ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

UNIVERSITY OF IOANNINA

DEPARTMENT OF MATHEMATICS

DEDICATED TO THE MEMORY OF
PROFESSOR M. K. GRAMMATIKOPOULOS

Number 16 Volume June-December 2006 & June 2007

Εις μνήμην

Ο Καθηγητής **Μύρων Κ. Γραμματικόπουλος** γεννήθηκε στη Γκάγκρα, Γεωργίας το 1935 και ήλθε στην Ελλάδα το 1966. Απέκτησε το Πτυχίο των Μαθηματικών το 1961 από το Πανεπιστήμιο του Καζακστάν, το Διδακτορικό Δίπλωμα το 1975 και το τίτλο του Υφηγητή το 1981 από το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων καθώς και το τιμητικό δίπλωμα του Επίτιμου Διδάκτορος το 1995 από το Πανεπιστήμιο Ρούσε, Βουλγαρίας.

Υπηρέτησε στο Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων επί μία τριακονταετία, 1972-2002, ως Βοηθός 1972-75, ως Επιμελητής 1975-81, ως Επίκουρος Καθηγητής 1981-85, ως Αναπληρωτής Καθηγητής 1985-86, και ως Καθηγητής 1986-2002, ενώ ενδιάμεσα το 1978-79 παρέμεινε ως Visiting Researcher στο Mississippi State University, το 1984-85 ως Visiting Assistant Professor στο University of Rhode Island και το 1991-92 και 1995-96 ως Visiting Professor στο Technical University of Rousse. Επίσης υπηρέτησε ως Αναπληρωτής Πρόεδρος του Τμήματος Μαθηματικών τη διετία 1989-91 και ως Διευθυντής του Τομέα Μαθηματικής Ανάλυσης την τριετία 1992-95.

Όσον αφορά στις ερευνητικές του δραστηριότητες, που διήρκεσαν μέχρι την τελευταία ημέρα της επίγειας ζωής του, ο **Μύρων Κ. Γραμματικόπουλος,** παρήγαγε πολύ ενδιαφέροντα συμπεράσματα στη Θεωρία Ταλάντωσης των Συνήθων και Συναρτησιακών Διαφορικών Εξισώσεων. Εκπόνησε μόνος ή σε συνεργασία με άλλους ερευνητές περισσότερες από 75 εργασίες που έχουν δημοσιευθεί σε γνωστά διεθνή περιοδικά και οι οποίες συχνά αναφέρονται από άλλους ερευνητές ανά τον κόσμο.

Ο εκλιπών ήταν μέλος της Συντακτικής Επιτροπής του Technical Report το διάστημα 1978-2002 και, ως ελάχιστος φόρος τιμής, ο Τόμος αυτός αφιερώνεται στη μνήμη του.

Αξέχαστε συνάδελφέ μας, ***αιωνία σου η μνήμη!***

In Memoriam

Professor **Myron K. Grammatikopoulos** was born in 1935 in Gagra, Georgia and moved to Greece in 1966. He received his Diploma in Mathematics in 1961 from the University of Kazakhstan and was awarded a Ph. D. in 1975 and a Docent Degree in 1981 from the University of Ioannina and a Doctor Honoris Causa in 1995 from the University of Rousse.

He served in the Department of Mathematics, University of Ioannina, from 1972-2002, as a Teaching Assistant 1972-75, as a Lecturer 1975-81, as an Assistant Professor 1981-85, as an Associate Professor 1985-86 and as a Professor 1986-2002, meanwhile he was appointed in 1978-79 as a Visiting Researcher, Mississippi State University, in 1984-85 as a Visiting Assistant Professor, University of Rhode Island, and as a Visiting Professor, Center of Mathematics, Technical University of Rousse, 1991-92 and 1995-96. He also served as a Deputy Chairman, Department of Mathematics, University of Ioannina, 1989-91 and as a Director, Section of Mathematical Analysis, 1992-95.

During the years of his scientific activity, which lasted until the last day of his earthly life, **Myron K. Grammatikopoulos,** obtained very interesting results on the Oscillation Theory of Ordinary and Functional Differential Equations. He authored and/or co-authored more than 75 papers published in well-known international journals, which are frequently cited by other researchers around the world.

He was a member of the Editorial Board of the Technical Report from 1978-2002 and this Volume is dedicated to his memory. ***May his memory be eternal!***

A queueing model with start-up/close-down times and retrial customers

Ioannis Dimitriou and Christos Langaris
Department of Mathematics, University of Ioannina,
45110 Ioannina, Greece.
E-mail: clagar@cc.uoi.gr

April 24, 2007

Abstract

We consider a queueing system with Poisson arrivals and arbitrarily distributed service times, vacation times, start up and close down times. The model accepts two types of customers, the ordinary and the retrial customers and the server takes a single vacation each time he becomes free. For such a model the stability conditions are investigated and the system state probabilities are obtained both in a transient and in a steady state and used to derive some important measures of the system performance.

Keywords: Poisson arrivals, Start up times, close down times, vacation, retrial queue, general services.

1 Introduction

Queueing models with vacation periods and start up/close down times have been proved very useful to model telecommunication systems and many other queueing situations containing "mechanical parts" that need a "preparation" (start up) before use and a switch off-maintenance period after use (computer systems, manufacturing systems e.t.c.). Applications of such kind of queueing models in SVC - based virtual LAN- emulation and in IP over ATM networks have been described in details in Sakai et al. [1], Niu & Takahashi [2] and in the references therein.

In all models, described and investigated above, the arriving "customers" are queued up and wait to be served. On the other hand, it is easy to realize that such kind of real situations accept, in many cases, and a second kind of "customers" that do not wait in a queue but instead, if they find upon arrival the server unavailable, they depart from the system and repeat their arrival later until succeed to be served. As a simple example of such a situation one can consider an X-ray unit or a tomographic unit, where the machine needs a special time to start working and to close down where there are no more patients waiting, while external phone calls of patients that ask for the results of their examinations or ask for medical advice, arrive and engage the "server ". Queueing systems with retrial customers are widely used in the literature to model telephone switching systems, telecommunication systems and computer networks. For a complete survey on past papers on such kind of models see Falin & Templeton [3], Kulkarni & Liang [4] and Artalejo [5].

In this paper, the two important features, i.e., the start up/close down feature and the retrial feature are combined together, for the first time in the literature. Thus here we study for the first time a queueing model of vacation-start up/close down nature accepting two types of customers, the ordinary customers that are queued up and wait for service and the retrial customers. The server needs a start up time before starts working on customers (different start up for each type of customers), a close down period upon finishing the job, and when he is free he departs for a single vacation. Moreover the ordinary customers have a kind of priority upon retrial customers, in the sense that the arrival of an ordinary customer interrupts the start up time of the retrial customer (if any), and the server starts to be prepared to serve the ordinary customer.

The article is organized as follows. A full description of the model and some, very useful for the analysis, preliminary results are given in section 2 and 3 respectively. The time dependent analysis of the system state probabilities is performed in section 4, while in section 5 the conditions for statistical equilibrium are investigated. Finally, the generating functions of the steady state probabilities are obtained in section 6 and used to obtain, in section 7, some important measures of the system performance.

2 The Model

Consider a single server queue accepting two types of customers. The P_1 customers (ordinary customers) arrive according to a *Poisson* distribution parameter λ_1 and queued up in an ordinary queue waiting to be served. The P_2 customers (retrial customers) arrive according to a *Poisson* distribution parameter λ_2 and, if find the server unavailable, they leave the system and join a retrial box from where they retry independently, after an exponential time parameter α, to find a position for service.

To start serving the P_1customers waiting in the queue or the P_2 customer who found a position for service, the server needs a start up period $S_i, \;\; i = 1,2$ (different for each type of customers), distributed according to a general distribution with distribution function (D.F.) $S_i(x)$, probability density function (p.d.f.) $s_i(x)$ and finite mean value $\bar{s}_i, \;\; i = 1,2$ respectively. Moreover the server, upon finishing all tasks in the queue and in the service area (the retrial box is not necessarily empty at this point), operates a close down period C arbitrarily distributed with D.F. $C(x)$, p.d.f. $c(x)$ and finite mean $\bar{c}$. During the close down period no retrial customer can access the service facility, while if a P_1 customer arrives during C, the server returns to the serving mode and starts serving P_1 customers but now with a different start up period S_3 with

D.F. $S_3(x)$, p.d.f. $s_3(x)$ and finite mean $\bar{s}_3$. This can be explained by the fact that after an incomplete close down period it is natural for the server to need a different start up time to transfer again the system in the serving mode.

When a close down period is successfully completed the server departs for a single vacation V which length is arbitrarily distributed with D.F. $V(x)$, p.d.f. $v(x)$, and finite mean $\bar{v}$. If, in the end of the vacation, there are P_1 customers waiting in the queue the server operates an S_1 period etc., while if the queue is empty he remains idle waiting for the first customer, from outside or from the retrial box, to start working again.

It is natural for the ordinary P_1 customers to have a kind of priority over the retrial P_2 customers. Thus, if a P_1 customer arrives during the start up time of a P_2 customer then this start up period is interrupted and a S_1 time followed by a busy period of P_1 customers and a close down period begins. The interrupted P_2 customer does not return to the retrial box but he restarts his start up time from the beginning when this close down period of P_1 customers is finished. On the other hand the arrival of a P_1 customer cannot interrupt the service time of a P_2 customer. In the later case, the service of the P_2 customer is completed and in the sequel the server starts working (start up plus busy period) on the P_1 customers.

Finally, the service times of both type of customers are arbitrarily distributed with D.F. $B_i(x)$, p.d.f. $b_i(x)$ and finite mean value $\bar{b}_i$ for $i = 1, 2$ respectively, while all random variables defined above are assumed to be independent.

3 Preliminary Results

We agree from here on to denote in general by $a^*(s)$ the Laplace Transform (L.T.) of any function $a(t)$. Let us denote now by $\mathcal{B}^{(i)}$ the duration of a busy period of P_1 customers which starts with $i = 1, 2, \ldots$ P_1 customers, and let $\mathcal{N}(\mathcal{B}^{(i)})$ be the number of P_2 customers arrive during $\mathcal{B}^{(i)}$. Define

$$g_m^{(i)}(t)dt = \Pr[t < \mathcal{B}^{(i)} \leq t + dt,\ \mathcal{N}(\mathcal{B}^{(i)}) = m].$$

Then it is known from Langaris & Katsaros [6] that

$$g^{*(i)}(s, z_2) = \int_0^\infty e^{-st} \sum_{m=0}^{\infty} g_m^{(i)}(t) z_2^m dt = x^i(s, z_2),$$

where $x(s, z_2)$ is the root in z_1 with the smallest absolute value of the equation

$$z_1 - b_1^*(s + \lambda_1(1 - z_1) + \lambda_2(1 - z_2)) = 0.$$

Let now denote by R the time interval from the beginning of a close down period until this period be successfully completed and let $N(R)$ be the number of P_2 customers arriving during R. If we define

$$r_j(t)dt = \Pr[t < R \leq t + dt,\ N(R) = j], \qquad r^*(s, z_2) = \int_0^\infty e^{-st} \sum_{j=0}^{\infty} r_j(t) z_2^j dt,$$

and denote

$$p_{im}(t) = e^{-\lambda_1 t}\frac{(\lambda_1 t)^i}{i!}e^{-\lambda_2 t}\frac{(\lambda_2 t)^m}{m!}, \qquad a(s, z_1, z_2) = s + \lambda_1(1 - z_1) + \lambda_2(1 - z_2),$$

then it is clear that

$$\begin{aligned} r_j(t) = \ & p_{0j}(t)c(t) + \{\lambda_1 \textstyle\sum_{k=0}^{j} p_{0k}(t)[1 - C(t)]\} \\ & *\{\textstyle\sum_{i=0}^{\infty} \sum_{m=0}^{j-k} \ p_{im}(t) \ s_3(t)\} * \textstyle\sum_{l=0}^{j-k-m} \ g_l^{(i+1)}(t) * r_{j-k-m-l}(t), \end{aligned} \tag{1}$$

and so finally

$$r^*(s, z_2) = \frac{c^*(a(s, 0, z_2))}{1 - M(s, x(s, z_2), z_2)},$$

with

$$M(s, z_1, z_2) = \lambda_1 z_1 s_3^*(a(s, z_1, z_2)\ )\frac{1 - c^*(a(s, 0, z_2))}{a(s, 0, z_2)}. \tag{2}$$

Let now Q be the time interval from the beginning of a vacation period until the epoch at which the server becomes idle, and let $N(Q)$ the number of P_2 customers arrive during Q. If we define

$$q_j(t)dt = \Pr[t < Q \leq t + dt,\ N(Q) = j], \qquad q^*(s, z_2) = \int_0^\infty e^{-st}\sum_{j=0}^{\infty} q_j(t) z_2^j dt,$$

then

$$\begin{aligned} q_j(t) = \ & p_{0j}(t)v(t) + \textstyle\sum_{i=1}^{\infty} \sum_{k=0}^{j} p_{ik}(t)v(t) * \sum_{l=0}^{\infty} \sum_{m=0}^{j-k} p_{lm}(t)s_1(t) \\ & * \textstyle\sum_{r=0}^{j-k-m} \ g_r^{(i+l)}(t) * \sum_{n=0}^{j-k-m-r} \ r_n(t) * q_{j-n-k-m-r}(t), \end{aligned} \tag{3}$$

and so

$$q^*(s, z_2) = \frac{v^*(a(s, 0, z_2))}{1 - \ [v^*(a(s,\ z_1, z_2)) - v^*(a(s, 0, z_2)]s_1^*(a(s,\ z_1, z_2))r^*(s, z_2)}. \tag{4}$$

Note here that if we denote

$$e(s, z_1, z_2) = \frac{c^*(a(s,0,z_2))v^*(a(s,0,z_2))}{1 - M(s,z_1,z_2) - c^*(a(s,0,z_2))[v^*(a(s,\ z_1,z_2)) - v^*(a(s,0,z_2)]s_1^*(a(s,\ z_1,z_2))}, \tag{5}$$

then

$$q^*(s, z_2) = \frac{e(s, x(s, z_2), z_2)}{r^*(s, z_2)}. \tag{6}$$

If finally $\rho_v = E(N(Q))$ then by differentiating (4) with respect to z_2 we arrive at $\rho_v = \bar{\rho}_v/(1 - \lambda_1 \bar{b}_1)$ where

$$\begin{aligned} \bar{\rho}_v \quad = \quad & \frac{\lambda_2}{\lambda_1 v^*(\lambda_1)c^*(\lambda_1)}\{(1 + \lambda_1\bar{s}_3)(1 - c^*(\lambda_1))(1 - v^*(\lambda_1)) \\ & + \lambda_1 c^*(\lambda_1)(\bar{v} + \bar{s}_1(1 - v^*(\lambda_1))\}. \end{aligned}$$

We are now ready to define the " service completion time" of a P_2 customer as the time $\bar{W}_2$ elapsed from the epoch at which this customer succeed to find a position for service until the time the server departs for a vacation. Let $N(\bar{W}_2)$ the number of new P_2 customers that arrive during $\bar{W}_2$ and

$$\bar{w}_j(t)dt = \Pr[t < \bar{W}_2 \le t+dt,\ N(\bar{W}_2) = j], \qquad \bar{w}^*(s, z_2) = \int_0^\infty e^{-st} \sum_{j=0}^{\infty} \bar{w}_j(t) z_2^j dt,$$

Then by writing for $\bar{w}_j(t)$ a similar expression as in (1) and (3) and taking Laplace transforms we arrive easily at

$$\bar{w}^*(s, z_2) = \frac{L(s, x(s, z_2), z_2) r^*(s, z_2)}{1 - K(s, x(s, z_2), z_2) r^*(s, z_2)}, \tag{7}$$

where

$$\begin{aligned} L(s, z_1, z_2) = \quad & s_2^*(a(s, 0, z_2))\{b_2^*(a(s, 0, z_2)) + s_1^*(a(s, z_1, z_2))[b_2^*(a(s, z_1, z_2)) \\ & -b_2^*(a(s, 0, z_2))]\}, \\ K(s, z_1, z_2) = \quad & \lambda_1 z_1 s_1^*(a(s, z_1, z_2)) \tfrac{1 - s_2^*(a(s,0,z_2))}{a(s,0,z_2)}. \end{aligned} \tag{8}$$

If finally $\rho_c = E(N(\bar{W}_2))$ then by differentiating (7) with respect to z_2 we arrive at $\rho_c = \bar{\rho}_c/(1 - \lambda_1 \bar{b}_1)$

$$\begin{aligned} \bar{\rho}_c \quad = \quad & \frac{\lambda_2}{\lambda_1 s_2^*(\lambda_1) c^*(\lambda_1)} \{(1 + \lambda_1 \bar{s}_3)(1 - c^*(\lambda_1)) \\ & + c^*(\lambda_1)[(1 + \lambda_1 \bar{s}_1)(1 - s_2^*(\lambda_1)) + \lambda_1 s_2^*(\lambda_1)(\bar{b}_2 + \bar{s}_1(1 - b_2^*(\lambda_1))]\}. \end{aligned}$$

The "generalized service completion time " of a P_2 customer, W_2 say, can be defined as the time elapsed from the epoch at which this customer succeed to find a position for service until the time the server is again idle and so free to accept the next customer (from outside or from the retrial box). Let $N(W_2)$ the number of new P_2 customers that arrive during W_2 and

$$w_j(t)dt = \Pr[t < W_2 \le t+dt,\ N(W_2) = j], \qquad w^*(s, z_2) = \int_0^\infty e^{-st} \sum_{j=0}^{\infty} w_j(t) z_2^j dt,$$

then it is clear that

$$w^*(s, z_2) = \bar{w}^*(s, z_2) q^*(s, z_2),$$

and from (6), (7)

$$w^*(s, z_2) = \frac{L(s, x(s, z_2), z_2) e(s, x(s, z_2), z_2)}{1 - K(s, x(s, z_2), z_2) r^*(s, z_2)}, \tag{9}$$

while, by suitable differentiations, the mean number of P_2 customers arriving during W_2 and the duration of W_2 are given by

$$\rho_w = E(N(W_2)) = \frac{\bar{\rho}_v + \bar{\rho}_c}{1 - \lambda_1 \bar{b}_1}, \qquad E(W_2) = \frac{1 - \rho_w}{\lambda_2}. \tag{10}$$

We define finally the "generalized busy period" of P_1 customers as the time interval, W_1 say, from the epoch at which a P_1 customer arrives in an idle server until the epoch at which the server remains idle again. If as before $N(W_1)$ is the number of new P_2 customers arrive during W_1 and

$$d_j(t)dt = \Pr[t < W_1 \leq t{+}dt,\ N(W_1) = j], \qquad d^*(s, z_2) = \int_0^\infty e^{-st} \sum_{j=0}^{\infty} d_j(t) z_2^j dt,$$

then

$$d_j(t) = \sum_{i=0}^{\infty} \sum_{m=0}^{j} p_{im}(t) s_1(t) * \sum_{l=0}^{j-m} g_l^{(i+1)}(t) * \sum_{k=0}^{j-m-l} r_k(t) * q_{j-m-l-k}(t),$$

and so

$$d^*(s, z_2) = x(s, z_2) s_1^*(a(s, x(s, z_2), z_2)) r^*(s, z_2) q^*(s, z_2). \tag{11}$$

If finally we differentiate (11), with respect to z_2 and s, we obtain

$$E(N(W_1)) = \bar{\rho}_d/(1 - \lambda_1 \bar{b}_1), \qquad E(W_1) = \frac{E(N(W_1))}{\lambda_2}, \tag{12}$$

where

$$\bar{\rho}_d = \frac{\lambda_2}{\lambda_1 v^*(\lambda_1) c^*(\lambda_1)} \{(1 + \lambda_1 \bar{s}_3)(1 - c^*(\lambda_1)) + c^*(\lambda_1)[\lambda_1(\bar{v} + \bar{s}_1) + \lambda_1 \bar{b}_1 v^*(\lambda_1)]\}.$$

4 Time Dependent Analysis

Let $N_i(t)$ $i = 1, 2$ be the number of P_i customers in the system at time t and denote by

$$\xi_t = \begin{cases} b_i & if\ a\ P_i\ customer\ in\ service\ at\ t \quad i = 1, 2 \\ s_i & if\ a\ P_i\ customer\ in\ start\ up\ at\ t \quad i = 1, 2 \\ s_3 & if\ a\ P_1\ customer\ in\ special\ start\ up\ at\ t \\ c & if\ the\ server\ on\ close\ down\ at\ t \\ v & if\ the\ server\ on\ vacation\ at\ t \\ id & if\ the\ server\ idle\ at\ t \end{cases}$$

and

$$u_t = \begin{cases} 1 & an\ interrupted\ P_2\ customer\ waits\ at\ t \\ 0 & no\ interrupted\ P_2\ customer\ waits\ at\ t \end{cases}$$

Let us denote also by $\bar{X}(t)$ the elapsed duration at time t of any random variable X . Define

$$p_{ij}^{(b_k)}(x, t)dx = \Pr[N_1(t) = i, N_2(t) = j, \xi_t = b_k, u_t = 0, x < \bar{B}_k(t) \leq x + dx], \quad k = 1, 2$$

$$p_{ij}^{(s_k)}(x, t)dx = \Pr[N_1(t) = i, N_2(t) = j, \xi_t = s_k, u_t = 0, x < \bar{S}_k(t) \leq x + dx], \quad k = 1, 2, 3$$

$$p_{ij}^{(c)}(x, t)dx = \Pr[N_1(t) = i, N_2(t) = j, \xi_t = c, u_t = 0, x < \bar{C}(t) \leq x + dx],$$

$$p_{ij}^{(v)}(x, t)dx = \Pr[N_1(t) = i, N_2(t) = j, \xi_t = v, u_t = 0, x < \bar{V}(t) \leq x + dx],$$

$$q_j^{(id)}(t) = \Pr[N_1(t) = 0, N_2(t) = j, \xi_t = id, u_t = 0],$$

$$\begin{aligned} P^{(\xi_t)}(s,z_1,z_2,x) &= \int_0^\infty e^{-st}\sum_{i=0}^\infty\sum_{j=0}^\infty p_{ij}^{(\xi_t)}(x,t)z_1^i z_2^j dt, \\ Q^*(s,z_2) &= \int_0^\infty e^{-st}\sum_{j=0}^\infty q_j^{(id)}(t)z_2^j dt, \end{aligned}$$

and denote by $\vec{p}_{ij}^{(\xi_t)}(x,t)$, $\vec{P}^{(\xi_t)}(s,z_1,z_2,x)$ the corresponding quantities for $u_t = 1$.Then by connecting as usual the probabilities at t and $t+dt$, forming Laplace Transforms and generating functions and solving the simple differential equations we arrive for $x>0$ at

$$\begin{aligned} &P^{(b_k)}(s,z_1,z_2,x) = P^{(b_k)}(s,z_1,z_2,0)(1-B_k(x))\exp[-a(s,z_1,z_2)x], \quad k=1,2 \\ &P^{(s_k)}(s,z_1,z_2,x) = P^{(s_k)}(s,z_1,z_2,0)(1-S_k(x))\exp[-a(s,z_1,z_2)x], \quad k=1,2,3 \\ &P^{(c)}(s,0,z_2,x) = P^{(c)}(s,0,z_2,0)(1-C(x))\exp[-a(s,0,z_2)x], \\ &P^{(v)}(s,z_1,z_2,x) = P^{(v)}(s,0,z_2,0)(1-V(x))\exp[-a(s,z_1,z_2)x], \end{aligned} \tag{13}$$

while

$$\begin{aligned} &\vec{P}^{(b_1)}(s,z_1,z_2,x) = \vec{P}^{(b_1)}(s,z_1,z_2,0)(1-B_1(x))\exp[-a(s,z_1,z_2)x], \\ &\vec{P}^{(s_k)}(s,z_1,z_2,x) = \vec{P}^{(s_k)}(s,z_1,z_2,0)(1-S_k(x))\exp[-a(s,z_1,z_2)x], \quad k=1,3 \\ &\vec{P}^{(c)}(s,0,z_2,x) = \vec{P}^{(c)}(s,0,z_2,0)(1-C(x))\exp[-a(s,0,z_2)x], \end{aligned} \tag{14}$$

and for the idle mode

$$az_2\frac{d}{dz_2}Q^*(s,z_2) + (s+\lambda)Q^*(s,z_2) = 1 + P^{(v)}(s,0,z_2,0)v^*(a(s,0,z_2)). \tag{15}$$

In a similar way we obtain, after algebraic manipulations, for the boundary conditions ($x=0$),

$$\begin{aligned} [z_1 - b_1^*(a(s,z_1,z_2))]P^{(b_1)}(s,z_1,z_2,0) &= P^{(s_1)}(s,z_1,z_2,0)s_1^*(a(s,z_1,z_2)) \\ &+P^{(s_3)}(s,z_1,z_2,0)s_3^*(a(s,z_1,z_2)) - P^{(b_1)}(s,0,z_2,0)b_1^*(a(s,0,z_2)) \end{aligned} \tag{16}$$

$$\begin{aligned} [z_1 - b_1^*(a(s,z_1,z_2))]\vec{P}^{(b_1)}(s,z_1,z_2,0) &= \vec{P}^{(s_1)}(s,z_1,z_2,0)s_1^*(a(s,z_1,z_2)) \\ &+\vec{P}^{(s_3)}(s,z_1,z_2,0)s_3^*(a(s,z_1,z_2)) - \vec{P}^{(b_1)}(s,0,z_2,0)b_1^*(a(s,0,z_2)), \end{aligned} \tag{17}$$

$$\begin{aligned} \vec{P}^{(s_1)}(s,z_1,z_2,0) &= \lambda_1 z_1 \tfrac{1-s_2^*(a(s,0,z_2))}{a(s,0,z_2)} P^{(s_2)}(s,0,z_2,0), \\ P^{(s_3)}(s,z_1,z_2,0) &= \lambda_1 z_1 \tfrac{1-c^*(a(s,0,z_2))}{a(s,0,z_2)} P^{(c)}(s,0,z_2,0), \\ \vec{P}^{(s_3)}(s,z_1,z_2,0) &= \lambda_1 z_1 \tfrac{1-c^*(a(s,0,z_2))}{a(s,0,z_2)} \vec{P}^{(c)}(s,0,z_2,0), \end{aligned} \tag{18}$$

$$
\begin{aligned}
P^{(v)}(s,0,z_2,0) &= P^{(c)}(s,0,z_2,0)c^*(a(s,0,z_2)),\\
\vec{P}^{(c)}(s,0,z_2,0) &= \vec{P}^{(b_1)}(s,0,z_2,0)b_1^*(a(s,0,z_2)),\\
P^{(b_2)}(s,0,z_2,0) &= P^{(s_2)}(s,0,z_2,0)s_2^*(a(s,0,z_2)),
\end{aligned}
\tag{19}
$$

while finally

$$
\begin{aligned}
P^{(c)}(s,0,z_2,0) &= P^{(b_1)}(s,0,z_2,0)b_1^*(a(s,0,z_2) + P^{(b_2)}(s,0,z_2,0)b_2^*(a(s,0,z_2)),\\
P^{(s_1)}(s,z_1,z_2,0) &= \lambda_1 z_1 Q^*(s,z_2) + P^{(v)}(s,0,z_2,0)[v^*(a(s,z_1,z_2))\\
&\quad -v^*(a(s,0,z_2))] + P^{(b_2)}(s,0,z_2,0)[b_2^*(a(s,z_1,z_2)) - b_2^*(a(s,0,z_2))],\\
P^{(s_2)}(s,0,z_2,0) &= a\tfrac{d}{dz_2}Q^*(s,z_2) + \lambda_2 Q^*(s,z_2) + \vec{P}^{(c)}(s,0,z_2,0)c^*(a(s,0,z_2)).
\end{aligned}
\tag{20}
$$

Let us define now

$$
T(s,z_1,z_2) = 1 - K(s,z_1,z_2)c^*(a(s,0,z_2)) - M(s,z_1,z_2),
$$

where the functions K and M have been defined in (8) and (2) respectively. Then by substituting from (18), (20) and (19) to (17) we arrive at

$$
\vec{P}^{(b_1)}(s,z_1,z_2,0) = \frac{K(s,z_1,z_2)\ Q_1^*(s,z_2) - T(s,z_1,z_2)\vec{P}^{(c)}(s,0,z_2,0)}{z_1 - b_1^*(a(s,z_1,z_2))}.
$$

with

$$
Q_1^*(s,z_2) = a\frac{d}{dz_2}Q^*(s,z_2) + \lambda_2 Q^*(s,z_2),
$$

and as the zero of the denominator in $|z_1| \le 1$, $x(s,z_2)$ say, must be zero of the numerator too, we obtain

$$
\vec{P}^{(c)}(s,0,z_2,0) = \frac{K(s,x(s,z_2),z_2)}{T(s,x(s,z_2),z_2)} Q_1^*(s,z_2)\ , \tag{21}
$$

$$
\vec{P}^{(b_1)}(s,z_1,z_2,0) = \frac{\vec{R}(s,z_1,z_2)}{z_1 - b_1^*(a(s,z_1,z_2))}\quad Q_1^*(s,z_2), \tag{22}
$$

with

$$
\vec{R}(s,z_1,z_2) = K(s,z_1,z_2) - K(s,x(s,z_2),z_2)\frac{T(s,z_1,z_2)}{T(s,x(s,z_2),z_2)}.
$$

Moreover from (18) and (20)

$$
\vec{P}^{(s_3)}(s,z_1,z_2,0) = \frac{M(s,z_1,z_2)K(s,x(s,z_2),z_2)}{T(s,x(s,z_2),z_2)s_3^*(a(s,z_1,z_2))}\ Q_1^*(s,z_2), \tag{23}
$$

$$
P^{(s_2)}(s,0,z_2,0) = \ R(s,z_2)\ Q_1^*(s,z_2), \tag{24}
$$

with

$$
R(s,z_2) = 1 + c^*(a(s,0,z_2))\frac{K(s,x(s,z_2),z_2)}{T(s,x(s,z_2),z_2)}.
$$

Now from (15)

$$P^{(v)}(s,0,z_2,0) = \frac{Q_2^*(s,z_2)-1}{v^*(a(s,0,z_2))}, \tag{25}$$

with

$$Q_2^*(s,z_2) = az_2 \frac{d}{dz_2} Q^*(s,z_2) + (s+\lambda) Q^*(s,z_2)$$

and substituting in (19)

$$P^{(c)}(s,0,z_2,0) = \frac{Q_2^*(s,z_2)-1}{v^*(a(s,0,z_2))c^*(a(s,0,z_2))}. \tag{26}$$

From (18), (19) and (24)

$$\begin{aligned}
\vec{P}^{(s_1)}(s,z_1,z_2,0) &= \tfrac{K(s,z_1,z_2)}{s_1^*(a(s,z_1,z_2))} R(s,z_2)\ Q_1^*(s,z_2),\\
P^{(b_2)}(s,0,z_2,0) &= s_2^*(a(s,0,z_2))\ R(s,z_2)\ Q_1^*(s,z_2),\\
P^{(s_3)}(s,z_1,z_2,0) &= \tfrac{M(s,z_1,z_2)}{s_3^*(a(s,z_1,z_2))}\ \tfrac{Q_2^*(s,z_2)-1}{v^*(a(s,0,z_2))c^*(a(s,0,z_2))}.
\end{aligned} \tag{27}$$

Substituting finally from (20) and (23) in (16) and denoting

$$\begin{aligned}
&h_1(s,z_1,z_2) = aL(s,z_1,z_2)\ R(s,z_2) - \ az_2/e(s,z_1,z_2),\\
&h_2(s,z_1,z_2) = \lambda_1 z_1 s_1^*(a(s,z_1,z_2)) + \lambda_2 L(s,z_1,z_2)\ R(s,z_2) - \ (s+\lambda)/e(s,z_1,z_2),
\end{aligned} \tag{28}$$

we arrive at

$$P^{(b_1)}(s,z_1,z_2,0) = \frac{h_1(s,z_1,z_2)\frac{d}{dz_2}Q^*(s,z_2)+\ h_2(s,z_1,z_2)Q^*(s,z_2)]+1/\ e(s,z_1,z_2)}{z_1-b_1^*(a(s,z_1,z_2))}, \tag{29}$$

and using the zero of the denominator in the unit disk we obtain

$$a(z_2 - D(s,z_2))\frac{d}{dz_2}Q^*(s,z_2) + F(s,z_2)Q^*(s,z_2) = 1, \tag{30}$$

where now

$$\begin{aligned}
D(s,z_2) &= \ L(s,x(s,z_2),z_2)\ R(s,z_2)e(s,x(s,z_2),z_2),\\
F(s,z_2) &= s+\lambda - \ \lambda_1 x(s,z_2) s_1^*(a(s,x(s,z_2),z_2))e(s,x(s,z_2),z_2) - \lambda_2 D(s,z_2)\\
&= s + \lambda_1(1-d^*(s,z_2)) + \lambda_2(1-D(s,z_2)).
\end{aligned} \tag{31}$$

We have to state here the following theorem

Theorem 1 *For (i) $Re(s) > 0$, $|w| \le 1$ (ii) $Re(s) \ge 0, |w| < 1$ and (iii) $Re(s) \ge 0, |w| \le 1$ and*

$$\rho = \lambda_1 \bar{b}_1 + \bar{\rho}_v + \bar{\rho}_c > 1 \tag{32}$$

the equation

$$z_2 - wD(s,z_2) = 0 \tag{33}$$

has one and only one root, $z_2 = \phi(s, w)$ say, inside the region $|z_2| < 1$. Specifically for $s = 0$ and $w = 1$, $\phi(0, 1)$ is the smallest positive real root of (33) with $\phi(0, 1) < 1$ if $\rho > 1$ and $\phi(0, 1) = 1$ for $\rho \leq 1$.

Proof: Comparing $D(s, z_2)$ in the first of (31) with the generating function $w^*(s, z_2)$ in (9) of section 3 one realizes easily that

$$D(s, z_2) = w^*(s, z_2) = \int_0^\infty e^{-st} \sum_{j=0}^{\infty} w_j(t) z_2^j dt,$$

i.e. $D(s, z_2)$ is in fact the Laplace transform of a generating function.

Thus for the closed contour $|z_2| = 1 - \epsilon$ ($\epsilon > 0$ is a small number) and under the assumptions (i) and (ii) we can always find a sufficiently small $\epsilon \geq 0$ such that

$$|wD(s, z_2)| \leq |w|\, D(Re(s), 1 - \epsilon) < 1 - \epsilon, \tag{34}$$

while for $Re(s) \geq 0$, $|w| \leq 1$ we need in addition

$$\frac{d}{d\epsilon} D(0, 1 - \epsilon) \mid_{\epsilon=0} < -1,$$

or $\rho > 1$ for the relation (34) to hold. A final reference to Rouche's theorem completes the first part of the proof.

Moreover for $s = 0$ and $w = 1$ the convex function $D(0, z_2)$ is a monotonically increasing function of z_2, for $0 \leq z_2 \leq 1$, taking the values $0 < D(0, 0) < 1$ and $D(0, 1) = 1$ and so $0 < \phi(0, 1) < 1$ if $\rho > 1$, while for $\rho \leq 1$, $\phi(0, 1)$ becomes equal to 1 and this completes the proof. □

Using the theorem above one can solve (see Falin & Fricker [7]) the differential equation (30) and obtain

$$Q^*(s, z_2) = \frac{1}{F(s, z_2)}, \qquad if\ z_2 = \phi(s, 1),$$

$$Q^*(s, z_2) = \int_{z_2}^{\phi(s,1)} \frac{1}{a(D(s, u) - u)} \exp\{\int_u^{z_2} \frac{F(s, x)}{a(D(s, x) - x)} dx\} du, \quad if\ z_2 \neq \phi(s, 1).$$

Thus the quantity $Q^*(s, z_2)$ is known and so from the second of (20) and (21)- (29) all generating functions are completely known. This completes the time-dependent analysis of the model.

5 Stability Conditions

For a stochastic process $(Y(t)\ ; t \geq 0)$ we will say that it is stable, if its limiting probabilities as $t \to \infty$ exist and form a distribution.

Consider now the points T_n in time at which, either a generalized busy period of P_1 customers, or a generalized completion time of a P_2 customer is finished, i.e. the points at which the server becomes idle. If

$$0 = T_0 < T_1 < T_2 < \cdots,$$

is the sequence of these points in ascending order and define $\zeta_n = N_2(T_n + 0)$, then it is easy to understand that the stochastic process $Z = (\zeta_n; n \geq 0)$ is an irreducible and aperiodic Markov chain. Then

Theorem 2 *For $\rho < 1$ the Markov chain Z is positive recurrent.*

Proof. To prove the theorem, we will use the following criterion (see Pakes [8]):

An irreducible and aperiodic Markov chain $(Y_n \; ; \; n \geq 0)$, with state space the nonnegative integers, is positive recurrent if $|\delta_k| < \infty$ for all $k = 0, 1, 2, \ldots$ and $\limsup_{k \to \infty} \delta_k < 0$, where $\delta_k = E[Y_{n+1} - Y_n \mid Y_n = k]$.

For the Markov chain Z of our model, let

$$h_{k,m}(t)dt = \Pr[t < T_{n+1} - T_n \leq t + dt, \; N_2(T_{n+1}) - N_2(T_n) = m | \; N_2(T_n) = k].$$

Then it is easy to see that for $m = 0, 1, 2, \ldots$

$$\begin{aligned} h_{k,m}(t) = \; & \lambda_1 e^{-(\lambda_1+\lambda_2+ka)t} * d_m(t) + \lambda_2 e^{-(\lambda_1+\lambda_2+ka)t} * w_m(t) \\ & + ka e^{-(\lambda_1+\lambda_2+ka)t} * w_{m+1}(t), \end{aligned}$$

while for $m = -1$

$$h_{k,-1}(t) = \; ka e^{-(\lambda_1+\lambda_2+ka)t} * w_0(t),$$

and so

$$\int_0^\infty e^{-st} \sum_{m=-1}^{\infty} h_{k,m}(t) z^m dt = \frac{\lambda_1 d^*(s,z) + \lambda_2 w^*(s,z) + \frac{ka}{z} w^*(s,z)}{s + \lambda_1 + \lambda_2 + ka}, \tag{35}$$

and by taking derivatives above with respect to z at the point $(z = 1, s = 0)$ we arrive at

$$\delta_k = \frac{\lambda_1 E(N(W_1)) + \lambda_2 E(N(W_2)) + ka[E(N(W_2)) - 1]}{\lambda_1 + \lambda_2 + ka}, \qquad k = 0, 1, \ldots .$$

where $E(N(W_1))$, $E(N(W_2))$ have been found in (10) and (12) respectively.

Thus for $\rho < 1$ we realize that $|\delta_k|$ is finite for all k and also $\limsup_{k \to \infty} \delta_k = E(N(W_2)) - 1 = \frac{\rho - 1}{1 - \lambda_1 b_1} < 0$, and the criterion is satisfied. □

Consider now the stochastic process

$$\mathbf{Z} = \{(N_1(t), \; N_2(t), \; \xi_t) : \; 0 \leq t < \infty\}$$

where $N_i(t)$, ξ_t have been defined in section 4. Then

Theorem 3 *For $\rho < 1$ the process $\mathbf{Z}$ is stable.*

Proof: Consider the quantity

$$m_k = E(T_1 |\ \zeta_0 = k)$$

By taking derivatives in (35) with respect to s (at $z = 1$) we obtain

$$m_k = \frac{\lambda_1 E(W_1) + \lambda_2 E(W_2) + kaE(W_2) + 1}{\lambda_1 + \lambda_2 + ka},$$

and if $q_k \quad k = 0, 1, 2, ...$ are the steady state probabilities of the positive recurrent Markov chain Z then

$$\mathbf{q} \cdot \mathbf{m} = \sum_{k=0}^{\infty} q_k m_k = E(W_2) + \{1 + \lambda_1 [E(W_1) - E(W_2)]\} \sum_{k=0}^{\infty} \frac{q_k}{\lambda_1 + \lambda_2 + ka}. \quad (36)$$

Now it is clear that there is always a finite integer k^* such that

$$\frac{1}{\lambda_1 + \lambda_2 + (k^* - 1)a} > 1 > \frac{1}{\lambda_1 + \lambda_2 + k^* a},$$

and so

$$\begin{aligned} \textstyle\sum_{k=0}^{\infty} \frac{q_k}{\lambda_1+\lambda_2+ka} = & \textstyle\sum_{k=0}^{k^*-1} \frac{q_k}{\lambda_1+\lambda_2+ka} + \sum_{k=k^*}^{\infty} \frac{q_k}{\lambda_1+\lambda_2+ka} < \sum_{k=0}^{k^*-1} \frac{q_k}{\lambda_1+\lambda_2+ka} \\ & \textstyle + \sum_{k=k^*}^{\infty} q_k = \sum_{k=0}^{k^*-1} \frac{q_k}{\lambda_1+\lambda_2+ka} + (1 - \sum_{k=0}^{k^*-1} q_k) < \infty \end{aligned}$$

and so from (36) using (10), (12) we understand that $\mathbf{q} \cdot \mathbf{m} < \infty$.

Consider finally the irreducible aperiodic and positive recurrent Markov Renewal Process $\{Z,\ T\} = \{(\zeta_n,\ T_n)\ :\ n = 0, 1, 2, ...\}$. It is easy to see that the stochastic process $\mathbf{Z}$ is a Semi-Regenerative Process with imbedded Markov Renewal Process $\{Z,\ T\}$ and as $\mathbf{q} \cdot \mathbf{m} < \infty$ it is clear that $\mathbf{Z}$ is stable (Cinlar [9], Theorem 6.12 p.347). $\square$

6 Steady State Probabilities

Suppose now that $\rho < 1$. Let

$$p_{ij}^{(\xi_t)}(x) = \lim_{t\to\infty} p_{ij}^{(\xi_t)}(x, t), \qquad q_j^{(id)} = \lim_{t\to\infty} q_j^{(id)}(t), \qquad Q^*(z_2) = \sum_{j=0}^{\infty} q_j^{(id)} z_2^j,$$

$$P^{(\xi_t)}(z_1, z_2, x) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p_{ij}^{(\xi_t)}(x) z_1^i z_2^j, \qquad P^{(\xi_t)}(z_1, z_2) = \int_0^{\infty} P^{(\xi_t)}(z_1, z_2, x)dx,$$

and denote by $\vec{p}_{ij}^{(\xi_t)}(x)$, $\vec{P}^{(\xi_t)}(z_1, z_2, x)$, $P^{(\xi_t)}(z_1, z_2)$ the corresponding quantities for $u_t = 1$. Then it is well known that

$$p_{ij}^{(\xi_t)}(x) = \lim_{t\to\infty} p_{ij}^{(\xi_t)}(x, t) = \lim_{s\to 0} s \int_0^{\infty} e^{-st} p_{ij}^{(\xi_t)}(x, t)dt,$$

$$q_j^{(id)} = \lim_{t\to\infty} q_j^{(id)}(t) = \lim_{s\to 0} s \int_0^\infty e^{-st} q_j^{(id)}(t)dt,$$

and so integrating with respect to x, multiplying by s and taking limits $s \to \infty$ in (13), (14) we arrive at

$$\begin{aligned}
P^{(b_k)}(z_1,z_2) &= P^{(b_k)}(z_1,z_2,0)[1-b_k^*(a(0,z_1,z_2))]/a(0,z_1,z_2), \qquad k=1,2\\
P^{(s_k)}(z_1,z_2) &= P^{(s_k)}(z_1,z_2,0)\ [1-s_k^*(a(0,z_1,z_2))]/a(0,z_1,z_2), \qquad k=1,2,3\\
P^{(c)}(0,z_2) &= P^{(c)}(0,z_2,0)\ [1-c^*(a(0,z_1,z_2))]/a(0,0,z_2),\\
P^{(v)}(z_1,z_2) &= P^{(v)}(0,z_2,0)\ [1-v^*(a(0,z_1,z_2))]/a(0,z_1,z_2),
\end{aligned} \tag{37}$$

and

$$\begin{aligned}
\vec{P}^{(b_1)}(z_1,z_2) &= \vec{P}^{(b_1)}(z_1,z_2,0)\ [1-b_1^*(a(0,z_1,z_2))]/a(0,z_1,z_2),\\
\vec{P}^{(s_k)}(z_1,z_2) &= \vec{P}^{(s_k)}(z_1,z_2,0)\ [1-s_k^*(a(0,z_1,z_2))]/a(0,z_1,z_2), \qquad k=1,3\\
\vec{P}^{(c)}(0,z_2) &= \vec{P}^{(c)}(0,z_2,0)\ [1-c^*(a(0,z_1,z_2))]/a(0,0,z_2).
\end{aligned} \tag{38}$$

In a similar way we obtain for the boundary conditions

$$\begin{aligned}
\vec{P}^{(c)}(0,z_2,0) &= \tfrac{K(0,x(0,z_2),z_2)}{T(0,x(0,z_2),z_2)}\ Q_1^*(z_2),\\
\vec{P}^{(b_1)}(z_1,z_2,0) &= \tfrac{\vec{R}(0,z_1,z_2)}{z_1-b_1^*(a(0,z_1,z_2))}\ Q_1^*(z_2),\\
\vec{P}^{(s_3)}(z_1,z_2,0) &= \tfrac{M(0,z_1,z_2)K(0,x(0,z_2),z_2)}{T(0,x(0,z_2),z_2)s_3^*(a(0,z_1,z_2))}\ Q_1^*(z_2),\\
P^{(s_2)}(0,z_2,0) &= R(0,z_2)\ Q_1^*(z_2),\\
\vec{P}^{(s_1)}(z_1,z_2,0) &= \tfrac{K(0,z_1,z_2)}{s_1^*(a(0,z_1,z_2))} R(0,z_2)\ Q_1^*(z_2),\\
P^{(b_2)}(0,z_2,0) &= s_2^*(a(0,0,z_2))\ R(0,z_2)\ Q_1^*(z_2),
\end{aligned} \tag{39}$$

$$\begin{aligned}
P^{(v)}(0,z_2,0) &= \tfrac{Q_2^*(z_2)}{v^*(a(0,0,z_2))},\\
P^{(c)}(0,z_2,0) &= \tfrac{Q_2^*(z_2)}{v^*(a(0,0,z_2))c^*(a(0,0,z_2))},\\
P^{(s_3)}(z_1,z_2,0) &= \tfrac{M(0,z_1,z_2)}{s_3^*(a(0,z_1,z_2))}\ \tfrac{Q_2^*(z_2)}{v^*(a(0,0,z_2))c^*(a(0,0,z_2))},
\end{aligned} \tag{40}$$

with

$$Q_1^*(z_2) = a\frac{d}{dz_2}Q^*(z_2) + \lambda_2 Q^*(z_2), \qquad Q_2^*(z_2) = az_2\frac{d}{dz_2}Q^*(z_2) + \lambda Q^*(z_2)$$

$$\begin{aligned}
P^{(s_1)}(z_1,z_2,0) = \ & \lambda_1 z_1 Q^*(z_2) + P^{(v)}(0,z_2,0)[v^*(a(0,z_1,z_2)) - v^*(a(0,0,z_2))]\\
& + P^{(b_2)}(0,z_2,0)[b_2^*(a(0,z_1,z_2)) - b_2^*(a(0,0,z_2))],
\end{aligned} \tag{41}$$

$$P^{(b_1)}(z_1, z_2, 0) = \frac{h_1(0, z_1, z_2)\frac{d}{dz_2}Q^*(z_2) + \ h_2(0, z_1, z_2)Q^*(z_2)]}{z_1 - b_1^*(a(0, z_1, z_2))}, \tag{42}$$

while the differential equation (30) becomes

$$a(z_2 - D(0, z_2))\frac{d}{dz_2}Q^*(z_2) + F(0, z_2)Q^*(z_2) = 0, \tag{43}$$

with $D(0, z_2) = w^*(0, z_2)$. From (31)

$$F(0, z_2) = \lambda_1(1 - d^*(0, z_2)) + \lambda_2(1 - D(0, z_2)) = \lambda[1 - G(z_2)],$$

where

$$G(z_2) = \frac{\lambda_1 d^*(0, z_2) + \lambda_2 w^*(0, z_2)}{\lambda}.$$

Let now

$$\omega(z_2) = \frac{1 - G(z_2)}{z_2 - w^*(0, z_2)},$$

then for $\rho < 1$ the quantity $z_2 - w^*(0, z_2)$ never becomes zero in $|z_2| < 1$ (Theorem 1) and also

$$\lim_{z_2 \to 1} \omega(z_2) = -\frac{\frac{\lambda_1}{\lambda}\bar{\rho}_d + \frac{\lambda_2}{\lambda}(\bar{\rho}_v + \bar{\rho}_c)}{1 - \rho} < \infty.$$

Thus $\omega(z_2)$ is an analytic function in $|z_2| < 1$ and a continuous one on the boundary and so for any $|z_2| \leq 1$ we can solve equation (43) and obtain

$$Q^*(z_2) = Q^*(1)\exp\{-\frac{\lambda}{a}\int_{z_2}^{1} \frac{1 - G(u)}{w^*(0, u) - u} du\},$$

Replacing finally $Q^*(z_2)$ back in the generating functions and asking for the total probabilities to sum to unity we arrive at

$$Q^*(1) = \frac{1 - \rho}{1 - \lambda_1\bar{b}_1 + \frac{\lambda_1}{\lambda_2}\bar{\rho}_d}.$$

and so the generating functions of the steady state probabilities are completely known.

The following theorem shows that the condition $\rho < 1$ is also necessary for a stable system.

Theorem 4 *If the stochastic process* $\mathbf{Z}$ *is stable then* $\rho < 1$.

Proof: Suppose that $\mathbf{Z}$ stable and $\rho > 1$. Then from theorem 1 the equation $z_2 - D(0, z_2) = 0$ has a root $0 < \phi(0, 1) < 1$ and

$$F(0, \phi(0, 1)) = \lambda_1(1 - d^*(0, \phi(0, 1))) + \lambda_2(1 - D(0, \phi(0, 1))) \neq 0.$$

By putting now $\phi(0,1)$ instead of z_2 in (43) we obtain

$$F(0,\phi(0,1))Q^*(\phi(0,1)) = 0,$$

and so $Q^*(\phi(0,1)) = \sum q_j^{(id)} \phi^j(0,1) = 0$ with $0 < \phi(0,1) < 1$. Thus $q_j^{(id)} = 0 \ \forall\, j$ and also from the generating functions in (37)- (42) it is clear that all probabilities become zero. This of course contradicts to the hypothesis that the system is stable.

Suppose finally that **Z** stable and $\rho = 1$. By taking derivatives with respect to z_2 in (43) (at $z_2 = 1$) we arrive (for $\rho = 1$) at

$$\frac{d}{dz_2}F(0,z_2)|_{z_2=1}Q^*(1) = -[\lambda_1 E(N(W_1)) + \lambda_2 E(N(W_2))]Q^*(1) = 0,$$

and so $Q^*(1) = \sum q_j^{id} = 0$ and this again contradicts to the hypothesis that the system is stable. $\square$

7 Performance Measures

7.1 Probabilities of server state

In this section we will use formulas for the generating functions obtained previously, to derive expressions for the probabilities of server state. Thus by putting $z_1 = z_2 = 1$ into relations (37)-(42) we obtain easily

$$\begin{aligned}
&P[\textit{server Idle}] = && P(\xi = id) = Q^*(1) = \tfrac{1-\rho}{1-\lambda_1\bar{b}_1 + \frac{\lambda_1}{\lambda_2}\bar{\rho}_d} \\
&P[\textit{a } P_1 \textit{ customer in service}] = && P(\xi = b_1) = \lambda_1\bar{b}_1 \\
&P[\textit{a } P_2 \textit{ customer in service}] = && P(\xi = b_2) = \lambda_2\bar{b}_2 \\
&P[\textit{server in } P_2 \textit{ start up}] = && P(\xi = s_2) = \tfrac{\lambda_2}{s_2^*(\lambda_1)}\tfrac{1-s_2^*(\lambda_1)}{\lambda_1} \\
&P[\textit{server in vacation}] = && P(\xi = v) = \tfrac{\lambda_2+\lambda_1 Q^*(1)}{v^*(\lambda_1)}\bar{v} \\
&P[\textit{server in close down}] = && P(\xi = c) = \tfrac{1-c^*(\lambda_1)}{\lambda_1 c^*(\lambda_1)}[\lambda_2\tfrac{1-s_2^*(\lambda_1)}{s_2^*(\lambda_1)} + \tfrac{\lambda_2+\lambda_1 Q^*(1)}{v^*(\lambda_1)}] \\
&P[\textit{server in special start up}] = && P(\xi = s_3) = \lambda_1\bar{s}_3 P(\xi = c) \\
&P[\textit{server in } P_1 \textit{ start up}] = && P(\xi = s_1) = \bar{s}_1[\tfrac{\lambda_2+\lambda_1 Q^*(1)}{v^*(\lambda_1)} - \lambda_2 b_2^*(\lambda_1) + \lambda_2\tfrac{1-s_2^*(\lambda_1)}{s_2^*(\lambda_1)}]
\end{aligned}$$

7.2 Mean number of ordinary customers

For any p.d.f. $a(t)$, let us denote now $\bar{a}^{(2)} = \int_0^\infty t^2 a(t)dt$, i.e., denote by $\bar{a}^{(2)}$ its second moment about zero. By differentiating the generating functions with respect to z_1 at the point $z_1 = z_2 = 1$ we obtain the mean number of P_1 customers, according to the server state, as following,

$$E(N_1;\ \xi = v) = \frac{\lambda_1\bar{v}^{(2)}}{2v^*(\lambda_1)}(\lambda_2 + \lambda_1 Q^*(1))$$

$$E\left(N_1;\xi=s_3\right)=\left(1-c^*\left(\lambda_1\right)\right)\left(\bar{s}_3+\frac{\lambda_1\bar{s}_3^{(2)}}{2}\right)\left(\frac{\lambda_2+\lambda_1Q^*\left(1\right)}{c^*\left(\lambda_1\right)v^*\left(\lambda_1\right)}+\lambda_2\frac{1-s_2^*\left(\lambda_1\right)}{c^*\left(\lambda_1\right)s_2^*\left(\lambda_1\right)}\right)$$

$$\begin{aligned}E\left(N_1;\xi=s_1\right) &= \lambda_1\bar{s}_1[Q^*\left(1\right)+\frac{\bar{v}}{v^*\left(\lambda_1\right)}\left(\lambda_2+\lambda_1Q^*\left(1\right)\right)+\lambda_2\bar{b}_2]+\frac{\lambda_1\bar{s}_1^{(2)}}{2}\\ &\times[\frac{\lambda_2+\lambda_1Q^*\left(1\right)}{v^*\left(\lambda_1\right)}-\lambda_2b_2^*\left(\lambda_1\right)]+\frac{\lambda_2\left(1-s_2^*\left(\lambda_1\right)\right)\left(\bar{s}_1+\lambda_1\frac{\bar{s}_1^{(2)}}{2}\right)}{s_2^*\left(\lambda_1\right)}\end{aligned}$$

$$E\left(N_1;\xi\ =b_2\right)=\frac{\lambda_1\lambda_2}{2}\bar{b}_2^{(2)}$$

$$E\left(N_1;\xi\ =b_1\right)=\frac{\lambda_1\bar{b}_1}{2(1-\lambda_1\bar{b}_1)}\left(\lambda_1\bar{b}_1^{(2)}/\bar{b}_1+A_1+A_2\right)$$

where

$$A_1=\frac{\lambda_2\left(1-s_2^*\left(\lambda_1\right)\right)}{s_2^*\left(\lambda_1\right)c^*\left(\lambda_1\right)}\left[\left(1-c^*\left(\lambda_1\right)\right)\left(\lambda_1\bar{s}_3^{(2)}+2\bar{s}_3\right)+c^*\left(\lambda_1\right)\left(\lambda_1\bar{s}_1^{(2)}+2\bar{s}_1\right)\right]$$

$$\begin{aligned}A_2 &= \lambda_1(\lambda_1\bar{s}_1^{(2)}+2\bar{s}_1)Q^*\left(1\right)+\frac{\lambda_2+\lambda_1Q^*\left(1\right)}{c^*\left(\lambda_1\right)v^*\left(\lambda_1\right)}\Big[\left(1-c^*\left(\lambda_1\right)\right)\left(\lambda_1\bar{s}_3^{(2)}+2\bar{s}_3\right)\\ &+\lambda_1c^*\left(\lambda_1\right)\left(\bar{s}_1^{(2)}\left(1-v^*\left(\lambda_1\right)\right)+2\bar{s}_1\bar{v}+\bar{v}^{(2)}\right)\Big]\\ &+\lambda_1\lambda_2(\bar{s}_1^{(2)}\left(1-b_2^*\left(\lambda_1\right)\right)+2\bar{s}_1\bar{b}_2+\bar{b}_2^{(2)})\end{aligned}$$

7.3 Mean number of retrial customers

To derive expressions for the mean number of customers in the retrial box we need firstly to calculate the derivatives of some functions defined in previous sections. Let

$$U\left(0,x\left(0,z_2\right),z_2\right)=\frac{K\left(0,x\left(0,z_2\right),z_2\right)}{T\left(0,x\left(0,z_2\right),z_2\right)},\quad U=U\left(0,1,1\right)=\frac{1-s_2^*(\lambda_1)}{c^*(\lambda_1)s_2^*(\lambda_1)},$$

$$Q^{*(1)}=\tfrac{d}{dz_2}\,Q^*\left(z_2\right)\big|_{z_2=1}=\tfrac{\lambda_2}{a}\left(\tfrac{\lambda_1\bar{p}_d+\lambda_2(\bar{p}_c+\bar{p}_v)}{\lambda_1\bar{p}_d+\lambda_2(1-\lambda_1\bar{b}_1)}\right),$$

$$Q^{*(2)}=\tfrac{d^2}{dz_2^2}\,Q^*\left(z_2\right)\big|_{z_2=1}=\tfrac{1-\lambda_1\bar{b}_1}{2a(1-\rho)}[\lambda_2\rho_w^{(2)}+\lambda_1\rho_d^{(2)}Q^*(1)],$$

$$H_1=aQ^{*(2)}+\lambda_2Q^{*(1)},\quad H_2=aQ^{*(2)}+\left(\lambda+a\right)Q^{*(1)}$$

where $\rho_d^{(2)}$ and $\rho_w^{(2)}$ are given below in (44) and (45). For any function $f^*(z_2)$ denote $f^{*(\nu)}=\frac{d^\nu}{dz_2^\nu}\,f^*(z_2)|_{z_2=1}$ and

$$\sigma_{f^*}\left(z_2\right)=\left(\tfrac{1-f^*(a(0,0,z_2))}{a(0,0,z_2)}\right),\qquad \sigma_{f^*}=\sigma_{f^*}\left(1\right)=\tfrac{1-f^*(\lambda_1)}{\lambda_1},$$

$$\sigma_{f^*}^{(1)}=\tfrac{d}{dz_2}\,\sigma_{f^*}\left(z_2\right)\big|_{z_2=1}=\tfrac{\lambda_1\lambda_2f^{*(1)}(\lambda_1)+\lambda_2(1-f^*(\lambda_1))}{\lambda_1^2},$$

$$\sigma_f^{(2)}=\tfrac{d^2}{dz_2^2}\,\sigma_{f^*}\left(z_2\right)\big|_{z_2=1}=\tfrac{-\lambda_1^2\lambda_2^2f^{*(2)}(\lambda_1)+2\lambda_2\left(\lambda_1\lambda_2f^{*(1)}(\lambda_1)+\lambda_2(1-f^*(\lambda_1))\right)}{\lambda_1^3}$$

while for any function $w\left(0,x\left(0,z_2\right),z_2\right)$ denote $\hat{w}^{(v)}=\frac{d^{v}}{dz_2^{v}}\ w\left(0,x\left(0,z_2\right),z_2\right)|_{z_2=1}$ .Then

$$\hat{K}^{(1)}=\lambda_2 s_2^{*(1)}\left(\lambda_1\right)+\tfrac{\lambda_2}{1-\lambda_1\bar{b}_1}\sigma_{s_2^*}\left(1+\lambda_1\bar{s}_1\right),$$

$$\hat{M}^{(1)}=\lambda_2 c^{*(1)}\left(\lambda_1\right)+\tfrac{\lambda_2}{1-\lambda_1\bar{b}_1}\sigma_{c^*}\left(1+\lambda_1\bar{s}_3\right),$$

$$\hat{T}^{(1)}=-\hat{K}^{(1)}c^*\left(\lambda_1\right)+\lambda_1\lambda_2 c^{*(1)}\left(\lambda_1\right)\ \sigma_{s_2^*}-\hat{M}^{(1)},$$

$$\hat{U}^{(1)}\ =\{\hat{K}^{(1)}c^*(\lambda_1)s_2^*(\lambda_1)-\hat{T}^{(1)}\lambda_1\ \sigma_{s_2^*}\}/\left(c^*(\lambda_1)s_2^*(\lambda_1)\right)^2,$$

$$\hat{e}^{(1)}=\tfrac{\lambda_2}{\lambda_1(1-\lambda_1\bar{b}_1)}\left[\tfrac{\lambda_1\ \sigma_{c^*}(1+\lambda_1\bar{s}_3)+\lambda_1 c^*(\lambda_1)(\bar{v}+\bar{s}_1\lambda_1\ \sigma_{v^*})}{c^*(\lambda_1)v^*(\lambda_1)}\right],$$

$$\hat{L}^{(1)}=-\lambda_2 s_2^{*(1)}\left(\lambda_1\right)+\tfrac{\lambda_2}{1-\lambda_1\bar{b}_1}s_2^*\left(\lambda_1\right)\left(\bar{b}_2+\bar{s}_1\lambda_1\ \sigma_{b_2^*}\right),$$

$$R^{(1)}=\tfrac{d}{dz_2}\ R\left(0,z_2\right)|_{z_2=1}=c^*\left(\lambda_1\right)\hat{U}^{(1)}-\tfrac{\lambda_2 c^{*(1)}(\lambda_1)\lambda_1\ \sigma_{s_2^*}}{c^*(\lambda_1)s_2^*(\lambda_1)}.$$

where the functions K, M, T, e, L, R, have been defined in sections 3 and 4. Using the above quantities we have for the retrial customers the following results.

$$E\left(N_2;\xi=v\right)=\frac{\bar{v}}{v^*\left(\lambda_1\right)}H_2+\frac{\lambda_2+\lambda_1 Q^*\left(1\right)}{v^*\left(\lambda_1\right)}\lambda_2[\frac{\bar{v}^{(2)}}{2}+\frac{v^{*(1)}\left(\lambda_1\right)}{v^*\left(\lambda_1\right)}\bar{v}],$$

$$E\left(N_2;\xi=s_2\right)=\sigma_{s_2^*}[\lambda_2 R^{(1)}+\frac{H_1}{s_2^*\left(\lambda_1\right)}]+\frac{\lambda_2}{s_2^*\left(\lambda_1\right)}\sigma_{s_2^*}^{(1)},$$

$$E\left(N_2;\xi=b_2\right)=\bar{b}_2[s_2^*\left(\lambda_1\right)\left(\lambda_2 R^{(1)}+\frac{H_1}{s_2^*\left(\lambda_1\right)}\right)-\lambda_2^2\frac{s_2^{*(1)}\left(\lambda_1\right)}{s_2^*\left(\lambda_1\right)}]+\frac{\lambda_2^2}{2}\bar{b}_2^{(2)},$$

$$\begin{aligned}E(N_2;\xi\ &=c)=\sigma_{c^*}\{[\lambda_2 U^{(1)}+H_1 U]+\tfrac{1}{c^*(\lambda_1)v^*(\lambda_1)}[H_2+\tfrac{(\lambda_2+\lambda_1 Q^*(1))}{c^*(\lambda_1)v^*(\lambda_1)}\lambda_2\\ &\times\left(c^{*(1)}\left(\lambda_1\right)v^*\left(\lambda_1\right)+c^*\left(\lambda_1\right)v^{*(1)}\left(\lambda_1\right)\right)]\}+\sigma_{c^*}^{(1)}[\lambda_2 U+\tfrac{(\lambda_2+\lambda_1 Q^*(1))}{c^*(\lambda_1)v^*(\lambda_1)}],\end{aligned}$$

$$E\left(N_2;\xi=s_3\right)=\lambda_1\{\bar{s}_3 E\left(N_2;\xi=c\right)+\frac{\lambda_2}{2}\bar{s}_3^{(2)}P\left(\xi=c\right)\}$$

$$E\left(N_2;\xi=s_1,u=1\right)=\ \ \lambda_1\bar{s}_1[\tfrac{\lambda_2}{s_2^*(\lambda_1)}\sigma_{s_2^*}^{(1)}+\sigma_{s_2^*}(\lambda_2 R^{(1)}+\tfrac{H_1}{s_2^*(\lambda_1)})]+\tfrac{\lambda_2^2\lambda_1}{2s_2^*(\lambda_1)}\bar{s}_1^{(2)}\sigma_{s_2^*},$$

$$\begin{aligned}E(N_2;\xi=s_1\quad,u=0)&=\tfrac{\lambda_2\bar{s}_1^{(2)}}{2}[\tfrac{\lambda_2+\lambda_1 Q^*(1)}{v^*(\lambda_1)}-\lambda_2 b_2^*\left(\lambda_1\right)]+\bar{s}_1\{\ \tfrac{\lambda_2+\lambda_1 Q^*(1)}{v^*(\lambda_1)}\lambda_2[\bar{v}\\ &+v^{*(1)}\left(\lambda_1\right)]+\lambda_1 Q^{*(1)}+\tfrac{\lambda_1\sigma_{v^*}}{v^*(\lambda_1)}[H_2+\tfrac{\lambda_2 v^{*(1)}(\lambda_1)}{v^*(\lambda_1)}\left(\lambda_2+\lambda_1 Q^*\left(1\right)\right)]\\ &+\lambda_1\sigma_{b_2^*}[s_2^*\left(\lambda_1\right)\left(\lambda_2 R^{(1)}+\tfrac{H_1}{s_2^*(\lambda_1)}\right)-\tfrac{\lambda_2^2 s_2^{*(1)}(\lambda_1)}{s_2^*(\lambda_1)}]+\lambda_2^2(\bar{b}_2+b_2^{*(1)}\left(\lambda_1\right))\}.\end{aligned}$$

Except for the quantities introduced before we need also the second derivatives to proceed to $E\left(N_2;\xi=b_1\right)$. Thus

$$\begin{aligned}
\hat{K}^{(2)} = \; & 2\tfrac{\lambda_2^2 s_2^{*(1)}(\lambda_1)}{\lambda_1\left(1-\lambda_1\bar{b}_1\right)}\left(1+\lambda_1\bar{s}_1\right)+\tfrac{2\lambda_2^2}{\lambda_1\left(1-\lambda_1\bar{b}_1\right)}\sigma_{s_2^*}-\lambda_2^2 s_2^{*(2)}\left(\lambda_1\right)\\
& +\tfrac{\lambda_2^2\sigma_{s_2^*}}{(1-\lambda_1\bar{b}_1)^2}[(\lambda_1\bar{s}_1^{(2)}+2\bar{s}_1)+\tfrac{\lambda_1\bar{b}_1^{(2)}}{1-\lambda_1\bar{b}_1}\left(1+\lambda_1\bar{s}_1\right)],\\
\hat{M}^{(2)} = \; & 2\tfrac{\lambda_2^2 c^{*(1)}(\lambda_1)}{\lambda_1\left(1-\lambda_1\bar{b}_1\right)}\left(1+\lambda_1\bar{s}_3\right)+2\tfrac{\lambda_2^2}{\lambda_1\left(1-\lambda_1\bar{b}_1\right)}\sigma_{c^*}-\lambda_2^2 c^{*(2)}\left(\lambda_1\right)\\
& +\tfrac{\lambda_2^2\sigma_{c^*}}{(1-\lambda_1\bar{b}_1)^2}[(\lambda_1\bar{s}_3^{(2)}+2\bar{s}_3)+\tfrac{\lambda_1\bar{b}_1^{(2)}}{1-\lambda_1\bar{b}_1}\left(1+\lambda_1\bar{s}_3\right)],\\
\hat{L}^{(2)} = \; & \lambda_2^2 s_2^{*(2)}\left(\lambda_1\right)-2\tfrac{\lambda_2^2 s_2^{*(1)}(\lambda_1)(\bar{b}_2+\bar{s}_1)}{1-\lambda_1\bar{b}_1}+2\tfrac{\lambda_2^2\bar{s}_1(s_2^{*(1)}(\lambda_1)b_2^*(\lambda_1)+s_2^*(\lambda_1)b_2^{*(1)}(\lambda_1))}{1-\lambda_1\bar{b}_1}\\
& +\tfrac{\lambda_2^2 s_2^*(\lambda_1)}{(1-\lambda_1\bar{b}_1)^2}[\tfrac{\lambda_1\bar{b}_1^{(2)}}{1-\lambda_1\bar{b}_1}\left(\bar{b}_2+\lambda_1\bar{s}_1\sigma_{b_2^*}\right)+(\bar{b}_2^{(2)}+\lambda_1\bar{s}_1^{(2)}\sigma_{b_2^*}+2\bar{s}_1\bar{b}_2)],\\
\hat{e}^{(2)} = \; & \hat{J}+2\left(\tfrac{\lambda_2\left(c^{*(1)}(\lambda_1)v^*(\lambda_1)+c^*(\lambda_1)v^{*(1)}(\lambda_1)\right)}{c^*(\lambda_1)v^*(\lambda_1)}+\hat{e}^{(1)}\right)\hat{e}^{(1)}
\end{aligned}$$

where

$$\begin{aligned}
\hat{J} \; = \; & \{\hat{M}^{(2)}+\lambda_2^2 c^{*(2)}\left(\lambda_1\right)-\frac{2\lambda_2^2 c^{*(1)}(\lambda_1)(\bar{v}+\bar{s}_1)}{1-\lambda_1\bar{b}_1}+\frac{2\lambda_2^2\bar{s}_1\left(c^{*(1)}(\lambda_1)v^*(\lambda_1)+c^*(\lambda_1)v^{*(1)}(\lambda_1)\right)}{1-\lambda_1\bar{b}_1}+\\
& \frac{\lambda_2^2 c^*(\lambda_1)}{(1-\lambda_1\bar{b}_1)^2}[\bar{s}_1^{(2)}\lambda_1\sigma_{v^*}+2\bar{s}_1\bar{v}+\bar{v}^{(2)}+\frac{\lambda_1\bar{b}_1^{(2)}}{1-\lambda_1\bar{b}_1}\left(\bar{v}+\bar{s}_1\lambda_1\sigma_{v^*}\right)]\}/\left[c^*\left(\lambda_1\right)v^*\left(\lambda_1\right)\right]
\end{aligned}$$

$$\begin{aligned}
\hat{T}^{(2)} = \; & -\hat{K}^{(2)}c^*\left(\lambda_1\right)+2\lambda_2 c^{*(1)}\left(\lambda_1\right)\hat{K}^{(1)}-\lambda_2^2 c^{*(2)}\left(\lambda_1\right)\lambda_1\sigma_{s_2^*}-\hat{M}^{(2)},\\
\hat{U}^{(2)} = \; & \tfrac{\hat{K}^{(2)}-2\hat{T}^{(1)}\hat{U}^{(1)}-\hat{T}^{(2)}U}{c^*(\lambda_1)s_2^*(\lambda_1)}.
\end{aligned}$$

Let now for any function $w\left(0,z_1,z_2\right)$ denote $\tilde{w}^{(\nu)}=\frac{d^\nu}{dz_2^\nu}\,w\left(0,1,z_2\right)\big|_{z_2=1}$. Then

$$\begin{aligned}
\tilde{K}^{(1)} &= \lambda_1\left[\lambda_2\bar{s}_1\sigma_{s_2^*}+\sigma_{s_2^*}^{(1)}\right], \quad \tilde{K}^{(2)}=\lambda_1\left[\sigma_{s_2^*}^{(2)}+\lambda_2^2\bar{s}_1^{(2)}\sigma_{s_2^*}+2\lambda_2\bar{s}_1\sigma_{s_2^*}^{(1)}\right],\\
\tilde{M}^{(1)} &= \lambda_1\left[\lambda_2\bar{s}_3\sigma_{c^*}+\sigma_{c^*}^{(1)}\right], \quad \tilde{M}^{(2)}=\lambda_1\left[\sigma_{c^*}^{(2)}+\lambda_2^2\bar{s}_3^{(2)}\sigma_{c^*}+2\lambda_2\bar{s}_3\sigma_{c^*}^{(1)}\right],\\
\tilde{T}^{(1)} &= -\tilde{K}^{(1)}c^*\left(\lambda_1\right)+\lambda_2 c^{*(1)}\left(\lambda_1\right)\lambda_1\sigma_{s_2^*}-\tilde{M}^{(1)},\\
\tilde{T}^{(2)} &= -\tilde{K}^{(2)}c^*\left(\lambda_1\right)+2\lambda_2 c^{*(1)}\left(\lambda_1\right)\tilde{K}^{(1)}-\lambda_2^2 c^{*(2)}\left(\lambda_1\right)\lambda_1\sigma_{s_2^*}-\tilde{M}^{(2)},\\
\vec{R}^{(1)} &= -\frac{d}{dz_2}\vec{R}(0,1,z_2)|_{z_2=1}=\tilde{T}^{(1)}U+c^*(\lambda_1)s_2^*(\lambda_1)\hat{U}^{(1)}-\tilde{K}^{(1)},\\
\vec{R}^{(2)} &= -\frac{d^2}{dz_2^2}\vec{R}(0,1,z_2)|_{z_2=1}=\tilde{T}^{(2)}U+2\tilde{T}^{(1)}\hat{U}^{(1)}+c^*(\lambda_1)s_2^*(\lambda_1)\hat{U}^{(2)}-\tilde{K}^{(2)},
\end{aligned}$$

and finally

$$E\left(\left(N_2;\xi=b_1,u=1\right)\right)=\frac{H_1}{\lambda_2}\vec{R}^{(1)}+\frac{\vec{R}^{(2)}}{2}.$$

For the computation of $E\left((N_2;\xi=b_1,u=0)\right)$ we need finally the derivatives, at the point $z_2=1$, of the functions $L(0,1,z_2)$, $h_i(0,1,z_2)$, $\ i=1,2$, and $e_1(0,1,z_2)=1/e(0,1,z_2)$ defined in (8), (28) and (5) respectively. Thus

$$\begin{aligned}
\tilde{L}^{(1)} = & \quad -\lambda_2 s_2^{*(1)}(\lambda_1)+\lambda_2 s_2^*(\lambda_1)\left(\bar{b}_2+\bar{s}_1\lambda_1\sigma_{b_2^*}\right), \\
\tilde{L}^{(2)} = & \quad \lambda_2^2 s_2^{*(2)}(\lambda_1)+2\lambda_2^2\bar{s}_1\left(s_2^{*(1)}(\lambda_1)\, b_2^*(\lambda_1)+s_2^*(\lambda_1)\, b_2^{*(1)}(\lambda_1)\right) \\
& +\lambda_2^2 s_2^*(\lambda_1)\left(\bar{s}_1^{(2)}\lambda_1\sigma_{b_2^*}+2\bar{s}_1\bar{b}_2+\bar{b}_2^{(2)}\right)-2\lambda_2^2 s_2^{*(1)}(\lambda_1)\left(\bar{b}_2+\bar{s}_1\right), \\
R^{(2)} = & \quad \tfrac{d^2}{dz_2^2}R(0,z_2)|_{z_2=1}=\lambda_2^2 c^{*(2)}(\lambda_1)\,U-2\lambda_2 c^{*(1)}(\lambda_1)\,\hat{U}^{(1)}+c^*(\lambda_1)\,\hat{U}^{(2)}, \\
\tilde{e}_1^{(1)} = & \quad -\tfrac{\lambda_2}{\lambda_1}\left[\tfrac{\lambda_1\sigma_{v^*}(1+\lambda_1\bar{s}_3)+\lambda_1 c^*(\lambda_1)(\bar{v}+\bar{s}_1\lambda_1\sigma_{v^*})}{c^*(\lambda_1)v^*(\lambda_1)}\right], \\
\tilde{e}_1^{(2)} = & \quad 2\lambda_2\tfrac{\left[c^{*(1)}(\lambda_1)v^*(\lambda_1)+c^*(\lambda_1)v^{*(1)}(\lambda_1)\right]}{c^*(\lambda_1)v^*(\lambda_1)}\tilde{e}_1^{(1)}-(1-\lambda_1\bar{b}_1)\hat{J}+\tfrac{\lambda_1\lambda_2^2}{(1-\lambda_1\bar{b}_1)^2} \\
& \times\left[\tfrac{\bar{b}_1(1-\lambda_1\bar{b}_1)(\sigma_{c^*}(\lambda_1\bar{s}_3^{(2)}+2\bar{s}_3)+c^*(\lambda_1)(\bar{s}_1^{(2)}\lambda_1\sigma_{v^*}+2\bar{s}_1\bar{v}+\bar{v}^{(2)})}{c^*(\lambda_1)v^*(\lambda_1)}\right. \\
& \left.+\tfrac{\bar{b}_1^{(2)}(\sigma_{c^*}(1+\lambda_1\bar{s}_3)+c^*(\lambda_1)(\bar{v}+\lambda_1\sigma_{v^*}\bar{s}_1))}{c^*(\lambda_1)v^*(\lambda_1)}\right]
\end{aligned}$$

and

$$\begin{aligned}
\tilde{h}_1^{(2)} &= a\{\tilde{L}^{(2)}/s_2^*(\lambda_1)+2\tilde{L}^{(1)}R^{(1)}+s_2^*(\lambda_1)\tilde{R}^{(2)}-2\tilde{e}_1^{(1)}-\tilde{e}_1^{(2)}\}, \\
\tilde{h}_2^{(2)} &= \lambda_1\lambda_2^2\bar{s}_1^{(2)}-\lambda_1\tilde{e}_1^{(2)}+\tfrac{\lambda_2}{a}\tilde{h}_1^{(2)}, \\
\tilde{h}_1^{(1)} &= a\{\tilde{L}^{(1)}/s_2^*(\lambda_1)+s_2^*(\lambda_1)R^{(1)}-1-\tilde{e}_1^{(1)}\}, \\
\tilde{h}_2^{(1)} &= \lambda_1\lambda_2\bar{s}_1+\tfrac{\lambda_2}{a}\tilde{h}_1^{(1)}-\lambda_1\tilde{e}_1^{(1)}.
\end{aligned}$$

Then finally

$$\begin{aligned}
E\left(N_2;\xi=b_1,u=0\right) = & \quad -\tfrac{1}{2\lambda_2}\{2\tilde{h}_1^{(1)}\tfrac{1-\lambda_1\bar{b}_1}{2a(1-\rho)}\left[\lambda_2\rho_w^{(2)}+\lambda_1\rho_d^{(2)}Q^*(1)\right]+ \\
& \tilde{h}_2^{(2)}Q^*(1)+\left(\tilde{h}_1^{(2)}+2\tilde{h}_2^{(1)}\right)\tfrac{Q^*(1)}{a}\left(\tfrac{\lambda_1\bar{\rho}_d+\lambda_2(\bar{\rho}_c+\bar{\rho}_v)}{1-\rho}\right)\},
\end{aligned}$$

with

$$\rho_r=\frac{\lambda_2\sigma_{c^*}\left(1+\lambda_1\bar{s}_3\right)}{\left(1-\lambda_1\bar{b}_1\right)c^*\left(\lambda_1\right)},$$

$$\begin{aligned}
\rho_r^{(2)} = & \quad \tfrac{2\lambda_2^2[c^{*(1)}(\lambda_1)(1+\lambda_1\bar{s}_3)+c^*(\lambda_1)\sigma_{c^*}]}{\lambda_1\left(1-\lambda_1\bar{b}_1\right)(c^*(\lambda_1))^2}+\tfrac{\lambda_2^2\sigma_{c^*}(\lambda_1\bar{s}_3^{(2)}+2\bar{s}_3)}{c^*(\lambda_1)(1-\lambda_1\bar{b}_1)^2}+2\rho_r^2+\tfrac{\lambda_1\lambda_2\bar{b}_1^{(2)}\rho_r}{\left(1-\lambda_1\bar{b}_1\right)^2}, \\
\rho_d^{(2)} = & \quad \tfrac{\lambda_2^2\bar{b}_1^{(2)}}{\left(1-\lambda_1\bar{b}_1\right)^3}\left(1+\lambda_1\bar{s}_1\right)+2\tfrac{\lambda_2\left(\bar{b}_1+\bar{s}_1\right)}{\left(1-\lambda_1\bar{b}_1\right)}\hat{e}^{(1)}+\tfrac{\lambda_2^2}{(1-\lambda_1\bar{b}_1)^2}(\bar{s}_1^{(2)}+2\bar{b}_1\bar{s}_1)\ +\hat{e}^{(2)},
\end{aligned} \tag{44}$$

$$\begin{aligned}
\rho_w^{(2)} = & \quad \tfrac{1}{s_2^*(\lambda_1)}\{\ \hat{L}^{(2)}+\hat{K}^{(2)}+2(\hat{L}^{(1)}\hat{e}^{(1)}+\hat{K}^{(1)}\rho_r)+s_2^*(\lambda_1)\,\hat{e}^{(2)}+\lambda_1\sigma_{s_2^*}\rho_r^{(2)} \\
& +2\lambda_2\rho_w\{s_2^{*(1)}(\lambda_1)+\tfrac{\sigma_{s_2^*}[\lambda_1\sigma_{c^*}(1+\lambda_1\bar{s}_3)+c^*(\lambda_1)(1+\lambda_1\bar{s}_1)]}{\left(1-\lambda_1\bar{b}_1\right)c^*(\lambda_1)}\ \}\}.
\end{aligned} \tag{45}$$

8 Conclusions

In this paper a queuing model with two kind of customers, ordinary and retrial customers, is studied. To start serving both type of customers, the server needs a start up time, while when there are no customers waiting service, the server performs a close down period and in the sequel he departs for a single vacation. Upon discovering a Markov Renewal Process at particular time epochs, we describe our system as a Semi Regenerating Process and use the theory of Markov Renewal Processes to derive conditions for the system stability. Moreover, using the supplementary variable technique, we obtain expressions for the generating functions of the system state probabilities, both in a transient and in a steady state, and use them to derive expressions for the mean number of customers in the system, and the proportion of time the server remains in a particular stage (idle, busy, in start up, in close down, in vacation). Although the model is quite general containing a large number of arbitrarily distributed random variables, the obtained expressions are easily computable and can be directly used to produce numerical results and to compare system performance, under different values of the parameters.

References

[1] Y. Sakai, Y. Takahashi, Y. Takahashi, T. Hasegawa, A composite queue with vacation/set-up/close-down times for SVCC in IP over ATM networks, J. Oper. Res. Soc. Japan 41 (1998) 68-80.networks, J. Oper. Res. Soc. Japan 41 (1998) 68-80.

[2] Z. Niu, Y. Takahashi, A finite-capacity queue with exhaustive vacation/close-down/set-up times and Markovian arrival processes, Queueing Systems 31 (1999) 1-23.

[3] G.I. Falin, J.G.C. Templeton, *Retrial Queues* (Chapman and Hall, London, 1997).

[4] V.G. Kulkarni, H.M. Liang, Retrial Queues Revisited, in J.H. Dshalalow (ed), *Frontiers in Queueing* 19-34 (CRP Press, 1997).

[5] J.R. Artalejo, A classified bibliography of research on retrial queues: Progress in 1990-1999, Top 7(2) (1999) 187-211.

[6] C. Langaris, A. Katsaros, Time-depended analysis of a queue with batch arrivals and N-levels of non-preemptive priority, Queueing Systems 19 (1995) 269-288

[7] G.I. Falin, C. Fricker, On the virtual waiting time in an M/G/1 retrial queue, J. Appl. Probab. 28 (1991) 446-460.

[8] A.G. Pakes, Some conditions for ergodicity and recurrence of Markov chains, Oper. Res. 17 (1969) 1058-1061

[9] E. Cinlar, *Introduction to Stochastic Processes* (Prentice-Hall, 1975).

GLOBAL SOLUTIONS APPROACHING LINES AT INFINITY TO SECOND ORDER NONLINEAR DELAY DIFFERENTIAL EQUATIONS

CH. G. PHILOS, I. K. PURNARAS AND P. CH. TSAMATOS

ABSTRACT. This article is concerned with second order nonlinear delay, and especially ordinary, differential equations. By the use of the fixed point technique based on the classical Schauder's theorem, for any given line, sufficient conditions are established in order that there exists at least one global solution which is asymptotic at ∞ to this line. In the special case of ordinary differential equations, via the Banach's Contraction Principle, for any given line, conditions are presented which guarantee that there exists a unique global solution that is asymptotic at ∞ to this line. The application of the results obtained to second order delay, and ordinary, differential equations of Emden-Fowler type is presented, and some examples demonstrating the applicability of the results are given. Finally, some supplementary results are obtained, which provide sufficient conditions for all global solutions belonging to a suitable class to be asymptotic at ∞ to lines.

1. INTRODUCTION

In the asymptotic theory of delay, and especially of ordinary, differential equations, an interesting problem is that of the study of solutions with prescribed asymptotic behavior. This problem has been the subject of many investigations; we restrict ourselves to mention the recent papers [2], [5], [10–20] and [22–26] as well as the older classical articles [8, 9] (for a more extensive bibliography, see [17,18]). It is of special interest to investigate global solutions, i.e. solutions on the whole given interval, with prescribed asymptotic behavior. On this problem there is an extensive bibliography (see, for example, [2], [5], [8, 9], [11-16] and [22–26]; for more references, see [15,17,18]). The present work deals with global solutions that are asymptotic at ∞ to lines for second order delay, and especially ordinary, differential equations. For the basic theory of delay differential equations, the reader is referred to the books [3,4,6].

In [17], the authors considered n-th order $(n > 1)$ nonlinear ordinary differential equations and studied solutions that behave asymptotically like polynomials at ∞. More precisely, for each given integer m with $1 \leq m \leq n-1$, sufficient conditions have been presented in order that, for any real polynomial of degree at most m, there exists a solution which is asymptotic at ∞ to this polynomial. Conditions

1991 *Mathematics Subject Classification.* Primary 34K25, 34E05, 34E10; Secondary 34D05, 47H10.

Key words and phrases. Nonlinear differential equation, delay differential equation, ordinary differential equation, asymptotic behavior, asymptotic properties, asymptotic expansions, global solutions, asymptotic to lines solutions, fixed point theory.

have been also given, which are sufficient for every solution to be asymptotic at ∞ to a real polynomial of degree at most $n-1$. The application of the results in [17] to the special case of second order nonlinear ordinary differential equations leads to improved versions of the ones contained in the recent paper by Lipovan [10] and of other related results existing in the literature. Note that the nonlinear term, in the differential equations considered in [17], depends only on the time t and the unknown function x.

In a subsequent paper [18], the first and the third author investigated solutions approaching polynomials at ∞ to the more general case of n-th order $(n > 1)$ nonlinear ordinary differential equations, in which the nonlinear term depends on the time t and on $x, x', ..., x^{(N)}$, where x is the unknown function and N is an integer with $0 \leq N \leq n-1$. The results obtained in [18] extend those in [17] concerning the particular case where $N = 0$.

It must be noted that, in [17,18] (as well as in [10]), only nonlinear *ordinary* differential equations are considered and that, in these recent works, solutions defined for all large t, but not always *global*, are investigated.

In the present work, we deal with second order nonlinear delay differential equations, and especially ordinary differential equations, and we study global solutions that are asymptotic at ∞ to lines. More precisely, for any given line $\xi t + \eta$ (ξ and η are real constants), we establish sufficient conditions for the existence of at least one global solution x such that $x(t) = \xi t + \eta + o(1)$ and $x'(t) = \xi + o(1)$, for $t \to \infty$. In the special case of second order nonlinear ordinary differential equations, for any given line $\xi t + \eta$ (with $\xi, \eta \in \mathbf{R}$), we also present conditions guaranteeing the existence and uniqueness of a global solution x satisfying $x(t) = \xi t + \eta + o(1)$ and $x'(t) = \xi + o(1)$, for $t \to \infty$. Moreover, we apply our results to the case of second order delay, and especially ordinary, differential equations of Emden-Fowler type, and we give some examples in order to demonstrate the applicability of the results. Finally, we provide sufficient conditions for every global solution x that belongs to a suitable class to satisfy $x(t) = \xi t + \eta + o(1)$ and $x'(t) = \xi + o(1)$, for $t \to \infty$, where ξ and η are real constants (depending on the solution x).

It is an open question whether the results of the present paper can be extended to the more general case of n-th order $(n > 1)$ nonlinear delay, and especially ordinary, differential equations. For such differential equations, it is an open problem to investigate the existence (and the uniqueness, in the special case of ordinary differential equations) of global solutions that are asymptotic at ∞ to real polynomials of degree at most m, where m is a given integer with $1 \leq m \leq n-1$.

Throughout the paper, for any interval I of the real line $\mathbf{R}$ and any subset Ω of $\mathbf{R}$, by $C(I, \Omega)$ we will denote the set of all continuous functions defined on I and having values in Ω. Moreover, *r will be a nonnegative real constant.* Furthermore, if t is a point in the interval $[0, \infty)$ and χ is a continuous real-valued function defined at least on $[t-r, t]$, the notation χ_t will be used for the function in $C([-r, 0], \mathbf{R})$ defined by the formula

$$\chi_t(\tau) = \chi(t+\tau) \quad \text{for } -r \leq \tau \leq 0.$$

We notice that the set $C([-r, 0], \mathbf{R})$ is a Banach space endowed with the usual sup-norm $\|\cdot\|$:

$$\|\psi\| = \max_{-r \leq \tau \leq 0} |\psi(\tau)| \quad \text{for } \psi \in C([-r, 0], \mathbf{R}).$$

Consider the second order nonlinear delay differential equation

$$\text{(E)} \qquad x''(t) + f(t, x_t, x'(t)) = 0,$$

where f *is a continuous real-valued function defined on the set* $[0,\infty)\times C([-r,0],\mathbf{R})\times\mathbf{R}$.

Consider also, in particular, the second order nonlinear delay differential equation

$$(\mathrm{E}_0) \qquad x''(t) + f_0(t, x_t) = 0,$$

where f_0 *is a continuous real-valued function defined on the set* $[0,\infty)\times C([-r,0],\mathbf{R})$.

We are interested in solutions of the delay differential equations (E) and (E_0) on the *whole* interval $[0,\infty)$. By a *solution on* $[0,\infty)$ of (E) [respectively, of (E_0)], we mean a function x in $C([-r,\infty),\mathbf{R})$ which is twice continuously differentiable on the interval $[0,\infty)$ and satisfies (E) [resp., (E_0)] for all $t\geq 0$.

Furthermore, let us concentrate on a particular class of delay differential equations. More precisely, let us consider the second order nonlinear delay differential equation

$$(\mathrm{E}') \qquad x''(t) + g(t, x(t-T_1(t)), ..., x(t-T_m(t)), x'(t)) = 0$$

and, in particular, the second order nonlinear delay differential equation

$$(\mathrm{E}'_0) \qquad x''(t) + g_0(t, x(t-T_1(t)), ..., x(t-T_m(t))) = 0,$$

where m *is a positive integer,* g *is a continuous real-valued function on* $[0,\infty)\times\mathbf{R}^{m+1}$, g_0 *is a continuous real-valued function on* $[0,\infty)\times\mathbf{R}^m$, *and* T_j $(j=1,...,m)$ *are nonnegative continuous real-valued functions on the interval* $[0,\infty)$ *with*

$$\max_{j=1,\dots,m}\sup_{t\geq 0} T_j(t) = r.$$

If the delay differential equation (E) or (E_0) is to be equivalent to (E') or (E'_0), respectively, we must define

$$f(t,\psi,z) = g(t, \psi(-T_1(t)), ..., \psi(-T_m(t)), z) \qquad \text{for any } (t,\psi,z)\in[0,\infty)\times C([-r,0],\mathbf{R})\times\mathbf{R}$$

or

$$f_0(t,\psi) = g_0(t, \psi(-T_1(t)), ..., \psi(-T_m(t))) \quad \text{for any } (t,\psi)\in[0,\infty)\times C([-r,0],\mathbf{R}),$$

respectively.

We restrict our attention only to solutions of the delay differential equations (E') and (E'_0) on the *whole* interval $[0,\infty)$. A *solution on* $[0,\infty)$ of (E') [resp., of (E'_0) is a function x in $C([-r,\infty),\mathbf{R})$, which is twice continuously differentiable on the interval $[0,\infty)$ and satisfies (E') [resp., (E'_0)] for all $t\geq 0$.

Now, let us consider the special case of *ordinary* differential equations. That is, consider the second order nonlinear ordinary differential equation

$$\text{(D)} \qquad x''(t) + h(t, x(t), x'(t)) = 0$$

and, especially, the second order nonlinear ordinary differential equation

$$(\mathrm{D}_0) \qquad x''(t) + h_0(t, x(t)) = 0,$$

where h *is a continuous real-valued function on* $[0,\infty)\times\mathbf{R}^2$, *and* h_0 *is is a continuous real-valued function on* $[0,\infty)\times\mathbf{R}$.

We confine our discussion only to solutions of the differential equations (D) and (D_0) on the *whole* interval $[0,\infty)$.

The results of the paper are stated in Section 2, while their proofs are given in Sections 3–5. Section 6 is devoted to the application of the results to second order (delay or, especially, ordinary) differential equations of Emden-Fowler type as well as to some examples demonstrating the applicability of our results. In the last section (Section 7) some supplementary results are given, which can be characterized as a complement of the results of the present work.

2. STATEMENT OF THE RESULTS

Our results in this paper are presented in the form of four theorems (Theorems 1–4) and four Corollaries (Corollaries 1–4). In Theorem 1 (respectively, Theorem 2), for given real constants ξ and η, sufficient conditions are established in order that the delay differential equation (E) [resp., ($\mathrm{E_0}$)] have at least one solution x on the interval $[0,\infty)$ such that $x(t) = \xi t + \eta + o(1)$ and $x'(t) = \xi + o(1)$, for $t \to \infty$. Corollary 1 (resp., Corollary 2) is the application of Theorem 1 (resp., Theorem 2) to the particular case of the delay differential equation (E$'$) [resp., (E'_0)], while Corollary 3 (resp., Corollary 4) is the specialization of Theorem 1 (resp., Theorem 2) to the ordinary differential equation (D) [resp., ($\mathrm{D_0}$)]. In Theorem 3 (resp., Theorem 4), for given real constants ξ and η, conditions are presented, which are suffficient for the ordinary differential equation (D) [resp., ($\mathrm{D_0}$)] to have exactly one solution x on the interval $[0,\infty)$ such that $x(t) = \xi t + \eta + o(1)$ and $x'(t) = \xi + o(1)$, for $t \to \infty$.

Theorem 1. *Assume that*

$$|f(t,\psi,z)| \le F(t,|\psi|,|z|) \quad \textit{for all } (t,\psi,z) \in [0,\infty) \times C([-r,0],\mathbf{R}) \times \mathbf{R}, \tag{2.1}$$

where F is a nonnegative real-valued function defined on $[0,\infty)\times C([-r,0],[0,\infty))\times [0,\infty)$, which satisfies the Continuity Condition:

(C) *$F(t,|\chi_t|,|\chi'(t)|)$ is continuous with respect to t in $[0,\infty)$ for each given function χ in $C([-r,\infty),\mathbf{R})$ which is continuously differentiable on the interval $[0,\infty)$.*

Suppose that:

(B) *For each $t \ge 0$, the function $F(t,\cdot,\cdot)$ is increasing on $C([-r,0],[0,\infty)) \times [0,\infty)$ in the sense that $F(t,\psi,z) \le F(t,\omega,v)$ for any ψ, ω in $C([-r,0],[0,\infty))$ with $\psi \le \omega$ (i.e. $\psi(\tau) \le \omega(\tau)$ for $-r \le \tau \le 0$) and any z, v in $[0,\infty)$ with $z \le v$.*

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|,|\eta|\}$ so that

$$\int_0^\infty tF(t,\gamma_t,c)dt \le c - |\eta| \tag{2.2}$$

and

$$\int_0^\infty F(t,\gamma_t,c)dt \le c - |\xi|, \tag{2.3}$$

where the function γ in $C([-r,\infty),[0,\infty))$ depends on c and is defined by

$$\gamma(t) = \begin{cases} c & \textit{for } -r \le t \le 0 \\ c(t+1) & \textit{for } t \ge 0. \end{cases} \tag{2.4}$$

Then the delay differential equation (E) *has at least one solution x on the interval* $[0,\infty)$ *such that*

$$x(t) = \xi t + \eta + o(1) \quad \text{for } t \to \infty \tag{2.5}$$

and

$$x'(t) = \xi + o(1) \quad \text{for } t \to \infty; \tag{2.6}$$

in addition, this solution x satisfies

$$x(t) = x(0) \quad \text{for } -r \le t \le 0, \tag{2.7}$$

$$\xi t + \eta - (c - |\eta|) \le x(t) \le \xi t + \eta + (c - |\eta|) \quad \text{for every } t \ge 0 \tag{2.8}$$

and

$$\xi - (c - |\xi|) \le x'(t) \le \xi + (c - |\xi|) \quad \text{for every } t \ge 0. \tag{2.9}$$

Theorem 2. *Assume that*

$$|f_0(t,\psi)| \le F_0(t,|\psi|) \quad \text{for all } (t,\psi) \in [0,\infty) \times C([-r,0],\mathbf{R}), \tag{2.10}$$

where F_0 is a nonnegative real-valued function defined on $[0,\infty)\times C([-r,0],[0,\infty))$, *which satisfies the Continuity Condition:*

(C_0) *$F_0(t,|\chi_t|)$ is continuous with respect to t in $[0,\infty)$ for each given function χ in* $C([-r,\infty),\mathbf{R})$.

Suppose that:

(B_0) *For each $t \ge 0$, the function $F_0(t,\cdot)$ is increasing on $C([-r,0],[0,\infty))$ in the sense that $F_0(t,\psi) \le F_0(t,\omega)$ for any ψ, ω in $C([-r,0],[0,\infty))$ with $\psi \le \omega$ (i.e. $\psi(\tau) \le \omega(\tau)$ for $-r \le \tau \le 0$).*

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|,|\eta|\}$ so that

$$\int_0^\infty t F_0(t,\gamma_t)dt \le c - \max\{|\xi|,|\eta|\}, \tag{2.11}$$

where the function γ in $C([-r,\infty),[0,\infty))$ depends on c and is defined by (2.4). *Then the delay differential equation* (E_0) *has at least one solution x on the interval* $[0,\infty)$ *such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.7) *and:*

$$\xi t + \eta - (c - \max\{|\xi|,|\eta|\}) \le x(t) \le \xi t + \eta + (c - \max\{|\xi|,|\eta|\}) \quad \text{for every } t \ge 0 \tag{2.12}$$

and

$$\xi - \int_0^\infty F_0(s,\gamma_s)ds \le x'(t) \le \xi + \int_0^\infty F_0(s,\gamma_s)ds \quad \text{for every } t \ge 0. \tag{2.13}$$

(*Note that, because of* (2.11), *$\int_0^\infty F_0(s,\gamma_s)ds$ is finite.*)

Corollary 1. *Assume that*

$$|g(t,y_1,\dots,y_m,z)| \le G(t,|y_1|,\dots,|y_m|,|z|) \quad \text{for } (t,y_1,\dots,y_m,z) \in [0,\infty) \times \mathbf{R}^{m+1},$$

where G is a nonnegative continuous real-valued function on $[0,\infty) \times [0,\infty)^{m+1}$. Suppose that:

(B$'$) *For each* $t \geq 0$, *the function* $G(t, \cdot, ..., \cdot, \cdot)$ *is increasing on* $[0, \infty)^{m+1}$ *in the sense that* $G(t, y_1, ..., y_m, z) \leq G(t, w_1, ..., w_m, v)$ *for any* $(y_1, ..., y_m, z)$, $(w_1, ..., w_m, v)$ *in* $[0, \infty)^{m+1}$ *with* $y_1 \leq w_1, ..., y_m \leq w_m$, $z \leq v$.

Let ξ *and* η *be given real constants, and let there exist a real number* c *with* $c > \max\{|\xi|, |\eta|\}$ *so that*

$$\int_0^\infty tG(t, \rho_1(t), ..., \rho_m(t), c)dt \leq c - |\eta|$$

and

$$\int_0^\infty G(t, \rho_1(t), ..., \rho_m(t), c)dt \leq c - |\xi|,$$

where, for each $j \in \{1, ..., m\}$, *the function* ρ_j *in* $C([0, \infty), [0, \infty))$ *depends on* c *and is defined by*

$$\rho_j(t) = \begin{cases} c, & \text{if } 0 \leq t \leq T_j(t) \\ c(t - T_j(t) + 1), & \text{if } t \geq T_j(t). \end{cases} \tag{2.14}$$

Then the delay differential equation (E$'$) *has at least one solution* x *on the interval* $[0, \infty)$ *such that* (2.5) *and* (2.6) *hold; in addition, this solution* x *satisfies* (2.7), (2.8) *and* (2.9).

Corollary 2. *Assume that*

$$|g_0(t, y_1, ..., y_m)| \leq G_0(t, |y_1|, ..., |y_m|) \quad \text{for } (t, y_1, ..., y_m) \in [0, \infty) \times \mathbf{R}^m,$$

where G_0 *is a nonnegative continuous real-valued function on* $[0, \infty) \times [0, \infty)^m$. *Suppose that:*

(B$'_0$) *For each* $t \geq 0$, *the function* $G_0(t, \cdot, ..., \cdot)$ *is increasing on* $[0, \infty)^m$ *in the sense that* $G_0(t, y_1, ..., y_m) \leq G_0(t, w_1, ..., w_m)$ *for any* $(y_1, ..., y_m)$, $(w_1, ..., w_m)$ *in* $[0, \infty)^m$ *with* $y_1 \leq w_1, ..., y_m \leq w_m$.

Let ξ *and* η *be given real constants, and let there exist a real number* c *with* $c > \max\{|\xi|, |\eta|\}$ *so that*

$$\int_0^\infty tG_0(t, \rho_1(t), ..., \rho_m(t))dt \leq c - \max\{|\xi|, |\eta|\}, \tag{2.15}$$

where, for each $j \in \{1, ..., m\}$, *the function* ρ_j *in* $C([0, \infty), [0, \infty))$ *depends on* c *and is defined by* (2.14). *Then the delay differential equation* (E$'_0$) *has at least one solution* x *on the interval* $[0, \infty)$ *such that* (2.5) *and* (2.6) *hold; in addition, this solution* x *satisfies* (2.7), (2.12), *and*

$$\xi - \int_0^\infty G_0(s, \rho_1(s), ..., \rho_m(s))ds \leq x'(t) \leq \xi + \int_0^\infty G_0(s, \rho_1(s), ..., \rho_m(s))ds$$

for every $t \geq 0$.

(*Note that, because of* (2.15), $\int_0^\infty G_0(s, \rho_1(s), ..., \rho_m(s))ds$ *is finite.*)

Corollary 3. *Assume that*

$$|h(t, y, z)| \leq H(t, |y|, |z|) \quad \text{for all } (t, y, z) \in [0, \infty) \times \mathbf{R}^2, \tag{2.16}$$

where H *is a nonnegative continuous real-valued function on* $[0, \infty) \times [0, \infty)^2$. *Suppose that:*

(A) *For each* $t \geq 0$, *the function* $H(t, \cdot, \cdot)$ *is increasing on* $[0, \infty)^2$ *in the sense that* $H(t, y, z) \leq H(t, w, v)$ *for any* (y, z), (w, v) *in* $[0, \infty)^2$ *with* $y \leq w$, $z \leq v$.

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that

$$\int_0^\infty tH(t, c(t+1), c)dt \le c - |\eta| \tag{2.17}$$

and

$$\int_0^\infty H(t, c(t+1), c)dt \le c - |\xi|. \tag{2.18}$$

Then the ordinary differential equation (D) *has at least one solution x on the interval $[0, \infty)$ such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.8) *and* (2.9).

Corollary 4. *Assume that*

$$|h_0(t, y)| \le H_0(t, |y|) \quad \textit{for all } (t, y) \in [0, \infty) \times \mathbf{R}, \tag{2.19}$$

where H_0 is a nonnegative continuous real-valued function on $[0, \infty) \times [0, \infty)$. Suppose that:

(A_0) *For each $t \ge 0$, the function $H_0(t, \cdot)$ is increasing on $[0, \infty)$ in the sense that $H_0(t, y) \le H_0(t, w)$ for any y, w in $[0, \infty)$ with $y \le w$.*

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that

$$\int_0^\infty tH_0(t, c(t+1))dt \le c - \max\{|\xi|, |\eta|\}. \tag{2.20}$$

Then the ordinary differential equation (D_0) *has at least one solution x on the interval $[0, \infty)$ such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.12) *and*

$$\xi - \int_0^\infty H_0(s, c(s+1))ds \le x'(t) \le \xi + \int_0^\infty H_0(s, c(s+1))ds \quad \textit{for every } t \ge 0. \tag{2.21}$$

(*Note that, because of* (2.20), $\int_0^\infty H_0(s, c(s+1))ds$ *is finite.*)

Theorem 3. *Let the following generalized Lipschitz condition be satisfied:*

$$|h(t, y, z) - h(t, w, v)| \le L(t) \max\{|y - w|, |z - v|\} \quad \textit{for all } (t, y, z), (t, w, v) \textit{ in } [0, \infty) \times \mathbf{R}^2, \tag{2.22}$$

where L is a nonnegative continuous real-valued function on the interval $[0, \infty)$ such that

$$\max\left\{\int_0^\infty t(t+1)L(t)dt, \ \int_0^\infty (t+1)L(t)dt\right\} < 1. \tag{2.23}$$

Moreover, assume that (2.16) *holds, where H is a nonnegative continuous real-valued function on $[0, \infty) \times [0, \infty)^2$. Suppose that* (A) *is satisfied.*

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that (2.17) *and* (2.18) *hold. Then the ordinary differential equation* (D) *has exactly one solution x on the interval $[0, \infty)$ with*

$$|x(0)| \le c \tag{2.24}$$

and

$$|x'(t)| \le c \quad \textit{for every } t \ge 0, \tag{2.25}$$

such that (2.5) *and* (2.6) *hold; in addition, this unique solution x satisfies* (2.8) *and* (2.9).

Theorem 4. *Let the following generalized Lipschitz condition be satisfied:*

$$|h_0(t,y) - h_0(t,w)| \le L_0(t)\,|y-w| \quad \textit{for all } (t,y),(t,w) \textit{ in } [0,\infty)\times \mathbf{R}, \tag{2.26}$$

where L_0 is a nonnegative continuous real-valued function on the interval $[0,\infty)$ such that

$$\int_0^\infty t(t+1)L_0(t)dt < 1. \tag{2.27}$$

Moreover, assume that (2.19) *holds, where H_0 is a nonnegative continuous real-valued function on $[0,\infty)\times[0,\infty)$. Suppose that* (A_0) *is satisfied.*

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|,|\eta|\}$ so that (2.20) *holds. Then the ordinary differential equation* (D_0) *has exactly one solution x on the interval $[0,\infty)$ with*

$$|x(t)| \le c(t+1) \quad \textit{for every } t \ge 0, \tag{2.28}$$

and such that (2.5) *and* (2.6) *hold; in addition, this unique solution x satisfies* (2.12) *and* (2.21).

Note: Inequalities (2.24) and (2.25) imply (2.28).

An important remark. (i) In the conclusions of Theorems 1 and 3 and of Corollaries 1 and 3, the solution x satisfies (2.8) and (2.9).

Assume that $\xi > 0$ and $\eta > 0$. Then (2.8) and (2.9) are written as

$$\xi t - (c - 2\eta) \le x(t) \le \xi t + c \quad \text{for every } t \ge 0 \tag{2.8$'$}$$

and

$$-(c - 2\xi) \le x'(t) \le c \quad \text{for every } t \ge 0, \tag{2.9$'$}$$

respectively. Furthermore, *in addition to the hypothesis $c > \xi$ and $c > \eta$, let us suppose that $c < 2\xi$ and $c \le 2\eta$. We have thus $0 < \xi < c < 2\xi$ and $0 < \eta < c \le 2\eta$.* Then (2.8$'$) guarantees that *the solution x is positive on the interval $(0,\infty)$ and such that $\lim_{t\to\infty} x(t) = \infty$* . Also, from (2.9$'$) it follows that $x'(t) > 0$ for $t \ge 0$ and so *x is strictly increasing on the interval $[0,\infty)$.*

Analogously, *in the case where $2\xi < -c < \xi < 0$ and $2\eta \le -c < \eta < 0$,* we can see that *the solution x is negative on the interval $(0,\infty)$ and such that $\lim_{t\to\infty} x(t) = -\infty$*, and that *$x$ is strictly decreasing on $[0,\infty)$.*

(ii) The solution x in the conclusion of Theorem 2 is such that (2.12) and (2.13) are satisfied. (Analogous inequalities are fulfilled for the solution x in the conclusions of Corollaries 2 and 4, and of Theorem 4).

Let ξ and η be positive. Then (2.12) becomes

$$\xi t - [c - (\eta + \max\{\xi,\eta\})] \le x(t) \le \xi t + [c - (-\eta + \max\{\xi,\eta\})] \quad \text{for every } t \ge 0. \tag{2.12$'$}$$

We have assumed that $c > \xi$ and $c > \eta$. *In addition to this assumption, let us suppose that $\xi > \int_0^\infty F_0(s,\gamma_s)ds$ and $c \le \eta + \max\{\xi,\eta\}$.* So, *we have* $0 \le$

$\int_0^\infty F_0(s,\gamma_s)ds < \xi < c$ *and* $0 < \eta < c \leq \eta + \max\{\xi, \eta\}$. It follows from (2.12′) that *the solution* x *is positive on the interval* $(0,\infty)$ *and satisfies* $\lim\limits_{t\to\infty} x(t) = \infty$. Moreover, (2.13) ensures that $x'(t) > 0$ for $t \geq 0$ and consequently x *is strictly increasing on the interval* $[0,\infty)$.

In a similar way, we can conclude that, *if* $-c < \xi < -\int_0^\infty F_0(s,\gamma_s)ds \leq 0$ *and* $\eta + \max\{\xi,\eta\} \leq -c < \eta < 0$, *then the solution* x *is negative on the interval* $(0,\infty)$ *with* $\lim\limits_{t\to\infty} x(t) = -\infty$, *and strictly decreasing on* $[0,\infty)$.

Before closing this section, we must point out the connection between Theorem 1 and Theorem 2. It is obvious that *Theorem 1 concerning the delay differential equation* (E) *is also applicable to the particular case of the delay differential equation* (E_0). It is remarkable that the result obtained by such an application is different from Theorem 2 dealing with the delay differential equation (E_0). As it is evident, the conclusion of Theorem 2 cannot be derived from the conclusion of Theorem 1. What is more, the spaces on which Schauder's theorem is applied in the proofs of these two theorems are different one another. Therefore, the proofs themselves are significantly different. Example 7 at the end of Section 6 illustrates the difference between the conlcusion deduced by Theorem 1 and the conlcusion deduced by Theorem 2.

Analogous remarks can be made for the connection between Corollary 1 and Corollary 2, between Corollary 3 and Corollary 4, and between Theorem 3 and Theorem 4.

3. PROOF OF THEOREM 1

The proof of Theorem 1 is based on the use of the following Schauder's fixed point theorem (see Schauder [21]).

The Schauder theorem. *Let* S *be a Banach space and* X *any nonempty convex and closed subset of* S. *If* M *is a continuous mapping of* X *into itself and* MX *is relatively compact, then the mapping* M *has at least one fixed point (i.e. there exists an* $x \in X$ *with* $x = Mx$).

Let $BC([0,\infty),\mathbf{R})$ be the Banach space of all bounded continuous real-valued functions on the interval $[0,\infty)$, endowed with the sup-norm $\|\cdot\|$ defined by

$$\|u\| = \sup_{t\geq 0} |u(t)| \quad \text{for } u \in BC([0,\infty),\mathbf{R}).$$

We need the following compactness criterion for subsets of $BC([0,\infty),\mathbf{R})$, which is a consequence of the well-known Arzelà-Ascoli theorem. This compactness criterion is an adaptation of a lemma due to Avramescu [1].

Compactness criterion. *Let* U *be an equicontinuous and uniformly bounded subset of the Banach space* $BC([0,\infty),\mathbf{R})$. *If* U *is equiconvergent at* ∞, *it is also relatively compact.*

Note that a set U of real-valued functions defined on the interval $[0,\infty)$ is called *equiconvergent at* ∞ if all functions in U are convergent in $\mathbf{R}$ at the point ∞ and, in addition, for each $\epsilon > 0$, there exists $T = T(\epsilon) > 0$ such that, for all functions u in U, it holds $\left|u(t) - \lim_{s\to\infty} u(s)\right| < \epsilon$ for every $t \geq T$.

Throughout the ramainder of this section, by S we will denote the set of all functions in $C([-r,\infty),\mathbf{R})$, which have bounded continuous derivatives on the interval $[0,\infty)$. The set S *is a Banach space endowed with the norm* $\|\cdot\|$ defined as follows

$$\| u \| = \max\left\{\max_{-r\leq t\leq 0}|u(t)|\,,\ \sup_{t\geq 0}|u'(t)|\right\} \quad \text{for } u\in S.$$

To prove Theorem 1, we first establish the following proposition.

Proposition 1. *Assume that* (2.1) *holds, where* F *is a nonnegative real-valued function defined on* $[0,\infty)\times C([-r,0],[0,\infty))\times[0,\infty)$, *which satisfies the Continuity Condition* (C). *Suppose that* (B) *is satisfied.*

Let ξ *and* η *be given real constants, and let* c *be a positive real number such that*

$$\int_0^\infty tF(t,\gamma_t,c)dt < \infty, \tag{3.1}$$

where the function γ *in* $C([-r,\infty),[0,\infty))$ *depends on* c *and is defined by* (2.4). *Let also* X *be the subset of* S *defined by*

$$X = \{x\in S:\ \| x \| \leq c\}. \tag{3.2}$$

Then the formula

$$(Mx)(t) = \begin{cases} \eta - \int_0^\infty sf(s,x_s,x'(s))ds & \text{for } -r\leq t\leq 0 \\ \xi t + \eta - \int_t^\infty (s-t)f(s,x_s,x'(s))ds & \text{for } t\geq 0 \end{cases} \tag{3.3}$$

makes sense for any function x *in* X, *and this formula defines a continuous mapping* M *of* X *into* S *such that* MX *is relatively compact.*

Proof of Proposition 1. Let x be an arbitrary function in X. From the definition of X, via (3.2), it follows that

$$|x(t)|\leq c \ \text{ for } -r\leq t\leq 0 \tag{3.4}$$

and

$$|x'(t)|\leq c \ \text{ for every } t\geq 0. \tag{3.5}$$

Inequality (3.4) gives, in particular, $|x(0)|\leq c$. So, by using this fact and (3.5), we obtain for $t\geq 0$

$$|x(t)| = \left|x(0) + \int_0^t x'(s)ds\right| \leq |x(0)| + \int_0^t |x'(s)|\,ds \leq c + ct,$$

i.e. we have

$$|x(t)|\leq c(t+1) \quad \text{for every } t\geq 0. \tag{3.6}$$

In view of (2.4), from (3.4) and (3.6) we conclude that

$$|x(t)|\leq \gamma(t) \ \text{ for } t\geq -r$$

and consequently

$$|x_t|\leq \gamma_t \ \text{ for all } t\geq 0. \tag{3.7}$$

By taking into account (3.7) and (3.5) and using the assumption (B), we get

$$F(t, |x_t|, |x'(t)|) \leq F(t, \gamma_t, c) \quad \text{for } t \geq 0.$$

On the other hand, because of (2.1), it holds

$$|f(t, x_t, x'(t))| \leq F(t, |x_t|, |x'(t)|) \quad \text{for } t \geq 0.$$

Thus, we find

$$|f(t, x_t, x'(t))| \leq F(t, \gamma_t, c) \quad \text{for every } t \geq 0. \tag{3.8}$$

Furthermore, by combining (3.1) and (3.8), we have

$$\int_0^\infty t\,|f(t, x_t, x'(t))|\,dt < \infty. \tag{3.9}$$

This, in particular, implies

$$\int_0^\infty |f(t, x_t, x'(t))|\,dt < \infty. \tag{3.10}$$

So, in view of (3.9) and (3.10), it is true that

$$\int_0^\infty tf(t, x_t, x'(t))dt \quad \text{and} \quad \int_0^\infty f(t, x_t, x'(t))dt \qquad \text{exist in } \mathbf{R}. \tag{3.11}$$

As (3.11) holds true for all functions x in X, we can immediately see that *the formula* (3.3) *makes sense for any function* x *in* X, *and this formula defines a mapping* M *of* X into $C([-r, \infty), \mathbf{R})$. We will show that M is a mapping of X *into* S, i.e. that $MX \subseteq S$. To this end, let us consider an arbitrary function x in X. Then, by taking into account (3.8), from (3.3) we obtain for $t \geq 0$

$$\begin{aligned} |(Mx)'(t)| &= \left|\xi + \int_t^\infty f(s, x_s, x'(s))ds\right| \leq |\xi| + \int_t^\infty |f(s, x_s, x'(s))|\,ds \\ &\leq |\xi| + \int_t^\infty F(s, \gamma_s, c)ds \leq |\xi| + \int_0^\infty F(s, \gamma_s, c)ds. \end{aligned}$$

Therefore,

$$|(Mx)'(t)| \leq Q \quad \text{for all } t \geq 0, \tag{3.12}$$

where

$$Q = |\xi| + \int_0^\infty F(s, \gamma_s, c)ds. \tag{3.13}$$

Note that (3.1) guarantees, in particular, that

$$\int_0^\infty F(t, \gamma_t, c)dt < \infty \tag{3.14}$$

and so Q is a nonnegative real constant. Inequality (3.12) means that $(Mx)'$ is always bounded on the interval $[0, \infty)$, and consequently Mx belongs to S. We have thus proved that, for any function x in X, $Mx \in S$, i.e. that $MX \subseteq S$.

Now, we shall prove that MX *is relatively compact.* From (3.3) it follows that, for each $x \in X$, the function Mx is constant on the interval $[-r, 0]$. By taking into account this fact as well as the definition of the norm $\nparallel \cdot \nparallel$, we can easily conclude that it suffices to prove that the set

$$U = \{((Mx) \mid [0, \infty))' : \ x \in X\}$$

is relatively compact in the Banach space $BC([0,\infty),\mathbf{R})$. Each function x in X satisfies (3.12), where the nonnegative real number Q is defined by (3.13) (and it is independent of x). This ensures that U is uniformly bounded. Furthermore, for any function x in X, it follows from (3.3) that

$$|(Mx)'(t)-\xi| = \left|\int_t^\infty f(s,x_s,x'(s))ds\right| \le \int_t^\infty |f(s,x_s,x'(s))|\,ds$$

for all $t \ge 0$, and consequently, by taking into account (3.8), we derive

$$|(Mx)'(t)-\xi| \le \int_t^\infty F(s,\gamma_s,c)ds \quad \text{for every } t \ge 0. \tag{3.15}$$

For any function x in X, (3.15) together with (3.14) imply that

$$\lim_{t\to\infty}(Mx)'(t) = \xi.$$

By using again (3.14) and (3.15), we immediately see that U is equiconvergent at ∞. Now, by (3.8), for any function x in X and every t_1, t_2 with $0 \le t_1 \le t_2$, from (3.3) we obtain

$$\begin{aligned}
&|(Mx)'(t_1)-(Mx)'(t_2)| \\
&= \left|\left[\xi + \int_{t_1}^\infty f(s,x_s,x'(s))ds\right] - \left[\xi + \int_{t_2}^\infty f(s,x_s,x'(s))ds\right]\right| \\
&= \left|\int_{t_1}^{t_2} f(s,x_s,x'(s))ds\right| \le \int_{t_1}^{t_2} |f(s,x_s,x'(s))|\,ds \\
&\le \int_{t_1}^{t_2} F(s,\gamma_s,c)ds.
\end{aligned}$$

Thus, by virtue of (3.14), it is easy to verify that U is equicontinuous. By the given compactness criterion, the set U is relatively compact in $BC([0,\infty),\mathbf{R})$. Hence, the relative compactness of MX (in S) has been established.

Next, we will show that *the mapping M is continuous.* For this purpose, let us consider an arbitrary function x in X and a sequence $\left(x^{[\nu]}\right)_{\nu\ge1}$ of functions in X with

$$\|\!\!\!\cdot\!\!\!\| - \lim_{\nu\to\infty} x^{[\nu]} = x.$$

It is not difficult to verify that

$$\lim_{\nu\to\infty} x^{[\nu]}(t) = x(t) \quad \text{uniformly in } t \in [-r,\infty)$$

and

$$\lim_{\nu\to\infty} (x^{[\nu]})'(t) = x'(t) \quad \text{uniformly in } t \in [0,\infty).$$

On the other hand, by (3.8), it holds

$$\left|f(t,x_t^{[\nu]},(x^{[\nu]})'(t))\right| \le F(t,\gamma_t,c) \quad \text{for every } t \ge 0 \quad \text{and} \quad \text{for all } \nu \ge 1.$$

Thus, because of (3.1) and (3.14), one can apply the Lebesgue dominated convergence theorem to obtain, for $t \ge 0$,

$$\lim_{\nu\to\infty}\int_t^\infty (s-t)f(s,x_s^{[\nu]},(x^{[\nu]})'(s))ds = \int_t^\infty (s-t)f(s,x_s,x'(s))ds.$$

So, from (3.3) it follows that

$$\lim_{\nu\to\infty}(Mx^{[\nu]})(t) = (Mx)(t) \quad \text{for } t \ge -r.$$

It remains to establish that this pointwise convergence is also $\nparallel \cdot \nparallel -$convergence, i.e. that

$$\nparallel \cdot \nparallel - \lim_{\nu\to\infty} Mx^{[\nu]} = Mx. \tag{3.16}$$

To this end, we consider an arbitrary subsequence $\left(Mx^{[\mu_\nu]}\right)_{\nu\geq 1}$ of $\left(Mx^{[\nu]}\right)_{\nu\geq 1}$. Since the set MX is relatively compact, there exist a subsequence $\left(Mx^{[\mu_{\lambda_\nu}]}\right)_{\nu\geq 1}$ of $\left(Mx^{[\mu_\nu]}\right)_{\nu\geq 1}$ and a function u in S so that

$$\nparallel \cdot \nparallel - \lim_{\nu\to\infty} Mx^{[\mu_{\lambda_\nu}]} = u.$$

As the $\nparallel \cdot \nparallel -$convergence implies the pointwise convergence to the same limit function, we must have $u = Mx$. That is, (3.16) holds true. Consequently, M is continuous.

The proof of the proposition has been completed.

Now, we proceed to the proof of Theorem 1.

Proof of Theorem 1. Let X be defined by (3.2). Clearly, X *is a nonempty convex and closed subset of* S. Assumption (2.2) guarantees, in particular, that (3.1) holds. So, by Proposition 1, the formula (3.3) makes sense for any function x in X, and this formula defines *a continuous mapping M of X into S such that MX is relatively compact.* We shall prove that M is a mapping of X *into itself,* i.e. that $MX \subseteq X$. Let us consider an arbitrary function x in X. Then, by taking into account (3.8), from (3.3) we obtain, for $-r \leq t \leq 0$,

$$\begin{aligned} |(Mx)(t) - \eta| &= \left| -\int_0^\infty sf(s, x_s, x'(s))ds \right| \leq \int_0^\infty s\,|f(s, x_s, x'(s))|\,ds \\ &\leq \int_0^\infty sF(s, \gamma_s, c)ds \end{aligned}$$

and consequently, in view of (2.2), we find

$$|(Mx)(t) - \eta| \leq c - |\eta| \quad \text{for } -r \leq t \leq 0. \tag{3.17}$$

Moreover, by using again (3.8), from (3.3) we derive for $t \geq 0$

$$\begin{aligned} |(Mx)'(t) - \xi| &= \left| \int_t^\infty f(s, x_s, x'(s))ds \right| \leq \int_t^\infty |f(s, x_s, x'(s))|\,ds \\ &\leq \int_t^\infty F(s, \gamma_s, c)ds \leq \int_0^\infty F(s, \gamma_s, c)ds \end{aligned}$$

and so, by (2.3), we get

$$|(Mx)'(t) - \xi| \leq c - |\xi| \quad \text{for every } t \geq 0. \tag{3.18}$$

Inequalities (3.17) and (3.18) give

$$|(Mx)(t)| \leq c \ \text{ for } -r \leq t \leq 0$$

and

$$|(Mx)'(t)| \leq c \ \text{ for every } t \geq 0,$$

respectively. From the last two inequalities it follows that $\nparallel Mx \nparallel \leq c$, which means that Mx belongs to X. We have thus proved that $Mx \in X$ for each $x \in X$, i.e. that $MX \subseteq X$.

Now, we apply the Schauder theorem to conclude that there exists at least one x in X with $x = Mx$, i.e.

$$x(t) = \begin{cases} \eta - \int_0^\infty sf(s, x_s, x'(s))ds & \text{for } -r \le t \le 0 \\ \xi t + \eta - \int_t^\infty (s-t) f(s, x_s, x'(s))ds & \text{for } t \ge 0. \end{cases} \tag{3.19}$$

From (3.19) we immediately obtain

$$x''(t) = -f(t, x_t, x'(t)) \quad \text{for all } t \ge 0$$

and so x is a solution on $[0, \infty)$ of the delay differential equation (E). As x belongs to X, (3.11) holds true and consequently

$$\lim_{t\to\infty} \int_t^\infty (s-t) f(s, x_s, x'(s))ds = 0 = \lim_{t\to\infty} \int_t^\infty f(s, x_s, x'(s))ds.$$

By using this fact, from (3.19) we can easily conclude that the solution x is such that (2.5) and (2.6) hold. Furthermore, (2.7) is an immediate consequence of (3.19). Moreover, by taking into account (3.8), from (3.19) we obtain for $t \ge 0$

$$\begin{aligned} |x(t) - (\xi t + \eta)| &= \left| -\int_t^\infty (s-t) f(s, x_s, x'(s))ds \right| \le \int_t^\infty (s-t)\, |f(s, x_s, x'(s))|\, ds \\ &\le \int_t^\infty (s-t) F(s, \gamma_s, c)ds \le \int_0^\infty sF(s, \gamma_s, c)ds. \end{aligned}$$

Thus, in view of (2.2), we have

$$|x(t) - (\xi t + \eta)| \le c - |\eta| \quad \text{for every } t \ge 0. \tag{3.20}$$

Also, since $x = Mx$, it follows from (3.18) that

$$|x'(t) - \xi| \le c - |\xi| \quad \text{for every } t \ge 0. \tag{3.21}$$

Finally, we see that (3.20) and (3.21) coincide with (2.8) and (2.9), respectively.

The proof of the theorem is complete.

4. PROOF OF THEOREM 2

The proof of Theorem 2 is also based on the use of the Schauder's theorem stated in the previous section. The compactness criterion for subsets of the Banach space $BC([0, \infty), \mathbf{R})$, which is given in Section 3, will also be needed in the present section.

In this section, S_0 stands for the set of all functions u in $C([-r, \infty), \mathbf{R})$ with $u(t) = O(t)$ for $t \to \infty$. The set S_0 *is a Banach space endowed with the norm* $\| \cdot \|_0$ defined by the formula

$$\| u \|_0 = \max\left\{ \max_{-r\le t\le 0} |u(t)|,\ \sup_{t\ge 0} \frac{|u(t)|}{t+1} \right\} \quad \text{for } u \in S_0.$$

The following proposition will be used in order to prove Theorem 2.

Proposition 2. *Assume that* (2.10) *holds, where* F_0 *is a nonnegative real-valued function defined on* $[0, \infty) \times C([-r, 0], [0, \infty))$, *which satisfies the Continuity Condition* (C_0). *Suppose that* (B_0) *is satisfied.*

Let ξ and η be given real constants, and let c be a positive real number such that

$$\int_0^\infty tF_0(t,\gamma_t)dt < \infty, \tag{4.1}$$

where the function γ in $C([-r,\infty),[0,\infty))$ depends on c and is defined by (2.4). Let also X_0 be the subset of S_0 defined by

$$X_0 = \{x \in S_0 : \ \|x\|_0 \leq c\}. \tag{4.2}$$

Then the formula

$$(M_0x)(t) = \begin{cases} \eta - \int_0^\infty sf_0(s,x_s)ds & \textit{for } -r \leq t \leq 0 \\ \xi t + \eta - \int_t^\infty (s-t)f_0(s,x_s)ds & \textit{for } t \geq 0 \end{cases} \tag{4.3}$$

makes sense for any function x in X_0, and this formula defines a continuous mapping M_0 of X_0 into S_0 such that M_0X_0 is relatively compact.

Proof of Proposition 2. Consider an arbitrary function x in X_0. By taking into account the definition, by (4.2), of the set X_0, we immediately see that x satisfies (3.4) and (3.6). These two inequalities together with (2.4) imply $|x(t)| \leq \gamma(t)$ for $t \geq -r$. Consequently, (3.7) holds true. By using (3.7) and the assumption (B_0), we find

$$F_0(t,|x_t|) \leq F_0(t,\gamma_t) \quad \text{for } t \geq 0.$$

But, in view of (2.10), it holds

$$|f_0(t,x_t)| \leq F_0(t,|x_t|) \quad \text{for } t \geq 0.$$

Hence, we have

$$|f_0(t,x_t)| \leq F_0(t,\gamma_t) \quad \text{for every } t \geq 0. \tag{4.4}$$

From (4.1) and (4.4) it follows that

$$\int_0^\infty t\,|f_0(t,x_t)|\,dt < \infty, \tag{4.5}$$

which ensures, in particular, that

$$\int_0^\infty |f_0(t,x_t)|\,dt < \infty. \tag{4.6}$$

Inequalities (4.5) and (4.6) guarantee that

$$\int_0^\infty tf_0(t,x_t)dt \quad \text{and} \quad \int_0^\infty f_0(t,x_t)dt \qquad \text{exist in } \mathbf{R}. \tag{4.7}$$

Since (4.7) holds true for every function x in X_0, we can immediately conclude that *the formula* (4.3) *makes sense for any function x in X_0, and this formula defines a mapping M_0 of X_0* into $C([-r,\infty),\mathbf{R})$. Furthermore, we shall prove that M_0 is a mapping of X_0 *into* S_0, i.e. that $M_0X_0 \subseteq S_0$. For this purpose, let us consider an arbitrary function x in X_0. Then, by taking into account (4.4), from (4.3) we derive for $t \geq 0$

$$\begin{aligned}
\frac{|(M_0x)(t)|}{t+1} &= \left|\frac{\xi t+\eta}{t+1} - \frac{1}{t+1}\int_t^{\infty}(s-t)f_0(s,x_s)ds\right| \\
&\leq \frac{|\xi|\, t+|\eta|}{t+1} + \frac{1}{t+1}\int_t^{\infty}(s-t)\,|f_0(s,x_s)|\,ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_t^{\infty}(s-t)\,|f_0(s,x_s)|\,ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_t^{\infty}(s-t)F_0(s,\gamma_s)ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_0^{\infty} sF_0(s,\gamma_s)ds.
\end{aligned}$$

So, if we set

$$Q_0 = \max\{|\xi|\,,|\eta|\} + \int_0^{\infty} sF_0(s,\gamma_s)ds, \tag{4.8}$$

then we have

$$\frac{|(M_0x)(t)|}{t+1} \leq Q_0 \quad \text{for all } t \geq 0. \tag{4.9}$$

We note that, because of (4.1), Q_0 is a nonnegative real constant. It follows from (4.9) that M_0x belongs to S_0. Thus, it has been established that $M_0x \in S_0$ for every function $x \in X_0$, i.e. that $M_0X_0 \subseteq S_0$.

Now, we will show that M_0X_0 *is relatively compact.* We observe that, for any function x in X_0, it follows from (4.3) that

$$(M_0x)(t) = (M_0x)(0) = \left.\frac{(M_0x)(s)}{s+1}\right|_{s=0} \qquad \text{for } -r \leq t \leq 0.$$

By taking into account this fact as well as the definition of the norm $\nparallel \cdot \nparallel_0$, we can easily see that it is enough to show that the set

$$U_0 = \left\{u: \text{ There exists } x \in X_0 \text{ such that } u(t) = \frac{(M_0x)(t)}{t+1} \text{ for } t \geq 0\right\}$$

is relatively compact in the Banach space $BC([0,\infty),\mathbf{R})$. Every function x in X_0 is such that (4.9) holds, where the nonnegative real constant Q_0 is defined by (4.8) (and it is independent of x). Thus, the set U_0 is uniformly bounded. Furthermore, let x be an arbitrary function in X_0. Then, from (4.3) we obtain for $t \geq 0$

$$\begin{aligned}
\left|\frac{(M_0x)(t)}{t+1} - \xi\right| &= \left|\frac{\xi t+\eta - \int_t^{\infty}(s-t)f_0(s,x_s)ds}{t+1} - \xi\right| \\
&= \frac{\left|-\xi+\eta - \int_t^{\infty}(s-t)f_0(s,x_s)ds\right|}{t+1} \\
&\leq \frac{|-\xi+\eta| + \int_t^{\infty}(s-t)\,|f_0(s,x_s)|\,ds}{t+1}.
\end{aligned}$$

Hence, in view of (4.4), it holds

$$\left|\frac{(M_0x)(t)}{t+1} - \xi\right| \leq \frac{|-\xi+\eta| + \int_t^{\infty}(s-t)F_0(s,\gamma_s)ds}{t+1} \qquad \text{for } t \geq 0. \tag{4.10}$$

It follows, in particular, from (4.1) that

$$\int_0^\infty F_0(t,\gamma_t)dt < \infty. \tag{4.11}$$

We get

$$\begin{aligned}\lim_{t\to\infty}\frac{|-\xi+\eta|+\int_t^\infty (s-t)F_0(s,\gamma_s)ds}{t+1} &= \lim_{t\to\infty}\left[|-\xi+\eta|+\int_t^\infty (s-t)F_0(s,\gamma_s)ds\right]' \\ &= \lim_{t\to\infty}\left[-\int_t^\infty F_0(s,\gamma_s)ds\right]\end{aligned}$$

and, consequently, by virtue of (4.11), we find

$$\lim_{t\to\infty}\frac{|-\xi+\eta|+\int_t^\infty (s-t)F_0(s,\gamma_s)ds}{t+1} = 0. \tag{4.12}$$

Inequality (4.10) together with (4.12) implies

$$\lim_{t\to\infty}\frac{(M_0x)(t)}{t+1} = \xi.$$

By using again (4.10) and (4.12), we can easily conclude that U_0 is equiconvergent at ∞. Now, let again x be an arbitrary function in X_0. From (4.3) we derive for every $t \ge 0$

$$\begin{aligned}\left|\left[\frac{(M_0x)(t)}{t+1}\right]'\right| &= \frac{1}{(t+1)^2}\left|(t+1)(M_0x)'(t)-(M_0x)(t)\right| \\ &\le \left|(t+1)(M_0x)'(t)-(M_0x)(t)\right| \\ &= \left|(t+1)\left[\xi+\int_t^\infty f_0(s,x_s)ds\right]\right. \\ &\quad \left. -\left[\xi t+\eta-\int_t^\infty (s-t)f_0(s,x_s)ds\right]\right| \\ &= \left|\xi-\eta+\int_t^\infty f_0(s,x_s)ds+\int_t^\infty sf_0(s,x_s)ds\right| \\ &\le |\xi-\eta|+\int_t^\infty |f_0(s,x_s)|\,ds+\int_t^\infty s\,|f_0(s,x_s)|\,ds \\ &\le |\xi-\eta|+\int_0^\infty |f_0(s,x_s)|\,ds+\int_0^\infty s\,|f_0(s,x_s)|\,ds.\end{aligned}$$

So, because of (4.4), we have

$$\left|\left[\frac{(M_0x)(t)}{t+1}\right]'\right| \le \Theta \quad \text{for every } t \ge 0, \tag{4.13}$$

where

$$\Theta = |\xi-\eta|+\int_0^\infty F_0(s,\gamma_s)ds+\int_0^\infty sF_0(s,\gamma_s)ds.$$

In view of (4.1) and (4.11), Θ is a nonnegative real number. By taking into account (4.13) and applying the mean value theorem, we find

$$\left|\frac{(M_0x)(t_1)}{t_1+1}-\frac{(M_0x)(t_2)}{t_2+1}\right| \le \Theta\,|t_1-t_2| \quad \text{for every } t_1 \ge 0,\ t_2 \ge 0.$$

Since the last inequality is fulfilled for all functions x in X_0 (and Θ is independent of x), we immediately see that U_0 is equicontinuous. By the given compactness criterion, U_0 is relatively compact in $BC([0,\infty),\mathbf{R})$. So, the relative compactness of M_0X_0 has been proved.

Next, we shall prove that *the mapping M_0 is continuous.* Let x be an arbitrary function in X_0 and $\left(x^{[\nu]}\right)_{\nu\geq 1}$ be any sequence of functions in X_0 with

$$\not\Vert \cdot \not\Vert_0 - \lim_{\nu\to\infty} x^{[\nu]} = x.$$

It is not difficult to verify that

$$\|\cdot\| - \lim_{\nu\to\infty} x_t^{[\nu]} = x_t \quad \text{for every } t \geq 0.$$

Moreover, (4.4) guarantees that

$$\left|f_0(t, x_t^{[\nu]})\right| \leq F_0(t, \gamma_t) \quad \text{for every } t \geq 0 \quad \text{and} \quad \text{for all } \nu \geq 1.$$

So, by taking into account (4.1) and (4.11), we can apply the Lebesgue dominated convergence theorem to obtain, for $t \geq 0$,

$$\lim_{\nu\to\infty} \int_t^\infty (s-t) f_0(s, x_s^{[\nu]}) ds = \int_t^\infty (s-t) f_0(s, x_s) ds.$$

Thus, from (4.3) it follows that

$$\lim_{\nu\to\infty} (M_0 x^{[\nu]})(t) = (M_0 x)(t) \quad \text{for } t \geq -r.$$

Since M_0X_0 is relatively compact and the $\not\Vert \cdot \not\Vert_0$ $-$convergence implies the pointwise convergence to the same limit function, we can follow the same procedure as in the proof of Proposition 1 to conclude that the above convergence is also $\not\Vert \cdot \not\Vert_0$ $-$convergence, i.e. to conclude that

$$\not\Vert \cdot \not\Vert_0 - \lim_{\nu\to\infty} M_0 x^{[\nu]} = M_0 x.$$

This shows that M_0 is continuous.

The proof of the proposition is now complete.

Now, we proceed to the proof of Theorem 2.

Proof of Theorem 2. Consider the set X_0 defined by (4.2). It is clear that *X_0 is a nonempty convex and closed subset of S_0.* It follows, in particular, from the hypothesis (2.11) that (4.1) holds. Hence, Proposition 2 guarantees that the formula (4.3) makes sense for any function x in X_0, and this formula defines *a continuous mapping M_0 of X_0 into S_0 such that M_0X_0 is relatively compact.* We will show that M_0 is a mapping of X_0 *into itself,* i.e. that $M_0X_0 \subseteq X_0$. For this purpose, let us consider an arbitrary function x in X_0. Then (4.9) is satisfied, where the nonnegative real number Q_0 is defined by (4.8). Assumption (2.11) ensures that $Q_0 \leq c$. So, (4.9) gives

$$\frac{|(M_0x)(t)|}{t+1} \leq c \quad \text{for every } t \geq 0. \tag{4.14}$$

In particular, (4.14) guarantees that $|(M_0x)(0)| \leq c$. But, from (4.3) it follows that M_0x is constant on the interval $[-r, 0]$. So, we always have

$$|(M_0x)(t)| \leq c \quad \text{for } -r \leq t \leq 0. \tag{4.15}$$

Inequalities (4.14) and (4.15) give $\nparallel M_0 x \nparallel_0 \leq c$, which means that $M_0 x$ belongs to X_0. So, we have proved that $M_0 x \in X_0$ for every function x in X_0, which ensures that $M_0 X_0 \subseteq X_0$.

Now, by applying the Schauder theorem, we conclude that there exists at least one x in X_0 with $x = M_0 x$, i.e.

$$x(t) = \begin{cases} \eta - \int_0^\infty s f_0(s, x_s) ds & \text{for } -r \leq t \leq 0 \\ \xi t + \eta - \int_t^\infty (s-t) f_0(s, x_s) ds & \text{for } t \geq 0. \end{cases} \tag{4.16}$$

We immediately obtain

$$x''(t) = -f_0(t, x_t) \quad \text{for } t \geq 0$$

and consequently x is a solution on $[0, \infty)$ of the delay differential equation (E_0). Since $x \in X_0$, (4.7) is true and so

$$\lim_{t\to\infty} \int_t^\infty (s-t) f_0(s, x_s) ds = 0 = \lim_{t\to\infty} \int_t^\infty f_0(s, x_s) ds.$$

By taking into account this fact, we can use (4.16) to see that the solution x is such that (2.5) and (2.6) hold. Next, we observe that (2.7) is an immediate consequence of (4.16). Furthermore, by taking into account (4.4), from (4.16) we get for $t \geq 0$

$$\begin{aligned} |x(t) - (\xi t + \eta)| &= \left| - \int_t^\infty (s-t) f_0(s, x_s) ds \right| \leq \int_t^\infty (s-t) \left| f_0(s, x_s) \right| ds \\ &\leq \int_t^\infty (s-t) F_0(s, \gamma_s) ds \leq \int_0^\infty s F_0(s, \gamma_s) ds. \end{aligned}$$

So, because of (2.11), it holds

$$|x(t) - (\xi t + \eta)| \leq c - \max\{|\xi|, |\eta|\} \quad \text{for every } t \geq 0. \tag{4.17}$$

Moreover, (4.16) gives, for $t \geq 0$,

$$|x'(t) - \xi| = \left| \int_t^\infty f_0(s, x_s) ds \right| \leq \int_t^\infty |f_0(s, x_s)| ds \leq \int_0^\infty |f_0(s, x_s)| ds.$$

Therefore, by (4.4), we have

$$|x'(t) - \xi| \leq \int_0^\infty F_0(s, \gamma_s) ds \quad \text{for every } t \geq 0. \tag{4.18}$$

Note that, because of (4.11), $\int_0^\infty F_0(s, \gamma_s) ds$ is finite. Finally, we see that (4.17) and (4.18) coincide with (2.12) and (2.13), respectively.

The proof of the theorem is complete.

5. PROOFS OF THEOREMS 3 AND 4

In order to prove Theorems 3 and 4, we will make use of the well-known Banach's Contraction Principle (see, e.g., Kartsatos [7]).

The Banach Contraction Principle. *Let P be a Banach space and Y any nonempty closed subset of P. If N is a contraction of Y into itself, then the mapping N has exactly one fixed point (i.e. there exists a unique $y \in Y$ with $y = Ny$).*

The following lemma provides a useful integral representation of Problem (D), (2.5), (2.6) (where ξ and η are given real constants), which will be used in proving Theorem 3.

Lemma 1. *Let ξ and η be given real constants. A real-valued function x, which is continuously differentiable on the interval $[0,\infty)$, is a solution on $[0,\infty)$ of the ordinary differential equation* (D) *such that* (2.5) *and* (2.6) *hold, if and only if it satisfies*

$$x(t) = \xi t + \eta - \int_t^\infty (s-t)h(s,x(s),x'(s))ds \quad \text{for } t \geq 0. \tag{5.1}$$

A particular case of Lemma 1 is Lemma 2 below concerning Problem (D_0), (2.5), (2.6); Lemma 2 will be used in the proof of Theorem 4.

Lemma 2. *Let ξ and η be given real constants. A function x in $C([0,\infty),\mathbf{R})$ is a solution on $[0,\infty)$ of the ordinary differential equation* (D_0) *such that* (2.5) *and* (2.6) *hold, if and only if it satisfies*

$$x(t) = \xi t + \eta - \int_t^\infty (s-t)h_0(s,x(s))ds \quad \text{for } t \geq 0. \tag{5.2}$$

Proof of Lemma 1. Let x be a real-valued function, which is continuously differentiable on the interval $[0,\infty)$.

Assume first that x satisfies (5.1). Then

$$\lim_{t\to\infty}[x(t) - (\xi t + \eta)] = -\lim_{t\to\infty}\int_t^\infty (s-t)h(s,x(s),x'(s))ds = 0$$

and so (2.5) holds true. Also, we immediately obtain

$$x'(t) = \xi + \int_t^\infty h(s,x(s),x'(s))ds \quad \text{for every } t \geq 0,$$

which gives

$$\lim_{t\to\infty}[x'(t) - \xi] = \lim_{t\to\infty}\int_t^\infty h(s,x(s),x'(s))ds = 0,$$

i.e. (2.6) is fulfilled. Moreover, we have

$$x''(t) = -h(t,x(t),x'(t)) \quad \text{for all } t \geq 0,$$

which means that x is a solution on $[0,\infty)$ of (D).

Conversely, let us suppose that x is a solution on $[0,\infty)$ of (D) such that (2.5) and (2.6) hold. Then from (D) it follows that

$$x'(T) - x'(t) = -\int_t^T h(s,x(s),x'(s))ds \quad \text{for all } T,\, t \text{ with } T \geq t \geq 0.$$

Consequently,

$$\lim_{T\to\infty} x'(T) - x'(t) = -\int_t^\infty h(s,x(s),x'(s))ds \quad \text{for every } t \geq 0.$$

But, in view of (2.6), we have $\lim_{T\to\infty} x'(T) = \xi$. Thus,

$$x'(t) = \xi + \int_t^\infty h(s, x(s), x'(s))ds \quad \text{for } t \geq 0.$$

This gives

$$x(T) - x(t) = \xi(T-t) + \int_t^T \int_s^\infty h(\sigma, x(\sigma), x'(\sigma))d\sigma ds \quad \text{for } T \geq t \geq 0,$$

which can equivalently be written as

$$[x(T) - (\xi T + \eta)] - [x(t) - (\xi t + \eta)] = \int_t^T \int_s^\infty h(\sigma, x(\sigma), x'(\sigma))d\sigma ds$$
$$\text{for } T \geq t \geq 0.$$

Hence,

$$\begin{aligned}\lim_{T\to\infty} [x(T) - (\xi T + \eta)] - [x(t) - (\xi t + \eta)] &= \int_t^\infty \int_s^\infty h(\sigma, x(\sigma), x'(\sigma))d\sigma ds \\ &= \int_t^\infty (s-t)h(s, x(s), x'(s))ds \quad \text{for } t \geq 0.\end{aligned}$$

But, because of (2.5), it holds $\lim_{T\to\infty} [x(T) - (\xi T + \eta)] = 0$. Therefore,

$$-x(t) + (\xi t + \eta) = \int_t^\infty (s-t)h(s, x(s), x'(s))ds \quad \text{for all } t \geq 0,$$

i.e. x satisfies (5.1).

The proof of the lemma has been finished.

Now, we are in a position to present the proofs of Theorems 3 and 4.

Proof of Theorem 3. Let P be the set of all real-valued functions on the interval $[0, \infty)$, which have bounded continuous derivatives on $[0, \infty)$. This set *is a Banach space endowed with the norm* $\|\cdot\|^*$ defined by

$$\|u\|^* = \max\left\{|u(0)|, \sup_{t\geq 0} |u'(t)|\right\} \quad \text{for } u \in P.$$

Let also Y be the *nonempty closed subset of* P defined by

$$Y = \{x \in P : \ \|x\|^* \leq c\}.$$

Clearly, Y is the subset of P consisting of all functions x in P which satisfy (2.24) and (2.25).

Consider now an arbitrary function x in Y. Then x satisfies (2.24) and (2.25), which imply (2.28). By using (2.25) and (2.28) as well as the hypothesis (A), we obtain

$$H(t, |x(t)|, |x'(t)|) \leq H(t, c(t+1), c) \quad \text{for } t \geq 0.$$

On the other hand, the assumption (2.16) guarantees that

$$|h(t, x(t), x'(t))| \leq H(t, |x(t)|, |x'(t)|) \quad \text{for } t \geq 0.$$

Thus, we have

$$|h(t, x(t), x'(t))| \leq H(t, c(t+1), c) \quad \text{for all } t \geq 0. \tag{5.3}$$

Furthermore, we observe that the hypothesis (2.17) ensures, in particular, that

$$\int_0^\infty tH(t, c(t+1), c)dt < \infty$$

and consequently, by taking into account (5.3), we obtain

$$\int_0^\infty t\,|h(t, x(t), x'(t))|\,dt < \infty.$$

So,

$$\int_0^\infty th(t, x(t), x'(t))dt \quad \text{and, in particular,} \quad \int_0^\infty h(t, x(t), x'(t))dt \qquad \text{exist in } \mathbf{R}.$$

This is true for all functions x in Y. Hence, the formula

$$(Nx)(t) = \xi t + \eta - \int_t^\infty (s-t)h(s, x(s), x'(s))ds \quad \textit{for } t \geq 0$$

makes sense for any function x in Y, and this formula defines *a mapping* N *of* Y into $C([0,\infty), \mathbf{R})$. We will show that N is a mapping of Y *into itself*, i.e. that $NY \subseteq Y$. To this end, let us consider an arbitrary function x in Y. Then, by taking into account (5.3), we obtain

$$\begin{aligned} |(Nx)(0)| &= \left|\eta - \int_0^\infty sh(s, x(s), x'(s))ds\right| \leq |\eta| + \int_0^\infty s\,|h(s, x(s), x'(s))|\,ds \\ &\leq |\eta| + \int_0^\infty sH(s, c(s+1), c)ds \end{aligned}$$

and consequently, in view of (2.17), it holds

$$|(Nx)(0)| \leq c. \tag{5.4}$$

Furthermore, by taking again into account (5.3), we derive for $t \geq 0$

$$\begin{aligned} |(Nx)'(t) - \xi| &= \left|\int_t^\infty h(s, x(s), x'(s))ds\right| \leq \int_t^\infty |h(s, x(s), x'(s))|\,ds \\ &\leq \int_t^\infty H(s, c(s+1), c)ds \leq \int_0^\infty H(s, c(s+1), c)ds \end{aligned}$$

and so, because of (2.18), we have

$$|(Nx)'(t) - \xi| \leq c - |\xi| \quad \text{for all } t \geq 0. \tag{5.5}$$

It follows from (5.5) that

$$|(Nx)'(t)| \leq c \;\; \text{for every } t \geq 0. \tag{5.6}$$

Inequalities (5.4) and (5.6) mean that Nx belongs to Y. It has been verified that, for each $x \in Y$, Nx belongs to Y. Thus, we always have $NY \subseteq Y$.

Now, let u be an arbitrary function in P. Then

$$|u(0)| \leq \|u\|^* \tag{5.7}$$

and

$$|u'(t)| \leq \|u\|^* \quad \text{for every } t \geq 0. \tag{5.8}$$

Furthermore, by using (5.7) and (5.8), we can immediately see that u is also such that

$$|u(t)| \leq \|u\|^*\,(t+1) \quad \text{for every } t \geq 0. \tag{5.9}$$

Next, let us consider two arbitrary functions x and y in Y. Then, by using the assumption (2.22) and taking into account (5.9) and (5.8), we obtain

$$\begin{aligned}
|(Nx)(0)-(Ny)(0)| &= \left|-\int_0^\infty s\left[h(s,x(s),x'(s))-h(s,y(s),y'(s))\right]ds\right| \\
&\leq \int_0^\infty s\left|h(s,x(s),x'(s))-h(s,y(s),y'(s))\right|ds \\
&\leq \int_0^\infty sL(s)\max\left\{|x(s)-y(s)|\,,|x'(s)-y'(s)|\right\}ds \\
&\leq \int_0^\infty sL(s)\max\left\{\|x-y\|^*(s+1),\|x-y\|^*\right\}ds \\
&= \left[\int_0^\infty sL(s)\max\{s+1,1\}\,ds\right]\|x-y\|^*.
\end{aligned}$$

That is,

$$|(Nx)(0)-(Ny)(0)| \leq \left[\int_0^\infty s(s+1)L(s)ds\right]\|x-y\|^*. \tag{5.10}$$

Furthermore, by using again (2.22) and taking again into account (5.9) and (5.8), we get for $t \geq 0$

$$\begin{aligned}
|(Nx)'(t)-(Ny)'(t)| &= \left|\int_t^\infty \left[h(s,x(s),x'(s))-h(s,y(s),y'(s))\right]ds\right| \\
&\leq \int_t^\infty \left|h(s,x(s),x'(s))-h(s,y(s),y'(s))\right|ds \\
&\leq \int_0^\infty \left|h(s,x(s),x'(s))-h(s,y(s),y'(s))\right|ds \\
&\leq \int_0^\infty L(s)\max\left\{|x(s)-y(s)|\,,|x'(s)-y'(s)|\right\}ds \\
&\leq \int_0^\infty L(s)\max\left\{\|x-y\|^*(s+1),\|x-y\|^*\right\}ds \\
&= \left[\int_0^\infty L(s)\max\{s+1,1\}\,ds\right]\|x-y\|^* \\
&= \left[\int_0^\infty (s+1)L(s)ds\right]\|x-y\|^*.
\end{aligned}$$

Thus, we find

$$\sup_{t\geq 0}|(Nx)'(t)-(Ny)'(t)| \leq \left[\int_0^\infty (s+1)L(s)ds\right]\|x-y\|^*. \tag{5.11}$$

Set

$$\theta = \max\left\{\int_0^\infty s(s+1)L(s)ds,\ \int_0^\infty (s+1)L(s)ds\right\}.$$

Then (5.10) and (5.11) give

$$\|Nx-Ny\|^* \leq \theta\,\|x-y\|^*.$$

This inequality holds true for all functions x and y in Y. On the other hand, from the hypothesis (2.23) it follows that $0 \leq \theta < 1$. We have thus proved that *the mapping $M : Y \longrightarrow Y$ is a contraction.*

Finally, by using the Banach Contraction Principle, we conclude that there exists exactly one function x in Y with $x = Nx$. We see that $x = Nx$ is equivalent to the fact that x satisfies (5.1). Hence, by Lemma 1, the ordinary differential equation (D) has exactly one solution x on the interval $[0, \infty)$ satisfying (2.24) and (2.25), and such that (2.5) and (2.6) hold. It remains to establish that this unique solution x satisfies (2.8) and (2.9). By taking into account (5.3), from (5.1) we obtain for $t \geq 0$

$$\begin{aligned}
&|x(t) - (\xi t + \eta)| \\
= \; & \left| -\int_t^{\infty} (s-t) h(s, x(s), x'(s)) ds \right| \leq \int_t^{\infty} (s-t) \left| h(s, x(s), x'(s)) \right| ds \\
\leq \; & \int_t^{\infty} (s-t) H(s, c(s+1), c) ds \leq \int_0^{\infty} s H(s, c(s+1), c) ds.
\end{aligned}$$

So, by using (2.17), we immediately arrive at (3.20). Moreover, as $x = Nx$, it follows from (5.5) that x satisfies (3.21). We see that (3.20) and (3.21) coincide with (2.8) and (2.9), respectively.

The proof of the theorem is complete.

Proof of Theorem 4. Consider the set P_0 of all continuous real-valued functions u on the interval $[0, \infty)$ with $u(t) = O(t)$ for $t \to \infty$. The set P_0 *is a Banach space endowed with the norm* $\|\cdot\|_0^*$ defined by

$$\|u\|_0^* = \sup_{t \geq 0} \frac{|u(t)|}{t+1} \quad \text{for } u \in P_0.$$

Consider also the set Y_0 defined by

$$Y_0 = \left\{ x \in P_0 : \; \|x\|_0^* \leq c \right\}.$$

It is clear that Y_0 is the subset of P_0 consisting of all functions x in P_0 which satisfy (2.28). The set Y_0 *is a nonempty closed subset of* P_0.

Now, let x be an arbitrary function in Y_0. Then x satisfies (2.28). By taking into account (2.28) and using the assumption (A_0), we get

$$H_0(t, |x(t)|) \leq H_0(t, c(t+1)) \quad \text{for } t \geq 0.$$

But, from the hypothesis (2.19) it follows that

$$|h_0(t, x(t))| \leq H_0(t, |x(t)|) \quad \text{for } t \geq 0.$$

By combining the last two inequalities, we obtain

$$|h_0(t, x(t))| \leq H_0(t, c(t+1)) \quad \text{for all } t \geq 0. \tag{5.12}$$

Next, we see that (2.20) implies, in particular,

$$\int_0^{\infty} t H_0(t, c(t+1)) dt < \infty.$$

Thus, because of (5.12), we have

$$\int_0^{\infty} t \left| h_0(t, x(t)) \right| dt < \infty,$$

which guarantees that

$$\int_0^{\infty} t h_0(t, x(t)) dt \quad \text{and, in particular,} \quad \int_0^{\infty} h_0(t, x(t)) dt \qquad \text{exist in } \mathbf{R}.$$

So, as the function x in Y_0 is arbitrary, we immediately see that the formula

$$(N_0x)(t) = \xi t + \eta - \int_t^\infty (s-t)h_0(s,x(s))ds \quad \text{for } t \geq 0$$

makes sense for any function x in Y_0, and this formula defines *a mapping* N_0 *of* Y_0 into $C([0,\infty),\mathbf{R})$. Furthermore, N_0 is a mapping of Y_0 *into itself*, i.e. it holds $N_0Y_0 \subseteq Y_0$. Indeed, by using (5.12) and the hypothesis (2.20), for any function x in Y_0, we obtain, for every $t \geq 0$,

$$\begin{aligned}
\frac{|(N_0x)(t)|}{t+1} &= \left|\frac{\xi t+\eta}{t+1} - \frac{1}{t+1}\int_t^\infty (s-t)h_0(s,x(s))ds\right| \\
&\leq \frac{|\xi|\,t+|\eta|}{t+1} + \frac{1}{t+1}\int_t^\infty (s-t)\,|h_0(s,x(s))|\,ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_t^\infty (s-t)\,|h_0(s,x(s))|\,ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_t^\infty (s-t)H_0(s,c(s+1))ds \\
&\leq \max\{|\xi|\,,|\eta|\} + \int_0^\infty sH_0(s,c(s+1))ds \\
&\leq c.
\end{aligned}$$

That is, for any $x \in Y_0$, N_0x belongs to Y_0. This proves our assertion.

Furthermore, let x and y be two arbitrary functions in Y_0. Then, by using the hypothesis (2.26), we get for $t \geq 0$

$$\begin{aligned}
\frac{|(N_0x)(t)-(N_0y)(t)|}{t+1} &= \frac{1}{t+1}\left|-\int_t^\infty (s-t)\left[h_0(s,x(s))-h_0(s,y(s))\right]ds\right| \\
&\leq \frac{1}{t+1}\int_t^\infty (s-t)\,|h_0(s,x(s))-h_0(s,y(s))|\,ds \\
&\leq \frac{1}{t+1}\int_t^\infty (s-t)L_0(s)\,|x(s)-y(s)|\,ds \\
&= \frac{1}{t+1}\int_t^\infty (s-t)(s+1)L_0(s)\frac{|x(s)-y(s)|}{s+1}ds \\
&\leq \left[\frac{1}{t+1}\int_t^\infty (s-t)(s+1)L_0(s)ds\right]\sup_{s\geq 0}\frac{|x(s)-y(s)|}{s+1}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\sup_{t\geq 0}\frac{|(N_0x)(t)-(N_0y)(t)|}{t+1} &\leq \left[\sup_{t\geq 0}\frac{1}{t+1}\int_t^\infty (s-t)(s+1)L_0(s)ds\right]\sup_{s\geq 0}\frac{|x(s)-y(s)|}{s+1} \\
&= \left[\int_0^\infty s(s+1)L_0(s)ds\right]\sup_{s\geq 0}\frac{|x(s)-y(s)|}{s+1}.
\end{aligned}$$

That is,

$$\|N_0x - N_0y\|_0^* \leq \theta_0\,\|x-y\|_0^*, \tag{5.13}$$

where

$$\theta_0 = \int_0^\infty s(s+1)L_0(s)ds.$$

Because of the assumption (2.27), we have $0 \leq \theta_0 < 1$. As (5.13) holds true for all functions x and y in Y_0, we conclude that *the mapping* $N_0 : Y_0 \longrightarrow Y_0$ *is a contraction.*

By the Banach Contraction Principle, there exists exactly one function x in Y_0 with $x = N_0 x$. Clearly, $x = N_0 x$ is equivalent to (5.2). So, from Lemma 2 it follows that the ordinary differential equation (D_0) has exactly one solution x on the interval $[0, \infty)$ satisfying (2.28), and such that (2.5) and (2.6) hold. Finally, we will show that this unique solution x satisfies also (2.12) and (2.21). By taking into account (5.12), from (5.2) we obtain for $t \geq 0$

$$\begin{aligned} |x(t) - (\xi t + \eta)| &= \left| -\int_t^\infty (s-t) h_0(s, x(s)) ds \right| \leq \int_t^\infty (s-t) \left| h_0(s, x(s)) \right| ds \\ &\leq \int_t^\infty (s-t) H_0(s, c(s+1)) ds \leq \int_0^\infty s H_0(s, c(s+1)) ds. \end{aligned}$$

Thus, by using (2.20), we arrive at (4.17). Furthermore, it follows from (5.2) that, for $t \geq 0$,

$$|x'(t) - \xi| = \left| \int_t^\infty h_0(s, x(s)) ds \right| \leq \int_t^\infty |h_0(s, x(s))| \, ds \leq \int_0^\infty |h_0(s, x(s))| \, ds$$

and consequently, in view of (5.12), we obtain

$$|x'(t) - \xi| \leq \int_0^\infty H_0(s, c(s+1)) ds \quad \text{for every } t \geq 0. \tag{5.14}$$

We notice that, because of (2.20), $\int_0^\infty H_0(s, c(s+1)) ds$ is finite. We immediately observe that (4.17) and (5.14) coincide with (2.12) and (2.21), respectively.

So, the proof of the theorem has been completed.

6. APPLICATION TO DIFFERENTIAL EQUATIONS OF EMDEN-FOWLER TYPE. EXAMPLES

Consider the second order nonlinear delay differential equations of Emden-Fowler type

$$x''(t) + a(t) |x(t-r)|^\alpha \operatorname{sgn} x(t-r) + b(t) |x'(t)|^\beta \operatorname{sgn} x'(t) = 0 \tag{6.1}$$

and

$$x''(t) + a(t) |x(t-r)|^\alpha \operatorname{sgn} x(t-r) = 0 \tag{6.2}$$

as well as the second order nonlinear ordinary Emden-Fowler differential equations

$$x''(t) + a(t) |x(t)|^\alpha \operatorname{sgn} x(t) + b(t) |x'(t)|^\beta \operatorname{sgn} x'(t) = 0 \tag{6.3}$$

and

$$x''(t) + a(t) |x(t)|^\alpha \operatorname{sgn} x(t) = 0, \tag{6.4}$$

where *a and b are continuous real-valued functions on the interval* $[0, \infty)$, *and α and β are positive real numbers.* Consider also the second order linear ordinary differential equations

$$x''(t) + a(t) x(t) + b(t) x'(t) = 0 \tag{6.5}$$

and

$$x''(t) + a(t)x(t) = 0. \tag{6.6}$$

By applying Theorem 1 (or, especially, Corollary 1) and Theorem 2 (or, especially, Corollary 2) to the delay differential equations (6.1) and (6.2), respectivelly, we are led to the following two results:

Result 1. *Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that*

$$c^{\alpha}\left[\int_0^r t\,|a(t)|\,dt + \int_r^{\infty} t(t-r+1)^{\alpha}\,|a(t)|\,dt\right] + c^{\beta}\int_0^{\infty} t\,|b(t)|\,dt \le c - |\eta| \tag{6.7}$$

and

$$c^{\alpha}\left[\int_0^r |a(t)|\,dt + \int_r^{\infty} (t-r+1)^{\alpha}\,|a(t)|\,dt\right] + c^{\beta}\int_0^{\infty} |b(t)|\,dt \le c - |\xi|\,. \tag{6.8}$$

Then the delay differential equation (6.1) *has at least one solution x on the interval $[0,\infty)$ such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.7), (2.8), *and* (2.9).

Result 2. *Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that*

$$c^{\alpha}\left[\int_0^r t\,|a(t)|\,dt + \int_r^{\infty} t(t-r+1)^{\alpha}\,|a(t)|\,dt\right] \le c - \max\{|\xi|, |\eta|\}. \tag{6.9}$$

Then the delay differential equation (6.2) *has at least one solution x on the interval $[0,\infty)$ such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.7), (2.12), *and*

$$\xi - c^{\alpha}\left[\int_0^r |a(s)|\,ds + \int_r^{\infty} (s-r+1)^{\alpha}\,|a(s)|\,ds\right] \le x'(t)$$
$$\le \xi + c^{\alpha}\left[\int_0^r |a(s)|\,ds + \int_r^{\infty} (s-r+1)^{\alpha}\,|a(s)|\,ds\right] \quad \text{for every } t \ge 0.$$

(*Note that, because of* (6.9), $\int_r^{\infty}(s-r+1)^{\alpha}\,|a(s)|\,ds$ *is finite.*)

Also, an application of Theorem 1 (or, especially, of Corollary 3) and of Theorem 2 (or, especially, of Corollary 4) to the ordinary differential equations (6.3) and (6.4), respectivelly, leads to the next two results:

Result 3. *Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that*

$$c^{\alpha}\int_0^{\infty} t(t+1)^{\alpha}\,|a(t)|\,dt + c^{\beta}\int_0^{\infty} t\,|b(t)|\,dt \le c - |\eta| \tag{6.10}$$

and

$$c^{\alpha}\int_0^{\infty} (t+1)^{\alpha}\,|a(t)|\,dt + c^{\beta}\int_0^{\infty} |b(t)|\,dt \le c - |\xi|\,. \tag{6.11}$$

Then the ordinary differential equation (6.3) *has at least one solution x on the interval $[0,\infty)$ such that* (2.5) *and* (2.6) *hold; in addition, this solution x satisfies* (2.8) *and* (2.9).

Result 4. *Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that*

$$c^\alpha \int_0^\infty t(t+1)^\alpha |a(t)| \, dt \le c - \max\{|\xi|, |\eta|\}. \tag{6.12}$$

Then the ordinary differential equation (6.4) has at least one solution x on the interval $[0, \infty)$ such that (2.5) and (2.6) hold; in addition, this solution x satisfies (2.12) and

$$\xi - c^\alpha \int_0^\infty (s+1)^\alpha |a(s)| \, ds \le x'(t) \le \xi + c^\alpha \int_0^\infty (s+1)^\alpha |a(s)| \, ds \quad \text{for every } t \ge 0.$$

(Note that, because of (6.12), $\int_0^\infty (s+1)^\alpha |a(s)| \, ds$ is finite.)

Moreover, if we apply Theorem 3 and Theorem 4 to the linear ordinary differential equations (6.5) and (6.6), respectivelly, then we can arrive at the following two results:

Result 5. *Assume that*

$$\max\left\{\int_0^\infty t(t+1) \left[|a(t)| + |b(t)|\right] dt, \ \int_0^\infty (t+1) \left[|a(t)| + |b(t)|\right] dt\right\} < 1. \tag{6.13}$$

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that

$$c\left[\int_0^\infty t(t+1) |a(t)| \, dt + \int_0^\infty t \, |b(t)| \, dt\right] \le c - |\eta| \tag{6.14}$$

and

$$c\left[\int_0^\infty (t+1) |a(t)| \, dt + \int_0^\infty |b(t)| \, dt\right] \le c - |\xi| \, . \tag{6.15}$$

Then the linear ordinary differential equation (6.5) has exactly one solution x on the interval $[0, \infty)$ satisfying (2.24) and (2.25), and such that (2.5) and (2.6) hold; in addition, this unique solution x satisfies (2.8) and (2.9).

Result 6. *Assume that*

$$\int_0^\infty t(t+1) |a(t)| \, dt < 1. \tag{6.16}$$

Let ξ and η be given real constants, and let there exist a real number c with $c > \max\{|\xi|, |\eta|\}$ so that

$$c \int_0^\infty t(t+1) |a(t)| \, dt \le c - \max\{|\xi|, |\eta|\}. \tag{6.17}$$

Then the linear ordinary differential equation (6.6) has exactly one solution x on the interval $[0, \infty)$ satisfying (2.28), and such that (2.5) and (2.6) hold; in addition, this unique solution x satisfies (2.12) and

$$\xi - c \int_0^\infty (s+1) |a(s)| \, ds \le x'(t) \le \xi + c \int_0^\infty (s+1) |a(s)| \, ds \quad \text{for every } t \ge 0.$$

(Note that, because of (6.17), $\int_0^\infty (s+1) |a(s)| \, ds$ is finite.)

Note: Provided that at least one of ξ and η is nonzero, the assumption (6.17) implies the hypothesis (6.16).

Now, we will give some examples to demonstrate the applicability of our results.

Example 1. Consider the delay differential equation (6.1) with $r = 1$, $\alpha = 2$, $\beta = 1$, and

$$a(t) = \frac{8}{21(t+1)^5} \quad \text{for } t \geq 0, \qquad b(t) = \frac{1}{3(t+1)^3} \quad \text{for } t \geq 0.$$

Take $\xi = \frac{5}{6}$ and $\eta = 1$. Inequality (6.7) becomes

$$c^2 \left[\int_0^1 t \frac{8}{21(t+1)^5} dt + \int_1^\infty t^3 \frac{8}{21(t+1)^5} dt \right] + c \int_0^\infty t \frac{1}{3(t+1)^3} dt \leq c - 1,$$

i.e.

$$\frac{1}{9}c^2 + \frac{1}{6}c \leq c - 1. \tag{6.18}$$

We immediately see that (6.18) holds if and only if

$$\frac{3}{2} \leq c \leq 6. \tag{6.19}$$

Furthermore, Inequality (6.8) is written as

$$c^2 \left[\int_0^1 \frac{8}{21(t+1)^5} dt + \int_1^\infty t^2 \frac{8}{21(t+1)^5} dt \right] + c \int_0^\infty \frac{1}{3(t+1)^3} dt \leq c - \frac{5}{6},$$

i.e.

$$\frac{1}{9}c^2 + \frac{1}{6}c \leq c - \frac{5}{6}. \tag{6.20}$$

We observe that (6.20) is satisfied if and only if

$$\frac{15 - \sqrt{105}}{4} \leq c \leq \frac{15 + \sqrt{105}}{4}. \tag{6.21}$$

Since

$$1 < \frac{15 - \sqrt{105}}{4} < \frac{3}{2} < 6 < \frac{15 + \sqrt{105}}{4},$$

both (6.19) and (6.21) are fulfilled if and only if c satisfies (6.19). That is, both Inequalities (6.7) and (6.8) hold if and only if c is such that (6.19) is satisfied. Thus, if we choose $c = \frac{3}{2}$, then Result 1 leads to the following result:

The delay differential equation

$$x''(t) + \frac{8}{21(t+1)^5} [x(t-1)]^2 \operatorname{sgn} x(t-1) + \frac{1}{3(t+1)^3} x'(t) = 0$$

has at least one solution x on the interval $[0, \infty)$ such that

$$x(t) = \frac{5}{6}t + 1 + o(1) \quad \textit{for } t \to \infty \tag{6.22}$$

and

$$x'(t) = \frac{5}{6} + o(1) \quad \textit{for } t \to \infty; \tag{6.23}$$

in addition, this solution x satisfies

$$x(t) = x(0) \quad \textit{for } -1 \leq t \leq 0, \tag{6.24}$$

$$\frac{5}{6}t+\frac{1}{2}\le x(t)\le \frac{5}{6}t+\frac{3}{2}\quad \textit{for every } t\ge 0 \tag{6.25}$$

and

$$\frac{1}{6}\le x'(t)\le \frac{3}{2}\quad \textit{for every } t\ge 0. \tag{6.26}$$

Example 2. Consider the delay differential equation (6.2) with $r=1$, $\alpha=2$, and

$$a(t)=\frac{16}{35(t+1)^5}\quad \text{for } t\ge 0.$$

Take $\xi=\frac{6}{5}$ and $\eta=1$. Inequality (6.9) is written

$$c^2\left[\int_0^1 t\frac{16}{35(t+1)^5}dt+\int_1^\infty t^3\frac{16}{35(t+1)^5}dt\right]\le c-\frac{6}{5},$$

i.e.

$$\frac{2}{15}c^2\le c-\frac{6}{5}. \tag{6.27}$$

We immediately see that (6.27) is satisfied if and only if (6.19) holds. That is, Inequality (6.9) holds true if and only if c satisfies (6.19). Choose $c=\frac{3}{2}$. Then, by applying Result 2, we arrive at the next result:

The delay differential equation

$$x''(t)+\frac{16}{35(t+1)^5}\left[x(t-1)\right]^2\operatorname{sgn}x(t-1)=0$$

has at least one solution x on the interval $[0,\infty)$ such that

$$x(t)=\frac{6}{5}t+1+o(1)\quad \textit{for } t\to\infty \tag{6.28}$$

and

$$x'(t)=\frac{6}{5}+o(1)\quad \textit{for } t\to\infty; \tag{6.29}$$

in addition, this solution x satisfies (6.24) and:

$$\frac{6}{5}t+\frac{7}{10}\le x(t)\le \frac{6}{5}t+\frac{13}{10}\quad \textit{for every } t\ge 0 \tag{6.30}$$

and

$$\frac{9}{10}\le x'(t)\le \frac{3}{2}\quad \textit{for every } t\ge 0. \tag{6.31}$$

Example 3. Let us consider the ordinary differential equation (6.3) with $\alpha=2$, $\beta=1$, and

$$a(t)=\frac{2}{9(t+1)^5}\quad \text{for } t\ge 0,\qquad b(t)=\frac{1}{3(t+1)^3}\quad \text{for } t\ge 0.$$

Let us take $\xi=\frac{5}{6}$ and $\eta=1$. Then, Inequality (6.10) becomes

$$c^2\int_0^\infty t\frac{2}{9(t+1)^3}dt+c\int_0^\infty t\frac{1}{3(t+1)^3}dt\le c-1,$$

which leads to (6.18). Also, Inequality (6.11) is written as

$$c^2 \int_0^\infty \frac{2}{9(t+1)^3} dt + c \int_0^\infty \frac{1}{3(t+1)^3} dt \le c - \frac{5}{6},$$

which is equivalent to (6.20). As in Example 1, we see that both (6.18) and (6.20) are satisfied if and only if (6.19) holds. So, both Inequalities (6.10) and (6.11) hold if and only if c satisfies (6.19). Thus, by applying Result 3 with $c = \frac{3}{2}$, we are led to the following result:

The ordinary differential equation

$$x''(t) + \frac{2}{9(t+1)^5} [x(t)]^2 \operatorname{sgn} x(t) + \frac{1}{3(t+1)^3} x'(t) = 0$$

has at least one solution x on the interval $[0, \infty)$ such that (6.22) *and* (6.23) hold; *in addition, this solution x satisfies* (6.25) *and* (6.26).

Example 4. Let us consider the ordinary differential equation (6.4) with $\alpha = 2$ and

$$a(t) = \frac{4}{15(t+1)^5} \quad \text{for } t \ge 0.$$

Let us take $\xi = \frac{6}{5}$ and $\eta = 1$. In this case, Inequality (6.12) is written as follows

$$c^2 \int_0^\infty t \frac{4}{15(t+1)^3} dt \le c - \frac{6}{5},$$

which is equivalent to (6.27). But, (6.27) holds if and only if c satisfies (6.19). That is, Inequality (6.12) is fulfilled if and only if c is such that (6.19) holds. So, an application of Result 4 with $c = \frac{3}{2}$ leads to the next result:

The ordinary differential equation

$$x''(t) + \frac{4}{15(t+1)^5} [x(t)]^2 \operatorname{sgn} x(t) = 0$$

has at least one solution x on the interval $[0, \infty)$ such that (6.28) *and* (6.29) *hold; in addition, this solution x satisfies* (6.30) *and* (6.31).

Example 5. Consider the linear ordinary differential equation (6.5) with

$$a(t) = b(t) = \frac{1}{2(t+1)^4} \quad \text{for } t \ge 0.$$

We find

$$\int_0^\infty t \frac{1}{(t+1)^3} dt = \int_0^\infty \frac{1}{(t+1)^3} dt = \frac{1}{2}$$

and hence (6.13) is always satisfied. Now, take $\xi = \frac{5}{6}$ and $\eta = 1$. Inequality (6.14) becomes

$$c \left[\int_0^\infty t \frac{1}{2(t+1)^3} dt + \int_0^\infty t \frac{1}{2(t+1)^4} dt \right] \le c - 1,$$

i.e

$$\frac{1}{3} c \le c - 1 \quad \text{or} \quad c \ge \frac{3}{2}.$$

Moreover, Inequality (6.15) is written as

$$c \left[\int_0^\infty \frac{1}{2(t+1)^3} dt + \int_0^\infty \frac{1}{2(t+1)^4} dt \right] \le c - \frac{5}{6},$$

i.e

$$\frac{5}{12}c \leq c - \frac{5}{6} \quad \text{or} \quad c \geq \frac{10}{7}.$$

Thus, both Inequalities (6.14) and (6.15) are satisfied if and only if $c \geq \frac{3}{2}$. So, by applying Result 5 with $c = \frac{3}{2}$, we are immediately led to the following result:

The linear ordinary differential equation

$$x''(t) + \frac{1}{2(t+1)^4}x(t) + \frac{1}{2(t+1)^4}x'(t) = 0$$

has exactly one solution x on the interval $[0, \infty)$ satisfying

$$|x(0)| \leq \frac{3}{2}$$

and

$$|x'(t)| \leq \frac{3}{2} \quad \textit{for every } t \geq 0,$$

and such that (6.22) *and* (6.23) *hold; in addition, this unique solution x satisfies* (6.25) *and* (6.26).

Example 6. Consider the linear ordinary differential equation (6.6) with

$$a(t) = \frac{2}{5(t+1)^4} \quad \text{for } t \geq 0.$$

Since

$$\int_0^\infty t\frac{2}{5(t+1)^3}dt = \frac{1}{5},$$

we see that (6.16) is always satisfied. Now, take $\xi = \frac{6}{5}$ and $\eta = 1$. Inequality (6.17) is written as

$$c\int_0^\infty t\frac{2}{5(t+1)^3}dt \leq c - \frac{6}{5},$$

i.e

$$\frac{1}{5}c \leq c - \frac{6}{5} \quad \text{or} \quad c \geq \frac{3}{2}.$$

So, Inequality (6.17) holds true if and only if $c \geq \frac{3}{2}$. Hence, an application of Result 6 with $c = \frac{3}{2}$ gives the next result:

The linear ordinary differential equation

$$x''(t) + \frac{2}{5(t+1)^4}x(t) = 0$$

has exactly one solution x on the interval $[0, \infty)$ satisfying

$$|x(t)| \leq \frac{3}{2}(t+1) \quad \textit{for every } t \geq 0,$$

and such that (6.28) *and* (6.29) *hold; in addition, this unique solution x satisfies* (6.30) *and* (6.31).

Finally, we give an example related to our comment at the end of Section 2.

Example 7. Consider the linear ordinary differential equation

$$x''(t) + \frac{e^{-6t+1}}{t+1}x(t) = 0. \tag{6.32}$$

This equation is of the form (6.4) with $\alpha = 1$ and

$$a(t) = \frac{e^{-6t+1}}{t+1} \quad \text{for } t \geq 0.$$

Take $\xi = \frac{7}{10}$ and $\eta = \frac{9}{10}$, and choose $c = 1$. Then, Inequality (6.12) becomes

$$e \int_0^\infty t e^{-6t} dt \leq \frac{1}{10}, \quad \text{i.e.} \quad \frac{e}{36} \leq \frac{1}{10}.$$

Thus, (6.12) holds true. Consequently, Result 4 guarantees the following:

The linear ordinary differential equation (6.32) *has at least one solution x on the interval* $[0, \infty)$ *such that*

$$x(t) = \frac{7}{10}t + \frac{9}{10} + o(1) \quad \textit{for } t \to \infty$$

and

$$x'(t) = \frac{7}{10} + o(1) \quad \textit{for} \quad t \to \infty;$$

in addition, this solution x satisfies

$$\frac{7}{10}t + \frac{4}{5} \leq x(t) \leq \frac{7}{10}t + 1 \quad \textit{for every } t \geq 0$$

and

$$\frac{7}{10} - \frac{e}{6} \leq x'(t) \leq \frac{7}{10} + \frac{e}{6} \quad \textit{for every } t \geq 0.$$

Now, we observe that Equation (6.32) can also be obtained from (6.3) by taking $\alpha = \beta = 1$, and

$$a(t) = \frac{e^{-6t+1}}{t+1} \quad \text{for } t \geq 0, \qquad b(t) = 0 \quad \text{for } t \geq 0.$$

Again, we take $\xi = \frac{7}{10}$ and $\eta = \frac{9}{10}$, and we choose $c = 1$. Then, Inequality (6.11) is written

$$e \int_0^\infty e^{-6t} dt \leq \frac{3}{10}, \quad \text{i.e.} \quad \frac{e}{6} \leq \frac{3}{10}.$$

So, (6.11) fails to hold. Thus, Result 3 is not applicable. Consequently, the above result for the linear ordinary differential equation (6.32) cannot be obtained from Result 3.

7. SOME SUPPLEMENTARY RESULTS

The results of this section are formulated as two theorems (Theorems I and II) and two corollaries (Corollaries I and II). Corollaries I and II are immediate consequences of Theorems I and II, respectively. Theorem I and Corollary I concern the delay differential equation (E), while Theorem II and Corollary II are dealing with the delay differential equation (E_0). It must be noted that Theorem I and Corollary I can be applied, in particular, to the delay differential equation (E') and, especially, to the ordinary differential equation (D); analogously, Theorem II and Corollary II are applicable to the particular case of the delay differential equation (E_0') as well as to the special case of the ordinary differential equation (D_0). These applications are left to the reader.

Theorem I. *Assume that* (2.1) *holds, where F is a nonnegative real-valued function defined on* $[0,\infty)\times C([-r,0],[0,\infty))\times[0,\infty)$*, which satisfies the Continuity Condition* (C). *Suppose that* (B) *is satisfied.*

Let c be a given positive real number such that (3.1) *holds, where the function γ in* $C([-r,\infty),[0,\infty))$ *depends on c and is defined by* (2.4). *Then every solution x on the interval* $[0,\infty)$ *of the delay differential equation* (E) *with*

$$\max\left\{\max_{-r\le t\le 0}|x(t)|\,,\ \sup_{t\ge 0}|x'(t)|\right\}\le c \tag{7.1}$$

satisfies (2.5) *and* (2.6), *where the real constants ξ and η depend on the solution x and are defined as follows:*

$$\xi = x'(0)-\int_0^\infty f(t,x_t,x'(t))dt \tag{7.2}$$

and

$$\eta = x(0)+\int_0^\infty tf(t,x_t,x'(t))dt. \tag{7.3}$$

Corollary I. *Assume that* (2.1) *holds, where F is a nonnegative real-valued function defined on* $[0,\infty)\times C([-r,0],[0,\infty))\times[0,\infty)$*, which satisfies the Continuity Condition* (C). *Suppose that* (B) *is satisfied.*

Assume that, for any positive real number c, (3.1) *holds, where the function γ in* $C([-r,\infty),[0,\infty))$ *depends on c and is defined by* (2.4). *Then every solution x on the interval* $[0,\infty)$ *of the delay differential equation* (E) *with bounded derivative on* $[0,\infty)$ *satisfies* (2.5) and (2.6), *where the real constants ξ and η depend on the solution x and are defined by* (7.2) *and* (7.3), *respectively.*

Theorem II. *Assume that* (2.10) *holds, where F_0 is a nonnegative real-valued function defined on* $[0,\infty)\times C([-r,0],[0,\infty))$*, which satisfies the Continuity Condition* (C_0). *Suppose that* (B_0) *is satisfied.*

Let c be a given positive real number such that (4.1) *holds, where the function γ in* $C([-r,\infty),[0,\infty))$ *depends on c and is defined by* (2.4). *Then every solution x on the interval* $[0,\infty)$ *of the delay differential equation* (E_0) *with*

$$\max\left\{\max_{-r\le t\le 0}|x(t)|\,,\ \sup_{t\ge 0}\frac{|x(t)|}{t+1}\right\}\le c \tag{7.4}$$

satisfies (2.5) *and* (2.6), *where the real constants ξ and η depend on the solution x and are defined as follows:*

$$\xi = x'(0)-\int_0^\infty f_0(t,x_t)dt \tag{7.5}$$

and

$$\eta = x(0)+\int_0^\infty tf_0(t,x_t)dt. \tag{7.6}$$

Corollary II. *Assume that* (2.10) *holds, where F_0 is a nonnegative real-valued function defined on* $[0,\infty)\times C([-r,0],[0,\infty))$*, which satisfies the Continuity Condition* (C_0). *Suppose that* (B_0) *is satisfied.*

Assume that, for any positive real number c, (4.1) *holds, where the function* γ *in* $C([-r,\infty),[0,\infty))$ *depends on c and is defined by* (2.4). *Then every solution x on the interval* $[0,\infty)$ *of the delay differential equation* (E_0) *with* $x(t)=O(t)$ *for* $t\to\infty$ *satisfies* (2.5) *and* (2.6), *where the real constants* ξ *and* η *depend on the solution x and are defined by* (7.5) *and* (7.6), *respectively.*

Proof of Theorem I. Let x be a solution on the interval $[0,\infty)$ of the delay differential equation (E) such that (7.1) is satisfied. It follows from (7.1) that x satisfies (3.4) and (3.5). As in the proof of Proposition 1, we can arrive at (3.11), which guarantees that (7.2) and (7.3) define two real constants ξ and η, respectively, depending on the solution x.

Now, from (E) it follows that

$$x(t)=x(0)+tx'(0)-\int_0^t (t-s)f(s,x_s,x'(s))ds \quad \text{for } t\geq 0. \tag{7.7}$$

For every $t\geq 0$, we obtain

$$\begin{aligned}
&-\int_0^t (t-s)f(s,x_s,x'(s))ds\\
=\ &\int_0^t (t-s)d\left[\int_s^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma\right]\\
=\ &-t\int_0^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma+\int_0^t\left[\int_s^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma\right]ds\\
=\ &-t\int_0^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma+\int_0^\infty\left[\int_s^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma\right]ds\\
&-\int_t^\infty\left[\int_s^\infty f(\sigma,x_\sigma,x'(\sigma))d\sigma\right]ds\\
=\ &-t\int_0^\infty f(s,x_s,x'(s))ds+\int_0^\infty sf(s,x_s,x'(s))ds-\int_t^\infty (s-t)f(s,x_s,x'(s))ds.
\end{aligned}$$

Thus, (7.7) gives

$$\begin{aligned}
x(t) &= x(0)+tx'(0)-t\int_0^\infty f(s,x_s,x'(s))ds+\int_0^\infty sf(s,x_s,x'(s))ds\\
&\quad -\int_t^\infty (s-t)f(s,x_s,x'(s))ds\\
&= \left[x'(0)-\int_0^\infty f(s,x_s,x'(s))ds\right]t+\left[x(0)+\int_0^\infty sf(s,x_s,x'(s))ds\right]\\
&\quad -\int_t^\infty (s-t)f(s,x_s,x'(s))ds.
\end{aligned}$$

Hence, in view of (7.2) and (7.3), we have

$$x(t)=\xi t+\eta-\int_t^\infty (s-t)f(s,x_s,x'(s))ds \quad \text{for } t\geq 0. \tag{7.8}$$

But, (3.11) ensures that

$$\lim_{t\to\infty}\int_t^\infty (s-t)f(s,x_s,x'(s))ds=0.$$

So, it follows from (7.8) that the solution x satisfies (2.5). Furthermore, from (7.8) we obtain

$$x'(t) = \xi + \int_t^{\infty} f(s, x_s, x'(s))ds \quad \text{for } t \geq 0. \tag{7.9}$$

But, because of (3.11), it holds

$$\lim_{t\to\infty} \int_t^{\infty} f(s, x_s, x'(s))ds = 0.$$

Thus, (7.9) implies that x satisfies (2.6).

The proof of the theorem has been completed.

Proof of Theorem II. Let x be a solution on the interval $[0, \infty)$ of the delay differential equation (E_0), which satisfies (7.4). We immediately observe that (7.4) guarantees that the solution x is such that (3.4) and (3.6) hold. As in the proof of Proposition 2, we can conclude that (4.7) holds true and hence (7.5) and (7.6) define two real constants ξ and η, respectively, which depend on the solution x.

The rest of the proof of the theorem is similar with the corresponding part of the proof of Theorem I, and will be omitted.

Before closing this section and ending the paper, we note that the problem of giving sufficient conditions for every solution to be asymptotic at ∞ to a line (depending on the solution) has recently been investigated in [10], for second order nonlinear ordinary differential equations. For the more general case of n-th order ($n > 1$) nonlinear ordinary differential equations, conditions have been established in [17,18], which guarantee that every solution is asymptotic at ∞ to a real polynomial of degree at most $n - 1$ (depending on the solution).

REFERENCES

[1] C. AVRAMESCU, Sur l' existence des solutions convergentès de systèmes d' équations différentielles non linéaires, *Ann. Mat. Pura Appl.* **81** (1969), 147-168.

[2] A. CONSTANTIN, On the existence of positive solutions of second order differential equations, *Ann. Mat. Pura Appl.* **184** (2005), 131-138.

[3] O. DIEKMANN, S. A. VAN GILS, S. M. VERDUYN LUNEL AND H.-O. WALTHER, *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Springer-Verlag, New York, 1995.

[4] R. D. DRIVER, *Ordinary and Delay Differential Equations*, Springer-Verlag, New York, 1977.

[5] S. G. DUBÉ AND A. B. MINGARELLI, Note on a non-oscillation theorem of Atkinson, *Electron. J. Differential Equations* **2004** (2004), No. 22, pp. 1-6.

[6] J. K. HALE AND S. M. VERDUYN LUNEL, *Introduction to Functional Differential Equations,* Springer-Verlag, New York, 1993.

[7] A. G. KARTSATOS, *Advanced Ordinary Differential Equations*, Mancorp Publishing, Tampa, 1993.

[8] T. KUSANO AND W. F. TRENCH, Global existence theorems for solutions of nonlinear differential equations with prescribed asymptotic behavior, *J. London Math. Soc.* **31** (1985), 478-486.

[9] T. KUSANO AND W. F. TRENCH, Existence of global solutions with prescribed asymptotic behavior for nonlinear ordinary differential equations, *Ann. Mat. Pura Appl.* **142** (1985), 381-392.

[10] O. LIPOVAN, On the asymptotic behaviour of the solutions to a class of second order nonlinear differential equations, *Glasg. Math. J.* **45** (2003), 179-187.

[11] Y. LIU, Existence and unboundedness of positive solutions for singular boundary value problems on half-line, *Appl. Math. Comput.* **144** (2003), 543-556.

[12] K. G. MAVRIDIS, CH. G. PHILOS AND P. CH. TSAMATOS, Existence of solutions of a boundary value problem on the half-line to second order nonlinear delay differential equations, *Arch. Math. (Basel)* **86** (2006), 163-175.

[13] K. G. MAVRIDIS, CH. G. PHILOS AND P. CH. TSAMATOS, Multiple positive solutions for a second order delay boundary value problem on the half-line, *Ann. Polon. Math.* **88** (2006), 173-191.

[14] O. G. MUSTAFA, Positive solutions of nonlinear differential equations with prescribed decay of the first derivative, *Nonlinear Anal.* **60** (2005), 179-185.

[15] O. G. MUSTAFA AND Y. V. ROGOVCHENKO, Global existence of solutions with prescribed asymptotic behavior for second-order nonlinear differential equations, *Nonlinear Anal.* **51** (2002), 339-368.

[16] O. G. MUSTAFA AND Y. V. ROGOVCHENKO, Global existence and asymptotic behavior of solutions of nonlinear differential equations, *Funkcial. Ekvac.* **47** (2004), 167-186.

[17] CH. G. PHILOS, I. K. PURNARAS AND P. CH. TSAMATOS, Asymptotic to polynomials solutions for nonlinear differential equations, *Nonlinear Anal.* **59** (2004), 1157-1179.

[18] CH. G. PHILOS AND P. CH. TSAMATOS, Solutions approaching polynomials at ∞ to nonlinear ordinary differential equations, *Electron. J. Differential Equations* **2005** (2005), No. 79, pp. 1-25.

[19] S. P. ROGOVCHENKO AND Y. V. ROGOVCHENKO, Asymptotic behavior of solutions of second order nonlinear differential equations, *Portugal. Math.* **57** (2000), 17-33.

[20] S. P. ROGOVCHENKO AND Y. V. ROGOVCHENKO, Asymptotic behavior of certain second order nonlinear differential equations, *Dynam. Systems Appl.* **10** (2001), 185-200.

[21] J. SCHAUDER, Der Fixpuktsatz in Funktionalräumen, *Studia Math.* **2** (1930), 171-180.

[22] B. YAN, Multiple unbounded solutions of boundary value problems for second-order differential equations on the half-line, *Nonlinear Anal.* **51** (2002), 1031-1044.

[23] B. YAN AND Y. LIU, Unbounded solutions of the singular boundary value problems for second order differential equations on the half-line, *Appl. Math. Comput.* **147** (2004), 629-644.

[24] Z. YIN, Monotone positive solutions of second-order nonlinear differential equations, *Nonlinear Anal.* **54** (2003), 391-403.

[25] E. WAHLÉN, Positive solutions of second-order differential equations, *Nonlinear Anal.* **58** (2004), 359-366.

[26] Z. ZHAO, Positive solutions of nonlinear second order ordinary differential equations, *Proc. Amer. Math. Soc.* **121** (1994), 465-469.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF IOANNINA, P.O. BOX 1186, 451 10 IOANNINA, GREECE

E-mail address: cphilos@cc.uoi.gr (Ch. G. Philos)

E-mail address: ipurnara@cc.uoi.gr (I. K. Purnaras)

E-mail address: ptsamato@cc.uoi.gr (P. Ch. Tsamatos)

Vector-valued p-adic Measures

A. K. Katsaras

Key words and phrases: Non-Archimedean fields, zero-dimensional spaces, p-adic measures, locally convex spaces.

2000 *Mathematics Subject Classification: 46S10, 46G10*

1 Preliminaries

Throughout this paper, $\mathbb{K}$ stands for a complete non-Archimedean valued field whose valuation is non-trivial. By a seminorm, on a vector space E over $\mathbb{K}$, we mean a non-Archimedean seminorm. Also by a locally convex space we will mean a non-Archimedean locally convex space over $\mathbb{K}$ (see [12] and [13]). For E a locally convex space, we denote by $cs(E)$ the collection of all continuous seminorms on E and by E' the topological dual of E. For a zero-dimensional Hausdorff topological space X, $\beta_o X$ is the Banachewski compactification of X, $C_b(X)$ the space of all continuous $\mathbb{K}$-valued functions on X and $C_{rc}(X)$ the space of all $f \in C_b(X)$ whose range is relatively compact. Every $f \in C_{rc}(X)$ has a continuous extension f^{β_o} to all of $\beta_o X$. For $f \in \mathbb{K}^X$ and $A \subset X$, we define

$$\|f\|_A = \sup\{|f(x)| : x \in A\} \quad and \quad \|f\| = \|f\|_X.$$

By $\overline{A}^{\beta_o X}$ we will denote the closure of A in $\beta_o X$.
Next we will recall the definition of the strict topology β on $C_b(X)$ which was given in [5]. Let Ω be the family of all compact subsets of $\beta_o X$ which are disjoint from X. For $Z \in \Omega$, let C_Z be the set of all $h \in C_{rc}(X)$ for which h^{β_o} vanishes on Z. We denote by β_Z the locally convex topology on $C_b(X)$ generated by the seminorms p_h, $h \in C_Z$, where $p_h(f) = \|hf\|$. The inductive limit of the topologies β_Z, $Z \in \Omega$, is the strict topology β. As it is shown in [7], Theorem 2.2, an absolutely convex subset W of $C_b(X)$ is a β_Z-neighborhood of zero iff, for each $r > 0$, there exist a clopen subset A of X, with $\overline{A}^{\beta_o X}$ disjoint from Z, and $\epsilon > 0$ such that

$$\{f \in C_b(X) : \|f\|_A \le \epsilon, \|f\| \le r\} \subset W.$$

Monna and Springer initiated in [11] non-Archimedean integration. In [13] and [14], van Rooij and Schikhof developed a non-Archimedean integration theory for scalar valued measures. Some results on measures with values in Banach spaces were given in [1], [2] and [3]. In this paper we will study measures with values in a locally convex space as well as integrals of scalar valued functions with respect to such measures.

2 Vector Measures

Let $\mathcal{R}$ be a separating algebra of subsets of a non-empty set X, i.e. $\mathcal{R}$ is a family of subsets of X with the following properties :

1. $X \in \mathcal{R}$ and, if $A, B \in \mathcal{R}$, then $A \cup B$, $A \cap B$, $A \setminus B$ are also in $\mathcal{R}$.

2. If x, y are distinct elements of X, then there exists a member of $\mathcal{R}$ containing x but not y.

We will call the members of $\mathcal{R}$ measurable sets. Clearly $\mathcal{R}$ is a base for a Hausdorff zero-dimensional topology $\tau_{\mathcal{R}}$ on X.
For a net (V_δ) of subsets of X we will write $V_\delta \downarrow \emptyset$ if it is decreasing and $\bigcap V_\delta = \emptyset$. Similarly we will write $V_n \downarrow \emptyset$ for a sequence (V_n) of sets which decreases to the empty set.
Let now E be a Hausdorff locally convex space. We denote by $M(\mathcal{R}, E)$ the space of all bounded finitely-additive measures $m : \mathcal{R} \to E$. For $m \in M(\mathcal{R}, E)$ and $p \in cs(E)$, we define

$$m_p : \mathcal{R} \to \mathbf{R}, \quad m_p(A) = \sup\{p(m(V)) : V \in \mathcal{R}, V \subset A\}$$

and $\|m\|_p = m_p(X)$. We also define

$$N_{m,p} : X \to \mathbf{R}, \quad N_{m,p}(x) = \inf\{m_p(V) : x \in V \in \mathcal{R}\}.$$

An element m of $M(\mathcal{R}, E)$ is called :

1. σ-additive if $m(V_n) \to 0$ if $V_n \downarrow \emptyset$.

2. τ-additive if $m(V_\delta) \to 0$ if $V_\delta \downarrow \emptyset$.

Let $M_\sigma(\mathcal{R}, E)$ (resp. $M_\tau(\mathcal{R}, E)$) be the space of all σ-additive (resp. τ-additive) members of $M(\mathcal{R}, E)$.

Theorem 2.1 *Let $m \in M(\mathcal{R}, E)$. Then*

1. m is τ-additive iff, for all $p \in cs(E)$, we have that $m_p(V_\delta) \to 0$ when $V_\delta \downarrow \emptyset$.

2. m is σ-additive iff, for all $p \in cs(E)$, we have that $m_p(V_n) \to 0$ when $V_n \downarrow \emptyset$.

Proof : (1). Clearly the condition is sufficient. Conversely, assume that m is τ-additive but the condition is not satisfied. Then there exist a $p \in cs(E)$, an $\epsilon > 0$ and a net $(V_\delta)_{\delta \in \Delta}$ of measurable sets which decreases to the empty set such that $m_p(V_\delta) > \epsilon$ for all δ.
Claim : For each $\delta \in \Delta$, there exist $\gamma \geq \delta$ and a measurable set A such that $V_\gamma \subset A \subset V_\delta$ and $p(m(A)) > \epsilon$. Indeed, there exists $B \subset V_\delta$ with $p(m(B)) > \epsilon$. For each $\gamma \in \Delta$, set $Z_\gamma = B \cap V_\gamma$, $W_\gamma = V_\gamma \setminus Z_\gamma$. Then $W_\gamma \downarrow \emptyset$. Since m is τ-additive, there exists $\gamma \geq \delta$ such that $p(m(W_\gamma)) < \epsilon$. The sets B and W_γ are disjoint. If $A = W_\gamma \cup B$, then $V_\gamma \subset A \subset V_\delta$ and

$$p(m(A)) = p(m(W_\gamma) + m(B)) = p(m(B)) > \epsilon,$$

which proves the claim.
Let now $\mathcal{F}$ be the family of all measurable sets A such that there are $\gamma \geq \delta$ with $V_\gamma \subset A \subset V_\delta$ and $p(m(A)) > \epsilon$. Since $\mathcal{F} \downarrow \emptyset$, we arrived at a contradiction. This proves (1).
(2). The proof is analogous to that of (1).

Theorem 2.2 *Let $m \in M_\tau(\mathcal{R}, E)$ and let $(V_i)_{i\in I}$ be a family of measurable sets. If $p \in cs(E)$, then for each measurable subset V of $\bigcup_{i\in I} V_i$, we have that*

$$m_p(V) \leq \sup_i m_p(V_i).$$

Proof : For each finite subset S of I, let $W_S = \bigcup_{i\in S} V_i$. Then $V \bigcap W_S^c \downarrow \emptyset$. If $m_p(V) > 0$, there exists a finite subset S of I such that $m_p(V \bigcap W_S^c) < m_p(V)$. Now

$$\begin{aligned} m_p(V) &= \max\{m_p(V \cap W_S), \quad m_p(V \cap W_S^c)\} \\ &= m_p(V \cap W_S) \leq m_p(W_S) = \max_{i\in S} m_p(V_i). \end{aligned}$$

Corollary 2.3 *Let $m \in M_\tau(\mathcal{R}, E)$, $p \in cs(E)$ and $V \in \mathcal{R}$. Then*

$$m_p(V) = \sup_{x\in V} N_{m,p}(x).$$

Proof : Clearly $m_p(V) \geq \alpha = \sup_{x\in V} N_{m,p}(x)$. On the other hand, if $\epsilon > 0$, then for each $x \in V$ there exists a measurable set V_x, with $x \in V_x \subset V$, such that $m_p(V_x) < N_{m,p}(x) + \epsilon \leq \alpha + \epsilon$. Since $V = \bigcup_{x\in V} V_x$, we have that

$$m_p(V) \leq \sup_{x\in V} m_p(V_x) \leq \alpha + \epsilon,$$

and the result follows as $\epsilon > 0$ was arbitrary.

Theorem 2.4 *Let $m \in M_\sigma(\mathcal{R}, E)$ and let (V_n) be a sequence of measurable sets. If $V \in \mathcal{R}$ is contained in $\bigcup V_n$, then $m_p(V) \leq \sup_n m_p(V_n)$.*

Proof : Let $W_n = \bigcup_{k=1}^n V_k$. Suppose that $m_p(V) > 0$. Since $V \bigcap W_n^c \downarrow \emptyset$, there exists an n such that $m_p(V \cap W_n^c) < m_p(V)$. Now

$$\begin{aligned} m_p(V) &= \max\{m_p(V \cap W_n^c), m_p(V \cap W_n)\} \\ &= m_p(V \cap W_n) \leq m_p(W_n) = \max_{1\leq k\leq n} m_p(V_k). \end{aligned}$$

Theorem 2.5 *If $m \in M(\mathcal{R}, E)$ and $p \in cs(E)$, then $N_{m,p}$ is upper semicontinuous.*

Proof: Let $\alpha > 0$ and $V = \{x : N_{m,p}(x) < \alpha\}$. For $x \in V$, there exists a measurable set A containing x and such that $m_p(A) < \alpha$. Now $x \in A \subset V$ and so V is open.

Theorem 2.6 *Let $m \in M_\tau(\mathcal{R}, E)$, $p \in cs(E)$ and $\epsilon > 0$. Then the set*

$$X_{p,\epsilon} = \{x : N_{m,p}(x) \geq \epsilon\}$$

is $\tau_{\mathcal{R}}$-compact.

Proof : Let $(V_i)_{i\in I}$ be a family of measurable sets covering $X_{p,\epsilon} = Y$. Since $N_{m,p}$ is upper semicontinuous, the set Y is closed. For each finite subset S of I, let $W_S = \bigcup_{i\in S} V_i$. Consider the family $\mathcal{F}$ of all measurable sets of the form $[W_S \cup V]^c$, where V is a measurable set disjoint from Y and S a finite subset of I. Then $\mathcal{F}$ is downwards directed and $\bigcap \mathcal{F} = \emptyset$. Since m is τ-additive, there are S and V such that $m_p([W_S \cup V]^c) < \epsilon$. But then $[W_S \cup V]^c \subset Y^c$, and thus $Y \subset W_S \cup V$, which implies that $Y \subset W_S$. This completes the proof.

Definition 2.7 *A subset G of X is said to be a support set of an $m \in M(\mathcal{R}, E)$ if $m(V) = 0$ for each measurable set V disjoint from G.*

Theorem 2.8 *Let $m \in M_\tau(\mathcal{R}, E)$. Then the set*

$$supp(m) = \overline{\bigcup_{p\in cs(E)} \{x : N_{m,p}(x) > 0\}}$$

is the smallest of all closed support sets of m.

Proof : If V is a measurable set disjoint from $supp(m)$, then for each $p \in cs(E)$ we have

$$p(m(V)) \le m_p(V) = \sup_{x\in V} N_{m,p}(x) = 0,$$

which proves that $supp(m)$ is a support set of m since E is Hausdorff. On the other hand, let F be a closed support set of m. Given $x \in F^c$, there exists $V \in \mathcal{R}$ with $x \in V \subset F^c$. Now, for each $p \in cs(E)$ and $y \in V$, we have that $N_{m,p}(y) \le m_pV) = 0$ and so the set

$$B = \bigcup_{p\in cs(E)} \{x : N_{m,p}(x) \neq 0\}$$

does not intersect V, which implies that $x \notin \overline{B} = supp(m)$. Thus $supp(m) \subset F$ and the result follows.

3 A Universal Measure

Let $\mathcal{R}$ be a separating algebra of subsets of X and let $S(\mathcal{R})$ be the vector space of all $\mathbb{K}$-valued $\mathcal{R}$-simple functions on X. Let

$$\chi : \mathcal{R} \to S(\mathcal{R}), \quad A \mapsto \chi_A.$$

Let E be a Hausdorff locally convex space. Every $m \in M(\mathcal{R}, E)$ induces a linear map

$$\hat{m} : S(\mathcal{R}) \to E, \quad \hat{m}\left(\sum_{k=1}^n \lambda_k \chi_{V_k}\right) = \sum_{k=1}^n \lambda_k m(V_k).$$

On $S(\mathcal{R})$ we consider the locally convex topologies ϕ, ϕ_σ, ϕ_τ defined as follows :

1. ϕ is the weakest locally convex topology for which, for each Hausdorff locally convex space E and each $m \in M(\mathcal{R}, E)$, the map $\hat{m} : S(\mathcal{R}) \to E$ is continuous.

2. ϕ_σ is the weakest locally convex topology for which, for each Hausdorff locally convex space E and each $m \in M_\sigma(\mathcal{R}, E)$, the map $\hat{m} : S(\mathcal{R}) \to E$ is continuous.

3. ϕ_τ is the weakest locally convex topology for which, for each Hausdorff locally convex space E and each $m \in M_\tau(\mathcal{R}, E)$, the map $\hat{m} : S(\mathcal{R}) \to E$ is continuous.

Clearly $\phi_\tau \subset \phi_\sigma \subset \phi$.

Lemma 3.1 *The topology ϕ_τ is Hausdorff.*

Proof: Every $x \in X$ defines a τ-additive measure

$$m_x : \mathcal{R} \to \mathbb{K}, \quad m_x(A) = \chi_A(x).$$

Let $g \in S(\mathcal{R})$, $g \neq 0$ and let $g(x) \neq 0$. Let $0 < \epsilon < |g(x)|$. The set

$$\{h \in S(\mathcal{R}) : |\hat{m}_x(h)| = |h(x)| < \epsilon\}$$

is a ϕ_τ-neighborhood of zero not containing g.

Theorem 3.2 *If $F = (S(\mathcal{R}), \rho)$, where $\rho = \phi$, ϕ_σ or ϕ_τ, then $\chi : \mathcal{R} \to F$ is a member of $M(\mathcal{R}, F)$, $M_\sigma(\mathcal{R}, F)$ or $M_\tau(\mathcal{R}, F)$, respectively.*

Proof : Assume that $F = (S(\mathcal{R}), \phi_\tau)$. Clearly χ is finitely additive. Let E be a Hausdorff locally convex space and let $m \in M_\tau(\mathcal{R}, E)$, $p \in cs(E)$. Let

$$W = \{s \in E : p(s) \leq 1\}.$$

Since $m \in M_\tau(\mathcal{R}, E)$, there exists $\lambda \in \mathbb{K}$ such that $m(\mathcal{R}) \subset \lambda W$. If

$$D = \{g \in S(\mathcal{R}) : \hat{m}(g) \in W\},$$

then $\chi(\mathcal{R}) \subset \lambda D$, which proves that $\chi : \mathcal{R} \to F$ is bounded. If (V_δ) is a net of measurable sets with $V_\delta \downarrow \emptyset$, then $m(V_\delta) \to 0$, and so $m(V_\delta) \in W$ eventually, which implies that $\chi_{V_\delta} \in D$ eventually. Thus $\chi \in M_\tau(\mathcal{R}, F)$. The proofs for the cases of ϕ and ϕ_σ are analogous.

Theorem 3.3 *Let E be a Hausdorff locally convex space. Then :*

1. *The map $m \mapsto \hat{m}$, from $M(\mathcal{R}, E)$ to the space $L((S(\mathcal{R}), \phi), E)$, of all continuous linear maps from $(S(\mathcal{R}), \phi)$ to E, is an algebraic isomorphism.*

2. *The map $m \mapsto \hat{m}$, from $M_\sigma(\mathcal{R}, E)$ to the space $L((S(\mathcal{R}), \phi_\sigma), E)$, is an algebraic isomorphism.*

3. *The map $m \mapsto \hat{m}$, from $M_\tau(\mathcal{R}, E)$ to the space $L((S(\mathcal{R}), \phi_\tau), E)$, is an algebraic isomorphism.*

Proof : (1) By the definition of ϕ, each $\hat{m}$ is continuous. On the other hand, let $u : (S(\mathcal{R},), \phi) \to E$ be a continuous linear map and take $m = u \circ \chi$. Then $m \in M(\mathcal{R}, E)$ and $\hat{m} = u$. The proofs of (2) and (3) are analogous.

Since, for every Hausdorff locally convex space E, every measure $m : \mathcal{R} \to E$ is of the form $m = u \circ \chi$, for some ϕ-continuous linear map u from $S(\mathcal{R})$ to E, we will refer to the measure $\chi : \mathcal{R} \to (S(\mathcal{R}), \phi)$ as a universal measure. Taking $\mathbb{K}$ in place of E and identifying each scalar measure μ on $\mathcal{R}$ by the corresponding linear functional $\hat{\mu}$, we get the following

Theorem 3.4 *The spaces $M(\mathcal{R}) = M(\mathcal{R}, \mathbb{K})$, $M_\sigma(\mathcal{R})$ and $M_\tau(\mathcal{R})$ are algebraically isomorphic with the spaces $(S(\mathcal{R}), \phi)'$, $(S(\mathcal{R}), \phi_\sigma)'$ and $(S(\mathcal{R}), \phi_\tau)'$, respectively.*

Theorem 3.5 *On the space $S(\mathcal{R})$, the topology ϕ is coarser than the topology τ_u of uniform convergence.*

Proof : Let E be a Hausdorff locally convex space and let $m \in M(\mathcal{R}, E)$. It suffices to show that $\hat{m} : (S(\mathcal{R}), \tau_u) \to E$ is continuous. Indeed, let $p \in cs(E)$. There exists $r > 0$ such that $p(m(A)) \leq r$ for all $A \in \mathcal{R}$. Now, for

$$V = \{g \in S(\mathcal{R}) : \|g\| \leq 1/r\},$$

we have that $p(\hat{m}(g)) \leq 1$ for all $g \in V$. Indeed, let $g \in V$, $g = \sum_{k=1}^{n} \lambda_k \chi_{A_k}$, where $A_1, \ldots, A_n$ are pairwise disjoint sets . Then $|\lambda_k| \leq 1/r$ and so

$$p(\hat{m}(g)) = p(\sum_{k=1}^{n} \lambda_k m(A_k)) \leq \max_k |\lambda_k| \cdot p(m(A_k)) \leq 1.$$

This completes the proof.

Theorem 3.6 *ϕ is the finest of all Hausdorff locally convex topologies ρ on $S(\mathcal{R})$ such that, for $F = (S(\mathcal{R}), \rho)$, the map $\chi : \mathcal{R} \to F$ is in $M(\mathcal{R}, F)$. Analogous results hold for ϕ_σ and ϕ_τ.*

Proof : Let ρ be a Hausdorff locally convex topology on $S(\mathcal{R})$ such that $\chi : \mathcal{R} \to (S(\mathcal{R}), \rho)$ is a bounded finitely additive measure. By the definition of ϕ , the linear map

$$\hat{\chi} : (S(\mathcal{R}), \phi) \to (S(\mathcal{R}), \rho)$$

is continuous. Since $\hat{\chi}$ is the identity map, it follows that ϕ is finer that ρ. Thus the result holds for ϕ. Analogous are the proofs for ϕ_σ and ϕ_τ.

Corollary 3.7 *On $S(\mathcal{R})$ the topology ϕ coincides with the topology τ_u of uniform convergence.*

Proof : It follows from Theorems 3.5 and 3.6 since $\chi : \mathcal{R} \to (S(\mathcal{R}), \tau_u)$ is a bounded finitely-additive measure.

Let $\sigma = \sigma(M(\mathcal{R}), S(\mathcal{R}))$. For a σ-bounded subset H of $M(\mathcal{R})$, we denote by H_σ the set H equipped with the topology induced by σ. Let $C_b(H_\sigma)$ be the space of all bounded continuous $\mathbb{K}$-valued functions on H_σ endowed with the sup norm topology. For $A \in \mathcal{R}$, the function $m \mapsto m(A)$, $m \in H$, is σ-continuous. Also this function is bounded because H is σ-bounded. Hence we get a map

$$\mu = \mu_H : \mathcal{R} \to C_b(H_\sigma), \quad < \mu(A), m >= m(A).$$

Theorem 3.8 *For a subset H of $M(\mathcal{R})$, the following are equivalent :*

1. *H is ϕ-equicontinuous.*
2. *H is σ-bounded and the map $\mu = \mu_H : \mathcal{R} \to F = C_b(H_\sigma)$ is in $M(\mathcal{R}, F)$.*

Proof : (1) $\Rightarrow$ (2). Since H is ϕ-equicontinuous, it is σ-bounded. Clearly μ is finitely additive. We need to show that $\mu(\mathcal{R})$ is a norm bounded subset of $C_b(H_\sigma)$. Indeed, let V be a ϕ-neighborhood of zero in $S(\mathcal{R})$ such that $H \subset V^o$. Since $\chi : \mathcal{R} \to (S(\mathcal{R}), \phi)$ is a bounded measure, there exists a non-zero element λ of $\mathbb{K}$ such that $\chi_A \in \lambda V$ for all $A \in \mathcal{R}$. Thus, for $A \in \mathcal{R}$ and $m \in H$, we have that $|m(A)| \leq |\lambda|$ and hence $\|\mu(A)\| \leq |\lambda|$. Thus, $\sup_{A \in \mathcal{R}} \|\mu(A)\| \leq |\lambda|$, which proves that $\mu \in M(\mathcal{R}, F)$.
(2) $\Rightarrow$ (1). Since $\mu : \mathcal{R} \to F = C_b(H_\sigma)$ is a bounded finitely-additive measure, it follows that $\hat{\mu} : (S(\mathcal{R}), \phi) \to F$ is continuous. Thus, there exists a ϕ-neighborhood V of zero such that $\|\hat{\mu}(g)\| \leq 1$ for all $g \in V$. Then $H \subset V^o$ and the result follows.

Theorem 3.9 *For a subset H of $M_\sigma(\mathcal{R})$, the following are equivalent :*

1. *H is ϕ_σ-equicontinuous.*
2. *H is σ-bounded and the map $\mu = \mu_H : \mathcal{R} \to C_b(H_\sigma)$ is a σ-additive measure.*
3. *H is σ-bounded and uniformly σ-additive.*
4. $\sup_{m \in H} \|m\| < \infty$ *and H is uniformly σ-additive.*

Proof : (1) $\Rightarrow$ (2). Since $\phi_\sigma \subset \phi$, it follows that H is ϕ-equicontinuous and thus (by the preceding Theorem) $\mu : \mathcal{R} \to C_b(H_\sigma)$ is a bounded finitely-additive measure. We need to show that μ is σ-additive. So let (V_n) be a sequence of measurable sets which decreases to the empty set. Since H is ϕ_σ-equicontinuous, there exists a ϕ_σ-neighborhood V of zero in $S(\mathcal{R})$ such that $H \subset V^o$. Let $\lambda \neq 0$. As $\chi : \mathcal{R} \to (S(\mathcal{R}), \phi_\sigma)$ is a σ-additive measure, there exists n_o such that $\chi_{V_n} \in \lambda V$, for all $n \geq n_o$. Thus, for $n \geq n_o$ and $m \in H$, we have $|m(V_n)| \leq |\lambda|$ and thus $\|\mu(A_n\| \leq |\lambda|$, which proves that μ is σ-additive.
(2) $\Rightarrow$ (3). Let $V_n \downarrow \emptyset$. Since $\mu(V_n) \to 0$ in $C_b(H_\sigma)$, given $\epsilon > 0$, there exists n_o such that $\|\mu(V_n)\| \leq \epsilon$ for all $n \geq n_o$. Thus, for $n \geq n_o$, we have that $|m(V_n| \leq \epsilon$ for all $m \in H$, which proves that H is uniformly σ-additive.
(3) $\Rightarrow$ (2). It is trivial.
(2) $\Rightarrow$ (1). Since $\mu = \mu_H : \mathcal{R} \to C_b(H_\sigma)$ is a σ-additive measure, the map $\hat{\mu} : (S(\mathcal{R}), \phi_\sigma) \to F$ is continuous. Hence, there exists a ϕ_σ-neighborhood V of zero such that $\|\hat{\mu}(g)\| \leq 1$ for all $g \in V$. But then $H \subset V^o$.
(1) $\Rightarrow$ (4). Since ϕ_σ is coarser than the topology τ_u of uniform convergence, it follows that H is τ_u-equicontinuous and hence $\sup_{m \in H} \|m\| < \infty$. Also H is uniformly

σ-additive since (1) implies (3). This clearly completes the proof.

The proof of the next Theorem is analogous to the one of the preceding Theorem.

Theorem 3.10 *For a subset H of $M_\tau(\mathcal{R})$, the following are equivalent :*

1. *H is ϕ_τ-equicontinuous.*
2. *H is σ-bounded and the map $\mu = \mu_H : \mathcal{R} \to C_b(H_\sigma)$ is a τ-additive measure.*
3. *H is σ-bounded and uniformly τ-additive.*
4. *$\sup_{m\in H} \|m\| < \infty$ and H is uniformly τ-additive.*

Theorem 3.11 *ϕ_τ is the weakest of all locally convex topologies ρ on $S(\mathcal{R})$ such that, for each non-Archimedean Banach space E and each $m \in M_\tau(\mathcal{R}, E)$, the map $\hat{m} : (S(\mathcal{R}), \rho) \to E$ is continuous.*

Proof : Let τ_o be the weakest of all locally convex topologies ρ on $S(\mathcal{R})$ having the property mentioned in the Theorem. Clearly τ_o is coarser than ϕ_τ. On the other hand, let W be a polar ϕ_τ-neighborhood of zero and let H be the polar of W in $M_\tau(\mathcal{R})$. By the preceding Theorem,

$$\mu = \mu_H : \mathcal{R} :\to E = C_b(H_\sigma)$$

is a τ-additive measure. If V is the unit ball of E, then $(\hat{\mu})^{-1}(V)$ ia a τ_o-neighborhood of zero. Since $(\hat{\mu})^{-1}(V) \subset H^o = W$, the result clearly follows.

4 Integration

Throughout the rest of the paper we will assume that E is a complete Hausdorff locally convex space (unless it is stated otherwise) and $\mathcal{R}$ a separating algebra of subsets of a non-empty set X. Let $m \in M(\mathcal{R}, E)$ and $A \in \mathcal{R}$. Let $\mathcal{D}_A$ be the family of all $\alpha = \{A_1, A_2, \ldots, A_n; x_1, x_2, \ldots, x_n\}$, where $\{A_1, A_2, \ldots, A_n\}$ is a finite $\mathcal{R}$-partition of A and $x_i \in A_i$. We make $\mathcal{D}_A$ into a directed set by defining $\alpha_1 \geq \alpha_2$ iff the partition of A in α_1 is a refinement of the one in α_2. For $\alpha = \{A_1, A_2, \ldots, A_n; x_1, x_2, \ldots, x_n\} \in \mathcal{D}_A$ and $f \in \mathbb{K}^X$, we define

$$\omega_\alpha(f, m) = \sum_{k=1}^{n} f(x_k) m(A_k).$$

If the $\lim_\alpha \omega_\alpha(f, m)$ exists in E, we will say that f is m-integrable over A and denote this limit by $\int_A f\,dm$. For $A = X$, we write simply $\int f\,dm$. It is easy to see that, if f is m-integrable over X, then it is m-integrable over every measurable subset A and $\int_A f\,dm = \int f\chi_A\,dm$. If f is bounded on A, then $p\left(\int_A f\,dm\right) \leq \|f\|_A \cdot m_p(A)$ for every $p \in cs(E)$.

Using an argument analogous to the one used in [6], Theorem 2.1 for scalar-valued measures, we get the following

Theorem 4.1 *If $m \in M(\mathcal{R}, E)$, then an $f \in \mathbb{K}^X$ is m-integrable iff, for each $p \in cs(E)$ and each $\epsilon > 0$, there exists an $\mathcal{R}$-partition $\{A_1, A_2, \ldots, A_n\}$ of X such that $|f(x) - f(y)| \cdot m_p(A_i) \leq \epsilon$, for all i, if the x, y are in A_i. Moreover, in this case we have that*

$$p\left(\int f\, dm - \sum_{i=1}^{n} f(x_i)m(A_i)\right) \leq \epsilon.$$

Theorem 4.2 *Let $m \in M(\mathcal{R}, E)$ and let $f \in \mathbb{K}^X$ be m-integrable. Then :*

1. *f is continuous at every x in the set*

$$D = \bigcup_{p \in cs(E)} \{x : N_{m,p}(x) \neq 0\}.$$

2. *For each $p \in cs(E)$, there exists a measurable set A, with $m_p(A^c) = 0$, such that f is bounded on A.*

Proof : (1). Suppose that $N_{m,p}(x) = d > 0$ and let $\epsilon > 0$. There exists an $\mathcal{R}$-partition $\{A_1, A_2, \ldots, A_n\}$ of X such that $|f(x) - f(y)| \cdot m_p(A_i) \leq d\epsilon$, if $x, y \in A_i$. If $x \in A_i$, then $|f(y) - f(x)| \leq \epsilon$ for all $y \in A_i$.
(2). Let $\{A_1, A_2, \ldots, A_n\}$ be an $\mathcal{R}$-partition of X such that $|f(x) - f(y)| \cdot m_p(A_i) \leq 1$, if $x, y \in A_i$. Let

$$A = \bigcup \{A_i : m_p(A_i) > 0\}.$$

It follows easily that f is bounded on A and that $m_p(A^c) = 0$.

Theorem 4.3 *Let $m \in M(\mathcal{R}, E)$. If $f, g \in \mathbb{K}^X$ are m-integrable, then $h = fg$ is also m-integrable.*

Proof : Let $p \in cs(E)$ and $\epsilon > 0$. There are measurable sets A, B such that $m_p(A^c) = m_p(B^c) = 0$ and f, g are bounded on A, B, respectively. Let $D = A \cap B$. Then $m_p(D^c) = 0$ and there exists a $d > 0$ such that $\|f\|_D, \|g\|_D \leq d$. Now there exists an $\mathcal{R}$-partition $\{A_1, A_2, \ldots, A_n\}$ of X, which is a refinement of $\{D, D^c\}$, such that

$$|f(x) - f(y)| \cdot m_p(A_i) < \epsilon/d \quad \textit{and} \quad |g(x) - g(y)| \cdot m_p(A_i) < \epsilon/d$$

if $x, y \in A_i$. Let now $x, y \in A_i$. If $A_i \subset D^c$, then $|h(x) - h(y)| \cdot m_p(A_i) = 0$. For $A_i \subset D$, we have that

$$\begin{aligned} |h(x) - h(y)| &= |[f(x) - f(y)]g(x) + f(y)[g(x) - g(y)]| \\ &\leq \max\{d \cdot |f(x) - f(y)|, d \cdot |g(x) - g(y)|\} \end{aligned}$$

and so $|h(x) - h(y)| \cdot m_p(A_i < \epsilon$. This completes the proof in view of Theorem 4.1.

Let now $m \in M(\mathcal{R}, E)$ and let $g \in \mathbb{K}^X$ be m-integrable. Define

$$m_g : \mathcal{R} \to E, \quad m_g(A) = \int_A g\, dm.$$

Clearly m_g is finitely-additive. Also, m_g is bounded. In fact, let $p \in cs(E)$. There exists a measurable set B such that $m_p(B^c) = 0$ and g is bounded on B. Let $d = \|g\|_B$. Let $A \in \mathcal{R}$, $W_1 = A \cap B$, $W_2 = A \cap B^c$. Since g is m-integrable, there exists an $\mathcal{R}$-partition $\{V_1, V_2, \ldots, V_n\}$ of A, which is a refinement of $\{W_1, W_2\}$ such that $|g(x) - g(y)| \cdot m_p(V_i) < 1$ if $x, y \in V_i$. Let $x_i \in V_i$. Then

$$p\left(\int_A g\, dm - \sum_{k=1}^{n} g(x_i)m(V_i)\right) < 1.$$

If $V_i \subset W_1$, then $p\,(g(x_i)m(V_i)) \leq d \cdot m_p(X)$, while for $V_i \subset W_2$ we have that $p\,(g(x_i)m(V_i)) = 0$. Thus

$$p\left(\int_A g\, dm\right) \leq \max\{1,\, d \cdot m_p(X)\}.$$

This proves that m_g is bounded and hence $m_g \in M(\mathcal{R}, E)$.

Theorem 4.4 *Let $m \in M(\mathcal{R}, E)$ and let $g \in \mathbb{K}^X$ be m-integrable. If $f \in \mathbb{K}^X$ is m-integrable, then f is m_g-integrable and $\int f\, dm_g = \int fg\, dm$.*

Proof : Let $p \in cs(E)$. There exists a measurable set D, with $m_p(D^c) = 0$, such that f, g are bounded on D. Let $d > \max\{\|f\|_D, \|g\|_D\}$. If V is a measurable set contained in D^c, then $p(m_g(V)) = 0$. This follows from the fact that, for $A \subset V$ we have that $p(g(x)m(A)) = 0$. Let now $\epsilon > 0$ be given. There exists an $\mathcal{R}$-partition $\{V_1, V_2, \ldots, V_N\}$ of X , which is a refinement of $\{D, D^c\}$, such that

$$|f(x) - f(y)| \cdot m_p(V_i) < \epsilon/d, \quad \textit{and} \quad |g(x) - g(y)| \cdot m_p(V_i) < \epsilon/d$$

if $x, y \in V_i$. We may assume that $\bigcup_{i=1}^{n} V_i = D$. For $A \in \mathcal{R}$, $A \subset V_i \subset D$, we have

$$p\left(\int_A g\, dm\right) \leq \|g\|_A \cdot m_p(A) \leq d \cdot m_p(V_i),$$

and hence $(m_g)_p(V_i) \leq d \cdot m_p(V_i)$. Thus, for $x, y \in V_i \subset D$, we have

$$|f(x) - f(y)| \cdot (m_g)_p(V_i) \leq d \cdot |f(x) - f(y)| \cdot m_p(V_i) \leq \epsilon.$$

The same inequality holds when $V_i \subset D^c$. This proves that f is m_g-integrable. If $x, y \in V_i \subset D$, then

$$p\left(\int f\, dm_g - \sum_{k=1}^{n} f(x_k)m_g(V_k)\right) \leq \epsilon.$$

Since, for $x, y \in V_k \subset D$, we have $|g(x) - g(y)| \cdot m_p(V_k) \leq \epsilon/d$, it follows that

$$p\,(m_g(V_k) - g(x_k)m(V_k)) \leq \epsilon/d.$$

For $x, y \in V_k \subset D$, we have

$$|f(x)g(x) - f(y)g(y)| \cdot m_p(V_k) \leq m_p(V_k) \cdot \max\{|g(x)| \cdot |f(x) - f(y)|, |f(y)| \cdot |g(x) - g(y)|\} \leq \epsilon.$$

Since $m_p(V_k) = 0$ if $V_k \subset D^c$, we get that

$$p\left(\int gf\,dm - \sum_{k=1}^{n} g(x_k)f(x_k)m(V_k)\right) \leq \epsilon.$$

Also, for $1 \leq k \leq n$, we have $p\,(f(x_k)g(x_k)m(V_k) - f(x_k)m_g(V_k)) \leq \epsilon$. It follows that

$$p\left(\int gf\,dm - \int f\,dm_g\right) \leq \epsilon.$$

This, being true for all $\epsilon > 0$, and the fact that E is Hausdorff, imply that

$$\int gf\,dm = \int f\,dm_g,$$

which completes the proof.

Theorem 4.5 *Let $m \in M(\mathcal{R}, E)$, $p \in cs(E)$ and $x \in X$. If $g \in \mathbb{K}^X$ is m-integrable, then*

$$N_{m_g,p}(x) = |g(x)| \cdot N_{m,p}(x).$$

Proof : Let $\epsilon > 0$. There exists an $\mathcal{R}$-partition $\{V_1, V_2, \cdots, V_n\}$ of X such that $|g(y) - g(z)| \cdot m_p(V_i) \leq \epsilon$ if $y, z \in V_i$.
Claim I : If V is a measurable subset of V_i containing x, then, for each $A \subset V$, we have

$$p(m_g(A)) \leq \max\{\epsilon,\ |g(x)| \cdot m_p(V)\} = \theta.$$

Indeed, if $x \in A$, then for each $y \in A$ we have that $|g(x) - g(y)| \cdot m_p(A) \leq \epsilon$, which implies that $p(m_g(A) - g(x)m(A)) \leq \epsilon$ and so

$$p(m_g(A)) \leq \max\{\epsilon, |g(x)| \cdot p(m(A))\} \leq \theta.$$

In case $x \in V \backslash A$, we get in the same way that $p(m_g(V \backslash A)) \leq \theta$. Also $p(m_g(V)) \leq \theta$, since $x \in V$. Thus

$$p(m_g(A)) = p(m_g(V) - m_g(V \setminus A)) \leq \theta,$$

and the claim follows.
Claim II. If W is a measurable subset of V_i containing x, then for each measurable set $A \subset W$, we have that

$$|g(x)| \cdot p(m(A)) \leq \max\{\epsilon, (m_g)_p(W)\} = d.$$

Indeed, if $x \in A \subset W$, then $p(m_g(A) - g(x)m(A)) \leq \epsilon$ and so

$$|g(x)| \cdot p(m(A)) \leq \max\{\epsilon, p(m_g(A))\} \leq d.$$

If $x \in W \setminus A$, then $|g(x)| \cdot p(m(W \setminus A)) \leq d$. Also $g(x)| \cdot p(m(W)) \leq d$, and so again $|g(x)| \cdot p(m(A)) \leq d$, which proves the claim.
Now there are measurable subsets V, W of V_i containing x such that

$$m_p(V) < N_{m,p}(x) + \epsilon, \quad \textit{and} \quad (m_g)_p(W) < \epsilon + N_{m_g,p}(x).$$

By claim I, we have

$$N_{m_g,p}(x) \leq (m_g)_p(V) \leq \max\{\epsilon,\ |g(x)| \cdot m_p(V)\} \leq \max\{\epsilon,\ |g(x)|[\epsilon + N_{m,p}(x)]\}.$$

Taking $\epsilon \to 0$, we get that

$$N_{m_g,p}(x) \leq |g(x)| \cdot N_{m,p}(x).$$

Also

$$|g(x)| \cdot N_{m,p}(x) \leq |g(x)| \cdot m_p(W) \leq \max\{\epsilon, (m_g)_p(W)\} < \epsilon + N_{m_g,p}(x).$$

Taking $\epsilon \to 0$, we get that

$$|g(x)| \cdot N_{m,p}(x) \leq N_{m_g,p}(x),$$

which completes the proof.

Theorem 4.6 *Let $m \in M(\mathcal{R}, E)$ and let $g \in \mathbb{K}^X$ be m-integrable. If m is τ-additive (resp. σ-additive), then m_g is τ-additive (resp. σ-additive).*

Proof : Assume that m is τ-additive and let $V_\delta \downarrow \emptyset$ and $p \in cs(E)$ There exists an $A \in \mathcal{R}$ such that $m_p(A^c) = 0$ and $\|f\|_A = d < \infty$. Given $\epsilon > 0$, there exists a δ_o such that $m_p(V_\delta) < \epsilon/d$ if $\delta \geq \delta_o$. For a measurable set V disjoint from A, we have $p(m_g(V)) = 0$. Thus, for $\delta \geq \delta_o$, we have

$$p(m_g(V_\delta)) = p(m_g(V_\delta \cap A)) \leq \|g\|_{V_\delta \cap A} \cdot m_p(V_\delta \cap A) \leq d \cdot m_p(V_\delta) < \epsilon.$$

This proves that m_g is τ-additive. The proof for the σ-additive case is analogous.

The proof of the next Theorem is analogous to the one given in [8], Theorem 2.16, for scalar-valued measures.

Theorem 4.7 *Let $m \in M(\mathcal{R}, E)$. For a subset Z of X, the following are equivalent:*

1. *χ_Z is m-integrable.*
2. *For each $p \in cs(E)$ and each $\epsilon > 0$, there are measurable sets V, W such that $V \subset Z \subset W$ and $m_p(W \setminus V) < \epsilon$.*

For $m \in M(\mathcal{R}, E)$, let $\mathbb{R}_m$ be the family of all $A \subset X$ such that χ_A is m-integrable. Using the preceding Theorem, we show easily that $\mathbb{R}_m$ is a separating algebra of subsets of X which contains $\mathcal{R}$. Define

$$\overline{m} : \mathbb{R}_m \to E, \quad \overline{m}(A) = \int \chi_A \, dm.$$

The proofs of the next two Theorems are analogous to the corresponding ones for scalar valued measures (see [8], Lemma 2.18, Theorems 2.22, 2.23, 2.24, 2.26 and Corollary 2.25).

Theorem 4.8 *1. For $A \in \mathcal{R}$ and $p \in cs(E)$, we have $m_p(A) = \overline{m}_p(A)$.*

2. $\overline{m}$ is σ-additive iff m is σ-additive.

3. $\overline{m}$ is τ-additive iff m is τ-additive.

4. $N_{m,p} = N_{\overline{m},p}$.

5. $\mathbb{R}_m = \mathbb{R}_{\overline{m}}$.

Theorem 4.9 *1. If $f \in \mathbb{K}^X$ is m-integrable, then f is also $\overline{m}$-integrable and $\int f\, dm = \int f\, d\overline{m}$.*

2. If $f \in \mathbb{K}^X$ is $\overline{m}$-integrable and bounded, then f is also m-integrable.

Lemma 4.10 *If $m \in M_\tau(\mathcal{R}, E)$, then every $\tau_{\mathcal{R}}$-clopen set A is in $\mathbb{R}_m$.*

Proof : Let $p \in cs(E)$ and $\epsilon > 0$. Consider the collection $\mathcal{F}$ of all $\mathcal{R}$-measurable sets of the form $W \setminus V$, where $V, W \in \mathcal{R}$ and $V \subset A \subset W$. Then $\mathcal{F} \downarrow \emptyset$. As m is τ-additive, there exists an $W \setminus V \in \mathcal{F}$ such that $m_p(W \setminus V) < \epsilon$, which proves that $A \in \mathbb{R}_m$.

Theorem 4.11 *Let $m \in M_\tau(\mathcal{R}, E)$ and $f \in \mathbb{K}^X$. If f is bounded and $\tau_{\mathcal{R}}$-continuous, then f is m-integrable (and hence $\overline{m}$-integrable).*

Proof : Without loss of generality, we may assume that $\|f\| \leq 1$. Let $p \in cs(E)$ and $\epsilon > 0$. The set $Y = \{x : N_{m,p}(x) \geq \epsilon\}$ is $\tau_{\mathcal{R}}$-compact. Choose $0 < \epsilon_1 < \epsilon$ such that $\epsilon_1 . m_p(X) < \epsilon$. There are $x_1, x_2, \cdots, x_n \in Y$ such that the sets

$$A_k = \{x : p(f(x) - f(x_k)) \leq \epsilon_1\}, \quad k = 1, \cdots, n.$$

are pairwise disjoint and cover Y. Each A_k is $\tau_{\mathcal{R}}$-clopen and hence it is a member of $\mathbb{R}_m$. Let $V_k, W_k \in \mathcal{R}$ be such that $V_k \subset A_k \subset W_k$ and $m_p(W_k \setminus V_k) < \epsilon$. Let $V_{n+1} = (\bigcup_{k=1}^n V_k)^c$. Then V_{n+1} is disjoint from Y. Indeed, if $x \in Y \cap V_{n+1}$, then $x \in W_k$, for some k, and so $N_{m,p}(x) \leq m_p(W_k \setminus V_k) < \epsilon$, a contradiction. As m is τ-additive, we have that $m_p(V_{n+1}) = \sup_{x \in V_{n+1}} N_{m,p}(x) \leq \epsilon$. If now $x, y \in V_i$, $i \leq n$, then

$$|f(x) - f(y)| \cdot m_p(V_i) \leq \epsilon_1 . \cdot m_p(X) < \epsilon.$$

Also, if $x, y \in V_{n+1}$, then $|f(x) - f(y)| \cdot m_p(V_{n+1}) \leq \epsilon$. This proves that f is m-integrable.

Theorem 4.12 *Let $m \in M_\tau(\mathcal{R}, E)$. For a subset A of X, the following are equivalent :*

1. $A \in \mathbb{R}_m$.

2. A is $\tau_{\mathbb{R}_m}$-clopen.

Proof : Clearly (1) $\Rightarrow$ (2). On the other hand let A be $\tau_{\mathbb{R}_m}$-clopen. Since $\overline{m}$ is τ-additive, it follows (by Theorem 4.11) that χ_A is $\overline{m}$-integrable and hence χ_A is m-integrable (by Theorem 4.9), which means that $A \in \mathbb{R}_m$.

Theorem 4.13 *Let $m \in M_\tau(\mathcal{R}, E)$ and consider on X the topology $\tau_{\mathcal{R}}$. Then the map*

$$u_m : C_b(X) \to E, \quad u_m(f) = \int f\, dm = \int f\, d\overline{m}$$

is β-continuous. Also, every β-continuous linear map $u : C_b(X) \to E$ is of the form $u = u_m$ for some $m \in M_\tau(\mathcal{R}, E)$.

Proof : Let $p \in cs(E)$ and $G \in \Omega$. We need to show that the set

$$V = \{f \in C_b(X) : p(u_m(f)) \leq 1\}$$

is a β_G-neighborhood of zero. Indeed, let $r > 0$. There exists a decreasing net (V_δ) of $\tau_{\mathcal{R}}$-clopen sets with $\bigcap_\delta \overline{V_\delta}^{\beta_o X} = G$. Since $V_\delta \in \mathbb{R}_m$ and $\overline{m}$ is τ-additive, there exists a δ such that $\overline{m}_p(V_\delta) < 1/r$. Now

$$V_1 = \{f \in C_b(X) : \|f\| \leq r, \quad \|f\|_{V_\delta^c} \leq 1/\|m\|_p\} \subset V.$$

In fact, let $f \in V_1$ and set $h = f\chi_{V_\delta}$, $g = f\chi_{V_\delta^c}$. Then

$$p\left(\int h\, dm\right) = p\left(\int h\, d\overline{m}\right) \leq \|h\| \cdot \overline{m}_p(V_\delta) \leq 1$$

and

$$p\left(\int g\, dm\right) = p\left(\int g\, d\overline{m}\right) \leq \|f\|_{V_\delta^c} \cdot \overline{m}_p(X) \leq 1.$$

Thus $p(\int f\, dm) \leq 1$, which shows that $V_1 \subset V$. Since the closure of V_δ^c in $\beta_o X$ is disjoint from G, this proves that V is a β_G-neighborhood of zero. This, being true for every $G \in \Omega$, implies that V is a β-neighborhood of zero and so u_m is β-continuous. Conversely let $u : (C_b(X), \beta) \to E$ be linear and continuous. Since β is coarser than the topology of uniform convergence, it follows that, for each $p \in cs(E)$, there exists a non-zero $\lambda \in \mathbb{K}$ such that

$$\{f \in C_b(X) : \|f\| \leq |\lambda|\} \subset \{f : p(u(f)) \leq 1\}.$$

Let $K(X)$ be the algebra of al $\tau_{\mathcal{R}}$-clopen subsets of X. Define

$$\mu : K(X) \to E, \quad \mu(A) = u(\chi_A).$$

Clearly μ is finitely-additive. Also, since $|\lambda\chi_A| \leq |\lambda|$, it follows that $p(\mu(A)) \leq |\lambda|^{-1}$, and so μ is bounded. If (V_δ) is a net of clopen sets which decreases to the empty set, then $\chi_{V_\delta} \to 0$ with respect to the topology β and so $\mu(V_\delta) \to 0$. Thus $\mu \in M_\tau(K(X), E)$. The restriction $m = \mu|_{\mathcal{R}}$ is in $M_\tau(\mathcal{R}, E)$. The subspace F of $C_b(X)$ spanned by the functions χ_A, $A \in K(X)$, is β-dense in $C_b(X)$. Since u and u_m are both β-continuous and they coincide in F, it follows that $u = u_m$ on $C_b(X)$. This completes the proof.

Theorem 4.14 *Let X be a zero-dimensional Hausdorff topological space and E a Hausdorff locally convex space. Then a linear map $u : C_b(X) \to E$ is β-continuous iff it is β_o-continuous.*

Proof : Let $\hat{E}$ be the completion of E and let $K(X)$ be the algebra of all clopen subsets of X. Suppose that u is β-continuous. Then $u : (C_b(X), \beta) \to \hat{E}$ is continuous. In view of the preceding Theorem, there exists an $m \in M_\tau(K(X), \hat{E})$ such that $u(f) = \int f\, dm$ for all $f \in C_b(X)$. Let $p \in cs(E)$ and

$$V = \{f : p(u(f)) \leq 1\}.$$

We need to show that V is a β_o-neighborhood of zero. By [4], Theorem 2.8, it suffices to show that, for each $r > 0$, there exists a compact subset Y of X and $\epsilon > 0$ such that

$$V_1 = \{f \in C_b(X) : \|f\| \leq r,\ \|f\|_Y \leq \epsilon\} \subset V.$$

Choose $\epsilon > 0$ such that $\epsilon \cdot m_p(X) \leq 1$ and $r \cdot \epsilon \leq 1$. The set $X_{p,\epsilon} = \{x : N_{m,p}(x) \geq \epsilon\}$ is compact. In the definition of V_1 take as Y the set $X_{p,\epsilon}$. Let $f \in V_1$ and $A = \{x : |f(x)| \leq \epsilon\}$. Then $m_p(A^c) = \sup_{x \in A^c} N_{m,p}(x) \leq \epsilon$. Now

$$p\left(\int_A f\, dm\right) \leq \epsilon \cdot m_p(X) \leq 1, \quad \textit{and} \quad p\left(\int_{A^c} f\, dm\right) \leq r \cdot m_p(A^c) \leq 1.$$

Thus $V_1 \subset V$ and the result follows.

Theorem 4.15 *Let $\mathcal{R}$ be a separating algebra of subsets of a set X and consider on X the topology $\tau_{\mathcal{R}}$. Then ϕ_τ coincides with the topology induced on $S(\mathcal{R})$ by β_o and by the topology induced by β.*

Proof : If (V_δ) is a net of measurable subsets of X which decreases to the empty set, then $\chi_{V_\delta} \downarrow 0$ and so $\chi_{V_\delta} \xrightarrow{\beta} 0$. Thus

$$\chi : \mathcal{R} \to (S(\mathcal{R}), \beta)$$

is a τ-additive measure. In view of Theorem 3.6 , it follows that ϕ_τ is finer than the topology induced on $S(\mathcal{R})$ by β. On the other hand, let E be a Hausdorff locally convex space and let $\hat{E}$ be its completion. If $m \in M_\tau(\mathcal{R}, E)$, then $m \in M_\tau(\mathcal{R}, \hat{E})$. The map

$$u : C_b(X) \to \hat{E}, \quad u(f) = \int f\, dm,$$

is β_o-continuous. Since $\hat{m} = u|_{S(\mathcal{R})}$, it follows that $\hat{m} : (S(\mathcal{R}), \beta_o) \to \hat{E}$ is continuous and hence $\hat{m} : (S(\mathcal{R}), \beta_o) \to E$ is continuous. This implies that ϕ_τ is coarser than the topology induced on $S(\mathcal{R})$ by β_o and the result follows.

Corollary 4.16 *The topology ϕ_τ is polar and locally solid.*

Lemma 4.17 *Let Z be a vector space over $\mathbb{K}$, D a subspace of Z and τ_1, τ_2 Hausdorff locally convex topologies on Z which induce the same topology on D and for both of which D is dense in Z. If τ_2 is finer than τ_1, then τ_1 and τ_2 coincide on Z.*

Proof: Let $G = (Z, \tau_2)$ and let $\hat{G}$ be its completion. The identity map $T : (Z, \tau_2) \to \hat{G}$ is clearly continuous. Let $S = T|_D$. Since τ_1 and τ_2 induce the same topology on D, it follows that $S : (D, \tau_1) \to \hat{G}$ is continuous. As D is τ_1-dense in Z, there exists a unique continuous extension $\hat{S} : (Z, \tau_1) \to \hat{G}$. Now $\hat{S} : (Z, \tau_2) \to \hat{G}$ is continuous. Since $\hat{S} = T$ on D and D is τ_2-dense in Z, it follows that $\hat{S} = T$ on Z. Thus

$$T = \hat{S} : (Z, \tau_1) \to \hat{G}$$

is continuous, which clearly implies that τ_1 is finer than τ_2 and the Lemma follows.

Theorem 4.18 *For any zero-dimensional Hausdorff topological space X, the topologies β and β_o coincide on $C_b(X)$.*

Proof : Let $K(X)$ be the algebra of all clopen subsets of X. Since $S(K(X))$ is β-dense in $C_b(X)$, the result follows from Theorem 4.15 and the preceding Lemma.

Theorem 4.19 *Let Δ be the family of all pairs (m, p) for which there exists a Hausdorff locally convex space E such that $p \in cs(E)$ and $m \in M_\tau(\mathcal{R}, E)$. To each $\delta = (m, p) \in \Delta$ corresponds the non-Archimedean seminorm $\|\cdot\|_{N_{m,p}}$ on $S(\mathcal{R})$. Then ϕ_τ coincides with the locally convex topology ρ generated by these seminorms.*

Proof : Let E be a Hausdorff locally convex space, $m \in M_\tau(\mathcal{R}, E)$ and $p \in cs(E)$. If $g = \sum_{k=1}^n \alpha_k \chi_{A_k} \in S(\mathcal{R})$, then

$$p(\hat{m}(g)) = p\left(\sum_{k=1}^n \alpha_k m(A_k)\right) \le \max_k |\alpha_k| \cdot p(m(A_k)) \le \|g\|_{N_{m,p}}.$$

Thus $\hat{m} : S(\mathcal{R}), \rho) \to E$ is continuous and so ϕ_τ is coarser than ρ. On the other hand, let $(m, p) \in \Delta$ and

$$V = \{g \in S(\mathcal{R}) : p(\hat{m}(g)) \le 1\}.$$

Since ϕ_τ is locally solid, there exists a solid ϕ_τ-neighborhood V_1 of zero contained in V. Now $V_1 \subset \{g : \|g\|_{N_{m,p}} \le 1\}$. In fact, assume that, for some $g = \sum_{k=1}^n \alpha_k \chi_{A_k} \in V_1$, we have that $\|g\|_{N_{m,p}} > 1$. There exists an x in some A_k such that

$$|g(x)| \cdot N_{m,p}(x) = |\alpha_k| \cdot N_{m,p}(x) > 1.$$

There is a measurable set A contained in A_k such that $|\alpha_k| \cdot p(m(A)) > 1$. If $h = \alpha_k \chi_A$, then $|h| \le |g|$ and so $h \in V_1$, which is a contradiction since $p(\hat{m}(h)) > 1$. This contradiction shows that

$$V_1 \subset \{g : \|g\|_{N_{m,p}} \le 1\}.$$

Thus ρ is coarser than ϕ_τ and the result follows.

5 (VR)-Integrals

Throughout this section, $\mathcal{R}$ will be a separating algebra of subsets of a set X, E a complete Hausdorff locally convex space and $m \in M_\tau(\mathcal{R}, E)$. For $p \in cs(E)$, and $f \in \mathbb{K}^X$, let

$$\|f\|_{N_{m,p}} = \sup_{x \in X} |f(x)| \cdot N_{m,p}(x).$$

Let G_m be the space of all $f \in \mathbb{K}^X$ for which $\|f\|_{N_{m,p}} < \infty$, for each $p \in cs(E)$. Each $\|.\|_{N_{m,p}}$ is a non-Archimedean seminorm on G_m. We will consider on G_m the locally convex topology generated by these seminorms.

Lemma 5.1 *If $g = \sum_{k=1}^n \alpha_k \chi_{A_k} \in S(\mathcal{R})$, then*

$$p\left(\sum_{k=1}^n \alpha_k m(A_k)\right) \le \|g\|_{N_{m,p}}.$$

Proof : We first observe that

$$\|g\|_{N_{m,p}} \le \|g\| \cdot m_p(X) < \infty.$$

If $g = \alpha \cdot \chi_A$, where $\alpha \in \mathbb{K}$ and $A \in \mathcal{R}$, then

$$\begin{aligned} p(\alpha \cdot m(A)) \le |\alpha| \cdot m_p(A) &= |\alpha| \cdot \sup_{x \in A} N_{m,p}(x) \\ &= \sup_{x \in X} |g(x)| \cdot N_{m,p}(x) = \|g\|_{N_{m,p}}. \end{aligned}$$

In the general case, we may assume that the sets A_k, $k = 1, \ldots, n$, are pairwise disjoint. Then

$$p\left(\sum_{k=1}^n \alpha_k \cdot m(A_k)\right) \le \max_k |\alpha_k| \cdot m_p(A_k) = \max_k \sup_{x \in A_k} |g(x)| \cdot N_{m,p}(x) = \|g\|_{N_{m,p}}.$$

Lemma 5.2 *If we consider on $S(\mathcal{R})$ the topology induced by the topology of G_m, then*

$$\omega : S(\mathcal{R}) \to E, \quad \omega(g) = \int g\, dm$$

is a continuous linear map.

Proof : It follows from the preceding Lemma.

Let now $\overline{S(\mathcal{R})}$ be the closure of $S(\mathcal{R})$ in G_m and let

$$\overline{\omega} : \overline{S(\mathcal{R})} \to E$$

be the unique continuous extension of ω.

Definition 5.3 *A function $f \in \mathbb{K}^X$ is said to be (VR)-integrable with respect to m if it belongs to $\overline{S(\mathcal{R})}$. In this case, $\overline{\omega}(f)$ is called the (VR)-integral of f, with respect to m, and will be denoted by $(VR) \int f\, dm$. We will denote by $L(m)$ the space $\overline{S(\mathcal{R})}$.*

Theorem 5.4 *If f is (VR)-integrable, then, for each $p \in cs(E)$, we have*

$$p\left((VR)\int f\,dm\right) \leq \|f\|_{N_{m,p}}.$$

Proof : There exists a ne (g_δ) in $S(\mathcal{R})$ such that $g_\delta \to f$ in $\overline{S(\mathcal{R})}$. Then

$$(VR)\int f\,dm = \lim_\delta \int g_\delta\,dm, \quad \textit{and} \quad \|g_\delta\|_{N_{m,p}} \to \|f\|_{N_{m,p}}.$$

Since

$$p\left(\int g_\delta\,dm\right) \leq \|g_\delta\|_{N_{m,p}},$$

the result follows.

Theorem 5.5 *The space G_m is complete and hence $L(m)$ is also complete.*

Proof : Let (f_δ) be a Cauchy net in G_m and let

$$A = \bigcup_{p \in cs(E)} \{x : N_{m,p}(x) > 0\}.$$

Let $x \in A$ and choose $p \in cs(E)$ such that $N_{m,p}(x) = d > 0$. Given $\epsilon > 0$, there exists a δ_o such that $\|f_\delta - f_{\delta'}\|_{N_{m,p}} < d\epsilon$ if $\delta, \delta' \geq \delta_o$. Now, for $\delta, \delta' \geq \delta_o$, we have $|f_\delta(x) - f_{\delta'}(x)| < \epsilon$. This proves that the net $(f_\delta(x))$ is Cauchy in $\mathbb{K}$. Define

$$f(x) = \lim_\delta f_\delta(x), \quad if \quad x \in A$$

and $f(x)$ arbitrarily if $x \notin A$. We will show that $f \in G_m$ and that $f_\delta \to f$. Indeed, given $p \in cs(E)$ and $\epsilon > 0$, there exists δ_o such that

$$|f_\delta(x) - f_{\delta'}(x)| \cdot N_{m,p}(x) < \epsilon$$

for all x and all $\delta, \delta' \geq \delta_o$. Let now $\delta \geq \delta_o$ be fixed. If $x \in A$, then taking the limits on δ', we get that $|f_\delta(x) - f(x)| \cdot N_{m,p}(x) \leq \epsilon$. The same inequality also holds when $x \notin A$. Thus, for all $\delta \geq \delta_o$, we have

$$\sup_{x \in X} |f_\delta(x) - f(x)| \cdot N_{m,p}(x) \leq \epsilon.$$

It follows from this that , for all $x \in X$, we have

$$|f(x)| \cdot N_{m,p}(x) \leq \max\{\epsilon, \|f_{\delta_o}\|_{N_{m,p}}\}$$

which proves that $f \in G_m$. Also, $\|f - f_\delta\|_{N_{m,p}} \leq \epsilon$ for $\delta \geq \delta_o$. Hence $f_\delta \to f$ and the proof is complete.

Theorem 5.6 *For a subset A of X, the following are equivalent:*

1. *χ_A is (VR)-integrable.*

2. *For each $p \in cs(E)$ and each $\epsilon > 0$, there exists $V \in \mathcal{R}$ such that $N_{m,p} < \epsilon$ on $A \triangle V$.*

3. *For each $p \in cs(E)$ and each $\epsilon > 0$, there exists $V \in \mathcal{R}$ such that*

$$V \cap X_{p,\epsilon} = A \cap X_{p,\epsilon}.$$

Proof : (1) $\Leftrightarrow$ (2). The proof is analogous to the one given in [13], Lemma 7.3 for scalar valued measures.

(2) $\Leftrightarrow$ (3). It follows from the fact that, for $V \in \mathcal{R}$, $V \cap X_{p,\epsilon} = A \cap X_{p,\epsilon}$ iff $N_{m,p} < \epsilon$ on $A \triangle V$.

Let now $\tilde{R}_m$ be the family of all subsets A of X for which χ_A is (VR)-integrable with respect to m. It is easy to see that $\tilde{R}_m$ is a separating algebra of subsets of X which contains $\mathcal{R}$. Let $\tau_{\tilde{R}_m}$ be the zero dimensional topology having $\tilde{R}_m$ as a basis. In view of Theorem 2.5, for all $p \in cs(E)$ and all $\epsilon > 0$, the set $X_{p,\epsilon} = \{x : N_{m,p}(x) \geq \epsilon\}$ is $\tau_{\mathcal{R}}$-compact. Since $A \in \tilde{R}_m$ iff , for all $p \in cs(E)$ and all $\epsilon > 0$, there exists $V \in \mathcal{R}$ such that $V \cap X_{p,\epsilon} = A \cap X_{p,\epsilon}$, it follows that $X_{p,\epsilon}$ is $\tau_{\tilde{R}_m}$-compact. Also, since $\tau_{\mathcal{R}}$ is Hausdorff, $\tau_{\mathcal{R}}$ and $\tau_{\tilde{R}_m}$ induce the same topology on $X_{p,\epsilon}$. Now we define

$$\tilde{m} : \tilde{R}_m \to E, \quad \tilde{m}(A) = (VR) \int_A \chi_A \, dm.$$

Clearly $\tilde{m}$ is finitely-additive. Also $\tilde{m}$ is bounded since, for each $p \in cs(E)$, we have

$$p(\tilde{m}(A)) \leq \sup_{x \in A} N_{m,p}(x) \leq m_p(X).$$

Thus $\tilde{m} \in M(\tilde{R}_m, E)$.

Lemma 5.7 *If $V \in \mathcal{R}$, then $m_p(V) = \tilde{m}_p(V)$.*

Proof : It is clear that $m_p(V) \leq \tilde{m}_p(V)$. Suppose that $\tilde{m}_p(V) > \theta > 0$. There exists $A \in \tilde{R}_m$, $A \subset V$, $p(\tilde{m}(A)) > \theta$. Since $p(\tilde{m}(A)) \leq \sup_{x \in A} N_{m,p}(x)$, there exists $x \in A$ such that $N_{m,p}(x) > \theta$ and so $m_p(V) \geq N_{m,p}(x) > \theta$. This proves that $m_p(V) \geq \tilde{m}_p(V)$ and the Lemma follows.

Lemma 5.8 $N_{m,p} = N_{\tilde{m},p}$.

Proof : Since $m_p(V) = \tilde{m}_p(V)$ for $V \in \mathcal{R}$, it follows that $N_{m,p} \geq N_{\tilde{m},p}$. Assume that there exists an x such that $N_{m,p}(x) > \theta > N_{\tilde{m},p}(x)$. Let $x \in A \in \tilde{R}_m$ be such that $\tilde{m}_p(A) < \theta$. Let $Y = X_{p,\theta}$ and let $V \in \mathcal{R}$ be such that $V \cap Y = A \cap Y$. Since $x \in A \cap Y$, we have that $x \in V$ and so $m_p(V) \geq N_{m,p}(x) > \theta$. Let $D \in \mathcal{R}$, $D \subset V$ be such that $p(m(D)) > \theta$. Now $p(\tilde{m}(D \cap A)) \leq \tilde{m}_p(A) < \theta$ and hence

$$p(m(D)) = p(\tilde{m}(D \cap A^c)) \leq \sup_{y \in D \setminus A} N_{m,p}(y).$$

But, for $y \in D \setminus A$, we have that $N_{m,p}(y) < \theta$ since $D \subset V$ and $A \cap Y = V \cap Y$. Thus $\theta < p(m(D)) \leq \theta$, a contradiction. This completes the proof.

Lemma 5.9 *For $A \subset X$, we have $A \in \tilde{R}_m$ iff A is $\tau_{\tilde{R}_m}$-clopen.*

Proof : Clearly every $A \in \tilde{R}_m$ is $\tau_{\tilde{R}_m}$-clopen. On the other hand let A be $\tau_{\tilde{R}_m}$-clopen and let $p \in cs(E)$, $\epsilon > 0$. Since $\tau_{\mathcal{R}}$ and $\tau_{\tilde{R}_m}$ induce the same topology on $X_{p,\epsilon}$, the set $G = A \bigcap X_{p,\epsilon}$ is clopen in $X_{p,\epsilon}$ for the topology induced by $\tau_{\mathcal{R}}$. For each $x \in G$, there exists an $A_x \in \mathcal{R}$ such that $x \in A_x \bigcap X_{p,\epsilon} \subset G$. As G is $\tau_{\mathcal{R}}$-compact, there are $x_1, x_2, \ldots, x_n \in G$ such that

$$G = \bigcup_{k=1}^{n} A_{x_k} \cap X_{p,\epsilon} = V \cap X_{p,\epsilon},$$

where $V = \bigcup_{k=1}^{n} A_{x_k} \in \mathcal{R}$. In view of Theorem 5.6, A is in $\tilde{R}_m$ and the result follows.

Theorem 5.10 $\tilde{m} \in M_\tau(\tilde{R}_m, E)$.

Proof : Let $\mathcal{A}$ be a family in $\tilde{R}_m$ which decreases to the empty set and let $p \in cs(E)$, $\epsilon > 0$, $Y = X_{p,\epsilon}$. For each A in $\mathcal{A}$, there exists $B \in \mathcal{R}$ such that $B \cap Y = A \cap Y$. Let

$$\mathcal{B} = \{B \in \mathcal{R} : \exists A \in \mathcal{A}, A \cap Y = B \cap Y\}.$$

It is easy to see that $\mathcal{B} \downarrow \emptyset$. Since $m \in M_\tau(\mathcal{R}, E)$, there exists $B \in \mathcal{B}$ such that $m_p(B) < \epsilon$. Let $A \in \mathcal{A}$ be such that $A \cap Y = B \cap Y$. If $x \in A$, then $x \notin Y$ and so $N_{m,p})x) < \epsilon$. If $G \in \tilde{R}_m$ is contained in A, then

$$p(\tilde{m}(G)) \leq \sup_{x \in G} N_{m,p}(x) \leq \epsilon$$

and so $\tilde{m}_p(A) \leq \epsilon$. This proves that

$$\lim_{A \in \mathcal{A}} \tilde{m}_p(A) = 0$$

and so $\tilde{m} \in M_\tau(\tilde{R}_m, E)$.

Lemma 5.11 *If $g \in S(\tilde{R}_m)$, then for each $p \in cs(E)$ and each $\epsilon > 0$, there exists an $h \in S(\mathcal{R})$ such that $\|h - g\|_{N_{m,p}} \leq \epsilon$.*

Proof : Assume that $g \neq 0$ and let $A_1, A_2, \ldots, A_n$ be pairwise disjoint members of $\tilde{R}_m$ and non-zero scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that $g = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$. Let $r = \max_k |\alpha_k|$. For each k, there exists a $B_k \in \mathcal{R}$ such that $N_{m,p} < \epsilon/r$ on $A_k \triangle B_k$. Since

$$\|\alpha_k \chi_{A_k} - \alpha_k \chi_{B_k}\|_{N_{m,p}} \leq |\alpha_k| \cdot \sup_{x \in A_k \triangle B_k} N_{m,p}(x) \leq \epsilon,$$

it follows that $\|h - g\|_{N_{m,p}} \leq \epsilon$.

Using Lemmas 5.7 and 5.11, we get the following

Theorem 5.12 *A function $f \in \mathbb{K}^X$ is (VR)-integrable with respect to m iff it is (VR)-integrable with respect to $\tilde{m}$. Moreover*

$$(VR)\int f\,dm = (VR)\int f\,d\tilde{m}.$$

Theorem 5.13 *If $f \in \mathbb{K}^X$ is m-integrable with respect to m, then it is also (VR)-integrable and*

$$\int f\,dm = (VR)\int f\,dm.$$

Proof : Let $p \in cs(E)$ and $\epsilon > 0$. There exists a $\mathcal{R}$-partition $\{V_1, V_2, \ldots, V_n\}$ of X such that $|f(x) - f(y)| \cdot m_p(V_i) < \epsilon$ if $x, y \in V_i$. Let $x_k \in V_k$ and $g = \sum_{k=1}^{n} f(x_k)\chi_{V_k}$. For $x \in V_k$, we have

$$|f(x) - g(x)| \cdot N_{m,p}(x) = |f(x) - f(x_k)| \cdot N_{m,p}(x) \leq |f(x) - f(x_k)| \cdot m_p(V_k) < \epsilon.$$

This proves that f is (VR)-integrable. Also $p\left(\int f\,dm - \int g\,dm\right) \leq \epsilon$ and

$$p\left((VR)\int f\,dm - \int g\,dm\right) = p\left((VR)\int (f-g)\,dm\right) \leq \|f - g\|_{N_{m,p}} \leq \epsilon.$$

Thus

$$p\left(\int f\,dm - (VR)\int f\,dm\right) \leq \epsilon.$$

Since E is Hausdorff, it follows that

$$\int f\,dm = (VR)\int f\,dm.$$

and the proof is complete.

Theorem 5.14 *Let Y be a zero-dimensional topological space and $f : X \to Y$. Then f is $\tau_{\tilde{R}_m}$-continuous iff, for each $p \in cs(E)$ and each $\epsilon > 0$, the restriction of f to $X_{p,\epsilon}$ is $\tau_{\mathcal{R}}$- continuous.*

Proof : Since $\tau_{\mathcal{R}}$ and $\tau_{\tilde{R}_m}$ induce the same topology on $X_{p,\epsilon}$, the necessity of the condition is clear. On the other hand, assume that the condition is satisfied and let Z be a clopen subset of Y. We need to show that $f^{-1}(A)$ is $\tau_{\tilde{R}_m}$-clopen, or equivalently that $f^{-1}(A) \in \tilde{R}_m$. Let $p \in cs(E)$ and $\epsilon > 0$. The restriction h of f to $X_{p,\epsilon}$ is $\tau_{\mathcal{R}}$-continuous. Thus

$$G = f^{-1}(A) \cap X_{p,\epsilon} = h^{-1}(A)$$

is clopen in $X_{p,\epsilon}$ for the topology induced by $\tau_{\mathcal{R}}$. For each $x \in G$, there exists $V_x \in \mathcal{R}$ such that $x \in V_x \bigcap X_{p,\epsilon} \subset G$. Since G is $\tau_{\mathcal{R}}$-compact, there are $x_1, x_2, \ldots, x_n$ in G such that

$$G = \bigcup_{k=1}^{n} V_{x_k} \cap X_{p,\epsilon}.$$

If $V = \bigcup_{k=1}^{n} V_{x_k} \in \mathcal{R}$, then

$$V \cap X_{p,\epsilon} = f^{-1}(A) \cap X_{p,\epsilon}.$$

In view of Lemma 5.9, we get that $f^{-1}(A) \in \tilde{R}_m$ and we are done.

Theorem 5.15 *Let $m \in M_\tau(\mathcal{R}, E)$ and $f \in \mathbb{K}^X$. Then, f is (VR)-integrable iff :*
a) f is $\tau_{\tilde{R}_m}$-continuous.
b) For each $p \in cs(E)$ and each $\epsilon > 0$, the set

$$D = \{x : |f(x)| \cdot N_{m,p}(x) \geq \epsilon\}$$

is $\tau_{\tilde{R}_m}$-compact.

Proof : Assume that f is (VR)-integrable and let $p \in cs(E)$ and $\epsilon > 0$. There exists a sequence (g_n) in $S(\mathcal{R})$ such that $\|f - g_n\|_{N_{m,p}} \to 0$. For $x \in X_{p,\epsilon}$, we have that

$$|f(x) - g_n(x)| \leq 1/\epsilon \cdot \|f - g_n\|_{N_{m,p}} \to 0$$

uniformly. Since each g_n is $\tau_{\mathcal{R}}$-continuous , it follows that $f|_{X_{p,\epsilon}}$ is $\tau_{\mathcal{R}}$-continuous and so f is $\tau_{\tilde{R}_m}$-continuous. Also, given $\epsilon > 0$, there exists a $g \in S(\mathcal{R})$ such that $\|f - g\|_{N_{m,p}} \leq \epsilon$. Let $\{V_1, V_2, \ldots, V_n\}$ be pairwise disjoint members of $\mathcal{R}$ and $\alpha_1, \alpha_2, \ldots, \alpha_n$ non-zero scalars such that $g = \sum_{k=1}^{n} \alpha_k \chi_{V_k}$. Now

$$D = \{x : |g(x)| \cdot N_{m,p}(x) \geq \epsilon\} = \bigcup_{k=1}^{n} [V_k \cap \{x : N_{m,p}(x) \geq \epsilon/|\alpha_k|\}],$$

and so D is $\tau_{\tilde{R}_m}$-compact. Moreover

$$D = \{x : |f(x)| \cdot N_{m,p}(x) \geq \epsilon\}.$$

Conversely, assume that the conditions (a), (b) are satisfied. Let $p \in cs(E)$, $\epsilon > 0$ and

$$D = \{x : |f(x)| \cdot N_{m,p}(x) \geq \epsilon\}.$$

For each $x \in D$, there exists an $A_x \in \tilde{R}_m$ such that

$$x \in A_x \subset \{y : |f(y) - f(x)| < \epsilon/m_p(X)\}.$$

By the $\tilde{R}_m$-compactness of D, there are $y_1, y_2, \ldots, y_n \in Y$ such that $D \subset \bigcup_{k=1}^{n} A_{y_k}$. Now, there are pairwise disjoint sets $V_1, V_2, \ldots, V_N$ in $\tilde{R}_m$ such that $D \subset \bigcup_{j=1}^{N} V_j$ and each V_j is contained in some A_{y_k}. Let

$$x_j \in V_j, \quad g = \sum_{j=1}^{N} f(x_j) \chi_{V_j}.$$

If $x \in V_j$, then

$$|f(x) - g(x)| \cdot N_{m,p}(x) = |f(x) - f(x_j)| \cdot N_{m,p}(x) \leq \|m\|_p \cdot \epsilon/\|m\|_p = \epsilon,$$

while, for $x \notin \bigcup_{j=1}^{N} V_j$ we have $g(x) = 0$ and $x \notin D$, which implies that

$$|f(x) - g(x)| \cdot N_{m,p}(x) = |f(x)| \cdot N_{m,p}(x) \leq \epsilon.$$

This proves that f is (VR)-integrable with respect to $\tilde{m}$ and hence it is (VR)-integrable with respect to m. This completes the proof.

6 The Measure $\tilde{m}_f$

In this section we will assume that E is a complete Hausdorff locally convex space, $\mathcal{R}$ a separating algebra of subsets of a set X and $m \in M_\tau(\mathcal{R}, E)$. Let $f \in \mathbb{K}^X$ be (VR)-integrable with respect to m and define

$$\tilde{m}_f : \tilde{R}_m \to E, \quad \tilde{m}_f(A) = (VR)\int_A f\, dm = (VR)\int \chi_A f\, dm.$$

Then, for each $p \in cs(E)$, we have

$$p(\tilde{m}_f(A)) \le \sup_{x\in A} |f(x)| \cdot N_{m,p}(x) \le \|f\|_{N_{m,p}},$$

and so $\tilde{m}_f$ is bounded and clearly finitely-additive. Also $\tilde{m}_f$ is τ-additive. Indeed, let (A_δ) be a net in $\tilde{R}_m$ which decreases to the empty set and let $p \in cs(E)$, $\epsilon > 0$. There exists a $g \in S(\mathcal{R})$ such that $\|f - g\|_{N_{m,p}} < \epsilon$. If $g = \sum_{k=1}^n \alpha_k \chi_{V_k}$, where $V_1, V_2, \ldots, V_n$ are pairwise disjoint members of $\mathcal{R}$, then

$$\tilde{m}_g(A_\delta) = \sum_{k=1}^n \alpha_k \tilde{m}(V_k \cap A_\delta).$$

Since $A_\delta \cap V_k \downarrow \emptyset$ and $\tilde{m}$ is τ-additive, there exists δ_o such that $p(\tilde{m}_g(A_\delta)) < \epsilon$ if $\delta \ge \delta_o$. Also

$$p(\tilde{m}_{f-g}(A_\delta)) \le \|f-g\|_{N_{m,p}} < \epsilon.$$

Thus, for $\delta \ge \delta_o$, we have that $p(\tilde{m}_f(A_\delta)) < \epsilon$, which proves that $\tilde{m}_f \in M_\tau(\tilde{R}_m, E)$.

Lemma 6.1 *If $g \in S(\mathcal{R})$, then $N_{\tilde{m}_g,p}(x) = |g(x)| \cdot N_{m,p}(x)$.*

Proof : Let $g = \sum_{k=1}^n \alpha_k \chi_{V_k}$, where $\{V_1, V_2, \ldots, V_n\}$ is an $\mathcal{R}$-partition of X. Let $x \in V_k$ and $h = \alpha_k \chi_{V_k}$. If $A \in \tilde{R}_m$ is contained in V_k, then

$$\tilde{m}_g(A) = \tilde{m}_h(A) = \alpha_k \cdot (VR)\int \chi_A\, dm = g(x)\tilde{m}(A).$$

Thus

$$N_{\tilde{m}_g,p}(x) = |g(x)| \cdot N_{\tilde{m},p}(x) = |g(x)| \cdot N_{m,p}(x).$$

Lemma 6.2 *Let $f, g \in \mathbb{K}^X$ be (VR)-integrable with respect to m. Then for each $V \in \tilde{R}_m$, we have*

$$|(\tilde{m}_f)_p(V) - (\tilde{m}_g)_p(V)| \le \|f - g\|_{N_{m,p}}.$$

Proof : Assume (say) that $(\tilde{m}_f)_p(V) - (\tilde{m}_g)_p(V)| \ge 0$. Given $\epsilon > 0$, there exists $A \in \tilde{R}_m$ contained in V such that $(\tilde{m}_f)_p(V) < p(\tilde{m}_f(A)) + \epsilon$. Now

$$\begin{aligned} 0 \le (\tilde{m}_f)_p(V) - (\tilde{m}_g)_p(V) &< \epsilon + p(\tilde{m}_f(A)) - p(\tilde{m}_g(A)) \\ &\le \epsilon + p(\tilde{m}_f(A) - \tilde{m}_g(A)) \\ &= \epsilon + p(\tilde{m}_{f-g}(A)) \le \epsilon + \|f-g\|_{N_{m,p}} \end{aligned}$$

and the Lemma follows taking $\epsilon \to 0$.

Lemma 6.3 *Let $f, g \in \mathbb{K}^X$ be (VR)-integrable with respect to m. Then*

$$|N_{\tilde{m}_f,p}(x) - N_{\tilde{m}_g,p}(x)| \leq \|f - g\|_{N_{m,p}}.$$

Proof : Suppose (say) that $0 \leq N_{\tilde{m}_f,p}(x) - N_{\tilde{m}_g,p}(x)$ and choose a $V \in \tilde{R}_m$ containing x such that $(\tilde{m}_g(V) < N_{\tilde{m}_g,p}(x) + \epsilon$. Now

$$0 \leq N_{\tilde{m}_f,p}(x) - N_{\tilde{m}_g,p}(x) \leq (\tilde{m}_f)_p(V) - [(\tilde{m}_g)_p(V) - \epsilon] \leq \epsilon + \|f - g\|_{N_{m,p}}.$$

Taking $\epsilon \to 0$, the Lemma follows.

Theorem 6.4 *If $f \in \mathbb{K}^X$ is (VR)-integrable with respect to m, then*

$$N_{\tilde{m}_f,p}(x) = |f(x)| \cdot N_{m,p}(x).$$

Proof : Given $\epsilon > 0$, there exists a $g \in S(\mathcal{R})$ such that $\|f - g\|_{N_{m,p}} < \epsilon$. By Lemma 6.1, we have $N_{\tilde{m}_g,p}(x) = |g(x)| \cdot N_{m,p}(x)$. Also

$$|\,|g(x)| \cdot N_{m,p}(x) - |f(x)| \cdot N_{m,p}(x)\,| \leq |g(x) - f(x)| \cdot N_{m,p}(x) < \epsilon.$$

Thus

$$|N_{\tilde{m}_f,p} - |f(x)| \cdot N_{m,p}(x)| \leq |N_{\tilde{m}_f,p}(x) - N_{\tilde{m}_g,p}(x)| + |\,|g(x)| \cdot N_{m,p}(x) - |f(x)| \cdot N_{m,p}(x)\,| \leq 2\epsilon.$$

As $\epsilon > 0$ was arbitrary, the Theorem follows.

Lemma 6.5 *If $f \in \mathbb{K}^X$ is (VR)-integrable with respect to m and $h \in S(\mathcal{R})$, then hf is (VR)-integrable.*

Proof : Let $\epsilon > 0$, $p \in cs(E)$, $d > \|h\|$. Choose $g \in S(\mathcal{R})$ such that $\|g - f\|_{N_{m,p}} < \epsilon/d$. Now $gh \in S(\mathcal{R})$ and $\|hf - gh\|_{N_{m,p}} < \epsilon$, which proves the Lemma.

Theorem 6.6 *Let $f \in \mathbb{K}^X$ be (VR)-integrable with respect to m. If $g \in \mathbb{K}^X$ is (VR)-integrable with respect to $\tilde{m}_f$, then gf is (VR)-integrable with respect to m and*

$$(VR)\int gf\,dm = (VR)\int g\,d\tilde{m}_f.$$

Proof : Given $p \in cs(E)$ and $\epsilon > 0$,let $h \in S(\tilde{R}_m)$ be such that $\|g - h\|_{N_{\tilde{m}_f,p}} < \epsilon$. Let $d > \|h\|$ and choose $f_1 \in S(\mathcal{R})$ such that $\|f - f_1\|_{N_{m,p}} < \epsilon/d$. If $V \in \tilde{R}_m$, then

$$\int \chi_V\,d\tilde{m}_f = \tilde{m}_f(V) = (VR)\int \chi_V f\,dm$$

and so $\int h\,d\tilde{m}_f = (VR)\int hf\,dm$. Now

$$p\left((VR)\int g\,d\tilde{m}_f - \int h\,d\tilde{m}_f\right) \leq \|g - h\|_{N_{\tilde{m}_f,p}} < \epsilon.$$

If $f_2 = f - f_1$, then

$$\|hf_2\|_{N_{m,p}} \leq \epsilon \quad and \quad \|g-h)f\|_{N_{m,p}} = \|g-h\|_{N_{\tilde{m}_f,p}} \leq \epsilon.$$

It follows that $\|gf - hf_1\|_{N_{m,p}} \leq \epsilon$. Since hf_1 is (VR)-integrable with respect to m, we get that gf is (VR)-integrable with respect to m. Also,

$$p\left((VR)\int fg\,dm - (VR)\int hf\,dm\right) \leq \|gf - hf\|_{N_{m,p}} \leq \epsilon.$$

It follows that

$$p\left((VR)\int fg\,dm - (VR)\int g\,d\tilde{m}_f\right) \leq \epsilon,$$

which clearly completes the proof.

Theorem 6.7 *Let $f, g \in \mathbb{K}^X$ be (VR)-integrable with respect to m. If g is bounded, then :*

1. *g is (VR)-integrable with respect to $\tilde{m}_f$.*
2. *gf is (VR)-integrable with respect to m.*
3. $\quad (VR)\int gf\,dm = (VR)\int g\,d\tilde{m}_f.$

The same result holds if we assume that f is bounded.

Proof : Assume that g is bounded. In view of the preceding Theorem, we only need to prove (1). By Theorem 5.15, g is $\tau_{\tilde{R}_m}$-continuous. As g was assumed to be bounded, we get that g is integrable with respet to $\tilde{m}_f$, which implies that it is (VR)-integrable with respect to the same measure (by Theorem 5.13). Thus (1) holds. In case f is bounded, let $d > \|f\|$ and choose $h \in S(\mathcal{R})$ such that $\|g-h\|_{N_{m,p}} < \epsilon/d$. Now

$$\|g-h\|_{N_{\tilde{m}_f,p}} = \|(g-h)f\|_{N_{m,p}} < \epsilon,$$

and so the result follows.

Theorem 6.8 *Let $f \in \mathbb{K}^X$ be (VR)-integrable with respect to m and let $g \in \mathbb{K}^X$ be m-integrable. Then :*

1. *g is (VR)-integrable with respect to $\tilde{m}_f$.*
2. *gf is (VR)-integrable with respect to m.*
3. $\quad (VR)\int gf\,dm = (VR)\int g\,d\tilde{m}_f.$

Proof : Let $p \in cs(E)$ and $\epsilon > 0$. Since g is m-integrable, there exists a $V \in S(\mathcal{R})$, with $m_p(V^c) = 0$, such that $\|g\|_V = d < \infty$. Let $g_1 = g\chi_V$. By the preceding Theorem, there exists an $h \in S(\tilde{R}_m)$ such that $\|g_1 - h\|_{N_{\tilde{m}_f,p}} < \epsilon$. For $x \in V^c$, we have

$$|g(x) - h(x)| \cdot N_{\tilde{m}_f,p}(x) = |f(x)(g(x)-h(x))| \cdot N_{m,p}(x) = 0.$$

Thus $\|g-h\|_{N_{\tilde{m}_f,p}} \leq \epsilon$. This proves (1) and the result follows.

7 The Completion of $(S(\mathcal{R}), \phi_\tau)$

In this section, $\mathcal{R}$ will be a separating algebra of subsets of a non-empty set X. We will equip X with the topology $\tau_{\mathcal{R}}$. As in [9], we will denote by $X^{(k)}$ the set X equipped with the zero-dimensional topology which has as a base the family of all subsets A of X such that $A \cap Y$ is clopen in Y for each compact subset Y of X. We will prove that $(C_b(X^{(k)}), \beta_o)$ coincides with the completion $\hat{F}$ of $F = (S(\mathcal{R}), \phi_\tau)$. As F is a polar Hausdorff space, its completion is the space of all linear functionals on $F' = M_\tau(\mathcal{R})$ which are $\sigma(F', F)$-continuous on ϕ_τ-equicontinuous subsets of $M_\tau(\mathcal{R})$ (see [10]). The topology of $\hat{F}$ is the one of uniform convergence on the ϕ_τ-equicontinuous subsets of $M_\tau(\mathcal{R})$. Since ϕ_τ is the topology induced on $S(\mathcal{R})$ by β_o and since β_o and the topology τ_u of uniform convergence have the same bounded sets, it follows that the strong topology on F' is the tropology given by the norm $m \mapsto \|m\|$.

Theorem 7.1 *The completion $\hat{F}$ of F is an algebraic subspace of the second dual F''. The topology of $\hat{F}$ is coarser than the topology induced on $\hat{F}$ by the norm topology of F''.*

Proof : Let u be a linear functionmal on $M_\tau(\mathcal{R})$ which is $\sigma(F', F)$-continuous on ϕ_τ-equicontinuous subsets of $M_\tau(\mathcal{R})$. Then u is norm-continuous. Indeed, let (m_n) be a sequence in $M_\tau(\mathcal{R})$ with $\|m_n\| \to 0$. The set $H = \{m_n : n \in \mathbf{N}\}$ is uniformly τ-additive. In fact, let (V_δ) be a net in $\mathcal{R}$ which decreases to the empty set and let $\epsilon > 0$. Choose n_o such that $\|m_n\| < \epsilon$ if $n > n_o$. If δ_o is such that $|m_n|(V_\delta) < \epsilon$ for all $\delta \geq \delta_o$ and all $n = 1, 2, \ldots, n_o$, then $|m|(V_\delta) < \epsilon$ for all $m \in H$ and all $\delta \geq \delta_o$. In view of Theorem 3.10, H is ϕ_τ-equicontinuous. As $\int g\, dm_n \to 0$ for all $g \in S(\mathcal{R})$, it follows that $u(m_n) \to 0$ and so $u \in F''$. The last assertion is a consequence of the fact that every ϕ_τ-equicontinuous subset of $M_\tau(\mathcal{R})$ is norm bounded.

Let $K(X)$ be the algebra of all $\tau_{\mathcal{R}}$- clopen subsets of X. For $m \in M_\tau(\mathcal{R}, E)$, let

$$\tilde{m} : K(X) \to \mathbb{K}, \quad \tilde{m}(A) = \int \chi_A \, dm.$$

Then $\tilde{m} \in M_\tau(K(X))$.

Lemma 7.2 *If H is a uniformly τ-additive subset of $M_\tau(\mathcal{R})$, then the set*

$$\tilde{H} = \{\tilde{m} : m \in H\}$$

is a uniformly τ-additive subset of $M_\tau(K(X))$.

Proof : Let (V_δ) be a net in $K(X)$ which decreases to the empty set. Consider the family $\mathcal{F}$ of all $A \in \mathcal{R}$ which contain some V_δ. Let $A_1, A_2 \in \mathcal{F}$ and let δ_1, δ_2 be such that $V_{\delta_i} \subset A_i$, for $i = 1, 2$. If $\delta \geq \delta_1, \delta_2$, then $V_\delta \subset A = A_1 \cap A_2$, which proves that $\mathcal{F}$ is downwards directed. Also, $\bigcap \mathcal{F} = \emptyset$. Indeed, let $x \in X$ and choose V_δ not containig x. There exists a $B \in S(\mathcal{R}$ such that $x \in B \subset V_\delta^c$. Now $V_\delta \subset A = B^c$ and $x \notin A$, which proves that $\bigcap \mathcal{F} = \emptyset$. As H is uniformly τ-additive, there exists $A \in \mathcal{F}$ with $|m|(A) < \epsilon$ for all $m \in H$. If V_δ is contained in A, then $|\tilde{m}|(V_\delta) \leq |\tilde{m}|(A) = |m|(A) < \epsilon$, for all $m \in H$, and the Lemma follows.

Theorem 7.3 *$(C_b(X), \beta_o)$ is a topological subspace of $\hat{F}$.*

Proof : Let $f \in C_b(X)$. Without loss of generality we may assume that $\|f\| \leq 1$. For each $m \in M_\tau(\mathcal{R})$, the integral $\int f\,dm$ exists. Thus f may be considered as a linear functional on $M_\tau(\mathcal{R}) = F'$. Let H be an absolutely convex ϕ_τ-equicontinuous subset of $M_\tau(\mathcal{R})$ and let (m_δ) be a net in H which is $\sigma(F', F)$-convergent to zero. We will show that $\int f\,dm_\delta \to 0$. As H is ϕ_τ-equicontinuous, we have that $d = \sup_{m \in H} \|m\| < \infty$. By the preceding Lemma, the set $\tilde{H}$ is a norm-bounded uniformly τ-additive subset of $M_\tau(K(X))$. By [4], Theorem 3.6, given $\epsilon > 0$, there exists a compact subset Y of X such that $|m|(V) = |\tilde{m}|(V) < \epsilon$ for all $m \in H$ and all $V \in \mathcal{R}$ disjoint from Y. For each $x \in Y$, there exists an $A_x \in \mathcal{R}$ containing x and such that

$$A_x \subset \{y : |f(y) - f(x)| < \epsilon/d\}.$$

By the compactness of Y, there are $x_1, x_2, \ldots, x_n$ in Y such that $Y \subset \bigcup_{k=1}^n A_{x_k}$. Now there are pairwise disjoint sets $B_1, B_2, \ldots, B_N$ in $\mathcal{R}$ covering Y such that each B_i is contained in some A_{x_k}. Let $y_i \in B_i$ and $g = \sum_{i=1}^N f(y_i)\chi_{B_i}$. For $x \in B = \bigcup_{i=1}^N B_i$, we have that $|f(x) - g(x)| < \epsilon/d$ and $|m|(B^c) < \epsilon$ for all $m \in H$. Let δ_o be such that $|\int g\,dm_\delta| < \epsilon$ if $\delta \geq \delta_o$. Since

$$\left|\int_B (f-g)\,dm_\delta\right| \leq d \cdot \|f-g\|_B \leq \epsilon \quad \text{and} \quad \left|\int_{B^c} (f-g)\,dm_\delta\right| \leq |m|(B^c) < \epsilon,$$

it follows that $|\int f\,dm_\delta| \leq \epsilon$ for all $\delta \geq \delta_o$. This proves that $f \in \hat{F}$.
Since β_o is polar, it follows from [4], Theorem 3.6, that β_o is the topology of uniform convergence on the family of all norm-bounded uniformly τ-additive subsets of $M_\tau(K(X))$. Let Z be such a subset of $M_\tau(K(X))$ and let $H = \{m|_\mathcal{R} : m \in Z\}$. Then H is uniformly τ-additive subset of $M_\tau(\mathcal{R})$ and

$$\sup_{\mu \in H} \|\mu\| = \sup_{m \in Z} \|m\| < \infty.$$

If H^o is the polar of H in $\hat{F}$ and Z^o the polar of Z in $C_b(X)$, then $Z^o = H^o \bigcap C_b(X)$. Now the result follows from this, the preceding Lemma and Theorem 3.10.

Theorem 7.4 *The completion of the space $F = (S(\mathcal{R}), \phi_\tau)$ coincides with the space $(C_b(X^{(k)}, \beta_o)$.*

Proof : By the preceding Theorem, $(C_b(X), \beta_o)$ is a topological subspace of $\hat{F}$. Thus $\hat{F}$ coincides with the completion of $(C_b(X), \beta_o)$. Now the result follows from [8], Theorem 4.3, in view of [9], Theorem 3.14

Let now E be a complete locally convex Hausdorff space and let $m \in M_\tau(\mathcal{R}, E)$. In view of the preceding Theorem, there exists a unique β_o-continuous extension u of $\hat{m}$ to all of $C_b(X^{(k)})$. We will show that, for all $f \in C_b(X^{(k)})$ we have $u(f) = (VR)\int f\,dm$.

Theorem 7.5 *Let $m \in M_\tau(\mathcal{R}, E)$, where E is a complete Hausdorff locally convex space. If*

$$u : (C_b(X^{(k)}), \beta_o) \to E$$

is the unique continuous extension of $\hat{m}$, then $u(f) = (VR)\int f\,dm$.

Proof: Let $f \in C_b(X^{(k)})$. Without loss of generality, we may assume that $\|f\| \leq 1$. Let Γ be the set of all $\gamma = (p, Y, n)$, where $p \in cs(E)$, $n \in \mathbf{N}$ and Y a compact subset of X. We make Γ into a directed set by defining $(p_1, Y_1, n_1) \geq (p_2, Y_2, n_2)$ iff $p_1 \geq p_2$, $Y_2 \subset Y_1$ and $n_1 \geq n_2$. Let

$$B = \{g \in C_b(X^{(k)}) : \|g\| \leq 1\}.$$

On B, β_o coincides with the topology of uniform convergence on the compact subsets of $X^{(k)}$ (equivalently on compact subsets of X by [9], Corollary 3.14).
Claim: For each $\gamma = (p, Y, n)$, there exists a $g_\gamma \in S(\mathcal{R})$, $g_\gamma \in B$, such that

$$\|f - g_\gamma\|_Y \leq 1/n, \quad \|f - g_\gamma\|_{N_{m,p}} \leq 1/n.$$

Indeed, choose $\epsilon > 0$ such that $\epsilon < 1/n$ and $\epsilon \cdot \|m\|_p < 1/n$. The set

$$Z = Y \bigcup \{x : N_{m,p}(x) \geq \epsilon\}$$

is compact. For each $y \in Z$, there exists $V_y \in S(\mathcal{R})$ containing y and such that

$$V_y \cap Z \subset \{z : |f(z) - f(y)| < \epsilon\}.$$

By the compactness of Z, there are pairwise disjoint $W_1, W_2, \ldots, W_N$ in $S(\mathcal{R})$ covering Z and such that each W_i is contained in some V_y. Choose $z_k \in W_k$ and take $g_\gamma = \sum_{k=1}^N f(z_k)\chi_{W_k}$. Then $g_\gamma \in B$. If $x \in Y$, then $|f(x) - g_\gamma(x)| \leq \epsilon < 1/n$ and so $\|f - g_\gamma\|_Y \leq 1/n$. Also, if $x \in W = \bigcup_{k=1}^N W_k$, then

$$|f(x) - g_\gamma(x)| \cdot N_{m,p}(x) \leq \epsilon \cdot \|m\|_p < 1/n,$$

while for $x \notin W$ we have that $N_{m,p}(x) \leq \epsilon < 1/n$. Thus $\|f - g_\gamma\|_{N_{m,p}} \leq 1/n$, which proves our claim.
Now the net (g_γ) is in B and converges to f with respect to the topology of uniform convergence on compact subsets of X and so (g_γ) is β_o-convergent to f, which implies that $u(f) = \lim u(g_\gamma)$. On the other hand, (g_γ) is contained in G_m and converges to f in the topology of G_m Thus

$$u(f) = \lim u(g_\gamma) = \lim \int g_\gamma\,dm = (VR)\int f\,dm.$$

This completes the proof.

Theorem 7.6 *Let X be a zero-dimensional Hausdorff space and let Δ be the family of all pairs (m, p) for which there exists a Hausdorff locally convex space E such that $p \in cs(E)$ and $m \in M_\tau(K(X), E)$, where $K(X)$ is the algebra of all clopen subsets of X. Then the topologies β and β_o on $C_b(X)$ coincide with the locally convex topology ρ generated by the seminorms $\|\cdot\|_{N_{m,p}}$, $(m, p) \in \Delta$.*

Proof : As it is shown in the proof of the preceding Theorem, the space $F = S(K(X))$ is ρ-dense in $C_b(X)$. Also F is dense in $C_b(X)$ for the topologies β and β_o. In view of Theorems 4.15, 4.18 and 4.19, the topologies β_o, β and ρ coincide on F. Also, ρ is coarser than β_o. Indeed, let $(m,p) \in \Delta$ and

$$V = \{f \in C_b(X) : \|f\|_{N_{m,p}} \leq 1\}.$$

Let $r > 0$ and choose $0 < \epsilon < 1/r$ such that $\epsilon \cdot m_p(X) < 1$. The set $Y = \{x : N_{m,p}(X) \geq \epsilon\}$ is compact. Moreover

$$V_1 = \{f \in C_b(X) : \|f\| \leq r, \ \|f\|_Y \leq \epsilon\}$$

is contained in V. In fact, let $f \in V_1$. If $x \in Y$, then $|f(x)| \cdot N_{m,p}(x) \leq \epsilon \cdot m_p(X) \leq 1$, while for $x \notin Y$ we have $|f(x)| \cdot N_{m,p}(x) \leq r\epsilon \leq 1$. Thus $\|f\|_{N_{m,p}} \leq 1$, i.e. $f \in V$. This, being true for each $r > 0$, implies that V is a β_o-neighborhood of zero. Now the result follows from Lemma 4.17.

References

[1] J. Aguayo, Vector measures and integral operators, in : Ultrametric Functional Analysis, Cont. Math., vol. **384**(2005), 1-13.

[2] J. Aguayo and T. E. Gilsdorf, Non-Archimedean vector measures and integral operators, in : p-adic Functional Analysis, Lecture Notes in Pure and Appplied Mathematics, vol **222**, Marcel Dekker, New York (2001), 1-11.

[3] J. Aguayo and M Nova, Non-Archimedean integral operators on the space of continuous functions, in : Ultrametric Functional Analysis, Cont. Math., vol **319**(2002), 1-15.

[4] A. K. Katsaras, The strict topology in non-Archimedean vector-valued function spaces, Proc. Kon. Ned. Akad. Wet. A **87** **(2)** (1984), 189-201.

[5] A. K. Katsaras, Strict topologies in non-Archimedean function spaces, Intern. J. Math. and Math. Sci., **7** **(1)**, (1984), 23-33.

[6] A. K. Katsaras, Separable measures and strict topologies on spaces of non-Archimedean valued functions, in : P-adic Numbers in Number Theory, Analytic Geometry and Functional Analysis, edided by S. Caenepeel, Bull. Belgian Math., (2002), 117-139.

[7] A. K. Katsaras, Strict topologies and vector measures on non-Archimedean spaces, Cont. Math. vol. **319** (2003), 109-129.

[8] A. K. Katsaras, Non-Archimedean integration and strict topologies, Cont. Math., vol. **384** (2005), 111-144.

[9] A. K. Katsaras, p-adic measures and p-adic spaces of continuous functions (preprint).

[10] A. K. Katsaras, The non-Archimedean Grothendieck's completeness theorem, Bull. Inst. Math. Acad. Sinica **19**(1991), 351-354.

[11] A. F. Monna and T. A. Springer, Integration non-archimedienne, Indag. Math. **25, no 4**(1963), 634-653.

[12] W. H. Schikhof, Locally convex spaces over non-spherically complete fields I, II, Bull. Soc. Math. Belg., Ser. B, **38** (1986), 187-224.

[13] A. C. M. van Rooij, Non-Archimedean Functional Analysis, New York and Bassel, Marcel Dekker, 1978.

[14] A. C. M. van Rooij and W. H. Schikhof, Non-Archimedean Integration Theory, Indag. Math., **31**(1969), 190-199.

Department of Mathematics, University of Ioannina, 45110 Ioannina, Greece
e-mail : akatsar@cc.uoi.gr

SUPERLINEAR CONVERGENCE FOR PCG USING BAND PLUS ALGEBRA PRECONDITIONERS FOR TOEPLITZ SYSTEMS*

D. NOUTSOS† AND P. VASSALOS‡

Abstract. This paper is concerned with the fast and efficient solution of $n \times n$ symmetric ill conditioned Toeplitz systems $T_n(f)x = b$ where the generating function f is a priori known and in particular is real valued, nonnegative, having isolated roots of even order. The preconditioner that we propose is a product of a band Toeplitz matrix and matrices that belong to a certain trigonometric algebra. The underline idea of the proposed scheme is to embody the well known advantages that each component of the product presents, when they are used alone at the same time which are minimized their disadvantages. As a result we obtain a flexible preconditioner which can be applied to the system $T_n(f)x = b$ infusing superlinear convergence to the PCG method. The important feature of the proposed technique is that it can be extended to cover the $2D$ case, i.e. ill-conditioned band Toeplitz with Toeplitz blocks (BTTB) matrices. We perform many numerical experiments and the results fully confirm the effectiveness of the proposed strategy and the adherence to the theoretical analysis.

Key words. Toeplitz, preconditioning, trigonometric algebras, PCG

AMS subject classifications. 65F10, 65F15, 65F35

1. Introduction.

In this paper we introduce and analyze a new approach for the solution, by means of the Preconditioned Conjugate Gradient (PCG) method, of ill conditioned linear systems $Tx = b$ where $T = T_n(f)$ is a Toeplitz matrix. A matrix is called Toeplitz matrix if its (i, j) entry depends only on the difference $i - j$ of the subscripts i.e. $t_{i,j} = t_{i-j}$. The function $f(x)$ whose Fourier coefficients give the diagonals of $T_n(f)$ i.e.

$$T_{j,k} = t_{j-k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-\mathrm{i}(j-k)x} dx, \qquad 1 \le j, k < n,$$

is called the generating function of $T_n(f)$ and in the rest of the paper we will assume that it is a priori known.

Such kind of matrices arise in a wide variety of fields of pure and applied mathematics such as signal theory, image processing, probability theory, harmonic analysis, control theory etc. Therefore, a fast and effective solver is not only welcome but also necessary.

Several direct methods for solving Toeplitz systems have been proposed; the most efficient algorithms are called "superfast" and require $O(n \log^2 n)$ operations to compute the solution. The stability properties of these direct methods are discussed in [6]. The main disadvantage of these kind of methods is that in 2D they can not exploit efficiently the Block Toeplitz structure of the matrices and as a consequence they are far away from being characterized a near optimal choice as they need $O(nm^2 \log nm)$.

We focus on the case where the generating function f is real-valued continuous 2π-periodic defined in $I = [-\pi, \pi]$, where the associated Toeplitz matrix is a Hermitian matrix.

*This research was co-funded by the European Union - European Social Fund (ESF) & National Sources, in the framework of the program "Pythagoras I" of the "Operational Program for Education and Initial Vocational Training" of the 3rd Community Support Framework of the Hellenic Ministry of Education.

†Department of Mathematics, University of Ioannina, GR45110 Greece. (dnoutsos@uoi.gr)

‡Department of Informatics, Athens University of Economics and Business, GR10434 Greece. (pvassal@aueb.gr)

In the case where f is a positive function the matrix becomes a well-conditioned Hermitian positive definite matrix. In addition, if f is also an even function, it becomes a well-conditioned symmetric positive definite (spd) matrix. For this case, preconditioners belonging to some trigonometric matrix algebra have been proposed to achieve superlinear convergence of the PCG method. Circulant preconditioners have been proposed by G. Strang [24], by R. Chan [7] and by R. Chan and M. Yeung [11] for well conditioned spd systems. τ preconditioners proposed for the same systems by D. Bini and F. Di Benedeto [2] and by by F. Di Benedeto [13]. To cover the well conditioned Hermitian positive definite case, Hartley prconditioners have been proposed by D. Bini and P. Favati [3] and by X.Q. Jin [16].

It is well known that matrices that belong to any trigonometric matrix algebra, when they are used as preconditioners, can not give superlinear convergence [17],[18]. Moreover, there are cases where the corespondent matrices are singular ones, as, e.g., in the case where f is a nonnegative function having roots of even order and the preconditioner matrix is chosen to be a circulant one of Strang type. In this specific case the system becomes an ill conditioned symmetric positive definite one. Problems with such kind of matrices arise in a variety of applications: signal and image processing, tomography, harmonic analysis and partial differential equations.

Band Toeplitz preconditioners are ideal to cover this case of ill conditioned systems. They succeed in making the condition number of the preconditioned system independent of the dimension n. First, R. Chan [8] proposed as preconditioner the band Toeplitz matrix generated by the trigonometric polynomial g that matches the roots of f. R. Chan and P. Tang [10] extended the previous preconditioner, to the ones based also to a kind of approximation of f and finally, S. Serra Capizzano [21] proposed the band Toeplitz preconditioner which is based on g that matches the roots and also on the best trigonometric Chebyshev approximation of the remaining positive part $\frac{f}{g}$.

Preconditioners based on τ algebra have studied by F. Di Benedeto , G. Fiorentino and S. Serra Capizzano [14], by F. Di Benedeto [12] and by Serra Capizzano [22], while ω-circulant preconditioners have been proposed by D. Potts and G. Steidl [20] and by R. Chan and W. K. Ching [9].

Finally, a mixed type preconditioner a product of band Toeplitz matrices and inverses of band Toeplitz matrices, based on the best rational approximation of the remaining positive part, has been studied and proposed by the authors in [19].

In this paper we study and propose as a preconditioner, a product of the band Toeplitz matrix generated by g and matrices that belong to any trigonometric algebra and correspond to an approximation of the positive part. The underline idea of the proposed scheme is to combine the well known advantages that each of the components of the product presents when it is used as a stand alone preconditioner. As a result we obtain a flexible preconditioner which can be applied to the system $T_n(f)x = b$ infusing superlinear convergence to the PCG method. Convergence theory of the proposed preconditioner is developed and an alternating technique is proposed in case where convergence is not achieved. Finally, we compare our method with the already known in the literature techniques.

The paper is organized as follows. In §2 we introduce the basic idea for the construction of the aforementioned preconditioners and study their computational cost. In §3 we develop the convergence theory in both cases of using band plus τ preconditioners and band plus circulant ones. In §4 we propose and study an alternating smoothing technique, for both cases, when the convergence properties studied in §3 do not hold. §5 is devoted to applications, to numerical experiments and to concluding remarks.

2. Band plus Algebra preconditioners. Let $f \in \mathcal{C}_{2\pi}$ be a 2π-periodic non-negative function with roots $x_0, x_1, \ldots, x_l$ of multiplicities $2k_1, 2k_2, \ldots, 2k_l$ respectively, with $k_1 + k_2 + \cdots + k_l = k$. Then f can be written as a product $g \cdot w$ where

$$g(x) = \prod_{i=1}^{l} (2 - 2\cos(x - x_i))^{k_i} \tag{2.1}$$

and with $w(x) > 0$ for every $x \in [-\pi, \pi]$

We define as a preconditioner for the system

$$T_n(f)x = b \tag{2.2}$$

the product of matrices

$$K_n^{\mathcal{A}}(f) = \mathcal{A}_n(\sqrt{w})T_n(g)\mathcal{A}_n(\sqrt{w}) = \mathcal{A}_n(h)T_n(g)\mathcal{A}_n(h) \tag{2.3}$$

with $\mathcal{A}_n \in \{\tau, \mathcal{C}, \mathcal{H}\}$, where $\{\tau, \mathcal{C}, \mathcal{H}\}$ is the set of matrices belonging to τ, Circulant and Hartley algebra, respectively. We have put for simplicity $h = \sqrt{w}$.

It is obvious from the construction of K, that it fulfils the fundamental properties that each preconditioner must have, i.e the positive definiteness and symmetry (Hermitian).

Although the idea of using as preconditioners for the system (2.2) a product of band Toeplitz matrices with τ, circulant or Hartley ones is not new (see e.g [9] or [23]), what we propose is more general and flexible in the sense that it can use as $\mathcal{A}_n$ any matrix belonging to $\{\tau, \mathcal{C}, \mathcal{H}\}$, can treat both symmetric and Hermitian systems ([23]) and can be efficiently extended to the $2D$ case.

2.1. Construction of the preconditioner-Computation cost. For the band Toeplitz matrix $T_n(g)$ things are straightforward. To construct $\mathcal{A}_n(h)$ we use the relation

$$\mathcal{A}_n(h) = Q_n \cdot Diag\,(h(\mathbf{u}^n)) \cdot Q_n^H$$

where the entries of the vector $\mathbf{u}^n$ are $\mathbf{u}_i^n = \frac{2\pi(i-1)}{n}$, $i = 1(1)n$ and Q_n is the Fourier matrix F_n for the circulant case or the matrix $Re(F_n) + Im(F_n)$ for the Hartley case. For the τ case we have $\mathbf{u}_i^n = \frac{\pi i}{(n+1)}$, $i = 1(1)n$ and $Q_n = \sqrt{\frac{2}{n+1}}[\sin(j\mathbf{u}_i^n)]_{i,j=1}^n$.

The evaluation of the function h at the points $\mathbf{u}^n$ requires the evaluation of the function w and the computation of real square roots, which can be done by a fast and simple algorithm based on "Newton's Method" and is a of $O(n)$ ops. In any case, the above procedure does not incur in the total asymptotic complexity of the method as it is implemented once per every n. The computation $Q \cdot \mathbf{v}$ is performed via Fast Fourier Transforms (or Fast Sine Transforms in the τ case) and requires $O(n \log n)$ ops. Finally, the 'inversion' of $T_n(g)$ can be done in $O(n \log p + p \log^2 p \log \frac{n}{p})$ ops, where p is its bandwidth, using the algorithm proposed in [4] or even better in $O(n)$ using the multigrid technique proposed in [15]. So, the total optimal cost of $O(n \log n)$ is preserved per each iteration of PCG.

3. Convergence Theory.

3.1. Convergence of the method: τ case. We start with the case where $\mathcal{A}_n \in \tau$. We will show that the main mass of the eigenvalues of the preconditioned matrix

$$(\tau_n(h)T_n(g)\tau_n(h))^{-1}T_n(f) \tag{3.1}$$

is clustered around unity. Before we give the main results for this case we give a definition and report a useful lemma.

DEFINITION 3.1. *The set of the continuous functions f for which the modulus of continuity $\omega(f,\delta)$ (see [25]) is $o(|log\delta|^{-1})$, is the Dini-Lipschitz class and is denoted by C^*.*

LEMMA 3.2. *Let $w \in C^*_{2\pi}$ be a positive and even function. Then, for any positive ϵ, there exist N and $M > 0$ such that for every $n > N$ at most M eigenvalues of the matrix $T_n(w) - \tau_n(w)$ have absolute value greater than ϵ.*

Proof. See [23], Theorem 2.1. □

THEOREM 3.3. *Let $T_n(f)$ be the Toeplitz matrix produced by a nonnegative function f in $C_{2\pi}$ which can be written as $f = g{\cdot}w$, where g the trigonometric polynomial of order k as it given by (2.1) and $w = h^2$ is a strictly positive even function belonging to C^*. Then, for every $\epsilon > 0$ there exist N and $\hat{M} > 0$ such that for every $n > N$ at most $\hat{M}$ eigenvalues of the preconditioned matrix (3.1) lie outside the interval $(1-\epsilon, 1+\epsilon)$.*

Proof. We begin with the observation that the matrix $T_n(f)$ can be written (see [5]) as $T_n(g)T_n(w) + L_1$, where L_1 is a low rank matrix. Taking into account the specific form of L_1, which contains only nonzero columns at the first and last k columns, we obtain that $rank(L_1) = rank(L_1^T) = 2k$ and $rank(L_1 + L_1^T) = 4k$. From the close relationship between τ matrices and band Toeplitz matrices we have that

$$\begin{aligned}
T_n(f) &= \frac{1}{2}(T_n(g)T_n(w) + L_1) + \frac{1}{2}(T_n(w)T_n(g) + L_1^T) \\
&= \frac{1}{2}((\tau_n(g) + L_2)T_n(w) + L_1) + \frac{1}{2}(T_n(w)(\tau_n(g) + L_2) + L_1^T) \\
&= \frac{1}{2}\tau_n(g)T_n(w) + \frac{1}{2}\tau_n(g)T_n(w) + L_3,
\end{aligned}$$

where L_2 and L_3 are low rank symmetric matrices. More specifically, as L_2 has nonzero elements only at the upper left and lower right corner, the factor $L_2T_n(\omega) + T_n(\omega)L_2$ has nonzero entries only in the $k-1$ first and last rows and columns, i.e it is a border matrix. So, the rank of the matrix L_3 is at most $4k$. To study the spectrum of the preconditioned matrix $K_n^\tau(f)^{-1}T_n(f)$ with $K_n^\tau(f)^{-1}$ as in (2.3), we consider the symmetric form of it $\hat{T}_n = T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}T_n(f)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}}$, which is similar to the first one. So

$$\begin{aligned}
\hat{T}_n &= T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}T_n(f)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}} \\
&= \frac{1}{2}T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}\left(\tau_n(g)T_n(w) + T_n(w)\tau_n(g) + L_3\right)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}} \\
&= \frac{1}{2}T_n(g)^{-\frac{1}{2}}\tau_n(g)\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}} \\
&+ \frac{1}{2}T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}\tau_n(g)T_n(g)^{-\frac{1}{2}} + L_4 \\
&= \frac{1}{2}T_n(g)^{-\frac{1}{2}}(T_n(g) - L_2)\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}(T_n(g)-L_2)T_n(g)^{-\frac{1}{2}}+L_4\\
&=\frac{1}{2}T_n(g)^{\frac{1}{2}}\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}T_n(g)^{-\frac{1}{2}}\\
&+\frac{1}{2}T_n(g)^{-\frac{1}{2}}\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1}T_n(g)^{\frac{1}{2}}+L_5,
\end{aligned}
$$

where L_3, L_4 and L_5 are dense symmetric matrices of low rank, with $rank(L_4) = rank(L_3)$ and therefore the rank of L_5 is at most $8k-4$ (the rank of L_4 plus twice the rank of L_2).

From Lemma 3.2 we obtain that for the choice of $\epsilon_h > 0$ there exist a low rank (of constant rank) matrix L_6 and a matrix E of small norm ($\|E\|_2 \le \epsilon_h$), such that

$$
\tau_n(h)^{-1}T_n(w)\tau_n(h)^{-1} = I + E + L_6, \tag{3.2}
$$

where I is the n-dimensional identity matrix. Hence

$$
\begin{aligned}
\hat{T}_n &= \frac{1}{2}T_n(g)^{\frac{1}{2}}(I+E+L_6)T_n(g)^{-\frac{1}{2}}+\frac{1}{2}T_n(g)^{-\frac{1}{2}}(I+E+L_6)T_n(g)^{\frac{1}{2}}\\
&+L_5 = I+\frac{1}{2}T_n(g)^{\frac{1}{2}}ET_n(g)^{-\frac{1}{2}}+\frac{1}{2}T_n(g)^{-\frac{1}{2}}ET_n(g)^{\frac{1}{2}}+L,
\end{aligned}
$$

where L is a symmetric low rank matrix with its rank being no greater than the sum of the rank of L_5 and the double of the one L_6.

The proof of the main issue that $\hat{T}_n$ has a clustering at one, is reduced to the proof that for every $\epsilon > 0$, there exists $\epsilon_h > 0$, with $\|E\|_2 \le \epsilon_h$, such that all the eigenvalues of the matrix

$$
\hat{A}_n = \frac{1}{2}T_n(g)^{\frac{1}{2}}ET_n(g)^{-\frac{1}{2}}+\frac{1}{2}T_n(g)^{-\frac{1}{2}}ET_n(g)^{\frac{1}{2}}
$$

belong in the interval $(-\epsilon,\epsilon)$. Equivalently, since $\hat{A}_n$ is symmetric, we have to prove that both matrices $\epsilon I + \hat{A}_n$ and $\epsilon I - \hat{A}_n$ are positive definite matrices.

First, we prove that $\epsilon I + \hat{A}_n$ is positive definite. This is equivalent to proving that

$$
T_n(g)^{\frac{1}{2}}(\epsilon I+\hat{A}_n)T_n(g)^{\frac{1}{2}} = \epsilon T_n(g)+\frac{1}{2}T_n(g)E+\frac{1}{2}ET_n(g)
$$

is a positive definite matrix. For this, we consider a normalized vector $x \in I\!R^n$, ($\|x\|_2 = 1$) and take the Rayleigh quotient

$$
\mathrm{r} = \epsilon x^T T_n(g)x+\frac{1}{2}x^T T_n(g)Ex+\frac{1}{2}x^T ET_n(g)x = \epsilon x^T T_n(g)x + x^T T_n(g)Ex.
$$

The norm of the vector $y = Ex$ is given by

$$
\hat{\epsilon} = \|y\|_2 = \|Ex\|_2 \le \|E\|_2\|x\|_2 \le \epsilon_h.
$$

Let z be the normalized vector of y, so $y = \hat{\epsilon}z$, then the Rayleigh quotient takes the form

$$
r = \epsilon x^T T_n(g)x + \hat{\epsilon}x^T T_n(g)z. \tag{3.3}
$$

The second term of (3.3) takes the minimum value for z being the normalized vector of $-T_n(g)x$. So,

$$
\begin{aligned}
r &\geq \epsilon x^T T_n(g)x - \hat{\epsilon}\frac{x^T T_n(g)^2 x}{\|T_n(g)x\|_2} \geq \epsilon\|T_n(g)^{\frac{1}{2}}x\|_2 - \epsilon_h\frac{\|T_n(g)x\|_2^2}{\|T_n(g)x\|_2} \\
&= \epsilon\|T_n(g)^{\frac{1}{2}}x\|_2 - \epsilon_h\|T_n(g)x\|_2 \geq \epsilon\|T_n(g)^{\frac{1}{2}}x\|_2 \\
&- \epsilon_h\|T_n(g)^{\frac{1}{2}}\|_2\|T_n(g)^{\frac{1}{2}}x\|_2 = \left(\epsilon - \epsilon_h\|T_n(g)^{\frac{1}{2}}\|_2\right)\|T_n(g)^{\frac{1}{2}}x\|_2.
\end{aligned}
$$

Since the operator $T(g)$ is bounded, we can choose the value of ϵ_h to be such that

$$
\epsilon > \epsilon_h\|T_n(g)^{\frac{1}{2}}\|_2, \tag{3.4}
$$

so that the Rayleigh quotient r will be positive since $\|x\|_2 = 1$. This holds true for every choice of x, so the matrix $\epsilon I + \hat{A}_n$ is a positive definite matrix.

To prove that the second matrix $\epsilon I - \hat{A}_n$ is positive definite we follow exactly the same argumentation and we end up with

$$
r = \epsilon x^T T_n(g)x - \hat{\epsilon}x^T T_n(g)z.
$$

in the place of (3.3). Then, the second term takes its maximum value for z being the normalized vector of $T_n(g)x$. After that, the proof follows the same step and the same conclusion is deduced. □

We will prove now the important feature that our preconditioner fulfils and leads to superlinear convergence of PCG. The clustering of the eigenvalues around 1 has been proven in Theorem 3.3. So, we have to prove that the outliers are uniformly far away from zero and from infinity. For this we will study Rayleigh quotients of the preconditioned matrix:

$$
\lambda_{\min}(K_n^{\tau^{-1}}T_n(f)) = \inf_{x\in I\!R^n}\frac{x^T K_n^\tau(f)^{-\frac{1}{2}}T_n(f)K_n^\tau(f)^{-\frac{1}{2}}x}{x^T x} = \inf_{x\in I\!R^n}\frac{x^T T_n(f)x}{x^T K_n^\tau(f)x} \tag{3.5}
$$

and

$$
\lambda_{\max}(K_n^{\tau^{-1}}T_n(f)) = \sup_{x\in I\!R^n}\frac{x^T K_n^\tau(f)^{-\frac{1}{2}}T_n(f)K_n^\tau(f)^{-\frac{1}{2}}x}{x^T x} = \sup_{x\in I\!R^n}\frac{x^T T_n(f)x}{x^T K_n^\tau(f)x}. \tag{3.6}
$$

Thus, we have to study the range of the Rayleigh quotient

$$
\frac{x^T T_n(f)x}{x^T K_n^\tau(f)x} = \frac{x^T T_n(f)x}{x^T \tau_n(h)T_n(g)\tau_n(h)x} = \frac{x^T T_n(f)x}{x^T T_n(g)x}\cdot\frac{x^T T_n(g)x}{x^T \tau_n(h)T_n(g)\tau_n(h)x}.
$$

It is well known that the range of the first Rayleigh quotient is contained in the range of the function $w = \frac{f}{g}$ which is positive and far from zero and infinity. Therefore, we have to prove that

$$
\begin{aligned}
&\liminf_{n\to\infty}\inf_{x\in I\!R^n}\frac{x^T T_n(g)x}{x^T \tau_n(h)T_n(g)\tau_n(h)x} > 0, \\
&\limsup_{n\to\infty}\sup_{x\in I\!R^n}\frac{x^T T_n(g)x}{x^T \tau_n(h)T_n(g)\tau_n(h)x} < \infty.
\end{aligned}
$$

We will prove only the first inequality of (3.7). The proof of the second one is similar. This is obtained from the observations that

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}^n} \frac{x^T T_n(g)x}{x^T \tau_n(h)T_n(g)\tau_n(h)x} = \infty \Leftrightarrow \liminf_{n\to\infty} \inf_{x\in\mathbb{R}^n} \frac{x^T \tau_n(h)T_n(g)\tau_n(h)x}{x^T T_n(g)x} = 0$$

and

$$\liminf_{n\to\infty} \inf_{x\in\mathbb{R}^n} \frac{x^T \tau_n(h)T_n(g)\tau_n(h)x}{x^T T_n(g)x} = \liminf_{n\to\infty} \inf_{x\in\mathbb{R}^n} \frac{x^T T_n(g)x}{x^T \tau_n(h^{-1})T_n(g)\tau_n(h^{-1})x}.$$

So, the proof of the second inequality of (3.7) is equivalent to the proof of the first one with the function h^{-1} in the place of h.

By inverting the ratio of the first inequality of (3.7) it is equivalent to proving that

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}^n} \frac{x^T \tau_n(h)T_n(g)\tau_n(h)x}{x^T T_n(g)x} < \infty, \tag{3.7}$$

so, we have to study the ratio

$$r_x = \frac{x^T \tau_n(h)T_n(g)\tau_n(h)x}{x^T T_n(g)x}. \tag{3.8}$$

It is well known that the band Toeplitz matrix $T_n(g)$ is written as a τ plus a Hankel matrix

$$T_n(g) = \tau_n(g) + H_n(g), \tag{3.9}$$

where $H_n(g)$ is the Hankel matrix of rank $2(k-1)$ of the form

$$H_n(g) = E_n(g) + E_n(g)^R, \tag{3.10}$$

with

$$E_n(g) = \begin{pmatrix} g_2 & g_3 & \cdots & g_k & \cdots & \cdots & 0 \\ g_3 & & \iddots & & \iddots & & \vdots \\ \vdots & \iddots & & & & & \\ g_k & & & & & & \vdots \\ \vdots & \iddots & & & & & \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & \cdots & & \cdots & 0 \end{pmatrix} \tag{3.11}$$

and $E_n(g)^R$ is obtained from the matrix $E_n(g)$ by taking all its rows and columns in reverse order. The entries g_i are the Fourier coefficients of the trigonometric polynomial g ($g(x) = g_0 + 2g_1\cos(x) + 2g_2\cos(2x) + \cdots + 2g_k\cos(kx)$). In the special case where the root is 0 of multiplicity $2k$ we have that $g_i = \begin{pmatrix} 2k \\ k-i \end{pmatrix}$. It is obvious that for $k=1$, $H_n(g) = 0$, which means that $T_n(g)$ (the Laplace matrix) is a τ matrix and the problem is solved. In case where $k=2$ we have that $H_n(g)$ is a semi-positive definite matrix of rank 2 with just ones in the positions $(1,1)$ and (n,n) and zeros elsewhere. In case $k > 2$, the matrix $H_n(g)$ becomes indefinite. We denote by Δ the $(k-1) \times (k-1)$ matrix formed by the first $k-1$ rows and columns of $E_n(g)$:

$$\Delta = \begin{pmatrix} g_2 & g_3 & \cdots & g_k \\ g_3 & & \iddots & 0 \\ \vdots & \iddots & & \vdots \\ g_k & 0 & \cdots & 0 \end{pmatrix} \tag{3.12}$$

and by Δ^R the matrix obtained from Δ by taking all its rows and columns in reverse order. For an n-dimensional vector x we denote by $\overline{x}^{(m)}$ and by $\underline{x}^{(m)}$ the m-dimensional vectors formed from the first and last m entries of x, respectively.

Recalling ratio (3.8), we get

$$\begin{aligned} r_x &= \frac{x^T\tau_n(h)T_n(g)\tau_n(h)x}{x^TT_n(g)x} = \frac{x^T\tau_n(h)\tau_n(g)\tau_n(h)x + x^T\tau_n(h)H_n(g)\tau_n(h)x}{x^T\tau_n(g)x + x^TH_n(g)x} \\ &= \frac{x^T\tau_n(h^2g)x + x^T\tau_n(h)H_n(g)\tau_n(h)x}{x^T\tau_n(g)x + \overline{x}^{(k-1)^T}\Delta\overline{x}^{(k-1)} + \underline{x}^{(k-1)^T}\Delta^R\underline{x}^{(k-1)}}. \end{aligned} \tag{3.13}$$

LEMMA 3.4. *Let x be a normalized n-dimensional vector ($\|x\|_2 = 1$) and the sequence of the vectors $\overline{x}^{(k-1)}$ is bounded i.e. $0 < c \le \|\overline{x}^{(k-1)}\|_2 \le 1$ for all n or the sequence of the vectors $\underline{x}^{(k-1)}$ is bounded i.e. $0 < c \le \|\underline{x}^{(k-1)}\|_2 \le 1$ for all n , with c being constant independent of n, then the ratio r_x is bounded.*

Proof. The assumption $0 < c \le \|\overline{x}^{(k-1)}\|_2 \le 1$ or $0 < c \le \|\underline{x}^{(k-1)}\|_2 \le 1$ means that $\|\overline{x}^{(k-1)}\|_2 = O(1) \bigcap \Omega(1)$ or $\|\underline{x}^{(k-1)}\|_2 = O(1) \bigcap \Omega(1)$, respectively. Without loss of generality, we suppose that $\|\overline{x}^{(k-1)}\|_2 = O(1) \bigcap \Omega(1)$, the proof for the case where $\|\underline{x}^{(k-1)}\|_2 = O(1) \bigcap \Omega(1)$ being the same. It is easily proved that there is a constant integer m independent of n such that $\|\overline{x}^{(m)}\|_2 = O(1) \bigcap \Omega(1)$ end $\|y^{(k)}\|_2 = o(1)$ where $y^{(k)}$ is the k-dimensional vector of the entries of x followed by the vector $\overline{x}^{(m)}$. This is true since otherwise there would be an infinitely large integer m, depending on n, such that every block of size k of the vector $\overline{x}^{(m)}$ should have constant norm independent of n. The latter is a contradiction since then $\|\overline{x}^{(m)}\|_2 \to \infty$. Since both the numerator and the denominator of the ratio in (3.8) are bounded from above, to prove that this ratio is bounded it is equivalent to prove that the denominator $x^TT_n(g)x$ is bounded from bellow far from zero for x of unit Euclidean norm. For this, we write the matrix $T_n(g)$ and the vector x in the following block form:

$$T_n(g) = \left(\begin{array}{c|c} T_m(g) & G \mid 0 \\ \hline \begin{array}{c} G^T \\ \hline 0 \end{array} & T_{n-m}(g) \end{array}\right), \quad x = \left(\begin{array}{c} \overline{x}^{(m)} \\ \hline y^{(k)} \\ \hline z \end{array}\right),$$

where G is an $m \times k$ Toeplitz matrix with nonzero entries only in the k diagonals in the left bottom corner. We take now the denominator:

$$\begin{aligned} x^TT_n(g)x &= (\overline{x}^{(m)^T} | y^{(k)^T} | z^T) \left(\begin{array}{c|c} T_m(g) & G \mid 0 \\ \hline \begin{array}{c} G^T \\ \hline 0 \end{array} & T_{n-m}(g) \end{array}\right) \left(\begin{array}{c} \overline{x}^{(m)} \\ \hline y^{(k)} \\ \hline z \end{array}\right) = \overline{x}^{(m)^T} \\ & T_m(g)\overline{x}^{(m)} + 2\overline{x}^{(m)^T}Gy^{(k)} + (y^{(k)^T}|z^T)T_{n-m}(g)\left(\begin{array}{c} y^{(k)} \\ \hline z \end{array}\right). \end{aligned} \tag{3.14}$$

Since $T_m(g)$ and $T_{n-m}(g)$ are positive definite matrices the first and the third terms in the sum of (3.14) are both positive numbers. The minimum value of the first term depends only on m, which is constant, and is of order $\frac{1}{m^{2k}}$ independently of n far from zero. The third term depends on n and may take small values near zero. The second term is the only one which may take negative values, but

$$|2\overline{x}^{(m)^T} G y^{(k)}| = 2\|\overline{x}^{(m)^T} G y^{(k)}\|_2 \leq 2\|\overline{x}^{(m)}\|_2 \|G\|_2 \|y^{(k)}\|_2 = o(1),$$

since $\|y^{(k)}\|_2 = o(1)$ and the other norms are constants. As a consequence, the first term is absolutely greater in order of magnitude than the second one, which characterizes the bounded behavior of all the sum, and our assertion has been proven.□

It remains to study the quantity r_x for vector sequences x such that

$$\|\overline{x}^{(m)}\|_2 = o(1) \quad \text{and} \quad \|\underline{x}^{(m)}\|_2 = o(1) \tag{3.15}$$

for each constant m independent of n. First, we write the vector x as a convex combination of the eigenvectors v_is of τ algebra, with entries $(v_i)_j = \sqrt{\frac{2}{n+1}} \sin(\frac{\pi i j}{n+1})$:

$$x = \sum_{i=1}^{n} c_i v_i, \qquad \sum_{i=1}^{n} |c_i|^2 = 1. \tag{3.16}$$

We denote by $\mathbf{D}$ the denominator and by $\mathbf{N}$ the numerator of the ratio r_x of (3.13). So the denominator is given by

$$\begin{aligned} \mathbf{D} &= \textstyle\sum_{i=1}^n c_i v_i^T \tau_n(g) \sum_{i=1}^n c_i v_i + \sum_{i=1}^n c_i v_i^T H_n(g) \sum_{i=1}^n c_i v_i \\ &= \textstyle\sum_{i=1}^n c_i^2 g_i + \sum_{i=1}^n c_i v_i^T H_n(g) \sum_{i=1}^n c_i v_i \\ &= \textstyle\sum_{i=1}^n c_i^2 g_i + \sum_{i=1}^n c_i \overline{v}_i^T \Delta \sum_{i=1}^n c_i \overline{v}_i + \sum_{i=1}^n c_i \underline{v}_i^T \Delta^R \sum_{i=1}^n c_i \underline{v}_i, \end{aligned} \tag{3.17}$$

while the numerator is given by

$$\begin{aligned} \mathbf{N} &= \textstyle\sum_{i=1}^n c_i v_i^T \tau_n(h^2 g) \sum_{i=1}^n c_i v_i + \sum_{i=1}^n c_i v_i^T \tau_n(h) H_n(g) \tau_n(h) \\ &\times \textstyle\sum_{i=1}^n c_i v_i = \sum_{i=1}^n c_i^2 h_i^2 g_i + \sum_{i=1}^n c_i h_i v_i^T H_n(g) \sum_{i=1}^n c_i h_i v_i \\ &= \textstyle\sum_{i=1}^n c_i^2 h_i^2 g_i + \sum_{i=1}^n c_i h_i \overline{v}_i^T \Delta \sum_{i=1}^n c_i h_i \overline{v}_i \\ &+ \textstyle\sum_{i=1}^n c_i h_i \underline{v}_i^T \Delta^R \sum_{i=1}^n c_i h_i \underline{v}_i, \end{aligned} \tag{3.18}$$

where $h_i = h(\frac{\pi i}{n+1}) > h_{\min} > 0$ and $g_i = g(\frac{\pi i}{n+1}) = (2 - 2\cos(\frac{\pi i}{n+1}))^k = (2\sin(\frac{\pi i}{2(n+1)}))^{2k}$. For simplicity, we have put $\overline{v}_i$ and $\underline{v}_i$ instead of $\overline{v}_i^{(k-1)}$ and $\underline{v}_i^{(k-1)}$, respectively. The first sum in both numerator and denominator is positive and we call it τ-term, since it corresponds to the Rayleigh quotient of a τ matrix. We call the other two terms, corresponding to the low rank correction matrices Δ and Δ^R, correction terms. The correction terms may take negative values. It is obvious that the τ-terms of the numerator and the denominator coincide with each other in order of magnitude for all the choices of the vector x, since

$$\sum_{i=1}^{n} c_i^2 h_i^2 g_i = \hat{h}^2 \sum_{i=1}^{n} c_i^2 g_i, \quad 0 < h_{\min} \leq \hat{h} \leq h_{\max} < \infty.$$

So, if the τ-terms are greater, in order of magnitude, than the associated correction terms, then r_x is bounded. The only case where r_x tends to infinity is that where the correction terms in the numerator exceed, in order of magnitude, either the associated τ-term and/or that of the denominator. We will try to find such cases, by comparing the τ-terms with the correction terms. Since the correction term corresponding to Δ^R

behaves exactly as the one corresponding to Δ, for simplicity we will compare only the τ-terms with the correction terms corresponding to Δ. In other words we consider that $|\overline{x}^T \Delta \overline{x}|$ is greater than or equal to $|\underline{x}^T \Delta^R \underline{x}|$, in order of magnitude. Given $\{N_n\}$ with $N_n = \{1, 2, \dots, n\}$ we define the sequence of subsets $\{S_n\}$ such that

$$(3.19) \quad \begin{array}{ll} 1) & S_n \subset N_n \forall n \\ 2) & \forall i_n \text{ sequence to which } i_k \in S_k \text{ we have } \lim_{n\to\infty} \frac{i_n}{n} = 0 \quad (i_n = o(n)). \end{array}$$

Accordingly the complementary sequence of subsets $\{Q_n\}$ is defined as

$$Q_n = N_n \setminus S_n. \tag{3.20}$$

It is obvious that the border of the above subsets S_n and Q_n is not clear, but this does not present any problem in the analysis that follows. However, we have to be careful to take only sequences belonging to $o(n)$ when dealing with $\{S_n\}$. We write the vector x as the sum $x = x_S + x_Q$ where

$$x_S = \sum_{i\in S_n} c_i v_i, \qquad x_Q = \sum_{i\in Q_n} c_i v_i. \tag{3.21}$$

We denote also by $\overline{x}_S = \sum_{i\in S_n} c_i \overline{v}_i$, $\underline{x}_S = \sum_{i\in S_n} c_i \underline{v}_i$, $\overline{x}_Q = \sum_{i\in Q_n} c_i \overline{v}_i$ and $\underline{x}_Q = \sum_{i\in Q_n} c_i \underline{v}_i$. In other words we separate the eigenvectors into those that correspond to "small" eigenvalues ($o(1)$) and those that correspond to "large" ones ($O(1) \bigcap \Omega(1)$).

We consider the sequences and

$$\{q_n\}_n = \{\sum_{i\in Q_n} c_i^2\}_n \quad \text{and} \quad \{s_n\}_n = \{\sum_{i\in S_n} c_i^2\}_n. \tag{3.22}$$

LEMMA 3.5. *Let x be such that $\|\overline{x}^{(k-1)}\|_2 = o(1)$ and $\|\underline{x}^{(k-1)}\|_2 = o(1)$ and the sequence $\{q_n\}_n$ of (3.22) is bounded, i.e. $0 < c \le q_n \le 1$, then the ratio r_x is bounded.*

Proof. In this case we have

$$x^T \tau_n(g) x = x_S^T \tau_n(g) x_S + x_Q^T \tau_n(g) x_Q = \sum_{i\in S_n} c_i^2 g_i + \sum_{i\in Q_n} c_i^2 g_i \sim c > 0,$$

since the eigenvalues of the second sum are bounded from bellow. On the other hand we have

$$|\overline{x}^{(k-1)^T} \Delta \overline{x}^{(k-1)}| \le \|\Delta\|_2 \|\overline{x}^{(k-1)}\|_2^2 = o(1),$$

since $\|\overline{x}^{(k-1)}\|_2 = o(1)$. We get the same conclusion for the term $|\underline{x}^{(k-1)^T} \Delta \underline{x}^{(k-1)}|$. So, the τ-term is the dominant term which is bounded from bellow. Since the numerator is bounded from above, r_x is bounded.□

LEMMA 3.6. *Let x be such that $\|\overline{x}^{(k-1)}\|_2 = o(1)$ and $\|\underline{x}^{(k-1)}\|_2 = o(1)$ and for the sequences $\{s_n\}_n$ and $\{q_n\}_n$ of (3.22) there hold $\lim_{n\to\infty} s_n = 1$, $\lim_{n\to\infty} q_n = 0$ with $\|\overline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$, then the ratio r_x is bounded.*

Proof. We suppose that the sequence $\{q_n\}_n$ tends to zero monotonically, since otherwise it can be split into monotonic subsequences.

The τ-term gives:

$$x^T \tau_n(g) x = \sum_{i\in S_n} c_i^2 g_i + \sum_{i\in Q_n} c_i^2 g_i, \tag{3.23}$$

while the correction term gives:

$$\overline{x}^T\Delta\overline{x} = (\overline{x}_S+\overline{x}_Q)^T\Delta(\overline{x}_S+\overline{x}_Q) = \overline{x}_S^T\Delta\overline{x}_S + 2\overline{x}_S^T\Delta\overline{x}_Q + \overline{x}_Q^T\Delta\overline{x}_Q. \tag{3.24}$$

For the vector $\overline{x}_Q$ we have

$$\|\overline{x}_Q\|_2 = \|\sum_{i\in Q_n} c_i\overline{v}_i\|_2 \le \sum_{i\in Q_n}|c_i|\|\overline{v}_i\|_2 \le \left(\sum_{i\in Q_n} c_i^2\right)^{\frac{1}{2}}\left(\sum_{i\in Q_n}\|\overline{v}_i\|_2^2\right)^{\frac{1}{2}} \sim (q_n)^{\frac{1}{2}},$$

since $\|\overline{v}_i\|_2^2 \sim \frac{1}{n}$, for all $i\in N_n$ and the cardinality of Q_n is $n-o(n)\sim n$. So, $\|\overline{x}_Q\|_2 = O\left((q_n)^{\frac{1}{2}}\right)$. Let $\|\overline{x}_Q\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$, then $|\overline{x}_Q^T\Delta\overline{x}_Q| \le \|\Delta\|_2\|\overline{x}_Q\|_2^2 = o(q_n)$, which means that the second sum of (3.23) exceeds the last one of (3.24) so,

$$x_Q^T T_n(g)x_Q = \sum_{i\in Q_n} c_i^2 g_i + \overline{x}_Q^T\Delta\overline{x}_Q + \underline{x}_Q^T\Delta^R\underline{x}_Q \sim q_n. \tag{3.25}$$

In the case where $\|\overline{x}_Q\|_2 \sim (q_n)^{\frac{1}{2}}$ we consider the quantity $x_Q^T T_n(g)x_Q$ and normalize the vector x_Q to the vector $\hat{x}_Q$ by multiplying by a number of order $(q_n)^{-\frac{1}{2}}$, such that $\|\hat{x}_Q\|_2 = 1$. If we consider the vector $\hat{x}_Q$ in the place of x, which means that there are no vectors of indices belonging to S_n in the convex combination, we get that $\sum_{i\in Q_n} c_i^2 = 1$ for the new coefficients c_is. Since $\|\overline{x}_Q\|_2 \sim (q_n)^{\frac{1}{2}}$ we obtain that $\|\overline{\hat{x}}_Q\|_2 \sim c > 0$. From Lemma 3.4, by replacing $\hat{x}_Q$ in the place of x, we obtain that $\hat{x}_Q^T T_n(g)\hat{x}_Q$ is bounded from bellow. If we come back to the quantity $x_Q^T T_n(g)x_Q$ by dividing the vector $\hat{x}_Q$ by the same number, we obtain the validity of (3.25). For the estimation of the associated term $x_Q^T\tau_n(h)^T T_n(g)\tau(h)x_Q$ of the numerator, we follow exactly the same steps in the proof by considering the vector $\tau_n(h)x$ in the place of x. So, we obtain

$$x_Q^T\tau_n(h)^T T_n(g)\tau_n(h)x_Q \sim x_Q^T T_n(g)x_Q \sim q_n. \tag{3.26}$$

Under the last assumption, $\|\overline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$, the remaining terms of (3.24) $\overline{x}_S^T\Delta\overline{x}_S$ and $2\overline{x}_S^T\Delta\overline{x}_Q$ are both absolutely smaller than q_n in order of magnitude. Exactly the same happens with the corresponding terms of the numerator. So, the order of the denominator of r_x is just the order of $\sum_{i\in S_n} c_i^2 g_i$ if it exceeds q_n or q_n otherwise, while the one of the numerator is just the order of $\sum_{i\in S_n} c_i^2 h_i^2 g_i$ if it exceeds q_n or q_n otherwise. In any case the numerator and the denominator coincide with each other, meaning that r_x is bounded.□

A useful definition is given here.

DEFINITION 3.7. *A positive and even function $h\in C_{2\pi}$ is said to be (m,ρ)-smooth function if it is an m times differentiable function in an open region of the point $\rho\in(-\pi,\pi)$ with $h^{(j)}(\rho)=0, j=1(1)m-1$ and $h^{(m)}(\rho)$ being bounded.*

LEMMA 3.8. *Let x be such that $\|\overline{x}^{(k-1)}\|_2 = o(1)$ and $\|\underline{x}^{(k-1)}\|_2 = o(1)$ and for the sequences $\{s_n\}_n$ and $\{q_n\}_n$ of (3.22) there hold $\lim_{n\to\infty} s_n = 1$, $\lim_{n\to\infty} q_n = 0$ with $\|\overline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$. Let also that h is a $(k-1,0)$-smooth function. Then, the ratio r_x is bounded.*

Proof. The proof follows exactly the same steps of Lemma 3.6 to obtain the same results until (3.26). In the sequel, we use the assumption that the function h is a

$(k-1,0)$-smooth function. By taking the Taylor expansion of h_is about the point zero we find

$$h_i = h\left(\frac{i\pi}{n+1}\right) = h_0 + \frac{(\frac{i\pi}{n+1})^{k-1}}{(k-1)!} h^{(k-1)}(\xi_i), \qquad \xi_i \in \left(0, \frac{i\pi}{n+1}\right). \tag{3.27}$$

Thus, the vector corresponding to $\overline{x}_S$ in the numerator is given by

$$\sum_{i\in S_n} h_i c_i \overline{v}_i = \sum_{i\in S_n} \left(h_0 + \frac{(\frac{i\pi}{n+1})^{k-1}}{(k-1)!} h^{(k-1)}(\xi_i)\right) c_i \overline{v}_i = h_0 \overline{x}_S + \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i,$$

where $\eta_i = \frac{h^{(k-1)}(\xi_i)}{(k-1)!}$, $i \in S_n$, bounded. The correction term of the numerator corresponding to Δ, is $\mathbf{Z} = \sum_{i=1}^n h_i c_i \overline{v}_i^T \Delta \sum_{i=1}^n h_i c_i \overline{v}_i$ which takes the form

$$\begin{aligned} \mathbf{Z} &= \textstyle\sum_{i\in S_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in S_n} h_i c_i \overline{v}_i + 2\sum_{i\in S_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i \\ &+ \textstyle\sum_{i\in Q_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i = \mathbf{Z}_1 + 2\mathbf{Z}_2 + \mathbf{Z}_3. \end{aligned} \tag{3.28}$$

We have proven that the third term $\mathbf{Z}_3$ coincides with q_n. The first term gives

$$\begin{aligned} \mathbf{Z}_1 &= \left(h_0 \overline{x}_S^T + \textstyle\sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i^T\right) \Delta \\ &\times \left(h_0 \overline{x}_S^T + \textstyle\sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i\right) \\ &= h_0^2 \overline{x}_S^T \Delta \overline{x}_S + 2h_0 \overline{x}_S^T \Delta \textstyle\sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \\ &+ \textstyle\sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i^T \Delta \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i, \end{aligned} \tag{3.29}$$

while the second one gives

$$\begin{aligned} \mathbf{Z}_2 &= \left(h_0 \overline{x}_S^T + \textstyle\sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i^T\right) \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i \\ &= h_0 \overline{x}_S^T \Delta \textstyle\sum_{i\in Q_n} h_i c_i \overline{v}_i + \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i^T \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i. \end{aligned} \tag{3.30}$$

First we will estimate the quantity $\mathbf{q} = \| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \|_2$. From $i \in S_n$ and the fact that $\overline{v}_i = \left(\sqrt{\frac{2}{n+1}} \sin\left(\frac{ij\pi}{n+1}\right)\right)_{j=1}^{k-1}$ we get that $\|\overline{v}_i\|_2 \sim \frac{i}{n^{\frac{3}{2}}}$. So,

$$\begin{aligned} \mathbf{q} &\leq \textstyle\sum_{i\in S_n} |\eta_i||c_i| \left(\frac{i\pi}{n+1}\right)^{k-1} \|\overline{v}_i\|_2 \sim \frac{\eta}{\sqrt{n}} \sum_{i\in S_n} |c_i| \left(\frac{i}{n}\right)^k \\ &\leq \textstyle\frac{\eta}{\sqrt{n}} \left(\sum_{i\in S_n} 1\right)^{\frac{1}{2}} \left(\sum_{i\in S_n} c_i^2 \left(\frac{i}{n}\right)^{2k}\right)^{\frac{1}{2}} \sim \sqrt{\frac{\#S_n}{n}} \left(\sum_{i\in S_n} c_i^2 g_i\right)^{\frac{1}{2}}, \end{aligned} \tag{3.31}$$

where $\eta \in (\min_i |\eta_i|, \max_i |\eta_i|)$ and $\#S_n$ means the cardinality of the set S_n. Since $\frac{\#S_n}{n} = o(1)$ we get that the quantity $\left(\sum_{i\in S_n} c_i^2 g_i\right)^{\frac{1}{2}}$, which is just the square root of the τ-term, exceeds $\| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \|_2$ in order of magnitude. Coming back to the terms $\mathbf{Z}_1$ and $\mathbf{Z}_2$ of the numerator we deduce that the order of the first term of $\mathbf{Z}_1$ in (3.29) is

$$|h_0^2 \overline{x}_S^T \Delta \overline{x}_S| \leq h_0^2 \|\overline{x}_S\|_2^2 \|\Delta\|_2 = \Omega(q_n),$$

which coincides with $\overline{x}_S^T \Delta \overline{x}_S$ of the denominator in (3.24). On the other hand we can prove that $|\overline{x}_S^T \Delta \overline{x}_S| \sim \|\overline{x}_S\|_2^2$ by taking into account the proof of Lemma 2.6 of [18]. In that work it was proved that

$$\overline{v}_i^T \Delta \overline{v}_j = \frac{2\sin^2(\theta)}{n+1} z_{ij}(\theta), \quad \theta = \frac{\pi}{n+1}, \quad i,j \in S_n$$

where

$$\lim_{\theta \to 0} z_{ij}(\theta) = ij \begin{pmatrix} 2k-4 \\ k-2 \end{pmatrix}.$$

Finally, we obtain that

$$\overline{x}_S^T \Delta \overline{x}_S = \frac{2\sin^2(\theta)}{n+1} \sum_{i \in S_n} \sum_{j \in S_n} c_i c_j z_{ij}(\theta) = \frac{2\sin^2(\theta)}{n+1} z(\theta),$$

where

$$\lim_{\theta \to 0} z(\theta) = \begin{pmatrix} 2k-4 \\ k-2 \end{pmatrix} \sum_{i \in S_n} \sum_{j \in S_n} i c_i j c_j = \begin{pmatrix} 2k-4 \\ k-2 \end{pmatrix} \left(\sum_{i \in S_n} i c_i \right)^2 \geq 0.$$

By applying the same considerations to the quantity $\|\overline{x}_S\|_2^2$, after a simple analysis, we have

$$\|\overline{x}_S\|_2^2 = \frac{2\sin^2(\theta)}{n+1} y(\theta),$$

where

$$\lim_{\theta \to 0} y(\theta) = \frac{(k-1)k(2k-1)}{6} \left(\sum_{i \in S_n} i c_i \right)^2 \geq 0.$$

From the relations above we conclude that the quantities $\overline{x}_S^T \Delta \overline{x}_S$ and $\|\overline{x}_S\|_2^2$ have the same order of magnitude.

The order of the second term of $\mathbf{Z}_1$ in (3.29) is

$$\left| 2h_0 \overline{x}_S^T \Delta \sum_{i \in S_n} \left(\frac{i\pi}{n+1} \right)^{k-1} \eta_i c_i \overline{v}_i \right| \leq 2h_0 \|\overline{x}_S\|_2 \|\Delta\|_2 \left\| \sum_{i \in S_n} \left(\frac{i\pi}{n+1} \right)^{k-1} \eta_i c_i \overline{v}_i \right\|_2$$

$$= \|\overline{x}_S\|_2 \times o\left(\left(\sum_{i \in S_n} c_i^2 g_i \right)^{\frac{1}{2}} \right).$$

This term is less than the first one, in order of magnitude, if $\sum_{i \in S_n} c_i^2 g_i = O\left(\|\overline{x}_S\|_2^2\right)$ while it is less than the corresponding τ-term, in order of magnitude, if $\sum_{i \in S_n} c_i^2 g_i = \Omega\left(\|\overline{x}_S\|_2^2\right)$. In any case it does not play a role in the order of magnitude of the numerator. We arrive at the same conclusion regarding the order of the third term of $\mathbf{Z}_1$ in (3.29) which is $o\left(\sum_{i \in S_n} c_i^2 g_i\right)$.

For the terms of $\mathbf{Z}_2$ in (3.30) we first estimate the term $\left\| \sum_{i \in Q_n} h_i c_i \overline{v}_i \right\|_2$:

$$\left\| \sum_{i \in Q_n} h_i c_i \overline{v}_i \right\|_2 \leq \sum_{i \in Q_n} h_i |c_i| \|\overline{v}_i\|_2 \leq \left(\sum_{i \in Q_n} c_i^2 \right)^{\frac{1}{2}} \left(\sum_{i \in Q_n} h_i^2 \|\overline{v}_i\|_2^2 \right)^{\frac{1}{2}} \sim (q_n)^{\frac{1}{2}}.$$

Therefore, the order of the first term of $\mathbf{Z}_2$ in (3.30) is given by

$$\left| h_0 \overline{x}_S^T \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i \right| \leq h_0 \|\overline{x}_S\|_2 \|\Delta\|_2 \left\| \sum_{i\in Q_n} h_i c_i \overline{v}_i \right\|_2 = \|\overline{x}_S\|_2 \times O\left((q_n)^{\frac{1}{2}}\right),$$

which is less, in order of magnitude, than $\overline{x}_S^T \Delta \overline{x}_S$ in the denominator of (3.24). The order of the second term of $\mathbf{Z}_2$ in (3.30) is given by

$$\left| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \Delta \sum_{i\in Q_n} h_i c_i \overline{v}_i \right| \leq \left\| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \right\|_2 \|\Delta\|_2$$

$$\times \left\| \sum_{i\in Q_n} h_i c_i \overline{v}_i \right\|_2 = o\left(\left(\sum_{i\in S_n} c_i^2 g_i \right)^{\frac{1}{2}} \right) \times O\left((q_n)^{\frac{1}{2}}\right),$$

which is less, in order of magnitude, than the same term $\overline{x}_S^T \Delta \overline{x}_S$, if $\sum_{i\in S_n} c_i^2 g_i = O\left(\|\overline{x}_S\|_2^2\right)$ while it is less than the corresponding τ-term, in order of magnitude, if $\sum_{i\in S_n} c_i^2 g_i = \Omega\left(\|\overline{x}_S\|_2^2\right)$, since $\|\overline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$. □

THEOREM 3.9. *Let $f \in C_{2\pi}^*$ be an even function with roots $x_1, x_2, \ldots, x_l$ with multiplicities $2k_1, 2k_2, \ldots, 2k_l$, respectively, g be the trigonometric polynomial of order $k = \sum_{j=1}^{l} k_j$ given by (2.1), that rises the roots and w be the remaining positive part of f ($f = g \cdot w$). If the function $h = \sqrt{w}$ is a $(k_j - 1, x_j)$-smooth function for all $j = 1(1)l$, then the spectrum of the preconditioned matrix $K_n^\tau(f)^{-1} T_n(f)$ is bounded from above as well as from bellow:*

$$c < \lambda_{\min}(K_n^\tau(f)^{-1} T_n(f)) < \lambda_{\max}(K_n^\tau(f)^{-1} T_n(f)) < C,$$

where c and C are constants independent of the size n.

Proof. For the case of one zero at 0, Lemmata 3.4, 3.5, 3.6 and 3.8 cover all possible choices of the vector $x \in I\!R^n$ to obtain that the Rayleigh quotient r_x is bounded. The case of one zero at a point different from 0 is simple since it can be transformed to zero by a shift transformation of the interval $[-\pi, \pi]$. The generalization to more roots is straightforward. The main difference concerns on the definition of the sets S_n and Q_n of (3.19). Under the assumption of l roots $x_1, x_2, \ldots, x_l$, we give the new definition of the above sets

(3.32)
1) $S_n \subset N_n \forall n$
2) $\forall i_n$ sequence to which $i_k \in S_k$ we have $\lim_{n\to\infty} \frac{i_n}{n} - x_j = 0$
$(i_n - n x_j = o(n)), \quad j = 1, 2, \ldots, l.$

and

$$Q_n = N_n \setminus S_n. \tag{3.33}$$

After that definition, Lemmata 3.4, 3.5, 3.6 and 3.8 work well to obtain our result that r_x is bounded, which completes the proof of the Theorem. □

As a subsequent result we have that the minimum eigenvalue of $[K_n^\tau(f)]^{-1} T_n(f)$ is bounded far away from zero. Hence, from the theorem of Axelsson and Lindskog [1], it follows immediately that the PCG method will have superlinear convergence.

We have to remark here that if the smoothing condition of the function h does not hold, the Rayleigh quotient r_x may not be bounded and consequently the PCG method may not have superlinear convergence. The worst case, where we get the maximum value of r_x, is that when choosing $x = x_S$. In that case the denominator coincides with $\frac{1}{n^{2k}}$ and so for the numerator to be of the same order the $(k-1,0)$-smoothness of the function h is necessary. Otherwise, if h is a $(k-2,0)$-smooth function, which is the best possible choice, we deduce that the numerator coincides with $\frac{1}{n^{2k-1}}$. As a consequence, r_x tends to infinity with a rate coinciding with n.

3.2. Convergence of the method:Circulant case. For circulant matrices, in order to show the clustering of the eigenvalues of the preconditioned matrix sequence

$$(C_n(h)T_n(g)C_n(h))^{-1}T_n(f) \tag{3.34}$$

around unity, we first remark that although a band Toeplitz matrix and a circulant one do not commute, they very nearly have the commutativity property since

$$rank(T_n(g) \cdot C - C \cdot T_n(g)) \leq 2k,$$

where k is the bandwidth of the band matrix and which is obviously independent of the dimension n of the problem. We will show that the main mass of the eigenvalues of the preconditioned matrix (3.34) is clustered around unity. Before giving the main results for this case, we report a useful lemma.

LEMMA 3.10. *Let $w \in C_{2\pi}^*$ be a positive and even function. Then, for any positive ϵ, there exist N and $M > 0$ such that for every $n > N$ at most M eigenvalues of the matrix $C_n^{-1}T_n(w)$ have absolute value greater than ϵ.*

Proof. See [23], Theorem 2.1 (The proof for circulant case is just the same as the one for τ case). □

THEOREM 3.11. *Let $T_n(f)$ be the Toeplitz matrix produced by a nonnegative function f in $C_{2\pi}$ which can be written as $f = g \cdot w$, where g is the even trigonometric polynomial as is defined in (2.1) and $w = h^2$ is a strictly positive even function belonging to C^*. Then for every $\epsilon > 0$ there exist N and $\hat{M} > 0$ such that for every $n > N$ at most $\hat{M}$ eigenvalues of the preconditioned matrix (3.34) lie outside the interval $(1-\epsilon, 1+\epsilon)$.*

Proof. We follow exactly the same steps and the same considerations as in the proof of Theorem 3.3 for the τ case, with the only difference being that the matrices $C_n(g)$ and $C_n(h)$ replace $\tau_n(g)$ and $\tau_n(h)$, respectively. First we obtain that

$$\begin{aligned} \hat{T}_n &= \tfrac{1}{2}T_n(g)^{\frac{1}{2}}C_n(h)^{-1}T_n(w)C_n(h)^{-1}T_n(g)^{-\frac{1}{2}} \\ &+ \tfrac{1}{2}T_n(g)^{-\frac{1}{2}}C_n(h)^{-1}T_n(w)C_n(h)^{-1}T_n(g)^{\frac{1}{2}} + L_5, \end{aligned} \tag{3.35}$$

with L_5 being symmetric and a low rank matrix (of constant rank). It is noted that we have used the same notation $\hat{T}_n$ for the associated symmetric form of the preconditioned matrix.

From Lemma 3.10 we obtain that for the choice of $\epsilon_h > 0$ there exist a low rank (of constant rank) matrix L_6 and a matrix E of small norm ($\|E\|_2 \leq \epsilon_h$), such that

$$C_n(h)^{-1}T_n(w)C_n(h)^{-1} = I + E + L_6. \tag{3.36}$$

Consequently, we obtain the relation

$$\hat{T}_n = I + \frac{1}{2}T_n(g)^{\frac{1}{2}}ET_n(g)^{-\frac{1}{2}} + \frac{1}{2}T_n(g)^{-\frac{1}{2}}ET_n(g)^{\frac{1}{2}} + L,$$

which is nothing but relation (3.3) for the τ case.

After the latter manipulations, the proof follows step by step the one given in Theorem 3.3 the same result is obtained. □

As in the case of τ matrices, we will prove the important feature that our preconditioner satisfies the and leads to superlinear convergence of PCG.

The clustering of the eigenvalues around 1 has been proven in Theorem 3.11. We have to prove now that there does not exist any eigenvalue, belonging to the outliers, that tends to zero or to infinity. For this we will study Rayleigh quotients of the preconditioned matrix, as in the τ case. It is easily proved that the previous analysis, from relation (3.5) to relation (3.7), for the τ case, holds also for the circulant case by simply replacing $\tau_n(h)$ by $C_n(h)$.

Therefore, we have to prove that

$$\limsup_{n\to\infty} \sup_{x\in \mathbb{R}^n} \frac{x^T C_n(h) T_n(g) C_n(h) x}{x^T T_n(g) x} < \infty. \tag{3.37}$$

For this, we have to study the ratio

$$r_x = \frac{x^T C_n(h) T_n(g) C_n(h) x}{x^T T_n(g) x}. \tag{3.38}$$

It is well known that the band Toeplitz matrix $T_n(g)$ is written as a circulant minus a low rank Toeplitz matrix

$$T_n(g) = C_n(g) - \tilde{T}_n(g), \tag{3.39}$$

where $\tilde{T}_n(g)$ is a Toeplitz matrix of rank $2k$ of the form

$$\tilde{T}_n(g) = \tilde{J}_n(g) + \tilde{J}_n(g)^T, \tag{3.40}$$

with

$$\tilde{J}_n(g) = \begin{pmatrix} 0 & \cdots & \cdots & g_k & \cdots & g_2 & g_1 \\ \vdots & & & & \ddots & & g_2 \\ & & & \ddots & & \ddots & \vdots \\ \vdots & & & & & & g_k \\ & & & & & \ddots & \vdots \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & \cdots & & \cdots & 0 \end{pmatrix}, \tag{3.41}$$

where the entries g_i are the Fourier coefficients of the trigonometric polynomial g ($g(x) = g_0 + 2g_1\cos(x) + 2g_2\cos(2x) + \cdots + 2g_k\cos(kx)$). It is obvious that $\tilde{T}_n(g)$ is an indefinite matrix, while C_n is a semi positive definite one. We define by Δ the $k\times k$ matrix formed by the first k rows and the last k columns of $\tilde{J}_n(g)$:

$$\Delta = \begin{pmatrix} g_k & \cdots & g_2 & g_1 \\ 0 & \ddots & & g_2 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & g_k \end{pmatrix} \tag{3.42}$$

We use the same notations $\overline{x}^{(m)}$ and $\underline{x}^{(m)}$ for the first and the last m-dimensional blocks of the vector x, respectively.

Recalling ratio (3.38), we find

$$
\begin{array}{rcl}
r_x & = & \frac{x^T C_n(h) T_n(g) C_n(h) x}{x^T T_n(g) x} = \frac{x^T C_n(h) C_n(g) C_n(h) x - x^T C_n(h) \tilde{T}_n(g) C_n(h) x}{x^T C_n(g) x - x^T \tilde{T}_n(g) x} \\
 & = & \frac{x^T C_n(h^2 g) x - x^T C_n(h) \tilde{T}_n(g) C_n(h) x}{x^T C_n(g) x - \overline{x}^{(k)T} \Delta \underline{x}^{(k)} - \underline{x}^{(k)T} \Delta^T \overline{x}^{(k)}} = \frac{x^T C_n(h^2 g) x - x^T C_n(h) \tilde{T}_n(g) C_n(h) x}{x^T C_n(g) x - 2\overline{x}^{(k)T} \Delta \underline{x}^{(k)}}.
\end{array}
\tag{3.43}
$$

LEMMA 3.12. *Let x be a normalized n-dimensional vector ($\|x\|_2 = 1$) and the sequence of the vectors $\overline{x}^{(k)}$ is bounded i.e. $0 < c \leq \|\overline{x}^{(k)}\|_2 \leq 1$ for all n or the sequence of the vectors $\underline{x}^{(k)}$ is bounded i.e. $0 < c \leq \|\underline{x}^{(k)}\|_2 \leq 1$ for all n , with c being constant independent of n, then the ratio r_x in (3.43) is bounded.*

Proof. The proof follows the same steps of the one of Lemma 3.4 □

It remains to study the quantity r_x for vectors x such that

$$
\|\overline{x}^{(m)}\|_2 = o(1) \quad \text{and} \quad \|\underline{x}^{(m)}\|_2 = o(1) \tag{3.44}
$$

for each constant m independent of n. First, we write the vector x as a convex combination of the eigenvectors v_is of circulant algebra, which are the Fourier vectors with entries $(v_i)_j = \frac{1}{\sqrt{n}} e^{\mathrm{i}\frac{2(i-1)(j-1)\pi}{n}}$. The eigenvectors v_i are complex vectors while we are interested in real vectors x. Without loss of generality, we assume that n is even. It is easily seen that only the vectors v_1 and $v_{\frac{n}{2}+1}$ are real vectors while all the others are complex ones, where v_{n-i+1} is conjugate with v_{i+1}, $i = 1, 2, \ldots, \frac{n}{2} - 1$. To form the real vector x, we have to chose real coefficients c_is in the convex combination with $c_{n-i+1} = c_{i+1}$, $i = 1, 2, \ldots, \frac{n}{2} - 1$ So,

$$
\begin{array}{rcl}
x & = & c_1 v_1 + \sum_{i=2}^{\frac{n}{2}} c_i v_i + c_{\frac{n}{2}+1} v_{\frac{n}{2}+1} + \sum_{i=2}^{\frac{n}{2}} c_i v_i^* \\
 & = & c_1 v_1 + 2\sum_{i=2}^{\frac{n}{2}} c_i Re(v_i) + c_{\frac{n}{2}+1} v_{\frac{n}{2}+1},
\end{array}
\tag{3.45}
$$

where $c_1^2 + 2\sum_{i=2}^{\frac{n}{2}} c_i^2 + c_{\frac{n}{2}+1}^2 = 1$ and $Re(v_i)$ being the real part of v_i, with

$$
Re(v_i)_j = \frac{1}{\sqrt{n}} \cos\left(\frac{2(i-1)(j-1)\pi}{n}\right). \tag{3.46}
$$

For simplicity, in what follows we write the convex combination in the form $x = \sum_{i=1}^{n} c_i v_i$, but we will have in mind that the coefficients c_is are as they are described in (3.45).

As in the τ case we symbolize by $\mathbf{D}$ the denominator and by $\mathbf{N}$ the numerator of the ratio r_x of (3.43). Therefore, the denominator is given by

$$
\begin{array}{rcl}
\mathbf{D} & = & \sum_{i=1}^{n} c_i v_i^T C_n(g) \sum_{i=1}^{n} c_i v_i - \sum_{i=1}^{n} c_i v_i^T \tilde{T}_n(g) \sum_{i=1}^{n} c_i v_i \\
 & = & \sum_{i=1}^{n} c_i^2 g_i - \sum_{i=1}^{n} c_i v_i^T \tilde{T}_n(g) \sum_{i=1}^{n} c_i v_i \\
 & = & \sum_{i=1}^{n} c_i^2 g_i - 2\sum_{i=1}^{n} c_i \overline{v}_i^T \Delta \sum_{i=1}^{n} c_i \underline{v}_i,
\end{array}
\tag{3.47}
$$

while the numerator is given by

$$
\begin{array}{rcl}
\mathbf{N} & = & \sum_{i=1}^{n} c_i v_i^T C_n(h^2 g) \sum_{i=1}^{n} c_i v_i - \sum_{i=1}^{n} c_i v_i^T C_n(h) \tilde{T}_n(g) C_n(h) \\
 & \times & \sum_{i=1}^{n} c_i v_i = \sum_{i=1}^{n} c_i^2 h_i^2 g_i - \sum_{i=1}^{n} c_i h_i v_i^T \tilde{T}_n(g) \sum_{i=1}^{n} c_i h_i v_i \\
 & = & \sum_{i=1}^{n} c_i^2 h_i^2 g_i - 2\sum_{i=1}^{n} c_i h_i \overline{v}_i^T \Delta \sum_{i=1}^{n} c_i h_i \underline{v}_i,
\end{array}
\tag{3.48}
$$

where $h_i = h(\frac{2(i-1)\pi}{n}) > h_{min} > 0$ and $g_i = g(\frac{2(i-1)\pi}{n}) = (2 - 2\cos(\frac{2(i-1)\pi}{n}))^k = (2\sin(\frac{(i-1)\pi}{n}))^{2k}$. For simplicity, we have put $\overline{v}_i$ and $\underline{v}_i$ instead of $\overline{v}_i^{(k)}$ and $\underline{v}_i^{(k)}$, respectively. The first sum in both the numerator and the denominator is positive and we call it circulant term, since it corresponds to the Rayleigh quotient of a circulant matrix. We call correction term, the second term which corresponds to the low rank correction matrix Δ. It is obvious that the circulant terms of the numerator and the denominator coincide with each other, in order of magnitude, for all the choices of the vector x, since

$$\sum_{i=1}^{n} c_i^2 h_i^2 g_i = \hat{h}^2 \sum_{i=1}^{n} c_i^2 g_i, \quad 0 < h_{\min} \le \hat{h} \le h_{\max} < \infty.$$

Thus, if the circulant term is greater in order of magnitude than the associated correction term, then r_x is bounded. The only case where it tends to infinity is the one in which the correction term in the numerator exceeds, in order of magnitude, that of the associated circulant term as well as the denominator. We will try to find such cases, by comparing the circulant term with the correction one.

In analogy with the τ case we define the sequences of subsets $\{S_n\}$ and $\{Q_n\}$ as follows

(3.49)
1) $S_n \subset N_n \forall n$
2) $\forall i_n$ sequence to which $i_k \in S_k$ we have $\lim_{n\to\infty} \frac{i_n}{n} = 0$, or $\lim_{n\to\infty} \frac{n-i_n}{n} = 0$,

$$Q_n = N_n \setminus S_n. \tag{3.50}$$

We use the same notations for the vectors x_S, x_Q, $\overline{x}_S$, $\underline{x}_S$, $\overline{x}_Q$ and $\underline{x}_Q$, and consider the subsequences $\{q_n\}_n = \{\sum_{i\in Q_n} c_i^2\}_n$ and $\{s_n\}_n = \{\sum_{i\in S_n} c_i^2\}_n$.

LEMMA 3.13. *Let x be such that $\|\overline{x}^{(k)}\|_2 = o(1)$ and $\|\underline{x}^{(k)}\|_2 = o(1)$ and the sequence $\{q_n\}_n$ is bounded, i.e. $0 < c \le q_n \le 1$, then the ratio r_x is bounded.*

Proof. As in the τ case

$$x^T C_n(g) x = x_S^T C_n(g) x_S + x_Q^T C_n(g) x_Q = \sum_{i\in S_n} c_i^2 g_i + \sum_{i\in Q_n} c_i^2 g_i \sim c > 0,$$

since the eigenvalues of the second sum are bounded from bellow. On the other hand we find

$$|\overline{x}^T \Delta \underline{x}| \le \|\Delta\|_2 \|\overline{x}\|_2 \|\underline{x}\|_2 = o(1),$$

since we have proven that both $\|\overline{x}\|_2 = o(1)$ and $\|\underline{x}\|_2 = o(1)$. Hence, the circulant term is the dominant term which is bounded from bellow. Since the numerator is bounded from above, r_x is bounded.□

LEMMA 3.14. *Let x be such that $\|\overline{x}^{(k)}\|_2 = o(1)$ and $\|\underline{x}^{(k)}\|_2 = o(1)$ and for the sequences $\{s_n\}_n$ and $\{q_n\}_n$ there hold $\lim_{n\to\infty} s_n = 1$, $\lim_{n\to\infty} q_n = 0$ with $\|\overline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$ and $\|\underline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$, then the ratio r_x is bounded.*

Proof. We suppose that the sequence $\{q_n\}_n$ tends to zero monotonically, since otherwise it can be split into monotonic subsequences.

The circulant term gives:

$$x^T C_n(g) x = \sum_{i\in S_n} c_i^2 g_i + \sum_{i\in Q_n} c_i^2 g_i, \tag{3.51}$$

while the correction term gives:

$$\overline{x}^T\Delta\underline{x} = (\overline{x}_S+\overline{x}_Q)^T\Delta(\underline{x}_S+\underline{x}_Q) = \overline{x}_S^T\Delta\underline{x}_S+\overline{x}_S^T\Delta\underline{x}_Q+\overline{x}_Q^T\Delta\underline{x}_S+\overline{x}_Q^T\Delta\underline{x}_Q. \tag{3.52}$$

For the vector sequences $\overline{x}_Q$ and $\underline{x}_Q$ we have

$$\|\overline{x}_Q\|_2 = \|\sum_{i\in Q_n} c_i\overline{v}_i\|_2 \le \sum_{i\in Q_n}|c_i|\|\overline{v}_i\|_2 \le \left(\sum_{i\in Q_n}c_i^2\right)^{\frac{1}{2}}\left(\sum_{i\in Q_n}\|\overline{v}_i\|_2^2\right)^{\frac{1}{2}} \sim (q_n)^{\frac{1}{2}},$$

since $\|\overline{v}_i\|_2^2 \sim \frac{1}{n}$, for all $i\in N_n$ and the cardinality of Q_n is $n-o(n)\sim n$, while

$$\|\underline{x}_Q\|_2 = \|\sum_{i\in Q_n} c_i\underline{v}_i\|_2 \le \sum_{i\in Q_n}|c_i|\|\underline{v}_i\|_2 \le \left(\sum_{i\in Q_n}c_i^2\right)^{\frac{1}{2}}\left(\sum_{i\in Q_n}\|\underline{v}_i\|_2^2\right)^{\frac{1}{2}} \sim (q_n)^{\frac{1}{2}},$$

for the same reason. So, $\|\overline{x}_Q\|_2 = O\left((q_n)^{\frac{1}{2}}\right)$ and $\|\underline{x}_Q\|_2 = O\left((q_n)^{\frac{1}{2}}\right)$. Let $\|\overline{x}_Q\|_2\|\underline{x}_Q\|_2 = o(q_n)$, then $|\overline{x}_Q^T\Delta\underline{x}_Q| \le \|\Delta\|_2\|\overline{x}_Q\|_2\|\underline{x}_Q\|_2 = o(q_n)$, which means that the second sum of (3.51) exceeds the last one of (3.52) so,

$$x_Q^T T_n(g)x_Q = \sum_{i\in Q_n} c_i^2 g_i - 2\overline{x}_Q^T\Delta\underline{x}_Q \sim q_n. \tag{3.53}$$

In the case where $\|\overline{x}_Q\|_2 \sim \|\underline{x}_Q\|_2 \sim (q_n)^{\frac{1}{2}}$ we consider the quantity $x_Q^T T_n(g)x_Q$ and normalize the vector x_Q to the vector $\hat{x}_Q$ by multiplying by a number of order $(q_n)^{-\frac{1}{2}}$, such that $\|\hat{x}_Q\|_2 = 1$. If we consider the vector $\hat{x}_Q$ in the place of x, which means that in the convex combination we do not have any vectors with indices belonging to S_n, we get that $\sum_{i\in Q_n} c_i^2 = 1$ for the new coefficients c_is. Since $\|\overline{x}_Q\|_2 \sim (q_n)^{\frac{1}{2}}$ we obtain that $\overline{\hat{x}}_Q \sim c > 0$. From Lemma 3.12, by replacing $\hat{x}_Q$ in the place of x, we obtain that $\hat{x}_Q^T T_n(g)\hat{x}_Q$ is bounded from bellow. If we come back to the quantity $x_Q^T T_n(g)x_Q$ by dividing the vector $\hat{x}_Q$ by the same number, we obtain the validity of (3.54). For the estimation of the associated term $x_Q^T C_n(h)^T T_n(g)C_n(h)x_Q$ of the numerator, we follow exactly the same proof by considering the vector $C_n(h)x$ in the place of x. Therefore

$$x_Q^T C_n(h)^T T_n(g)C_n(h)x_Q \sim x_Q^T T_n(g)x_Q \sim q_n. \tag{3.54}$$

Under the assumptions $\|\overline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$ and $\|\underline{x}_S\|_2 = o\left((q_n)^{\frac{1}{2}}\right)$, the remaining terms $\overline{x}_S^T\Delta\underline{x}_S$, $\overline{x}_S^T\Delta\underline{x}_Q$ and $\overline{x}_Q^T\Delta\underline{x}_S$ of (3.52) are all absolutely smaller than q_n in order of magnitude. Exactly the same happens to the corresponding terms of the numerator. So, the order of the denominator of r_x is just the order of $\sum_{i\in S_n} c_i^2 g_i$ if it exceeds q_n or q_n otherwise, while the one of the numerator is just the order of $\sum_{i\in S_n} c_i^2 h_i^2 g_i$ if it exceeds q_n or q_n otherwise. In any case the numerator and the denominator coincide with each other, meaning that r_x is bounded.□

LEMMA 3.15. *Let x be such that $\|\overline{x}^{(k)}\|_2 = o(1)$ and $\|\underline{x}^{(k)}\|_2 = o(1)$ and for the sequences $\{s_n\}_n$ and $\{q_n\}_n$ there hold $\lim_{n\to\infty} s_n = 1$, $\lim_{n\to\infty} q_n = 0$ with $\|\overline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$ or $\|\underline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$. Let also that h is a $(k,0)$-smooth function. Then, the ratio r_x is bounded.*

Proof. The proof follows exactly the same steps of Lemma 3.14 to obtain the same results until (3.54). First, we will prove that $\|\overline{x}_S\|_2 \sim \|\underline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$, otherwise $\|\overline{x}_S\|_2, \|\underline{x}_S\|_2 = o\left(\sum_{i\in S_n} c_i^2 g_i^2\right)$. For this we assume, without loss of generality, that $\|\underline{x}_S\|_2 = o(\|\overline{x}_S\|_2)$ and are looking for a contradiction. From the considerations (3.45) and (3.46) it is easily seen that

$$(3.55)\quad (\overline{x}_S)_j = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i \cos\left(\frac{2(i-1)(j-1)\pi}{n}\right) = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i\cos((j-1)y_i)$$

$$(3.56)\quad (\underline{x}_S)_j = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i\cos((n-k+j-1)y_i) = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i\cos((k+1-j)y_i),$$

for all $j = 1,2,\cdots k$, where we have put $y_i = \frac{2(i-1)\pi}{n}$. It is obvious that $(\overline{x}_S)_j = (\underline{x}_S)_{k-j}$, $j = 2,3,\dots k$, which means that the above vectors have common entries with possible different orderings except for the first ones, i.e. $(\overline{x}_S)_1 = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i$ and $(\underline{x}_S)_1 = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i\cos(ky_i)$. To have different orders of magnitude in the vectors $\overline{x}_S$ and $\underline{x}_S$, it should be

$$(3.57)\qquad (\overline{x}_S)_1 = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i \sim \|\overline{x}_S\|_2$$

and

$$(3.58)\qquad (\overline{x}_S)_j = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i\cos(jy_i) = o\left(\|\overline{x}_S\|_2\right),\quad j = 1,2,\dots,k.$$

We consider now the vector $z = (z_1\ z_2\ \cdots z_k)^T$ which is bounded $\|z\| < \infty$, independent of n. From the difference in the order of magnitude of the entries in (3.57) and (3.58) we deduce that, for all such vectors, there holds

$$(3.59)\quad (\overline{x}_S)_1 = \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i \sim \frac{1}{\sqrt{n}}\sum_{i\in S_n} c_i - \frac{1}{\sqrt{n}}\sum_{j=1}^{k} z_j \sum_{i\in S_n} c_i\cos(jy_i) \sim \|\overline{x}_S\|_2.$$

The Taylor expansion with $k+1$ terms of $\cos(jy_i)$ gives

$$(3.60)\quad \cos(jy_i) = 1 - \frac{(jy_i)^2}{2} + \cdots (-1)^{(k-1)}\frac{(jy_i)^{2(k-1)}}{2(k-1)!} + (-1)^k\frac{(jy_i)^{2k}}{2k!}\cos(j\hat{y}_i),$$

where $\hat{y}_i \in (0, y_i)$. By replacing in (3.59) we find

$$(3.61)\quad \begin{array}{rcl} (\overline{x}_S)_1 & \sim & \frac{1}{\sqrt{n}}\left[\sum_{i\in S_n} c_i - \sum_{j=1}^k z_j\sum_{i\in S_n} c_i\cos(jy_i)\right] \\ & = & \frac{1}{\sqrt{n}}\left[\sum_{i\in S_n} c_i - \sum_{j=1}^k z_j\sum_{i\in S_n} c_i + \sum_{j=1}^k z_j\frac{j^2}{2}\sum_{i\in S_n} c_i y_i^2\right. \\ & - & \cdots + (-1)^{(k-1)}\sum_{j=1}^k z_j\frac{j^{2(k-1)}}{2(k-1)!}\sum_{i\in S_n} c_i y_i^{2(k-1)} \\ & + & \left.(-1)^k\sum_{j=1}^k z_j\frac{j^{2(k-1)}}{2(k-1)!}\sum_{i\in S_n} c_i y_i^{2k}\cos(j\hat{y}_i)\right]. \end{array}$$

If we choose the vector z such that

$$(3.62)\quad \sum_{j=1}^k z_j = 1,\quad \sum_{j=1}^k z_j j^2 = 0,\quad \sum_{j=1}^k z_j j^4 = 0,\quad \dots,\quad \sum_{j=1}^k z_j j^{2(k-1)} = 0,$$

all the terms in (3.61) are zero except the last one. Thus the order of $\|\overline{x}_S\|_2$ is given by

$$
(3.63) \qquad
\begin{array}{rcl}
\|\overline{x}_S\|_2 & \sim & |(\overline{x}_S)_1| \sim \frac{1}{\sqrt{n}} |\sum_{j=1}^{k} z_j \frac{j^{2k}}{2k!} \sum_{i\in S_n} c_i y_i^{2k} \cos(j\hat{y}_i)| \\
& \leq & \frac{1}{\sqrt{n}} \sum_{j=1}^{k} |z_j| \frac{j^{2k}}{2k!} \sum_{i\in S_n} |c_i| y_i^{2k} \sim \frac{1}{\sqrt{n}} \sum_{i\in S_n} |c_i| y_i^{2k} \\
& \leq & \frac{1}{\sqrt{n}} \left(\sum_{i\in S_n} 1\right)^{\frac{1}{2}} \left(\sum_{i\in S_n} c_i^2 y_i^{4k}\right)^{\frac{1}{2}} \sim \sqrt{\frac{\#S_n}{n}} \left(\sum_{i\in S_n} c_i^2 g_i^2\right)^{\frac{1}{2}} \\
& = & o\left(\sum_{i\in S_n} c_i^2 g_i^2\right)^{\frac{1}{2}},
\end{array}
$$

which constitutes a contradiction. In the case where $\|\overline{x}_S\|_2 = o\left(\sum_{i\in S_n} c_i^2 g_i^2\right)^{\frac{1}{2}}$, the ratio is bounded since the circulant term exceeds all the others. The choice (3.62) can be obtained from the solution of the $k \times k$ linear system

$$
(3.64) \qquad
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & 2^2 & 3^2 & \cdots & (k-1)^2 \\
1 & 2^4 & 3^4 & \cdots & (k-1)^4 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 2^{2(k-1)} & 3^{2(k-1)} & \cdots & (k-1)^{2(k-1)}
\end{pmatrix}
\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_k \end{pmatrix}
=
\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.
$$

This is a Vandermonde system which has a unique solution different from zero and bounded since it depends only on k and not on n.

In the same way we can prove that $\|\overline{x}_S^T \underline{x}_S\|_2 \sim \|\overline{x}_S\|_2^2 \sim \|\underline{x}_S\|_2^2$. By taking into account Lemma 2.9 and Lemma 2.4 of [18] we can prove that $|\overline{x}_S^T \Delta \underline{x}_S| \sim |\overline{x}_S^T \underline{x}_S|$, as we have done in the τ case.

As a consequence,

$$
(3.65) \qquad \|\overline{x}_S\|_2^2 \sim \|\underline{x}_S\|_2^2 \sim |\overline{x}_S^T \underline{x}_S| \sim |\overline{x}_S^T \Delta \underline{x}_S| = \Omega\left(q_n\right).
$$

In that case we use the assumption that the function h is a $(k,0)$-smooth function. By taking the Taylor expansion of h_is about the point zero we deduce

$$
h_i = h\left(\frac{2(i-1)\pi}{n}\right) = h_0 + \frac{\left(\frac{2(i-1)\pi}{n}\right)^k}{k!} h^{(k)}(\xi_i), \qquad \xi_i \in \left(0, \frac{2(i-1)\pi}{n}\right).
$$

Hence, the vector corresponding to $\overline{x}_S$ in the numerator is given by

$$
\sum_{i\in S_n} h_i c_i \overline{v}_i = \sum_{i\in S_n} \left(h_0 + \frac{\left(\frac{2(i-1)\pi}{n}\right)^k}{k!} h^{(k)}(\xi_i)\right) c_i \overline{v}_i = h_0 \overline{x}_S + \sum_{i\in S_n} \left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i \overline{v}_i,
$$

where $\eta_i = \frac{h^{(k)}(\xi_i)}{k!}$, $\quad i \in S_n$, is bounded, while the one corresponding to $\underline{x}_S$ is

$$
\sum_{i\in S_n} h_i c_i \underline{v}_i = h_0 \underline{x}_S + \sum_{i\in S_n} \left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i \underline{v}_i.
$$

The correction term of the numerator, $\mathbf{Z} = \sum_{i=1}^n h_i c_i \overline{v}_i^T \Delta \sum_{i=1}^n h_i c_i \underline{v}_i$ takes the form

$$
(3.66) \qquad
\begin{array}{rcl}
\mathbf{Z} & = & \sum_{i\in S_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in S_n} h_i c_i \underline{v}_i + \sum_{i\in S_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in Q_n} h_i c_i \underline{v}_i \\
& + & \sum_{i\in Q_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in S_n} h_i c_i \underline{v}_i + \sum_{i\in Q_n} h_i c_i \overline{v}_i^T \Delta \sum_{i\in Q_n} h_i c_i \underline{v}_i \\
& = & \mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3 + \mathbf{Z}_4.
\end{array}
$$

As a conclusion, we have proved that the fourth term $\mathbf{Z}_4$ does not exceed q_n. The other terms give

$$
\begin{aligned}
\mathbf{Z}_1 &= \left(h_0\overline{x}_S^T + \sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i \overline{v}_i^T\right)\Delta \\
(3.67)\qquad &\times \left(h_0\underline{x}_S^T + \sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i \underline{v}_i\right) = h_0^2\overline{x}_S^T\Delta\underline{x}_S \\
&+ h_0\overline{x}_S^T\Delta\sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i + h_0\sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \\
&\times \eta_i c_i\overline{v}_i\Delta\underline{x}_S^T + \sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i\overline{v}_i^T\Delta\sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i,
\end{aligned}
$$

$$
(3.68)\mathbf{Z}_2 = h_0\overline{x}_S^T\Delta\sum_{i\in Q_n} h_i c_i\underline{v}_i + \sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i\overline{v}_i^T\Delta\sum_{i\in Q_n} h_i c_i\underline{v}_i,
$$

$$
(3.69)\mathbf{Z}_3 = h_0\sum_{i\in Q_n} h_i c_i\overline{v}_i^T\Delta\underline{x}_S + \sum_{i\in Q_n} h_i c_i\overline{v}_i^T\Delta\sum_{i\in S_n}\left(\tfrac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i.
$$

First we estimate the quantities

$$
\overline{q} = \|\sum_{i\in S_n}\left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i\overline{v}_i\|_2 \quad\text{and}\quad \underline{q} = \|\sum_{i\in S_n}\left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i\|_2.
$$

From $i\in S_n$ and the fact that $\overline{v}_i = \left(\sqrt{\frac{1}{n}}e^{\frac{(i-1)(j-1)\pi}{n}}\right)_{j=1}^k$ we get that $\|\overline{v}_i\|_2 \sim \frac{1}{n^{\frac{1}{2}}}$. Therefore,

$$
\begin{aligned}
\overline{q} &\le \sum_{i\in S_n}|\eta_i||c_i|\left(\tfrac{2(i-1)\pi}{n}\right)^k\|\overline{v}_i\|_2 \sim \tfrac{\eta}{\sqrt{n}}\sum_{i\in S_n}|c_i|\left(\tfrac{i}{n}\right)^k \\
&\le \tfrac{\eta}{\sqrt{n}}\left(\sum_{i\in S_n}1\right)^{\frac{1}{2}}\left(\sum_{i\in S_n}c_i^2\left(\tfrac{i}{n}\right)^{2k}\right)^{\frac{1}{2}} \sim \sqrt{\tfrac{\#S_n}{n}}\left(\sum_{i\in S_n}c_i^2 g_i\right)^{\frac{1}{2}},
\end{aligned}
$$

where $\eta\in(\min_i|\eta_i|,\max_i|\eta_i|)$. Since $\frac{\#S_n}{n} = o(1)$ we deduce that $\overline{q} = o\left(\left(\sum_{i\in S_n}c_i^2 g_i\right)^{\frac{1}{2}}\right)$. For the same reason we get $\underline{q} = o\left(\left(\sum_{i\in S_n}c_i^2 g_i\right)^{\frac{1}{2}}\right)$. Coming back to the terms $\mathbf{Z}_1$, $\mathbf{Z}_2$ and $\mathbf{Z}_3$ of the numerator we get that the order of the first term of $\mathbf{Z}_1$ in (3.67) is $|h_0^2\overline{x}_S^T\Delta\underline{x}_S| = \Omega(q_n)$, given by (3.65), which coincides with $\overline{x}_S^T\Delta\underline{x}_S$ of the denominator in (3.52). The order of the second and the third term of $\mathbf{Z}_1$ in (3.67) are

$$
p_1 = \left|h_0\overline{x}_S^T\Delta\sum_{i\in S_n}\left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i\right| \le h_0\|\overline{x}_S\|_2\|\Delta\|_2\left\|\sum_{i\in S_n}\left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i\underline{v}_i\right\|_2
$$
$$
= \|\overline{x}_S\|_2\times o\left(\left(\sum_{i\in S_n}c_i^2 g_i\right)^{\frac{1}{2}}\right),
$$

$$
p_2 = \left|h_0\sum_{i\in S_n}\left(\frac{2(i-1)\pi}{n}\right)^k \eta_i c_i\overline{v}_i^T\Delta\underline{x}_S\right| = \|\overline{x}_S\|_2\times o\left(\left(\sum_{i\in S_n}c_i^2 g_i\right)^{\frac{1}{2}}\right).
$$

If $\sum_{i\in S_n}c_i^2 g_i = O\left(\|\overline{x}_S\|_2\right)$, then both terms are less than the first one of $\mathbf{Z}_1$, in order of magnitude. If $\sum_{i\in S_n}c_i^2 g_i = \Omega\left(\|\overline{x}_S\|_2\right)$, then both p_1 and p_2 are less than the corresponding circulant term, in order of magnitude. In any case they do not play any role in the order of magnitude of the numerator. We arrive at the same conclusion for the order of the third term of $\mathbf{Z}_1$ in (3.67) which is $o\left(\sum_{i\in S_n} c_i^2 g_i\right)$. For the terms of $\mathbf{Z}_2$ in (3.68) we find that the order of the first one is

$$\left| h_0 \overline{x}_S^T \Delta \sum_{i\in Q_n} h_i c_i \underline{v}_i \right| \le h_0 \|\overline{x}_S\|_2 \|\Delta\|_2 \left\| \sum_{i\in Q_n} h_i c_i \underline{v}_i \right\|_2 \sim \|\overline{x}_S\|_2 (q_n)^{\frac{1}{2}},$$

which coincides with $\overline{x}_S^T \Delta \underline{x}_Q$ of the denominator in (3.52). The order of the second term of $\mathbf{Z}_2$ in (3.68) is

$$\begin{aligned} \left| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \Delta \sum_{i\in Q_n} h_i c_i \underline{v}_i \right| &\le \left\| \sum_{i\in S_n} \left(\frac{i\pi}{n+1}\right)^{k-1} \eta_i c_i \overline{v}_i \right\|_2 \|\Delta\|_2 \\ &\times \left\| \sum_{i\in Q_n} h_i c_i \underline{v}_i \right\|_2 = o\left(\left(\sum_{i\in S_n} c_i^2 g_i \right)^{\frac{1}{2}} \right) \times (q_n)^{\frac{1}{2}}, \end{aligned}$$

which is less than the first one, in order of magnitude, if $\sum_{i\in S_n} c_i^2 g_i = O\left(\|\overline{x}_S\|_2^2\right)$ while it is less than the corresponding circulant term, in order of magnitude, if $\sum_{i\in S_n} c_i^2 g_i = \Omega\left(\|\overline{x}_S\|_2^2\right)$, since $\|\overline{x}_S\|_2 = \Omega\left((q_n)^{\frac{1}{2}}\right)$. Exactly the same happens with the terms of $\mathbf{Z}_3$ in (3.69).□

THEOREM 3.16. *Let $f \in C_{2\pi}^*$ be an even function with roots $x_0, x_1, \ldots, x_l$ with multiplicities $2k_1, 2k_2, \ldots, 2k_l$, respectively, g the trigonometric polynomial of order $k = \sum_{j=1}^{l} k_j$ given by (2.1), that rises the roots and w the remaining positive part of f ($f = g \cdot w$). If the function $h = \sqrt{w}$ is a (k_j, x_j)-smooth function for all x_js, $j = 1(1)l$, then the spectrum of the preconditioned matrix $K_n^C(f)^{-1}T_n(f)$ is bounded from above as well as from bellow:*

$$c < \lambda_{\min}([K_n^C(f)]^{-1}T_n(f)) < \lambda_{\max}([K_n^C(f)]^{-1}T_n(f)) < C, \tag{3.70}$$

where c and C are constants independent of the size n.

Proof. For the case of one zero at 0, Lemmata 3.12, 3.13, 3.14 and 3.15 cover all possible choices of the vector $x \in \mathbb{R}^n$ to obtain that the Rayleigh quotient r_x is bounded. The case of one zero at a point different from 0 is covered by a shift transformation of the interval $[-\pi, \pi]$. The generalization to more roots is straightforward. The main difference concerns on the definition of the sets S_n and Q_n of (3.49) and (3.50). We give the new definition of the above sets

(3.71)
1) $S_n \subset N_n \forall n$
2) $\forall i_n$ sequence to which $i_k \in S_k$ we have $\lim_{n\to\infty} \frac{i_n}{n} - x_j = 0$ or $\lim_{n\to\infty} \frac{n-i_n}{n} - x_j = 0, \quad j = 1, 2, \ldots, l.$

and

$$Q_n = N_n \setminus S_n. \tag{3.72}$$

After that definition, Lemmata 3.12, 3.13, 3.14 and 3.15 work well to obtain our result that r_x is bounded, which completes the proof of the Theorem. □

As a subsequent result we have that the minimum eigenvalue of $[K_n^C(f)]^{-1}T_n(f)$ is bounded far away from zero. Hence, from the theorem of Axelsson and Lindskog [1] it follows immediate that the PCG method will have superlinear convergence.

We have to remark here that if the smoothing condition of the function h does not hold, the Rayleigh quotient r_x may not be bounded and consequently the PCG method may not have superlinear convergence. The worst case, where we get the maximum value of r_x, is the one of choosing $x = x_S$. In that case the denominator coincides with $\frac{1}{n^{2k}}$ and so for the numerator to be of the same order the $(k,0)$-smoothness of the function h is necessary. Otherwise, if h is a $(k-1,0)$-smooth function, which is the best possible choice, we find that the numerator coincides with $\frac{1}{n^{2k-1}}$. Consequently, r_x tends to infinity with a rate coinciding with n.

REMARK 3.1. *Following a theory closely related to that just developed, band plus Hartley preconditioners could be applied for the solution of ill-conditioned Hermitian Toeplitz systems. In this paper, we do not study this case. We simply remark that a similar analysis could be applied to obtain analogous results for the superlinearity of the convergence. Since Hartley matrices are closely related to circulant matrices, we believe that $(k,0)$-smoothness, for the function h, is needed.*

4. Smoothing technique. Our analysis brings up the following question: Is the condition of smoothing valid for most of the applications? The answer to this question is not positive. There are problems where the positive part h is smooth enough but in most of them we are not guaranteed. In some of the problems the function h is not differentiable at 0, nor continuous. In the following two subsections we propose a smoothing technique which approximates h with a $(k-1,0)$-smooth function for the τ case and with a $(k,0)$-smooth function for the Circulant case, respectively, in order to get superlinear convergence.

4.1. Smoothing technique: τ case. Let assume that the factor h of the generating function f is not a $(k-1,0)$-smooth function. We define the function $\hat{h}$ as follows

$$\hat{h}(x) = \begin{cases} P_k[h](x) & \text{if} \quad x \in (-\epsilon, \epsilon) \\ h(x) & if \quad x \in [-\pi, -\epsilon] \cup [\epsilon, \pi] \end{cases} \tag{4.1}$$

where ϵ is a small positive constant and $P_k[h]$ is an even and a $(k-1,0)$-smooth function which interpolates h at the points $-\epsilon, 0, \epsilon$. It is obvious that we can choose as $P_k[h]$ the function

$$P_k[h](x) = \frac{h(\epsilon) - h_0}{\epsilon^k}|x|^k + h_0, \tag{4.2}$$

which is a k degree interpolation polynomial on the interval $(0, \epsilon)$, or the function

$$P_k[h](x) = \frac{h(\epsilon) - h_0}{(2 - 2\cos(\epsilon))^{\frac{k}{2}}}(2 - 2\cos(x))^{\frac{k}{2}} + h_0, \tag{4.3}$$

which, for even k, is a k degree interpolation trigonometric polynomial on the interval $(-\epsilon, \epsilon)$. For small ϵ the function $P_k[h]$ is a very good approximation of h on the interval $(-\epsilon, \epsilon)$. For this reason we propose as preconditioner the matrix

$$K_n^\tau(\hat{f}) = \tau_n(\hat{h})T_n(g)\tau_n(\hat{h}). \tag{4.4}$$

The smoothness identity of the function $\hat{f} = g \cdot \hat{h}^2$ is valid and Theorem 3.9 guarantees superlinear convergence of the PCG method with preconditioned matrix sequence $K_n^\tau(\hat{f})^{-1}T_n(f)$. We state here the generalization of Theorem 3.9.

THEOREM 4.1. *Let $f \in C_{2\pi}^*$ be an even function with roots $x_0, x_1, \ldots, x_l$ with multiplicities $2k_1, 2k_2, \ldots, 2k_l$, respectively, g the trigonometric polynomial of order $k = \sum_{j=1}^{l} k_j$ given by (2.1), that rises the roots, w the remaining positive part of f ($f = g \cdot w$) and $h = \sqrt{w}$. We define the function $\hat{h}$ as follows:*

$$\hat{h}(x) = \begin{cases} P_{k_j}[h](x) & \text{if } x \in (x_j - \epsilon_j, x_j + \epsilon_j), j = 1, 2, \ldots, l \text{ and} \\ & h \text{ is not a } (k_j - 1, x_j)\text{-smooth function} \\ h(x) & \text{elsewhere} \end{cases}, \tag{4.5}$$

where $\epsilon_j, j = 1, 2, \ldots, l$ are small positive constants and

$$P_{k_j}[h](x) = \frac{(x - x_j + \epsilon_j)h(x_j + \epsilon_j) - (x - x_j - \epsilon_j)h(x_j - \epsilon_j) - 2\epsilon_j h(x_j)}{2\epsilon_j^{k+1}} |x - x_j|^k + h(x_j) \qquad \text{or}$$

$$\begin{aligned} P_{k_j}[h](x) &= \frac{(2 - 2\cos(x - x_j + \epsilon_j))h(x_j + \epsilon_j) + (2 - 2\cos(x - x_j - \epsilon_j))h(x_j - \epsilon_j) - (2 - 2\cos(2\epsilon_j))h(x_j)}{(2 - 2\cos(2\epsilon_j)(2 - 2\cos(\epsilon_j))^{\frac{k}{2}}} \\ &\times \ (2 - 2\cos(x - x_j))^{\frac{k}{2}} + h(x_j). \end{aligned}$$

Then, the spectrum of the preconditioned matrix $K_n^\tau(\hat{f})^{-1}T_n(f)$ ($\hat{f} = g \cdot \hat{h}^2$) is bounded from above as well as from bellow:

$$c < \lambda_{\min}(K_n^\tau(\hat{f})^{-1}T_n(f)) < \lambda_{\max}(K_n^\tau(\hat{f})^{-1}T_n(f)) < C,$$

where c and C are constants independent of the size n. We have to remark here that the functions $P_{k_j}[h]$ have been taken to be interpolation functions of the function f at the points $x_j - \epsilon_j, x_j, x_j + \epsilon_j$ as we have done in relations (4.2) and (4.3) for the points $-\epsilon, 0, \epsilon$.

4.2. Smoothing technique: Circulant case. Let us assume that the factor h of the generating function f is not a $(k, 0)$-smooth function. Then, we define the function $\hat{h}$ given in (4.1), in analogy with the τ case. $P_k[h]$ is an even and a $(k, 0)$-smooth function which interpolates h at the points $-\epsilon, 0, \epsilon$ and could be chosen as

$$P_k[h](x) = \frac{h(\epsilon) - h_0}{\epsilon^k} |x|^{k+1} + h_0, \tag{4.6}$$

or

$$P_k[h](x) = \frac{h(\epsilon) - h_0}{(2 - 2\cos(\epsilon))^{\frac{k+1}{2}}} (2 - 2\cos(x))^{\frac{k+1}{2}} + h_0. \tag{4.7}$$

For small ϵ the function $P_k[h]$ is a very good approximation of h on the interval $(-\epsilon, \epsilon)$. Then, we propose as preconditioner the matrix

$$K_n^C(\hat{f}) = C_n(\hat{h})T_n(g)C_n(\hat{h}). \tag{4.8}$$

The smoothing identity of the function $\hat{f} = g \cdot \hat{h}^2$ is valid and Theorem 3.16 insures superlinear convergence of the PCG method with preconditioned matrix sequence $K_n^C(\hat{f})^{-1}T_n(f)$. We state here the generalization of Theorem 3.16.

THEOREM 4.2. *Let $f \in C_{2\pi}$ be an even function with roots $x_0, x_1, \ldots, x_l$ with multiplicities $2k_1, 2k_2, \ldots, 2k_l$, respectively, g the trigonometric polynomial of order*

$k = \sum_{j=1}^{l} k_j$ given by (2.1), that rises the roots, w the remaining positive part of f ($f = g \cdot w$) and $h = \sqrt{w}$. We define the function $\hat{h}$ as follows:

$$\hat{h}(x) = \begin{cases} P_{k_j}[h](x) & \textit{if } x \in (x_j - \epsilon_j, x_j + \epsilon_j), j = 1, 2, \ldots, l \textit{ and} \\ & h \textit{ is not a } (k_j, x_j)\textit{-smooth function} \\ h(x) & \textit{elsewhere} \end{cases}, \tag{4.9}$$

where $\epsilon_j, j = 1, 2, \ldots, l$ are small positive constants and

$$P_{k_j}[h](x) = \frac{(x-x_j+\epsilon_j)h(x_j+\epsilon_j)-(x-x_j-\epsilon_j)h(x_j-\epsilon_j)-2\epsilon_j h(x_j)}{2\epsilon_j^{k+2}}|x - x_j|^{k+1} + h(x_j) \qquad \textit{or}$$

$$\begin{aligned} P_{k_j}[h](x) &= \frac{(2-2\cos(x-x_j+\epsilon_j))h(x_j+\epsilon_j)+(2-2\cos(x-x_j+\epsilon_j))h(x_j-\epsilon_j)-(2-2\cos(2\epsilon_j))h(x_j)}{(2-2\cos(2\epsilon_j))(2-2\cos(\epsilon_j))^{\frac{k+1}{2}}} \\ &\times \ (2 - 2\cos(x - x_j))^{\frac{k+1}{2}} + h(x_j). \end{aligned}$$

Then, the spectrum of the preconditioned matrix $K_n^C(\hat{f})^{-1}T_n(f)$ ($\hat{f} = g \cdot \hat{h}^2$) is bounded from above as well as from bellow:

$$c < \lambda_{\min}(K_n^C(\hat{f})^{-1}T_n(f)) < \lambda_{\max}(K_n^C(\hat{f})^{-1}T_n(f)) < C, \tag{4.10}$$

where c and C are constants independent of the size n.

REMARK 4.1. *The same smoothing technique could be applied for the band plus Hartley preconditioners, when the function h is not a $(k, 0)$-smooth function.*

5. Numerical Experiments. In this section we report some numerical examples to show the efficiency of the proposed preconditioners and to confirm the validity of the presented theory. The experiments were carried out using Matlab. In all the examples the righthand side of the system was $(1\ 1\ \cdots\ 1)^T$ in order to compare our method with methods proposed by other researchers. We have run also our examples with the righthand side being random vectors and we have obtained results with the same behavior. The zero vector was as initial guess for the PCG method and as stopping criterion was taken the validity of the inequality $\frac{\|r^{(k)}\|_2}{\|r^{(0)}\|_2} \leq 10^{-7}$, where $r^{(k)}$ is the residual vector in the kth iteration.

EXAMPLE 5.1. We consider the function $f_1(x) = x^4$ as generating function. The associated function $h = \frac{x^2}{2-2\cos(x)}$ is a $(2, 0)$-smooth function and so, smoothing technique is not needed for both band plus τ and band plus circulant preconditioners. In Table 5.1 the number of iterations needed to achieve the predefined accuracy are illustrated. We compare the performance of our preconditioners with a variety of other well known and optimal preconditioners: R is the pioneering one proposed by R. Chan [8]. S^{*3} is the proposal of S. Serra Capizzano in [21] using best Chebyshev approximation (3 is the degree of the polynomial). $M^{(1,2)}$ is the preconditioner proposed by D. Noutsos and P. Vassalos in [19], which is based on best rational approximation with 1, 2 being the degrees of the numerator and denominator, respectively. W is the ω circulant preconditioner proposed by D. Potts and G. Steidl in [20]. Finally, by τ and $\mathcal{C}$, we denote the proposed in this paper band plus τ and band plus circulant preconditioners, respectively. The efficiency of our preconditioners is clearly shown.

EXAMPLE 5.2. Let

$$f_2(x) = \begin{cases} x^2|x+1| & |x| \leq \frac{\pi}{2} \\ (\frac{\pi}{2}+2)x^2 & x \in [-\pi, \pi] \setminus [-\frac{\pi}{2}, \frac{\pi}{2}] \end{cases}$$

TABLE 5.1
Number of iterations for $f_1(x) = x^4$

n	R	S^{*3}	$M^{1,2}$	W	τ	$\mathcal{C}$
32	15	11	6	7	5	6
64	20	11	8	8	5	6
128	24	12	10	8	6	6
256	27	12	11	9	7	7
512	29	13	11	9	7	7
1024	30	13	12	9	7	7

TABLE 5.2
$f_2(x)$

n	$\lambda_{\max}\mathcal{T}$	$\lambda_{\min}\mathcal{T}$	τ	$\lambda_{\max}\mathcal{C}$	$\lambda_{\min}\mathcal{C}$	$\mathcal{C}$	B
32	1.7612	0.9003	6	4.2123	0.7960	9	8
64	1.7694	0.8925	7	4.2465	0.8027	10	24
128	1.7736	0.8869	7	4.2648	0.8070	10	27
256	1.7758	0.8825	7	4.2742	0.8098	11	29
512	1.7771	0.8791	7	4.2791	0.8116	12	30
1024	1.7778	0.8764	7	4.2815	0.8127	12	31

be the generating function. The corresponding function h is $\sqrt{\frac{f_2(x)}{2-2\cos(x)}}$, which is an $(1,0)$-smooth function. Hence, our preconditioners ensure superlinear convergence without any smoothing technique. In Table 5.2 we give the minimum and the maximum eigenvalues of the preconditioned matrix and the iterations of the PCG method needed for both τ and circulant cases. In the last column, denoted by B, we give for comparison the iterations needed if we use the band Toeplitz preconditioner generated by the trigonometric polynomial which rises the roots.

EXAMPLE 5.3. For the generated function

$$f_3(x) = \begin{cases} x^4|x+1| & |x| \le \frac{\pi}{2} \\ (\frac{\pi}{2}+2)x^4 & x \in [-\pi,\pi] \setminus [-\frac{\pi}{2},\frac{\pi}{2}] \end{cases}$$

we have that $k = 2$. It is easily checked that the corresponding function $h(x) = \frac{\sqrt{f_3(x)}}{2-2\cos(x)}$, is an $(1,0)$-smooth function. Consequently, the τ plus band preconditioner works well without smoothing technique, while the circulant plus band one needs a further smoothing step. In Table 5.3 we give the corresponding results, as in Table 5.2 for the τ case without smoothing, while in Table 5.4 we give the results for the circulant case without and with smoothing technique. The band plus circulant preconditioner is denoted by $\hat{\mathcal{C}}$. It is easily seen that the smoothing technique is required for the circulant case to achieve superlinearity.

EXAMPLE 5.4. Finally, we consider the function

$$f_4(x) = \begin{cases} x^6|x+1| & |x| \le \frac{\pi}{2} \\ (\frac{\pi}{2}+2)x^6 & x \in [-\pi,\pi] \setminus [-\frac{\pi}{2},\frac{\pi}{2}] \end{cases}$$

as generating function. In this example we have $k = 3$ and moreover the corresponding function $h(x) = \sqrt{\frac{f_4(x)}{(2-2\cos(x))^3}}$ is also an $(1,0)$-smooth function. Thus, the smoothing

TABLE 5.3
$f_3(x)$ τ without smoothing

n	$\lambda_{\max}\tau$	$\lambda_{\min}\tau$	τ	B
16	4.977	0.854	7	8
32	5.5929	0.843	8	17
64	6.049	0.835	10	34
128	6.3624	0.8291	11	45
256	6.5669	0.8249	11	54
512	6.6955	0.8221	11	61
1024	6.7744	0.8205	12	67

TABLE 5.4
$f_3(x)$ Circulant and smoothing circulant in $[-.5, .5]$

n	$\lambda_{\max}C$	$\lambda_{\min}C$	C	$\lambda_{\max}\hat{C}$	$\lambda_{\min}\hat{C}$	$\hat{C}$	B
16	29.893	0.3498	11	28.433	0.37039	11	8
32	49.417	0.2286	13	32.369	0.34827	13	17
64	83.835	0.1386	15	34.260	0.34001	14	34
128	146.42	0.0789	18	35.552	0.3328	15	45
256	263.63	0.0428	23	36.218	0.3292	17	54
512	488.33	0.0224	26	36.556	0.3273	18	61
1024	926.19	0.0115	29	36.725	0.3265	18	67

technique is necessary for both cases to achieve superlinearity. In Table 5.5 we give the results for the τ case without and with smoothing technique, while in Table 5.6 we give the associated results for the circulant case. The meaning of stars is that the iterations required are over 100. The presented numerical results fully confirm the theory developed in the previous Sections.

In Figure 5.1, the smoothing technique is shown graphically for the function $h(x) = \frac{x^2(1+|x|)}{2-2cos(x)}$. We have to remark that h is not a differentiable function at zero.

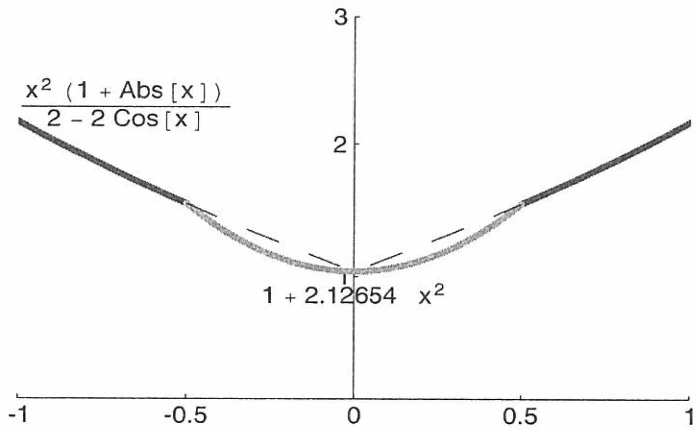

FIG. 5.1. *Smoothing of $h(x) = \frac{x^2(1+|x|)}{2-2cos(x)}$, by interpolation.*

TABLE 5.5
$f_4(x)$ τ *and* τ *with smoothing in* $[-.5, .5]$

n	$\lambda_{\max}\tau$	$\lambda_{\min}\tau$	τ	$\lambda_{\max}\hat{\tau}$	$\lambda_{\min}\hat{\tau}$	$\hat{\tau}$	B
16	24.416	0.6582	10	31.832	0.4781	8	9
32	40.853	0.4729	14	57.051	0.3312	10	20
64	68.551	0.3134	20	63.556	0.3281	11	48
128	116.29	0.1929	33	65.301	0.2965	13	*
256	201.33	0.1096	53	66.761	0.2897	14	*
512	358.56	0.0581	*	67.102	0.2813	15	*
1024	698.12	0.0246	*	67.289	0.2794	15	*

TABLE 5.6
$f_4(x)$ *Circulant and smoothing circulant in* $[-.5, .5]$

n	$\lambda_{\max}\mathcal{C}$	$\lambda_{\min}\mathcal{C}$	$\mathcal{C}$	$\lambda_{\max}\hat{\mathcal{C}}$	$\lambda_{\min}\hat{\mathcal{C}}$	$\hat{\mathcal{C}}$	B
16	371.96	0.0953	12	338.15	0.1073	11	9
32	1525.2	0.0239	17	517.36	0.0863	13	20
64	7855.2	0.0041	25	653.94	0.0756	16	48
128	48497	0.0006	43	743.32	0.0699	19	*
256	$3.3E5$	$7.5E-5$	79	792.62	0.0672	21	*
512	$2.5E6$	$1.6E-5$	*	818.57	0.0669	22	*
1024	$1.7E7$	$2.7E-6$	*	829.61	0.066710	23	*

Acknowledgement. The authors would like to thank professor S. Serra Cappizano for the fruitful discussion during his visit in Greece.

REFERENCES

[1] O. AXELSSON AND G. LINDSKOG, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523.

[2] D. BINI AND F. D. BENEDETTO, *A new preconditioner for the parallel solution of positive definite Toeplitz systems*, Proc. 2nd ACM SPAA, Crete, Greece, (1990), pp. 220–223.

[3] D. BINI AND P. FAVATI, *On a matrix algebra related to th discrete Hartley transform*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 500–507.

[4] D. BINI AND B. MEINI, *Effective methods for solving banded Toeplitz systems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 700–719.

[5] A. BOTTCHER AND B. SILBERMANN, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, 1nd ed., 2001.

[6] J. BUNCH, *Stability of methods for solving Toeplitz systems of equations*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 349–364.

[7] R. CHAN, *Circulant preconditioners for Hermitian Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.

[8] ———, *Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.

[9] R. CHAN AND W. CHING, *Toeplitz-circulant preconditioners for Toeplitz systems and their applications on queueing networks with batch arrivals.*, SIAM J. Sci. Comput., 17 (1996), pp. 762–772.

[10] R. CHAN AND P. TANG, *Fast band-Toeplitz preconditioners for Hermitian Toeplitz systems*, SIAM J. Sci. Comput., 15 (1994), pp. 164–171.

[11] R. CHAN AND M. YEUNG, *Circulant preconditioners from kernels*, SIAM J. Numer. Anal., 29 (1992).

[12] F. DI BENEDETTO, *Analysis of preconditioning techniques for ill-conditioned Toeplitz matrices*,

SIAM J. Sci. Comput., 16 (1995), pp. 682–697.

[13] ———, *Preconditioning of block Toeplitz matrices by sine transform.*, SIAM J. Sci. Comput., 18 (1997), pp. 499–515.

[14] F. Di Benedetto, G. Fiorentino, and S. Serra Capizzano, *C.G. preconditioning of Toeplitz matrices*, Comput. Math. Appl., 25 (1993), pp. 35–45.

[15] G. Fiorentino and S. Serra Capizzano, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions.*, SIAM J. Sci. Comput., 17 (1996), pp. 1068–1081.

[16] X. Q. Jin, *Hartley preconditioners for Toeplitz systems generated by positive continuous functions*, BIT, 34 (1994), pp. 367–371.

[17] D. Noutsos, S. Serra Capizzano, and P. Vassalos, *Spectral equivalence and matrix algebra preconditioners for multilevel Toeplitz systems: a negative result.*, Structured Matrices in Mathematics, Computer Science, and Engineering, Comtemporary Math., 323 (2003), pp. 313–322.

[18] ———, *Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate.*, Theoret. Computer Sci., 315 (2004), pp. 557–579.

[19] D. Noutsos and P. Vassalos, *New band Toeplitz preconditioners for ill-conditioned symmetric positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 728–743.

[20] D. Potts and G. Steidl, *Preconditioners for ill-conditioned Toeplitz matrices*, BIT, 39 (1999), pp. 513–533.

[21] S. Serra Capizzano, *Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems*, Math. Comput., 66 (1997), pp. 651–665.

[22] ———, *A Korovkin-type theory for finite Toeplitz operators via matrix algebras*, Numer. Math., 82 (1998), pp. 117–142.

[23] ———, *Superlinear PCG methods for symmetric Toeplitz systems*, Math. Comput., 68 (1999), pp. 793–803.

[24] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[25] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge, 2nd ed., 1959.

The conditioning of FD matrix sequences coming from semi-elliptic Differential Equations [1]

D. Noutsos [a] S. Serra Capizzano [b] P. Vassalos [c]

[a] *University of Ioannina, Department of Mathematics, T.K 45110 Ioannina, Greece*

[b] *Università dell'Insubria - Sede di Como, Dipartimento di Fisica e Matematica, Via Valleggio 11 - 22100 COMO, Italy.*

[c] *Athens University of Business and Economics, Department of Informatics, Patission 76, T.K 10434, Athens, Greece*

Abstract

In this paper we are concerned with the study of spectral properties of the sequence of matrices $\{A_n(a)\}$ coming from the discretization, using centered finite differences of minimal order, of elliptic (or semielliptic) differential operators $L(a,u)$ of the form

$$\begin{cases} -\frac{d}{dx}\left(a(x)\frac{d}{dx}u(x)\right) = f(x) & \text{on } \Omega = (0,1), \\ \text{Dirichlet B.C. on } \partial\Omega, \end{cases} \tag{1}$$

where the nonnegative, bounded coefficient function $a(x)$ of the differential operator may have some isolated zeros in $\overline{\Omega} = \Omega \cup \partial\Omega$. More precisely, we state and prove the explicit form of the inverse of $\{A_n(a)\}$ and some formulas concerning the relations between the orders of zeros of $a(x)$ and the asymptotic behavior of the minimal eigenvalue (condition number) of the related matrices. As a conclusion, and in connection with our theoretical findings, first we extend the analysis to higher order (semi-elliptic) differential operators, and then we present various numerical experiments, showing that similar results must hold true in 2D as well.

Key words: Finite Differences, Toeplitz matrices, Boundary Value Problems, Spectral Distribution.
1991 MSC: 65N22, 65F10, 15A12.

Email addresses: dnoutsos@uoi.gr (D. Noutsos), stefano.serrac @uninsubria.it (S. Serra Capizzano), pvassal@aueb.gr (P. Vassalos).

[1] This research was co-funded by the European Union - European Social Fund (ESF) & National Sources, in the framework of the program "Pythagoras I" of the

1 Introduction

The numerical solution of elliptic 1D and 2D Boundary Value Problems (BVPs) is a classical topic arising from a wide range of applications such as elasticity problems, nuclear and petroleum engineering etc. [31]. In these contexts, the coefficient function can be continuous or discontinuous, but its positivity guarantees the ellipticity of the continuous problem. On the other hand, for the calculation of special functions or for applications to mathematical biology and mathematical finance, the strict ellipticity is lost and indeed the function may have isolated zeros generally located at the boundary $\partial\Omega$ of the definition domain (see [16,32,1] and references therein).

Since the arising linear systems are of large size, fast and efficient resolution methods are always welcome and, for stability reasons, iterative techniques have to be preferred. However, in order to devise efficient and accurate iterative procedures, crucial spectral properties of $\{A_n(a)\}$ must be understood. In particular, we are interested in spectral localization results and especially in the asymptotic behavior of the extreme eigenvalues (which implies the knowledge of the asymptotical conditioning). Furthermore, the characterization and understanding of the subspace where the ill-conditioning occurs would be also useful, at least in a certain approximate sense. In fact the latter information represents a theoretical basis for the construction of effective preconditioners for classical and Krylov based iterative methods or in designing good prolongation/restriction operators for multigrid methods (see [15,30] and references therein). In the specific case of elliptic and semi-elliptic non-necessarily symmetric BVPs and positive definite ill-conditioned non-necessarily Hermitian Toeplitz sequences, this approach has been quite successful, both in sequential and parallel models of computation (see [11,12,21,27,24,26,4])

In this paper, we study the asymptotic conditioning with special attention to the minimal eigenvalue, since it is easy to prove that the maximal eigenvalue is bounded by a pure constant (see e.g. [10,25]). From the viewpoint of the mathematical tools, we widely use three notions of positivity: componentwise positivity (so that the Perron-Frobenius theory [31] can be invoked), positive definiteness (so that the evaluation of the spectral norm, induced by the Euclidean vector norm, is reduced to an eigenvalue analysis i.e. to study of the spectral radius), and operator positivity (so that powerful equivalence results can be applied, see [23]).

For problem (1) and for strictly positive coefficient function $a(x)$, in [10] it has been proved that the Euclidean condition number of $A_n(a)$ grows as n^2. For the degenerate case of $a(x)$ with some isolated zeros, in [21], the second author

"Operational Program for Education and Initial Vocational Training" of the 3rd Community Support Framework of the Hellenic Ministry of Education.

argues that the condition number of the arising sequence $\{A_n\}_n$ is affected by two factors (see also [25] and Subsection 2.2): the order of the differential operator which causes a growth of order n^2 (for second order problems) and the order α of the unique zero of the coefficient $a(x)$ which gives a contribution of order n^α.

The main goal of this paper is to give an explicit formula for the inverse of A_n and an asymptotical study of its condition number, for every nonnegative bounded function $a(x)$, not necessarily regular (see the beginning of Section 3 for the precise hypotheses), and with a unique zero: in particular, we show that the conditioning grows as $n^{\max\{\alpha,2\}}$, up at most to the factor $\log(n)$ only in the case where $\alpha = 2$.

The analysis is then extended to the case of several zeros and to the case of higher order operators: more specifically, when more than one zero is involved the behavior of the conditioning becomes less regular and resonance effects appear, increasing the order of the conditioning; on the other hand, for $2k$th order BVPs, $k \geq 1$, and with a unique zero of order α in the nonnegative coefficient, the quantity $n^{\max\{\alpha,2\}}$ is simply (and naturally) replaced by $n^{\max\{\alpha,2k\}}$. Finally, even though we focus our attention on 1D problems, we should stress that an interesting side-effect of this paper is to provide a theoretical framework which can be exploited to cover the less explored and highly interesting multidimensional case.

The paper is organized as follows: in Section 2 we set the problem in more detail, we set notations, and we report in a organized way some more or less known results from the relevant literature; Section 3 is devoted to give the explicit form for the inverse of the matrix A_n, a fundamental tool for our derivations, while, in Section 4, we determine the asymptotic behavior of the spectral radius of A_n^{-1}, for the second order problem in (1). Section 5 is addressed to the extension of our findings in the case of arbitrary order elliptic BVPs. Furthermore, in Section 6 we discuss the extension of our main theorem in 2D, something which is ascertained numerically in Section 7, where several 1D and 2D numerical experiments are presented and discussed. Section 8 is finally devoted to conclusions and perspectives.

2 Definition of the problem, notations, and preliminary results

Let us consider the second order BVP (1) and its approximation by using centered finite differences, of minimal bandwidth, of precision order two, and of stepsize $h = (n+1)^{-1}$ on the grid-points $x_0 = 0, x_1, x_2, \ldots, x_n, x_{n+1} = 1$. If x_t denotes $\dfrac{t}{n+1}$, $t \in [0, n+1]$, $a_t = a(x_t)$, $f_t = f(x_t)$, and u_t represents an approximation of $u(\cdot)$ at x_t, then the considered numerical scheme leads to the following set of equations

$$-a_{i-\frac{1}{2}}u_{i-1} + (a_{i-\frac{1}{2}} + a_{i+\frac{1}{2}})u_i - a_{i+\frac{1}{2}}u_{i+1} = h^2 f_i, \quad i = 1, 2, \ldots, n.$$

Then, by collecting the above formulae and by taking into account the boundary conditions, we arrive to $n \times n$ linear system whose coefficient matrix $A_n(a)$ shows the form

$$\begin{bmatrix} a_{\frac{1}{2}} + a_{\frac{3}{2}} & -a_{\frac{3}{2}} & & & & \\ -a_{\frac{3}{2}} & a_{\frac{3}{2}} + a_{\frac{5}{2}} & -a_{\frac{5}{2}} & & & \\ & -a_{\frac{5}{2}} & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & -a_{n-\frac{1}{2}} \\ & & & & -a_{n-\frac{1}{2}} & a_{n-\frac{1}{2}} + a_{n+\frac{1}{2}} \end{bmatrix} \qquad (2)$$

Let $T_n = A_n(1)$ be the Toeplitz matrix (i.e. constant along diagonals) discretizing problem (1) with $a = 1$, that is the matrix in (2) with $a = 1$. The matrices $A_n(a)$ can be expressed as

$$A_n(a) = \sum_{i=1}^{n+1} a_{i-1/2} Q_n(i), \qquad (3)$$

where the matrices $Q_n(i)$ are symmetric nonnegative definite dyads given by

$$Q_n(i) = \begin{array}{c} \begin{array}{cccccc} & & & i & & \end{array} \\ \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & -1 & \cdots & 0 \\ 0 & \cdots & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \end{array}, \qquad i = 2, \ldots, n,$$

and $Q_n(1) = e_1 e_1^T$, $Q_n(n+1) = e_n e_n^T$, with e_j, $j = 1, \ldots, n$, representing the jth column of the identity matrix.

Therefore the matrix

$$T_n = \sum_{i=1}^{n+1} Q_n(i), \qquad (4)$$

is the sum of all the dyads $Q_n(i)$ and $A_n(a)$ is a weighted sum of the same dyads according to the weights $a_{i-1/2}$, $i = 1, 2, \ldots, n+1$. Moreover each dyad has a "local structure" with respect to the canonical basis of $\mathbb{R}^{n \times n}$ so that each weight $a_{i-1/2}$ contributes in the matrix $A_n(a)$ to $E_{i-1,i-1}$, $E_{i,i-1}$, $E_{i-1,i}, E_{i,i}$ where $E_{s,t} = e_s e_t^T$. Furthermore, this notion of "locality" is geometrical as well, since vectors of the canonical basis that are close (e_s and e_t are close if $|s-t|/n = o(1)$) correspond to dyads

$$Q_n(s-1), \quad Q_n(s), \quad Q_n(t-1), \quad Q_n(t)$$

such that the related weights come from close points in the interval $[0, 1]$. Therefore we can say that the matrices $\{A_n(a)\}_n$ have a local decomposition with respect to the Toeplitz matrices $\{T_n = A_n(1)\}_n$: this locality principle is important for obtaining global distribution results for the spectra of the related matrix sequences (see e.g. [29,22]). However, again thanks to (3) and to the nonnegative definiteness of the basic dyads $Q_n(i)$, an other important aspect is that $A_n(\cdot)$, regarded as an operator from a suitable function space $\mathcal{S}$ into $\mathbb{R}^{n \times n}$, is linear and positive i.e. $A_n(\alpha a + \beta b) = \alpha A_n(a) + \beta A_n(b)$, $\alpha, \beta \in \mathbb{R}$, $a, b \in \mathcal{S}$ and $A_n(a)$ is nonnegative definite if a is nonnegative, as a function in $\mathcal{S}$ (see [23,28] for a general discussion and several results on matrix-valued linear positive operators). In Subsection 2.2 , we will use (3), (4), and this notion of operator positivity for obtaining preliminary results on the eigenvalues of $A_n(a)$.

Finally we should emphasize that the latter dyadic decompositions have a much broader interest and, in actuality, they apply to general differential operators approximated by general finite differences (see [25, Theorem 4.1] and also Lemma 2.1, Corollary 3.3, and Theorem 3.5 in the same paper) and by finite elements (see Sections 3 and 4 in [3]).

2.1 Notations

We introduce symbols that we will use throughout the paper. Let us consider two nonnegative functions $\alpha(\cdot)$ and $\beta(\cdot)$ defined over a domain D with accumulation point $\bar{x}$ (if $D = \mathbb{N}$ then $\bar{x} = \infty$, if $D = [0, 1]^d$, $d = 1, 2$, then $\bar{x}$ can be any point of D). We write

- $\alpha(\cdot) = O(\beta(\cdot))$ if and only if there exists a pure positive constant K such that $\alpha(x) \leq K\beta(x)$, for every (or for almost every) $x \in D$ (here and in the following for pure or universal constant we mean a quantity not depending on the variable $x \in D$);
- $\alpha(\cdot) = \Omega(\beta(\cdot))$ if and only if there exists a pure positive constant K such that $\alpha(x) \geq K\beta(x)$, for every (or for almost every) $x \in D$;

- $\alpha(\cdot) = o(\beta(\cdot))$ if and only if $\alpha(\cdot) = O(\beta(\cdot))$ and $\lim_{x \to \bar{x}} \alpha(x)/\beta(x) = 0$ with $\bar{x}$ given accumulation point of D which will be clear from the context;
- $\alpha(\cdot) \sim \beta(\cdot)$ if and only if $\alpha(\cdot) = O(\beta(\cdot))$ and $\beta(\cdot) = O(\alpha(\cdot))$ (or, equivalently, if and only if $\alpha(\cdot) = O(\beta(\cdot))$ and $\alpha(\cdot) = \Omega(\beta(\cdot))$;
- $\alpha(\cdot) \approx \beta(\cdot)$ if and only if $\alpha(\cdot) \sim \beta(\cdot)$ and $\lim_{x \to \bar{x}} \alpha(x)/\beta(x) = 1$ with $\bar{x}$ given accumulation point of D (the latter can be rewritten as $\alpha(x) = \beta(x)(1+o(1))$ with $1 + o(1)$ uniformly positive in D).

2.2 Preliminary results

In the following, with respect to problem (1) and hence with respect to the matrix structure in (2), we assume that the functional coefficient $a(x)$ is bounded, piece-wise continuous, nonnegative, and with a unique zero at 0 of order α i.e. $a(x) \sim x^\alpha$ on $D = [0, 1]$.

Since $A_n(\cdot)$ can be regarded as a matrix-valued linear positive operator, it is clear that it is also monotone (see [23]) that is $A_n(b) \geq A_n(a)$ if $b \geq a$ where, as usual, the ordering is the partial ordering in the space of symmetric real matrices and that of the function space $\mathcal{S}$, respectively. Therefore, since in our context the coefficient $a(x)$ is nonnegative and bounded, it follows that $A_n(a) \leq \|a\|_\infty A_n(1) = \|a\|_\infty T_n$. From the latter, from the monotonicity of the eigenvalues (i.e. $A \leq B$ implies $\lambda_j(A) \leq \lambda_j(B)$, for every pair of $n \times n$ Hermitian matrices and for every index $j = 1, 2, \ldots, n$, where $\lambda_1(X) \leq \lambda_2(X) \leq \cdots \leq \lambda_n(X)$, $X \in \{A, B\}$, see [6]) and from the known expression of the eigenvalues of T_n, we deduce that

$$\lambda_{\min}(A_n) \leq \|a\|_\infty 4 \sin^2\left(\frac{\pi}{2(n+1)}\right) \sim n^{-2}. \tag{5}$$

On the other hand, if $a(x)$ has a unique zero at zero of order α, then the minimal eigenvalue of $A_n = A_n(a)$ tends to zero at least as $n^{-\alpha}$ (see also [21, Proof of Theorem 4.1]). In fact, from (2) and from the Courant-Fisher characterization (see e.g. [6]), we have

$$\lambda_{\min}(A_n) \leq \frac{e_1^T A_n e_1}{e_1^T e_1} = a_{\frac{1}{2}} + a_{\frac{3}{2}} \sim n^{-\alpha}. \tag{6}$$

Therefore the latter bounds imply

$$\lambda_{\min}(A_n) \leq C n^{-\max\{\alpha, 2\}}, \tag{7}$$

with C universal constant independent of n (indeed depending only on the coefficient $a(x)$, see (6)). Conversely, by exploiting again the monotonicity of the operator $A_n(\cdot)$ and of the eigenvalues, and by using the dyadic decomposition in (3), it follows that

$$\lambda_{\min}(A_n) \geq \min_{1 \leq i \leq n+1} a_{i-1/2} \lambda_{\min}(T_n) \sim n^{-(\alpha+2)}. \tag{8}$$

Here we are interested in filling the gap between (7) and (8) and in fact, in Section 4, we will prove via Perron-Frobenius tools (see e.g. [31]) that the order of the true behavior of the minimal eigenvalue is described by $n^{-\max\{\alpha,2\}}$, with, at most, an additional factor $\log(n)$ in the case where $\alpha = 2$: that factor could be motivated as a kind of resonance typical of finite differences in presence of multiple zeros in the characteristic polynomial. The latter statement has also important implications concerning eigenvectors: indeed the two sources of ill-conditioning, the low frequencies coming from the constant coefficient Laplacian, and the space spanned by few canonical vectors related to the position of the zero of $a(x)$, do not interfere. There is only a superposition effect so that the size of the degenerating subspace (i.e. that related to small eigenvalues) becomes larger, but the order of ill-conditioning is not worse than that of the two factors separately. Therefore, both for designing multigrid methods or preconditioners, we can treat the two ill-conditioned spaces separately in a multi-iterative sense [20], as already done e.g. in [21] by considering a multiplicative diagonal plus Toeplitz preconditioner: more precisely, the diagonal part takes care of the ill-conditioning induced by the zero of $a(x)$ and the Toeplitz part takes care of that induced by the Laplacian (a similar idea is adapted in [26] in a multigrid setting). Finally we just mention that other results of this type can be found in [21, Theorem 4.1] and [25, Corollary 4.1 and the third item of Theorem 4.3].

3 Explicit form for the inverse of the matrix A_n

Let us consider the second order BVP (1) discretized as described in Section 2. We assume that the functional coefficient $a(x)$ is bounded, piece-wise continuous, nonnegative, and with a unique zero at 0 of order α i.e. $a(x) \sim x^\alpha$ on $D = [0, 1]$. The matrix coming from the considered approximation is $A_n = A_n(a)$ as displayed in (2). In the quoted literature, we find several contributions discussing the form of the inverse of a tridiagonal matrix, or more generally, on the one of a band matrix. First in 1960, F. Gantmacher and M. Krein [13] proved that the inverse of a symmetric nonsingular tridiagonal matrix is a Green matrix which is defined by the Hadamard product of a weak type D and a flipped weak type D matrices as follows:

$$C = U \circ V = \begin{bmatrix} u_1 & u_1 & \cdots & u_1 \\ u_1 & u_2 & \cdots & u_2 \\ \vdots & \vdots & \ddots & \vdots \\ u_1 & u_2 & \cdots & u_n \end{bmatrix} \circ \begin{bmatrix} v_1 & v_2 & \cdots & v_n \\ v_2 & v_2 & \cdots & v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n & v_n & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_1v_2 & u_2v_2 & \cdots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_1v_n & u_2v_n & \cdots & u_nv_n \end{bmatrix}. \quad (9)$$

Conversely, the same authors have proven that the inverse of a Green matrix is a symmetric tridiagonal matrix. In 1970, M. Capovani [8] stated and derived relations which give the entries of the inverse of a tridiagonal matrix in terms of its entries and its subdeterminants. In the same paper he gave the form of the inverse of some particular cases of tridiagonal and block tridiagonal matrices. One year later the same author [9], extended the result of F. Gantmacher and M. Krein [13] for nonsymmetric matrices. R. Bevilacqua and M. Capovani [5] in 1976, gave structural properties to determine the coefficients of the inverse of a (block) as a function of its (blocks) entries. In 1979, W. Barrett [2] proved that a matrix R with $R_{22}, \ldots, R_{n-1,n-1} \neq 0$ has the triangle property if and only if its inverse is a tridiagonal matrix: more in detail, a matrix R has this useful property if $R_{ij} = \frac{R_{ik}R_{kj}}{R_{kk}}$ for all $i < k < j$ and all $i > k > j$. In 1987, P. Rózsa [19], using properties of Green's matrices and of semi-separable matrices, proposed an algorithm to determine the elements of the inverse of a band matrix by solving some difference equations. Later in 1998, J. McDonald, R. Nabben, M. Neumann, H. Schneider and M. Tsatsomeros [17] generalized the result of F. Gantmacher and M. Krein [13] for nonsymmetric tridiagonal Z-matrices and they proved properties for the inverse of a tridiagonal M-matrix. They gave also properties for the inverse of such matrices in terms of special structured matrices called cyclopses (see again [17] for a formal definition). More recently, i.e. in 1999, R. Nabben [18] proved properties for the inverse of tridiagonal M, positive definite and diagonally dominant matrices.

The matrix A_n in (2) has most of the above "good" properties: it is an irreducible nonsingular tridiagonal Z-matrix, an M-matrix, and also a symmetric positive definite matrix. Hence, we can combine the above results for characterizing its inverse. However, the matrix A_n has an additional property that all row sums are zeros except the first and the last one. Taking into account Corollary 3.6 of [17] or Corollary 2.6 of [18], concerning properties of the inverse of an M-matrix, and Corollary 2.7 of [18], concerning on properties of the inverse of a positive definite matrix, we obtain that the numbers $u_i, v_i,\ i = 1, 2, \ldots, n$, appearing in the Hadamard product (9), can be chosen to be positive and such that

$$0 < \frac{u_1}{v_1} < \frac{u_2}{v_2} < \cdots < \frac{u_n}{v_n}. \quad (10)$$

In the sequel we will find an explicit form for the matrix A_n^{-1} by using the forms of A_n and C in (2) and (9), respectively and inequalities (10). We take the product $A_n C$ which should be the identity matrix I.

For $k < j$, the inner product of the kth row of A_n with the jth column of C gives

$$\begin{aligned} 0 = (A_n C)_{kj} &= v_j \left(-a_{k-\frac{1}{2}} u_{k-1} + (a_{k-\frac{1}{2}} + a_{k+\frac{1}{2}}) u_k - a_{k+\frac{1}{2}} u_{k+1} \right) \\ &= v_j \left(a_{k-\frac{1}{2}} (u_k - u_{k-1}) - a_{k+\frac{1}{2}} (u_{k+1} - u_k) \right). \end{aligned}$$

We observe that this equality holds true if we chose, up to a constant factor,

$$u_k - u_{k-1} = \frac{1}{a_{k-\frac{1}{2}}}, \quad k = 2, 3, \ldots, n.$$

One solution of this difference equation, up to a constant factor, is

$$u_k = \sum_{i=1}^{k} \frac{1}{a_{i-\frac{1}{2}}}, \quad k = 1, 2, 3, \ldots, n.$$

For $k = 1$ we have

$$0 = (A_n C)_{1j} = v_j \left((a_{\frac{1}{2}} + a_{\frac{3}{2}}) \frac{1}{a_{\frac{1}{2}}} - a_{\frac{3}{2}} \left(\frac{1}{a_{\frac{1}{2}}} + \frac{1}{a_{\frac{3}{2}}} \right) \right)$$

which holds true.

For $k > j$, the associated inner products give

$$\begin{aligned} 0 = (A_n C)_{kj} &= u_j \left(-a_{k-\frac{1}{2}} v_{k-1} + (a_{k-\frac{1}{2}} + a_{k+\frac{1}{2}}) v_k - a_{k+\frac{1}{2}} v_{k+1} \right) \\ &= u_j \left(-a_{k-\frac{1}{2}} (v_{k-1} - v_k) + a_{k+\frac{1}{2}} (v_k - v_{k+1}) \right). \end{aligned}$$

We observe also here that we can chose, up to a constant factor,

$$v_{k-1} - v_k = \frac{1}{a_{k-\frac{1}{2}}}, \quad k = 2, 3, \ldots, n.$$

One solution of this difference equation, up to a constant factor, is

$$v_k = \sum_{i=k}^{n} \frac{1}{a_{i+\frac{1}{2}}}, \quad k = 1, 2, 3, \ldots, n.$$

For $k = n$ we have

$$0 = (A_n C)_{nj} = u_j \left(-a_{n-\frac{1}{2}} \left(\frac{1}{a_{n-\frac{1}{2}}} + \frac{1}{a_{n+\frac{1}{2}}} \right) + (a_{n-\frac{1}{2}} + a_{n+\frac{1}{2}}) \frac{1}{a_{n+\frac{1}{2}}} \right)$$

which holds true. We define by s_k and by s the sums $\sum_{i=k}^{n} \frac{1}{a_{i+\frac{1}{2}}}$ and $\sum_{i=0}^{n} \frac{1}{a_{i+\frac{1}{2}}}$, respectively. It is obvious that with the above choices, up to a constant factor, we have $v_k = s_k$, $u_k = s - s_k$, $k = 1, 2, \ldots, n$. We observe also that the sequence v_k strictly decreases while u_k strictly increases, so the inequalities (10) are satisfied.

It remains to check the inner products for $k = j$.

$$\begin{aligned}
(A_nC)_{kk} &= -a_{k-\frac{1}{2}}u_{k-1}v_k + (a_{k-\frac{1}{2}} + a_{k+\frac{1}{2}})u_kv_k - a_{k+\frac{1}{2}}u_kv_{k+1} \\
&= -a_{k-\frac{1}{2}}(s - s_{k-1})s_k + (a_{k-\frac{1}{2}} + a_{k+\frac{1}{2}})(s - s_k)s_k - a_{k+\frac{1}{2}}(s - s_k)s_{k+1} \\
&= a_{k-\frac{1}{2}}(s_{k-1} - s_k)s_k + a_{k+\frac{1}{2}}(s_k - s_{k+1})(s - s_k) \\
&= s_k + (s - s_k) = s, \quad k = 2, 3, \ldots, n-1,
\end{aligned}$$

$$\begin{aligned}
(A_nC)_{11} &= (a_{\frac{1}{2}} + a_{\frac{3}{2}})u_1v_1 - a_{\frac{3}{2}}u_1v_2 = a_{\frac{1}{2}}(s - s_1)s_1 + a_{\frac{3}{2}}(s_1 - s_2)(s - s_1) \\
&= s_1 + (s - s_1) = s,
\end{aligned}$$

$$\begin{aligned}
(A_nC)_{nn} &= -a_{n-\frac{1}{2}}u_{n-1}v_n + (a_{n-\frac{1}{2}} + a_{n+\frac{1}{2}})u_nv_n \\
&= a_{n-\frac{1}{2}}(s_{n-1} - s_n)s_n + a_{n+\frac{1}{2}}(s - s_n)s_n = s_n + (s - s_n) = s.
\end{aligned}$$

As a consequence $A_nC = sI$. To eliminate s we have to chose the constant factors of the matrices U and V, in such a way that the relative product equals $\frac{1}{s}$. Then, the inverse of A_n is obtained by dividing C by s which gives us the explicit form:

$$A_n^{-1} = \begin{bmatrix}
\frac{s_1(s-s_1)}{s} & \frac{s_2(s-s_1)}{s} & \frac{s_3(s-s_1)}{s} & \ldots & \frac{s_n(s-s_1)}{s} \\
\frac{s_2(s-s_1)}{s} & \frac{s_2(s-s_2)}{s} & \frac{s_3(s-s_2)}{s} & \ldots & \frac{s_n(s-s_2)}{s} \\
\frac{s_3(s-s_1)}{s} & \frac{s_3(s-s_2)}{s} & \frac{s_3(s-s_3)}{s} & \ldots & \frac{s_n(s-s_3)}{s} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{s_n(s-s_1)}{s} & \frac{s_n(s-s_2)}{s} & \frac{s_n(s-s_3)}{s} & \ldots & \frac{s_n(s-s_n)}{s}
\end{bmatrix}. \tag{11}$$

It follows another proof to obtain the explicit form of A_n^{-1} independent of the previous approach. It is not related to the theory appearing in the referred literature, but only depends on the form (2) of A_n and on a tricky use of the Sherman-Morrison formula.

We consider the matrix

$$\tilde{A}_n = \text{tridiag}[-1 \;\; 1 \;\; 0]\text{diag}[a_{\frac{3}{2}} \;\; a_{\frac{5}{2}} \;\; \cdots \;\; a_{n+\frac{1}{2}}]\text{tridiag}[0 \;\; 1 \;\; -1]$$

i.e.

$$\tilde{A}_n = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ 0 & -1 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \begin{bmatrix} a_{\frac{3}{2}} & & & & \\ & a_{\frac{5}{2}} & & & \\ & & a_{\frac{7}{2}} & & \\ & & & \ddots & \\ & & & & a_{n+\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ & 1 & -1 & \cdots & 0 \\ & & 1 & \ddots & \vdots \\ & & & \ddots & -1 \\ & & & & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{\frac{3}{2}} & -a_{\frac{3}{2}} & & & \\ -a_{\frac{3}{2}} & a_{\frac{3}{2}} + a_{\frac{5}{2}} & -a_{\frac{5}{2}} & & \\ & -a_{\frac{5}{2}} & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -a_{n-\frac{1}{2}} \\ & & & -a_{n-\frac{1}{2}} & a_{n-\frac{1}{2}} + a_{n+\frac{1}{2}} \end{bmatrix}.$$

We observe that the matrices $\tilde{A}_n$ and A_n differ only in position $(1,1)$, where the term $a_{\frac{1}{2}}$ does not appear in the matrix $\tilde{A}_n$. By (3), it follows that

$$A_n = \tilde{A}_n + a_{\frac{1}{2}} e_1 e_1^T,$$

e_1 being the first column of the identity matrix. Our aim is to find an explicit form of the inverse of A_n and hence, from the above relation, we write

$$A_n^{-1} = (\tilde{A}_n + a_{\frac{1}{2}} e_1 e_1^T)^{-1} = (I + a_{\frac{1}{2}} \tilde{A}_n^{-1} e_1 e_1^T)^{-1} \tilde{A}_n^{-1}. \tag{12}$$

As a consequence, in order to determine a formula for the inverse of A_n, it suffices to compute the inverses of the two factors appearing in (12). For that purpose, we start our analysis by studying the inverse of $\tilde{A}_n$ (in this respect, we should acknowledge that the expression of $\tilde{A}_n^{-1}$ can be found, for the specific case of $a = 1$, in [7, Chapter 4, Exercise 8, p. 108]).

From the above factorization of $\tilde{A}_n$, we find

$$\tilde{A}_n^{-1} = (\text{tridiag}[0 \;\; 1 \;\; -1])^{-1} (\text{diag}[a_{\frac{3}{2}} \;\; a_{\frac{5}{2}} \;\; \cdots \;\; a_{n+\frac{1}{2}}])^{-1} (\text{tridiag}[-1 \;\; 1 \;\; 0])^{-1}$$

that is

$$\tilde{A}_n^{-1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ & 1 & 1 & \cdots & 1 \\ & & 1 & \ddots & \vdots \\ & & & \ddots & 1 \\ & & & & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{a_{\frac{3}{2}}} & & & & \\ & \frac{1}{a_{\frac{5}{2}}} & & & \\ & & \frac{1}{a_{\frac{7}{2}}} & & \\ & & & \ddots & \\ & & & & \frac{1}{a_{n+\frac{1}{2}}} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{a_{\frac{3}{2}}} & \frac{1}{a_{\frac{5}{2}}} & \frac{1}{a_{\frac{7}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \\ & \frac{1}{a_{\frac{5}{2}}} & \frac{1}{a_{\frac{7}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \\ & & \frac{1}{a_{\frac{7}{2}}} & \ddots & \vdots \\ & & & \ddots & \frac{1}{a_{n+\frac{1}{2}}} \\ & & & & \frac{1}{a_{n+\frac{1}{2}}} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=2}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=3}^n \frac{1}{a_{i+\frac{1}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \\ \sum_{i=2}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=2}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=3}^n \frac{1}{a_{i+\frac{1}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \\ \sum_{i=3}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=3}^n \frac{1}{a_{i+\frac{1}{2}}} & \sum_{i=3}^n \frac{1}{a_{i+\frac{1}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{n+\frac{1}{2}}} & \frac{1}{a_{n+\frac{1}{2}}} & \frac{1}{a_{n+\frac{1}{2}}} & \cdots & \frac{1}{a_{n+\frac{1}{2}}} \end{bmatrix}$$

$$= \begin{bmatrix} s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_2 & s_3 & \cdots & s_n \\ s_3 & s_3 & s_3 & \cdots & s_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_n & s_n & s_n & \cdots & s_n \end{bmatrix}.$$

The matrix $a_{\frac{1}{2}}\tilde{A}_n^{-1}e_1e_1^T$ has nonzero elements only in the first column whose expression is given by $a_{\frac{1}{2}}s_k$, $k = 1, 2, \ldots, n$. Therefore the analysis of A_n^{-1} is equivalently transformed into the inversion of the matrix

$$I + a_{\frac{1}{2}} \tilde{A}_n^{-1} e_1 e_1^T = \begin{bmatrix} a_{\frac{1}{2}} s & & & & \\ a_{\frac{1}{2}} s_2 & 1 & & & \\ a_{\frac{1}{2}} s_3 & & 1 & & \\ \vdots & & & \ddots & \\ a_{\frac{1}{2}} s_n & & & & 1 \end{bmatrix},$$

where the entry in position $(1,1)$ has been obtained by $1 + a_{\frac{1}{2}} s_1 = a_{\frac{1}{2}}(\frac{1}{a_{\frac{1}{2}}} + s_1) = a_{\frac{1}{2}} s$. It is well-known that the inverse of the above matrix maintains the same structure (since it is a slight variation of an elementary Gauss matrix, see [14]), and it is easily obtained as

$$(I + a_{\frac{1}{2}} \tilde{A}_n^{-1} e_1 e_1^T)^{-1} = \begin{bmatrix} \frac{1}{a_{\frac{1}{2}}} \frac{1}{s} & & & & \\ -\frac{s_2}{s} & 1 & & & \\ -\frac{s_3}{s} & & 1 & & \\ \vdots & & & \ddots & \\ -\frac{s_n}{s} & & & & 1 \end{bmatrix}.$$

In conclusion

$$A_n^{-1} = \begin{bmatrix} \frac{s-s_1}{s} & & & & \\ -\frac{s_2}{s} & 1 & & & \\ -\frac{s_3}{s} & & 1 & & \\ \vdots & & & \ddots & \\ -\frac{s_n}{s} & & & & 1 \end{bmatrix} \begin{bmatrix} s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_2 & s_3 & \cdots & s_n \\ s_3 & s_3 & s_3 & \cdots & s_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_n & s_n & s_n & \cdots & s_n \end{bmatrix},$$

where we have replaced $\frac{1}{a_{\frac{1}{2}}}$ by $s - s_1$. The above matrix product leads to the following results: the entries of the first row of A_n^{-1} are

$$(A_n^{-1})_{1j} = \frac{s_j(s - s_1)}{s}, \quad j = 1, 2, \ldots, n.$$

The entries $(A_n^{-1})_{ij}, \quad i = 2, 3, \ldots, n$ for $i \geq j$, are given by

$$(A_n^{-1})_{ij} = -\frac{s_i s_j}{s} + s_i = \frac{s_i(s - s_j)}{s},$$

while, for $i < j$, we find

$$(A_n^{-1})_{ij} = -\frac{s_i s_j}{s} + s_j = \frac{s_j(s - s_i)}{s}.$$

Consequently a compact formula for the inverse of the matrix A_n is given by the explicit form (11)

4 The spectral radius of A_n^{-1}

For determining the asymptotic behavior of the condition number of the matrix A_n, we have to estimate the smallest eigenvalue since its maximal eigenvalue is bounded by $4\|a\|_\infty$, since $A_n = A_n(a) \leq \|a\|_\infty T_n$ by operator positivity of $A_n(\cdot)$ (see Subsection 2.2) and since $\lambda_{\max}(T_n) < 4$ by Gershgorin's theorem (see e.g. [6,31]). Instead of this, we study the spectral radius of the inverse of A_n. The matrix A_n^{-1} is a symmetric positive definite matrix with positive elements. Thus we make use of the Perron Frobenius theory (see e.g. [31]) for positive (nonnegative) matrices. Our analysis is obtained via a series of preliminary results.

Lemma 4.1. *Let $\{A_n\}_n$, $A_n \in \mathbb{R}^{n \times n}$, be a sequence of symmetric positive definite, irreducible, and nonnegative matrices. If there exists a number $g(n)$ of rows such that their row sums are greater than or equal to $f(n)$, then the order of the spectral radius $\rho(A_n)$ is greater than or equal to $\frac{g(n)f(n)}{n}$, so that*

$$\rho(A_n) = \Omega\left(\frac{g(n)f(n)}{n}\right).$$

Proof. Without loss of generality, we suppose that the row sums are in decreasing order (otherwise this can be obtained by a proper permutation similarity transformation). By using the Courant-Fisher characterization [6], we find

$$\begin{aligned}
\rho(A_n) &= \sup_{x \in \mathbb{R}^n, \|x\|=1} x^T A_n x \geq \frac{1}{n} e^T(n) A_n e(n) \\
&= \frac{1}{n} e^T(n) \begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \vdots \\ \tilde{S}_n \end{pmatrix} \geq \frac{1}{n} \sum_{i=1}^{g(n)} \tilde{S}_i \geq \frac{1}{n} g(n) f(n),
\end{aligned}$$

where the normalized vector $\frac{1}{\sqrt{n}} e(n)$ has replaced x, with $e(n)$ being the vector of all ones, and where we have denoted by $\tilde{S}_i$ the ith row sum of the matrix A_n. □

We introduce the following definition.

Definition 4.1. *A symmetric, positive definite, irreducible, and nonnegative matrix $A \in \mathbb{R}^{n\times n}$, given in decreasing order of row sums, is dominated by the first $g(n) \times g(n)$ block if $\tilde{S}_i \sim \tilde{S}_{B_i}$, where $\tilde{S}_{B_i} = \sum_{j=1}^{g(n)} a_{ij}$ and the symbol $\sim$ is that defined in Subsection 2.1.*

Lemma 4.2. *Let $\{A_n\}_n$, $A_n \in \mathbb{R}^{n\times n}$, be a sequence of symmetric positive definite, nonnegative matrices, which are dominated by their first $g(n) \times g(n)$ block. If $f(n)$ is the smallest row sum of the first $g(n)$ rows, then the order of the spectral radius $\rho(A_n)$ is greater than or equal to $f(n)$, so that*

$$\rho(A_n) = \Omega(f(n)).$$

Proof. The proof follows the same procedure as of in Lemma 4.1. For that we take the normalized vector $\frac{1}{\sqrt{g(n)}}e(g(n))$, with $e(g(n))$ being the vector of ones in the first $g(n)$ entries and zeros otherwise. Thus

$$\begin{aligned}
\rho(A_n) &= \sup_{x\in\mathbb{R}^n, \|x\|=1} x^T A_n x \\
&\geq \tfrac{1}{g(n)} e^T(g(n)) A_n e(g(n)) \\
&\simeq \tfrac{1}{g(n)} e^T(g(n)) \begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \vdots \\ \tilde{S}_{g(n)} \\ \vdots \end{pmatrix} \\
&= \tfrac{1}{g(n)} \sum_{i=1}^{g(n)} \tilde{S}_i \geq \tfrac{g(n) f(n)}{g(n)} = f(n)
\end{aligned}$$

and the proof is complete. □

Lemma 4.3. *Let $\{A_n\}_n$, $A_n \in \mathbb{R}^{n\times n}$, be a sequence of symmetric positive definite, nonnegative matrices, which are dominated by their first $g(n) \times g(n)$ block. If all the first $g(n)$ rows are of the same order of $f(n)$, then the spectral radius $\rho(A_n)$ is exactly of order $f(n)$.*

Proof. From Lemma 4.2 we deduce $\rho(A_n) = \Omega(f(n))$. On the other hand, from the Perron-Frobenius theory, we obtain that $\rho(A_n) \leq \max_i \tilde{S}_i$. As a consequence, $\rho(A_n) = O(f(n))$ and the proof is complete. □

Now, we are ready to state and prove the main theorem of this section concerning the relation between the order of the zero of the coefficient function $a(x)$ and the condition number of the matrix A_n.

Theorem 4.4. *Let $\{A_n\}_n$, $A_n \in \mathbb{R}^{n\times n}$, be the sequence of matrices derived from the discretization of the Semielliptic Differential Equation (1) with the* *bounded coefficient function $a(x)$ having a unique root at 0 of order α i.e. $a(x) \sim x^\alpha$ on $D = [0,1]$. Then, for the spectral condition number $\kappa(A_n)$ of the matrix A_n which coincides in order with the spectral radius of A_n^{-1}, we find*

$$\kappa(A_n) \sim \rho(A_n^{-1}) \sim \begin{cases} n^2, & 0 \le \alpha < 2, \\ O(n^2 \log(n)) \cap \Omega(n^2), & \alpha = 2, \\ n^\alpha, & \alpha > 2. \end{cases} \tag{13}$$

Proof. The part $\kappa(A_n) \sim \rho(A_n^{-1})$ simply follows from the relations $\|A_n^{-1}\| = \rho(A_n^{-1})$, $\|A_n\| = \rho(A_n) \le 4\|a\|_\infty$, and $\lim_{n\to\infty} \rho(A_n) = 4\|a\|_\infty$, where the positive definiteness of A_n and the distribution results in [29] come into the play.

The fact that $a(x) \sim x^\alpha$ means that there exist positive constants c and C far from zero and infinity such that, uniformly with respect to $x \in [0,1]$, we have

$$cx^\alpha \le a(x) \le Cx^\alpha.$$

From the positivity of the operator $A_n(\cdot)$ we obtain

$$cA_n(x^\alpha) \le A_n(a(x)) \le CA_n(x^\alpha)$$

where the meaning of the inequalities is in the sense of the partial ordering in the real space of Hermitian (real symmetric) matrices. The latter implies

$$c\lambda_i(A_n(x^\alpha)) \le \lambda_i(A_n(a(x))) \le C\lambda_i(A_n(x^\alpha)), \quad i = 1, 2, \ldots, n,$$

and, in particular, this holds also for the minimal eigenvalue, which means that the minimal eigenvalue of $A_n(x^\alpha)$ and the minimal eigenvalue of $A_n(a(x))$ coincide in order of magnitude. As a consequence, it is enough to reduce our study to the matrix $A_n(x^\alpha)$, instead of $A_n(a(x))$.

For the remaining part, since the matrix A_n^{-1} is a symmetric positive definite matrix with positive elements, we will prove our assertion, by estimating the row sums of the matrix A_n^{-1} with functional coefficient x^α, given in its explicit form (11), and by using the previous lemmas. For this we study the following cases:

Case 1: $\alpha = 0$.

The result related to this case is well-known [10], since the matrix T_n coincides exactly with tridiag$[-1 \quad 2 \quad -1]$, i.e. the Laplace matrix with eigenvalues $4\sin^2\left(\frac{j\pi}{2(n+1)}\right)$, $j = 1, \ldots, n$. Hence

$$\kappa(A_n) \sim \rho(A_n^{-1}) \sim n^2. \tag{14}$$

We remark here that this result could be obtained also by following the reasoning we will use in the subsequent cases.

Case 2: $0 < \alpha < 1$.

We estimate the kth row sum $\tilde{S}_k$ of A_n^{-1}

$$\tilde{S}_k = \frac{S_k}{S}\sum_{i=1}^{k}(S - S_i) + \frac{S - S_k}{S}\sum_{i=k+1}^{n} S_i. \tag{15}$$

First we consider the quantity S_i:

$$\begin{aligned} S_i &= \sum_{j=i}^{n} \frac{1}{a_{j+\frac{1}{2}}} = \sum_{j=i}^{n} \frac{1}{\left(\frac{j+\frac{1}{2}}{n+1}\right)^{\alpha}} = \sum_{j=i}^{n} \left(\frac{2j+1}{2(n+1)}\right)^{-\alpha} \\ &= (n+1)\sum_{j=i}^{n} \left(\frac{2j+1}{2(n+1)}\right)^{-\alpha} \frac{1}{n+1}. \end{aligned}$$

Taking into account that we have uniformly discretized the interval $[0,1]$ in $n+1$ subintervals, we get that the value $\left(\frac{2j+1}{2(n+1)}\right)^{-\alpha} \frac{1}{n+1}$ is the (Lebesgue) measure of the rectangle with x-edges $[\frac{j}{n+1}, \frac{j+1}{n+1}]$ and y-edges $\left[0, \left(\frac{2j+1}{2(n+1)}\right)^{-\alpha}\right]$. Therefore, the above sum is approximated by an integral as follows:

$$S_i \approx (n+1)\int_{\frac{i}{n+1}}^{1} x^{-\alpha} dx = \frac{n+1}{1-\alpha}\left[x^{1-\alpha}\right]_{\frac{i}{n+1}}^{1} = \frac{n+1}{1-\alpha}\left[1 - \left(\frac{i}{n+1}\right)^{1-\alpha}\right]. \tag{16}$$

It is easily checked that the error of the above approximation is less than S_i in order of magnitude. If we substitute $i = 0$ in relation (16), then we estimate the quantity S as

$$S \approx (n+1)\int_0^1 x^{-\alpha} dx = \frac{n+1}{1-\alpha}\left[x^{1-\alpha}\right]_0^1 = \frac{n+1}{1-\alpha}. \tag{17}$$

From (16) and (17) we find

$$S - S_i \approx \frac{n+1}{1-\alpha} - \frac{n+1}{1-\alpha}\left[1 - \left(\frac{i}{n+1}\right)^{1-\alpha}\right] = \frac{(n+1)^{\alpha}\ i^{1-\alpha}}{1-\alpha}. \tag{18}$$

By taking the sum of the coefficients in (16), we deduce

$$\begin{aligned} \textstyle\sum_{i=k+1}^{n} S_i &\approx \textstyle\sum_{i=k+1}^{n} \frac{n+1}{1-\alpha} - \sum_{i=k+1}^{n} \frac{n+1}{1-\alpha}\left(\frac{i}{n+1}\right)^{1-\alpha} \\ &\approx \frac{(n+1)(n-k)}{1-\alpha} - \frac{(n+1)^2}{1-\alpha}\int_{\frac{k+1}{n+1}}^{1} x^{1-\alpha} dx \\ &= \frac{(n+1)(n-k)}{1-\alpha} - \frac{(n+1)^2}{(1-\alpha)(2-\alpha)}\left[1 - \left(\frac{k+1}{n+1}\right)^{2-\alpha}\right] \\ &= \frac{(n+1)^2}{2-\alpha} + \frac{(n+1)^{\alpha}(k+1)^{2-\alpha}}{(1-\alpha)(2-\alpha)} - \frac{(n+1)(k+1)}{(1-\alpha)}, \end{aligned} \tag{19}$$

where we used $\sum_{i=k+1}^{n}\left(\frac{i}{n+1}\right)^{1-\alpha}\frac{1}{n+1}\approx\int_{\frac{k+1}{n+1}}^{1}x^{1-\alpha}dx$. Similarly, by taking the sum of the coefficients in (18), we find

$$\begin{aligned}\textstyle\sum_{i=1}^{k}(S-S_i) &\approx \tfrac{(n+1)^\alpha}{1-\alpha}\textstyle\sum_{i=1}^{k}i^{1-\alpha}\approx\tfrac{(n+1)^2}{1-\alpha}\int_0^{\frac{k}{n+1}}x^{1-\alpha}dx\\ &=\tfrac{(n+1)^2}{(1-\alpha)(2-\alpha)}\left(\tfrac{k}{n+1}\right)^{2-\alpha}=\tfrac{(n+1)^\alpha k^{2-\alpha}}{(1-\alpha)(2-\alpha)},\end{aligned}\tag{20}$$

where $\sum_{i=1}^{k}\left(\frac{i}{n+1}\right)^{1-\alpha}\frac{1}{n+1}\approx\int_0^{\frac{k}{n+1}}x^{1-\alpha}dx$. By replacing the explicit formulae (16), (17), (18), (19), and (20) in relation (15), we arrive to estimate $\tilde{S}_k$ that is

$$\begin{aligned}\tilde{S}_k &= \left[1-\left(\tfrac{k}{n+1}\right)^{1-\alpha}\right]\tfrac{(n+1)^\alpha k^{2-\alpha}}{(1-\alpha)(2-\alpha)}\\ &+\left(\tfrac{k}{n+1}\right)^{1-\alpha}\left[\tfrac{(n+1)^2}{2-\alpha}+\tfrac{(n+1)^\alpha(k+1)^{2-\alpha}}{(1-\alpha)(2-\alpha)}-\tfrac{(n+1)(k+1)}{(1-\alpha)}\right].\end{aligned}\tag{21}$$

We plainly observe that $\tilde{S}_k$ does not exceed, in order of magnitude, the value $\max\{(n+1)^\alpha k^{2-\alpha},\ (n+1)^{1+\alpha}k^{1-\alpha},\ (n+1)^{2\alpha-1}k^{3-2\alpha}\}$. In any case this maximum is of order of n^2. On the other hand, by studying (21) for $\frac{n}{4}+1\le k\le\frac{3n}{4}$, we obtain that

$$\tilde{S}_k\sim n^2,\qquad \frac{n}{4}+1\le k\le\frac{3n}{4}.\tag{22}$$

We consider now the matrix $B_{\frac{n}{2}}$, the $\frac{n}{2}\times\frac{n}{2}$ block of A_n^{-1} formed by deleting the first and the last $\frac{n}{4}$ rows and columns. We denote by $\tilde{S}_{B_k}$ the kth row sum of the matrix $B_{\frac{n}{2}}$, where the index k ranges from $\frac{n}{4}+1$ to $\frac{3n}{4}$. Taking into account (15), we infer

$$\tilde{S}_{B_k}=\frac{S_k}{S}\sum_{i=\frac{n}{4}+1}^{k}(S-S_i)+\frac{S-S_k}{S}\sum_{i=k+1}^{\frac{3n}{4}}S_i.\tag{23}$$

By making analogous calculations, as in the estimation of $\tilde{S}_k$, we find

$$\begin{aligned}\tilde{S}_{B_k} &= \left[1-\left(\tfrac{k}{n+1}\right)^{1-\alpha}\right]\tfrac{(n+1)^\alpha\left(k^{2-\alpha}-(\frac{n}{4})^{2-\alpha}\right)}{(1-\alpha)(2-\alpha)}\\ &+\left(\tfrac{k}{n+1}\right)^{1-\alpha}\left[\tfrac{(n+1)(\frac{3n}{4}-k)}{1-\alpha}+\tfrac{(n+1)^\alpha\left((k+1)^{2-\alpha}-(\frac{3n}{4}+1)^{2-\alpha}\right)}{(1-\alpha)(2-\alpha)}\right].\end{aligned}\tag{24}$$

It is easy to understand that (24) implies

$$\tilde{S}_{B_k}\sim n^2,\qquad \frac{n}{4}+1\le k\le\frac{3n}{4}.\tag{25}$$

We apply now a permutation transformation to the matrix A_n^{-1} in such a way that its block $B_{\frac{n}{2}}$ will appear in the first $\frac{n}{2}\times\frac{n}{2}$ rows and columns. Then, the permuted matrix is dominated to the first $\frac{n}{2}\times\frac{n}{2}$ block, with all the first $\frac{n}{2}$ row sums being of order n^2. In this case Lemma 4.3 is applied to obtain relation (14) that is

$$\kappa(A_n)\sim\rho(A_n^{-1})\sim n^2.$$

Case 3: $\alpha = 1$.

We follow the same steps as in the previous case:

$$\begin{aligned}
S_i =& \sum_{j=i}^{n} \frac{1}{a_{j+\frac{1}{2}}} = \sum_{j=i}^{n} \frac{2(n+1)}{2j+1} \\
&\approx (n+1) \int_{\frac{i}{n+1}}^{1} x^{-1} dx = (n+1) \log\left(\frac{n+1}{i}\right); & (26) \\
S =& 2(n+1) + S_1 \approx (n+1)\left(2 + \log(n+1)\right); & (27) \\
S - S_i \approx& (n+1)\left(2 + \log(n+1)\right) - (n+1)\log\left(\frac{n+1}{i}\right) \\
=& (n+1)\left(2 + \log(i)\right). & (28)
\end{aligned}$$

Now by substituting (26), (27), (28) in relation (15), we infer the following estimate for $\tilde{S}_k$:

$$\begin{aligned}
\tilde{S}_k \approx& \tfrac{(n+1)\log\left(\frac{n+1}{k}\right)}{(n+1)(2+\log(n+1))} \textstyle\sum_{i=1}^{k}(n+1)\left(2+\log(i)\right) \\
+& \tfrac{(n+1)(2+\log(k))}{(n+1)(2+\log(n+1))} \textstyle\sum_{i=k+1}^{n}(n+1)\log\left(\frac{n+1}{i}\right) \\
=& \tfrac{(n+1)\log\left(\frac{n+1}{k}\right)}{2+\log(n+1)} \left(2k + \textstyle\sum_{i=1}^{k}\log(i)\right) \\
+& \tfrac{(n+1)(2+\log(k))}{2+\log(n+1)} \left((n-k)\log(n+1) - \textstyle\sum_{i=k+1}^{n}\log(i)\right).
\end{aligned} \qquad (29)$$

On the other hand

$$\sum_{i=1}^{k} \log(i) \approx \int_{1}^{k} \log(x)dx = k\log(k) - k + 1$$

and

$$\sum_{i=k+1}^{n} \log(i) \approx \int_{k+1}^{n} \log(x)dx = n\log(n) - n - k\log(k) + k + 1.$$

By replacing the latter terms in (29), we obtain

$$\begin{aligned}
\tilde{S}_k \approx& \tfrac{(n+1)\log\left(\frac{n+1}{k}\right)}{2+\log(n+1)}(k + k\log(k) + 1) \\
+& \tfrac{(n+1)(2+\log(k))}{2+\log(n+1)} \left(n\log\left(\tfrac{n+1}{n}\right) - k\log\left(\tfrac{n+1}{k}\right) + n - k - 1\right),
\end{aligned} \qquad (30)$$

and hence the quantity $\tilde{S}_k$ does not exceed n^2 in order of magnitude. Furthermore, by analyzing (30) for $\frac{n}{4} + 1 \leq k \leq \frac{3n}{4}$ we obtain relation (22) that is

$$\tilde{S}_k \sim n^2, \qquad \frac{n}{4} + 1 \leq k \leq \frac{3n}{4}.$$

As in the previous case, we consider the matrix $B_{\frac{n}{2}}$, the same $\frac{n}{2} \times \frac{n}{2}$ block of A_n^{-1}, and we estimate the row sums $\tilde{S}_{B_k}$, $\quad \frac{n}{4}+1 \leq k \leq \frac{3n}{4}$, *i.e.*,

$$\begin{aligned} \tilde{S}_{B_k} \approx & \tfrac{(n+1)\log\left(\frac{n+1}{k}\right)}{2+\log(n+1)} \left(k - \tfrac{n}{4} + k\log(k) - \tfrac{n}{4}\log\left(\tfrac{n}{4}\right)\right) \\ & + \tfrac{(n+1)(2+\log(k))}{2+\log(n+1)} \left(\tfrac{3n}{4} - k + \tfrac{3n}{4}\log\left(\tfrac{n+1}{\frac{3n}{4}}\right) + k\log(\tfrac{n+1}{k})\right). \end{aligned} \tag{31}$$

It is easily checked that (31) implies the same conclusion (25), as in the previous case. Applying again Lemma 4.3 as in (14), we find

$$\kappa(A_n) \sim \rho(A_n^{-1}) \sim n^2.$$

Case 4: $1 < \alpha < 2$.

In analogy with the previous cases we estimate

$$S_i = \sum_{j=i}^{n} \left(\frac{2j+1}{2(n+1)}\right)^{-\alpha} \approx (n+1)\int_{\frac{i}{n+1}}^{1} x^{-\alpha}dx = \frac{n+1}{\alpha-1}\left[\left(\frac{n+1}{i}\right)^{\alpha-1} - 1\right]; \tag{32}$$

$$S = 2^\alpha(n+1)^\alpha + S_1 \approx (n+1)^\alpha\left[2^\alpha + \frac{1}{\alpha-1}\right] - \frac{n+1}{\alpha-1}; \tag{33}$$

$$S - S_i \approx (n+1)^\alpha\left[2^\alpha + \frac{1}{\alpha-1}\left(1 - \frac{1}{i^{\alpha-1}}\right)\right]; \tag{34}$$

$$\begin{aligned} \textstyle\sum_{i=k+1}^{n} S_i &\approx \tfrac{(n+1)^\alpha}{\alpha-1}\textstyle\sum_{i=k+1}^{n} i^{1-\alpha} - \tfrac{(n+1)(n-k)}{\alpha-1} \\ &\approx \tfrac{(n+1)^2}{\alpha-1}\textstyle\int_{\frac{k+1}{n+1}}^{1} x^{1-\alpha}dx - \tfrac{(n+1)(n-k)}{\alpha-1} \\ &= \tfrac{1}{\alpha-1}\left[\tfrac{1}{2-\alpha}(n+1)^2 + (n+1)(k+1) - \tfrac{(n+1)^\alpha(k+1)^{2-\alpha}}{(2-\alpha)}\right]; \end{aligned} \tag{35}$$

$$\begin{aligned} \textstyle\sum_{i=1}^{k}(S - S_i) &\approx (n+1)^\alpha k\left(2^\alpha + \tfrac{1}{\alpha-1}\right) - \tfrac{(n+1)^\alpha}{\alpha-1}\textstyle\sum_{i=1}^{k} i^{1-\alpha} \\ &\approx (n+1)^\alpha k\left(2^\alpha + \tfrac{1}{\alpha-1}\right) - \tfrac{(n+1)^2}{\alpha-1}\textstyle\int_0^{\frac{k}{n+1}} x^{1-\alpha}dx \\ &= (n+1)^\alpha k\left(2^\alpha + \tfrac{1}{\alpha-1} - \tfrac{k^{1-\alpha}}{(\alpha-1)(2-\alpha)}\right). \end{aligned} \tag{36}$$

Substituting the explicit quantities (32), (33), (34), (35), and (36) in relation (15), we deduce that

$$\begin{aligned} \tilde{S}_k \approx & \frac{\frac{n+1}{\alpha-1}\left[\left(\frac{n+1}{k}\right)^{\alpha-1} - 1\right]}{(n+1)^\alpha\left[2^\alpha + \frac{1}{\alpha-1}\right] - \frac{n+1}{\alpha-1}}(n+1)^\alpha k\left(2^\alpha + \tfrac{1}{\alpha-1} - \tfrac{k^{1-\alpha}}{(\alpha-1)(2-\alpha)}\right) \\ & + \frac{(n+1)^\alpha\left[2^\alpha + \frac{1}{\alpha-1}\left(1 - \frac{1}{k^{\alpha-1}}\right)\right]}{(n+1)^\alpha\left[2^\alpha + \frac{1}{\alpha-1}\right] - \frac{n+1}{\alpha-1}}\left[\tfrac{(n+1)^2}{(\alpha-1)(2-\alpha)} + \tfrac{(n+1)(k+1)}{\alpha-1} - \tfrac{(n+1)^\alpha(k+1)^{2-\alpha}}{(\alpha-1)(2-\alpha)}\right]. \end{aligned} \tag{37}$$

A plain analysis of the main terms of (37) shows that the order of $\tilde{S}_k$ does not exceed n^2. Moreover, the study of (30) for $\frac{n}{4}+1 \leq k \leq \frac{3n}{4}$ leads to relation (22), i.e.,

$$\tilde{S}_k \sim n^2, \qquad \frac{n}{4}+1 \leq k \leq \frac{3n}{4}.$$

We consider once again the matrix $B_{\frac{n}{2}}$. Then

$$\begin{aligned}\tilde{S}_{B_k} \approx\ & \frac{\frac{n+1}{\alpha-1}\left[\left(\frac{n+1}{k}\right)^{\alpha-1}-1\right](n+1)^{\alpha}}{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\right]-\frac{n+1}{\alpha-1}}\left[\left(k-\frac{n}{4}\right)\left(2^{\alpha}+\frac{1}{\alpha-1}\right)-\frac{k^{2-\alpha}-\left(\frac{n}{4}\right)^{2-\alpha}}{(\alpha-1)(2-\alpha)}\right] \\ +\ & \frac{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\left(1-\frac{1}{k^{\alpha-1}}\right)\right]}{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\right]-\frac{n+1}{\alpha-1}}\left[\frac{(n+1)^{\alpha}\left((\frac{3n}{4}+1)^{2-\alpha}-k^{2-\alpha}\right)}{(\alpha-1)(2-\alpha)}-\frac{(n+1)(\frac{3n}{4}-k)}{\alpha-1}\right].\end{aligned} \tag{38}$$

The analysis of (38) gives the same conclusion as (25), and then, by Lemma 4.3, we obtain relation (14), i.e., $\kappa(A_n) \sim \rho(A_n^{-1}) \sim n^2$.

Case 5: $\alpha = 2$.

As in the preceding cases we have:

$$S_i = \sum_{j=i}^{n}\left(\frac{2j+1}{2(n+1)}\right)^{-2} \approx (n+1)\int_{\frac{i}{n+1}}^{1} x^{-2}dx = (n+1)\left(\frac{n+1}{i}-1\right); \tag{39}$$

$$S = 4(n+1)^2 + S_1 \approx (n+1)(5n+4); \tag{40}$$

$$S - S_i \approx (n+1)(5n+4) - (n+1)\left(\frac{n+1}{i}-1\right) = (n+1)^2\left(5-\frac{1}{i}\right); \tag{41}$$

$$\begin{aligned}\textstyle\sum_{i=k+1}^{n} S_i &\approx (n+1)\textstyle\sum_{i=k+1}^{n}\left(\frac{i}{n+1}\right)^{-1} - (n+1)(n-k) \\ &\approx (n+1)^2\textstyle\int_{\frac{k+1}{n+1}}^{1} x^{-1}dx - (n+1)(n-k) \\ &= (n+1)^2\log\left(\frac{n+1}{k+1}\right) - (n+1)(n-k);\end{aligned} \tag{42}$$

$$\begin{aligned}\textstyle\sum_{i=1}^{k}(S-S_i) &\approx 5(n+1)^2k - (n+1)\textstyle\sum_{i=1}^{k}\left(\frac{i}{n+1}\right)^{-1} \\ &\approx 5(n+1)^2k - (n+1)^2\textstyle\int_{\frac{1}{n+1}}^{\frac{k+1}{n+1}} x^{-1}dx \\ &= (n+1)^2\left(5k-\log(k+1)\right).\end{aligned} \tag{43}$$

For the estimation of $\tilde{S}_k$ we employ (39), (40), (41), (42), and (43) in relation (15):

$$\begin{aligned}\tilde{S}_k \approx\ & \tfrac{\frac{n+1}{k}-1}{5n+4}(n+1)^2\left(5k-\log(k+1)\right) \\ +\ & \tfrac{(n+1)\left(5-\frac{1}{k}\right)}{5n+4}(n+1)\left((n+1)\log\left(\tfrac{n+1}{k+1}\right)-(n-k)\right) \\ =\ & \tfrac{(n+1)^2}{5n+4}[5+\tfrac{n-k}{k}+5(n+1)\log(n+1) \\ -\ & (5n+4)\log(k+1)-\tfrac{n+1}{k}\log(n+1)].\end{aligned} \tag{44}$$

A straightforward conclusion is that $\tilde{S}_k$ does not exceed $n^2\log(n)$ in order of magnitude. On the other hand, by exploiting (44) for $1 \le k \le m$, where m is a constant integer independent of n, we obtain

$$\tilde{S}_k \sim n^2\log(n), \qquad 1 \le k \le m. \tag{45}$$

By the Perron Frobenius theory on nonnegative matrices, we find

$$\rho(A_n^{-1}) = O(n^2\log(n)). \tag{46}$$

We consider the $m \times m$ matrix B_m, which is the submatrix of A_n^{-1} formed by the first m rows and columns. The estimation of the row sums $\tilde{S}_{B_k}$, $\quad 1 \leq k \leq m$, leads to

$$\begin{aligned}\tilde{S}_{B_k} &\approx \frac{\frac{n+1}{k}-1}{5n+4}(n+1)^2\,(5k-\log(k+1)) \\ &+ \frac{(n+1)\left(5-\frac{1}{k}\right)}{5n+4}(n+1)\left((n+1)\log\left(\frac{m+1}{k+1}\right)-(m-k)\right) \\ &= \frac{(n+1)^2}{5n+4}[5(n+1-m)+\frac{m-k}{k}+5(n+1)\log(m+1) \\ &-(5n+4)\log(k+1)-\frac{n+1}{k}\log(m+1)].\end{aligned} \tag{47}$$

Since m and k are constant independent of n, it follows that

$$\tilde{S}_{B_k} \sim n^2, \qquad 1 \leq k \leq m. \tag{48}$$

By the interlacing law we obtain $\rho(A_n^{-1}) \geq \rho(B_m) \sim n^2$ and therefore

$$\rho(A_n^{-1}) = \Omega(n^2). \tag{49}$$

In conclusion, from (49) and (46), we deduce that

$$\kappa(A_n) \sim \rho(A_n^{-1}) = O(n^2\log(n)) \cap \Omega(n^2).$$

Case 6: $\alpha > 2$.

It is easily seen that the estimation of the quantities S_i, S, $S-S_i$ and $\sum_{i=k+1}^{n} S_i$ is just the same as in Case 4, when dealing with relations (32), (33), (34), and (35), respectively. The only modification we need is to estimate the quantity $\sum_{i=1}^{k} S - S_i$, by exploiting an alternative approximation since $\int_0^{\frac{k}{n+1}} x^{1-\alpha}dx$ diverges for $\alpha > 2$. More in detail we have

$$\begin{aligned}\sum_{i=1}^{k}(S-S_i) &\approx (n+1)^{\alpha}k\left(2^{\alpha}+\frac{1}{\alpha-1}\right)-\frac{(n+1)^{\alpha}}{\alpha-1}\sum_{i=1}^{k} i^{1-\alpha} \\ &\approx (n+1)^{\alpha}k\left(2^{\alpha}+\frac{1}{\alpha-1}\right)-\frac{(n+1)^{\alpha}}{\alpha-1}\int_{\frac{1}{n+1}}^{\frac{k+1}{n+1}} x^{1-\alpha}dx \\ &= (n+1)^{\alpha}\left[k\left(2^{\alpha}+\frac{1}{\alpha-1}\right)-\frac{1-\frac{1}{(k+1)^{\alpha-2}}}{(\alpha-1)(\alpha-2)}\right].\end{aligned} \tag{50}$$

We estimate $\tilde{S}_k$ by replacing (32), (33), (34), (35), and (50) in relation (15):

$$\begin{aligned}\tilde{S}_k \approx\ & \frac{\frac{n+1}{\alpha-1}\left[\left(\frac{n+1}{k}\right)^{\alpha-1}-1\right]}{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\right]-\frac{n+1}{\alpha-1}}(n+1)^{\alpha}\left[k\left(2^{\alpha}+\frac{1}{\alpha-1}\right)-\frac{1-\frac{1}{(k+1)^{\alpha-2}}}{(\alpha-1)(\alpha-2)}\right] \\ &+ \frac{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\left(1-\frac{1}{k^{\alpha-1}}\right)\right]}{(n+1)^{\alpha}\left[2^{\alpha}+\frac{1}{\alpha-1}\right]-\frac{n+1}{\alpha-1}}\left[\frac{(n+1)^{\alpha}(k+1)^{2-\alpha}}{(\alpha-1)(\alpha-2)}-\frac{(n+1)^2}{(\alpha-1)(\alpha-2)}+\frac{(n+1)(k+1)}{\alpha-1}\right].\end{aligned} \tag{51}$$

Again we deduce that $\tilde{S}_k$ grows in order as n^{α}. Moreover, by studying (51) for $1 \leq k \leq \bar{k}$, where $\bar{k}$ is a constant independent of n, we find that both terms of (51) are of order n^α. Thus

$$\tilde{S}_k \sim n^\alpha, \qquad\qquad 1 \leq k \leq \bar{k}. \tag{52}$$

By considering the matrix $B_{\bar{k}}$, the submatrix of A_n^{-1} formed by the first $\bar{k}$ rows and columns, in the formula of the kth row sum of $B_{\bar{k}}$ the first term of (51) appears unchanged, while the changes appear only in the second term. Thus,

$$\tilde{S}_{B_k} \sim n^\alpha, \qquad\qquad 1 \leq k \leq \bar{k}. \tag{53}$$

Finally, by Lemma 4.3 we obtain

$$\kappa(A_n) \sim \rho(A_n^{-1}) \sim n^\alpha, \tag{54}$$

and the proof of the theorem is completed. □

5 The case of higher order BVPs

The results of Theorem 4.4 can be extended in a straightforward manner to cover the case where the BVP is of order higher than 2, i.e., our equations are of the form

$$\begin{cases} (-1)^k \frac{d^k}{dx^k}\left(a(x)\frac{d^k}{dx^k}u(x)\right) = f(x) & \text{on } \Omega = (0,1), \quad k = 2,3,\ldots \\ \text{homogeneous B.C. on } \partial\Omega, & \end{cases} \tag{55}$$

where the function $a(x)$ has a root at $\tilde{x}_0 \in \overline{\Omega}$ of order α. In analogy to the case of second order operators, we approximate (55) on a uniform grid of stepsize $h = (n+1)^{-1}$, using centered finite differences of minimal precision order 2. As a consequence we find $2k+1$ band $n \times n$ linear systems $A_n(a)x = b$.

The generalization of Theorem 4.4 takes the following form:
Theorem 5.1. *Let $\{A_n\}_n$, $A_n \in \mathbb{R}^{n\times n}$, be the sequence of matrices derived from the descritization of the Semielliptic Differential Equation (55) with the bounded coefficient function $a(x)$ having a unique root at 0 of order α i.e. $a(x) \sim x^\alpha$ on $D = [0,1]$. Then, for the spectral condition number $\kappa(A_n)$ of the matrix A_n which coincides in order with the spectral radius of A_n^{-1}, there holds:*

$$\kappa_2(A_n) \sim \begin{cases} n^{2k}, & 0 \leq \alpha < 2k, \\ O(n^{2k}\log(n)) \cap \Omega(n^{2k}), & \alpha = 2k, \\ n^\alpha, & \alpha > 2k. \end{cases} \tag{56}$$

The proof follows exactly the same governing ideas as the proof of Theorem 4.4 but with the mathematical manipulation becoming more and more complicated and tricky as the order of the BVP increases. The reason for that concerns essentially the formulation of the explicit form for the inverse of the coefficient matrix A_n. In Section 7 we give many numerical examples regarding the case of BVPs with order higher than two, with all of them fully confirming the theoretical results given in Theorem 5.1.

Remark 5.1. *The assumption for the coefficient function regarding the uniqueness of its root cannot be relaxed to "many isolated roots". The reason, which has been mentioned also in [25], is that in this case the condition number grows in an unpredictable (nonmonotone) way as the dimension of the problem tends to infinity: in reality, the matrix A_n may happen to be also singular for certain dimensions. More specifically, performing various numerical experiments (see Tables 9, 10, and 11 for a partial account on our findings), we have observed that:*

- *For k=1,2,3 there exists $a(x)$ such that $\max\{\alpha_i\} < 2k$ and $\kappa_2(A_n) \not\sim n^{2k}$; more precisely, $\kappa_2(A_n) = \Omega(n^{2k+\delta})$, for some $\delta > 0$;*
- *For k=1,2,3 there exists $a(x)$ such that $\max\{\alpha_i\} = 2k$ and $\kappa_2(A_n) \not\sim O(n^{2k}\log(n)) \cap \Omega(n^{2k})$; more in detail, $\kappa_2(A_n) = \Omega(n^{2k+\delta})$, for some $\delta > 0$;*
- *For k=1,2,3 there exists $a(x)$ such that $\max\{\alpha_i\} > 2k$ and $\kappa_2(A_n) \not\sim n^{\max\{\alpha_i\}}$; more precisely, $\kappa_2(A_n) = \Omega(n^{\max\{\alpha_i\}+\delta})$, for some $\delta > 0$.*

In Section 7 we report some examples concerning this case, and the conclusion is that the condition numbers grow faster, when compared with the bounds in Theorem 4.4 and Theorem 5.1: the reason is a kind of interference between the sources of ill-conditioning represented by the different zeros (for a nice contrast with the case of a unique zero, see the discussion at the end of Subsection 2.2).

6 Remarks on the 2D case

We consider the 2D problem

$$-\frac{\partial}{\partial x}\left(a(x,y)\frac{\partial}{\partial x}u\right) - \frac{\partial}{\partial y}\left(b(x,y)\frac{\partial}{\partial y}u\right) = f(x,y) \qquad (57)$$

with Dirichlet boundary conditions. Using the well-known five points formula and by ordering the unknowns in the classic manner, we arrive to the $n^2 \times n^2$ linear system

$$A_{nn}x = b,$$

where A_{nn} is a symmetric positive definite block tridiagonal matrix, with the diagonal blocks being tridiagonal matrices and the off diagonal blocks being diagonal ones.

As we have mentioned from the beginning of this paper, the main contribution of this work will be to give a guideline and to establish a theoretical framework for dealing with the more interesting 2D case, which is of great importance from both, theoretical and practical point of view. A trivial but immediate application of our estimation of the condition number to the 2D case, is the circumstance where the coefficient functions are of separable variables. In addition, we perform various numerical experiments and it clearly emerges that the results of Theorem 4.4, under suitable assumptions, can be analogously extended to cover also the 2D case. The following definition is useful.

Definition 6.1. *Let $f(x,y)$ be a nonnegative bounded function having a zero at (x_0,y_0). We say that the order of zero is $\alpha \in (0,\infty)$ if there exists a finite number p of curves C_i, $i=1,\ldots,p$, defined by $l_i(x,y)=0$, passing through (x_0,y_0) and regular in it such that $f \sim \hat{f}$ and*

$$\hat{f}(x,y) = \sum_{i=1}^{p} |l_i(x,y)|^{\alpha} + g(x,y),$$

where g has a zero at (x_0,y_0) of order at least $\beta > \alpha$.

We are ready to state our conjecture concerning the relation of the condition of A_{nn} and the order of the zeros of the coefficient functions:

Statement 6.1. *Let us assume that the coefficient functions $a(x,y), b(x,y)$ have zeros $(x_0,y_0), (x_1,y_1)$ of orders α_a, α_b, respectively. Then for the spectral condition number $\kappa_2(A_{nn})$ of the matrix A_{nn} there holds:*

$$\kappa_2(A_{nn}) \sim \begin{cases} n^2, & 0 \leq \min\{\alpha_a,\alpha_b\} < 2; \\ O(n^2\log(n)) \cap \Omega(n^2), & \min\{\alpha_a,\alpha_b\} = 2, \\ n^{\min\{\alpha_a,\alpha_b\}}, & \min\{\alpha_a,\alpha_b\} > 2. \end{cases}$$

7 Numerical experiments

In this section we present several numerical tests concerning both 1D and 2D BVPs. We will start by discussing experiments on univariate BVPs of order 2, 4, and 6, respectively.

The quantity which is of main interest in our context is the estimation of

$$\rho_m = \log_2\left(\frac{\lambda_{\min}(A_{2^m})}{\lambda_{\min}(A_{2^{(m+1)}})}\right).$$

We observe that ρ_m reflects the decrement rate of the minimal eigenvalue of the coefficient matrix A_n.

For the second order BVP in (1) we have used as coefficient functions the following test functions:

$$a_1(x) = \left|x - \frac{1}{\sqrt{2}}\right|, \quad a_2(x) = (x - .3)^2, \quad a_3(x) = \left|x - \frac{\pi}{4}\right|^{\frac{5}{2}}$$

and the results are given in Tables 1, 2, and 3, respectively. Regarding the fourth order BVP i.e. (55) with $k = 2$, we use the functions

$$a_4(x) = \left(x - \frac{1}{\sqrt{3}}\right)^2, \quad a_5(x) = \sin(x)^4, \quad a_6(x) = x^5,$$

with associated results in Tables 4, 5, and 6, while, for the sixth order BVP i.e. (55) with $k = 3$, we have chosen as coefficient functions

$$a_7(x) = \sin(x)^4, \quad a_8(x) = x^7,$$

with related results in Tables 7, and 8.

Obviously, in order to perform a meaningful test for our theoretical derivations, the considered coefficient functions have different analytical behaviors, and with roots of order less, equal or greater than the order of the differential equation. In all cases, we ascertain numerically the theoretical findings in Theorems 4.4 and 5.1.

Table 1
1D, $k = 1$:$a(x) = \left|x - \frac{1}{\sqrt{2}}\right|$.

m	5	6	7	8	9
$\lambda_{\min}$	2.263×10^{-3}	5.483×10^{-4}	1.324×10^{-4}	3.2×10^{-5}	7.76×10^{-6}
ρ_m	2.045	2.05	2.048	2.044	2.04
m	10	11	12	13	14
$\lambda_{\min}$	1.889×10^{-6}	4.612×10^{-7}	1.129×10^{-7}	2.773×10^{-8}	6.824×10^{-9}
ρ_m	2.034	2.03	2.026	2.023	

For the case where $a(x)$ has multiple roots in $[0, 1]$ things completely change as reported in Remark 5.1. Tables 9, 10, and 11 show this "irregular" behavior for the quantity ρ_m.

a) $k = 2$ $\quad a(x) = x^3|x - .3|^{\frac{5}{2}}$,
b) $k = 2$ $\quad a(x) = (x - \frac{1}{\sqrt{2}})^2(x - \frac{1}{\sqrt{3}})^4$,
c) $k = 1$ $\quad a(x) = (x - .5)^2x^3$.

For the 2D case, we consider the following four examples:

a) $a(x, y) = b(x, y) = x + y$,

Table 2
1D, $k = 1{:}a(x) = (x - .3)^2$.

m	5	6	7	8	9
$\lambda_{\min}$	3.251×10^{-4}	9.608×10^{-5}	2.063×10^{-5}	4.882×10^{-6}	1.179×10^{-6}
ρ_m	1.759	2.22	2.079	2.050	1.898
m	10	11	12	13	14
$\lambda_{\min}$	3.163×10^{-7}	7.282×10^{-8}	1.762×10^{-8}	4.318×10^{-9}	1.123×10^{-9}
ρ_m	2.119	2.047	2.029	1.943	

Table 3
1D, $k = 1{:}a(x) = (x - \pi/4)^{\frac{5}{2}}$.

m	5	6	7	8	9
$\lambda_{\min}$	5.206×10^{-5}	9.278×10^{-6}	2.46×10^{-6}	3.296×10^{-7}	5.446×10^{-8}
ρ_m	2.488	1.915	2.9	2.597	2.549
m	10	11	12	13	14
$\lambda_{\min}$	9.305×10^{-9}	2.282×10^{-9}	3.633×10^{-10}	6.529×10^{-11}	1.193×10^{-11}
ρ_m	2.028	2.651	2.476	2.452	

Table 4
1D, $k = 2{:}a(x) = \left|x - \frac{1}{\sqrt{3}}\right|$.

m	5	6	7	8	9	10	11	12	13
ρ_m	3.646	3.961	3.996	4.038	3.995	4.151	3.851	4.022	4.034

Table 5
1D, $k = 2$: $a(x) = \sin(x)^4$.

m	5	6	7	8	9	10	11	12	13
ρ_m	4.208	4.189	4.161	4.135	4.113	4.094	4.078	4.066	4.056

Table 6
1D, $k = 2$: $a(x) = x^5$.

m	5	6	7	8	9	10	11	12	13
ρ_m	4.911	4.951	4.974	4.987	4.993	4.997	4.998	4.999	5

b) $a(x,y) = x^3 + y^4$, $b(x,y) = x^5 + y^6$,
c) $a(x,y) = x^2 + y^2$, $b(x,y) = (x+y)^2$,
d) $a(x,y) = |x-y|^3$, $b(x,y) = |x - \frac{1}{2}|^3 + |y - \frac{1}{2}|^3$.

The results in Tables 12, 13, 14, and 15 fully confirm the statements formulated

Table 7
1D, $k=3$: $a(x)=\sin(x)^4$.

m	5	6	7	8	9	10	11	12	13
ρ_m	5.853	5.929	5.965	5.983	5.992	5.996	5.999	5.999	6

Table 8
1D, $k=3$: $a(x)=x^7$.

m	5	6	7	8	9	10	11	12	13
ρ_m	6.846	6.922	6.961	6.980	6.990	6.995	6.998	6.999	6.999

Table 9
1D, $k=2$, multiple root case: $a(x)=x^3|x-.3|^{\frac{5}{2}}$.

m	5	6	7	8	9	10	11	12	13
ρ_m	5.51	2.982	2.278	3.574	7.025	1.811	1.779	3.52	6.947

Table 10
1D, $k=2$, multiple root case : $a(x)=\left(x-\frac{1}{\sqrt{2}}\right)^2\left(x-\frac{1}{\sqrt{3}}\right)^4$.

m	5	6	7	8	9	10	11	12	13
ρ_m	7.264	4.516	4.601	4.316	5.192	4.719	6.447	3.725	4.876

Table 11
1D, $k=2$ multiple root case: $a(x)=(x-.5)^2x^3$.

m	5	6	7	8	9	10	11	12	13
ρ_m	3.89	3.944	3.972	3.986	3.993	3.997	3.998	3.999	4

at the end of Section 6.

Table 12
2D case: $a(x,y)=b(x,y)=x+y$

m	3	4	5	6	7	8
ρ_m	1.826	1.911	1.956	1.978	1.989	1.995

Table 13
2D case: $a(x,y)=x^2+y^2,\ b(x,y)=(x+y)^2$

m	3	4	5	6	7	8
ρ_m	1.921	1.967	1.990	2.001	2.005	2.008

Table 14
2D case: $a(x,y) = x^3 + y^4,\ b(x,y) = x^5 + y^6$

m	3	4	5	6	7	8
ρ_m	2.885	2.949	2.979	2.992	2.997	2.999

Table 15
2D Case: $a(x,y) = |x-y|^3,\ b(x,y) = |x-\frac{1}{2}|^3 + |y-\frac{1}{2}|^3$

m	3	4	5	6	7	8
ρ_m	2.757	2.871	2.934	2.967	2.983	2.991

8 Conclusions

In this paper we have studied the conditioning of semi-elliptic differential problems (the elliptic case is plain thanks to monotonicity arguments). As a main tool we have employed the notion of positivity in three different aspects: definite positivity, operator positivity (especially in Subsection 2.2), and component-wise positivity (especially in Section 4). Our main result is that the two sources of ill-conditioning, the low frequencies coming from the constant coefficient Laplacian, and the space spanned by few canonical vectors related to the position of the zero of $a(x)$, do not interfere; conversely, we numerically observe a bad interference, a kind of resonance, in presence of distinct zeros in the coefficient $a(x)$. Therefore, when a unique zero is considered, there is only a superposition effect so that the size of the degenerating subspace, i.e. that related to small eigenvalues, becomes larger, but the order of ill-conditioning is not worse than that of the two factors separately. As a consequence, both for designing multigrid methods or preconditioners, we can treat the two ill-conditioned spaces separately and this of course implies a simplification in the practical programming and in the theoretical convergence analysis (see e.g. [21,24,26,4])). Finally, there is still the open problem of completing our study in three directions: we would like to identify the constants hidden in the equivalence relations of the main Theorems 4.4 and 5.1, we would like to add more terms if the asymptotic expansion of the condition number of A_n, and, more important, we would like to include the more challenging multidimensional setting. Indeed, as a final remark, we stress that partial results are easily available, by repeating e.g. the same derivations as in Subsection 2.2 in a multilevel setting: however a complete analysis is still missing.

References

[1] D. Aregba-Driollet, R. Natalini, and S. Tang, *Explicit diffusive kinetic schemes for nonlinear degenerate parabolic systems*, Math. Comp. **73-245** (2004), 63–94.

[2] W. Barrett, *A theorem on inverse of tridiagonal matrices*, Linear Algebra Appl. **27** (1979), 211–217.

[3] B. Beckermann and S. Serra Capizzano, *On the asymptotic spectrum of Finite Elements matrices*, SIAM J. Numer. Anal. (to appear).

[4] D. Bertaccini, G. Golub, S. Serra Capizzano, and C. Tablino Possio, *Preconditioned HSS methods for the solution of non-Hermitian positive definite linear systems and applications to the discrete convection-diffusion equation*, Numer. Math. **99** (2005), 441–484.

[5] R. Bevilacqua and M. Capovani, *Proprietà delle matrici a banda ad elementi ed a blocchi*, Bolletino UMI **13-B** (1976), 844–861.

[6] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[7] A. Böttcher and S. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*, SIAM Publ., Philadelphia, PA, 2005.

[8] M. Capovani, *Sulla determinazione della inversa delle matrici tridiagonali e tridiagonali a blocchi*, Calcolo **7** (1970), 295–303.

[9] ———, *Su alcune propiertà delle matrici tridiagonali e pentadiagonali*, Calcolo **8** (1971), 149–159.

[10] R. Chan and T. Chan, *Circulant preconditioners for elliptic problems*, J. Numer. Linear Algebra Appl. **1** (1992), 77–101.

[11] G. Fiorentino and S. Serra Capizzano, *Fast parallel solvers for elliptic problems*, Comput. Math. Appl. **32** (1996), 61–68.

[12] ———, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comput. **17** (1996), no. 4, 1068–1081.

[13] F. Gantmacher and M. Krein, *Oszillationsmatrizen, oszillationskerne und kleine schwingungen mechanischer systeme*, Akademie-Verlag, Berlin, 1960.

[14] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, London, 1996.

[15] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM Publ., Philadelphia, PA, 1997.

[16] D. Marini and P. Pietra, *Mixed finite element approximation of a degenerate elliptic problem*, Numer. Math. **71** (1995), 225–236.

[17] J. McDonald, R. Nabben, M. Neumann, Schneider H., and M. Tsatsomeros, *Inverse tridiagonal Z-matrices*, Linear and Multilinear Algebra **45** (1998), 75–97.

[18] R. Nabben, *Decay rates of the inverse of nonsymmetric tridiagonal and band matrices*, SIAM J. Matrix Anal. Appl. **20** (1999), 820–837.

[19] P. Rózsa, *On the inverse of band matrices*, Integral equations Operator Theory **10** (1987), 82–95.

[20] S. Serra Capizzano, *Multi-iterative methods*, Comput. Math. Appl. **26** (1993), 65–87.

[21] ______, *The rate of convergence of Toeplitz based PCG methods for second order nonlinear boundary value problems*, Numer. Math. **81** (1999), 461–495.

[22] ______, *Locally X matrices, spectral distributions, preconditioning, and applications*, SIAM J. Matrix Anal. Appl. **21** (2000), 1354–1388.

[23] ______, *Some theorems on linear positive operators and functionals and their applications*, Comput. Math. Appl. **39** (2000), 139–167.

[24] ______, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math. **92** (2002), 433–465.

[25] S. Serra Capizzano and C. Tablino Possio, *Spectral and structural analysis of high precision finite difference matrices for elliptic operators*, Linear Algebra Appl. **293** (1999), 85–131.

[26] ______, *Multigrid methods for multilevel circulant matrices*, SIAM J. Sci. Comput. **26** (2004), 55–85.

[27] S. Serra Capizzano and P. Tilli, *Extreme singular values and eigenvalues of non Hermitian Toeplitz matrices*, J. Computat. Appl. Math. **108** (1999), no. 1-2, 113–130.

[28] ______, *On unitarily invariant norms of matrix valued linear positive operators*, J. Ineq. Appl. **7** (2002), no. 3, 309–330.

[29] P. Tilli, *Locally Toeplitz sequences: spectral properties and applications*, Linear Algebra Appl. **278** (1998), 91–120.

[30] U. Trottenberg, C.W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.

[31] R. Varga, *Matrix Iterative Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1962.

[32] P. Wilmott, S. Howison, and J. Dewynne, *The Mathematics of Financial Derivatives*, Cambridge Univ. Press, Cambridge, MA, 1998.

Oscillation Criteria of First Order Linear Difference Equations with Delay Argument

G. E. Chatzarakis,[a] R. Koplatadze,[b] and I. P. Stavroulakis[a]

Abstract

This paper presents new sufficient conditions for the oscillation of all proper solutions of the first order linear difference equation with delay argument

$$\Delta u(k) + p(k)\, u(\tau(k)) = 0, \quad k \in N,$$

where $\Delta u(k) = u(k+1) - u(k)$, $p : N \to R_+$, $\tau : N \to N$ and $\lim\limits_{k \to +\infty} \tau(k) = +\infty$. Examples illustrating the results are given. It is to be pointed out that this is the first paper dealing with the oscillatory behaviour of the equation in the case of a general delay argument $\tau(k)$.

[a] Department of Mathematics, University of Ioannina, 451 10, Greece, ipstav@cc.uoi.gr
[b] A. Razmadze Mathematical Institute Georgia Academy of Sciences, 1, M. Aleksidze st., Tbilisi, Georgia

1 Introduction

Consider the first order linear difference equation with delay argument

$$\Delta u(k) + p(k)\, u(\tau(k)) = 0, \quad k \in N, \tag{E}$$

where $\Delta u(k) = u(k+1) - u(k)$, $p : N \to R_+$, $\tau : N \to N$ and $\lim\limits_{k \to +\infty} \tau(k) = +\infty$.

Strong interest in equation (E) is motivated by the fact, that it represents a discrete analogue of the delay differential equation (see [12] and the references cited therein)

$$x'(t) + p(t)\, x(\tau(t)) = 0, \quad p(t) \geq 0, \quad \tau(t) \leq t \quad \text{for} \quad t \geq 0.$$

By a *proper solution* of Eq.(E) we mean a function $u : N_{n_0} \longrightarrow R$, $n_0 = \min\{\tau(k) : k \in N_n\}$, $N_n = \{n, n+1, \dots\}$ which satisfies Eq.(E) on N_n and $\sup\{|u(i)| : i \geq k\} > 0$ for $k \in N_{n_0}$.

A proper solution $u : N \longrightarrow R$ of Eq.(E) is said to be *oscillatory* (around zero) if for every positive integer n there exist $n_1, n_2 \in N_n$, such that $u(n_1)u(n_2) \leq 0$. Otherwise, the solution is said to be *non-oscillatory.* In other words, a proper solution u is oscillatory if it is neither eventually positive nor eventually negative.

In the last few decades the oscillation theory of delay differential equations has been extensively developed. The oscillation theory of discrete analogues of delay differential equations has also attracted growing attention in the recent few years. In particular, the problem of establishing sufficient conditions for the oscillation of all solutions of the equation

$$\Delta u(k) + p(k)\, u(k-n) = 0, \quad k \in N \tag{E$_1$}$$

has been the subject of many recent investigations. See for example [2–11, 13–16] and the references cited therein.

In 1989, Erbe and Zhang [6], proved that, if $p(k) \geq 0$, then either one of the following conditions

$$\liminf_{k \longrightarrow +\infty} p(k) > \frac{n^n}{(n+1)^{n+1}} \tag{1.1}$$

or

$$\limsup_{k \longrightarrow +\infty} \sum_{i=k-n}^{k} p(i) > 1 \tag{1.2}$$

implies that all solutions of Eq.(E$_1$) oscillate.

In the same year, Ladas, Philos and Sficas [9], proved that the same conclusion holds if $p(k) \geq 0$ and

$$\liminf_{k \longrightarrow +\infty} \left(\frac{1}{n} \sum_{i=k-n}^{k-1} p(i) \right) > \frac{n^n}{(n+1)^{n+1}}\,. \tag{1.3}$$

It is interesting to establish sufficient conditions for the oscillation of all solutions of Eq.(E_1) when the conditions (1.2) and (1.3) are not satisfied. Many researchers focused on the improvement of the upper bound of the ratio $u(k-n)/u(k)$ for possible non-oscillatory solutions u of Eq.(E_1). In 1993, Yu, Zhang and Qian [16], and Lalli and Zhang [10], trying to improve (1.2) established some false oscillation conditions due to the fact that both were based on an erroneous discrete version of the Koplatadze-Chanturia lemma [8]. For more details the reader is referred to [5,3].

In 1995, Stavroulakis [14], proved that if

$$0 < \alpha := \liminf_{k \to +\infty} \sum_{i=k-n}^{k-1} p(i) \le \left(\frac{n}{n+1}\right)^{n+1} \quad \text{and} \quad \limsup_{k \to +\infty} p(k) > 1 - \frac{\alpha^2}{4}$$

then all solutions of Eq.(E_1) oscillate.

In 1999, Domshlak [5], and in 2000, Cheng and Zhang [3], established the following lemmas respectively, which may be looked upon as discrete versions of Koplatadze-Chanturia lemma [8].

Lemma 1.1 *([5]) Assume that u is an eventually positive solution of Eq.*(E_1) *and that*

$$\sum_{i=k-n}^{k-1} p(i) \ge \alpha > 0 \quad \textit{for large} \quad k.$$

Then

$$u(k) > \frac{\alpha^2}{4} u(k-n) \quad \textit{for large} \quad k. \tag{1.4}$$

Lemma 1.2 *([3]) Assume that u is an eventually positive solution of Eq.*(E_1) *and that*

$$\sum_{i=k-n}^{k-1} p(i) \ge \alpha > 0 \quad \textit{for large} \quad k.$$

Then

$$u(k) > \alpha^n u(k-n) \quad \textit{for large} \quad k. \tag{1.5}$$

In 2004, Stavroulakis [15], based on the above two lemmas, established the following theorem.

Theorem 1.1 *([15]) Assume that*

$$0 < \alpha := \liminf_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) \le \left(\frac{n}{n+1}\right)^{n+1}.$$

Then either one of the conditions

$$\limsup_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) > 1 - \frac{\alpha^2}{4} \tag{1.6}$$

or

$$\limsup_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) > 1 - \alpha^n \tag{1.7}$$

implies that all solutions of Eq.(E_1) *oscillate.*

In 2006, Chatzarakis and Stavroulakis [2], established the following lemma.

Lemma 1.3 *([2]) Assume that u is an eventually positive solution of Eq.*(E_1) *and that*

$$\sum_{i=k-n}^{k-1} p(i) \ge \alpha > 0 \quad \textit{for large} \quad k.$$

Then

$$u(k) > \frac{\alpha^2}{2(2-\alpha)} u(k-n) \quad \textit{for large} \quad k. \tag{1.8}$$

Based on the above lemma, they established the following theorem.

Theorem 1.2 *([2]) Assume that*

$$0 < \alpha := \liminf_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) \le \left(\frac{n}{n+1}\right)^{n+1}$$

and

$$\limsup_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) > 1 - \frac{\alpha^2}{2(2-\alpha)}. \tag{1.9}$$

Then all solutions of Eq. (E_1) *oscillate.*

In this paper, the authors improve the upper bound of the ratio $u(\tau(k))/u(k+1)$ for possible non-oscillatory proper solutions u of Eq.(E) and derive new sufficient oscillation conditions. It is to be emphasized that this is the first paper dealing with the oscillatory behaviour of Eq.(E) in the case of a general delay argument $\tau(k)$.

2 Oscillation Criteria for Eq. (E)

In this section we first establish two lemmas which will be used in the proof of our main results.

Consider the difference inequality

$$\Delta u(k) + q(k)\, u(\sigma(k)) \le 0, \quad k \in N, \tag{2.1}$$

where

$$q : N \to R_+, \quad \sigma : N \to N \quad \text{and} \quad \lim_{k\to+\infty} \sigma(k) = +\infty. \tag{2.2}$$

Lemma 2.1 *Let*

$$\liminf_{k\to+\infty} \sum_{i=\tau(k)}^{k-1} p(i) = \alpha > 0, \tag{2.3}$$

$$\sigma(k) \le \tau(k) \le k-1, \quad p(k) \le q(k) \quad \textit{for} \quad k \in N \tag{2.4}$$

and $u : N_{n_0} \to (0, +\infty)$ *be a positive proper solution of* (2.1) *for a certain* $n_0 \in N$. *Then Eq.*(E) *has a proper solution* $u_* : N_{n_1} \to (0, +\infty)$ *such that*

$$0 < u_*(k) \le u(k) \quad \textit{for} \quad k \in N_{n_1}, \tag{2.5}$$

where $n_1 > n_0$ *is a sufficiently large natural number.*

Proof. Let $u : N_{n_0} \to (0, +\infty)$ be a positive proper solution of (2.1). By (2.2) and (2.3), it is clear that there exists $n_1 \in N_{n_0}$ such that

$$u(\sigma(k)) > 0 \quad \text{and} \quad \sum_{i=\tau(k)}^{k-1} p(i) > 0 \quad \text{for} \quad k \in N_{n_1}. \tag{2.6}$$

From (2.1), we have

$$u(k) \ge \sum_{i=k}^{+\infty} q(i)\, u(\sigma(i)) \quad \text{for} \quad k \in N_{n_1}. \tag{2.7}$$

Assume that $n_* = \min\{\tau(k) : k \in N_{n_1}\}$ and consider the sequence of functions $u_i : N_{n_*} \to R$ $(i = 1, 2, \dots\}$ defined as follows

$$u_1(k) = u(k) \quad \text{for} \quad k \in N_{n_*},$$

$$u_j(k) = \begin{cases} \displaystyle\sum_{i=k}^{+\infty} p(i)\, u_{j-1}(\tau(i)) & \text{for} \quad k \in N_{n_1}, \\ u(k) & \text{for} \quad k \in [n_*, n_1) \quad (j = 2, 3, \dots). \end{cases} \tag{2.8}$$

By (2.4), (2.7) and using the fact that the function u is nonincreasing, we have

$$u_2(k) = \sum_{i=k}^{+\infty} p(i)\, u_1(\tau(i)) \le \sum_{i=k}^{+\infty} q(i)\, u(\sigma(i)) \le u(k) = u_1(k) \quad \text{for} \ \ k \ge n_1.$$

Thus

$$u_j(k) \le u_{j-1}(k) \quad \text{for} \quad k \in N_{n_1} \quad (j = 2, 3, \dots). \tag{2.9}$$

Denote $\lim\limits_{j\to+\infty} u_j(k) = u_*(k)$ (according to (2.9) this limit exists). Therefore, from (2.8), we get

$$u_*(k) = \sum_{i=k}^{+\infty} p(i)\, u_*(\tau(i)) \quad \text{for} \quad k \in N_{n_1}. \tag{2.10}$$

Now, we will show that $u_*(k) > 0$ for $k \ge n_1$. Assume, for the sake of contradiction, that there exists $n_2 \ge n_1$ such that $u_*(k) = 0$ for $k \in N_{n_2}$ and $u_*(k) > 0$ for $k \in [n_*, n_2)$. Denote by N^* the set of natural numbers n for which $\tau(k) = n_2$ and $n^* = \min N^*$. By (2.10) and (2.4) it is clear that $n^* \ge n_2$. Therefore, if $c = \min\{u_*(\tau(i)) : \tau(n^*) \le i \le n^* - 1\} > 0$, by (2.4) and (2.6), we have

$$u_*(n_2) = \sum_{i=n_2}^{+\infty} p(i)\, u_*(\tau(i)) \ge \sum_{i=\tau(n^*)}^{n^*-1} p(i)\, u_*(\tau(i)) \ge c \sum_{i=\tau(n^*)}^{n^*-1} p(i) > 0,$$

which, in view of $u_*(n_2) = 0$, leads to a contradiction. Therefore, $u_*(k) > 0$ for $k \ge n_1$.

Hence Eq.(E) has a proper solution u_* satisfying $0 < u_*(k) \le u(k)$ for $k \in N_{n_1}$. The proof is complete.

Lemma 2.2 *Assume that u is a positive proper solution of Eq.(E), where*

$$\begin{gathered} p : N \to R_+, \quad \tau : N \to N \ \ \textit{is nondecreasing function,} \\ \tau(k) \le k-1, \quad \textit{for} \quad k \in N, \quad \lim_{k\to+\infty} \tau(k) = +\infty \end{gathered} \tag{2.11}$$

and

$$\liminf_{k\to+\infty} \sum_{i=\tau(k)}^{k-1} p(i) = \alpha \in (0, 1]. \tag{2.12}$$

Then

$$\limsup_{k\to+\infty} \frac{u(\tau(k))}{u(k+1)} \le \left(\frac{1+\sqrt{1-\alpha}}{\alpha}\right)^2. \tag{2.13}$$

If, additionally, $p(k) \ge 1-\sqrt{1-\alpha}$ *for large* k*, then*

$$\limsup_{k\to+\infty} \frac{u(\tau(k))}{u(k+1)} \le \frac{1-\alpha+\sqrt{1-\alpha}}{\alpha^2}. \tag{2.14}$$

Proof. By (2.12), it is clear that, for any $\varepsilon \in (0,\alpha)$ there exists $n_0 = n_0(\varepsilon) \in N$ such that

$$\sum_{i=\tau(k)}^{k-1} p(i) \ge \alpha - \varepsilon \quad \text{for} \quad k \in N_{n_0}. \tag{2.15}$$

Since u is a positive proper solution of Eq.(E), then there exists $n_1 \in N_{n_0}$ such that

$$u(\tau(k)) > 0 \quad \text{for} \quad k \in N_{n_1}.$$

Thus, from Eq.(E), we have

$$u(k+1) - u(k) = -p(k)\, u(\tau(k)) \le 0$$

and so u is an eventually nonincreasing function of positive numbers.

From (2.15), it is clear that, if $\omega \in (0, \alpha-\varepsilon)$, there exists $k^* \ge k$ such that

$$\sum_{i=k}^{k^*-1} p(i) < \omega \quad \text{and} \quad \sum_{i=k}^{k^*} p(i) \ge \omega. \tag{2.16}$$

This is because in the case where $p(k) < \omega$, it is clear that, there exists $k^* > k$ such that (2.16) is satisfied, while in the case where $p(k) \ge \omega$, then $k^* = k$, and therefore

$$\sum_{i=k}^{k^*-1} p(i) = \sum_{i=k}^{k-1} p(i) \text{ (by which we mean)} = 0 < \omega \quad \text{and} \quad \sum_{i=k}^{k^*} p(i) = \sum_{i=k}^{k} p(i) = p(k) \ge \omega.$$

That is, in both cases (2.16) is satisfied. Thus

$$\sum_{i=\tau(k^*)}^{k-1} p(i) = \sum_{i=\tau(k^*)}^{k^*-1} p(i) - \sum_{i=k}^{k^*-1} p(i) \ge (\alpha-\varepsilon) - \omega.$$

Now, summing up Eq.(E) first from k to k^* and then from $\tau(k^*)$ to $k-1$, and using the fact that the function u is nonincreasing and the function τ is nondecreasing, we have

$$u(k) - u(k^*+1) = \sum_{i=k}^{k^*} p(i)\, u(\tau(i)) \geq \left(\sum_{i=k}^{k^*} p(i)\right) u(\tau(k^*)) \geq \omega\; u(\tau(k^*))$$

or

$$u(k) \geq u(k^*+1) + \omega u(\tau(k^*)) \tag{2.17}$$

and then

$$u(\tau(k^*)) - u(k) = \sum_{i=\tau(k^*)}^{k-1} p(i)\, u(\tau(i)) \geq \left(\sum_{i=\tau(k^*)}^{k-1} p(i)\right) u(\tau(k-1)) \geq [(\alpha-\varepsilon)-\omega]u(\tau(k-1))$$

or

$$u(\tau(k^*)) \geq u(k) + [(\alpha-\varepsilon)-\omega]u(\tau(k-1)). \tag{2.18}$$

Combining inequalities (2.17) and (2.18), we obtain

$$u(k) \geq u(k^*+1) + \omega[u(k) + ((\alpha-\varepsilon)-\omega)u(\tau(k-1))]$$

or

$$u(k) \geq \frac{\omega[(\alpha-\varepsilon)-\omega]}{1-\omega} u(\tau(k-1)). \tag{2.19}$$

Observe that the function $f : (0,\alpha) \to (0,1)$ defined as

$$f(\omega) : \;= \frac{\omega[(\alpha-\varepsilon)-\omega]}{1-\omega} \tag{2.20}$$

attains its maximum at $\omega = 1 - \sqrt{1-(\alpha-\varepsilon)}$, which is equal to

$$f_{\max} = \left(1 - \sqrt{1-(\alpha-\varepsilon)}\right)^2.$$

Thus, for $\omega = 1 - \sqrt{1-(\alpha-\varepsilon)} \in (0, \alpha-\varepsilon)$ inequality (2.19) becomes

$$u(k) \geq \left(1 - \sqrt{1-(\alpha-\varepsilon)}\right)^2 u(\tau(k-1))$$

or

$$\frac{u(\tau(k-1))}{u(k)} \leq \left(\frac{1+\sqrt{1-(\alpha-\varepsilon)}}{\alpha-\varepsilon}\right)^2 \tag{2.21}$$

and, for large k, we have

$$\frac{u(\tau(k))}{u(k+1)} \leq \left(\frac{1+\sqrt{1-(\alpha-\varepsilon)}}{\alpha-\varepsilon}\right)^2.$$

Hence,

$$\limsup_{k\to+\infty} \frac{u(\tau(k))}{u(k+1)} \leq \left(\frac{1+\sqrt{1-(\alpha-\varepsilon)}}{\alpha-\varepsilon}\right)^2,$$

which, for arbitrarily small values of ε, implies (2.13).

Next we consider the particular case where $p(k) \geq 1-\sqrt{1-\alpha}$.

In this case, from Eq.(E) we have

$$u(k) = u(k+1) + p(k)\,u(\tau(k)) \geq (1-\sqrt{1-\alpha})\,u(\tau(k)). \tag{2.22}$$

Now, summing up Eq.(E) from $\tau(k)$ to $k-1$, and using the fact that the function u is nonincreasing and the function τ is nondecreasing, we have

$$u(\tau(k))-u(k) = \sum_{i=\tau(k)}^{k-1} p(i)\,u(\tau(i)) \geq \left(\sum_{i=\tau(k)}^{k-1} p(i)\right) u(\tau(k-1)) \geq (\alpha-\varepsilon)u(\tau(k-1))$$

or

$$u(\tau(k)) \geq u(k) + (\alpha-\varepsilon)\;u(\tau(k-1)). \tag{2.23}$$

Combining inequalities (2.22) and (2.23), we obtain

$$u(k) \geq (1-\sqrt{1-\alpha})[u(k) + (\alpha-\varepsilon)u(\tau(k-1))]$$

or

$$\frac{u(\tau(k-1))}{u(k)} \leq \frac{1-\alpha+\sqrt{1-\alpha}}{\alpha(\alpha-\varepsilon)} \tag{2.24}$$

and, for large k,

$$\frac{u(\tau(k))}{u(k+1)} \leq \frac{1-\alpha+\sqrt{1-\alpha}}{\alpha(\alpha-\varepsilon)}.$$

Hence

$$\limsup_{k\to+\infty} \frac{u(\tau(k))}{u(k+1)} \leq \frac{1-\alpha+\sqrt{1-\alpha}}{\alpha(\alpha-\varepsilon)}.$$

The last inequality, for arbitrarily small values of ε, implies (2.14). The proof is complete.

Theorem 2.1 *Assume that* $\tau(k) \leq k$ *and*

$$\sigma(k) = \max\left\{\tau(s) : 1 \leq s \leq k, \quad s \in N\right\}. \tag{2.25}$$

If

$$\limsup_{k \to +\infty} \sum_{i=\sigma(k)}^{k} p(i) > 1, \tag{2.26}$$

then all proper solutions of Eq.(E) *oscillate.*

Proof. Assume, for the sake of contradiction, that $u_0 : N_{n_0} \to (0, +\infty)$ is a positive proper solution of Eq.(E).

Since the function u_0 is nonincreasing and $\sigma(k) = \max\left\{\tau(s) : 1 \leq s \leq k, \; s \in N\right\}$ then, for sufficiently large $k \in N_{n_0}$, u_0 satisfies the following inequality

$$\Delta u_0(k) + p(k)\, u_0(\sigma(k)) \leq 0.$$

Summing up the last inequality from $\sigma(k)$ to k, and using the fact that the function u_0 is nonincreasing and the function σ is nondecreasing, we have

$$u_0(\sigma(k))\left(\sum_{i=\sigma(k)}^{k} p(i) - 1\right) \leq 0.$$

Therefore, for sufficiently large k

$$\sum_{i=\sigma(k)}^{k} p(i) \leq 1,$$

which contradicts (2.26). The proof is complete.

Remark 2.1 *In the special case of Eq.*(E_1) *the above condition* (2.26) *leads to the condition* (1.2) *presented in [6].*

Theorem 2.2 *Assume that*

$$\liminf_{k \to +\infty} \sum_{i=\tau(k)}^{k-1} p(i) = \alpha \in (0, 1] \tag{2.12}$$

and

$$\limsup_{k\to+\infty}\sum_{i=\sigma(k)}^{k}p(i)>1-\left(1-\sqrt{1-\alpha}\right)^2, \tag{2.27}$$

where

$$\sigma(k)=\max\left\{\tau(s):1\le s\le k,\quad s\in N\right\}. \tag{2.25}$$

Then all proper solutions of Eq.(E) *oscillate.*

If, additionally, $p(k)\ge 1-\sqrt{1-\alpha}$ *for large* k, *and*

$$\limsup_{k\to+\infty}\sum_{i=\sigma(k)}^{k}p(i)>1-\alpha\frac{1-\sqrt{1-\alpha}}{\sqrt{1-\alpha}} \tag{2.28}$$

then all proper solutions of Eq.(E) *oscillate.*

Proof. We will first show that

$$\liminf_{k\to+\infty}\sum_{i=\sigma(k)}^{k-1}p(i)=\alpha. \tag{2.29}$$

Indeed, since $\tau(k)\le\sigma(k)$, then by (2.12), it is obvious that

$$\liminf_{k\to+\infty}\sum_{i=\sigma(k)}^{k-1}p(i)\le\liminf_{k\to+\infty}\sum_{i=\tau(k)}^{k-1}p(i)=\alpha. \tag{2.30}$$

Thus, there exists a subsequence $\{k_i\}_{i=1}^{+\infty}$ of natural numbers such that $k_i\uparrow +\infty$ for $i\to+\infty$ and

$$\liminf_{k\to+\infty}\sum_{i=\sigma(k)}^{k-1}p(i)=\lim_{k\to+\infty}\sum_{j=\sigma(k_i)}^{k_i-1}p(j). \tag{2.31}$$

On the other hand, from the definition of the function σ and taking into account that $\lim_{k\to+\infty}\tau(k)=+\infty$ for any k_i $(i=1,2,\dots)$ there exists $k_i'\le k_i$ such that $\sigma(k)=\sigma(k_i)$ since $k_i'\le k\le k_i$, $\lim_{i\to+\infty}k_i'=+\infty$, and $\sigma(k_i')=\tau(k_i')$ $(i=1,2,\dots)$. Thus,

$$\sum_{j=\sigma(k_i)}^{k_i-1}p(j)\ge\sum_{j=\sigma(k_i)}^{k_i'-1}p(j)=\sum_{j=\sigma(k_i')}^{k_i'-1}p(j)=\sum_{j=\tau(k_i')}^{k_i'-1}p(j). \tag{2.32}$$

Combining inequalitites (2.31) and (2.32), we obtain

$$\liminf_{k\to+\infty} \sum_{i=\sigma(k)}^{k-1} p(i) \geq \liminf_{k\to+\infty} \sum_{i=\tau(k)}^{k-1} p(i) = \alpha. \tag{2.33}$$

The last inequality together with (2.30) imply (2.29).

Now assume, for the sake of contradiction, that u is a positive proper solution of Eq.(E). Then, for sufficiently large k, the function u is a positive proper solution of

$$\Delta u(k) + p(k)\, u(\sigma(k)) \leq 0.$$

By Lemma 2.1, the equation

$$\Delta u(k) + p(k)\, u(\sigma(k)) = 0 \tag{2.34}$$

has a positive proper solution $u_* : N_{n_0} \to (0, +\infty)$, where $n_0 \in N$ is sufficiently large.

Since (2.29) is satisfied, inequality (2.13) becomes

$$\limsup_{k\to+\infty} \frac{u_*(\sigma(k))}{u_*(k+1)} \leq \left(\frac{1+\sqrt{1-\alpha}}{\alpha}\right)^2 \tag{2.35}$$

or, if $p(k) \geq 1-\sqrt{1-\alpha}$ for sufficiently large k, then inequality (2.14) becomes

$$\limsup_{k\to+\infty} \frac{u_*(\sigma(k))}{u_*(k+1)} \leq \frac{1-\alpha+\sqrt{1-\alpha}}{\alpha^2}. \tag{2.36}$$

In the case that (2.35) holds, for any $\varepsilon \in (0, (1-\sqrt{1-\alpha})^2)$ and for sufficiently large k, we have

$$u_*(k+1) \geq \big((1-\sqrt{1-\alpha})^2 - \varepsilon\big) u_*(\sigma(k)). \tag{2.37}$$

Now, summing up Eq.(2.34) from $\sigma(k)$ to k, and using the fact that the function u_* is nonincreasing and the function σ is nondecreasing, we have

$$u_*(\sigma(k)) \geq u_*(k+1) + \left(\sum_{i=\sigma(k)}^{k} p(i)\right) u_*(\sigma(k)). \tag{2.38}$$

Combining inequalities (2.38) and (2.37), we obtain

$$u_*(\sigma(k)) \geq \left(\left(1 - \sqrt{1-\alpha}\right)^2 - \varepsilon + \sum_{i=\sigma(k)}^{k} p(i) \right) u_*(\sigma(k)).$$

Hence

$$\limsup_{k\to+\infty} \sum_{i=\sigma(k)}^{k} p(i) \leq 1 - \left(1 - \sqrt{1-\alpha}\right)^2 + \varepsilon,$$

which, for arbitrarily small values of ε, becomes

$$\limsup_{k\to+\infty} \sum_{i=\sigma(k)}^{k} p(i) \leq 1 - \left(1 - \sqrt{1-\alpha}\right)^2.$$

This, contradicts (2.27).

In the case that (2.36) holds, following a similar procedure, we are led to the inequality

$$\limsup_{k\to+\infty} \sum_{i=\sigma(k)}^{k} p(i) \leq 1 - \alpha \frac{1 - \sqrt{1-\alpha}}{\sqrt{1-\alpha}}$$

which contradicts (2.28). The proof is complete.

Remark 2.2 *If $\alpha > 1$, by (2.3), it is obvious, that the conditions of Theorem 2.1 are satisfied and therefore all proper solutions of Eq.(E) oscillate.*

Corollary 2.1 *Assume that*

$$0 < \alpha : \;= \liminf_{k\to+\infty} \sum_{i=k-n}^{k-1} p(i) \leq \left(\frac{n}{n+1}\right)^{n+1}$$

and

$$\limsup_{k\to+\infty} \sum_{i=k-n}^{k} p(i) > 1 - \left(1 - \sqrt{1-\alpha}\right)^2. \tag{2.27$'$}$$

Then all proper solutions of Eq.(E_1) *oscillate.*

If, additionally, for sufficiently large k, $p(k) \geq 1 - \sqrt{1-\alpha}$, and

$$\limsup_{k\to+\infty} \sum_{i=k-n}^{k} p(i) > 1 - \alpha \frac{1 - \sqrt{1-\alpha}}{\sqrt{1-\alpha}}, \tag{2.28$'$}$$

then all proper solutions of Eq.(E_1) *oscillate.*

Now we present an example in which the condition (2.27′) of the above Corollary is satisfied, while none of the conditions (1.2), (1.3), (1.6), (1.7) and (1.9), is satisfied.

Example 1 *Consider the equation*

$$x(k+1) - x(k) + p(k)x(k-12) = 0, \quad k = 0,1,2,...,$$

where

$$p(13k+1) = ... = p(13k+12) = \frac{35}{1200}, \; p(13k+13) = \frac{35}{1200} + \frac{6}{10}, \quad k = 0,1,2,....$$

Here $n = 12$ *and it is easy to see that*

$$\alpha = \liminf_{k\to\infty} \sum_{i=k-12}^{k-1} p(i) = \frac{35}{100} < \left(\frac{12}{13}\right)^{13} \simeq 0.3532$$

$$\limsup_{k\to\infty} \sum_{i=k-12}^{k-1} p(i) = \frac{35}{100} + \frac{6}{10} = 0.950$$

and

$$\limsup_{k\to\infty} \sum_{i=k-12}^{k} p(i) = \frac{35}{1200} + \frac{950}{1000} = 0.9791 > 1 - \left(1 - \sqrt{1-\alpha}\right)^2 \simeq 0.9624.$$

We see that the condition (2.27') of Corollary 2.1 is satisfied and therefore all solutions oscillate. Observe, however, that

$$0.9791 < 1,$$

$$\alpha = 0.35 < \left(\frac{12}{13}\right)^{13} \simeq 0.3532,$$

$$0.950 < 1 - \frac{\alpha^2}{4} \simeq 0.9693,$$

$$0.950 < 1 - \alpha^{12} \simeq 0.9999,$$

and

$$0.950 < 1 - \frac{\alpha^2}{2(2-\alpha)} \simeq 0.9628.$$

Therefore none of the conditions (1.2), (1.3), (1.6), (1.7) and (1.9), is satisfied.

Theorem 2.3 *Assume that $\alpha \in (0,1]$ and there exist $n_0 \in N$ and a function $\widetilde{p} \in L_{\mathrm{loc}}(R_+, R_+)$ such that*

$$t^2\,\widetilde{p}(t) \text{ is nondecreasing function, } \widetilde{p}(i) \le p(i) \text{ for } i \in N_{n_0} \tag{2.39}$$

and

$$\liminf_{k\to+\infty} \int_{\tau(k)-1}^{k-1} \widetilde{p}(s)ds \ge \alpha. \tag{2.40}$$

Then condition (2.27) (*or, if for sufficiently large k, $p(k) \ge 1-\sqrt{1-\alpha}$, condition* (2.28)) *is sufficient for all proper solutions of Eq.*(E) *to oscillate.*

Proof. In view of Lemma 2.1 and Theorem 2.2, to prove Theorem 2.3, it sufficies to show that

$$\liminf_{k\to+\infty} \sum_{i=\tau(k)}^{k-1} p(i) \ge \alpha. \tag{2.41}$$

By (2.39) and (2.40), we have

$$\sum_{i=\tau(k)}^{k-1} p(i) \ge \sum_{i=\tau(k)}^{k-1} \frac{(i-1)i^2}{i}\widetilde{p}(i)\int_{i-1}^{i}\frac{ds}{s^2} \ge \sum_{i=\tau(k)}^{k-1}\frac{i-1}{i}\int_{i-1}^{i}\widetilde{p}(s)\,ds \ge$$

$$\ge \frac{\tau(k)-1}{\tau(k)}\sum_{i=\tau(k)}^{k-1}\int_{i-1}^{i}\widetilde{p}(s)\,ds = \frac{\tau(k)-1}{\tau(k)}\int_{\tau(k)-1}^{k-1}\widetilde{p}(s)\,ds. \tag{2.42}$$

Since $\tau(k) \to +\infty$ for $k \to +\infty$, inequality (2.42), in view of (2.40), implies (2.41). The proof is complete.

Corollary 2.2 *Consider Eq.(E) and let $c \in (0,+\infty)$, $\beta \in (0,1)$, $c\ln\beta \ge -1$ and for large k*

$$p(k) \ge \frac{c}{k}, \quad \tau(k) \le [\beta k],$$

and

$$\limsup_{k\to+\infty} \sum_{i=[\beta k]}^{k} p(i) > 1-\left(1-\sqrt{1-\alpha}\right)^2,$$

where $\alpha = \ln\beta^{-c}$ and $[\beta k]$ denotes the integer part of βk. Then all proper solutions of Eq.(E) *oscillate.*

Proof. Take $\widetilde{p}(t) = \frac{c}{t}$ and $\alpha = \ln \beta^{-c}$. Then it is easily shown that the conditions of Theorem 2.3 are satisfied.

Analogously, if we take $\widetilde{p}(t) = \frac{c}{t \ln t}$, we have the following

Corollary 2.3 *Consider Eq.(E) and let $c \in (0, +\infty)$, $\beta \in (0, 1)$, $c \ln \beta \geq -1$ and for large k*

$$p(k) \geq \frac{c}{k \ln k}, \quad \tau(k) \leq [k^\beta].$$

and

$$\limsup_{k \to +\infty} \sum_{i=[k^\beta]}^{k} p(i) > 1 - \left(1 - \sqrt{1 - \alpha}\right)^2,$$

where $\alpha = \ln \beta^{-c}$ and $[k^\beta]$ denotes the integer part of k^β. Then all proper solutions of Eq.(E) *oscillate.*

References

[1] R. P. Agarwal and P. J. Y. Wong, Advanced topics in difference equations. *Kluwer Academic Publishers*, 1997

[2] G. E. Chatzarakis and I. P. Stavroulakis, Oscillations of First Order Linear Delay Difference Equations. *The Australian J. Math. Analysis and Appl.*, 2006 (in press).

[3] S. S. Cheng and G. Zhang, "Virus" in several discrete oscillation theorems. *Appl. Math. Lett.* **13** (2000), No. 3, 9–13.

[4] Y. Domshlak, A discrete analogue of the Sturm comparison theorem for non-symmetric equations. (Russian) *Dokl. Akad. Nauk Az. SSR* **37** (1981), No. 10, 12–15.

[5] Y. Domshlak, What should be a discrete version of the Chanturia-Koplatadze lemma? *Funct. Differ. Equ.* **6** (1999), No. 3-4, 299–304.

[6] L. Erbe and B. G. Zhang, Oscillation of discrete analogues of delay equations. *Differ. Integral Equ.* **2** (1989), No. 3, 300–309.

[7] I. Győri and G. Ladas, Oscillation theory of delay differential equations with applications. *Clarendon Press. Oxford*, 1991.

[8] R. G. Koplatadze and T. A. Chanturia, On oscillatory and monotone solutions of first order differential equations with retarded argument. (Russian) *Differentsial'nye Uravneniya* **18** (1982), No. 8, 1463–1465.

[9] G. Ladas, C. Philos and Y. Sficas, Sharp conditions for the oscillation of delay difference equations. *J. Appl. Math. Simulation* **2** (1989), No. 2, 101–112.

[10] B. Lalli and B. G. Zhang, Oscillation of difference equations. *Colloquium Math.* **65** (1993), No. 1, 25–32.

[11] Ch. G. Philos, I.K. Pournaras and I. P. Stavroulakis, Sufficient conditions for the oscillation of delay difference equations, *J. Difference Equ. Appl.*, **10** (2004), 419-435.

[12] Y. G. Sficas and I. P. Stavroulakis, Oscillation criteria for first-order delay equations. *Bull. Lond. Math. Soc.* **35** (2003), No. 2, 239–246.

[13] J. H. Shen and I. P. Stavroulakis, Oscillation criteria for delay difference equations. *Electron. J. Differ. Equ.* **2001** (2001), No. 10, 1–15.

[14] I. P. Stavroulakis, Oscillations of delay difference equations. *Comput. Math. Appl.* **29** (1995), No. 7, 83–88.

[15] I. P. Stavroulakis, Oscillation criteria for first order delay difference equations. *Mediterr. J. Math.* **1** (2004), No. 2, 231–240.

[16] J. S. Yu, B. G. Zhang and X. Z. Qian, Oscillations of delay difference equations with oscillating coefficients. *J. Math. Anal. Appl.* **177** (1993), No. 2, 432–444.

OSCILLATIONS OF FIRST ORDER IMPULSIVE DELAY DIFFERENTIAL EQUATIONS

by

M.K. Grammatikopoulos[(1)], M.B.Dimitrova[(2)] and V.I. Donev[(2)]

First order impulsive delay differential equations are studied, where the fixed moments of impulsive effect (the jump points) are considered as up-jump points. Sufficient integral conditions for all solutions of these type of equations to be oscillatory are established.

Key words: Oscillation, impulsive differential equations, up-jump points, retarded argument.

AMS (MOS) subject classifications: Primary 34A37, 34K11.

1. Introduction

Impulsive delay differential equations can model various processes and phenomena which depend on their prehistory and are subject to short-time disturbances. Such processes occur in the theory of optimal control, theoretical physics, population dynamics, pharmacokinetics, biotechnologies, industrial robotics, economics, etc. Starting from the work of Mil'man and Myshkis [9], in resent years there has been much current interest in studying of impulsive differential equations. Among numerous publications, we choose to refer to [1]-[12].

Consider the first order impulsive delay differential equations of the form

$$x'(t) + q(t)x(t) + p(t)x(t-h) = 0, \; t \neq \tau_k \qquad (E_1)$$

with the impulsive condition

$$\Delta x(\tau_k) = x(\tau_k + 0) - x(\tau_k - 0) = I_k(x(\tau_k), x(\tau_k - h)), \; k \in N$$

and with the initial condition

$$x(t) = \varphi(t), \; -h \leq t \leq 0; \; \varphi \in C([-h, 0]; R).$$

Here the delay $h > 0$ is a constant and $\tau_k \in (0, +\infty)$, $k \in N$ are fixed moments of impulsive effect (the jump points), which we characterize as *down-jumps* when $\Delta x(\tau_k) < 0$, $k \in N$ and as *up-jumps* when $\Delta x(\tau_k) > 0$, $k \in N$).

Denote by $PC(R, R)$ the set of all piecewise continuous on the intervals $(\tau_k, \tau_{k+1}], k \in N$ functions $u: R \to R$ which at the jump points $\tau_k, k \in N$ are continuous from the left, i.e. $u(\tau_k - 0) = \lim_{t \to \tau_k - 0} u(t) = u(\tau_k)$, and may have discontinuities of first kind at the jump points $\tau_k, k \in N$.

We also denote by $i[\tau_0, t)$ the number of fixed jump points $\tau_k \in [\tau_0, t),\ k \in N$, for $t > \tau_0$.

We clarify that

$$i[\tau_0, t) = \begin{cases} 0, \text{ for } \ \tau \in [\tau_0, \tau_1), \\ 0, \text{ for } \ \tau \in [\tau_1, \tau_2), \\ \ldots.. \\ k, \text{ for } \ \tau \in [\tau_k, \tau_{k+1}), \ k \in N. \end{cases}$$

Our aim is to establish sufficient conditions under which the equation $(E_1)$is oscillatory. In order to obtain our results, we need the following

Lemma 1 *Let $\tau_k,\ k \in N$ be fixed moments of impulsive effect (the jump points) with the property*

$$0 < \tau_1 < \tau_2 < \ldots < \tau_k < \ldots, \quad \lim_{t \to +\infty} \tau_k = +\infty$$

Then for every fixed $h > 0$ and for every $t \in [h, +\infty)$

$$M = \max_{t \in [h, +\infty]} i[t-h, t) < +\infty,$$

i.e. the number of the fixed moments of impulse effect $\tau_k \in [t-h, t),\ k \in N$ is finite.

Proof. Since, by the properties of the sequence $\tau_k,\ k \in N$, it follows that $\limsup_{k \to +\infty} \tau_k = +\infty$, we conclude that the only accumulation point of this sequence is that $+\infty$. Accordingly, for any number $T > 0$ there is an $n_0 \in N$ such that for every $n \geq n_0$ we have $\tau_n > T$. That is, the number of the fixed jump points in every finite interval of the form $[T-h, T)$ is a finite number. The proof of the lemma is complete.

Throughout this paper, unless otherwise mentioned, we will assume that the following hypotheses are satisfied:

(H_1) $0 < \tau_1 < \tau_2 < \ldots < \tau_k < \ldots, \lim_{k \to \infty} \tau_k = +\infty$ and

$0 < \min\{\tau_{k+1} - \tau_k\} \leq \max\{\tau_{k+1} - \tau_k\} < +\infty, k \in N$;

(H_2) The function $p : PC([0, \infty), (0, \infty))$ (resp. the function $q : PC([-h, \infty), R)$ with points of discontinuity $\tau_k, k \in N$, where it is continuous from the left, i.e. $p(\tau_k - 0) = p(\tau_k), k \in N$ (resp. $q(\tau_k - 0) = q(\tau_k), k \in N$);

(H_3) The function $I_k \in C(R^2; R)$ for all $v \in R$ and $k \in N$ has the following sign property

$$uI_k(u,v) > 0 \ for \ u \neq 0.$$

Moreover, the following notions will be used throughout this paper.

A continuous real valued function u defined on an interval of the form $[a, +\infty)$ eventually has some property if there is a number $b \geq a$ such that u has this property on the interval $[b, +\infty)$.

A real valued function u piecewise continuous on the set $[-h, \infty) \setminus \{\tau_k\}_{k=1}^{\infty}$ and continuous from the left at the jump points $\tau_k, k \in N$ with initial function $\varphi \in C([-h, 0]; R)$ is said to be a solution to $Eq.(E_1)$ if $u(t) = \varphi(t)$ for every $t \in [-h, 0]$ and $u(t)$ satisfies $Eq.(E_1)$ for all sufficiently large $t \geq 0$.

Without other mention, we will assume throughout that every solution $u(t)$ of $Eq.(E_1)$, that is under consideration here, is continuable to the right and is nontrivial. That is, $u(t)$ is defined on some ray of the form $[T_u, +\infty)$ and

$$\sup \{|u(t)| : \ t \geq T\} > 0 \text{ for each } T \geq T_u.$$

Such a solution is called a *regular solution* of $Eq.(E_1)$.

As usual, a regular solution of $Eq.(E_1)$ is called *nonoscillatory* if it is *eventually of constant sign* , i.e. if it is *eventually positive* or *eventually negative*. Otherwise, it is called *oscillatory*. Furthermore, $Eq.(E_1)$ is called *oscillatory* if every its regular solution is oscillatory. Otherwise, it is called *non-oscillatory*.

2. Main results

In order to achieve our goal, we begin our investigation with a special case of $Eq.(E_1)$. Namely, we consider the first order impulsive delay differential equation

$$x'(t) + p(t)x(t-h) = 0, \ t \neq \tau_k \tag{E_2}$$

with the impulsive condition

$$\Delta x(\tau_k) = x(\tau_k + 0) - x(\tau_k - 0) = I_k(x(\tau_k), x(\tau_k - h)), \ k \in N$$

and with the initial condition

$$x(t) = \varphi(t), -h \leq t \leq 0; \varphi \in C([-h, 0]; R),$$

which results from $Eq.(E_1)$ in the case where the function q is identically zero on the interval $[-h, \infty)$.

We start with the following

Lemma 2 *Let $x(t)$ be a non-oscillatory solution of $Eq.(E_2)$ and assume that the hypotheses (H_1) - (H_3) are satisfied. Suppose also that:*

(C_1) There is a positive constant L such that $\mid I_k(u,v) \mid \leq L \mid u \mid$ *for* $u \neq 0, v \in R,\ k \in N$ *and*

(C_2) $\liminf\limits_{t\to\infty} \int\limits_{t-h}^{t} p(s)ds \geq \frac{1}{e}(1+L)^{M},\ M = \max\limits_{t\in[h,+\infty]} i[t-h,t).$

Then $w(t)=\frac{x(t-h)}{x(t)}$ *is an eventually bounded function.*

Proof. Since the negative of a solution of $Eq.(E_2)$ is again a solution of $Eq.(E_2)$, it suffices to prove the lemma in the case of an eventually positive solution. So, suppose that $x(t)$ is an eventually positive solution of $Eq.(E_2)$. That is, there is a $t_0 \geq 0$ such that $x(t) > 0$ for $t \geq t_0$, while $x(t-h) > 0$ for $t \geq t_0 + h = t_1$. Therefore, from the impulsive condition of $Eq.(E_2)$, in view of the hypotheses (H_2) and (H_3), it follows that $x'(t) < 0$ and $\Delta x(\tau_k) > 0$ for $t, \tau_k \geq t_1, k \in N$ .Thus, $x(t)$ is a decreasing function on every interval $(\tau_k, \tau_{k+1}], \tau_k \geq t_1$, $k \in N$ and it has discontinuities of the first kind at the points of impulse effect $\tau_k \in R_+$, $k \in N$, considered as up-jumps.

Remark that, from the impulsive condition of $Eq.(E_2)$, using (C_1), we find

$$\frac{x(\tau_k+0)}{x(\tau_k)} = 1 + \frac{I_k(x(\tau_k), x(\tau_k - h))}{x(\tau_k)} \leq 1 + \frac{Lx(\tau_k)}{x(\tau_k)} = 1 + L,\ k \in N. \tag{1}$$

In order to prove our lemma, consider now the interval of integration $(t, t+\frac{h}{2})$, $t \geq t_1$ of $Eq.(E_2)$ and the number of the discontinuity points $i[t, t+\frac{h}{2})$ in it. Depending on the location of the points $t-h$ and t with respect to the jump points τ_k, $k \in N$, we distinguish the following five possible cases.

Case 1. When $t-h,\ t \in (\tau_k, \tau_{k+1}]$, $k \in N$ and exactly one of the following holds: either $i[t, t+\frac{h}{2}) = 0$ or else $i[t, t+\frac{h}{2}) = 1$.

Remark that, if $i[t, t+\frac{h}{2}) = 1$, then the only possible point of discontinuity in the interval $(t, t+\frac{h}{2})$ is the point τ_{k+1}.

In this case, integrating $Eq.(E_2)$ from t to $t+\frac{h}{2}$, $t \geq t_1 + \frac{h}{2}$, we obtain

$$x(t+\frac{h}{2}) - x(t) - \sum_{t\leq\tau_n\leq t+\frac{h}{2}} I_n(x(\tau_n), x(\tau_n - h)) + \int\limits_{t}^{t+\frac{h}{2}} p(s)x(s-h)ds = 0,$$

and hence we find

$$x(t) + \sum_{t\leq\tau_n\leq t+\frac{h}{2}} I_n(x(\tau_n), x(\tau_n - h)) \geq \int\limits_{t}^{t+\frac{h}{2}} p(s)x(s-h)ds \geq \int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)x(s-h)ds,$$

where $t_{M_1} = \max(t, max_{t\leq\tau_n\leq t+\frac{h}{2}}\tau_n)$. Observe that, if $i[t,t+\frac{h}{2}) = 0$, then $t_{M_1} = t$, while if $i[t,t+\frac{h}{2}) = 1$, then $t_{M_1} = \tau_{k+1}$. Now, applying the assumption *(C_1)* to the last inequality, we see that

$$x(t) + L \sum_{t\leq\tau_n\leq t+\frac{h}{2}} x(\tau_n) \geq x(t-\frac{h}{2}) \int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)ds$$

which implies that

$$x(t) + Lx(t)i[t,t+\frac{h}{2}) \geq x(t) + L \sum_{t\leq\tau_n\leq t+\frac{h}{2}} x(\tau_n) \geq x(t-\frac{h}{2}) \int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)ds,$$

and hence we get

$$\frac{x(t-\frac{h}{2})}{x(t)} \leq \frac{1+Li[t,t+\frac{h}{2})}{\int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)ds}. \tag{2}$$

Next, integrating $Eq.(E_2)$ from $t-\frac{h}{2}$ to t, $t-\frac{h}{2} \geq t_1$, we see that

$$x(t) - x(t-\frac{h}{2}) + \int\limits_{t-\frac{h}{2}}^{t} p(s)x(s-h)ds = 0$$

from where we obtain

$$x(t-\frac{h}{2}) \geq x(s-h) \int\limits_{t-\frac{h}{2}}^{t} p(s)\ ds$$

and so we see that

$$\frac{x(t-h)}{x(t-\frac{h}{2})} \leq \frac{1}{\int\limits_{t-\frac{h}{2}}^{t} p(s)ds} \tag{3}$$

In view of (2) and (3) and using the decreasing character of the function $x(t)$ on every interval $(\tau_k, \tau_{k+1}], \tau_k \geq t_1,\ k \in N$, we easily conclude that

$$1 < \frac{x(t-h)}{x(t)} \leq \frac{1+Li[t,t+\frac{h}{2})}{\int\limits_{t-\frac{h}{2}}^{t} p(s)ds \int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)ds}. \tag{4}$$

This shows that the function $w(t)$, $t \geq t_1$ is bounded and proves our assertion in *Case 1.*

Case 2. When $t-h, t \in (\tau_k, \tau_{k+1}],\ k \in N$ *and* $i[t,t+\frac{h}{2}) > 1$.

In this case, it is always possible to choose a sequence of points $\xi_l \in (\tau_k, \tau_{k+1}],\ l = 1,2,...,r$ with $\xi_1 = t-h$ and $\xi_r = t$, where for $h_{\xi_l} = \xi_l - \xi_{l-1}, l = 2,...,r$, as in *Case 1*, exactly one of the following holds: either $i[\xi_l, \xi_l + \frac{1}{2}h_{\xi_l}) = 0$ or else $i[\xi_l, \xi_l + \frac{1}{2}h_{\xi_l}) = 1$. Then, for each pair ξ_{l-1}, ξ_l, $l = 2, ..., r$, as in the proof of *Case 1*, we obtain

$$1 < \frac{x(\xi_{l-1})}{x(\xi_l)} \leq \frac{1 + Li[\xi_l, \xi_l + \frac{1}{2}h_{\xi_l})}{\int\limits_{\xi_l - \frac{1}{2}h_{\xi_l}}^{\xi_l} p(s)ds \int\limits_{t_{M_{2,l}}}^{\xi_l + \frac{1}{2}h_{\xi_l}} p(s)ds}, \tag{5}$$

where $t_{M_{2,l}} = \max(\xi_l, max_{\xi_l \leq \tau_n \leq \xi_l + \frac{1}{2}h_{\xi_l}} \tau_n)$. In view of (5), we easily conclude that

$$1 \leq \frac{x(\xi_1)}{x(\xi_2)} \frac{x(\xi_2)}{x(\xi_3)} ... \frac{x(\xi_{r-1})}{x(\xi_r)} = \frac{x(t-h)}{x(t)} \leq \prod_{1 \leq l \leq r} \frac{1 + Li[\xi_l, \xi_l + \frac{1}{2}h_{\xi_l})}{\int\limits_{\xi_l - \frac{1}{2}h_{\xi_l}}^{\xi_l} p(s)ds \int\limits_{t_{M_{2,l}}}^{\xi_l + \frac{1}{2}h_{\xi_l}} p(s)ds}$$

which proves our assertion in *Case 2*.

Case 3. *When $t \in (\tau_{k+1}, \tau_{k+2}]$, $t - h \in (\tau_k, \tau_{k+1}]$, $k \in N$ and exactly one of the following holds: either $i[t, t + \frac{h}{2}) = 0$ or else $i[t, t + \frac{h}{2}) = 1$.*

Remark that, if $i[t, t + \frac{h}{2}) = 1$, then the only possible point of discontinuity in the interval $(t, t + \frac{h}{2})$ is the point τ_{k+2}.

In this case, because of the up-jump at the point τ_{k+1} ($\Delta x(\tau_{k+1}) > 0$ for $\tau_{k+1} \geq t_1$), depending on the value of $h > 0$ it is possible to have either (a) $x(t-h) \leq x(t)$ or (b) $x(t-h) \geq x(t)$.

If (a) holds, then (1) implies that

$$\frac{1}{1+L} \leq \frac{x(\tau_{k+1})}{x(\tau_{k+1}+0)} \leq \frac{x(t-h)}{x(t)} \leq 1 \tag{6}$$

which proves our claim in this case.

Assume now that (b) holds. In this case integrating $Eq.(E_2)$ from t to $t + \frac{h}{2}$, $t \geq t_1 + \frac{h}{2}$, and then from $t - \frac{h}{2}$ to t, $t - \frac{h}{2} \geq t_1$, as in the proof of *Case 1*, we derive (2) and

$$\frac{x(t-h)}{x(t-\frac{h}{2})} \leq \frac{1 + Li[t - \frac{h}{2}, t)}{\int\limits_{t_{M_3}}^{t} p(s)ds} \tag{7}$$

respectively, where $t_{M_3} = \max(t - \frac{h}{2}, max_{t - \frac{h}{2} \leq \tau_n \leq t} \tau_n)$. Remark that, if $i[t - \frac{h}{2}, t) = 0$, then $t_{M_3} = t - \frac{h}{2}$, while if $i[t - \frac{h}{2}, t) = 1$, then $t_{M_3} = \tau_{k+1}$.

By (2) and (7), taking into account the fact that $x(t-h) \geq x(t)$, we conclude that

$$1 \leq \frac{x(t-h)}{x(t)} \leq \frac{(1 + Li[t - \frac{h}{2}, t))(1 + Li[t, t + \frac{h}{2}))}{\int\limits_{t_{M_3}}^{t} p(s)ds \int\limits_{t_{M_1}}^{t+\frac{h}{2}} p(s)ds},$$

which is similar to (4) and proves our assertion in *Case 3*.

Case 4. *When* $t \in (\tau_{k+1}, \tau_{k+2}],\ \ t-h \in (\tau_k, \tau_{k+1}],\ k \in N$ *and* $i[t, t+\frac{h}{2}) > 1$.

Here, as in *Case 3,* it is possible to have either (a) $x(t-h) \le x(t)$ or (b) $x(t-h) \ge x(t)$. If (a) holds, then we derive (6) which proves our assertion. So, assume that (b) holds. In this case, it is always possible to choose a sequence $\eta_i \in (\tau_k, \tau_{k+1}],\ i = 1, 2, ..., s$ with $\eta_0 = t-h$, and such that for $h_{\eta_i} = \eta_i - \eta_{i-1},\ i = 1, 2, ..., s$, as in *Case 1,* exactly one of the following to be hold: either $i[\eta_i, \eta_i + \frac{1}{2}h_{\eta_i}) = 0$ or else $i[\eta_i, \eta_i + \frac{1}{2}h_{\eta_i}) = 1$. Then, as in the proof of *Case 1,* for each pair η_{i-1} and $\eta_i,\ i = 1, 2, ..., s$ we obtain

$$1 < \frac{x(\eta_{i-1})}{x(\eta_i)} \le \frac{1 + Li[\eta_i, \eta_i + \frac{1}{2}h_{\eta_i})}{\int\limits_{\eta_i - \frac{1}{2}h_{\eta_i}}^{\eta_i} p(s)ds \int\limits_{t_{M_{4,i}}}^{\eta_i + \frac{1}{2}h_{\eta_i}} p(s)ds} \tag{8}$$

where $t_{M_{4,i}} = \max(\eta_i, max_{\eta_i \le \tau_n \le \eta_i + \frac{1}{2}h_{\eta_i}} \tau_n)$.

Since $x(\eta_s) \in (\tau_k, \tau_{k+1}]$, we may choose a point $\xi_1 < t$ such that $\xi_1 \in (\tau_{k+1}, \tau_{k+2}]$, and, as in *Case 3,* η_s and ξ_1 for $h_1 = \xi_1 - \eta_s$ to satisfy exactly one of the following: either $i[\xi_1, \xi_1 + \frac{1}{2}h_1) = 0$ or else $i[\xi_1, \xi_1 + \frac{1}{2}h_1) = 1$.Then, for the pair η_s and ξ_1 , as in the proof of *Case 3,* we obtain

$$1 < \frac{x(\eta_s)}{x(\xi_1)} \le \frac{(1 + Li[\xi_1 - \frac{1}{2}h_1, \xi_1))(1 + Li[\xi_1, \xi_1 + \frac{1}{2}h_1))}{\int\limits_{t_{M_{4a}}}^{t} p(s)ds \int\limits_{t_{M_{4b}}}^{\xi_1 + \frac{1}{2}h_1} p(s)ds} = L_{\xi_1}(t) \tag{9}$$

where $t_{M_{4a}} = \max(\xi_1 - \frac{h_1}{2}, max_{\xi_1 - \frac{1}{2}h_1 \le \tau_n \le \xi_1} \tau_n)$, $t_{M_{4b}} = \max(\xi_1, max_{\xi_1 \le \tau_n \le \xi_1 + \frac{1}{2}h_1} \tau_n)$.

Now, in view of (8) and (9), we conclude that

$$1 \le \frac{x(\eta_0)}{x(\eta_1)} \frac{x(\eta_1)}{x(\eta_2)} ... \frac{x(\eta_s)}{x(\xi_1)} = \frac{x(\eta_0)}{x(\xi_1)} = \frac{x(t-h)}{x(\xi_1)} \le L_{\xi_1}(t) \prod_{1 \le i \le s} \frac{1 + Li[\eta_i, \eta_i + \frac{1}{2}h_{\eta_i})}{\int\limits_{\eta_i - \frac{1}{2}h_{\eta_i}}^{\eta_i} p(s)ds \int\limits_{t_{M_{4,i}}}^{\eta_i + \frac{1}{2}h_{\eta_i}} p(s)ds},$$

i.e. the function $\frac{x(t-h)}{x(\xi_1)}$ is bounded.

Finally, since the points ξ_1 and t with $\xi_1 < t$ belong to the same interval $(\tau_{k+1}, \tau_{k+2}]$, applying *Cases 1* or *Cases 2,* we prove that the function $\frac{x(\xi_1)}{x(t)}$ is also bounded.

So, from the above observation, it follows that the function $\frac{x(t-h)}{x(\xi_1)} \frac{x(\xi_1)}{x(t)} = \frac{x(t-h)}{x(t)} = w(t)$ for $t \ge t_1$ is bounded.

Case 5. *When* $t \in (\tau_{k+1}, \tau_{k+2}]$, *while* $t - h \in (\tau_{k-m}, \tau_{(k-m)+1}]$, $k \in N$ *for some fixed* $m \in \{1, 2, ..., M\}$, *where* $M = \max\limits_{t \in [h, +\infty]} i[t-h, t)$.

In this case for some fixed $m \in \{1, 2, 3, ...M\}$ we see that

$$\tau_{k-m} < t-h < \tau_{(k-m)+1)} < \tau_{(k-m)+2)} < ...\tau_k < \tau_{k+1} < t < \tau_{k+2},\ k \in N$$

Let $\xi_0 = t-h$. Let also $\xi_j \in (\tau_{(k-m)+j}, \tau_{(k-m)+j+1}]$, $j = 1,2,...,m+1$ be a sequence of points with $\xi_{m+1} = t$, for which exactly one of the previous cases holds. Then, for each pair ξ_{j-1}, ξ_j, $j = 1,2,...,m+1$, as in the proofs of the previous cases considered above, we derive that each of the functions $\frac{x(\xi_{j-1})}{x(\xi_j)}, j = 1,2,...,m+1$ is bounded. Therefore, the function $\frac{x(\xi_0)}{x(\xi_1)}\frac{x(\xi_1)}{x(\xi_2)}\cdots\frac{x(\xi_{m-1})}{x(\xi_m)}\frac{x(\xi_m)}{x(\xi_{m+1})} = \frac{x(t-h)}{x(t)} = w(t)$ for $t \geq t_1$ is also bounded.

The proof of the lemma is complete.

Now we state our first theorem which ensure that all solutions of $Eq.(E_2)$ are oscillatory.

Theorem 1 *Assume that the hypotheses $(H_1) - (H_3)$ are satisfied. Suppose also that:*

(C_1) There is a positive constant L such that $| I_k(u,v) | \leq L | u |$ *for* $u \neq 0, v \in R$, $k \in N$

and

(C_2) $\liminf_{t\to\infty} \int_{t-h}^{t} p(s)ds \geq \frac{1}{e}(1+L)^{M}$, $M = \max_{t\in[h,+\infty]} i[t-h,t)$.

Then the equation (E_2) is oscillatory.

Proof. As in the proof of Lemma 2, we consider an eventually positive solution $x(t)$ of $Eq.(E_2)$ and a $t_1 \geq t_0 + h > 0$ such that $x(t) > 0$ and $x(t-h) > 0$ for $t \geq t_1$. Then, again as in the proof of Lemma 2, from the impulsive condition of $Eq.(E_2)$, using (C_1), we find

$$\frac{x(\tau_k+0)}{x(\tau_k)} = 1 + \frac{I_k(x(\tau_k), x(\tau_k-h))}{x(\tau_k)} \leq 1 + \frac{Lx(\tau_k)}{x(\tau_k)} = 1+L, \; k \in N. \tag{1}$$

Next, divide $Eq.(E_2)$ by $x(t)$, $t \geq t_1$ and integrate from $t-h$ to t to derive

$$\ln\frac{x(t-h)}{x(t)} + \sum_{t-h\leq\tau_k<t} \ln\frac{x(\tau_k+0)}{x(\tau_k)} = \int_{t-h}^{t} p(s)\frac{x(s-h)}{x(s)}ds,$$

where, by Lemma 2, $w(t) = \frac{x(t-h)}{x(t)}$, $t \geq t_1$ is a bounded function. From the above expression, in view of (1), we find

$$\ln w(t)(1+L)^{M} \geq \ln[w(t) \prod_{t-h\leq\tau_k<t} (1+L)] > w_l \int_{t-h}^{t} p(s)ds \tag{10}$$

where

$$w_l = \liminf_{t\to\infty} w(t), \qquad t \geq t_1.$$

Clearly, (10) implies that

$$\liminf_{t\to\infty} \int_{t-h}^{t} p(s)ds < \frac{1}{e}(1+L)^{M},$$

which contradicts *(C_2)*. The proof of the theorem is complete.

As an immediate consequence of Theorem 1, we have the following

Corollary 1 *Suppose that all assumptions of Theorem 1 are satisfied. Then the corresponding to the equation* (E_2)*:*

(a) inequality

$$x'(t) + p(t)x(t-h) \leq 0, \ t \neq \tau_k \tag{$N_{2,\leq}$}$$

$$\Delta x(\tau_k) \leq I_k(x(\tau_k), x(\tau_k - h)), \ k \in N$$

has no eventually positive solutions;

(b) inequality

$$x'(t) + p(t)x(t-h) \geq 0, \ t \neq \tau_k \tag{$N_{2,\geq}$}$$

$$\Delta x(\tau_k) \geq I_k(x(\tau_k), x(\tau_k - h)), \ k \in N$$

has no eventually negative solutions.

The proof of Corollary 1 is similar to the proof of Theorem 1 and so it is omitted.

Our next result concerns the oscillatory character of $Eq.(E_1)$. More precisely, we establish the following

Theorem 2 *Assume that the hypotheses (H_1) – (H_3) and (C_1) are satisfied. Suppose also that*

$$(C_3) \ \liminf_{t\to\infty} \int_{t-h}^{t} p(s)\exp(\int_{s-h}^{s} q(u)du)ds > \frac{1}{e}(1+L)^{\ M}.$$

Then the equation (E_1) *is oscillatory.*

Proof. Since the negative of a solution of $Eq.(E_1)$ is again a solution of $Eq.(E_1)$, it suffices to prove the theorem in the case of an eventually positive solution. So, suppose that $x(t)$ is an eventually positive solution of $Eq.(E_1)$. That is, there is a $t_0 \geq 0$ such that $x(t) > 0$ for $t \geq t_0$, while $x(t-h) > 0$ for $t \geq t_0 + h = t_1$. Set

$$x(t) = z(t)exp(\int_{0}^{t} q(s)ds), \ t \geq t_1. \tag{11}$$

Substituting (11) into $Eq.(E_1)$, we obtain

$$z'(t) + p_1(t)z(t-h) = 0, \ t \neq \tau_k; \tag{12}$$

with the impulsive condition

$$\Delta z(\tau_k) = J_k(z(\tau_k), z(\tau_k - h)), \ k \in N$$

where

$$p_1(t) = p(t) \ exp(\int_{t-h}^{t} q(s)ds), \ t \ \geq t_1$$

and

$$J_k(z(\tau_k), z(\tau_k-h)) = I_k(z(\tau_k)exp(-\int_0^{\tau_k} q(s)ds) \, , z(\tau_k-h)exp[-(\int_0^{\tau_k-h} q(s)ds)exp(\int_0^{\tau_k} q(s)ds)]), k \in N.$$

Since $Eq.(12)$ is of the form of $Eq.(E_2)$ and the functions p_1 and J_k, $k \in N$ satisfy the assumptions of Theorem 1, the conclusion of Theorem 2 is obvious.

Theorem 2 furnish the following

Corollary 2 *Suppose that all assumptions of Theorem 2 are satisfied. Then the corresponding to the equation* (E_1)*:*

(a) inequality

$$x'(t) + q(t)x(t) + p(t)x(t-h) \leq 0, \ t \neq \tau_k \tag{$N_{1,\leq}$}$$

$$\Delta x(\tau_k) \leq I_k(x(\tau_k), x(\tau_k - h)), \ k \in N$$

has no eventually positive solutions;

(b) inequality

$$x'(t) + q(t)x(t) + p(t)x(t-h) \geq 0, \ t \neq \tau_k \tag{$N_{1,\geq}$}$$

$$\Delta x(\tau_k) \geq I_k(x(\tau_k), x(\tau_k - h)), \ k \in N$$

has no eventually negative solutions.

The proof of Corollary 2 is similar to that of Theorem 2 and therefore it is omitted.

3. Examples

In order to illustrate the obtained results, we offer the following two examples .

Example 1 *Consider the impulsive delay differential equation*

$$x'(t) + \frac{5}{4}x(t-1) = 0, \; t \neq \tau_k, \; k \in N,$$

$$\Delta x(\tau_k) = \frac{1}{2}x(\tau_k) + x(\tau_k - 1), \; k \in N,$$

where $h = 1$ *and* $\tau_{k+1} - \tau_k = 1$. *In this case we have* $M = \max_{t \in [h, +\infty]} i[t-h, t) = 1$ *and especially for the assumptions* (C_1) *and* (C_2) *it is fulfilled*

$$\liminf_{t \to \infty} \int_{t-h}^{t} p(s)ds = \frac{5}{4} \geq \frac{1}{e}(1+L)^{\;M} \approx 1.1.$$

when

$$|\frac{1}{2}x(\tau_k) + x(\tau_k - 1)| \;\leq L|x(\tau_k)|, \; k \in N \;\; for \; L \leq 2.$$

So, the assumptions of Theorem 1 are satisfied. Therefore, by Theorem 1, all solutions of the above equation are oscillatory. For example, the function

$$x(t) = e^{-\lambda_* t} A^{i[\tau_0, t)} \; with \; the \; initial \; function \; \varphi(t) = e^{-\lambda_* . t}, t \in [\tau_0 - 1, \tau_0],$$

where $\lambda_* = -1.9834$ *and* $A = -0.087$ *is an oscillatory solution of this equation.*

Example 2 *Consider the impulsive retarded differential equation*

$$x'(t) + \frac{1}{4}x(t-1) = 0, \; t \neq \tau_k, \; k \in N,$$

$$\Delta x(\tau_k) = -\frac{2}{10}x(\tau_k) + \frac{1}{10}x(\tau_k - 1), \; k \in N,$$

with $h = 1$ *and* $\tau_{k+1} - \tau_k = 1$. *In this case we have* $M = \max_{t \in [h, +\infty]} i[t-h, t) = 1$ *and it is easy to check that the assumption* (C_2) *is not satisfied, i.e.*

$$\liminf_{t \to \infty} \int_{t-h}^{t} p(s)ds = \frac{1}{4} < \frac{1}{e}(1+L)^{\;M} \;\approx\; 0.405$$

when

$$| - \frac{2}{10}x(\tau_k) + \frac{1}{10}x(\tau_k - 1)| \;\leq L|x(\tau_k)|, \; k \in N \;\; for \; L \leq 0.1.$$

Hence, the above equation is non-oscillatory. That means that among its solutions at least one is non-oscillatory. In fact, the function

$$x(t) = e^{-\lambda_* t} A^{i[\tau_0, t)} \; with \; the \; initial \; function \; \varphi(t) = e^{-\lambda_* . t}, t \in [\tau_0 - 1, \tau_0], \tau_0 > 0$$

where $\lambda_* = 0.385$ *and* $A = 0.954$, *is a non-oscillatory solution of this equation. Remark that the above equation admits also oscillatory solutions. Such a solution is the function* $x(t) = e^{-\lambda_* t} A^{i[\tau_0, t)}$, *where* $\lambda_* = -2.04$ *and* $A = -0.016$.

References

[1] BAINOV D.D., SIMEONOV P.S., *Systems with Impuls Effect: Stability, Theory and Applications.* Ellis Horwood, Chichester, 1989.

[2] BAINOV D.D., SIMEONOV P.S., *Theory of Impulsive Differential Equations: Periodic solutions and Applications.* Longman, Harlow, 1993.

[3] DIMITROVA M.B., *Oscillation Criteria for Impulsive Differential Equations of First Order with Retarded Argument* , Southeast Asian Bull. of Math., **23**(1999), 599-605.

[4] DIMITROVA M.B., *Criteria for oscillation of Impulsive Differential Equations of first order with deviating arguments,* Kyungpook Math.J.,**40**(2000), 29-37.

[5] GOPALSAMY K., ZHANG B.G., *On delay differential equations with impulses*, J. Math. Anal. Appl., **139** (1989), No. 1, 110–122.

[6] GYORI I., LADAS G., *Oscillation Theory of Delay Differential Equations with Applications*, Clarendon Press, Oxford, 1991.

[7] KRUGER-THIMER E., *Formal Theory of drug dosage regiments,* I. J. Theo. Biol. **13**(1966).

[8] LAKSMIKANTHAM V., BAINOV D.D., SIMEONOV P.S., *Theory of Impulsive Differential Equations.* World Scientific Publishing Company, Singapore, 1989.

[9] MIL'MAN V.D., MYSHKIS A.D., *On the Stability of Motio in the Presence of Impuls (Russian)*, Sibirian Math. J. **1**(1960), 233-237.

[10] SAMOILENKO A.M., PERESTYUK N.A., *Differential Equations with Impulse Effect (Russian).* Vishcha Shkola, Kiev, 1987.

[11] YAN J., ZHAO A., *Oscillation and stability of linear impulsive delay differential equations,* J.Math.Anal.Appl. **227**(1998), 187-194.

[12] ZHANG, A. ZHAO, J. YAN, *Oscillation criteria for Impulsive Delay Differential Equations,* J.Math.Anal.Appl. **205**(1997), 461-470.

(1)Department of Mathematics, University of Ioannina,

451 10 Ioannina, Greece

(2)Department of Mathematics, Technical University of Sliven,

8800 Sliven, Bulgaria

OSCILLATIONS OF FIRST ORDER NEUTRAL IMPULSIVE DELAY DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

by

M.K. Grammatikopoulos[(1)], M.B. Dimitrova[(2)] and V.I. Donev[(2)]

Abstract. This paper is dealing with the oscillatory properties of first order delay neutral impulsive differential equations and corresponding to them inequalities with constant coefficients. The established sufficient conditions ensure the oscillation of every solution of this type of equations.

Key words and phrases: oscillation of solutions, neutral impulsive delay differential equations and inequalities, constant coefficients.

AMS (MOS) Subject Classifications: 34K11, 34K40, 34A37.

1. Introduction

Impulsive differential equations with deviating arguments (IDEDA) are adequate mathematical models for the simulation of processes that depend on their history and are subject to short-time disturbances. Such processes occur in the theory of optimal control, theoretical physics, population dynamics, biotechnology, industrial robotics, etc. In contrast to the theory of ordinary impulsive differential equations (see, [1] - [3] and [20]) and differential equations with deviating arguments (see, [11], [13], [14]and [18]), the theory of IDEDA admits some theoretical and practical difficulties. We note here that [12] is the first work where IDEDA were considered. For more results, concerning IDEDA, we choose to refer to [4]-[6],[8],[23] and [24]. Much less we know about the neutral impulsive differential equations, i.e. equations in which the highest-order derivative of the unknown function appears in the equation with the argument t (the present state of the system), as well as with one or more retarded and/or advanced arguments (the past and/or the future state of the system). Note that equations of this type appear in networks, containing lossless transmission lines. Such networks arise , for example, in high speed computers, where lossless transmission lines are used to interconnect switching circuits (see, [7] and [21]).

As it is known (see [11]), the appearance of the neutral term in a differential equation can cause or destroy the oscillation of its solutions. Moreover, the study of neutral differential equations in general , presents complications which are unfamiliar for non-neutral differential equations. As far as for a discussion on some more applications and some drastic differences in behavior of the solution of neutral differential equations see, for example, [15],[16] and [22].

2. Preliminaries

In this article we consider the first order delay neutral impulsive differential equation with constant coefficients of the form

$$\frac{d}{dt}[y(t) - cy(t-h)] + qy(t-\sigma) = 0, \ t \neq \tau_k \qquad (E_1)$$

$$\Delta[y(\tau_k) - cy(\tau_k - h)] + p_k y(\tau_k - \sigma) = 0, \ k \in N$$

as well as the corresponding to it inequalities

$$\frac{d}{dt}[y(t) - cy(t-h)] + qy(t-\sigma) \leq 0, \ t \neq \tau_k \qquad (N_{1,\leq})$$

$$\Delta[y(\tau_k) - cy(\tau_k - h)] + p_k y(\tau_k - \sigma) \leq 0, \ k \in N$$

and

$$\frac{d}{dt}[y(t) - cy(t-h)] + qy(t-\sigma) \geq 0, \ t \neq \tau_k \qquad (N_{1,\geq})$$

$$\Delta[y(\tau_k) - cy(\tau_k - h)] + p_k y(\tau_k - \sigma) \geq 0, \ k \in N$$

where $c \in (0,1)$, q, $p_k \in [0,+\infty), k \in N$ and $h, \sigma \in (0,+\infty)$.

Moreover, we consider a special case of the equation (E_1) and the corresponding to it inequalities which are of the form

$$y'(t) + qy(t-\sigma) = 0, \ t \neq \tau_k \qquad (E_2)$$

$$\Delta y(\tau_k) + p_k y(\tau_k - \sigma) = 0, \ k \in N$$

$$y'(t) + qy(t-\sigma) \leq 0, \ t \neq \tau_k \qquad (N_{2,\leq})$$

$$\Delta y(\tau_k) + p_k y(\tau_k - \sigma) \leq 0, \ k \in N$$

and

$$y'(t) + qy(t-\sigma) \geq 0, \ t \neq \tau_k \qquad (N_{2,\geq})$$

$$\Delta y(\tau_k) + p_k y(\tau_k - \sigma) \geq 0, \ k \in N$$

respectively.

Here the deviations h and/or σ are positive constants and $\tau_k \in (0,+\infty), k \in N$ are fixed moments of impulsive effect (the jump points), which we characterize as *down-jumps* when $\Delta x(\tau_k) < 0$, $k \in N$ and as *up-jumps* when $\Delta x(\tau_k) > 0$, $k \in N$.

Denote by $PC(R,R)$ the set of all piecewise continuous on the intervals $(\tau_k, \tau_{k+1}]$, $k \in N$ functions $u: R \longrightarrow R$ which at the jump points τ_k, $k \in N$ are continuous from the left, i.e. $u(\tau_k - 0) = \lim_{t \to \tau_k - 0} u(t) = u(\tau_k)$, and may have discontinuities of first kind at the jump points τ_k, $k \in N$.

Suppose that the fixed moments of impulsive effect (the jump points) τ_k, $k \in N$ have the properties

$$t_0 = \tau_0 < \tau_1 < \tau_2 < \ldots < \tau_k < \ldots, \lim_{k \to +\infty} \tau_k = +\infty, \max\{\tau_{k+1} - \tau_k\} < +\infty, \ k \in N$$

Moreover, the following notions will be used throughout this paper.

A continuous real valued function u defined on an interval of the form $[a, +\infty)$ eventually has some property if there is a number $b \geq a$ such that u has this property on the interval $[b, +\infty)$.

Let $\rho = \max\{\sigma, h\}$. We will say that a function $y(t)$ is a solution of $Eq.(E_1)$, if there exists a number $T_0 \in R$ such that $y \in PC([T_0 - \rho, +\infty], R)$, the function $z(t) = y(t) - cy(t-h)$ is continuously differentiable for $t \geq T_0$, $t \neq \tau_k$, $k \in N$ and $y(t)$ satisfies $Eq.(E_1)$ for all $t \geq T_0$.

Furthermore, our results here pertain only to the nontrivial continuable solutions $y(t)$ of the equation (E_1), i.e. $y(t)$ is defined on an interval of the form $[T_y, +\infty)$ for some $T_y \geq T_0$ and

$$\sup\{|y(t)|: \ t \geq T\} > 0 \text{ for each } T \geq T_y.$$

Such a solution of $Eq.(E_1)$ is called *regular.* A regular solution $y(t)$ of $Eq.(E_1)$, is said to be *nonoscillatory,* if there exists a number $t_0 \geq 0$ such that $y(t)$ is of constant sign for every $t \geq t_0$. Otherwise, it is called *oscillatory*. Also, note that a *nonoscillatory* solution is called *eventually positive* (*eventually negative*), if the constant sign that determines its *nonoscillation* is positive (negative). Equation (E_1) is called *oscillatory,* if all its solutions are oscillatory. Otherwise, it is called *nonoscillatory.*

In what follows we will consider $Eq.(E_1)$, only in the cases, where it is a **neutral** ($h \neq 0$, $c \neq 0$) and an **impulsive** ($p_k \neq 0$ or $p_k = 0$ with $\tau_{k+1} - \tau_k = h$, $k \in N$) **differential equation** with two different deviations ($\sigma \neq 0$, $h \neq 0$, $\sigma \neq h$) or with a single deviation ($\sigma = h \neq 0$). So, in what follows, without further mention, we will assume that

$$c \in (0,1), \ q, p_k \in [0, +\infty), k \in N \text{ and } h, \sigma \in (0, +\infty)$$

Finally, in this article, when we write a functional expression, we will mean that it holds for all sufficiently large values of the argument.

Our aim is to establish sufficient conditions under which the equation (E_1) is oscillatory.

To this end, we need the following two lemmas.

The first lemma (see, [9],[10] and[13]) describes the asymptotic behavior of the functions $z(t) = y(t) - cy(t-h)$ and $y(t)$, where $y(t)$ is an eventually positive solution of $Eq.(E_1)$.

Lemma 1 *Let $y(t)$ be an eventually positive solution of $Eq.(E_1)$. Then:*

(a) $z(t) > 0$ for all large t with $\lim_{t \to +\infty} z(t) = 0$ and $\lim_{\tau_k \to +\infty} |\Delta z(\tau_k)| = 0$;

(b) $\lim_{t \to +\infty} y(t) = 0$ and $\lim_{\tau_k \to +\infty} |\Delta y(\tau_k)| = 0$.

Lemma 1, applied to the differentiable function $z(t) = y(t) - cy(t-h)$ and to twice differentiable function $w(t) = z(t) - cz(t-h)$, where $y(t)$ is an eventually positive solution of $Eq.(E_1)$, leads to the following proposition which is useful for our purposes.

Lemma 2 *Let $y(t)$ be an eventually positive solution of $Eq.(E_1)$. Then the functions $z(t) = y(t) - cy(t-h)$ and $w(t) = z(t) - cz(t-h)$ are also solutions of $Eq.(E_1)$ with the properties:*

(a) $z(t) > 0$, $z'(t) < 0$ eventually and

$$\lim_{t \to +\infty} z(t) = 0, \lim_{\tau_k \to +\infty} |\Delta z(\tau_k)| = 0;$$

(b) $w(t) > 0$, $w'(t) < 0$ and $w''(t) > 0$ eventually and

$$\lim_{t \to +\infty} w(t) = 0, \lim_{\tau_k \to +\infty} |\Delta w(\tau_k)| = 0.$$

Proof. As the negative of a solution of (E_1) is also a solution of the same equation, it suffices to prove the lemma for an eventually positive solution y(t) of (E_1). Thus, assume, for the sake of contradiction, that $y(t)$ is an eventually positive solution of (E_1). Then, since the equation (E_1) is an autonomous one, it follows that $y(t-h)$ is also a solution of (E_1). Therefore, $z(t)$ as a linear combination of solutions of (E_1) is itself a solution of (E_1). By similar arguments we easily conclude that $w(t)$ is also a solution of (E_1). Now, using Lemma 1, it is easy to see that for all large t

$$z(t) > 0, \; z'(t) < 0$$

and that

$$\lim_{t \to +\infty} z(t) = 0, \lim_{\tau_k \to +\infty} |\Delta z(\tau_k)| = 0.$$

By the same manner we conclude that for all large t

$$w(t) > 0, \; w'(t) < 0 \text{ and } w(t)'' = [z(t) - cz(t-h)]' = -qz'(t-\sigma) > 0, \; t \neq \tau_k$$

and that

$$\lim_{t \to +\infty} w(t) = 0, \lim_{\tau_k \to +\infty} |\Delta w(\tau_k)| = 0.$$

This completes the proof of the lemma.

3. Oscillation of all solutions of (E_2)

The results of this section will be used in the study of the oscillatory properties of (E_1) and the corresponding to it inequalities $(N_{1,\leq})$ and $(N_{1,\geq})$ respectively.

Consider the first order ordinary impulsive delay differential equation (E_2) and the corresponding to it inequalities $(N_{2,\leq})$ and $(N_{2,\geq})$, which are special cases of the equation (E_1).

Note that, as it is well-known (see, for example, [20] and [18]), a necessary and sufficient condition for the oscillation of all solutions of the delay differential equation (E_2), without impulsive effects, is that $q\sigma > \frac{1}{e}$. On the other hand, if the condition $q\sigma \leq \frac{1}{e}$ holds, then, according to a result in [17] (see also [18]), the delay differential equation (E_2), without impulsive effects, is non-oscillatory. Our results below, demonstrate the influence of impulsive effects on the behavior of solutions of (E_2). Indeed, Corollary 1 below shows the fact that the delay differential equation (E_2), subject to impulsive effects, is oscillatory even in the case, where $q\sigma \leq \frac{1}{e}$.

Theorem 1 *Assume that*

$$\liminf_{t\to+\infty}(q\sigma + \sum_{t-\sigma\leq\tau_k\leq t} p_k) \geq 1.$$

Then:

(a) the equation (E_2) is oscillatory;

(b) the inequality $(N_{2,\leq})$ has no eventually positive solutions;

(c) the inequality $(N_{2,\geq})$, has no eventually negative solutions.

Proof. Since the proofs of (a),(b) and (c) can be carried out by similar arguments, it suffices to prove only the case (a). To this end, as in the proof of Lemma 2, we assume that y(t) is an eventually positive solution of (E_2). Then there exists a $t_0 > 0$ such that $y(t) > 0$ for every $t > t_0$. Also, there is a $t_1 \geq t_0+\sigma$ such that $y(t-\sigma) > 0$, $y'(t) < 0$ and $\Delta y(\tau_k) = -p_k y(\tau_k-\sigma) < 0$, $k \in N$ for every $t \geq t_1$. That means that y is decreasing function with down-jumps $(\Delta y(\tau_k) < 0)$, $k \in N$.

Integrating (E_2) from $t-\sigma$ to t, we find

$$y(t) - y(t-\sigma) - \sum_{t-\sigma\leq\tau_k\leq t} \Delta y(\tau_k) + \int_{t-\sigma}^{t} qy(s-\sigma)ds = 0.$$

Remark that, because $y(t)$ is a positive decreasing function of t , from last equality we derive

$$-y(t-\sigma) + \sum_{t-\sigma\leq\tau_k\leq t} p_k y(\tau_k-\sigma) + q\sigma y(t-\sigma) \leq 0. \tag{1}$$

as well as

$$y(\tau_k-\sigma) > y(t-\sigma) > 0, \quad \text{when } \tau_k-\sigma < t-\sigma.$$

Hence, (1) yields

$$y(t-\sigma)(-1+q\sigma+\sum_{t-\sigma\leq\tau_k<t} p_k)<0$$

and finally we conclude that

$$q\sigma+\sum_{t-\sigma\leq\tau_k\leq t} p_k<1.$$

But the last inequality contradicts our assumptions and the conclusion of the theorem is evident.

As a consequence of the above theorem, we have the following important

Corollary 1 *Let* $0\leq q\sigma\leq\frac{1}{e}$ *and assume that* $\liminf_{t\to+\infty}\sum_{t-\sigma\leq\tau_k\leq t} p_k\geq 1$.

Then the conclusion of Theorem 1 holds.

We conclude this section with the following

4. Oscillation of all solutions of Eq. (E_1).

Having in mind the results of the previous section, we establish our main result which ensure the oscillation of all solutions of the equation (E_1).

Theorem 2 *Assume that* $\sigma>h$ *and that*

$$\liminf_{t\to+\infty}[q(\sigma-h)+\sum_{t-(\sigma-h)\leq\tau_k\leq t} p_k]\geq 1-c.$$

Then:

(a) the equation (1) is oscillatory;

(b) the inequality (2) has now eventually positive solutions;

(c) the inequality (3) has no eventually negative solutions.

Proof. As in the proof of Theorem 1, we prove only the case (a). To do that, as in the proof of Lemma 2, we assume, for the sake of contradiction, that Eq. (E_1) has an eventually positive solution $y(t)$. Then there exists a $t_0>0$ such that $y(t)>0$ for every $t>t_0$. Also, there is a $t_1\geq t_0+\sigma$ such that $y(t-\sigma)>0$, $y'(t)<0$ and $\Delta[y(\tau_k)-cy(\tau_k-h)]=-p_k y(\tau_k-\sigma)<0$, $k\in N$ for every $t\geq t_1$. Now, by Lemma 2, it follows that for every $t\geq t_1$ the functions $z(t)=y(t)-cy(t-h)>0$ and $w(t)=z(t)-cz(t-h)>0$ are solutions to the equation (E_1). That is, $w(t)$ satisfies the equation

$$\frac{d}{dt}[w(t)-cw(t-h)]+qw(t-\sigma)=0,\ t\neq\tau_k, \tag{2}$$

$$\Delta[w(\tau_k) - cw(\tau_k - h)] + p_k w(\tau_k - \sigma) = 0, \;\; k \in N$$

Note that, by Lemma 2, $w(t)$ is an eventually positive strongly decreasing, while $w'(t)$ is an eventually negative strongly increasing function. Therefore, it is easy to see that

$$\begin{aligned} w'(t-h) - cw'(t-h) + qw(t-\sigma+h) &\leq w'(t) - cw'(t-h) + qw(t-\sigma) \\ &= \frac{d}{dt}[w(t) - cw(t-h)] + qw(t-\sigma) = 0 \end{aligned} \tag{3}$$

Moreover, since $z(t)$ is a decreasing function, we see that $z(\tau_k - \sigma) < z(\tau_k - \sigma - h)$ and so, using the definitions of the functions $z(t)$ and $w(t)$, it is easy to conclude that

$$\Delta w(\tau_k) = -p_k z(\tau_k - \sigma) > -p_k z(\tau_k - \sigma - h) = \Delta w(\tau_k - h), \;\; k \in N$$

So, in view of the above observation, from (2) it follows that for each $k \in N$

$$\begin{aligned} \Delta w(\tau_k - h) - c\Delta w(\tau_k - h) + p_k w(\tau_k - \sigma + h) &\leq \Delta w(\tau_k) - c\Delta w(\tau_k - h) + p_k w(\tau_k - \sigma) \\ &= \Delta[w(\tau_k) - cw(\tau_k - h)] + p_k w(\tau_k - \sigma) = 0 \end{aligned} \tag{4}$$

Now, by (3) and (4) , it follows that $w(t)$ is an eventually positive function for which

$$(1-c)w'(t-h) + qw(t-\sigma+h) \leq 0, \;\; t \neq \tau_k$$

$$(1-c)\Delta w(\tau_k - h) + p_k w(\tau_k - \sigma + h) \leq 0, \;\; k \in N$$

Hence, we conclude that $w(t)$ is an eventually positive solution to the inequality

$$w'(t) + \frac{q}{1-c} w(t-\sigma+h) \leq 0, \;\; t \neq \tau_k, \tag{5}$$

$$\Delta w(\tau_k) + \frac{p_k}{1-c} w(\tau_k - \sigma + h) \leq 0, \;\; k \in N$$

which is a contradiction. Indeed, the inequality (5) is of the form $(N_{2,\leq})$. But, by Theorem 1(b), the inequality (5) can not have eventually positive solutions.

The proof of the theorem is complete.

As consequences of the above theorem, we formulate the following propositions, the first of which is an analogous to Corollary 1.

Corollary 2 *Assume that* $0 \leq q(\sigma - h) \leq \frac{1}{e}$ *and that*

$$\liminf_{t \to +\infty} \sum_{t-(\sigma-h) \leq \tau_k \leq t} p_k \geq 1 - c.$$

Then the conclusion of Theorem 2 holds.

Corollary 3 *Assume that $q(\sigma - h) \geq \frac{1}{e}$ and that*

$$\liminf_{t \to +\infty} \sum_{t-(\sigma-h) \leq \tau_k \leq t} p_k \geq 1 - c - \frac{1}{e}.$$

Then the conclusion of Theorem 2 holds.

Next will be the result in the case of single deviation of Eq.(E_1).

Theorem 3 *Assume that $\sigma = h$ and that*

$$\liminf_{t \to +\infty} [qh + \sum_{t \leq \tau_k \leq t+h} p_k] \geq 1 + c.$$

Then Eq. (E_1) is oscillatory.

Proof. Let, for the sake of contradiction, $y(t)$ be an eventually positive solution solution of the equation (E_1). Then, in view of Lemma 2, the function $z(t) = y(t) - cy(t-h)$ and $w(t) = z(t) - cz(t-h)$ are eventually positive solutions to the equation (E_1). That is, w(t) satisfies

$$[w(t) - cw(t-h)]' + qw(t-h) = 0 \ , t \neq \tau_k \tag{6}$$

$$\Delta[w(\tau_k) - cw(\tau_k - h)] + p_k w(\tau_k - h) = 0, \ k \in N.$$

Integrating Eq. (6) from t to $t+h$, we obtain

$$w(t+h) - w(t) - c[w(t) - w(t-h)] + \sum_{t \leq \tau_k \leq t+h} p_k w(\tau_k - h) + q \int_t^{t+h} w(s-h)ds = 0,$$

and equivalently

$$w(t+h) - w(t) - c[w(t) - w(t-h)] + \sum_{t \leq \tau_k \leq t+h} p_k w(\tau_k - h) + qhw(t+h-h) \leq 0. \tag{7}$$

Since w(t) is a decreasing function of t, we see that $w(\tau_k - h) > w(t)$ for $t \leq \tau_k \leq t+h$ and so, from (7) we derive

$$w(t)(-1 - c + \sum_{t \leq \tau_k \leq t+h} p_k + qh) < 0,$$

which implies that

$$qh + \sum_{t \leq \tau_k \leq t+h} p_k < 1 + c.$$

The obtained contradiction proves the theorem.

As a consequence of the above theorem, we formulate the following proposition, that is an analogous to Corollary 1.

Corollary 4 *Assume that* $\sigma = h$ *and* $0 \le qh \le \frac{1}{e}$ *and that*

$$\liminf_{t\to+\infty} \sum_{t\le\tau_k\le t+h} p_k \ge 1 + c.$$

Then:

(i) the equation (1) is oscillatory;

(ii) the inequality (2) has no eventually positive solutions;

(iii) the inequality (3) has no eventually negative solutions.

We conclude with an example, which illustrates Theorem 3 and its Corollary 4.

Example 1 *The neutral impulsive differential equation with* $\tau_{k+1} - \tau_k = 1, \ k \in N$

$$[y(t) - \frac{1}{2}y(t-1)]' = 0, \ t \ne \tau_k$$

$$\Delta[y(\tau_k) - \frac{1}{2}y(\tau_k - 1)] + \frac{3}{2}y(\tau_k - 1) = 0 \ , k \in N,$$

for every $t > \tau_0 = 0$ *satisfies the assumptions of Corollary 4 of Theorem 3, i.e.*

$\liminf_{t\to+\infty} [qh + \sum_{t\le\tau_k\le t+h} p_k] \ge 1 + c$, where $\sigma = h = 1, q = 0, \ c = \frac{1}{2}, \ p_k = p = \frac{3}{2}$.

Hence, it has only oscillatory solutions. It is obvious that these solutions will be in the form of piece-vice constant functions $y(t) = A_k$, *for* $t \in (\tau_{k-1}, \tau_k], \ k \in N, \ t > \tau_0 = 0$ *with initial function*

$$\varphi(t) = A_0, \ t \in [\tau_0 - 1, \tau_0,], \ A_0 \in R$$

where the "pulsatile" coefficients A_k *are determined by the difference scheme*

$$\Delta[y(\tau_k) - \frac{1}{2}y(\tau_k - 1)] + \frac{3}{2}y(\tau_k - 1) = 0 \ , k \in N,$$

$$i.e. \quad A_{k+1} = y(\tau_{k+1}) = (1 + c)A_k - (p + c)A_{k-1},$$

where

$$A_{k-1} = y(\tau_{k-1}), \quad A_k = y(\tau_k), \quad A_{-1} = A_0.$$

References

[1] Bainov D.D., Lakshmikantham V., Simeonov P.S., *Theory of Impulsive Differential Equations,* World Scientific, Singapore, 1989.

[2] Bainov D.D., Simeonov P.S., *Systems with impulse effect: Stability, Theory and Applications,* Ellis Horwood Ltd, Chichester, 1989.

[3] Bainov D.D., Simeonov P.S., *Impulsive Differential Equations: Periodic solutions and Applications,* Longman, Harlow, 1993.

[4] BAINOV D.D., DIMITROVA M.B., *Sufficient conditions for oscillation of the solutions of a class of Impulsive Differential Equations with deviating arguments*, J.Appl.Math., Stochastic Analysis, 9(1996), 33-42.

[5] BAINOV D.D., DIMITROVA M.B., DISHLIEV A., *Necessary and sufficient conditions for existence of non-oscillatory solutions of a class of Impulsive Differential Equations of second order with retarded argument*, Applicable Analysis 63(1996), 287-297.

[6] BAINOV D.D., DIMITROVA M.B., SIMEONOV P.S., *Sufficient conditions for the oscillation of the solutions of a class of impulsive differential equations with advanced argument,* Publ. Inst.Math.(Beograd) 59(1996), 39-48.

[7] BRAYTON R.K., WILLOUGHBY R.A., *On the numerical integration of a symmetric system of difference-differential equations of neutral type* , J. Math.Anal. Appl. 18(1967), 182-189.

[8] DIMITROVA M.B., *Criteria for oscillation of Impulsive Differential Equations of first order with deviating arguments,* Kyungpook Math.J.,40 (2000), 29-37.

[9] DONEV V.I. , *Existence and asymptotic behavior of the positive solutions of a class of neutral impulsive differential equations in their spacial cases,* 28-th IMSS, Sozopol, 2002.

[10] DONEV V.I., DIMITROVA M.B., *Oscillatory properties of the solutions of impulsive differential equations with deviating arguments,* Announcements ot the Union of scientists - Rousse, 4 (2003), 88-93.

[11] ERBE L.H., KONG Q., ZHANG B.G., *Oscillation Theory for Functional Differential Equations,* Marcel Decker Inc., New York, 1995.

[12] GOPALSAMY K., ZHANG B.G., *On delay Differential Equations with impulses,*J.Math.Anal.Appl. 139(1989), 110-122.

[13] GRAMMATIKOPULOS M.K., GROVE E.A., LADAS G., *Oscillation and asymptotic behavior of Neutral Differential Equations with Deviating Arguments,* Applicable Analysis 22(1986),1-19.

[14] GYORI I., LADAS G.,*Oscillation Theory of Delay Differential Equations with Applications,* Clarendon Press, Oxford, 1991.

[15] HALE J.K., *Stability of linear systems with delays, Stability problems,* C.I.M.E., June 1974 (Cremoneze, Roma), pp. 19-35.

[16] HALE J.K., *Theory of functional differential equations,* Springer Verlag, New York, 1977.

[17] KOPLATADZE R.G. AND CHANTURIA T.A., *On oscillatory and monotonic solutions of first order differential equations with deviating arguments,* Differentcial'nye Uravnenija 18(1982),1463-1465(Russian).

[18] LADDE G.S., LAKSHMIKANTHAM V., ZHANG B.G., *Oscillation theory of differential equations with deviating arguments*, Pure and Applied Mathematics, Vol. 110, Marcel Decker, New York, 1987.

[19] MYSHKIS A.D., *Linear differential equations with retarded argument (Russian),* Moskow, Nauka, 1972.

[20] SAMOILENKO A.M., PRESTYUK N.A., *Impulsive differential equations ,* World Sci. Ser. A: Monograph and Treatises - 1995.- 14

[21] SLEMROD M., INFANTE E.F., *Asymptotic stability criteria for linear system of difference-differential equations of neutal type and their discrete analogues*, J.Math.Anal.Appl. 38, 388-415,1972

[22] SNOW W., *Existence, uniqueness and stability for nonlinear differential-difference equation in the neutral case*, N.Y.U. Curant Inst. Math.Sci.Rep. IMM-NYU 328 (February) 1965.

[23] Zhang Y., Zhao A., Yan J., *Oscillation criteria for Impulsive Delay Differential Equations*,J.Math.Anal.Appl. 205, 461-470, 1997.

[24] Zhao A., Yan J., *Existence of positive solutions of Delay Differential Equations with impulses,* J.Math.Anal.Appl., 210 (667-678), 1997.

(1)Department of Mathematics
University of Ioannina
451 10 Ioannina, Greece

(2)Department of Mathematics
Technical University
8800 Sliven, Bulgaria