \mathrm{Ai}\left(\frac{x}{C}\right) = \frac{1}{\pi}\int_0^\infty \cos\left(\frac{1}{3}\zeta^3 + \frac{x}{C}\zeta\right) d\zeta,$$

where $C =$ constant, describes the dynamics of the Airy packet in a constant electrostatic field. We remind (see [40]) that the solution is reduced to the translation of the packet, so that its shape remains unchanged under the action of a constant force, which subjects the packet to constant acceleration. This apparently contradictory statement (see [2]) is resolved once we note that the Airy function is not square integrable and we cannot define its centroid

$$\langle x\rangle = \int_{-\infty}^{\infty} \frac{xf(x)}{\int_{-\infty}^{\infty} f(\xi)d\xi}\, dx.$$

In the context of physical applications it is important to underline that without the square integrability we cannot refer to the average of the coordinate and of other observables, while we can still talk about the dynamics of some abstract points, for example, about the acceleration of the point, in which the value of the Airy function equals zero.

The evolution of the Airy and of the Gauss packets, governed by the following generalized heat type equation with the linear term:

$$\partial_t F(x,t) = \alpha\partial_x^2 F(x,t) + \beta x F(x,t) \tag{99}$$

with the initial condition

$$F(x,0) = f(x)$$

and for $\alpha = 1$, was investigated in [40]. The solution of Eq. (99) goes along the lines of the solution of the Schrödinger equation and for $\alpha = 1$ we end up with

$$F(x,t) = \mathrm{e}^{\Phi(x,t;\beta)}\ \hat{\Theta}\hat{S}f(x) = \mathrm{e}^{\Phi(x,t;\beta)} f\left(x + \beta t^2, t\right), \tag{100}$$

where

$$\Phi(x,t;\beta) = \frac{1}{3}\beta^2 t^3 + \beta t x,$$

$$\hat{\Theta} = \exp\left(\beta t^2 \partial_x\right)$$

and

$$\hat{S} = \exp\left(t\partial_x^2\right).$$

The solution (100) consists in the action of the evolution operator on the initial condition:

$$F(x,0) = f(x),$$

which is transformed by $\hat{\hat{S}}$ and $\hat{\hat{\Theta}}$. For the Airy initial condition given by

$$f(x) = \mathrm{Ai}\left(\frac{x}{C}\right),$$

the operational method readily yields

$$\mathrm{Ai}\,(x,t) = \hat{S}\mathrm{Ai}\left(\frac{x}{C}\right) = \exp\left(t\partial_x^2\right)\mathrm{Ai}\left(\frac{x}{C}\right)$$

and the following solution of Eq. (99):

$$\begin{aligned} F(x,t) &= \mathrm{e}^{\Phi(x,t,\beta)}\,\mathrm{Ai}\left(x+\beta,t^2,t\right) \\ &= \mathrm{e}^{\Phi(x,t,\beta)}\,\frac{1}{\pi}\int_0^\infty \cos\left[\frac{1}{3}\left(\zeta^3 + \frac{\zeta(x+\beta t^2)}{C}\right)\right]\exp\left(-\frac{\zeta^2 t}{C^2}\right)d\zeta. \end{aligned} \quad (101)$$

Without going further into the details (see [40]), we just note here that the aforementioned solution (101) exhibits distinguished fading of the oscillations without any spread, while the evolution of the Gauss function lowers the curve and spreads the packet, too.

Let us choose the initial condition given by

$$f\,(x) = x^n$$

for the heat conduction type equation (99). Then, upon the action of the heat diffusion operator on it and according to the following operational definition of the Hermite polynomials (39):

$$\exp\left(a\partial_x^2\right)x^n = H_n\,(x,a)\,,$$

we obtain the solution:

$$F(x) \propto H_n\,(x+ab,a)\,,$$

where $a = \alpha t$ and $b = \beta t$, and we end up with

$$F\,(x,t) = \mathrm{e}^{\Phi}\,H_n\left(x+\alpha\beta t^2,\alpha t\right). \quad (102)$$

Evidently, if the initial function can be expanded in series as follows:

$$f(x) = \sum_{n=0}^{\infty} c_n x^n,$$

then the solution appears in the form of the series:

$$F(x,t) = \mathrm{e}^{\Phi} \sum_{n=0}^{\infty} c_n H_n\left(x + \alpha\beta t^2, \alpha t\right).$$

Let us now choose another initial condition such as (for example)

$$f(x) = x^k\,\mathrm{e}^{\delta x}.$$

Then, by virtue of the operational rule (72), we obtain

$$\hat{S} f(x) = \exp\left[\delta\left(x + \delta a\right)\right] H_k\left(x + 2\delta a, a\right) = f(x,t).$$

The consequent action of the translation operator $\hat{\Theta}$ yields the shift along the argument x and, thus, we obtain the required solution in the following form:

$$F(x,t) = \mathrm{e}^{\Phi + \Delta_1} H_k\left(x + \frac{\alpha\delta^2}{\beta}\left(\frac{2t\beta}{\delta} + t^2\left(\frac{\beta}{\delta}\right)^2\right), \alpha t\right), \tag{103}$$

where

$$\Delta_1 = \delta\left(x + \delta\alpha t + \alpha\beta t^2\right).$$

For $\delta = 0$, it immediately leads us to the result (102). Studying the evolution at prolonged times, such that $t \gg \delta/\beta$, we notice that $\Delta_1 \ll \Phi$ and, provided that (at so long times) the special condition $x \ll \delta^2\alpha/\beta$ is also fulfilled, that is the coordinate travel is limited, we end up with the separation of the dependence of the solution on time and coordinate. The coordinate dependence is all given by the exponential factor $\exp(\beta t x)$, while the time is contained in the Hermite polynomial arguments $H_k\left(2\delta\alpha t + \alpha\beta t^2, \alpha t\right)$. For the short times of the evolution of the system, such that $t \ll \delta/\beta$ the phase approximately reads as follows:

$$\Phi + \Delta_1 \cong x\delta + \alpha\delta^2 t$$

and the Hermite polynomials depend on both coordinate and time: $H_k(x, \alpha t)$. Thus, for relatively short times we have the solution approximated by

$$F(x,t)\big|_{t \ll \delta/\beta} \cong \mathrm{e}^{x\delta + \alpha\delta^2 t} H_k(x, \alpha t)$$

and for infinitely short times $\alpha t \to 0$ the Hermite polynomials become

$$H_k(x,0) = x^k,$$

which is in perfect agreement with our initial condition:

$$f(x) = x^k\, e^{\delta x}.$$

Let us now obtain the solution for the following equation:

$$\partial_t F(x,t) = \partial_x^2 F(x,t) + 2\delta\partial_x F(x,t) + \beta x F(x,t) + \gamma F(x,t) \tag{104}$$

with

$$F(x,0) = f(x),$$

which can be considered as the generalization of the Eq. (99) upon the substitution of the derivative $\partial_x \to \partial_x + \delta$, where we set $\alpha = 1$ without substantial sacrifice of the generality. Distinguishing the perfect square of the operator $\partial_x + \delta$ and following the general lines for solution of such equations, developed in the beginning of this chapter, we can search for the desired solution of Eq. (104) in the following form:

$$F(x,t) = \exp\left[t\left(\gamma - \delta^2\right) - \delta x\right] G(x,t), \tag{105}$$

where $G(x,t)$ satisfies Eq. (99) for G with the initial condition:

$$G(x,0) = g(x) \qquad \text{and} \qquad g(x) = \exp(\delta x) f(x).$$

Let us choose the initial condition for (104), for example, in the form of the powers:

$$f(x) = x^k.$$

Then

$$g(x) = x^k\, e^{\delta x}$$

and the solution of the equation:

$$\partial_t G = (\partial_x^2 + \beta x)G$$

is given by (103) upon the substitution of $F \to G$. Eventually, we end up with the desired solution of the Eq. (104), written in terms of the Hermite polynomials as follows:

$$F(x,t) = e^{\Phi + \Delta_2} H_k(x + 2t\delta + t^2\beta, t), \tag{106}$$

with the additional phase given by

$$\Delta_2 = t\gamma + t^2\delta\beta.$$

We note that the generalization of the above considered equations of the heat conduction, unifying both the Laguerre diffusion and the Laguerre heat conduction equations, is, in essence, the Black–Scholes equation with the Laguerre derivative, which was considered above in the beginning of the present chapter.

Eventually, let us complement the above considered heat diffusion type equations by the following two-dimensional heat conduction equation with the linear terms:

$$\partial_t F(x,y,t) = \left\{\left(\alpha\partial_x^2 + \beta\partial_x\partial_y + \gamma\partial_y^2\right) + bx + cy\right\} F(x,y,t) \tag{107}$$

$$(\min\{\alpha,\beta,\gamma\} > 0)$$

and the initial condition:

$$F(x,y,0) = f(x,y).$$

It can be solved in the way, similar to that in [31, 40], where one-dimensional case was studied, or with the help of the famous Baker–Campbell–Hausdorf formula:

$$\exp\left(\hat{A} + \hat{B}\right) = \exp\left(-\frac{[\hat{A},\hat{B}]}{2}\right)\exp\left(\hat{A}\right)\exp\left(\hat{B}\right).$$

We apply it to the operators given by

$$\hat{A} = \alpha\partial_x^2 + \beta\partial_x\partial_y + \gamma\partial_y^2$$

and

$$\hat{B} = bx + cy,$$

which are distinguished in (107), and we make consequent use of the operational identities, such as

$$e^{\beta\partial_x\partial_y}\, e^{bx+cy} = e^{\beta b\partial_y + \beta c\partial_x}\, e^{bx+cy}\, e^{\beta\partial_x\partial_y},$$

$$e^{\alpha\partial_x^2}\, e^{\beta x}\, g(x,y) = e^{b^2\alpha}\, e^{bx}\, e^{2b\alpha\partial_x}\, e^{\alpha\partial_x^2}$$

and

$$e^{\gamma\partial_x}\, e^{\alpha x}\, g(x) = e^{\alpha x + \alpha\gamma}\, e^{\gamma\partial_x}\, g(x),$$

to obtain (after some manipulations with the operators in the exponential) the following two-dimensional generalization of the solution (100):

$$F(x,y,t) = e^{\Psi}\,\hat{\Theta}_x\hat{\Theta}_y\hat{\mathrm{E}}f(x,y), \tag{108}$$

where

$$\Psi = (\alpha b^2 + \gamma c^2 + \beta bc)t^3/3 + t(bx + cy)$$

is the phase,

$$\hat{\Theta}_x = \mathrm{e}^{t^2(\alpha b+\beta c/2)\partial_x}$$

and

$$\hat{\Theta}_y = \mathrm{e}^{t^2(\gamma c+\beta b/2)\partial_y}$$

are the diffusion operators for each of the two coordinates and

$$\hat{\mathrm{E}} = \exp\left[t\left(\alpha\partial_x^2 + \beta\partial_x\partial_y + \gamma\partial_y^2\right)\right] \tag{109}$$

represents the heat diffusion operator for the two-dimensional case, analogous to $\hat{S}$ operator in (25) in the one-dimensional case. These operators consequently transform the initial function $f(x, y)$ exactly as in the one-dimensional case:

$$F(x,t) \propto \hat{\Theta}\hat{S}f(x)\,.$$

The result of the action of the heat diffusion operator $\hat{\mathrm{E}}$ on the initial condition is given by

$$f(x,y,t) = \hat{\mathrm{E}}f(x,y)$$

and, consequently, the commuting diffusion operators $\hat{\Theta}_x$ and $\hat{\Theta}_y$ shift the argument of the function f as follows:

$$\hat{\Theta}_x\hat{\Theta}_y f(x,y,t) = f\left(x + t^2\left(\alpha b + \frac{\beta c}{2}\right), y + t^2\left(\gamma c + \frac{\beta b}{2}\right), t\right).$$

Thus, in complete analogy with the one-dimensional case, we obtain the solution of the two-dimensional heat conduction equation with linear terms (107) in the following form:

$$\begin{aligned} F(x,y,t) &= \mathrm{e}^{\Psi}\ \hat{\Theta}_x\hat{\Theta}_y\hat{\mathrm{E}}f(x,y) \\ &\propto f\left(x + t^2\left(\alpha b + \frac{\beta c}{2}\right), y + t^2\left(\gamma c + \frac{\beta b}{2}\right), t\right). \end{aligned} \tag{110}$$

The explicit double integral form of the operator $\hat{\mathrm{E}}$ was obtained in [15] and it is of the Gauss type integral, which we omit here for the sake of conciseness. It is easy to demonstrate that in the case of $\beta = 0$ the heat diffusion is executed by the one-dimensional operators (25) $\hat{S}_x\hat{S}_y$ instead of the more general operator $\hat{\mathrm{E}}$. The solution in this case reads as follows:

$$F(x,y,t) = \mathrm{e}^{\Psi}\ \hat{\Theta}_x\hat{\Theta}_y\hat{S}_x\hat{S}_y f\,(x,y) \propto f\left(x + t^2\alpha b, y + t^2\gamma c, t\right). \tag{111}$$

We note that the operational definition (109) allows for the direct computation of the result in many cases, avoiding the necessity to calculate the double integral for the operator $\hat{\mathrm{E}}$. Let us, for example, choose the initial function in the form of the powers of x and y:

$$f\,(x,y) = x^m y^n.$$

Then, according to the operational definition of the Hermite polynomials of four variables and two indices $H_{m,n}\,(x, t\alpha, y, t\gamma|\,\beta)$ (see, for example, [11, 15] and [10]), we obtain

$$\hat{\mathrm{E}}\,\{x^m y^n\} = H_{m,n}\,(x, t\alpha, y, t\gamma|\,t\beta)\,, \tag{112}$$

where $H_{m,n}\,(x, t\alpha, y, t\gamma|\,\beta)$ are the above-mentioned Hermite polynomials with the following generating exponent:

$$\sum_{m,n}^{\infty} \frac{u^m v^n}{m!n!} H_{m,n}\,(x, \alpha, y, \gamma|\,\beta) = \exp\left(xu + \alpha u^2 + yv + \gamma v^2 + \beta uv\right). \tag{113}$$

The presentation of $H_{m,n}\,(x, t\alpha, y, t\gamma|\,\beta)$ in the form of sums (see [15]) of the two-variable Hermite polynomials $H_m\,(x,y)$, defined in (39), reads as follows:

$$H_{m,n}\,(x,\alpha; y, \gamma|\,\beta) = m!n! \sum_{s=0}^{\min(m,n)} \frac{\beta^s}{s!(m-s)!(n-s)!} H_{m-s}\,(x,\alpha)\,H_{n-s}\,(y,\gamma)\,. \tag{114}$$

The action of the translation operators $\hat{\Theta}_x\hat{\Theta}_y$ on the Hermite polynomials

$$H_{m,n}\,(x, t\alpha, y, t\gamma|\,\beta)$$

yields the solution of the two-dimensional heat type equation, with the linear terms (107) and with the initial condition in the form of powers given by

$$f\,(x,y) = x^m y^n,$$

as follows:

$$F\,(x,t) = \mathrm{e}^{\Psi}\, H_{m,n}\left(x + t^2\left(\alpha b + \frac{\beta c}{2}\right), t\alpha; y + t^2\left(\gamma c + \frac{\beta b}{2}\right), t\gamma\,\middle|\, t\beta\right). \tag{115}$$

It appears evident that the obtained solution (115) of the two-dimensional heat conduction problem (107) represents a direct generalization of the solution (102) for the one-dimensional heat conduction analog (99).

Fokker-Planck Type Equations

The equations of the Fokker-Planck type are frequent in many physical problems. It is enough to mention the propagation of beams of charged particles in accelerators and in insertion devices, such as undulators, and the relevant problems in FEL modeling. For example, the following operator:

$$\left(\frac{2t}{\tau}\right)\left(\sigma_{\varepsilon}^{2}\partial_{x}^{2}+\partial_{x}x\right)$$

appears, when the electron beams in accelerators and in storage rings are modeled with account for the diffusive and for the damping effects in them, where τ is the typical damping time of the electron beam due to the synchrotron radiation of the electrons in the bending magnets and σ_{ε} is their deviation from the so-called uniform distribution. The diffusion and the damping processes, acting together, equilibrate each other with time passing, and, thus, the stationary solution eventually appears. The evolution operator for the above-mentioned problem consists of the operators

$$\hat{A}=\left(\frac{2t}{\tau}\right)\sigma_{\varepsilon}^{2}\partial_{x}^{2}$$

and

$$\hat{B}=\left(\frac{2t}{\tau}\right)\partial_{x}x.$$

Their commutator reads as follows:

$$\left[\hat{A},\hat{B}\right]=\left(\frac{4t}{\tau}\right)\sigma_{\varepsilon}^{2}\hat{A}=m\hat{A}$$

and (upon their ordering in the exponential) we obtain

$$U(t)=\exp\left(\hat{A}+\hat{B}\right)=\exp\left(\frac{1-\exp(-m)}{m}\hat{A}\right)\exp(\hat{B}). \tag{116}$$

After the above-considered examples, the operational solution of the following one-dimensional Fokker–Plank equation:

$$\partial_{t}F(x,t)=\alpha\partial_{x}^{2}F(x,t)+\beta x\partial_{x}F(x,t),F\left(x,0\right)=f\left(x\right) \tag{117}$$

seems elementary. Indeed, we construct

$$F(x) = \hat{U} f(x), \hat{U} f(x) = \exp\left(t\alpha\partial_x^2 + t\beta x\partial_x\right) f(x) = \exp\left(\sigma\partial_x^2\right) f\left(\mathrm{e}^{\beta t} x\right), \quad (118)$$

where

$$\sigma = \left(1 - \mathrm{e}^{-2\beta t}\right) \frac{\alpha}{\beta},$$

and we proceed along the lines of the solution of Schrödinger equation. Upon the trivial change of variables, we obtain

$$\hat{U} f(x) = \hat{S} f(y),$$
$$y = x \exp(b)$$

and

$$\hat{S} = \exp\left(\frac{\rho}{2}\partial_y^2\right),$$

where

$$\rho(t) = \frac{\alpha}{\beta}\left(\mathrm{e}^{2\beta t} - 1\right)$$

and we end up with the following simple solution of the Eq. (117):

$$F(x,t) = \frac{1}{\sqrt{2\pi\rho}} \int_{-\infty}^{\infty} \exp\left(-\frac{\left(\mathrm{e}^{\beta t}\; x - \xi\right)^2}{2\rho}\right) f(\xi) d\xi. \quad (119)$$

For example, for the initial Gaussian distribution given by

$$f(x) = \exp\left(-x^2\right)$$

by means of the Gauss transforms, we obtain the solution in the following form:

$$F(x,t)|_{f(x)=\exp\left(-x^2\right)} = \frac{1}{\sqrt{1 + 2\rho(t)}} \exp\left(-\frac{\mathrm{e}^{2\beta t}\, x^2}{1 + 2\rho(t)}\right), \quad (120)$$

where

$$\rho(t) = \frac{\alpha}{\beta}\left(\mathrm{e}^{2\beta t} - 1\right).$$

The solution of the following equation very similar to (117):

$$\partial_t F(x,t) = \alpha\partial_x^2 F(x,t) + \beta\partial_x x F(x,t) \quad (121)$$

differs from (120) just by the common phase factor. For the initial Gaussian function

$$f(x) = \exp\left(-x^2\right),$$

common for charged particle beams in accelerators, the solution of the Eq. (121) reads as follows:

$$\begin{aligned} F(x,t)|_{f(x)=\exp\left(-x^2\right)} &= \frac{e^{\beta t}}{\sqrt{1+2\rho(t)}} \exp\left(-\frac{e^{2\beta t}\, x^2}{1+2\rho(t)}\right) \\ &= \frac{1}{\sqrt{\eta(t)}} \exp\left(-\frac{x^2}{\eta(t)}\right), \end{aligned} \tag{122}$$

where

$$\eta(t) = \frac{2\alpha}{\beta}\left(1 - e^{-2\beta t} + \frac{\beta}{2\alpha}\, e^{-2\beta t}\right).$$

Differently from the Schrödinger equation solution, the function f in the Fokker–Plank type equation solutions is transformed by the sole operator $\hat{S}$ [compare Eqs. (118) and (122) with Eqs. (100) and (97)]. Moreover, the dependence on time is more complicated in the Fokker–Plank type solutions.

Let us consider the following generalization of the Fokker–Plank type equation:

$$\partial_t F(x,t) = \left[\alpha \partial_x^2 + (\beta x + 2\alpha\delta)\partial_x + \beta\delta x + \gamma\right] F(x,t) \tag{123}$$

with the initial condition of the Gaussian type:

$$f(x) = \exp\left(-x^2\right).$$

By distinguishing the operator $\partial_x + \delta$, we conclude that the solution of the generalized Fokker–Plank equation (123) reduces to the solution of Eq. (117) for the function G with the initial condition:

$$g\,(x) = G\,(x,0) = e^{\delta x} f\,(x) = \exp\left(\delta x - x^2\right).$$

In its turn, the solution for $G\,(x,t)$ can be obtained through the following Gauss transforms:

$$G\,(x,t) = \frac{1}{\sqrt{2\pi\rho}} \int_{-\infty}^{\infty} g\,(\xi) \exp\left[-\left(\frac{e^{\beta t}\, x - \xi}{\sqrt{2\rho}}\right)^2\right] d\xi.$$

Then, for

$$F\,(x,t) = e^{\gamma - \alpha\delta^2}\, e^{-\delta x}\, G\,(x,t),$$

we readily obtain the solution of Eq. (123) in the following form:

$$F(x,t)|_{f(x)=\exp\left(-x^2\right)}$$
$$= \frac{\exp\left[\gamma - \left(\frac{\delta}{2}\right)^2 \ (4\alpha - 1) - \delta x\right]}{\sqrt{1 + 2\rho(t)}} \ \exp\left(-\frac{\left(e^{\beta t}\, x - \delta/2\right)^2}{1 + 2\rho(t)}\right). \tag{124}$$

Operational Solutions for Some Second Order of Time Partial Differential Equations

Some Second Order of Time Partial Differential Equations with Constant Coefficients

The operational method for solution of differential equations can be successfully applied for the second order of time partial differential equations as well. Let us consider the equations of the following type:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon \frac{\partial}{\partial t}\right) F\,(x,t) = \hat{D}\,(x)\, F\,(x,t)\,, \tag{125}$$

where $\hat{D}\,(x)$ is a differential operator, acting over the coordinate, such as, the heat diffusion operator ∂_x^2 or the Laguerre derivative ${}_LD_x$ or any other. General solution of Eq. (125) reads as follows:

$$F\,(x,t) = e^{-\frac{t\varepsilon}{2}} \left(e^{-\frac{t}{2}\sqrt{\varepsilon^2 + 4\hat{D}(x)}}\, C_1\,(x) + e^{\frac{t}{2}\sqrt{\varepsilon^2 + 4\hat{D}(x)}}\, C_2\,(x)\right), \tag{126}$$

where $C_{1,2}\,(x)$ are to be determined from the initial conditions. Suppose that the initial condition

$$F\,(x,0) = f\,(x)$$

is given and, for example, we can require for the second-order equation (125) that its solution converges at infinite time $t \to \infty$, which is reasonable for physical applications. Other initial conditions are, of course, possible, but they will be considered elsewhere. The above choice sets

$$C_2\,(x) = 0$$

and the remaining branch of the solution is subject to the Laplace transforms as follows:

$$e^{-t\sqrt{V}} = \frac{t}{2\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}}\, e^{-\frac{t^2}{4\xi}-\xi V}, \quad t > 0. \tag{127}$$

Thus, we obtain for Eq. (125) the following fading at infinite time solution:

$$F(x,t) = e^{-\frac{\varepsilon t}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}}\, e^{-\frac{t^2}{16\xi}}\, e^{-\xi\varepsilon^2}\, e^{-4\xi\hat{D}(x)} f(x), \tag{128}$$

provided that the integral converges. The particular form of the solution depends on the operator $\hat{D}(x)$ and on the initial function $f(x)$. Let us now consider a few examples.

Laguerre-Type Diffusion with the Second-Order Time Derivative

We first choose

$$\hat{D}(x) = \partial_x x\, \partial_x \equiv_L D_x,$$

so that we seek a solution of the following equation:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon \frac{\partial}{\partial t}\right) F(x,t) = \frac{\partial}{\partial x} x \frac{\partial}{\partial x} F(x,t), \tag{129}$$

where

$$F(x,0) = f(x)$$

and

$$F(x,\infty) < \infty.$$

To this end, we have to calculate the result of the action of the exponential Laguerre derivative operator on the initial function:

$$e^{-4\xi_L D_x} f(x).$$

The simplest example of the initial function, for which the integral surely converges, is, perhaps, the power function

$$f(x) = \frac{(-x)^n}{n!}.$$

Then

$$e^{-4\xi_L D_x} f(x) = L_n(x, 4\xi)$$

and the fading at $t \to \infty$ solution of the second order of time Laguerre diffusion equation (129) with the initial condition:

$$f(x) = \frac{(-x)^n}{n!}$$

reads as follows:

$$F(x,t) = \mathrm{e}^{-\frac{\varepsilon t}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \mathrm{e}^{-\frac{t^2}{16\xi}-\xi\varepsilon^2} L_n(x, 4\xi). \tag{130}$$

Further calculations can be done with account for the series presentation (43), while the machine computations favour the form (44) with one argument of $L_n(z)$. Calculation of the integral in

$$\frac{t}{4}\int_0^\infty \frac{(4u)^m\, du}{u\sqrt{u}} \exp\left(-\frac{t^2}{16u} - u\varepsilon^2\right) = \sqrt{t\varepsilon}\left(\frac{t}{\varepsilon}\right)^m K_{\frac{1}{2}-m}\left(\frac{t\varepsilon}{2}\right), \tag{131}$$

where $K_n(y)$ are the modified Bessel functions of the second kind, yields the following solution:

$$F(x,t) = \mathrm{e}^{-\frac{t\varepsilon}{2}} \sqrt{\frac{t\varepsilon}{\pi}}\, n! \sum_{r=0}^{n} \frac{(-x)^r}{(n-r)!\,(r!)^2} \left(\frac{t}{\varepsilon}\right)^{n-r} K_{\frac{1}{2}-(n-r)}\left(\frac{t\varepsilon}{2}\right). \tag{132}$$

The above sum can be easily computed, since the Bessel functions with half-integer indices have explicit expression in terms of elementary functions:

$$K_\nu(z) = \sqrt{\frac{\pi}{2}} \frac{\mathrm{e}^{-z}}{\sqrt{z}} \sum_{j=0}^{[|\nu|-1/2]} \frac{(j+|\nu|-1/2)!}{j!(-j+|\nu|-1/2)!} (2z)^{-j} \qquad \left(\nu - \frac{1}{2} \in \mathbb{Z}\right). \tag{133}$$

For example, for the initial function

$$f(x) = x^2$$

and for

$$F(x,\infty) < \infty,$$

we obtain the following solution of Eq. (129):

$$F(x,t)|_{n=2} = \mathrm{e}^{-t\varepsilon}\left(2t^2\varepsilon + x^2\varepsilon^3 + 4t\left(1 - x\varepsilon^2\right)\right)/2\varepsilon^3. \tag{134}$$

We note that (134) diverges for $\varepsilon \to 0$:

$$\lim_{\varepsilon \to 0} F(x,t)\Big|_{n=2} \propto (1/0).$$

Evidently, in the case of $\varepsilon = 0$, we have the following equation:

$$\frac{\partial^2}{\partial t^2} F(x,t) = \hat{D}(x) F(x,t) \tag{135}$$

and, for the differential operator $\hat{D}$ given by

$$\hat{D}(x) = {}_L D_x,$$

it reads as follows:

$$\frac{\partial^2}{\partial t^2} F(x,t) = \partial_x x \partial_x F(x,t). \tag{136}$$

In order to solve it, we compute the following integral:

$$\frac{t}{2}\int_0^\infty \frac{u^m du}{u\sqrt{u}} \exp\left(-\frac{t^2}{4u}\right) = 4^{-m} t^{2m} \Gamma\left(\frac{1}{2} - m\right),$$

arising instead of (131). With its help, we obtain the solution of (136) for the initial function given by

$$f(x) = x^n$$

as follows:

$$\begin{aligned} F(x,t) &= \frac{n!}{\sqrt{\pi}} \sum_{r=0}^{n} \frac{(-x)^r\, 4^{-(n-r)} t^{2(n-r)} \Gamma\left[1/2 - (n-r)\right]}{(n-r)!\,(r!)^2} \\ &= n! \frac{(-x)^n\, C_{2n}^{(-2n)}\left(\frac{t}{2\sqrt{x}}\right)}{\Gamma(1+2n)}, \end{aligned} \tag{137}$$

where $C_n^k(x)$ is the Gegenbauer polynomial of degree k in x. Formula (137) yields the following particular result for $n = 2$:

$$F(x,t)|_{n=2} = \frac{1}{12}\left(t^4 + 12t^2 x + 6x^2\right).$$

More scrupulous study may consist, for example, in defining a new variable εt and then seeking a limit of $\varepsilon \to 0$ such that

$$\varepsilon t|_{t\to\infty} = \text{constant}.$$

Black–Scholes Type Second Order of Time Differential Equations

Let us consider the Black–Scholes type differential operator in the right-hand side of the second-order differential equation, namely

$$\left(\frac{\partial^2}{\partial t^2}+\varepsilon\frac{\partial}{\partial t}\right)G(x,t)=\left(x^2\frac{\partial^2}{\partial x^2}+\lambda x\frac{\partial}{\partial x}-\mu\right)G(x,t), \tag{138}$$

where

$$G(x,0)=g(x)$$

and

$$G(x,\infty)<\infty.$$

Distinguishing the full square of the operator $\bar{D}$ given by

$$\bar{D}=x\partial_x,$$

we have to compute the result of the action of the exponential operator:

$$\exp\left[-4\xi\left(\left(\bar{D}+\frac{\lambda}{2}\right)^2-\nu\right)\right]g(x),$$

where

$$\nu=\mu+\left(\frac{\lambda}{2}\right)^2.$$

In order to achieve it, we exploit the operational identity (17), by applying it to the exponential operator:

$$\exp\left[\left(\bar{D}+\frac{\lambda}{2}\right)^2\right].$$

We thus obtain

$$\mathrm{e}^{-4\xi\left(\bar{D}+\frac{\lambda}{2}\right)^2}=\int_{-\infty}^{\infty}\exp\left(-u^2+2iu\lambda\sqrt{\xi}+4iu\sqrt{\xi}\,\bar{D}\right)\frac{du}{\sqrt{\pi}}.$$

Upon the action on $g(x)$, we get

$$\exp\left(4iu\sqrt{\xi}x\partial_x\right)g(x)=g\left(e^{4iu\sqrt{\xi}}x\right),$$

which yields the function:

$$g(x,\xi) = \mathrm{e}^{-4\xi\left(\bar{D}+\frac{\lambda}{2}\right)^2} f(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-u^2+2iu\lambda\sqrt{\xi}}\, g\left(\mathrm{e}^{4iu\sqrt{\xi}}\, x\right) du, \tag{139}$$

and the solution of Eq. (138) now takes the following form:

$$G(x,t) = \mathrm{e}^{-t\varepsilon/2} \frac{t}{4\sqrt{\pi}} \int_0^{\infty} \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi} - \varepsilon^2\xi + 4\xi\nu\right) g(x,\xi). \tag{140}$$

Eventually, we consider the following rather complicated differential equation of the second order in time and coordinate:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon\frac{\partial}{\partial t}\right) F(x,t) = \left(x^2\frac{\partial^2}{\partial x^2} + 2\alpha x^2\frac{\partial}{\partial x} + \lambda x\frac{\partial}{\partial x} + (\alpha x)^2 - \mu\right) F(x,t). \tag{141}$$

With the initial condition:

$$F(x,0) = f(x),$$

we seek non-diverging at infinite times solution:

$$F(x,\infty) < \infty.$$

The solution arises directly from (140) and (139) and from (138). By noting that Eq. (141) is, in fact, Eq. (138) with $\partial_x + \alpha$ instead of ∂_x, we write our solution

$$F(x,t) = \mathrm{e}^{-\alpha x}\, G(x,t),$$

where $G(x,t)$ satisfies Eq. (138) with the initial condition:

$$g(x) = G(x,0) = \mathrm{e}^{\alpha x}\, F(x,0) = \mathrm{e}^{\alpha x} f(x),$$

and we demand

$$F(x,\infty) < \infty \qquad \text{and} \qquad G(x,\infty) < \infty,$$

respectively. Now the expression (140) for $G(x,t)$ with account for (139), where

$$g(x) = \mathrm{e}^{\alpha x} f(x),$$

provides our solution

$$F(x,t) = \mathrm{e}^{-\alpha x}\, G(x,t).$$

Consider the simple example of the initial function:

$$f(x) = \mathrm{e}^{-\alpha x} x^n,$$

which illustrates the above-described technique. It returns $g(x) = x^n$ and we easily obtain upon the integration over du and $d\xi$ the function $G(x,t)$ given by

$$G(x,t) = x^n \exp\left[-\frac{t}{2}\left(\varepsilon + \sqrt{\varepsilon^2 + 4\left[n(n-1+\lambda)\right] - \mu}\right)\right]. \tag{142}$$

It, in turn, yields the desired solution:

$$F(x,t)|_{t\to\infty} < \infty$$

of (141) with

$$f(x) = \mathrm{e}^{-\alpha x} x^n$$

as follows:

$$F(x,t) = x^n \exp\left[-\alpha x - \frac{t}{2}\left(\varepsilon + \sqrt{\varepsilon^2 + 4\left[n(n-1+\lambda)\right] - \mu}\right)\right]. \tag{143}$$

The solution for the particular case $\varepsilon = 0$ reads as follows:

$$F(x,t)|_{\varepsilon=0} = x^n \,\mathrm{e}^{-\alpha x - t\sqrt{n(n-1+\lambda)-\mu/4}}.$$

Hyperbolic Heat Conduction Equation

Another example of the operator $\hat{D}(x)$ in the right-hand side of the Eq. (125) is given by the following telegraph type equation with the first order of time derivative term, known also as a hyperbolic heat conduction equation:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon\frac{\partial}{\partial t}\right) F(x,t) = \left(\alpha\frac{\partial^2}{\partial x^2} + \kappa\right) F(x,t), \tag{144}$$

where

$$F(x,0) = f(x) \qquad \text{and} \qquad F(x,\infty) < \infty.$$

The non-diverging solution for the initial function $f(x)$ evidently reads as follows:

$$F(x,t) = \mathrm{e}^{-\frac{\varepsilon t}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi} - \xi\left(\varepsilon^2 + 4\kappa\right)\right) \mathrm{e}^{-4\alpha\xi\partial_x^2} f(x). \tag{145}$$

In this case, we have to compute the result of the action of the operator $\hat{S}$:

$$e^{-4\alpha\xi\partial_x^2} f(x),$$

which can be accomplished with the help of the identity (17), resulting in

$$F(x,t) = e^{-\frac{\varepsilon t}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{du}{u\sqrt{u}} \exp\left(-\frac{t^2}{16u} - u\left(\varepsilon^2 + 4\kappa\right)\right) \cdot \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-v^2} f\left(x + 2iv\sqrt{\zeta}\right) dv, \tag{146}$$

where $\zeta = 4u\alpha$. For example, for the initial function:

$$f(x) = e^{\gamma x} x^n,$$

we make use of the operational identity (72) to obtain the exact form of the solution:

$$lF(x,t) = \frac{t\, e^{-\frac{t\varepsilon}{2}+\gamma x}}{4\sqrt{\pi}} \int_0^\infty \frac{H_n(x - 2\gamma\zeta, -\zeta)}{u\sqrt{u}} \exp\left(-\frac{t^2}{16u} - u\delta\right) du, \tag{147}$$

where

$$\delta = \varepsilon^2 + 4\left(\kappa + \alpha\gamma^2\right).$$

The Hermite polynomials possess the sum-presentation (39), which allows the integration, but the result is rather cumbersome and we omit it for brevity. The simpler example of

$$f(x) = x^n$$

produces more compact Hermite polynomials upon the action of the operator $\hat{S}$ on it according to the operational definition (39). Thus, the solution of the Eq. (144) reads as follows:

$$\begin{aligned} F(x,t) &= e^{-\frac{t\varepsilon}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{H_n(x, -4\alpha u)}{u\sqrt{u}} \exp\left(-\frac{t^2}{16u} - u\left(\varepsilon^2 + 4\kappa\right)\right) du \\ &= e^{-\frac{t\varepsilon}{2}} \sqrt{\frac{t\sqrt{\varepsilon^2 + 4\kappa}}{\pi}}\, n! \sum_{r=0}^{[n/2]} \frac{(-\alpha)^r (x)^{n-2r}}{(n-r)!\, r!} \left(\frac{t}{\sqrt{\varepsilon^2 + 4\kappa}}\right)^r \\ &\quad \cdot K_{\frac{1}{2}-r}\left(\frac{t}{2}\sqrt{\varepsilon^2 + 4\kappa}\right). \end{aligned} \tag{148}$$

For $\alpha = 1, \kappa = 0$ and $n \in \mathbb{N}\backslash\{0\} = \{2, 3, 4, \cdots\}$, we obtain the following solutions:

$$F(x,t)|_{n=2} = \frac{e^{-t\varepsilon}\left(-2t + x^2\varepsilon\right)}{\varepsilon},$$

$$F(x,t)|_{n=3} = \frac{e^{-t\varepsilon}x\left(-6t + x^2\varepsilon\right)}{\varepsilon},$$

$$F(x,t)|_{n=4} = \frac{e^{-t\varepsilon}\left(12t^2\varepsilon + x^4\varepsilon^3 - 12t\left(-2 + x^2\varepsilon^2\right)\right)}{\varepsilon^3},$$

and so on.

Eventually, let us choose the initial function:

$$f(x) = e^{ix},$$

which, upon the double integration in (146), produces the following simple solution for the telegraph equation (144):

$$F(x,t) = \exp\left[ix - \frac{t}{2}\left(\varepsilon + \sqrt{V}\right)\right], \tag{149}$$

where

$$V = \varepsilon^2 + 4(\kappa - \alpha).$$

We observe that the above solution presents no spread, but just the fading of the initial function with time. The physical meanings of the above equation and its solution become transparent upon considering a signal in a section of a transmission line in terms of voltage and current. We assume that the cable is imperfectly insulated so that there are both capacitance and current leakage to ground; the current in the electrical circuit is schematically shown in Fig. 1.

The equation of the voltage $u(x,t)$ in a cable line at any point and at any time is given by the following one-dimensional hyperbolic second-order telegraph equation

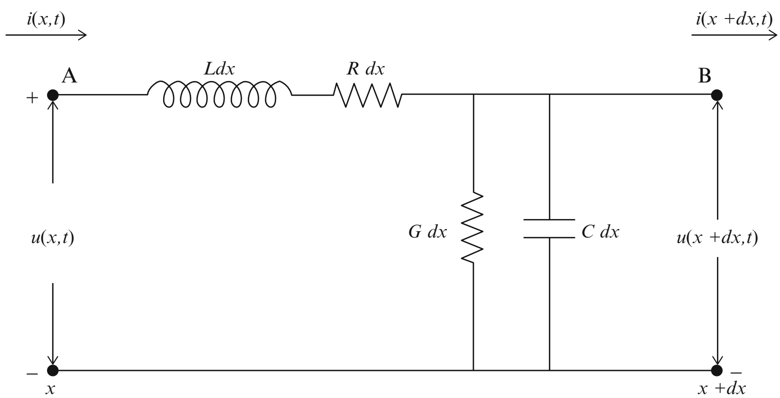


Fig. 1 Schematic diagram of a telegraphic transmission line with leakage

(see [25]):

$$\left(\frac{\partial^2}{\partial t^2}+(a+b)\,\frac{\partial}{\partial t}\right)u\,(x,t)=\left(c^2\frac{\partial^2}{\partial x^2}-ab\right)u\,(x,t)\,, \tag{150}$$

where

$$a=\frac{G}{C},\quad b=\frac{R}{L}\qquad\text{and}\qquad c^2=\frac{1}{LC}.$$

Evidently, in our notations

$$\alpha=c^2,\quad \kappa=-ab,\quad \varepsilon=(a+b)$$

and

$$V=(a-b)^2-(2c)^2\,,$$

and for

$$\left(\frac{G}{C}-\frac{R}{L}\right)^2>\frac{4}{LC},$$

the solution fades without shift:

$$u\,(x,t)=\exp\left[ix-\frac{t}{2}\left(\varepsilon+\sqrt{V}\right)\right]\qquad (V>0).$$

In the case when

$$\left(\frac{G}{C}-\frac{R}{L}\right)^2<\frac{4}{LC},$$

the voltage behaviour in the circuit exhibits the spatial shift and minor fading

$$u\,(x,t)=\exp\left[i\left(x-\frac{t}{2}\sqrt{|V|}\right)-\frac{t}{2}\varepsilon\right]\qquad (V<0),$$

which is determined exclusively by ε. Thus, we have obtained the exact solution for the voltage in an imperfect electric cable at any time and any point for the given initial voltage $u(x,t=0)$. The example of $u\,(x,t)$ for the values of $G=C=R=L=1$ in the cable line and for the initial voltage distribution given by

$$u\,(x,0)=\cos x$$

demonstrates the temporal fading and shift in Fig. 2.

In the context of heat conduction, the following equation:

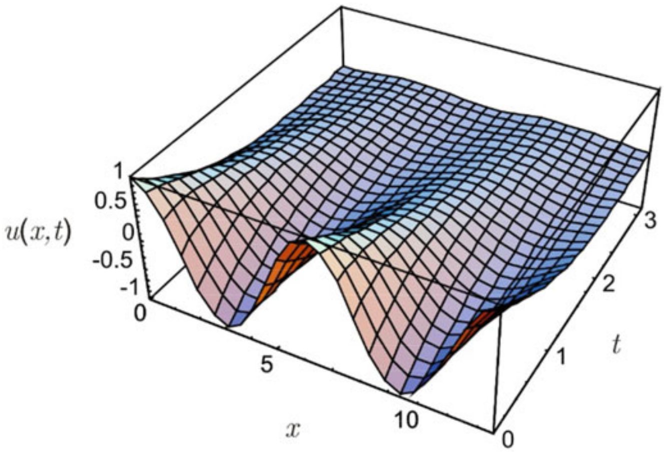


Fig. 2 Time and space distribution of the voltage $u(x,0) = \cos x$ in the cable for $G = C = R = L = 1$

$$\nabla^2 T = \frac{1}{D_T}\frac{\partial T}{\partial t} + \frac{\tau}{D_T}\frac{\partial^2 T}{\partial t^2}$$

arises, where D_T is the thermal diffusivity. The ratio given by

$$v_t = \sqrt{D_T/\tau}$$

is a velocity-like quantity, representing the speed of the heat wave in the medium, which characterizes the thermal wave propagation the same way as the diffusion behaviour is characterized by the diffusivity. The material parameter τ is an intrinsic thermal property, representing the build-up period for the initiation of a heat flow after a temperature gradient was imposed at the boundary of the domain. The heat flow does not start instantaneously, but rises gradually with a relaxation time τ after the application of the temperature gradient. Thus, there is a phase lag for the disappearance of the heat flow after removal of the temperature gradient. This relaxation time is associated with the linkage time of phonon–phonon collision, necessary for the initiation of a heat flow and is a measure of the thermal inertia of the medium. The telegraph equation (144) with $\kappa = 0$ then describes a one-dimensional case of the above non-Fourier heat conduction. The solution of the hyperbolic heat equation (144) in three-dimensional case with ∇^2 instead of ∂_x^2 assumes the following form:

$$F(x,y,z,t) = e^{-\frac{\varepsilon t}{2}} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi} - \xi\left(\varepsilon^2 + 4\kappa\right)\right) \cdot \hat{S}_x \hat{S}_y \hat{S}_z f(x,y,z), \tag{151}$$

where

$$f(x,y,z) = F(x,y,z,0),$$

and consists in the action of the heat diffusion operators for each coordinate on the initial function:

$$\hat{S}_x \hat{S}_y \hat{S}_z f(x,y,z),$$

$$\hat{S}_\ell = e^{-4\alpha\xi\partial_\ell^2} f(\ell) \qquad (\ell = x,y,z).$$

Their explicit action can be obtained in complete analogy with (146) and contains multiple integrals of Gauss type. The operational definition (25) is also useful in this context.

It is now easy to modify the telegraph equation (144) by adding non-commuting with ∂_x^2 terms such as x and other, in the right-hand side of (144) and, accounting for the commutators, arising in the exponential, obtain the solution. The examples of such generalization will be considered elsewhere.

Some Second Order of Time Partial Differential Equations with Mixed Derivatives and Non-constant Coefficients

On the way of generalization of the above-developed operational technique for solution of higher order partial differential equations, let us consider Eq. (125), where the coefficient ε is an operator $\hat{\varepsilon}(x)$:

$$\left(\frac{\partial^2}{\partial t^2} + \hat{\varepsilon}(x)\frac{\partial}{\partial t}\right) F(x,t) = \hat{D}(x) F(x,t). \tag{152}$$

Its non-diverging at the $t \to \infty$ solution, developing from

$$F(x,0) = f(x),$$

follows from (128):

$$F(x,t) = e^{-\frac{t}{2}\hat{\varepsilon}(x)} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi}\right) e^{-\xi[\hat{\varepsilon}(x)]^2} e^{-4\xi\hat{D}(x)} f(x). \tag{153}$$

The way we proceed with the calculation of the above integral totally depends on the explicit form of the operators $\hat{\varepsilon}(x)$ and $\hat{D}(x)$. Dependently on the value of their commutator $[\hat{\varepsilon}^2, \hat{D}]$, we may apply the appropriate identity to decompose them. Perhaps, the simplest example is given by the set of the operators:

$$\hat{\varepsilon} = \varepsilon \partial_x \qquad \text{and} \qquad \hat{D} = \partial_x^2 + \kappa^2,$$

which yield the following modified hyperbolic heat conduction equation with the mixed derivative term:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon \frac{\partial^2}{\partial x \partial t}\right) F(x,t) = \left(\frac{\partial^2}{\partial x^2} + \kappa^2\right) F(x,t) \qquad (\varepsilon, \kappa = \text{constant}). \tag{154}$$

Its non-diverging solution for

$$F(x,0) = f(x) \qquad \text{and} \qquad F(x,\infty) < \infty$$

reads as follows:

$$\begin{aligned} F(x,t) &= \mathrm{e}^{-\frac{t}{2}\varepsilon\partial_x} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi} - 4\kappa^2\xi\right) \exp\left[-\xi\left(4+\varepsilon^2\right)\partial_x^2\right] f(x) \\ &= \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi} - 4\kappa^2\xi\right) \hat{\Theta}\hat{S} f(x). \end{aligned} \tag{155}$$

Hence, the solution consists in the action of the operator $\hat{S}$, given by

$$\hat{S} = \exp\left[-\xi\left(4+\varepsilon^2\right)\partial_x^2\right] f(x)$$

and, in the shift along the x-axis, executed by the operator $\hat{\Theta}$ given by

$$\hat{\Theta} = \mathrm{e}^{-\frac{t}{2}\varepsilon\partial_x}$$

with the consequent integration of the resulting function with the kernel of the exponential-power product. For example, for the initial power function x^n, we readily obtain the Hermite polynomials according to their operational definition (39), so that the solution of Eq. (154) reads as follows:

$$\begin{aligned} F(x,t)|_{f(x)=x^n} = \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} H_n\left(x - \frac{t\varepsilon}{2}, -\xi\left(4+\varepsilon^2\right)\right) \\ \cdot \exp\left(-\frac{t^2}{16\xi} - 4\kappa^2\xi\right). \end{aligned} \tag{156}$$

The integration can be accomplished if we expand $H_n(x,y)$ in the series in (39), resulting in

$$F(x,t)\big|_{f(x)=x^n} = n!\sqrt{\frac{2t\kappa}{\pi}}\sum_{r=0}^{[n/2]}\frac{(x-t\varepsilon/2)^{n-2r}}{2^{3r}(n-2r)!r!}\left(-\frac{t\left(4+\varepsilon^2\right)}{\kappa}\right)^r K_{\frac{1}{2}-r}(t\kappa). \quad (157)$$

For example, with the initial function:

$$f(x) = x^2$$

for (154), its solution (157) reduces to

$$F(x,t)\big|_{n=2} = \mathrm{e}^{-t\kappa}\left(\frac{4x^2\kappa + t\big(-4+\varepsilon\big[-\varepsilon+\kappa(-4x+t\varepsilon)\big]\big)}{4\kappa}\right)$$

and for

$$f(x) = x^3,$$

we obtain

$$F(x,t)\big|_{n=3} = \mathrm{e}^{-t\kappa}\left(\frac{(-2x+t\varepsilon)\left(4x^2\kappa + t\big(-12+\varepsilon\big[-3\varepsilon+\kappa(-4x+t\varepsilon)\big]\big)\right)}{8\kappa}\right).$$

In the particular case when $\varepsilon = 1$ and $\kappa = 0$, we end up with

$$\begin{aligned} F(x,t)\big|_{\varepsilon=1,\kappa=0} &= \frac{n!\,t}{4\sqrt{\pi}}\int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}}\exp\left(-\frac{t^2}{16\xi}\right)\sum_{r=0}^{[n/2]}\frac{(-5\xi)^r\left(x-\frac{t}{2}\right)^{n-2r}}{(n-2r)!\,r!} \\ &= \frac{n!}{\sqrt{\pi}}\sum_{r=0}^{[n/2]}\frac{(x-t/2)^{n-2r}}{(n-2r)!\,r!}\left(-\frac{5}{16}\right)^r t^{2r}\Gamma\left(\frac{1}{2}-r\right). \end{aligned} \quad (158)$$

Eventually, let us explore the solution of the following differential equation of the second order:

$$\left(\frac{\partial^2}{\partial t^2} + \varepsilon\frac{\partial^2}{\partial x\partial t}\right)F(x,t) = \left(\frac{\partial^2}{\partial x^2} + \beta x\partial_x\right)F(x,t) \qquad (\varepsilon, \beta = \text{constant}), \quad (159)$$

where the right-hand side is that of the Fokker–Plank equation with the operator given by

$$\hat{D} = \partial_x^2 + \beta x\partial_x$$

and the left-hand side contains, apart the second order time derivative ∂_t^2, the operator given by

$$\hat{\varepsilon} = \varepsilon \partial_x.$$

The solution (153) of Eq. (159) involves the operator given by

$$\hat{\varepsilon}^2 = \varepsilon^2 \partial_x^2,$$

commuting with the second-order derivative in $\hat{D}$. Thus, the exponential of (153) now includes

$$-\xi \hat{\varepsilon}^2 - 4\xi \hat{D} = -a\partial_x^2 - bx\partial_x,$$

where

$$a = \xi \left(\varepsilon^2 + 4\right) > 0, \quad b = 4\xi\beta > 0 \qquad \text{and} \qquad \frac{a}{b} = \frac{\varepsilon^2 + 4}{4\beta}$$

and

$$F(x,t) = \mathrm{e}^{-\frac{t}{2}\varepsilon\partial_x} \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi}\right) \mathrm{e}^{-a\partial_x^2 - bx\partial_x} f(x). \tag{160}$$

The exponential in

$$\hat{U} = \mathrm{e}^{-a\partial_x^2 - bx\partial_x}$$

is not different from that in a common Fokker–Plank equation and we proceed along the lines, designated in [10] and [31, 35] to obtain

$$\hat{U} = \hat{S} f(y),$$
$$\hat{S} = \mathrm{e}^{-\frac{1}{2}\rho\partial_y^2}$$

and

$$y = \mathrm{e}^{-b} x,$$

so that the solution of Eq. (159) reads as follows:

$$F(x,t) = \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi}{\xi\sqrt{\xi}} \exp\left(-\frac{t^2}{16\xi}\right) \hat{\Theta}\hat{S} f(y), \tag{161}$$

where

$$\hat{\Theta} = \mathrm{e}^{-\frac{t}{2}\varepsilon\partial_x}.$$

For the initial Gaussian function given by

$$f(x) = \exp\left(-x^2\right),$$

we have

$$\hat{U}f(x) = \hat{S}f(y) = \frac{1}{\sqrt{1-2\rho}} \exp\left(-\frac{e^{-2b}\,x^2}{1-2\rho}\right), \tag{162}$$

where

$$\rho = \frac{a}{b}\left(1 - e^{-2b}\right) = \frac{\varepsilon^2 + 4}{4\beta}\left(1 - e^{-8\beta\xi}\right).$$

Thus, for the initial Gaussian function:

$$f(x) = \exp\left(-x^2\right),$$

the limited at infinite times solution (153) of Eq. (159) becomes

$$F(x,t)|_{f(x)=e^{-x^2}} = \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{d\xi e^{-\frac{t^2}{16\xi} - \frac{e^{-8\xi\beta}(x-t\varepsilon/2)^2}{1-2(1-e^{-8\beta\xi}(1-2\beta\xi\varepsilon^2))/\beta}}}{\xi\sqrt{\xi}\sqrt{1-2(1-e^{-8\beta\xi}(1-2\beta\xi\varepsilon^2))/\beta}}. \tag{163}$$

Evidently, the integral converges regularly, provided that

$$\beta > \frac{\varepsilon^2}{2} + 2,$$

and (for $\varepsilon = 0$) it becomes

$$F(x,t;\varepsilon=0)|_{f(x)=e^{-x^2}} = \frac{t}{4\sqrt{\pi}} \int_0^\infty \frac{\exp\left(-\frac{t^2}{16\xi} - \frac{e^{-8\xi\beta}(x)^2}{1-2(1-e^{-8\xi\beta})/\beta}\right)}{\xi\sqrt{\xi}\sqrt{1-2(1-e^{-8\xi\beta})/\beta}}\, d\xi. \tag{164}$$

The above solution was validated numerically; it satisfies Eq. (159). The function

$$F(x,t)|_{f(x)=e^{-x^2}}$$

for $\varepsilon = 0$ and $\beta = 4$ is plotted in Fig. 3.

We observe the temporal fading of the Gaussian. Non-zero values of ε break the coordinate symmetry of the solution as seen in Fig. 4 for $\varepsilon = 1.5$, $\beta = 4$. Observe the temporal fading, accompanied by the shift of the initial Gaussian in Fig. 4.

The above-developed operational technique can be applied for other equations, involving more complicated differential operators. The examples will be considered in the forthcoming publications.

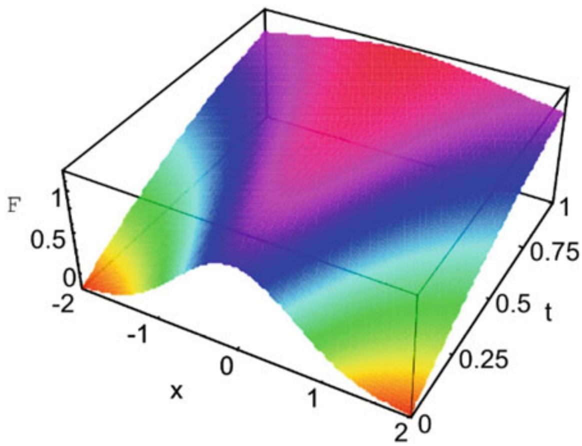


Fig. 3 Solution of equation $\left(\frac{\partial^2}{\partial t^2}\right) F(x,t) = \left(\frac{\partial^2}{\partial x^2} + \beta x \partial_x\right) F(x,t)$ for $f(x) = \exp(-x^2)$, $\beta = 4$

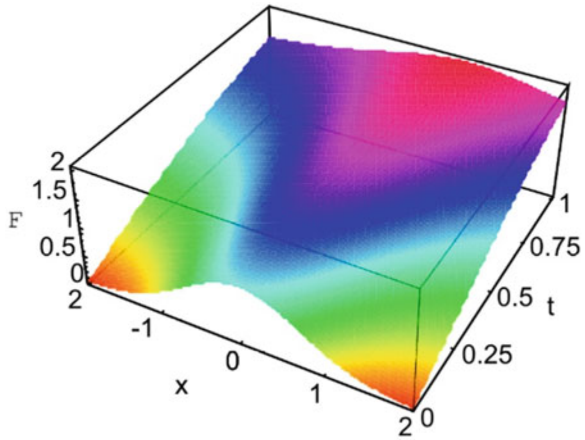


Fig. 4 Solution of equation $\left(\frac{\partial^2}{\partial t^2} + \varepsilon \frac{\partial^2}{\partial x \partial t}\right) F(x,t) = \left(\frac{\partial^2}{\partial x^2} + \beta x \partial_x\right) F(x,t)$ for $f(x) = \exp(-x^2)]$, $\varepsilon = 1.5$, $\beta = 4$

General Families of Special Functions and Orthogonal Polynomials for the UR Studies

Concluding our survey, we note that many equations in physics and, in particular, in physics of accelerators, as well as the equations, related to the propagation and to the radiation of charged particles, assume compact form in operational notations. For example, the Colson equation, common in the physics of FEL, in the limit of low FEL gain, has the following integro-differential form (see [13]):

$$\frac{da}{d\tau} = i\pi g_0 \int_0^{\tau} \tau' a(\tau - \tau') \exp\left(-i\nu\tau'\right) d\tau', \tag{165}$$

where g_0 is the FEL gain in the low gain limit, a is the dimensionless strength of the Colson's field. In terms of the inverse derivative operator (2) this equation reads as follows:

$$\mathrm{e}^{i\nu\tau} D_\tau a(\tau) = i\pi g_0 D_\tau^{-2} \left\{\mathrm{e}^{i\nu\tau}\, a(\tau)\right\}. \tag{166}$$

The above-developed technique, involving exponential operators, can be applied for modeling the dynamics and the radiation processes in FEL and in UR devices. With account for inhomogeneous losses, for example, due to the energy spread of electrons in the beam, the beam divergency etc., the Colson's equation (165) is modified accordingly (see [4]):

$$\frac{da}{d\tau} = i\pi g_0 \int_0^\tau \tau' a(\tau - \tau')\hat{O} \exp\left(-i\nu\tau'\right) d\tau', \tag{167}$$

where the operator $\hat{O}$ describes the losses. For example, the energy spread is accounted for by

$$\hat{O} = \mathrm{e}^{\frac{1}{2}\pi\mu_\varepsilon\partial_\nu^2}.$$

The non-periodic constituents in the undulator magnetic field, the energy spread and the beam divergency can be accurately accounted for by the generalization of the Airy functions (see [22] and [32]). The solutions obtained in [14, 33, 34, 36, 37] and [38, 39] analytically describe the spectrum, the intensity and the emission line shape of the emitted harmonics. The emission line shape differs from the commonly used function given by

$$\mathrm{sinc}\nu_n \equiv \frac{\sin \nu_n}{\nu_n},$$

where

$$\nu_n = \pi n N \left(\frac{\omega}{\omega_n} - 1\right)$$

is the detuning parameter, describing the deviation of the radiation frequency ω from the central frequencies of each harmonic ω_n. The accurate account is possible by the generalized Airy functions, which have the following integral form:

$$\begin{aligned} S(\alpha, \beta, \varepsilon) &\equiv \int_0^1 \exp\left[i\left(\alpha\,\tau + \varepsilon\,\tau^2 + \beta\,\tau^3\right)\right] d\tau \\ &= \sum_{n=0}^{\infty} \frac{i^n}{(n+1)!}\, H_n(\alpha, -\beta, -i\varepsilon). \end{aligned} \tag{168}$$

They can also be rewritten as the following series of the generalized Hermite polynomials of three indices:

$$H_n(x,y,z) = n! \sum_{r=0}^{[n/3]} \frac{z^{n-3r}}{(n-3r)!r!} H_n(x,y), \tag{169}$$

where the polynomials $H_n(x,y)$ are given either by series or by the operational definition (39) and are reduced to the common Hermite polynomials (41). Although less useful for practical purposes, this form reveals the link of the generalized Airy function $S(x,y,z)$ to the polynomials with the following exponential generating function:

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x,y,z) = \exp\left(xt + yt^2 + zt^3\right). \tag{170}$$

Considering the fact that the expression for the UR intensity involves series of generalized Bessel functions, which may have two, and more, arguments and indices, it is definitely preferable to substitute the second involved series of the Hermite polynomials by their operational definition for the easy of manipulation. Indeed, the operational exponential definition given by

$$H_n(x,y,z) = \exp\left(y\,\partial_x^2 + z\,\partial_x^3\right)\{x^n\} \tag{171}$$

yields the following operational presentation for the generalized Airy function:

$$\begin{aligned} S(x,y,z) &= \mathrm{e}^{-i\,y\,\partial_x^2 - z\,\partial_x^3} \int_0^1 \mathrm{e}^{i\,x\,\tau}\,d\tau \\ &= \mathrm{e}^{-i\,y\,\partial_x^2 - z\,\partial_x^3} \left\{\frac{\sin x}{x}\,\mathrm{e}^{i\,x}\right\}. \end{aligned} \tag{172}$$

It reveals the operational relation between functions $S(x,y,z)$, which describe the emission line shape with account for the losses in undulators, and the function sinc$x$, describing the line of the ideal device without losses. A rather detailed description of the generalized Airy functions behaviour can be found in [8]. On the undulator axis, where the radiation is at its maximum, the UR line shape is described by the two-variable Airy type function

$$\begin{aligned} S(\alpha,\beta) &= \int_0^1 \mathrm{e}^{i\left(\alpha\,\tau + \beta\,\tau^3\right)}\,d\tau \\ &= \sum_{m=0}^{\infty} \frac{i^m}{(m+1)!} H_m(\alpha,-\beta), \end{aligned} \tag{173}$$

and (in the ideal undulator) it is reduced to

$$S(\alpha,0,0) = \mathrm{e}^{i\alpha}\,\mathrm{sinc}\alpha. \tag{174}$$

Simplifications follow in the operational presentation (172) of $S(x, y, z)$ accordingly. Not going into the details involved (see, for example, [3, 5, 20] and [23]), we just note that, according to Madey's theorem (see [21]), the derivative of the UR line shape given by

$$-\frac{\partial(\mathrm{sinc}^2 \nu_n)}{\partial \nu_n}$$

describes the FEL gain in the so-called Compton or low gain regime (see Fig. 5).

In the presence of the constant magnetic field it becomes (see Fig. 6)

$$-\frac{\partial \left(S^2(\nu_n)\right)}{\partial \nu_n}.$$

In this case, the dependence on the constant magnetic component is all contained in the parameter β given by

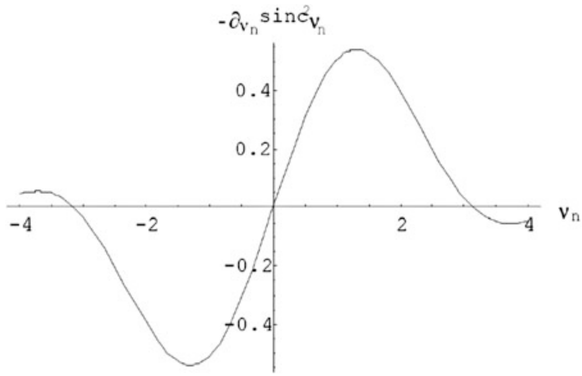


Fig. 5 Function $-\partial(\mathrm{sinc}^2 \nu_n)/\partial \nu_n$, which describes the FEL gain with an ideal undulator

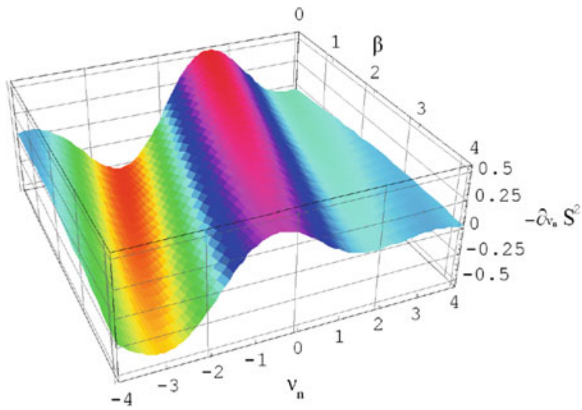


Fig. 6 Function $-\partial S^2\,(\nu_n, \beta)\,/\partial \nu_n$, which describes the FEL gain with account for the constant magnetic field and for the beam divergency

$$\beta = (2\pi n N) \frac{(\gamma\theta_H)^2}{1 + \frac{k^2}{2} + (\gamma\theta_H)^2}, \tag{175}$$

where

$$\theta_H = \frac{2}{\sqrt{3}} \frac{k}{\gamma} \pi N \kappa_1$$

is the effective bending angle due to the constant magnetic field given by

$$H_\text{d} = H_0 \kappa_1,$$

H_0 is the amplitude value of the strength of the periodic magnetic field, N is the number of periods in the undulator and n is the radiated harmonic number. It is evident from the Fig. 6 the constant field ($\beta > 0$) lowers the FEL gain and shifts the spectral maximum of the curve back, respectively, to the common value

$$\nu_n = \frac{\pi}{2}.$$

The increase of the strength of constant magnetic component indiscriminately lowers the gain through the whole spectrum range because of the coherence length is no more maintained for the electrons. The values of the order of

$$H_\text{d} \approx H_0 \cdot 10^{-4}$$

already may cause serious disruption in the coherence of the electrons oscillations (see [32] and [33]).

Conclusions

We obtained solutions for differential equations by means of operational method. We demonstrated that the usages of the operational method together with integral transforms constitute a reasonably simpler approach to a wide spectrum of differential equations. In our present investigation, not only have we obtained solutions of relevant physical problems, but we have also demonstrated their transparent meaning and distinguished effect of each term in the initial equation on the solution. Moreover, we have obtained some solutions by employing generalized forms of the Laguerre and Hermite orthogonal polynomials, which give rise to a possibility to write solutions, in some general cases, in remarkably compact series-forms, facilitating the analysis considerably. Frequently, they also give links to special functions, which (on the other hand) can be generated by the operational exponents. Such operational approaches can be even more advantageous in calculus. We demonstrated how inverting the derivative operators of various orders and their combinations could be a viable way to obtain easy and straightforward solutions of various types of differential equations and mathematical problems, describing wide range of physical processes as well.

Acknowledgements The authors are grateful to A. Borisov, A. Lobanov and A. Bubnov for several useful discussions about the subject-matter of this investigation.

References

1. Appell, P., Kampé de Fériet, J.: Fonctions Hypergéométriques et Hypersphériques; Polynômes d'Hermite. Gauthier-Villars, Paris (1926)
2. Berry, M.V., Balazs, N.J.: Nonspreading wave packets. Am. J. Phys. **47**, 264–266 (1979)
3. Brau, C.A.: Free-Electron Lasers. Academic, Boston (1990)
4. Colson, W.B., Gallardo, L.C., Bosco, P.M.: Free electron laser gain degradation and electron beam quality. Phys. Rev. A **34**, 4875–4883 (1986)
5. Dattoli, G.: Lectures on Free Electron Lasers. World Scientific, Singapore (1993)
6. Dattoli, G.: Generalized polynomials, operational identities and their applications. J. Comput. Appl. Math. **118**, 111–123 (2000)
7. Dattoli, G., Zhukovsky, K.V.: Appell polynomial series expansion. Int. Math. Forum **5**, 649–662 (2010)
8. Dattoli, G., Srivastava, H.M., Zhukovsky, K.V.: Orthogonality properties of the Hermite and related polynomials. J. Comput. Appl. Math. **182**, 165–172 (2005)
9. Dattoli, G., Mikhailin, V.V., Ottaviani, P.L., Zhukovsky, K.V.: Two-frequency undulator and harmonic generation by an ultrarelativistic electron. J. Appl. Phys. **100**(8), Article ID 084507, 1–12 (2006)
10. Dattoli, G., Srivastava, H.M., Zhukovsky, K.V.: Operational methods and differential equations with applications to initial-value problems. Appl. Math. Comput. **184**, 979–1001 (2007)
11. Dattoli, G., Srivastava, H.M., Zhukovsky, K.V.: A new family of integral transforms and their applications. Integral Transform. Spec. Funct. **17**, 31–37 (2007)
12. Dattoli, G., Mikhailin, V.V., Zhukovsky, K.: Undulator radiation in a periodic magnetic field with a constant component. J. Appl. Phys. **104**, 124507-1–124507-8 (2008)
13. Dattoli, G., Migliorati, M., Zhukovsky, K.V.: Summation formulae and Stirling numbers. Int. Math. Forum **4**, 2017–2040 (2009)
14. Dattoli, G., Mikhailin, V.V., Zhukovsky, K.V.: Influence of a constant magnetic field on the radiation of a planar undulator. Mosc. Univ. Phys. Bull. **64**, 507–512 (2009); see also Vestn. Mosk. Univ. Fiz. Astron. **5**, 33–38 (2009)
15. Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F.G.: Higher Transcendental Functions, vol. II. McGraw-Hill, New York (1953)
16. Gould, H.W., Hopper, A.T.: Operational formulas connected with two generalizations of Hermite polynomials. Duke Math. J. **29**, 51–63 (1962)
17. Haimo, D.T., Markett, C.: A representation theory for solutions of a higher-order heat equation, I. J. Math. Anal. Appl. **168**, 89–107 (1992)
18. Haimo, D.T., Markett, C.: A representation theory for solutions of a higher-order heat equation, II. J. Math. Anal. Appl. **168**, 289–305 (1992)
19. Hirschman, I.I., Widder, D.V.: The Convolution Transform. Princeton University Press, Princeton (1955)
20. Luchini, P., Motz, H.: Undulators and Free-Electron Lasers. Clarendon Press, Oxford (1990)
21. Madey, J.M.J.: Stimulated emission of Bremsstrahlung in a periodic magnetic field. J. Appl. Phys. **42**, 1906–1913 (1971)

22. Mikhailin, V.V., Zhukovsky, K.V., Kudiukova, A.I.: On the radiation of a planar undulator with constant magnetic field on its axis taken into account. J. Surf. Invest. X-Ray Synchrotron Neutron Tech. **8**, 422–428 (2014)
23. Murphy, J.B., Pellegrini, C.: Introduction to the physics of the free electron laser. In: Month, M., Turner, S. (eds.) Frontiers of Particle Beams. Proceedings of a Topical Course, Held by the Joint US-CERN School on Particle Accelerators, at South Padre Island, Texas; October 23–29, 1986). Lecture Notes in Physics, vol. 296, pp. 163–219. Springer, Berlin.Heidelberg (1988); see also Colson, W.B., Pellegrini, C., Renieri, A. (eds.): Laser Handbook, vol. 6. North-Holland, Amsterdam (1990)
24. Srivastava, H.M., Manocha, H.L.: A Treatise on Generating Functions. Halsted Press (Ellis Horwood Limited, Chichester). Wiley, New York (1984)
25. Terman, F.E.: Radio Engineers' Handbook, 1st edn. McGraw-Hill, New York (1943)
26. Vallée, O., Soares, M.: Airy Functions and Application to Physics. World Scientific, Singapore (2004)
27. Watson, G.N.: A Treatise on the Theory of Bessel Functions, 2nd edn. Cambridge University Press, Cambridge (1944)
28. Zhukovsky, K.V.: Exact solution of Guyer-Krumhansl type heat equation by operational method. Int. J. Heat Mass Transf. **96**, 132–144 (2016)
29. Zhukovsky, K.: Violation of the maximum principle and negative solutions for pulse propagation in Guyer-Krumhansl model. Int. J. Heat Mass Transf. **98**, 523–529 (2016).
30. Zhukovsky, K.V.: Undulator radiation in multiple magnetic fields. In: Chua, D.M., Toh, H.F. (eds.) Synchrotron: Design, Properties and Applications, pp. 39–92, Nova Science, New York (2012)
31. Zhukovsky, K.V.: Solution of some types of differential equations: operational calculus and inverse differential operators. Sci. World J. **2014**, Article ID 454865, 1–8 (2014)
32. Zhukovsky, K.V.: Inhomogeneous and homogeneous losses and magnetic field effect in planar undulator radiation. Prog. Electromagn. Res. B **59**, 245–256 (2014)
33. Zhukovsky, K.V.: Analytical account for a planar undulator performance in a constant magnetic field. J. Electromagn. Waves Appl. **28**, 1869–1887 (2014)
34. Zhukovsky, K.V.: A model for analytical description of magnetic field effects and losses in a planar undulator radiation. J. Surf. Invest. X-Ray Synchrotron Neutron Tech. **8**, 1068–1081 (2014)
35. Zhukovsky, K.V.: A method of inverse differential operators using orthogonal polynomials and special functions for solving some types of differential equations and physical problems. Mosc. Univ. Phys. Bull. **70**(2), 93–100 (2015). doi:10.3103/S0027134915020137
36. Zhukovsky, K.V.: Harmonic generation by ultrarelativistic electrons in a planar undulator and the emission-line broadening. J. Electromagn. Waves Appl. **29**, 132–142 (2015)
37. Zhukovsky, K.V.: High harmonic generation in the undulators for free electron lasers. Opt. Commun. **353**, 35–41 (2015)
38. Zhukovsky K.V.: Harmonic radiation in a double-frequency undulator with account for broadening, Mosc. Univ. Phys. Bull. **70**(4), 232–239 (2015). doi:10.3103/S0027134915040177
39. Zhukovsky, K.: High harmonic generation in undulators for FEL, Nucl. Instrum. Methods Phys. Res. B **369**, 9–14 (2016), doi:10.1016/j.nimb.2015.10.041
40. Zhukovsky, K.V., Dattoli, G.: Evolution of non-spreading Airy wavepackets in time dependent linear potentials. Appl. Math. Comput. **217**, 7966–7974 (2011)

A Modified Pointwise Estimate on Simultaneous Approximation by Bernstein Polynomials

Gancho Tachev

Abstract We prove pointwise estimate for the approximation of the k-th derivative of the function f by the k-th derivative of its Bernstein operator B_n. We compare our result with similar estimates obtained earlier.

Keywords Bernstein polynomials • Simultaneous approximation • Moduli of smoothness

Introduction

For every function $f \in C[0,1]$ the Bernstein polynomial operator at the point $x \in [0,1]$ is given by

$$B_n(f;x) = \sum_{k=0}^{n} f(\frac{k}{n}) \cdot \binom{n}{k} x^k (1-x)^{n-k}. \tag{1}$$

The k-th derivative of $B_n(f,x)$ can be expressed as

$$B_n^{(k)}(f,x) = n(n-1)\dots(n-k+1) \cdot \sum_{\nu=0}^{n-k} \Delta^k f\left(\frac{\nu}{n}\right) \cdot \binom{n-k}{\nu} \cdot x^\nu \cdot (1-x)^{n-\nu-k} \tag{2}$$

with

$$\Delta f\left(\frac{\nu}{n}\right) = f\left(\frac{\nu+1}{n}\right) - f\left(\frac{\nu}{n}\right)$$

and the k-th iterate of Δ is denoted by Δ^k. This formula was established by Lorentz in [11]. From (2) we can derive the following convergence result (see Lorentz-[11, 12]), formulated in the book of DeVore and Lorentz (see Theorem 2.1 in Chap. 10 in [1]) as:

G. Tachev (✉)
Department of Mathematics, University of Architecture, BG-1046 Sofia, Bulgaria
e-mail: gtt_fte@uacg.bg

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_27

Theorem A. *(i) If $f \in C[0,1]$, then $\|B_n^{(k)}(f) - f^{(k)}\| \to 0$ if $n \to \infty$. (ii) if $f \in C[0,1]$, $0 \le x_0 \le 1$ and f is $(k-1)$ times continuously differentiable in some neighborhood of x_0, then $B_n^{(k)}(f, x_0) \to f^{(k)}(x_0)$ whenever $f^{(k)}(x_0)$ exists.*

The first quantitative form of Theorem A, which gives us the possibility to measure the rate of uniform convergence $\|B_n^{(k)}(f) - f^{(k)}\| \to 0$ in terms of the first modulus of continuity $\omega_1(f^{(r)}, \cdot)$ was proved by Knoop and Pottinger in [9] in 1976:

Theorem B. *Let $K = [a,b]$ and $K' = [c,d]$ be a subinterval of K, $r \in N$ and $L : C^r(K) \to C^r(K')$ be almost convex of order $r-1$. If $L(\pi_{r-1}) \subseteq \pi_{r-1}$, then for $f \in C^r(K)$, $x \in K'$, $\delta > 0$ there holds*

$$\begin{aligned} |D^r f(x) - D^r Lf(x)| &\le \tfrac{1}{r!}|D^r f(x)| \cdot |D^r e_r(x) - D^r Le_r(x)| + \\ &+ \left[\tfrac{1}{r!} D^r Le_r(x) + \delta^{-2} \cdot \beta_L^2(x)\right] \cdot \omega_1(f^{(r)}, \delta), \end{aligned} \tag{3}$$

where

$$\beta_L^2(x) := D^r L\left(\frac{2}{(r+2)!} \cdot e_{r+2} - \frac{2}{(r+1)!} \cdot x \cdot e_{r+1} + \frac{1}{r!} x^2 \cdot e_r\right)(x).$$

For more details and denotations see [9]. In his dissertation—[3] H. Gonska improved the estimate in Theorem B showing the following (see also Theorem 3.1 in [4])

Theorem C. *By the same conditions as in Theorem B there holds*

$$\begin{aligned} |D^r f(x) - D^r Lf(x)| &\le 3\|D^r f\|_{K'} \cdot |\tfrac{1}{r!} \cdot D^r Le_r(x) - 1| + \\ &+ 2\gamma_L(x) \cdot \max\{\tfrac{1}{\delta}, \tfrac{1}{b-a}\} \cdot \omega_1(D^r f, \delta) + \\ &+ \left[\tfrac{3}{2} \cdot \left(\|\tfrac{1}{r!} D^r Le_r\| + 1\right) + \beta_L^2(x) \cdot \max\{\tfrac{1}{\delta}, \tfrac{1}{b-a}\}\right] \cdot \omega_2(D^r f, \delta), \end{aligned} \tag{4}$$

where β_L^2 is the same quantity as in Theorem B and

$$\gamma_L(x) = \left| D^r L\left(\frac{1}{(r+1)!} \cdot e_{r+1} - \frac{1}{r!} \cdot x \cdot e_r\right)(x)\right|.$$

Generalizing the ideas from [3, 4] Kacsó proved in [8] the following:

Theorem D. *For n sufficiently large one has for $r \ge 0$:*

$$\begin{aligned} |D^r B_n f(x) - D^r f(x)| &\le \tfrac{r(r-1)}{2n} \cdot |D^r f(x)| + \\ &+ \tfrac{r}{2\sqrt{n}} \omega_1(D^r f, \tfrac{1}{\sqrt{n}}) + \tfrac{9}{8} \omega_2(D^r f, \tfrac{1}{\sqrt{n}}). \end{aligned} \tag{5}$$

Very recently, in order to derive both an error bound and an asymptotic formula for the derivatives of Bernstein polynomial, M. Floater proved in [2]

Theorem E. *If $f \in C^{k+2}[0,1]$, for some $k \geq 0$, then*

$$\left|(B_nf)^{(k)}(x) - f^{(k)}(x)\right| \leq \tfrac{1}{2n}\Big[k(k-1)\cdot\|f^{(k)}\| + k|1-2x|\cdot\|f^{(k+1)}\| + +x(1-x)\cdot\|f^{(k+2)}\|\Big]. \tag{6}$$

Similar estimate was proved by López-Moreno et al. in [10] using a completely different approach. We point out that the idea for the proof of Theorem E is quite different from the methods, used in Theorems B–D—where the emphasis is on the evaluation of the moments of the operator L and their derivatives. We recall that $e_i : x \to x^i$, $x \in [0,1]$, $i = 0, 1, \ldots$ are the monomial functions. If we compare Theorem D with Theorem E by the assumption that $f \in C^{r+2}[0,1]$ and using the properties of ω_1 and ω_2 we may obtain the same order $O(\frac{1}{n})$, $n \to \infty$ of the convergence for $|D^rB_nf(x) - D^rf(x)|$. More detailed analysis shows that none of the last two upper bounds (5) and (6) can be delivered from the other one, i.e. each of the both estimates has its advantages and disadvantages. Some recent results for Bernstein-type operators are listed in [5, 7] and the references therein. Our main result states the following:

Theorem 1. *For $f \in C^k[0,1]$ and $k = 1, 2, \ldots, n-1$ one has for $x \in [0,1]$*

$$\left|(B_nf)^{(k)}(x) - f^{(k)}(x)\right| \leq \left|f^{(k)}(x) - (B_{n-k}f^{(k)})(x)\right|\cdot(1-\tfrac{1}{n})^{k-1} + + \min\{1, \tfrac{k(k-1)}{2n}\}\cdot|f^{(k)}| + \omega_1(f^{(k)}, \tfrac{k}{n})\cdot(1-\tfrac{1}{n})^{k-1}. \tag{7}$$

The method of the proof is different from all the above-mentioned results and relies on the representation (2) and some simple calculations. In section "Proof of Theorem 1" we give the proof of Theorem 1. We end the paper with some remarks and comparisons of our result with the previous one.

Proof of Theorem 1

We start with the following:

Lemma 1. *For $k \geq 2$, $n \geq 1$ the following*

$$1 - \prod_{i=1}^{k-1}(1 - \frac{i}{n}) \leq \frac{k(k-1)}{2n} \tag{8}$$

holds true.

Proof. We use induction w.r.t. k. For $k = 2$ (8) is obvious and we have equality. Suppose (8) holds for $k-1$. We calculate

$$1-\prod_{i=1}^{k}(1-\tfrac{i}{n}) = 1-\left[\prod_{i=1}^{k-1}(1-\tfrac{i}{n})\right]\cdot(1-\tfrac{k}{n}) =$$
$$= 1-\prod_{i=1}^{k-1}(1-\tfrac{i}{n})+\tfrac{k}{n}\cdot\prod_{i=1}^{k-1}(1-\tfrac{i}{n}) \le \tfrac{k(k-1)}{2n}+\tfrac{k}{n}\cdot 1 = \tfrac{k(k+1)}{2n}.$$

The proof is completed.

Remark 1. To obtain an upper bound in (8) we may proceed also as follows:

$$1-\prod_{i=1}^{k-1}(1-\frac{i}{n}) \le 1-\left(1-\frac{k-1}{n}\right)^{k-1} \begin{cases} \le \frac{(k-1)^2}{n} \\ \le 1. \end{cases} \tag{9}$$

In the last we have used the Bernoulli inequality. The estimate (9) was used in Lemma 5 in [6].

Proof of Theorem 1. The formula (2) and the properties of $\Delta^k f(\frac{\nu}{n})$ imply

$$B_n^{(k)}(f,x) = \prod_{i=1}^{k-1}(1-\frac{i}{n})\cdot\sum_{\nu=0}^{n-k} f^{(k)}\left(\frac{\nu}{n}+\theta_\nu\frac{k}{n}\right)\cdot\binom{n-k}{\nu}x^\nu(1-x)^{n-\nu-k},$$

where $0 < \theta_\nu < 1$. Therefore,

$$B_n^{(k)}(f,x) = \prod_{i=1}^{k-1}(1-\tfrac{i}{n})\cdot$$
$$\left\{B_{n-k}(f^{(k)},x)+\sum_{\nu=0}^{n-k}\left[f^{(k)}\left(\tfrac{\nu}{n}+\theta_\nu\tfrac{k}{n}\right)-f^{(k)}\left(\tfrac{\nu}{n-k}\right)\right]\binom{n-k}{\nu}x^\nu(1-x)^{n-k-\nu}\right\}. \tag{10}$$

Hence

$$B_n^{(k)}(f,x)-f^{(k)}(x) = \prod_{i=1}^{k-1}(1-\tfrac{i}{n})\cdot\left\{B_{n-k}(f^{(k)},x)-f^{(k)}(x)\right\}-$$
$$-\left[1-\prod_{i=1}^{k-1}(1-\tfrac{i}{n})\right]\cdot f^{(k)}(x)+ \tag{11}$$
$$+\prod_{i=1}^{k-1}(1-\tfrac{i}{n})\cdot\sum_{\nu=0}^{n-k}\left[f^{(k)}\left(\tfrac{\nu}{n}+\theta_\nu\cdot\tfrac{k}{n}\right)-f^{(k)}\left(\tfrac{\nu}{n-k}\right)\right]\binom{n-k}{\nu}x^\nu(1-x)^{n-k-\nu}.$$

In the right-hand side of (11) we have three summands. To obtain an upper bound for the third summand in terms of $\omega_1(f^{(k)},\cdot)$ we evaluate the difference

$$A_{\nu,k,n} := (\frac{\nu}{n}+\theta_\nu\cdot\frac{k}{n})-(\frac{\nu}{n-k}),\ 0<\theta_\nu<1,\ 0\le\nu\le n-k.$$

Simple calculations yield

$$|A_{\nu,k,n}| = \frac{k}{n} \cdot \left|\frac{\nu - \theta_\nu(n-k)}{n-k}\right| \leq \frac{k}{n} \cdot 1, \tag{12}$$

for all $0 < \theta_\nu < 1$, $0 \leq \nu \leq n-k$. Consequently (10)–(12), and Lemma 1 imply

$$\begin{aligned} \left|(B_nf)^{(k)}(x) - f^{(k)}(x)\right| &\leq \left|f^{(k)}(x) - (B_{n-k}f^{(k)})(x)\right| \cdot (1-\tfrac{1}{n})^{k-1} + \\ &+ \min\{1, \tfrac{k(k-1)}{2n}\} \cdot |f^{(k)}| + \omega_1(f^{(k)}, \tfrac{k}{n}) \cdot (1-\tfrac{1}{n})^{k-1}. \end{aligned} \tag{13}$$

The proof of Theorem 1 is completed.

Corollary 1. *For $f \in C^{k+2}[0,1]$ and $k = 1, 2, \ldots, n-1$ one has for $x \in [0,1]$*

$$\begin{aligned} \left|(B_nf)^{(k)}(x) - f^{(k)}(x)\right| &\leq (1-\tfrac{1}{n})^{k-1} \cdot \tfrac{x(1-x)}{2(n-k)} \cdot \|f^{(k+2)}\| + \\ &+ \min\{1, \tfrac{k(k-1)}{2n}\} \cdot \|f^{(k)}\| + (1-\tfrac{1}{n})^{k-1} \cdot \tfrac{k}{n} \cdot \|f^{(k+1)}\|. \end{aligned} \tag{14}$$

Proof. With $\|\cdot\|$-the max norm on [0, 1], the error bound

$$|B_n(g,x) - g(x)| \leq \frac{1}{2n} x(1-x) \cdot \|g''\| \tag{15}$$

given in Chap. 10 in [1] holds true for all $g \in C^2[0,1]$. We apply this estimate for $g = f^{(k)}$. Also the following property of ω_1 is well known

$$\omega_1(g,\delta) \leq \delta\|g'\| \text{ for all } g \in C^1[0,1]. \tag{16}$$

We apply this estimate for $g = f^{(k)}$. Last (13), (14), and (16) complete the proof.

Remark 2. If we estimate

$$\min\{1, \frac{k(k-1)}{2n}\} \leq \frac{k(k-1)}{2n}$$

we observe that the second summand in the right hand side of (2.7) is the same as appearing in Theorems D and E.

Remark 3. The main applications of Theorems D, E and Theorem 1 is for a given function $f \in C^k[0,1]$ to measure the rate of convergence for $|B_n^{(k)}(f,x) - f^{(k)}(x)|$. On the other hand, the estimate (7) in Theorem 1 is valid for all $k = 1, 2, \ldots, n-1$. Therefore if we take $k = n^{\frac{1}{2}+\alpha}$, for some $0 < \alpha < 1$ and again take $n \to \infty$, then we obtain

$$\frac{k(k-1)}{2n} \to \infty, \; n \to \infty$$

in Theorems D and E. Instead of this in Theorem 1 we may now evaluate

$$\min\{1, \frac{k(k-1)}{2n}\} \leq 1$$

and thus obtain better upper bound for the factor in front of $\|f^{(k)}\|$. For example, if we set $f_0(x) = e^x$, and $k_0 = n^{\frac{1}{2}+\alpha}$, $\alpha \in (0,1)$, then Corollary 1 implies

$$|B_n^{(k_0)}(f_0, x) - f_0^{(k_0)}(x)| = O(1),\ n \to \infty.$$

The last conclusion is not possible to be obtained neither from Theorem D nor from Theorem E.

Remark 4. If now we set $k_1 = n-1$ and compare our main result with Theorem E, we conclude that the factors in front of $\|f^{(k+2)}\|$ and $\|f^{(k+1)}\|$ are better in Theorem E. But for any function, infinitely differentiable on $[0,1]$ such that $\|g^{(n)}\| \leq C$ for all $n = 1, 2, \ldots,$ (for example $g(x) = e^x$) according to Corollary 1 we would have

$$|B_n^{(k_1)}(g, x) - g^{(k_1)}(x)| = O(1),\ n \to \infty,$$

while Theorems E and D would imply

$$\lim_{n\to\infty} |B_n^{(k_1)}(g, x) - g^{(k_1)}(x)| = \infty.$$

To summarize our remarks we may conclude that there are cases, where the estimates in Theorems E and D are more appropriate than those in Theorem 1 and vice versa.

References

1. DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Springer, New York (1993)
2. Floater, M.S.: On the convergence of derivatives of Bernstein approximation. J. Approx. Theory **134**, 130–135 (2005)
3. Gonska, H.H.: Quantitative Aussagen zur Approximation durch positive lineare Operatoren. Dissertation, Universität Duisburg (1979)
4. Gonska, H.H.: Quantitative Korovkin type theorems on simultaneous approximation. Math. Z. **186**, 419–433 (1984)
5. Gonska, H.H.: Positive operators and approximation of functions: selected topics. In: Proceedings of the Summer School “Operator Methods for Evolution Equations and Approximation Problems” (OMEEAP 2002), Monopoli, September 2002
6. Gonska, H.H., Prestin, J., Tachev, G., Zhou, D.X.: Simultaneous approximation by Bernstein operator in Hölder norms. Math. Nachr. **286**, 349–359 (2013)
7. Gupta, V., Agarwal, R.P.: Convergence Estimates in Approximation Theory. Springer International Publishing, Switzerland (2014)

8. Kacsó, D.P.: Simultaneous approximation by almost convex operators. Rend. Circ. Mat. Palermo (2) **68**(Suppl. 68), 523–538 (2002)
9. Knoop, H.B., Pottinger, P., von Korovkin-Typ für C^k -Räume, E.S.: Math. Z. **148**, 23–32 (1976)
10. López-Moreno, A.J., Martínez-Moreno, J., Muñoz-Delgado, F.J.: Asymptotic expression of derivatives of Bernstein type operators. Rend. Circ. Mat. Palermo, Ser. II **68**, 615–624 (2002)
11. Lorentz, G.G.: Zur Theorie der Polynome von S. Bernstein. Matem. Sbornik **2**, 543–556 (1937)
12. Lorentz, G.G.: Bernstein Polynomials. University of Toronto Press, Toronto (1953)

Structural Fixed Point Results in Metric Spaces

Mihai Turinici

Abstract In Part 1, a class of anticipative contractions over quasi-ordered metric spaces is introduced and a corresponding lot of metrical fixed point theorems is formulated. The obtained facts include some well-known statements in the area due to Boyd and Wong or Matkowski, as well as a recent contribution due to Choudhury and Kundu (Demonstr Math 46:327–334, 2013). Further, in Part 2, a relative type version is given for the fixed point result in Leader (Math Jpn 24:17–24, 1979). Finally, in Part 3, an almost metric version is established for the 2008 Jachymski fixed point result (Proc Am Math Soc 136:1359–1373, 2008) involving Banach contractions over metric spaces endowed with a graph.

Keywords Quasi-ordered metric space • (Globally strong) Picard operator • Fixed point • Admissible function • (Anticipative) Meir–Keeler contraction • Relative Leader contraction • Left-lsc function • Admissible point • Generalized (almost) metric space • Banach contraction • Equivalence class • Graph • Separation • Completeness

Functional Anticipative Contractions in Quasi-Ordered Metric Spaces

Introduction

Let X be a nonempty set. Call the subset $Y \in 2^X$, *almost singleton* (in short: *asingleton*) provided [$y_1, y_2 \in Y$ implies $y_1 = y_2$]; and *singleton* if, in addition, $Y \in (2)^X$; note that in this case, $Y = \{y\}$, for some $y \in X$. [As usual, 2^X denotes the class of all subsets in X; and $(2)^X$ stands for the subclass of all nonempty members in 2^X]. Further, let $d : X \times X \to R_+ := [0, \infty[$ be a *metric* over it; the couple (X, d) will be termed a *metric space*. Finally, let $T \in \mathscr{F}(X)$ be a self-map of X. [Here, for each couple A, B of nonempty sets, $\mathscr{F}(A, B)$ stands for the class of all

M. Turinici (✉)
"A. Myller" Mathematical Seminar, "A. I. Cuza" University, 700506 Iaşi, Romania
e-mail: mturi@uaic.ro

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_28

functions from A to B; when $A = B$, we write $\mathscr{F}(A)$ in place of $\mathscr{F}(A, A)$]. Denote $\mathrm{Fix}(T) = \{x \in X; x = Tx\}$; each point of this set is referred to as *fixed* under T. In the metrical fixed point theory, such points are to be determined according to the context below, comparable with the one in Rus [32, Chap. 2, Sect. 2.2]:

(**pic-1**) We say that T is a *Picard operator* (modulo d) if, for each $x \in X$, the iterative sequence $(T^n x; n \geq 0)$ is d-convergent; and a *globally Picard operator* (modulo d) if, in addition, $\mathrm{Fix}(T)$ is an asingleton

(**pic-2**) We say that T is a *strong Picard operator* (modulo d) if, for each $x \in X$, $(T^n x; n \geq 0)$ is d-convergent with $\lim_n(T^n x) \in \mathrm{Fix}(T)$; and a *globally strong Picard operator* (modulo d) if, in addition, $\mathrm{Fix}(T)$ is an asingleton (hence, a singleton).

In this perspective, two basic answers to the posed question were obtained. Given $\alpha, \beta \geq 0$, let us say that T is *Banach* $(d; \alpha)$*-contractive*, provided

(a01) $d(Tx, Ty) \leq \alpha d(x, y)$, for all $x, y \in X$;

and *Kannan* $(d; \beta)$*-contractive*, provided

(a02) $d(Tx, Ty) \leq \beta[d(x, Tx) + d(y, Ty)]$, for all $x, y \in X$.

Theorem 1. *Suppose that one of the conditions below is fulfilled:*

(i) *T is Banach $(d; \alpha)$-contractive, for some $\alpha \in [0, 1[$*
(ii) *T is Kannan $(d; \beta)$-contractive, for some $\beta \in [0, 1/2[$.*

In addition, let X be d-complete. Then, T is globally strong Picard (modulo d).

The former of these is Banach's contraction principle [3]; and the latter one is Kannan's fixed point statement [17]. These results found a multitude of applications in operator equations theory; so, they were the subject of many extensions. A natural way of doing this is by considering (implicit) "functional" contractive conditions

(a03) $F(d(Tx, Ty), d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)) \leq 0$,
for all $x, y \in X$ with $x \leq y$;

where $F : R_+^6 \to R$ is a function and $(\leq)$ is a *quasi-order* on X. When $(\leq) = X \times X$ (the *trivial quasi-order* on X), some outstanding results in the area have been established by Boyd and Wong [6], Matkowski [23], and Leader [21]; for more details about other possible choices of F, we refer to the 1977 paper by Rhoades [31]. On the other hand, in the case of $(\leq)$ being a *(partial) order* on X, an appropriate extension of Matkowski's result was obtained in the 1986 paper by Turinici [36]. (Note that, two decades later, these results have been re-discovered—at the level of Banach contractive maps—by Ran and Reurings [30]; see also Nieto and Rodriguez-Lopez [28]). Further, an extension—to the same setting—of the above Leader's contribution was performed in Agarwal et al. [1]; and, since then, the number of such papers increased rapidly.

Note that all these conditions are *non-anticipative*; i.e., the left member of (a03) does not contain terms like $d(T^i u, T^j v), u, v \in \{x, y\}$, where $i + j \geq 3$; so, the question arises as to what extent it is possible to have a unitary treatment of such requirements so as to include *anticipative* contractions (in the above sense). It is our aim in the following to give a partial answer to this, via Meir–Keeler techniques [25]. Further aspects will be delineated elsewhere.

Preliminaries

In the following, some preliminary facts about convergent/Cauchy sequences in a metric space and admissible functions are being discussed.

(A) Let (X, d) be a metric space. By a *sequence* in X, we mean any mapping $x : N \to X$ where $N := \{0, 1, \ldots\}$ is the set of *natural* numbers. For simplicity reasons, it will be useful to denote it as $(x(n); n \geq 0)$, or $(x_n; n \geq 0)$; moreover, when no confusion can arise, we further simplify this notation as $(x(n))$ or (x_n), respectively. Also, any sequence $(y_n := x_{i(n)}; n \geq 0)$ with $(i(n); n \geq 0)$ being strictly ascending [hence, $i(n) \to \infty$ as $n \to \infty$], will be referred to as a *subsequence* of $(x_n; n \geq 0)$.

We say that the sequence (x_n) in X, *d-converges* to $x \in X$ (and write: $x_n \xrightarrow{d} x$) iff $[d(x_n, x) \to 0$ as $n \to \infty]$; that is

$$\forall \varepsilon > 0, \exists p = p(\varepsilon), \forall n\colon\ (p \leq n \Longrightarrow d(x_n, x) < \varepsilon).$$

The subset $\lim_n(x_n)$ of all such x is an asingleton, because d is triangular and sufficient; when it is nonempty, (x_n) is called *d-convergent*. In this case, $\lim_n(x_n)$ is a singleton, $\{z\}$; as usual, we write $\{z\} = \lim_n(x_n)$ as $z = \lim_n(x_n)$. Further, call (x_n), *d-Cauchy* provided $[d(x_m, x_n) \to 0$ as $m, n \to \infty, m < n]$; that is

$$\forall \varepsilon > 0, \exists q = q(\varepsilon), \forall (m, n)\colon\ (q \leq m < n \Longrightarrow d(x_m, x_n) < \varepsilon).$$

Clearly, any d-convergent sequence is d-Cauchy too; when the reciprocal holds as well, X is called *d-complete*. Note that any d-Cauchy sequence $(x_n; n \geq 0)$ is *d-semi-Cauchy*; i.e.,

(b01) $\quad d(x_n, x_{n+1}) \to 0$, as $n \to \infty$.

Concerning these concepts, the following fact is useful.

Lemma 1. *The mapping $(x, y) \mapsto d(x, y)$ is d-Lipschitz, in the sense*

$$|d(x, y) - d(u, v)| \leq d(x, u) + d(y, v),\ \ \forall (x, y), (u, v) \in X \times X. \tag{1}$$

As a consequence, this map is d-continuous; i.e.,

$$x_n \xrightarrow{d} x, y_n \xrightarrow{d} y\ \ imply\ \ d(x_n, y_n) \to d(x, y). \tag{2}$$

The proof is immediate, by the properties of $d(.,.)$; so, we omit the details.

(B) Let $\mathscr{F}(re)(R_+)$ denote the class of all $\varphi \in \mathscr{F}(R_+)$ with the *regressive* property: $[\varphi(0) = 0;\ \varphi(t) < t, \forall t > 0]$. Call $\varphi \in \mathscr{F}(re)(R_+)$, *Meir–Keeler admissible*, when it satisfies

(b02) $\forall \gamma > 0, \exists \beta \in]0, \gamma[, (\forall t)\colon (\gamma < t < \gamma + \beta \Longrightarrow \varphi(t) \leq \gamma)$;
or, equivalently (by the choice of our function):
$\forall \gamma > 0, \exists \beta \in]0, \gamma[, (\forall t)\colon (0 \leq t < \gamma + \beta \Longrightarrow \varphi(t) \leq \gamma)$.

This convention is related to the developments in Meir and Keeler [25]. Some important classes of such functions are given below.

(**I**) For any $\varphi \in \mathscr{F}(re)(R_+)$ and any $s \in R^0_+ :=]0, \infty[$, put

(b03) $\Lambda_+\varphi(s) = \inf_{\varepsilon>0} \Phi(s+)(\varepsilon)$; where $\Phi(s+)(\varepsilon) = \sup \varphi(]s, s + \varepsilon[)$
(b04) $\Lambda^+\varphi(s) = \sup\{\varphi(s), \Lambda_+\varphi(s)\}$.

By this very definition, we have the representation (for all $s \in R^0_+$)

$$\Lambda^+\varphi(s) = \inf_{\varepsilon>0} \Phi[s+](\varepsilon);\ \text{where } \Phi[s+](\varepsilon) = \sup\{\varphi([s, s + \varepsilon[), \varepsilon > 0. \quad (3)$$

From the regressive property of φ, these limit quantities are finite; precisely,

$$0 \leq \varphi(s) \leq \Lambda^+\varphi(s) \leq s, \quad \forall s \in R^0_+. \quad (4)$$

The following consequence of this will be useful. Given the sequence $(r_n; n \geq 0)$ in R and the point $r \in R$, let us write

$r_n \to r+$ (respectively, $r_n \to r++$), if $r_n \to r$ and
$r_n \geq r$ (respectively, $r_n > r$), for all $n \geq 0$ large enough.

Lemma 2. *Let $\varphi \in \mathscr{F}(re)(R_+)$ and $s \in R^0_+$ be arbitrary fixed. Then,*

(i) $\limsup_n(\varphi(t_n)) \leq \Lambda^+\varphi(s)$, *for each sequence (t_n) in R^0_+ with $t_n \to s+$; hence, in particular, for each sequence (t_n) in R^0_+ with $t_n \to s++$*
(ii) *there exists a sequence (r_n) in R^0_+ with $r_n \to s+$ and $\varphi(r_n) \to \Lambda^+\varphi(s)$.*

Proof. (i) Given $\varepsilon > 0$, there exists a rank $p(\varepsilon) \geq 0$ such that $s \leq t_n < s + \varepsilon$, for all $n \geq p(\varepsilon)$; hence

$$\limsup_n(\varphi(t_n)) \leq \sup\{\varphi(t_n); n \geq p(\varepsilon)\} \leq \Phi[s+](\varepsilon).$$

It suffices taking the infimum over $\varepsilon > 0$ in this relation to get the desired fact.

(ii) When $\Lambda^+\varphi(s) = 0$, the written conclusion is clear, with $(r_n = s; n \geq 0)$; for, in this case, $\varphi(s) = 0$. Suppose now that $\Lambda^+\varphi(s) > 0$. By definition,

$$\forall \varepsilon \in]0, \Lambda^+\varphi(s)[,\ \exists \delta \in]0, \varepsilon[\colon\ \Lambda^+\varphi(s) - \varepsilon < \Lambda^+\varphi(s) \leq \Phi[s+](\delta) < \Lambda^+\varphi(s) + \varepsilon.$$

This tells us that there must be some r in $[s, s + \delta[$ with

$$\Lambda^+\varphi(s) - \varepsilon < \varphi(r) < \Lambda^+\varphi(s) + \varepsilon.$$

Taking a sequence (ε_n) in $]0, \Lambda^+\varphi(s)[$ with $\varepsilon_n \to 0$, there exists a corresponding sequence (r_n) in R^0_+ with $r_n \to s+$ and $\varphi(r_n) \to \Lambda^+\varphi(s)$.

Call $\varphi \in \mathscr{F}(re)(R_+)$, *Boyd–Wong admissible* [6], if

(b05) $\Lambda^+\varphi(s) < s$ (or, equivalently: $\Lambda_+\varphi(s) < s$), for all $s > 0$.

In particular, $\varphi \in \mathscr{F}(re)(R_+)$ is Boyd–Wong admissible provided it is upper semicontinuous at the right on R^0_+:

$\Lambda^+\varphi(s) = \varphi(s)$ (or, equivalently: $\Lambda_+\varphi(s) \leq \varphi(s)$), $\forall s \in R^0_+$.

Note that this is fulfilled when φ is continuous at the right on R^0_+; for, in such a case, $\Lambda_+\varphi(s) = \varphi(s), \forall s \in R^0_+$. Another example is furnished by a preceding auxiliary fact. Call $\varphi \in \mathscr{F}(re)(R_+)$, *Geraghty admissible* [10], provided

(b06) $(t_n; n \geq 0) \subset R^0_+$ and $\varphi(t_n)/t_n \to 1$ imply $t_n \to 0$.

Lemma 3. *Let $\varphi \in \mathscr{F}(re)(R_+)$ be Geraghty admissible. Then, φ is necessarily Boyd–Wong admissible.*

Proof. Suppose that $\varphi \in \mathscr{F}(re)(R_+)$ is not Boyd–Wong admissible. From a previous relation, there exists $s \in R^0_+$ with $\Lambda^+\varphi(s) = s$. Combining with a preceding fact, there exists a sequence (r_n) in R^0_+ with $r_n \to s+$ and $\varphi(r_n) \to s$; whence $\varphi(r_n)/r_n \to 1$; i.e.: φ is not Geraghty admissible. The obtained contradiction proves our claim.

The reciprocal of this is not in general true. In fact, the function

$(\varphi : R_+ \to R_+)$: $\varphi(t) = t(1 - e^{-t}), t \geq 0$

belongs to $\mathscr{F}(re)(R_+)$ and is continuous; hence, necessarily, Boyd–Wong admissible (see above). On the other hand, taking the sequence $(t_n = n + 1; n \geq 0)$ in R^0_+, we have $\varphi(t_n)/t_n \to 1$ and $t_n \to \infty$; hence, φ is not Geraghty admissible.

(**II**) Call $\varphi \in \mathscr{F}(re)(R_+)$, *Matkowski admissible* [23], provided

(b07) φ is increasing and $\varphi^n(t) \to 0$ as $n \to \infty$, for all $t > 0$.

[Here, φ^n stands for the n-th iterate of φ]. Note that the obtained class of functions is distinct from the above introduced one, as simple examples show.

Now, let us say that $\varphi \in \mathscr{F}(re)(R_+)$ is *Boyd–Wong–Matkowski admissible* (abbreviated: BWM-admissible) if it is either Boyd–Wong admissible or Matkowski admissible. The following auxiliary fact will be useful (cf. Jachymski [13]):

Lemma 4. *Let $\varphi \in \mathscr{F}(re)(R_+)$ be a BWM-admissible function. Then, φ is Meir–Keeler admissible (see above).*

Proof. (i) Suppose that $\varphi \in \mathscr{F}(re)(R_+)$ is Boyd–Wong admissible; and fix $\gamma > 0$. As $\Lambda^+\varphi(\gamma) < \gamma$, we can choose the number $\eta > 0$, with $\Lambda^+\varphi(\gamma) < \eta < \gamma$; whence $\Lambda_+\varphi(\gamma) < \eta < \gamma$. By definition, there exists $\beta = \beta(\eta) > 0$ such that $\gamma < t < \gamma + \beta$ implies $\varphi(t) < \eta < \gamma$; and conclusion follows.

(ii) Assume that $\varphi \in \mathscr{F}(re)(R_+)$ is Matkowski admissible. If the underlying property fails, then (for some $\gamma > 0$):

$$\forall \beta > 0, \exists t \in [0, \gamma + \beta[, \text{ such that } \varphi(t) > \gamma \text{ (hence, } \gamma < t < \gamma + \beta).$$

As φ is increasing, this yields $\varphi(t) > \gamma, \forall t > \gamma$. By induction, we get

$$\varphi^n(t) > \gamma, \text{ for each } n, \text{and each } t > \gamma;$$

hence, taking some $t > \gamma$ and passing to limit as $n \to \infty$, one gets $0 \geq \gamma$; contradiction. This ends the argument.

Statement of the Problem

Let (X, d) be a metric space. Further, let $(\leq)$ be a *quasi-order*; i.e.: a *reflexive* ($x \leq x, \forall x \in X$) and *transitive* ($x \leq y, y \leq z \Longrightarrow x \leq z$) relation over X; then, $(X, d, \leq)$ will be referred to as a *quasi-ordered metric space*. Call the subset Y of X, $(\leq)$-*asingleton*, if $[y_1, y_2 \in Y, y_1 \leq y_2]$ imply $y_1 = y_2$; and $(\leq)$-*singleton* if, in addition, Y is nonempty. Clearly, in the amorphous case (characterized as: $(\leq) = X \times X$), $(\leq)$-asingleton (resp., $(\leq)$-singleton) is identical with asingleton (resp., singleton); but, in general, this cannot be true.

(A) Further, take some $T \in \mathscr{F}(X)$; supposed to satisfy

(c01) T is semi-progressive: $X(T, \leq) := \{x \in X; x \leq Tx\} \neq \emptyset$
(c02) T is increasing: $x \leq y$ implies $Tx \leq Ty$.

We have to determine circumstances under which $\text{Fix}(T)$ be nonempty; and, if this holds, to establish whether T is *fix-*$(\leq)$*-asingleton* (i.e.: $\text{Fix}(T)$ is $(\leq)$-asingleton); or, equivalently: T is *fix-*$(\leq)$*-singleton* (in the sense: $\text{Fix}(T)$ is $(\leq)$-singleton). The basic directions under which the investigations be conducted are shown in our list below, comparable with the one in Turinici [37]:

(**opic-1**) We say that T is a *Picard operator* (modulo $(d, \leq)$) when, for each point $x \in X(T, \leq)$, $(T^n x; n \geq 0)$ is d-convergent; and a *globally Picard operator* (modulo $(d, \leq)$) if, in addition, T is fix-$(\leq)$-asingleton
(**opic-2**) We say that T is a *strong Picard operator* (modulo $(d, \leq)$) when, for each point $x \in X(T, \leq)$, $(T^n x; n \geq 0)$ is d-convergent with $\lim_n(T^n x) \in \text{Fix}(T)$; and a *globally strong Picard operator* (modulo $(d, \leq)$) if, in addition, T is fix-$(\leq)$-asingleton (hence, fix-$(\leq)$-singleton)
(**opic-3**) We say that T is a *Bellman Picard operator* (modulo $(d, \leq)$) when, for each point $x \in X(T, \leq)$, $(T^n x; n \geq 0)$ is d-convergent with $T^n x \leq \lim_n(T^n x) \in \text{Fix}(T), \forall n \geq 0$; and a *globally Bellman Picard operator* (modulo $(d, \leq)$) if, in addition, T is fix-$(\leq)$-asingleton (hence, fix-$(\leq)$-singleton).

In particular, when $(\leq) = X \times X$, the list of such notions is identical with the standard one in Rus [32, Chap. 2, Sect. 2.2]; because, in this case, $X(T, \leq) = X$.

The regularity conditions for such properties are being founded on *ascending orbital* concepts (in short: (a-o)-concepts). Call the sequence $(z_n; n \geq 0)$ in X, *ascending*, if $z_i \leq z_j$ for $i \leq j$; and *T-orbital*, when it is a subsequence of $(T^n x; n \geq 0)$, for some $x \in X$; the intersection of these notions is just the precise one.

(reg-1) Call X, *(a-o,d)-complete*, provided (for each (a-o)-sequence) d-Cauchy $\Longrightarrow$ d-convergent

(reg-2) We say that T is *(a-o,d)-continuous*, if [(z_n)=(a-o)-sequence and $z_n \xrightarrow{d} z$] imply $Tz_n \xrightarrow{d} Tz$

(reg-3) Call $(\leq)$, *(a-o,d)-self-closed*, when [(z_n)=(a-o)-sequence and $z_n \xrightarrow{d} z$] imply [$z_n \leq z, \forall n$]; or, equivalently: the d-limit of each d-convergent (a-o)-sequence is an upper bound of it.

Finally, when the orbital properties are ignored, all (a-o)-concepts above become *a-concepts*; and, if the ascending properties are ignored too, these conventions may be written in the way encountered at introductory part; we do not give details.

(B) Let $(X, d, \leq)$ be a quasi-ordered metric space; and T be a self-map of X; supposed to be semi-progressive and increasing. The fixed points of T are to be determined in a setting we just exposed. It remains to discuss the contractive type conditions to be used. Let $(<)$ stand for the relation

(c03) $x < y$ iff $x \leq y$ and $x \neq y$.

Clearly, $(<)$ is *irreflexive* [$x < x$ is false, for each $x \in X$]; but not transitive, as long as $(\leq)$ is not antisymmetric. Further, denote for each $x, y \in X$,

$A_1(x,y) = d(x,y), A_2(x,y) = (1/2)[d(x,Tx) + d(y,Ty)],$
$A_3(x,y) = \max\{d(x,Tx), d(y,Ty)\}, A_4(x,y) = (1/2)[d(x,Ty) + d(Tx,y)],$
$A_5(x,y) = \max\{d(Tx,T^2x), d(T^2x,y)\}, A_6(x,y) = (1/2)[d(x,T^2x) + d(T^2x,Ty)],$
$\mathscr{A} = \{A_1, A_2, A_3, A_4, A_5, A_6\}.$

By taking all possible maxima between these, one gets $2^6 - 1 = 63$ functions of this type (including the ones we just listed); these may be written as

$\max(H)$, where $H \in (2)^{\mathscr{A}}$ (=the class of all nonempty subsets in $\mathscr{A}$).

Note that, as $A_2 \leq A_3$, some of these maxima are identical; precisely,

$$(\forall H \in (2)^{\mathscr{A}}, A_2, A_3 \in H): \max(H) = \max(H \setminus \{A_2\});$$

however, this is not important for us.

Having these precise, let $P : X \times X \to R_+$ be a map. [The standard choices for it are from the described ones; but, some other choices are also possible]. We say that T is *Meir–Keeler contractive* (modulo $(d, \leq; P)$), if

(c04) $x < y, P(x,y) > 0$ imply $d(Tx,Ty) < P(x,y)$
(T is strictly nonexpansive (modulo $(d, \leq; P)$))

(c05) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, \varepsilon < P(x,y) < \varepsilon + \delta] \Longrightarrow d(Tx, Ty) \leq \varepsilon$
(T has the Meir–Keeler property (modulo $(d, \leq; P)$)).

Note that, by the former of these, the Meir–Keeler property may be written as

(c06) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, 0 < P(x,y) < \varepsilon + \delta] \Longrightarrow d(Tx, Ty) \leq \varepsilon$.

In particular, when $P = A_1$, this convention is just the one due to Meir and Keeler [25], under the technical improvements in Matkowski [24]; see also Ćirić [9].

In the following, two basic examples of such contractions are to be constructed.

(B-1) Given the map $P : X \times X \to R_+$ and the function $\varphi \in \mathscr{F}(R_+)$, call T, $(d, \leq; P, \varphi)$-*contractive*, if

(c07) $d(Tx, Ty) \leq \varphi(P(x,y)), \forall x, y \in X, x < y, P(x,y) > 0$.

Proposition 1. *Assume that T is $(d, \leq; P, \varphi)$-contractive, where $\varphi \in \mathscr{F}(re)(R_+)$ is Meir–Keeler admissible. Then, T is Meir–Keeler contractive (modulo $(d, \leq; P)$).*

Proof. (i) Let $x, y \in X$ be such that $x < y, P(x,y) > 0$. By the contractive condition [and φ=regressive], one has $d(Tx, Ty) < P(x,y)$; so that T is strictly nonexpansive (modulo $(d, \leq; P)$).

(ii) Let $\varepsilon > 0$ be arbitrary fixed; and $\delta > 0$ be the number assured by the Meir–Keeler admissible property of φ. Further, let $x, y \in X$ be such that $x < y$ and $\varepsilon < P(x,y) < \varepsilon + \delta$. By the contractive condition and admissible property,

$$d(Tx, Ty) \leq \varphi(P(x,y)) \leq \varepsilon;$$

so that, T has the Meir–Keeler property (modulo $(d, \leq; P)$).

(B-2) Let us say that (ψ, φ) is a *pair of generalized altering functions* in the class $\mathscr{F}(R_+)$, provided

(c08) ψ is increasing, and φ is reflexive sufficient
(in the sense: $\varphi(0) = 0$ and $\varphi(R^0_+) \subseteq R^0_+$)

(c09) $(\forall \varepsilon > 0)$: there are no sequences $(t_n; n \geq 0)$ in R_+ with $t_n \to \varepsilon++$, such that $\varphi(t_n) \leq \psi(t_n) - \psi(\varepsilon)$, for all $n \geq 0$ large enough

(c10) $(\forall \varepsilon > 0)$: there are no sequences $(t_n; n \geq 0)$ in R_+ with $t_n \to \varepsilon--$, such that $\varphi(\varepsilon) \leq \psi(\varepsilon) - \psi(t_n)$, for all $n \geq 0$ large enough.

Here, for each sequence $(r_n; n \geq 0)$ in R and each $r \in R$, we denoted

$r_n \to r-$ (respectively, $r_n \to r--$), if $r_n \to r$ and
$r_n \leq r$ (respectively, $r_n < r$), for all $n \geq 0$ large enough.

Given the map $P : X \times X \to R_+$ and the couple (ψ, φ) of functions in $\mathscr{F}(R_+)$, let us say that T is $(d, \leq; P, (\psi, \varphi))$-*contractive*, provided

(c11) $\psi(d(Tx, Ty)) \leq \psi(P(x,y)) - \varphi(P(x,y))$,
for all $x, y \in X$ with $x < y, P(x,y) > 0$.

Proposition 2. *Suppose that T is $(d, \leq; P, (\psi, \varphi))$-contractive, for a pair (ψ, φ) of generalized altering functions in $\mathscr{F}(R_+)$. Then, necessarily, T is Meir–Keeler contractive (modulo $(d, \leq; P)$).*

Proof. (i) Let $x, y \in X$ be such that $x < y, P(x, y) > 0$. By the contractive condition (and the choice of our pair), $\psi(d(Tx, Ty)) < \psi(P(x, y))$. This, via [$\psi$=increasing] yields $d(Tx, Ty) < P(x, y)$; so that T is strictly nonexpansive (modulo $(d, \leq; P)$).

(ii) Assume by contradiction that T does not have the Meir–Keeler property (modulo $(d, \leq; P)$); i.e., for some $\varepsilon > 0$,

$$\forall \delta > 0, \exists (x_\delta, y_\delta) \in (<) : \ [\varepsilon < P(x_\delta, y_\delta) < \varepsilon + \delta, \ d(Tx_\delta, Ty_\delta) > \varepsilon].$$

Taking a zero converging sequence (δ_n) in R^0_+, we get a couple of sequences $(x_n; n \geq 0)$ and $(y_n; n \geq 0)$ in X, so as

$$(\forall n) : \ \ x_n < y_n, \ \varepsilon < P(x_n, y_n) < \varepsilon + \delta_n, \ d(Tx_n, Ty_n) > \varepsilon. \tag{5}$$

By the contractive condition (and ψ=increasing), we get

$$\psi(\varepsilon) \leq \psi(P(x_n, y_n)) - \varphi(P(x_n, y_n)), \ \forall n;$$

or, equivalently,

$$(0 <)\ \varphi(P(x_n, y_n)) \leq \psi(P(x_n, y_n)) - \psi(\varepsilon), \ \ \forall n. \tag{6}$$

By (5), the sequence $(t_n := P(x_n, y_n); n \geq 0)$ fulfills $t_n \to \varepsilon + +$; so that the hypothesis about (ψ, φ) will be contradicted. Hence, T has the Meir–Keeler property (modulo $(d, \leq; P)$); and the proof is complete.

(C) These contractive properties must be accompanied with regularity conditions involving the mapping P. The basic ones are described as below.

(**I**) We say that P is $(\leq; T)$-*inf-bounded*, provided

(c12) $\quad x < Tx, Tx < T^2x, x < y, y \neq Ty$ imply $P(x, y) > 0$.

This property holds for $P = A_i$, where $i \in \{1, 2, 3, 5\}$. For, if $x, y \in X$ are taken as in the premise above, then

$$A_1(x, y) = d(x, y) > 0, \ A_2(x, y) \geq \min\{d(x, Tx), d(y, Ty)\} > 0,$$
$$A_3(x, y) \geq \min\{d(x, Tx), d(y, Ty)\} > 0, \ A_5(x, y) \geq d(Tx, T^2x) > 0.$$

As a consequence, this property holds for all $P = \max(H)$ where $H \in (2)^{\mathscr{A}}$ fulfills $H \cap \{A_1, A_2, A_3, A_5\} \neq \emptyset$.

Further, the mappings $P \in \{A_4, A_6, \max(A_4, A_6)\}$ may be sometimes $(\leq; T)$-inf-bounded. In fact, let $x, y \in X$ be again taken as in the premise above:

$$x < Tx, Tx < T^2x, x < y, y \neq Ty.$$

If $P(x, y) = 0$ (where P is taken as before) we anyway have $x = T^2x$; so, replacing in the condition above,

$$x < Tx, \text{ and } Tx < x (= T^2x). \tag{7}$$

Now, if $(<)$ is transitive—which, e.g., happens when $(\leq)$ is antisymmetric (hence, a (partial) ordering on X)—the obtained relation is contradictory; wherefrom, any such mapping $P(.,.)$ is $(\leq; T)$-inf-bounded. But, when $(<)$ is not transitive, a conclusion like this does not seem to be retainable. So, from our general perspective, any maximum function $P = \max(H)$ where $H \in (2)^{\mathscr{A}}$ fulfills $H \subseteq \{A_4, A_6\}$ (hence, 3 such elements) is not endowed with the $(\leq; T)$-inf-bounded property; but, the remaining $63 - 3 = 60$ maximum functions fulfill the $(\leq; T)$-inf-bounded property. [As precise, some of these are identical; but this is not essential for us].

(**II**) Let us say that P is $(\leq; T)$-*orbitally bounded*, provided

(c13) $\quad x < Tx, Tx < T^2x \Longrightarrow P(x, Tx) \leq A_3(x, Tx)$.

For example, any mapping $P \in \{A_1, A_2, A_3, A_5\}$ fulfills such a condition, as it can be directly seen. But, the maps $P \in \{A_4, A_6\}$ also fulfill it. Indeed, for each $x \in X$,

$$A_4(x, Tx) = A_6(x, Tx) = (1/2)d(x, T^2x) \leq \max\{d(x, Tx), d(Tx, T^2x)\};$$

and this proves our claim. Hence, summing up, all functions in $\mathscr{A}$ have such a property; this continues to be valid for each maximum function $P = \max(H)$, where H is a nonempty subset of $\mathscr{A}$.

(**III**) Let us introduce the mapping

$$M(x, y) = \text{diam}\{x, Tx, T^2x, y, Ty\}, x, y \in X.$$

Here, $\text{diam}(Z) = \sup\{d(x, y); x, y \in Z\}$, is the *diameter* of the subset Z in X. We say that P is $(\leq; T)$-*diametral*, provided

(c14) $\quad x < Tx, Tx < T^2x, x < y, y \neq Ty \Longrightarrow P(x, y) \leq M(x, y)$.

Note that all functions $P \in \mathscr{A} = \{A_1, A_2, A_3, A_4, A_5, A_6\}$ have this property, as it can be directly seen. This continues to hold for all maximum functions $P = \max(H)$, where H is a nonempty subset of $\mathscr{A}$.

(**IV**) Finally, let us say that the map $P : X \times X \to R_+$ is $(\leq; T)$-*asymptotic*, provided it fulfills one of the conditions

(c15) $\quad P$ is singular $(\leq; T)$-asymptotic:
for each (a-o)-sequence $(u_n; n \geq 0)$ in X, and each $z \in X$ with
$(u_n \xrightarrow{d} z, Tu_n \xrightarrow{d} z, T^2u_n \xrightarrow{d} z, d(z, Tz) > 0)$,
the relation $[P(x_n, z) \to d(z, Tz)]$ is false

(c16) $\quad P$ is regular $(\leq; T)$-asymptotic:
for each (a-o)-sequence $(u_n; n \geq 0)$ in X, and each $z \in X$ with

$(u_n \xrightarrow{d} z, Tu_n \xrightarrow{d} z, T^2u_n \xrightarrow{d} z, d(z,Tz) > 0, P(u_n,z) \to d(z,Tz))$,
there exists a subsequence $(v_n; n \geq 0)$ of $(u_n; n \geq 0)$ such that
$(P(v_n,z) = d(z,Tz)$, for all $n \geq 0)$.

Proposition 3. *Under these conventions,*

(i) each maximum function $P = \max(H)$*, where* $H \in (2)^{\mathscr{A}}$ *fulfills* $A_3 \notin H$ *is singular* $(\leq; T)$*-asymptotic*

(ii) each maximum function $P = \max(H)$*, where* $H \in (2)^{\mathscr{A}}$ *fulfills* $A_3 \in H$ *is regular* $(\leq; T)$*-asymptotic.*

Proof. (i) Let the (a-o)-sequence $(u_n; n \geq 0)$ in X, and the point $z \in X$ be such that the relations below hold

$$u_n \xrightarrow{d} z, Tu_n \xrightarrow{d} z, T^2u_n \xrightarrow{d} z, d(z,Tz) > 0.$$

By the triangular inequality and continuity of $d(.,.)$ (see above), one derives (under the notation $b := d(z,Tz)$)

$$d(u_n,z), d(Tu_n,z), d(T^2u_n,z) \to 0,$$
$$d(u_n,Tu_n), d(Tu_n,T^2u_n), d(u_n,T^2u_n) \to 0,$$
$$d(u_n,Tz), d(Tu_n,Tz), d(T^2u_n,Tz) \to b.$$

This, by definition, gives

$$A_1(u_n,z) \to 0,\ A_2(u_n,z) \to b/2,\ A_3(u_n,z) \to b,$$
$$A_4(u_n,z) \to b/2, A_5(u_n,z) \to 0, A_6(u_n,z) \to b/2;$$

whence (as $b > 0$): each maximum function $P = \max(H)$ where $H \in (2)^{\mathscr{A}}$ fulfills $A_3 \notin H$, is singular $(\leq; T)$-asymptotic.

(ii) By the same relations we have $P(u_n,z) \to b$, for each maximum function $P = \max(H)$ where $H \in (2)^{\mathscr{A}}$ fulfills $A_3 \in H$. The above convergence properties of $(u_n; n \geq 0)$ tell us that, for a certain rank $n(z) \geq 0$, we must have $\forall n \geq n(z)$

$$d(u_n,z), d(Tu_n,z), d(T^2u_n,z) < b/2,$$
$$d(u_n,Tu_n), d(Tu_n,T^2u_n), d(u_n,T^2u_n) < b/2.$$

This, by the d-Lipschitz property of $d(.,.)$, gives [for all $n \geq n(z)$]

$$|d(u_n,Tz) - b| \leq d(u_n,z) < b/2,$$
$$|d(Tu_n,Tz) - b| \leq d(Tu_n,z) < b/2,$$
$$|d(T^2u_n,Tz) - b| \leq d(T^2u_n,z) < b/2;$$

wherefrom (for the same ranks)

$$b/2 < d(u_n,Tz), d(Tu_n,Tz), d(T^2u_n,Tz) < 3b/2.$$

Combining these, yields, for all $n \geq n(z)$

$$A_1(u_n, z) < b/2 < b,\ A_2(u_n, z) < 3b/4 < b,\ A_3(u_n, z) = b,$$
$$A_4(u_n, z) < b,\ A_5(u_n, z) < b/2 < b,\ A_6(u_n, z) < b;$$

whence (for the same ranks)

$$P(u_n, z) = b, \text{ whenever the maximum function}$$
$$P = \max(H) \text{ with } H \in (2)^{\mathscr{A}} \text{ is such that } A_3 \in H.$$

This shows that these maps are regular $(\leq; T)$-asymptotic; as claimed.

(V) These concepts may be used to solve a structural question concerning our mapping. Namely, let us say that $P : X \times X \to R_+$ is *fix-$(\leq; T)$-separated*, if

(c17) for each $x, y \in \mathrm{Fix}(T)$ with $x < y$, we get $P(x, y) > 0$.

The following auxiliary statement is useful in the sequel.

Proposition 4. *Assume that T is Meir–Keeler contractive (modulo $(d, \leq; P)$), where $P : X \times X \to R_+$ is fix-$(\leq; T)$-separated, $(\leq; T)$-diametral. Then, necessarily, T is fix-$(\leq)$-asingleton.*

Proof. Let $z_1, z_2 \in \mathrm{Fix}(T)$ be such that $z_1 \leq z_2$; and assume (by contradiction) that $z_1 \neq z_2$; hence, $z_1 < z_2$. As P is fix-$(\leq; T)$-separated, we must have $P(z_1, z_2) > 0$. This, by the Meir–Keeler contractive condition, yields $d(z_1, z_2) = d(Tz_1, Tz_2) < P(z_1, z_2)$. On the other hand, as P is $(\leq; T)$-diametral, we must have $P(z_1, z_2) \leq M(z_1, z_2) = d(z_1, z_2)$. The obtained relations are, however, contradictory. As a consequence, $z_1 = z_2$; and our assertion is proved.

For example, $P = A_1$ is fix-$(\leq; T)$-separated; because

$$x, y \in \mathrm{Fix}(T),\ x < y \Longrightarrow A_1(x, y) = d(x, y) > 0.$$

[This property continues to hold for all maximum functions $P = \max(H)$, where $H \in (2)^{\mathscr{A}}$ fulfills $A_1 \in H$. The associated Meir–Keeler contractions (based on such separated maps) will be called *separated*; for example, this is the case with the Banach contractions [3].

It remains now to discuss the alternative of [$P(.,.)$ is fix-$(\leq; T)$-separated] being not accepted. This, technically speaking, amounts to assume that

(c18) P is fix-$(\leq; T)$-nonseparated:
for each $x, y \in \mathrm{Fix}(T)$ with $x < y$, we get $P(x, y) = 0$.

Note that, any $P \in \{A_2, A_3\}$ has such a property; for, evidently,

$$P(x, y) = 0, \text{ for all } x, y \in \mathrm{Fix}(T), \text{ and all such } P(.,.);$$

hence the claim. In this case, the only chance to have the required property is to derive it from contractive type conditions, *against* the singular property imposed upon $P(.,.)$. A possible way of solving this may be described as below. Given $P : X \times X \to R_+$, call T, $(d, \leq; P)$-*compatible* when

(c19) $(\forall x, y \in X, x < y)$: $P(x, y) = 0 \Longrightarrow (Tx = Ty)$.

Lemma 5. *Suppose that T is $(d, \leq; P)$-compatible and P is fix-$(\leq; T)$-nonseparated. Then, T is fix-$(\leq)$-asingleton.*

Proof. Let $z_1, z_2 \in \mathrm{Fix}(T)$ be such that $z_1 \leq z_2$; and assume that $z_1 \neq z_2$; hence $z_1 < z_2$. As P is fix-$(\leq; T)$-singular, $P(z_1, z_2) = 0$; so, by the compatible condition $Tz_1 = Tz_2$; or, equivalently, $z_1 = z_2$; contradiction; hence, the claim.

Concerning the effectiveness of this approach, a direct analysis of the argument we just exposed tells us that a compatible condition like before is without object along the points $x, y \in \mathrm{Fix}(T)$ with $x < y$; for, in such a case, the fix-$(\leq; T)$-nonseparated property of P gives us $x = y$; contrary to the choice of our data. In other words, at the level of fix-$(\leq; T)$-nonseparated maps $P(.,.)$, the compatible type approach is, practically, against their "natural" property; and must be replaced—as an extra condition—with the fix-$(\leq)$-asingleton property of T. On the other hand, under (c18), the Meir-Keeler contractive conditions we just exposed are of no use in deducing the fix-$(\leq)$-asingleton property of T; since these are non-applicable to any pair (x, y) of points in $\mathrm{Fix}(T)$ with $x < y$. The associated Meir–Keeler contractions (based on such nonseparated maps) will be called *nonseparated*; for example, this is the case with Kannan's contraction [17]. Further aspects will be discussed elsewhere.

Main Result

Let $(X, d, \leq)$ be a quasi-ordered metric space. Further, let T be a self-map of X; supposed to be semi-progressive and increasing. We have to determine whether $\mathrm{Fix}(T)$ is nonempty; and, if this holds, to establish whether T is *fix-$(\leq)$-asingleton*; or, equivalently: T is *fix-$(\leq)$-singleton*. The specific directions under which this problem is to be solved were already listed. Sufficient conditions for getting such properties are being founded on the ascending orbital concepts (in short: (a-o)-concepts) we just introduced. Finally, the specific contractive properties to be used have already been analyzed.

The first main result of this exposition is

Theorem 2. *Assume that T is Meir–Keeler contractive (modulo $(d, \leq; P)$), for some mapping $P : X \times X \to R_+$ endowed with the properties: $(\leq; T)$-inf-bounded, $(\leq; T)$-orbitally bounded, and $(\leq; T)$-diametral. Also, suppose that X is (a-o,d)-complete, and T is fix-$(\leq)$-asingleton. Then,*

(I) *T is a globally strong Picard operator (modulo $(d, \leq)$) when, in addition, T is (a-o,d)-continuous*

(II) *T is a globally Bellman Picard operator (modulo $(d, \leq)$) when, in addition, one of the conditions below holds:*

(i) *$(\leq)$ is (a-o,d)-self-closed and P is singular $(\leq; T)$-asymptotic*

(ii) *$(\leq)$ is (a-o,d)-self-closed, P is $(\leq; T)$-asymptotic and T is taken as $(d, \leq; P, \varphi)$-contractive, for a certain Meir–Keeler admissible function $\varphi \in \mathscr{F}(re)(R_+)$*

(iii) *$(\leq)$ is (a-o,d)-self-closed, P is $(\leq; T)$-asymptotic, and T is $(d, \leq; P, (\psi, \varphi))$-contractive, for a certain pair (ψ, φ) of generalized altering functions in $\mathscr{F}(R_+)$.*

Proof. There are several steps to be passed.

Part 1. We claim that, under the conditions in the statement

$$d(Tx, T^2x) < d(x, Tx), \text{ whenever } x < Tx, Tx < T^2x. \tag{8}$$

In fact, let $x \in X$ be as in the premise above. Then,

$$x < Tx, Tx < T^2x, x < y, y < Ty, \text{ where } y = Tx.$$

As P is $(\leq; T)$-inf-bounded, this yields $P(x, Tx) > 0$; so that, by the Meir–Keeler contractive property,

$$d(Tx, T^2x) < P(x, Tx).$$

On the other hand, as P is $(\leq; T)$-orbitally bounded, we must have

$$(0 <)P(x, Tx) \leq A_3(x, Tx) = \max\{d(x, Tx), d(Tx, T^2x)\}.$$

Combining with the preceding relation gives

$$d(Tx, T^2x) < \max\{d(x, Tx), d(Tx, T^2x)\};$$

wherefrom, $d(Tx, T^2x) < d(x, Tx)$; as claimed.
It remains now to establish the strong/Bellman Picard property (modulo $(d, \leq)$). Take some $x_0 \in X(T, \leq)$; and put $(x_n = T^n x_0; n \geq 0)$; this is an ascending orbital sequence. If $x_n = x_{n+1}$ for some $n \geq 0$, we are done; so, without loss, one may assume that, for all $n \geq 0$,

(d01) $x_n \neq x_{n+1}$; hence, $x_n < x_{n+1}$, $\rho_n := d(x_n, x_{n+1}) > 0$.

Part 2. By the preceding developments, we have

$$\rho_{n+1} = d(Tx_n, Tx_{n+1}) < d(x_n, x_{n+1}) = \rho_n, \ \forall n;$$

so, the sequence $(\rho_n; n \geq 0)$ is strictly descending in R_+. As a consequence, $\rho := \lim_n \rho_n$ exists as an element of R_+. Assume by contradiction that $\rho > 0$; and let $\delta > 0$ be the number given by the Meir–Keeler contractive condition (modulo $(d, \leq; P)$) upon T. By definition, there exists a rank $n(\delta) \geq 0$, such that $n \geq n(\delta)$ implies $\rho < \rho_n < \rho + \delta$. This, by the $(\leq, T)$-inf-bounded and $(\leq, T)$-orbitally bounded properties, yields (for the same ranks)

$$0 < P(x_n, x_{n+1}) \leq \max\{\rho_n, \rho_{n+1}\} = \rho_n;$$

wherefrom

$$(x_n < x_{n+1} \text{ and) } 0 < P(x_n, x_{n+1}) < \rho + \delta, \ \forall n \geq n(\delta).$$

Combining with the (variant of) Meir–Keeler property gives

$$\rho < \rho_{n+1} = d(Tx_n, Tx_{n+1}) \leq \rho, \ \forall n \geq n(\delta);$$

a contradiction. Hence, $\rho = 0$; so that, (x_n) is d-semi-Cauchy; in the sense

$$\rho_n := d(x_n, x_{n+1}) = d(x_n, Tx_n) \to 0, \ \text{as } n \to \infty. \tag{9}$$

Part 3. Suppose that

(d02) there exist $i, j \in N$ such that $i < j, x_i = x_j$.

Denoting $p = j - i$, we thus have $p > 0$ and $x_i = x_{i+p}$; so that

$$x_i = x_{i+p}, x_{i+1} = x_{i+p+1}; \text{whence } \rho_i = \rho_{i+p}.$$

This is in contradiction with the strict decreasing property of $(\rho_n; n \geq 0)$. Hence, our working hypothesis cannot hold; wherefrom

$$0 \leq i < j \Longrightarrow x_i < x_j; \ \text{so that, } \ d(x_i, x_j) > 0. \tag{10}$$

Note that, as a consequence of this, the map $n \mapsto x_n$ is injective.

Part 4. We now establish that $(x_n; n \geq 0)$ is a d-Cauchy sequence. Let $\gamma > 0$ be arbitrary fixed; and $\beta > 0$ be the number associated by the Meir–Keeler contractive property (modulo $(d, \leq; P)$); without loss, one may assume that $\beta < \gamma$. As $(x_n; n \geq 0)$ is d-semi-Cauchy, there exists some rank $j(\beta) \geq 0$, with

$$d(x_n, x_{n+1}) < \beta/4 \text{ for all } n \geq j(\beta); \tag{11}$$

whence, by the triangular inequality,

$$d(x_n, x_{n+i}) < \beta/2 \ (< \gamma + \beta/2), \ \ \forall n \geq j(\beta), \ \forall i \in \{1, 2\}. \tag{12}$$

We now claim that

$$(\forall p \geq 1): \ [d(x_n, x_{n+p}) < \gamma + \beta/2, \ \ \forall n \geq j(\beta)]; \tag{13}$$

wherefrom, the desired property of (x_n) follows. To do this, an induction argument upon p is applied. The case $p \in \{1, 2\}$ is clear. Assume that our evaluation holds for $p \in \{1, \ldots, q\}$, where $q \geq 2$; we show that it holds as well for $p = q + 1$. So, let $n \geq j(\beta)$ be arbitrary fixed. Since (cf. a previous relation)

$$x_n < x_{n+1},\ x_{n+1} < x_{n+2},\ x_n < x_{n+q},\ x_{n+q} < x_{n+q+1},$$

we derive (as P is $(\leq; T)$-inf-bounded), $P(x_n, x_{n+q}) > 0$. By the d-semi-Cauchy property and inductive hypothesis,

$$\begin{aligned}
&d(x_{n+i}, x_{n+q}) < \gamma + \beta/2 < \gamma + \beta,\ \forall i \in \{0, 1, 2\}\\
&d(x_{n+i}, x_{n+q+1}) < \gamma + \beta/2 < \gamma + \beta,\ \forall i \in \{1, 2\}\\
&d(x_n, x_{n+1}), d(x_n, x_{n+2}), d(x_{n+1}, x_{n+2}) < \beta/2 < \gamma + \beta\\
&d(x_{n+q}, x_{n+q+1}) < \beta/2 < \gamma + \beta.
\end{aligned}$$

Moreover, from the triangular inequality,

$$d(x_n, x_{n+q+1}) \leq d(x_n, x_{n+q}) + d(x_{n+q}, x_{n+q+1}) < \gamma + \beta;$$

and this, by the $(\leq; T)$-diametral property of P, yields

$$(0 <)P(x_n, x_{n+q}) \leq \text{diam}\{x_n, x_{n+1}, x_{n+2}, x_{n+q}, x_{n+q+1}\} < \gamma + \beta.$$

An application of the Meir–Keeler contractive condition to (x_n, x_{n+q}) (possible, by the above relations) yields

$$d(x_{n+1}, x_{n+q+1}) = d(Tx_n, Tx_{n+q}) \leq \gamma;$$

so that, by the triangular inequality,

$$d(x_n, x_{n+q+1}) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+q+1}) < \gamma + \beta/2;$$

and the assertion follows.

Part 5. As X is (a-o,d)-complete, $x_n \stackrel{d}{\longrightarrow} z$, for some (uniquely determined) $z \in X$. If the alternative below holds

(d03) for each n, there exists $m > n$, with $x_m = z$,

there must be a strictly ascending sequence of ranks $(i(n); n \geq 0)$ such that $x_{i(n)} = z$ (hence, $x_{i(n)+1} = Tz$) for all n; and then [as $(x_{i(n)+1}; n \geq 0)$ is a subsequence of $(x_n; n \geq 0)$], one gets $x_{i(n)+1} \stackrel{d}{\longrightarrow} z$; whence (as d is sufficient), $z = Tz$. So, in the following, we may assume that the opposite alternative is true: there exists $h \geq 0$, such that

(d04) $n \geq h \Longrightarrow x_n \neq z$; hence $0 < d(x_n, z)$.

There are two cases to discuss.

Case A. Suppose that T is (a-o,d)-continuous. Then

$$y_n := Tx_n \xrightarrow{d} Tz \text{ as } n \to \infty.$$

On the other hand, $(y_n = x_{n+1}; n \geq 0)$ is a subsequence of $(x_n; n \geq 0)$; whence $y_n \xrightarrow{d} z$; and this yields (as d is sufficient), $z = Tz$.

Case B. Now, let us assume that $(\leq)$ is (a-o,d)-self-closed. By the convergence property above, it results (via (d04))

$$x_n \leq z, \forall n; \text{ hence } x_n < z, \text{ for all } n \geq h. \tag{14}$$

We now intend to show that the alternative $z \neq Tz$ (hence $b := d(z, Tz) > 0$) yields a contradiction. In fact, the imposed conditions give, $\forall n \geq h$,

$$x_n < x_{n+1}, x_{n+1} < x_{n+2}, x_n < z, z \neq Tz;$$

so that (as P is $(\leq; T)$-inf-bounded),

$$P(x_n, z) > 0, \text{ for all } n \geq h.$$

The Meir–Keeler contractive condition is therefore applicable to (x_n, z), for each $n \geq h$; and yields (for these ranks)

$$d(Tx_n, Tz) < P(x_n, z) \leq \text{diam}\{x_n, Tx_n, T^2x_n, z, Tz\}. \tag{15}$$

As $(Tx_n = x_{n+1}; n \geq 0)$ and $(T^2x_n = x_{n+2}; n \geq 2)$ are subsequences of $(x_n; n \geq 0)$, we have $Tx_n \xrightarrow{d} z, T^2x_n \xrightarrow{d} z$ as $n \to \infty$; so, combining with triangle inequality and the d-continuous property of distance function,

$$\begin{aligned} &d(x_n, Tx_n), d(x_n, T^2x_n), d(Tx_n, T^2x_n) \to 0, \\ &d(x_n, z), d(Tx_n, z), d(T^2x_n, z) \to 0, \\ &d(x_n, Tz), d(Tx_n, Tz), d(T^2x_n, Tz) \to b. \end{aligned} \tag{16}$$

Note that, as a consequence of this

$$\text{diam}\{x_n, Tx_n, T^2x_n, z, Tz) \to b, \text{ as } n \to \infty; \tag{17}$$

and this yields (via (15) above)

$$P(x_n, x) \to b, \text{ as } n \to \infty.$$

Now, two cases are to be discussed.

Case 1. Suppose that P is singular $(\leq; T)$-asymptotic. As a consequence, the convergence relation we just obtained is false. Hence, $b > 0$ is impossible; wherefrom, the conclusion follows.

Case 2. Suppose that T fulfills one of the contractive conditions in the statement. The case of P being singular $(\leq; T)$-asymptotic was already clarified; so, without loss, one may assume that P is regular $(\leq; T)$-asymptotic. By definition, there must be a subsequence $(y_n = x_{i(n)}; n \geq 0)$ of $(x_n; n \geq 0)$ (where $(i(n); n \geq 0)$ is strictly ascending), such that

$$P(y_n, z) = b, \text{ for all } n \geq 0. \tag{18}$$

Note that, as $i(n) \to \infty$ as $n \to \infty$, there is no loss in generality if we suppose that

(d05) $i(n) \geq h$ (hence, $y_n = x_{i(n)} < z$), $\forall n \geq 0$.

Two possibilities occur.

Case 2a. Suppose that T is $(d, \leq; P, \varphi)$-contractive, for a certain Meir–Keeler admissible function $\varphi \in \mathscr{F}(re)(R_+)$. Combining with (18) gives

$$d(Ty_n, z) \leq \varphi(b), \ \forall n \geq 0.$$

So, passing to limit as $n \to \infty$, one gets $b \leq \varphi(b)$; contradiction. Hence, necessarily, $b = 0$; and the claim follows.

Case 2b. Suppose that T is $(d, \leq; P, (\psi, \varphi))$-contractive, for a certain pair (ψ, φ) of generalized altering functions in $\mathscr{F}(R_+)$. Again combining with (18), yields

$$0 < \varphi(b) \leq \psi(b) - \psi(d(Ty_n, Tz)), \ \forall n \geq n(b). \tag{19}$$

But then, the sequence $(t_n := d(Ty_n, Tz); n \geq 0)$ fulfills $t_n \to b--$; in contradiction with the choice of (ψ, φ). So necessarily, $b = 0$; and the proof is complete.

Note that further extensions of these facts to *dislocated metric spaces* (cf. Hitzler [11, Chap. I, Sect. 3]) and *conical metric spaces* taken as in Hyers et al. [12, Chap. 3] are available; we do not give details.

Particular Aspects

Let $(X, d, \leq)$ be a quasi-ordered metric space. Denote by $(<)$ the relation

(e01) $x < y$ iff $x \leq y$ and $x \neq y$;

remember that it is irreflexive but, in general, not transitive. Further, let T be a self-map of X; supposed to be semi-progressive and increasing. We have to determine whether Fix(T) is nonempty; and, if this holds, to establish whether T is *fix-*($\leq$)*-asingleton*; or, equivalently: T is *fix-*($\leq$)*-singleton*. The specific directions under which this problem is to be solved were already listed. Sufficient conditions for getting such properties are being founded on the ascending orbital concepts (in short: (a-o)-concepts) we just introduced. Finally, the specific contractive properties to be used were already considered. As a by-product of all these, we stated our first main result in this exposition. It is our aim in the sequel to expose certain particular cases of it, with a practical relevance.

Remember that, for each $x, y \in X$, we denoted

$$A_1(x,y) = d(x,y), A_2(x,y) = (1/2)[d(x,Tx) + d(y,Ty)],$$
$$A_3(x,y) = \max\{d(x,Tx), d(y,Ty)\}, A_4(x,y) = (1/2)[d(x,Ty) + d(Tx,y)],$$
$$A_5(x,y) = \max\{d(Tx,T^2x), d(T^2x,y)\}, A_6(x,y) = (1/2)[d(x,T^2x) + d(T^2x,Ty)],$$
$$\mathscr{A} = \{A_1, A_2, A_3, A_4, A_5, A_6\}.$$

By taking all possible maxima between these, one gets $2^6 - 1 = 63$ functions of this type (including the ones we just listed); these may be written as

$$P = \max(H), H \in (2)^{\mathscr{A}} \text{ (=the class of all nonempty subsets in } \mathscr{A}\text{)}.$$

[Note that, as $A_2 \leq A_3$, some of these maxima are identical; but, this is not important for us]. Denote also, for simplicity

(e02) $\mathscr{H} = \{H \in (2)^{\mathscr{A}}; H \cap \{A_1, A_2, A_3, A_5\} \neq \emptyset\}$;

this according to our previous developments may be viewed as the generating class of *admissible* (maximum) functions. Precisely, the maximum functions $P = \max(H)$ where $H \in \mathscr{H}$ fulfill all regularity conditions appearing in our main result. Technically speaking, there are two subclasses of $\mathscr{H}$ to be discussed further.

(A) Let in the following denote

(e03) $\mathscr{H}_1 = \{H \in \mathscr{H}; A_1 \in \mathscr{H}\}$;

i.e.: $\mathscr{H}_1$ is the subclass of all elements in $\mathscr{H}$ that contain A_1. As d is sufficient, we have, for each $P = \max(H)$, with $H \in \mathscr{H}_1$ the sufficiency property

$$x, y \in X, x < y \Longrightarrow P(x,y) \geq A_1(x,y) > 0.$$

Given such a mapping $P(.,.)$, call T, *Meir–Keeler contractive* (modulo $(d, \leq; P)$), if

(e04) $x < y$ implies $d(Tx,Ty) < P(x,y)$
(T is strictly nonexpansive (modulo $(d, \leq; P)$))
(e05) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, \varepsilon < P(x,y) < \varepsilon + \delta] \Longrightarrow d(Tx,Ty) \leq \varepsilon$
(T has the Meir–Keeler property (modulo $(d, \leq; P)$)).

Note that, by the former of these, the Meir–Keeler property may be written as

(e06) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, P(x,y) < \varepsilon + \delta] \Longrightarrow d(Tx,Ty) \leq \varepsilon$.

Clearly, the conventions above are equivalent with the previous ones, by means of the sufficiency property we just described. Some basic examples of such contractions were already given in a previous place; we do not give details.

The second main result of this exposition is

Theorem 3. *Assume that T is Meir–Keeler contractive (modulo $(d,\leq;P)$), for some $P=\max(H)$, where $H\in\mathscr{H}_1$. Moreover, let X be (a-o,d)-complete. Then,*

- **(I)** *T is a globally strong Picard operator (modulo $(d,\leq)$) when, in addition, T is (a-o,d)-continuous*
- **(II)** *T is a globally Bellman Picard operator (modulo $(d,\leq)$) when, in addition, one of the conditions below holds:*
- **(i)** *$(\leq)$ is (a-o,d)-self-closed and $H\in\mathscr{H}_1$ fulfills $A_3\notin H$*
- **(ii)** *$(\leq)$ is (a-o,d)-self-closed, and T is $(d,\leq;P,\varphi)$-contractive, for a certain Meir–Keeler admissible function $\varphi\in\mathscr{F}(re)(R_+)$*
- **(iii)** *$(\leq)$ is (a-o,d)-self-closed, and T is $(d,\leq;P,(\psi,\varphi))$-contractive, for a certain pair (ψ,φ) of generalized altering functions in $\mathscr{F}(R_+)$.*

Proof. It will suffice verifying that the first main result is applicable, with the choice $P=\max(H)$, where $H\in\mathscr{H}_1$. There are several steps to be passed.

Step 1. From the sufficiency property, we get that any $P=\max(H)$ with $H\in\mathscr{H}_1$ is fix-$(\leq;T)$-separated; this, along with the $(\leq;T)$-diametral property, assures us, by a preceding auxiliary fact, that T is fix-$(\leq)$-asingleton.

Step 2. By the same sufficiency property,

$$x<Tx, Tx<T^2x, x<y, y\neq Ty \Longrightarrow P(x,y)\geq A_1(x,y)=d(x,y)>0;$$

so that P is $(\leq;T)$-inf-bounded. Moreover, as (for all $x\in X$)

$$A_1(x,Tx), A_2(x,Tx), A_5(x,Tx)\leq A_3(x,Tx),$$
$$A_4(x,Tx)=A_6(x,Tx)=(1/2)d(x,T^2x)\leq A_3(x,Tx),$$

we get, for all such $P=\max(H)$ with $H\in\mathscr{H}_1$,

$$P(x,Tx)\leq A_3(x,Tx),\ \forall x\in X;$$

whence P is $(\leq;T)$-orbitally bounded. Finally, by this very definition, we have [for the same maximum functions P like before]

$$P(x,y)\leq \operatorname{diam}\{x,Tx,T^2x,y,Ty\},\ \forall x,y\in X;$$

and this shows that P is $(\leq;T)$-diametral.

Step 3. By a previous auxiliary fact, we have that

(i) the maximum function $P=\max(H)$ where $H\in\mathscr{H}_1$ fulfills $A_3\notin H$ is singular $(\leq;T)$-asymptotic

(ii) the maximum function $P=\max(H)$ where $H\in\mathscr{H}_1$ fulfills $A_3\in H$ is regular $(\leq;T)$-asymptotic.

Putting these together, it results that the first main result is indeed applicable to these data; and, from this, we are done.

In particular, if φ is Matkowski admissible, this result includes the one in Agarwal et al. [1] for $P = \max(H)$ where $H = \{A_1, A_3, A_4\}$; and, respectively, the one in Turinici [36], in case of $P = \max(H)$, where $H = \{A_1\}$; i.e.: $P = A_1$. Its counterpart, obtained upon φ being Boyd–Wong admissible, seems to be new.

(B) Let in the following denote

(e07) $\mathscr{H}_2 = \mathscr{H} \setminus \mathscr{H}_1$; i.e.: $\mathscr{H}_2 = \{H \in \mathscr{H}; A_1 \notin \mathscr{H}\}$

($\mathscr{H}_2$ is the subclass of all elements of $\mathscr{H}$ that do not contain A_1). Given $P = \max(H)$ where $H \in \mathscr{H}_2$, call T, *Meir–Keeler* $(d, \leq; P)$*-contractive*, if

(e08) $x < y, P(x, y) > 0$ implies $d(Tx, Ty) < P(x, y)$
(T is strictly nonexpansive (modulo $(d, \leq; P)$))
(e09) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, \varepsilon < P(x, y) < \varepsilon + \delta] \Longrightarrow d(Tx, Ty) \leq \varepsilon$
(T has the Meir–Keeler property (modulo $(d, \leq; P)$)).

Note that, by the former of these, the Meir–Keeler property may be written as

(e10) $\forall \varepsilon > 0, \exists \delta > 0$: $[x < y, 0 < P(x, y) < \varepsilon + \delta] \Longrightarrow d(Tx, Ty) \leq \varepsilon$.

The proposed conventions are identical with the old ones. As before, some basic examples of such contractions were given in a previous place.

Our third main result in this exposition is

Theorem 4. *Assume that T is Meir–Keeler contractive (modulo $(d, \leq; P)$), for some $P = \max(H)$, where $H \in \mathscr{H}_2$. In addition, let X be (a-o,d)-complete and T be fix-$(\leq)$-asingleton. Then,*

- **(I)** *T is a globally strong Picard operator (modulo $(d, \leq)$) when, in addition, T is (a-o,d)-continuous*
- **(II)** *T is a globally Bellman Picard operator (modulo $(d, \leq)$) when, in addition, one of the conditions below holds:*
- **(i)** *$(\leq)$ is (a-o,d)-self-closed and $H \in \mathscr{H}_2$ fulfills $A_3 \notin H$*
- **(ii)** *$(\leq)$ is (a-o,d)-self-closed, and T is $(d, \leq; P, \varphi)$-contractive, for a certain Meir–Keeler admissible function $\varphi \in \mathscr{F}(re)(R_+)$*
- **(iii)** *$(\leq)$ is (a-o,d)-self-closed, and T is $(d, \leq; P, (\psi, \varphi))$-contractive, for a certain pair (ψ, φ) of generalized altering functions in $\mathscr{F}(R_+)$.*

Proof. It will suffice verifying that the first main result is applicable, under this choice of our data. There are two steps to be passed.

Step 1. By definition, any $H \in \mathscr{H}$ fulfills $H \cap \{A_1, A_2, A_3, A_5\} \neq \emptyset$. On the other hand, any $H \in \mathscr{H}_2$ fulfills $A_1 \notin H$; so that $H \cap \{A_2, A_3, A_5\} \neq \emptyset$. This tells us that, whenever $x, y \in X$ are such that

$$x < Tx, Tx < T^2x, x < y, y \neq Ty,$$

then, any $P = \min(H)$ where $H \in \mathscr{H}_2$, fulfills one of the inequalities below

$$P(x, y) \geq A_2(x, y) \geq \min\{d(x, Tx), d(y, Ty)\} > 0$$
$$P(x, y) \geq A_3(x, y) \geq \min\{d(x, Tx), d(y, Ty)\} > 0$$
$$P(x, y) \geq A_5(x, y) \geq d(Tx, T^2x) > 0;$$

wherefrom, P is $(\leq; T)$-inf-bounded. Moreover, as (for each $x \in X$)

$$A_2(x, Tx), A_5(x, Tx) \leq A_3(x, Tx)$$
$$A_4(x, Tx) = A_6(x, Tx) = (1/2)d(x, T^2x) \leq A_3(x, Tx),$$

we derive that each P like before is $(\leq; T)$-orbitally bounded. Finally, by the very definition of the composed functions,

$$P(x, y) \leq \text{diam}\{x, Tx, T^2x, y, Ty\}, \ \forall x, y \in X;$$

and this shows that all such P are $(\leq; T)$-diametral.

Step 2. By a previous auxiliary fact, we have that

(i) the maximum function $P = \max(H)$ where $H \in \mathscr{H}_2$ fulfills $A_3 \notin H$ is singular $(\leq; T)$-asymptotic
(ii) the maximum function $P = \max(H)$ where $H \in \mathscr{H}_2$ fulfills $A_3 \in H$ is regular $(\leq; T)$-asymptotic.

Putting these together, the first main result applies in this framework; and the conclusion follows.

Note that, by the developments in a preceding place, the Boyd–Wong admissible functions φ are allowed here. In particular, as established there, any Geraghty admissible function φ is endowed with such a property. In this case, the obtained result includes the one due to Choudhury and Kundu [8], when $P = A_2$. Further aspects may be found in Berzig et al. [4]; see also Turinici [39].

Relative Leader Contractions in Metric Spaces

Introduction

Let X be a nonempty set. Call the subset Y of X, *almost singleton* (in short: *asingleton*), provided $[y_1, y_2 \in Y$ implies $y_1 = y_2]$; and *singleton* if, in addition, Y is nonempty; note that in this case $Y = \{y\}$, for some $y \in X$. Take a metric $d : X \times X \to R_+ := [0, \infty[$ over it; the couple (X, d) will be then termed a *metric space*. Further, let $T \in \mathscr{F}(X)$ be a self-map of X. [Here, for each couple A, B of nonempty sets, $\mathscr{F}(A, B)$ denotes the class of all functions from A to B; when $A = B$, we write $\mathscr{F}(A)$ in place of $\mathscr{F}(A, A)$]. Denote $\text{Fix}(T) = \{x \in X; x = Tx\}$;

each point of this set is referred to as *fixed* under T. Determining of such points is to be performed in the context below, comparable with the one in Rus [32, Chap. 2, Sect. 2.2]:

(**pic-1**) We say that T is a *Picard operator* (modulo d) if, for each $x \in X$, the iterative sequence $(T^n x; n \geq 0)$ is d-convergent; and a *globally Picard operator* (modulo d) if, in addition, $\text{Fix}(T)$ is an asingleton

(**pic-2**) We say that T is a *strong Picard operator* (modulo d) if, for each $x \in X$, $(T^n x; n \geq 0)$ is d-convergent with $\lim_n(T^n x) \in \text{Fix}(T)$; and a *globally strong Picard operator* (modulo d) if, in addition, $\text{Fix}(T)$ is an asingleton (hence, a singleton).

Sufficient metrical conditions for such properties may be given under the lines below. Let $\varphi \in \mathscr{F}(R_+)$ be a function; we say that T is (d, φ)-*contractive*, if

(a01) $d(Tx, Ty) \leq \varphi(d(x, y))$, for all $x, y \in X$.

The problem to be solved is that of establishing the remaining conditions upon $(X, d), T$ and φ, so that T be a (globally) strong Picard operator (modulo d). In this direction, a basic (absolute) result was established in 1968 by Browder [7].

Theorem 5. *Suppose that T is (d, φ)-contractive, where*

(a02) *φ is regressive over R_+: $\varphi(0) = 0$ and $(\varphi(t) < t, \forall t > 0)$*
(a03) *φ is increasing and right continuous on $R_+^0 :=]0, \infty[$ (in the sense: $\varphi(s) = \varphi(s + 0)$, for all $s \in R_+^0$).*

In addition, let X be d-complete. Then, T is globally strong Picard (modulo d).

An outstanding extension of this result is the 1975 one due to Matkowski [23], based on (a03) being substituted by the weaker condition

(a04) φ is increasing and asymptotic on R_+^0 ($\varphi^n(s) \to 0$ as $n \to \infty$,for each $s \in R_+^0$).

Further aspects may be found in Rhoades [31] and the references therein.

In the following, we will be interested in *relative* type versions of Theorem 5. These start from the observation that our contractive condition (a01) is formulated in terms of $P = \{d(x, y); x, y \in X\}$; and the (sequential) iterative techniques to be followed involve its closure, $Q := \text{cl}(P)$. Clearly, the relational chain is available

$$0 \in P \subseteq Q \subseteq R_+;$$

but, the alternatives

(alt-1) $P_0 := P \setminus \{0\}$ is empty (i.e.: $P = \{0\}$) or $P \neq R_+$
(alt-2) $Q_0 := Q \setminus \{0\}$ is empty (i.e.: $Q = \{0\}$) or $Q \neq R_+$

cannot be avoided, in general. In this perspective, a basic answer to the posed question was provided in 1969 by Boyd and Wong [6].

Theorem 6. *Suppose that T is (d, φ)-contractive, where*

(a05) φ *is regressive over* Q: $\varphi(0) = 0$ *and* ($\varphi(t) < t, \forall t \in Q_0$)
(a06) φ *is usc on* Q_0: $\limsup_{t \to c} \varphi(t) \leq \varphi(c)$, *for all* $c \in Q_0$.

In addition, let X be d-complete. Then, T is globally strong Picard (modulo d).

Note that Theorem 6 is a "relative" type result; because the regularity conditions upon φ are stated in terms of the "reduced" codomain P (of $d(.,.)$) and/or its closure, $Q = \text{cl}(P)$. It is our aim in the following to show (in section "Main Result") that a corresponding version of this result is available, by starting from the Leader type contractions [21]

(a07) T is *left* (d, ψ)*-contractive*: $\psi(d(Tx, Ty)) \leq d(x, y), x, y \in X$;

where $\psi \in \mathscr{F}(R_+)$ is a function. Further, in section "Implicit Version," an implicit version of these techniques is proposed. Finally, section "Preliminaries" is preliminary. Some other aspects will be considered elsewhere.

Preliminaries

In the following, some preliminary facts about convergent/Cauchy sequences in a metric space, and admissible functions are being discussed.

(I) Let (X, d) be a metric space. By a *sequence* in X, we mean any mapping $x : N \to X$; where $N := \{0, 1, \ldots\}$ is the set of *natural* numbers. For simplicity reasons, it will be useful to denote it as $(x(n); n \geq 0)$, or $(x_n; n \geq 0)$; moreover, when no confusion can arise, we further simplify this notation as $(x(n))$ or (x_n), respectively. Also, any sequence $(y_n := x_{i(n)}; n \geq 0)$ with $(i(n); n \geq 0)$ being strictly ascending [hence, $i(n) \to \infty$ as $n \to \infty$] will be referred to as a *subsequence* of $(x_n; n \geq 0)$.

We say that the sequence $(x_n; n \geq 0)$ in X, *d-converges* to $x \in X$ (and write: $x_n \xrightarrow{d} x$) iff $d(x_n, x) \to 0$ as $n \to \infty$; that is

(b01) $\forall \varepsilon > 0, \exists p = p(\varepsilon), \forall n$: $(p \leq n \Longrightarrow d(x_n, x) < \varepsilon)$;
or, equivalently,
$\forall \varepsilon > 0, \exists p = p(\varepsilon), \forall n$: $(p < n \Longrightarrow d(x_n, x) \leq \varepsilon)$.

The subset $\lim_n(x_n)$ of all such x is an asingleton, because $d(.,.)$ is triangular and sufficient; when it is nonempty (hence, a singleton), $(x_n; n \geq 0)$ is called *d-convergent*. Then, $\lim_n(x_n) = \{z\}$ (for some $z \in X$); as usual, this relation writes $\lim_n(x_n) = z$. Further, let us say that $(x_n; n \geq 0)$ is *d-Cauchy*, provided $d(x_m, x_n) \to 0$ as $m, n \to \infty, m < n$; that is

(b02) $\forall \varepsilon > 0, \exists q = q(\varepsilon), \forall (m, n)$: $(q \leq m < n \Longrightarrow d(x_m, x_n) < \varepsilon)$;
or, equivalently,
$\forall \varepsilon > 0, \exists q = q(\varepsilon), \forall (m, n)$: $(q < m < n \Longrightarrow d(x_m, x_n) \leq \varepsilon)$.

Clearly, any d-convergent sequence is d-Cauchy too; when the reciprocal holds as well, X is called *d-complete*. On the other hand, each d-Cauchy sequence $(x_n; n \geq 0)$ is *d-semi-Cauchy*; i.e.,

(b03) $d(x_n, x_{n+1}) \to 0$, as $n \to \infty$;
or, equivalently (as $d(.,.)$ is triangular):
$d(x_n, x_{n+i}) \to 0$, as $n \to \infty$, for each $i \geq 1$.

The converse implication is not in general true.

In the following, a useful property is deduced for the d-semi-Cauchy sequences in X which are not d-Cauchy. Let us say that the subset Θ of R^0_+ is $(>)$-*cofinal* in R^0_+, when:

for each $\varepsilon \in R^0_+$, there exists $\theta \in \Theta$ with $\varepsilon > \theta$.

Further, given the sequence $(r_n; n \geq 0)$ in R and the point $r \in R$, let us write

$r_n \to r+$ (respectively, $r_n \to r++$), if $r_n \to r$ and
$r_n \geq r$ (respectively, $r_n > r$), for all $n \geq 0$ large enough.

Proposition 5. *Suppose that the sequence $(x_n; n \geq 0)$ in X with*

(b04) $(x_n; n \geq 0)$ is d-semi-Cauchy ($r_n := d(x_n, x_{n+1}) \to 0$, as $n \to \infty$)

is not endowed with the d-Cauchy property. Further, let Θ be a $(>)$-cofinal part of R^0_+. There exist then a number $b \in \Theta$, a rank $j(b) \geq 0$, and a couple of rank-sequences $(m(j); j \geq 0)$, $(n(j); j \geq 0)$, with

$$(\forall j \geq 0): \ j < m(j) < n(j), \ \textit{and} \quad \alpha_j := d(x_{m(j)}, x_{n(j)}) > b, \ \beta_j := d(x_{m(j)}, x_{n(j)-1}) \leq b \tag{20}$$

$$(\forall j \geq j(b)): \ n(j) - m(j) \geq 2, \ r_{m(j)-1}, r_{n(j)-1} < b/3; \quad \textit{hence,} \ s_j := r_{m(j)-1} + r_{n(j)-1} < 2b/3 \tag{21}$$

$$(\forall j \geq j(b)): \ b < \alpha_j \leq \beta_j + r_{n(j)-1} \leq b + r_{n(j)-1}, \quad \textit{and} \ \alpha_j - s_j \leq \gamma_j := d(x_{m(j)-1}, x_{n(j)-1}) \leq \alpha_j + s_j \tag{22}$$

$$\alpha_j \to b++, \ \textit{and} \ \beta_j, \gamma_j \to b, \ \textit{as} \ j \to \infty. \tag{23}$$

Proof. By a previous remark, the d-Cauchy property of our sequence writes:

$\forall \varepsilon \in R^0_+, \exists k = k(\varepsilon)$: $k < m < n \Longrightarrow d(x_m, x_n) \leq \varepsilon$.

As Θ is $(>)$-cofinal in R^0_+, this property may be also written as

$\forall \theta \in \Theta, \exists k = k(\theta)$: $k < m < n \Longrightarrow d(x_m, x_n) \leq \theta$.

The negation of this property means: there exists $b \in \Theta$ such that

$$(\forall j \geq 0): \ A(j) := \{(m,n) \in N \times N; j < m < n, d(x_m, x_n) > b\} \neq \emptyset.$$

Having this precise, denote, for each $j \geq 0$,

$$m(j) = \min \mathrm{Dom}(A(j)),\ n(j) = \min A(j)(m(j)).$$

The couple of rank-sequences $(m(j); j \geq 0)$, $(n(j); j \geq 0)$ fulfills (20). On the other hand, letting $j(b) \geq 0$ be such that

$$r_i = d(x_i, x_{i+1}) < b/3, \text{ for all } i \geq j(b);$$

it is clear that (21) holds too. Moreover, by the triangular inequality, we have for all $j \geq j(b)$

$$b < \alpha_j \leq \beta_j + r_{n(j)-1} \leq b + r_{n(j)-1} < 4b/3$$
$$|\gamma_j - \alpha_j| \leq s_j = r_{m(j)-1} + r_{n(j)-1} < 2b/3;$$

and this yields (22). Finally, passing to limit as $j \to \infty$ in these relations gives (23). The proof is complete.

In particular, when $\Theta = R^0_+$, this result is, essentially, the one due to Khan et al. [18]; but, the line of argument goes back to Boyd and Wong [6].

(II) Let $\mathscr{F}(pro)(R_+)$ denote the class of all $\psi \in \mathscr{F}(R_+)$, endowed with the *progressive* property: $[\psi(0) = 0;\ \psi(t) > t, \forall t > 0]$. For any $\psi \in \mathscr{F}(pro)(R_+)$ and any $s \in R^0_+$, put

$$\Lambda_+\psi(s) = \sup_{\varepsilon>0} \Psi(s+)(\varepsilon);\text{ where } \Psi(s+)(\varepsilon) = \inf \psi(]s, s+\varepsilon[), \varepsilon > 0;$$
$$\Lambda^+\psi(s) = \inf\{\psi(s), \Lambda_+\psi(s)\}.$$

By this very definition, we have the representation (for all $s \in R^0_+$)

$$\Lambda^+\psi(s) = \sup_{\varepsilon>0} \Psi[s+](\varepsilon);\text{ where } \Psi[s+](\varepsilon) = \inf \psi([s, s+\varepsilon[), \varepsilon > 0.$$

From the progressive property of ψ, these limit quantities are positive; precisely,

$$s \leq \Lambda^+\psi(s) \leq \psi(s), \quad \forall s \in R^0_+.$$

The following consequence of this will be useful.

Proposition 6. *Let $\psi \in \mathscr{F}(pro)(R_+)$ and $s \in R^0_+$ be arbitrary fixed. Then,*

(i) $\liminf_n(\psi(t_n)) \geq \Lambda^+\psi(s)$, *for each sequence (t_n) in R^0_+ with $t_n \to s+$; hence, in particular, for each sequence (t_n) in R^0_+ with $t_n \to s++$*

(ii) there exists a sequence $(r_n; n \geq 0)$ in R^0_+ such that $r_n \to s+$ and $\psi(r_n) \to \Lambda^+\psi(s)$ as $n \to \infty$.

Proof. (i) Given $\varepsilon > 0$, there exists a rank $p(\varepsilon) \geq 0$ such that $s \leq t_n < s + \varepsilon$, for all $n \geq p(\varepsilon)$; hence

$$\liminf_n(\psi(t_n)) \geq \inf\{\psi(t_n); n \geq p(\varepsilon)\} \geq \Psi[s+](\varepsilon).$$

It suffices taking the supremum over $\varepsilon > 0$ in this relation to get the desired fact.

(ii) When $\Lambda^+\psi(s) = \psi(s)$, the written conclusion is clear, with $(r_n = s; n \geq 0)$. Suppose now that $\Lambda^+\psi(s) < \psi(s)$; and let $\varepsilon \in]0, \psi(s) - \Lambda^+\psi(s)[$ be arbitrary fixed. By definition,

$$\exists \delta \in]0, \varepsilon[: \;\; \Lambda^+\psi(s) - \varepsilon < \Psi[s+](\delta) \leq \Lambda^+\psi(s) < \Lambda^+\psi(s) + \varepsilon;$$

wherefrom, there must be some $r = r(\varepsilon)$ in $[s, s + \delta[$, with

$$\Lambda^+\psi(s) - \varepsilon < \psi(r) < \Lambda^+\psi(s) + \varepsilon.$$

Taking a sequence $(\varepsilon_n; n \geq 0)$ in $]0, \psi(s) - \Lambda^+\psi(s)[$ with $\varepsilon_n \to 0$, there exists a corresponding sequence $(r_n = r(\varepsilon_n); n \geq 0)$ in R^0_+ with $r_n \to s+$ and $\varphi(r_n) \to \Lambda^+\psi(s)$; hence, the conclusion.

Call $\psi \in \mathscr{F}(pro)(R_+)$, *Leader admissible* [21], if

(b05) $\quad \Lambda^+\psi(s) > s$ (or, equivalently: $\Lambda_+\psi(s) > s$), for all $s > 0$.

In particular, $\psi \in \mathscr{F}(pro)(R_+)$ is Leader admissible provided it is lower semicontinuous at the right on R^0_+:

$\Lambda^+\psi(s) = \psi(s)$ (or, equivalently: $\Lambda_+\psi(s) \geq \psi(s)$), $\forall s \in R^0_+$.

This, e.g., is fulfilled when ψ is continuous at the right on R^0_+; for, in such a case,

$\Lambda_+\psi(s) = \psi(s)$, for all $s \in R^0_+$.

A useful property of these functions is contained in

Proposition 7. *Assume that $\psi \in \mathscr{F}(pro)(R_+)$ is Leader admissible. Then,*

ψ is compatible: for each sequence $(r_n; n \geq 0)$ in R^0_+ with $(\psi(r_{n+1}) \leq r_n, \forall n)$, one gets $r_n \to 0$.

Proof. Let $(r_n; n \geq 0)$ be as in the premise of this assertion. As ψ is progressive,

$$r_{n+1} < \psi(r_{n+1}) \leq r_n, \text{ for all } n.$$

The sequence $(r_n; n \geq 0)$ is therefore strictly descending; hence, $\gamma := \lim_n(r_n)$ exists in R_+ and $r_n \to \gamma{+}{+}$. In addition, by the same double inequality, $\psi(r_n) \to \gamma$ as $n \to \infty$. Suppose by contradiction that $\gamma > 0$. From a previous auxiliary fact,

$$\gamma = \lim_n \psi(r_n) = \liminf_n \psi(r_n) \geq \Lambda^+\psi(\gamma) > \gamma.$$

The contradiction at which we arrived shows that our working assumption cannot be accepted; whence, $\gamma = 0$; and the proof is complete.

Main Result

Let (X, d) be a metric space. Denote, for simplicity,

$$P = \{d(x, y); x, y \in X\}, Q = \mathrm{cl}(P); P_0 = P \setminus \{0\}, Q_0 := Q \setminus \{0\}.$$

Further, take some $T \in \mathscr{F}(X)$. As precise, the problem to be solved is that of determining sufficient conditions upon these data so that T be a (globally) strong Picard operator (modulo d). The basic concept involved here is to be described as follows. Any (real) sequence $(\xi_n; n \geq 0)$ in P, having the form

(c01) $(\xi_n = d(T^{h(n)}x, T^{k(n)}x); n \geq 0)$,

where $x \in X$ and $(h(n); n \geq 0), (k(n); n \geq 0)$ are rank-sequences with $[h(n) < k(n), \forall n]$, will be called (d, T)-*orbital*; the class of all these will be denoted as $\omega(d, T)$.

Having this precise, we may now list the regularity conditions upon our elements to be used in our investigations. The global one is being founded on *orbital* concepts (in short: o-concepts). Namely, call the sequence $(z_n; n \geq 0)$ in X, T-*orbital*, when it is a subsequence of $(T^n x; n \geq 0)$, for some $x \in X$.

(**reg-1**) Call X, *(o,d)-complete*, provided (for each o-sequence) d-Cauchy $\Longrightarrow$ d-convergent

The local (specific) regularity conditions to be used are constructed with the aid of (d, T)-orbital sequences; and write

(**reg-2**) Call the function $\psi \in \mathscr{F}(R_+)$, *compatible over P* (modulo (d, T)), if

(c02) for each (d, T)-orbital sequence $(\xi_n; n \geq 0)$ in P_0 fulfilling $(\psi(\xi_{n+1}) \leq \xi_n, \forall n)$ we must have $\xi_n \to 0$ as $n \to \infty$.

(**reg-3**) Let Θ be a $(>)$-cofinal part in R_+^0. We say that $\psi \in \mathscr{F}(R_+)$ is *sequentially admissible* over P (modulo $(d, T; \Theta)$), when

(c03) for each (d, T)-orbital sequence $(\xi_n; n \geq 0)$ in P_0, and each point $b \in Q \cap \Theta$ with $\xi_n \to b++$, the relation $[\liminf_n \psi(\xi_n) = b]$ is impossible.

Likewise, $\psi \in \mathscr{F}(R_+)$ is called *sequentially admissible* over P (modulo (d, T)), when it satisfies

(c04) ψ is sequentially admissible over P (modulo $(d, T; \Theta)$), for at least one $(>)$-cofinal part Θ in R_+^0.

To get concrete examples of such functions, note that both these properties hold under the global condition

(c05) ψ is progressive (on R_+) and Leader admissible (see above).

In fact, the former of these follows by a previous auxiliary fact; and the latter one holds with respect to each $(>)$-cofinal subset Θ of R_0. For, let the (d, T)-orbital sequence $(\xi_n; n \geq 0)$ in P_0, and the point $b \in Q \cap \Theta$ be such that $\xi_n \to b++$. Then, according to a previous relation

$$\liminf_n(\psi(\xi_n)) \geq \Lambda^+\psi(b) > b;$$

and the claim follows.

A local version of these facts is to be given in the class $\mathscr{F}(in)(R_+)$ of all increasing functions belonging to $\mathscr{F}(R_+)$. Namely, given $\psi \in \mathscr{F}(R_+)$, let us consider the couple of (local) conditions

(c06) ψ is *progressive on* P: $\psi(0) = 0$ and $(\psi(t) > t, \forall\, t \in P_0)$
(c07) ψ is *progressive on* Q: $\psi(0) = 0$ and $(\psi(t) > t, \forall\, t \in Q_0)$.

Note that, by the former of these, one derives that

$$\psi \text{is } \textit{weakly progressive} \text{ over } P: \ \psi(t) \geq t, \ \text{for all } t \in P. \tag{24}$$

Similarly, from the latter of these, one gets

$$\psi \text{is } \textit{weakly progressive} \text{ over } Q: \ \psi(t) \geq t, \ \text{for all } t \in Q. \tag{25}$$

We are now in position to state the announced answer to our question.

Proposition 8. *Suppose that $\psi \in \mathscr{F}(in)(R_+)$ is progressive on Q. Then,*

(i) *ψ is compatible on P (modulo (d, T))*
(ii) *ψ is sequentially admissible over P (modulo (d, T)).*

Proof. (i) Let the (d, T)-orbital sequence $(\xi_n; n \geq 0)$ in P_0 be such that

$\psi(\xi_{n+1}) \leq \xi_n$, for all n.

As ψ is progressive on P, $(\xi_n; n \geq 0)$ is strictly descending; hence, $\xi := \lim_n \xi_n$ exists (in Q) and $\xi_n > \xi$, for all n. Assume by contradiction that $\xi > 0$. Passing to limit in the above relation gives (via ψ=increasing and $\xi_n \to \xi++$)

$$\psi(\xi + 0) = \lim_n \psi(\xi_n) \leq \xi.$$

This, along with the property (deductible under ψ=progressive on Q)

$$\xi < \psi(\xi) \leq \psi(\xi + 0)$$

yields a contradiction. Hence, $\xi = 0$; and the claim follows.

(ii) By the increasing property of $\psi \in \mathscr{F}(R_+)$, the subset

$\Gamma = \{s \in R_+^0; \psi \text{ is (right or left) discontinuous at } s\}$

appears as (at most) countable; cf. Natanson [27, Chap. 8, Sect. 1]; whence, necessarily, $\Theta := R_+^0 \setminus \Gamma$ is $(>)$-cofinal in R_+^0. Let the (d, T)-orbital sequence

(ξ_n) in P_0 and the point $b \in Q \cap \Theta$ be such that $\xi_n \to b{+}{+}$. As ψ is (bilaterally) continuous at b, we have (by the regressive on Q assumption)

$$(\liminf_n \psi(\xi_n) =) \lim_n \psi(\xi_n) = \psi(b) > b;$$

wherefrom, ψ is sequentially admissible over P (modulo $(d, T; \Theta)$).

Having these precise, we may now state our first main result. Letting $\psi \in \mathscr{F}(R_+)$ be some function, remember that T is called *left* (d, ψ)*-contractive*, provided

(c08) $\psi(d(Tx, Ty)) \leq d(x, y)$, for all $x, y \in X$.

Theorem 7. *Suppose that T is left (d, ψ)-contractive, where $\psi \in \mathscr{F}(R_+)$ is progressive on P and compatible, sequentially admissible over P (modulo (d, T)). In addition, let X be (o, d)-complete. Then, T is globally strong Picard (modulo d).*

Proof. By the contractive property, we have (via ψ=progressive over P)

$$d(Tx, Ty) < d(x, y), \text{ for all } x, y \in X, x \neq y; \tag{26}$$

i.e.: T is *d-strictly-nonexpansive*; and, from this, the asingleton property of $\mathrm{Fix}(T)$ follows. It remains to establish the strong Picard property for T. Fix $x_0 \in X$; and put $(x_n = T^n x_0; n \geq 0)$. If $x_n = x_{n+1}$, for some $n \geq 0$, we are done (in view of $x_n \in \mathrm{Fix}(T)$); so, it remains to discuss the opposite case

(c09) $x_n \neq x_{n+1}$ (i.e., $\tau_n := d(x_n, x_{n+1}) > 0$), $\forall n$.

There are several parts to be passed.

Part 1. By this very condition, $(\tau_n; n \geq 0)$ is a (d, T)-orbital sequence in P_0. On the other hand, from the contractive condition (and ψ=progressive on P), one gets (by the working condition above)

$$(\tau_{n+1} <) \ \psi(\tau_{n+1}) \leq \tau_n, \ \forall n.$$

As ψ is compatible over P, it results that

$$(\tau_n =) d(x_n, x_{n+1}) \to 0, \text{ as } n \to \infty; \tag{27}$$

i.e.: $(x_n; n \geq 0)$ is d-semi-Cauchy.

Part 2. By the imposed conditions, ψ is sequentially admissible over P (modulo $(d, T; \Theta)$), for a certain $(>)$-cofinal subset Θ in R^0_+. We now show that $(x_n; n \geq 0)$ is d-Cauchy. Suppose by contradiction that this is not true. From a previous auxiliary fact, there exist then a number $b \in \Theta$, a rank $j(b) \geq 0$, and a couple of rank-sequences $(m(j); j \geq 0)$, $(n(j); j \geq 0)$, with the properties (20)–(23). In particular, this gives

$$(j < m(j) < n(j),\ \forall j),\ (n(j) - m(j) \geq 2,\ \forall j \geq j(b)) \tag{28}$$

$$\alpha_j := d(x_{m(j)}, x_{n(j)}) \to b++,\ \text{as}\ j \to \infty \tag{29}$$

$$\gamma_j := d(x_{m(j)-1}, x_{n(j)-1}) \to b,\ \text{as}\ j \to \infty. \tag{30}$$

Note that, by the properties above, $(\alpha_j; j \geq 0)$ and $(\gamma_j; j \geq 0)$ are (d, T)-orbital sequence in P_0 and P, respectively; moreover, by definition, $b \in Q \cap \Theta$. From the contractive condition (and ψ=progressive),

$$\alpha_j < \psi(\alpha_j) \leq \gamma_j,\ \forall j.$$

So, passing to limit as $j \to \infty$, one gets

$$b \leq \lim_j \psi(\alpha_j) \leq b;\ \text{hence},\ (\liminf_j \psi(\alpha_j) =) \lim_j \psi(\alpha_j) = b; \tag{31}$$

in contradiction with ψ being sequentially admissible over P (modulo $(d, T; \Theta)$). Hence, $(x_n; n \geq 0)$ is a d-Cauchy sequence in X.

Part 3. As X is (o, d)-complete, $x_n \xrightarrow{d} z$ for some (uniquely determined) $z \in X$. Since $(y_n := x_{n+1}; n \geq 0)$ is a subsequence of $(x_n; n \geq 0)$, we also have $y_n \xrightarrow{d} z$ as $n \to \infty$. On the other hand, by the d-strict-nonexpansive property of T, we get

$$T \text{ is } d\text{-nonexpansive: } d(Tx, Ty) \leq d(x, y), \text{ for all } x, y \in X.$$

A direct application of this to our data yields

$$d(y_n, Tz) \leq d(x_n, z),\ \text{for all}\ n.$$

Combining with $d(x_n, z) \to 0$, gives

$$d(y_n, Tz) \to 0 \text{ (whence, } y_n \xrightarrow{d} Tz\text{), as } n \to \infty.$$

Combining these gives (as d=metric) $z = Tz$. The proof is complete.

Note that, by a preceding discussion, Theorem 7 is just the (orbital) "left" variant of Theorem 6. This is also true for the related contribution in Akkouchi [2]; as results from the developments in Jachymski [15]. On the other hand, Theorem 7 includes as well the (orbital) "left" (global) version of Matkowski's fixed point result [23]. However, the contractive condition in Meir and Keeler [25] does not seem to be included here; further aspects will be discussed in a separate paper.

Implicit Version

A close analysis of the reasoning above shows that, ultimately, the contradictory character of sequential relations appearing at different stages of the proof is given by the very existence of admissible sequences constructed there; and not by the limit properties deduced from them. This remark allows us to establish an implicit extension of Theorem 7 above.

Let (X, d) be a metric space; and fix some $T \in \mathscr{F}(X)$. Denote, as before

$$P = \{d(x, y); x, y \in X\}, Q = \text{cl}(P), P_0 = P \setminus \{0\}, Q_0 := Q \setminus \{0\}.$$

Further, let Δ be a nonempty subset of $P \times P$; hence, a relation over P. We say that T is an *implicit Δ-contraction*, if

(d01) $(d(Tx, Ty), d(x, y)) \in \Delta, \forall x, y \in X.$

It is our aim in the following to get sufficient regularity conditions under which the fixed point problem (under the lines we just described) has a positive solution. The general one writes

(**ireg-1**) Let us say that (the relation) Δ is *super-diagonal*, provided

(d02) $(s, t) \in \Delta$ and $t > 0$ imply $s < t$.

Note that the case $t = 0$ is not entering in this convention.

For the local (specific) ones, remember that any (real) sequence $(\xi_n; n \geq 0)$ in P, having the form

$$(\xi_n = d(T^{h(n)}x, T^{k(n)}x); n \geq 0),$$

where $x \in X$ and $(h(n); n \geq 0), k(n); n \geq 0)$ are rank-sequences with $[h(n) < k(n), \forall n]$, will be called *(d, T)-orbital*; the class of all these will be denoted as $\omega(d, T)$. Now, the announced regularity conditions upon our data write

(**ireg-2**) Let us say that (the relation) Δ is *compatible* (modulo (d, T)), when

(d03) each (d, T)-orbital sequence $(\xi_n; n \geq 0)$ in P_0 with $((\xi_{n+1}, \xi_n) \in \Delta, \forall n)$, fulfills $\xi_n \to 0$ as $n \to \infty$.

(**ireg-3**) Let Θ be a $(>)$-cofinal part in R^0_+. We say that (the relation) Δ is *distinguished* (modulo $(d, T; \Theta)$), when

(d04) for each couple $(\xi_n; n \geq 0), (\zeta_n; n \geq 0)$ of (d, T)-orbital sequences in P_0 and P respectively, and each $b \in Q \cap \Theta$ with $(\xi_n \to b++, \zeta_n \to b)$, the property $(\xi_n, \zeta_n) \in \Delta$ cannot hold for infinitely many n.

Likewise, Δ will be termed *distinguished* (modulo (d, T)), when

(d05) Δ is distinguished (modulo $(d, T; \Theta)$), for at least one $(>)$-cofinal subset Θ of R^0_+.

Some concrete examples of relations Δ for which these conditions hold will be given a bit further. For the moment, we are interested to state our second main result of this exposition.

Theorem 8. *Assume that T is an implicit Δ-contraction, where the relation $\Delta \subseteq P \times P$ is super-diagonal and compatible, distinguished (modulo (d, T)). In addition, let X be (o, d)-complete. Then, T is globally strong Picard (modulo d).*

Proof. We follow the basic lines of our previous reasoning.

By the contractive condition and super-diagonal property of Δ, we have that T is d-strictly-nonexpansive (see above); wherefrom, the singleton property of $\text{Fix}(T)$ follows. It remains then to establish the strong Picard property (modulo d). Fix $x_0 \in X$; and put $(x_n = T^n x_0; n \geq 0)$. If $x_n = x_{n+1}$, for some $n \geq 0$, we are done (in view of $x_n \in \text{Fix}(T)$); so, it remains to discuss the opposite case

$x_n \neq x_{n+1}$ (i.e., $\tau_n := d(x_n, x_{n+1}) > 0$), $\forall n$.

Three steps must be passed.

Step 1. By this very condition, $(\tau_n; n \geq 0)$ is (d, T)-orbital in P_0. On the other hand, by the contractive condition,

$$(\tau_{n+1}, \tau_n) \in \Delta, \text{ for all } n.$$

These, along with Δ being compatible (modulo (d, T)), give

$$\tau_n \to 0 \text{ as } n \to \infty; \text{ whence, } (x_n; n \geq 0) \text{ is } d\text{-semi-Cauchy.}$$

Step 2. By the imposed condition, there exists a $(>)$-cofinal subset Θ in R^0_+, such that Δ is distinguished (modulo $(d, T; \Theta)$). We now show that $(x_n; n \geq 0)$ is d-Cauchy. Suppose by contradiction that this is not true. From a previous auxiliary fact, there exist then a number $b \in \Theta$, a rank $j(b) \geq 0$, and a couple of rank-sequences $(m(j); j \geq 0)$, $(n(j); j \geq 0)$, with the properties (20)–(23). In particular, this gives

$$(j < m(j) < n(j),\ \forall j),\ (n(j) - m(j) \geq 2,\ \forall j \geq j(b)) \tag{32}$$

$$\alpha_j := d(x_{m(j)}, x_{n(j)}) \to b++,\ \text{as } j \to \infty \tag{33}$$

$$\gamma_j := d(x_{m(j)-1}, x_{n(j)-1}) \to b,\ \text{as } j \to \infty. \tag{34}$$

Note that, by the properties above, $(\alpha_j; j \geq 0)$ and $(\gamma_j; j \geq 0)$ are (d, T)-orbital sequences in P_0 and P, respectively; moreover, by definition, $b \in Q \cap \Theta$. From the contractive condition, we have

$$(\alpha_j, \gamma_j) \in \Delta,\ \text{for all } j \geq 0. \tag{35}$$

This, along with Δ being distinguished (modulo $(d, T; \Theta)$), yields a contradiction. Hence, $(x_n; n \geq 0)$ is d-Cauchy, as claimed.

Step 3. As X is (o,d)-complete, $x_n \overset{d}{\longrightarrow} z$ for some $z \in X$. Since $(y_n := x_{n+1}; n \geq 0)$ is a subsequence of $(x_n; n \geq 0)$, we also have $y_n \overset{d}{\longrightarrow} z$. On the other hand, by the strict d-nonexpansive property of T, we have that T is d-nonexpansive (see above). A direct application of this to our data yields

$$d(y_n, Tz) \le d(x_n, z), \text{ for all } n.$$

Combining with $d(x_n, z) \to 0$ gives $y_n \xrightarrow{d} Tz$. Putting these convergence relations together yields (as d=metric) $z = Tz$. The proof is thereby complete.

In particular, when $\Delta \subseteq P \times P$ is taken as

(d06) $(u, v) \in \Delta$ iff $\psi(u) \le v$,

where $\psi \in \mathscr{F}(R_+)$ is progressive on P and compatible, sequentially admissible (modulo (d, T)), this result is nothing else than Theorem 7. Another choice is

(d07) $(u, v) \in \Delta$ iff $u \le \varphi(v)$,

where $\varphi \in \mathscr{F}(R_+)$ fulfills the regularity conditions in Theorem 6. Then, Theorem 8 may be viewed as an orbital extension of the quoted result.

Note, finally, that Theorem 8 may be extended to contractive conditions like

(d08) $(d(Tx, Ty), d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx)) \in \Delta$;

where Δ is a subset of P^6, subjected to appropriate conditions. In this case, the obtained statement includes the related one in Turinici [34]; as well as the fixed point result (involving altering metrics) due to Khan et al. [18]. Further aspects may be found in the 2001 survey paper due to Kirk [19].

Banach Contractions in Relational Metric Spaces

Introduction

Let X be a nonempty set. By a *sequence* in X, we mean any mapping $x : N \to X$; where $N := \{0, 1, \ldots\}$ is the set of *natural* numbers. For simplicity reasons, it will be useful to denote it as $(x(n); n \ge 0)$, or $(x_n; n \ge 0)$; moreover, when no confusion can arise, we further simplify this notation as $(x(n))$ or (x_n), respectively. Also, any sequence $(y_n := x_{i(n)}; n \ge 0)$ with $(i(n); n \ge 0)$ being strictly ascending [hence, $i(n) \to \infty$ as $n \to \infty$] will be referred to as a *subsequence* of $(x_n; n \ge 0)$.

Call the subset $Y \in 2^X$, *almost singleton* (in short: *asingleton*) provided [$y_1, y_2 \in Y$ implies $y_1 = y_2$]; and *singleton* if, in addition, $Y \in (2)^X$; note that, in this case, $Y = \{y\}$, for some $y \in X$. (Here, 2^X stands for the class of all subsets in X; and $(2)^X$, for the subclass of all nonempty members in 2^X). Further, let $d : X \times X \to R_+ := [0, \infty[$ be a *metric* over it (in the usual sense); the couple (X, d) will be termed a *metric space*. Finally, let $T \in \mathscr{F}(X)$ be a self-map of X. [Here, for each couple A, B of nonempty sets, $\mathscr{F}(A, B)$ stands for the class of all functions from A to B; when $A = B$, we write $\mathscr{F}(A)$ in place of $\mathscr{F}(A, A)$]. Denote $\text{Fix}(T) = \{x \in X; x = Tx\}$;

each point of this set is referred to as *fixed* under T. In the metrical fixed point theory, such points are to be determined according to the context below, comparable with the one in Rus [32, Chap. 2, Sect. 2.2]:

(**pic-1**) Call $x \in X$, a *Picard point* (modulo (d, T)), provided the iterative sequence $(T^n x; n \geq 0)$ is d-convergent; if this holds for all $x \in X$, we say that T is a *Picard operator* (modulo d); and, if (in addition) $\mathrm{Fix}(T)$ is an asingleton, then T is referred to as a *globally Picard operator* (modulo d)

(**pic-2**) Call $x \in X$, a *strong Picard point* (modulo (d, T)), provided the iterative sequence $(T^n x; n \geq 0)$ is d-convergent with $\lim_n (T^n x) \in \mathrm{Fix}(T)$; if this holds for all $x \in X$, we say that T is a *strong Picard operator* (modulo d); and, if (in addition) $\mathrm{Fix}(T)$ is an asingleton (hence, a singleton), then T is referred to as a *globally strong Picard operator* (modulo d).

In this perspective, a basic answer to the posed question is the 1922 one due to Banach [3]. Given $\alpha > 0$, let us say that T is $(d; \alpha)$-*contractive*, provided

(a01) $d(Tx, Ty) \leq \alpha d(x, y)$, for all $x, y \in X$.

Theorem 9. *Suppose that T is Banach $(d; \alpha)$-contractive, for some $\alpha \in]0, 1[$. In addition, let $d(., .)$ be complete on X. Then, T is a globally strong Picard operator (modulo d).*

This result (referred to as: Banach's contraction principle; in short: BCP) found a multitude of applications in operator equations theory; so, it was the subject of many extensions. A natural way of doing this is by considering "functional" contractive conditions like

(a02) $d(Tx, Ty) \leq F(d(x, y), d(x, Tx), d(y, Ty), d(x, Ty), d(y, Tx))$, for all $x, y \in X$;

where $F : R_+^5 \to R_+$ is an appropriate function; for more details about the possible choices of F, we refer to the 1977 paper by Rhoades [31]. Another way of extension is that of conditions imposed upon d being modified; this, technically speaking, is to be described in a generalized metrical context.

Let again X be a nonempty set. By a *generalized metric* over X we shall mean any map $e : X \times X \to R_+ \cup \{\infty\}$, endowed with the properties: *reflexive* $[e(x, x) = 0, \forall x \in X]$, *triangular* $[e(x, z) \leq e(x, y) + e(y, z), \forall x, y, z \in X]$, *sufficient* $[x, y \in X, e(x, y) = 0 \implies x = y]$, and *symmetric* $[e(x, y) = e(y, x), \forall x, y \in X]$; the couple (X, e) will be then referred to as a *generalized metric space*. Fix such an object $e(., .)$ in the sequel. The relation (over X)

(a03) $(x, y \in X)$; $x \approx y$ iff $e(x, y) < \infty$

is reflexive, symmetric and transitive; hence, an *equivalence*. For each $a \in X$, let $X(a, \approx) := \{x \in X; a \approx x\}$ stand for the *equivalence class* of a; note that the restriction of e to $X(a, \approx)$ is a standard metric. Conversely, let $(\sim)$ be an equivalence relation over X; as usual, we identify it with its graph in $X \times X$. Further, let $\Delta : (\sim) \to R_+$ be a mapping; endowed with the properties: $(\sim)$-*reflexive* $[\Delta(x, x) = 0, \forall x \in X]$, $(\sim)$-*triangular* $[\Delta(x, z) \leq \Delta(x, y) + \Delta(y, z)$, if $x \sim y, y \sim z]$, $(\sim)$-*sufficient*

$[x \sim y, \Delta(x,y) = 0 \Longrightarrow x = y]$, and $(\sim)$-*symmetric* $[\Delta(x,y) = \Delta(y,x)$, whenever $x \sim y]$. In this case, the mapping (from $X \times X$ to $R_+ \cup \{\infty\}$)

(a04) $e(x,y) = \Delta(x,y$, if $x \sim y$; $e(x,y) = \infty$, otherwise,

is a generalized metric on X, as it can be directly seen.

Now, given a generalized metric space (X,e), we may introduce an e-convergence and e-Cauchy structure on X as follows. For each sequence $(x_n; n \geq 0)$ in X and $x \in X$, put $x_n \xrightarrow{e} x$ if $e(x_n, x) \to 0$ as $n \to \infty$; i.e.:

$$\forall \delta > 0, \exists i = i(\delta) \colon i \leq n \Longrightarrow e(x_n, x) < \delta.$$

The set $\lim_n(x_n)$ of all such $x \in X$ is an asingleton, as it can be directly seen; when it is nonempty—hence, a singleton, $\{z\}$—we say that (x_n) is *e-convergent*; then, as usual, $\{z\} = \lim_n(x_n)$ will be written as $z = \lim_n(x_n)$. Likewise, the *e-Cauchy* property of (x_n) is depicted as: $e(x_m, x_n) \to 0$ as $m, n \to \infty, m < n$; i.e.:

$$\forall \delta > 0, \exists j = j(\delta) \colon j \leq m < n \Longrightarrow e(x_m, x_n) < \delta.$$

By the properties of $e(.,.)$, any e-convergent sequence is e-Cauchy too; when the reciprocal of this holds too, we say that $e(.,.)$ is *complete* on X.

The following local result is our starting point. Let Y be a subset of X. We say that $u \in X$ is *e-adherent* to Y, when

(a05) $(\forall \delta > 0, \exists x_\delta \in Y \colon e(x_\delta, u) < \delta)$; or, equivalently:
$u = \lim_n(x_n)$, for some sequence $(x_n; n \geq 0)$ in Y.

The set of all such points will be called the *e-closure* of Y; and denoted as $\mathrm{cl}_e(Y)$; when $Y = \mathrm{cl}_e(Y)$, then Y will be called *e-closed*.

Proposition 9. *Let (X,e) be a generalized metric space; and $a \in X$ be arbitrary fixed. Then,*

- **(i)** *The equivalence class $X_a := X(a, \approx)$ is e-closed (see above)*
- **(ii)** *The restriction of $e(.,.)$ to X_a is a standard metric (on X_a)*
- **(iii)** *The same restriction is complete on X_a, whenever the initial generalized metric $e(.,.)$ is complete on X.*

[A direct proof of these assertions is immediately available; so, we do not give further details].

As a consequence of this statement, it results that a *local* version of the (global) Picard conventions above is needed here.

Let (X,e) be a generalized metric space. Call $U \in 2^X$, *e-asingleton* provided $[y_1, y_2 \in U, e(y_1,y_2) < \infty$ imply $y_1 = y_2]$; and *e-singleton* if, in addition, $U \in (2)^X$. [Note that, in this last case, $\mathrm{card}(U) > 1$ is not avoidable; for example, the (nonempty) subset $U = \{u_1, u_2\}$ of X where $e(u_1,u_2) = \infty$ (hence, $u_1 \neq u_2$) is an e-singleton with $\mathrm{card}(U) = 2$]. Further, let $T \in \mathscr{F}(X)$ be a self-map of X. The following *initial* regularity conditions for these data are imposed:

(**reg-1**) The generalized metric $e(.,.)$ is called *complete*, provided: each e-Cauchy sequence in X is e-convergent. Note that this concept is equivalent with its local version. Precisely, call $e(.,.)$, *locally complete* when, for each $a \in X$, its restriction to the equivalence class $X(a, \approx)$ is complete. By the preceding statement, we have (for the generalized metric e): complete $\Longrightarrow$ locally complete. The converse inclusion is also true; so that (summing up) we have

$$(\forall \text{ generalized metric } e)\text{: complete} \Longleftrightarrow \text{locally complete.}$$

(**reg-2**) Call T, *e-semi-progressive* provided

(a06) $X(T; e) := \{x \in X; e(x, Tx) < \infty\}$ is not empty.

This may be also referred to as

$X(T; \approx) := \{x \in X; x \approx Tx\}$ is not empty;

for, evidently, $X(T, e) = X(T, \approx)$.

(**reg-3**) Call T, *e-increasing* provided

(a07) $e(x, y) < \infty$ implies $e(Tx, Ty) < \infty$.

To explain the terminology, note that it may be also written as

T is $(\approx)$-increasing: $x \approx y$ implies $Tx \approx Ty$.

Fix in the following some $Y \in (2)^X$; and denote $\text{Fix}(T; Y) = \text{Fix}(T) \cap Y$ (the trace over Y of $\text{Fix}(T)$). In addition, take some $W \in (2)^X$. Sufficient conditions for existence of such points involve the local concepts below:

(**loc-pic-1**) Call $x \in X$, a *Picard point* (modulo (e, T)) relative to Y, provided $(T^n x; n \geq 0)$ is e-convergent with $\lim_n(T^n x) \in Y$. If such a property holds for all $x \in W$, we say that T is a *Picard operator* (modulo e) relative to (W, Y); and, if (in addition) $\text{Fix}(T; Y)$ is an e-asingleton, then T is referred to as a *globally Picard operator* (modulo e) relative to the couple (W, Y)

(**loc-pic-2**) Call $x \in X$, a *strong Picard point* (modulo (e, T)), relative to Y, provided $(T^n x; n \geq 0)$ is e-convergent with $\lim_n(T^n x) \in \text{Fix}(T; Y)$. If this holds for all $x \in W$, we say that T is a *strong Picard operator* (modulo e) relative to (W, Y); and, if (in addition) $\text{Fix}(T; Y)$ is an e-asingleton (hence, an e-singleton), then T is referred to as a *globally strong Picard operator* (modulo e) relative to the couple (W, Y).

The contractive condition upon our data may be written as:

(**con-gm**) Given $\alpha > 0$, let us say that T is $(e; \alpha)$-*contractive*, provided

(a08) $e(Tx, Ty) \leq \alpha e(x, y)$, for each $x, y \in X$ with $e(x, y) < \infty$.

Note that, as a consequence, T is $(\approx)$-increasing; but, this is not important for us.

The following fixed point result due to Luxemburg [22] and Jung [16] (referred to as: Luxemburg-Jung contraction principle; in short: LJCP) is then available:

Theorem 10. *Assume that the (e-semi-progressive, e-increasing) self-map $T \in \mathscr{F}(X)$ is $(e;\alpha)$-contractive, for some $\alpha \in]0,1[$. In addition, let $e(.,.)$ be complete, and fix some point $x_0 \in X(T,e) = X(T,\approx)$. Then, the following conclusions hold:*

(Con-1) the equivalence class $X(x_0,\approx)$ is e-closed, T-invariant, and the generalized metric $e(.,.)$ is complete on $X(x_0,\approx)$
(Con-2) x_0 is a strong Picard point (modulo (e,T)) relative to $X(x_0,\approx)$
(Con-3) T is a globally strong Picard operator (modulo e), relative to the couple $(X(x_0,\approx), X(x_0,\approx))$.

In particular, when $e(.,.)$ is a standard metric on X, this result is just the Banach contraction principle we already stated. Some functional extensions of these facts may be found in Turinici [35].

Now, further extensions of the Banach contraction principle were obtained by Nieto and Rodriguez-Lopez [29] (the (partially) ordered case) or Jachymski [14] (the relational case). So, we may ask of which relationships exist between these and the Luxemburg-Jung contraction principle above. It is our aim in the following to establish (in section "Particular Aspects") that these are reducible to certain "non-symmetric" versions of LJCP (exposed in section "Main Result"). The specific tools for deducing these conclusions are described in section "Preliminaries." Some other aspects of these developments will be delineated elsewhere.

Preliminaries

Let X be a nonempty set. By a *relation* over it we mean any (nonempty) part $\mathscr{R}$ of $X \times X$; sometimes, we write $x\mathscr{R}y$ in place of $(x,y) \in \mathscr{R}$, for simplicity. In particular, $\mathscr{I} := \{(x,x); x \in X\}$ (the *diagonal* of X) is such an object, referred to as: the *identity* relation. Any relation $\mathscr{R}$ with

(b01) $\mathscr{I} \subseteq \mathscr{R}$ (i.e.: $x\mathscr{R}x$, $\forall x \in X$)

is called *reflexive*; note that, its x-section $X(x,\mathscr{R}) = \{y \in X; x\mathscr{R}y\}$ is nonempty, for each $x \in X$. Given the reflexive relations $\mathscr{A}$ and $\mathscr{B}$, denote their *product* as

$$\mathscr{A} \circ \mathscr{B} = \{(x,y) \in X \times X; \exists z \in X : (x,z) \in \mathscr{A}, (z,y) \in \mathscr{B}\};$$

clearly, $\mathscr{A} \circ \mathscr{B}$ is reflexive too. Finally, for each reflexive relation $\mathscr{R}$, define

$$\mathscr{R}^1 = \mathscr{R}; R^{n+1} = \mathscr{R}^n \circ \mathscr{R}, n \geq 1.$$

By the reflexive property of $\mathscr{R}$, we have

$$\mathscr{R}^n\text{=reflexive}, \forall n \geq 1; \text{ hence, } \mathscr{R}^n \subseteq R^m \text{ if } n \leq m. \tag{36}$$

In this case, its associated relation $\mathscr{S} = \cup\{\mathscr{R}^n; n \geq 1\}$ is reflexive too and transitive (i.e.: $\mathscr{S}^2 \subseteq \mathscr{S}$); hence, it is a *quasi-order* on X.

(B) Let $(\perp)$ be a reflexive relation over X. Denote $(\leq) = \cup\{\perp^n; n \geq 1\}$; as noticed, this is a *quasi-order* (reflexive and transitive relation) over X; hence, the notation is meaningful. A characterization of this object may be given as follows. Given $x, y \in X$, any k-tuple $(z_1, \ldots, z_k)$ (for $k \geq 2$) in X^k with $z_1 = x, z_k = y$, and $[z_i \perp z_{i+1}, i \in \{1, \ldots, k-1\}]$ will be referred to as a $(\perp)$-*chain* between x and y; the class of all these will be denoted as $C(x, y; \perp)$. In this case,

$$(x, y \in X)\colon\ x \leq y \text{ iff } C(x, y; \perp) \text{ is nonempty.} \tag{37}$$

A basic construction involving this relation is to be made under the consideration of an a-metric structure on X. Let $d : X \times X \to R_+$ be a mapping, endowed with the properties: reflexive, triangular, and sufficient. Note that $d(.,.)$ has all properties of a metric, excepting symmetry; it will be referred to as an *almost metric* (in short: *a-metric*) on X; and the couple (X, d) will be termed an *a-metric space*. Having this precise, we may introduce a d-convergence and d-Cauchy structure on X, as follows. For each sequence $(x_n; n \geq 0)$ in X and each $x \in X$, put $x_n \overset{d}{\longrightarrow} x$ if $d(x_n, x) \to 0$ as $n \to \infty$; i.e.:

$$\forall \varepsilon > 0, \exists i = i(\varepsilon)\colon i \leq n \Longrightarrow d(x_n, x) < \varepsilon.$$

The class of all such x will be denoted as $\lim_n(x_n)$; when it is nonempty, we say that (x_n) is *d-convergent*. Note that $\lim_n(x_n)$ is not in general asingleton; but, when d is *separated*, in the sense

(b02) $\quad (\forall \text{ sequence } (x_n) \text{ in } X)\colon x_n \overset{d}{\longrightarrow} u \text{ and } x_n \overset{d}{\longrightarrow} v \text{ imply } u = v,$

this property is available. In this case, given a d-convergent sequence (x_n) of X, we must have $\{z\} = \lim_n(x_n)$ (for $z \in X$); as usual, we write this last relation as $z = \lim_n(x_n)$. Note that sufficiency of $d(.,.)$ follows from the separated condition. In fact, let $u, v \in X$ be such that $d(u, v) = 0$. Then, the constant sequence $(x_n = u; n \geq 0)$ fulfills $x_n \overset{d}{\longrightarrow} u, x_n \overset{d}{\longrightarrow} v$; whence (by the separated property) $u = v$. Likewise, the *d-Cauchy* property of a sequence (x_n) in X is depicted as: $d(x_m, x_n) \to 0$ as $m, n \to \infty, m < n$; or, equivalently:

$$\forall \varepsilon > 0, \exists j = j(\varepsilon)\colon j \leq m < n \Longrightarrow d(x_m, x_n) < \varepsilon.$$

It is to be stressed that, by the non-symmetry of $d(.,.)$, a d-convergent sequence need not be d-Cauchy; however, when the reciprocal inclusion holds (each d-Cauchy is d-convergent), we say that $d(.,.)$ is *complete* on X.

Now, starting from the couple $(\leq, d)$ introduced as before, the following a-metrical type construction is proposed. Define a mapping $\Delta : (\leq) \to R_+$ as: for each $x, y \in X$ with $x \leq y$,

(b03) $\Delta(x, y) = \inf[d(z_1, z_2) + \ldots + d(z_{k-1}, z_k)]$,
where $(z_1, \ldots, z_k)$ (for $k \geq 2$) is a $(\perp)$-chain between x and y.

The following properties are valid:

(p-1) Δ is $(\leq)$-reflexive $[\Delta(x, x) = 0, \forall x \in X]$,
(p-2) Δ is $(\leq)$-triangular $[\Delta(x, z) \leq \Delta(x, y) + \Delta(y, z)$, if $x \leq y, y \leq z]$.

In addition, the triangular property of d gives

$(\forall x, y \in X, x \leq y)$: $d(x, y) \leq d(z_1, z_2) + \ldots + d(z_{k-1}, z_k)$,
for any $(\perp)$-chain $(z_1, \ldots, z_k)$ (where $k \geq 2$) between x and y.

So, passing to infimum,

$$d(x, y) \leq \Delta(x, y), \text{ for all } x, y \in X \text{ with } x \leq y; \tag{38}$$

referred to as: d is $(\leq)$-*subordinated* to Δ. Note that, in such a case

(p-3) Δ is $(\leq)$-sufficient $[x \leq y, \Delta(x, y) = 0 \Longrightarrow x = y]$.

Finally, by the very definition of Δ, we get the *reduction formula*

$$d(x, y) \geq \Delta(x, y) \text{ (hence } d(x, y) = \Delta(x, y)), \text{ whenever } x \perp y. \tag{39}$$

Having these precise, define the mapping $e : X \times X \to R_+ \cup \{\infty\}$, as

(b04) $e(x, y) = \Delta(x, y$, if $x \leq y$; $e(x, y) = \infty$, otherwise,

Clearly, $e(.,.)$ is endowed with the properties: reflexive, triangular, and sufficient. We then say that $e(.,.)$ is a *generalized almost metric* [in short: *generalized a-metric*] on X; and the couple (X, e) will be referred to as a *generalized a-metric space*. This convention comes from the fact that the mapping $e(.,.)$ has all properties of a generalized metric (over X), excepting symmetry.

Now, as in the generalized metric case, we may introduce an e-convergence and e-Cauchy structure on X. Namely, for each sequence $(x_n; n \geq 0)$ in X and each $x \in X$, put $x_n \overset{e}{\longrightarrow} x$ if $e(x_n, x) \to 0$ as $n \to \infty$; i.e.:

$\forall \delta > 0, \exists i = i(\delta)$: $i \leq n \Longrightarrow e(x_n, x) < \delta$.

The class of all such x will be denoted as $\lim_n(x_n)$; when it is nonempty, we say that (x_n) is *e-convergent*. Note that $\lim_n(x_n)$ is not in general asingleton; but, when $e(.,.)$ is *separated*, in the sense

(b05) ($\forall$ sequence (x_n) in X): $x_n \overset{e}{\longrightarrow} u$ and $x_n \overset{e}{\longrightarrow} v$ imply $u = v$,

this property is available. Then, for an e-convergent sequence (x_n) of X, we have $\{z\} = \lim_n(x_n)$ (where $z \in X$); for simplicity, we write this last relation as $z = \lim_n(x_n)$. As before, the sufficiency of $e(.,.)$ follows from this. In fact, let $u, v \in X$ be such that $e(u, v) = 0$. Then, the constant sequence $(x_n = u; n \geq 0)$ fulfills $x_n \overset{e}{\longrightarrow} u, x_n \overset{e}{\longrightarrow} v$; whence (by the separated property) $u = v$. Likewise, given the sequence (x_n) of X, call it *e-Cauchy*, provided: $e(x_m, x_n) \to 0$ as $m, n \to \infty, m < n$; or, equivalently:

$\forall \delta > 0, \exists j = j(\delta)$: $j \leq m < n \Longrightarrow e(x_m, x_n) < \delta$.

Note that, by the non-symmetry of $e(.,.)$, an e-convergent sequence need not be e-Cauchy; however, when the reciprocal inclusion holds, (each e-Cauchy is e-convergent) we say that $e(.,.)$ is *complete* on X.

Having these precise, it would be useful for us to establish some basic properties of $e(.,.)$, in terms of the couple $(\perp, d)$ (or, equivalently: $(\leq, d)$).

(A) Concerning the separated property, a simple answer to it is contained in

Proposition 10. *Suppose that the a-metric $d(.,.)$ is separated. Then, the generalized a-metric $e(.,.)$ is separated too.*

Proof. Let the sequence $(x_n; n \geq 0)$ in X and the points $u, v \in X$ be such that $x_n \xrightarrow{e} u, x_n \xrightarrow{e} v$; i.e.: $e(x_n, u) \to 0, e(x_n, v) \to 0$. Without loss, one may assume that $x_n \leq u, x_n \leq v$, for all n; and then (by definition) we must have $\Delta(x_n, u) \to 0, \Delta(x_n, v) \to 0$. This, according to the subordination property, gives $d(x_n, u) \to 0, d(x_n, v) \to 0$; so that (as $d(.,.)$ is separated), $u = v$.

(B) Concerning the completeness question, an appropriate answer to it needs certain conventions. Call the sequence $(x_n; n \geq 0)$ in X, $(\perp)$-*ascending* provided $x_n \perp x_{n+1}, \forall n$. Then, let us say that $(\perp)$ is *almost-self-closed* (modulo d), when

(b06) if $(x_n; n \geq 0)$ is $(\perp)$-ascending and $x_n \xrightarrow{d} x$,
there exists a subsequence $(y_n; n \geq 0)$ of it with $(y_n \perp x, \forall n)$.

The following auxiliary fact is to be noted.

Proposition 11. *Let the e-Cauchy sequence $(x_n; n \geq 0)$ in X and the point $v \in X$ be such that*

$y_n \xrightarrow{e} v$, for some subsequence $(y_n = x_{k(n)}; n \geq 0)$ of $(x_n; n \geq 0)$.

Then, necessarily $x_n \xrightarrow{e} v$ as $n \to \infty$.

Proof. Let $\delta > 0$ be arbitrary fixed. From the e-Cauchy property, there exists $m(\delta) \geq 0$, such that

$$m(\delta) \leq i \leq j \Longrightarrow e(x_i, x_j) < \delta/2.$$

On the other hand, by the sub-sequential convergence property, we have that, for the same $\delta > 0$, there exists some $n(\delta) \geq m(\delta)$, such that

$$n(\delta) \leq n \Longrightarrow e(y_n, v) = e(x_{k(n)}, v) < \delta/2.$$

Finally, as $(k(n); n \geq 0)$ is strictly ascending, we must have

$$k(n) \geq n, \text{ for all } n \geq 0; \text{ hence, } k(n) \geq n(\delta), \text{ for all } n \geq n(\delta).$$

Combining these we have, for each $n \geq n(\delta)(\geq m(\delta))$

$$d(x_n, v) \leq d(x_n, x_{k(n)}) + d(x_{k(n)}, v) < \delta/2 + \delta/2 = \delta;$$

and the conclusion follows.

We are now in a position to state the desired answer.

Proposition 12. *Suppose that $d(.,.)$ is (separated) complete and $(\perp)$ is almost-self-closed (modulo d). Then, $e(.,.)$ is a (separated and) complete generalized a-metric.*

Proof. Let $(x_n; n \geq 0)$ be an e-Cauchy sequence in X; i.e.:

$\forall \delta > 0, \exists k(\delta)$: $k(\delta) \leq n \leq m \Longrightarrow e(x_n, x_m) < \delta$.

Without loss, one may assume that

$x_n \leq x_m$ (i.e.: $e(x_n, x_m) = \Delta(x_n, x_m)$), whenever $n \leq m$;

so that the working hypothesis writes:

$\forall \delta > 0, \exists k(\delta)$: $k(\delta) \leq n \leq m \Longrightarrow \Delta(x_n, x_m) < \delta$.

As a direct consequence of this, there exists a strictly ascending sequence of ranks $(k(n); n \geq 0)$, in such a way that

$$(\forall n) : k(n) < m \Longrightarrow \Delta(x_{k(n)}, x_m) < 2^{-n}.$$

Denoting $(y_n := x_{k(n)}, n \geq 0)$, we therefore have $\Delta(y_n, y_{n+1}) < 2^{-n}$, $\forall n$. Moreover, by the imposed e-Cauchy property (and a previous auxiliary fact)

$(x_n; n \geq 0)$ is e-convergent iff $(y_n; n \geq 0)$ is e-convergent.

To establish this last property, one may proceed as follows. As $\Delta(y_0, y_1) < 2^{-0}$, there exists (for the starting $p(0) = 0$), a $(\perp)$-chain $(z_{p(0)}, \ldots, z_{p(1)})$ between y_0 and y_1 (hence $p(1) - p(0) \geq 1, z_{p(0)} = y_0, z_{p(1)} = y_1$), such that

$$d(z_{p(0)}, z_{p(0)+1}) + \ldots + d(z_{p(1)-1}, z_{p(1)}) < 2^{-0}.$$

Further, as $\Delta(y_1, y_2) < 2^{-1}$, there exists a $(\perp)$-chain $(z_{p(1)}, \ldots, z_{p(2)})$ between y_1 and y_2 (hence $p(2) - p(1) \geq 1, z_{p(1)} = y_1, z_{p(2)} = y_2$), such that

$$d(z_{p(1)}, z_{p(1)+1}) + \ldots + d(z_{p(2)-1}, z_{p(2)}) < 2^{-1};$$

and so on. The procedure may continue indefinitely; it gives us a $(\perp)$-ascending (hence, $(\leq)$-ascending) sequence $(z_n; n \geq 0)$ in X with (cf. the reduction formula)

$$\sum_n \Delta(z_n, z_{n+1}) = \sum_n d(z_n, z_{n+1}) < \sum_n 2^{-n} < \infty. \tag{40}$$

In particular, $(z_n; n \geq 0)$ is d-Cauchy; wherefrom (as d is separated complete), $z_n \xrightarrow{d} z$ as $n \to \infty$, for some uniquely determined $z \in X$. Further, by the choice of $(\perp)$, there must be a subsequence $(t_n := z_{q(n)}; n \geq 0)$ of $(z_n; n \geq 0)$, with

$$t_n \perp z \text{ (hence, } t_n \leq z), \forall n.$$

This firstly gives (by the above convergence property), $t_n \xrightarrow{d} z$ as $n \to \infty$. Secondly (again combining with the reduction formula),

$$\Delta(t_n, z) = d(t_n, z), \ \forall n;$$

so that (by the above relation),

$$\Delta(t_n, z) \to 0 \text{ (i.e.: } t_n \xrightarrow{e} z), \text{ as } n \to \infty.$$

On the other hand, (40) tells us that $(z_n; n \geq 0)$ is Δ-Cauchy (i.e.: e-Cauchy). Adding the e-convergence property of $(t_n; n \geq 0)$, gives $z_n \xrightarrow{e} z$ as $n \to \infty$ (see above); wherefrom [as $(y_n; n \geq 0)$ is a subsequence of $(z_n; n \geq 0)$], $y_n \xrightarrow{e} z$ as $n \to \infty$. This finally gives $x_n \xrightarrow{e} z$; and concludes the argument.

Main Result

Let X be a nonempty set. By a *pseudometric* (resp., *generalized pseudometric*) over X we shall mean any map $(x, y) \mapsto e(x, y)$ from $X \times X$ to R_+ (resp., $R_+ \cup \{\infty\}$). Any reflexive, triangular, sufficient (generalized) pseudometric $e(.,.)$ will be referred to as a *(generalized) a-metric*; and the couple (X, e) will be referred to as a *(generalized) a-metric space*. This convention comes from the fact that the mapping $e(.,.)$ has all properties of a (generalized) metric (over X), excepting symmetry.

Let (X, e) be a generalized a-metric space. Denote for simplicity

(c01) $(x, y \in X)$: $x \leq y$ iff $e(x, y) < \infty$.

This relation is reflexive and transitive; hence, a quasi-order on X. Conversely, let $(\leq)$ be a quasi-order on X; and $\Delta : (\leq) \to R_+$ be a mapping endowed with the properties: $(\leq)$-reflexive, $(\leq)$-triangular, and $(\leq)$-sufficient. In this case, the generalized pseudometric (on X)

(c02) $e(x, y) = \Delta(x, y)$, if $x \leq y$; $e(x, y) = \infty$, otherwise,

is a generalized a-metric on X, as it can be directly seen.

Now, according to our previous conventions, we may introduce an e-convergence and e-Cauchy structure on X. This allows us defining an e-closure operator on X as follows. Let Y be a subset of X. We say that $u \in X$ is *e-adherent* to Y, when

(c03) $\forall \delta > 0, \exists x_\delta \in Y$: $e(x_\delta, u) < \delta$; or, equivalently:
$u \in \lim_n(x_n)$, for some sequence $(x_n; n \geq 0)$ in Y.

The set of all such points will be called the *e-closure* of Y; and denoted as: $\text{cl}_e(Y)$.

The basic properties of the mapping $Y \mapsto \text{cl}_e(Y)$ are contained in

Lemma 6. *Under the above conventions, we have*

(clo-1) *(progressiveness)* $Y \subseteq \text{cl}_e(Y)$, $\forall Y \in 2^X$,
(clo-2) *(identity)* $\emptyset = \text{cl}_e(\emptyset), X = \text{cl}_e(X)$,
(clo-3) *(monotonicity)* $\text{cl}_e(Y_1) \subseteq \text{cl}_e(Y_2)$, if $Y_1 \subseteq Y_2$,
(clo-4) *(additivity)* $\text{cl}_e(U \cup V) = \text{cl}_e(U) \cup \text{cl}_e(V)$, $\forall U, V \in 2^X$,
(clo-5) *(idempotence)* $\text{cl}_e(\text{cl}_e(Y)) = \text{cl}_e(Y)$, $\forall Y \in 2^X$.

Proof. (clo-1), (clo-2), (clo-3): Evident.

(clo-4) The right to left inclusion is clear, by the monotone property; so, it remains to verify the left to right inclusion. Let $w \in \text{cl}_e(U \cup V)$ be arbitrary fixed:

$$w \in \lim_n(z_n), \text{ for some sequence } (z_n; n \geq 0) \text{ in } U \cup V.$$

If the alternative below holds

for each (index) h, there exists (another index) $k > h$, with $x_k \in U$

then, a strictly ascending rank-sequence $(i(n); n \geq 0)$ may be found, such that

$$w \in \lim_n(z_{i(n)}), \ (z_{i(n)}; n \geq 0) = \text{sequence in } U;$$

wherefrom, $w \in \text{cl}_e(U)$. Otherwise (if the opposite alternative holds), we must have (by the choice of our sequence)

there exists an index h, fulfilling: $z_k \in V$, for all $k > h$;

and this tells us that

$$w \in \lim_n(z_{h+1+n}), \ (z_{h+1+n}; n \geq 0) = \text{sequence in } V;$$

whence, $w \in \text{cl}_e(V)$.

(clo-5) As before, the right to left inclusion is clear (by progressiveness and monotone properties); so, it remains to establish the left to right inclusion. Let $w \in \text{cl}_e(\text{cl}_e(Y))$ be arbitrary fixed. Given $\delta > 0$, there exists (by definition) some $y \in \text{cl}_e(Y)$ with $e(y, w) < \delta$. On the other hand, for this $y \in \text{cl}_e(Y)$ and the same $\delta > 0$, there exists $x \in Y$ with $e(x, y) < \delta$. Combining these yields (by the triangular inequality)

$$e(x, w) \leq e(x, y) + e(y, w) < 2\delta;$$

and since $\delta > 0$ was arbitrarily chosen, $w \in \text{cl}_e(Y)$. The proof is complete.

Note that, as a consequence of this, the mapping $Y \mapsto \mathrm{cl}_e(Y)$ is a Kuratowski closure operator [20, Chap. I, Sect. 4]; so, it induces a topology $\mathscr{T}(e)$ on X, according to the lines in Bourbaki [5, Chap. I, Sect. 1]. In particular, the $\mathscr{T}(e)$-closed subsets Y of X are characterized as $Y = \mathrm{cl}_e(Y)$; then Y is referred to as *e-closed*. For example, $Y = \emptyset$ and $Y = X$ have this property; further examples are to be determined via

Proposition 13. *Let (X, e) be a generalized a-metric space; and $(\leq)$ stand for the associated to $e(.,.)$ quasi-order (see above). Further, let $a \in X$ be arbitrary fixed. Then,*

(i) *the section $X(a, \leq) := \{x \in X; a \leq x\}$ is e-closed*
(ii) *the restriction of $e(.,.)$ to $X(a, \leq)$ is complete, whenever $e(.,.)$ is complete.*

Proof. (i) Let $(x_n; n \geq 0)$ be a sequence in $X(a, \leq)$; hence (by hypothesis)

$e(a, x_n) < \infty$, for all n.

In addition, suppose that $x \in X$ is such that $x_n \xrightarrow{e} x$. Letting $\delta > 0$ be arbitrary fixed, there must be some rank $n(\delta)$, with

$$e(x_n, x) < \delta < \infty, \text{ for all } n \geq n(\delta).$$

Combining with the above relation yields (by the triangular property)

$$(\forall n \geq n(\delta)) : \; e(a, x) \leq e(a, x_n) + e(x_n, x) < \infty;$$

wherefrom $x \in X(a, \leq)$.

(ii) Let $(x_n; n \geq 0)$ be an e-Cauchy sequence in $X(a, \leq)$. By the completeness hypothesis, $x_n \xrightarrow{e} z$ as $n \to \infty$, for some uniquely determined $z \in X$. On the other hand, from the previous part, we must have $z \in X(a, \leq)$; and this ends the argument.

Remark 1. The property discussed here is referred to as: the associated to $e(.,.)$ quasi-order $(\leq)$ is *semi-closed*; cf. Nachbin [26, Appendix]. Note that, unlike the generalized metric setting, the restriction of $e(.,.)$ to such sections $X(a, \leq)$ (with $a \in X$) is no longer a standard a-metric.

As a consequence of this statement, it results—like in the generalized metric setting—that Picard conventions to be introduced must have a local character; their precise content is described below.

Let (X, e) be a generalized a-metric space. Call the subset U of X, *e-asingleton* provided $[y_1, y_2 \in U, e(y_1, y_2) < \infty$ imply $y_1 = y_2]$; and *e-singleton* if, in addition, U is nonempty. (Remember that, in this last case, the alternative $\mathrm{card}(U) \geq 2$ is not avoidable). Further, let $T \in \mathscr{F}(X)$ be a self-map of X. The following *initial* regularity conditions about these data are accepted:

(**ureg-1**) The generalized a-metric $e(.,.)$ is *separated*; i.e.:

(c04) ($\forall$ sequence (x_n) in X): $x_n \stackrel{e}{\longrightarrow} u$ and $x_n \stackrel{e}{\longrightarrow} v$ imply $u = v$.

[Remember that the sufficiency of $e(.,.)$ follows from this].

(**ureg-2**) Let us say that $e(.,.)$ is *complete*, when

(c05) (for each sequence in X): e-Cauchy implies e-convergent.

Note that—like in the generalized metric setting—this concept is equivalent with its local version for generalized a-metrics. Precisely, call $e(.,.)$, *locally complete* when, for each $a \in X$, its restriction to the section $X(a, \le)$ is complete. By the preceding statement, we have (for each generalized a-metric): complete $\Longrightarrow$ locally complete. The converse inclusion is also true; hence

$$[\forall \text{ generalized a-metric}]: \text{complete} \Longleftrightarrow \text{locally complete}.$$

(**ureg-3**) Call T, *e-semi-progressive*, provided

(c06) $X(T; e) := \{x \in X; e(x, Tx) < \infty\}$ is nonempty.

Clearly, according to

$$e(x, Tx) < \infty \text{ if and only if } x \le Tx,$$

this e-semi-progressive condition may be written as:

T is $(\le)$-semi-progressive: $X(T, \le) := \{x \in X; x \le Tx\} \ne \emptyset$.

(**ureg-4**) Call T, *e-increasing* provided

(c07) $e(x, y) < \infty$ implies $e(Tx, Ty) < \infty$.

This, according to our conventions, may be also referred to as

T is $(\le)$-increasing: $x \le y$ implies $Tx \le Ty$.

Fix in the following some $Y \in 2^X$; and denote $\text{Fix}(T; Y) = \text{Fix}(T) \cap Y$ (the trace over Y of $\text{Fix}(T)$). Also, take some $W \in (2)^X$. Sufficient conditions for existence of such points are to be discussed along the "unilateral" versions of our previous local concepts.

(**uloc-pic-1**) Call $x \in X$, a *Picard point* (modulo (e, T)) relative to Y, provided $(T^n x; n \ge 0)$ is e-convergent, with $\lim_n(T^n x) \in Y$; if such a property holds for all $x \in W$, we say that T is a *Picard operator* (modulo e) relative to (W, Y); and, if (in addition) $\text{Fix}(T; Y)$ is an e-asingleton, then T is referred to as a *globally Picard operator* (modulo e) relative to (W, Y)

(**uloc-pic-2**) Call $x \in X$, a *strong Picard point* (modulo (e, T)) relative to Y, provided $(T^n x; n \ge 0)$ is e-convergent with $\lim_n(T^n x) \in \text{Fix}(T; Y)$; if this holds for all $x \in W$, we say that T is a *strong Picard operator* (modulo e) relative to

(W, Y); and, if (in addition) $\text{Fix}(T; Y)$ is an e-asingleton (hence, an e-singleton), then T is referred to as a *globally strong Picard operator* (modulo e) relative to (W, Y).

Finally, the contractive condition upon our data may be written as:

(**con-gam**) Given $\alpha > 0$, let us say that T is $(e; \alpha)$*-contractive*, provided

(c08) $e(Tx, Ty) \leq \alpha e(x, y)$, for each $x, y \in X$ with $e(x, y) < \infty$.

Note that, as a direct consequence of this, we have

T is $(\leq)$-increasing: $x \leq y$ implies $Tx \leq Ty$.

For other consequences of the same, we need a lot of preliminary concepts and auxiliary facts. Define some other relation $(\preceq)$ over X, as

(c09) $(u, v \in X)$: $u \preceq v$ iff $e(T^n u, T^n v) \to 0$, as $n \to \infty$.

Clearly, $(\preceq)$ is reflexive and transitive; hence, it is a quasi-order on X. By definition, $u \preceq v$ will be read as: the couple (u, v) is (e, T)*-asymptotic*.

The following simple property will be useful in the sequel

Lemma 7. *Suppose that* $e(.,.)$ *is separated; and let the couple* (u, v) *in* X *be* (e, T)*-asymptotic (i.e.:* $u \preceq v$*). Then,*

$$\lim_n(T^n u) = \lim_n(T^n v), \text{ provided } \lim_n(T^n v) \text{ exists.} \tag{41}$$

Proof. Denote $b = \lim_n(T^n v)$; hence, $e(T^n v, b) \to 0$ as $n \to \infty$. By the triangular inequality, we derive

$$e(T^n u, b) \leq e(T^n u, T^n v) + e(T^n v, b), \ \forall n.$$

Passing to limit as $n \to \infty$ gives $T^n u \xrightarrow{e} b$; and this, by the separated property, yields $b = \lim_n(T^n u)$; hence the conclusion.

We may now pass to the posed problem. Concerning the relationship between $(\preceq)$ and the initial quasi-order $(\leq)$, the following assertion is available.

Lemma 8. *Suppose that* T *is* $(e; \alpha)$*-contractive, for some* $\alpha \in]0, 1[$*. Then,*

$$u, v \in X, \ u \leq v \Longrightarrow u \preceq v; \tag{42}$$

i.e.: $(\preceq)$ *is finer than* $(\leq)$*.*

Proof. Let $u, v \in X$ be such that $u \leq v$; i.e.: $e(u, v) < \infty$. By the contractive condition, we have

$$e(T^n u, T^n v) \leq \alpha^n e(u, v), \ \forall n \geq 0.$$

And this, along with the choice of α, gives the desired fact.

Remark 2. The reciprocal inclusion is not in general true. In fact, let the points $u, v \in X$ be such that: $u \leq v$ is false, but $Tu \leq Tv$. Then, by the result above $\lim_n e(T^n(Tu), T^n(Tv)) = \lim_n e(T^{n+1}u, T^{n+1}v) = 0$; so that, $u \preceq v$; but, from this very choice, $u \leq v$ is not true.

Finally, as a last auxiliary fact about our data, we get

Lemma 9. *Suppose that T is $(e;\alpha)$-contractive, for some $\alpha \in]0,1[$. Then,* Fix(T) *is e-asingleton; hence,* Fix(T, Y) *is e-asingleton, for each (nonempty) part Y of X.*

Proof. Let $z_1, z_2 \in$ Fix(T) be such that $e(z_1, z_2) < \infty$. By the contractive condition,

$$e(z_1, z_2) = e(Tz_1, Tz_2) \leq \alpha e(z_1, z_2);$$

wherefrom, $e(z_1, z_2) = 0$. This, along with the sufficiency property of $e(.,.)$, gives $z_1 = z_2$; and concludes the argument.

We are now in position to state the main result of this exposition.

Theorem 11. *Assume that T is $(e;\alpha)$-contractive, for some $\alpha \in]0,1[$ (hence (see above): e-increasing). In addition, let $e(.,.)$ be separated, complete, and T be e-semi-progressive. Finally, take some point $x_0 \in X(T,e) = X(T,\leq)$. Then*

(Ko-1) $X(x_0, \leq)$ *is e-closed, T-invariant, and* $e(.,.)$ *is complete over* $X(x_0, \leq)$
(Ko-2) x_0 *is a strong Picard point (modulo (e,T)) relative to* $X(x_0, \leq)$
(Ko-3) *T is globally strong Picard (modulo e), relative to* $(X(x_0, \geq), X(x_0, \leq))$
(Ko-4) *T is globally strong Picard (modulo e), relative to* $(X(x_0, \leq) \cap X(T,e), X(x_0, \leq))$.

Proof. (I) Evident, by our previous developments.

(II) Denote for simplicity $(x_n = T^n x_0; n \geq 0)$. By definition, we have $x_0 \leq x_1$; i.e., $e(x_0, x_1) < \infty$. This, via T=increasing, tells us that $(x_n; n \geq 0)$ is an ascending sequence in $X(x_0, \leq)$. Moreover, by a repeated application of contractive condition,

$$e(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1),\ \forall n;$$

wherefrom, $(x_n; n \geq 0)$ is e-Cauchy. By completeness, there must be some $z \in X$ such that $x_n \overset{e}{\longrightarrow} z$ as $n \to \infty$. This firstly gives $z \in X(x_0, \leq)$ [as $X(x_0, \leq)$ is e-closed]. Secondly, by the very definition of convergence, there must be some rank $n(z) \geq 0$, such that

$$e(x_n, z) < 1 \text{ (hence, } x_n \leq z\text{), for all } n \geq n(z).$$

Combining with the contractive condition yields

$$e(x_{n+1}, Tz) \leq \alpha e(x_n, z),\ \forall n \geq n(z);$$

whence (passing to limit as $n \to \infty$), $x_n \overset{e}{\longrightarrow} Tz$. This, along with the separated property, gives $z = Tz$; hence, $z \in$ Fix$(T, X(x_0, \leq))$.

(III) Let $u_0 \in X(x_0, \geq)$ be arbitrary fixed; hence, $u_0 \leq x_0$. By a previous auxiliary fact, we must have

$$u_0 \preceq x_0 \text{ [i.e.: } d(T^n u_0, x_n) \to 0, \text{ as } n \to \infty];$$

i.e.: the couple (u_0, x_0) is (e, T)-asymptotic. Then (see above)

$$\lim_n T^n u_0 = \lim_n (x_n) = z;$$

whence u_0 is strong Picard (modulo (e, T)) relative to $X(x_0, \leq)$. On the other hand, by another auxiliary fact, $\text{Fix}(T)$ is e-asingleton; hence, so is $\text{Fix}(T; X(x_0, \leq))$. Putting these together gives the desired conclusion.

(IV) Let $v_0 \in X(x_0, \leq) \cap X(T, e)$ be arbitrary fixed; hence, $x_0 \leq v_0 \in X(T, e)$. It will suffice to apply the conclusions **(Ko-1)**+ **(Ko-2)** to the starting point v_0 (with the associated section $X(v_0, \leq)$) and conclusion **(Ko-3)** to the pair (x_0, v_0) (in place of (u_0, x_0)) to get the desired assertion. The proof is thereby complete.

In particular, when $e(.,.)$ is a generalized metric on X, Theorem 11 is just LJCP. Further aspects may be found in Turinici [38].

Particular Aspects

Let X be a nonempty set; and $(\perp)$ be a reflexive relation over it. Further, let $(\leq) := \cup\{\perp^n; n \geq 1\}$ stand for the minimal quasi-order that includes $(\perp)$. Remember that a characterization of this object may be obtained as follows. Given $x, y \in X$, any k-tuple $(z_1, \ldots, z_k)$ (for $k \geq 2$) in X^k with $z_1 = x, z_k = y$, and $[z_i \perp z_{i+1}, i \in \{1, \ldots, k-1\}]$ will be referred to as a $(\perp)$-*chain* between x and y; the class of all these will be denoted as $C(x, y; \perp)$. In this case, we have

$$(x, y \in X):\ x \leq y \text{ iff } C(x, y; \perp) \text{ is nonempty.} \tag{43}$$

(A) Let $T \in \mathscr{F}(X)$ be a self-map of X. We call it $(\perp)$-*increasing*, provided

(d01) (for all $x, y \in X$): $x \perp y \Longrightarrow Tx \perp Ty$.

Concerning the corresponding property for associated quasi-order $(\leq)$, the following simple answer is available:

Lemma 10. *Under these conventions, the generic relation holds*

$$(\forall T \in \mathscr{F}(X)):\ (\perp)\text{-increasing} \Longrightarrow (\leq)\text{-increasing.} \tag{44}$$

Proof. Let $x, y \in X$ be such that $x \leq y$. By a previous characterization, there must be a $(\perp)$-chain $(z_1, \ldots, z_k)$ (for $k \geq 2$) in X^k, joining x and y. Combining with the imposed hypothesis, $(Tz_1, \ldots, Tz_k)$ (for $k \geq 2$) is a $(\perp)$-chain in X^k, which connects Tx and Ty; wherefrom, $Tx \leq Ty$.

Remark 3. The converse inclusion is not in general true. In fact, put $X = \{1, 2, 3\}$ and take the reflexive relation $(\perp) = \{(1, 1), (2, 2), (3, 3), (1, 3), (3, 2)\}$. Its associated quasi-order is

$$(\leq) = \{(1, 1), (2, 2), (3, 3), (1, 2), (1, 3), (3, 2)\}.$$

Define $T \in \mathscr{F}(X)$ as: $[T1 = 1, T2 = 2, T3 = 2]$. Clearly, T is $(\leq)$-increasing, since $(T1, T2) = (1, 2), (T1, T3) = (1, 2), (T3, T2) = (2, 2)$. On the other hand, T is not $(\perp)$-increasing, because $(1, 3) \in (\perp)$, but $(T1, T3) = (1, 2) \notin (\perp)$.

(B) Further conventions are to be introduced under a (generalized) a-metric framework. Let $d : X \times X \to R_+$ be a reflexive, triangular, sufficient mapping; referred to as an *a-metric*) on X; then, the couple (X, d) will be termed an *a-metric space*. Define a mapping $\Delta : (\leq) \to R_+$ as: for each $x, y \in X$ with $x \leq y$,

$\Delta(x, y) = \inf[d(z_1, z_2) + \ldots + d(z_{k-1}, z_k)]$,
where $(z_1, \ldots, z_k)$ (for $k \geq 2$) is a $(\perp)$-chain between x and y.

As precise, this mapping is $(\leq)$-reflexive $(\leq)$-triangular and $(\leq)$-sufficient. Finally, define the mapping $e : X \times X \to R_+ \cup \{\infty\}$, as

$e(x, y) = \Delta(x, y$, if $x \leq y$; $e(x, y) = \infty$, otherwise,

Clearly, $e(.,.)$ is reflexive triangular and sufficient. We then say that $e(.,.)$ is the *generalized a-metric* on X induced by Δ; and the couple (X, e) will be referred to as a *generalized a-metric space*.

Remember that some basic properties of $e(.,.)$—such as separation and completeness—may be formulated in terms of corresponding properties for $(\perp, d)$. In the following, a corresponding formulation of e-contractive properties is to be given. Precisely, call $T \in \mathscr{F}(X)$, $(e; \alpha)$*-contractive* (for some $\alpha > 0$), provided

(d02) $e(Tx, Ty) \leq \alpha e(x, y)$, whenever $e(x, y) < \infty$.

The associated property in terms of $(\perp, d)$ may be introduced as: we say that $T \in \mathscr{F}(X)$ is $(d, \perp; \alpha)$*-contractive* (where $\alpha > 0$), provided

(d03) $x, y \in X, x \perp y$ implies $d(Tx, Ty) \leq \alpha d(x, y)$.

We may now give an appropriate answer to the posed question.

Lemma 11. *Let the* $(\perp)$*-increasing map* $T \in \mathscr{F}(X)$ *and the number* $\alpha > 0$ *be such that* T *is* $(d, \perp; \alpha)$*-contractive. Then,* T *is* $(e; \alpha)$*-contractive.*

Proof. Suppose that $x, y \in X$ are such that $e(x, y) < \infty$; i.e.: $x \leq y$. For each $(\perp)$-chain $(z_1, \ldots, z_k)$ (where $k \geq 2$) joining x and y, we have (as T is

$(\perp)$-increasing) that $(Tz_1, \ldots, Tz_k)$ is a $(\perp)$-chain between Tx and Ty. So, by the contractive condition, we get

$$\Delta(Tx, Ty) \leq \sum_{i=1}^{k-1} d(Tz_i, Tz_{i+1}) \leq \alpha \sum_{i=1}^{k-1} d(z_i, z_{i+1}),$$

for all such $(\perp)$-chains; wherefrom, passing to infimum (over $C(x, y; \perp)$),

$$\Delta(Tx, Ty) \leq \alpha\Delta(x, y); \text{ or, equivalently: } e(Tx, Ty) \leq \alpha e(x, y).$$

The proof is complete.

(C) Under these preliminaries, we may now state the announced result. Let X be a nonempty set; and $(\perp)$ be a reflexive relation over X. Further, let $(\leq)$ stand for the minimal quasi-order including $(\perp)$ (see above); and $d : X \times X \to R_+$ be an a-metric over X. Let also $\Delta : (\leq) \to R_+$ be the associated to $(\leq, d)$ mapping; and $e : X \times X \to R_+ \cup \{\infty\}$ stand for the attached generalized a-metric. Finally, let T be a self-map of X; supposed to satisfy

(d04) T is $(\leq)$-semi-progressive: $X(T; \leq) := \{x \in X; x \leq Tx\} \neq \emptyset$.

Our second main result in this exposition is

Theorem 12. *Suppose that T is $(d, \perp; \alpha)$-contractive, for some $\alpha \in]0, 1[$. In addition, let d be separated, complete, $(\perp)$ be almost-self-closed (modulo d), and suppose that T is $(\perp)$-increasing, $(\leq)$-semi-progressive. Finally, take some point $x_0 \in X(T, \leq)$. Then, conclusions of the first main result are being retainable.*

Proof. There are several steps to be passed.

Step 1. From a couple of auxiliary statements, e is separated and complete; because d is separated, complete and $(\perp)$ is almost-self-closed (modulo d).

Step 2. According to

$$x \leq Tx \text{ if and only if } e(x, Tx) < \infty,$$

the $(\leq)$-semi-progressive condition may be written as:

(d05) T is e-semi-progressive: $X(T, e) := \{x \in X; e(x, Tx) < \infty\} \neq \emptyset$.

In addition (cf. an auxiliary fact), the $(d, \perp; \alpha)$-contractive and $(\perp)$-increasing conditions upon T assure us that this self-map is $(e; \alpha)$-contractive.

Step 3. Summing up, the first main result applies to the generalized a-metric structure (X, e), the self-map T and the number $\alpha \in]0, 1[$. In this case, by the conclusions listed there, we get all desired facts. The proof is complete.

In particular, our progressive condition is fulfilled when

(d06) T is $(\perp)$-semi-progressive: $X(T; \perp) := \{x \in X; x \perp Tx\} \neq \emptyset$.

Then, our second main result above reduces to the statement in Jachymski [14]; but, the methods used there are quite different from the above ones. Further aspects may be found in Samet and Turinici [33].

References

1. Agarwal, R.P., El-Gebeily, M.A., O'Regan, D.: Generalized contractions in partially ordered metric spaces. Appl. Anal. **87**, 109–116 (2008)
2. Akkouchi, M.: On a fixed point theorem of D.W. Boyd and J.S. Wong. Acta Math. Vietnam. **27**, 231–237 (2002)
3. Banach, S.: Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundam. Math. **3**, 133–181 (1922)
4. Berzig, M., Karapinar, E., Roldan, A.: Discussion on generalized-($\alpha\psi, \beta\varphi$)-contractive mappings via generalized altering distance function and related fixed point theorems. Abstr. Appl. Anal. **2013**, Article ID 259768 (2013)
5. Bourbaki, N.: General Topology (Chapters 1–4). Springer, Berlin (1989)
6. Boyd, D.W., Wong, J.S.W.: On nonlinear contractions. Proc. Am. Math. Soc. **20**, 458–464 (1969)
7. Browder, F.E.: On the convergence of successive approximations for nonlinear functional equations. Indag. Math. **30**, 27–35 (1968)
8. Choudhury, B.S., Kundu, A.: A Kannan-like contraction in partially ordered spaces. Demonstr. Math. **46**, 327–334 (2013)
9. Cirić, L.B.: A new fixed-point theorem for contractive mappings. Publ. Inst. Math. **30**(44), 25–27 (1981)
10. Geraghty, M.A.: On contractive mappings. Proc. Am. Math. Soc. **40**, 604–608 (1973)
11. Hitzler, P.: Generalized metrics and topology in logic programming semantics. Ph.D. thesis, University College Cork (Ireland) (2001)
12. Hyers, D.H., Isac, G., Rassias, T.M.: Topics in Nonlinear Analysis and Applications. World Scientific Publishing, Singapore (1997)
13. Jachymski, J.: Common fixed point theorems for some families of mappings. Indian J. Pure Appl. Math. **25**, 925–937 (1994)
14. Jachymski, J.: The contraction principle for mappings on a metric space with a graph. Proc. Am. Math. Soc. **136**, 1359–1373 (2008)
15. Jachymski, J.: Equivalent conditions for generalized contractions on (ordered) metric spaces. Nonlinear Anal. **74**, 768–774 (2011)
16. Jung, C.F.K.: On generalized complete metric spaces. Bull. Am. Math. Soc. **75**, 113–116 (1969)
17. Kannan, R.: Some results on fixed points. Bull. Calcutta Math. Soc. **60**, 71–76 (1968)
18. Khan, M.S., Swaleh, M., Sessa, S.: Fixed point theorems by altering distances between the points. Bull. Aust. Math. Soc. **30**, 1–9 (1984)
19. Kirk, W.A.: Contraction mappings and extensions. In: Kirk, W.A., Sims, B. (eds.) Handbook of Metric Fixed Point Theory, pp. 1–34. Kluwer Academic Publishers, Dordrecht (2001)
20. Kuratowski, C.: Topologie (vol. I). Mathematics Monographs, vol. 20. Scientific Publishers (1958)
21. Leader, S.: Fixed points for general contractions in metric spaces. Math. Jpn. **24**, 17–24 (1979)
22. Luxemburg, W.A.J.: On the convergence of successive approximations in the theory of ordinary differential equations (II). Indag. Math. **20**, 540–546 (1958)
23. Matkowski, J.: Integrable solutions of functional equations. Diss. Math. **127**, 1–68 (1975)
24. Matkowski, J.: Fixed point theorems for contractive mappings in metric spaces. Časopis Pěst. Mat. **105**, 341–344 (1980)

25. Meir, A., Keeler, E.: A theorem on contraction mappings. J. Math. Anal. Appl. **28**, 326–329 (1969)
26. Nachbin, L.: Topology and Order. D. van Nostrand Comp., Princeton (1965)
27. Natanson, I.P.: Theory of Functions of a Real Variable, vol. I. Frederick Ungar Publishing Co., New York (1964)
28. Nieto, J.J., Rodriguez-Lopez, R.: Contractive mapping theorems in partially ordered sets and applications to ordinary differential equations. Order **22**, 223–239 (2005)
29. Nieto, J.J., Rodriguez-Lopez, R.: Existence and uniqueness of fixed point in partially ordered sets and applications to ordinary differential equations. Acta Math. Sinica (Engl. Ser.) **23**, 2205–2212 (2007)
30. Ran, A.C.M., Reurings, M.C.: A fixed point theorem in partially ordered sets and some applications to matrix equations. Proc. Am. Math. Soc. **132**, 1435–1443 (2004)
31. Rhoades, B.E.: A comparison of various definitions of contractive mappings. Trans. Am. Math. Soc. **226**, 257–290 (1977)
32. Rus, I.A.: Generalized Contractions and Applications. Cluj University Press, Cluj-Napoca (2001)
33. Samet, B., Turinici, M.: Fixed point theorems on a metric space endowed with an arbitrary binary relation and applications. Commun. Math. Anal. **13**, 82–97 (2012)
34. Turinici, M.: Fixed points of implicit contraction mappings. An. Şt. Univ. "A. I. Cuza" Iaşi (S. I-a: Mat.) **22**, 177–180 (1976)
35. Turinici, M.: Nonlinear contractions and applications to Volterra functional equations. An. Şt. Univ. "A. I. Cuza" Iaşi (S I-a: Mat) **23**, 43–50 (1977)
36. Turinici, M.: Abstract comparison principles and multivariable Gronwall-Bellman inequalities. J. Math. Anal. Appl. **117**, 100–127 (1986)
37. Turinici, M.: Ran-Reurings theorems in ordered metric spaces. J. Indian Math. Soc. **78**, 207–214 (2011)
38. Turinici, M.: Ran-Reurings fixed point results in ordered metric spaces. Lib. Math. **31**, 49–55 (2011)
39. Turinici, M.L Contractive maps in Mustafa-Sims metric spaces. Int. J. Nonlinear Anal. Appl. **5**, 26–53 (2014)

Models of Fuzzy Linear Regression: An Application in Engineering

Christos Tzimopoulos, Kyriakos Papadopoulos, and Basil K. Papadopoulos

Abstract The classical Linear Regression is an approximation for the creation of a model which connects a dependent variable y with one or more independent variables X, and it is subjected to some assumptions. The violation of these assumptions can influence negatively the power of the use of statistical regression and its quality. Nowadays, to exceed this problem, a new method has been introduced, and is in use, that is called *fuzzy regression*. Fuzzy regression is considered to be possibilistic, with the distribution function of the possibility to be connected with the membership function of fuzzy numbers. In this article, we examine three models of fuzzy regression, for confidence level h=0, for the case of crisp input values, fuzzy output values and fuzzy regression parameters, with an application to Hydrology, in two rainfall stations in Northern Greece.

Keywords Fuzzy linear regression models • Tanaka model • Savic and Pedrycz model • Least squares model application

Introduction

In Statistics, Linear Regression is an approximation for the creation of a model connecting a dependent variable y with one or more independent variables X. In the case of one independent variable we have simple Linear Regression. Linear Regression was the first type of analysis of Regression which was studied knowingly and was used in practical applications. Today, Linear Regression is used

C. Tzimopoulos (✉)
School of Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: ctzimop@gmail.com

K. Papadopoulos
College of Engineering and Technology, American University of the Middle East, Egaila, Kuwait
e-mail: Kyriakos.P@aum.edu.kw

B.K. Papadopoulos
School of Engineering, Democritus University of Thrace, Xanthi, Greece
e-mail: papadob@civil.duth.gr

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_29

in many problems of applied science (Geology, Meteorology, Physical Sciences, Engineering, Medicine, etc.) and most of the applications refer to the following two categories [24]:

- If the target is the prediction, Linear Regression is used for creating a prediction model which will simulate a set of observed times y and $X_i, i = 1, \ldots n$. After creating such a model, if it is given an addition random value X_i, without the existence of a corresponding value for y, this model is used for giving a prediction value for y.
- For a given set of values y and variables $X_1, \ldots X_n$ which are connected with y, the Linear Regression is applied to quantify the closeness of the relationship between y and X_i, to evaluate which X_i have no relationship with y, and to identify which subsets of X_i contain redundant information for y.

The classical model of Linear Regression has the form:

$$y_i = \beta_1 x_{i1} + \ldots + \beta_n x_{in} + \varepsilon_i, i = 1, \ldots, m$$

or, in vector form:

$$\overrightarrow{y} = X \cdot \overrightarrow{\beta} + \overrightarrow{\varepsilon}$$

where $\overrightarrow{y}$ is a vector that is called dependent variable, X is the matrix of independent variables, $\overrightarrow{\beta}$ is the vector of the parameters of the regression, and $\overrightarrow{\varepsilon}$ is the vector of error terms. The estimation of the parameters $\overrightarrow{\beta}$ is being done with the classic method of the least squares (OLS), which minimizes the sum of squares of the remainders and leads to the below estimated value for the parameter $\overrightarrow{\beta}$:

$$\hat{\beta} = (X^T X)^{-1} X^T \cdot \overrightarrow{y}$$

This estimator is unbiased and consistent, since the errors are not dependent from the independent variables. It should be noted that an application of Linear Regression requires the following:

1. Linear functional form of the relation, where the linearity refers to the parameters and not the variables.
2. We suppose that there is no precise linear relation between independent variables.
3. The errors are supposed to have expected value zero, for each observation, i.e. $E[\varepsilon_i] = 0$, and the violations of the independence $E[\varepsilon_i, \varepsilon_j] = 0$ are considered very serious.
4. The variance of errors is: $\text{Var}[\varepsilon_i] = \sigma^2$ for all $i = 1, \ldots, n$, and the covariance of the errors is: $\text{Cov}[\varepsilon_i, \varepsilon_j] = 0$ for all $i \neq j$. If the error has constant variance, it is called homoscedastic.
5. Nonstochastic independent variables

6. The remainders are distributed normally, with zero mean and constant variance. The ε_i are $N(0, \sigma^2)$.

The violation of the above basic assumptions can influence negatively the power of the use of the linear regression and also its quality. Nowadays, in order to handle this problem, the fuzzy regression has been introduced and is in use. It was developed by Tanaka et al. [22], Tanaka [18], Tanaka and Watada [21], Tanaka and Hayashi [19], Tanaka and Ishibuchi [20] and is considered more effective from the statistical regression, with respect to the above facts being violated or not being used properly. For example, when human judgment is involved, fuzzy procedures are introduced, which should be explained, or when we have a small number of crisp data or even fuzzy data, the use of statistical regression is excluded. In addition, the descriptive ability of fuzzy regression sets is better than the statistical regression when the size of the sample space is small and the ability of statistical regression is diminished. Thus, fuzzy regression can be a useful tool for the users and researchers in different areas, when they study the relationship between variables with fuzzy, deficient, and restricted information. So, while the classical regression is considered probabilistic, fuzzy regression is considered possibilistic and the membership function $\mu_{\widetilde{F}}$, of a fuzzy number $\widetilde{F}$, is linked with the possibility distribution function. In other words, for a given fuzzy number $\widetilde{F}$ and membership function $\mu_{\widetilde{F}}$, the possibility distribution function of a possibility is defined as:

$$\pi_x(x) \triangleq \mu_{\widetilde{F}}(x)$$

In addition, the fuzzy regression analysis uses fuzzy function for the regression coefficients, in contrast with the classical regression analysis. The above problem [16, 19] usually meets the three cases below:

1. Crisp input values x_{ij} and output y_j, fuzzy estimated values $\widetilde{Y}_i$.
2. Crisp input values x_{ij} and fuzzy output values $\tilde{y}_i$, fuzzy estimated values $\widetilde{Y}_i$.
3. Fuzzy input values $\tilde{x}_{ij}$ and output $\tilde{y}_i$, fuzzy estimated values $\widetilde{Y}_i$.

In all of the above cases, the parameters of the regression are considered fuzzy. The symbols $(x_{ij}, y_j, \tilde{x}_{ij}, \tilde{y}_i)$ refer to the data of the system while $\widetilde{Y}_i$ refers to the estimated value.
The adjustment of a model of fuzzy regression is done via two general methods [9]:

– The **possibilistic model** ([13, 14, 16, 17, 22], etc.). One minimizes the fuzziness of the model by taking into account the minimum of the spreads around the center of the fuzzy parameters, by considering as important to include the experimental values of every sample within a specific interval of possible data.
– The **model of the classic method of least squares** [4, 7, 8, 26]. The distance between the estimated output value of the model $\widetilde{Y}_i$ and of the observed output value $\tilde{y}_i$ is being minimized.

In the initial model of Tanaka et al. [22], the objective function used as a criterion is the minimilization of the total fuzziness *s*, which was determined as the sum of the individual spreads of the fuzzy variables:

$$\min J(c) = \sum_i c_i$$

Because of the criticism of the above model ($c_i = 0$, [11]), Tanaka [18], modified this model by considering the sum of the spreads of the estimated value of the model $\widetilde{Y}_i$:

$$\min J(c) = \sum_i c^T |x_i|$$

In addition, in all of his models, the membership functions used were triangular and in the constraints, the subset-hood of fuzzy numbers with a confidence level h was used. This subset-hood included three cases:

1. **Minimum Problem**: The measured values for h degree of confidence should be included within the estimated values:

$$[\tilde{y}_j^\ell, \tilde{y}_j^r]^h \subseteq [\widetilde{Y}_j^\ell, Y_j^r]^h$$

2. **Maximum Problem**: The estimated values for h degree of confidence will be contained within the measured values:

$$[\widetilde{Y}_j^\ell, Y_j^r]^h \subseteq [\tilde{y}_j^\ell, \tilde{y}_j^r]^h$$

3. **Dual Problem**: The intersection of measured values and estimated values is considered different from 0:

$$[\tilde{y}_j^\ell, \tilde{y}_j^r]^h \cap [\widetilde{Y}_j^\ell, Y_j^r]^h \neq 0$$

Tanaka [18] believes that there is always a solution to the minimum problem and the dual problem. Furthermore, Tanaka and Watada [21] showed that the maximum problem has a solution if and only if the objective function of the dual problem is equal to 0. Hereafter, in all problems of possibilistic regression, the first case was used, namely the case of the minimum problem.

Apart from the above named researchers, the problem of fuzzy linear regression was studied by Bardossy et al. [3] who applied the theory in hydrology problems using the minimum problem model, Savic and Pedrycz [17] who used the dual form of the minimum problem model, Moskowitz and Kim [12] who carried out a sensitivity analysis of the confidence level h in the minimum problem model, Redden and Woodall [16] who used fuzzy independent variables and introduced improvements to similar problems. Chen [6] modified the Tanaka method [18] to include measured points with large errors (outliers) and considered that the difference between the estimated spread and the measurement error should be less than a certain amount. Hung and Yang [10] improved the method of Chen, implementing a process for measuring the influence of an observation in the value of the objective function of Tanaka, when the observation is omitted. We should also mention Pasha et al. [15] who used the entropy in the objective function with the same constraints.

Regarding the method of least squares, we must mention Diamond [7, 8] who first introduced this method and for symmetric fuzzy numbers defined the distance between two fuzzy numbers ds^2 as equal to the sum of squared deviations of the supports, and proved that it is equivalent to the standard norm that is equal to the sum of squared deviations of both the supports and also the kernel.Chang and Lee [4] introduced a generalized weighted least-squares regression algorithm and used iterative methods to determine the supports and the kernel. Xu and Li [25] used a balanced norm as distance between two fuzzy numbers and Gauss functions as membership functions. Subsequently, a system of fuzzy equations was formed and solved by analytical method. Bargiela et al. [2] proposed a recursive algorithm for multiple regression by the method of least squares, and they solved the problem as an optimization problem with steepest descent. Arabpour and Tata [1] used a least squares model similar to the model of Diamond [7] and generalized the metric used by Diamond, using it as an optimization criterion for estimating the model parameters which are the coefficients of the supports and the kernel.

This article examines: (a) a possibilistic model by Tanaka, (b) a possibilistic model by Savic and Pedrycz, (c) a model of least squares. These models are tested for a confidence level h=0 and in the case of crisp input values, fuzzy output values and fuzzy parameters in Hydrology problems. Symmetric triangular fuzzy numbers are used for the parameters and the measured output values. The application involves the existing correlation of monthly rainfall average heights for two rainfall stations in Serres (Greece): Aggistro (X) and Upper Vrontous (Y) for 12 periods.

Mathematical Model

Possibilistic Model of Tanaka

Consider a fuzzy dependent variable $\widetilde{Y}_j$ and x_{ij} the independent variables influencing the variable $\widetilde{Y}_j$. The result of a fuzzy linear regression is an equation of the form ([18–21], etc.):

$$\widetilde{Y}_j = \widetilde{A}_0 + \widetilde{A}_1 x_{1j} + \widetilde{A}_2 x_{2j} + \ldots + \widetilde{A}_n x_{nj} = \sum_{i=0}^{n} \widetilde{A}_i x_{ij}, x_{0j} = 1 \qquad (1)$$

where the measured input values x_{ij} are classical crisp numbers and the measured output values, fuzzy numbers. The parameters $\widetilde{A}_i = (r_i, c_i)$ are considered symmetric triangular fuzzy numbers, with membership function given as follows:

$$\mu_{\widetilde{A}_i}(x) = \begin{cases} 0 & x < r_i - c_i \text{ or } x > r_i + c_i \\ 1 - \frac{|x-r_i|}{c_i} & r_i - c_i \leq x \leq r_i + c_i \end{cases} \tag{2}$$

The parameters r_i, c_i are, respectively, the means and spread of the parameter $\widetilde{A}_i$. The membership functions are normally written using the L, R types as follows:

1. $L(\frac{r_i-x}{c_i})$, for $r_i - c_i \leq x \leq r_i$ and 0 elsewhere
2. $R(\frac{x-r_i}{c_i})$, for $r_i \leq x \leq r_i + c_i$ and 0 elsewhere

Functions L, R have the following properties: $L(0) = R(0) = 1$, $L(1) = R(1) = 0$ and $L^{-1}(h) = R^{-1}(h) = 1 - h$.

The h-cuts confidence [4] of the parameter $\widetilde{A}_i$ are given as follows:

$$[\widetilde{A}_i]^h = [r_i - L^{-1}(h)c_i, r_i + L^{-1}(h)c_i] \tag{3}$$

while the h-cuts confidence of the fuzzy dependent variable $\widetilde{Y}_j$ are given as follows:

$$[\widetilde{Y}_j]^h = [R_j - C_j = \sum_{i=0}^{n} r_i x_{ij} - L^{-1}(h) \sum_{i=0}^{n} c_i x_{ij},\ R_j + C_j \tag{4}$$
$$= \sum_{i=0}^{n} r_i x_{ij} + L^{-1}(h) \sum_{i=0}^{n} c_i x_{ij}], x_{0j} = 1.$$

the mean m_i and the spread c_i of the fuzzy number $\widetilde{Y}_j$ are given as follows:

$$m_i = \sum_{i=0}^{n} r_i x_{ij}, c_i(h) = L^{-1}(h) \sum_{i=0}^{n} c_i x_{ij} \tag{5}$$

In the above equations the m_i is called the kernel of the fuzzy number $[\widetilde{Y}_j]$, while the cuts $[\widetilde{Y}_j^{\ell}, Y_j^r]^{h=0}$ are called its supports.

The observed output value $\tilde{y}_i$ is fuzzy, with mean value y_i and spread e_i (Fig. 1) and must be included in the estimated value $[\widetilde{Y}_j]^h$ [5, 12]. Therefore it will hold that $[\tilde{y}_j]^h \subseteq [Y_L^h, Y_U^h]$ and the above problem is formulated as a linear programming problem:

$$\min(c) = mc_0 + \sum_{j=1}^{m} \sum_{i=1}^{n} c_i x_{ij}, c_i \geq 0 \tag{6}$$

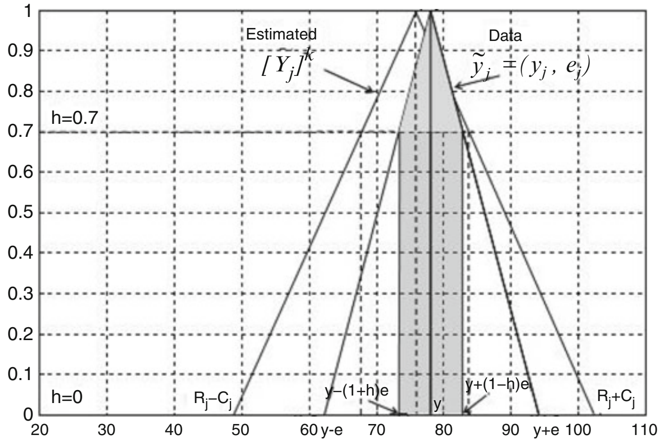


Fig. 1 Case for fuzzy measured output value

s.t.

$$\sum_{i=0}^{n} r_i x_{ij} + L^{-1}(h) \sum_{i=0}^{n} c_i x_{ij} \geq y_j + (1-h)e_j, \rightarrow \vec{r}^{\,T} + L^{-1}(h)\vec{c}^{\,T} \cdot X \geq \vec{y} + L^{-1}(h)\vec{e},$$

$$\sum_{i=0}^{n} r_i x_{ij} - L^{-1}(h) \sum_{i=0}^{n} c_i x_{ij} \leq y_j - (1-h)e_j, \rightarrow (\vec{r}^{\,T} - L^{-1}(h)\vec{c}^{\,T}) \cdot X \leq \vec{y} - L^{-1}(h)\vec{e},$$

$$\vec{r}^{\,T} = [r_0, r_1, \ldots, r_n], \vec{c}^{\,T} = [c_0, c_1, \ldots, c_n],$$
$$\vec{y}^{\,T} = [y_0, y_1, \ldots, y_n], \vec{e}^{\,T} = [e_0, e_1, \ldots, e_n], x_{0j} = 1. \tag{7}$$

The Possibilistic Model of Savic and Pedrycz

The model of Savic and Pedrycz [17] is known as FLSLR (Fuzzy Least-Squares Linear Regression). In this model, Savic and Pedrycz used the method of Tanaka in two phases:

First Phase In the first phase they consider the input values x_{ij} and output y_j as crisp values and also that their model is probabilistic. They apply the method of least squares to estimate the values of parameters $\vec{r}$. The normal equations are written as follows:

$$(X^T X) \cdot \vec{r} = X^T \cdot \vec{y} \tag{8}$$

and least squares estimators as he solution of (8) are:

$$\vec{r}^{*} = (X^{T} X)^{-1} X^{T} \cdot \vec{y} \tag{9}$$

Second Phase In this step it is again a linear programming problem, like that of Tanaka, as in Eqs. (6) and (7). The only difference lies in vector $\vec{r}$ that has already been calculated in the first phase and is now regarded as the known vector $\vec{r}^{*}$. The equations are presented with a little modification to the originals of Savic and Pedrycz. That is, the objective function in the original article of Savic and Pedrycz had the expression:

$$\min(c) = \sum_{i=0}^{n} C_i,$$

while here it is used with the subsequent expression of Tanaka:

$$\min(c) = mc_0 + \sum_{j=1}^{m} \sum_{i=1}^{n} c_i x_{ij}, c_i \geq 0.$$

The problem of Savic and Pedrycz is now defined as follows:

$$\min(c) = mc_0 + \sum_{j=1}^{m} \sum_{i=1}^{n} c_i x_{ij}, c_i \geq 0. \tag{10}$$

s.t.

$$\begin{aligned} &(\vec{r}^{T*} + L^{-1}(j)\vec{c}^{T}) \cdot X \geq \vec{y} + L^{-1}(h)\vec{e}, \\ &(\vec{r}^{T*} - L^{-1}(h)\vec{c}^{T}) \cdot X \leq \vec{y} - L^{-1}(h)\vec{e} \end{aligned} \tag{11}$$

Vector $\vec{r}^{*}$ is now known and is transferred to the right-hand side of the inequalities.

The Least Squares Model

General

Diamond [7] was the first who worked on the model of least squares. For this purpose he introduces into the space $F(R)$ of triangular fuzzy numbers:

$$\widetilde{X}_1^{a=0} = (\vec{r}_1 - \vec{c}_1^{\ell}, \vec{r}_1, \vec{r}_1 + \vec{c}_1^{r}) \cdot X, \widetilde{X}_2^{a=0} = (\vec{r}_2 - \vec{c}_2^{\ell}, \vec{r}_2, \vec{r}_2 + \vec{c}_2^{r}) \cdot X \tag{12}$$

the following metric d:

$$d(\widetilde{X}_1,\widetilde{X}_2)^2 = ((\overrightarrow{r}_2 - \overrightarrow{r}_1)\cdot X)^2 + ((\overrightarrow{r}_2 - \overrightarrow{c}_2^{\ell})\cdot X - (\overrightarrow{r}_1 - \overrightarrow{c}_1^{\ell})\cdot X)^2 + ((\overrightarrow{r}_2 + \overrightarrow{c}_2^{\ell})\cdot X - (\overrightarrow{r}_1 + \overrightarrow{c}_1^{\ell})\cdot X)^2 \quad (13)$$

and proves that the metric space $(F(R), d)$ is complete. When the fuzzy numbers are symmetric: $\widetilde{X}_1 = (\overrightarrow{r}_1, \overrightarrow{c}_1)\cdot X, \widetilde{X}_2 = (\overrightarrow{r}_2, \overrightarrow{c}_2)\cdot X$ he introduces a function d_s as follows:

$$d_s(\widetilde{X}_1,\widetilde{X}_2)^2 = (X_2^{\ell} - X_1^{\ell})^2 + (X_2^r - X_1^r)^2 = ((\overrightarrow{r}_2 - \overrightarrow{c}_2^{\ell})\cdot X - (\overrightarrow{r}_1 - \overrightarrow{c}_1^{\ell})\cdot X)^2 + ((\overrightarrow{r}_2 + \overrightarrow{c}_2^{\ell})\cdot X - (\overrightarrow{r}_1 + \overrightarrow{c}_1^{\ell})\cdot X)^2 \quad (14)$$

and proves it to be an equivalent metric in the space of compact spaces. Here the metric of Diamond [7] is selected and the estimate of the symmetric number $\widetilde{Y}_1$ is examined:

$$\widetilde{Y}_i = \widetilde{A}_0 + \widetilde{A}_1 x_i, \widetilde{A}_0(r_0, c_0), \widetilde{A}_1 = (r_1, c_1)$$

or:

$$[\widetilde{Y}_{i=0}]^{a=0} = [r_0 - c_0 + (r_1 - c_1)x_i, r_0 + c_0 + (r_1 + c_1)x_i]^{a=0} \quad (15)$$

Symmetric Fuzzy Numbers

Now consider the case, where both the fuzzy numbers $(\widetilde{Y}_l^{a=0}, \tilde{y}^{a=0_l})$ are symmetric. Then the minimization problem becomes:

$$minimize S = \sum d_s(\widetilde{Y}_i^{a=0}, \tilde{y}_i^{a=0})^2 = \\ = \sum_{i=1}^{n}\{(y_i^{\ell} - (r_0 - c_0 + (r_1 - c_1)x_i))^2 + (y_i^r - (r_0 + c_0 + (r_1 + c_1)x_i))^2\} \quad (16)$$

Minimizing the quantity S is obtained by equating the derivatives of this quantity with respect to the parameters $(r_0 - c_0), (r_0 + c_0), (r_1 - c_1), (r_1 + c_1)$ to 0.

$$\frac{\partial S}{\partial(r_0 - c_0)} = 0 \rightarrow \bar{y}^{\ell} - (r_0 - c_0) - (r_1 - c_1)\bar{x} = 0, \\ \frac{\partial S}{\partial(r_1 - c_1)} = 0 \rightarrow \sum_{i=1}^{n} x_i y_i^{\ell} - n\bar{x}(r_0 - c_0) - (r_1 - c_1)\sum_{i=1}^{n} x_i^2 = 0 \quad (17)$$

$$r_1 - c_1 = \frac{\sum_{i=1}^{n} x_i y_i^\ell - n\overline{x}\overline{y}^\ell}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 - c_0 = \overline{y}^\ell - (r_1 - c_1)\overline{x} \tag{18}$$

$$\begin{aligned} &\frac{\partial S}{\partial(r_0 + c_0)} = 0 \to \overline{y}^r - (r_0 + c_0) - (r_1 + c_1)\overline{x} = 0, \\ &\frac{\partial S}{\partial(r_1 + c_1)} = 0 \to \sum_{i=1}^{n} x_i y_i^r - n\overline{x}(r_0 + c_0) - (r_1 + c_1)\sum_{i=1}^{n} x_i^2 = 0 \end{aligned} \tag{19}$$

$$r_1 + c_1 = \frac{\sum_{i=1}^{n} x_i y_i^r - n\overline{x}\overline{y}^r}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 + c_0 = \overline{y}^r - (r_1 + c_1)\overline{x} \tag{20}$$

In the above, the quantities $\overline{y}^\ell, \overline{x}, \overline{y}^r$ are given as follows:

$$\overline{y}^\ell = \frac{1}{n}\sum_{i=1}^{n} y_i^\ell, \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \overline{y}^r = \frac{1}{n}\sum_{i=1}^{n} y_i^r, y_i^\ell = y_i - e_i, y_i^r = y_i + e_i \tag{21}$$

considering that the fuzzy member $\tilde{y}_i = (y_i, e_i)$ is symmetric. From the above expressions, the quantities r_0, c_0, r_1, c_1 are derived and subsequently the values of the parameters $\widetilde{A}_0 = (r_0, c_0), \widetilde{A}_1 = (r_1, c_1)$:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, c_1 = \frac{\sum_{i=1}^{n} x_i e_i - n\overline{x}\overline{e}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} \tag{22}$$

$$r_0 = \overline{y} - r_1\overline{x}, c_0 = \overline{e} - c_1\overline{x}, \overline{e} = \frac{1}{n}\sum_{i=1}^{n} e_i \tag{23}$$

In case we have crisp experimental output data, it holds:

$$r_0 = \overline{y} - 2r_1\overline{x}, c_0 = 0, \widetilde{A}_0 = (r_0, 0), \widetilde{A}_1 = (r_1, 0) \cdot \tilde{y}_i \equiv y_i, y_i^\ell = y_i^r = y_i, e_i = 0 \tag{24}$$

and therefore we get:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, c_1 = 0. \tag{25}$$

Thus for this case (crisp experimental output) the least squares method gives crisp numbers for the parameters $\widetilde{A}_0, \widetilde{A}_1$ and cannot be applied, whereas the possibilistic method gives fuzzy numbers and can be applied.

Non-symmetric Triangular Fuzzy Numbers

In this case the triangular fuzzy number estimate and the measured triangular fuzzy number output are respectively:

$$\begin{aligned}
[\widetilde{Y}_i]^{a=0} &= [r_0 - c_0^{\ell} + (r_1 - c_1^{\ell})x_i, r_0 + c_0^r + (r_1 + c_1^r)x_i]^{a=0},\\
[\tilde{y}_i]^{a=0} &= [y_i - e_i^{\ell}, y_i + e_i^r]^{a=0}
\end{aligned} \tag{26}$$

The problem of minimization for both non-symmetric fuzzy numbers $(\widetilde{Y}_l^{a=0}, \tilde{y}_l^{a=0})$ is written as follows:

$$\begin{aligned}
minimize S &= d(\widetilde{Y}_l^{a=0}, \tilde{y}_l^{a=0})^2\\
&= \sum_{i=1}^{n}\{(y_i - (r_0 + r_1 x_i))^2 + (y_i^{\ell} - (r_0 - c_0^{\ell} + (r_1 - c_1^{\ell})x_i))^2\\
&\quad +(y_i^r - (r_0 + c_0^r + (r_1 + c_1^r)x_i))^2\}
\end{aligned} \tag{27}$$

Minimizing the quantity S is achieved by equating the derivatives of this quantity with respect to the parameters $r_0, r_1, (r_0 - c_0), (r_0 + c_0), (r_1 - c_1)$ and $(r_1 + c_1)$ to 0.

$$\begin{aligned}
&\frac{\partial S}{\partial r_0} = 0 \to \overline{y} - r_0 - r_1\overline{x} = 0,\\
&\frac{\partial S}{\partial r_1} = 0 \to \sum_{i=1}^{n} x_i y_i - n\overline{x}r_0 - r_1 \sum_{i=1}^{n} x_i^2 = 0
\end{aligned} \tag{28}$$

$$r_0 = \overline{y} - r_1\overline{x}, r_1 = \frac{\sum\limits_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum\limits_{i=1}^{n} x_i^2 - n\overline{x}^2} \tag{29}$$

$$\begin{aligned}
&\frac{\partial S}{\partial(r_0 - c_0^{\ell})} = 0 \to \overline{y}^{\ell} - (r_0 - c_0^{\ell}) - (r_1 - c_1^{\ell})\overline{x} = 0\\
&\frac{\partial S}{\partial(r_1 - c_1^{\ell})} = 0 \to \sum_{i=1}^{n} x_i y_i^{\ell} - n\overline{x}(r_0 - c_0^{\ell}) - (r_1 - c_1^{\ell}) \sum_{i=1}^{n} x_i^2 = 0
\end{aligned} \tag{30}$$

$$r_0 - c_0^\ell = \overline{y}^\ell - (r_1 - c_1^\ell)\overline{x}, r_1 - c_1^\ell = \frac{\sum_{i=1}^{n} x_i y_i^\ell - n\overline{x}\overline{y}^\ell}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2} \tag{31}$$

$$c_1^\ell = \frac{\sum_{i=1}^{n} x_i e_i^\ell - n\overline{x}\overline{e}^\ell}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2} \tag{32}$$

$$\begin{aligned} \frac{\partial S}{\partial (r_0 + c_0^r)} &= 0, \rightarrow \overline{y}^r - (r_0 + c_0^r) - (r_1 + c_1^r)\overline{x} = 0 \\ \frac{\partial S}{\partial (r_1 + c_1^r)} &= 0, \rightarrow \sum_{i=1}^{n} x_i y_i^r - n\overline{x}(r_0 + c_0^r) - (r_1 + c_1^r)\sum_{i=1}^{n} x_i^2 = 0 \end{aligned} \tag{33}$$

$$r_1 + c_1^r = \frac{\sum_{i=1}^{n} x_i y_i^r - n\overline{x}\overline{y}^r}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2}, r_0 + c_0^r = \overline{y}^r - (r_1 + c_1^r)\overline{x} \tag{34}$$

$$c_1^r = \frac{\sum_{i=1}^{n} x_i e_i^r - n\overline{x}\overline{e}^r}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2}, c_0^r = \overline{e}^r - c_1^r\overline{x} \tag{35}$$

Compatibility Condition

Diamond [7] defines a compatibility condition for the above model which is defined as follows:

$$\overline{e}\sum_{i=1}^{M}(x_{i1} - \overline{x})(y_i - \overline{y}_i) \geq \overline{x}\sum_{i=1}^{M}(x_{i1} - \overline{x})(e_i - \overline{e}) \geq 0 \tag{36}$$

The data set is called compatible and means that the trends between the center line and the spreads are the same, and therefore the values r_0, r_1, c_0, c_1 should be positive. Wang and Tsaur [23] generalized the above compatibility condition as follows:

(a) If $r_0 > 0, r_1 > 0$ and c_0, c_1 unconstrained, then:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 = \overline{y} - r_1\overline{x}, c_1 = \frac{\sum_{i=1}^{n} x_i e_i - n\overline{x}\overline{e}}{\sum_{i=1}^{n} x_i - n(\overline{x})^2}, c_0 = (\overline{e} - c\overline{x}) \tag{37}$$

(b) If $r_0 < 0, r_1 > 0, c_0 < 0$ and c_1 unconstrained, then:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 = \overline{y} - r_1\overline{x}, c_1 = \frac{\sum_{i=1}^{n} x_i e_i - n\overline{xe}}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2}, c_0^r = -(\overline{e} - c_1\overline{x}) \quad (38)$$

(c) If $r_0 > 0, r_1 < 0, c_0$ unconstrained and $c_1 < 0$, then:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 = \overline{y} - r_1\overline{x}, c_1 = -(\frac{\sum_{i=1}^{n} x_i e_i - n\overline{xe}}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2}), c_0 = (\overline{e} - c_1\overline{x}) \quad (39)$$

(d) If $r_0 < 0, r_1 < 0$ and $c_0 < 0, c_1 < 0$, then:

$$r_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}, r_0 = \overline{y} - r_1\overline{x}, c_1 = -(\frac{\sum_{i=1}^{n} x_i e_i - n\overline{xe}}{\sum_{i=1}^{n} x_i^2 - n(\overline{x})^2}), c_0 = -(\overline{e} - c_1\overline{x}) \quad (40)$$

Moreover, the case (d) is divided into two sub-cases:

(d1) $r_0 < 0, r_1 > 0, c_0 < 0$ and c_1 unconstrained. This is equal to (b).
(d2) $r_0 > 0, r_1 < 0, c_0$ unconstrained and $c_1 < 0$. This is equal to (c).

The procedure for fuzzy least squares is now defined as follows:

1. Compute the values of r_0, r_1.
2. If $r_0 > 0, r_1 > 0$, then the solution is equal to case (a)
3. If $r_0 < 0, r_1 > 0$, then if

$$\sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} x_{i1} e_i - \sum_{i=1}^{M} x_{i1}^2 \sum_{i=1}^{M} e_i \geq 0$$

the solution is (a), otherwise it is equal to (b).

4. If $r_0 > 0, r_1 < 0$, then if

$$\sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} x_{i1} e_i - \sum_{i=1}^{M} x_{i1}^2 \sum_{i=1}^{M} e_i \geq 0$$

the solution is (a), otherwise it is equal to (c).

5. If $r_0 < 0, r_1 < 0$, then if

$$\sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} x_{i1} e_i - \sum_{i=1}^{M} x_{i1}^2 \sum_{i=1}^{M} e_i \geq 0, \sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} e_i - N \sum_{i=1}^{M} x_{i1} e_i \geq 0$$

the solution is (a). If all are negative, then it is equal to (d). If it is either:

$$\sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} x_{i1} e_i - \sum_{i=1}^{M} x_{i1}^2 \sum_{i=1}^{M} e_i < 0$$

or

$$\sum_{i=1}^{M} x_{i1} \sum_{i=1}^{M} e_i - N \sum_{i=1}^{M} x_{i1} e_i \geq 0$$

it is equal to (b) or (c), respectively.

In the case of multiple regression, according to [23], the method of least squares is not convenient.

Application

We consider rainfall measurement stations of Aggistro and Upper Vrontou with the following data:

T	1929-30	1930-31	1931-32	1932-33	1933-34	1934-35	1935-36	1936-37	1937-38	1938-39	1939-40	1940-41
Aggistro-X	47,6	65,6	40,3	30,8	45,1	54,1	58,7	60,6	55,1	45,9	54,7	47,5
Upper Vrontou-Y	65,8	118,9	57	52,5	61,1	73,9	104,4	83,1	78,8	79	106,5	77,5
e	13,16	23,78	11,4	10,5	12,22	14,78	20,88	16,62	15,76	15,8	21,3	15,5

The Upper Vrontou data is considered a symmetric fuzzy number $\tilde{y} = (y_i, e_i)$.

Tanaka Model-Minimum Problem

For $h = 0$, the model is written as follows:

$\min(12c_0 + 606c_1)$
s.t.
$r_0 - c_0 + 47.6r_1 - 47.6c_1 \leq 52.64$
...
...
$r_0 - c_0 + 47.5r_1 - 47.5c_1 \leq 62$
$r_0 + c_0 + 47.6r_1 + 47.6c_1 \geq 78.96$

...
...
$r_0 + c_0 + 47.5r_1 + 47.5c_1 \geq 93$

The above is a linear programming problem and its solution is:

$$\widetilde{Y}_j = (0, 2.33032) + (1.71463, 0.57915)x_{1j},$$

where the supports are given by the equations:

$$[\widetilde{Y}_j^\ell = -2.33032 + 1.135484x_j, \widetilde{Y}_j^r = 2.330323 + 2.293778x_j],$$

while the kernel is given by the equation:

$$\widetilde{Y}_j^m = 1.714631x_j.$$

Both parameters $\widetilde{A}_0$ and $\widetilde{A}_1$ are given in Fig. 2 and the model of Tanaka is given in Fig. 3.

Method of Savic–Pedrycz

For $h = 0$, the model is written as follows:

First phase: The data are considered crisp and the model crisp. The regression line is equal to:

$$y = r_0 + r_1x, r_0 = -11.9413, r_1 = 1.818145.$$

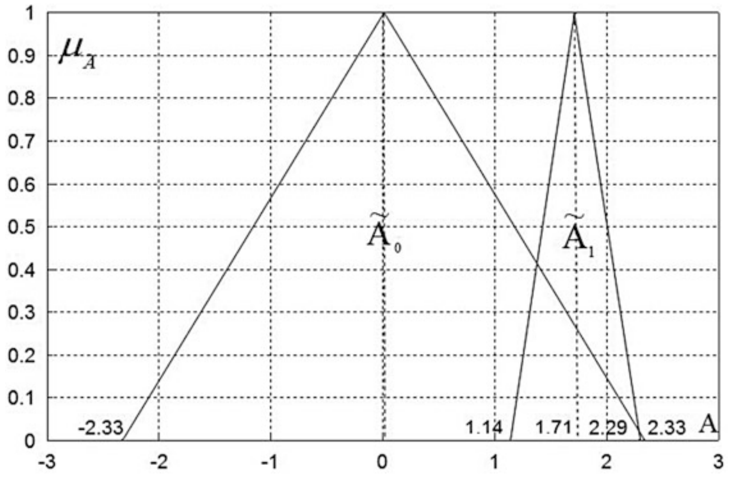


Fig. 2 Parameters $\widetilde{A}_0$ and $\widetilde{A}_1$

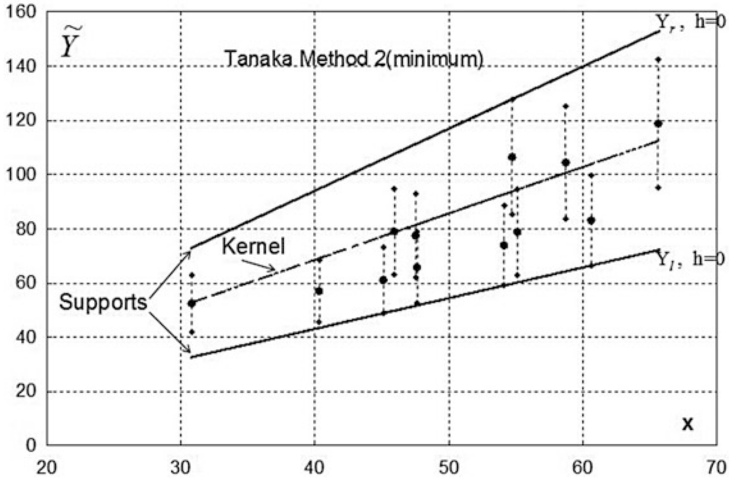


Fig. 3 Tanaka's model

Second phase: The model is written as follows:

$$
\begin{aligned}
&\min(12c_0 + 606c_1)\\
&\text{s.t.}\\
&-c_0 - 47.6c_1 \le -21.96\\
&\dots\\
&\dots\\
&-c_0 - 47.5c_1 \le -12.42\\
&c_0 + 47.6c_1 \ge 4.36\\
&\dots\\
&\dots\\
&c_0 + 47.5c_1 \ge 18.57
\end{aligned}
$$

The solution is: $c_0 = 0, c_1 = 0.73654$
The final regression line will be:

$$\widetilde{Y}_j = (r_0, 0) + (r_1, c_1)x = (-11.9413, 0) + (1.818145, 0.73654)x,$$

where the supports are given by the equations:

$$[\widetilde{Y}_j^{\ell} = -11.9413 + 1.081604x_j, \widetilde{Y}_j^{r} = -11.9413 + 2.554469x_j],$$

while the kernel is given by the equation:

$$\widetilde{Y}_j^{m} = -11.9413 + 1.818145x_j.$$

Parameter $\widetilde{A}_0$ is a crisp number while parameter $\widetilde{A}_1$ is given in Fig. 4. The model of Savic–Pedrycz is shown in Fig. 5.

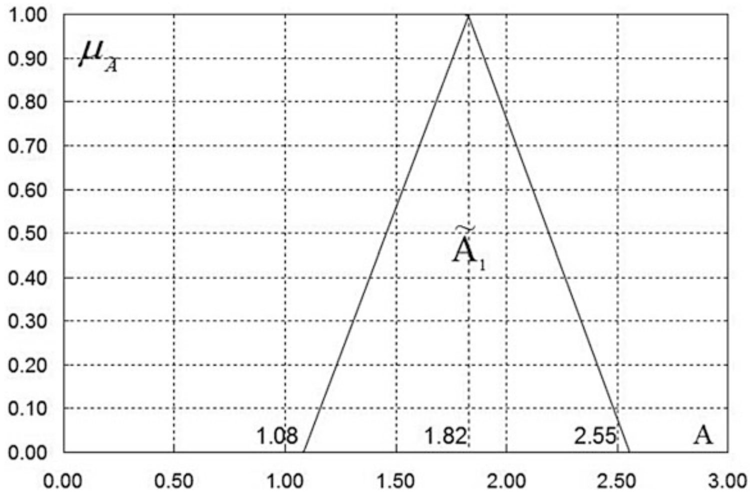


Fig. 4 Parameter $\widetilde{A}_1$

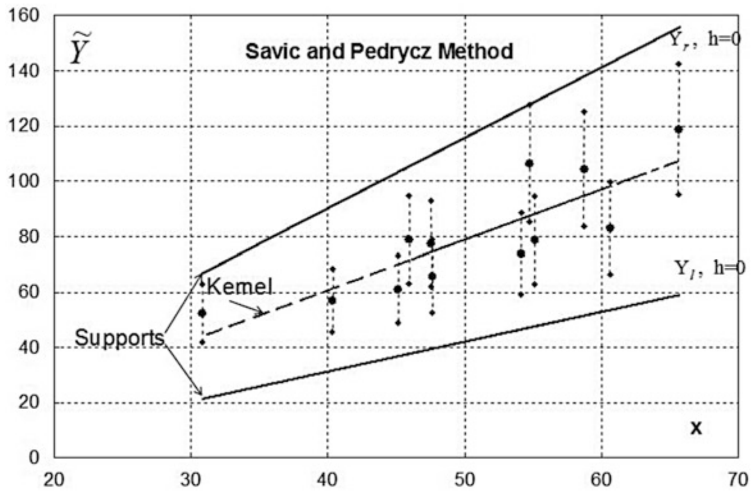


Fig. 5 Model of Savic–Pedrycz

Method of Least Squares

For h=0, the values of r_0, r_1, c_1 are equal to:

$$r_0 = -11.94113, r_1 = 1.818145, c_1 = 0.363629$$

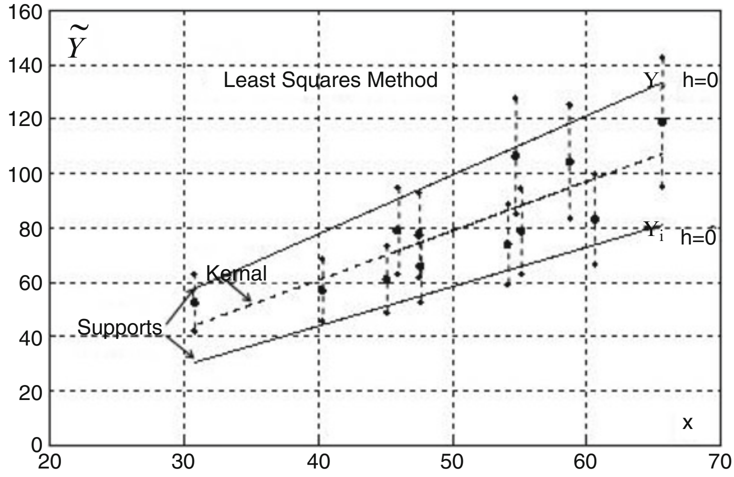


Fig. 6 The model of least squares

Because we have $r_0 < 0, r_1 > 0, c_1 < 0$, we belong to the case (b) and we get: $c_0^r = -(\overline{e}^r) - c_1^r\overline{x})$ or $c_0 = 2.388263$. The final regression line will be:

$$\widetilde{Y}_j = (r_0, c_0) + (r_1, c_1) = (-11.9413, 2.388263) + (1.818145, 0.363629),$$

the supports are given by:

$$[\widetilde{Y}_j^\ell = -14.329579 + 1.45452x_j, \widetilde{Y}_j^r = -9.55305 + 2.18177x_j],$$

while the kernel is given by the equation:

$$\widetilde{Y}_j^m = -11.9413 + 1.818145x_j$$

The model of least squares is shown in Fig. 6.

Distance Between Measured and Estimated Values

Since both the measured and estimated values are symmetric, the distance between the two number $(\widetilde{Y}_i^{a=0}, \tilde{y}_1^{a=0})$ is given by Diamond [8] as follows:

$$d_s(\widetilde{Y}_i^{a=0}, \tilde{y}_i^{a=0})^2 = \{y_i^\ell - (r_0 - c_0 + (r_1 - c_1)x_i)\}^2 + \{y_i^r - (r_0 + c_0 + (r_1 + c_1)x_i)\}^2$$

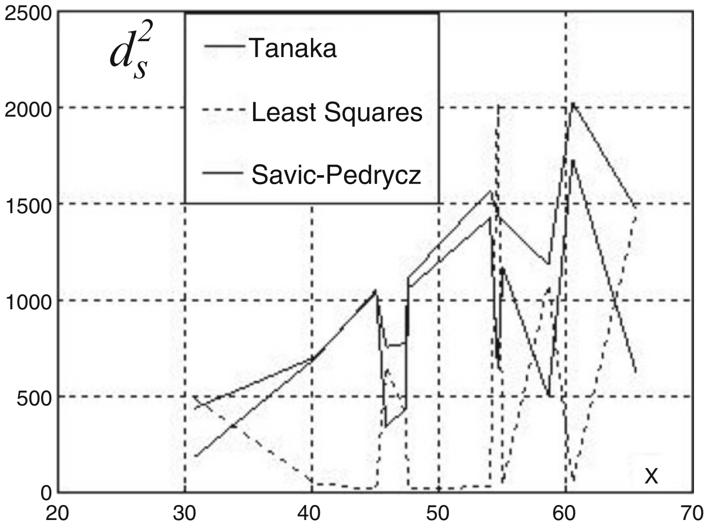


Fig. 7 The distance d_S^2

This distance is the sum of squared deviations of the supports of the two numbers $(\widetilde{Y}^{a=0}{}_i, \tilde{y}_i^{a=0})$. In Fig. 7 above, the distance d_S^2 is given as a function of x for the cases of Tanaka, Savic–Pedrycz and the least squares model.

Comments-Conclusions

With regard to the two models of Tanaka we make the following observations:

– These models both belong to the category of possibilistic models. In the case of the Tanaka model, the kernel is expressed by the equation $\widetilde{Y}_j^m = 1.714631x_j$ and it is the center line of the two supports, but does not coincide with the kernel of Savic–Pedrycz which is expressed by the equation $\widetilde{Y}_j^m = -11.9413 + 1.818145x_j$. The kernel of Savic–Pedrycz is also the middle (center) line of the two supports and it is a result of the linear regression of the mean values of the experimental output values y.
– In both models, their supports are the envelope of the fuzzy experimental values $\tilde{y}_j$. However, in the case of the Tanaka model, the distances of the supports from the experimental values (as shown in Fig. 7) are smaller than those of Savic–Pedrycz. The sum of these distances is:

$$\sum_{Tanaka} d_S(\widetilde{Y}_i^{a=0}, \tilde{y}_i^{a=0})^2 = 9900 < \sum_{Savic--Pedrycz} d_S(\widetilde{Y}_i^{a=0}, \tilde{y}_i^{a=0})^2 = 13848$$

- The model of least squares shows the smallest distance of the supports from the experimental values compared to the other two models as shown in Fig. 7. The sum of these distances is:

$$\sum_{LeastSquares} d_S(\widetilde{Y}_i^{a=0}, \tilde{y}_i^{a=0})^2 = 3405$$

- In the model of least squares, the kernel has the same equation as that of Savic–Pedrycz, that is, it is also as a result of the linear regression of the mean values of experimental output values y. But as we observed from Fig. 6, its supports are not the envelope of the fuzzy experimental values because in this model, the subset-hood of measured values is not included.
- Also in the case where we have crisp experimental output data, the least squares method gives crisp numbers for the parameters $\widetilde{A}_0, \widetilde{A}_1$ and cannot be applied while the possibilistic method gives fuzzy numbers and can be applied.
- On the basis of these models, it is possible to predict the values for the station of Upper Vrontous (Y) when fed with the values of Aggistro station, and therefore the possibility of replacing the missing values of the station Y.

References

1. Arabpour, A.R., Tata, M.: Estimating the parameters of a fuzzy linear regression model. Iran. J. Fuzzy Syst. **5**(2), 1–19 (2008)
2. Bargiela, A., Pedrycz, W., Nahashina, T.: Multiple regression with fuzzy data. Fuzzy Sets Syst. **158**, 2169–2188 (2007)
3. Bardossy, A., Bogardi, I., Ducstein, L.: Fuzzy regression in hydrology. Water Resour. Res. **26**(7), 1497–1508 (1990)
4. Chang, P.-T., Lee, E.S.: A generalized fuzzy weighted least-squares regression. Fuzzy Sets Syst. **82**, 289–298 (1995)
5. Charfeddine, S., Mora-Camin F., De Coligny, M.: Fuzzy linear regression: Application to the estimation of air transport demand. In: International Conference of Fuzzy Sets and Soft Computing in Economics and Finance, Saint-Petersburg (2004)
6. Chen, Y.S.: Outliers detection and confidence interval modification in fuzzy regression. Fuzzy Sets Syst. **119**, 259–272 (2001)
7. Diamond, P.: Fuzzy least squares. Inf. Sci. **46**, 141–157 (1988)
8. Diamond, P.: Higher level fuzzy numbers arising from fuzzy regression models. Fuzzy Sets Syst. **36**, 265–275 (1990)
9. Diamond, P., Tanaka, H.: Fuzzy regression analysis. In: Fuzzy Sets in Decision Analysis, Operations Research and Statistics, vol. 387. Kluwer Academic Publishers, Boston (1999)
10. Hung, W.-L., Yang, M.-S.: An omission approach for detecting outliers in fuzzy regression models. Fuzzy Sets Syst. **157**, 3109–3122 (2006)
11. József, S.: On the effect of linear data transformations in possibilistic fuzzy linear regression. Fuzzy Sets Syst. **45**, 185–188 (1992)
12. Moskowitz, H., Kim, K.: On assessing the H value in fuzzy linear regression. Fuzzy Sets Syst. **58**, 303–327 (1993)
13. Papadopoulos, B., Sirpi, M.: Similarities in fuzzy regression models. J. Optim. Theory Appl. **102**(2), 373–383 (1999)

14. Papadopoulos, B., Sirpi, M.: Similarities and distances in fuzzy regression models. Soft Comput. **8**(8), 556–561 (2004)
15. Pasha, E., Razzaghnia T., Allahvirando T., Yari G., Mostafaei, H.R.: Fuzzy linear regression models with fuzzy entropy. Appl. Math. Sci. **1**(35), 1715–1724 (2007)
16. Redden, D.T., Woodall, W.H.: Further examination of fuzzy linear regression. Fuzzy Sets Syst. **79**, 203–211 (1996)
17. Savic, D.A., Pedrycz, W.: Further examination of fuzzy linear regression. Fuzzy Sets Syst. **39**, 51–63 (1991)
18. Tanaka, H.: Fuzzy data analysis by possibilistic linear models. Fuzzy Sets Syst. **24**, 363–375 (1987)
19. Tanaka H., Hayashi, I.: Possibilistic linear regression analysis for fuzzy data. Eur. J. Oper. Res. **40**, 389–396 (1989)
20. Tanaka, H., Ishibuchi, H.: Identification of possibilistic linear systems by quadratic membership functions of fuzzy parameters. Fuzzy Sets Syst. **41**, 145–160 (1991)
21. Tanaka, H., Watada, J.: Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets Syst. **27**, 275–289 (1988)
22. Tanaka, H., Uejima S., Asai, K.: Fuzzy linear regression models. IEEE Trans. Man Cybern. **12**, 903–907 (1982)
23. Wang, H.-F., Tsaur, R.-C.: Theory and methodology-resolution of fuzzy regression model. Eur. J. Oper. Res. **126**, 637–650 (2000)
24. Xin, Y., Su, X.G.: Linear Regression Analysis: Theory and Computing. World Scientific Publishing Company, Singapore; Hackensack, NJ (2009)
25. Xu, R., Li, C.: Multidimensional least-squares fitting with a fuzzy model. Fuzzy Sets Syst. **119**, 215–223 (2001).
26. Yang, M.-S., Liu, H.-H.: Fuzzy least-squares algorithms for interactive fuzzy linear regression models. Fuzzy Sets Syst. **135**, 305–316 (2003)

Properties of Functions of Generalized Bounded Variations

Rajendra G. Vyas

Abstract Looking to the features of functions of bounded variation, the notion of bounded variation is generalized in many ways and different classes of functions of generalized bounded variations are introduced. In the present chapter introducing different classes of functions of generalized bounded variations their main interesting properties are discussed. Inter-relations between them and the classes related to them are given in the next section. Finally we try to present overall picture of Fourier analysis of these classes.

Keywords Absolutely continuous function • Banach space • Banach Algebra • Bounded variation • Convolution product • Fourier series • Fourier coefficients • Module • Ring

Introduction

In 1881, while studying the convergence of Fourier series, Jordan [14] introduced the concept of bounded variation. This class BV, of functions of bounded variation, has lots of applications in almost all branches of pure and applied mathematics. One of the features which distinguished the class BV from other classes is that its analysis can be carried in most elegant ways. The class has many interesting applications in theory of functions, harmonic analysis, approximation theory, functional analysis, measure and integration, partial differential equations, numerical analysis, special functions,fuzzy theory and digital image processing.

The class $\mathrm{BV}([0, 2\pi])$ is a Banach algebra with suitable variation norm, Fourier series of a function of this class converges everywhere pointwise, the class can be regarded as a module over the ring $L^1([0, 2\pi])$, the class of all 2π-periodic Lebesgue integrable functions over $[0, 2\pi]$, with respect to the convolution product. Other important aspect of functions of BV is that they form an algebra of discontinuous functions whose first derivative exists almost everywhere; due to this fact, they can

R.G. Vyas (✉)
Faculty of Science, Department of Mathematics, The M. S. University of Baroda, Vadodara, India
e-mail: drrgvyas@yahoo.com

T.M. Rassias, V. Gupta (eds.), *Mathematical Analysis, Approximation Theory and Their Applications*, Springer Optimization and Its Applications 111,
DOI 10.1007/978-3-319-31281-1_30

be frequently used to define generalized solutions of nonlinear problems involving functionals, ordinary and partial differential equations in mathematics, physics, and engineering. By suitable choice of functions of the class BV it is possible to obtain solutions to many integral equations which describe some physical phenomena. The class BV includes the Sobolev class $W^{1,1}$ and can be equivalently defined in terms of measure distributional derivatives or L^1-limits of bounded sequences in $W^{1,1}$, which has many interesting properties in digital image processing.

Objective of the second section is to recall some basic definitions and important properties of the class BV and some other standard classes like: AC, the class of functions of absolutely continuous functions; the class of functions which satisfy Lipschitz condition etc.

Properties of Functions of Bounded Variation

Definition 1. Given an interval $I = [a, b]$, we say that a complex function f defined on I is said to be of Bounded variations (that is, $f \in BV(I, C)$) if

$$V(f, I) = \sup_P \{V(P, f, I)\} < \infty, \quad \textit{where}$$

$$V(P, f, I) = (\sum_{m=1}^{n} \mid \triangle f(I_m) \mid), \ \textit{in which}$$

$\triangle f(I_m) = f(x_m) - f(x_{m-1})$ and $P = \{a = x_0, x_1, \ldots, x_n = b\}$ is a partition of I.

For real valued functions, the class of functions of bounded variation over $[a, b]$ is denoted by $BV(I)$, that is we omit writing C. Similarly, one can define BV for a function from a compact subset of **R** into a Banach space or a metric space.

Some of the well-known properties of functions of bounded variation are as follows:

(i) Every function of the class BV(I) is bounded but the converse is not true.
(ii) The class BV(I) is a commutative unital Banach algebra with respect to the pointwise operations and the variation norm $|||f||| = ||f||_\infty + V(f, I)$.
(iii) It is a module over the ring $L^1([0, 2\pi])$ with respect to the convolution product.
(iv) The class $NBV([a, b])$ of functions of normalized bounded variation on $[a, b]$ is the Dual of the Banach space $C([a, b])$, of all continuous functions on $[a, b]$.

Theorem 1. *If $f \in BV[a, b]$, then for any $a < c < b$, $V(f, [a, b]) = V(f, [a, c]) + V(f, [c, b])$.*

Proof. Let $P* = P \cup \{c\}$ be the refinement of a partition P of $[a, b]$. Then $P* = P_1 \cup P_2$, where P_1 and P_2 are partitions of $[a, c]$ and $[c, b]$, respectively. Then
$V(f, P) \le V(f, P_1) + V(f, P_2) \le V(f, [a, c]) + V(f, [c, b])$ implies

$$V(f, [a, b]) \le V(f, [a, c]) + V(f, [c, b]). \tag{1}$$

For any $\epsilon > 0$ there exist partitions P_3 and P_4 of $[a, c]$ and $[c, b]$, respectively, such that

$$V(f, [a, c]) - \frac{\epsilon}{2} < V(f, P_1) \text{ and } V(f, [c, b]) - \frac{\epsilon}{2} < V(f, P_2).$$

Thus

$$V(f, [a, c]) + V(f, [c, b]) - \epsilon < V(f, P_1) + V(f, P_2)$$
$$= V(f, P_1 \cup P_2) \leq V(f, [a, b]), \ \forall \epsilon > 0.$$

Hence, the result follows from (1).

Theorem 2. *If $f \in BV[a, b]$, then the function $g(x) = V(f, [a, x]), \ \forall x \in [a, b]$, is an increasing function. Moreover for any $x \in (a, b)$ we have $g(x+0) - g(x) = |f(x+0) - f(x)|$ and $g(x) - g(x-0) = |f(x) - f(x-0)|$.*

Proof. Let $S = |f(t+0) - f(t)|$ for any fixed but arbitrary $t \in (a, b)$. Given any $\epsilon > 0$, there exist $\delta > 0$ such that $t < x < t + \delta$ implies $S - \frac{\epsilon}{2} < |f(t+0) - f(t)| < S + \frac{\epsilon}{2}$.

Let $P = \{t = x_0, x_1, x_2, \ldots, x_n = b\}$ be a partition of $[t, b]$ such that $V(f, [t, b]) - \frac{\epsilon}{2} < V(f, P)$. Without the loss of generality we assume that the point x_1 can be treated as arbitrary point satisfying $t < x_1 < t + \delta$. For $P* = P - \{t\}$ a partition of $[x_1, b]$ we have

$$V(f, p) - V(f, P*) = V(f, [t, x_1]) = |f(x_1) - f(t)| \geq V(f, [t, b]) - V(f, P^*) - \frac{\epsilon}{2}.$$

Thus,

$$S - \epsilon < S - \frac{\epsilon}{2} < |f(x_1) - f(t)| \leq g(x_1) - g(t) \leq |f(x_1) - f(t)| + \frac{\epsilon}{2} < S + \epsilon .$$

Hence, the result follows.

Thus f is continuous at a point if and only if g is continuous at that point.

Theorem 3 (Jordon Decomposition Theorem). *Every function of bounded variation can be expressed as difference of two monotonic functions.*

Proof. For any $f \in BV[a, b]$ and for any $x, y \in I, x < y$, it is known that

$$V(f, [a, y]) = V(f, [a, x]) + V(f, [x, y]).$$

Define $g(x) = V(f, [a, x])$ *and* $k(x) = g(x) - f(x), \forall x \in [a, b]$.

Obviously g is monotonically increasing and $g(y) - g(x) = V(f, [x, y]) \geq |f(y) - f(x)| \geq f(y) - f(x)$ implies k is also increasing. Thus f can be expressed as difference of two monotonic functions g and k.

This characterization has lots of applications in Analysis. It's well known direct application is: "If $f \in BV([a, b])$ then f is differentiable a.e. in $[a, b]$". One of its important applications is Helly's theorem which has lots of applications in function spaces and is generalized in many ways. The following Lemma is used to prove Helly's theorem.

Lemma 1. *If an infinite family of functions $F = \{f(x)\}$, defined on I, is uniformly bounded by a constant M that is $||f||_\infty \leq M, \quad \forall f \in F$, then the following hold.*

(i) For any countable subset E of I, it is possible to find a sequence $\{f_n(x)\}$ in the family F which converges pointwise in E.
(ii) Whenever F is an increasing family of functions, there exists a sequence of functions $\{f_n(x)\}$ in F which converges to an increasing function $\phi(x)$ at every point of I.

Proof. Let $E = \{x_k\}$ be any countable subset of $[a,b]$. Consider a set $\{f(x_1)\}$ of values taken on by the functions of the family F at the point x_1. This set is bounded and, by Bolzano–Weierstrass Theorem, it has a convergent subsequence:

$$f_1^{(1)}(x_1), f_2^{(1)}(x_1), f_3^{(1)}(x_1), \cdots; \quad \lim_{n\to\infty} f_n^{(1)}(x_1) = A_1.$$

Similarly, the sequence of functions $\{f_n^{(1)}\}$ gives a convergent subsequence

$$f_1^{(2)}(x_2), f_2^{(2)}(x_2), f_3^{(2)}(x_2), \cdots; \quad \lim_{n\to\infty} f_n^{(2)}(x_2) = A_2.$$

Continuing this process indefinitely, we construct a countable set of convergent sequences:

$$f_1^{(k)}(x_k), f_2^{(k)}(x_k), f_3^{(k)}(x_k), \cdots; \quad \lim_{n\to\infty} f_n^{(k)}(x_k) = A_k, \ \textit{for } k = 1,2,3\ldots. \ .$$

Form the sequence of diagonal elements of the infinite matrix, construct the sequence $\{f_n^{(n)}(x_k)\}$. This sequence converges at every point of the set E. In fact, for every fixed k, the sequence $\{f_n^{(n)}(x_k)\}$ converges to A_k. Hence, (i) follows.

Taking E as the set consisting all rational of I. In view of (i) of the lemma, there exist a sequence of functions of the family F, say $F_0 = \{f^{(n)}(x)\}$ such that $\lim_{n\to\infty} f^{(n)}(x_k)$ exists and is finite for all x_k in E.

Define $\psi(x) = \lim_{n\to\infty} f^{(n)}(x_k)$, for all $x_k \in E$. Then, $\psi(x)$ is increasing function on E, that is, if $x_i, x_k \in E$ and $x_k \leq x_i$, then $\psi(x_k) \leq \psi(x_i)$.

For $x \in [a,b] - E$ define $\psi(x)$ by $\psi(x) = \sup_{x_k<x}\{\psi(x_k)\}$, for all $x_k \in E$. Obviously, $\psi(x)$ is an increasing function on $[a,b]$ and it has at the most countable simple discontinuity. Show that

$$\lim_{n\to\infty} f^{(n)}(x) = \psi(x), \ \textit{at every point x where } \psi(x) \textit{ is continuous.} \qquad (2)$$

Let $\varepsilon > 0$, and let x_k and x_i be points in E such that $x_k < x_0 < x_i, \quad \psi(x_i) - \psi(x_k) < \frac{\varepsilon}{2}$. Fixing the points x_k and x_i, select the natural number $n > n_0$,

$$|f^{(n)}(x_k) - \psi(x_k)| < \frac{\varepsilon}{2}, \quad |f^{(n)}(x_i) - \psi(x_i)| < \frac{\varepsilon}{2}.$$

It is easy to see that $\psi(x_0) - \varepsilon < f^{(n)}(x_k) \leq f^{(n)}(x_i) < \psi(x_0) + \varepsilon$, for $n > n_0$. Since $f^{(n)}(x_k) \leq f^{(n)}(x_0) \leq f^{(n)}(x_i)$, we have $\psi(x_0) - \varepsilon < f^{(n)}(x_0) < \psi(x_0) + \varepsilon$, for $n > n_0$. This proves (2). Thus, the equality (2) can fail only at finite or countable set Q, where $\psi(x)$ is discontinuous.

We now apply Lemma 1(i) to the sequence F_0, taking for the set E the set of those points of Q where (2) does not hold. This yields a subsequence $\{f_n(x)\}$ of F_0, which converges at all points of I. Setting $\phi(x) = \lim_{n\to\infty} f_n(x)$, we obtain a function which is obviously an increasing function. Hence, the Lemma follows.

Theorem 4 (Helly's First Theorem). *Let an infinite family of functions $F = \{f(x)\}$ be defined on I such that each member of the family and its total variation are bounded by one and the same constant M; that is, $|f(x)| \leq M,\ V(f; I) \leq M,\ \forall f \in F$; then, there is a sequence of functions $\{f_n(x)\}$, in F, converges pointwise, to some function $\phi \in BV(I)$, on I.*

Proof. For any $f \in F$, set $h(x) = V(f; [a, x])$ and $g(x) = h(x) - f(x)$. Then $h(x)$ and $g(x)$ are increasing as well as bounded functions. In view of Lemma 1, the family $\{h(x)\}$ has a convergent sequence $\{h_k(x)\}$. Define $\alpha(x) = \lim_{k\to\infty} h_k(x)$. Similarly, the family $\{g_k(x)\}$ has a convergent subsequence $\{g_{k_i}(x)\}$. Define $\beta(x) = \lim_{i\to\infty} g_{k_i}(x)$. Then the sequence of functions $f_{k_i}(x) = h_{k_i}(x) - g_{k_i}(x)$, in F, converges to the function $\phi(x) = \alpha(x) - \beta(x)$. This proves the Helly's theorem.

Definition 2. A complex function f defined on $[a, b]$ is said to be a regulated function if it has right-hand and left-hand limit at every point of the interval $[a, b)$ and $(a, b]$, respectively.

Let $\mathbf{G}(I, C)$ be the class of all regulated complex functions, for a real function we write omitting C. $\mathbf{G}(I, C)$ is a commutative C^*-algebra with unity, with respect to the pointwise operations and the sup norm. Moreover it is closed $*$-subalgebra of the commutative C^*-algebra of $\mathbf{B}(I, C)$, class of all bounded complex functions on I. Since, every monotonic function has at the most countable simple discontinuity, $BV(I) \subset \mathbf{G}(I) \subset \mathbf{B}(I)$.

A complex function on I is called a step function if there exist a partition of I into finitely many disjoint subintervals on each of which f is constant. Let $S(I, C)$ be the class of all step functions. Clearly $S(I, C) \subset BV(I, C) \subset \mathbf{G}(I, C) \subset \mathbf{B}(I, C)$. Moreover, $S(I, C)n$ is dense subset of $\mathbf{G}(I, C)$ with respect to the sup-norm. $\mathbf{G}(I, C)$ is the completion of both $BV(I, C)$ and $S(I, C)$. Note that $BV(I, C)$ is not complete with respect to the sup-norm. Thus, every regulated function can be uniformly approximated by a step function as well as by a function of bounded variation.

Observe that the class of all real continuous functions, $C([a, b])$ is a closed subspace of $\mathbf{G}(I)$. Properties of these classes motivated to introduce the Stieltjes integral as framework for solving ODEs.

Definition 3 (Stieltjes Integral). For given real functions f and g on $[a, b]$ and a partition $P = \{a = x_0, x_1, \ldots, x_n = b\}$ of $[a, b]$.

The Stieltjes integral is defined as

$$\int_a^b f(t)dg(t) = \lim_{\delta \to 0} \sum_{i=1}^{n} f(c_i)(g(x_i) - g(x_{i-1})), \quad c_i \in [x_{i-1}, x_i],$$

if the limit exists, where $\delta = Min\{|x_i - x_{i-1}| : i = 1, 2, \ldots n\}$.

Theorem 5 (Properties of the Stieltjes Integral).

(i) If $f \in \mathbf{G}(I)$ *and* $g \in BV(I)$*, then* $\int_a^b f(t)dg(t)$ *exist and satisfy*

$$|\int_a^b f(t)dg(t)| \leq ||f||_\infty V(g, I)$$

(ii) If $\{f_n\}$ *and* $\{g_n\}$ *are sequences in* $\mathbf{G}(I)$ *and* $f_n \in BV(I)$*, respectively, such that they converge uniformly to* f *and* g*, respectively, and* $V(g_n, [I]) \leq M$*,* $\forall n$*, then*

$$\lim_{n \to \infty} \int_a^b f_n(t)dg_n(t) = \int_a^b f(t)dg(t).$$

Let $BV_C(I) = BV(I) \cap C(I)$. Obviously $BV_C(I)$ is closed in $BV(I)$ with respect to the variation norm. Hence $BV_C(I)$ is a Banach algebra with respect to the variation norm. It is observed that the variation function inherits smoothness properties of the function.

Theorem 6. *If* $f \in BV_C([a, b])$*, then*

$$V(f; [a, b]) = \lim_{h \to 0+} \frac{1}{h} \int_a^{b-h} |f(x) - f(x + h)| \, dx.$$

Proof. Since $|f(x) - f(x + h)| \leq V(f; [a, x + h]) - V(f; [a, x])$, it follows that

$\int_a^{b-h} |f(x) - f(x + h)| \, dx$
$\leq \int_a^{b-h} (V(f; [a, x + h]) - V(f; [a, x])) \, dx$
$= \int_a^{b-h} V(f; [a, x + h]) \, dx - \int_a^{b-h} V(f; [a, x]) \, dx$

$$= \int_{a+h}^{b} V(f; [a, x]) \, dx - \int_a^{b-h} V(f; [a, x]) \, dx$$

$$= \int_{b-h}^{b} V(f; [a, x]) \, dx - \int_a^{a+h} V(f; [a, x]) \, dx$$

$\leq h \, V(f; I), \quad \forall \, h > 0.$

Thus, $\lim_{h \to 0+} \sup \frac{1}{h} \int_a^{b-h} |f(x) - f(x + h)| \, dx \leq V(f; I)$.

Now let $P : a = x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_n = b$ be any partition of I. For any h small enough to ensure that $x_{n-1} < b - h$ and for $i \in \mathbb{N}$ such that $0 \leq i \leq n-2$ the following is true.

$$\begin{aligned}\frac{1}{h}\int_{x_i}^{x_{i+1}} |f(x) - f(x+h)|\, dx &\geq |\frac{1}{h}\int_{x_i}^{x_{i+1}} (f(x) - f(x+h))\, dx| \\ &= |\frac{1}{h}\int_{x_i}^{x_{i+1}} f(x)\, dx - \frac{1}{h}\int_{x_i+h}^{x_{i+1}+h} f(x)\, dx| \\ &= |\frac{1}{h}\int_{x_i}^{x_i+h} f(x)\, dx - \frac{1}{h}\int_{x_{i+1}}^{x_{i+1}+h} f(x)\, dx|.\end{aligned}$$

The right-hand side tends to $|f(x_{i+1}) - f(x_i)|$ as $h \to 0^+$ according to the fundamental theorem of calculus. So,

$$\lim_{h\to 0^+} \inf \frac{1}{h}\int_{x_i}^{x_{i+1}} |f(x) - f(x+h)|\, dx \geq |f(x_{i+1}) - f(x_i)|, \textit{ for } 0 \leq i \leq n-2.$$

Similarly, one gets $\lim_{h\to 0^+} \inf \frac{1}{h}\int_{x_{n-1}}^{b-h} |f(x) - f(x+h)|\, dx \geq |f(b) - f(x_{n-1})|$.

Summing these inequality,

$$\lim_{h\to 0^+} \inf \frac{1}{h}\int_{a}^{b-h} |f(x) - f(x+h)|\, dx \geq \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| = V(f, P).$$

Thus, $V(f; I) \leq \lim_{h\to 0^+} \inf \frac{1}{h}\int_{a}^{b-h} |f(x) - f(x+h)|\, dx$

$$\leq \lim_{h\to 0^+} \sup \frac{1}{h}\int_{a}^{b-h} |f(x) - f(x+h)|\, dx \leq V(f; I).$$

This proves the Theorem.

Important subspace of the space $BV([a, b])$ are the class $AC([a, b])$, the class of absolutely continuous functions on $[a, b]$ and $\mathrm{Lip}(\alpha, [a, b])$, the class of functions satisfying the Lipschitz condition of order α, where $0 < \alpha \leq 1$.

Definition 4. A function $f : I \to \mathbb{R}$ is said to be absolutely continuous on I if for each $\epsilon > 0$ there exists $\delta > 0$ such that $\sum_{i=1}^{n} |f(d_i) - f(c_i)| < \epsilon$ whenever $\{[c_i, d_i] : 1 \leq i \leq n\}$ is a finite collection of non-overlapping subintervals of I such that $\sum_{i=1}^{n}(d_i - c_i) < \delta$.

It is observed that every absolutely continuous functions is of bounded variation but the converse is not true. Obviously, an absolutely continuous function is uniformly continuous but the converse is not true. This can be followed from the following well-known example of the Cantor function.

Example 1. Define a sequence of functions $\{f_n\}$ from $[0,1]$ in to $[0,1]$ converges uniformly to the function call Cantor function, which is defined as follows.

Let $f_0(x) = x,\ \forall\ x \in [0,1]$. Then for every integer $n \geq 0$ the next function f_{n+1} is defined in terms of f_n as follows.

$$f_{n+1}(x) = \begin{cases} f_n(3x)/2, & \text{if x}\in[0,1/3], \\ 1/2, & \text{if x}\in[1/3,2/3], \\ 1/2 + f_n(3x-2)/2, & \text{if x}\in[2/3,1]. \end{cases}$$

Uniform convergence of the sequence $\{f_n\}$ can be followed from the following inequalities.

$$||f_{n+1} - f_n||_\infty \leq \frac{1}{2}||f_n - f_{n-1}||_\infty \leq \frac{1}{2^n}||f_1 - f_0||_\infty \to 0\ as\ n \to \infty.$$

The Cantor function f is also defined as:

For $x \in [0,1]$ consider its ternary expansion $\{a_i\}$, that is $x = \sum_{i=0}^{\infty} \frac{a_i}{3^i}, \quad a_i \in \{0,1,2\}$. Then

$f(x) = \sum_{i=0}^{\infty} \frac{a_i}{2^{i+1}}$, if none of the coefficients a_i takes the value 1 and
$f(x) = \sum_{i=0}^{n(x)-1} \frac{a_i}{2^{i+1}} + \frac{1}{2^{n(x)}}$ where $n(x)$ is the smallest integer n such that $a_n = 1$, otherwise.

Now, we show that the Cantor function is not absolutely continuous:
Consider the set of intervals

$\{[0,1/3],\ [2/3,1]\}$
$\{[0,1/9],\ [2/9,1/3],\ [2/3,7/9],\ [8/9,1]\}$
$\{[0,1/27],[2/27,1/9],[2/9,7/27],[8/27,1/3],[2/3,19/27],[20/27,7/9],$
$[8/9,25/27],[26/27,1]\}$

Label these intervals by $\{[c_i,d_i]:\ 1 \leq i \leq n\}$.
Then $f(d_i) = f(c_{i+1})$ implies

$$\sum_{i=1}^{n} |f(d_i) - f(c_i)| = \sum_{i=1}^{n} f(d_i) - f(c_i) = 1.$$

Here, the sum of the length of these intervals is $\left(\frac{2}{3}\right)^n \to 0\ as\ n \to \infty$.

Form the definition, it follows that f is locally constant on the Ternary set E (that is complement of the Cantor set) and the measure of E is 1.

Since the Cantor function is continuous and monotonically increasing on $[0,1]$, $f \in BV([0,1])$.

Observe that the derivative of f vanishes almost everywhere and

$$f(1) - f(0) \neq \int_0^1 f'(t)\,dt.$$

Obviously, the class of all absolutely continuous functions on I is a Banach subalgebra of BV(I).

Definition 5. A real or a complex function f defined in I is said to be Lipschitz function of the order $\alpha, 0 < \alpha \leq 1$, (that is, $f \in Lip(\alpha, I)$), if

$$|f(x) - f(y)| = O(|x - y|^{\alpha}, \quad \forall\, x, y \in I.$$

For $\alpha = 1$, we omit writing α.

Thus, we have

$$C^1(I) \subset Lip(I) \subset AC(I) \subset BV(I) \subset L^p(I) \subset L^1(I), \quad for \quad p > 1.$$

Problem 1. A function satisfies Lipschitz condition on $[a, b]$, if and only if it is the integral of a bounded function that is

$$f(x) = f(a) + \int_a^x g(t)\, dt, \quad \forall x \in [a, b], \quad with \quad g \in L^{\infty}[a, b].$$

Problem 2. The vector space $Lip([a, b])$ is a Banach space with respect to the pointwise linear operations and the norm $||f||_{Lip} = |f(a)| + M(f), f \in Lip([a, b])$, where

$$M(f) = \sup_{a \leq x < y \leq b} \left\{ \frac{|f(y) - f(x)|}{|y - x|} \right\}.$$

Problem 3. The vector space $C^1([a, b])$, of all real valued functions on $[a, b]$ with first order continuous derivatives, is a Banach space with respect to the norm

$$|||f|||_1 = ||f||_{\infty} + ||f'||_{\infty}.$$

It is easy to prove that every Lipschitz function is absolutely continuous. But the converse need not be true.

Example 2. The function $f(x) = \sqrt{x},\ x \in [0, 1]$, is absolutely continuous but $f \notin Lip([0, 1])$.

For any $\epsilon > 0$ taking $\{[a_i, b_i];\ 1 \leq i \leq n\}$ to be non-overlapping collection of intervals in [0,1] such that $\sum_i (b_i - a_i) < \delta = \epsilon^2/4$. Choose $A = \epsilon^2/4$, break the sum into two parts over [0,A] and over [A,1]. Without the loss of generality, assume that $A = b_m$. Then first part $\sum_{i=1}^m |f(b_i) - f(a_i)| \leq \sqrt{A} = \epsilon/2$,

and second part

$$\sum_{i=m+1}^{n} |f(b_i) - f(a_i)| = \sum_{i=m+1}^{n} \frac{b_i - a_i}{\sqrt{b_i} + \sqrt{a_i}} \leq \frac{1}{2\sqrt{A}} (\sum_{i=m+1}^{n} (b_i - a_i)) < \epsilon .$$

Hence, $f \in AC[0,1]$.

For any $x, y \in (0,1]$,

$|f(x) - f(y)| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \to \infty$ as either x, or y, or both x and y tends to 0.

This implies that $f \notin Lip([0,1])$.

Note that $f \in AC[0,1]$ has unbounded derivative on (0,1). That is, $f \notin C^1[0,1]$.

Theorem 7. *Let f be a real function on [a,b] and f′ exists on (a,b). Then f′ is bounded on (a,b) if and only if f satisfies Lipschitz condition on [a,b].*

$f \in Lip([a,b])$ implies its derivative is bounded follows from the definition of Lipschitz function.

Converse part can be proved by indirect method, using the mean value theorem.

From the last example, we say that a continuous function with an unbounded derivatives may be absolutely continuous.

It is easy to show that if f is absolutely continuous (or bounded variation) function then so is $|f|$.

Composition of two absolutely continuous functions need not be absolutely continuous can be observed from the following example.

Example 3. Let $f(x) = \sqrt{x}$, $x \in [0,1]$ and $g(x) = x^2 cos(\frac{\pi}{x})$, $x \in (0,1]$, and $g(0) = 0$.

Obviously, $f \in AC[0,1]$. As g is differentiable and its derivative is bounded on $(0,1)$, $g \in Lip([0,1])$. Thus, $g \in AC[0,1]$ implies $h = |g| \in AC[0,1]$. it is easy to show that f composition h that is $(foh)(x) = x\sqrt{|cos(\frac{\pi}{x})|} \notin BV[0,1]$.

In the above example $g \in Lip([0,1])$ but $g \notin C^1([0,1])$.

Looking to the feature of the class $BV[a,b]$, the notion of bounded variation has been generalized in many ways and many generalized bounded variations are introduced.

Generalized Bounded Variations

A mathematician's desire for more elegant and/or more generality in treating a particular problem leads further to interesting generalizations of the concept of bounded variation in many ways. Consequently, different classes of functions of generalized bounded variations are introduced.

In 1924, Wiener introduced the class $BV^{(p)}([a,b])$, $p \geq 1$, of functions of p-bounded variation over $[a,b]$. The concept of p-bounded variation was sub-sequently generalized by Young [37], in 1937, to the class $\phi - BV([a,b])$ of functions of ϕ-bounded variation over $[a,b]$. Another class that was directly influenced by the study of the convergence problems in the theory of Fourier series, namely $\Lambda - BV([a,b])$ of functions of Λ-bounded variation over $[a,b]$ appeared in 1972 in Waterman's paper. Subsequently in 1980, Shiba [21] introduced the class $\Lambda - BV^{(p)}([a,b])$ of functions of p-Λ-bounded variation over $[a,b]$. The class $\phi - \Lambda - BV([a,b])$ of functions of ϕ-Λ-bounded variation over $[a,b]$ was introduced by Schramm and Waterman [20] in 1982. . In 1990, Kita and Yoneda [16] defined the class $BV(p(n)\uparrow p,[a,b])$ $(1\le p\le\infty)$ of functions of $p(n)$-bounded variation over $[a,b]$. It was generalized to the class $BV(p(n)\uparrow p,\varphi,[a,b])$ by Akhobadze [1] in 2000. Finally in 2011, the class $\Lambda BV(p(n)\uparrow p,\varphi,[a,b])$ of functions of $p(n)$-Λ-bounded variation over $[a,b]$ appeared in [27]. By considering the differences of order $r\ge 2$ the class $r-BV([a,b])$ of functions of bounded rth-variation over $[a,b]$ [34] is one of the important generalizations of the Jordan's class.

While investigating the convergence of Fourier series in the $L^1[0,2\pi]$-norm, in 1996 Moricz and Siddiqi [17] introduced the class $BVM([0,2\pi])$ of functions of bounded variation in the mean. The concept of bounded variation in the mean was subsequently generalized by R. E. Castillo in 2005 and it was to the class $BV^{(p)}M([0,2\pi])$ of functions of p-bounded variation in the mean.

Notations and Definitions

In the sequel $\mathbb{T}=[0,2\pi)$ $\mathbb{L}$ is a class of non-decreasing sequences $\Lambda=\{\lambda_n\}_{n=1}^{\infty}$ of positive numbers such that $\sum_n\frac{1}{\lambda_n}$ diverges; ϕ is an increasing non-negative convex function defined on $[0,\infty)$ such that $\phi(0)=0$, $\frac{\phi(x)}{x}\to 0$ as $x\to 0$ and $\frac{\phi(x)}{x}\to\infty$ as $x\to\infty$; $\varphi(n)$ is a real sequence such that $\varphi(1)\ge 2$ and $\varphi(n)\to\infty$ as $n\to\infty$; and $\mathbb{C}$ is a set of complex numbers.

Definition 1. Given sequence $\Lambda=\{\lambda_n\}_{n=1}^{\infty}\in\mathbb{L}$ and $p\ge 1$, a function f defined on an interval $I:=[a,b]$ is said to be of p-Λ-bounded variation (that is, $f\in\Lambda BV^{(p)}(I)$) if

$$V_{\Lambda_p}(f,I)=\sup_{\{I_i\}}\left\{\left(\sum_i\frac{|f(I_i)|^p}{\lambda_i}\right)^{\frac{1}{p}}\right\}<\infty,$$

where $\{I_i\}=\{[a_i,b_i]\}$ is a finite collection of non-overlapping subintervals in I and $f(I_i)=f(b_i)-f(a_i)$.

In the Definition 1, for $\Lambda=\{1\}$ (that is, $\lambda_n=1$, for all n) and $p=1$ one gets the class $BV(I)$; for $p=1$ one gets the class $\Lambda BV(I)$; for $\Lambda=\{n\}$ (that is, $\lambda_n=n$, for all n); and $p=1$ one gets the class $HBV(I)$ (the class of function of Harmonic bounded variation); and for $\Lambda=\{1\}$ one gets the class $BV^{(p)}(I)$.

For any $\Lambda\in\mathbb{L}$ and $p\ge 1$, we have

$$\left(\sum_i\frac{|f(I_i)|^p}{\lambda_i}\right)^{\frac{1}{p}}\le\left(\frac{1}{\lambda_1}\right)^{\frac{1}{p}}\left(\sum_i|f(I_i)|^p\right)^{\frac{1}{p}}\le\left(\frac{1}{\lambda_1}\right)^{\frac{1}{p}}\sum_i|f(I_i)|.$$

This implies

$$BV(I) \subset BV^{(p)}(I) \subset \Lambda BV^{(p)}(I).$$

It is observed that if $f \in \Lambda BV^{(p)}(I)$ then it is a regulated function over I [23, Theorem 2, p. 92] (that is, f has right-hand and left-hand limits at every point of the intervals $[a,b)$ and $(a,b]$, respectively). If f is a regulated function over I, then $f \in \Lambda BV^{(p)}(I)$ for some sequence $\Lambda \in \mathbb{L}$ [23, Theorem 6, p. 92]. Thus, the union of $\Lambda BV^{(p)}(I)$ functions over all sequences Λ is the class of regulated functions over I.

Moreover, if $f \in \Lambda BV^{(p)}(I)$ for every sequence Λ, then $f \in BV^{(p)}(I)$ [23, Theorem 3, p. 92]. Thus, the intersection of $\Lambda BV^{(p)}(I)$ functions over all sequences Λ is the class $BV^{(p)}(I)$ (that is, $\bigcap_\Lambda \Lambda BV^{(p)}(I) = BV^{(p)}(I)$).

Therefore, we can say that the class $\Lambda BV^{(p)}(I)$ lies between the class of regulated functions over I and the class $BV^{(p)}(I)$.

The following example shows that a continuous function need not be of p-Λ-bounded variation.

Example 1. Given a $p \geq 1$. Let $f : [0,1] \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x^{\frac{1}{p}} \, cos\left(\frac{\pi}{2x}\right), & \text{if } x \in (0,1], \\ 0, & \text{otherwise.} \end{cases}$$

Obviously, $f \in \mathbf{C}([0,1])$. For any $m = 2k$, where $k \in \mathbb{N}$, if we consider the points $x_0 = 0$ and $x_i = \frac{1}{m+1-i}$, for $i = 1, 2, \cdots, m$, then we have $0 = x_0 \leq x_1 \leq \cdots \leq x_m = 1$ and

$$f(x_i) = \begin{cases} 0, & \text{if } i \text{ is even}, \\ \pm(x_i)^{\frac{1}{p}}, & \text{if } i \text{ is odd}. \end{cases}$$

Therefore,

$$\sum_{i=0}^{m-1} |f(x_{i+1}) - f(x_i)|^p = \frac{1}{m} + \frac{1}{m} + \frac{1}{m-2} + \frac{1}{m-2} + \cdots + \frac{1}{2} + \frac{1}{2}$$
$$= \sum_{i=1}^{k} \frac{1}{i} \to \infty \; as \; k \to \infty.$$

Thus, $f \notin BV^{(p)}([0,1])$. Since $\bigcap_\Lambda \Lambda BV^{(p)}([0,1]) = BV^{(p)}([0,1])$, $f \notin \Lambda BV^{(p)}([0,1])$ for at lest one sequence $\Lambda \in \mathbb{L}$.

Since f is regulated function over $[0,1]$, $f \in \Lambda^{'} BV^{(p)}([0,1])$, for some $\Lambda^{'} \in \mathbb{L}$ (in view of the result [23, Theorem 6, p. 92]).

Hence, $BV^{(p)}(I) \subsetneqq \Lambda^{'} BV^{(p)}(I)$.

Definition 2. Given a continuous function ϕ defined on $[0,\infty)$ and strictly increasing from 0 to ∞, a function f defined on an interval I is said to be of ϕ-Λ-bounded variation (that is, $f \in \phi\Lambda BV(I)$) if

$$V_{\Lambda_\phi}(f,I) = \sup_{\{I_i\}} \left\{ \sum_i \frac{\phi(|f(I_i)|)}{\lambda_i} \right\} < \infty,$$

where I, Λ, $\{I_i\}$, and $f(I_i)$ are as defined above in the Definition 1.

Here, function ϕ is said to have property Δ_2 if there is a constant $d \geq 2$ such that $\phi(2x) \leq d\phi(x)$, for all $x \geq 0$.

In the Definition 2, for $\phi(x) = x$ and $\Lambda = \{1\}$ one gets the class $BV(I)$; for $\phi(x) = x$, one gets the class $\Lambda BV(I)$; for $\phi(x) = x^p$, one gets the class $\Lambda BV^{(p)}(I)$; and for $\Lambda = \{1\}$, one gets the class $\phi BV(I)$.

Definition 3. Given sequence $\varphi(n)$ and $1 \leq p(n) \uparrow p$ as $n \to \infty$, where $1 \leq p \leq \infty$, a function f defined on an interval I is said to be of $p(n)$-Λ-bounded variation (that is, $f \in \Lambda BV(p(n) \uparrow p, \varphi, I)$) if

$$V_{\Lambda_{p(n)}}(f,\varphi,I) = \sup_{n\geq 1} \sup_{\{I_i\}} \left\{ \left(\sum_i \frac{|f(I_i)|^{p(n)}}{\lambda_i} \right)^{\frac{1}{p(n)}} : \delta\{I_i\} \geq \frac{b-a}{\varphi(n)} \right\} < \infty,$$

where I, Λ, $\{I_i\}$, and $f(I_i)$ are as defined earlier in the Definition 1, and

$$\delta\{I_i\} = \inf_i \{|a_i - b_i|\}.$$

In the Definition 3, for $\varphi(n) = 2^n$, for all n, and $\Lambda = \{1\}$, one gets the class $BV(p(n) \uparrow p, I)$; for $\Lambda = \{1\}$ one gets the class $BV(p(n) \uparrow p, \varphi, I)$; for $p = \infty$, one gets the class $\Lambda BV(p(n) \uparrow \infty, \varphi, I)$; for $\Lambda = \{1\}$, and $p = \infty$ one gets the class $BV(p(n) \uparrow \infty, \varphi, I)$; and for $p(n) = p$, for all n, one gets the class $\Lambda BV^{(p)}(I)$.

It is observed that [28, Lemma 2.7, p. 226], for $1 \leq p < \infty$,

$$BV^{(p)}(I) \subseteq BV(p(n) \uparrow \infty, \varphi, I)$$

and

$$\bigcup_{1\leq q<p} BV^{(q)}(I) \subseteq BV(p(n) \uparrow p, \varphi, I) \subseteq BV^{(p)}(I).$$

Definition 4. Given a positive integer r, a function f defined on an interval I is said to be of bounded rth-variation (that is, $f \in r - BV(I)$) if

$$V_r(f,I) = \sup_P \left\{ \sum_{i=0}^{m-r} |\Delta^r f(x_i)| \right\} < \infty,$$

where I is as defined earlier in the Definition 1, $P : a = x_0 < x_1 < \cdots < x_m = b$,

$$\Delta f(x_i) = f(x_{i+1}) - f(x_i)$$

and

$$\Delta^k f(x_i) = \Delta^{k-1}(\Delta f(x_i)),\ k \geq 2,$$

so that

$$\Delta^r f(x_i) = \sum_{u=0}^{r} (-1)^u \binom{r}{u} f(x_{i+r-u}).$$

Obviously, $BV(I) \subset r - BV(I) \subset B(I)$, where $B(I)$ is a class of all bounded functions on I.

The following example shows that $BV(I) \neq r - BV(I)$.

Example 2. Consider everywhere continuous but nowhere differentiable function of Weierstrass [11], defined as

$$f(x) = \sum_{n=1}^{\infty} b^{-n} \cos(b^n x),\ \ b \text{ an integer} > 1,$$

satisfies the condition

$$|f(x+h) + f(x-h) - 2f(x)| = O(|h|) \text{ as } h \to 0$$

uniformly in x in $\overline{\mathbb{T}}$ and, therefore, it is of bounded second variation over $\overline{\mathbb{T}}$(that is, $f \in 2 - BV(\overline{\mathbb{T}})$) [38]. However, f being a nowhere differentiable function, it is not of bounded variation over $\overline{\mathbb{T}}$(that is, $f \notin BV(\overline{\mathbb{T}})$).

Definition 5. Given a function $f \in L^p(\overline{\mathbb{T}})$, where $p \geq 1$, the p-integral modulus of continuity of f of higher differences of order $r \geq 1$ is defined as

$$\omega_r^{(p)}(f;\delta) = \sup\left\{\left(\frac{1}{2\pi}\int_{\overline{\mathbb{T}}} |\Delta^r f(x;h)|^p\,dx\right)^{\frac{1}{p}} :\ 0 < h \leq \delta\right\},$$

where

$$\Delta^r f(x;h) = \sum_{u=0}^{r} (-1)^u \binom{r}{u} f(x + (r-u)h).$$

In the Definition 5, for $r = 1$, we omit writing r, one gets $\omega^{(p)}(f;\delta)$, the p-integral modulus of continuity of f.

For $p \geq 1$ and $\alpha \in (0, 1]$, we say that $f \in Lip(p; \alpha)(\overline{\mathbb{T}})$ if

$$\omega^{(p)}(f; \delta) = O(\delta^{\alpha}).$$

In the Definition 5, for $p = \infty$ and $r = 1$, we omit writing p and r, one gets $\omega(f; \delta)$, the modulus of continuity of f, and in that case the class $Lip(p; \alpha)(\overline{\mathbb{T}})$ reduces to the Lipschitz class $Lip(\alpha)(\overline{\mathbb{T}})$.

Some basic notion as bounded variation, Lipschitz function, absolute continuity etc., have been generalized to integral settings in the literature to obtain various classes of functions which are of interest in studying Fourier analysis. Hardy and Littlewood studied functions which satisfy Lipschitz condition in L^pnorm, and obtain results about convergence of their Fourier series. While studying the convergence of a Fourier series in L^1 norm, Moricz and Siddiqi [17] introduce the class bounded variation in mean, which is further generalized to the class p-bounded variation in mean [4].

For a 2π-periodic function, the notion of p-bounded variation in the sense of L^p-norm is defined as follows.

Definition 6. Let $f \in L^p(\overline{\mathbb{T}})$ with $p \geq 1$. We say $f \in BV^{(p)}M(\overline{\mathbb{T}})$(that is, f is a function of p-bounded variation in the mean over $\overline{\mathbb{T}}$) if

$$V_p^m(f, \overline{\mathbb{T}}) = \sup_{\{I_i\}} \left\{ \sum_i \int_{\overline{\mathbb{T}}} \frac{|f(I_{ix})|^p}{|I_{ix}|^{p-1}}\, dx \right\} < \infty,$$

where $\{I_i\} = \{[x_i, x_{i+1}]\}$ is a finite collection of non-overlapping subintervals in $\overline{\mathbb{T}}$, $\{I_{ix}\} = \{[x + x_i, x + x_{i+1}]\}$, $f(I_{ix}) = f(x + x_{i+1}) - f(x + x_i)$ and $|I_{ix}| = |x_{i+1} - x_i|$.

In the Definition 6, for $p = 1$ one gets the class $BVM(\overline{\mathbb{T}})$.

The class of one variable functions of bounded variation as well as the classes of one variable functions of generalized bounded variations are of great interest because of their valuable properties like additivity, differentiability, measurability, integrability, etc. Because of all such properties, functions of these classes owe their important role in the study of Operator theory, Fourier series, Walsh–Fourier series, Fourier–Haar series, Fourier–Jacobi series and other orthogonal series, Stieltjes and other integrals, and the calculus of variations.

Functions of generalized bounded variation play an important role in various aspects of mathematical analysis. Their main influence is in connection with the study of Fourier series. Many of these classes are linear spaces and become Banach spaces when they are equipped with suitable norms involving generalized variations. In many cases, the norms are also sub-multiplicative, and so the function spaces carry an additional structure of Banach algebras with respect to the pointwise operations. In 1976, D. Waterman proved that the class $\Lambda BV([a, b])$ is a Banach space with respect to the pointwise operations and the Λ-variation norm. This result was extended in 2006 by proving that the class $\Lambda BV^{(p)}([a, b])$ is a Banach space with respect to the pointwise operations and the Λ_p-variation norm, [23, Theorem 1, p. 92]. In 2010, Kantrowitz [15, Theorem 1, p. 171] observed that the Λ_p-variation norm is submultiplicative. Thus, the class $\Lambda BV^{(p)}([a,b])$ carries an additional structure of Banach algebra with respect to the pointwise operations. While studying the spectral theory of linear operators on Banach spaces, extending the usual interval definition of a function of bounded variation to new definition of a function of bounded variation on a non-empty compact subset σ of $\mathbb{R}$, Ashton and Doust [2] in 2005 generalized the class $BV([a,b])$ to the class $BV(\sigma)$. They observed that the extended class $BV(\sigma)$ forms a Banach algebra, with respect to the pointwise operations and the variation norm and it has some interesting applications to the Operator theory.

The summary of the important similarities and differences between the class $BV(\sigma)$ and the class $BV(I)$ are listed in the paper [2]. Some of the important properties of the class $\Lambda BV^{(p)}(I)$ are listed in the paper of Vyas [23].

Some basic properties of the class $\Lambda BV^{(p)}(I)$ are as followed.

Theorem 1. *If $f, g \in \Lambda BV^{(p)}(I)$, then the following hold:*

(*i*) *f and g are bounded.*
(*ii*) $V_{\Lambda_p}(f+g, I) \le V_{\Lambda_p}(f, I) + V_{\Lambda_p}(g, I)$.
(*iii*) $V_{\Lambda_p}(\alpha f, I) = |\alpha| V_{\Lambda_p}(f, I)$, *for any* $\alpha \in \mathbb{R}$.
(*iv*) $V_{\Lambda_p}(fg, I) \le \|f\|_\infty V_{\Lambda_p}(g, I) + \|g\|_\infty V_{\Lambda_p}(f, I)$, *where* $\|f\|_\infty$.

Proof of Theorem. Let $P = \{x_i\}_{i=1}^m$ be any partition of I.

For any $x \in I$,

$$|f(x)| \le |f(x) - f(a)| + |f(a)|$$

$$= (\lambda_1)^{\frac{1}{p}} \left(\frac{(|f(x) - f(a)|)^p}{\lambda_1} \right)^{\frac{1}{p}} + |f(a)|$$

$$\le (\lambda_1)^{\frac{1}{p}} V_{\Lambda_p}(f, I) + |f(a)| < \infty$$

Thus, the Theorem (*i*) follows.

Proof of Theorem (*ii*). Since, $f, g \in \Lambda BV^{(p)}(I)$, by Minkowski's inequality we have

$$\left(\sum_i \frac{(|\Delta(f+g)(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}} = \left(\sum_i \frac{(|\Delta f(x_i) + \Delta g(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}}$$

$$\le \left(\sum_i \frac{(|\Delta f(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}} + \left(\sum_i \frac{(|\Delta g(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}}$$

$$\le V_{\Lambda_p}(f, I) + V_{\Lambda_p}(g, I).$$

Thus, $V_{\Lambda_p}(f+g, I) \le V_{\Lambda_p}(f, I) + V_{\Lambda_p}(g, I)$.

Proof of Theorem (iv).

$$\left(\sum_i \frac{(|\Delta fg(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}} = \left(\sum_i \frac{(|g(x_{i+1})\Delta f(x_i)+f(x_i)\Delta g(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}}$$

$$\le \left(\sum_i \frac{(\|g\|_\infty)^p(|\Delta f(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}} + \left(\sum_i \frac{(\|f\|_\infty)^p(|\Delta g(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}}$$

$$= \|g\|_\infty \left(\sum_i \frac{(|\Delta f(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}} + \|f\|_\infty \left(\sum_i \frac{(|\Delta g(x_i)|)^p}{\lambda_i}\right)^{\frac{1}{p}}$$

$$\le \|g\|_\infty \, V_{\Lambda_p}(f,I) + \|f\|_\infty \, V_{\Lambda_p}(g,I).$$

Thus,

$$V_{\Lambda_p}(fg,I) \le \|f\|_\infty \, V_{\Lambda_p}(g,I) + \|g\|_\infty \, V_{\Lambda_p}(f,I).$$

This completes the proof of the theorem.

It is easy to observe that the space $\Lambda BV^{(p)}(I)$ is a normed linear space with respect to the variation norm $\|\cdot\|_{\Lambda_p}$ defined as

$$\|f\|_{\Lambda_p} = \|f\|_\infty + V_{\Lambda_p}(f,I), \quad f \in \Lambda BV^{(p)}(I).$$

Theorem 2. $\left(\Lambda BV^{(p)}(I), \|\cdot\|_{\Lambda_p}\right)$ *is a commutative unital Banach algebra with respect to the pointwise operations.*

Proof of the Theorem. Let $\{f_k\}_{k=1}^{\infty}$ be a Cauchy sequence in $\Lambda BV^{(p)}(I)$. Therefore, it converges uniformly to some function say f on I. For any partition P of I, we get

$$V_{\Lambda_p}(f_k,I,P) \le V_{\Lambda_p}(f_k - f_l,I,P) + V_{\Lambda_p}(f_l,I,P)$$
$$\le V_{\Lambda_p}(f_k - f_l,I) + V_{\Lambda_p}(f_l,I).$$

This implies

$$V_{\Lambda_p}(f_k,I) \le V_{\Lambda_p}(f_k - f_l,I) + V_{\Lambda_p}(f_l,I)$$

and

$$|V_{\Lambda_p}(f_k,I) - V_{\Lambda_p}(f_l,I)| \le V_{\Lambda_p}(f_k - f_l,I) \to 0 \text{ as } k,l \to \infty.$$

Hence, $\{V_{\Lambda_p}(f_k,I)\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$ and it is bounded by some constant say $M > 0$. Therefore,

$$V_{\Lambda_p}(f,I,P) = \lim_{k\to\infty} V_{\Lambda_p}(f_k,I,P)$$
$$\le \lim_{k\to\infty} V_{\Lambda_p}(f_k,I) \le M < \infty.$$

Thus, $f \in \Lambda BV^{(p)}(I)$.

Since $\{f_k\}_{k=1}^{\infty}$ is a Cauchy sequence, for any $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$V_{\Lambda_p}(f_k - f_l, I, P) < \epsilon, \quad \textit{for all } k, l \geq n_0.$$

Letting $l \to \infty$ and taking supremum on both sides of the above inequality, we get $V_{\Lambda_p}(f_k - f, I) < \epsilon, \quad \textit{for all } k \geq n_0.$

Thus, $\|f_k - f\|_{\Lambda_p} \to 0 \ \ as \ \ k \to \infty.$

Hence, $\left(\Lambda BV^{(p)}(I), \|\cdot\|_{\Lambda_p}\right)$ is a Banach space.

For any $g_1, g_2 \in \Lambda BV^{(p)}(I)$,

$$\|g_1 g_2\|_{\Lambda_p} = \|g_1 g_2\|_\infty + V_{\Lambda_p}(g_1 g_2, I)$$

$$\leq \|g_1\|_\infty \|g_2\|_\infty + \|g_1\|_\infty V_{\Lambda_p}(g_2, I) + \|g_2\|_\infty V_{\Lambda_p}(g_1, I) \quad (\textit{by previous Theorem})$$

$$\leq \|g_1\|_\infty(\|g_2\|_\infty + V_{\Lambda_p}(g_2, I)) + \|g_2\|_\infty V_{\Lambda_p}(g_1, I) + V_{\Lambda_p}(g_1, I)\, V_{\Lambda_p}(g_2, I)$$

$$= (\|g_1\|_\infty + V_{\Lambda_p}(g_1, I))\, (\|g_2\|_\infty + V_{\Lambda_p}(g_2, I))$$

$$= \|g_1\|_{\Lambda_p} \|g_2\|_{\Lambda_p}.$$

This completes the proof of the theorem.

Similarly, it is easy to observe that the space $\phi - \Lambda BV(I)$ is a normed linear space with respect to the variation norm $\|\cdot\|_{\phi_\Lambda}$ defined as

$$\|f\|_{\phi_\Lambda} = \|f\|_\infty + V_{\phi_\Lambda}(f, I), \quad f \in \phi - \Lambda BV(I).$$

Like previous theorem, one can prove the following theorems.

Theorem 3. *The class* $\left(\phi - \Lambda BV^{(p)}(I), \|\cdot\|_{\phi_\Lambda}\right)$ *is a commutative unital Banach space with respect to the pointwise operations.*

Theorem 4 ([27]). *The class* $\Lambda BV(p(n) \uparrow p, \varphi, I)$ $(1 \leq p \leq \infty)$ *is a commutative unital Banach algebra with respect to the pointwise operations and the variation norm defined as*

$$||f||_{var} = ||f||_\infty + V_{\Lambda_{p(n)}}(f, \varphi, I), \quad f \in \Lambda BV(p(n), \varphi).$$

We need the following property to prove this result.

Lemma 1. *Let* $1 \leq p \leq \infty$. *Then* $\Lambda BV(p(n) \uparrow p, \varphi) \subseteq B[0, 2\pi]$.

Proof. Let $f \in \Lambda BV(p(n), \varphi)$. Suppose f is not bounded. Then for a sequence $\{M_i\}$ increasing to ∞, there exists a sequence $\{x_i\}$ in $[0, 2\pi]$ such that $|f(x_i)| \geq M_i$ for all i. Since $\{x_i\}$ is bounded, it has a convergent subsequence $\{x_{i_k}\}$ converging to some x_0 in $[0, 2\pi]$ and we may assume without loss of generality that $|x_{i_k} - x_0| < \frac{\pi}{\varphi(1)}$ for all i_k.

Thus, we do have a sequence $\{x_m\}$ in $[0, 2\pi]$ converging to $x_0 \in [0, 2\pi]$ such that $|x_m - x_0| < \frac{\pi}{\varphi(1)}$ for all m and $|f(x_m)| \to \infty$ as $m \to \infty$.

Consider any x_m and put

$$t_0 = x_0 - \frac{2\pi}{\varphi(1)}, \quad t_1 = x_m, \quad t_2 = x_0 + \frac{2\pi}{\varphi(1)}.$$

Then for sufficiently large n,

$$min\,\{t_1 - t_0,\ t_2 - t_1\} \geq \frac{\pi}{\varphi(1)} \geq \frac{2\pi}{\varphi(n)},$$

and hence

$$V_\Lambda(f, p(n), \varphi) \geq \frac{|f(x_m) - f(t_0)|}{\lambda_2} \geq \frac{|f(x_m)| - |f(t_0)|}{\lambda_2}.$$

Since this holds for all x_m, we get $V_\Lambda(f, p(n), \varphi) = \infty$ - a contradiction. Thus the result follows.

It is known that the space $L^1[0, 2\pi]$, the class of all 2π-periodic Lebesgue integrable functions over $[0, 2\pi]$ is not closed with respect to the pointwise product.

Example 3. Let $f[0, 2\pi] \to R$ be defined as

$$f(x) = \begin{cases} \frac{1}{\sqrt{x}}, & \text{if } x \in (0, 1], \\ 0, & \text{otherwise.} \end{cases}$$

Then $f \in L^1[0, 2\pi]$ but $f^2 \notin L^1[0, 2\pi]$.

The problem of finding a binary operation on the integrable function that forms a Banach algebra as well as that would correspond to pointwise multiplications of their Fourier transforms motivated to introduce the convolution product $L^1[0, 2\pi]$ in Fourier analysis, which is defined as follows.

Definition 7. For any f, g $\in L^1[0,2\pi]$ their convolution product, denoted by $f * g$, is defined as

$$(f * g)(x) = \frac{1}{2\pi} \int_0^{2\pi} (f(x - y)g(y))dy, \quad \forall x \in [0, 2\pi].$$

It is easy to show that the convolution product $* : L^1[0, 2\pi] \times L^1[0, 2\pi] \to L^1[0, 2\pi]$ is a continuous bilinear map. Moreover $L^1[0, 2\pi]$ is a commutative complex Banach algebra without identity with respect to the convolution product.

Since $e_n * e_m = 0$, if $m \neq n$, where $e_m = e^{inx}$, $\forall m \in \mathbb{Z}$, $L^1[0, 2\pi]$ is a commutative ring without unity but not integral domain with respect to the convolution product.

It is well known that

$$L^1 * E = \{f * g : f \in L^1, g \in E\} \subset E, \tag{3}$$

for $E = L^p[0, 2\pi]$, or $E = BV([0, 2\pi])$, or $E = AC([0, 2\pi])$, or $E = C^k[0, 2\pi]$.

Thus, the convolution product is inheriting the best smoothness (or variation) property from each parent function. Because of these properties the convolution product is used in digital image processing for filtering an image, for smoothing an edge of image as well as for an edge detection of an image.

It is observed that, the classes of functions of generalized bounded variations share many properties of functions of bounded variation. Therefore it is interesting to know whether these notions of generalized bounded variations are hereditary under the convolution or not.

Theorem 5. *If* $f \in L^1[0, 2\pi]$ *and* $g \in Lip(\alpha, p)$ *over* $[0, 2\pi], 0 \le \alpha \le 1, p \ge 1$, *then* $f * g \in Lip(\alpha, p)$ *over* $[0, 2\pi]$.

The Theorem can be easily proved from the following lemma.

Lemma 2. *If* $f \in L^1[0, 2\pi]$ *and* $g \in L^p[0, 2\pi]$ $(p \ge 1)$, *then*
$\| T_h f * g - f * g \|_{p,[0,2\pi]} \le \| f \|_1 \ \| T_h g - g \|_{p,[0,2\pi]}$.

Proof of the Lemma. For any $h \in L^q[0, 2\pi]$, where q satisfies $\frac{1}{p} + \frac{1}{q} = 1$, from Fubini–Tonelli theorem we get

$$\begin{aligned}
&\left| \int_0^{2\pi} [T_h f * g(x) - f * g(x)] h(x) dx \right| \\
&\le \int_0^{2\pi} | h(x) | \{ \int_0^{2\pi} | f(y) | \ | (T_h g - g)(x - y) | dy \} dx \\
&= \int_0^{2\pi} | f(y) | \{ \int_0^{2\pi} | h(x) | \ | (T_h g - g)(x - y) | dx \} dy \\
&\le \| f \|_1 \ \| h \|_q \ \| T_h g - g \|_p, \text{ from Holder's inequality.}
\end{aligned}$$

Hence the result follows from the converse of the Holder's inequality [7, Exercise 3.6, p. 65].

Theorem 6. *The class* $\Lambda BV(p(n) \uparrow p, \varphi, [0, 2\pi])$ $(1 \le p \le \infty)$ *is a Banach algebra with respect to the pointwise linear operations and the convolution product.*

Proof. We know that the space $\Lambda BV(p(n), \varphi)$ is a Banach space.

Let $\{I_m\}$ be any collection of non-overlapping sequence of subintervals in $[0, 2\pi]$. For any $f, g \in \Lambda BV(p(n), \varphi)$.

$$|(f * g)(I_m)| \le ||f||_\infty \left(\frac{1}{2\pi} \int_0^{2\pi} |g(b_m - y) - f(a_m - y)| dy \right).$$

Since $p(n) \geq 1$ for all n and $\psi(x) = x^{p(n)}$ is convex for all n. In view of Jensen's inequality, one gets

$$|(f * g)(I_m)|^{p(n)} \leq ||f||_\infty^{p(n)} \left(\frac{1}{2\pi}\int_0^{2\pi} |g(b_m - y) - f(a_m - y)|^{p(n)} dy\right).$$

Multiplying both the sides of the above inequality by $\frac{1}{\lambda_m}$ and summing over m, we have

$\sum_m \left(\frac{|(f*g)(I_m)|^{p(n)}}{\lambda_m}\right)$

$$\leq ||f||_\infty^{p(n)} \left(\frac{1}{2\pi}\int_0^{2\pi} \left(\sum_m \frac{|g(b_m - y) - f(a_m - y)|^{p(n)}}{\lambda_m}\right) dy\right).$$

It follows that

$$V_\Lambda(f * g, p(n), \varphi) \leq ||f||_\infty \, V_\Lambda(g, p(n), \varphi),$$

which together with

$$||f * g||_\infty \leq ||f||_\infty \, ||g||_\infty$$

implies

$$||f * g||_{\Lambda_{p(n)}} \leq ||f||_{\Lambda_{p(n)}} \, ||g||_{\Lambda_{p(n)}}.$$

This completes the proof.

Theorem 7. *If $f \in L[0, 2\pi]$ and $g \in BV(p(n) \uparrow p, \varphi, [0, 2\pi])$ $(1 \leq p \leq \infty)$, then f*$g \in BV(p(n) \uparrow p, \varphi, [0, 2\pi])$.*

Proof. If one of f and g is 0, then $f * g$=0 and the result follows.

We may therefore assume $||f||_1 = 1$ without the loss of generality.

Let $\{I_m\}$ be any collection of non-overlapping sequence of subintervals in $[0, 2\pi]$, where $I_m = [a_m, b_m]$. Then by Holder's inequality, one gets

$(|f * g(I_m|)^{p(n)}$

$$\leq \left[\left(\frac{1}{2\pi}\int_0^{2\pi} (|f(y)|^{1/q(n)})(|f(y)|^{1/p(n)} \, |g(b_m - y) - g(a_m - y)|)dy\right)\right]^{p(n)}$$

$$\leq \frac{1}{2\pi}\int_0^{2\pi} |f(y)||g(b_m - y) - g(a_m - y)|^{p(n)} dy,$$

where in $q(n)$ is the index conjugate to $p(n)$ for each n.

Dividing both the sides of the above inequality by λ_m and summing over m, we have

$$\sum_m \frac{|f*g(I_m)|^{p(n)}}{\lambda_k}$$

$$\leq (\frac{1}{2\pi}\int_0^{2\pi} |f(y)|(\sum_m \frac{|g(b_m - y) - g(a_m - y)|^{p(n)}}{\lambda_m})dy\ )$$

$$\leq\ V_\Lambda(g, p(n), \varphi)^{p(n)};$$

and hence

$$V_\Lambda(f * g, p(n), \varphi) \leq\ V_\Lambda(g, p(n), \varphi).$$

This proves the theorem.

Similarly, we can prove the following theorems [24].

Theorem 8. *If* $f \in L^1[0, 2\pi]$ *and* $g \in \bigwedge BV[0, 2\pi]$*, then* $f * g \in \bigwedge BV[0, 2\pi]$.

Theorem 9. *If* $f \in L^1[0, 2\pi]$ *and* $g \in \bigwedge BV^{(p)}[0, 2\pi], p \geq 1$*, then* $f * g \in \bigwedge BV^{(p)}[0, 2\pi]$.

Remark: Since L^1 is a ring with respect to convolution as a ring product. From Theorem 9 the class $\bigwedge BV^{(p)}$ can be regarded as a module over the ring L^1.

Theorem 10. *If* $f \in L^1[0, 2\pi]$ *and* $g \in$ *r-BV[0,2π], then* $f * g \in$ *r-BV[0,2π].*

Open Factorization Problems

It is known that [7, Edward, p. 127]

$$C * AC = C^1,\ \ L^1 * AC = AC \ \ \ and \ \ \ L^1 * BV([0, 2\pi]) = AC([0, 2\pi]).$$

So the obvious question is: Can we say that for different class of generalized bounded variations say E

$$L^1 * E = E \ \ \ or \ \ \ L^1 * E = AC?$$

Interrelations Between Functions of Generalized Bounded Variations

Given an arbitrary nondecreasing continuous function ϑ defined on $[0, 2\pi]$, $\vartheta(0) = 0$ and ϑ subadditive, define

$$H^{\vartheta} = \{f \in C : \ \omega(f, \delta) = O(\vartheta(\delta)) \ \ as\ \delta \to 0\}.$$

Note that if $\vartheta(\delta) = \delta^\alpha,\ 0 < \alpha \leq 1$ then this class is $Lip(\alpha)$ class.

Let $h = \{h_n\}$ be a concave sequence of positive numbers converges to 0. The Chanturiya class $V[h]$ is defined to be the class of all functions

$$f : [0, 2\pi] \to \mathbb{R} \;\; such\ that \;\; \upsilon(n,f) = sup \sum_{i=k}^{n} |f(I_k)|,$$

the supremum is taken over family $\{I_k\}_{k=1}^{n}$ of non-overlapping subintervals of $[0, 2\pi]$.

Using the simple averaging arguments, the following inclusion relations between generalized bounded variations are obtained.

Theorem 1 ([10, Theorem 1]). *$H^{\omega} \subset BV(p(n) \uparrow \infty)$ if and only if*

$$\omega(t) = O\left(t^{1/p([log_2 1/t])}\right) \quad as \quad t \to 0+.$$

Theorem 2 ([10, Theorem 2]). *$V[\upsilon(n)] \subset BV(p(n) \uparrow \infty)$ if and only if*

$$\overline{lim}_{n \to \infty} \left(\sum_{k=1}^{2^n} (\upsilon(k) - \upsilon(k-1))^{p(n)}\right)^{1/p(n)} < \infty.$$

Theorem 3 ([12, Theorem 3.1]). *If $f \in \Lambda BV^{(p)}$, then*

$$\upsilon(n,f) \le n\, V[\frac{V(f)}{(\sum_{i=1}^{n} \frac{1}{\lambda_i})^{1/p}}].$$

Theorem 4 ([3, Theorem 4.1]). $\phi\Lambda BV \subset V[n\phi^{-1}(\frac{1}{\sum_{i=1}^{n}\frac{1}{\lambda_i}})]$

Relation between $\Lambda BV^{(p)}$ and generalized Wiener class is as follows.

Theorem 5 ([13]). *The inclusion $\Lambda BV^{(p)} \subset BV(p(n) \uparrow p)$ holds if and only if*

$$\limsup_{n \to \infty} \left\{ \max_{1 \le k \le 2^n} \frac{k^{\frac{1}{p(n)}}}{\sum_{i=1}^{k}(\frac{1}{\lambda_i})} \right\} < \infty.$$

Fourier Coefficients Properties of Functions of Generalized Bounded Variations

These classes of functions of generalized bounded variation possess many interesting properties in Fourier analysis. In the last couple of decades many mathematicians like Waterman, Schramm, Chanturiya, Miriz [18], Patadia [19, 34] and many others have studied different Fourier coefficients properties like estimation of the order of magnitude Fourier coefficients, conditions for the convergence of Fourier series, conditions for the absolute convergence of Fourier series, conditions for the different type of summabilities of Fourier series, conditions for the different type of absolute summabilities of Fourier series, etc. of functions of these generalized bounded variations are studied in detail.

For any complex valued, 2π-periodic, function $f \in L^1(\overline{\mathbb{T}})$, where $\mathbb{T} := [-\pi, \pi)$ is the one-dimensional torus, its Fourier series is defined as

$$f(x) \sim \sum_{m\in\mathbb{Z}} \hat{f}(m)e^{imx}, \qquad x \in \overline{\mathbb{T}},$$

where

$$\hat{f}(m) = \left(\frac{1}{2\pi}\right) \int_{\overline{\mathbb{T}}} f(x)e^{-imx}\, dx$$

denote the mth Fourier coefficient of f.

Fourier series of a function f is said to be β $(0 < \beta \leq 2)$ absolutely convergent if

$$\sum_{m\in\mathbb{Z}} |\hat{f}(m)|^{\beta} < \infty.$$

For $\beta = 1$, one gets the absolute convergence of a Fourier series.

Some of the interesting properties Fourier coefficients properties of functions of generalized bounded variations are listed below.

Theorem 1 ([27]). *If $f \in \Lambda BV(p(n) \uparrow \infty, \varphi)$, then*

$$\hat{f}(m) = O(1/(\sum_{i=1}^{|m|} \frac{1}{\lambda_i})^{1/p(\tau(m))}),$$

where

$$\tau(m) = min\{k : k \in \mathbb{N}, \varphi(k) \geq m\},\ m \geq 1. \tag{4}$$

Theorem 2 ([20]). *If $f \in \phi\Lambda BV^{(p)}[0, 2\pi]$, then*

$$\hat{f}(m) = O\left(1/(\sum_{i=1}^{|m|} \frac{1}{\lambda_i})^{1/p}\right).$$

Theorem 3 ([20]). *If $f \in \Lambda BV^{(p)}[0, 2\pi]$, then*

$$\hat{f}(m) = O\left(\phi^{-1}\left(1/(\sum_{i=1}^{|m|} \frac{1}{\lambda_i})\right)\right).$$

Theorem 4 ([25]). *Let $f \in \Lambda BV^{(p)}([0, 2\pi])$. If $1 \le p < 2r,\ 1 < r < \infty$ and*

$$\sum_{m=1}^{\infty} \left(\frac{(\omega^{((2-p)s+p)}(f; \frac{\pi}{m}))^{2-\frac{p}{r}}}{m(\sum_{i=1}^{m} \frac{1}{\lambda_i})^{\frac{1}{r}}} \right)^{\frac{\beta}{2}} < \infty,$$

then the Fourier series of f is β absolutely convergent.

Theorem 5 ([25]). *Let $f \in \phi\Lambda BV([0, 2\pi])$. If $1 \le p < 2r,\ 1 < r < \infty$ and*

$$\sum_{m=1}^{\infty} \left(\frac{\left(\phi^{-1}\left(\frac{(\omega^{((2-p)s+p)}(f; \frac{\pi}{m}))^{2r-p}}{(\sum_{i=1}^{m} \frac{1}{\lambda_i})}\right)\right)^{1/r}}{m} \right)^{\frac{\beta}{2}} < \infty,$$

then the Fourier series of f is β absolutely convergent.

Convergence and uniform convergence of Fourier series of functions generalized bounded variations are studied in detail by Salem, Chanturiya, Watermans, and many others. The concept of Λ-BV (and harmonic variation, in particular) has originated from Goffman's and Waterman's investigation of the conditions under which a Fourier series converges everywhere for every change variable. Waterman proved the following result in this direction.

Theorem 6 ([36]). *If $f \in HBV[0, 2\pi]$, then the partial sums of its Fourier series are uniformly bounded. The series converges everywhere and converges uniformly on closed intervals of points of continuity. If $\Lambda BV \supsetneq HBV$, then there is a continuous $f \in \Lambda BV$ whose Fourier series diverges at a point.*

Theorem 7 ([3, Corollary 3, p. 231]). *If a continuous function $f \in \phi\Lambda BV[0, 2\pi]$ and $\sum_{n=1}^{\infty} \left(\frac{1}{n}\phi^{-1}\left(\frac{1}{\sum_{k=1}^{n} \frac{1}{\lambda_k}}\right)\right) < \infty$, then the Fourier series of f converges uniformly.*

Corollary 7.1. *If a continuous function $f \in \Lambda BV^{(p)}[0, 2\pi]$ and*

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} \left(\frac{1}{\sum_{k=1}^{n} \frac{1}{\lambda_k}} \right)^{1/p} \right) < \infty,$$

then the Fourier series of f converges uniformly.

Chanturiya proved the most general result, in this direction, as follows.

Theorem 8 ([5]). *If $f \in C[0, 2\pi]$, then*

$$||f - S_n(f)||_\infty \leq C \min_{1 \leq m \leq [\frac{n-1}{2}]} \left(\omega(f, \frac{1}{n}) \sum_{k=1}^{m} \frac{1}{k} + \sum_{k=m+1}^{[(n-1)/2]} \frac{\upsilon(f, k)}{k^2} \right), \quad n \geq 3.$$

Similarly type of results are also studied for Walsh–Fourier series, Fourier–Haar series, and many other such orthogonal Fourier series in detail.

These notions of generalized bounded variations are extended for higher dimension spaces as well as for functions of several variables. Recently, some of the interesting properties of N-variables functions of such generalized classes are studied in function analysis [2, 6, 8, 9, 26, 29], approximation theory, harmonic analysis [22, 30–33, 35, 39], digital image processing, and many other branches of mathematics.

References

1. Akhobadze, T.: Functions of generalized Wiener classes $BV(p(n) \uparrow \infty, \varphi)$ and their Fourier coefficients. Georgian Math. J. **7**(3), 401–416 (2000)
2. Ashton, B., Doust, I.: Functions of bounded variation on compact subsets of the plane. Stud. Math. **169**(2), 163–188 (2005)
3. Avdispahić, M.: Concepts of generalized bounded variation and the theory of Fourier series. Int. J. Math. Math. Sci. **9**(2), 223–244 (1986)
4. Castillo, R.E.: Bounded $p-$variation in the mean. East-West J. Math. **7**(1), 61–67 (2005)
5. Chanturiya, Z.A.: On uniform convergence of Fourier series Math. Sb. **100**, 534–554 (1976); Math. USSR Sb. **29**, 475–495 (1976)
6. Darji, K.N., Vyas, R.G.: A note on Walsh−Fourier coefficients. Bull. Math. Anal. Appl. **4**(2), 116–119 (2012). MR2955925 (2013:42C10)
7. Edwards, R.E.: Fourier Series- A Mordan Introduction Volume I, 2nd edn. Springer, New York (1979)
8. Fülöp, V., Móricz, F.: Order of magnitude of multiple Fourier coefficients of functions of bounded variation. Acta Math. Hung. **104**(1–2), 95–104 (2004)
9. Ghorpade, S.R., Limaye, B.V.: A Course in Multivariable Calculus and Analysis. Springer, Berlin (2010)
10. Goginava, U.: Relations between some classes of functions. Scientiae Math. Jpn. **4**, 105–114 (2001)
11. Hardy, G.H.: Weierstrass's non-differentiable function. Trans. Am. Math. Soc. **17**(3), 301–325 (1916)
12. Hormozi, M., Ledaric, A.A., Prus-Wisniowski, F.: On $p - \Lambda-$bounded variation. Bell. Iranian Math. Soc. **37**(4), 35–49 (2011)
13. Hormozi, M., Prus-Wisniowski, F., Rosengren, H.: Inclusions of Waterman-Shibe spaces into Kita-Yoneda classes. J. Math. Anal. Appl. **419**, 428–432 (2014)
14. Jordan, C.: Sur la série de Fourier. C. R. Acad. Sci. Paris **92**, 228–230 (1881)
15. Kantrowitz, R.: Submultiplicativity of norms for spaces of generalized $BV-$functions. Real Anal. Exch. **36**(1), 169–176 (2010)
16. Kita, H., Yoneda, K.: A generalization of bounded variation. Acta Math. Hung. **56**(3–4), 229–238 (1990)

17. Móricz, F., Siddiqi, A.H.: A quantified version of the Dirichlet-Jordan test in L^1−norm. Rend. Circ. Mat. Palermo **45**(1) , 19–24 (1996)
18. Móricz, F., Veres, A.: Absolute convergence of multiple Fourier series revisited. Anal. Math. **34**(2), 145–162 (2008)
19. Patadia, J.R., Vyas, R.G.: Fourier series with small gaps and functions of generalized variations. J. Math. Anal. Appl. **182**(1), 113–126 (1994)
20. Schramm, M., Waterman, D.: On the magnitude of Fourier coefficients. Proc. Am. Math. Soc. **85**(3), 407–410 (1982)
21. Shiba, M. , On absolute convergence of Fourier series of functions of class $\Lambda BV^{(p)}$. Sci. Rep. Fukushima Univ. **30**, 7–10 (1980)
22. Vyas, R.G.: A note on Fourier coefficients of functions of generalized Wiener class. Georgian Math. J. **12**(3), 535–538 (2005)
23. Vyas, R.G.: Properties of functions of generalized bounded variation. Mat. Vesnik **58**(3–4), 91–96 (2006)
24. Vyas, R.G.: On the convolution of functions of generalized bounded variations. Georgian Math. J. **13**(1), 193–197 (2006)
25. Vyas, R.G.: On the absolute convergence of Fourier series of functions of $\Lambda BV^{(p)}$ and $\varphi\Lambda BV$. Georgian Math. J. **14**(4), 769–774 (2007)
26. Vyas, R.G.: On variation topology. Mat. Vesnik **62**(1), 47–50 (2010)
27. Vyas, R.G.: A note on functions of $p(n)-\Lambda$−bounded variation. J. Indian Math. Soc. **78**(1–4), 215–220 (2011)
28. Vyas, R.G.: Small gaps Fourier series and generalized variations. Adv. Pure Appl. Math. **3**(2), 223–230 (2012)
29. Vyas, R.G.: Convolution functions of several variables with generalized bounded variation. Anal. Math. **39**(2), 153–161 (2013)
30. Vyas, R.G., Darji, K.N.: On Banach algebra valued functions of bounded generalized variation of one and several variables. Bull. Math. Anal. Appl. **4**(1), 181–189 (2012). MR2955887 (2013:47B40)
31. Vyas, R.G., Darji, K.N.: On multiple Walsh Fourier coefficients. J. Indian Math. Soc. **79**(1–4), 219–228 (2012). MR2976017 (2013:42C10). Zbl 1267.42033
32. Vyas, R.G., Darji, K.N.: On absolute convergence of multiple Fourier series. Math. Notes **94**(1), 71–81 (2013); Russian Translation Mat. Zametki **94**(1) , 81–93 (2013). Zbl 1278.42009
33. Vyas, R.G., Darji, K.N.: Order of magnitude of multiple Fourier coefficients. Anal. Theory Appl. **29**(1), 27–36 (2013) . MR3109878 (2013:42B05)
34. Vyas, R.G., Patadia, J.R.: On the absolute convergence of Fourier series of functions of generalized bounded variations. J. Indian Math. Soc. **62**(1–4), 129–136 (1996)
35. Vyas, R.G., Darji, K.N.: On multiple Fourier coefficients of functions of $\phi - \Lambda$−bounded variation, Math. Inequal. Appl. **17**(3), 1153–1160 (2014)
36. Waterman, D.: On convergence of Fourier series of functions of generalized bounded variation. Studia Math. **44**, 107–117 (1972)
37. Young, L.C.: Sur une generalization de la notion de variation de puissance p-ieme boranee au sense de M. Wiener, et sur la convergence de series de Fourier. C. R. Acad. Sci. Paris **204**, 470–472 (1937)
38. Zygmund, A.: Smooth functions. Duke Math. J. **12**, 47–76 (1945)
39. Zygmund, A.: Trigonometric Series, vols. I and II. Cambridge University Press, Cambridge (1977)